Analysis of Interconnected Systems with Nonlinear and Hybrid Dynamics

Analyse des Systèmes Interconnectés avec Dynamiques Nonlinéaires et Hybrides

Mémoire d'Habilitation à Diriger les Recherches

Université de Toulouse Paul Sabatier

par

Aneel Tanwani
Chargé de Recherche au CNRS

Soutenu le 14 Mars, 2025

Rapporteurs:

ANTOINE GIRARD, Directeur de Recherche au CNRS, France
LARS GRÜNE, Professeur, Universität Bayreuth, Allemagne
JOAO HESPANHA, Professeur, University of California at Santa Barbara, États-Unis

Jury:

BERNARD BROGLIATO, Directeur de Recherche à l'INRIA, France
ANTOINE GIRARD, Directeur de Recherche au CNRS, France
LARS GRÜNE, Professeur, Universität Bayreuth, Allemagne
DANIEL LIBERZON, Professeur, University of Illinois at Urbana-Champaign, États-Unis
CHRISTOPHE PRIEUR, Directeur de Recherche au CNRS, France
HYUNGBO SHIM, Professeur, Seoul National University, Corée du Sud
SOPHIE TARBOURIECH, Directrice de Recherche au CNRS, France

Dedicated to . . .

Neetu, for her love and support

Anaya and Arvin, for all the joy they bring

Contents

Preface

The field of systems and control theory continues to evolve at a rapid pace and as a result, the mathematical models used for modern engineering systems are becoming increasingly complex. For analyzing these systems, one viewpoint that helps for analysis is to regard these systems as an interconnection of several subsystems via some coupled terms. This allows us to reduce the complexity of the analysis and develop a general methodology for synthesis of control systems. With this motivation, this manuscript is aimed at the study of interconnections of dynamical systems, where the particular focus is on systems with nonlinear and hybrid dynamics.

The results presented in this manuscript are based on my collaborations since my recruitment at LAAS – CNRS in Toulouse, but I will mostly focus on my work from last 5 years. A major part of this manuscript is based on analyzing switched and hybrid systems in deterministic and stochastic setting. I was first exposed to these topics during my PhD and so, I want to first thank Daniel Liberzon for not only being the primary contact for most of my scientific knowledge, but also for nurturing a decision making ability which plays an important role in shaping the career of a researcher. I am grateful for the fact that my interaction with Daniel has continued after the PhD and the support I have received from him over the years. The material presented in Section 2.1 is based on a collaboration with him and Shenyu Liu [140], which partly builds upon the work initiated in [233].

An important result based on the interconnections involving nonlinear and hybrid systems can be found in Chapter 3, which is based on collaboration with Hyungbo Shim and Andrew Teel. My interactions with Andy have been rather brief but, at the same time, very inspiring. His work speaks volumes of his impact on the research topic studied in this manuscript and I have always admired the feedback received while interacting with him. With Hyungbo, I basically started working as a PhD student at University of Illinois. It is really amazing how our small conversations have led to so much collaborations. His experience on analysis of nonlinear control systems has helped me in several ways over the years. After our initial work during PhD, we could resume our collaboration by securing a grant financed by CNRS and Korean funding agency (NRF). I particularly thank him for making a trip to Toulouse in 2021 when it was difficult to travel due to Covid-19 pandemic. This trip led to the initial work [212], and also resulted in an amazing excursion to the vineyards of Saint-Émilion. The results reported here are based on [213], and stem from the aforementioned grant.

In Chapter 4, we study a class of nonsmooth systems which I first came across during my postdoc with Bernard Brogliato and Christophe Prieur at INRIA, Grenoble. Bernard, with his colleagues at INRIA, has studied several variants of this system class in depth from analysis and simulation viewpoint, with a particular focus on their relevance in mechanics. During this collaboration with Bernard, I started learning different properties of this system class which are mostly captured by looking at the underlying maximal monotone mapping that describes the nonsmooth relation. In particular, we adopt the viewpoint that a large class of such systems can be seen as an interconnection of an ordinary differential equation with a maximal monotone mapping. This even allows us to make connections with analysis of optimization algorithms and other topics. I could partly investigate the later topics with Kanat Camlibel, but this needs further investigation. The results based on these collaborations are only mentioned briefly but the research on these topics has provided a platform which I continue to explore in several directions. In particular, my first ANR grant, that I received in 2017, is closely related to these topics. Some of the results developed during this project as a part of jointly supervised PhD thesis with Didier Henrion, have been reported in Chapter 4. In fact, we could use tools based on polynomial optimization to provide computational methods for stability analysis for a class of state-constrained systems, and we continue to explore these computational methods for design problems like optimal control in our ongoing jointly supervised PhD program.

Moving away from the deterministic setting, we study the important problem of state estimation in stochastic systems in Chapter 5. The hybrid nature of the problem comes from the fact that we mostly focus on systems with continuous-time dynamics and discrete (in time) observation process. My first exposure to studying state estimation problems for hybrid systems goes back to my PhD thesis, and seeing a certain interest in the study of estimation problems with time-sampled observations, I became curious to study the stochastic version of the problem. In relevance to the unifying theme of interconnections studied in this manuscript, we actually focus on two classes of estimation algorithms which involve interaction of multiple estimators. The first one is based on the notion of ensemble filters, or feedback particle filters. The other one is based on the idea of using multi-estimators to extract knowledge about some unknown parameters. In general, I greatly appreciate the interactions with Debasish Chatterjee, Lars Grüne, Joao Hespanha, Prashant Mehta, and Serdar Yuksel at various stages who either provided some technical feedback or the encouragement to study this topic.

I must acknowledge the interactions with all my PhD students: Matteo Della Rossa, Marianne Souaiby, Olga Yufereva and Saroj Chhatoi; and MS interns: Guang-Xue Zhang and Anubhab Dasgupta. Supervising these students has been an extremely refreshing experience. I particularly admire the professionalism with which they carried out their tasks despite several hardships around the time. I first studied the cascade interconnections with Guang-Xue in [233] which we also study in Section 2.1 but with different approach. Section 2.2 is based on the use of appropriate notion of derivative for the nonsmooth Lyapunov functions and it comes from the work done with Matteo during his PhD thesis [67]. The result in Section 4.4 comes from a joint publication with Marianne [197]. The work of ensemble filters in Section 5.3 comes from the PhD thesis of Olga [215].

Finally, all the work reported here has been carried out during my stay at LAAS, while being a part of the MAC team. I thank my colleagues for the learning curve that I have experienced as a member of this team.

1

Introduction

The study of dynamical systems has evolved over the past few decades so that it allows for analysis of complex behaviors in interconnected networks, such as communication channels coupled with dynamical systems, cyber-physical systems, biological ecosystems, social networks, power grids, and many more. In its current form, the research in systems and control theory relies on the analysis of a dynamical system described by (ordinary, stochastic, or partial) differential equations and to model the complex phenomenon in aforementioned applications, it is often necessary to consider the interaction of several such dynamical systems. The resulting interconnected system may manifest complex behavior which one needs to analyze to aid the design process.

In this manuscript, we consider the interconnections which comprise nonlinear and hybrid dynamics. Nonlinear systems, described by ordinary differential equation (ODEs), present a challenge from analytical and numerical perspective due to their sensitivity to initial conditions and the potential for chaotic behavior. Hybrid dynamical systems, which combine continuous and discrete dynamics, provide the most suitable framework for logic-based decision making and describing different modes of operation. Their combination allows us to model and study a wide variety of applications as the ones mentioned above. The term *interconnection* is used in a rather broad sense here and it basically captures how one system can be seen as a composition of several (possibly memoryless) subsystems and how the properties of individual elements of a system can describe the behavior of the overall system. In the subsequent chapters, we report several works which can be commonly seen from the lens of analyzing interconnected dynamical systems. The techniques used in these works are rather different from each and they basically present my three active areas of research. To provide an overview, Chapter 2 and Chapter 3 deal with switched and hybrid systems; Chapter 2 studies the classical interconnections with two subsystems either in the cascade form or feedback configuration, and Chapter 3 considers dynamical systems connected by a graph where the questions related to consensus are considered. Chapter 4 deals with systems where the vector field has monotonicity structure, and these are inherently motivated by nonlinearities which can be expresses as subdifferentials of convex functions (for which the gradient may not exist in the classical sense). Analysis here is mostly based on the theory of passivity and the interconnections here either refer to a dynamical system coupled to a static memoryless system with monotone properties, or the feedback interconnection of two dynamical systems with monotone properties. Finally, in Chapter 5, we refer to interconnections in stochastic setting while studying estimation, or filtering, problem for continuous-time stochastic systems with time-sampled observations. The interconnections basically result from the particular filtering algorithms that we study here, and we study the analysis techniques for such continuous-discrete interacting filters.

In this chapter, we first recall the different types of interconnections from the literature to explain the basic concepts that appear later in the manuscript. Section 1.1 and 1.2 present the basic terminology in this regard, and we summarize the contribution from each chapter in Section 1.3.

1.1 Interconnected Nonlinear ODEs

To set the stage for the different types of interconnections presented in this section, let us first consider the nonlinear system

$$\dot{x} = f(x, \omega), \quad x(0) \in \mathbb{R}^n, \tag{1.1}$$

where the vector field $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is assumed to be locally Lipschitz, and $\omega : \mathbb{R}_{\geq 0} \in \mathbb{R}^m$ is the external input.

For this conventional class of systems (1.1), several books are now available which treat topics ranging from existence of solutions, qualitative behavior of trajectories, sensitivity of solutions with respect to parameters, stability analysis with and without inputs, design of controllers and observers, and many more. The reader may refer to [219, 121, 84, 170] for fundamental treatment on these topics.

Comparison functions. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be *positive definite* if it is continuous, $\alpha(0) = 0$ and $\alpha(s) > 0$ for all $s > 0$. If α is also strictly increasing, then it is said to be of class $\mathcal{K}$. In addition if α is also unbounded, then it is said to be of class $\mathcal{K}_\infty$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to be of class $\mathcal{KL}$ if $\beta(\cdot, t)$ is of class $\mathcal{K}$ for each $t \in \mathbb{R}_{\geq 0}$, $\beta(s, \cdot)$ is continuously decreasing and $\beta(s, t) \to 0$ as $t \to \infty$ for each fixed $s \in \mathbb{R}_{\geq 0}$. In addition, we require class $\mathcal{KLL}$ functions: A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a class $\mathcal{KLL}$ function if $\beta(\cdot, \cdot, j)$ is a class $\mathcal{KL}$ function for each $j \geq 0$ and $\beta(\cdot, t, \cdot)$ is a class $\mathcal{KL}$ function for each $s \geq 0$. A comprehensive reference that collects several interesting properties of these comparison functions with focus on dynamical systems is [120].

Using the comparison functions, we can formulate some basic definitions of interest for system (1.1) and later explain their relevance for studying interconnections.

Definition 1.1. We say that the system (1.1) is globally asymptotically stable (GAS) with respect to the origin, or the origin is GAS, if there exists a class $\mathcal{KL}$ function β such that, for every $t \geq 0$, it holds that

$$|x(t)| \leq \beta(|x(0)|, t). \tag{1.2}$$

It is a common practice to describe asymptotic stability using a class $\mathcal{KL}$ function, as done in (1.2). However, it follows from [194, Proposition 7] that for every class $\mathcal{KL}$ function β, we can always find class $\mathcal{K}_\infty$ functions η, θ, and a scalar $k > 0$ such that $\beta(r, t) \leq \eta(e^{-kt}\theta(r))$. We say that the origin is GAS(η) for system (1.1) if, for every $t \geq 0$, it holds that

$$|x(t)| \leq \eta\Big(e^{-k\,t}\theta(|x(0)|)\Big).$$

In the presence of input ω, we consider the notion of *input-to-state stability* (ISS) and *integral input-to-state stability* (iISS) pioneered in [193] and [194], respectively.

Definition 1.2. The system (1.1) is said to be *input-to-state stable* (ISS) with respect to the origin if there exists $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ such that, for every $t \geq 0$, it holds that

$$|x(t)|_{\mathcal{A}} \leq \beta(|x(0)|, t) + \gamma(\operatorname*{ess\,sup}_{s\in[0,t]} |\omega(s)|), \tag{1.3}$$

where $\operatorname{ess\,sup}_{s\in I} |\omega(s)|$ denotes the essential supremum norm of ω over a set I (that is, the supremum of ω on I except for a set of measure zero).

Definition 1.3. The system (1.1) is said to be *integral input-to-state stable* (iISS) with respect to the origin if there exist $\alpha \in \mathcal{K}_\infty, \beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ such that, for every $t \geq 0$, it holds that

$$\alpha(|x(t)|) \leq \beta(|x(0)|, t) + \int_0^t \gamma(|\omega(s)|)\mathrm{d}s. \tag{1.4}$$

With appropriate generalizations, these two properties present the basic tools that are required for studying interconnections involving hybrid systems, and this will be done later in Chapter 2 and Chapter 3. For the time being, let us present the basic methodology and some references that describe the use of these properties in studying interconnections with nonlinear ODEs.

1.1.1 Cascade connections

The study of cascade interconnections has been instrumental in improving the design techniques for nonlinear control systems. The cascades are modeled by following differential equations:

$$\dot{x} = f_x(x, z) \tag{1.5a}$$
$$\dot{z} = f_z(z) \tag{1.5b}$$

where $f_z : \mathbb{R}^m \to \mathbb{R}^m$ describes the driving system, and $f_x : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ characterizes the driven system with input z. Unlike linear systems, an interesting aspect of the cascade configuration is that the globally asymptotic stability (GAS) of the driving system, and that of the driven system with 0 input, does not ensure stability of the interconnection. To ensure stability, we need to characterize the response of the driven system to the input z appropriately, and make it compatible with the convergence rate of the driving system. One possible way to do so is to impose an input-to-state stability (ISS) assumption for the x-subsystem (1.5a) and GAS assumption on z-dynamics, to ensure that the overall cascade is GAS. We can also derive more relaxed conditions which only involve the iISS property. We state one such result below which appeared in [10]:

Theorem 1.4. *Consider the cascade interconnection in* (1.5) *and suppose that*

- *the system* (1.5a) *is iISS with respect to the origin and the gain* γ
- *the system* (1.5b) *is GAS*(η) *with respect to the origin*

and it holds that

$$\int_0^1 \frac{(\gamma \circ \eta)(s)}{s}\mathrm{d}s < \infty, \tag{1.6}$$

then the interconnected system (1.5) *is GAS with respect to the origin.*

This result describes the system properties on the driving system (1.5b), the driven system (1.5a) and the coupling condition (1.6). It can be checked that if (1.5a) satisfies the stronger property of being ISS, then the gain μ can be chosen such that the condition (1.6) holds for any class $\mathcal{K}_\infty$ function η. However, the criteria involving iISS property, stated in Theorem 1.4, provides a more generic result. In Chapter 2 of this manuscript, we will explore the cascade interconnections when the x-dynamics and z-dynamics are both switching using the appropriate generalization of iISS property.

1.1.2 Small-gain configuration

One can think of cascades from the previous section as a particular case of a more general interconnection of nonlinear systems in feedback configuration which are analyzed using small-gain theorems. In particular, small-gain conditions are used to analyze the stability of two nonlinear systems in feedback configuration described by following equations:

$$\dot{x} = f_x(x, z) \tag{1.7a}$$

$$\dot{z} = f_z(z, x). \tag{1.7b}$$

We associate ISS property with each of the two constituent subsystems. More formally, we say that the system (1.7a) is ISS with respect to input z if there exists a class $\mathcal{KL}$ function β_x and a class $\mathcal{K}_\infty$ function γ_x such that,

$$|x(t)| \leq \beta_x(|x_0|, t) + \gamma_x(\sup_{s\in[0,t]} |z(s)|). \tag{1.8}$$

Similarly, we say that the z-subsystem is ISS with respect to the input x if there exists a class $\mathcal{KL}$ function β_z and a class $\mathcal{K}_\infty$ function γ_z such that,

$$|z(t)| \leq \beta_z(|z_0|, t) + \gamma_z(\sup_{s\in[0,t]} |x(s)|). \tag{1.9}$$

The functions γ_x and γ_z are called the gain functions and they are used to describe a small-gain condition. To write this condition, let us consider the identity function id, so that $\mathrm{id}(s) = s$, and suppose that there exist $\mathcal{K}_\infty$ functions ρ_x and ρ_z such that

$$(\mathrm{id} + \rho_z) \circ \gamma_z \circ (\mathrm{id} + \rho_x) \circ \gamma_x(s) \leq \mathrm{id}(s),$$
$$(\mathrm{id} + \rho_x) \circ \gamma_x \circ (\mathrm{id} + \rho_z) \circ \gamma_z(s) \leq \mathrm{id}(s),$$

then the coupled system (1.7) is GAS. Intuitively speaking, the small gain condition simply states that the effect of the external inputs is not amplified too much. Because of the rather general form of the interconnection in (1.7), we have seen a broad utility of such feedback configurations ranging from feedback control of a single system to large-scale systems. Consequently, many generalizations of small-gain theorems have been developed over the years. Some of the earlier results on small-gain can be found in [68, 103]. The treatment given here is based on the notion of ISS and the paper [111] provides several results for feedback connections using ISS and its variants. Building on this work, [112] provides a construction of a max-separable Lyapunov function under ISS like conditions on individual subsystems and the small gain condition. Their generalization to interconnections with more than two subsystems was considered in [58], and for networks of infinite size in [118, 119]. The discrete version of these results, while using ISS property, can be found in [113, 126] for two subsystems, and in [141] for more than two subsystems. Some results have also been developed for the case when the individual subsystems are not necessarily ISS using techniques based on sum separable Lyapunov functions [105, 106], or some geometric techniques [5].

Going beyond the feedback interconnections involving nonlinear systems, results have also been developed for interconnections with hybrid dynamics. These results partly stem from the deriving studying feedback stabilization of nonlinear systems with control using sampled-data [157], quantized measurements [133], or event-based sampling [143]. This has later developed into general theory for Lyapunov based formulation of small-gain theorems for hybrid systems with two subsystems in [134]. For more than two subsystems, with not necessarily ISS subsystems, these results have been generalized in [151]. In Chapter 2, we also provide a generalization of a small-gain theorem for state-dependent switched systems using non-differentiable Lyapunov functions.

1.1.3 Dissipativity-based analysis

Another important tool for analyzing the feedback interconnections is based on the theory of *dissipativity.* This technique describes the concepts that are later used (to some extent) in Chapter 4. For the sake of exposition, here we allow the interconnection of a passive dynamical system either with another passive dynamical system, or with a memoryless function. To elaborate further, we say that a memoryless system $y = h(u)$ is *passive* if $\langle u, y \rangle \geq 0$. Next, consider a dynamical system with input u and output y,

$$\begin{aligned} \dot{x} &= f(x, u) \\ y &= h(x) \end{aligned} \tag{1.10}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ locally Lipschitz, and $h : \mathbb{R}^n \to \mathbb{R}^m$ is continuous, with $f(0, 0) = 0$ and $h(0, 0) = 0$. We say that the system (1.10) is *passive* if there exists a continuously differentiable positive semidefinite function $V(x)$ (called the storage function) such that

$$\dot{V} = \langle \nabla V(x), f(x, u) \rangle \leq \langle u, y \rangle, \quad \forall\, (x, u) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Let us now consider the interconnection of two (possibly memoryless) passive systems:

$$\Sigma_i : \begin{cases} \dot{x}_i = f_i(x_i, e_i), \\ y_i = h_i(x_i) \end{cases} \tag{1.11}$$

for $i = 1, 2$. The feedback interconnection is described by taking, $e_1 = u_1 - h_2(e_1)$ and $e_2 = u_2 - h_1(e_1)$. In the first instance, we want to study the passivity of the closed-loop system taking the form

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

where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \qquad \text{and} \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

To analyze such systems, we have the following fundamental result about the preservation of passivity for the interconnection [121, Theorem 6.1]: "*The feedback interconnection of two passive systems is passive.*" This result can be proved by taking $V = V_1 + V_2$ as the storage function and checking that we indeed have the inequality $\dot{V} \leq \langle u, y \rangle$.

Passivity is an important property for analysis of dynamical systems. Combined with several other standard structural properties, we can establish finiteness of input-output gain, and also use different variants of passivity for synthesis of controllers. The notion of passivity and its generalized notion, dissipativity, comes from the pioneering work in [107] and continues to develop as a fundamental building block for analysis and design of control systems, as evident by several monographs. We see its utility in control design [191, 175], analyzing interconnections [11, 26], input-output gains [186], robustness [188], optimal control [86] and many more system-theoretic properties. One can even investigate the use of passivity in analysis of large-scale networks [12]. For our purposes, we will see that, in Chapter 4, similar ideas are useful in analyzing a class of systems that are described by maximal monotone mappings, as the definition of monotonicity has close resemblance with the notion of passivity.

1.1.4 Singular perturbation methods

The final tool which we want to present for the analysis of coupled nonlinear systems is the so-called *singular perturbation* method. This is specially useful when a parameter in the system takes a very small value and we want to understand the properties of the system that hold uniformly for small enough values of the parameter. As the problem description suggests, singular perturbation techniques are studied from the mathematical side for analyzing dynamical systems [71], and also from control theoretic perspective in design applications involving high gain [123, 124, 202], [121, Chapter 11]. One can also consult a more recent reference [160] for an overview and historical developments on the subject.

The basic idea of this technique is to decompose the system in two time-scales and analyze the stability of the overall system by associating desired properties with the subsystems resulting from the decomposition. The standard singular perturbation model can be seen as an interconnection of two dynamical systems, where the derivatives of the state for one of the subsystems is multiplied by a small positive parameter ε, that is,

$$\dot{x} = f(x, z, \varepsilon) \tag{1.12}$$
$$\varepsilon \dot{z} = g(x, z). \tag{1.13}$$

We assume that $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ are regular enough over a certain domain. By setting $\varepsilon = 0$, equation (1.13) reduces to

$$0 = g(x, z). \tag{1.14}$$

In standard singular perturbation models, it is assumed that the equation (1.14) has isolated roots for each value of x in the admissible domain. Focusing on one such root of the equation (1.14), we consider a function $h : \mathbb{R}^n \to \mathbb{R}^m$, so that (1.14) results in $z = h(x)$, for every x. Substituting this value of z in (1.12) results in the following reduced order model

$$\dot{x} = f(x, h(x), 0). \tag{1.15}$$

In other words, the reduced model (1.15) is obtained by substituting the value of z at the *quasi-steady-state* $h(x)$, and for this reason, we often call (1.15) the quasi-steady-state system (QSS). On the other hand, by introducing the stretched time variable $\tau = t/\varepsilon$, the *boundary-layer* (BL) system is described as

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = g(x, z). \tag{1.16}$$

For each x, we want to ensure that the solution of the BL system (1.16) converges to $h(x)$ with x a frozen variable. The study of singularly perturbed system, therefore, corresponds to describing the qualitative behavior of the coupled system (1.12)–(1.13) under appropriate assumptions on the quasi-steady-state system and the boundary-layer system.

Depending on whether the problem is being studied over a finite or an infinite interval, the solutions of (1.12) and (1.13) can be approximated using the solution of (1.15) and (1.16), respectively.

In our work presented in Chapter 3, we will use singular perturbation techniques for analyzing a specific class of hybrid systems. In the literature, we have seen the use of singular perturbation methods for a class of hybrid systems where fast evolution appears in continuous-time dynamics only [184, 223]. The solutions resulting from a continuous-time boundary-layer system generate an average vector field which approximates the slow subsystem. Some studies have also been carried out in the context of switched systems [146, 177, 46]. A recent overview on singular perturbation and averaging methods for systems with continuous and discrete dynamics appears in [1]. In our case, the singular perturbation parameter is introduced to control the frequency of jumps in the state trajectories, and such a setup is used to model the rapid communication between agents which update their states by exchanging information with neighboring agents at discrete times.

1.2 Interacting Stochastic Systems

The later part of this manuscript considers some work on interconnections arising in stochastic setting, and in particular, the filtering problem. Estimation theory in stochastic dynamical systems is a problem of great interest because of its relevance in several engineering related disciplines. Over the decades, we have seen development of several important concepts and elegant results in this field, stemming from the pioneering work of Kalman and Bucy [115]. If we take a look at the classical filtering problem with linear dynamics and Gaussian noise, the elegant closed form solution proposed in [115] requires us to simulate a stochastic differential equation (SDE) driven by the observation process with the injection gain depending on the error covariance process obtained by solving a Riccati differential equation. However, such a solution has its limitations as it cannot be readily applied to nonlinear systems and requires perfect knowledge of the model. We will study two different classes of algorithms in hybrid setting which potentially overcome these limitations. Let us first present the underlying concepts that go behind the design of these algorithms.

1.2.1 Mean-field approximations

The theory of mean-field approximations addresses the question of developing macroscopic models for a collection of particles, where the evolution of each particle is influenced by another particle. When the number of particles is large, the macroscopic model provides an approximation for the overall population, and one can deduce the qualitative behaviour of the particles directly from the macroscopic model. Some basic results on this topic can be found in the monographs [203], [199], and [83].

To provide a quick overview, let us consider a system of M particles interacting with each other and described via following SDEs, for $i = 1, \dots, M$:

$$\mathrm{d}X_t^i = f(X_t^i)\mathrm{d}t + u_i\mathrm{d}t + \mathrm{d}\omega_t^i, \qquad X_0^i = Z^i, \tag{1.17}$$

where ω_t^i are independent copies of the standard Wiener process, and the initial condition Z^i are independent, identically distributed (i.i.d.) random variables in $\mathbb{R}^n$. The mapping $f : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be globally Lipschitz continuous. The only source of coupling among the particles arises due to the input u_i, and without this coupling, each particle would evolve independently due to independent copies of the Brownian motion and i.i.d. initial conditions. For each $i = 1, \dots, M$, we choose

$$u_i = \frac{b}{M} \sum_{j=1}^{M} X_t^j$$

which results in

$$\mathrm{d}X_t^i = f(X_t^i)\mathrm{d}t + \frac{b}{M} \sum_{j=1}^{M} X_t^j \mathrm{d}t + \mathrm{d}\omega_t^i, \qquad X_0^i = Z^i. \tag{1.18}$$

We are interested in analyzing the system as the number of particles M gets large. To develop an intuition for what happens in that case, we see that as the number of particles gets large, the individual processes X_t^i become independent and identically distributed. Using the law of large numbers, we consider the following approximations

$$\frac{1}{M} \sum_{j=1}^{M} X_t^j \simeq \mathsf{E}[\overline{X}_t] = \int_{\mathbb{R}^n} y \eta_{\overline{X}_t}(\mathrm{d}y)$$

where $\overline{X}_t$ is an independent copy of X_t^i, and E denotes the expectation with respect to the law of $\overline{X}_t$. We can therefore approximate the evolution of X_t^i by $\overline{X}_t$ which is obtained by solving the following McKean-Vlasov differential equation:

$$\mathrm{d}\overline{X}_t = f(\overline{X}_t)\mathrm{d}t + b \int_{\mathbb{R}^n} y \, \eta_{\overline{X}_t}(\mathrm{d}y)\mathrm{d}t + \mathrm{d}\omega_t, \qquad \overline{X}_0 = Z$$

where Z is an independent copy of X_0^i for some $i = 1, \dots, N$.

Mean field approximation techniques have been studied in the context of filtering as well [21, 205], where the basic idea is to replace u_i in (1.17) with an appropriate term that may depend on the other particles in the population and the innovation obtained from the observation process. The key is to choose the input such that the resulting McKean-Vlasov differential equation can describe the optimal posterior distribution that we seek to compute in the first place. Even in linear case, the advantage of taking this route is that by simulating (a relatively small number of) particles, we can get very good approximations of the optimal distribution and the simulation of these particles does not require solving Ricatti differential equations because the injection gains are obtained from the empirical covariance of the population. We will explore these ideas further in Chapter 5.

1.2.2 Bayesian adaptive filtering

Another limitation of the standard filtering techniques comes from the fact that it requires exact knowledge of the model parameters (including the noise statistics) to get the optimal distribution. In general, this information may not be available or alternatively, we would want to develop estimation techniques which perform reasonably well in such cases. One solution to this problem comes from the literature on adaptive filtering with Bayesian inference.

Using standard results in filtering theory, see for example [55], it follows that the minimum mean square error (MMSE) estimate $\widehat{X}_t$ subject to the the information $\mathcal{Y}_{N_t}$, is given by,

$$\widehat{X}_t = \mathsf{E}[X_t \,|\, \mathcal{Y}_{N_t}] = \int X_t \mathsf{p}(X_t \,|\, \mathcal{Y}_{N_t}) \mathrm{d}X_t, \tag{1.19}$$

where $\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})$ denotes the probability density of θ conditioned upon the observations received up to time t. If Θ is a given set of parameters, and we are interested in computing the optimal estimate for a given a distribution of the parameters over the set Θ, then the foregoing equation actually yields,

$$\begin{aligned}\widehat{X}_t &= \int X_t \int_\Theta \mathsf{p}(X_t \,|\, \theta, \mathcal{Y}_{N_t})\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})\mathrm{d}\theta \,\mathrm{d}X_t, \\ &= \int_\Theta \widehat{x}_t(\theta)\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})\mathrm{d}\theta,\end{aligned} \tag{1.20}$$

where $\widehat{x}_t(\theta) := \mathsf{E}[X_t \,|\, \theta, \mathcal{Y}_{N_t}]$, for $\theta \in \Theta$, is the minimum mean square estimate with θ given.

If Θ is a discrete set, and is described as $\Theta := \{\theta_1, \theta_2, \cdots, \theta_K\}$ with K being the total number of unknown parameters, then we get

$$\widehat{X}_t := \sum_{i=1}^{K} \widehat{x}_t(\theta_i)\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t}). \tag{1.21}$$

That is, $\widehat{X}_t$ is obtained by computing $\widehat{x}_t(\theta_i)$ for each value of $\theta_i \in \Theta$ and taking their weighted sum as determined by the conditional probability mass function $\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t})$. As we collect more and more observations, the conditional distribution $\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t})$ is updated as a function of $\widehat{X}_t$. Thus, the filter obtained from (1.21) represents an interaction of different filters that generate $\widehat{x}_t(\theta_i)$, and is designed for the case when there are unknown parameters. We will explore this estimation algorithm in further detail in Chapter 5 where the objective is to quantify the asymptotic performance of the filters.

1.3 Contributions

This manuscript collects our works on studying interconnections along the directions that we highlighted in the previous section, where the systems in particular comprise continuous and discrete dynamics. The four subsequent chapters have been organized so that they can be read independently of each other.

Chapter 2 is basically aimed at studying some basic interconnections involving switched systems, a specific class of hybrid systems. In particular, the focus is on developing iISS and ISS results and using these results for studying cascade and small-gain configurations with switching dynamics. The two broad classes of switched systems that we consider are described by: a) time-dependent switching signal and b) state-dependent switching signals. For time-dependent switching systems, we develop conditions for ISS and iISS in case of nonlinear dynamics with inputs, resets and possibly unstable subsystems. For the dissipation inequalities associated with the Lyapunov function of each subsystem, it is assumed that the supply functions, which characterize the decay rate and ISS/iISS gains of the subsystems, are nonlinear. Using the notion of average dwell-time (ADT) to limit the number of switching instants on an interval, and the notion of average activation time (AAT) to limit the activation time for unstable systems, a formula relating ADT and AAT is derived to guarantee ISS/iISS of the switched system. Using the iISS property, we also discuss the stability of cascade interconnections with switching dynamics. These developments are primarily based on [233] and [140].

Secondly, when working with state-dependent switched systems, we first carry out the Fillipov regularization to embed such systems in differential inclusions. We then carry out robustness analysis for feedback interconnection of two systems described by upper semicontinuous differential inclusions using a generalized notion of derivatives associated with locally Lipschitz Lyapunov functions obtained from a finite family of differentiable functions. The analysis builds on sufficient conditions for ISS of differential inclusions using a class of nonsmooth (but locally Lipschitz) candidate Lyapunov functions and the concept of Lie generalized derivative. In general, our conditions are less conservative than the more common Clarke derivative-based conditions. These results have been published in [67].

Chapter 3 considers a theoretical framework of singularly perturbed hybrid systems which we have developed in [213], and it allows us to analyze the stability of a network of agents described by nonlinear oscillators which only exchange partial information about their state with some of its neighbors from time to time. We describe such systems as an interconnection of two hybrid subsystems, a timer which triggers the jumps, and some discrete variables to determine the index of the jump maps. The flow equations of these variables are singularly perturbed differential equations, and in particular, smaller value of the singular perturbation parameter leads to increase in the frequency of the jump instants. For the limiting value of this parameter, we consider a decomposition which comprises a quasi-steady-state system modeled by a differential equation without any jumps, and a boundary-layer system described by purely discrete dynamics. Under appropriate stability assumptions on the quasi-steady-state system and the boundary-layer system, we derive results showing practical stability of a compact attractor when the information exchange between the agents occurs frequently often.

Chapter 4 considers a class of dynamical systems for which the interconnections are studied using passivity-like properties. Inspired by the fact that the subgradient of non-differentiable convex functions have some monotonicity properties, we consider differential inclusions where the right-hand side has a maximal monotone mapping (with a negative sign) and possibly some drift term. Due to the relevance of these mappings in wide range of applications related to optimization, it is of interest to study which operations preserve maximal monotonicity. We review some results from [27], [36], and [208] which relate to preserving maximal monotonicity for interconnections involving passive dynamical systems. We then provide some sufficient conditions under which we can establish maximal monotonicity for two dynamical systems with maximal monotone systems in feedback configuration. In case, the dynamics cannot be reduced to a maximal monotone map and contain a drift term in addition, we consider the converse Lyapunov theorem for such systems and provide some algorithms for computing the Lyapunov functions in the class of rational polynomial functions. The latter developments in this chapter have been published in [197].

Chapter 5 makes the transition from deterministic to stochastic hybrid systems in the setting of filtering problem. For continuous-time linear stochastic systems, we consider two classes of filtering algorithms which involve interaction of multiple filters. The first such algorithm considers the design of ensemble filters when the observation process is randomly sampled. We propose a continuous-discrete McKean–Vlasov type diffusion process with additive Gaussian noise in observation model, which is used to describe the evolution of the particles. These particles are coupled through the empirical covariance and require less computations for implementation than the optimal ones based on solving Riccati differential equations. Using appropriate analysis tools, we show that the empirical mean and the sample covariance of the particle filters converges to the mean and covariance of the optimal filter if the mean sampling rate of the observation process satisfies certain inequality, and as the number of particles tends to infinity. The paper corresponding to these results is [215].

The second class of algorithms proposed in Chapter 5 is aimed at studying the filtering problem with unknown parameters in the system model. In particular, we assume that the unknown parameters belong to a discrete set and the posterior distribution of the unknown parameters conditioned upon available observations is computed from Bayes' rule. The resulting state estimate is a weighted sum of the state estimates generated by multiple Kalman filters, where the weights are determined by the posterior distribution of the unknown parameters. We analyze the performance of the algorithm by looking at its asymptotic behavior and establishing boundedness of the error covariance matrix. Our preliminary investigations along these lines have been reported in [207].

2

Switched Systems

Switched systems – comprising a family of dynamical subsystems orchestrated by a piecewise constant switching signal – provide a mathematical framework for modeling systems where the state trajectories exhibit sudden transitions, either due to instant change in the vector field or due to jumps in some of the state variables [132]. Depending upon how the switching signal evolves, whether it is time-dependent or state-dependent, the approach for analyzing such systems may be different. In this chapter, we are mainly interested in studying cascade and feedback interconnections of switched systems where the switching signal is either time-dependent or state-dependent. In the previous chapter, we saw that integral input-to-state stability (iISS) and input-to-state stability (ISS) are important properties for the study of such interconnections. Thus, most of the developments in this chapter are focused on establishing the iISS or ISS property for the two aforementioned classes of switched systems and the results for the interconnections are basically a consequence of these properties.

2.1 Time-dependent Switched Systems

We first start with the switched systems where the switching signals are time-dependent. In particular, to provide stability conditions, we put some restrictions on how fast or slow the switching can occur between certain subsystems. This allows us to identify the classes of switching signals for which the desired stability properties hold.

2.1.1 Models and stability notions

In this section, we introduce some basic elements such as the definition of switched systems and switching signals, and the definitions of ISS and iISS of switched systems.

Switched systems. Let $\mathsf{P} \subset \mathbb{N}$. For each $p \in \mathsf{P}$, there is a vector field $f_p : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, which is jointly locally Lipschitz in its arguments. The differential equations

$$\dot{x} = f_p(x, \omega), \quad p \in \mathsf{P} \tag{2.1}$$

are the dynamics of the *subsystems* or *modes* of the switched system with input $\omega : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$. For each pair $(p, q) \in \mathsf{P} \times \mathsf{P}$, there is also a continuous *jump map* $g_{p,q} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. It is assumed that for each $p \in \mathcal{P}$, $x \in \mathbb{R}^n$ and $\omega \in \mathbb{R}^m$, the set $G_p(x, \omega) := \{g_{p,q}(x, \omega) \,:\, q \in \mathcal{P}\} \subset \mathbb{R}^n$ is closed. Let $\overline{\Sigma}$ be the set of all right-continuous, piecewise constant mappings from $\mathbb{R}_{\geq 0}$ to P with a locally finite number of discontinuities, called *switching signals*. For each switching signal $\sigma \in \overline{\Sigma}$, define

$$\mathcal{T}(\sigma) := \{t > 0 : \sigma(t) \neq \sigma(t^-)\}. \tag{2.2}$$

In other words, $\mathcal{T}(\sigma)$ is the collection of switching instants. With this data, we consider switched dynamical systems with inputs and resets, described by

$$\dot{x}(t) = f_{\sigma(t)}(x(t), \omega(t)) \qquad \text{if } t \notin \mathcal{T}(\sigma), \tag{2.3a}$$

$$x(t) = g_{\sigma(t^-),\sigma(t)}(x(t^-), \omega(t)) \quad \text{if } t \in \mathcal{T}(\sigma), \tag{2.3b}$$

where $\omega : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ is locally essentially bounded on $\mathbb{R}_{\geq 0}$ and bounded on $\mathcal{T}(\sigma)$. We note that, over an interval $[t_0, t_1)$ where σ is constant, (2.3a) is seen as a time-varying differential equation with possible discontinuities due to inputs. The solution of (2.3a) is, therefore, interpreted in Carathéodory sense over such an interval $[t_0, t_1)$, see [91, Section I.5]. With $f_p(\cdot, \cdot)$ locally Lipschitz in both arguments, the conditions for existence of Carathéodory solutions hold [91, Theorem 5.1, Page 28]. We call (2.3a) the *continuous flow* and (2.3b) the *discrete jumps*. We denote the solution of (2.3) with initial state x_0, input ω and switching signal σ by $x(\cdot; x_0, \omega, \sigma)$. When x_0, ω, σ are clear from the context, the solution is also abbreviated by $x(\cdot)$.

Switching signals. We now specify the class of switching signals for which we study the stability of the system (2.3). We write the set P as a disjoint union $\mathsf{P} = \mathsf{P}_\mathrm{s} \cup \mathsf{P}_\mathrm{u}$, and introduce two different notions which constrain the class of switching signals. Conceptually, the class of switching signals with ADT constraints puts an upper bound on the average number of switches over any time interval. According to [99], we say that a switching signal σ satisfies *average dwell-time* (ADT) condition with parameters $\tau_a > 0$ (average dwell-time) and $N_0 \geq 1$ if

$$\forall t_2 \geq t_1 \geq 0 : \quad N_\sigma(t_1, t_2) \leq N_0 + \frac{t_2 - t_1}{\tau_a}, \tag{2.4}$$

where $N_\sigma(t_1, t_2)$ is the number of switches occurring in the interval $(t_1, t_2]$: $N_\sigma(t_1, t_2) := |(t_1, t_2] \cap \mathcal{T}(\sigma)|$. We denote by $\Sigma_{\mathrm{ADT}}(\tau_a, N_0)$ to be the collection of all ADT switching signals with parameters τ_a, N_0.

The second constraint provides a bound on the activation time for the modes contained in P_u. Following the definition in [155], we say that a switching signal σ satisfies *average activation time* (AAT) condition with parameters $\eta \in [0, 1]$ (percentage activation time), $T_0 \geq 0$, if

$$\forall t_2 \geq t_1 \geq 0 : \quad T_{\sigma,\mathsf{P}_\mathrm{u}}(t_1, t_2) \leq T_0 + \eta(t_2 - t_1), \tag{2.5}$$

where $T_{\sigma,\mathsf{P}_\mathrm{u}}(t_1, t_2)$ is the activation time of modes in P_u over the interval $(t_1, t_2]$. More precisely, if we consider the function $\mathbf{1}_{\mathsf{P}_\mathrm{u}} : \mathsf{P} \to \{0, 1\}$, defined as

$$\mathbf{1}_{\mathsf{P}_\mathrm{u}}(\sigma) = \begin{cases} 1, & \text{if } \sigma \in \mathsf{P}_\mathrm{u}, \\ 0, & \text{otherwise,} \end{cases} \tag{2.6}$$

then $T_{\sigma,\mathsf{P}_\mathrm{u}}(t_1, t_2) := \int_{t_1}^{t_2} \mathbf{1}_{\mathsf{P}_\mathrm{u}}(\sigma(\tau))d\tau$. We denote by $\Sigma_{\mathrm{AAT}}(\mathsf{P}_\mathrm{u}, \eta, T_0)$ to be the collection of all AAT switching signals of modes in P_u with parameters η, T_0. Notice that the AAT condition is always with respect to the set P_u, which we later associate with the collection of subsystems with not necessarily stable dynamics. In addition, in the special case when $\eta = 1$, (2.5) imposes no constraints on the activation time and hence the switching signal can be arbitrary. On the other hand, when $T_0 = 0$ and $\eta < 1$, (2.5) implies that modes in P_u are not activated at all.

Stability definitions. Let $\Sigma \subset \overline{\Sigma}$ be a set of switching signals. In this chapter, we will focus on the following two uniform external stability properties of switched systems with respect to all switching signals in Σ. In what follows, the essential supremum norm of ω over a set I (that is, the supremum of ω on I except for a set of measure zero) is denoted by $\operatorname{ess\,sup}_{s\in I} |\omega(s)|$.

Definition 2.1. Consider switched system (2.3) with $\omega \equiv 0$. The system is uniformly 0-GAS, or simply UGAS, over Σ if there exist $\beta \in \mathcal{KLL}$ such that

$$|x(t; x_0, 0, \sigma)| \leq \beta(|x_0|, t, N_\sigma(0, t)). \tag{2.7}$$

We say that the system is UGAS(η) over Σ, if there exist $\theta, \eta \in \mathcal{K}$ and $k > 0$ such that

$$|x(t; x_0, 0, \sigma)| \leq \eta(e^{-k(t+N_\sigma(0,t))}\theta(|x_0|_{\mathcal{A}})). \tag{2.8}$$

In Definition 2.1, the inequalities (2.7) and (2.8) are equivalent. This follows from [194, Proposition 7]. While the use of (2.7) is more common in the literature, we will use (2.8) to emphasize the role of the function η. A characterization similar to (2.8) has also been used in our recent work [64] to unify certain stability notions for switched systems and for proving a converse Lyapunov result.

Definition 2.2. A switched system (2.3) is uniformly *input-to-state stable* (ISS) over Σ if there exist $\beta \in \mathcal{KL}, \gamma \in \mathcal{K}$ such that

$$|x(t; x_0, \omega, \sigma)| \leq \beta(|x_0|, t, N_\sigma(0, t)) + \gamma(\operatorname*{ess\,sup}_{s\in[0,t]\setminus\mathcal{T}(\sigma)} |\omega(s)|) + \gamma(\sup_{s\in[0,t]\cap\mathcal{T}(\sigma)} |\omega(s)|) \tag{2.9}$$

for all $t \geq 0, x_0 \in \mathbb{R}^n$ and $\sigma \in \Sigma$.

Definition 2.3. A switched system (2.3) is uniformly *integral input-to-state stable* (iISS) over Σ if there exist $\alpha_0 \in \mathcal{K}_\infty, \beta \in \mathcal{KL}, \gamma \in \mathcal{K}$ such that

$$\alpha_0(|x(t; x_0, \omega, \sigma)|) \leq \beta(|x_0|, t, N_\sigma(0, t)) + \int_0^t \gamma(|\omega(s)|)ds + \sum_{s\in[0,t]\cap\mathcal{T}(\sigma)} \gamma(|\omega(s)|) \tag{2.10}$$

for all $t \geq 0, x_0 \in \mathbb{R}^n$, and $\sigma \in \Sigma$.

Our definitions of ISS and iISS are adopted from [32] and [158]. Notice that the last terms in (2.9), and in (2.10) do not appear in standard definitions of ISS and iISS for non-switched systems; these additional terms are needed in our framework to capture the growth of state trajectories at switching instants due to the presence of inputs in the jump dynamics (2.3b).

2.1.2 Stability of cascades

As the main application of the results to follow, let us now consider the cascade interconnection of two time-dependent switched systems as illustrated in Figure 2.1. We will state a result about the stability of cascades using certain stability properties for the individual subsystems, and in the following section, we provide sufficient conditions that allow us to establish such properties.

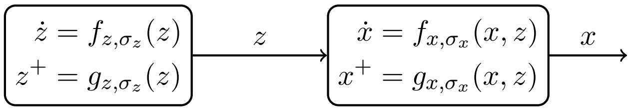


Figure 2.1: Two-stage serial cascade system with switching dynamics.

Let us describe the model of switched systems in cascade configuration. The driving system is an autonomous switched system with jumps at switching times, and is described as:

$$\dot z(t) = f_{\sigma_z(t)}(z(t)), \qquad \text{if } t \notin \mathcal{T}(\sigma_z), \tag{2.11a}$$

$$z(t) = g_{z,\sigma_z(t^-),\sigma_z(t)}(z(t^-)), \quad \text{if } t \in \mathcal{T}(\sigma_z), \tag{2.11b}$$

where $\sigma_z : \mathbb{R}_{\geq 0} \to \mathsf{P}_z$ is the switching signal belonging to a class Σ_z. For each $p, q \in \mathsf{P}_z$, the mappings $f_{z,p} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_z}$ and $g_{z,p,q} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_z}$ are assumed to be continuous and satisfy $f_{z,p}(0) = 0$ and $g_{z,p,q}(0) = 0$.

The driven system is also another switched impulsive system with a different switching signal, and is described as,

$$\dot x(t) = f_{\sigma_x(t)}(x(t), z(t)) \qquad \text{if } t \notin \mathcal{T}(\sigma), \tag{2.12a}$$

$$x(t) = g_{x,\sigma_x(t^-),\sigma_x(t)}(x(t^-), z(t)) \quad \text{if } t \in \mathcal{T}(\sigma), \tag{2.12b}$$

where $\sigma_x : \mathbb{R}_{\geq 0} \to \mathsf{P}_x$ is the switching signal belonging to a class Σ_x. For each $p, q \in \mathsf{P}_x$, the mappings $f_{x,p} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \to \mathbb{R}^{n_x}$ and $g_{x,p,q} : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \to \mathbb{R}^{n_x}$ are assumed to be continuous and satisfy $f_{x,p}(0,0) = 0$ and $g_{x,p,q}(0,0) = 0$.

We have the following main result:

Theorem 2.4. *Consider the cascade interconnection in* (2.11)–(2.12) *and suppose that*

- *the driving system* (2.11) *is UGAS*(η) *over* Σ_z *with respect to the origin in* $\mathbb{R}^{n_z}$*,*
- *the driven system* (2.11) *is uniformly iISS over* Σ_x *with respect to the origin in* $\mathbb{R}^{n_x}$ *and the gain* γ*.*

If it holds that

$$\int_0^1 \frac{(\gamma \circ \eta)(s)}{s} \mathrm{d}s < \infty, \tag{2.13}$$

then the interconnected switched system (2.11), (2.12) *is UGAS over the set* $\Sigma_x \times \Sigma_z$ *with respect to the set origin in* $\mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$*.*

Proof. The proof follows the template used in [10, Theorem 1], and we just need to show that the origin in $\mathbb{R}^{n_x}$ is asymptotically stable for (2.12) under the stated hypotheses. The primary difference here is that we have to handle an additional term due to state resets in the inequality that characterizes the iISS property. With the iISS assumption, we have the following bound on the trajectories of (2.12):

$$\alpha_0(|x(t;x_0)|) \leq \beta(|x_0|, t) + \int_0^t \gamma(|z(s)|) ds + \sum_{s \in [0,t] \cap \mathcal{T}(\sigma)} \gamma(|z(s)|) \tag{2.14}$$

for all $t \geq 0, x_0 \in \mathbb{R}^n$, and $\sigma_x \in \Sigma_x$. We first show that, under the condition (2.13), the last two terms in the foregoing inequality are bounded by a class $\mathcal{K}$ function of $z(0)$, which in turn implies stability of the origin in $\mathbb{R}^{n_x}$.

We note that, using (2.8),

$$\begin{aligned}\int_0^\infty \gamma(|z(s)|)ds &\leq \int_0^\infty (\gamma\circ\eta)(\mathrm{e}^{-k(s+N_{\sigma_z}(0,s))}\theta(|z_0|))\mathrm{d}s\\ &\leq \int_0^\infty (\gamma\circ\eta)(\mathrm{e}^{-ks}\theta(|z_0|))\mathrm{d}s\\ &= \frac{1}{k}\int_0^{\theta(|z_0|)} \frac{(\gamma\circ\eta)(\hat{s})}{\hat{s}}\mathrm{d}\hat{s}\end{aligned}$$

where $\hat{s} := \mathrm{e}^{-ks}\theta(|z_0|)$. Similarly, we note that

$$\begin{aligned}\sum_{s\in\mathbb{R}_{\geq 0}\cap\mathcal{T}(\sigma)} \gamma(|z(s)|) &\leq \sum_{s\in\mathbb{R}_{\geq 0}\cap\mathcal{T}(\sigma)} (\gamma\circ\eta)(\mathrm{e}^{-k(s+N_{\sigma_z}(0,s))}\theta(|z_0|))\\ &\leq (\gamma\circ\eta)(\theta(|z_0|)) + \sum_{s\in[t_1,\infty)\cap\mathcal{T}(\sigma)} (\gamma\circ\eta)(\mathrm{e}^{-kN_{\sigma_z}(0,s)}\theta(|z_0|))\\ &\leq (\gamma\circ\eta)(\theta(|z_0|)) + \int_0^\infty (\gamma\circ\eta)(\mathrm{e}^{-ks}\theta(|z_0|))\mathrm{d}s\\ &= (\gamma\circ\eta)(\theta(|z_0|)) + \frac{1}{k}\int_0^{\theta(|z_0|)} \frac{(\gamma\circ\eta)(\hat{s})}{\hat{s}}\mathrm{d}\hat{s}\end{aligned}$$

where we once agin choose $\hat{s} := \mathrm{e}^{-ks}\theta(|z_0|)$. Under condition (2.13), the function $\mu : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$,

$$\mathbb{R}_{\geq 0} \ni t' \mapsto \mu(t') = \int_0^{t'} \frac{(\gamma\circ\eta)(\hat{s})}{\hat{s}}\mathrm{d}\hat{s}$$

is a class $\mathcal{K}$ function because $\mu(0) = 0$ and the integrand is strictly positive. Thus,

$$\alpha_0(|x(t;x_0)|) \leq \beta(|x_0|,0,0) + \frac{2}{k}\mu(\theta(|z_0|)) + (\gamma\circ\eta)(\theta(|z_0|)).$$

We can make the right-hand side arbitrarily small by choosing $|z_0|$ and $|x_0|$ to be small, which ensures the stability (x,z)-dynamics. Similarly, the attractivity of $x(\cdot)$ can also be proven by putting all the bounds together. In particular, we can write a bound on $|x(t)|$ similar to (2.14), over the interval $[T,\infty)$ for T sufficiently large and use the fact that the boundedness of $\int_0^\infty \gamma(|z(s)|)\mathrm{d}s$ implies that the tail $\int_T^\infty \gamma(|z(s)|)\mathrm{d}s$ is arbitrarily small for large enough T. The reasoning given in [194, Proposition 6] can be used here to formalize this argument. □

Theorem 2.4 provides rather relaxed conditions based on iISS under which we can establish stability of cascade interconnections. We can also develop a Lyapunov function based approach under ISS assumption on driven system dynamics and this approach has been presented in [233]. The essential ingredients in studying these interconnections is to prove UGAS, or ISS/iISS, of a switched system for a class of switching signals. This is the topic of the first part of this chapter where our main result in Theorem 2.5 provides a unified approach to establish these properties. Before presenting that result, let us present an overview of the literature on this topic.

2.1.3 Literature overview

A canonical approach for stability analysis of switched systems is to develop characterizations based on the stability properties of the individual subsystems, and the effects observed at switching times. For a switched system to be GAS, ISS or iISS uniformly over *arbitrary switching signals*, not only all the subsystems are required to have these properties, but in addition all subsystems must share a common Lyapunov function. While the existence of a common Lyapunov function turns out to be the necessary and sufficient condition for GAS, ISS or iISS of switched systems under arbitrary switching [60, 89], such conditions are very restrictive in practice and there are well-known classes of switched systems which are GAS, ISS or iISS uniformly over some set of switching signals but not all switching signals [132]. For such systems, different approaches can be used to describe the sets of switching signals over which the desired stability property is ensured uniformly. The most straightforward method is to introduce slow switching. Intuitively, the actions of switching are treated as the source of instability and hence, as long as switching occurs sufficiently infrequently, the stability of the switched systems can be ensured. Well-known slow switching conditions include *dwell-time* (DT) condition and *average dwell-time* (ADT) condition. An autonomous switched system with a common asymptotically stable equilibrium for all subsystems is shown to be GAS, uniformly for switching signals with sufficiently large DT or ADT [153, 99]. For switched systems with inputs, ISS can also be studied similarly with respect to slow switching signals subjected to DT or ADT conditions [222].

Building on these results, there has been significant research to generalize the class of switching signals for which one can establish GAS, ISS, or iISS uniformly. These generalizations are developed by modifying the framework from aforementioned references in certain directions: a) by modifying the definition of ADT, b) by allowing some unstable subsystems and/or stabilizing discrete jump maps, and c) by relaxing the assumptions on the supply functions in the dissipation inequalities associated to each subsystem. For the first case, the notion of ADT is generalized by replacing the affine function of time in the definition of ADT by a more generic function in [125]. Similar generalization is also seen in the recent work [90]. Meanwhile, in [147], the authors propose impulse sequences whose impulse frequencies are eventually uniformly bounded, but iISS-related questions have not been investigated for such a class of signals. For the second case, one may consider unstable subsystems with stabilizing impulsive maps and frequent switching, so that the impulsive maps dominate the continuous flow and stabilize the system [227, 229]. Conditions describing how fast such stabilizing impulsive effects must occur are captured under the notion of *reverse average dwell-time* (rADT), introduced in [102]. Alternatively, since the activation of unstable subsystems is treated as the source of instability, stability of a switched system with unstable subsystems can be preserved by limiting the time that the system dwells in the unstable modes. Following this approach, *average activation time* (AAT) conditions are introduced, which are usually imposed together with ADT conditions in the characterization of switching signals, uniformly over which an impulsive switched system is shown GAS, ISS or iISS [155, 140]. The third research direction comes from relaxing the linear decay on the Lyapunov functions of individual subsystems to nonlinear decay functions. The papers [63, 137] provide counterexamples of switched systems with all stable subsystems where the dissipation inequalities have nonlinear supply functions and the resulting switched system is not asymptotically stable no matter how large the DT or ADT is chosen. Sufficient conditions for impulsive systems with nonlinear decay functions in dissipation inequalities associated to continuous and discrete dynamics appear in [181, 57], which have even been extended to infinite dimensional setup [59]. These were generalized to switched systems in [233] while using a Lyapunov function for the stability analysis, and the construction of this Lyapunov function builds on some earlier work done in [171]. Inspired by [233], the paper [140] allows for unstable systems and develops conditions for GAS, ISS, iISS which relate ADT and AAT in a rather elegant manner. In fact, it is possible to move beyond the framework of constant ADT/AAT which allows us to propose a class of switching signals that can handle subsystems with nonlinear decay functions and are not uniformly stable with constant ADT; such studies have appeared recently in [138]. Here, we also mention that the approach used in [233, 140, 138] is based on reformulating the problem in hybrid systems framework of [80] and using the ISS, or iISS, characterizations developed in [32, 33, 158].

As already mentioned, the ISS notion has been useful in analyzing interconnections of two dynamical systems, either in cascade form [195], or in feedback by using the small-gain condition [111, 112]. Moving away from the framework of conventional nonlinear systems, the ISS notion has been generalized to systems with continuous and discrete dynamics. More recently, we have seen ISS results for interconnections of hybrid systems [134, 182], and time-dependent switched systems [228, 151, 233].

2.1.4 Uniform stability under constrained switching

In line with the stability properties required for analyzing the cascade interconnection in Section 2.1.2, we now state a result that provides sufficient conditions for establishing such properties for a switched system. We start with some assumptions on the data of system (2.3) required for the statement of the main result.

Assumption 2.1. There exist $\mathcal{C}^1$ Lyapunov functions $V_p : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, $p \in \mathsf{P}$, satisfying the conditions:

(TD-L1) There exist $\underline{\alpha}, \overline{\alpha} \in \mathcal{K}_\infty$ such that

$$\underline{\alpha}(|x|) \leq V_p(x) \leq \overline{\alpha}(|x|), \quad \forall x \in \mathbb{R}^n, p \in \mathsf{P}. \tag{2.15}$$

(TD-L2) There exist a disjoint partition $\mathsf{P} = \mathsf{P}_\mathrm{s} \cup \mathsf{P}_\mathrm{u}$, two positive definite functions $\alpha_\mathrm{s}, \alpha_\mathrm{u}$, and $\gamma \in \mathcal{K}$ such that

$$\Big\langle \frac{\partial}{\partial x} V_p(x), f_p(x,\omega) \Big\rangle \leq -\alpha_\mathrm{s}(V_p(x)) + \gamma(|\omega|) \quad \forall x \in \mathbb{R}^n, \omega \in \mathbb{R}^m, p \in \mathsf{P}_\mathrm{s}, \tag{2.16a}$$

and

$$\Big\langle \frac{\partial}{\partial x} V_p(x), f_p(x,\omega) \Big\rangle \leq \alpha_\mathrm{u}(V_p(x)) + \gamma(|\omega|) \quad \forall x \in \mathbb{R}^n, \omega \in \mathbb{R}^m, p \in \mathsf{P}_\mathrm{u}. \tag{2.16b}$$

(TD-L3) There exist $\chi \in \mathcal{K}_\infty, \rho \in \mathcal{K}$ such that

$$V_q(g_{p,q}(x,\omega)) \leq \chi(V_p(x)) + \rho(|\omega|) \quad \forall x \in \mathbb{R}^n, \omega \in \mathbb{R}^m, p, q \in \mathsf{P}. \tag{2.17}$$

Remark 2.1. For each subsystem in P_s, **(TD-L1)** and the condition (2.16a) in **(TD-L2)** provide a necessary and sufficient condition for it to be iISS [6]. In addition if it is assumed that α_s in (2.16a) is $\mathcal{K}_\infty$, then each subsystem in P_s is ISS. △

Remark 2.2. The partition $\mathsf{P}_\mathrm{s}, \mathsf{P}_\mathrm{u}$ is not necessarily an indicator whether a subsystem is iISS or not. Indeed, when a subsystem is not iISS, (2.16a) will not hold for any choice of V_p but (2.16b) may still hold. However, even if a subsystem is iISS, it may still lead to the inequality (2.16b) due to the improper choice of V_p. In this case this subsystem needs to be categorized to the set P_u. △

Based on Assumption 2.1, we introduce the function $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as

$$\psi(t) := \min_{r \in [0,t]} \{\alpha_\mathrm{s}(r) + c(t-r)\} \tag{2.18}$$

where $c > 0$ is some constant. The next assumption bounds the divergence rate for the subsystems in P_u:

Assumption 2.2. The supremum

$$\kappa := \sup_{r>0} \frac{\alpha_{\mathrm{u}}(r)}{\psi(r)} \tag{2.19}$$

is finite, where α_{u} comes from **(TD-L2)** in Assumption 2.1 and ψ is defined in (2.18).

