University of California
Santa Barbara

Design and Analysis of Hybrid and Hybrid-Inspired Control Systems in Stochastic and Non-Stochastic Settings

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

Doctor of Philosophy
in
Electrical and Computer Engineering

by

Matina Baradaran Hosseini

Committee in charge:

Professor Andrew R. Teel, Chair
Professor Francesco Bullo
Professor Katie Byl
Professor Joao Hespanha

December 2022

The Dissertation of Matina Baradaran Hosseini is approved.

Professor Francesco Bullo

Professor Katie Byl

Professor Joao Hespanha

Professor Andrew R. Teel, Committee Chair

November 2022

Design and Analysis of Hybrid and Hybrid-Inspired Control Systems in Stochastic and Non-Stochastic Settings

by

Matina Baradaran Hosseini

To my dearest parents, little brother,
and my beloved partner

Acknowledgements

Many wonderful people played significant roles in my life as a Ph.D. student during the past 5 years, and if it was not for their support I would have had little success in completing this dissertation.

First and foremost, I would like to express my gratitude to my advisor, Andrew Teel, for his patience with me, his consistent guidance, mentorship, availability, and his integrity to his students. Prof. Teel has taught me the art of asking the right questions when it comes to addressing problems and critical thinking. He has equipped me with the ability to form/solve problems with rigorous mathematical and analytical tools as a researcher and an engineer. Working under Andy has enabled me to grow both personally and professionally. Andy's attention to detail and his rigor in communication are invaluable lessons for me, that I only hope to be able to resemble for the rest of my career as one of his students. His flexibility and utmost understanding of my personal experience as a Ph.D. student at UCSB made my journey smooth and memorable despite many obstacles on the way.

I am very grateful for the existence of the Center for Control, Dynamical Systems, and Computation (CCDC), which kept me informed of the ongoing research in my field and allowed me to meet many distinguished researchers in person. I am also grateful for having Prof. Hespanha, Prof. Bullo, and Prof. Byl, on my committee. I could not have asked for a better committee of experts from the controls field and I appreciate their invaluable feedback.

I would like to thank my collaborators, Justin Le, and Jorge Poveda. Without their discerning questions and feedback, our works might not have been successfully published. I am grateful for the support and camaraderie of my peers, whose presence at the department was a comfort for me in distress. In particular, Sina Zamani and Murat Kaan Erdal, without whose true friendships, and corroborating presence through hard times since day one I could not have made it out of graduate school. Not only I have learned from so many research-related discus-

sions we had, but also have I learned from their generosity in helping others unconditionally and picking their peers up.

I am also lucky to have been surrounded by fantastic controls and dynamic systems community, with special mention to Guosong Yang, Sharad Shankar, Henrique Farrez, Bryce Ferguson, Daniel Lazar, and Maurice Filo. During my first 6 months of starting graduate school, Maurice shared his useful insights on the Ph.D. journey helping me take the first steps right. I'd like to also highlight my gratitude to Guosong for our intuitive discussions regarding research. Guosong's insight into nonlinear control systems could always shed light on the dead-end in my research. Finally, I want to use this opportunity to appreciate my fellow female comrades Elizabeth Huang and Sana Amani for speaking up about their experiences as women in Engineering departments. Their experiences taught me that it is not solely about including women in STEM, but about cultivating a culture of respecting women as peers and allowing them to be their authentic selves without having to fit in or stay silent to keep the status quo. I am also appreciative of Valerie de Veyra for always bringing the paperwork in order and for her impeccable memory.

I want to acknowledge my beloved Santa Barbara family, Majid, Pita, Scott, and Lisa, without whom my life in this city would not have been as enjoyable. I am also very thankful to my aunt Catherine for her heartening encouragement during the past years.

Finally, and most significantly, I want to thank my mom, Laya for her selfless support through all these years from afar, my Dad Amir who is my first inspiration for following engineering, my brother Moeed, and of course my handsome partner for his unconditional love.

Curriculum Vitæ

Matina Baradaran Hosseini

Education

2022	Ph.D. in Electrical and Computer Engineering (Expected), University of California, Santa Barbara.
2021	M.S. in Electrical and Computer Engineering, University of California, Santa Barbara.
2016	B.E. in Electrical Engineering, Aachen University of Applied Sciences, Aachen, Germany

Experience

Summer 2022	Modeling/Control Intern at Bloom Energy, San Jose, CA
Winter 2022	Teacher Assistant, UCSB, CA
Summer 2021	Battery Technology Intern at Romeo Power, Los Angeles, CA
Summer 2019	Control Research Intern at Robert Bosch LLC, Sunnyvale, CA
2017 - 2022	Graduate Student Researcher, UCSB, CA
Fall 2017	Teacher Assistant, UCSB, CA
2015 - 2017	Research Assistant at Lucas Nuelle GmbH, Kerpen, Germany
2015 - 2017	Research Assistant at RWTH University (Institute of Automatic Control), Aachen, Germany
2015 - 2016	Research Assistant at RWTH University (Institute of High Voltage Technology), Aachen, Germany
Summer 2015	Engineering Intern at TenneT TSO, Bayreuth, Germany

Publications

- Matina Baradaran, Justin H. Le, Andrew R. Teel, "Input-to-state stability of soft-reset systems with nonlinear data" submitted for "A Tribute to Eduardo Sontag for his 70th Birthday in Mathematics of Control, Signals, and Systems", Springer (2022)
- Matina Baradaran, Justin H. Le, Andrew R. Teel, "Analyzing the Persistent Asset Switches in Continuous Hybrid Optimization Algorithms" presented in "American Control Conference, New Orleans, Louisiana, USA" (2021)
- Matina Baradaran, Andrew R. Teel, "Omega-limit sets and robust stability for switched systems with distinct equilibria" presented in "IFAC World Congress of the International Federation of Automatic Control, Berlin, Germany" (2020)
- Matina Baradaran, Andrew R. Teel, "Global Optimization on the Sphere with Half-space Constraints: A Stochastic Hybrid Systems Approach" presented in "IEEE Conference on Decision and