Our last assumption will be instrumental in deriving a lower bound for ADT in a similar fashion as in [233]:

Assumption 2.3. The supremum

$$\zeta^* := \sup_{s>0} \int_s^{\chi(s)} \frac{1}{\psi(r)} dr \tag{2.20}$$

is finite, where χ comes from **(TD-L3)** in Assumption 2.1 and ψ is defined in (2.18).

We are now ready to state the first main theorem of this chapter:

Theorem 2.5. *Consider the switched system* (2.3) *and suppose that Assumptions 2.1, 2.2, and 2.3 hold with the function ψ constructed via* (2.18) *for some $c > 0$. If $\tau_a > \zeta^*$, then for every $\eta \in [0, 1]$ satisfying the inequality*

$$(1 + \kappa)\eta + \frac{\zeta^*}{\tau_a} < 1, \tag{2.21}$$

the system (2.3) *is uniformly iISS over $\Sigma_{\mathrm{ADT}}(\tau_a, N_0) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$ for any $N_0 \geq 1$ and $T_0 \geq 0$. Moreover, if in* **(TD-L2)** *$\alpha_{\mathrm{s}} \in \mathcal{K}_\infty$, and $\tau_a > \zeta^*$, then system* (2.3) *is uniformly ISS over $\Sigma_{\mathrm{ADT}}(\tau_a, N_0) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$ for every $\eta \in [0, 1]$ satisfying* (2.21) *and any $N_0 \geq 1, T_0 \geq 0$.*

To provide some insight about Theorem 2.5, let us consider its application to some special types of switched systems. We first consider the case when all the subsystems are ISS/iISS. In this case, we can always find V_p satisfying (2.16a) in **(TD-L2)** for all $p \in \mathsf{P}$ so that $\mathsf{P}_{\mathrm{u}} = \emptyset$. Hence the AAT condition vanishes and we could simply set $\eta = 0$. As a result, (2.21) reduces to $\tau_a > \zeta^*$. In other words, in this case the switched system (2.3) is uniformly ISS/iISS over $\Sigma_{\mathrm{ADT}}(\tau_a, N_0)$ for any $\tau_a > \zeta^*, N_0 \geq 1$.

We next consider the case when the switching does not introduce jumps in the Lyapunov functions evaluated along the solutions of the unforced system, so that $\chi = \mathrm{id}$. We then conclude from (2.20) that $\zeta^* = 0$, so by Theorem 2.5, the switched system is ISS/iISS with arbitrarily small values of the average dwell-time τ_a. In addition, (2.21) results in $\eta < \frac{1}{1+\kappa}$ and the switched system (2.3) is uniformly ISS/iISS over $\left(\cup_{\tau_a>0} \Sigma_{\mathrm{ADT}}(\tau_a, N_0)\right) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$.

Lastly, we consider the case when the decaying rates and growth rates in **(TD-L2)**, **(TD-L3)** of Assumption 2.1 are all linear; that is, there exist $\lambda_{\mathrm{s}}, \lambda_{\mathrm{u}}, \mu > 0$ such that $\alpha_{\mathrm{s}}(r) = \lambda_{\mathrm{s}} r, \alpha_{\mathrm{u}}(r) = \lambda_{\mathrm{u}} r, \chi(r) = \mu r$ for all $r \geq 0$. In this case, the formula (2.20) yields

$$\zeta^* = \int_s^{\mu s} \frac{1}{\lambda_{\mathrm{s}} r} dr = \frac{\ln \mu}{\lambda_{\mathrm{s}}}$$

and consequently the inequality (2.21) becomes

$$\left(1 + \frac{\lambda_{\mathrm{u}}}{\lambda_{\mathrm{s}}}\right)\eta + \frac{\ln \mu}{\lambda_{\mathrm{s}} \tau_a} < 1.$$

After rearranging the terms, we recover the conditions on τ_a and η which guarantee ISS of switched system as stated in [155, Theorem 2]. Thus our work turns out to be a generalization from linear to nonlinear setting of the known results in the literature.

Discussion of the assumptions. Once we impose some characterizations of the stability or instability of individual subsystems and the reset maps in Assumption 2.1, it is seen that the stability condition in (2.21) primarily depends on finiteness of κ in Assumption 2.2 and ζ^* in Assumption 2.3. In both the assumptions, the function ψ constructed in (2.18) plays a key role.

1. Note that (2.18) immediately implies that $\psi(t) \leq \overline{\psi}(t) := \min\{\alpha_{\mathrm{s}}(t), ct\}$. In [139], we proposed stability conditions using the function $\overline{\psi}$ while assuming that it is globally one-sided Lipschitz. That assumption is necessary in [139]; for example, take $\alpha_{\mathrm{s}}(r) = re^{-r} + 1 - \cos(r^2)$, then $\overline{\psi}(t)$ is not one-sided Lipschitz for any $c > 0$. In this work, the Lipschitzness assumption is no longer required because the function ψ in (2.18) has this property by construction.

2. The function ψ is essentially a global one-sided Lipschitz modification of α_{s}. In case α_{s} is already globally Lipschitz, ψ can be chosen equal to α_{s} by setting c larger than its Lipschitz constant. Functions like $\alpha_{\mathrm{s}}(r) = r, \alpha_{\mathrm{s}}(r) = \ln(1 + r)$ and $\alpha_{\mathrm{s}}(r) = re^{-r}$ are examples of such a case.

3. Combining (2.18) and (2.19), we conclude from (2.16b) that for all $p \in \mathsf{P}_{\mathrm{u}}$,
$$\Big\langle \frac{\partial}{\partial x} V_p(x), f_p(x, \omega)\Big\rangle \leq \kappa c V_p(x) + \gamma(|\omega|).$$
According to [7, Theorem 1], this property in turn implies that these subsystems are forward complete, which is a very reasonable assumption on the subsystems as one cannot expect the switched system to be ISS/iISS if an active subsystem has finite escape time.

4. In order to get a qualitative answer to the question of when the switched system is ISS/iISS under slow switching, we need to better understand Assumption 2.3. The next lemma states that, in order to show that ζ^* is finite, we only need to examine the values of $\psi(s), \chi(s)$ for small and large s without computing the integral in (2.20). To this end, we say a positive definite function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is *initially increasing* if there exists $r > 0$ such that $\alpha(s) \leq \alpha(t)$ for all $0 \leq s \leq t \leq r$, and *eventually increasing* (resp. *eventually decreasing*) if there exists $R > 0$ such that $\alpha(s) \leq \alpha(t)$ (resp. $\alpha(s) \geq \alpha(t)$) for all $R \leq s \leq t$.

   Lemma 2.6. *The value ζ^* defined in* (2.20) *is finite when both conditions (a) and (b) hold:*

   (a) ψ is initially increasing and $\limsup_{s\to 0^+} \frac{\chi(s)-s}{\psi(s)} < \infty$*, and*

   (b) ψ is eventually increasing and $\limsup_{s\to\infty} \frac{\chi(s)-s}{\psi(s)} < \infty$*, or*
   ψ is eventually decreasing and $\limsup_{s\to\infty} \frac{\chi(s)-s}{\psi(\chi(s))} < \infty$.

   Proof. We only study the non-trivial case when $\chi(s) > s$ for all $s > 0$. By inspecting the definition of ζ^* in (2.20), we see that it suffices to show that $\limsup_s \int_s^{\chi(s)} \frac{1}{\psi(r)} dr$ is finite when s approaches both 0^+ and infinity in order for ζ^* to be finite. When ψ is initially increasing, we have $\overline{r} > 0$ such that for all $s > 0$ with $\chi(s) \leq \overline{r}$,
$$\int_s^{\chi(s)} \frac{1}{\psi(r)} dr \leq \int_s^{\chi(s)} \frac{1}{\psi(s)} dr = \frac{\chi(s) - s}{\psi(s)}.$$
   Take the limit as $s \to 0^+$, and it follows from the first condition in Lemma 2.6 that $\limsup_{s\to 0^+} \int_s^{\chi(s)} \frac{1}{\psi(r)} dr$ is bounded. Similarly, the second condition results in $\limsup_{s\to\infty} \int_s^{\chi(s)} \frac{1}{\psi(r)} dr$ being bounded. □

It is not hard to see that ζ^* is finite for the linear functions when $\psi(s) = \lambda s, \chi(s) = \mu s$. In addition, using Lemma 2.6, it is straightforward to verify that the nonlinear pairs such as $(\psi(s), \chi(s)) = (1 - e^{-s}, s + 1 - e^{-s})$, $(e^{-s}(1 - e^{-s}), \ln(e^s + 1))$ also ensure finite ζ^*.

2.1.5 Hybrid systems framework

Our proof of Theorem 2.5 relies on rewriting the switched system into a hybrid system in the framework of [80] and then showing ISS/iISS of that hybrid system. In this section we provide a brief description of hybrid systems, which also includes the formulation of ADT and AAT constrained switching signals in the framework of hybrid systems. In particular, we consider the hybrid dynamical system with inputs, which is described as:

$$\begin{cases} \dot{\xi} \in \mathcal{F}(\xi, d), & (\xi, d) \in \mathcal{C}, \\ \xi^+ \in \mathcal{G}(\xi, d), & (\xi, d) \in \mathcal{D}, \end{cases} \tag{2.22}$$

where the state trajectory ξ evolves in $\mathcal{X} \subset \mathbb{R}^n$, and the disturbance d takes values in $\mathbb{R}^m$. The sets $\mathcal{C}$, $\mathcal{D}$ are subsets of $\mathcal{X} \times \mathbb{R}^m$ and are called flow and jump sets, respectively. The evolution of the state is thus described by $\mathcal{F}$ (during flows) and by $\mathcal{G}$ (at jump instants), which are set-valued mappings from $\mathcal{X} \times \mathbb{R}^m$ to $\mathcal{X}$. The solutions of the hybrid system (2.22) are defined on *hybrid time domains*. A set $I \subset \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ is called a compact hybrid time domain if $I = \cup_{j=0}^{J-1}([t_j, t_{j+1}], j)$ for some finite sequence of times $0 = t_0 \leq t_1 \cdots \leq t_J$; and I is a hybrid time domain if for all $(T, J) \in I$, $I \cap ([0, T] \times \{0, 1, \ldots, J\})$ is a compact hybrid time domain. The domain of the hybrid input d in (2.22) is a hybrid time domain, and $d : \operatorname{dom} d \to \mathbb{R}^m$ is such that $d(\cdot, j)$ is Lebesgue measurable and locally essentially bounded for each $j \in \mathbb{Z}_{\geq 0}$. When the system data $(\mathcal{F}, \mathcal{G}, \mathcal{C}, \mathcal{D})$ satisfies certain basic assumptions [80, Assumption 6.5], [32, Page 48], it holds that, for each input $d : \operatorname{dom} d \to \mathbb{R}^m$, the system (2.22) admits a local solution, called *hybrid arc*, $\xi : \operatorname{dom} \xi \to \mathcal{X}$. A hybrid arc $\xi : \operatorname{dom} \xi \to \mathcal{X}$ and a hybrid input $d : \operatorname{dom} d \to \mathbb{R}^m$ are a solution pair to the hybrid model (2.22) if:

- $\operatorname{dom} \xi = \operatorname{dom} d$;
- for all $j \in Z_{\geq 0}$ and almost all $t \in \mathbb{R}_{\geq 0}$ such that $(t, j) \in \operatorname{dom} \xi$, we have $(\xi(t, j), d(t, j)) \in \mathcal{C}$ and $\dot{\xi}(t, j) \in \mathcal{F}(\xi(t, j), d(t, j))$; and
- for $(t, j) \in \operatorname{dom} \xi$ such that $(t, j + 1) \in \operatorname{dom} \xi$ we have $(\xi(t, j), d(t, j)) \in \mathcal{D}$ and $x(t, j + 1) \in \mathcal{G}(\xi(t, j), d(t, j))$.

For an initial state ξ_0 and an input d, a hybrid arc of (2.22) is denoted by $\xi(\cdot, \cdot; \xi_0, d)$. In what follows, we denote the distance between a vector $\xi \in \mathcal{X}$ and a compact set $\mathcal{A} \subset \mathcal{X}$ by $|\xi|_{\mathcal{A}}$, that is, $|\xi| := \inf_{\zeta \in \mathcal{A}} |\xi - \zeta|$. Following [32], for a hybrid signal d, we use the notation $\|d\|_{(t,j)}$ to denote the maximum between $\operatorname*{ess\,sup}_{(\hat{t},\hat{j}+1) \notin \operatorname{dom} d, \hat{t}+\hat{j} \leq t+j} |d(\hat{t}, \hat{j})|$ and $\sup_{(\hat{t},\hat{j}+1) \in \operatorname{dom} d, \hat{t}+\hat{j} \leq t+j} |d(\hat{t}, \hat{j})|$.

Definition 2.7. A hybrid system (2.22) is said to be *input-to-state stable* (ISS) with respect to $\mathcal{A}$ if there exist $\beta_h \in \mathcal{KLL}$ and $\gamma_h \in \mathcal{K}$ such that for all $\xi_0 \in \mathcal{X}$ and all $(t, j) \in \operatorname{dom} \xi$, each solution pair (ξ, d) satisfies

$$|\xi(t, j; \xi_0, d)|_{\mathcal{A}} \leq \beta_h(|\xi_0|_{\mathcal{A}}, t, j) + \gamma_h(\|d\|_{(t,j)}). \tag{2.23}$$

Definition 2.8. A hybrid system (2.22) is said to be *integral input-to-state stable* (iISS) with respect to $\mathcal{A}$ if there exist $\alpha_h \in \mathcal{K}_\infty, \beta_h \in \mathcal{KLL}$ and $\gamma_h \in \mathcal{K}$ such that for all $\xi_0 \in \mathcal{X}$ and all $(t,j) \in \operatorname{dom} \xi$, each solution pair (ξ, d) satisfies

$$\alpha_h(|\xi(t,j;\xi_0,d)|_{\mathcal{A}}) \leq \beta_h(|\xi_0|_{\mathcal{A}}, t, j) + \int_0^t \gamma_h(|d(s, j_{d,s})|)ds + \sum_{\substack{(\hat{t},\hat{j}+1)\in \operatorname{dom} d \\ \hat{t}+\hat{j}\leq t+j}} \gamma_h(|d(\hat{t},\hat{j})|) \tag{2.24}$$

where $j_{d,s} := \max\{i \in \mathbb{N}^* \,|\, (s,i) \in \operatorname{dom} d\}$.

We can also characterize ISS and iISS using Lyapunov functions. The following result provides sufficient conditions in terms of existence of Lyapunov functions with certain properties which guarantee ISS/iISS.

Lemma 2.9 ([32, 158])**.** *A hybrid system* (2.22) *is iISS with respect to* $\mathcal{A}$ *if there exists a smooth function* $V : \mathcal{X} \to \mathbb{R}_{\geq 0}$*, positive definite functions* α_c, α_d*, functions* $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ *and* $\gamma_c, \gamma_d \in \mathcal{K}$ *such that*

$$\alpha_1(|\xi|_{\mathcal{A}}) \leq V(\xi) \leq \alpha_2(|\xi|_{\mathcal{A}}) \tag{2.25}$$

for all $\xi \in \mathcal{X}$*,*

$$\left\langle \frac{\partial}{\partial \xi} V(\xi), \overline{f} \right\rangle \leq -\alpha_c(|\xi|_{\mathcal{A}}) + \gamma_c(|d|) \tag{2.26}$$

for all $(\xi, d) \in \mathcal{C}, \overline{f} \in \mathcal{F}(\xi, d)$*, and*

$$V(\overline{g}) - V(\xi) \leq -\alpha_d(|\xi|_{\mathcal{A}}) + \gamma_d(|d|) \tag{2.27}$$

for all $(\xi, d) \in \mathcal{D}, \overline{g} \in \mathcal{G}(\xi, d)$*. In addition if* $\alpha_c, \alpha_d \in \mathcal{K}_\infty$*, then the hybrid system* (2.22) *is ISS with respect to* $\mathcal{A}$*.*

We call a function V satisfying the inequalities in Lemma 2.9 with positive definite functions α_c, α_d (resp. $\alpha_c, \alpha_d \in \mathcal{K}_\infty$) a hybrid iISS (resp. ISS) Lyapunov function.

Remark 2.3. In papers [32, 158], the authors work with a single-valued version of (2.22), that is, $\mathcal{F}(\xi, d) = f(\xi, d)$ and $\mathcal{G}(\xi, d) = g(\xi, d)$. The statements in Lemma 2.9 are obvious extensions of those results. Additionally, we see a condition on system data in the results of [32, 158], which says that, for each $\xi \in \mathcal{X}$ and each $\epsilon \geq 0$, the set $\{f(\xi, d) \,|\, d \in \mathbb{R}^m \cap \epsilon\overline{\mathbb{B}}\}$ is convex. However, this condition is only required for proving the converse implication that ISS/iISS implies the existence of an appropriate Lyapunov function, and is not needed for our purposes.

One can write the switched system (2.3), along with the ADT and AAT constraints on the switching signal, more naturally in the form of (2.22). The reader may also consult [80, Example 3.22] and [167] for such reformulations also. For our purposes, we let

$$\xi := \begin{pmatrix} y \\ p \\ \tau \\ \tau_u \end{pmatrix} \in \mathbb{R}^n \times \mathsf{P} \times [0, N_0] \times [0, T_0] =: \mathcal{X}. \tag{2.28}$$

The hybrid model capturing the dynamics of the switched system, driven by an $\mathbb{R}^m$-valued disturbance d and a switching signal $\sigma \in \Sigma_{\mathrm{ADT}}(\tau_a, N_0) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$, is compactly written as

$$\dot{\xi} \in \mathcal{F}(\xi, d) : \begin{cases} \dot{y} = f_p(y, d), \\ \dot{p} = 0, \\ \dot{\tau} \in [0, \frac{1}{\tau_a}], \\ \dot{\tau}_u \in [\mathbf{1}_{\mathsf{P}_{\mathrm{u}}}(p) - \eta, \mathbf{1}_{\mathsf{P}_{\mathrm{u}}}(p)], \end{cases} \qquad (\xi, d) \in \mathcal{C}, \tag{2.29a}$$

$$\xi^+ \in \mathcal{G}(\xi, d) : \begin{cases} y^+ \in G_p(y,d), \\ p^+ \in \mathsf{P}, \\ \tau^+ = \tau - 1, \\ \tau_u^+ = \tau_u, \end{cases} \qquad (\xi, d) \in \mathcal{D}, \tag{2.29b}$$

where $\mathcal{C} := \mathcal{X} \times \mathbb{R}^m, \mathcal{D} := \mathbb{R}^n \times \mathsf{P} \times [1, N_0] \times [0, T_0] \times \mathbb{R}^m$, and $G_p(y,d) = \bigcup_{q \in \mathsf{P}} g_{p,q}(y,d)$. For the system (2.29), it is of interest to study whether the compact set

$$\mathcal{A} := \{0\}^n \times \mathsf{P} \times [0, N_0] \times [0, T_0] \tag{2.30}$$

is an attractor. We observe that system (2.29) satisfies basic assumptions listed in [80, Assumption 6.5]:

- The sets $\mathcal{C}$ and $\mathcal{D}$ are closed.
- The mapping $\mathcal{F}(\cdot,\cdot)$ is outer semi-continuous, convex-valued and locally bounded relative to $\mathcal{C}$, and $\mathcal{C} \subseteq \operatorname{dom}(\mathcal{F})$. Here, we used the fact that, for each $p \in \mathcal{P}$, the mapping $(y,d) \mapsto f_p(y,d)$ is locally Lipschitz continuous by assumption, and the set-valued dynamics for (p, τ, τ_u) are outer semi-continuous.
- The mapping $\mathcal{G}(\cdot,\cdot)$ is outer semi-continuous and locally bounded relative to $\mathcal{D}$. This is because, for each $p \in \mathsf{P}$, $y \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, we assumed that $G_p(y,d)$ is closed and hence the mapping $(y,d) \mapsto G_p(y,d)$ is outer semi-continuous. Because of the discrete topology of the set $\mathcal{P}$, it then follows that $(p,y,d) \mapsto G_p(y,d)$, as a subset of $\mathcal{P} \times \mathbb{R}^n \times \mathbb{R}^m$, is closed, and outer semi-continuity of $\mathcal{G}(\cdot,\cdot)$ follows by checking that its graph is closed.

Comparing (2.23) and (2.24) with (2.9) and (2.10) respectively, we have the following straightforward result, which shows that studying ISS (resp. iISS) for system (2.3) is equivalent to studying ISS (resp. iISS) for system (2.29).

Proposition 2.10. *The switched system* (2.3) *is uniformly ISS (resp. iISS) over the class of switching signals* $\Sigma_{\mathrm{ADT}}(\tau_a, N_0) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$ *if the hybrid system* (2.29)*, with augmented state variable defined in* (2.28)*, is ISS (resp. iISS) with respect to* $\mathcal{A}$*.*

2.1.6 Main elements of proof

Let us now have a look into the proof of Theorem 2.5. The main idea is to construct an (i)ISS-Lyapunov function for the overall switched system, and show that under the conditions listed in the theorem statement, we get the inequalities that characterize the (i)ISS property for the hybrid system. In the case of linear supply functions, no reset maps, and ADT constrained switching signals, this idea initially appeared in [99].

To construct the (i)ISS Lyapunov function for the hybrid system, we define $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as follows:

$$\varphi(s) := \begin{cases} \exp\left(\int_1^s \frac{2c}{\psi(r)} dr\right), & \text{if } s > 0, \\ 0, & \text{if } s = 0, \end{cases} \tag{2.31}$$

where we recall that c and ψ are introduced in (2.18).

Since $\tau_a > \zeta^*$ and (2.21) holds, we pick $\zeta \in (\zeta^*, \tau_a)$ which is sufficiently close to ζ^* so that

$$(\kappa + 1)\eta + \frac{\zeta}{\tau_a} < 1. \tag{2.32}$$

We prove the theorem by showing that

$$V(\xi) := \varphi^{-1}\Big(e^{2c\big(\zeta\tau+(\kappa+1)(T_0-\tau_u)\big)}\varphi\big(V_p(x)\big)\Big) \tag{2.33}$$

is an (i)ISS Lyapunov function for the hybrid system (2.29) corresponding to the switched system (2.3) with ADT and AAT constrained switching signals.

Some properties of functions ψ and φ. In analyzing the function (2.33), we need to have a closer look at the functions ψ constructed in (2.18) and φ in (2.31), and in particular, the following properties are of interest:

- The function ψ is positive definite. It is also of class $\mathcal{K}_\infty$ if α_s is of class $\mathcal{K}_\infty$.
- The function ψ constructed in (2.18) is a globally one-sided Lipschitz function on $\mathbb{R}_{\geq 0}$ with constant c; that is, for any $t_2 \geq t_1 \geq 0$, $\psi(t_2) - \psi(t_1) \leq c(t_2 - t_1)$.
- It holds that $\varphi \in \mathcal{K}_\infty$.
- Let $c_0 > 0$ and $s \geq t > 0$ such that $\varphi(s) \leq c_0\varphi(t)$. It follows that $\frac{\psi(s)}{\psi(t)} \leq \sqrt{c_0}$.
- The function $\theta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by $\theta(s) := \varphi^{-1}(c_0\varphi(s))$ with any $c_0 \geq 1$ is globally one-sided Lipschitz with constant $\sqrt{c_0}$.

We next show that, using the aforementioned properties, we can prove the desired inequalities for the function (2.33). Firstly, to verify the condition (2.25), we set

$$\alpha_1(s) = \underline{\alpha}(s), \quad \text{and} \quad \alpha_2(s) = \varphi^{-1}\Big(e^{2c\big(\zeta N_0+(\kappa+1)T_0\big)}\varphi(\overline{\alpha}(s))\Big)$$

and the bounds in (2.25) clearly hold.

We next verify the conditions (2.26) and (2.27). We recall from the proof of [6, Theorem 1] that the inequality (2.26) holds for some positive definite function α_c if there exists a positive definite function $\tilde{\alpha}_c$ such that

$$\Big\langle \frac{\partial}{\partial\xi}V(\xi), \overline{f}\Big\rangle \leq -\tilde{\alpha}_c(V(\xi)) + \gamma_c(|d|). \tag{2.34}$$

Moreover, $\alpha_c \in \mathcal{K}_\infty$ if $\tilde{\alpha}_c \in \mathcal{K}_\infty$, by setting $\alpha_c := \tilde{\alpha}_c \circ \alpha_1$. Similarly (2.27) holds for some positive definite function α_d if there exists a positive definite function $\tilde{\alpha}_d$ such that

$$V(\overline{g}) - V(\xi) \leq -\tilde{\alpha}_d(V(\xi)) + \gamma_d(|d|) \tag{2.35}$$

and $\alpha_d \in \mathcal{K}_\infty$ if $\tilde{\alpha}_d \in \mathcal{K}_\infty$. To verify the condition (2.34) (and therefore (2.26)) for the continuous flow, we define

$$W(\xi) := e^{2c\big(\zeta\tau+(\kappa+1)(T_0-\tau_u)\big)}\varphi(V_p(x)). \tag{2.36}$$

Picking $(\xi, d) \in \mathcal{C}$ and $\overline{f} \in \mathcal{F}(\xi, d)$, we have

$$\begin{aligned}\Big\langle \frac{\partial}{\partial\xi}W(\xi), \overline{f}\Big\rangle &= 2c\zeta W(\xi)\dot{\tau} - 2c(\kappa+1)W(\xi)\dot{\tau}_u \\ &\qquad + \Big\langle e^{2c(\zeta\tau+(\kappa+1)(T_0-\tau_u))}\varphi'(V_p(x))\frac{\partial}{\partial x}V_p(x), f_p(x,d)\Big\rangle \\ &= 2c\left(\zeta\dot{\tau} - (\kappa+1)\dot{\tau}_u + \frac{1}{\psi(V_p(x))}\Big\langle \frac{\partial}{\partial x}V_p(x), f_p(x,d)\Big\rangle\right)W(\xi).\end{aligned}$$

When $p \in \mathsf{P}_\mathrm{s}$, it follows from (2.16a) and (2.18) that $\left\langle \frac{\partial}{\partial x} V_p(x), f_p(x,d) \right\rangle \leq -\psi(V_p(x)) + \gamma(|d|)$. In addition $\dot{\tau} \in [0, \frac{1}{\tau_a}]$ and $\dot{\tau}_u \in [-\eta, 0]$, so

$$\left\langle \frac{\partial}{\partial \xi} W(\xi), \overline{f} \right\rangle \leq 2c \left(\frac{\zeta}{\tau_a} + (\kappa+1)\eta - 1 \right) W(\xi) + \frac{2cW(\xi)}{\psi(V_p(x))} \gamma(|d|).$$

When $p \in \mathsf{P}_\mathrm{u}$, it follows from (2.16b) and (2.19) that $\left\langle \frac{\partial}{\partial x} V_p(x), f_p \right\rangle \leq \kappa\psi(V_p(x)) + \gamma(|d|)$. In addition $\dot{\tau} \in [0, \frac{1}{\tau_a}]$ and $\dot{\tau}_u \in [1-\eta, 1]$, so

$$\left\langle \frac{\partial}{\partial \xi} W(\xi), \overline{f} \right\rangle \leq 2c \left(\frac{\zeta}{\tau_a} + (\kappa+1)(\eta-1) + \kappa \right) W(\xi) + \frac{2cW(\xi)}{\psi(V_p(x))} \gamma(|d|)$$

and we reach the same upper bound as in the first case. Appealing to (2.32), we conclude that

$$\left\langle \frac{\partial}{\partial \xi} W(\xi), \overline{f} \right\rangle \leq -2caW(\xi) + \frac{2cW(\xi)}{\psi(V_p(x))} \gamma(|d|)$$

where

$$a := 1 - \frac{\zeta}{\tau_a} - (\kappa+1)\eta > 0. \tag{2.37}$$

Consequently it can be computed that along the differential inclusion (2.29a),

$$\begin{aligned}
\left\langle \frac{\partial}{\partial \xi} V(\xi), \overline{f} \right\rangle &= (\varphi^{-1})'(W(\xi)) \left\langle \frac{\partial}{\partial \xi} W(\xi), \overline{f} \right\rangle \\
&\leq \left(\frac{\psi(\varphi^{-1}(W(\xi)))}{2cW(\xi)} \right) \left(-2caW(\xi) + \frac{2cW(\xi)}{\psi(V_p(x))} \gamma(|d|) \right) \\
&= -a\psi(\varphi^{-1}(W(\xi))) + \frac{\psi(\varphi^{-1}(W(\xi)))}{\psi(V_p(x))} \gamma(|d|) \\
&= -a\psi(V(\xi)) + \frac{\psi(V(\xi))}{\psi(V_p(x))} \gamma(|d|).
\end{aligned}$$

To further simplify the right-hand side, let $v := V_p(x)$, $w := V(\xi)$. Then we have $w \geq v \geq 0$ and $\varphi(w) = W(\xi) = e^{2c\left(\zeta\tau + (\kappa+1)(T_0 - \tau_u)\right)} \varphi(V_p(x)) \leq e^{2c\left(\zeta N_0 + (\kappa+1)T_0\right)} \varphi(v)$. Hence,

$$\left\langle \frac{\partial}{\partial \xi} V(\xi), \overline{f} \right\rangle \leq -a\psi(V(\xi)) + e^{c\left(\zeta N_0 + (\kappa+1)T_0\right)} \gamma(|d|)$$

and (2.34) holds with

$$\tilde{\alpha}_c(s) := a\psi(s) \tag{2.38a}$$

$$\gamma_c(s) := e^{c\left(\zeta N_0 + (\kappa+1)T_0\right)} \gamma(s). \tag{2.38b}$$

It now follows that if α_s is positive definite (resp. $\alpha_\mathrm{s} \in \mathcal{K}_\infty$), then $\tilde{\alpha}_c$ is also positive definite (resp. $\alpha_c \in \mathcal{K}_\infty$) and so is α_c. In addition, $\gamma_c \in \mathcal{K}$.

Lastly, to verify the condition (2.35) (and therefore (2.27)) for the discrete jumps, we pick $(\xi, d) \in \mathcal{D}$ and $\bar{g} \in \mathcal{G}(\xi, d)$. Notice that

$$\begin{aligned}
\varphi(\chi(V_p(x))) &= \exp\left(\int_1^{\chi(V_p(x))} \frac{2c}{\psi(r)} dr \right) \\
&= \exp\left(\int_{V_p(x)}^{\chi(V_p(x))} \frac{2c}{\psi(r)} dr \right) \exp\left(\int_1^{V_p(x)} \frac{2c}{\psi(r)} dr \right) \leq e^{2c\zeta^*} \varphi(V_p(x))
\end{aligned}$$

where we recall that ζ^* is defined in (2.20). Hence, if we define

$$\theta(s) := \varphi^{-1}\left(e^{2c(\zeta(\tau-1)+(\kappa+1)(T_0-\tau_u))}\varphi(s)\right),$$

and recall the definition of W from (2.36), we have the estimate that $\theta\big(\chi(V_p(x))\big) \leq \varphi^{-1}\left(e^{2c(\zeta(\tau-1)+(\kappa+1)(T_0-\tau_u))+2c\zeta^*}\varphi\big(V_p(x)\big)\right) = \varphi^{-1}\left(e^{2c(\zeta^*-\zeta)}W(\xi)\right)$. Therefore, using the assumption **(TD-L3)**, we get

$$\begin{aligned}
V(\overline{g}) &= \varphi^{-1}\left(e^{2c(\zeta\tau^+ +(\kappa+1)(T_0-\tau_u^+))}\varphi(V_{p^+}(x^+))\right)\\
&= \varphi^{-1}\left(e^{2c(\zeta(\tau-1)+(\kappa+1)(T_0-\tau_u))}\varphi\big(V_q(g_{p,q}(x,d))\big)\right)\\
&\leq \varphi^{-1}\left(e^{2c(\zeta(\tau-1)+(\kappa+1)(T_0-\tau_u))}\varphi\big(\chi(V_p(x))+\rho(|d|)\big)\right)\\
&= \theta\big(\chi(V_p(x))+\rho(|d|)\big)\\
&\leq \theta\big(\chi(V_p(x))\big)+e^{c(\zeta(\tau-1)+(\kappa+1)(T_0-\tau_u))}\rho(|d|)\\
&\leq \varphi^{-1}\left(e^{2c(\zeta^*-\zeta)}W(\xi)\right)+e^{c(\zeta(N_0-1)+(\kappa+1)T_0)}\rho(|d|)\\
&= \varphi^{-1}\left(e^{2c(\zeta^*-\zeta)}\varphi(V(\xi))\right)+e^{c(\zeta(N_0-1)+(\kappa+1)T_0)}\rho(|d|).
\end{aligned}$$

Thus the inequality (2.35) is seen to hold with

$$\tilde{\alpha}_d(s) := s-\varphi^{-1}\left(e^{2c(\zeta^*-\zeta)}\varphi(s)\right) \tag{2.39a}$$

$$\gamma_d(s) := e^{c(\zeta(N_0-1)+(\kappa+1)T_0)}\rho(s). \tag{2.39b}$$

Recall that $\zeta > \zeta^*$ and $\varphi \in \mathcal{K}_\infty$; thus by construction $\tilde{\alpha}_d$ is positive definite. In addition, $\gamma_d \in \mathcal{K}$. Lastly we show that $\tilde{\alpha}_d \in \mathcal{K}_\infty$ when $\alpha_{\mathrm{s}} \in \mathcal{K}_\infty$. It follows from (2.39a) that $\varphi\big((s-\tilde{\alpha}_d(s)\big) = e^{2c(\zeta^*-\zeta)}\varphi(s)$. Using the definition of φ in (2.31), and taking logarithm, it follows that

$$\int_1^{s-\tilde{\alpha}_d(s)} \frac{2c}{\psi(r)}dr = 2c(\zeta^*-\zeta)+\int_1^s \frac{2c}{\psi(r)}dr.$$

Hence, we have

$$\int_{s-\tilde{\alpha}_d(s)}^{s} \frac{1}{\psi(r)}dr = \zeta-\zeta^* > 0. \tag{2.40}$$

From the listed properties of function ψ, when $\alpha_{\mathrm{s}} \in \mathcal{K}_\infty$, $\psi \in \mathcal{K}_\infty$ as well and thus $\frac{1}{\psi(r)}$ decreases to 0 as r increases to infinity. Therefore, (2.40) holds for all $s > 0$ only when the length of the interval of integration grows to infinity as $s \to \infty$. In other words, we conclude that $\tilde{\alpha}_d \in \mathcal{K}_\infty$.

With all the three conditions (2.25), (2.26) and (2.27) verified, we appeal to Lemma 2.9 and conclude that the hybrid system (2.29) is ISS (resp. iISS) when $\alpha_s \in \mathcal{K}_\infty$ (resp. α_s is only positive definite). Then by Proposition 2.10, the switched system (2.3) is uniformly ISS (resp. uniformly iISS) over the set of signals $\Sigma_{\mathrm{ADT}}(\tau_a, N_0) \cap \Sigma_{\mathrm{AAT}}(\mathsf{P}_{\mathrm{u}}, \eta, T_0)$. This completes the proof of Theorem 2.5.

2.2 State-Dependent Switched Systems

Differently from the class of switched systems studied in Section 2.1, we now consider systems where the switching rule is defined by a function of state. This kind of switching rule is also seen in the design of stabilizing feedback laws, such as, uniting local and global controllers [172], or the design of switching rule using Lyapunov functions. Our end goal, once again, is to analyze interconnections of switched systems, but we will develop different tools in this section towards that end. In particular, we will use Lyapunov function based approach, but the peculiar element of our approach is that we will use a certain notion of derivative for nonsmooth Lyapunov functions which provides relaxed conditions for stability in state-dependent switched systems. We will develop an ISS result using this notion of derivative and later use it for feedback interconnections.

More specifically, given a family of vector fields $\{f_1, \dots, f_K\} \subset \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^n)$ and a switching signal $\sigma : \mathbb{R}^n \to \{1, \dots, K\}$, we now consider the system

$$\dot{x} = f_{\sigma(x)}(x, u). \tag{2.41}$$

For studying generalized solutions of such discontinuous systems, we extend the map $f_{\sigma(\cdot)}(\cdot, u)$, considering the Filippov regularization of (2.41) (see [75]). This leads to a differential inclusion of the form (see Section 2.2.3 for details)

$$\dot{x} \in F(x, u), \tag{2.42}$$

where F satisfies some regularity assumptions. Our first objective is to study asymptotic stability and robustness with respect to u for system (2.42). We then apply our results to the analysis of interconnected systems of the form

$$\begin{cases} \dot{x}_1 \in F_1(x_1, x_2, u), \\ \dot{x}_2 \in F_2(x_1, x_2, u). \end{cases} \tag{2.43}$$

In the previous section, we saw how multiple Lyapunov functions can be used to analyze the stability of time-dependent switched systems. This is essentially done by switching from one Lyapunov function to another when a switch occurs. When dealing with state-dependent switched systems, the *patching* of the Lyapunov functions may make the resulting common Lyapunov function non-differentiable, but locally Lipschitz in most cases. This element is seen in the analysis of asymptotic stability using piecewise differentiable functions [114, 16], and to some extent for establishing ISS in [96], [95].

The subsequent sections of this chapter are about developing sufficient conditions for ISS using locally Lipschitz Lyapunov functions for the class of differential inclusions in (2.42). The concept of set-valued derivatives for locally Lipschitz functions, introduced in [48], [14], is crucial to properly define the notion of *derivatives along the system's trajectories*. In particular, we focus on two different notions of set-valued derivatives that we call *Clarke* and *Lie derivatives*, each of them being a set-valued map from the state space $\mathbb{R}^n$ to the real numbers, see [41] and [53] for the formal definitions. For a large class of locally Lipschitz functions called *non-pathological functions* (as phrased in [217]), the notion of Lie derivative leads to less conservative stability conditions, see [41] and our recent papers [65] and [66]. As another example, the Lie derivative concept has been recently used in [116] to identify and remove infeasible directions of a differential inclusion of the form (2.42), and for stability analysis using an invariance principle for state-dependent switched systems [117], based on the ideas already introduced in [183].

2.2.1 Generalized derivatives

Our first aim is to prove ISS of a differential inclusion with inputs (2.42) via non-smooth Lyapunov functions, and thus in the following we collect various notions of generalized derivatives and gradients. Given a locally Lipschitz function $V : \mathbb{R}^n \to \mathbb{R}$ we have the following characterization of the Clarke generalized gradient [48, Theorem 2.5.1, page 63] which is taken here as a definition. Let $V : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function, Clarke generalized gradient of V at x is

$$\partial V(x) := \mathrm{co}\left\{\lim_{k\to\infty} \nabla V(x_k) \,|\, x_k \to x,\ x_k \notin \mathcal{N}_V\right\}, \tag{2.44}$$

where $\mathcal{N}_V \subseteq \mathbb{R}^n$ is the set where ∇V is not defined, which has zero Lebesgue measure by Rademacher's Theorem, and $\mathrm{co}(S)$ denotes the convex hull of a set $S \subseteq \mathbb{R}^n$. We now introduce two different notions of generalized directional derivatives for locally Lipschitz functions with respect to differential inclusion (2.42), which appeared firstly in [14].

Definition 2.11 (Set-valued directional derivatives [14, 53])**.** Consider the unperturbed differential inclusion (2.42) with $u \equiv 0$; given a locally Lipschitz continuous function $V : \mathbb{R}^n \to \mathbb{R}$, the *Clarke generalized derivative* of V with respect to F, denoted $\dot{V}_F(x)$, is defined as

$$\dot{V}_F(x) := \{\langle p, f\rangle \mid p \in \partial V(x),\ f \in F(x,0)\,\}.$$

Additionally, we define the *Lie generalized derivative* of V with respect to F, denoted $\dot{\overline{V}}_F$, as

$$\dot{\overline{V}}_F(x) := \{a \in \mathbb{R} \,|\, \exists f \in F(x,0) : \langle p, f\rangle = a,\ \forall p \in \partial V(x)\}.$$

These concepts can be extended to the case of a perturbed differential inclusion with input (2.42) as follows:

$$\begin{aligned} \dot{V}_F(x,u) &:= \{\langle p, f\rangle \mid p \in \partial V(x),\ f \in F(x,u)\,\}, \\ \dot{\overline{V}}_F(x,u) &:= \{a \in \mathbb{R} \,|\, \exists f \in F(x,u) : \langle p, f\rangle = a,\ \forall p \in \partial V(x)\}. \end{aligned} \tag{2.45}$$

For each $(x,u) \in \mathbb{R}^n \times \mathbb{R}^m$ the sets $\dot{V}_F(x,u)$ and $\dot{\overline{V}}_F(x,u)$ are closed and bounded intervals, with $\dot{\overline{V}}_F(x,u)$ possibly empty, see [41]. In particular

$$\dot{\overline{V}}_F(x,u) \subseteq \dot{V}_F(x,u). \tag{2.46}$$

Moreover, if V is continuously differentiable at x, one has $\partial V(x) = \{\nabla V(x)\}$ and thus

$$\dot{\overline{V}}_F(x,u) = \dot{V}_F(x,u) = \{\langle \nabla V(x), f\rangle \,|\, f \in F(x,u)\}.$$

2.2.2 Non-pathological functions

We now introduce a class of locally Lipschitz functions (and not necessarily $\mathcal{C}^1$), firstly introduced in [217].

Definition 2.12. [217] A locally Lipschitz function $V : \mathbb{R}^n \to \mathbb{R}$ is said to be *non-pathological* if, given any absolutely continuous function $\varphi \in AC(\mathbb{R}_+, \mathbb{R}^n)$, we have that for almost every $t \in \mathbb{R}_+$ there exists $a_t \in \mathbb{R}$ such that

$$\langle v, \dot{\varphi}(t)\rangle = a_t, \quad \text{for all } v \in \partial V(\varphi(t)).$$

In other words, $\partial V(\varphi(t))$ is a subset of an affine subspace orthogonal to $\dot{\varphi}(t)$, for almost every $t \in \mathbb{R}_+$. △

The usefulness of non-pathological functions is mainly given by the following result.

Proposition 2.13. *[217] If $V : \mathbb{R}^n \to \mathbb{R}$ is non-pathological and $\varphi : \mathbb{R}_+ \to \mathbb{R}^n$ is an absolutely continuous function, then the set*

$$\{\langle p, \dot{\varphi}(t)\rangle \mid p \in \partial V(\varphi(t))\},$$

is equal to the singleton $\{\frac{d}{dt}V(\varphi(t))\}$ for almost every $t \in \mathbb{R}_+$.

Remark 2.4. Given a non-pathological function $V : \mathbb{R}^n \to \mathbb{R}$, for any $u \in \mathcal{U}$, any initial condition $x_0 \in \mathbb{R}^n$ and any solution $x : \mathrm{dom}(x(\cdot)) \to \mathbb{R}^n$ of (2.42), we have that

$$\frac{d}{dt}V(x(t)) \in \dot{\overline{V}}_F(x(t), u(t)) \tag{2.47}$$

for almost every $t \in \mathrm{dom}(x(\cdot))$. In fact, by Proposition 2.13, for almost every $t \in \mathrm{dom}(x(\cdot))$, we have that

$$\begin{aligned}\left\{\frac{d}{dt}V(x(t))\right\} &= \{\langle p, \dot{x}(t)\rangle \mid p \in \partial V(x(t))\} \\ &\subseteq \{a \in \mathbb{R} \mid \exists f \in F(x(t), u(t)) : \langle p, f\rangle = a, \forall p \in \partial V(x(t))\} \\ &= \dot{\overline{V}}_F(x(t), u(t)).\end{aligned}$$

Non-pathological functions form a large class of functions which clearly includes $\mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$, we recall here some important properties of this family of functions, for the proofs we refer to [217] and [15].

Lemma 2.14. *The set of non-pathological functions is closed under addition, multiplication by scalars and pointwise maximum operator. More precisely, if $V_1, V_2 : \mathbb{R}^n \to \mathbb{R}$ are non-pathological then $V_1 + V_2$, $\max\{V_1(x), V_2(x)\}$ and λV ($\lambda \in \mathbb{R}$) are non-pathological.*

2.2.3 ISS with generalized derivatives

We now provide sufficient conditions for ISS of system (2.42). In what follows, due to the fact that the set $\dot{\overline{V}}_F(x, u)$ is possibly empty, we adopt the convention $\max\emptyset = -\infty$. The novelty of the following ISS-Lyapunov result lies in the fact that we require *Lie generalized derivative* of the Lyapunov function to be negative definite. Recalling the inclusion (2.47), this statement can be seen as a generalization of the existing results on ISS of differential inclusions relying on the notion of Clarke derivative, in particular [134].

Theorem 2.15. *Let $V : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz and non-pathological function such that there exist $\underline{\alpha}, \overline{\alpha} \in \mathcal{K}_\infty$, $\rho \in \mathcal{PD}$ and $\gamma \in \mathcal{K}$ such that*

$$\underline{\alpha}(|x|) \le V(x) \le \overline{\alpha}(|x|), \tag{2.48}$$

$$V(x) > \gamma(|u|) \;\;\Rightarrow\;\; \max \dot{\overline{V}}_F(x, u) \le -\rho(|x|), \tag{2.49}$$

then system (2.42) *is ISS w.r.t. u, and V is called a* non-pathological ISS-Lyapunov function *for system* (2.42).

Proof. Let us consider any initial condition $x_0 \in \mathbb{R}^n$ and any input $u \in \mathcal{U}$. Let $x : \mathrm{dom}(x(\cdot)) \to \mathbb{R}^n$ be a solution of system (2.42) starting at x_0 and with input u. Let us note that the function $V \circ x : \mathrm{dom}(x(\cdot)) \to \mathbb{R}$ is absolutely continuous because it is the composition of a locally Lipschitz continuous function and an absolutely continuous function. Then $\frac{d}{dt}V(x(t))$ exists almost everywhere. Since V is non-pathological, recalling (2.47) in Remark 2.4, we have that $\frac{d}{dt}V(x(t)) \in \dot{\overline{V}}_F(x(t), u(t))$ for almost every $t \in \mathrm{dom}(x(\cdot))$. From equation (2.49), we have that

$$\frac{d}{dt}V(x(t)) \leq -\rho(|x(t)|),$$

for almost every $t \in \mathrm{dom}(x(\cdot))$ such that $V(x(t)) > \gamma(|u(t)|)$. The proof is completed by following standard approaches as those in [196], [121, Theorem 4.18]. □

Remark 2.5. As already observed in the literature, for example in [32], under some assumptions the existence of $\rho \in \mathcal{PD}$ and $\gamma \in \mathcal{K}$ such that condition (2.49) holds is equivalent to the existence of two functions $\widehat{\rho} \in \mathcal{K}_\infty$ and $\widehat{\gamma} \in \mathcal{K}$ such that

$$\max \dot{\overline{V}}_F(x, u) \leq -\widehat{\rho}(|x|) + \widehat{\gamma}(|u|) \ \ \forall (x, u) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{2.50}$$

Indeed, implication (2.50) $\Rightarrow$ (2.49) holds by choosing $\rho := \frac{1}{2}\widehat{\rho}$ and $\gamma = \widehat{\rho}^{-1} \circ 2\widehat{\gamma}$. The converse implication (2.49) $\Rightarrow$ (2.50) holds if $F(0, 0) = \{0\}$ and $\rho \in \mathcal{K}_\infty$. Indeed, it suffices to choose $\widehat{\rho} := \rho$ and $\widehat{\gamma} \in \mathcal{K}$ such that $\widehat{\gamma}(r) \geq \max\{0, \widehat{\gamma}_0(r)\}$, for all $r \in \mathbb{R}_+$, where

$$\widehat{\gamma}_0(r) := \max\left\{\max \dot{\overline{V}}_F(x, u) + \rho(|x|) \,|\, |x| \leq \gamma(r), \ |u| \leq r\right\}.$$

It is possible to show that $\widehat{\gamma}_0$ above is well defined and $\widehat{\gamma}_0 \in \mathcal{K}$, using local boundedness of ∂V and regularity of $F : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$. Moreover, using (2.48), another equivalent formulation of condition (2.50) corresponds to asking that there exist two functions $\widetilde{\rho} \in \mathcal{K}_\infty$ and $\widetilde{\gamma} \in \mathcal{K}$ such that

$$\max \dot{\overline{V}}_F(x, u) \leq -\widetilde{\rho}(V(x)) + \widetilde{\gamma}(|u|), \ \ \forall (x, u) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{2.51}$$

The advantage of (2.51) is that in this formulation the function $\widetilde{\rho}$ represents the decay rate of V along the solutions. △

2.2.4 Embedding as differential inclusion

For the class of state-dependent switched systems that we consider, we introduce the concept of a *well-behaved* partition of the state space for the associated switched system.

Definition 2.16 (Proper State-Space Partition)**.** Given a finite set of indexes $\mathcal{I} := \{1, \dots K\}$, let us consider closed sets $X_1, \dots, X_K \subseteq \mathbb{R}^n$ and open sets $\mathcal{O}_1, \dots \mathcal{O}_K \subseteq \mathbb{R}^n$ such that

a) $\bigcup_{i=1}^K X_i = \mathbb{R}^n$,

b) $X_i \subseteq \mathcal{O}_i$, for all $i \in \mathcal{I}$,

c) $\overline{\mathrm{int}(X_i)} = X_i$, for all $i \in \mathcal{I}$,

d) For every $i \in \mathcal{I}$, $\mathrm{bd}(X_i)$ has zero Lebesgue measure,

e) $X_i \cap X_j = \mathrm{bd}(X_i) \cap \mathrm{bd}(X_j)$, for all $i, j \in \mathcal{I}$, $i \neq j$,

In this situation, we say that $\mathcal{X} := \{X_i, \mathcal{O}_i\}_{i\in\mathcal{I}}$ is a *proper partition* of $\mathbb{R}^n$. We define $\partial X := \cup_{i\in\mathcal{I}} \mathrm{bd}(X_i)$, and we underline that ∂X has zero Lebesgue measure. △

Given a proper partition $\mathcal{X}$ of $\mathbb{R}^n$, we can introduce an *index indicator map*, that is a set valued map $\mathcal{I}_\mathcal{X} : \mathbb{R}^n \rightrightarrows \mathcal{I}$ defined as

$$\mathcal{I}_\mathcal{X}(x) := \{i \in \mathcal{I} \mid x \in X_i\}. \tag{2.52}$$

We underline that $\mathcal{I}_\mathcal{X}$ is almost everywhere single valued. In fact, by Definition 2.16, item e), if $x \in \mathrm{int}(X_\ell)$ for some $\ell \in \mathcal{I}$ then $\mathcal{I}_\mathcal{X}(x) = \{\ell\}$.

Definition 2.17 (State-Dependent Switched System)**.** Given $\mathcal{X} = \{X_i, \mathcal{O}_i\}_{i\in\mathcal{I}}$ a proper partition of $\mathbb{R}^n$, consider $f_i \in \mathcal{C}^1(\mathcal{O}_i \times \mathbb{R}^m, \mathbb{R}^n)$, $i \in \mathcal{I}$. A *state-dependent switching signal* associated to $\mathcal{X}$ is a function $\sigma : \mathbb{R}^n \to \mathcal{I}$ such that

$$\sigma(x) = i, \quad \text{if } x \in \mathrm{int}(X_i), \tag{2.53}$$

and the (perturbed) *state-dependent switched system* associated to $\{X_i, \mathcal{O}_i, f_i\}_{i\in\mathcal{I}}$ is the differential equation

$$\dot{x} = f_{\sigma(x)}(x, u). \tag{2.54}$$

We note here that, given a proper partition $\mathcal{X} = \{X_i, \mathcal{O}_i\}$, a state dependent switching signal associated to it is not uniquely defined: the value of σ remains unspecified on the null-measure set $\partial\mathcal{X}$. We now clarify why this ambiguity does not affect the solution set of the corresponding state-dependent switched system.

System (2.54) has a discontinuous right-hand side in the first argument thus it may not have any Carathéodory solutions at the discontinuity points of $f_{\sigma(\cdot)}(\cdot, u)$, see [53]. Many possible definitions of "generalized solutions" for discontinuous dynamical system are possible (see for example [41] or [53]); we consider the concept of *Filippov solutions*, introduced firstly in [75]. More formally, we define Fil, the Filippov regularization of the discontinuous map f_σ, as

$$\begin{aligned}\mathrm{Fil}(f_\sigma)(x,u) &:= \bigcap_{\delta>0}\bigcap_{\mu(S)=0} \overline{\mathrm{co}}\{f_{\sigma(y)}(y,u) \mid y \in \mathbb{B}(x,\delta) \setminus S\}\\ &=: F^{\mathrm{sw}}(x,u),\end{aligned}$$

where μ denotes the Lebesgue measure. Under the hypotheses in Definitions 2.16 and 2.17, it can be proven that

$$F^{\mathrm{sw}}(x,u) = \mathrm{co}\{f_i(x,u) \mid i \in \mathcal{I}_{\mathcal{X}}(x)\},$$

see for example [75], [53] and [95]. Summarizing, we consider the regularized differential inclusion

$$\dot{x} \in F^{\mathrm{sw}}(x,u) = \mathrm{co}\{f_i(x,u) \mid i \in \mathcal{I}_{\mathcal{X}}(x)\} \tag{2.55}$$

considering again signals $u \in \mathcal{U}$.

Since it is easily verified that $f_{\sigma(x)}(x,u) \in F^{\mathrm{sw}}(x,u)$ for all $(x,u) \in \mathbb{R}^n \times \mathbb{R}^m$, we have that any solution of (2.54) is also a solution of (2.55). Moreover, it can be seen that $F^{\mathrm{sw}} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has non empty, compact and convex values, the map $x \mapsto F^{\mathrm{sw}}(x,u)$ is upper semicontinuous for any $u \in \mathbb{R}^m$ and the map $u \mapsto F^{\mathrm{sw}}(x,u)$ is continuous for any $x \in \mathbb{R}^n$, and thus we have existence of solutions of (2.55) from *any* initial condition and any input $u \in \mathcal{U}$ (which was not the case for (2.54), see [53]). Then, characterizing desirable (stability/convergence) properties of the solutions of (2.55) indirectly also characterizes the solutions of (2.54). In particular, a *Filippov solution* of system (2.54) is by definition a solution of the differential inclusion (2.55).

2.2.5 ISS for state-dependent switched system

In this section, we apply our ISS result to a specific differential inclusion with inputs arising from a suitable regularization of state-dependent switched systems. We introduce here a family of locally Lipschitz functions that we propose as candidate Lyapunov functions for system (2.55).

Definition 2.18 (Piecewise $\mathcal{C}^1$ Functions)**.** Consider $\mathcal{J} = \{1, \dots, N\}$ and $\mathcal{Y} = \{Y_j, \mathcal{V}_j\}_{j\in\mathcal{J}}$, a proper partition of $\mathbb{R}^n$. A continuous function $V : \mathbb{R}^n \to \mathbb{R}$ is called a *piecewise* $\mathcal{C}^1$ *function* with respect to the proper partition $\mathcal{Y}$ (and we write $V \in \mathscr{P}(\mathcal{Y})$) if there exist real-valued functions $V_1, \dots, V_N$ such that

1. $V_j \in \mathcal{C}^1(\mathcal{V}_j, \mathbb{R})$ for each $j \in \mathcal{J}$,
2. $V(x) = V_j(x)$, if $x \in Y_j$. △

Piecewise $\mathcal{C}^1$ functions with respect to a proper partition are a particular kind of *"piecewise* $\mathcal{C}^1$ *functions"* defined in [190].

Proposition 2.19. *Consider $V \in \mathscr{P}(\mathcal{Y})$, with respect to a proper partition $\mathcal{Y} = \{Y_j, \mathcal{V}_j\}_{j\in\mathcal{J}}$, in the sense of Definition 2.18. Then the following hold:*

1. V is locally Lipschitz, non-pathological and

$$\partial V(x) = \operatorname{co}\{\nabla V_j(x) \mid j \in \mathcal{I}_{\mathcal{Y}}(x)\}. \tag{2.56}$$

2. Given $F^{\mathrm{sw}} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined as in (2.55)*, we have*

$$\dot{\overline{V}}_{F^{\mathrm{sw}}}(x,u) = \Big\{a \in \mathbb{R} \,\Big|\, \exists f \in F^{\mathrm{sw}}(x,u) : \langle \nabla V_j(x), f\rangle = a,\, \forall j \in \mathcal{I}_{\mathcal{Y}}(x)\Big\}. \tag{2.57}$$

The reader may look at the proof this proposition in our paper [67]. We can now specialize the ISS results in this setting. First, we consider a candidate Lyapunov function in the class of *piecewise* $\mathcal{C}^1$ *functions* $\mathcal{P}(\mathcal{Y})$, where the proper partition $\mathcal{Y}$ does not necessarily coincide with $\mathcal{X}$, the proper partition associated with the considered switched system. Then we present specifically the case $\mathcal{X} = \mathcal{Y}$, see the subsequent Remark 2.6 for further discussion.

Corollary 2.20 (ISS for state-dependent switching)**.** *Consider a proper partition $\mathcal{X} = \{X_i, \mathcal{O}_i\}_{i\in\mathcal{I}}$ and an associated switched system* (2.55)*. Let us consider another proper partition $\mathcal{Y} = \{Y_j, \mathcal{V}_j\}_{j\in\mathcal{J}}$ and let $V \in \mathscr{P}(\mathcal{Y})$. Suppose that there exist $\underline{\alpha}, \overline{\alpha} \in \mathcal{K}_\infty$, $\rho \in \mathcal{PD}$ and $\gamma \in \mathcal{K}$ such that:*

(SD-L1a) *for each $j \in \mathcal{J}$, for each $x \in Y_j$,*

$$\underline{\alpha}(|x|) \le V_j(x) \le \overline{\alpha}(|x|);$$

(SD-L2a) *for each $j \in \mathcal{J}$, for each $x \in \operatorname{int}(Y_j) \setminus \partial X$ and for each $u \in \mathbb{R}^m$,*

$$V(x) > \gamma(|u|) \ \ \Rightarrow \ \ \langle \nabla V_j(x), f_{\sigma(x)}(x,u)\rangle \le -\rho(|x|);$$

(SD-L3a) *for each $(x,u) \in \partial X \times \mathbb{R}^m$,*

$$V(x) > \gamma(|u|) \ \ \Rightarrow \ \ \max \dot{\overline{V}}_{F^{\mathrm{sw}}}(x,u) \le -\rho(|x|);$$

then system (2.55) *is ISS.*

Proof. By the non-pathological property of V established in Proposition 2.19, we can apply Theorem 2.15. Getting inequality (2.48) from **(SD-L1a)** is straightforward. We check inequality (2.49) decomposing $\mathbb{R}^n$ as follows:

$$\mathbb{R}^n = \Big(\bigcup_{j\in\mathcal{J}} \operatorname{int}(Y_j) \setminus \partial X\Big) \cup \partial X \cup (\partial Y \setminus \partial X).$$

Consider first a point $x \in \mathrm{int}(Y_j) \setminus \partial X$ for some $j \in \mathcal{J}$. Function $V \in \mathscr{P}(\mathcal{Y})$ is $\mathcal{C}^1$ at x, and $F^{\mathrm{sw}}(x,u) = \{f_{\sigma(x)}(x,u)\}$ is single-valued. Thus, for any $u \in \mathbb{R}^m$, we have

$$\dot{V}_{F^{\mathrm{sw}}}(x,u) = \{\langle \nabla V_j(x), f_{\sigma(x)}(x,u)\rangle\},$$

and by **(SD-L2a)**, the implication in (2.49) holds.
If $x \in \partial X$, the assertion follows directly by **(SD-L3a)**.
As the last step, consider a point $x \in \partial Y \setminus \partial X$, and thus $x \in \mathrm{int}(X_i)$ for some $i \in \mathcal{I}$. In particular at x, we have $F^{\mathrm{sw}}(x,u) = \{f_i(x,u)\}$. Consider $u \in \mathbb{R}^m$ and suppose that $V(x) > \gamma(|u|)$. Recalling (2.56) in Lemma 2.19 and by Definition 2.16, for each $j \in \mathcal{I}_{\mathcal{Y}}(x)$, there exists a sequence $x_k^j \to x$ such that $x_k^j \in \mathrm{int}(Y_j) \cap \mathrm{int}(X_i)$ for all $k \in \mathbb{N}^*$. By continuity of V and γ, we can suppose $V(x_k^j) > \gamma(|u|)$. At these points, from **(SD-L2a)**, we get $\langle \nabla V_j(x_k^j), f_i(x_k^j,u)\rangle \le -\rho(|x_j^k|)$. By continuity of ∇V_j, f_i and ρ we have

$$\langle \nabla V_j(x), f_i(x,u)\rangle \le -\rho(|x|),$$

for each $j \in \mathcal{I}_{\mathcal{Y}}(x)$. We have proved (2.49) for all $x \in \mathbb{R}^n$, and this concludes the proof. □

We underline that we are not explicitly checking (2.49) on the zero Lebesgue measure set $\partial Y \setminus \partial X$: this is possible thanks to the continuity of F^{sw} when restricted to $\mathrm{int}(X_i)$ for some $i \in \mathcal{I}$.

As a special case of Corollary 2.20, we present the situation $\mathcal{X} = \mathcal{Y}$.

Corollary 2.21. *Consider a proper partition $\mathcal{X} = \{X_i, \mathcal{O}_i\}_{i\in\mathcal{I}}$ and the associated switched system* (2.55)*. Consider $V \in \mathscr{P}(\mathcal{X})$ such that there exist $\underline{\alpha}, \overline{\alpha} \in \mathcal{K}_\infty$, $\rho \in \mathcal{PD}$ and $\gamma \in \mathcal{K}$ satisfying*

(SD-L1b) *for each $i \in \mathcal{I}$, for each $x \in X_i$*

$$\underline{\alpha}(|x|) \le V_i(x) \le \overline{\alpha}(|x|);$$

(SD-L2b) *for each $i \in \mathcal{I}$, for each $(x,u) \in \mathrm{int}(X_i) \times \mathbb{R}^m$,*

$$V(x) > \gamma(|u|) \;\;\Rightarrow\;\; \langle \nabla V_i(x), f_i(x,u)\rangle \le -\rho(|x|);$$

(SD-L3b) *for each $(x,u) \in \partial X \times \mathbb{R}^m$,*

$$V(x) > \gamma(|u|) \;\;\Rightarrow\;\; \max \dot{V}_{F^{\mathrm{sw}}}(x,u) \le -\rho(|x|);$$

then system (2.55) *is ISS.*

Remark 2.6 (Comparison between Corollary 2.20 and Corollary 2.21)*.* To check the conditions of Corollary 2.21 for each $i \in \mathcal{I}$, we must find a smooth Lyapunov function V_i for the system f_i on the set $\mathcal{O}_i \supset X_i$. Then we must construct a continuous function V by *gluing together* the V_i's on ∂X and finally check the Lie-based condition **(SD-L3b)** on the switching surface ∂X. On the other hand, in some situations, it can be difficult to construct a smooth Lyapunov function even for a single subsystem in its region of activation if the subsystem is unstable. For this reason, in Corollary 2.20 we allow the candidate Lyapunov functions to be possibly non-differentiable in the interior of the X_i, and we do not need to check the conditions at the point of non-differentiability of V, since $F(\cdot,u)$ is continuous in a neighborhood of $\partial Y \setminus \partial X$. △

In the following example, mainly inspired by [114, Example 1], we apply the result presented in Corollary 2.21 to a particular state-dependent switched system.

Example 2.1. Consider the proper state-space partition $\mathcal{X} = \{X_i, \mathbb{R}^2\}_{i\in\{1,2\}}$ of $\mathbb{R}^2$ defined by

$$X_i := \{x \in \mathbb{R}^2 \mid x^\top Q_i x \geq 0\},$$

with $Q_1 = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, and $Q_2 = -Q_1$. Define the switching system

$$\dot{x} = \begin{cases} A_1 x + Bu, & \text{if } x \in X_1, \\ A_2 x + Bu, & \text{if } x \in X_2, \end{cases}$$

with $A_1 = \left[\begin{smallmatrix} -0.1 & 1 \\ -5 & -0.1 \end{smallmatrix}\right]$, $A_2 = \left[\begin{smallmatrix} -0.1 & 5 \\ -1 & -0.1 \end{smallmatrix}\right]$ and $B \in \mathbb{R}^{2\times m}$ arbitrary. We want to study the arising Filippov regularization defined as in (2.55); for a graphical representation of the unperturbed case $u \equiv 0$, see [65, Figure 1]. Consider the function $V : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$V(x) = \begin{cases} V_1(x) = x^\top P_1 x & \text{if } x \in X_1, \\ V_2(x) = x^\top P_2 x & \text{if } x \in X_2, \end{cases}$$

with $P_1 = \left[\begin{smallmatrix} 5 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $P_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 5 \end{smallmatrix}\right]$. We now show that V is an ISS Lyapunov function, in the sense of Corollary 2.21. Since $P_2 - P_1 = 4Q$, V is continuous (and thus $V \in \mathscr{P}(\mathcal{X})$) and moreover it satisfies item **(SD-L1b)** of Corollary 2.21 with $\underline{\alpha}(s) = s^2$ and $\overline{\alpha}(s) = 5s^2$. Moreover it can be seen that

$$A_i^\top P_i + P_i A_i + 0.1I \prec 0, \quad \text{for any } i \in \{1, 2\};$$

thus, following the reasoning of [121, Lemma 4.6], it can be proven that, for any $0 < \varepsilon < 0.1$ we have

$$|x| \geq b|u|, \Rightarrow x^\top (A_i^\top P_i + P_i A_i)x + 2x^\top P_i Bu \leq -\varepsilon |x|^2, \tag{2.58}$$

with $b > 0$ depending on A_1, A_2, B, P_1, P_2 and $0 < \varepsilon < 0.1$.
It is easy to check $\partial X = \{x \in \mathbb{R}^2 \mid x^\top Q_1 x = 0\} = \mathcal{R}_1 \cup \mathcal{R}_2$ where $\mathcal{R}_1 = \{[x_1\ x_2]^\top \in \mathbb{R}^2 \mid x_1 = x_2\}$ and $\mathcal{R}_2 = \{[x_1\ x_2]^\top \in \mathbb{R}^2 \mid x_1 = -x_2\}$. Consider $z \in \partial X$, and firstly, let us suppose that $z \in \mathcal{R}_1$, and define $\mu := \frac{|z|}{\sqrt{2}}$; we can write $z = \mu[1\ 1]^\top$ (w.l.o.g. supposing that z lies in the first quadrant). By Proposition 2.19 and (2.55), we have that

$$\partial V(z) = 2\mu \operatorname{co}\{v_1, v_2\} \quad \text{and} \quad F(z, u) = \mu \operatorname{co}\{f_1, f_2\} + Bu,$$

where $v_1 = [5\ 1]^\top$, $v_2 = [1\ 5]^\top$ and $f_i = A_i[1\ 1]^\top$ for $i \in \{1, 2\}$. Recalling Definition 2.11 we have that $\dot{\overline{V}}_{F^{\mathrm{sw}}}(z, u) \neq \emptyset$ if and only if there exists $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$, such that

$$[\mu(\lambda f_1 + (1-\lambda) f_2) + Bu]^\top v_1 = [\mu(\lambda f_1 + (1-\lambda) f_2) + Bu]^\top v_2.$$

Inserting the numerical values and simplifying, we obtain

$$6\mu = [-1\ \ 1]^\top Bu,$$

(note that the dependence on λ cancels out). Now, it is clear that if $6\mu \geq \sqrt{2}\|B\||u|$, this equation has no solution[1]. Recalling the definition of μ we have proved that

$$z \in \mathcal{R}_1 \text{ and } |z| \geq \frac{1}{3}\|B\||u| \quad \Rightarrow \quad \dot{\overline{V}}_{F^{\mathrm{sw}}}(z, u) = \emptyset. \tag{2.59}$$

[1]Here and in what follows we consider the 2-induced matrix norm, i.e., given $M \in \mathbb{R}^{n\times m}$ we define

$$\|M\| := \sup_{|x|=1} |Mx|$$

Since $A_i = R^\top A_j R$ and $P_i = R^\top P_j R$ with $R = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ for any $i, j \in \{1, 2\}$, $i \neq j$, it can be seen that the same holds on the line $\mathcal{R}_2$. Fixing $\varepsilon < 0.1$ in (2.58) and defining $\gamma(s) := 5 \max\{bs, (\|B\|/3)\, s\}$ and $\rho(s) := \varepsilon s^2$, we have proved Items **(SD-L2b)** and **(SD-L3b)** of Corollary 2.21, establishing that V is a (non-pathological) ISS Lyapunov function and thus the system is ISS. Note that the same reasoning holds when replacing the input term "Bu" with a continuous function $g : \mathbb{R}^2 \times \mathbb{R}^m \to \mathbb{R}^2$ such that $|g(x, u)| \leq L|u|$ for some $L > 0$, for every $(x, u) \in \mathbb{R}^2 \times \mathbb{R}^m$. △

2.2.6 Interconnected differential inclusions

In this section, we use Theorem 2.15 to study stability of feedback interconnection of two systems modeled by differential inclusions based on our work in [67]. For the system shown in Figure 2.2, we establish ISS of the interconnected system by constructing a Lyapunov function from two (nonsmooth) ISS-Lyapunov functions associated with the two subsystems.

Consider $F_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \rightrightarrows \mathbb{R}^{n_1}$ and $F_2 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^m \rightrightarrows \mathbb{R}^{n_2}$, and suppose that they have non empty, compact and convex values and are upper semicontinuous in the first two arguments and continuous in the third one. Consider the interconnection

$$\dot{x}_1 \in F_1(x_1, x_2, u), \tag{2.60a}$$

$$\dot{x}_2 \in F_2(x_1, x_2, u). \tag{2.60b}$$

We introduce the notation $x = (x_1, x_2) := (x_1^\top, x_2^\top)^\top \in \mathbb{R}^n = \mathbb{R}^{n_1+n_2}$ and the augmented differential inclusion

$$\dot{x} \in F(x, u) := \begin{pmatrix} F_1(x_1, x_2, u) \\ F_2(x_1, x_2, u) \end{pmatrix}. \tag{2.61}$$

We start our construction by assuming the existence of ISS-Lyapunov functions for the two subsystems, in order to conclude ISS of the overall interconnection (2.61). These types of assumptions also appear in the classical ISS-based approaches [111, 112], and have also been used in the hybrid setting [95], [134]. The novelty introduced here lies in the fact that we consider *non-pathological* ISS Lyapunov functions satisfying the "relaxed" conditions involving the Lie derivative presented in Theorem 2.15, as formalized in the following statement.

Assumption 2.4. Suppose that there exist non-pathological functions $V_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ and $V_2 : \mathbb{R}^{n_2} \to \mathbb{R}$ such that

1a) There exist $\underline{\alpha}_1, \overline{\alpha}_1 \in \mathcal{K}_\infty$ satisfying

$$\underline{\alpha}_1(|x_1|) \leq V_1(x_1) \leq \overline{\alpha}_1(|x_1|), \quad \forall\, x_1 \in \mathbb{R}^{n_1}.$$

1b) There exist $\underline{\alpha}_2, \overline{\alpha}_2 \in \mathcal{K}_\infty$ satisfying

$$\underline{\alpha}_2(|x_2|) \leq V_2(x_2) \leq \overline{\alpha}_2(|x_2|), \quad \forall\, x_2 \in \mathbb{R}^{n_2}.$$

2a) There exist $\rho_1 \in \mathcal{PD}$, and $\chi_1, \gamma_1 \in \mathcal{K}$ satisfying

$$\begin{gathered} V_1(x_1) > \max\{\chi_1(V_2(x_2)), \gamma_1(|u|)\} \\ \Downarrow \\ \max \dot{\overline{V}}_{1,F_1}(x_1, x_2, u) \leq -\rho_1(V_1(x_1)) \end{gathered}$$

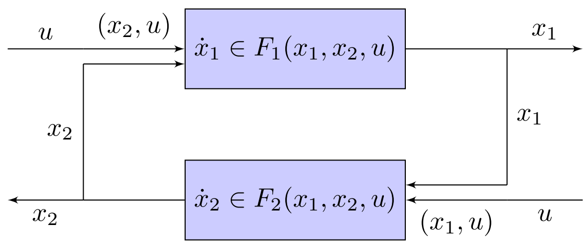


Figure 2.2: The interconnected system in (2.61).

2b) There exist $\rho_2 \in \mathcal{PD}$, and $\chi_2, \gamma_2 \in \mathcal{K}$ satisfying

$$
\begin{gathered}
V_2(x_2) > \max\{\chi_2(V_1(x_1)), \gamma_2(|u|)\} \\
\Downarrow \\
\max \overline{\dot{V}}_{2,F_2}(x_1, x_2, u) \leq -\rho_2(V_2(x_2))
\end{gathered}
$$

Since we want to combine the functions V_1 and V_2 to obtain a non-pathological ISS function $W : \mathbb{R}^n \to \mathbb{R}$ for the interconnected system (2.61), we need the following results from non-smooth analysis.

Fact 2.22. *[48, Theorem 2.6.6] Consider a locally Lipschitz function $V : \mathbb{R}^k \to \mathbb{R}$ and $\sigma \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$, and define $U := \sigma \circ V$. We have*

$$
\partial U(x) = \sigma'(V(x))\partial V(x), \quad \forall\, x \in \mathbb{R}^k,
$$

where $\sigma'(s)$ denotes the derivative of σ at $s \in \mathbb{R}$.

Fact 2.23. *[48, Proposition 2.3.12] Given two locally Lipschitz functions $V_1 : \mathbb{R}^k \to \mathbb{R}$ and $V_2 : \mathbb{R}^k \to \mathbb{R}$ consider the function $V(x) := \max\{V_1(x), V_2(x)\}$. Given any $z \in \mathbb{R}^k$ such that $V_1(z) = V_2(z)$, it holds that*

$$
\partial V(z) \subseteq \operatorname{co}\{\partial V_1(z), \partial V_2(z)\}.
$$

Moreover, we need this well-known comparison result, for the proof, see [112, Theorem 3.1].

Fact 2.24. *Given $\chi_1, \chi_2 \in \mathcal{K}$ such that $\chi_1 \circ \chi_2(r) < r$, $\forall r > 0$, there exists a continuously differentiable $\sigma \in \mathcal{K}_\infty$ with $\sigma'(s) > 0$ for all $s \in [0, \infty)$, such that*

$$
\chi_2(r) < \sigma(r), \quad \text{and } \chi_1(\sigma(r)) < r, \quad \forall\, r > 0. \tag{2.62}
$$

The geometrical intuition behind (2.62) is that the graph of the function σ lies between the graphs of χ_2 and χ_1^{-1}, see for example Fig.1 in [112]. Finally, the following lemma will be used in the proof of forthcoming Theorem 2.26.

Lemma 2.25. *Suppose $V_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ and $V_2 : \mathbb{R}^{n_2} \to \mathbb{R}$ are two non-pathological functions satisfying Assumption 2.4. Consider $\sigma \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}_+)$ such that $\sigma'(s) > 0$ for all $s > 0$ and the composite function $U_1 := \sigma \circ V_1$. Let $W(x_1, x_2) := \max\{U_1(x_1), V_2(x_2)\}$, and consider a point $z = (z_1, z_2) \neq (0, 0)$, $z_1 \in \mathbb{R}^{n_1}$, $z_2 \in \mathbb{R}^{n_2}$ such that $U_1(z_1) = V_2(z_2)$. It holds that*

$$
\partial W(z_1, z_2) = \operatorname{co}\left\{\partial \widehat{U}_1(z_1), \partial \widehat{V}_2(z_2)\right\} \tag{2.63}
$$

where

$$\partial\widehat{U}_1(z_1) := \begin{pmatrix}\partial U_1(z_1)\\ 0\end{pmatrix} = \begin{pmatrix}\sigma'(V_1(z_1))\partial V_1(z_1)\\ 0\end{pmatrix} \quad \textit{and}$$
$$\partial\widehat{V}_2(z_2) := \begin{pmatrix}0\\ \partial V_2(z_2)\end{pmatrix}.$$

Proof. Consider a point $z = (z_1, z_2) \neq (0,0)$ such that $U_1(z_1) = V_2(z_2)$, the inclusion $\partial W(z) \subseteq \mathrm{co}\{\partial\widehat{U}_1(z_1), \partial\widehat{V}_2(z_2)\}$ is obtained by Fact 2.23.

For the converse inclusion, due to convexity of $\partial W(z_1, z_2)$, it suffices to show that $\partial\widehat{U}_1(z_1) \subseteq \partial W(z_1, z_2)$ and $\partial\widehat{V}_2(z_2) \subseteq \partial W(z_1, z_2)$. We only prove the first inclusion, as the other one can be proved with a similar reasoning. Consider any $\hat{v} = \begin{pmatrix}v_1\\ 0\end{pmatrix} \in \partial\widehat{U}_1(z_1)$, where $v_1 \in \partial U_1(z_1)$. Recalling the definition of the generalized gradient (2.44), there exists a sequence $x_1^k \in \mathbb{R}^{n_1}$ such that $x_1^k \to z_1$ and such that $\nabla U_1(x_1^k)$ is defined and $\nabla U_1(x_1^k) \to v_1$. Moreover, from *1b)* and *2b)* of Assumption 2.4, the function V_2 has no local minima other than 0 because V_2 is a Lyapunov function for the unperturbed system $\dot{x}_2 \in F_2(0, x_2, 0)$. Thus, considering any point $x_2 \neq 0$, V_2 is decreasing along the solutions starting at $(0, x_2)$ with zero input. By local existence of solutions from any initial point, we have that x_2 cannot be a local minimum of V_2. Thus z_2 is not a local minima for V_2, and we can consider a sequence $x_2^\ell \to z_2$ such that $V_2(z_2) > V_2(x_2^\ell)$, for all $\ell \in \mathbb{N}^*$. Now, by continuity of U_1 and V_2, for each $\ell \in \mathbb{N}^*$, there exists $k_\ell \in \mathbb{N}^*$ such that

$$U_1(x_1^{k_\ell}) > V_2(x_2^\ell). \tag{2.64}$$

Consider the sequence $\overline{x}_\ell := \begin{pmatrix}x_1^{k_\ell}\\ x_2^\ell\end{pmatrix} \in \mathbb{R}^n$. We have $\overline{x}_\ell \to z = (z_1, z_2)$, and from equation (2.64)

$$W(\overline{x}_\ell) = \max\{U_1(x_1^{k_\ell}), V_2(x_2^\ell)\} = U_1(x_1^{k_\ell}), \quad \forall\, \ell \in \mathbb{N}^*.$$

Thus, W is differentiable at all $\overline{x}_\ell \in \mathbb{R}^n$ and

$$\lim_{\ell\to\infty} \nabla W(\overline{x}_\ell) = \begin{pmatrix}\lim_{\ell\to\infty} \nabla U_1(x_1^{k_\ell})\\ 0\end{pmatrix} = \begin{pmatrix}v_1\\ 0\end{pmatrix} = \hat{v}.$$

By definition of $\hat{v}$ and the generalized gradient, it follows that $\hat{v} \in \partial W(z_1, z_2)$ and hence $\partial\widehat{U}_1(z_1) \subseteq \partial W(z_1, z_2)$. Similarly, one can prove that $\partial\widehat{V}_2(z_2) \subseteq \partial W(z_1, z_2)$, and thus the equality (2.63) holds. □

We have now all the necessary tools to present a small gain theorem involving non-pathological ISS functions, adapting the idea firstly proposed in [112].

Theorem 2.26 (Generalized Small Gain Theorem)**.** *Consider the non-pathological functions V_1, V_2 satisfying Assumption 2.4 and suppose that*

$$\chi_1 \circ \chi_2(r) < r, \quad \forall r > 0. \tag{2.65}$$

Considering a function $\sigma \in \mathcal{K}_\infty \cap \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}_+)$ satisfying property (2.62) *in Fact 2.24, define $W : \mathbb{R}^n \to \mathbb{R}$ as*

$$W(x_1, x_2) := \max\{\sigma(V_1(x_1)), V_2(x_2)\}. \tag{2.66}$$

Then W is a non-pathological ISS function and thus system (2.61) *is ISS.*

Proof. We want to show that $W : \mathbb{R}^n \to \mathbb{R}$ satisfies all the conditions of Theorem 2.15. To this end, it is enough to show that

A) $\sigma \circ V_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ is non-pathological, and $W : \mathbb{R}^n \to \mathbb{R}$ is non-pathological.

B) There exist $\rho \in \mathcal{PD}$ and $\gamma \in \mathcal{K}$ such that

$$W(x) > \gamma(|u|) \;\; \Rightarrow \;\; \max \overline{\dot{W}}_F(x,u) \leq -\rho(|x|). \tag{2.67}$$

Proof of A): We recall that $V_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ is non-pathological and $\sigma \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ and $\sigma'(r) > 0$ for all $r > 0$. Defining $U_1 := \sigma \circ V_1$, by Fact 2.22 we have $\partial U_1(x) = \sigma'(V_1(x))\partial V_1(x)$ for all $x \in \mathbb{R}^{n_1}$. Moreover for any absolutely continuous function $\varphi : [0,T) \to \mathbb{R}^{n_1}$, by Definition 2.12 we have that $\partial V_1(\varphi(t))$ is a subset of an affine subspace orthogonal to $\dot{\varphi}(t)$, for almost every $t \in [0,T)$, and the same holds for $\partial U_1(\varphi(t)) = \sigma'(V_1(\varphi(t))\partial V_1(\varphi(t))$. Thus $U_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ is non-pathological. The non-pathology of $W : \mathbb{R}^n \to \mathbb{R}$ follows from the fact that pointwise maximum of non-pathological functions is non-pathological, as stated in Lemma 2.14.
Proof of B): We proceed by considering three cases. Let us define the sets

$$\begin{aligned}
\mathcal{O}_1 :=& \{(x_1,x_2) \in \mathbb{R}^n \mid V_2(x_2) < \sigma(V_1(x_1))\},\\
\mathcal{O}_2 :=& \{(x_1,x_2) \in \mathbb{R}^n \mid V_2(x_2) > \sigma(V_1(x_1))\},\\
\Gamma :=& \{(x_1,x_2) \in \mathbb{R}^n \mid V_2(x_2) = \sigma(V_1(x_1))\}.
\end{aligned}$$

For $z = (z_1,z_2) \in \mathcal{O}_1$, by continuity there exists a neighborhood $\mathcal{U}$ of z where $W(x) = \sigma(V_1(x_1))$, for all $x = (x_1,x_2) \in \mathcal{U}$. By Fact 2.22, we have that $\partial W(z) = \sigma'(V(z_1))\partial V_1(z_1) \times \{0\}$. Thus $f = (f_1,f_2) \in F(z,u)$ is such that $\langle p, f\rangle = a$, $\forall p \in \partial W(z)$ if and only if f_1 satisfies $\langle p_1, f_1\rangle = a$, $\forall p_1 \in \sigma'(V(z_1))\partial V_1(z_1)$. From (2.57), we thus get

$$\overline{\dot{W}}_F(z,u) = \sigma'(V(z_1))\overline{\dot{V}}_{1,F_1}(z_1,z_2,u). \tag{2.68}$$

Recalling that $z \in \mathcal{O}_1$ and equation (2.62), we have $\chi_1(V_2(z_2)) < \chi_1(\sigma(V_1(z_1))) < V_1(z_1)$. Thus, by condition *2a)* of Assumption 2.4, we have from (2.68) that

$$W(z) > \widehat{\gamma}_1(|u|) \Rightarrow \max \overline{\dot{W}}_F(z,u) \leq -\widehat{\rho}_1(W(z)),\; \forall z \in \mathcal{O}_1, \tag{2.69}$$

where $\widehat{\rho}_1(s) := \sigma'(\sigma^{-1}(s))\,\rho_1(\sigma^{-1}(s))$ is a positive definite function and $\widehat{\gamma}_1(s) := \sigma(\gamma_1(s))$ is class $\mathcal{K}$.
For $z = (z_1,z_2) \in \mathcal{O}_2$, following the same reasoning (but without the complications introduced by σ), one has that

$$W(z) > \gamma_2(|u|) \Rightarrow \max \overline{\dot{W}}_F(z,u) \leq -\rho_2(W(z)),\; \forall z \in \mathcal{O}_2. \tag{2.70}$$

Before addressing $z = (z_1,z_2) \in \Gamma$, using an idea proposed in [116], we introduce the following notation motivated by definition (2.45): Given $F : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and a locally Lipschitz function $V : \mathbb{R}^n \to \mathbb{R}$ we define

$$F^V(z,u) := \{f \in F(z,u) | \exists a \in \mathbb{R} : \langle v, f\rangle = a,\, \forall v \in \partial V(z)\}\,.$$

By Definition 2.11, it is clear that

$$\overline{\dot{V}}_F(z,u) = \left\{\langle v, f\rangle \mid v \in \partial V(z), f \in F^V(z,u)\right\}. \tag{2.71}$$

We continue by using the following set inclusion whose proof is postponed a few lines to avoid breaking the flow of the exposition:

$$F^W(z,u) \subseteq F_1^{V_1}(z_1,z_2,u) \times F_2^{V_2}(z_1,z_2,u). \tag{2.72}$$

Consider $z = (z_1, z_2)^\top \in \Gamma$ and take any $w \in \partial W(z)$, by Lemma 2.25, there exist $v_1 \in \partial V_1(z_1)$, $v_2 \in \partial V_2(z_2)$ and $\lambda \in [0,1]$ such that

$$w = \begin{pmatrix} \lambda\sigma'(V_1(z_1))v_1 \\ (1-\lambda)v_2 \end{pmatrix}.$$

Consider $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in F^W(z,u)$, so that, from (2.72), $f_1 \in F_1^{V_1}(z_1, z_2, u)$ and $f_2 \in F_2^{V_2}(z_1, z_2, u)$. Using (2.71), we may proceed as in (2.69) and (2.70) and use continuity of W to get, for all $z \in \Gamma = \mathrm{bd}(\mathcal{O}_1) \cap \mathrm{bd}(\mathcal{O}_2)$

$$\begin{aligned} W(z) > \widehat{\gamma}_1(|u|) &\Rightarrow \max_{\substack{f_1 \in F_1^{V_1}(z,u) \\ v_1 \in \partial V_1(z_1)}} \sigma'(V_1(z_1))\langle v_1, f_1\rangle \le -\widehat{\rho}_1(W(z)) \\ W(z) > \gamma_2(|u|) &\Rightarrow \max_{\substack{f_2 \in F_2^{V_2}(z,u) \\ v_2 \in \partial V_2(z_2)}} \langle v_2, f_2\rangle \le -\rho_2(W(z)). \end{aligned} \tag{2.73}$$

Using (2.73) we finally get that $W(z) > \max\{\widehat{\gamma}_1(|u|), \gamma_2(|u|)\}$ implies

$$\begin{aligned} \langle w, f\rangle &= \lambda\sigma'(V_1(z_1))\langle v_1, f_1\rangle + (1-\lambda)\langle v_2, f_2\rangle \\ &\le -\lambda\widehat{\rho}_1(W(z)) - (1-\lambda)\rho_2(W(z)) \\ &\le -\min\{\widehat{\rho}_1(W(z)), \rho_2(W(z))\}. \end{aligned}$$

Thus, letting $\gamma(s) := \max\{\widehat{\gamma}_1(s), \gamma_2(s)\}$ and $\rho(s) := \min\{\widehat{\rho}_1(s), \rho_2(s)\}$, we have

$$W(z) > \gamma(|u|) \ \Rightarrow \max \dot{\overline{W}}_F(z,u) \le -\rho(W(z)), \ \forall z \in \Gamma. \tag{2.74}$$

Collecting (2.69), (2.70) and (2.74) we can conclude (2.67), and prove item B).

We complete the proof by proving (2.72). To this, take any $f \in F^W(z,u)$. By definition of F in (2.61), we have that $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ for some $f_1 \in F_1(z_1, z_2, u)$ and $f_2 \in F_2(z_1, z_2, u)$. By Lemma 2.25 and Fact 2.22, for any $v_1 \in \partial V_1(z_1)$, the vector $w = \begin{pmatrix} \sigma'(V_1(z_1))v_1 \\ 0 \end{pmatrix} \in \partial W(z)$ and thus

$$\langle w, f\rangle = \sigma'(V_1(z_1))\langle v_1, f_1\rangle.$$

By varying v_1 in $\partial V_1(z_1)$ and recalling that $f \in F^W(z,u)$ (and thus $\langle w, f\rangle$ is constant for all $w \in \partial W(z)$), we obtain that $f_1 \in F_1^{V_1}(z_1, z_2, u)$, that is $\langle v_1, f_1\rangle$ constant w.r.t. $v_1 \in \partial V_1(z_1)$. The same reasoning applies to f_2, considering a vector $w = \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \in \partial W(z)$, with $v_2 \in \partial V_2(z_2)$, concluding the proof of the claim. $\square$

To summarize our main contribution behind the result presented in Theorem 2.26, we note that the idea of analyzing the derivative of the composite function W in the three sets $\mathcal{O}_1, \mathcal{O}_2, \Gamma$, appeared firstly in [112], and is the common idea of many results on small-gain theorems for interconnected systems, see for example [134]. The analysis in $\mathcal{O}_1$ and $\mathcal{O}_2$ was straightforward, but because of non-differentiability of V_1 and V_2, the analysis in the set Γ is different from [112]. In particular the additional tools from nonsmooth analysis have been used to study the Lie-derivative of W along F on the set Γ.

2.3 Perspectives

In this chapter, we provided sufficient conditions for the ISS/iISS property of switched systems for the two cases where the switching is either time-dependent or state-dependent. In the two cases, even though the approach was different and we used different tools for analysis, the common methodology was to make use of Lyapunov functions associated with individual subsystems to provide conditions for UGAS, ISS or iISS of the overall switched system while taking into consideration the characteristics of the switching signal. We then used these properties to analyze the interconnections. We will study more complex interconnections in hybrid setting in the next chapter. Since the primary contribution of this chapter is to study Lyapunov function based conditions for stability of switched systems, let us conclude this chapter by pointing out two research directions that we have partly investigated recently, which one could associate with an open problem formulated in [101]; see also [47] for a recent overview.

Converse results for constrained switching

We only presented our work describing sufficient conditions where the switching is either constrained by ADT/AAT, or as some function of the state. For these cases, there are in general few converse results in the literature [225, 45]. In the case of time-dependent constrained switching, the interest for studying necessary conditions lies in quantifying the best possible bounds that are necessary for stability of the switched systems. The sufficient condition proposed in this chapter are somewhat easier to derive for a broad class of systems but, in general, they tend to be conservative. For example, with ADT constrained switching signals, we may get very large lower bounds on ADT that ensures stability. While some work has been done to compute tight lower bounds in the linear case [225, 45], there are still several questions that need to be explored. In a recent work [64], we develop a converse counterpart of Theorem 2.5 when the switching signal is only constrained by ADT. Further investigation is being carried out about using such results to get tight lower bounds numerically.

For the case of state-dependent switched systems, we saw the use of a generalized notion of derivative for developing ISS characterization. This relies on finding a nonpathological Lyapunov function and a particular notion of derivative. We do not yet know if the given conditions are also necessary, that is, the converse results using this generalized notion of derivative have not yet been developed. These converse results also provide an indicator for developing algorithms that can be used for computing Lyapunov functions, which also remains an interesting direction of research.

Computing Lyapunov functions

The statement of Theorem 2.5 says that, a certain stability property of interest holds if the inequality (2.21) is satisfied. The parameters describing this inequality are actually obtained from the Lyapunov functions and the system data. So, a somewhat natural question that arises here is to compute Lyapunov functions that provide us the optimal values of the parameters appearing in (2.21). Once again, we have investigated this question for ADT constrained switching signals in our paper [88], where we use linear programming based techniques to find Lyapunov functions, while simultaneously minimizing the lower bounds on ADT which is expressed as a nonlinear function of the variables in the optimization problem. At the same time, we have recently seen different numerical techniques using neural networks [85] for computation of Lyapunov functions. Interestingly, the paper [85] proposes a decomposition of an ODE into several subsystems and uses small-gain conditions to find Lyapunov functions in compositional form. Such works open up the possibility of using neural network architectures for computation of Lyapunov function and it needs to be seen whether the kind of Lyapunov functions studied in this chapter could also benefit from such computational techniques.

Here, we also consider the possibility of computing Lyapunov functions for the state-dependent switched systems. The proposed class of piecewise differentiable functions includes as a subclass the functions obtained by taking pointwise maximum or minimum of Lyapunov functions associated to the subsystems, which is commonly used in analysis and design of switched systems. To develop a framework for computing such Lyapunov functions, one could consider studying algorithms which implement Clarke generalized gradient but this can be conservative for certain classes of switched systems. The other approach is to use the generalize Lie derivative proposed in Definition 2.11.

3

Singularly Perturbed Hybrid Systems and Networks

Decision making in multiagent systems has been an important topic of research over the years and different tools have been developed in the research community for analysis and control of such complex systems [108, 150, 29, 131]. One important question in studying these networks is to ensure certain level of performance despite the uncertainty in communication between the agents. In the work presented in this chapter, we are particularly concerned with the case where the communication graph between dynamic agents in a network may switch from one configuration to another, and we develop tools which allow us to understand the asymptotic behavior of such systems. This way we make transition from somewhat simpler interconnections presented in Chapter 2 to more complex interconnections where the coupling rule between the agents and the structure of the underlying graph plays a role in the analysis as well.

We adopt the framework of hybrid systems to model the system with aforementioned characteristics [80]. The proposed model describes an interconnected hybrid system where one subsystem has stable continuous dynamics, and the other subsystem is modeled by a discrete-time switching system and possibly unstable continuous dynamics. The particular structure that we consider in our work is analyzed using singular perturbation methods. The basic idea is to see the system as a perturbation of a nominal system that comprises the boundary-layer system (describing the limit of fast dynamics) and the quasi-steady-state system (described by slow system with limiting value of fast dynamics). This sort of decomposition into slow and fast dynamics has been used for analyzing robustness of input-to-state stability to slowly varying parameters and rapidly varying signals in [216].

For the class of interconnected hybrid systems analyzed in this chapter, we analyze stability with respect to a parameter in system description which governs the frequency of jump dynamics and plays the role of a singular perturbation parameter. In particular, as the parameter gets smaller, the frequency of jumps increases. This viewpoint was adopted by the authors in [212] and our recent submission [213] generalizes that framework in several directions with new proof techniques. For such systems, it is natural to consider the decomposition of dynamics in the so-called *quasi-steady-state* (QSS) system and a *boundary-layer* (BL) system. With appropriate stability assumptions on these subsystems, we can analyze practical stability of the interconnected system for small enough value of the perturbation parameter. An interesting element of this decomposition is that our boundary-layer system is described by a discrete-time switching system, and the stability of this system is formulated in terms of Lyapunov functions which is consistent with several studies in the literature [3, 78].

As an application of theoretical results, we also provide a case study on practical synchronization of nonlinear oscillators connected via an undirected graph, where the agents may reset their states based on the information received from their neighbors. Dynamical systems with resetting controllers have been studied for certain applications [173]. For multi-agent systems with diffusive coupling and continuous-time heterogeneous agents [122], [161], it is observed that the agents achieve practical convergence if the coupling strength is large enough. One can view the coupling parameter as a singular perturbation parameter. More recently, we have also seen the use of discontinuous coupling for synchronization of networked system in [52]. These discontinuities in the couplings can be interpreted as changes in the graph topology being used to describe the connections between agents, and consensus problems with time-varying graphs have gathered significant interest in the literature; see [108, 39] and references therein. In particular, our work provides analysis of networked systems where the discontinuity in coupling arises due to time-sampled information exchange between neighbors and the average value of the sampling interval plays the role of the singular perturbation parameter. The change in the graph topology at communication times is conveniently modeled using the framework of discrete-time switched systems. Under natural assumptions on the connectivity of the switching communication graph and dynamics of the agents, we show practical synchronization of the agents' trajectories under sufficiently frequent communication.

3.1 System Class

In this chapter, we consider hybrid dynamical systems described by ordinary differential inclusions over a flow set, and difference inclusions over a jump set. The peculiar feature of our models is that they are singularly perturbed, that is, there is a parameter $\epsilon > 0$ which affects the qualitative behavior of the trajectories and we analyze stability of the system for small enough values of this parameter.

3.1.1 Timer with frequent jumps

An important modeling aspect for us is to consider a timer with the following dynamics

$$\begin{aligned} \epsilon\dot{\tau} &\in [\sigma_1, \sigma_2], && \tau \in [0, N_0] \\ \tau^+ &= \tau - 1, && \tau \in [1, N_0]. \end{aligned} \tag{3.1}$$

The evolution of the timer τ, determined by the parameters $N_0 \in \mathbb{N}^*$, $0 < \epsilon \leq 1$, and $0 < \sigma_1 \leq \sigma_2$, describes the interplay between the flow and the jump. In our work, we will consider hybrid systems where the jumps are triggered only by the timer τ. We are particularly interested in the case where the parameter ϵ is very small, and we refer to ϵ as the *singular perturbation parameter*.

In the literature on hybrid systems, we find different setups where time-triggered jumps play a crucial role in the stability of the hybrid system. As we saw in the previous chapter, the notion of (average) dwell-time [153, 99] is used for putting an upper bound on how frequently the jumps can occur, whereas the notion of reverse average dwell-time [102] is used for imposing a lower bound on the frequency of jumps. The timer in (3.1) has the property that it simultaneously puts a lower and upper bound on the number of jumps over an interval $(s,t]$, that is,

$$1 + \frac{\sigma_1(t-s)}{\epsilon} - N_0 \leq J_{(s,t]} \leq N_0 + \frac{\sigma_2(t-s)}{\epsilon} \tag{3.2}$$

where $J_{(s,t]}$ denotes the number of jumps in the variable τ over the interval $(s,t]$. It readily follows that, with such a timer, there is no accumulation time for jumps (i.e., there are finitely many jumps in any finite continuous-time interval) and the domain of every complete solution has unbounded continuous-time element.

3.1.2 Overall model

We are interested in modeling the evolution of two variables $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$. The dynamics of x variable are modeled by an ordinary differential equation. For the evolution of y-dynamics, in addition to a flow map described by a differential equation, we have a collection of jump maps to define the jump dynamics. The times at which these jumps occur is modeled by the timer τ in (3.1) (that is, the jumps occur whenever τ jumps), and the jump map for y is indexed by a variable p belonging to a finite set $\mathcal{P}$.

The state variables for the overall system are $(x,y,\tau,\mathsf{p}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathcal{P}$, and we consider a closed set C_{xy} contained in an open domain $\mathscr{D}_{xy} \subset \mathbb{R}^n \times \mathbb{R}^m$. The state variables evolve inside the flow set

$$\mathcal{C} := C_{xy} \times [0, N_0] \times \mathcal{P},$$

and go through an instantaneous change inside the jump set

$$\mathcal{D} := C_{xy} \times [1, N_0] \times \mathcal{P}.$$

The underlying equations that describe the evolution of the state variables are:

$$\begin{cases} \dot{x} &= f_x(x,y,\epsilon), \\ \dot{y} &= f_y(x,y,\epsilon), \\ \epsilon\dot{\tau} &\in [\sigma_1, \sigma_2], \\ \dot{\mathsf{p}} &= 0, \end{cases} \qquad \text{for } (x,y,\tau,\mathsf{p}) \in \mathcal{C}, \tag{3.3a}$$

$$\begin{cases} x^+ &= x, \\ y^+ &= g_y(x,y,\mathsf{p}), \\ \tau^+ &= \tau - 1, \\ \mathsf{p}^+ &\in G_{\mathsf{p}}(\mathsf{p}), \end{cases} \qquad \text{for } (x,y,\tau,\mathsf{p}) \in \mathcal{D}. \tag{3.3b}$$

For the timer dynamics, the positive constants σ_1 and σ_2 are such that $\sigma_2 \geq \sigma_1 > 0$. We assume that the vector fields $f_x : \mathscr{D}_{xy} \times \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, $f_y : \mathscr{D}_{xy} \times \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ and the jump maps $g_y : \mathscr{D}_{xy} \times \mathcal{P} \to \mathbb{R}^m$ are continuous. The set-valued map $G_{\mathsf{p}} : \mathcal{P} \rightrightarrows \mathcal{P}$ allows non-deterministic behavior for the variable p at jumps. The solutions of system (3.3) are defined over a subset of $\mathbb{R}_{\geq 0} \times \mathbb{N}^*$, called *hybrid time domain* [80, Chap. 2]. In what follows, we use the notation $(x_0, y_0) := (x(0,0), y(0,0))$ to denote the initial condition for (x,y)-dynamics.

For the problems studied in this chapter, the parameter ϵ appearing in (3.3) is taken to be small, which implies that the jumps occur frequently, and thus, the evolution of y relies heavily on the jump dynamics. Since the jumps correspond to rapid evolution, we call x the *slow variable* and y the *fast variable*.

3.1.3 System decomposition

We are interested in studying the behavior of system (3.3) when the parameter ϵ is sufficiently small and takes values close to zero. It turns out that the evolution of trajectories $(x(\cdot), y(\cdot))$ can be approximated by two different dynamical systems. The first system is called the *boundary-layer* system, which is given by

$$\begin{aligned} \mathsf{x}^+ &= \mathsf{x} \\ \mathsf{y}^+ &= g_y(\mathsf{x}, \mathsf{y}, \mathsf{p}) \\ \mathsf{p}^+ &\in G_{\mathsf{p}}(\mathsf{p}) \end{aligned} \tag{3.4}$$

for $(\mathsf{x}, \mathsf{y}) \in C_{xy}$, and $\mathsf{p} \in \mathcal{P}$. Recalling the expression of the timer dynamics, we observe that a small value of ϵ corresponds to many jumps in the trajectories of system (3.3) within a small continuous-time interval. This means that, from an initial condition $(x(0,0), y(0,0))$, while the solution $x(\cdot)$ remains close to $x(0,0)$, the solution $y(\cdot)$ evolves like a solution to the discrete-time system (3.4). In order to define an equilibrium point for the discrete-time system (3.4), we let C_x denote the natural projection of C_{xy} on $\mathbb{R}^n$, that is, $C_x := \{x \in \mathbb{R}^n : \exists\, y \text{ s.t. } (x, y) \in C_{xy}\}$; and we stipulate that there is a continuously differentiable function $h : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$g_y(\mathsf{x}, h(\mathsf{x}), \mathsf{p}) = h(\mathsf{x}), \quad \forall\, \mathsf{x} \in \overline{C}_x, \quad \forall\, \mathsf{p} \in \mathcal{P}, \tag{3.5}$$

where $\overline{C}_x$ denotes the closure of C_x. We assume that $(\mathsf{x}, h(\mathsf{x})) \in C_{xy}$ for all $\mathsf{x} \in C_x$ in order for the problem to be well-posed. Imagine that $y(t, j)$ rapidly approaches $h(x(0, 0))$ after some jumps, and remains near $h(x(t, j))$ as time goes on. Then, the behavior of $x(t, j)$ can be approximated by the *quasi-steady-state* system defined as

$$\dot{\mathsf{x}} = f_x(\mathsf{x}, h(\mathsf{x}), 0), \quad \mathsf{x} \in \overline{C}_x, \tag{3.6}$$

which is the evolution of $x(\cdot)$ on the *slow manifold* $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y = h(x)\}$, when $\epsilon = 0$. We associate a compact forward invariant set $\mathcal{A}_x$ with (3.6) so that if $\mathsf{x}(0) \in \mathcal{A}_x$, then for every solution of (3.6), we have $\mathsf{x}(t) \in \mathcal{A}_x$.

3.1.4 Stability notions

The primary focus of this chapter is the stability analysis of system (3.3). In particular, we are interested in developing bounds on the (x, y) trajectories, when the initial conditions belong to a given compact set and the perturbation parameter ϵ tends to 0. This stability notion is formalized in the following definition:

Definition 3.1. For the hybrid system (3.3), with boundary-layer system (3.4) and quasi-steady-state system (3.6), the compact set $\mathcal{A} := \{(\mathsf{x}, \mathsf{y}) \in C_{xy} : \mathsf{x} \in \mathcal{A}_x, \mathsf{y} = h(\mathsf{x})\} \times [0, N_0] \times \mathcal{P}$ is *two-time-scale semiglobally practically stable*, or TSPS for short, if for a given compact set $I_{xy} \subset C_{xy}$, there are a positive number ϵ^*, class $\mathcal{KL}$ functions β_x and β_y, and class $\mathcal{K}$ functions χ_x and χ_y such that, for any $0 < \epsilon \leq \epsilon^*$ and for any $(x_0, y_0) \in I_{xy}$, the solution $(x(t, j), y(t, j))$ to (3.3) satisfies the following,

$$|x(t, j)|_{\mathcal{A}_x} \leq \beta_x(|x_0|_{\mathcal{A}_x}, t) + \chi_x(\epsilon), \tag{3.7a}$$

$$|y(t, j) - h(x(t, j))| \leq \beta_y(|y_0 - h(x_0)|, t/\epsilon) + \chi_y(\epsilon), \tag{3.7b}$$

for all $(t, j) \in \operatorname{dom}(x, y, \tau, \mathsf{p})$.

For the results developed in this chapter, we impose some stability assumptions on the boundary-layer system and the quasi-steady-state system. This will allow us to establish the property stated in Definition 3.1 for the coupled system (3.3).

3.2 Main Result

To establish the property described in Definition 3.1 for system (3.3), we impose some stability assumptions on the quasi-steady-state system and the boundary-layer system.

Assumption 3.1. There exists a continuously differentiable function $V_x : \mathscr{D}_x = \{x : \exists\, y \text{ s.t. } (x, y) \in \mathscr{D}_{xy}\} \to \mathbb{R}$ such that, for all $\mathsf{x} \in \overline{C}_x$,

$$\underline{\alpha}_x(|\mathsf{x}|_{\mathcal{A}_x}) \leq V_x(\mathsf{x}) \leq \overline{\alpha}_x(|\mathsf{x}|_{\mathcal{A}_x}), \tag{3.8}$$

$$\nabla_x V_x(\mathsf{x}) \cdot f_x(\mathsf{x}, h(\mathsf{x}), 0) \leq -\alpha_x(V_x(\mathsf{x})), \tag{3.9}$$

where $\underline{\alpha}_x$ and $\overline{\alpha}_x$ are class $\mathcal{K}_\infty$ functions, and α_x is a class $\mathcal{K}$ function.

Following Assumption 3.1, we use [135, Lemma 4.4] to get a class $\mathcal{KL}$ function β_x^0 from α_x appearing in (3.9). Let[1]

$$\eta_x(s) := -\int_1^s \frac{dr}{\min\{r, \alpha_x(r)\}}, \qquad \text{for } s > 0,$$

and

$$\beta_x^0(s,t) := \begin{cases} \eta_x^{-1}(\eta_x(s) + t), & \text{for } s > 0, \\ 0, & \text{for } s = 0. \end{cases} \tag{3.10}$$

For certain applications, such as the one considered later in Section 3.3, the switching may be constrained and we may only have Lyapunov functions which decrease after a certain number of jumps. In this section, we formulate the stability result which is particularly suited for such cases. More precisely, for each $(\mathsf{x}, \mathsf{y}, \mathsf{p}) \in C_{xy} \times \mathcal{P}$, let us introduce the notation

$$(g_y(\mathsf{x}, \mathsf{y}, \mathsf{p}), G_{\mathsf{p}}(\mathsf{p}))^r := \{(y, \mathsf{q}) : (y, \mathsf{q}) = (y_r, \mathsf{p}_r), (y_0, \mathsf{p}_0) = (\mathsf{y}, \mathsf{p}), (y_{i+1}, \mathsf{p}_{i+1}) \in (g_y(\mathsf{x}, y_i, \mathsf{p}_i), G_{\mathsf{p}}(\mathsf{p}_i)),\ 0 \leq i \leq r-1\}.$$

Assumption 3.2. For every compact set $K_x \subset C_x$, there exist a positive integer $\mathfrak{n}$, continuously differentiable functions $V_y : \mathscr{D}_{xy} \times \mathcal{P} \to \mathbb{R}$, class $\mathcal{K}_\infty$ functions $\underline{\alpha}_y$ and $\overline{\alpha}_y$, and a non-negative constant $\gamma_m < 1$ that satisfy the following items:

(BL-M1) for every $(\mathsf{x}, \mathsf{y}, \mathsf{p}) \in \{(\mathsf{x}, \mathsf{y}) \in C_{xy} : \mathsf{x} \in K_x\} \times \mathcal{P}$,

$$\underline{\alpha}_y(|\mathsf{y} - h(\mathsf{x})|) \leq V_y(\mathsf{x}, \mathsf{y}, \mathsf{p}) \leq \overline{\alpha}_y(|\mathsf{y} - h(\mathsf{x})|),$$

(BL-M2) there exists a set $P \subset \mathcal{P}$, such that,

- every solution of $(\mathsf{p} \in \mathcal{P}, \mathsf{p}^+ \in G_{\mathsf{p}}(\mathsf{p}))$ reaches P in at most $\mathfrak{n}$ steps,
- for every $(\mathsf{x}, \mathsf{y}, \mathsf{p}) \in \{(\mathsf{x}, \mathsf{y}) \in C_{xy} : \mathsf{x} \in K_x\} \times P$, there exists $r \in \{1, \ldots, \mathfrak{n}\}$, such that for every $(g, \mathsf{q}) \in (g_y(\mathsf{x}, \mathsf{y}, \mathsf{p}), G_{\mathsf{p}}(\mathsf{p}))^r$, we have that $\mathsf{q} \in P$ and

$$V_y(\mathsf{x}, g, \mathsf{q}) \leq \gamma_m V_y(\mathsf{x}, \mathsf{y}, \mathsf{p}). \tag{3.11}$$

With the two aforementioned assumptions, we arrive at the following main result of this chapter.

[1] One can also take $\eta_x(s) = -\int_1^s \frac{dr}{\alpha_x(r)}$, for $s > 0$, if it holds that $\lim_{s \to 0^+} \eta_x(s) = +\infty$.

Theorem 3.2. *For the hybrid system* (3.3)*, suppose that the boundary-layer system* (3.4) *satisfies Assumption 3.2 and the quasi-steady-state system* (3.6) *satisfies Assumption 3.1. Then, the set $\mathcal{A}$ is TSPS, and the estimate* (3.7a) *holds with*

$$\beta_x(s,t) := \underline{\alpha}_x^{-1}\Big(\max\Big\{\beta_x^0\left(2\beta_x^0(\overline{\alpha}_x(s),1), \max\{0,t-1\}\right), \beta_x^0(\overline{\alpha}_x(s),t)\Big\}\Big) \tag{3.12}$$

and (3.7b) *holds with*

$$\beta_y(s,\tau) := C_g \circ \underline{\alpha}_y^{-1}\Big(2\gamma_m^{\frac{\sigma_1}{N_0\mathfrak{n}}\tau}\overline{\alpha}_y(s)\Big) \tag{3.13}$$

where γ_m appears in (3.11)*, and C_g is a class $\mathcal{K}$ function.*

The expressions for class $\mathcal{K}$ functions χ_x and χ_y are derived in the proof of Theorem 3.2 which has been partly developed in Section 3.2.2. The complete proof can be found in [213].

3.2.1 Boundary-layer system with constrained switching

A natural motivation for Assumption 3.2 is to analyze boundary-layer systems with constrained switching where we do not have a common Lyapunov function. To explain how Assumption 3.2 can be used in such cases, we consider a finite discrete set $\mathcal{Q}$, and a family of continuous mappings $g_y(\cdot,\cdot,q) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$, with $q \in \mathcal{Q}$. The mappings $g_y(\cdot,\cdot,q)$ indexed by $q \in \mathcal{Q}$ would be used in the description of the discrete dynamics, and we are particularly interested in the case where the transition between the jump maps respects some pre-specified constraints. To describe these constraints, we consider a set $\mathcal{S}$ which comprises *strings* (also called *words*) of finite length formed by the alphabet $\mathcal{Q}$. For a string $\mathfrak{s} \in \mathcal{S}$, we use the notation $\mathfrak{s}_\mathfrak{r} \in \mathcal{Q}$ to denote the $\mathfrak{r}$-th letter of that string $\mathfrak{s}$, with $1 \le \mathfrak{r} \le |\mathfrak{s}|$, and $|\mathfrak{s}|$ denoting the total number of letters in $\mathfrak{s}$. We also consider the set-valued mapping $\mathrm{succ} : \mathcal{S} \rightrightarrows \mathcal{S}$, where $\mathrm{succ}(\mathfrak{s}) \subset \mathcal{S}$ denotes a nonempty set of admissible successors of $\mathfrak{s}$. We emphasize that the set $\mathcal{Q}$ is used as index for different jump maps, whereas the set $\mathcal{S}$ is used to encode the constraints on switching sequence which respect certain assumptions related to stability. We assume that $\mathcal{S}$ is a finite set, and let $\mathfrak{n} := \max_{\mathfrak{s}\in\mathcal{S}} |\mathfrak{s}|$.

Now, we can model our discrete dynamics with constraints on the switching sequence in the form of a conventional one-step difference equation by using certain discrete variables. Indeed, we suppose the jump map $g_y(\cdot,\cdot,q)$ is selected by the rule $q = \mathfrak{s}_\mathfrak{r}$, where the string $\mathfrak{s} \in \mathcal{S}$. To embed the variables $(\mathfrak{s},\mathfrak{r})$ in a Euclidean space of fixed dimension, we consider the mapping $\mathsf{a} : \mathcal{S} \to \mathbb{R}^\mathfrak{n}$ that appends some symbols not in $\mathcal{Q}$ to a string $\mathfrak{s} \in \mathcal{S}$ to make it a word of length $\mathfrak{n}$, and we let $\bar{\mathfrak{s}} = \mathsf{a}(\mathfrak{s})$. The inverse of this operation is considered to be well-defined and allows us to recover the string $\mathfrak{s} \in \mathcal{S}$ from $\bar{\mathfrak{s}} \in \mathbb{R}^\mathfrak{n}$, that is, $\mathfrak{s} = \mathsf{a}^{-1}(\bar{\mathfrak{s}})$. We can now define the evolution of $\bar{\mathfrak{s}}$ and an indexing variable $\mathfrak{r} \in \mathbb{N}^*$ as follows:

$$\bar{\mathfrak{s}}^+ \in S_1(\bar{\mathfrak{s}},\mathfrak{r}) = \begin{cases} \{\bar{\mathfrak{s}}\}, & \text{if } \mathfrak{r} < |\mathsf{a}^{-1}(\bar{\mathfrak{s}})| \\ \mathsf{a}(\mathrm{succ}(\mathsf{a}^{-1}(\bar{\mathfrak{s}}))), & \text{otherwise} \end{cases} \tag{3.14a}$$

$$\mathfrak{r}^+ \in S_2(\bar{\mathfrak{s}},\mathfrak{r}) = \begin{cases} \{\mathfrak{r}+1\}, & \text{if } \mathfrak{r} < |\mathsf{a}^{-1}(\bar{\mathfrak{s}})| \\ \{1\}, & \text{otherwise.} \end{cases} \tag{3.14b}$$

We note that the cardinality of the variable $\bar{\mathfrak{s}}$ remains fixed over time, and $(\bar{\mathfrak{s}},\mathfrak{r}) \in \mathbb{R}^\mathfrak{n} \times [1,\mathfrak{n}]$. We can now apply Assumption 3.2 to the boundary-layer system (3.4) with $\mathsf{p} = (\mathfrak{s},\mathfrak{r})$, $\mathcal{P} = \mathcal{S} \times \{1,\dots,\mathfrak{n}\}$, and choosing $P = \mathcal{S} \times \{1\}$. This will provide us the sufficient conditions for asymptotic stability of switched systems with multiple Lyapunov functions in the same spirit as [3].

3.2.2 Proof of Theorem 3.2

For the proof, it is convenient to introduce $z := y - h(x)$ and consider the closed set

$$C_X := \{(x, z) : (x, z + h(x)) \in C_{xy}\}.$$

With the state $(x, z, \tau, \mathsf{p}) \in C_X \times [0, N_0] \times \mathcal{P}$, we can rewrite the flow dynamics of (3.3) as

$$\begin{aligned} \dot{x} &= f_x(x, h(x), 0) + [f_x(x, z + h(x), \epsilon) - f_x(x, h(x), 0)] \\ &= f_x(x, h(x), 0) + \widetilde{f}_x(x, z + h(x), \epsilon), \\ \dot{z} &= f_y(x, z + h(x), \epsilon) - \frac{\partial h}{\partial x}(x) f_x(x, z + h(x), \epsilon) \\ &=: f_z(x, z, \epsilon), \\ \epsilon\dot{\tau} &\in [\sigma_1, \sigma_2], \\ \dot{\mathsf{p}} &= 0, \end{aligned} \tag{3.15a}$$

and the jump dynamics of (3.3), for $(x, z, \tau, \mathsf{p}) \in C_X \times [1, N_0] \times \mathcal{P}$, as

$$\begin{aligned} x^+ &= x, \\ z^+ &= g_y(x, z + h(x), \mathsf{p}) - h(x) =: \widetilde{g}(x, z, \mathsf{p}), \\ \tau^+ &= \tau - 1, \\ \mathsf{p}^+ &\in G_{\mathsf{p}}(\mathsf{p}). \end{aligned} \tag{3.15b}$$

For simplicity, define $X := \mathrm{col}(x, z)$ so that parts of (3.15) can be written concisely as

$$\dot{X} = F(X, \epsilon) := \begin{bmatrix} f_x(x, z + h(x), \epsilon) \\ f_z(x, z, \epsilon) \end{bmatrix}$$
$$X^+ = G(X, \mathsf{p}) := \begin{bmatrix} x \\ \widetilde{g}(x, z, \mathsf{p}) \end{bmatrix}.$$

Let $P \subset \mathcal{P}$ be the set introduced in Assumption 3.2. Without loss of generality, we assume that $\mathsf{p}(0, 0) \in P$. For a solution of (3.15), we can extract a sequence from the values of p-trajectory which corresponds to the change in values of p and also when it enters the set P. Letting $\mathsf{p}(0, 0) = \mathsf{p}_1^0$, we label this sequence as $\mathsf{p}_1^0, \mathsf{p}_2^0, \ldots, \mathsf{p}_{r_0}^0,\ \mathsf{p}_1^1, \mathsf{p}_2^1, \ldots, \mathsf{p}_{r_1}^1, \ldots, \mathsf{p}_1^k, \mathsf{p}_2^k, \ldots, \mathsf{p}_{r_k}^k, \mathsf{p}_1^{k+1}, \ldots$. In this sequence, $\mathsf{p}_1^k \in P$ for each $k \in \mathbb{N}^*$ and after $r_k \leq \mathfrak{n}$ transitions, we arrive at $\mathsf{p}_1^{k+1} \in P$ so that **(BL-M2)** in Assumption 3.2 holds. We let $J_k = \sum_{j=0}^{k} r_j$. Consequently, we will use the notation $\widetilde{g}_r^k(x, z) := \widetilde{g}(x, z, \mathsf{p}_r^k)$ and $G_r^k(X) := G(X, \mathsf{p}_r^k)$ with $r \leq r_k \leq \mathfrak{n}$. For multiple jumps indexed by the discrete variable p, we will use the notation $G_{r,1}^k(X) := G_r^k \circ G_{r-1}^k \circ \cdots \circ G_1^k$.

Now we look for a compact set over which the trajectories evolve for a while from the initial conditions. This allows us to find quantitative measures used for ultimate bound and convergence rate of z, with which we look at the asymptotic behavior of the overall coupled system.

Recalling our switching model, it follows that, for any j,

$$t_{j+1} - t_j \leq \frac{N_0 \epsilon}{\sigma_1} \leq \frac{N_0}{\sigma_1} =: T_0. \tag{3.16}$$

Now let $\phi_\epsilon(t, X_0)$ be the solution X at time t to the flow $\dot{X} = F(X, \epsilon)$ with $X(0) = X_0$. Define

$$\Phi_{\mathsf{p}}(A) := \{G(\phi_\epsilon(t, X_0), \mathsf{p}) \in C_X : X_0 \in A, \epsilon \in [0, 1], t \in [0, T_0]\},$$

which is a set of reachable states from a set A within the time T_0 when ϵ belongs to $[0,1]$, followed by one jump through $G(\cdot,\mathsf{p})$. Note that $\Phi_{\mathsf{p}}(A)$ is compact if A is compact. Let $\mathcal{P}^*$ be the set of all sequences of elements of $\mathcal{P}$ of length at most $\mathfrak{n}$ for which **(BL-M2)** holds, and denote a generic element of $\mathcal{P}^*$ by $\mathsf{p}^k = \mathsf{p}_1^k \cdot \mathsf{p}_2^k \cdots \mathsf{p}_{r_k}^k$. Define, for $\mathsf{p}^k \in \mathcal{P}^*$

$$\Phi_{\mathsf{p}^k}(A) := \Phi_{\mathsf{p}_r^k} \circ \cdots \circ \Phi_{\mathsf{p}_2^k} \circ \Phi_{\mathsf{p}_1^k}(A),$$

which is compact when A is compact. Finally, let

$$\Phi_{\mathcal{P}}(A) := \cup_{\mathsf{p}^k \in \mathcal{P}^*} \Phi_{\mathsf{p}^k}(A),$$

which is a reachable set from A under the jumps by any sequence in $\mathcal{P}^*$ and the flows up to the time $T_0\mathfrak{n}$ with $\epsilon \in [0,1]$. Since $\mathcal{P}^*$ is finite, the set $\Phi_{\mathcal{P}}(A)$ is compact whenever A is compact.

From the given compact set $I_{xy} \subset C_{xy}$, we consider the compact set $I_x := \{x \in C_x : \exists\, y \text{ s.t. } (x,y) \in I_{xy}\}$, and the continuously differentiable function V_x obtained from Assumption 3.1. We let

$$\mu := \max_{x \in I_x} V_x(x)$$

and define

$$K_x := \{x \in C_x : V_x(x) \leq \mu + \delta_\mu\} \tag{3.17}$$

with a small $\delta_\mu > 0$.

Now, with K_x in (3.17), Assumption 3.2 yields V_y, and using V_y, define

$$\nu := \max_{(x,y) \in I_{xy}, \mathsf{p} \in \mathcal{P}} V_y(x,y,\mathsf{p}). \tag{3.18}$$

Then, it is obvious that

$$\begin{aligned} I_{xy} &\subset \{(x,y) \in C_{xy} : x \in I_x, V_y(x,y,\mathsf{p}) \leq \nu, \forall \mathsf{p} \in \mathcal{P}\} \\ &\subset \{(x,y) \in C_{xy} : x \in K_x, \exists\, \mathsf{p} \in \mathcal{P} \text{ s.t. } V_y(x,y,\mathsf{p}) \leq \nu\} \\ &=: K_{xy0}. \end{aligned}$$

Let $K_X := \{(x,z) : (x, z + h(x)) \in K_{xy0}\}$. For convenience, let $V_z(x,z,\mathsf{p}) := V_y(x, z + h(x), \mathsf{p})$, $\mathsf{p} \in \mathcal{P}$. Then,

$$K_X = \cup_{\mathsf{p} \in \mathcal{P}} \{(x,z) \in C_X : x \in K_x, V_z(x,z,\mathsf{p}) \leq \nu\}. \tag{3.19}$$

Finally, define

$$\Omega_X := \Phi_{\mathcal{P}}(K_X).$$

And we keep in mind that, by the construction of Ω_X,

$$X(t,j) \in \Omega_X, \quad \forall (t_{J_{k-1}}, J_{k-1}) \leq (t,j) \leq (t_{J_k}, J_k), \tag{3.20}$$

whenever $X(t_{J_{k-1}}, J_{k-1}) \in K_X$ for any $k \geq 1$.[2]

On the compact set $\Omega_X \times [0,1]$ for (x,z,ϵ), one can find M_x and M_z such that

$$|f_x(x, z + h(x), \epsilon)| \leq M_x, \quad |f_z(x,z,\epsilon)| \leq M_z. \tag{3.21}$$

[2]In the proof, the time inequality such as $(t_a, a) \leq (t,j) < (t_b, b)$ should be understood as $t_a \leq t < t_b$, $a \leq j < b$, and $(t,j) \in \operatorname{dom} X$, even if $\operatorname{dom} X$ is often omitted.

Define $M_{xz} := \max\{M_x, M_z\}$. Due to continuity of $\widetilde{f}$, there is a class $\mathcal{K}_\infty$ function λ_f such that

$$|\widetilde{f}_x(x, z + h(x), \epsilon)| \leq \lambda_f(|z|) + \lambda_f(\epsilon), \tag{3.22}$$

for all $(x, z, \epsilon) \in \Omega_X \times [0, 1]$. By construction, the mapping $X \mapsto G(X, \mathsf{p})$ is continuous and $G(0, \mathsf{p}) = 0$, for all $\mathsf{p} \in \mathcal{P}$. Over the compact set Ω_X, we can find a class $\mathcal{K}$ function λ_G such that $|G(X_2, \mathsf{p}) - G(X_1, \mathsf{p})| \leq \lambda_G(|X_2 - X_1|)$, for all $X_1, X_2 \in \Omega_X$ and all $\mathsf{p} \in \mathcal{P}$. We will use the notation $\lambda_G^1 = \lambda_G$ and $\lambda_G^\ell := \lambda_G \circ \lambda_G^{\ell-1}$, for each $\ell \geq 2$. Also, we let L_V^x be Lipschitz constant for V_x on Ω_X and L_V^z be Lipschitz constant for $V_z(\cdot, \cdot, \mathsf{p})$ on Ω_X for all $\mathsf{p} \in \mathcal{P}$.

Suppose that ϵ is sufficiently small. In particular, ϵ is small enough to satisfy the inequalities (3.24) given below, for which we define:

$$\nu^* := \lambda_f^{-1}\left(\frac{\alpha_x(\mu + \delta_\mu)}{2L_V^x}\right),$$

$$C_0 := \sqrt{2}M_{xz}(N_0/\sigma_1), \qquad C_{\mathfrak{n}}(\epsilon) := \sum_{\ell=1}^{\mathfrak{n}} \lambda_G^\ell(C_0\epsilon),$$

$$C_a(\epsilon) := C_0\epsilon + C_{\mathfrak{n}}(\epsilon), \qquad C_g(s) := \max\{s, \lambda_G^{\mathfrak{n}-1}(s)\},$$

and choose a positive integer N such that

$$\gamma_m^N \nu \leq \frac{1}{2}\underline{\alpha}_y\left(C_g^{-1}(\nu^*/2)\right). \tag{3.23}$$

Then, the inequalities for ϵ are:

$$0 < \epsilon \leq 1, \tag{3.24a}$$
$$L_V^z C_{\mathfrak{n}}(\epsilon) \leq (1 - \gamma_m)\nu, \tag{3.24b}$$
$$L_V^x M_x \mathfrak{n}\frac{N_0}{\sigma_1}\epsilon \leq \frac{\delta_\mu}{N}, \tag{3.24c}$$
$$L_V^x \lambda_f(\epsilon) < \frac{\alpha_x(\mu + \delta_\mu)}{2}, \tag{3.24d}$$
$$\frac{L_V^z C_{\mathfrak{n}}(\epsilon)}{1 - \gamma_m} \leq \frac{1}{2}\underline{\alpha}_y\left(C_g^{-1}(\nu^*/2)\right), \tag{3.24e}$$
$$C_a(\epsilon) \leq \frac{\nu^*}{2}. \tag{3.24f}$$

With such ϵ, our first claim is the following.

Claim 1: For any $k \geq 1$, suppose that $V_x(x(t_{J_{k-1}}, J_{k-1})) \leq \mu + \delta_\mu$ and $X(t_{J_{k-1}}, J_{k-1}) \in K_X$. If $V_x(x(t_{J_k}, J_k)) \leq \mu + \delta_\mu$, then $X(t_{J_k}, J_k) \in K_X$.

Proof of Claim 1: By the construction of Ω_X, (3.20) holds, and so, with (3.16), the evolution of (3.15) satisfies

$$|X(t, j) - X(t_j, j)| \leq C_0\epsilon \tag{3.25}$$

in the time interval $t_j \leq t \leq t_{j+1}$ for each j such that $J_{k-1} \leq j < J_k$. Now, with

$$\widetilde{Y}_k := X(t_{J_k}, J_k) - G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1})),$$

we claim that

$$\begin{aligned}|\widetilde{Y}_k| &= |X(t_{J_k}, J_k) - G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1}))| \\ &\leq \sum_{\ell=1}^{r_k} \lambda_G^\ell(C_0\epsilon) \leq C_{\mathfrak{n}}(\epsilon).\end{aligned} \tag{3.26}$$

The last inequality can be shown by induction. To see the underlying argument with notational simplicity, let us consider the case when $k = 1$. Note that

$$|X(t_1, 1) - G_1^0(X(0,0))| = |G_1^0(X(t_1,0)) - G_1^0(X(0,0))| \le \lambda_G(C_0\epsilon),$$

and with $r_1 \ge 2$,

$$\begin{aligned}|X(t_2, 2) - G_{2,1}^0(X(0,0))| &\le |G_2^0(X(t_2,1)) - G_2^0(X(t_1,1))| + |G_2^0(X(t_1,1)) - G_{2,1}^0(X(0,0))| \\ &= |G_2^0(X(t_2,1)) - G_2^0(X(t_1,1))| + |G_{2,1}^0(X(t_1,0)) - G_{2,1}^0(X(0,0))| \\ &\le \lambda_G(C_0\epsilon) + \lambda_G^2(C_0\epsilon).\end{aligned}$$

Repeating the arguments, we obtain that

$$|X(t_{J_1}, J_1) - G_{r_1,1}^0(X(0,0))| \le (\lambda_G + \lambda_G^2 + \cdots + \lambda_G^{r_1})(C_0\epsilon) \tag{3.27}$$

and (3.26) can be proved similarly. With (3.26) at hand, we see from Assumption 3.2, (3.18), and (3.24b), that

$$\begin{aligned}V_z(X(t_{J_k}, J_k), \mathsf{p}_1^k) &= V_z(G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1})) + \widetilde{Y}_k, \mathsf{p}_1^k) \\ &\le V_z(G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1})), \mathsf{p}_1^k) + L_V^z|\widetilde{Y}_k| \\ &\le \gamma_m V_z(X(t_{J_{k-1}}, J_{k-1}), \mathsf{p}_1^{k-1}) + L_V^z|\widetilde{Y}_k| \\ &\le \gamma_m \nu + L_V^z C_{\mathfrak{n}}(\epsilon) \le \nu.\end{aligned} \tag{3.28}$$

Therefore, considering the definition of K_x in (3.17) and K_X in (3.19), we conclude that $X(t_{J_k}, J_k) \in K_X$ under the assumption $V_x(x(t_{J_k}, J_k)) \le \mu + \delta_\mu$. //

Claim 2: With N chosen in (3.23),

$$V_x(x(t,j)) \le \mu + \delta_\mu, \quad \forall (0,0) \le (t,j) \le (t_{J_N}, J_N), \tag{3.29}$$

and for $k = 0, \ldots, N$,

$$|z(t_{J_k}, J_k)| \le \underline{\alpha}_y^{-1}\left(\gamma_m^k \nu + \frac{L_{V_z} C_{\mathfrak{n}}(\epsilon)}{1-\gamma_m}\right) \quad \text{and} \tag{3.30}$$

$$|z(t,j)| \le C_a(\epsilon) + C_g(|z(t_{J_k}, J_k)|), \qquad t_{J_k} \le t \le t_{J_{k+1}},\ J_k \le j < J_{k+1}. \tag{3.31}$$

Proof of Claim 2: We first note that, as long as $(x, z, \epsilon) \in \Omega_X \times [0,1]$,

$$|\dot{V}_x| = \left|\frac{\partial V_x}{\partial x}(x) f_x(x, z + h(x), \epsilon)\right| \le L_V^x M_x. \tag{3.32}$$

Therefore, it follows from $V_x(x(0,0)) \le \mu$ and $X(0,0) \in K_X$, and from (3.24c) and (3.16), that $V_x(x(t,j)) \le \mu + \delta_\mu/N$ for $(0,0) \le (t,j) \le (t_{J_1}, J_1)$. Moreover, from Claim 1, we also have that $X(t_{J_1}, J_1) \in K_X$. Repeating this argument, we get

$$V_x(x(t,j)) \le \mu + \frac{k}{N}\delta_\mu, \quad \forall (0,0) \le (t,j) \le (t_{J_k}, J_k),$$

which implies (3.29). Therefore, $X(t,j) \in \Omega_X$ for $(0,0) \le (t,j) \le (t_{J_N}, J_N)$.

In order to prove (3.30), we proceed further with (3.28) as

$$\begin{aligned}V_z(X(t_{J_k}, J_k), \mathsf{p}_1^k) &= V_z(G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1})) + \widetilde{Y}_k, \mathsf{p}_1^k) \\ &\le V_z(G_{r_k,1}^{k-1}(X(t_{J_{k-1}}, J_{k-1})), \mathsf{p}_1^k) + L_V^z|\widetilde{Y}_k|\end{aligned}$$

$$
\begin{aligned}
&\le \gamma_m V_z(X(t_{J_{k-1}}, J_{k-1}), \mathsf{p}_1^{k-1}) + L_V^z |\widetilde{Y}_k| \\
&\le \gamma_m V_z(G^{k-2}_{r_{k-2},1}(X(t_{J_{k-2}}, J_{k-2})), \mathsf{p}_1^{k-1}) + \gamma_m L_V^z |\widetilde{Y}_{k-1}| + L_V^z |\widetilde{Y}_k| \\
&\le \gamma_m^2 V_z(X(t_{J_{k-2}}, J_{k-2}), \mathsf{p}_1^{k-2}) + \gamma_m L_V^z |\widetilde{Y}_{k-1}| + L_V^z |\widetilde{Y}_k| \\
&\quad \vdots \\
&\le \gamma_m^{k-1} V_z(X(t_{J_1}, J_1), \mathsf{p}_1^1) + \sum_{\ell=0}^{k-2} \gamma_m^\ell L_V^z |\widetilde{Y}_{k-\ell}| \\
&\le \gamma_m^k V_z(X(0,0), \mathsf{p}_1^0) + \sum_{\ell=0}^{k-1} \gamma_m^\ell L_V^z |\widetilde{Y}_{k-\ell}|.
\end{aligned}
$$

This means that, since $\sum_{\ell=0}^\infty \gamma_m^\ell = 1/(1-\gamma_m)$ and $|\widetilde{Y}_k| \le C_{\mathfrak{n}}(\epsilon)$,

$$
\begin{aligned}
|z(t_{J_k}, J_k)| &\le \underline{\alpha}_y^{-1} \left(V_z(X(t_{J_k}, J_k), \mathsf{p}_1^k) \right) && (3.33) \\
&\le \underline{\alpha}_y^{-1} \left(\gamma_m^k V_z(X(0,0), \mathsf{p}_1^0) + \frac{L_V^z C_{\mathfrak{n}}(\epsilon)}{1-\gamma_m} \right) && (3.34) \\
&\le \underline{\alpha}_y^{-1} \left(\gamma_m^k \nu + \frac{L_V^z C_{\mathfrak{n}}(\epsilon)}{1-\gamma_m} \right), && (3.35)
\end{aligned}
$$

which yields (3.30).

While (3.35) indicates the decrease of z as k increases, let us look at the behavior of z where we may go outside the set P. For this, we note that for every $\mathsf{p} \in \mathcal{P}$, it holds that $\widetilde{g}(x, 0, \mathsf{p}) = 0$ and $|\widetilde{g}(x, z, \mathsf{p})| \le \lambda_G(|z|)$ for all $(x, z) \in \Omega_X$. Thus, it follows that

$$
|\widetilde{g}_r^k \circ G_{r-1}^k \circ \cdots \circ G_1^k(X)| \le \lambda_G^r(|z|), \quad 1 \le r \le r_k,
$$

for $X = (x, z) \in K_X$ because, e.g., $|\widetilde{g}_2^k(G_1^k(x, z))| \le \lambda_G(|\widetilde{g}_1^k(x, z)|) \le \lambda_G^2(|z|)$. Then, for $k = 0, \ldots, N$, if $X(t_{J_k}, J_k) \in K_X$, an upper bound of $|z(t, j)|$ for $t_j \le t \le t_{j+1}$, where $J_k \le j < J_{k+1}$, can be found as

$$
\begin{aligned}
|z(t, j)| &\le |z(t, j) - z(t_j, j)| + |z(t_j, j)| \\
&\le C_0 \epsilon + |z(t_j, j) - \widetilde{g}_j^k \circ G_{j-1}^k \circ \cdots \circ G_1^k(X(t_{J_k}, J_k))| + \lambda_G^j(|z(t_{J_k}, J_k)|).
\end{aligned}
$$

An upper bound of $|z(t_j, j) - \widetilde{g}_j^k \circ G_{j-1}^k \circ \cdots \circ G_1^k(X(t_{J_k}, J_k))|$ can be obtained similar to (3.27) as $(\lambda_G + \lambda_G^2 + \cdots + \lambda_G^j)(C_0\epsilon)$. Hence, for $t_j \le t \le t_{j+1}$ and $J_k \le j < J_{k+1}$,

$$
\begin{aligned}
|z(t, j)| &\le C_0 \epsilon + \sum_{\ell=1}^{j} \lambda_G^\ell(C_0 \epsilon) + \lambda_G^j(|z(t_{J_k}, J_k)|) \\
&\le C_a(\epsilon) + C_g(|z(t_{J_k}, J_k)|).
\end{aligned}
$$

Similar inequality holds for $t_{J_k} \le t \le t_{J_{k+1}}$ and $J_k \le j < J_{k+1}$ as long as $X(t_{J_k}, J_k) \in K_X$, which is (3.31). //

Claim 3: Thanks to (3.24d), (3.24e), and (3.24f), the inequality (3.29) holds for all $(t, j) \in \operatorname{dom} X$, and thus, inequalities (3.30) and (3.31) holds for all $k \ge 0$.

Proof of Claim 3: Let us look at the derivative of V_x on the set Ω_X with $\epsilon \in [0, 1]$, while recalling (3.22):

$$
\dot{V}_x = \nabla_x V_x \cdot f_x(x, h(x), 0) + \nabla_x V_x \cdot \widetilde{f}_x(x, z + h(x), \epsilon)
$$

$$\leq -\alpha_x(V_x(x)) + L_V^x \lambda_f(|z|) + L_V^x \lambda_f(\epsilon). \tag{3.36}$$

Then it is seen that, if ϵ satisfies (3.24d) and $|z| \leq \nu^*$, then $\dot{V}_x < 0$ on the level set $\{x : V_x(x) = \mu + \delta_\mu\}$. In fact, with (3.23) and (3.24e), the inequality (3.30) implies that $|z(t_{J_N}, J_N)| \leq C_g^{-1}(\nu^*/2) \leq \nu^*$, and from there, with (3.24f), the inequality (3.31) implies that $|z(t,j)| \leq \nu^*$ for $t_{J_N} \leq t \leq t_{J_{N+1}}$ and $J_N \leq j < J_{N+1}$. This implies that $V_x(x(t,j)) \leq \mu + \delta_\mu$ for $(0,0) \leq (t,j) \leq (t_{J_{N+1}}, J_{N+1})$, and this argument repeats. //

Putting altogether, we choose ϵ^* in Theorem 3.2 such that all the inequalities (3.24a)–(3.24f) hold with ϵ^*. Then, the following arguments lead us to the estimate in (3.7b).

For $(t,j) \in \operatorname{dom} X$, it follows from (3.16) that $j \geq (\sigma_1/N_0)(t/\epsilon)$, and therefore, for k such that $J_k \leq j < J_{k+1}$, we have

$$k \geq \frac{\sigma_1}{N_0 \mathfrak{n}} \frac{t}{\epsilon}. \tag{3.37}$$

Then, from (3.34), (3.30), (3.31), and Assumption 3.2, we get

$$|z(t,j)| \leq C_g \circ \underline{\alpha}_y^{-1} \left(\gamma_m^{\frac{\sigma_1}{N_0 \mathfrak{n}} \frac{t}{\epsilon}} \overline{\alpha}_y(|z(0,0)|) + \frac{L_V^z C_{\mathfrak{n}}(\epsilon)}{1-\gamma_m} \right) + C_a(\epsilon). \tag{3.38}$$

Consider the class $\mathcal{KL}$ function $\beta_y(s,\tau)$ in (3.13) and the class $\mathcal{K}$ function

$$\chi_y(\epsilon) := C_g \circ \underline{\alpha}_y^{-1} \left(2 \frac{L_V^z C_{\mathfrak{n}}(\epsilon)}{1-\gamma_m} \right) + C_a(\epsilon).$$

Then (3.38) yields the inequality in (3.7b) and the derivation of (3.13) in Theorem 3.2 is complete.

For deriving the inequality (3.12), we note that

$$\begin{aligned} \dot{V}_x &= \nabla_x V_x \cdot f_x(x, h(x), 0) + \nabla_x V_x \cdot \widetilde{f}_x(x, z+h(x), \epsilon) \\ &\leq -\alpha_x(V_x(x)) + L_V^x(\lambda_f(|z|) + \lambda_f(\epsilon)). \end{aligned} \tag{3.39}$$

The inequality (3.39) provides a Lyapunov characterization for ISS with z and ϵ as inputs, but we need to analyze it carefully to get the estimate on x-trajectory independently of the initial conditions on y-dynamics. From this point onwards, the derivation of (3.12) is carried out by using the bound from (3.38) in (3.39) and it involves (somewhat lengthy) arguments based on analysis of comparison functions. The remaining details are being skipped for keeping the presentation concise.

3.3 Network of Oscillators

As an application of Theorem 3.2, we study the problem of synchronization of coupled nonlinear oscillators. In particular, we consider N continuous-time controlled oscillators, which are connected to each other through an undirected graph. The control action for each oscillator is chosen as a function of the states of their neighbors at some discrete times. If this control is updated sufficiently often then we show that all the trajectories of the oscillators converge to each other in practical sense.

Each of the N oscillators (the agent) has a two dimensional state $(\omega_i, \xi_i) \in \mathbb{R}^2$, for $i = 1, \ldots, N$, whose dynamic equation is written in the form of Liénard system as

$$\begin{aligned} \dot{\omega}_i &= -\omega_i + \xi_i \\ \dot{\xi}_i &= \phi_i(\omega_i, \xi_i) := (1 - f_i(\omega_i))(-\omega_i + \xi_i) - g_i(\omega_i) \end{aligned} \tag{3.40}$$

where the functions f_i and g_i are assumed to be locally Lipschitz. For example, if the agent i has $f_i(\omega_i) = \mu_i(\omega_i^2 - 1)$ and $g_i(\omega_i) = \omega_i$ with a constant $\mu_i > 0$, then the agent i is a well-known Van der Pol oscillator [121]. If we look at an individual oscillator (3.40), it may have a stable or unstable limit cycle. In what follows, we consider a network of such oscillators, where each oscillator exchanges information about its state ξ only with its neighbors at some discrete times, and consequently, ξ gets reset whenever some new information is received from the neighbors. Our objective is to show that, with an appropriate reset rule, if the information between the neighbors is exchanged sufficiently often then all the oscillators in the network practically synchronize and converge practically to the same limit cycle.

3.3.1 Connectivity graphs

We consider a set of N vertices $\mathcal{V} = \{1, \cdots, N\}$, and let $\mathcal{Q}$ be the index set for different collections of edges associated to these nodes, that is, for each $q \in \mathcal{Q}$, $\mathcal{E}_q$ is a set of *unordered* pair of vertices that represents one possible collection of edges. An element in $\mathcal{E}_q$ is represented by (i, j), for some choice of $i, j \in \mathcal{V}$, $i \neq j$. We call $\mathcal{G}_q = (\mathcal{V}, \mathcal{E}_q)$ a *graph* with nodes $\mathcal{V}$ and edges $\mathcal{E}_q$. Here, the graph $\mathcal{G}_q$ is *undirected.* Agents i and j are *neighbors* of each other in graph $\mathcal{G}_q$ if $(i, j) \in \mathcal{E}_q$. The neighbors of a node $i \in \mathcal{V}$ in a graph $\mathcal{G}_q$ are denoted by $\mathcal{N}_i^q$. Let $\mathcal{S}$ be a set of words obtained from the alphabet $\mathcal{Q}$ and denote an element of $\mathcal{S}$ by $\mathfrak{s}$. We say that the graphs $\{\mathcal{G}_q\}_{q\in\mathfrak{s}}$ are *jointly connected* if $\mathcal{G}_\mathfrak{s} := (\mathcal{V}, \mathcal{E}_\mathfrak{s})$ is connected where $\mathcal{E}_\mathfrak{s} := \cup_{q\in\mathfrak{s}}\mathcal{E}_q$, in which '$q \in \mathfrak{s}$' implies $q \in \{\mathfrak{s}_1, \cdots, \mathfrak{s}_{|\mathfrak{s}|}\}$.

Assumption 3.3. The set $\mathcal{S}$ is finite, and for each $\mathfrak{s} \in \mathcal{S}$, the collection of graphs $\{\mathcal{G}_q\}_{q\in\mathfrak{s}}$ is jointly connected.[3]

In our setup, the continuous-time oscillators described in (3.40) represent the nodes which communicate with another neighboring oscillator only at some discrete time instants. The edges of a graph $\mathcal{G}_q$ represent communication link between the corresponding nodes. The motivation for introducing different graphs $\mathcal{G}_q$ is to allow for different communication patterns at different times among the oscillators.

With each $\mathcal{G}_q$, we associate a doubly stochastic matrix $\mathsf{W}_q \in \mathbb{R}^{N\times N}$. If at a given time instant, the graph $\mathcal{G}_q$ is active, then the state ξ_i is updated at that instant as a function of its neighbors' state ξ_j, $j \in \mathcal{N}_i^q$. In particular, we describe the update model as

$$\begin{aligned} \omega_i^+ &= \omega_i \\ \xi_i^+ &= \sum_{j\in\mathcal{N}_i^q} \mathsf{w}_{ij}^q \xi_j \end{aligned} \tag{3.41}$$

where w_{ij}^q is the (i, j)-th entry of W_q. We assume that the graph $\mathcal{G}_q$ contains the self-loop for every node, so that $i \in \mathcal{N}_i^q$, $\forall i$, and all the diagonal entries of W_q are positive.

Example 3.1. If the graph $\mathcal{G}_q$ is modeled by a symmetric adjacency matrix A_q, and D_q is the diagonal matrix that describes the degree of each node, then W_q can be given by $I_N - \varphi D_q + \varphi A_q$, where $\varphi > 0$ is chosen small enough such that W_q becomes a non-negative matrix. Then, every row and column sums of W_q is 1 by construction.

Example 3.2. Consider a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Let E denote the total number of edges in $\mathcal{E}$, and let $\mathcal{E}_q$ be a singleton which corresponds to a single edge contained

[3] We consider the case of an undirected graph for simple presentation, but the study in this section can be transparently extended to the case where $\{\mathcal{G}_q\}_{q\in\mathfrak{s}}$ is jointly strongly connected and balanced.

in $\mathcal{E}$. We let $\mathcal{Q} = \{1, \cdots, E\}$ and consider strings $\mathfrak{s}$ from the alphabet $\mathcal{Q}$ such that the collection of graphs $\{\mathcal{G}_q = (\mathcal{V}, \mathcal{E}_q)\}_{q\in\mathfrak{s}}$ is jointly connected. Under connectivity assumption on $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the set of such strings is nonempty and one possible choice is to take the strings $\mathfrak{s}$ as words of length E with distinct letters from the alphabet $\mathcal{Q}$. In this case, we can choose $\mathsf{W}_q \in \mathbb{R}^{N\times N}$, $q \in \mathcal{Q}$ to be a doubly stochastic matrix such that, for the index $(i, j) \in \mathcal{E}_q$,

$$\mathsf{w}^q_{ij} = \mathsf{w}^q_{ji} = \mathsf{w}^q_{ii} = \mathsf{w}^q_{jj} = \frac{1}{2}$$

and $\mathsf{w}^q_{kk} = 1$ for all $k \notin \{i, j\}$, and all other elements of W_q are zero.

Thus far, in this subsection, we have introduced a family of graphs $\mathcal{G}_q = (\mathcal{V}, \mathcal{E}_q)$ and a reset rule (3.41) for individual agents which depends on the neighbors described by $\mathcal{G}_q$. To specify the times at which an agent resets its state and to describe the index of the active graph, we consider a timer τ and a discrete state variable p, whose evolution is described as,

$$(\tau, \mathsf{p}) \in [0, N_0] \times \mathcal{P} : \begin{cases} \epsilon\dot{\tau} \in [\sigma_1, \sigma_2] \\ \dot{\mathsf{p}} = 0 \end{cases} \tag{3.42}$$

$$(\tau, \mathsf{p}) \in [1, N_0] \times \mathcal{P} : \begin{cases} \tau^+ = \tau - 1 \\ \mathsf{p}^+ \in G_{\mathsf{p}}(\mathsf{p}) \end{cases} \tag{3.43}$$

where N_0, σ_1, and σ_2 are some positive constants that determine the frequency at which the graphs switch from one configuration to another. The mapping G_{p} describes the evolution of discrete variables according to equation (3.14).

3.3.2 Coordinate transformation

We now collect the dynamics of individual agents from (3.40) and the connectivity structure from Section 3.3.1, to write them in the form of the hybrid system (3.3). This requires us to introduce a certain coordinate transformation for agents' dynamics which we describe next.

Let $\mathsf{R} \in \mathbb{R}^{N\times(N-1)}$ be a matrix such that

$$\mathsf{R}^\top \mathsf{R} = I, \quad \text{and} \quad \mathbf{1}_N^\top \mathsf{R} = 0,$$

where $\mathbf{1}_N \in \mathbb{R}^N$ is a vector with all entries equal to 1. Associated to each matrix W_q, let us introduce the matrix Λ_q as follows:

$$\Lambda_q := \mathsf{R}^\top \mathsf{W}_q \mathsf{R} \in \mathbb{R}^{(N-1)\times(N-1)}.$$

Next, we define a matrix Q as

$$\mathsf{Q}^{-1} := \begin{bmatrix} \mathbf{1}_N & \mathsf{R} \end{bmatrix}, \quad \text{so that} \quad \mathsf{Q} = \begin{bmatrix} \frac{1}{N}\mathbf{1}_N^\top \\ \mathsf{R}^\top \end{bmatrix}.$$

With the matrix Q, we can write (3.40) and (3.41) in the form of (3.3). Let us introduce the new coordinates (s, ζ) for $\xi := \mathrm{col}(\xi_1, \ldots, \xi_N) \in \mathbb{R}^N$ as

$$\begin{pmatrix} s \\ \zeta \end{pmatrix} := \mathsf{Q}\xi = \begin{bmatrix} \frac{1}{N}\mathbf{1}_N^\top \\ \mathsf{R}^\top \end{bmatrix} \xi \tag{3.44}$$

or, in other words

$$s = \frac{1}{N}\sum_{i=1}^{N}\xi_i \ \in \mathbb{R} \qquad \text{and} \qquad \zeta = \mathsf{R}^\top \xi \ \in \mathbb{R}^{N-1}.$$

It also follows that $\xi = \mathbf{1}_N s + \mathsf{R}\zeta$. Let R_i be the i-th row of R. Then, the system (3.40), (3.41) can be written as

$$\begin{aligned}
\dot\omega_i &= -\omega_i + s + \mathsf{R}_i\zeta, \qquad i = 1,\dots,N,\\
\dot s &= \frac{1}{N}\sum_{i=1}^{N}(1 - f_i(\omega_i))(-\omega_i + s + \mathsf{R}_i\zeta) - g_i(\omega_i)\\
\dot\zeta &= \mathsf{R}^\top\phi(\omega, \mathbf{1}_N s + \mathsf{R}\zeta)\\
\epsilon\dot\tau &\in [\sigma_1, \sigma_2]\\
\dot{\mathsf{p}} &= 0,
\end{aligned} \tag{3.45a}$$

where $\omega := \mathrm{col}(\omega_1, \dots, \omega_N)$ and

$$\phi(\omega, \mathbf{1}_N s + \mathsf{R}\zeta) = \begin{bmatrix} \phi_1(\omega_1, s + \mathsf{R}_1\zeta) \\ \vdots \\ \phi_N(\omega_N, s + \mathsf{R}_N\zeta) \end{bmatrix}$$

with the flow set $\mathcal{C} := \mathbb{R}^{2N} \times [0, N_0] \times \mathcal{P}$, and

$$\begin{aligned}
\omega_i^+ &= \omega_i, \quad \text{for} \quad i = 1,\dots,N\\
s^+ &= s\\
\zeta^+ &= \Lambda_{\mathsf{p}}\zeta\\
\tau^+ &= \tau - 1\\
\mathsf{p}^+ &\in G_{\mathsf{p}}(\mathsf{p})
\end{aligned} \tag{3.45b}$$

where Λ_{p} is the notation for $\Lambda_{\mathfrak{s}_\mathfrak{r}}$ considering $\mathsf{p} = (\mathfrak{s}, \mathfrak{r})$, with the jump set $\mathcal{D} := \mathbb{R}^{2N} \times [1, N_0] \times \mathcal{P}$. Here, if we treat $(\omega_1, \dots, \omega_N, s) = (\omega, s)$ as the slow variable and ζ as the fast variable with both the flow set and the jump set for (ω, s, ζ) being the same as $\mathbb{R}^{2N}$, then the system is in the form of (3.3).

3.3.3 Quasi-steady-state system

For stability analysis, we now consider the QSS system associated to (3.45). The QSS system is

$$\begin{aligned}
\dot\omega_i &= -\omega_i + s, \qquad i = 1,\dots,N,\\
\dot s &= \frac{1}{N}\sum_{i=1}^{N}(1 - f_i(\omega_i))(-\omega_i + s) - g_i(\omega_i),
\end{aligned} \tag{3.46}$$

which corresponds to (3.6). The following assumption and lemma assure that this QSS system satisfies the assumption of Theorem 3.2.

Assumption 3.4. The second order dynamical system

$$\begin{aligned}
\dot\chi_0 &= -\chi_0 + s_0\\
\dot s_0 &= (1 - f(\chi_0))(-\chi_0 + s_0) - g(\chi_0)
\end{aligned} \tag{3.47}$$

where

$$f(\chi_0) := \frac{1}{N}\sum_{i=1}^{N} f_i(\chi_0), \qquad g(\chi_0) := \frac{1}{N}\sum_{i=1}^{N} g_i(\chi_0) \tag{3.48}$$

has a locally asymptotically stable limit cycle $\mathcal{A}_0 \subset \mathbb{R}^2$.

The limit cycle $\mathcal{A}_0$ in Assumption 3.4 refers to an isolated periodic orbit, that is, every solution (χ_0, s_0) of system (3.47) with $(\chi_0(0), s_0(0)) \in \mathcal{A}_0$ is periodic, and $(\chi_0(t), s_0(t)) \in \mathcal{A}_0$, for all $t \geq 0$. Moreover, $\mathcal{A}_0$ is an isolated and compact set. Local asymptotic stability of $\mathcal{A}_0$ ensures that the nearby trajectories converge to this limit cycle in some neighborhood of $\mathcal{A}_0$. A sufficient condition on f and g for Assumption 3.4 can be found in [166, Sec. 3.8]. We use this assumption to show the existence of a stable limit cycle for system (3.46).

Lemma 3.3. *Under Assumption 3.4, the QSS system* (3.46) *has a locally asymptotically stable compact set $\mathcal{A}_x \subset \mathbb{R}^{N+1}$, which is a stable limit cycle for* (3.46), *given by*

$$\mathcal{A}_x = \{(\chi_0, \ldots, \chi_0, s_0) \in \mathbb{R}^{N+1} : (\chi_0, s_0) \in \mathcal{A}_0\}. \tag{3.49}$$

In particular, there exists V_x that satisfies (3.8) *and* (3.9) *on a neighborhood Ω_x of $\mathcal{A}_x$.*

Proof. The basic idea behind the proof is to use a coordinate transformation and combine it with the cascade arguments [191, Proposition 4.1]. Let

$$\chi := \frac{1}{N}\mathbf{1}_N^\top \omega \in \mathbb{R}, \quad \widetilde{\chi} := \mathsf{R}^\top \omega \in \mathbb{R}^{N-1}$$

so that $\omega_i = \chi + \mathsf{R}_i\widetilde{\chi}$. Then, system (3.46) can be rewritten as

$$\begin{aligned}
\dot{\widetilde{\chi}} &= -\widetilde{\chi}, \\
\dot{\chi} &= -\chi + s, \\
\dot{s} &= \frac{1}{N}\sum_{i=1}^{N}(1 - f_i(\chi + \mathsf{R}_i\widetilde{\chi}))(s - \chi - \mathsf{R}_i\widetilde{\chi}) - g_i(\chi + \mathsf{R}_i\widetilde{\chi}) \\
&= (1 - f(\chi))(-\chi + s) - g(\chi) + \frac{1}{N}\sum_{i=1}^{N}\psi_i(\widetilde{\chi}, \chi, s),
\end{aligned} \tag{3.50}$$

where ψ_i vanishes when $\widetilde{\chi} = 0$ because

$$\psi_i = f_i(\chi + \mathsf{R}_i\widetilde{\chi})\mathsf{R}_i\widetilde{\chi} + (f_i(\chi + \mathsf{R}_i\widetilde{\chi}) - f_i(\chi))(\chi - s) + (g_i(\chi) - g_i(\chi + \mathsf{R}_i\widetilde{\chi})).$$

Let 0_{N-1} denote the vector in $\mathbb{R}^{N-1}$ with all zero entries. Using Assumption 3.4, we next show that the set $\{0_{N-1}\} \times \mathcal{A}_0 \subset \mathbb{R}^{N+1}$ is locally asymptotically stable for (3.50), which corresponds to the local asymptotic stability of the set $\mathcal{A}_x$ for system (3.46). Indeed, due to Assumption 3.4, the set $\{0_{N-1}\} \times \mathcal{A}_0$ is locally asymptotically stable when the dynamics of (3.50) are restricted to the set $\{(\widetilde{\chi}, \chi, s) \in \mathbb{R}^{N+1} : \widetilde{\chi} = 0\}$. From [191, Theorem 2.24], it follows that $\{0_{N-1}\} \times \mathcal{A}_0$ is a stable compact set for system (3.50). To show that the set is also attractive, let $\Omega_0 \subset \mathbb{R}^2$ be the region of attraction for the set $\mathcal{A}_0$ of system (3.47), and let $\Omega_{\widetilde{\chi}} \subset \mathbb{R}^{N-1}$ be an arbitrary compact set containing the origin. Due to stability of $\{0_{N-1}\} \times \mathcal{A}_0 \subset \mathbb{R}^{N+1}$, there exists a set $\Omega_x \subset \mathbb{R}^{N+1}$ such that for all initial conditions in Ω_x, the solutions stay within the set $\Omega_{\widetilde{\chi}} \times \Omega_0$. With all solutions bounded, and $\widetilde{\chi}(t)$ converging to $\{0_{N-1}\}$ as $t \to \infty$, it follows that $(\widetilde{\chi}(t), \chi(t), s(t))$ converges to the largest invariant set contained in $\{0_{N-1}\} \times \Omega_0$ which is the set $\{0_{N-1}\} \times \mathcal{A}_0$. The existence of a continuously differentiable Lyapunov function follows from classical converse arguments. Finally, it is seen from (3.49) that the behavior of (3.46) on $\mathcal{A}_x$ is governed by the behavior of (3.47) on $\mathcal{A}_0$, which is a periodic motion. Therefore, $\mathcal{A}_x$ is a stable limit cycle. □

Remark 3.1. In Lemma 3.3, we did not provide an explicit construction of the Lyapunov function V_x for the QSS system (3.46). However, in the proof we saw that the asymptotic stability of the set $\mathcal{A}_x$ is essentially established by using a cascade argument between (χ, s) dynamics and $\widetilde{\chi}$ dynamics. In some instances, one can explicitly provide a construction of Lyapunov function V_x from the individual Lyapunov functions associated with (χ, s)-dynamics with respect to $\mathcal{A}_0$, and $\widetilde{\chi}$ dynamics with respect to $\{0_{N-1}\}$. In our conference paper [212], we provided such a construction under the assumption that $\mathcal{A}_0$ is locally *exponentially* stable for system (3.47). △

3.3.4 Boundary-layer system

The boundary-layer system associated to (3.45) is

$$\begin{aligned} \omega_i^+ &= \omega_i, \qquad \text{for} \;\; i = 1, \dots, N \\ s^+ &= s \\ \zeta^+ &= \Lambda_{\mathsf{p}} \zeta \\ \mathsf{p}^+ &\in G_{\mathsf{p}}(\mathsf{p}). \end{aligned} \tag{3.51}$$

We note that ζ-dynamics are completely decoupled from (ω_i, s) for each $i = 1, \dots, N$. Assumption 3.2, with $h \equiv 0$, therefore corresponds to checking the stability of (ζ, p) dynamics with respect to the set $\{0_{N-1}\} \times \mathcal{P}$. This follows from the following lemma.

Lemma 3.4. *Under Assumption 3.3, Assumption 3.2 holds for the boundary-layer system* (3.51) *with*

$$V_y(\omega, s, \zeta, \mathsf{p}) = \zeta^\top \zeta$$

and with $\gamma_m = \max_{\mathfrak{s} \in \mathcal{S}} \overline{\lambda}_{\mathfrak{s}}^2$, *where* $\overline{\lambda}_{\mathfrak{s}}$ *denotes the largest singular value of the matrix* $\overline{\Lambda}_{\mathfrak{s}} := \Lambda_{\mathfrak{s}_{|\mathfrak{s}|}} \cdots \Lambda_{\mathfrak{s}_1}$.

Proof. Since $\{\mathcal{G}_q\}_{q \in \mathfrak{s}}$ is jointly connected and the matrix W_q is associated with the graph $\mathcal{G}_q$, it follows that

$$\mathsf{W}_{\mathfrak{s}_{|\mathfrak{s}|}} + \cdots + \mathsf{W}_{\mathfrak{s}_1}$$

is a primitive matrix. On the other hand, by [108, Lemma 2], it holds with $c_{\mathfrak{s}} > 0$ that

$$\overline{\mathsf{W}}_{\mathfrak{s}} := \mathsf{W}_{\mathfrak{s}_{|\mathfrak{s}|}} \cdots \mathsf{W}_{\mathfrak{s}_1} \succcurlyeq c_{\mathfrak{s}} (\mathsf{W}_{\mathfrak{s}_{|\mathfrak{s}|}} + \cdots + \mathsf{W}_{\mathfrak{s}_1}), \tag{3.52}$$

where $\succcurlyeq$ implies component-wise inequality. This implies $\overline{\mathsf{W}}_{\mathfrak{s}}$ is primitive (because $A \succcurlyeq B$ with a primitive matrix B implies that A is primitive). Moreover, we have

$$\begin{aligned} \mathsf{Q} \overline{\mathsf{W}}_{\mathfrak{s}} \mathsf{Q}^{-1} &= \mathsf{Q} \mathsf{W}_{\mathfrak{s}_{|\mathfrak{s}|}} \mathsf{Q}^{-1} \cdots \mathsf{Q} \mathsf{W}_{\mathfrak{s}_2} \mathsf{Q}^{-1} \mathsf{Q} \mathsf{W}_{\mathfrak{s}_1} \mathsf{Q}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \Lambda_{\mathfrak{s}_{|\mathfrak{s}|}} \cdots \Lambda_{\mathfrak{s}_1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \overline{\Lambda}_{\mathfrak{s}} \end{bmatrix}. \end{aligned}$$

Now we note that $\overline{\mathsf{W}}_{\mathfrak{s}}$ is primitive and doubly stochastic whose diagonal elements are all positive. This implies that $\lim_{k \to \infty} \overline{\mathsf{W}}_{\mathfrak{s}}^k = \mathbf{1}_N c^T$ with a vector $c \in \mathbb{R}^N$ [29]. Since the matrix $\mathbf{1}_N c^T$ has rank one, the matrix $\lim_{k \to \infty} \overline{\Lambda}_{\mathfrak{s}}^k$ should be the zero matrix; that is, $\|\overline{\Lambda}_{\mathfrak{s}}\|_2 < 1$ where $\| \cdot \|_2$ represents the induced-two-norm, and thus, $\gamma_m < 1$. This in turn implies that $\zeta^\top \overline{\Lambda}_{\mathfrak{s}}^\top \overline{\Lambda}_{\mathfrak{s}} \zeta \leq \gamma_m \zeta^\top \zeta$, for all $\mathfrak{s} \in \mathcal{S}$, and Assumption 3.2 holds. □

3.3.5 Stability result

Combining the statements of Lemma 3.3 and Lemma 3.4, we arrive at the following result about synchronization of trajectories modeled by (3.45).

Theorem 3.5. *Suppose that Assumptions 3.3 and 3.4 hold and consider a compact set $I_x \subset \Omega_x$ and a compact set $I_y \subset \mathbb{R}^{N-1}$, for which $(\omega(0,0), (1/N)\sum_{i=1}^N \xi_i(0,0)) \in I_x$ and $\mathsf{R}^\top \xi(0,0) \in I_y$. Then, there are $\epsilon^* > 0$, class $\mathcal{KL}$ functions β_x and β_y, and class $\mathcal{K}$ functions χ_x and χ_y such that*

$$|(\omega(t,j), s(t,j))|_{\mathcal{A}_x} \leq \beta_x(|(\omega(0,0), s(0,0))|_{\mathcal{A}_x}, t) + \chi_x(\epsilon) \tag{3.53}$$

$$|\zeta(t,j)| \leq \beta_y(|\zeta(0,0)|, t/\epsilon) + \chi_y(\epsilon), \tag{3.54}$$

for all $\epsilon \in (0, \epsilon^*]$ and $(t,j) \in$ dom (ω, s, ζ)*. In particular, for each $\rho > 0$, there is ϵ^* such that, for each $\epsilon \in (0, \epsilon^*]$, the solutions to the multi-agent system* (3.40) *with* (3.41) *satisfy, for all $i, j = 1, \ldots, N$,*

$$\limsup_{t\to\infty} |(\omega_i(t,j), \xi_i(t,j)) - (\omega_j(t,j), \xi_j(t,j))| \leq \rho \tag{3.55}$$

$$\limsup_{t\to\infty} |(\omega_i(t,j), \xi_i(t,j))|_{\mathcal{A}_0} \leq \rho \tag{3.56}$$

if dom (ω, ξ) *is unbounded in t-axis.*

Proof. Proof of (3.53) and (3.54) is a consequence of Theorem 3.2 by taking $I_{xy} = I_x \times I_y$. To show (3.55), we note that $\xi = \mathbf{1}_N s + \mathsf{R}\zeta$. For any i and j, we have that $|\xi_i - \xi_j| \leq |\xi_i - s| + |s - \xi_j| = |\mathsf{R}_i\zeta| + |\mathsf{R}_j\zeta|$. This in turn implies from (3.54) that $\limsup_{t\to\infty} |\xi_i(t,j) - \xi_j(t,j)|$ can be made arbitrarily small by taking ϵ^* sufficiently small. Since $\limsup_{t\to\infty} |(\omega_1, \ldots, \omega_N, s)|_{\mathcal{A}_x}$ can be made arbitrarily small by taking sufficiently small ϵ^* by (3.53), we infer that $\limsup_{t\to\infty} |\omega_i(t,j) - \omega_j(t,j)|$ can also be made arbitrarily small due to the definition of the set $\mathcal{A}_x$. This proves (3.55). The inequality in (3.56) is also a direct consequence of the foregoing analysis and the definition of the set $\mathcal{A}_x$. □

Remark 3.2. It can be seen from the proof of Theorem 3.2 that, in the current case-study, the function C_g is given by $C_g(s) = C_\Lambda s$ with $C_\Lambda := \max_{q\in\mathcal{Q}} \|\text{blockdiag}(I_{N+1}, \Lambda_q^{\mathfrak{n}-1})\|$. Combined with the fact that the functions V_y are quadratic, the inequality (3.54) takes the form

$$|\zeta(t,j)| \leq \sqrt{2} C_\Lambda \gamma_m^{\frac{\sigma_1}{2N_0 \mathfrak{n}} \frac{t}{\epsilon}} |\zeta(0,0)| + \chi_y(\epsilon)$$

with some class $\mathcal{K}$ function χ_y. Also, from the proof, we see that, $|\xi_i - \xi_j| \leq |\mathsf{R}_i\zeta| + |\mathsf{R}_j\zeta| \leq 2\max_{1\leq i\leq N} \|R_i\| \cdot |\zeta|$, which implies that the rate of synchronization for ξ_i is much faster than that for ω_i when ϵ is small enough. This behavior can be seen in the bottom left and bottom right plots of Fig. 3.1. △

Implication of (3.55) is that the coupled oscillator (3.40) practically synchronizes under the impulsive coupling (3.41). Moreover, (3.56) implies that the limit cycle $\mathcal{A}_0$ is practically attractive for each oscillator under the coupling. In fact, the limit cycle $\mathcal{A}_0$ is emergent and collective one for all the oscillators in the sense that it is the limit cycle of the blended dynamics (3.47) with (3.48). This can be seen from the illustrations reported in the following example.

Example 3.3. We have simulated[4] a network with four Van der Pol oscillators described by (3.40) with $f_i(\omega_i) = \mu_i(\omega_i^2 - 1)$ where $\mu_1 = -.1$, $\mu_2 = 1$, $\mu_3 = 2$, and

[4]Source code for the simulation in Example 3.3 is available at `https://github.com/aneeltanwani/coupled_osc_switching.git`.

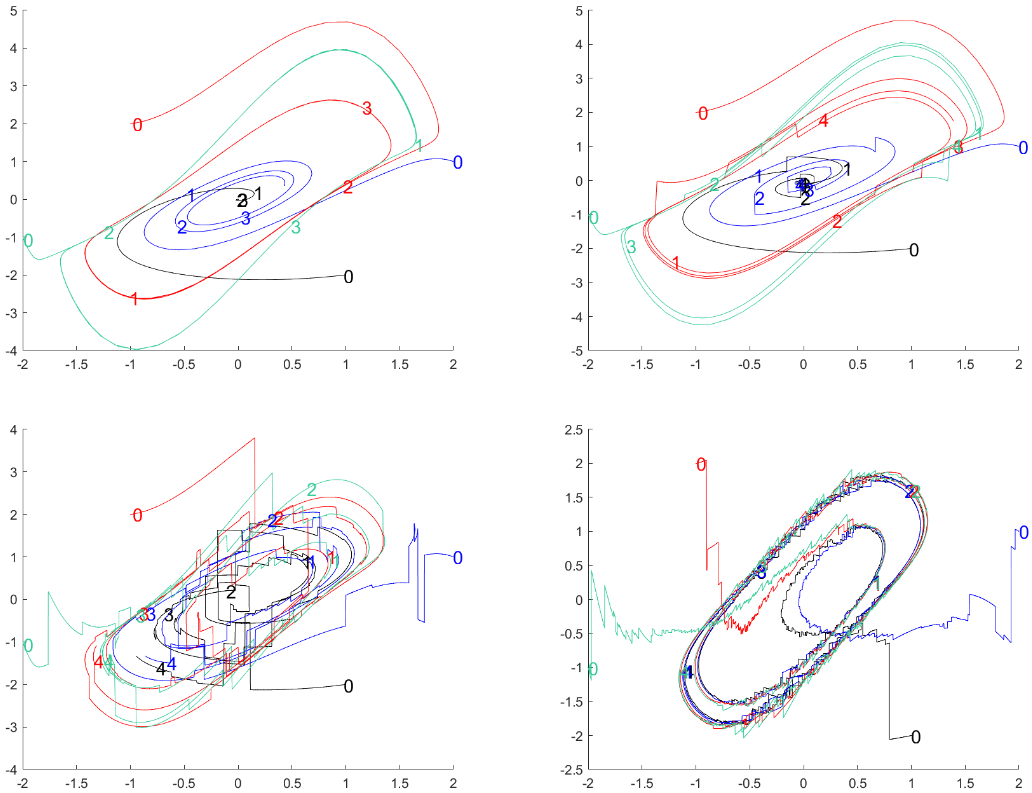


Figure 3.1: Plot of $(\omega_i(t), \xi_i(t))$ for four agents. Numbers are time stamps in increasing order, and all trajectories circulate clock-wise. (Top left): Case of no coupling; Started from four different initial conditions (time stamps '0'), their motions are not synchronized, and the black agent (agent 1) does not have a stable limit cycle so that it converges to the origin. (Top right): Coupled with $\epsilon = 1$; Jumps are observed but synchrony is weak. (Bottom left): Coupled with $\epsilon = 0.1$. The overlapped numbers imply that they are synchronized to some extent. (Bottom right): Coupled with $\epsilon = 0.01$: approximate synchronization.

$\mu_4 = 3$, and $g_i(\omega_i) = \omega_i$ for $i = 1, \ldots, 4$. The oscillators are connected through an undirected cyclic (and hence connected) graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, 3, 4\}$ and $\mathcal{E} = \{(1,2), (2,3), (3,4), (4,1)\}$. To define the reset rule for updating the state ξ of the oscillators according to (3.41), we split the graph into four subgraphs $\mathcal{G}_q = (\mathcal{V}, \mathcal{E}_q)$, $q \in \mathcal{Q} = \{1, 2, 3, 4\}$, with $\mathcal{E}_1 = \{(1,2)\}$, $\mathcal{E}_2 = \{(2,3)\}$, $\mathcal{E}_3 = \{(3,4)\}$, $\mathcal{E}_4 = \{(4,1)\}$. The set $\mathcal{S}$ comprises strings of length $\mathfrak{n} \geq 4$ such that each string contains at least a single occurrence of the letters 1, 2, 3, and 4. The weight matrix W_q is constructed as in Example 3.2. For the timer, we work with, $\sigma_1 = 0.5$, $\sigma_2 = 1.5$, $N_0 = 2$, and study different cases for $\epsilon \in \{1, 0.1, 0.01\}$. The state trajectories of all the agents are plotted in Fig. 3.1 for three different values of ϵ. Our proposed reset rule only updates the state ξ_i of the i-th agent whenever an edge connecting this agent with one of its neighbors is activated. Also, as we make ϵ smaller, we see that all the agents converge to the same limit cycle and the distance between their states gets smaller as well.

3.4 Perspectives

Motivated by the applications of hybrid dynamical systems where frequent jumps are necessary for stability, we propose analysis of such systems based on singular perturbation viewpoint. The dynamics are decomposed into a continuous-time quasi-steady-state system and a discrete-time boundary-layer system. Under the assumption that the subsystems resulting from this decomposition satisfy appropriate stability assumptions, we can derive estimates on the norm of the trajectories showing practical convergence. In the process, we see the trade-offs between the ultimate bound and the parameter describing the frequency of jumps. An important application of our results is seen in the study of network control with heterogenous agents where the graph describing the information exchange between agents is switching.

Several interesting research directions may emanate from the framework adopted in this work. Firstly, it may be possible to improve the estimates on system trajectories reported in Theorem 3.2 since the proof involves computation of certain constants over the compact sets for the general class of nonlinear systems. For some specific system class such as linear systems, or globally Lipschitz dynamics, one may modify the proof techniques for getting tighter estimates. Also, it may be possible to identify conditions on system data under which we can establish asymptotic stability for the limiting value of the perturbation parameter. Naturally, these questions open up the possibility of investigating more applications where the proposed setup of singularly perturbed hybrid systems is useful for analysis. This includes especially the control design problem related to stabilization over networks where frequent jumps in the state of the system contribute to stability of the system.

4

Maximally Monotone Systems

In this chapter, we make the transition towards a particular class of dynamical systems involving maximal monotone mappings, and broadly speaking, these systems are described as,

$$\dot{x} \in f(x) - \mathcal{M}(x), \qquad x(0) \in \mathrm{dom}(\mathcal{M}), \tag{4.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is sometimes called the drift term, $\mathcal{M} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone mapping, and $\mathrm{dom}(\mathcal{M})$ is the effective domain of $\mathcal{M}$, that is, $\{x \in \mathbb{R}^n \,|\, \mathcal{M}(x) \neq \emptyset\}$. For several applications, and some of the results presented later in this chapter, we would often look at the case where the drift term is not present. The structure of interest in these systems arises due to the presence of maximal monotone mapping $\mathcal{M}$, which are defined as follows:

Definition 4.1 (Maximal monotone mappings)**.** Consider a set-valued map $\mathcal{M} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. We say that $\mathcal{M}$ is *monotone* if it has the property that

$$\langle y_1 - y_0, x_1 - x_0 \rangle \geq 0, \quad \text{for all} \;\; y_0 \in \mathcal{M}(x_0), y_1 \in \mathcal{M}(x_1). \tag{4.2}$$

We say that $\mathcal{M}$ is *maximal monotone* if no enlargement of its graph is possible in $\mathbb{R}^n \times \mathbb{R}^n$ without destroying monotonicity. In other words, for every pair $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^n \setminus \mathrm{graph}(\mathcal{M})$, there exists $(\tilde{x}, \tilde{y}) \in \mathrm{graph}(\mathcal{M})$ with $\langle \hat{y} - \tilde{y}, \hat{x} - \tilde{x} \rangle < 0$.

Furthermore, we say that $\mathcal{M} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is α-strongly maximal monotone, if it is maximal monotone and

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq \alpha |x_1 - x_2|^2, \quad \forall\, y_i \in \mathcal{M}(x_i), i = 1, 2. \tag{4.3}$$

Intuitively speaking, maximal monotone mappings are set-valued mappings which satisfy a monotonicity property (as one observes in several control systems), and in addition, the maximality refers to some sort of regularity implying that the graph of $\mathcal{M}$ does not have any “holes” in it. These two properties provide the necessary ingredients for proving the existence and uniqueness of solutions for system (4.1); a classic treatment of such fundamental issues can be found in [24]. Moreover, for α-strongly maximal monotone systems, we immediately obtain the following boundedness/convergence property for the solutions of system (4.1):

Proposition 4.2. *Consider the solutions $x_1(\cdot)$, $x_2(\cdot)$ to system* (4.1) *with $f \equiv 0$ and $x_1(0), x_2(0) \in \mathrm{dom}(\mathcal{M})$. Then, for all $t \geq 0$, we have*

$$|x_1(t) - x_2(t)| \leq e^{-\alpha t} |x_1(0) - x_2(0)|.$$

Proof. Let us compute the time derivate of $|x_1(t) - x_2(t)|^2$ as follows:

$$\begin{aligned}\frac{d}{dt}|x_1(t) - x_2(t)|^2 &= \frac{d}{dt}\langle x_1(t) - x_2(t), x_1(t) - x_2(t)\rangle \\ &= 2\langle \dot{x}_1(t) - \dot{x}_2(t), x_1(t) - x_2(t)\rangle \\ &\leq -2\alpha|x_1(t) - x_2(t)|^2\end{aligned}$$

and hence $|x_1(t) - x_2(t)|^2 \leq e^{-2\alpha t}|x_1(0) - x_2(0)|^2$. □

Thus, if the right-hand side of a differential inclusion has a strongly maximal monotone mapping, with a negative sign, all trajectories asymptotically converge to each other regardless of the initial condition.

4.1 Motivation Based on Constrained Optimization

To see a motivation for the system class (4.1) in engineering and optimization related problems, let us consider the following optimization problem:

$$\begin{aligned} \min_x \quad & g(x) \\ \text{subject to} \quad & x \in \mathcal{C} \end{aligned}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable convex function, and $\mathcal{C} \subset \mathbb{R}^n$ is a closed convex set. We can write the constrained optimization problem as an unconstrained optimization problem

$$\min_x \quad g(x) + \psi_{\mathcal{C}}(x)$$

where $\psi_{\mathcal{C}}$ is an indicator function associated with the set $\mathcal{C}$, and is defined as,

$$\psi_{\mathcal{C}}(x) := \begin{cases} 0, & \text{if } x \in \mathcal{C}, \\ +\infty, & \text{if } x \notin \mathcal{C}. \end{cases} \tag{4.4}$$

The *proximal gradient method* for solving this optimization problem is

$$x^{k+1} := \operatorname{prox}_{\epsilon\psi_{\mathcal{C}}}(x^k - \epsilon\nabla g(x^k)) = \operatorname{proj}_{\mathcal{C}}(x^k - \epsilon\nabla g(x^k)). \tag{4.5}$$

Here, $\operatorname{prox}_{\epsilon h}$ denotes the proximal operator for the scaled function ϵh, for $\epsilon > 0$, that is,

$$\operatorname{prox}_{\epsilon h}(v) = \arg\min_x \left\{ h(x) + \frac{1}{2\epsilon}\|x - v\|^2 \right\}.$$

When taking $h = \psi_{\mathcal{C}}$, we observe that $\operatorname{prox}_{\epsilon h}(v)$ equals the projection of v on the set $\mathcal{C}$, denoted by $\operatorname{proj}_{\mathcal{C}}(v)$.

We can characterize the optimal solution using Karush-Kuhn-Tucker (KKT) conditions. That is, if $x^\star$ is the minimizer, then $0 \in \nabla g(x^\star) + \partial\psi_{\mathcal{C}}(x^\star)$. In other words, $x^\star$ is the stationary point of the differential inclusion:

$$\dot{x} \in -\nabla g(x) - \partial\psi_{\mathcal{C}}(x). \tag{4.6}$$

Here, $\partial\psi_{\mathcal{C}}$ denotes the subdifferential of $\psi_{\mathcal{C}}$, which is defined later in (4.8).

Thus, differential inclusion (4.6) represents the generalized gradient descent algorithm. When g is a convex function, and the constraint set $\mathcal{C}$ is also convex, it can be checked that the mapping $(\nabla g + \partial\psi_{\mathcal{C}})(\cdot)$ is maximal monotone and presents a special case of (4.1). Algorithm (4.5) can be seen as one possible time-discretization of the differential inclusion (4.6) and the first question of interest is to study whether the solutions of (4.6) can be approximated by algorithms of the form (4.5) with appropriate interpolation and small enough step size. Secondly, a natural question of interest is to study the convergence of solutions of (4.6) to a stationary point, which corresponds to studying the asymptotic stability of the system. We will discuss these two directions in the following sections.

Relation to complementarity systems: As a special case, let us consider the following quadratic program

$$\begin{aligned} \min_x \quad & \frac{1}{2}x^\top Ax + b^\top x \\ \text{subject to} \quad & Cx \geq c \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, and $C \in \mathbb{R}^{m\times n}$. We assume that A is symmetric. In addition, if A is positive-semidefinite, then the cost function is convex, and we have a convex quadratic program.

Using KKT conditions, if $x^\star$ is a locally optimal solution of the quadratic program, then there exists $\lambda^\star \in \mathbb{R}^m$ such that

$$\begin{gathered} Ax^\star + b - C^\top\lambda^\star = 0, \\ \lambda^\star \geq 0, \qquad Cx^\star - c \geq 0, \qquad \lambda^{\star\top}(Cx^\star - c) = 0. \end{gathered}$$

Such points can be described as the stationary points of the linear complementarity system

$$\begin{aligned} &\dot{x} = -Ax - b + C^\top\lambda \\ &0 \leq \lambda \perp Cx - c \geq 0, \end{aligned}$$

which is actually a particular case of (4.6). We will study formal connections between the differential inclusion (4.6) and the complementarity systems in the following sections.

Gradient descent with coupled variables: Consider the following constrained optimization problem:

$$\min_{x\in\mathcal{X}} g(x)$$

where $\mathbb{R}^n \supseteq \mathcal{X} \triangleq \Omega \cap \{x \,|\, Ax \leq b\}$, for some $A \in \mathbb{R}^{m\times n}$. We assume that $g : \mathbb{R}^n \to \mathbb{R}$ is strictly convex and $\Omega \subseteq \mathbb{R}^n$ is convex, hence $\mathcal{X}$ is convex. Using the KKT conditions, the optimization problem is solved by finding a pair (x^*, λ^*) such that

$$\begin{aligned} &0 \in \nabla g(x^*) + \mathcal{N}_\Omega(x^*) + A^\top\lambda^* \\ &0 \in b + \mathcal{N}_{\mathbb{R}^m_+}(\lambda^*) - Ax^*, \end{aligned}$$

where the notation $\mathcal{N}_S(x)$ denotes the outward normal cone to a closed convex set S at $x \in S$. We can study this problem by considering the solution of the following dynamical system and analyzing the convergence of the state variables to the origin,

$$\begin{aligned} &\dot{x} \in -(\nabla g(x) + \mathcal{N}_\Omega(x)) - A^\top\lambda \\ &\dot{\lambda} \in -(b + \mathcal{N}_{\mathbb{R}^m_+}(\lambda)) + Ax. \end{aligned} \tag{4.7}$$

Note that, due to convexity assumption, the operators $(\nabla g(x) + \mathcal{N}_\Omega(x))$ and $(b + \mathcal{N}_{\mathbb{R}_+}(\lambda))$ are both maximal monotone. System (4.7) therefore defines an interconnection of x- and λ-dynamics, and each of these dynamics contain a maximal monotone operator and a coupling term. Using the properties of the operators that describe subsystem dynamics, and the interconnection term, we would like to know if the overall system is maximal monotone. For this example, the answer is indeed affirmative as we will see from later developments.

4.2 Nonsmooth Relations

For the system class considered in (4.1), and as we saw in the motivating examples in Section 4.1, the set-valued maximal monotone mapping often arises from taking the subdifferential of convex functions, and could thus be related to a nonsmooth mapping in the system description. In this section, let us explore some instances of nonsmooth relations with which we can associate the maximal monotone property. There are several books which treat nonsmooth, convex, and variational analysis in depth [180, 152, 104]. We will focus on a particular class of nonsmooth relations which are mainly covered in [54, 74].

Let us start with defining a generalization of the derivative operator for convex functions, which are not differentiable in the conventional sense.

Definition 4.3 (Proper, lower semicontinuous functions)**.** Consider an extended real-valued function $f : \mathbb{R}^n \to [-\infty, \infty]$. The *effective domain* of f is $\mathrm{dom}(f) := \{x \in \mathbb{R}^n \,|\, f(x) < +\infty\}$. We say that f is proper if $f(x) < +\infty$ for at least one $x \in \mathbb{R}^n$ and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$. The function f is called *lower semicontinuous* at a point $x \in \mathbb{R}^n$ if $f(x) \leq \liminf_{k\to\infty} f(x_k)$ for every sequence $\{x_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ with $x_k \to x$.

Definition 4.4 (Subdifferential of a convex function)**.** Consider an extended real-valued function $f : \mathbb{R}^n \to [-\infty, \infty]$, which is proper. Furthermore, assume that[1] f is *convex.* We say that v is a subgradient of f at $x \in \mathrm{dom}(f)$ if

$$\langle v, y - x\rangle \leq f(y) - f(x), \qquad \text{for all } y \in \mathbb{R}^n$$

The subdifferential of f at $x \in \mathrm{dom}(f)$ is $\partial f(x) := \{v \in \mathbb{R}^n \mid v \text{ is a subgradient of } f \text{ at } x\}$. We say that f is subdifferentiable at $x \in \mathbb{R}^n$ if $\partial f(x) \neq \emptyset$.

As an illustration of this definition, let us consider a closed convex set $\mathcal{C} \subset \mathbb{R}^n$, with its indicator function $\psi_{\mathcal{C}}$ as in (4.4). If $\mathcal{C}$ is nonempty, then $\psi_{\mathcal{C}}$ is proper, and lower semicontinuous. It can be readily checked that

$$\partial\psi_{\mathcal{C}}(x) = \{v \in \mathbb{R}^n \mid \langle v, y - x\rangle \leq 0, \text{ for all } y \in \mathcal{C}\} = \mathcal{N}_{\mathcal{C}}(x). \tag{4.8}$$

That is, the set-valued operator $x \mapsto \mathcal{N}_{\mathcal{C}}(x)$ is the subdifferential of the discontinuous function $\psi_{\mathcal{C}}$ at x, that is, the outward normal cone to a set equals the subdifferential of the indicator function of that set. Let us now relate this nonsmooth relation $x \mapsto \mathcal{N}_{\mathcal{C}}(x)$ with some other frequently occurring notions.

Definition 4.5 (Variational Inequality)**.** Given a set $\mathcal{C} \subset \mathbb{R}^n$ and a mapping $F : \mathcal{C} \to \mathbb{R}^n$, the variational inequality, denoted by $\mathrm{VI}(\mathcal{C}, F)$, is to find a vector $x \in \mathcal{C}$ such that

$$\langle y - x, F(x)\rangle \geq 0, \qquad \forall\, y \in \mathcal{C}.$$

It follows from the definition that x solves $\mathrm{VI}(\mathcal{C}, F)$, if and only if,

$$F(x) \in -\mathcal{N}_{\mathcal{C}}(x).$$

Definition 4.6 (Complementarity Problem)**.** Given a cone $K \subset \mathbb{R}^n$ and a mapping $F : K \to \mathbb{R}^n$, the *complementarity problem*, denoted by $\mathrm{CP}(K, F)$, is to find a vector $\lambda \in K^\star$ satisfying the following conditions:

$$K \ni \lambda \perp F(\lambda) \in K^\star,$$

[1] Note that, for a *proper* function f, we say that f is convex if for every $\alpha \in [0, 1]$, $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$, for all $x, y \in \mathbb{R}^n$.

where $K^\star$ is the dual cone of K, and is defined as $K^\star := \{w \in \mathbb{R}^n \mid \langle w, \lambda \rangle \geq 0, \text{ for all } \lambda \in K\}$. If $F(\lambda) = M\lambda + q$ for some $M \in \mathbb{R}^{n\times n}$, $q \in \mathbb{R}^n$, and $K = \mathbb{R}^n_{\geq 0}$ is the positive orthant of $\mathbb{R}^n$, then we call it a *linear complementarity problem*, which is denoted by $\mathrm{LCP}(M, q)$.

Proposition 4.7. *Let $K \subset \mathbb{R}^n$ be a cone. A vector x solves* $\mathrm{VI}(K, F)$ *if and only if x solves* $\mathrm{CP}(K, F)$.

Thus, we have seen some instances of nonsmooth relations that can be seen as maximal monotone mappings. These include the normal cone mapping, subdifferential of indicator functions, variational inequalities, and complementarity problems. The foregoing provided some discussions about some connections between these relations.

4.2.1 Modeling piecewise linear relations

Piecewise linear relations, such as, saturation or quantization, appear quite often in the modeling of control systems. Many of these nonlinearities have a certain structure that can be exploited for analysis. In this section, we will show that we can model piecewise linear relations using the framework of linear complementarity problems (LCP) and this allows us to study the properties of nonlinearities in a systematic manner. Also, one of the advantage of using LCP framework is that there are several numerical solvers available now for solving complementarity relations.

Definition 4.8. Consider a collection of open sets $\mathcal{O}_i \subset \mathbb{R}^n$, $i \in \mathcal{I}$, for some finite index set $\mathcal{I}$, with the properties that $\mathcal{O}_i \cap \mathcal{O}_j = \emptyset$, and $\cup_{i\in\mathcal{I}}\overline{\mathcal{O}}_i = \mathbb{R}^n$. We say that a *continuous* function $f : \mathbb{R}^n \to \mathbb{R}^m$ is piecewise linear (PWL) with respect to $\{\mathcal{O}_i\}_{i\in\mathcal{I}}$ if

$$f_i(x) = A_i x + b_i, \quad x \in \mathcal{O}_i$$

for some matrices $A_i \in \mathbb{R}^{m\times n}$, and vectors $b_i \in \mathbb{R}^m$.

On the other hand, the linear complementarity problem that we consider, can be described by following relations:

$$y = Ax + B\lambda + b \tag{4.9a}$$

$$w = Cx + D\lambda + c \tag{4.9b}$$

$$0 \leq \lambda \perp w \geq 0, \tag{4.9c}$$

where, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{m\times k}$, $b \in \mathbb{R}^m$, and $C \in \mathbb{R}^{k\times m}$, $D \in \mathbb{R}^{k\times k}$, $c \in \mathbb{R}^k$.

From complementarity to PWL mappings

When the matrix D in (4.9) is a P-matrix, we can explicitly solve for λ, which will be a piecewise linear function of x. To see this, let us first consider a scalar example.

Example 4.1. For $d > 0$, consider the complementarity relation

$$w = x + d\lambda + c$$
$$0 \leq \lambda \perp w \geq 0$$

If $x + c > 0$, then $\lambda = 0$. If $x + c \leq 0$, then $\lambda = -\frac{(x+c)}{d}$. Thus, λ is a piecewise linear function, where the first piece takes the value $-\frac{(x+c)}{d}$ over the interval $(-\infty, -c)$, and the second piece is identically zero on the interval $(-c, +\infty)$.

If we introduce the notation

$$\mathbb{R} \ni z \mapsto f_{\text{pos}}(z) = (|z| + z)/2,$$

then, we can compactly write our solution as

$$\lambda = f_{\text{pos}}(-(x+c)/d).$$

Let us next consider a second order example with an aim to provide an intuition for analysis in higher dimensions.

Example 4.2. Consider the following complementarity relation

$$0 \leq \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} q_1(x) \\ q_2(x) \end{pmatrix} \perp \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \geq 0.$$

We have a unique solution if and only if D is a P-matrix. To see the need for this particular assumption on D matrix, let us solve the LCP.

- If $\lambda_1 = \lambda_2 = 0$, then we are in the region $q_1(x) \geq 0$ and $q_2(x) \geq 0$.
- If $\lambda_1 \geq 0$, $\lambda_2 = 0$, then essentially $q_1(x) \leq 0$ and we have $\lambda_1 = -\frac{q_1(x)}{d_{11}}$. For the corresponding region, we also have, $w_2(x) = d_{21}\lambda_1 + q_2(x) = -\frac{d_{21}}{d_{11}}q_1(x) + q_2(x) \geq 0$.
- If $\lambda_2 \geq 0$, $\lambda_1 = 0$, then $q_2(x) \leq 0$ and we obtain $\lambda_2 = -\frac{q_2(x)}{d_{22}}$. To define the corresponding region, we also have that, $w_1(x) = d_{12}\lambda_2 + q_1(x) = -\frac{d_{12}}{d_{22}}q_2(x) + q_1(x) \geq 0$.
- If $\lambda_2 > 0$, $\lambda_1 > 0$, then the corresponding region is $q_1(x) \leq 0$ and $q_2(x) \leq 0$. We solve for λ_1 and λ_2 using the relation that $w_1 = 0$ and $w_2 = 0$. That is,

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = -\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}^{-1} \begin{pmatrix} q_1(x) \\ q_2(x) \end{pmatrix}$$

In other words,

$$\begin{aligned} \lambda_1 &= \frac{d_{12}q_2(x) - d_{22}q_1(x)}{d_{11}d_{22} - d_{21}d_{12}} \geq 0, \quad \text{if} \quad q_2(x) \geq \frac{d_{22}}{d_{12}}q_1(x) \\ \lambda_2 &= \frac{d_{21}q_1(x) - d_{11}q_2(x)}{d_{11}d_{22} - d_{21}d_{12}} \geq 0, \quad \text{if} \quad \frac{d_{21}}{d_{11}}q_1(x) \geq q_2(x) \end{aligned}$$

In all the above cases, we see that, for existence of nonnegative λ_1 and λ_2, it is necessary that the principal minors of D matrix are strictly positive.

From PWL mappings to complementarity

Let us consider some basic examples of piecewise linear relations modeled via complementarity frameworks. Some further examples can be found in [25].

Example 4.3. For some $a \in \mathbb{R}$, consider the following piecewise linear function, which also models the ReLU activation function:

$$x \mapsto y = f(x) = \max\{a, x\},$$

which basically models ReLU activation function. We can write it as

$$y = x + \lambda, \qquad w = x + \lambda - a, \qquad 0 \leq \lambda \perp w \geq 0.$$

or

$$y = a + \lambda, \qquad w = -x + a + \lambda, \qquad 0 \leq \lambda \perp w \geq 0.$$

Example 4.4. Now, let us consider the saturation function:

$$y = f(x) = \mathrm{sat}(x) = \begin{cases} 1, & \text{if } x \geq 1 \\ x, & \text{if } -1 \leq x \leq 1 \\ -1, & \text{if } x \leq -1 \end{cases}$$

We can write it in complementarity form as

$$y = -1 + \lambda_1 - \lambda_2, \qquad w_1 = -x + \lambda_1 - 1, \qquad w_2 = -x + \lambda_2 + 1, \qquad 0 \leq \lambda \perp w \geq 0.$$

Example 4.5. The so-called static backlash hysteresis contains functions like [109, Section 6],

$$y = \max(u, \min(v, w)) = \max(u, -\max(-v, -w))$$

which can be rewritten using complementarity relations as

$$\begin{aligned} &y = w - \lambda_1 + \lambda_2, \\ &0 \leq \lambda_1 \perp \lambda_1 + v - w \geq 0, \\ &0 \leq \lambda_2 \perp -\lambda_1 + \lambda_2 - u + w \geq 0. \end{aligned}$$

Remark 4.1. One can potentially use the framework of linear complementarity problem to model piecewise linear functions of the following form:

$$y = A_i x + b_i, \quad \text{if } C_i x + c_i \geq 0.$$

Such discussions appear in [130]. An algorithm for encoding PWL relations into more generic complementarity framework also appears in [218]. However, some questions remain unaddressed when seeking such generalizations (see [27, Section 3.11]).

4.2.2 Projected dynamical systems

The so-called *projected dynamical systems* model the trajectories confined to a given set. A formal treatment of this class of dynamical systems can be found in [156]. For describing the models of these systems, we will first introduce the *tangent cones* and then make connections with related system classes.

Tangent cones

Detailed discussions about different notions of tangent cones can be found in [13]. Here, we provide some fundamentals that are necessary for our purposes.

Let S be a subset of $\mathbb{R}^n$. The following notions of tangent cones are widely used:

Definition 4.9 (Contingent Cone)**.** The Bouligand tangent cone, or contingent cone, to a set $S \subset \mathbb{R}^n$ at a point $x \in S$, denoted $\mathcal{T}_S(x)$, is

$$\begin{aligned} \mathcal{T}_S(x) &= \{z \in \mathbb{R}^n | \exists\, \{t_n\} \subset \mathbb{R}_+, \exists\, \{z_n\} \subset \mathbb{R}^n, \text{such that } t_n \searrow 0, z_n \to z, x + t_n z_n \in S, \forall\, n \in \mathbb{N}\} \\ &= \{z \in \mathbb{R}^n | \exists\, \{x_n\} \subset S, \exists\, \{\lambda_n\} \subset \mathbb{R}_+, \lambda_n \searrow 0, \text{such that } x_n \to x \text{ and } \frac{x_n - x}{\lambda_n} \to z\}. \end{aligned}$$

Definition 4.10 (Clarke Tangent Cone)**.** The Clarke tangent cone to the set S, denoted by $\mathcal{T}_S^C(x)$, is defined as follows:

$$\mathcal{T}_S^C(x) = \Big\{ z \in \mathbb{R}^n | \forall \{x_n\} \subset S, \forall \{t_n\} \subset \mathbb{R}_+ \text{ with } \; x_n \to x, t_n \searrow 0, \\ \exists \{z_n\} \in \mathbb{R}^n \text{ with } z_n \to z, x_n + t_n z_n \in S, \forall\, n \in \mathbb{N} \Big\}.$$

In fact, Clarke tangent cone at a point is the inner limit of contingent cone of the neighboring points, that is, $\mathcal{T}_S^C(x) = \liminf_y \mathcal{T}_S(y)$. In general, it holds that $\mathcal{T}_S^C(x) \subseteq \mathcal{T}_S(x)$. We say that S is *tangentially regular* if $\mathcal{T}_S(x) = \mathcal{T}_S^C(x)$ for every $x \in S$. Convex sets are tangentially regular but those are not the only ones.

Projection on tangent cone

Let us now define the dynamical systems using the framework of tangent cones. For that, we first introduce the projection of a vector $v \in \mathbb{R}^n$ onto S, denoted by $\mathcal{P}_S(v)$, which is defined as

$$\mathcal{P}_S(v) = \arg\min_{x \in S} \|v - x\|. \tag{4.10}$$

where $\|\cdot\|$ denotes the Euclidean distance, and we note that $\arg\min$ may not have a unique solution if S is not convex.

Intuition behind projected dynamical systems: Consider the evolution of the system $\dot{x} = v$, with $x(0) = x_0$ on the boundary of S. The first-order approximation of the resulting trajectory gives $x(\delta) = x_0 + \delta v$, where $\delta > 0$ is very small. If we want $x(\delta)$ to be in the set S, one way is to project $x_0 + \delta v$ on the set S.

To write a differential equation for the resulting motion, we therefore need to consider the directional derivative of the operator $\mathcal{P}_S(\cdot)$. For a given vector $v \in \mathbb{R}^n$, $\Pi(x; v)$ denotes the directional derivative of $\mathcal{P}_S(x)$ and is defined as

$$\Pi_S(x; v) = \lim_{\delta \to 0} \frac{\mathcal{P}_S(x + \delta v) - \mathcal{P}_S(x)}{\delta}. \tag{4.11}$$

We note that

$$\mathcal{P}_S(x + \delta v) = x + \delta \mathcal{P}_{\mathcal{T}_S(x)}(v) + o(\delta)$$

where $o(\delta)$ is such that $\lim_{\delta \to 0^+} \frac{o(\delta)}{\delta} = 0$, and hence

$$\Pi_S(x; v) = \mathcal{P}_{\mathcal{T}_S(x)}(v).$$

System Class: With this motivation, we therefore consider the class of *projected dynamical systems*, defined as,

$$\dot{x} = \Pi_S(x; f(x)) \tag{4.12}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable.

More recently, we have seen extensions of this conventional framework which are motivated by applications in optimization and implementing reset control. The paper [92] considers such system class using a state-dependent projection metric. An interesting application using projection based controllers has been reported in [61, 94], which not only involves a different projection metric but also considers sets that are neither convex nor tangentially regular.

Connections with maximal monotone differential inclusion

We can describe the solutions of the differential equation (4.12) using a differential inclusion with a normal cone operator, so that the resulting system takes the form (4.1). In particular, the *least-norm* element of the set-valued mapping describing the differential inclusion coincides with the right-hand side of (4.12). The key to this connection comes from the following elegant result:

Theorem 4.11 (Moreau's two-cone decomposition). *Let $K \subset \mathbb{R}^n$ be a closed convex cone, and let $K^\circ := \{w \in \mathbb{R}^n \mid \langle w, v\rangle \leq 0,\ \forall v \in K\}$ denote the polar cone. For every $z \in \mathbb{R}^n$, it holds that*

$$z = \mathcal{P}_K(z) + \mathcal{P}_{K^\circ}(z).$$

To use this result, we recall that for a closed, convex set $S \subset \mathbb{R}^n$, the tangent cone $\mathcal{T}_S(x)$ is closed and convex, and $\mathcal{T}_S^\circ(x) = \mathcal{N}_S(x)$, for each $x \in \mathbb{R}^n$. We can then write,

$$f(x) = \mathcal{P}_{\mathcal{T}_S(x)}(f(x)) + \mathcal{P}_{\mathcal{N}_S(x)}(f(x))$$

or equivalently,

$$\mathcal{P}_{\mathcal{T}_S(x)}(f(x)) = f(x) - \mathcal{P}_{\mathcal{N}_S(x)}(f(x)).$$

Note that, $f(x) - \mathcal{P}_{\mathcal{N}_S(x)}(f(x))$ is just the least norm element of the set $f(x) - \mathcal{N}_S(x)$, that is,

$$f(x) - \mathcal{P}_{\mathcal{N}_S(x)}(f(x)) = \operatorname*{arg\,min}_{w \in f(x) - \mathcal{N}_S(x)} \|w\|.$$

In other words, $\Pi_S(x; f(x))$ is just the least-norm element of $f(x) - \mathcal{N}_S(x)$. Therefore, the solutions of (4.12) coincide with the slow solutions of

$$\dot{x} \in f(x) - \mathcal{N}_S(x). \tag{4.13}$$

Since $\mathcal{N}_S(\cdot)$ is a maximal monotone mapping, system (4.13) is in the form of (4.1).

Connections with complementarity systems: In the differential inclusion (4.13), if we take S to be a closed convex cone K, then the differential inclusion becomes a cone complementarity system. Recall that,

$$\lambda \in -\mathcal{N}_K(x) \quad \Leftrightarrow \quad K^\star \ni \lambda \perp x \in K,$$

where $K^\star$ is the dual cone to K. We can thus write (4.13) equivalently as

$$\begin{aligned} \dot{x} &= f(x) + \lambda, \\ K^\star &\ni \lambda \perp x \in K. \end{aligned}$$

We have thus seen that a well-known class of projected dynamical systems can be rewritten in the form of differential inclusion involving maximal monotone mapping, or dynamical system with complementarity problem. This helps us in better understanding the motivation for the system class proposed at the beginning of this chapter in (4.1). We will now move in the direction of analysing this system class in the remaining sections.

4.3 Monotonicity Preserving Interconnections

Until now, in this chapter, we saw some instances of nonsmooth relations which are maximally monotone. The examples in Section 4.1 provided some examples of dynamical systems where these relations appear, and the overall system described by the sum of a drift term and the nonsmooth mapping is still maximal monotone because of the monotonicity of the drift term. Thus, the sum of two maximal monotone mappings is one example of an operation that preserves maximal monotonicity (under some appropriate constraint qualifications). In the previous section, we looked at projected dynamical systems, and for f monotone, these systems provide another instance of an operation (described by the projection on to the tangent cone of a closed convex set) that preserves maximal monotonicity.

In general, it is an interesting question to investigate whether certain operations preserve maximal monotonicity or not. For our purposes, when dealing with dynamical systems, this question naturally translates to studying interconnections involving (possibly memoryless) systems with maximal monotone mappings and providing conditions which preserve maximal monotonicity.

4.3.1 Passive systems with maximal monotone mappings

Let us first consider the feedback interconnection of a passive system with maximal monotone mappings, and show that under appropriate assumptions, the interconnection is also maximal monotone. The system that we consider are described by following equations:

$$\begin{aligned} \dot{x} &= Ax + B\lambda \\ w &= Cx + D\lambda \\ w &\in \mathcal{M}(-\lambda) \end{aligned} \tag{4.14}$$

where the quadruple (A, B, C, D) describes a passive system with input $\lambda \in \mathbb{R}^m$ and output $w \in \mathbb{R}^m$. System (4.14) therefore describes the interconnection of a passive dynamical system with a memoryless maximal monotone mapping $\mathcal{M} : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$. This system can be equivalently written in the following form:

$$\dot{x} \in -\mathcal{H}(x) := Ax - B(\mathcal{M} + D)^{-1}(Cx) \tag{4.15}$$

where $\mathcal{H} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued mapping with $\operatorname{dom}(\mathcal{H}) = C^{-1}(\operatorname{im}(\mathcal{M} + D))$.

Theorem 4.12 ([38, Theorem 2]). *Suppose that*

1. *The quadruple* (A, B, C, D) *is passive with storage function* $x \mapsto x^\top x$
2. *The mapping* $\mathcal{M}$ *is maximal monotone*
3. *It holds that* $\operatorname{im} C \cap \operatorname{rint}(\operatorname{im}(\mathcal{M} + D)) \neq \emptyset$

then the mapping $\mathcal{H}$ *in* (4.15) *is maximal monotone.*

A time-varying version of this result has been developed in [36]. In particular, if the variable w in (4.14) is of the form $w(t) = Cx(t) + D\lambda(t) + v(t)$, with $v : [0, T] \to \mathbb{R}^m$ an external absolutely continuous function, then the mapping $\mathcal{H}$ defined above is time-dependent and in particular, the domain of this mapping varies with time as well. The resulting differential inclusion is then described by

$$\dot{x}(t) \in -\mathcal{H}(t, x), \tag{4.16}$$

where $\mathcal{H}(t,\cdot)$ is a maximal monotone mapping for each $t \in [0,T]$. Even though the interconnection of the passive time-invariant system with this time-varying complementarity relation preserves maximal monotonicity for each $t \in [0,T]$, studying the existence of solutions for this system class is challenging and requires careful analysis. In our work [36], we provide a generalization of proximal-type algorithms (that were presented in Section 4.1) that can be used to simulate the dynamical system (4.16) and we can study the convergence of those algorithms with respect to the size of step for discretization.

4.3.2 Maximal monotone systems in feedback

Moving away from the setup of a passive dynamical system with a memoryless maximal monotone mapping, we now consider the interconnections of the form:

$$\dot{y} \in -\mathcal{M}_y(y,z) \tag{4.17a}$$

$$\dot{z} \in -\mathcal{M}_z(y,z) \tag{4.17b}$$

where $\mathcal{M}_y(\cdot,z) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, and $\mathcal{M}_z(y,\cdot) : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ define set-valued mappings for a given $z \in \mathbb{R}^m$, and a given $y \in \mathbb{R}^n$, respectively. We impose the following conditions on these mappings:

(M1) There exist $\alpha_y \geq 0$ and a function $\gamma_y : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ such that, for every $y_1^\star \in \mathcal{M}_y(y_1,z_1)$, $y_2^\star \in \mathcal{M}_y(y_2,z_2)$, we have

$$\begin{aligned}\langle y_1^\star - y_2^\star, y_2 - y_1\rangle \leq -\alpha_y\|y_2-y_1\|^2 + \gamma_y(y_1,z_2) - \gamma_y(y_1,z_1)\\ + \gamma_y(y_2,z_1) - \gamma_y(y_2,z_2).\end{aligned} \tag{4.18}$$

(M2) There exist $\alpha_z \geq 0$ and a function $\gamma_z : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ such that, for every $z_1^\star \in \mathcal{M}_z(y_1,z_1)$, $z_2^\star \in \mathcal{M}_z(y_2,z_2)$, we have

$$\begin{aligned}\langle z_1^\star - z_2^\star, z_2 - z_1\rangle \leq -\alpha_z\|z_2-z_1\|^2 + \gamma_z(y_2,z_1) - \gamma_z(y_1,z_1)\\ + \gamma_z(y_1,z_2) - \gamma_z(y_2,z_2).\end{aligned} \tag{4.19}$$

(M3) For each $z \in \mathbb{R}^m$, and every $(y_1^\star, y_1) \in \mathbb{R}^n \times \mathbb{R}^n$ that satisfies $y_1^\star \notin \mathcal{M}_y(y_1,z)$, there exists $(y_2^\star, y_2)$ with $y_2^\star \in \mathcal{M}_y(y_2,z)$ such that

$$\langle y_2^\star - y_1^\star, y_2 - y_1\rangle < 0.$$

(M4) For each $y \in \mathbb{R}^n$, and every $(z_1^\star, z_1) \in \mathbb{R}^m \times \mathbb{R}^m$ that satisfies $z_1^\star \notin \mathcal{M}_z(y,z_1)$, there exists $(z_2^\star, z_2)$ with $z_2^\star \in \mathcal{M}_z(y,z_2)$ such that

$$\langle z_2^\star - z_1^\star, z_2 - z_1\rangle < 0.$$

(M5) For each $z \in \mathbb{R}^m$, $\mathrm{dom}(\mathcal{M}_y(\cdot,z)) \neq \emptyset$, and for each $y \in \mathbb{R}^n$, $\mathrm{dom}(\mathcal{M}_z(y,\cdot)) \neq \emptyset$.

The conditions **(M1)** and **(M2)** basically describe the (strong) monotonicity of the mappings $\mathcal{M}_y(\cdot,z)$ and $\mathcal{M}_z(y,\cdot)$, respectively, for a given $z \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. We will present some case studies in this section which would clarify these inequalities further. At this moment, in **(M1)**, it could be readily seen that if $y_i^\star \in \mathcal{M}_y(y_i,z)$ for some $z \in \mathbb{R}^m$ and $i = 1,2$, then the inequality (4.18) reduces to the inequality in (4.3). The same observation can be made for **(M2)**. Conditions **(M3)** and **(M4)** basically refer to the maximality of $\mathcal{M}_y(\cdot,z)$ and $\mathcal{M}_z(y,\cdot)$, respectively, for a given $z \in \mathbb{R}^m$, $y \in \mathbb{R}^n$. In this chapter, we use **(M5)** for simplifying some technical arguments used in the proof of the main result and it would be natural to think about relaxing this assumption in future work.

Theorem 4.13. *Consider the interconnected dynamical system* (4.17) *under the conditions* **(M1)**, **(M2)**, **(M3)**, **(M4)**, **(M5)**. *Suppose that the function $\gamma_y + \gamma_z$ is bilinear, so that there exist $C_\gamma \geq 0$ that satisfies*

$$\|(\gamma_y + \gamma_z)(y, z)\| \leq C_\gamma(\|y\|^2 + \|z\|^2), \tag{4.20}$$

for all $(y, z) \in \mathbb{R}^n \times \mathbb{R}^m$. If, it holds that,

$$\min\{\alpha_y, \alpha_z\} \geq C_\gamma \tag{4.21}$$

then $(\mathcal{M}_y, \mathcal{M}_z) \triangleq \mathcal{M} : \mathbb{R}^{n+m} \rightrightarrows \mathbb{R}^{n+m}$ in system (4.17) *is maximal monotone. If the inequality in* (4.21) *is strict, then $\mathcal{M}$ is strongly maximal monotone.*

Proof. Let $x_1 \triangleq (y_1, z_1)$, and $x_2 \triangleq (y_2, z_2)$, where $y_i \in \operatorname{dom} \mathcal{M}_y(\cdot, z_i)$ and $z_i \in \operatorname{dom} \mathcal{M}_z(y_i, \cdot)$, for $i = 1, 2$. Also, let $\alpha \triangleq \min\{\alpha_y, \alpha_z\}$, and $\gamma \triangleq \gamma_y + \gamma_z$. In the sequel, we will prove monotonicity and maximality separately.

Monotonicity: To check the monotonicity of the mapping $\mathcal{M}(\cdot)$, we consider $x_1^\star \in \mathcal{M}(x_1)$ and $x_2^\star \in \mathcal{M}(x_2)$. We can then write $x_1^\star = (y_1^\star, z_1^\star)$ with $y_1^\star \in \mathcal{M}_y(y_1, z_1)$, and $z_1^\star \in \mathcal{M}_z(y_1, z_1)$. Similarly, we can write $x_2^\star = (y_2^\star, z_2^\star)$ with $y_2^\star \in \mathcal{M}_y(y_2, z_2)$, and $z_2^\star \in \mathcal{M}_z(y_2, z_2)$. It then follows that

$$\begin{aligned}
\langle x_2^\star - x_1^\star, x_2 - x_1 \rangle &= \langle y_2^\star - y_1^\star, y_2 - y_1 \rangle + \langle z_2^\star - z_1^\star, z_2 - z_1 \rangle \\
&\geq \alpha_y \|y_2 - y_1\|^2 + \alpha_z \|z_2 - z_1\|^2 + \gamma(y_1, z_1) - \gamma(y_1, z_2) + \gamma(y_2, z_2) - \gamma(y_2, z_1) \\
&= \alpha_y \|y_2 - y_1\|^2 + \alpha_z \|z_2 - z_1\|^2 + \gamma(y_1 - y_2, z_1 - z_2) \\
&\geq \alpha \|y_2 - y_1\|^2 + \alpha \|z_2 - z_1\|^2 - C_\gamma \left(\|y_1 - y_2\|^2 + \|z_1 - z_2\|^2 \right) \\
&\geq (\alpha - C_\gamma) \left(\|y_2 - y_1\|^2 + \|z_2 - z_1\|^2 \right)
\end{aligned}$$

and hence, under the condition (4.21), the mapping $\mathcal{M}(\cdot)$ is monotone, or strongly monotone if the inequality in (4.21) is strict.

Maximality: Next, we use **(M3)**, **(M4)** and **(M5)** to show the maximality of $\mathcal{M}$. Towards this end, fix any $(y_1, z_1) \in \mathbb{R}^n \times \mathbb{R}^m$ and any $(y_1^\star, z_1^\star) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $(y_1^\star, z_1^\star) \notin \mathcal{M}(y_1, z_1)$. One of the two statements is true: either $y_1^\star \notin \mathcal{M}_y(y_1, z_1)$, or $z_1^\star \notin \mathcal{M}_z(y_1, z_1)$. For definiteness, let us suppose the former, that is, $y_1^\star \notin \mathcal{M}_y(y_1, z_1)$. From **(M5)**, $\operatorname{dom}(\mathcal{M}_y(\cdot, z_1)) \neq \emptyset$. It then follows from **(M3)** that there exists $(y_2^\star, y_2) \in \operatorname{graph}(\mathcal{M}_y(\cdot, z_1))$ such that

$$\langle y_2^\star - y_1^\star, y_2 - y_1 \rangle < 0.$$

Consider the point (y_2, z_1), and once again, using **(M5)**, we have that $\operatorname{dom}(\mathcal{M}_z(y_2, \cdot)) \neq \emptyset$. If $z_1 \in \operatorname{dom}(\mathcal{M}_z(y_2, \cdot))$, then we pick $z_2 = z_1$ and $z_2^\star \in \mathcal{M}_z(y_2, z_1)$, so that

$$\langle z_2^\star - z_1^\star, z_2 - z_1 \rangle = 0.$$

Else, if $z_1 \notin \operatorname{dom}(\mathcal{M}_z(y_2, \cdot))$, then using **(M4)**, there exists $(z_2^\star, z_2) \in \operatorname{graph}(\mathcal{M}_z(y_2, \cdot))$ such that

$$\langle z_2^\star - z_1^\star, z_2 - z_1 \rangle < 0.$$

We thus have $(y_2^\star, z_2^\star) \in \mathcal{M}(y_2, z_2)$ such that

$$\langle y_2^\star - y_1^\star, y_2 - y_1 \rangle + \langle z_2^\star - z_1^\star, z_2 - z_1 \rangle < 0,$$

which shows the maximality of the mapping $\mathcal{M}(\cdot)$. □

4.3.3 Saddle-point dynamics

As a first instance of the illustration of the interconnection considered in (4.17), and the result presented in Theorem 4.13, let us consider a saddle function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{-\infty, +\infty\}$, and the dynamics defined by its subdifferential. In particular, we assume that $f(\cdot,\cdot)$ satisfies the following:

(SF1) For each $z \in \mathbb{R}^m$, the function $f(\cdot, z)$ is concave, and upper semicontinuous.

(SF2) For each $y \in \mathbb{R}^n$, the function $f(y, \cdot)$ is convex, and lower semicontinuous.

(SF3) The function $f(\cdot,\cdot)$ is finite-valued everywhere on $\mathbb{R}^n \times \mathbb{R}^m$.

For such functions, we let $\partial_z f(y, z)$ denote the subdifferential of the convex function $f(y, \cdot)$ at $z \in \mathbb{R}^m$, which is well-defined due to **(SF2)**. Similarly, for each $z \in \mathbb{R}^m$, we let $\partial_y f(y, z)$ denote the superdifferential of the concave function $f(\cdot, z)$ at $y \in \mathbb{R}^n$, which is well-defined due to **(SF1)**. Condition **(SF3)** is somewhat restrictive in the context of saddle functions as considered in [176], but it simplifies the analysis. We can now introduce the saddle-point dynamics as follows [79]:

$$\dot{y} \in \partial_y f(y, z) \tag{4.22a}$$
$$\dot{z} \in -\partial_z f(y, z) \tag{4.22b}$$

Proposition 4.14. *Consider the dynamical system* (4.22) *with* $f(\cdot,\cdot)$ *satisfying* **(SF1)**, **(SF2)**, **(SF3)**. *Then, the set-valued mapping* $(x, y) \mapsto (-\partial_y f(y, z), \partial_z f(y, z))$ *is maximal monotone.*

Proof. The proof basically relies on showing that the conditions **(M1)**–**(M5)** hold, and so the result follows due to Theorem 4.13.

Consider the points $(y_i, z_i) \in \mathbb{R}^n \times \mathbb{R}^m$ and $(y_i^\star, z_i^\star) \in -\partial_y f(y_i, z_i) \times \partial_z f(y_i, z_i)$. By definition,

$$\langle y_1^\star, y - y_1\rangle - f(y_1, z_1) \le -f(y, z_1), \quad \forall\, y \in \mathbb{R}^n$$
$$\langle y_2^\star, y - y_2\rangle - f(y_2, z_2) \le -f(y, z_2), \quad \forall\, y \in \mathbb{R}^n$$

This results in

$$\langle y_1^\star - y_2^\star, y_2 - y_1\rangle \le f(y_1, z_1) - f(y_1, z_2) - f(y_2, z_1) + f(y_2, z_2) \tag{4.23}$$

so that **(M1)** holds with $\alpha_y = 0$ and $\gamma_y(\cdot,\cdot) = -f(\cdot,\cdot)$.

Similarly, using the definition of the subdifferential of the convex function, we get

$$\langle z_1^\star, z - z_1\rangle + f(y_1, z_1) \le f(y_1, z), \quad \forall\, z \in \mathbb{R}^m$$
$$\langle z_2^\star, z - z_2\rangle + f(y_2, z_2) \le f(y_2, z), \quad \forall\, z \in \mathbb{R}^m$$

This results in

$$\langle z_1^\star - z_2^\star, z_2 - z_1\rangle \le f(y_1, z_2) - f(y_1, z_1) + f(y_2, z_1) - f(y_2, z_2) \tag{4.24}$$

so that **(M2)** holds with $\alpha_z = 0$ and $\gamma_z(\cdot,\cdot) = f(\cdot,\cdot)$.

Recalling the fact that the subdifferential of a proper, convex, lower semicontinuous function, and the superdifferential of a proper, concave, upper semicontinuous function are maximal monotone, the statements in **(M3)**, **(M4)**, **(M5)** follow from **(SF1)**, **(SF2)**, and **(SF3)**. □

4.3.4 Dissipative interconnection

Consider the dynamical system with state x, input u and output ω, described by

$$\begin{aligned} \dot{x} &= F(x,u) \\ \omega &= h(x) \end{aligned} \tag{4.25}$$

for some continuous functions $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, and $h : \mathbb{R}^n \to \mathbb{R}^m$. We say that the system (4.25) is *incrementally passive* if there exists a continuously differentiable storage function V, $\mathbb{R}^{2n} \ni (x_1, x_2) \mapsto V(x_1, x_2) \in \mathbb{R}_+$ such that, $\frac{\partial V}{\partial x_1} F(x_1, u_1) + \frac{\partial V}{\partial x_2} F(x_2, u_2) \leq \langle \omega_1 - \omega_2, u_1 - u_2 \rangle$, where $\omega_1 = h(x_1)$ and $\omega_2 = h(x_2)$.

Let us now consider a negative feedback interconnection of two incrementally passive systems. The first system is described as

$$\dot{y} = F_y(y, u_y), \quad \omega_y = y \tag{4.26}$$

with $F_y : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ continuous. The dynamics of the second system are given by

$$\dot{z} = F_z(z, u_z), \quad \omega_z = z \tag{4.27}$$

with $F_y : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ continuous. The interconnection between these systems is described via the negative feedback as:

$$u_y = -\omega_z, \qquad u_z = \omega_y. \tag{4.28}$$

The interconnection of incrementally passive systems has been widely studied in the literature, see for example [232], [68, p. 184], [165], [186, Prop. 2.2.21]. Here we show that such interconnections are actually maximally monotone. We choose to work with continuous single-valued functions so that the maximality holds [180, Example 12.7].

Proposition 4.15. *Suppose that the system* (4.26) *is incrementally passive with output* ω_y*, input* u_y*, the storage function* $V_y(y_1, y_2) = \|y_1 - y_2\|^2$*, and that the system* (4.27) *is incrementally passive with output* ω_z*, input* u_z*, the storage function* $V_z(z_1, z_2) = \|z_1 - z_2\|^2$*. Then, the system defined by their interconnection via negative feedback* (4.28) *is maximal monotone.*

Proof. The proof follows once again by an application of Theorem 4.13, where we can show that **(M1)**–**(M5)** hold. The statement **(M1)** basically follows from the incremental passivity of (4.26) with $\alpha_y = 0$ and $\gamma_y(y, z) = \langle y, z \rangle$. Similarly, **(M2)** follows from the incremental passivity of (4.27) with $\alpha_z = 0$ and $\gamma_z(y, z) = -\langle y, z \rangle$. Thus, $\gamma_y + \gamma_z = 0$. Statements **(M3)**, **(M4)**, **(M5)** basically follow from the continuity of the mappings $F_y(\cdot, \cdot)$ and $F_z(\cdot, \cdot)$ over the entire space $\mathbb{R}^n \times \mathbb{R}^n$. □

4.4 Stability of Constrained Systems

In the previous sections, we saw that certain classes of nonsmooth systems can equivalently be written as differential inclusions with maximal monotone mappings. In that form, asymptotic behavior of trajectories, or convergence to an equilibrium can be analyzed thanks to Proposition 4.2. In the general case, for system class (4.1), the drift term f is not necessarily maximal monotone and and the stability of the system needs to be analyzed differently. When the mapping $\mathcal{M}(\cdot)$ is of the form $\mathcal{N}_S(\cdot)$ for some set S, the dynamics model certain constraints and we are interested in studying stability of such systems in this section. In particular, we consider the system class:

$$\dot{x} \in f(x) - \mathcal{N}_S(x) \tag{4.29}$$

where S is a closed convex set, and $\mathcal{N}_S$ denotes the outward normal cone. We have already seen connections between this system class and other models such as projected dynamical systems, and systems with variational inequalities. Some of the earlier references in this direction which study stability problem are due to [81, 37]. The results presented in this section are mostly borrowed from our paper [197].

Example 4.6. To provide a motivation for the primary question addressed in our stability analysis, consider the linear complementarity system

$$\dot{x} \in \begin{bmatrix} -1 & -2 \\ -1 & -1 \end{bmatrix} x - \mathcal{N}_{\mathbb{R}^2_+}(x)$$

which is of the form (4.29). For this system, there **does not** exist a quadratic positive definite Lyapunov function matrix P such that its derivative along the system dynamics (for all choices of the vector field) is negative definite. This is because A is not Hurwitz and $0 \in \mathcal{N}_{\mathbb{R}^2_+}(x)$ for each $x \in \mathbb{R}^2_+$. However, the constrained system is asymptotically stable. △

In general, constrained systems may be unstable even if A is Hurwitz stable. Our goal in this section is to develop appropriate conditions for stability of the systems with constraints given in the form (4.29).

Theorem 4.16 (Sufficient conditions with constraints [82])**.** *For a proper, convex, lower semicontinuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$, consider the system*

$$\dot{x} \in f(x) - \partial\varphi(x), \quad x(0) \in \operatorname{dom}(\partial\varphi), \tag{4.30}$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a proper, lower semicontinuous, and convex function with $0 \in \partial\varphi(0)$. Assume that there exists a continuously differentiable, positive definite function $V(\cdot)$ such that

- *$V(0) = 0$, and $V(x) \geq c\|x\|^r$ for $x \in \operatorname{dom}(\partial\varphi)$,*
- *For every $x \in \operatorname{dom}(\partial\varphi)$, it holds that*

$$\langle f(x), \nabla V(x)\rangle - \langle w, \nabla V(x)\rangle \leq -\lambda V(x), \quad \forall\, w \in \partial\varphi(x), \tag{4.31}$$

  for some $\lambda > 0$.

Then the origin $\{0\}$ is asymptotically stable.

One way to arrive at (4.31) is the following: if we can ensure that for every $x \in \operatorname{dom}(\partial\varphi)$, we have $x - \nabla V(x) \in \operatorname{dom}(\varphi)$, and

$$\langle f(x), \nabla V(x)\rangle + \varphi(x - \nabla V(x)) - \varphi(x) \leq -\lambda V(x), \quad \forall\, x \in \operatorname{dom}(\partial\varphi), \tag{4.32}$$

then the inequality (4.32) implies (4.31). This is because, by the definition of subdifferentials, $w \in \partial\varphi(x)$, satisfies

$$\langle w, v - x\rangle \leq \varphi(v) - \varphi(x), \quad \forall v \in \mathbb{R}^n, x \in \operatorname{dom}(\partial\varphi),$$

and by taking $v = x - \nabla V(x)$, we get,

$$\langle w, -\nabla V(x)\rangle \leq \varphi(x - \nabla V(x)) - \varphi(x), \quad \forall v \in \mathbb{R}^n, x \in \operatorname{dom}(\partial\varphi).$$

We can use this later observation to write somewhat simpler conditions for the case when φ is taken to be the indicator function of a set, that is, $\varphi = \psi_S$ for some closed and convex set S. In that case, $w \in \partial\varphi(x)$ is nonzero only when x is on the boundary of the set.

Corollary 4.17. *Take φ to be the indicator function of a closed convex set S, that is, $\varphi = \psi_S$ with $0 \in S$ and $f(0) = 0$. Then the origin of system* (4.30) *is asymptotically stable if there exists a continuously differentiable, positive definite function $V(\cdot)$ that satisfies the following*

- $V(0) = 0$, *and* $V(x) \geq c\|x\|^r$ *for* $x \in S$,
- *For each* $x \in S$, $\langle f(x), \nabla V(x)\rangle \leq -\lambda V(x)$,
- $(x - \nabla V(x)) \in S$ *whenever x is on the boundary of S.*

4.4.1 Copositive Lyapunov functions

Taking S to be the closed convex cone in $\mathbb{R}^n$, and taking V to be a quadratic function, we can formulate the conditions in (4.17) in terms of copositive matrices.

Definition 4.18 (Copositive Matrices)**.** A matrix $P \in \mathbb{R}^{n\times n}$ is said to be copositive on the convex cone K if $\langle Px, x\rangle \geq 0$, for every $x \in K$. A matrix $P \in \mathbb{R}^{n\times n}$ is said to be strictly copositive on K if there exists $c > 0$ such that

$$\langle Px, x\rangle \geq c\,\|x\|^2, \qquad \text{for every } x \in K.$$

Proposition 4.19 (Cone-membership conditions for matrices [81])**.** *Consider the system* (4.30) *with $\varphi = \psi_{\mathbb{R}^n_+}$. Assume that there exists a matrix $P = P^\top \in \mathbb{R}^{n\times n}$ such that*

- P *is strictly copositive.*
- $x - (P + P^\top)x \geq 0$, *whenever* $x_i = 0$.
- $-(A^\top P + PA)$ *is (strictly) copositive,*

then the origin is (asymptotically) stable.

4.4.2 Algorithmic detection

In Proposition 4.19, we provided sufficient conditions for asymptotic stability. By reducing our search to the Lyapunov functions of the form $V(x) = x^\top Px$, the proposed sufficient conditions boil down to checking whether P and some other matrices obtained from P are copositive or not. Checking whether a given matrix is copositive or not is a problem that has attracted a lot of attention in the optimization community, and in particular, *Copositive Programming.* For an overview of this class of conic optimization problems, the reader may consult [22, 23].

Let us now provide an algorithm for checking whether a matrix is copositive, or not. This will help us in writing an algorithm for computing (quadratic) copositive Lyapunov functions and the developments carried out in this section are based on the papers [30, 31].

In what follows, it is important to recall the notion of a simplex: An n-simplex is an n-dimensional polytope Δ which is the convex hull of its $n+1$ affinely independent vertices $\{x_0, x_1, \ldots, x_n\}$, that is, $x_1 - x_0, \cdots, x_n - x_{n-1}$ are linearly independent, and

$$\Delta := \left\{ \theta_0 x_0 + \cdots + \theta_n x_n \middle| \sum_{i=0}^{n} \theta_i = 1 \text{ and } \theta_i \geq 0 \text{ for all } i \in \{0, \ldots, n\} \right\}.$$

The algorithm builds on the following observation:

Proposition 4.20. *A matrix P is copositive, if and only if $x^\top P x \geq 0$ for all $x \in \mathbb{R}^n_+$ with $\|x\| = 1$.*

In the above statement, we may take any norm. If we choose 1-norm, then the problem of checking copositivity on $\mathbb{R}^n_+$ boils down to checking copositivity on the standard simplex, $\Delta^S := \{x \in \mathbb{R}^n_+ \,:\, \|x\|_1 = 1\}$.

Proposition 4.21. *Let $\Delta = \operatorname{conv}\{v_1, \ldots, v_n\}$ be a simplex in $\mathbb{R}^n$. If*

$$v_i^\top P v_j \geq 0, \qquad \text{for all } i, j \in \{1, \cdots, n\} \tag{4.33}$$

then $x^\top P x \geq 0$, for all $x \in \Delta$.

The condition in Proposition 4.21 provides a test for checking copositivity over a simplex. If the matrix P satisfies such a condition on the set Δ^S, then we are done. However, this is a very restrictive condition because the vertices v_i and v_j may be too far. Therefore, we try to provide a test for copositivity on smaller simplices where the distance between vertices is not too big. We therefore need to cut our standard simplex Δ^S into small simplices.

Definition 4.22. Let Δ be a simplex in $\mathbb{R}^n$. A family $\mathcal{P}_m := \{\Delta^1, \ldots, \Delta^m\}$ of simplices satisfying

$$\Delta = \bigcup_{i=1}^{m} \Delta^i \text{ and } \operatorname{int} \Delta^i \cap \operatorname{int} \Delta^j = \emptyset \text{ for } i \neq j$$

is called a simplicial partition of Δ.

With this definition, we can divide the task of checking copositivity on smaller simplices.

Proposition 4.23. *Let $P \in \mathbb{R}^{n\times n}$ be symmetric, and let $\mathcal{P}_m := \{\Delta^1, \ldots, \Delta^m\}$ be a simplical partition of Δ^S, where $v_1^k, \ldots, v_n^k$ denote the vertices of simplex Δ^k. If for each $k \in \{1, \cdots, m\}$, it holds that $(v_i^k)^\top P v_j^k \geq 0$ for all $i, j \in \{1, \ldots, n\}$, then P is copositive.*

So, if we start with a standard simplex Δ^S, and matrix P does not satisfy (4.33) for all the vertices in Δ^S, then we cut Δ^S into smaller simplices, and check condition (4.33) only for the vertices of the smaller simplices. However, this needs to be done for all simplices.

Definition 4.24. For a simplicial partition $\mathcal{P}_m = \{\Delta^1, \ldots, \Delta^m\}$ of a simplex Δ, where $v_1^k, \cdots, v_n^k$ denote the vertices of simplex Δ^k, the maximum diameter of a simplex in $\mathcal{P}_m$ is defined as

$$\delta(\mathcal{P}_m) := \max_{k\in\{1,\ldots,m\}} \max_{i,j\in\{1,\ldots,n\}} \|v_i^k - v_j^k\|.$$

Thus, the diameter of the partition quantifies the distance between vertices in each simplex contained in the partition. Finally, we need to make sure that by making the partition finer and finer, we are increasing our chances of finding the right answer to whether P is copositive or not. In other words, we have to be sure that if the matrix P is copositive then there exists a simplical partition with small enough diameter such that (4.33) holds for the vertices of each simplex contained in the partition.

Theorem 4.25. *Let $P \in \mathbb{R}^{n\times n}$ be strictly copositive. Then there exists $\varepsilon > 0$ such that for all finite simplicial partitions $\mathcal{P}_m = \{\Delta^1, \cdots, \Delta^m\}$ of Δ^S with $\delta(\mathcal{P}) < \epsilon$, it holds that for each $k \in \{1, \ldots, m\}$,*

$$(v_i^k)^\top P v_j^k \geq 0, \qquad \text{for all } i, j \in \{1, \cdots, n\}.$$

where $v_1^k, \ldots, v_n^k$ denote the vertices of simplex Δ^k.

4.4.3 Generalizations and converse result

In Example 4.6, we saw that a positive definite Lyapunov function may not exist for complementarity systems even if it is asymptotically stable. In other words, the class of positive definite functions is too restrictive for complementarity systems. We then saw some sufficient conditions which allow us to find a quadratic copositive Lyapunov function. One may ask the question whether there always exists a copositive Lyapunov function for asymptotically stable complementarity systems. We provide an affirmative answer to this question in [197] and here we recall some results based on that work. In particular, the notion of copositive Lyapunov functions is generalized for the systems of the form

$$\dot{x} \in f(x) - \mathcal{N}_K(x) \tag{4.34}$$

where K is a closed convex cone in $\mathbb{R}^n$. One can equally write the system as a complementarity system, or a projected dynamical system.

Definition 4.26 (Cone-Copositive Lyapunov Function)**.** For system (4.34) with the equilibrium at the origin, we say that $V \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ is a (global) cone-copositive Lyapunov function with respect to K, if

- There exist class $\mathcal{K}_\infty$ functions $\underline{\alpha}$, $\overline{\alpha}$ such that

$$\underline{\alpha}(|x|) \leq V(x) \leq \overline{\alpha}(|x|), \quad \forall\, x \in K$$

- It holds that

$$\langle \nabla V(x), f(x) \rangle < 0, \quad \forall\, x \in \operatorname{int}(K)$$
$$\langle \nabla V(x), f(x) + \eta \rangle < 0, \quad \forall x \in \operatorname{bd}(K),$$

  and η solves the complementarity relation

$$K^\star \ni \eta \perp f(x) + \eta \in \mathcal{T}_K(x), \qquad \eta \in \operatorname{bd}(K). \tag{4.35}$$

For brevity, we will use the notation $\eta \in \mathrm{LCCP}(f(x), I, \mathcal{T}_K(x))$ for the solution of the complementarity problem (4.35).

Theorem 4.27. *Consider system* (4.34) *having an equilibrium point at the origin, and suppose that f is Lipschitz continuous, and $K \subset \mathbb{R}^n$ is a closed convex cone. If the origin is globally exponentially stable for system* (4.34)*, then there exists a continuously differentiable cone-copositive Lyapunov function.*

For the proof of Theorem 4.27, we construct the Lyapunov function for (4.34) by introducing a function $V : \mathbb{R}^n \to \mathbb{R}$, defined as

$$V(z) = \int_0^\infty \|x(\tau; \operatorname{proj}_K(z))\|^{\frac{2L}{\alpha}+1} \, d\tau, \tag{4.36}$$

where $x(\tau; \overline{z})$ denotes the solution to system (4.34) at time $\tau \geq 0$ with $x(0^+) = \overline{z} \in K$, L is the Lipschitz modulus of $f(\cdot)$, and α is the uniform decay rate for system trajectories. Note that V is defined for each $z \in \mathbb{R}^n$ and not just for $z \in K$. When $z \notin K$, the term $x(\tau; \operatorname{proj}_K(z))$ can be interpreted as the solution obtained by projecting the initial condition on K, and then propagating it continuously according to the system vector field. Thus, for $\tau > 0$, we have $x(\tau; \operatorname{proj}_K(z)) = x(\tau; z)$ for each $z \in \mathbb{R}^n$.

Following main steps are involved in proving Theorem 4.27:

- Get the appropriate lower bounds
- Show that V is Lipschitz continuous
- Compute the infinitesimal change in the value of V when evaluated along the system trajectories
- Smoothen V to make it continuously differentiable

Bounds on Solutions

The following lemma demonstrates the continuity of solutions with respect to the initial conditions and plays an important role in the remainder of the proof.

Lemma 4.28. *Let L be the Lipschitz modulus of f. If x and $\widehat{x}$ are two solutions to system* (4.34) *that satisfy $x(0) = z \in K$ and $\widehat{x}(0) = \widehat{z} \in K$, then it holds that, for each $\tau > 0$,*

$$\|x(\tau; z) - \widehat{x}(\tau; \widehat{z})\| \leq e^{L\tau} \|z - \widehat{z}\| \tag{4.37a}$$

and for some $C > 0$,

$$\|x(\tau; z)\| \geq e^{-C\tau} \|z\|. \tag{4.37b}$$

Proof. It will be assumed without loss of generality that $z \in K$ and $\widehat{z} \in K$ since $\|\operatorname{proj}_K(z) - \operatorname{proj}_K(\widehat{z})\| \leq \|z - \widehat{z}\|$. By definition of the solution to (4.34) and monotonicity of the normal cone operator, it follows that, for each $y \in K$,

$$\left\langle \frac{dx}{dt}(t) - f(x(t)), y - x(t) \right\rangle \geq 0$$

and similarly, for each $\widehat{y} \in K$,

$$\left\langle \frac{d\widehat{x}}{dt}(t) - f(\widehat{x}(t)), \widehat{y} - \widehat{x}(t) \right\rangle \geq 0,$$

where we have suppressed the dependence of x and $\widehat{x}$ on the initial condition for brevity. Letting $y = \widehat{x}(t) \in K$, and $\widehat{y} = x(t) \in K$, we get the following by adding the last two inequalities:

$$\left\langle \frac{d}{dt}(x(t) - \widehat{x}(t)), x(t) - \widehat{x}(t) \right\rangle \leq \langle f(x(t)) - f(\widehat{x}(t)), x(t) - \widehat{x}(t) \rangle, \tag{4.38}$$

or equivalently,

$$\frac{d}{dt}\|x(t)-\widehat{x}(t)\|^2 \leq 2\left\langle f(x(t))-f(\widehat{x}(t)), x(t)-\widehat{x}(t)\right\rangle.$$

Because of the Lipschitz continuity assumption, $\|f(x(t))-f(\widehat{x}(t))\| \leq L\|x(t)-\widehat{x}(t)\|$, and hence,

$$\frac{d}{dt}\|x(t)-\widehat{x}(t)\|^2 \leq 2L\|x(t)-\widehat{x}(t)\|^2.$$

The bound in (4.37a) now follows by integrating both sides, or invoking the so-called comparison lemma [121, Lemma 3.4]. To get the bound in (4.37b), we make use of [197, Proposition 19] which ensures that there exists a constant $C_\eta > 0$ such that $|\eta| \leq C_\eta |f(x(t))|$. We therefore get

$$\left|\frac{d}{dt}\|x(t)\|^2\right| = 2\left|\langle x(t), f(x(t))+\eta\rangle\right| \leq 2L(1+C_\eta)\|x(t)\|^2.$$

In particular, $\frac{d}{dt}\|x(t)\|^2 \geq -2L(1+C_\eta)\|x(t)\|^2$, and hence, the inequality in (4.37b) follows by taking $C = L(1+C_\eta)$. □

To show that V satisfies item 1) of Definition 4.26, let us first use the bound in (4.37b) from Lemma 4.28, so that

$$\begin{aligned} V(z) &\geq \int_0^\infty e^{-(2L+\alpha)C\tau/\alpha}\|\operatorname{proj}_K(z)\|^{\frac{2L}{\alpha}+1}\,d\tau \\ &\geq \underline{C}\|\operatorname{proj}_K(z)\|^{\frac{2L}{\alpha}+1}, \end{aligned}$$

for some $\underline{C} > 0$. Also, exponential stability of the origin implies that $\|x(\tau;z)\| \leq c_0 e^{-\alpha\tau}\|\operatorname{proj}_K(z)\|$ and hence there exists $\overline{C} > 0$

$$\begin{aligned} V(z) &\leq c_0\int_0^\infty e^{-(2L+\alpha)\tau}\|\operatorname{proj}_K(z)\|^{\frac{2L}{\alpha}+1}\,d\tau \\ &\leq \overline{C}\,\|\operatorname{proj}_K(z)\|^{\frac{2L}{\alpha}+1}. \end{aligned}$$

Local Lipschitz Continuity of V

To show that V is locally Lipschitz continuous, we need the following two properties [50]:

- V is continuous; and
- its Dini subderivative[2] satisfies

$$DV(z;v) \leq \phi(z)\|v\| \tag{4.39}$$

  for every $v \in \mathbb{R}^n$, every $z \in \mathbb{R}^n$, and some locally bounded function $\phi : \mathbb{R}^n \to \mathbb{R}$, with $\phi(z) > 0$ for $z \neq 0$.

[2]The Dini subderivative of V at x in the direction v is defined as

$$DV(x;v) := \liminf_{w\to v,\varepsilon\to 0^+}\frac{V(x+\varepsilon w)-V(x)}{\varepsilon}.$$

The continuity of V follows directly from Lemma 4.28 as the exponential bound on the solutions of the system makes V a composition of continuous functions. These properties can again be shown using Lemma 4.28. Fix $v \in \mathbb{R}^n$. Consider a sequence of initial conditions $\widehat{z}_k = z + \varepsilon_k v$. We get

$$\begin{aligned} DV(z;v) &\leq \liminf_{\varepsilon_k \to 0} \frac{V(z+\varepsilon_k v) - V(z)}{\varepsilon_k} \\ &= \liminf_{\varepsilon_k \to 0} \frac{1}{\varepsilon_k} \left(\int_0^\infty (\|\widehat{x}_k(\tau;\widehat{z}_k)\|^{\frac{2L}{\alpha}+1} - \|x(\tau;z)\|^{\frac{2L}{\alpha}+1}) d\tau \right). \end{aligned} \tag{4.40}$$

Using the mean-value theorem, for each $s \geq 0$, there exists $\xi(s)$ between $\|\widehat{x}_k(s;\widehat{z}_k)\|$ and $\|x(s;z)\|$ such that

$$\begin{aligned} \|\widehat{x}_k(s;\widehat{z}_k)\|^{\frac{2L}{\alpha}+1} - \|x(s;z)\|^{\frac{2L}{\alpha}+1} &\leq \left| \, \|\widehat{x}_k(s;\widehat{z}_k)\|^{\frac{2L}{\alpha}+1} - \|x(s;z)\|^{\frac{2L}{\alpha}+1} \right| \\ &= \left| \, \xi(s)^{2L/\alpha} (\|\widehat{x}_k(s;\widehat{z}_k)\| - \|x(s;z)\|) \right| \\ &\leq |\xi(s)|^{2L/\alpha} \|\widehat{x}_k(s;\widehat{z}_k) - x(s;z)\|. \end{aligned} \tag{4.41}$$

It follows from Lemma 4.28 that $\|\widehat{x}_k(s;\widehat{z}_k) - x(s;z)\| \leq e^{Ls}\varepsilon_k \|v\|$. Substituting these bounds in (4.40), we get

$$DV(z;v) \leq \|v\| \int_0^\infty e^{Ls} |\xi(s)|^{2L/\alpha} \, ds.$$

Due to the exponential stability assumption, $\|\xi(s)\| \leq \widehat{c}\, e^{-\alpha s}\|z\|$, for some $\widehat{c} > 0$, and hence we choose

$$\phi(z) = \widehat{c}\, \|z\|^{2L/\alpha} \int_0^\infty e^{-Ls} \, ds,$$

so that the bound (4.39) is seen to hold. Thus, V is locally Lipschitz continuous.

Infinitesimal Decrease in V

As the next step, we show that the function V decreases along the system vector field. In what follows, we will denote the right-hand side of (4.34) by $F(z)$, so that

$$F(z) \in f(z) - \mathcal{N}_K(z).$$

The function V in (4.36) is differentiable almost everywhere because it is locally Lipschitz continuous. We next show that the Dini subderivative of V, along $F(z)$ is negative definite.

Lemma 4.29. *For the function $V : \mathbb{R}^n \to \mathbb{R}$ in* (4.36)*, and $z \in K$,*

$$DV(z;F(z)) \leq -\|z\|^{\frac{2L}{\alpha}+1}. \tag{4.42}$$

Proof of Lemma 4.29. To prove (4.42), we need a bound on $V(z) - V(z + tF(z))$ for $t \geq 0$ sufficiently small. We will get the desired bounds by rewriting the difference as

$$V(z+tF(z)) - V(z) = \big[V(z+tF(z)) - V(x(t;z))\big] + \big[V(x(t;z)) - V(z)\big] \tag{4.43}$$

and getting a bound on each of the two difference terms on the right-hand side. The first term $V(z + tF(z)) - V(x(t;z))$ can be analyzed from the following lemma:

Lemma 4.30. *For $t > 0$ sufficiently small, it holds that*

$$\|z + tF(z) - x(t; z)\| \leq o(t) \tag{4.44}$$

for each $z \in K$.

The proof of Lemma 4.30 will follow momentarily. Using the estimate (4.44), and the inequalities (4.40) and (4.41), we get

$$\begin{aligned} V(z + tF(z)) - V(x(t; z)) &\leq C_\phi \|z + tF(z) - x(t; z)\| \\ &= o(t), \end{aligned}$$

for a fixed $z \in K$, and some $C_\phi > 0$. For the second term on the right-hand side of (4.43), it follows from the definition of V in (4.36), with $x(0) = z$, that

$$V(z) \geq \int_0^t \|x(\tau; z)\|^{\frac{2L}{\alpha}+1} d\tau + \int_0^\infty \|x(\tau; x(t; z))\|^{\frac{2L}{\alpha}+1} d\tau$$

and hence

$$V(x(t; z)) - V(z) \leq -\int_0^t \|x(\tau; z)\|^{\frac{2L}{\alpha}+1} d\tau. \tag{4.45}$$

Substituting the bounds from (4.44) and (4.45) in (4.43), we get

$$\begin{aligned} \liminf_{t\to 0^+} \frac{V(z + tF(z)) - V(z)}{t} &\leq \liminf_{t\to 0^+} \frac{-\int_0^t \|x(\tau; z)\|^{\frac{2L}{\alpha}+1} d\tau}{t} \\ &= \liminf_{t\to 0^+} -\|x(t; z)\|^{\frac{2L}{\alpha}+1} = -\|z\|^{\frac{2L}{\alpha}+1}, \end{aligned}$$

and hence the Dini subderivative of V is negative definite for almost every $z \in K$. $\square$

Proof of Lemma 4.30. By definition, the solution x of system (4.34), with $x(0) = z \in K$, satisfies

$$\langle \dot{x}(t) - f(x(t)), x(t) - y \rangle \leq 0, \quad \forall\, y \in K, \tag{4.46}$$

for almost all $t \geq 0$. For $h > 0$ small enough, introduce the function $\widetilde{z} : [0, h] \to \mathbb{R}^n$ given by

$$\widetilde{z}(t) := z + tF(z) = z + tf(z) + t\eta_z,$$

where η_z is such that $\eta_z = 0$ for $z \in \operatorname{int}(K)$ and $\eta_z \in \operatorname{LCCP}(f(z), I, \mathcal{T}_K(z))$ for $z \in \operatorname{bd}(K)$. It is readily checked that $\widetilde{z}(t) \in K$ for all $t \in [0, h]$, and $\dot{\widetilde{z}} = \frac{d}{dt}\widetilde{z}(t) = F(z) = f(z) + \eta_z$. From the definition of $F(z)$, it follows that $\left\langle f(z) - \dot{\widetilde{z}}\,, \widetilde{y} - z \right\rangle \leq 0$, for all $\widetilde{y} \in K$, or equivalently,

$$\left\langle f(z) - \dot{\widetilde{z}}\,, \widetilde{y} - \widetilde{z}(t) \right\rangle \leq \langle F(z) - f(z), \widetilde{z}(t) - z \rangle\,, \forall\, \widetilde{y} \in K.$$

For $t > 0$ small enough, we have $(x(t) - \widetilde{z}(t) + t\widetilde{z}(t)) \in K$. Since K is a cone, we can take $\widetilde{y} = \frac{1}{t}\left(x(t) - \widetilde{z}(t) + t\widetilde{z}(t)\right) \in K$ to get

$$\left\langle f(z) - \dot{\widetilde{z}}\,, x(t) - \widetilde{z}(t) \right\rangle \leq t\, \langle \eta_z, \widetilde{z}(t) - z \rangle\,.$$

Taking $y = \widetilde{z}(t)$ in (4.46), and adding it to the last inequality, we get

$$\left\langle \dot{x}(t) - \dot{\widetilde{z}}, x(t) - \widetilde{z}(t) \right\rangle \leq \langle f(x(t)) - f(z), x(t) - \widetilde{z}(t) \rangle + t\, \langle \eta_z, \widetilde{z}(t) - z \rangle\,. \tag{4.47}$$

To bound the terms on the right-hand side, we observe that

$$\begin{aligned}\|f(z) - f(x(t))\| &\leq L\|z - x(t)\| \\ &\leq L\|z - \widetilde{z}(t)\| + L\|\widetilde{z}(t) - x(t)\| \\ &\leq Lt\|F(z)\| + L\|\widetilde{z}(t) - x(t)\|.\end{aligned}$$

Using [197, Proposition 19], there is some constant C_η such that

$$\langle \eta_z, \widetilde{z}(t) - z\rangle = \langle \eta_z, tF(z)\rangle \leq t\,\|\eta_z\|\,\|F(z)\| \leq C_\eta t\|F(z)\|.$$

Consequently, we get

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|x(t;z) - \widetilde{z}(t)\|^2 = \left\langle \dot{\widetilde{z}} - \dot{x}(t), \widetilde{z}(t) - x(t)\right\rangle \\ &\leq \|f(z) - f(x(t))\| \cdot \|\widetilde{z}(t) - x(t)\| + C_\eta t^2\|F(z)\| \\ &\leq L\|z - x(t)\| \cdot \|\widetilde{z}(t) - x(t)\| + C_\eta t^2\|F(z)\| \\ &\leq Lt\|F(z)\|\|\widetilde{z}(t) - x(t)\| + L\|\widetilde{z}(t) - x(t)\|^2 + C_\eta t^2\|F(z)\| \\ &\leq C_1\|\widetilde{z}(t) - x(t)\|^2 + C_{2,z}t^2\end{aligned}$$

where we used Young's inequality for the product term $Lt\|F(z)\| \cdot \|\widetilde{z}(t) - x(t)\|$, and chose $C_1 = (L + 0.5L^2)$ and $C_{2,z} = \max\{0.5\|F(z)\|^2, C_\eta\|F(z)\|\}$. Solving the differential inequality, and using the fact that, $\widetilde{z}(0) = x(0)$, we get

$$\|\widetilde{z}(t) - x(t)\|^2 \leq 2C_{2,z}\int_0^t \exp(2C_1(t-s))s^2\,ds.$$

Solving the integral on the right, we get

$$\|\widetilde{z}(t) - x(t)\|^2 \leq C_{3,z}\,t^3 + o(t^3),$$

for some $C_{3,z} > 0$, whence the estimate in (4.44) follows. □

Regularization of V

The next step is to regularize V so that we obtain a continuously differentiable Lyapunov function. The construction given next yields a $\mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ Lyapunov function. This regularization technique is inspired by [49], and we omit the details. It has also been used for smoothening of locally Lipschitz Lyapunov functions for hybrid systems [80, Chapter 7] and switched systems [65].

4.4.4 Computational aspects

Under some additional assumptions on the system vector field f, we can get additional properties for the Lyapunov function which are desirable for computation. In particular, if f is homogenous, then one can find a homogenous Lyapunov function as well.

Proposition 4.31. *Consider dynamical system* (4.34) *with f homogeneous and the origin locally exponential stable. Then, there exists a cone-copositive Lyapunov function $\overline{W}$ for* (4.34)*, of class $\mathcal{C}^{k-1}$ on $\mathbb{R}^n\backslash\{0\}$, and it satisfies*

$$\overline{W}(sx) = s^k\overline{W}(x)$$

for all $x \in \mathbb{R}^n\backslash\{0\}$ and $s > 0$.

The advantage of the homogenous Lyapunov function is that we can reduce our search to compute a Lyapunov function over a unit ball. Moreover, we can approximate its value, and that of its derivative, by a homogenous polynomial Lyapunov function, which can be parameterized by finitely many unknowns if the degree is fixed.

Proposition 4.32. *Let $W \in \mathcal{C}^1(\mathbb{R}^n; \mathbb{R}_+)$ be a homogeneous function of degree d and $\epsilon > 0$ be a given scalar. There exist an even integer r and a homogeneous polynomial p of degree $r + d$, such that*

$$\max\left\{\max_{x\in S^{n-1}} \left|\widetilde{W}(x)\right|, \max_{x\in S^{n-1}} \left\|\nabla\widetilde{W}(x)\right\|\right\} \leq \epsilon$$

where S^{n-1} denotes the unit sphere in $\mathbb{R}^n$ and $\widetilde{W}(x) = W(x) - \frac{p(x)}{\|x\|^r}$.

With such a rational function in hand which approximates the homogeneous function from Proposition 4.31 (in terms of value and gradient) to desired accuracy, one can establish the existence of a rational homogeneous cone-copositive Lyapunov function.

We can establish the inequalities which will be used in our algorithms to find copositive homogeneous Lyapunov functions. We restrict our attention to full-dimensional polyhedral cones, that is, $K = \{x \in \mathbb{R}^n | Fx \geq 0\}$ with non-empty interior. By using Proposition 4.32, let

$$V(x) = \frac{h(x)}{(\sum_{i=1}^n x_i^2)^r} = \frac{h(x)}{\|x\|_2^{2r}}$$

where r is a nonnegative integer, and $h(\cdot)$ is a homogeneous polynomial of degree at least $2r + 1$, copositive on K. Here, we used the notation that $x = (x_1, x_2, \ldots, x_n)^\top \in K$.

We now establish the inequalities which will be used in our algorithms to find copositive homogeneous Lyapunov functions. We restrict our attention to full-dimensional polyhedral cones, that is, $K = \{x \in \mathbb{R}^n | Fx \geq 0\}$ with non-empty interior. By using Proposition 4.32, let

$$V(x) = \frac{h(x)}{(\sum_{i=1}^n x_i^2)^r} = \frac{h(x)}{\|x\|_2^{2r}}$$

where r is a nonnegative integer, and $h(\cdot)$ is a homogeneous polynomial of degree at least $2r + 1$, copositive on K. Here, we used the notation that $x = (x_1, x_2, \ldots, x_n)^\top \in K$.

As we know, finding such Lyapunov function is equivalent to finding V that satisfies the inequalities:

$$V(x) = \frac{h(x)}{\|x\|_2^{2r}} \geq 0, \;\; \forall x \in K\backslash\{0\} \tag{4.48a}$$

$$-\left\langle \nabla V(x), f(x) + \eta \right\rangle \geq 0, \;\; \forall x \in K\backslash\{0\} \tag{4.48b}$$

where

$$\begin{aligned} &-\left\langle \nabla V(x), f(x) + \eta \right\rangle \\ &= \frac{-\|x\|_2^2 \left\langle \nabla h(x), f(x) + \eta \right\rangle + 2rh(x) \left\langle x, f(x) + \eta \right\rangle}{\|x\|_2^{2(r+1)}}. \end{aligned}$$

with $\eta = \mathrm{LCCP}(f(x), I, \mathcal{T}_K(x))$. The numerator is denoted by

$$s(x) := -\|x\|_2^2 \left\langle \nabla h(x), f(x) + \eta \right\rangle + 2rh(x) \left\langle x, f(x) + \eta \right\rangle$$

which is a homogeneous polynomial if h and f are homogeneous polynomials. So we have

$$V(x) = \frac{h(x)}{\|x\|_2^{2r}} \geq 0, \;\; \forall x \in K\backslash\{0\} \tag{4.49a}$$

$$-\langle \nabla V(x), f(x)+\eta \rangle = \frac{s(x)}{\|x\|_2^{2(r+1)}} \geq 0, \ \ \forall x \in K\backslash\{0\}. \tag{4.49b}$$

Thus, finding a copositive V for system (4.34) with the structure imposed in Proposition 4.32 boils down to finding h and s such that

$$h(x) \geq 0, \ \ \forall x \in K\backslash\{0\} \tag{4.50a}$$
$$s(x) \geq 0, \ \ \forall x \in K\backslash\{0\}. \tag{4.50b}$$

Since η is nonzero only on the boundaries of K, we replace the second inequality in (4.50) by inequalities with respect to each face of polyhedron K. Let $S_i := \{x \in K \,|\, (Fx)_i = 0\}$, $i \in \{1, \ldots, n_K\}$ denote the faces of K. Let

$$s_i(x) = -\|x\|_2^2 \langle \nabla h(x), f(x)+\eta_i \rangle + 2rh(x) \langle x, f(x)+\eta_i \rangle$$

for all $x \in S_i$ where $\eta_i = \mathrm{LCCP}(f(x), I, \mathcal{T}_K(x))$. In the interior of K, we have $\eta = 0$ so let

$$s_0(x) = -\|x\|_2^2 \langle \nabla h(x), f(x) \rangle + 2rh(x) \langle x, f(x) \rangle .$$

Consequently, the inequalities used for finding V can be rewritten as follows:

$$h(x) \geq 0, \ \ \forall x \in K\backslash\{0\} \tag{4.51a}$$
$$s_0(x) \geq 0, \ \ \forall x \in \mathrm{int}(K\backslash\{0\}) \tag{4.51b}$$
$$s_i(x) \geq 0, \ \ \forall x \in S_i, \ \ i \in \{1, \ldots, n_K\}, \tag{4.51c}$$

where $S_i := \{x \in K \,|\, (Fx)_i = 0\}$, $i \in \{1, \ldots, n_K\}$ denote the faces of K.

For computing cone-copositive Lyapunov functions V, we study two techniques for finding polynomials that satisfy (4.51). In what follows, we assume that $K = \mathbb{R}^n_+$, that is, K is the positive orthant with n faces so that $S_i = \{x \in \mathbb{R}^n_+ \,|\, x_i = 0\}$, $i = 1, \ldots, n$. The more general case of polyhedral cones can be covered with state transformations or decompositions but such details are being avoided for the ease of exposition.

Discretization Method

The basic idea behind the discretization methods is to select a certain number of points in the cone $\mathbb{R}^n_+$ and evaluate the inequalities (4.51) with a certain polynomial function parameterized by finitely many unknowns. This allows us to construct an inner approximation of copositive polynomials with respect to cone $\mathbb{R}^n_+$. In the literature, there exist algorithms for checking copositivity of a matrix using discretization methods [30, 31, 70] which we presented briefly in Section 4.4.2. Here, we restrict ourselves to discretization schemes and generalize the existing algorithms for arbitrary nonlinear polynomials (not necessarily quadratic functions).

To describe this discretization algorithm, let us first consider the convex cone of copositive polynomials

$$\mathcal{C} := \left\{ g \in \mathbb{R}^d[x] \,\middle|\, \begin{array}{l} g \text{ is homogeneous and} \\ g(x) \geq 0 \ \text{ for all } \ x \in \mathbb{R}^n_+ \end{array} \right\}, \tag{4.52}$$

where $\mathbb{R}^d[x]$ denotes the ring of polynomials of degree d, over the field of reals, in $x \in \mathbb{R}^n$. We will establish an inner approximation of $\mathcal{C}$ based on simplicial partitions inside cone $\mathbb{R}^n_+$. To do so, we first need to introduce *tensors*, which generalize the notion of a matrix, and will be used for compact representation of polynomials of our interest.

Definition 4.33. A tensor $\mathcal{A}$ of order d over $\mathbb{R}^n$ is a multilinear form

$$\begin{array}{ccc} \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{d \text{ times}} & \to & \mathbb{R} \\ (x^1, x^2, \ldots, x^d) & \mapsto & \mathcal{A}[x^1, x^2, \ldots, x^d] \end{array}$$

where

$$\mathcal{A}[x^1, x^2, \ldots, x^d] = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_d=1}^{n} a_{i_1,i_2,\ldots,i_d} x^1_{i_1} \cdots x^d_{i_d}$$

and $a_{i_1,i_2,\ldots,i_d}$ corresponds to a real number from a table with n^d entries, indexed by $i_1, i_2, \ldots, i_d \in \{1, \ldots, n\}$. We say that $\mathcal{A}$ is symmetric if

$$a_{i_1,i_2,\ldots,i_d} = a_{j_1,j_2,\ldots,j_d}$$

whenever $i_1 + i_2 + \cdots + i_d = j_1 + j_2 + \cdots + j_d$, for all possible permutations $i_1, i_2, \ldots, i_d$ and $j_1, j_2, \ldots, j_d$ of $\{1, \ldots, n\}$.

A matrix $A \in \mathbb{R}^{n \times n}$ describes a tensor of order 2 over $\mathbb{R}^n$, also called a quadratic form, where the coefficients of the quadratic form belong to a table with n^2 entries $a_{i,j}$ with $i, j = \{1, \ldots, n\}$. A general homogeneous polynomial $g \in \mathbb{R}^d[x]$, with $d \geq 2$, can be written as

$$g(x) = g(x_1, \ldots, x_n) = \sum_{\substack{i=(i_1,\ldots,i_n) \\ i_1+\cdots+i_n=d}} a_i x_1^{i_1} \cdots x_n^{i_n}.$$

Using the tensor representation, g can also be compactly written in the form $g(x) = G[\underbrace{x, x, \ldots, x}_{d \text{ times}}]$, where G is a symmetric tensor. One can now use the techniques presented in Section 4.4.2 to compute the polynomial function g which satisfies the inequalities characterizing stability over the positive orthant.

4.4.5 Sum-of-squares method

A commonly employed tool for checking the positivity of a polynomial is to write it in the form of a sum-of-squares (SOS) of other polynomials. While testing positivity is a computationally hard problem, the question of finding an SOS decomposition of a polynomial is actually a semidefinite program [169]. The crux of such ideas can be found in [163] and its application to copositivity is sketched in [164].

The basic idea is to get rid of the constraint $x \in \mathbb{R}^n_+$. We let $x_i = y_i^2$, $i \in \{1, \cdots, n\}$ be the change of variable where y^2 is the short-hand for $(y_1^2, \ldots, y_n^2)$. Clearly we have

$$h(x) \geq 0, \ \forall x \in \mathbb{R}^n_+ \iff h(y^2) \geq 0, \ \forall y \in \mathbb{R}^n.$$

Then, the inequalities (4.51) are rewritten as follows

$$P_h(y) := h(y^2) \geq 0, \ \forall y \in \mathbb{R}^n \tag{4.53a}$$

$$P_{s_0}(y) := s_0(y^2) \geq 0, \ y_i \neq 0, \forall i \tag{4.53b}$$

$$P_{s_i}(y) := s_i(y^2) \geq 0, \ y_i = 0, \ i \in \{1, ..., n\} \tag{4.53c}$$

where h, s_o, s_i are homogeneous polynomials.

Next, we define the polynomials

$$P_h^{(d)}(y) := \|y\|^{2d} P_h(y) \tag{4.54a}$$

$$P_{s_0}^{(d)}(y) := \|y\|^{2d} P_{s_0}(y) \tag{4.54b}$$
$$P_{s_i}^{(d)}(y) := \|y\|^{2d} P_{s_i}(y) \tag{4.54c}$$

where d is an integer. It is obvious that inequalities (4.53) are satisfied if and only if

$$P_h^{(d)}(y) \geq 0, \;\; \forall y \in \mathbb{R}^n \tag{4.55a}$$
$$P_{s_0}^{(d)}(y) \geq 0, \;\; y_i \neq 0, \forall i \tag{4.55b}$$
$$P_{s_i}^{(d)}(y) \geq 0, \;\; y_i = 0, \;\; i \in \{1, ..., n\}\,. \tag{4.55c}$$

Proposition 4.34. *For the homogeneous copositive functions h, s_0 and s_i, $i \in \{1, \ldots, n\}$, there exists $d \in \mathbb{N}^*$ sufficiently large such that the polynomials $P_h^{(d)}$, $P_{s_0}^{(d)}$ and $P_{s_i}^{(d)}$ are SOS.*

Proof. We will carry out the proof only for h and it will be similar for the other polynomials. Let

$$K_n^d := \left\{ h \in \mathbb{R}[x] |\; P_h^{(d)} \;\text{ SOS} \right\}$$
$$C_n^d := \left\{ h \in \mathbb{R}[x] |\; P_h^{(d)} \;\text{ has positive coefficients} \right\}.$$

We notice that $C_n^d \subseteq K_n^d$ because if $P_h^{(d)}(y)$ has only positive coefficients then the polynomial $P_h(y) = h(y^2)$ is SOS and since $P_h(y)$ is multiplied by $\|y\|^{2d}$, it follows that $P_h^{(d)}(y)$ is SOS. So we just need to prove that $P_h^{(d)}(y)$ has positive coefficients.

The copositivity of h is equivalent to the positivity of P_h. And since h is homogeneous, this will be also equivalent to the positivity of P_h on the unit ball which means $P_h(y) \geq 0$, $\forall y \in \mathbb{R}^n$, $\sum_{i=1}^n y_i^2 = 1$. By substituting y_i^2 by z_i, we obtain $P_h(z) \geq 0$, $\forall z \geq 0$, $\sum_{i=1}^n z_i = 1$.

Let us now recall Pólya's Theorem, see [129] and [168] for the proof.

Theorem 4.35. (Pólya's Theorem) *Let $f \in \mathbb{R}[x]$ be homogeneous. If $f \geq 0$ on the simplex $\{x \geq 0 |\; \sum_{i=1}^n x_i = 1\}$, then there exists a sufficiently large $\ell \in \mathbb{N}$ for which the polynomial $(\sum_{i=1}^n x_i)^\ell f(x)$ has all its coefficients nonnegative.*

Applying this theorem to the homogeneous polynomial $P_h(z)$, we obtain that for sufficiently large $d \in \mathbb{N}$, all the coefficients of the polynomial $P_h^{(d)}(z) = (\sum_{i=1}^n z_i)^d P_h(z)$ are positive. Then $P_h^{(d)}$ is SOS in view of the fact that $C_n^d \subseteq K_n^d$. □

To sum up this section, the foregoing result allows us to write an algorithm to compute the polynomials $P_h^{(d)}$, $P_{s_0}^{(d)}$ and $P_{s_i}^{(d)}$ in the form of SOS, the result is then used to get a homogeneous copositive Lyapunov function.

4.5 Perspectives

In this chapter, we considered a class of dynamical systems which allows us to model constraints on the state trajectories. We broadly use the framework of differential inclusion that includes a drift term and a maximal monotone mapping. In particular, we saw their utility in modeling piecewise affine relations, relations involving variational inequalities, and projection-based dynamics. The interconnections of such systems have also appeared in the literature in different forms, and we aim to develop analysis techniques for interconnections of such systems. The stability analysis itself has some unique characteristics due to the role of the maximal monotone mapping, and we proposed some numerical algorithms to compute Lyapunov functions for a specific subclass of such systems. Several interesting research directions emanate from the work reported in this chapter, which could be grouped as follows:

Stability for broader class of complementarity relations

This chapter addressed the stability analysis for a class of complementarity systems using the method of Lyapunov functions. Questions pertaining to the existence of cone-copositive Lyapunov functions were answered in the affirmative for exponentially stable systems. Some refinements of this result, under certain conditions on the vector field in the system dynamics, allow us to restrict our search for cone-copositive Lyapunov functions within the class of homogeneous and rational polynomials. These statements indeed bring tractability to the numerical methods that have been proposed in this chapter for computing Lyapunov functions. In particular, two hierarchies of convex optimization problems are obtained using the methods based on discretization and SOS approximation, respectively.

The first direction of research is to extend our results to broader classes of systems with maximal monotone mappings as reported in [27]. Lately, we have seen several cases where the control action is modeled as the outcome of an optimization algorithm, or involves some neural network architecture based on Rectified Linear (ReLU) activation functions. We presented several instances of modeling such static relations using the complementarity framework, or as a particular case of maximal monotone mapping. A natural question therefore would be to extend the framework of stability analysis to be able to address such feedback control systems.

Here, we also point out that some extensions at the level of designing algorithms are also of potential interest. At this moment our algorithms are specifically adapted to the constraint sets of the form positive orthant or an invertible linear transformation of such sets. Adapting this technique to more generic sets remains to be seen. Also, in our current treatment, we have considered discretization algorithms in $\mathbb{R}^2$ and $\mathbb{R}^3$, where it is relatively straightforward to write algorithms for partition of simplices. It remains to be seen how the algorithms for simplicial partition in higher dimensions perform in computing Lyapunov functions.

Optimal control of constrained systems

We did not address the question of designing controllers in this chapter. Knowing that the proposed framework allows us to model constraints, the question of designing controllers with constraints on state trajectories and inputs has also been addressed to some extent in the literature [209, 142, 94]. At the same time, there has been a certain interest in designing optimal controllers for the systems with normal cone mappings from analysis perspective [28, 51, 97], as well as numerical side [200, 220, 159].

Our efforts in this direction are based on considering a certain relaxation of the optimal control problem in the space of measures [20, 128, 221]. This relies on studying the evolution of proposed system class in the space of measures where the initial condition is described by a probability measure. This leads to a description in the form of a continuity equation and for the particular class of systems addressed in this chapter, we have developed this framework in [198, 44]. Our initial investigations about the questions related to optimal control have been reported in [43]. We are currently investigating several questions which relate to the existence of optimal solutions, and how one can construct their approximation via appropriate discretization schemes. Even for the discretized version, the problem has to be solved numerically in the space of moments, which is infinite dimensional. We are currently working on using certain algorithms based on semidefinite programming that allow us to compute the moments of the optimal solution.

Optimization algorithms

The question of interconnections that we studied in this chapter was primarily motivated by the idea of looking at the continuous-time version of certain optimization algorithms, which in steady state converge to the equilibrium points characterized by KKT conditions for certain programs. Several interesting connections have appeared recently between the use of optimization algorithms, maximal monotone mappings, and robust control [187, 226]. It remains to be seen whether we can adopt the viewpoint of looking at these algorithms from the viewpoint of interconnections and develop the analysis tools that allow us to study the convergence of optimization algorithms and lead to control synthesis techniques.

5

Ensemble Filters and Bayesian Inference

For several applications ranging from robotics to meteorology, filtering algorithms have been widely used as a reliable tool for estimation purposes with uncertain models and noisy observations. The early developments in the area of filtering can be traced back to [115], where the authors provided an elegant recursive closed form solution for optimal state estimation in linear stochastic systems. Since this pioneering work, the field has evolved enormously to develop algorithms for nonlinear dynamical systems, address different information structures, and provide computationally efficient methods for large-scale systems. An excellent compilation of different developments on the topic of filtering appears in [55].

At the same time, stochastic control under information constraints is another relevant direction of research which has grown considerably in the past two decades due to common use of communication channels between different components of the system. In such cases, it becomes important to adapt the policies to achieve optimality under given information structure and analyze the performance of such policies [145, 231, 18]. Modern control systems often require us to carry out the estimation task remotely, when the measurements from the sensors are communicated over a network. Thus, an important research direction is to study the filtering problem with constraints on the information available for computing the optimal distribution. In particular, motivated by the idea of implementing filters subject to observations transmitted over networks through some communication protocols, it is natural to stipulate that the observations arrive at some random time instants [192, 98]. Some of these algorithms can be expressed by differential equations with resets at random time instants, and stability analysis for such dynamical systems has been studied in [9], [8]. We have also seen some results developed for the continuous-time stochastic optimal control problems subject to randomly sampled state measurements [2].

With this motivation, in this chapter, we will address some filtering algorithms for continuous-time stochastic dynamical systems, where the continuous measurements of the observation process are not available at the remote estimation site. In particular, we take the abstraction that the transmission over a network can be modeled by a random sampling process, and we take that process to be a Poisson counter. For such a setup, we have studied continuous-discrete filters where the estimator resets its state whenever a new observation is received [110]. We provide the setup in Section 5.1 and recall some results about the evolution of expected error covariance and analyzing its asymptotic behavior. With Poisson sampling, several calculations become tractable and we are able to characterize the effect of mean sampling rate on the error covariance and convergence rates of the filter [214, 56]. We then extend this line of thought to discuss two problem settings where multiple filters interact with each other: the first one being a particular instance of ensemble filtering where the injection gains are computed from the empirical distribution of the particles, and the second one involves Bayesian inference to compute the probability mass of unknown parameters so that the resulting estimate is a weighted sum of the individual filters that correspond to the possible parameter values.

5.1 System Class and Problem Setup

Let us begin with the description of the system class and the formulation of the basic filtering problem studied in this chapter. We consider dynamical systems modeled by linear stochastic differential equations of the form

$$\mathrm{d}x_t = Ax_t\,\mathrm{d}t + G\,\mathrm{d}\omega_t \tag{5.1}$$

where $(x(t))_{t\geq 0}$ is an $\mathbb{R}^n$-valued diffusion process describing the state. Let $(\Omega, \mathcal{F}, \mathsf{P})$ denote the underlying probability space. It is assumed that, for each $t \geq 0$, $(\omega(t))_{t\geq 0}$ is a zero mean $\mathbb{R}^m$-valued standard Wiener process adapted to the filtration $\mathcal{F}_t \subset \mathcal{F}$, with the property that $\mathsf{E}[\mathrm{d}\omega(t)\,\mathrm{d}\omega(t)^\top] = I_m\mathrm{d}t$, for each $t \geq 0$. The matrices $A \in \mathbb{R}^{n\times n}$ and $G \in \mathbb{R}^{n\times m}$ are taken as constant with (A, G) controllable, and the process $(\omega(t))_{t\geq 0}$ does not depend on the state. The solutions of the stochastic differential equation (5.1) are interpreted in the sense of Itô stochastic integral.

Measurement process

Our goal is to study the state estimation problem when the output measurements are available only at random times. The motivation to work with randomly time-sampled measurements comes from several applications, such as, communication over networks which allow information packets to be sent at some discrete randomly distributed time instants. Thus, we consider a monotone nondecreasing sequence $(\tau_n)_{n\in\mathbb{N}}$ taking values in $\mathbb{R}_{\geq 0}$ which denote the time instants at which the measurements are available for estimation. We introduce the process N_t defined as

$$N_t \coloneqq \sup\{n \in \mathbb{N} \mid \tau_n \leq t\} \quad \text{for } t \in \mathbb{R}, \tag{5.2}$$

and it is assumed that $(N_t)_{t\geq 0}$ is a Poisson process of intensity $\lambda > 0$ and it is independent of the noise and the state processes. Recall [201, Theorem 2.3.2] that the Poisson process of intensity $\lambda > 0$ is a continuous-time random process $\left(N_t\right)_{t\geq 0}$ taking values in $\mathbb{N}^* := \mathbb{N}\cup\{0\}$, with $N_0 = 0$, for every $n \in \mathbb{N}^*$ and $0 =: t_0 < t_1 < \cdots < t_n < +\infty$ the increments $\{N_{t_k} - N_{t_{k-1}}\}_{k=1}^n$ are independent, and $N_{t_k} - N_{t_{k-1}}$ is distributed as a Poisson-$\lambda(t_k - t_{k-1})$ random variable for each k. The Poisson process is among the most well-studied processes, and standard results (see, e.g., [201, §2.3]) show that it is memoryless and Markovian.

The discretized, and noisy, observation process is thus defined as

$$y_{\tau_{N_t}} = Cx(\tau_{N_t}) + \nu_{\tau_{N_t}}, \qquad t \geq 0. \tag{5.3}$$

where $C \in \mathbb{R}^{p\times n}$ is a constant matrix, and ν_k is a sequence of i.i.d. Gaussian noise processes and $\nu_0 \sim \mathcal{N}(0, V)$. Equation (5.3) is motivated by the fact that a continuous observation process $\mathrm{d}z = Cx\mathrm{d}t + d\eta$ for a Wiener process η is formally equivalent to $y_t = Cx_t + \nu_t$, with the identifications $y_t \sim \frac{\mathrm{d}z_t}{\mathrm{d}t}$ and $\nu_t \sim \frac{\mathrm{d}\eta_t}{\mathrm{d}t}$, so that ν_t is a Gaussian process; see [110, Chapter 4] for further details. Our goal is to construct the estimate $\widehat{X}_t$, which minimizes the mean square estimation error, using the observations $\mathcal{Y}_t := \{y_{\tau_k} \mid k \leq N_t\}$.

Optimal filter

The basic problem in filter design is to find an estimate of the state process which minimizes the mean square estimation error, and is described by the expectation of the state process $(x_t)_{t\geq 0}$ conditioned upon the measurements observed over the interval $[0, t]$, that is, $\mathcal{Y}_t$. In particular, with the structure imposed on the system dynamics in this section, the conditional expectation is Gaussian and the two moments are simulated through ordinary differential equations with updates at times when a new measurement arrives. For an arbitrary strictly increasing real-valued sequence $(\tau_k)_{k\in\mathbb{N}^*}$, this procedure is also proposed in [110, Thm. 7.1]. If we specify a sequence $(\tau_k)_{k\in\mathbb{N}^*}$ so that it corresponds to the arrival times of a Poisson process, we simulate the mean of the conditional distribution as:

$$\dot{\widehat{X}}_t = A\widehat{X}_t \mathrm{d}t, \quad t \in [\tau_{N_t}, \tau_{1+N_t}[\tag{5.4a}$$

$$\widehat{X}_t^+ = \widehat{X}_t + K_t(y_t - C\widehat{X}_t), \quad t = \tau_{N_t} \tag{5.4b}$$

where the injection gain $K_t = P_t C^\top (C P_t C^\top + V_t)^{-1}$, and the error covariance process $(P_t)_{t\geq 0}$ is described as

$$\dot{P}_t = (AP_t + P_t A^\top + GG^\top), \quad t \in [\tau_{N_t}, \tau_{1+N_t}[\tag{5.5a}$$

$$P_t^+ = P_t - P_t C^\top (C P_t C^\top + V_t)^{-1} C P_t, \quad t = \tau_{N_t}. \tag{5.5b}$$

To write things more compactly later on, we adopt the formalism of writing the continuous-discrete equations (5.4a) and (5.4b) together in a single differential equation, when the jumps are driven by a Poisson counter N_t :

$$\mathrm{d}\widehat{X}_t = A\widehat{X}_t \mathrm{d}t + K_t(y_t - C\widehat{X}_t)\mathrm{d}N_t. \tag{5.6}$$

Similarly, using this formalism, (5.5a) and (5.5b) can be written in a combined form as,

$$\mathrm{d}P_t = (AP_t + P_t A^\top + GG^\top)\mathrm{d}t - P_t C^\top (C P_t C^\top + V_t)^{-1} C P_t \mathrm{d}N_t. \tag{5.7}$$

In the foregoing discussion, one makes the observation that the optimal conditional distribution is Gaussian for each realization of $(N_t)_{t\geq 0}$ despite the fact that the mean and covariance are discontinuous along each sample path.

Expectation of moments for the error process

From analysis viewpoint, it is important to look at the expectation of the process $(P_t)_{t\geq 0}$ with respect to the sampling times $(\tau_{N_t})_{t\geq 0}$. The basic toolset that goes into computing this expectation is to compute the infinitesimal generator for the function of the error process to obtain a differential equation for the evolution of moments in expectation, and this technique is reviewed in Section 5.2. In our previous work [214], we showed that the expectation of piecewise deterministic process P_t, denoted by $\mathcal{P}_t$, is described by the following differential equation:

$$\dot{\mathcal{P}}_t = A\mathcal{P}_t + \mathcal{P}_t A^\top + GG^\top - \lambda \mathcal{P}_t C^\top (C\mathcal{P}_t C^\top + V)^{-1} C\mathcal{P}_t. \tag{5.8}$$

We can now provide conditions in terms of the lower bounds on the mean sampling rate $\lambda > 0$ and the structural assumptions on controllability and observability of the pairs (A, G) and (A, C) that guarantee boundedness of $\mathcal{P}_t$. Such results have been reported in [214, 56].

For the analysis of the first moment of the estimation error, we consider the dynamics of the error process $\chi_t := (x_t - \widehat{X}_t)_{t\geq 0}$, which are described as,

$$\begin{aligned}
\mathrm{d}\chi_t &= A\chi_t \mathrm{d}t + G\mathrm{d}\omega_t, \\
\chi_{\tau^+_{N_t}} &= \chi_{\tau^-_{N_t}} - K_{\tau^-_{N_t}}\left(y_{\tau^-_{N_t}} - C\widehat{X}_{\tau^-_{N_t}}\right), \\
&= \left(I - K_{\tau^-_{N_t}} C\right)\chi_{\tau^-_{N_t}} - K_{\tau^-_{N_t}}\nu_{\tau^-_{N_t}}.
\end{aligned}$$

Let us denote by $\overline{\chi}_t := \mathsf{E}[\chi_t]$ the expected error with the expectation taken with respect to the sampling process N_t. Using [100, Theorem 1], or Proposition 5.3 appearing in the next section, the evolution of the expected value of the error are given by:

$$\dot{\overline{\chi}}_t = (A - \lambda\mathcal{K}_t C)\overline{\chi}_t, \tag{5.9}$$

where $\mathcal{K}_t = \mathcal{P}_t C^\top (C\mathcal{P}_t C^\top + V)^{-1}$.

5.2 Tools for Analysis

In this section, we discuss some analysis tools for stochastic differential equations, which provide the essential concepts that have been used for deriving the main results in the following section. To do so, we consider the following differential equation for a stochastic process $(x_t)_{t\geq 0}$ evolving in $\mathbb{R}^n$:

$$\mathrm{d}x_t = f(x_t)\mathrm{d}t + g(x_t)\mathrm{d}\omega + h(x_t, \nu_{N_t})\mathrm{d}N_t \tag{5.10}$$

where N_t is a Poisson process with intensity $\lambda > 0$, ω is a standard Wiener process, and ν_{N_t} is a sequence of i.i.d. random variables with probability law μ.

To study the evolution of a function of the random process $(x_t)_{t\geq 0}$, we make use of the Ito's chain rule. The reader may consult [174, Chapter II, Section 7] for detailed exposition on this topic. Here, the particular form we adopt is tailored for the differential equations appearing in earlier sections.

Proposition 5.1 (Ito's chain rule)**.** *For a twice continuously differentiable function $\psi : \mathbb{R}^n \to \mathbb{R}$, it holds that*

$$\begin{aligned}
\mathrm{d}\psi(x_t) = \Big[\langle\nabla\psi(x_t), f(x_t)\rangle + \frac{1}{2}\mathrm{tr}\Big(\frac{\partial^2\psi(x_t)}{\partial x}g(x_t)g(x_t)^\top\Big)\Big]\mathrm{d}t \\
+ \langle\nabla\psi(x_t), g(x_t)\rangle\mathrm{d}\omega + \Big[\psi(x_t + h(x, \nu_{N_t})) - \psi(x_t)\Big]\mathrm{d}N_t.
\end{aligned} \tag{5.11}$$

Ito's chain rule describes the evolution of the function ψ evaluated along the solution of the stochastic differential equation (5.10). However, to describe the evolution of the expectation of ψ in differential form, we need to consider the extended generator as defined below:

Definition 5.2 (Extended generator)**.** Given a real-valued function $\psi : \mathbb{R}^n \to \mathbb{R}$, the *extended generator* of the process $(x_t)_{t\geq 0}$ described by (5.10) is the linear operator $\psi \mapsto \mathcal{L}\psi$ defined by

$$\begin{aligned} &\mathbb{R}^n \ni z \mapsto \mathcal{L}\psi(z) \in \mathbb{R} \\ &\mathcal{L}\psi(z) := \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\Big(\mathsf{E}\big[\psi(x(t+\varepsilon)) \,|\, x(t) = z\big] - \psi(z)\Big). \end{aligned} \tag{5.12}$$

We obtain the expected value of ψ by integrating the generator, which can be seen as a generalization of the classical Dynkin's formula:

$$\mathsf{E}[\psi(x(t))] = \mathsf{E}[\psi(x(0))] + \mathsf{E}\Big[\int_0^t \mathcal{L}\psi(x(s))\,\mathrm{d}s\Big]. \tag{5.13}$$

For our purposes, it is useful to compute an explicit expression of the generator which can then be analyzed for studying the qualitative behavior of $\mathsf{E}[\psi(x)]$. Several references in the literature derive generator equations for stochastic processes with jumps, see for example [100] for a derivation in the context of Wiener process driven differential equations with renewal processes. Here, we provide a statement specifically tailored for our purposes.

Proposition 5.3. *If the sampling process $(N_t)_{t\geq 0}$ is Poisson with intensity $\lambda > 0$, then the process $(x(t))_{t\geq 0}$ described in (5.10) is Markovian. Moreover, for any function $\mathbb{R}^n \ni z \mapsto \psi(z) \in \mathbb{R}$ with at most polynomial growth as $\|z\| \to +\infty$, we have*

$$\begin{aligned} \mathcal{L}\psi(z) = \langle \nabla\psi(z), f(z)\rangle + \frac{1}{2}\mathrm{tr}\Big(\frac{\partial^2\psi(z)}{\partial z^2} g(z)g(z)^\top\Big) \\ + \lambda\Big[\int \psi(z + h(z,\nu))\mu(\mathrm{d}\nu) - \psi(z)\Big]. \end{aligned} \tag{5.14}$$

5.3 Ensemble Filters

In the literature, we find several techniques that provide good approximation of the filtering distribution [4]. Among these, the use of Monte Carlo integration methods for approximating the optimal distribution have gained significant interest in the literature [69, 179]. In the same spirit, [72] introduced the technique of ensemble Kalman filters, or particle filters, to develop filtering methods for large scale applications related to geophysical sciences, so that the error covariance is computed from a *collection* of state estimators rather than from a single (Riccati) differential equation. Since then, the use of particle filters, or ensemble Kalman filters, have had a notable impact where estimation with noisy data is required in large-scale models. This approach is based on simulating the evolution of different particles through differential equations that are coupled through the associated empirical mean and the empirical error covariance. Several review articles [42, 35] and the books [73] provide an overview of developments in that area. In most of these works, we do not find much details about the theoretical analysis of the proposed filtering techniques, and this area of mathematical analysis of the ensemble Kalman filters has gathered attention only very recently.

Also, over the past decade, we see that the ensemble filtering methods can be viewed from the lens of mean-field models described by stochastic differential equations, and the particles are simply the approximations of these mean field models. Recent review articles which elaborate on this viewpoint are [21, 205]. In fact, the limiting behavior of these particles is described by a McKean–Vlasov type diffusion process, which is also referred to as the mean-field process. In the literature, this mean-field process is seen to be chosen in different ways, e.g., by adding noise in the prediction term and the correction term of the Kalman–Bucy process [62], or as a non-diffusion equation that is optimal in the measure transportation sense [204].

Our primary objective in is to develop particle filters for continuous-time stochastic processes subject to randomly time-sampled observations. In the literature, we do find some variants of *continuous-discrete particle filters* in different settings. The paper [19] provides one (and possibly the first) such example, where the authors use mollifiers in the particle equations to smoothen the dynamics, but no statements about the limiting process are provided. The paper [230] develops particle filters for nonlinear systems using the time-discretization procedure as a part of the derivation and studies convergence as the length of the sampling interval converges to zero.

For our purposes, the state process is modeled by linear continuous-time Ornstein-Uhlenbeck process and the sampling process for the observations is a Poisson counter. We develop particle filters which update their estimate whenever the Poisson counter increments, which refers to the arrival of a new measurement from the observation process. Our objective here is to study a generalization of the so-called *vanilla filters*: It involves noise terms in the description of the mean-field process. The presence of noise is motivated by the fact that it leads to positive definite (and hence invertible) error covariance matrix that is important for good performance of the particles. However, the presence of noise terms makes the analysis more challenging and we develop appropriate tools which allow us to analyze the resulting stochastic processes. In essence, we show that for appropriate sampling rate, if the number of the particles is large enough then the empirical mean and the empirical covariance converge to the optimal mean and the covariance.

5.3.1 Mean field process and ensemble filters

In the previous section, we saw that the implementation of conventional optimal filter, even for linear systems, is computationally heavy since it requires simulating a Riccati differential equation to compute the injection gains. For a state process evolving in $\mathbb{R}^n$, this optimal filter involves solving $(n^2+n)/2$ differential equations for the covariance process $(P_t)_{t\geq 0}$ and n differential equations for the first moment of the estimate $(\widehat{X}_t)_{t\geq 0}$. This can be quite cumbersome for large values of n especially when we take into consideration the additional operations involved in computing these processes. For this reason, there has been extensive research for other methods to implement optimal filter, or its approximation. One such technique is based on the use of the *ensemble*, or particle, filters. The basic idea of the particle filters is to simulate a collection of particles through stochastic differential equation which are coupled to each other through joint statistics of the ensemble.

Description of McKean-Vlasov type process

For the definition of particles in the ensemble, we first describe a diffusion process of McKean-Vlasov type differential equation. This process is, in particular, continuous-discrete because our observation process is discrete. In particular, we consider $(S_t)_{t\geq 0}$ described as,

$$
\begin{aligned}
\mathrm{d}S_t &:= AS_t\mathrm{d}t + G\mathrm{d}\bar{\omega}_t + Q_tC^\top(CQC^\top+V_t)^{-1}\left[y_t-(CS_t+\bar{\nu}_t)\right]\mathrm{d}N_t && (5.15a)\\
Q_t &:= \mathsf{E}[(S_t-\widehat{S}_t)(S_t-\widehat{S}_t)^\top\,|\,\mathcal{Y}_t] && (5.15b)
\end{aligned}
$$

$$\widehat{S}_t := \mathsf{E}[S_t \mid \mathcal{Y}_t], \tag{5.15c}$$

where the processes $\bar{\omega}_t$ and $\bar{\nu}_t$ are independent copies of the processes ω_t and ν_t respectively, that were introduced in (5.1) and (5.3). The motivation of working with this particular diffusion process comes from the fact that the conditional mean $\widehat{S}_t$ and the conditional covariance Q_t satisfy the same differential equations as the ones obtained for their optimal counterparts, that is, it can be checked that

$$\mathrm{d}\widehat{S}_t = A\widehat{S}_t \mathrm{d}t + Q_t C^\top (CQ_tC^\top + V_t)^{-1}(y_t - C\widehat{S}_t)\mathrm{d}N_t \tag{5.16}$$

and

$$\mathrm{d}Q_t = (AQ_t + Q_tA^\top + GG^\top)\mathrm{d}t - Q_tC^\top(CQ_tC^\top + V_t)^{-1}CQ_t\mathrm{d}N_t, \tag{5.17}$$

so that, one sees the resemblance with the processes described in (5.6) and (5.7), respectively.

Particle filters

From the mean-field process (5.15), we describe an ensemble of M particles. For $\ell = 1, \dots, M$, the particle S_t^ℓ is described by the following stochastic differential equation conditioned upon the observation process and the Poisson sampling process:

$$\mathrm{d}S_t^\ell = AS_t^\ell \mathrm{d}t + G\mathrm{d}\omega_t^\ell + K_t^M(y_t - CS_t^\ell - \nu_t^\ell)\,\mathrm{d}N_t, \tag{5.18}$$

where, for each ℓ, ω_t^ℓ and ν_t^ℓ represent the independent copies of the processes ω_t and ν_t respectively. The term K_t^M denotes the empirical gain computed at time instants $t = \tau_{N_t}$ as follows:

$$K_t^M = Q_t^M C^\top (CQ_t^M C^\top + V)^{-1} \tag{5.19}$$

in which Q_t^M denotes the empirical covariance, which is computed using the empirical mean $\widehat{S}_t^M$ as follows:

$$Q_t^M = \frac{1}{M}\sum_{\ell=1}^{M}(S_t^\ell - \widehat{S}_t^M)(S_t^\ell - \widehat{S}_t^M)^\top \tag{5.20}$$

$$\widehat{S}_t^M = \frac{1}{M}\sum_{\ell=1}^{M} S_t^\ell. \tag{5.21}$$

One clearly sees that the computational effort in implementing the particle filters (5.18) involves solving Mn differential equations. For computing the gain, the empirical mean and the sample covariance only needs to be updated at discrete times using relatively simpler arithmetic operations.

Convergence result

The obvious question concerning the proposed filters in (5.18) is whether they approximate the optimal filter (5.6)–(5.7) reasonably well. In this regard, our main result shows that if the mean sampling rate satisfies certain conditions, then the difference (in expectation) between the empirical covariance of the particles (5.20) and the optimal covariance (5.21) is $O(\frac{1}{M})$. Similarly, we can show that, as the number of particles increase, the expectation of the empirical mean of the particles approaches the expected value of the optimal estimate $\widehat{X}$.

Our main result formalizing these observations is stated next. To avoid technicalities in the proof, we present the convergence result for the scalar case only, where A is a real number and G, C are nonzero scalars (so that the pair (A, G) is controllable and (A, C) is observable). This allows us to give explicit bounds on the approximation error, and the conditions on mean sampling rate required for establishing these bounds.

Theorem 5.4. *Consider system* (5.1) *with optimal estimator* (5.6) *and the ensemble filter of M particles described by* (5.18). *There exists $0 < \gamma < 1$, such that if the inequality*

$$\lambda > 2A/(1-\gamma^2), \tag{5.22}$$

then it holds that

$$\limsup_{t\to\infty} \|\mathsf{E}[P_t] - \mathsf{E}[Q_t^M]\| \leq \frac{G^2 + \lambda V/C^2}{M(\lambda(1-\gamma^2) - 2A)} \tag{5.23}$$

and moreover,

$$\|\mathsf{E}[\widehat{S}_t^M] - \mathsf{E}[\widehat{X}_t]\| \to 0, \qquad \textit{as } M \to \infty, \textit{and } t \to \infty.$$

The proof of this result is carried out in Section 5.3.2. The constant γ appearing in the theorem statement actually corresponds to the lower bound on the expectation of the covariance matrices P_t and Q_t^M, which actually depend on λ. The expression on the right-hand side of (5.22) therefore depends on λ and it turns out that this inequality yields upper and lower bounds on the mean sampling rate. Nonzero value of G (and hence the controllability of (A, G)) is important for establishing this lower bound. The lower bound on the mean sampling rate is consistent with what we observe in [211] for stabilization of linear systems with Poisson-sampled full state feedback, but includes an additional factor of $(1-\gamma^2)$. We note that, in terms of the quality of approximation of the optimal covariance by empirical covariance, the bound in (5.23) increases with increase in the covariance of process and sensor noises, and decreases as number of particles M gets large.

Simulation of an academic example

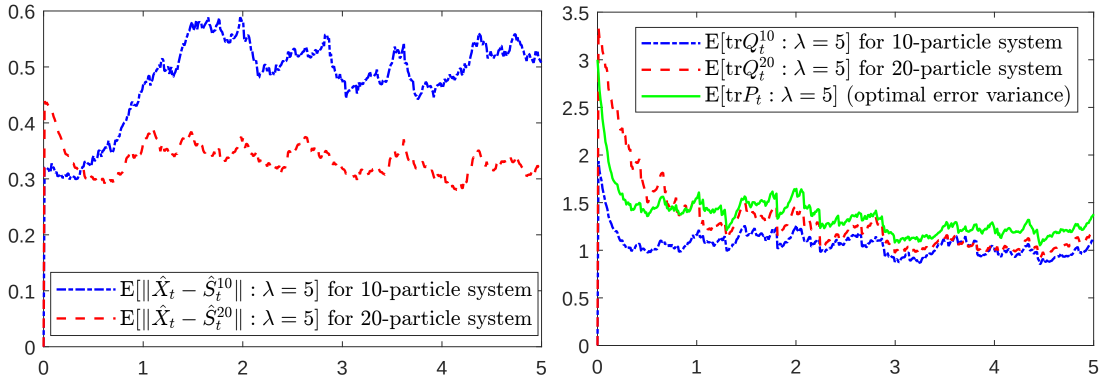
To illustrate the simulation of our proposed particle filters in (5.18) and the result reported in Theorem 5.4, we consider an academic example where the system is described by the equations:

$$\mathrm{d}x_t = Ax_t\mathrm{d}t + G\mathrm{d}\omega_t \tag{5.24a}$$

$$y_{\tau_k} = Cx_{\tau_k} + \nu_{\tau_k}, \tag{5.24b}$$

with $A = \begin{bmatrix} 0 & 3 & 1 \\ 2 & -2 & 1 \\ -2 & 1 & -3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$, $G = \begin{bmatrix} 0.5 & 0.5 & 0.5 \end{bmatrix}^\top$, and for each $k \in \mathbb{N}^*$, ν_{τ_k} is normally distributed with mean $(0,0)^\top$ and the constant variance $V = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}$.

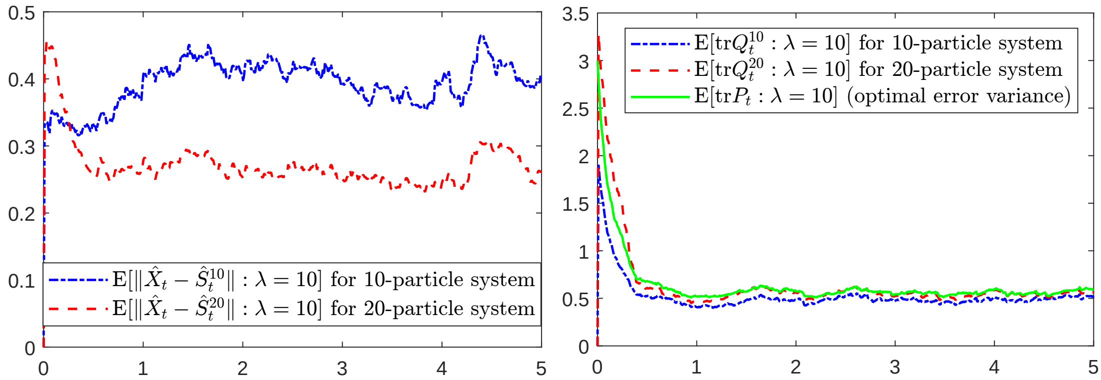
To measure the effectiveness of the particle system, we compare it with the optimal estimator, and look at the entities that are being compared in Theorem 5.4 for different values of mean sampling rate. The simulation results reported in Figure 5.1 and Figure 5.2 are consistent with the results in Theorem 5.4.



(a) The difference of optimal mean and empirical mean is plotted for 10-particles (blue curve) and 20 particles (red curve).

(b) The plot shows the optimal covariance (green); empirical covariance for 10-particles (blue) and 20 particles (red).

Figure 5.1: Comparison of mean and covariance for the optimal and ensemble filters with mean sampling rate $\lambda = 5$.



(a) The difference of optimal mean and empirical mean is plotted for 10-particles (blue curve) and 20 particles (red curve).

(b) The plot shows the optimal covariance (green); empirical covariance for 10-particles (blue) and 20 particles (red).

Figure 5.2: Comparison of mean and covariance for the optimal and ensemble filters with mean sampling rate $\lambda = 10$.

5.3.2 Convergence of ensemble filters

This section essentially carries out the proof of Theorem 5.4. We first derive the differential equations for empirical mean and covariance using Ito's chain rule given in Proposition 5.1 in the process. We then compute the expectations of the processes using the generator equation stated in Proposition 5.3, which are then used for studying convergence results stated in Theorem 5.4. We emphasize that the developments in this section are carried out for the scalar case only.

Differential equations for empirical mean and covariance

We recall that the evolution of the particles is described by the differential equation (5.18) and the corresponding empirical mean is defined in (5.21). We now write down the evolution of the empirical mean of the particles in differential form as follows:

$$\mathrm{d}\widehat{S}_t^M = A\widehat{S}_t^M + \frac{1}{\sqrt{M}}G\mathrm{d}\widetilde{\omega}_t^M + K_t^M\left(y_t - C\widehat{S}_t^M - \frac{1}{\sqrt{M}}\widetilde{\nu}_t^M\right)\mathrm{d}N_t$$

where we used the notation

$$\widetilde{\omega}_t^M := \frac{1}{\sqrt{M}}\sum_{\ell=1}^M \omega_t^\ell \qquad \text{so that } \mathsf{E}[\mathrm{d}\widetilde{\omega}_t^M \mathrm{d}\widetilde{\omega}_t^{M\top}] = \mathrm{d}t$$

and

$$\widetilde{\nu}_t^M := \frac{1}{\sqrt{M}}\sum_{\ell=1}^M \nu_t^\ell \qquad \text{so that } \widetilde{\nu}_t^M \sim \nu.$$

Note that the process $\widetilde{\nu}_t^M$ is also i.i.d. Gaussian with covariance equal to V.

Next we write the empirical covariance of the particles, defined in (5.20), in differential form. To do so, we let

$$q_t^\ell := S_t^\ell - \widehat{S}_t^M.$$

It then follows that

$$\mathrm{d}q_t^\ell = Aq_t^\ell\mathrm{d}t + G\left(\mathrm{d}\omega_t^\ell - \frac{1}{\sqrt{M}}\mathrm{d}\widetilde{\omega}_t^M\right) - K_t^M C q_t^\ell \mathrm{d}N_t - K_t^M\left(\nu_t^\ell - \frac{1}{\sqrt{M}}\widetilde{\nu}_t^M\right)\mathrm{d}N_t. \tag{5.25}$$

Consequently, we can use Proposition 5.1 to get

$$\begin{aligned}\mathrm{d}(q_t^\ell)^2 = &\left(2A(q_t^\ell)^2 + G^2 - \frac{1}{M}G^2\right)\mathrm{d}t + 2Gq_t^\ell\left(\mathrm{d}\omega_t^\ell - \frac{1}{\sqrt{M}}\mathrm{d}\widetilde{\omega}_t^M\right)\\ &+ \Big[(q_t^\ell - K_t^M C q_t^\ell - K_t^M\nu_t^\ell + \frac{1}{\sqrt{M}}K_t^M\widetilde{\nu}_t^M)^2 - (q_t^\ell)^2\Big]\mathrm{d}N_t.\end{aligned}$$

Straightforward simplification yields

$$\begin{aligned}\mathrm{d}(q_t^\ell)^2 = &\left(2A(q_t^\ell)^2 + G^2 - \frac{1}{M}G^2\right)\mathrm{d}t + 2Gq_t^\ell\left(\mathrm{d}\omega_t^\ell - \frac{1}{\sqrt{M}}\mathrm{d}\widetilde{\omega}_t^M\right)\\ &+ \Big[(K_t^M)^2C^2(q_t^\ell)^2 - 2K_t^M C(q_t^\ell)^2 + (K_t^M)^2(\nu_t^\ell)^2 + \frac{1}{M}(K_t^M)^2(\widetilde{\nu}_t^M)^2 - \frac{2}{\sqrt{M}}(K_t^M)^2\nu_t^\ell\widetilde{\nu}_t^M\\ &- 2(1 - K_t^M C)q_t^\ell K_t^M(\nu_t^\ell - \frac{1}{\sqrt{M}}K_t^M\widetilde{\nu}_t^M)\Big]\mathrm{d}N_t.\end{aligned}$$

Note that, we can write $Q_t^M = \sum_{\ell=1}^M (q_t^\ell)^2$. Furthermore, recalling the expression for empirical gain in (5.19) we can check that

$$(K_t^M)^2C^2Q_t^M - 2K_t^M CQ_t^M + (K_t^M)^2V = -K_t^M CQ_t^M.$$

The last two identities allow us to write

$$\begin{aligned}\mathrm{d}Q_t^M = &\left(2AQ_t^M + G^2 - \frac{1}{M}G^2\right)\mathrm{d}t2G\sum_{\ell=1}^M q_t^\ell\left(\mathrm{d}\omega_t^\ell - \frac{1}{\sqrt{M}}\mathrm{d}\widetilde{\omega}_t^M\right)\\ &+ \Big[-K_t^M CQ_t^M - (K_t^M)^2V + \frac{1}{M}(K_t^M)^2\sum_{\ell=1}^M(\nu_t^\ell)^2\\ &+ \frac{1}{M}(K_t^M)^2(\widetilde{\nu}_t^M)^2 - \frac{2}{M\sqrt{M}}(K_t^M)^2\sum_{\ell=1}^M \nu_t^\ell\widetilde{\nu}_t^M\\ &- \frac{2}{M}(1 - K_t^M C)K_t^M\sum_{\ell=1}^M q_t^\ell(\nu_t^\ell - \frac{1}{\sqrt{M}}K_t^M\widetilde{\nu}_t^M)\Big]\mathrm{d}N_t.\end{aligned} \tag{5.26}$$

We next use this expression to compute the expectation of $\mathsf{E}[Q_t^M]$ with respect to the noise terms and, in addition, the sampling times of the Poisson counter.

Convergence of expected covariance

Using Proposition 5.3, we now turn our attention to computing the expectation of optimal covariance P_t (with respect to sampling times), which was described by a piecewise deterministic Markov process in (5.7). Similarly, we will use Proposition 5.3 to describe the expectation of the empirical covariance (with respect to noise in the particle filters and the underlying sampling process), which is described by a jump diffusion process in (5.26). In what follows, we use the notation $\mathcal{P}_t := \mathsf{E}[P_t]$, $\mathcal{Q}_t^M = \mathsf{E}[\mathcal{Q}_t^M]$, and similarly, for the corresponding injection gains, we use $\mathcal{K}_t := \mathcal{P}_t C(\mathcal{P}_t C^2 + V)^{-1}$ and $\mathcal{K}_t^M := \mathcal{Q}_t^M C(\mathcal{Q}_t^M C^2 + V)^{-1}$.

By applying Proposition 5.3 to the process (5.7), we first get

$$\begin{aligned} \dot{\mathcal{P}}_t &= 2A\mathcal{P}_t + G^2 - \lambda \mathcal{P}_t^2 C^2 (\mathcal{P}_t C^2 + V)^{-1} \\ &= (2A - \lambda)\mathcal{P}_t + G^2 + \lambda \mathcal{P}_t (1 - \mathcal{P}_t C^2 (\mathcal{P}_t C^2 + V)^{-1}) \\ &= (2A - \lambda)\mathcal{P}_t + G^2 + \lambda \mathcal{P}_t \phi(\mathcal{P}_t) \end{aligned} \tag{5.27}$$

where the mapping $\mathcal{P}_t \mapsto \phi(\mathcal{P}_t)$ is defined as

$$\phi(\mathcal{P}_t) := (1 - \mathcal{P}_t C^2 (\mathcal{P}_t C^2 + V)^{-1}).$$

On the other hand, using the fact that ν_t^ℓ and $\widetilde{\nu}_t^M$ are mean zero processes, the application of Proposition 5.3 to the process Q_t^M in (5.26) yields,

$$\begin{aligned} \dot{\mathcal{Q}}_t^M &= 2A\mathcal{Q}_t^M + G^2 - \frac{1}{M}G^2 - \lambda \mathcal{Q}_t^M C^2 (\mathcal{Q}_t^M C^2 + V)^{-1} \mathcal{Q}_t^M + \frac{\lambda}{M}(\mathcal{K}_t^M)^2 V \\ &= (2A - \lambda)\mathcal{Q}_t^M + G^2 + \lambda \mathcal{Q}_t^M \phi(\mathcal{Q}_t^M) - \frac{1}{M}G^2 + \frac{\lambda}{M}(\mathcal{K}_t^M)^2 V. \end{aligned} \tag{5.28}$$

To analyze the difference between optimal and empirical covariance, let us introduce $\mathcal{E}_t^M := \mathcal{Q}_t^M - \mathcal{P}_t$, which satisfies the following equation:

$$\dot{\mathcal{E}}_t^M = (2A - \lambda)\mathcal{E}_t^M + \lambda \mathcal{Q}_t^M \phi(\mathcal{Q}_t^M) - \lambda \mathcal{P}_t \phi(\mathcal{P}_t) - \frac{1}{M}G^2 + \frac{\lambda}{M}(\mathcal{K}_t^M)^2 V. \tag{5.29}$$

The following lemma states some important properties that we need in the sequel, and is stated without proof.

Lemma 5.5. *For the equation* (5.27) *and equation* (5.28)*, we have the following relations:*

1. *There exists $C_0 > 0$ (depending on the initial conditions and noise covariance levels) such that $\mathcal{P}_t > C_0$ and $\mathcal{Q}_t^M > C_0$, for each $t \geq 0$.*

2. *There exists $0 < \gamma < 1$ (depending on C_0) such that for each $t \geq 0$, we have*

$$\begin{aligned} 1 - \gamma &\leq \|\mathcal{P}_t C^2 (\mathcal{P}_t C^2 + V)^{-1}\| < 1, \\ 1 - \gamma &\leq \|\mathcal{Q}_t^M C^2 (\mathcal{Q}_t^M C^2 + V)^{-1}\| < 1. \end{aligned}$$

   Consequently, it holds that $\|\phi(\mathcal{P}_t)\| \leq \gamma$ and $\|\phi(\mathcal{Q}_t^M)\| \leq \gamma$.

3. *The function ϕ has the property that for each P_1 and P_2, it satisfies*

$$\phi(P_1)P_1 - \phi(P_2)P_2 \leq \phi(P_1)(P_1 - P_2)\phi(P_2).$$

Proof. Let us provide a proof for the third item, which is independent of the first two items. To do so, we introduce the notation

$$K_1 := P_1 C(P_1 C^2 + V)^{-1}; \quad K_2 := P_2 C(P_2 C^2 + V)^{-1}.$$

It can be easily checked that

$$\phi(P_1) = 1 - K_1C; \quad \phi(P_2) = 1 - K_2C,$$

and moreover, we have the identities

$$\phi(P_1)P_1C = K_1V; \quad \phi(P_2)P_2C = K_2V.$$

The last two equations result in

$$\begin{aligned}
&\phi(P_1)P_1 - \phi(P_2)P_2 \\
&= \phi(P_1)P_1 + K_1VK_2 - K_1VK_2 - \phi(P_2)P_2 \\
&= \phi(P_1)P_1 + \phi(P_1)P_1CK_2 - K_1\phi(P_2)P_2C - \phi(P_2)P_2 \\
&= \phi(P_1)P_1(1 - K_2C) - (1 - K_1C)\phi(P_2)P_2 \\
&= \phi(P_1)P_1\phi(P_2) - \phi(P_1)\phi(P_2)P_2 \\
&= \phi(P_1)(P_1 - P_2)\phi(P_2)
\end{aligned}$$

which is the desired statement in Lemma 5.5. □

One can now use the statements in this lemma to get a bound on the evolution of $\mathcal{E}_t^M$. This is done by looking at the derivative of $(\mathcal{E}_t^M)^2$, that is,

$$\frac{d}{dt}(\mathcal{E}_t^M)^2 \le 2(2A - \lambda)(\mathcal{E}_t^M)^2 + 2\lambda\gamma^2(\mathcal{E}_t^M)^2 + \frac{2}{M}\|\mathcal{E}_t^M\|G^2 + \frac{2\lambda}{MC^2}V\|\mathcal{E}_t^M\|.$$

In the last inequality, the quadratic term in $(\mathcal{E}_t^M)^2$ is negative if

$$\lambda - 2A > \lambda\gamma^2$$

or equivalently,

$$\lambda > \frac{2A}{1 - \gamma^2}. \tag{5.30}$$

The desired result about ultimate bound on $\mathcal{E}_t^M$ now follows for every finite value of λ satisfying (5.30). In particular,

$$\limsup_{t\to\infty} \|\mathcal{E}_t^M\| \le \frac{G^2 + \lambda V/C^2}{M(\lambda(1 - \gamma^2) - 2A)}$$

where we note that the denominator is strictly positive due to the bound on λ in (5.30).

Convergence of expected mean

The last step in the proof of Theorem 5.4 is to analyze the difference between the expected values of the optimal mean and empirical mean. In this regard, we let $\widehat{\mathcal{S}}_t^M := \mathsf{E}[\widehat{S}_t^M]$, and $\widehat{\mathcal{X}}_t = \mathsf{E}[\widehat{X}_t]$. Applying Proposition 5.3, we first observe that

$$\dot{\widehat{\mathcal{X}}}_t = A\widehat{\mathcal{X}}_t + \lambda\mathcal{K}_t(\mathsf{E}[y_t] - C\widehat{\mathcal{X}}_t)$$

and

$$\dot{\widehat{\mathcal{S}}}_t^M = A\widehat{\mathcal{S}}_t^M + \lambda\mathcal{K}_t^M(\mathsf{E}[y_t] - C\widehat{\mathcal{S}}_t^M)$$

We now look at the asymptotic behavior of $\widehat{\chi}_t^M := \widehat{\mathcal{X}}_t - \widehat{\mathcal{S}}_t^M$, which satisfies

$$\begin{aligned}\dot{\widehat{\chi}}_t^M &= A\widehat{\chi}_t^M + \lambda\mathcal{K}_t(y_t - C\widehat{\mathcal{X}}_t) - \lambda\mathcal{K}_t^M(\mathsf{E}[y_t] - C\widehat{\mathcal{S}}_t^M) \\ &= (A - \lambda\mathcal{K}_t^M C)\widehat{\chi}_t^M + \lambda(\mathcal{K}_t - \mathcal{K}_t^M)(\mathsf{E}[y_t] - C\widehat{\mathcal{X}}_t).\end{aligned}$$

The bounds on asymptotic value of $\widehat{\chi}_t^M$ can now be established by analyzing the derivative of the following weighted quadratic function:

$$\widehat{\chi}_t^M \mapsto (\mathcal{Q}_t^M)^{-1}(\widehat{\chi}_t^M)^2.$$

We observe that

$$\begin{aligned}\frac{d}{dt}(\mathcal{Q}_t^M)^{-1} &= -(\mathcal{Q}_t^M)^{-1}\dot{\mathcal{Q}}_t^M(\mathcal{Q}_t^M)^{-1} \\ &= -2(A - \lambda\mathcal{K}_t^M C)(\mathcal{Q}_t^M)^{-1} - \lambda C^2(\mathcal{Q}_t^M C^2 + V)^{-1} \\ &\quad - (\mathcal{Q}_t^M)^{-2}\left(G^2 + \frac{1}{M}G^2 + \frac{\lambda}{M}\mathcal{K}_t^M V \mathcal{K}_t^M\right).\end{aligned}$$

Consequently, this last expression yields

$$\begin{aligned}\frac{d}{dt}\widehat{\chi}_t^M(\mathcal{Q}_t^M)^{-1}\widehat{\chi}_t^M = -(\widehat{\chi}_t^M)^2\Big[&\lambda C^2(\mathcal{Q}_t^M C^2 + V)^{-1} \\ &+ (\mathcal{Q}_t^M)^{-2}\left(G^2 + \frac{1}{M}G^2 + \frac{\lambda}{M}\mathcal{K}_t^M V\mathcal{K}_t^M\right)\Big] \\ &+ 2\lambda(\mathcal{Q}_t^M)^{-1}\widehat{\chi}_t^M(\mathcal{K}_t - \mathcal{K}_t^M)(\mathsf{E}[y_t] - C\widehat{\mathcal{X}}_t).\end{aligned}$$

The right-hand side can be seen as a sum of two terms: the first one is negative definite and is quadratic in $\widehat{\chi}_t^M$; the second term is linear in $\widehat{\chi}_t^M$ but contains $(\mathcal{K}_t - \mathcal{K}_t^M)$ which is converging to zero as M gets large. Using simple algebraic manipulations, and noting that $\mathcal{Q}_t^M$ is positive definite, it can be shown that $\widehat{\chi}_t^M$ converges to zero, as $t \to \infty$ and as M gets large. This completes the proof of Theorem 5.4.

Bounds on sampling rate

It can be seen that the constant γ used in the derivations actually depends on the mean sampling rate λ. Here we derive the bounds explicitly on λ. From the definition of γ, it follows that

$$\gamma \le 1 - \frac{\mathcal{P}_t C^2}{\mathcal{P}_t C^2 + V} = \frac{V}{\mathcal{P}_t C^2 + V}.$$

On the other hand, it can be shown that

$$\mathcal{P}_t \ge \min\left\{P_0, \frac{G^2}{\lambda - 2A}\right\}.$$

and, for $G_M := G\sqrt{(1 - 1/M)}$

$$\mathcal{Q}_t \ge \min\left\{Q_0^M, \frac{G_M^2}{\lambda - 2A}\right\}.$$

Letting $\widetilde{\lambda} := \lambda - 2A$, we get

$$\gamma \le \max\left\{\frac{V}{\min\{P_0, Q_0^M\}C^2 + V}, \frac{V\widetilde{\lambda}}{G_M^2 C^2 + V\widetilde{\lambda}}\right\}.$$

Consequently, the condition

$$\widetilde{\lambda} > \lambda\gamma^2$$

holds if

$$\widetilde{\lambda} > \max\left\{\frac{V^2}{(\min\{P_0, Q_0\}C^2 + V)^2}, \frac{V^2\widetilde{\lambda}^2}{(G_M^2 C^2 + V\widetilde{\lambda})^2}\right\}(\widetilde{\lambda} + 2A) \ .$$

Using simple algebraic manipulations, the last inequality holds if and only if

$$2V(\lambda - 2A)(AV - G_M^2 C^2) < G_M^4 C^4,$$

and

$$\lambda > 2A\left(1 - \frac{V^2}{(\min\{P_0, Q_0\}C^2 + V)^2}\right)^{-1} .$$

The first inequality holds if $AV \leq G_M^2 C^2$, or the measurement noise covariance V is relatively small.

Summary

In this section, we presented our work on the approximation of optimal posterior distribution using vanilla ensemble filters for continuous-time linear Gaussian systems with Poisson-sampled observations. The description of the particles in the ensemble is based on a continuous-discrete mean field process that contains flow equations with Wiener processes and jump equations with additive Gaussian noise in the modeled observation process. Using an extended generator, we compute the expectation of the empirical mean and empirical covariance of the particles with respect to the sampling process. The main result shows that for appropriate values of the mean sampling rate, the empirical mean and covariance converge to the optimal mean and variance, respectively, as the number of particles tend to infinity.

5.4 Bayesian Inference and Multi-Estimators

We next consider the problem of filtering for continuous-time linear Gaussian systems with unknown parameters, while the observation process is time-sampled as seen earlier in this chapter. Kalman filter provides a rather intuitive way of weighing the information contained in the model and the measurements depending upon the covariance of the noise driving the state process and the covariance of the measurement noise. The accurate information of the model and noise statistics is therefore very important for optimal state estimation, and if this information is not available accurately, the filter can even generate diverging state estimates [76]. With unknown parameters, the basic solution proposed in Section 5.1 cannot be directly implemented in that case. Working under the assumption that the unknown parameters belong to a finite set, we revisit the idea of using multiple filters, each of which corresponds to the possible value of the unknown parameter. The posterior distribution of the unknown parameters conditioned upon available observations is computed from Bayes' rule. The resulting state estimate is a weighted sum of the state estimates generated by multiple filters, where the weights are determined by the posterior distribution of the unknown parameters. We analyze the performance of the algorithm by looking at its asymptotic behavior and establishing boundedness of the error covariance matrix.

5.4.1 Background on adaptive filtering

For systems with unknown parameters, the state estimation for stochastic systems has been classically studied under the framework of adaptive state estimation [149]. Several approaches have been developed in this regard and several monographs, such as [148], [224], and the survey article [234] provide an overview of the developments in this area. Quite naturally, one possible way to study this problem is to compute the estimates of the unknown parameters from the noisy measurements and use them to update the filtering equations. One possible approach relies on inference using Bayes' rule to compute the posterior distribution of the unknown parameters [144], [127]. The approach has the drawback that computing this posterior, despite being in recursive form, is computationally heavy. Another approach (seen as a continuation of the Bayes' estimation) relates to maximizing different likelihood functions [148, Chapter 10]. This involves computing a tractable form for the likelihood function and setting the gradient with respect to parameters equal to zero to find the values which maximize the likelihood function. Analytical solutions can be derived in certain restrictive cases for steady state operation. Some of these techniques appear in the book [185], while providing an overview of Bayesian methods in filtering problem. A rather comprehensive collection of statistical algorithms used for parameter inference can be found in [40]. The recent work [162] provides one possible solution which relies on computing gradients backwards to speed up certain computations.

Most of the aforementioned references deal with systems in discrete-time and provide *only* the algorithms for optimal state estimation with unknown parameters, based on tools from information theory and control systems. There are relatively few works that establish theoretical guarantees of such algorithms, see for example [4, Chapter 10] and [17], which in turn builds on computing the variance of Bayes' estimates [136] and proving the consistency of maximum likelihood methods [34]. The book [224] collects some of these earlier developments in a succinct form with different assumptions on the set of unknown parameters. The problem of optimal state estimation under limited information continues to attract attention but lately, for systems with unknown dynamics, this problem has been studied using different tools. The paper [178] uses the form of optimal Kalman filter to compute the injection gain directly from the data using neural networks, and several other works also try to compute the optimal gain using different techniques [77]. However, more often than not, these techniques work under very restrictive cases that the full state measurements are available as ground truth which essentially violates the spirit of state estimation problems from system-theoretic viewpoint. Some preliminary work based on the use of stochastic gradient descent methods to compute linear gains minimizing the square norm of the residual has been presented in [206].

The problem of state estimation for stochastic systems with unknown parameters, despite being classic, continues to remain relevant as we seek algorithms which require less information with theoretical guarantees about their long-term behavior. In the remainder of this chapter, we take a step in the direction of studying the filter design for linear stochastic systems with unknown parameters, and the observations process being time sampled by a Poisson counter. We work under the assumption that all the unknown parameters belong to a discrete set. We construct an optimal filter for each value of the parameter which basically resembles Kalman filter but is driven by randomly-sampled observations, as proposed in Section 5.1. The optimal state estimate minimizing the variance of the estimation error is then described by a weighted sum of the estimates generated by individual filters, where the weights are determined by the posterior probability density function of the unknown parameters conditioned upon the observation process. For the discrete parameter set, we propose a recursive algorithm which allows for computing these conditional posterior distributions using the Bayes' rule. While such methods are computationally heavy, there are some consistency results in the literature that advocate their broad applicability. Once the algorithm is designed, we analyze the performance by looking at the asymptotic value of the expectation of the total error variance. We show that, if the underlying Bayes' inference rule is consistent, then the error covariance converges to the optimal value that arises in the case with completely known parameters. This chapter is based on our initial investigations into this problem.

5.4.2 Model with unknown parameters

In what follows, we will denote the unknowns by a vector θ, which belongs to a set Θ in some Euclidean space. These unknowns may appear in the system data or the noise statistics that are used later on for designing optimal filters. We consider the dynamical systems modeled by linear stochastic differential equations of the form:

$$\mathrm{d}X_t = A(\theta)X_t \,\mathrm{d}t + G(\theta)\,\mathrm{d}\omega_t \tag{5.31}$$

where $(X_t)_{t\geq 0}$ is an $\mathbb{R}^n$-valued diffusion process describing the state. The matrices $A(\theta) \in \mathbb{R}^{n\times n}$ and $G(\theta) \in \mathbb{R}^{n\times m}$ are constant but may depend on the unknown parameter $\theta \in \Theta$. Let $(\Omega, \mathcal{F}, \mathsf{P})$ denote the underlying probability space. It is assumed that, for each $t \geq 0$, $(\omega_t)_{t\geq 0}$ is a zero mean $\mathbb{R}^m$-valued standard Wiener process adapted to the filtration $\mathcal{F}_t \subset \mathcal{F}$, with the property that $\mathsf{E}[\mathrm{d}\omega_t\,\mathrm{d}\omega_t^\top] = I_m \mathrm{d}t$, for each $t \geq 0$. The process $(\omega_t)_{t\geq 0}$ does not depend on the state, and since we allow G to be unknown, there is no loss of generality in assuming that the statistics of random process $\mathrm{d}\omega$ are known. The solutions of the stochastic differential equation (5.31) are interpreted in the sense of Itô stochastic integral.

Our goal is to study the state estimation problem when the output measurements are available only at random times. We consider the process $(N_t)_{t\geq 0}$ as defined in (5.2) which denote the time instants at which the measurements are available for estimation. The discretized, and noisy, observation process is thus defined as

$$Y_{\tau_{N_t}} = C(\theta)X(\tau_{N_t}) + \nu_{\tau_{N_t}}, \qquad t \geq 0. \tag{5.32}$$

where $C(\theta) \in \mathbb{R}^{p\times n}$ is a constant matrix, and ν_k is a sequence of i.i.d. Gaussian noise processes and $\nu_0 \sim \mathcal{N}(0, V(\theta))$. Our goal is to construct the estimate $\widehat{X}_t$, and compute a probabilistic estimate of the unknown parameter θ, which minimizes the mean square estimation error, using the observations $\mathcal{Y}_{N_t} := \{Y_{\tau_k} \,|\, k \leq N_t\}$. For brevity, we will sometimes write Y_{N_t} in place of $Y_{\tau_{N_t}}$.

5.4.3 Filtering algorithm

The estimate of the state process which minimizes the mean square estimation error is described by the expectation of the state process $(X_t)_{t\geq 0}$ conditioned upon the measurements observed over the interval $[0, t]$, that is, $\mathcal{Y}_{N_t}$. In the current setting, our task is complicated by the fact that the system description contains some unknown parameters which need to be estimated as well.

Using standard results in filtering theory, see for example [55], it follows that the minimum mean square error (MMSE) estimate $\widehat{X}_t$ subject to the the information $\mathcal{Y}_{N_t}$, is given by,

$$\begin{aligned}\widehat{X}_t &= \mathsf{E}[X_t \,|\, \mathcal{Y}_{N_t}] = \int X_t \mathsf{p}(X_t \,|\, \mathcal{Y}_{N_t}) \mathrm{d}X_t, \\ &= \int X_t \int_\Theta \mathsf{p}(X_t \,|\, \theta, \mathcal{Y}_{N_t}) \mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t}) \mathrm{d}\theta \, \mathrm{d}X_t, \\ &= \int_\Theta \widehat{x}_t(\theta) \mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t}) \mathrm{d}\theta,\end{aligned} \tag{5.33}$$

where $\widehat{x}_t(\theta) := \mathsf{E}[X_t \,|\, \theta, \mathcal{Y}_{N_t}]$ is the minimum mean square estimate with θ given; and $\mathsf{p}(X_t \,|\, \mathcal{Y}_{N_t})$, respectively $\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})$, denotes the probability density of X_t, respectively θ, conditioned upon the observations received up to time t.

In the sequel, it will be assumed that Θ is a discrete set, and is described as $\Theta := \{\theta_1, \theta_2, \cdots, \theta_L\}$ with L being the total number of unknown parameters. We also stipulate that the pair $(A(\theta), C(\theta))$ is observable, and $(A(\theta), G(\theta))$ is controllable, for each $\theta \in \Theta$. In this case, the foregoing equation results in

$$\widehat{X}_t := \sum_{i=1}^{L} \widehat{x}_t(\theta_i) \mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t}). \tag{5.34}$$

That is, $\widehat{X}_t$ is obtained by computing $\widehat{x}_t(\theta_i)$ for each value of $\theta_i \in \Theta$ and taking their weighted sum as determined by the conditional probability mass function $\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t})$. The remainder of this section describes how to compute each of the two terms on the right-hand side of (5.34) for a fixed value of $\theta_i \in \Theta$.

State estimate for given parameters

Note that the conditional density $\mathsf{p}(X_t \,|\, \theta, \mathcal{Y}_{N_t})$ is Gaussian and $\widehat{x}_t(\theta_i)$ represents its first moment. It is described by a differential equation with updates at times when a new measurement arrives. If we specify a sequence $(\tau_k)_{k \in \mathbb{N}^*}$ so that it corresponds to the arrival times of a Poisson process, we compute $\widehat{x}_t(\theta)$ for $\theta \in \Theta$ as:

$$\dot{\widehat{x}}_t(\theta) = A(\theta)\widehat{x}_t(\theta)\mathrm{d}t, \quad t \in [\tau_{N_t}, \tau_{1+N_t}[, \tag{5.35a}$$

$$\widehat{x}_t^+(\theta) = \widehat{x}_t(\theta) + K_t(\theta)(y_t - C(\theta)\widehat{x}_t(\theta)), \quad t = \tau_{N_t}, \tag{5.35b}$$

where the injection gain $K_t(\theta)$ is defined as

$$K_t(\theta) = P_t(\theta) C(\theta)^\top M(\theta)^{-1}, \tag{5.36}$$

$$M(\theta) := C(\theta) P_t(\theta) C^\top(\theta) + V_t(\theta), \tag{5.37}$$

and the error covariance process $(P_t(\theta))_{t \geq 0}$ is described by the following equations: for $t \in [\tau_{N_t}, \tau_{1+N_t}[$, we solve

$$\dot{P}_t(\theta) = A(\theta) P_t(\theta) + P_t(\theta) A(\theta)^\top + G(\theta) G(\theta)^\top \tag{5.38a}$$

and for $t = \tau_{N_t}$, we solve

$$P_t^+(\theta) = P_t(\theta) - P_t(\theta) C(\theta)^\top M(\theta)^{-1} C(\theta) P_t(\theta). \tag{5.38b}$$

To write things more compactly later on, we adopt the formalism of writing the continuous-discrete equations (5.35a) and (5.35b) together in a single differential equation, when the jumps are driven by a Poisson counter N_t :

$$\mathrm{d}\widehat{x}_t(\theta) = A(\theta)\widehat{x}_t(\theta)\mathrm{d}t + K_t(\theta)(y_t - C(\theta)\widehat{x}_t(\theta))\mathrm{d}N_t. \tag{5.39}$$

Similarly, using this formalism, (5.38a) and (5.38b) can be written in a combined form as,

$$\begin{aligned} \mathrm{d}P_t(\theta) = (A(\theta)P_t(\theta) + P_t(\theta)A(\theta)^\top + G(\theta)G(\theta)^\top)\mathrm{d}t \\ - P_t(\theta)C(\theta)^\top M(\theta)^{-1}C(\theta)P_t(\theta)\mathrm{d}N_t. \end{aligned} \tag{5.40}$$

In the foregoing discussion, one makes the observation that the optimal conditional distribution with given θ is Gaussian for each realization of $(N_t)_{t\geq 0}$ despite the fact that the mean and covariance are discontinuous along each sample path.

As a performance metric, we now look at the expectation of the process $(P_t)_{t\geq 0}$ with respect to the sampling times $(\tau_{N_t})_{t\geq 0}$. In the previous section, when the system parameters are known, we have computed the expected covariance for Poisson sampling. That result can be applied in the current setting to get the expectation of $P_t(\theta)$, for a known value of θ. In particular, we let $\mathcal{P}_t(\theta) = \mathsf{E}[P_t(\theta)]$ and a direct application of [214, Proposition 4.1] yields the following result:

$$\dot{\mathcal{P}}_t(\theta) = A(\theta)\mathcal{P}_t(\theta) + \mathcal{P}_t(\theta)A(\theta)^\top + G(\theta)G(\theta)^\top - \lambda\mathcal{P}_t(\theta)C(\theta)^\top\mathcal{M}(\theta)^{-1}C(\theta)\mathcal{P}_t(\theta), \tag{5.41}$$

where $\mathcal{M}(\theta) := (C(\theta)\mathcal{P}_t(\theta)C^\top(\theta) + V(\theta))$. We can now provide conditions in terms of the lower bounds on the mean sampling rate $\lambda > 0$ and the structural assumptions on controllability and observability of the pairs $(A(\theta), G(\theta))$ and $(A(\theta), C(\theta))$ that guarantee boundedness of $\mathcal{P}_t(\theta)$, see [56, 214]. The boundedness is also important for asymptotic analysis of the first moment of the error process $(X_t - \widehat{X}_t)_{t\geq 0}$ as its dynamics depends on the error covariance.

Recursive expression for parameter estimation

The next element required for the computation of (5.34) is the posterior distribution $\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})$ for each $\theta \in \Theta$. A recursive formula for the computation of the conditional density $\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t})$ is obtained from the Bayes' rule as follows:

$$\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t}) = \mathsf{p}(\theta \,|\, Y_{N_t}, \mathcal{Y}_{N_t-1}) \tag{5.42a}$$

$$= \frac{\mathsf{p}(Y_{N_t} \,|\, \theta, \mathcal{Y}_{N_t-1})\mathsf{p}(\theta \,|\, \mathcal{Y}_{N_t-1})}{\sum_{i=1}^{L} \mathsf{p}(Y_{N_t} \,|\, \theta_i, \mathcal{Y}_{N_t-1})\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t-1})} \tag{5.42b}$$

We next find the expression for $\mathsf{p}(Y_{N_t} \,|\, \theta, \mathcal{Y}_{N_t-1})$ which corresponds to predicting the probability density of Y_{N_t} using the past measurement $\mathcal{Y}_{N_t-1}$ for a fixed value of θ. Let $\widehat{y}_{N_t}(\theta) := C(\theta)\widehat{x}_{\tau_{N_t}}(\theta)$ and $\widetilde{y}_{N_t}(\theta) := Y_{N_t} - \widehat{y}_{N_t}(\theta)$. When the signal is generated by the parameter θ, we denote the covariance due to innovation by $\Omega_{N_t}(\theta) := \mathsf{E}[\widetilde{y}_{N_t(\theta)}\widetilde{y}_{N_t}^\top(\theta)]$. For each $\theta_i \in \Theta$, the expression for $\Omega_{N_t}(\theta)$ can be computed offline and we have,

$$\Omega_{N_t}(\theta) = \mathsf{E}[\widetilde{y}_{N_t(\theta)}\widetilde{y}_{N_t}^\top(\theta)] = C(\theta)P_t(\theta)C^\top(\theta) + V(\theta), \quad t = \tau_{N_t}.$$

It can be readily checked that $\mathsf{p}(Y_{N_t} \,|\, \theta_i, \mathcal{Y}_{N_t-1})$ is Gaussian with mean $C(\theta_i)\widehat{x}_{\tau_{N_t}}(\theta_i)$ and covariance $\Omega_{N_t}(\theta_i)$, so that

$$\mathsf{p}(Y_{N_t} \,|\, \theta_i, \mathcal{Y}_{N_t-1}) = (2\pi)^{-p/2}|\Omega_{N_t}^{-1}(\theta_i)|^{1/2}e^{-\frac{1}{2}\widetilde{y}_{N_t}^\top(\theta_i)\Omega_{N_t}(\theta_i)\widetilde{y}_{N_t}(\theta_i)}.$$

Substituting this last expression in (5.42b), we calculate $\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t})$ recursively as

$$\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t}) = c|\Omega_{N_t}^{-1}(\theta_i)|^{1/2}e^{-\frac{1}{2}\widetilde{y}_{N_t}^\top(\theta_i)\Omega_{N_t}(\theta_i)\widetilde{y}_{N_t}(\theta_i)}\mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t-1}) \tag{5.43}$$

where c is a normalizing constant.

5.4.4 Performance analysis

For the estimate proposed in (5.34), we now analyze the performance and in particular the asymptotic behavior of the error covariance matrices and *a posterior* probability density function of the unknown parameters conditioned upon the observation process.

Derivation of error covariance

We recall that the state estimate $(\widehat{X}_t)_{t\geq 0}$ is defined by equation (5.34). The following proposition characterizes the error covariance resulting from this estimate:

Proposition 5.6. *Let us consider the error covariance matrix,*

$$P_t := \mathsf{E}[(X_t - \widehat{X}_t) \cdot (X_t - \widehat{X}_t)^\top \,|\, \mathcal{Y}_{N_t}] \tag{5.44}$$

Then, it holds that

$$P_t = \sum_{i=1}^{L} \{P_t(\theta_i) + P_t^{\mathrm{m}}(\theta_i)\} \mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t}), \tag{5.45}$$

where for each $\theta_i \in \Theta$, $P_t(\theta_i)$ is obtained by the equation (5.40), and $P_t^{\mathrm{m}}(\theta_i)$ is defined as,

$$P_t^{\mathrm{m}}(\theta_i) := (\widehat{X}_t - \widehat{x}_t(\theta_i)) \cdot (\widehat{X}_t - \widehat{x}_t(\theta_i))^\top. \tag{5.46}$$

Proof. By definition, we have

$$\begin{aligned} P_t &= \mathsf{E}[(X_t - \widehat{X}_t) \cdot (X_t - \widehat{X}_t)^\top \,|\, \mathcal{Y}_{N_t}] \\ &= \sum_{i=1}^{L} P_t^{\mathrm{a}}(\theta_i) \mathsf{p}(\theta_i \,|\, \mathcal{Y}_{N_t}), \end{aligned}$$

where $P_t^{\mathrm{a}}(\theta)$ is obtained as follows:

$$\begin{aligned} P_t^{\mathrm{a}}(\theta) &:= \mathsf{E}[(X_t - \widehat{X}_t) \cdot (X_t - \widehat{X}_t)^\top \,|\, \theta, \mathcal{Y}_{N_t}] \\ &= \mathsf{E}[X_t X_t^\top \,|\, \theta, \mathcal{Y}_{N_t}] + \widehat{X}_t \widehat{X}_t^\top - \widehat{X}_t \mathsf{E}[X_t^\top \,|\, \theta, \mathcal{Y}_{N_t}] - \mathsf{E}[X_t \,|\, \theta, \mathcal{Y}_{N_t}] \widehat{X}_t^\top \\ &= \mathsf{E}[X_t X_t^\top \,|\, \theta, \mathcal{Y}_{N_t}] + \widehat{X}_t \widehat{X}_t^\top - \widehat{X}_t \widehat{x}_t^\top(\theta) - \widehat{x}_t(\theta) \widehat{X}_t^\top \\ &= \mathsf{E}[(X_t - \widehat{x}_t(\theta)) \cdot (X_t - \widehat{x}_t(\theta))^\top \,|\, \theta, \mathcal{Y}_{N_t}] + (\widehat{X}_t - \widehat{x}_t(\theta)) \cdot (\widehat{X}_t - \widehat{x}_t(\theta))^\top \\ &= P_t(\theta) + P_t^{\mathrm{m}}(\theta), \end{aligned}$$

which proves the desired statement. □

In the statement of Proposition 5.6, P_t is introduced as a performance metric for the state estimation process with unknown parameters. Equation (5.45) provides a decomposition of this quantity in two terms, and the weights of these terms are determined by the posterior distribution of the parameters. The first of these terms $P_t(\theta_i)$ is obtained from the Kalman-like filter by setting $\theta = \theta_i$ in the system equations. The second term $P_t^{\mathrm{m}}(\theta_i)$ denotes the cost of simulating multiple Kalman-like filters with mismatch in the parameters of the true system. Note that, for each $i = 1, \dots, L$, both terms are multiplied by the posterior distribution of θ_i.

Asymptotic optimality

We can now state the primary result which looks at the asymptotic behavior of expectation of P_t in (5.45) over sampling times (governed by a Poisson counter). To state our result, we need to introduce two main assumptions. The first of these assumptions corresponds to the convergence of posterior distribution for large t.

Assumption 5.1. (Consistency of Posterior) There exists $\theta^* \in \Theta$ such that

$$\lim_{t\to\infty} \mathsf{p}(\theta^* \mid \mathcal{Y}_{N_t}) = 1, \quad \text{a.s.} \tag{5.47}$$

Although, not written explicitly, the condition (5.47) ensures that, for $\theta_i \neq \theta^*$, the posterior distribution $\mathsf{p}(\theta_i \mid \mathcal{Y}_{N_t})$ goes to zero as t gets large. Assumption 5.1 basically guarantees that the probability that our algorithm finds the true parameter eventually converges to 1 as we get more and more data. Such statements about the Bayesian update rules, or maximum likelihood technique are usually called *consistency results* in information theory. Some references do appear in the literature which develop conditions for checking consistency of the maximum likelihood method [34] and Bayes' inference rule [17], [4, Chapter 10].

Under Assumption 5.1, we can now develop a qualitative result about the asymptotic behavior of the total error covariance P_t given in (5.45).

Theorem 5.7. *Consider the system* (5.31) *with observation process* (5.32) *and N_t a Poisson counter of intensity $\lambda > 0$. The optimal state estimate in mean square sense is given by $\widehat{X}_t$ in* (5.34)*, where $\widehat{x}_t(\theta_i)$ is obtained from* (5.35) *and $\mathsf{p}(\theta_i \mid \mathcal{Y}_{N_t})$ is obtained from* (5.43)*. If Assumption 5.1 holds, then the estimation error covariance $\mathsf{E}[P_t]$ defined in* (5.44) *converges to the solution of the following equation:*

$$\begin{aligned}\dot{\mathcal{P}}_t(\theta^*) = A(\theta^*)\mathcal{P}_t(\theta^*) + \mathcal{P}_t(\theta^*)A(\theta^*)^\top + G(\theta^*)G(\theta^*)^\top \\ - \lambda\mathcal{P}_t(\theta^*)C(\theta^*)^\top M(\theta^*)^{-1}C(\theta^*)\mathcal{P}_t(\theta^*)\end{aligned} \tag{5.48}$$

where $\mathcal{P}_t(\theta^*) = \mathsf{E}[P_t(\theta^*)]$, and θ^* is such that (5.47) *holds.*

Theorem 5.7 basically states that, in the long run, as our knowledge of the unknown parameter increases, we can approximate the error covariance in the same manner as in the case of known parameter. Further analysis about the rate of convergence and impact on performance due to good/poor estimation of unknown parameter is a topic of ongoing investigation.

5.5 Perspectives

In this chapter, we basically considered two classes of filtering algorithms for linear stochastic systems with Poisson-sampled observations. The results reported in this chapter present our first approach at analyzing these algorithms from system-theoretic perspective. The primary aim being to relate the expectation of error moments in these suboptimal filters with the mean sampling rate of the sampling process, to quantify certain performance metric for these filters with the available information. There are several questions that are currently being investigated to formalize the work presented here.

Generalization of the ensemble filters

In the first part of this chapter, we considered the analysis of so-called vanilla ensemble filters for continuous-time systems. The main result describing the convergence of empirical moments to optimal moments was stated in the scalar case only. The immediate extension of our work would be to generalize the proof of Theorem 5.4 to the systems of higher dimension.

Moreover, it is also interesting to investigate whether the tools developed in this chapter can be used to study broader class of ensemble filters with stochastic noise. At this moment, we only discussed the vanilla case where the mean field process is driven by a copy of the observation noise and process noise. There are other variants of ensemble filtering which consider mean field process without the observation noise, and it needs to be seen how the analysis shapes up in that case and what are the advantages, if any, that can be obtained from different techniques. Based on our initial investigation, we see that the trade-offs appear between the asymptotic performance and computational burden required for implementing these algorithms.

Bayesian inference

In the second part of the chapter, we considered the problem of joint state estimation and detection of unknown parameters for a class of linear stochastic systems subject to randomly sampled observation process. We use the techniques based on Bayesian inference to compute the posterior distribution of the unknown parameters conditioned upon the available observations. For each value of the unknown parameter, a standard Kalman-like filter is designed, and the overall estimate is the weighted sum of individual filters. We use the covariance of the estimation error between the true state and the computed estimate as a measure of performance. Expectation of this covariance matrix over the sampling process is computed using the infinitesimal generator, and this allows us to analyze its asymptotic behavior.

Several interesting questions appear from this initial investigation. One needs to understand carefully the conditions under which the inference update rule for the unknown parameters can converge, and provide easy-to-check tests for such conditions. The other immediate direction of research is to study the case where the set of unknown parameters is continuous and not necessarily discrete. Some work on adaptive state estimation under such hypothesis involves quantization of the parameter space and showing that the inference rule converges to a value closest to the unknown parameter in some appropriate metric [93, Section 5]. At the same time, one could revisit the question of investigating other approaches for the search of unknown parameters, for example, the use of machine learning techniques to address the Bayesian inference question. It could also be interesting to see whether the ideas presented in this chapter could be merged with ongoing research on developing computationally efficient filtering algorithms, so that the proposed methodology could be applied to nonlinear systems as well.

Other suboptimal estimation techniques

In this chapter, we studied a suboptimal estimation technique inspired by ensemble filtering. Based on the principle of model predictive control (MPC) [87], we can also consider an optimal estimation problem and solve it by computing optimal estimates over finite horizon. Such techniques also provide a suboptimal estimate where the distance to optimality depends on the length of chosen horizon [154]. We have developed some preliminary results about stability performance of MPC under randomly sampled measurements in [210]. Meanwhile, there has been some progress on characterization of suboptimality using turnpike phenomenon for stochastic MPC [189]. This naturally opens up the question of investigating MPC based schemes for constructing suboptimal estimation algorithms, and it remains to be seen how these techniques compare with the ensemble based techniques presented in this manuscript.

Bibliography

[1] M. Abdelgalil, D.E. Ochoa and J.I. Poveda. “Multi-time scale control and optimization via averaging and singular perturbation theory: From ODEs to hybrid dynamical systems”. In: *Annual Reviews in Control* 56 (2023), p. 100926.

[2] M. Adès, P. E. Caines and R. P. Malhamé. “Stochastic optimal control under Poisson-distributed observations”. In: *IEEE Transactions on Automatic Control* 45.1 (2000), pp. 3–13.

[3] A.A. Ahmadi, R.M. Jungers, P.A. Parrilo and M. Roozbehani. “Joint spectral radius and path-complete graph Lyapunov functions”. In: *SIAM Journal on Control and Optimization* 52.1 (2014), pp. 687–717.

[4] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Dover Publications, 1979.

[5] D. Angeli and A. Astolfi. “A tight small gain theorem for not necessarily ISS systems”. In: *Systems & Control Letters* 56 (2007), pp. 87–91.

[6] D. Angeli, E. D. Sontag and Y. Wang. “A characterization of integral input-to-state stability”. In: *IEEE Trans. Automatic Control* 45.6 (2000), pp. 1082–1097.

[7] D. Angeli and E.D. Sontag. “Forward completeness, unboundedness observability, and their Lyapunov characterizations”. In: *Systems & Control Letters* 38.4 (1999), pp. 209–217.

[8] D. Antunes, J. Hespanha and C. Silvestre. “Stability of networked control systems with asynchronous renewal links: An impulsive systems approach”. In: *Automatica* 49.2 (2013), pp. 402–413.

[9] D. J. Antunes, J. P. Hespanha and C. J. Silvestre. “Volterra integral approach to impulsive renewal systems: Application to networked control”. In: *IEEE Transactions on Automatic Control* 57.3 (2012), pp. 607–619.

[10] M. Arcak, D. Angeli and E.D. Sontag. “A unifying integral ISS framework for stability of nonlinear cascades”. In: *SIAM J. Control Optimization* 40.6 (2002), pp. 1888–1904.

[11] M. Arcak and N.C. Martins. “Dissipativity tools for convergence to Nash equilibria in population games”. In: *IEEE Transactions on Control of Network Systems* 8.1 (2021), pp. 39–50.

[12] M. Arcak, C. Meissen and A. Packard. *Networks of Dissipative Systems: Compositional Certification of Stability, Performance and Safety*. Boston: SpringerBriefs, 2016.

[13] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Boston Basel Berlin: Birkhäuser, 1990.

[14] A. Bacciotti and F.M. Ceragioli. “Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions”. In: *ESAIM: Control, Optimization and Calculus of Variations* 4 (1999), pp. 361–376.

[15] A. Bacciotti and F.M. Ceragioli. “Nonsmooth optimal regulation and discontinuous stabilization”. In: *Abstract and Applied Analysis* 2003.20 (2003), pp. 1159–1195.

[16] R. Baier, L. Grüne and S.F. Hafstein. “Linear programming based Lyapunov function computation for differential inclusions”. In: *Discrete & Continuous Dynamical Systems - B* 17.1 (2012), pp. 33–56.

[17] Y. Baram and N.R. Sandell. “Consistent estimation on finite parameter set with application to linear systems identification”. In: *IEEE Transactions on Automatic Control* AC-23.3 (1976), pp. 451–454.

[18] T. Başar, J. Moon, J. Yong and H. Zhang. “Nonstandard linear-quadratic decision making”. In: *Proceedings of 62nd IEEE Conference on Decision and Control (CDC)*. 2023, pp. 8901–8920.

[19] K. Bergemann and S. Reich. “A mollified ensemble Kalman filter”. In: *Q.J. Royal Meteorological Society* 136.651 (2010), pp. 1636–1643.

[20] A.G. Bhatt and V.S. Borkar. “Occupation measures for controlled Markov processes: characterization and optimality”. In: *Annals of Probability* 24 (1996), pp. 1531–1562.

[21] A.N. Bishop and P. Del Moral. “On the mathematical theory of ensemble (linear-Gaussian) Kalman–Bucy filtering”. In: *Math. Control, Signals and Systems* (2023).

[22] I.M. Bomze. “Copositive optimization – Recent developments and applications”. In: *European Journal of Operational Research* 216.3 (2012), pp. 509–520.

[23] I.M. Bomze, M. Dür and C. Teo. “Copositive optimization – A survey”. In: *Optima 89, Mathematical Opimization Society Newsletter* (2012).

[24] H. Brézis. *Opérateurs Maximaux Monotones et Semi-Groupes des Contractions dans les Espaces de Hilbert.* North-Holland, Mathematics Studies, 1973.

[25] B. Brogliato. “Some perspectives on the analysis and control of complementarity systems”. In: *IEEE Transactions on Automatic Control* 48.6 (2003), pp. 918–935.

[26] B. Brogliato, R. Lozano, B. Maschke and O. Egeland. *Dissipative Systems Analysis and Control.* Third. Communications and Control Engineering. London: Springer Nature Switzerland AG, 2020.

[27] B. Brogliato and A. Tanwani. “Dynamical systems coupled with monotone set-valued operators: Formalisms, applications, well-posedness, and stability”. In: *SIAM Review* 62.1 (2020), pp. 3–129.

[28] M. Brokate and P. Krejci. “Optimal control of ODE systems involving a rate independent variational inequality”. In: *Discrete and Continuous Dynamical Systems - Series B* 18.2 (2013), pp. 331–348.

[29] F. Bullo. *Lectures on network systems.* Kindle Direct Publishing Seattle, DC, USA, 2020.

[30] S. Bundfuss and M. Dür. “Algorithmic copositivity detection by simplicial partition”. In: *Linear Algebra and its Applications* 428 (2008), pp. 1511–1523.

[31] S. Bundfuss and M. Dür. “An adaptive linear approximation algorithm for copositive programs”. In: *SIAM J. Optimization* 20.1 (2009), pp. 30–53.

[32] C. Cai and A. R. Teel. “Characterizations of input-to-state stability for hybrid systems”. In: *Systems & Control Letters* 58.1 (2009), pp. 47–53.

[33] C. Cai and A.R. Teel. “Robust input-to-state stability for hybrid systems”. In: *SIAM J. Control Optimization* 51.2 (2013), pp. 1651–1678.

[34] P. Caines and J. Rissanen. “Maximum likelihood estimation of parameters in multivariate Gaussian stochastic processes”. In: *IEEE Transactions on Information Theory* 20.1 (1974), pp. 102–104.

[35] E. Calvello, S. Reich and A.M. Stuart. “Ensemble Kalman methods: A mean field perspective”. Available online; arXiv:2209.11371v1. 2022.

[36] M.-K. Camlibel, L. Iannelli and A. Tanwani. “Convergence of proximal solutions for evolution inclusions with time-dependent maximal monotone operators”. In: *Mathematical Programming* 194 (2022), pp. 1017–1059.

[37] M.K. Camlibel, J.-S. Pang and J. Shen. “Lyapunov stability of complementarity and extended systems”. In: *SIAM J. Optimization* 17.4 (2006), pp. 1056–1101.

[38] M.K. Camlibel and J.M. Schumacher. “Linear passive systems and maximal monotone mappings”. In: *Mathematical Programming, Ser. B* 157.2 (2016), pp. 397–420.

[39] M. Cao, A.S. Morse and B.D.O. Anderson. “Reaching a consensus in a dynamically changing environment: A graphical approach”. In: *SIAM Journal on Control and Optimization* 47.2 (2008), pp. 575–600.

[40] O. Cappé, E. Moulines and T. Rydén. *Inference in Hidden Markov Models*. Springer New York, NY, 2005.

[41] F.M. Ceragioli. “Discontinuous ordinary differential equations and stabilization”. PhD thesis. Univ. Firenze, Italy, 2000.

[42] Z. Chen. “Bayesian filtering: From Kalman filters to particle filters, and beyond”. In: *Statistics* 182.1 (2003), pp. 1–69.

[43] S.P. Chhatoi, A. Tanwani and D. Henrion. “Optimal control of nonsmooth dynamical systems using measure relaxations”. In: *Proceedings of 63rd IEEE Conf. on Decision & Control*. 2024.

[44] S.P. Chhatoi, A. Tanwani and D. Henrion. “Evolution of measures in nonsmooth dynamical systems: Formalisms and computation”. In: *Automatica* 179 (2025), p. 112402.

[45] Y. Chitour, N. Guglielmi, V.Yu. Protasov and M. Sigalotti. “Switching systems with dwell time: Computing the maximal Lyapunov exponent”. In: *Nonlinear Analysis: Hybrid Systems* 40 (2021), p. 101021.

[46] Y. Chitour, I. Haidar, P. Mason and M. Sigalotti. “Upper and lower bounds for the maximal Lyapunov exponent of singularly perturbed linear switching systems”. In: *Automatica* 155 (2023), p. 11151.

[47] Y. Chitour, P. Mason and M. Sigalotti. “A characterization of switched linear control systems with finite L_2-gain”. In: *IEEE Transactions on Automatic Control* 62.4 (2017), pp. 1825–1837.

[48] F.H. Clarke. *Optimization and nonsmooth analysis*. Classics in Applied Mathematics. SIAM, 1990.

[49] F.H. Clarke, Y.S. Ledyaev and R.J. Stern. “Asymptotic stability and smooth Lyapunov functions”. In: *Journal of Differential Equations* 149 (1998), pp. 69–114.

[50] F.H. Clarke, R.J. Stern and P.R. Wolenski. “Subgradient criteria for monotonicity, the Lipschitz condition, and convexity”. In: *Canadian Journal of Mathematics* 45 (1993), pp. 1167–1183.

[51] G. Colombo, R. Henrion, D.-H. Hoang and S.-B. Mordukhovich. “Optimal control of the sweeping process over polyhedral controlled sets”. In: *Journal of Differential Equations* 260.4 (2016), pp. 3397–3447.

[52] M. Coraggio, P. DeLellis and M. di Bernardo. “Distributed discontinuous coupling for convergence in heterogeneous networks”. In: *IEEE Control Systems Letters* 5.3 (2020), pp. 1037–1042.

[53] J. Cortes. “Discontinuous dynamical systems”. In: *IEEE Control Systems Magazine* 28.3 (2008), pp. 36–73.

[54] R.-W. Cottle, J.-S. Pang and R.-E. Stone. *The Linear Complementarity Problem.* Computer Science and Scientific Computing. Academic Press, 1992.

[55] D. Crisan and B. Rozovskii. *The Oxford Handbook of Nonlinear Filtering.* Oxford University Press, 2011.

[56] A. Dasgupta and A. Tanwani. “Steady state analysis for Kalman filters with Poisson-sampled observations”. In: *Systems & Control Letters* (2025). To appear. Available online: `https://laas.hal.science/hal-04798886`.

[57] S. Dashkovskiy and A. Mironchenko. “Input-to-state stability of nonlinear impulsive systems”. In: *SIAM J. Control Optim.* 51.3 (2013), pp. 1962–1987.

[58] S. Dashkovskiy, B.S. Rüffer and F.R. Wirth. “Small gain theorems for large scale systems and construction of ISS Lyapunov functions”. In: *SIAM J. Control Optimization* 48.6 (2010), pp. 4089–4118.

[59] S. Dashkovskiy and V. Slynko. “Dwell-time stability conditions for infinite dimensional impulsive systems”. In: *Automatica* 147 (2023), p. 110695.

[60] W.P. Dayawansa and C.F. Martin. “A converse Lyapunov theorem for a class of dynamical systems which undergo switching”. In: *IEEE Transactions on Automatic Control* 44.4 (1999), pp. 751–760.

[61] D.A. Deenen, B. Sharif, S.J.A.M. van den Eijnden, H. Nijmeijer, W.P.M.H. Heemels and M.F. Heertjes. “Projection-based integrators for improved motion control: Formalization, well-posedness and stability of HIGS”. In: *Automatica* 133 (2021), p. 109830.

[62] P. Del Moral and J. Tugaut. “On the stability and the uniform propagation of chaos properties of ensemble Kalman–Bucy filters”. In: *The Annals of Applied Probability* 28.2 (2018), pp. 790–850.

[63] M. Della Rossa and A. Tanwani. “Instability of dwell-time constrained switched nonlinear systems”. In: *Systems & Control Letters* 162 (2022), p. 105164.

[64] M. Della Rossa and A. Tanwani. “Converse Lyapunov results for stability of switched systems with average dwell-time”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 31.15 (2025).

[65] M. Della Rossa, A. Tanwani and L. Zaccarian. “Smooth Approximation of Patchy Lyapunov Functions for Switching Systems”. In: *Proc. 11th IFAC Symposium on Nonlinear Control Systems (NolCoS).* 2019.

[66] M. Della Rossa, A. Tanwani and L. Zaccarian. “Max–min Lyapunov functions for switched systems and related differential inclusions”. In: *Automatica* 120 (2020), p. 109123.

[67] M. Della Rossa, A. Tanwani and L. Zaccarian. “Nonpathological ISS-Lyapunov functions for interconnected differential inclusions”. In: *IEEE Transactions on Automatic Control* 67.8 (2022), pp. 3774–3789.

[68] C.A. Desoer and M. Vidyasagar. *Feedback Systems: Input-Output Properties*. New York: Academic Press, 1975.

[69] A. Doucet, N. De Freitas and N. Gordon, eds. *Sequential Monte Carlo methods in practice*. 1st ed. Springer, 2001.

[70] M. Dür. “Copositive programming - A survey”. In: *Recent Advances in Optimization and its Application in Engineering, Springer Verlag*. Ed. by M. Diehl, F. Glineur, E. Jarlebring and W. Michiels. 2010, pp. 3–20.

[71] W. Eckhaus. *Asymptotic Analysis of Singular Perturbations*. Vol. 9. North Holland Publishing Company, 1979.

[72] G. Evensen. “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics”. In: *Journal of Geophysical Research: Oceans* 99 (1994), pp. 10143–10162.

[73] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. 2nd. Springer Science & Business Media, 2009.

[74] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Series in Operations Research and Financial Engineering. Springer, 2003.

[75] A.F. Filippov. *Differential Equations with Discontinuous Right-Hand Side*. Kluwer Academic Publisher, 1988.

[76] R. Fitzgerald. “Divergence of the Kalman filter”. In: *IEEE Transactions on Automatic Control* 16.6 (1971), pp. 736–747.

[77] G.H.W. Gebhardt, A. Kupcsik and G. Neumann. “The kernel Kalman rule”. In: *Maching Learning* 108 (2019), pp. 2113–2157.

[78] A. Girard and P. Mason. “Lyapunov functions for shuffle asymptotic stability of discrete-time switched systems”. In: *IEEE Control Systems Letters* 3.3 (2019), pp. 499–504.

[79] R. Goebel. “Stability and robustness for saddle-point dynamics through monotone mappings”. In: *Systems and Control Letters* 108 (2017), pp. 16–22.

[80] R. Goebel, R. G. Sanfelice and A. R. Teel. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. Princeton University Press, 2012.

[81] D. Goeleven and B. Brogliato. “Stability and instability matrices for linear evolution variational inequalities”. In: *IEEE Transactions on Automatic Control* 49.4 (2004), pp. 521–534.

[82] D. Goeleven, D. Motreanu and V. Motreanu. “On the stability of stationary solutions of first-order evolution variational inequalities”. In: *Advances in Nonlinear Variational Inequalities* 6 (2003), pp. 1–30.

[83] F. Golse. “On the dynamics of large Particle systems in the mean field limit”. In: *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*. Ed. by A. Muntean, J. Rademacher and A. Zagaris. Cham: Springer International Publishing, 2016, pp. 1–144.

[84] L. Grüne. *Asymptotic Behavior of Dynamical and Control Systems under Perturbation and Discretization*. Springer Berlin Heidelberg, 2002.

[85] L. Grüne. “Computing Lyapunov functions using deep neural networks”. In: *Journal of Computational Dynamics* 8.2 (2021), pp. 131–152.

[86] L. Grüne. “Dissipativity and optimal control”. arXiv:2101.12606. 2021.

[87] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control: Theory and Algorithms*. 2nd. London: Springer-Verlag, 2017.

[88] S. Hafstein and A. Tanwani. “Linear programming based lower bounds on average dwell-time via multiple Lyapunov functions”. In: *European Journal of Control* 74 (2023), p. 100838.

[89] H. Haimovich and J.L. Mancilla-Aguilar. “A characterization of integral ISS for switched and time-varying systems”. In: *IEEE Transactions on Automatic Control* 63.2 (2018), pp. 578–585.

[90] H. Haimovich and J.L. Mancilla-Aguilar. “Strong ISS implies strong iISS for time-varying impulsive systems”. In: *Automatica* 122 (2020), p. 109224.

[91] J. K. Hale. *Ordinary Differential Equations*. 2nd. Malabar, Florida: Krieger Publishing Company, 1980.

[92] A. Hauswirth, S. Bolognani and F. Dörfler. “Projected dynamical systems on irregular, non-Euclidean domains for nonlinear optimization”. In: *SIAM Journal on Control and Optimization* 59.1 (2021), pp. 635–668.

[93] R.M. Hawkes and J.B. Moore. “Performance bounds for adaptive estimation”. In: *Proceedings of the IEEE* 64.8 (1976), pp. 1143–1150.

[94] W.P.M.H. Heemels and A. Tanwani. “Existence and completeness of solutions to extended projected dynamical systems and sector-bounded projection-based controllers”. In: *IEEE Control Systems Letters* 7 (2023), pp. 1590–1595.

[95] W.P.M.H. Heemels and S. Weiland. “Input-to-state stability and interconnections of discontinuous dynamical systems”. In: *Automatica* 44.12 (2008), pp. 3079–3086.

[96] W.P.M.H. Heemels, S. Weiland and A.Lj. Juloski. “Input-to-state stability of discontinuous dynamical systems with an observer-based control application”. In: *Proceedings of ACM Workshop on Hybrid Systems: Computation and Control*. 2007, pp. 259–272.

[97] C. Hermosilla and M. Palladino. “Optimal control of the sweeping process with a nonsmooth moving set”. In: *SIAM Journal on Control and Optimization* 60.5 (2022), pp. 2811–2834.

[98] J. P. Hespanha. “Modeling and analysis of networked control systems using stochastic hybrid systems”. In: *Annual Reviews in Control* 38.2 (2014), pp. 155–170.

[99] J. P. Hespanha and A. S. Morse. “Stability of switched systems with average dwell-time”. In: *Proceedings of 38th IEEE Conf. on Decision and Control*. 1999, pp. 2655–2660.

[100] J. P. Hespanha and A. R. Teel. “Stochastic impulsive systems driven by renewal processes”. In: *Proc. 17th International Symposium on Mathematical Theory of Networked Systems*. 2006, pp. 606–618.

[101] J.P. Hespanha. “L_2-induced gains of switched linear systems”. In: *Unsolved Problems in Mathematical Systems and Control Theory*. Ed. by Vincent D. Blondel and Alexandre Megretski. Princeton: Princeton University Press, 2004, pp. 131–133.

[102] J.P. Hespanha, D. Liberzon and A.R. Teel. “Lyapunov conditions for input-to-state stability of impulsive systems”. In: *Automatica* 44 (2008), pp. 2735–2744.

[103] D.J. Hill. “A generalization of the small-gain theorem for nonlinear feedback systems”. In: *Automatica* 27.6 (1991), pp. 1043–1045.

[104] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis.* Grundlehren Text Editions. Springer Verlag, 2001.

[105] H. Ito. “State-dependent scaling problems and stability of interconnected iISS and ISS systems”. In: *IEEE Trans. Automatic Control* 51 (2006), pp. 1626–1643.

[106] H. Ito and Z.-P. Jiang. “Necessary and sufficient small gain conditions for integral input-to-state stable systems: A Lyapunov perspective”. In: *IEEE Trans. Automatic Control* 54 (2009), pp. 2389–2404.

[107] J.C. Willems. “Dissipative dynamical systems; Part I: General theory; Part II: Linear systems with quadratic supply rates”. In: *Archive for Rational Mechanics and Analysis* 45.5 (1972), pp. 321–393.

[108] A. Jadbabaie, J. Lin and A.S. Morse. “Coordination of groups of mobile autonomous agents using nearest neighbor rules”. In: *IEEE Transactions on Automatic Control* 48.6 (2003), pp. 988–1001.

[109] B. Jayawardhana, H. Logemann and E. Ryan. “Input-to-state stability of differential inclusions with applications to hysteretic and quantized feedback systems”. In: *SIAM Journal on Control and Optimization* 48.2 (2009), pp. 1031–1054.

[110] A.H. Jazwinski. *Stochastic processes and filtering theory.* New York: Dover Publications, Inc., 2007.

[111] Z. -P. Jiang, A. R. Teel and L. Praly. “Small-gain theorem for ISS systems and applications”. In: *Math. Control, Signals, and Systems* 7 (2 1994), pp. 95–120.

[112] Z.-P. Jiang, I.M.Y. Mareels and Y. Wang. “A Lyapunov formulation of the nonlinear small-gain theorem for interconnected ISS systems”. In: *Automatica* 32.8 (1996), pp. 1211–1215.

[113] Z.-P. Jiang and Y. Wang. “Input-to-state stability for discrete-time nonlinear systems”. In: *Automatica* 37.6 (2001), pp. 857–869.

[114] M. Johansson and A. Rantzer. “Computation of piecewise quadratic Lyapunov functions for hybrid systems”. In: *IEEE Transactions on Automatic Control* 43.4 (1998), pp. 555–559.

[115] R.E. Kalman and R.S. Bucy. “New results in linear filtering and prediction theory”. In: *Journal of Basic Engineering* 83.1 (1961), pp. 95–108.

[116] R. Kamalapurkar, W.E. Dixon and A.R. Teel. “On reduction of differential inclusions and Lyapunov stability”. In: *ESAIM: Control, Optimization and Calculus of Variations* 26.24 (2020).

[117] R. Kamalapurkar, J.A. Rosenfeld, A. Parikh, A.R. Teel and W.E. Dixon. “Invariance-like results for nonautonomous switched systems”. In: *IEEE Transactions on Automatic Control* 64.2 (2019), pp. 614–627.

[118] C. Kawan, A. Mironchenko, A. Swikir, N. Noroozi and M. Zamani. “A Lyapunov-based small-gain theorem for infinite networks”. In: *IEEE Transactions on Automatic Control* 66.12 (2021), pp. 5830–5844.

[119] C. Kawan, A. Mironchenko and M. Zamani. “A Lyapunov-based ISS small-gain theorem for infinite networks of nonlinear systems”. In: *IEEE Transactions on Automatic Control* 68.3 (2023), pp. 1447–1462.

[120] C.M. Kellett. “A compendium of comparison function results”. In: *Math. Control, Signals, and Systems* 26 (2014), pp. 339–374.

[121] H.K. Khalil. *Nonlinear Systems.* 3rd. Prentice-Hall, Inc., 2002.

[122] J. Kim, J. Yang, H. Shim, J.-S. Kim and J.H. Seo. “Robustness of synchronization of heterogeneous agents by strong coupling and a large number of agents”. In: *IEEE Transactions on Automatic Control* 61.10 (2016), pp. 3096–3102.

[123] P.V. Kokotovic. “Applications of singular perturbation techniques to control problems”. In: *SIAM Review* 26.4 (1984), pp. 501–550.

[124] P.V. Kokotovic, H.K. Khalil and J. O’Reilly. *Singular Perturbations Methods in Control: Analysis and Design.* New York: Academic Press, 1986.

[125] A. Kundu, D. Chatterjee and D. Liberzon. “Generalized switching signals for input-to-state stability of switched systems”. In: *Automatica* 64 (2016), pp. 270–277.

[126] D.S. Laila and D. Nesic. “Discrete-time Lyapunov-based small-gain theorem for parameterized interconnected ISS systems”. In: *IEEE Transactions on Automatic Control* 48.10 (2003), pp. 1783–1788.

[127] D.G. Lainiotis. “Partitioning: A unifying framework for adaptive systems, I: Estimation”. In: *Proceedings of the IEEE* 64.8 (1976), pp. 1126–1143.

[128] J.B. Lasserre, D. Henrion, C. Prieur and E. Trélat. “Nonlinear optimal control via occupation measures and LMI-relaxations”. In: *SIAM Journal on Control and Optimization* 47.4 (2008), pp. 1643–1666.

[129] M. Laurent. “Sums of squares, moment matrices and optimization over polynomials”. In: *Emerging Applications of Algebraic Geometry.* Vol. 149. The IMA Volumes in Mathematics and its Applications Series. Springer Verlag, 2009, pp. 157–270.

[130] D.M.W. Leenaerts and W.M.G. van Bokhoven. *Piecewise Linear Modeling and Analysis.* Springer Verlag, 1998.

[131] N.E. Leonard, A. Bizyaeva and A. Franci. “Fast and flexible multiagent decision-making”. In: *Annual Review of Control, Robotics, and Autonomous Systems* 7 (2024), pp. 12.1–12.27.

[132] D. Liberzon. *Switching in Systems and Control.* Boston: Birkhäuser, 2003.

[133] D. Liberzon. “Quantization, time delays, and nonlinear stabilization”. In: *IEEE Trans. Automatic Control* 51.7 (2006), pp. 1190–1195.

[134] D. Liberzon, D. Nesic and A.R. Teel. “Lyapunov-based small-gain theorems for hybrid systems”. In: *IEEE Trans. Automatic Control* 59.6 (2014), pp. 1395–1410.

[135] Y. Lin, E.D. Sontag and Y. Wang. “A smooth converse Lyapunov theorem for robust stability”. In: *SIAM Journal on Control and Optimization* 34.1 (1996), pp. 124–160.

[136] L. Liporace. “Variance of Bayes estimates”. In: *IEEE Transactions on Information Theory* 17.6 (1971), pp. 665–669.

[137] S. Liu, A. Russo, D. Liberzon and A. Cavallo. “Integral-input-to-state stability of switched nonlinear systems under slow switching”. In: *IEEE Transactions on Automatic Control* 67.11 (2021), pp. 5841–5855.

[138] S. Liu and A. Tanwani. “Impulsive switching signals with functional inequalities: Stability analysis using hybrid systems framework”. In: *Automatica* 171 (2025), p. 111928.

[139] S. Liu, A. Tanwani and D. Liberzon. “Average dwell-time bounds for ISS and integral ISS of switched systems using Lyapunov functions”. In: *Proc. 59th IEEE Conf. on Decision and Control.* 2020, pp. 6291–6296.

[140] S. Liu, A. Tanwani and D. Liberzon. "ISS and integral-ISS of switched systems with nonlinear supply functions". In: *Mathematics of Controls, Signals, and Systems* 34 (2022), pp. 297–327.

[141] T. Liu, Z.-P. Jiang and D.J. Hill. "Lyapunov formulation of the ISS cyclic-small-gain theorem for hybrid dynamical networks". In: *Nonlinear Analysis: Hybrid Systems* 6.4 (2012), pp. 988–1001.

[142] P. Lorenzetti and G. Weiss. "PI control of stable nonlinear plants using projected dynamical systems". In: *Automatica* 146.110606 (2022).

[143] A.I. Maass, W. Wang, D. Nesic, R. Postoyan and M. Heemels. "Event-triggered control through the eyes of a hybrid small-gain theorem". In: *IEEE Trans. Automatic Control* 68.10 (2023), pp. 5906–5921.

[144] D. Magill. "Optimal adaptive estimation of sampled stochastic processes". In: *IEEE Transactions on Automatic Control* 10.4 (1965), pp. 434–439.

[145] A. Mahajan and D. Teneketzis. "Optimal performance of networked control systems with nonclassical information structures". In: *SIAM Journal on Control and Optimization* 48.3 (2009), pp. 1377–1404.

[146] I. Malloci, J. Daafouz and C. Iung. "Stabilization of continuous-time singularly perturbed switched systems". In: *Proceedings of the 48th IEEE Conference on Decision and Control.* 2009, pp. 6371–6376.

[147] J.L. Mancilla-Aguilar, H. Haimovich and P. Feketa. "Uniform stability of nonlinear time-varying impulsive systems with eventually uniformly bounded impulse frequency". In: *Nonlinear Analysis: Hybrid Systems* 38 (2020), p. 100933. ISSN: 1751-570X.

[148] P.S. Maybeck. *Stochastic Models, Estimation, and Control.* Vol. 1. Academic Press, 1982.

[149] R. Mehra. "Approaches to adaptive filtering". In: *IEEE Transactions on Automatic Control* 17.5 (1972), pp. 693–698.

[150] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks.* Princeton University Press, 2010.

[151] A. Mironchenko, G. Yang and D. Liberzon. "Lyapunov small-gain theorems for networks of not necessarily ISS hybrid systems". In: *Automatica* 88 (2018), pp. 10–20.

[152] B.-S. Mordukhovich. *Variational Analysis and Applications.* Springer Monographs in Mathematics. Cham, CH: Springer International Publishing AG, 2018.

[153] A.S. Morse. "Supervisory control of families of linear set-point controllers Part I: Exact matching". In: *IEEE Transactions on Automatic Control* 41.10 (1996), pp. 1413–1431.

[154] M.A. Müller. "Nonlinear moving horizon estimation in the presence of bounded disturbances". In: *Automatica* 79 (2017), pp. 306–314.

[155] M.A. Müller and D. Liberzon. "Input/output-to-state stability and state-norm estimators for switched nonlinear systems". In: *Automatica* 48.9 (2012), pp. 2029–2039.

[156] A. Nagurney and D. Zhang. *Projected Dynamical Systems and Variational Inequalities with Applications.* Vol. 2. Springer Science & Business Media, 1996.

[157] D. Nesic and A.R. Teel. "Input–output stability properties of networked control systems". In: *IEEE Trans. Automatic Control* 49.10 (2004), pp. 1560–1667.

[158] N. Noroozi, A. Khayatian and R. Geiselhart. “A characterization of integral input-to-state stability for hybrid systems”. In: *Math. Control, Signals, and Systems* 29.3 (2017).

[159] A. Nurkanović and M. Diehl. “NOSNOC: A software package for numerical optimal control of nonsmooth systems”. In: *IEEE Control Systems Letters* 6 (2022), pp. 3110–3115.

[160] R.E. O’Malley. *Historical Developments in Singular Perturbations*. Springer Cham, 2014.

[161] E. Panteley and A. Loria. “Synchronization and dynamic consensus of heterogeneous networked systems”. In: *IEEE Transactions on Automatic Control* 62.8 (2017), pp. 3758–3773.

[162] C. Parellier, A. Barrau and S. Bonnabel. “Speeding-up backpropagation of gradients through the Kalman filter via closed-form expressions”. In: *IEEE Transactions on Automatic Control* 68.12 (2023), pp. 8171–8177.

[163] P.A. Parrilo. “Semidefinite programming based tests for matrix copositivity”. In: *Proceedings of the 39th IEEE Conf. on Decision and Control*. Vol. 96. 2000, pp. 4624–4629.

[164] P.A. Parrilo. “Structured semidefinite programs and semi-algebraic geometry methods in robustness and optimization”. PhD thesis. California Institute of Technology, Pasadena, CA, May 2000.

[165] A. Pavlov and L. Marconi. “Incremental passivity and output regulation”. In: *Systems and Control Letters* 57.5 (2008), pp. 400–409.

[166] L. Perko. *Differential equations and dynamical systems*. third. Newyork: Springer, 2000.

[167] J.I. Poveda and A.R. Teel. “A framework for a class of hybrid extremum seeking controllers with dynamic inclusions”. In: *Automatica* 76 (2017), pp. 113–126.

[168] V. Powers and B. Reznick. “A new bound for Polya’s Theorem with applications to polynomials positive on polyhedra”. In: *Journal of Pure and Applied Algebra* 164.1-2 (2001), pp. 221–229.

[169] V. Powers and T. Wörmann. “An algorithm for sums of squares of real polynomials”. In: *J. of Pure and Applied Algebra* 127.1 (1998), pp. 99–104.

[170] L. Praly and D. Bresch-Pietri. *Fonctions de Lyapunov, Stabilité, Stabilisation et Atténuation de perturbations*. Spartacus IDH, 2022.

[171] L. Praly and Y. Wang. “Stabilization in spite of matched unmodeled dynamics and an equivalent definition of input-to-state stability”. In: *Math. Control, Signals, and Systems* 9 (1996), pp. 1–33.

[172] C. Prieur. “Uniting local and global controllers with robustness to vanishing noise”. In: *Math. Control, Signals, and Systems* 14.2 (2001), pp. 143–172.

[173] C. Prieur, I. Queinnec, S. Tarbouriech and L. Zaccarian. “Analysis and synthesis of reset control systems”. In: *Foundations and Trends in Systems and Control* 6.2-3 (2018), pp. 117–338.

[174] P.E. Protter. *Stochastic Integration and Differential Equations*. 2nd. Springer, 2005.

[175] R. Ortega, J. Perez, A. Loria, P.J. Nicklasson and H.J. Sira-Ramirez. *Passivity-based Control of Euler-Lagrange Systems*. London: Springer, 2013.

[176] R.T. Rockafellar. “Monotone operators associated with saddle functions and minimax problems”. In: *Nonlinear Functional Analysis, Vol. I, Proceedings of Symposia in Pure Math.* Ed. by Felix E. Browder. American Mathematical Society, 1970.

[177] J.B. Rejeb, I.-C. Morarescu, A. Girard and J. Daafouz. “Stability analysis of a general class of singularly perturbed linear hybrid systems”. In: *Automatica* 90 (2018), pp. 98–108.

[178] G. Revach, N. Shlezinger, X. Ni, A.L. Escoriza, R.J.G. van Sloun and Y.C. Eldar. “KalmanNet: Neural network aided Kalman filtering for partially known dynamics”. In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 1532–1547.

[179] B. Ristic, S. Arulampalam and N. Gordon. *Beyond the Kalman filter: Particle filters for tracking applications.* Artech house, 2003.

[180] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis.* 3rd printing, 2009. Vol. 317. Gundlehren der mathematischen Wissenchaften. Berlin: Springer-Verlag, 1998.

[181] A.M. Samoilenko and N.A. Perestyuk. *Impulsive Differential Equations.* World Scientific Publishing, 1995.

[182] R. G. Sanfelice. “Input-output-to-state stability tools for hybrid systems and their interconnections”. In: *IEEE Trans. Automatic Control* 59.5 (2014), pp. 1360–1366.

[183] R.G. Sanfelice, R. Goebel and A.R. Teel. “Invariance principles for hybrid systems with connections to detectability and asymptotic stability”. In: *IEEE Transactions on Automatic Control* 52.12 (2007), pp. 2282–2297.

[184] R.G. Sanfelice and A.R. Teel. “On singular perturbations due to fast actuators in hybrid control systems”. In: *Automatica* 47 (2011), pp. 692–701.

[185] S. Särkkä. *Bayesian Filtering and Smoothing.* Cambridge, UK: Cambridge University Press, 2013.

[186] A. van der Schaft. *L_2-Gain and Passivity Techniques in Nonlinear Control.* Third. Communications and Control Engineering. Springer International Publishing, 2017.

[187] C. Scherer and C. Ebenbauer. “Convex synthesis of accelerated gradient algorithms”. In: *SIAM Journal on Control and Optimization* 59.6 (2021), pp. 4615–4645.

[188] C.W. Scherer. “Dissipativity and integral quadratic constraints: Tailored computational robustness tests for complex interconnections”. In: *IEEE Control Systems Magazine* 42.3 (2022), pp. 115–139.

[189] J. Schießl, M.H. Baumann, T. Faulwasser and L. L. Grüne. “On the relationship between stochastic turnpike and dissipativity notions”. In: *IEEE Transactions on Automatic Control* (2024).

[190] S. Scholtes. *Introduction to Piecewise Differentiable Equations.* Springer Briefs in Optimization. New York: Springer-Verlag, 2012.

[191] R. Sepulchre, M. Janković and Petar V. Kokotović. *Constructive Nonlinear Control.* Communications and Control Engineering Series. Berlin: Springer-Verlag, 1997.

[192] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan and S. S. Sastry. “Kalman filtering with intermittent observations”. In: *IEEE Transactions on Automatic Control* 49.9 (2004), pp. 1453–1464.

[193] E.D. Sontag. “Smooth stabilization implies coprime factorization”. In: *IEEE Trans. Automatic Control* 34.4 (1989), pp. 435–443.

[194] E.D. Sontag. “Comments on integral variants of ISS”. In: *Systems & Control Letters* 34.1 (1998), pp. 93–100.

[195] E.D. Sontag and A.R. Teel. “Changing supply functions in input/state stable systems”. In: *IEEE Trans. Automatic Control* 40.8 (1995), pp. 1476–1478.

[196] E.D. Sontag and Y. Wang. “On characterizations of the input-to-state stability property”. In: *Systems & Control Letters* 24.5 (1995), pp. 351–359.

[197] M. Souaiby, A. Tanwani and D. Henrion. “Cone-copositive Lyapunov functions for complementarity systems: Converse result and polynomial approximation”. In: *IEEE Transactions on Automatic Control* 67.3 (2022), pp. 1253–1268.

[198] M. Souaiby, A. Tanwani and D. Henrion. “Ensemble approximations for constrained dynamical systems using Liouville equation”. In: *Automatica* 149 (2023), p. 110836.

[199] H. Spohn. *Large scale dynamics of interacting particles*. Springer Science & Business Media, 2012.

[200] D. Stewart and M. Anitescu. “Optimal control of systems with discontinuous differential equations”. In: *Numerische Mathematik* 114 (2012), pp. 653–695.

[201] Y. Suhov and M. Kelbert. *Probability and Statistics by Example. II.* Markov chains: a primer in random processes and their applications. Cambridge University Press, Cambridge, 2008, pp. x+487.

[202] H.J. Sussmann and P.V. Kokotovic. “The peaking phenomenon and the global stabilization of nonlinear systems”. In: *IEEE Transactions on Automatic Control* 36.4 (1991), pp. 424–440.

[203] A.-S. Sznitman. “Topics in propagation of chaos”. In: *Ecole d’été de probabilités de Saint-Flour XIX—1989* 1464 (1991), pp. 165–251.

[204] A. Taghvaei and P.G. Mehta. “An optimal transport formulation of the ensemble Kalman filter”. In: *IEEE Transactions on Automatic Control* 66.7 (2021), pp. 3052–3067.

[205] A. Taghvaei and P.G. Mehta. “A survey of feedback particle filter and related controlled interacting particle systems (CIPS)”. In: *Annual Reviews in Control* 55 (2023), pp. 356–378.

[206] S. Talebi, A. Taghvaei and M. Mesbahi. “Data-driven optimal filtering for linear systems with unknown noise covariances”. In: *Poster at NeurIPS*. 2023.

[207] A. Tanwani. “On analyzing filters with Bayesian parameter inference and Poisson-sampled observations”. In: *Proceedings of 63rd IEEE Conf. on Decision & Control*. 2024.

[208] A. Tanwani and B. Brogliato. “Analyzing feedback interconnections of maximal monotone systems using dissipativity approach”. In: *IFAC-PapersOnLine* 55.30 (2022). Proceedings of 25th Intl. Symp. on Mathematical Theory of Networks and Systems, pp. 156–161.

[209] A. Tanwani, B. Brogliato and C. Prieur. “Well-Posedness and output regulation for implicit time-varying evolution variational inequalities”. In: *SIAM J. Control and Optimization* 56.2 (2018), pp. 751–781.

[210] A. Tanwani, D. Chatterjee and L. Grüne. “Performance bounds for stochastic receding horizon control with randomly sampled measurements”. In: *Proceedings of 58th IEEE Conf. on Decision & Control*. 2019, pp. 2330–2335.

[211] A. Tanwani, D. Chatterjee and D. Liberzon. “Stabilization of continuous-time deterministic systems under random sampling: Overview and recent developments”. In: *Uncertainty in Complex Networked Systems*. Ed. by T. Başar. Springer Nature, 2018, pp. 209–246.

[212] A. Tanwani and H. Shim. “Lyapunov functions for singularly perturbed hybrid systems with frequent jump dynamics”. In: *Proceedings of 60th IEEE Conference on Decision & Control*. 2021.

[213] A. Tanwani, H. Shim and A.R. Teel. “Singularly perturbed hybrid systems for analysis of networks with frequently switching graphs”. In: *IEEE Transactions of Automatic Control* 70.7 (2025), pp. 4344–4359.

[214] A. Tanwani and O. Yufereva. “Error covariance bounds for suboptimal filters with Lipschitzian drift and Poisson-sampled measurements”. In: *Automatica* 122 (2020).

[215] A. Tanwani and O. Yufereva. “Convergence analysis of ensemble filters for linear stochastic systems with Poisson-sampled observations”. In: *IFAC-PapersOnLine* 58.17 (2024). Proceedings of 25th Intl. Symp. on Mathematical Theory of Networks and Systems, pp. 202–207.

[216] A.R. Teel, L. Moreau and D. Nesic. “A unified framework for input-to-state stability in systems with two time scales”. In: *IEEE Transactions on Automatic Control* 48.9 (2003), pp. 1526–1544.

[217] M. Valadier. “Entrainement unilateral, lignes de descente, fonctions Lipschitziennes non pathologiques”. In: *CRAS Paris* 308 (1989), pp. 241–244.

[218] L. Vandenberghe, B.-L. de Moor and J. Vandewalle. “The generalized linear complementarity problem applied to the complete analysis of resistive piecewise-linear circuits”. In: *IEEE Transactions on Circuits and Systems* 36 (1989), pp. 1382–1391.

[219] M. Vidyasagar. *Nonlinear Systems Analysis*. 2nd. Englewood Cliffs, NJ, 1993.

[220] A. Vieira, B. Brogliato and C. Prieur. “Quadratic optimal control of linear complementarity systems: First-Order necessary conditions and numerical analysis”. In: *IEEE Transactions on Automatic Control* 65.6 (2020), pp. 2743–2750.

[221] R. Vinter. *Optimal Control*. Birkhäuser Boston, 2010.

[222] L. Vu, D. Chatterjee and D. Liberzon. “Input-to-state stability of switched systems and switching adaptive control”. In: *Automatica* 43.4 (2007), pp. 639–646.

[223] W. Wang, A.R. Teel and D. Nesic. “Analysis for a class of singularly perturbed hybrid systems via averaging”. In: *Automatica* 48 (2012), pp. 1057–1068.

[224] K. Watanabe. *Adaptive Estimation and Control: Partitioning Approach*. Prentice Hall International (UK), 1992.

[225] F. Wirth. “A converse Lyapunov theorem for linear parameter-varying and linear switching systems”. In: *SIAM Journal on Control and Optimization* 44.1 (2005), pp. 210–239.

[226] W. Wu, J. Chen, M.R. Jovanović and T.T. Georgiou. “Tannenbaum’s gain-margin optimization meets Polyak’s heavy-ball algorithm”. arXiv:2409.19882. 2024.

[227] W. Xiang and J. Xiao. “Stabilization of switched continuous-time systems with all modes unstable via dwell time switching”. In: *Automatica* 50.3 (2014), pp. 940–945.

[228] G. Yang and D. Liberzon. “A Lyapunov-based small-gain theorem for interconnected switched systems”. In: *Systems & Control Letters* 78 (2015), pp. 47–54.

[229] H. Yang, V. Cocquempot and B. Jiang. "On stabilization of switched nonlinear systems with unstable modes". In: *Systems & Control Letters* 58.10 (2009), pp. 703–708.

[230] T. Yang, P.G. Mehta and S.P. Meyn. "Feedback particle filter". In: *IEEE Transactions on Automatic Control* 58.10 (2013), pp. 2465–2480.

[231] S. Yüksel and T. Basar. *Stochastic Networked Control Systems*. Vol. 10. New York, NY: Springer, 2013.

[232] G. Zames. "On the input–output stability of nonlinear time-varying feedback systems–Part I". In: *IEEE Transactions on Automatic Control* 11.2 (1966), pp. 228–238.

[233] G. X. Zhang and A. Tanwani. "ISS Lyapunov functions for cascade switched systems and sampled-data control". In: *Automatica* 105 (2019), pp. 216–227.

[234] L. Zhang, D. Sidoti, A. Bienkowski, K.R. Pattipati, Y. Bar-Shalom and D.L. Kleinman. "On the identification of noise covariances and adaptive Kalman filtering: A new look at a 50 year-old problem". In: *IEEE Access* 8 (2020), pp. 59362–59388.

Publications (by author)

Journal Papers

[J1] A. Dasgupta and A. Tanwani. “Steady state analysis for Kalman filters with Poisson-sampled observations”. In: *Systems & Control Letters* (2025). To appear. Available online: `https://laas.hal.science/hal-04798886`.

[J2] S.P. Chhatoi, A. Tanwani and D. Henrion. “Evolution of measures in nonsmooth dynamical Systems: Formalisms and computation”. In: *Automatica* 179 (2025), p. 112402.

[J3] A. Tanwani, H. Shim and A.R. Teel. “Singularly perturbed hybrid systems for analysis of networks with frequently switching graphs”. In: *IEEE Transactions on Automatic Control* 70.7 (2025), pp. 4344–4359.

[J4] M. Della Rossa and A. Tanwani. “Converse Lyapunov results for stability of switched systems with average dwell-time”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 31.15 (2025).

[J5] M.Z. Almuzakki, B. Jayawardhana, A. Tanwani and A.I. Vakis. “Exponential stabilization of linear systems using nearest-action control with countable input set”. In: *Systems & Control Letters* 196.105992 (2025).

[J6] S. Liu and A. Tanwani. “Impulsive switching signals with functional inequalities: Stability analysis using hybrid systems framework”. In: *Automatica* 171 (2025), p. 111928.

[J7] W.P.M.H. Heemels and A. Tanwani. “Existence and completeness of solutions to extended projected dynamical systems and sector-bounded projection-based controllers”. In: *IEEE Control Systems Letters* 7 (2023), pp. 1590–1595.

[J8] S. Hafstein and A. Tanwani. “Linear programming based lower bounds on average dwell-time via multiple Lyapunov functions”. In: *European Journal of Control* 74 (2023), p. 100838.

[J9] M. Souaiby, A. Tanwani and D. Henrion. “Ensemble approximations for constrained dynamical systems using Liouville equation”. In: *Automatica* 149 (2023), p. 110836.

[J10] A. Tanwani. “Suboptimal filtering over sensor networks with random communication”. In: *IEEE Transactions on Automatic Control* 67.10 (2022), pp. 5456–5463.

[J11] S. Liu, A. Tanwani and D. Liberzon. “ISS and integral-ISS of switched systems with nonlinear supply functions”. In: *Mathematics of Controls, Signals, and Systems* 34 (2022), pp. 297–327.

[J12] M. Souaiby, A. Tanwani and D. Henrion. “Cone-copositive Lyapunov functions for complementarity systems: Converse result and polynomial approximation”. In: *IEEE Transactions on Automatic Control* 67.3 (2022), pp. 1253–1268.

[J13] M.-K. Camlibel, L. Iannelli and A. Tanwani. “Convergence of proximal solutions for evolution inclusions with time-dependent maximal monotone operators”. In: *Mathematical Programming, Ser. A* 194 (2022), pp. 1017–1059.

[J14] M.Z. Almuzakki, B. Jayawardhana and A. Tanwani. “Nearest neighbor control for practical stabilization of passive nonlinear systems”. In: *Automatica* 141 (2022), p. 110278.

[J15] M. Della Rossa, A. Tanwani and L. Zaccarian. “Nonpathological ISS-Lyapunov functions for interconnected differential inclusions”. In: *IEEE Transactions on Automatic Control* 67.8 (2022), pp. 3774–3789.

[J16] M. Della Rossa and A. Tanwani. “Instability of dwell-time constrained switched nonlinear systems”. In: *Systems & Control Letters* 162 (2022), p. 105164.

[J17] M. Della Rossa, R. Goebel, A. Tanwani and L. Zaccarian. “Piecewise structure of Lyapunov functions and densely checked decrease conditions for hybrid systems”. In: *Mathematics of Control, Signals, and Systems* 33 (2021), pp. 123–149.

[J18] A. Tanwani and O. Yufereva. “Error covariance bounds for suboptimal filters with Lipschitzian drift and Poisson-sampled measurements”. In: *Automatica* 122 (2020).

[J19] M. Della Rossa, A. Tanwani and L. Zaccarian. “Max–min Lyapunov functions for switched systems and related differential inclusions”. In: *Automatica* 120 (2020), p. 109123.

[J20] B. Brogliato and A. Tanwani. “Dynamical systems coupled with monotone set-valued operators: Formalisms, applications, well-posedness, and stability”. In: *SIAM Review* 62.1 (2020), pp. 3–129.

[J21] A. Tanwani and Q. Zhu. “Feedback Nash equilibrium for randomly switching differential-algebraic games”. In: *IEEE Transactions on Automatic Control* 65.8 (2020), pp. 3286–3301.

[J22] G.X. Zhang and A. Tanwani. “ISS Lyapunov functions for cascade switched systems and sampled-data control”. In: *Automatica* 105 (2019), pp. 216–227.

[J23] A. Tanwani and S. Trenn. “Detectability and observer design for switched differential–algebraic equations”. In: *Automatica* 99 (2019), pp. 289–300.

[J24] A. Tanwani, B. Brogliato and C. Prieur. “Well-Posedness and output regulation for implicit time-varying evolution variational inequalities”. In: *SIAM J. Control and Optimization* 56.2 (2018), pp. 751–781.

[J25] A. Tanwani and S. Trenn. “Determinability and state-estimation for switched differential-algebraic equations”. In: *Automatica* 76 (2017), pp. 17–31.

[J26] A. Tanwani, B. Brogliato and C. Prieur. “Observer design for unilaterally constrained Lagrangian systems: A passivity-based approach”. In: *IEEE Transactions on Automatic Control* 61.9 (2016), pp. 2386–2401.

[J27] A. Tanwani, C. Prieur and M. Fiacchini. “Observer-based feedback stabilization of linear systems with event-triggered sampling and dynamic quantization”. In: *Systems & Control Letters* 94 (2016), pp. 46–56.

[J28] A. Tanwani, B. Brogliato and C. Prieur. “Stability notions for a class of nonlinear systems with measure controls”. In: *Mathematics of Control, Signals, and Systems* 27.2 (2015), pp. 245–275.

[J29] A. Tanwani, H. Shim and D. Liberzon. "Comments on "Observability of Switched Linear Systems: Characterization and Observer Design"". In: *IEEE Transactions on Automatic Control* 60.12 (2015), pp. 3396–3400.

[J30] A. Tanwani, B. Brogliato and C. Prieur. "Stability and observer design for Lur'e systems with multivalued, non-monotone, time-varying nonlinearities and state jumps". In: *SIAM J. Control and Optimization* 56.6 (2014), pp. 3639–3672.

[J31] H. Shim and A. Tanwani. "Hybrid-type observer design based on a sufficient condition for observability in switched nonlinear systems". In: *Int. J. Robust & Nonlinear Control: Special Issue on High Gain Observers and Nonlinear Output Feedback Control* 24.6 (2014), pp. 1064–1089.

[J32] A. Tanwani, H. Shim and D. Liberzon. "Observability for switched linear systems: Characterization and observer design". In: *IEEE Transactions on Automatic Control* 58.4 (2013), pp. 891–904.

[J33] A. Tanwani, A.D. Dominguez-Garcia and D. Liberzon. "An inversion-based approach to fault detection and isolation in switching electrical networks". In: *IEEE Transactions on Control Systems Technology* 19.5 (2011), pp. 1059–1074.

[J34] A. Tanwani and D. Liberzon. "Invertibility of switched nonlinear systems". In: *Automatica* 46.12 (2010), pp. 1962–1973.

Book Chapters

[BC1] A. Tanwani, D. Chatterjee and D. Liberzon. "Stabilization of Deterministic Control Systems Under Random Sampling: Overview and Recent Developments". In: *Uncertainty in Complex Networked Systems*. Ed. by T. Başar. Systems & Control: Foundations & Applications. Switzerland: SpringerNature, 2018, pp. 209–246.

[BC2] A. Tanwani, C. Prieur and S. Tarbouriech. "Stabilization of linear hyperbolic systems of balance laws with measurement errors". In: *Control Subject to Computational and Communication Constraints: Current Challenges*. Vol. 475. Springer, Cham, 2018, pp. 357–374.

[BC3] C. Prieur and A. Tanwani. "Asymptotic stabilization of some finite and infinite dimensional systems by means of dynamic event-triggered output feedbacks". In: *Feedback Stabilization of Controlled Dynamical Systems*. Vol. 473. Springer, Cham, 2017, pp. 201–230.

[BC4] A. Tanwani, H. Shim and D. Liberzon. "Observer design for switched linear systems with state jumps". In: *Hybrid Dynamical Systems*. Springer, 2015, pp. 179–203.

[BC5] M. Petreczky, A. Tanwani and S. Trenn. "Observability of switched linear systems". In: *Hybrid dynamical systems*. Springer, 2015, pp. 205–240.

Conference Papers (Published and Peer-Reviewed)

[C1] A. Tanwani. "On analyzing filters with Bayesian parameter inference and Poisson-sampled observations". In: *Proceedings of 63rd IEEE Conf. on Decision & Control*. 2024.

[C2] S.P. Chhatoi, A. Tanwani and D. Henrion. “Optimal control of nonsmooth dynamical systems using measure relaxations”. In: *Proceedings of 63rd IEEE Conf. on Decision & Control*. 2024.

[C3] A. Tanwani and O. Yufereva. “Convergence analysis of ensemble filters for linear stochastic systems with Poisson-sampled observations”. In: *Proceedings of 26th Intl. Symp. on Mathematical Theory of Networks and Systems*. 2024.

[C4] O. Yufereva and A. Tanwani. “Transport inspired particle filters with Poisson-sampled observations in Gaussian setting”. In: *Proceedings of 62nd IEEE Conf. on Decision & Control*. 2023, pp. 7695–7700.

[C5] O. Yufereva and A. Tanwani. “Approximation of nonlinear filters for continuous-time Markov chains under randomly-sampled observations”. In: *Proceedings of 61st IEEE Conf. on Decision & Control*. 2022, pp. 6620–6625.

[C6] A. Tanwani and B. Brogliato. “Analyzing feedback interconnections of maximal monotone systems using dissipativity approach”. In: *IFAC-PapersOnLine* 55.30 (2022). Proceedings of 25th Intl. Symp. on Mathematical Theory of Networks and Systems, pp. 156–161.

[C7] A. Tanwani and H. Shim. “Lyapunov functions for singularly perturbed hybrid systems with frequent jump dynamics”. In: *Proceedings of 60th IEEE Conf. on Decision & Control*. 2021, pp. 5382–5387.

[C8] A. Tanwani. “Error bounds for locally optimal distributed filters with random communication graphs”. In: *Proceedings of 60th IEEE Conf. on Decision & Control*. 2021, pp. 1609–1614.

[C9] A. Tanwani. “Poster: Filtering over networks with random communication”. In: *Proceedings of 24th Intl. Conf. on Hybrid Systems: Computation and Control*. 2021.

[C10] K. Camlibel and A. Tanwani. “A discretization algorithm for time-varying composite gradient flow dynamics”. In: *IFAC-PapersOnLine* 54.9 (2021). Proceedings of 24th Intl. Symp. on Mathematical Theory of Networks and Systems, pp. 558–563.

[C11] S. Liu, A. Tanwani and D. Liberzon. “Average dwell-time bounds for ISS and integral ISS of switched systems using Lyapunov functions”. In: *Proceedings of 59th IEEE Conf. on Decision & Control*. 2020, pp. 6291–6296.

[C12] M. Souaiby, A. Tanwani and D. Henrion. “Computation of Lyapunov functions under state constraints using semidefinite programming hierarchies”. In: *IFAC-PapersOnLine* 53.2 (2020). Proceedings of 21st IFAC World Congress, pp. 6281–6286.

[C13] A. Tanwani, D. Chatterjee and L. Grüne. “Performance bounds for stochastic receding horizon control with randomly sampled measurements”. In: *Proceedings of 58th IEEE Conf. on Decision & Control*. 2019, pp. 2330–2335.

[C14] M. Della Rossa, R. Goebel, A. Tanwani and L. Zaccarian. “Almost everywhere conditions for hybrid Lipschitz Lyapunov Functions”. In: *Proceedings of 58th IEEE Conf. on Decision & Control*. 2019, pp. 8148–8153.

[C15] B. Jayawardhana, M.Z. Almuzakki and A. Tanwani. “Practical stabilization of passive nonlinear systems with limited control”. In: *IFAC-PapersOnLine* 52.16 (2019). Proceedings of 11th IFAC Symposium on Nonlinear Control Systems (NolCoS), pp. 460–465.

[C16] M. Della Rossa, A. Tanwani and L. Zaccarian. “Smooth approximation of patchy Lyapunov functions for switching systems”. In: 52.16 (2019). Proceedings of 11th IFAC Symposium on Nonlinear Control Systems (NolCoS), pp. 364–369.

[C17] M. Della Rossa, A. Tanwani and L. Zaccarian. “Max-Min Lyapunov functions for switching differential inclusions”. In: *Proceedings of 57th IEEE Conf. on Decision & Control*. 2018, pp. 5664–5669.

[C18] A. Tanwani, R. Jungers and W.P.M.H Heemels. “Observability of discrete-time linear systems with communication protocols and dropouts”. In: *Proceedings of 57th IEEE Conf. on Decision & Control*. 2018, pp. 4194–4199.

[C19] A. Tanwani and Q. Zhu. “Feedback Nash equilibrium for Markov jump games under differential-algebraic constraints with application to robust control”. In: *Proceedings of 57th IEEE Conf. on Decision & Control*. 2018, pp. 1124–1129.

[C20] A. Tanwani, S. Marx and C. Prieur. “Local input-to-state stabilization of 1-D linear reaction-diffusion equation with bounded feedback”. In: *Proceedings of 23rd Intl. Symp. on Mathematical Theory of Networks and Systems*. 2018, pp. 576–581.

[C21] G.-X. Zhang and A. Tanwani. “Stabilization of switched nonlinear systems with dynamic output feedback and event-based sampling”. In: *Proceedings of American Control Conference*. 2018, pp. 2198–2203.

[C22] A. Tanwani and A.R. Teel. “Stabilization with event-driven controllers over a digital communication channel with random transmissions”. In: *Proceedings of 56th IEEE Conf. on Decision & Control*. 2017, pp. 6063–6068.

[C23] N. Espitia, A. Tanwani and S. Tarbouriech. “Stabilization of boundary controlled hyperbolic PDEs via Lyapunov-based event triggered sampling and quantization”. In: *Proceedings of 56th IEEE Conf. on Decision & Control*. 2017, pp. 1266–1271.

[C24] A. Tanwani and S. Trenn. “Observer design for detectable switched differential-algebraic equations”. In: *IFAC-PapersOnLine* 50.1 (2017). Proceedings of 20th IFAC World Congress, pp. 2953–2958.

[C25] K. Camlibel, L. Iannelli, A. Tanwani and S. Trenn. “Differential-algebraic inclusions with maximal monotone operators”. In: *Proceedings of 55th IEEE Conf. on Decision & Control*. 2016, pp. 610–615.

[C26] A. Tanwani, C. Prieur and S. Tarbouriech. “Input-to-state stabilization in H^1-norm for boundary controlled linear hyperbolic PDEs with application to quantized control”. In: *Proceedings of 55th IEEE Conf. on Decision & Control*. 2016, pp. 3112–3117.

[C27] A. Tanwani, A.R. Teel and C. Prieur. “On using norm estimators for event-triggered control with dynamic output feedback”. In: *Proceedings of 54th IEEE Conf. on Decision & Control*. 2015, pp. 5500–5505.

[C28] A. Tanwani and S. Trenn. “On detectability of switched linear differential-algebraic equations”. In: *Proceedings of 54th IEEE Conf. on Decision & Control*. 2015, pp. 2957–2962.

[C29] A. Tanwani. “Output regulation in differential variational inequalities using internal model principle and passivity-based approach”. In: *Proceedings of Applied Mathematics and Mechanics (PAMM)* 15.1 (2015), pp. 31–34.

[C30] A. Tanwani, B. Brogliato and C. Prieur. “On output regulation in systems with differential variational inequalities”. In: *Proceedings of 53rd IEEE Conf. on Decision & Control*. 2014, pp. 3077–3082.

[C31] A. Tanwani, B. Brogliato and C. Prieur. “On output regulation in state-constrained systems: An application to polyhedral case”. In: *IFAC Proceedings Volumes* 47.3 (2014). Proceedings of 19th IFAC World Congress, pp. 1513–1518.

[C32] A. Tanwani, B. Brogliato and C. Prieur. “Passivity-based observer design for a class of Lagrangian systems with perfect unilateral constraints”. In: *Proceedings of 52nd IEEE Conf. on Decision & Control.* 2013, pp. 3338–3343.

[C33] A. Tanwani and S. Trenn. “An observer for switched differential-algebraic equations based on geometric characterization of observability”. In: *Proceedings of 52nd IEEE Conf. on Decision & Control.* 2013, pp. 5981–5986.

[C34] A. Tanwani, B. Brogliato and C. Prieur. “On stability of measure differential equations”. In: *IFAC Proceedings Volumes* 46.23 (2013). Proceedings of 9th IFAC Symposium on Nonlinear Control Systems (NolCoS), pp. 241–246.

[C35] H. Shim, A. Tanwani and Z. Ping. “Back-and-forth operation of state observers and norm estimation of estimation error”. In: *Proceedings of 51st IEEE Conf. on Decision & Control.* 2012, pp. 3221–3226.

[C36] A. Tanwani and S. Trenn. “Observability of switched differential-algebraic equations for general switching signals”. In: *Proceedings of 51st IEEE Conf. on Decision & Control.* 2012, pp. 2648–2653.

[C37] A. Tanwani and D. Liberzon. “Robust Invertibility of Switched Linear Systems”. In: *Proceedings of 50th IEEE Conf. on Decision and Control.* 2011, pp. 441–446.

[C38] H. Shim and A. Tanwani. “On a sufficient condition for observability of nonlinear switched systems and observer design strategy”. In: *Proceedings of American Control Conference.* 2011, pp. 441–446.

[C39] A. Tanwani, H. Shim and D. Liberzon. “Observability implies observer design for switched linear systems”. In: *Proceedings of ACM Conf. on Hybrid Systems: Computation and Control.* 2011, pp. 3–12.

[C40] A. Tanwani and S. Trenn. “On observability of switched differential-algebraic equations”. In: *Proceedings of 49th IEEE Conf. on Decision & Control.* 2010, pp. 5656–5661.

[C41] A. Tanwani and D. Liberzon. “Invertibility of nonlinear switched systems”. In: *Proceedings of 47th IEEE Conf. on Decision & Control.* 2008, pp. 286–291.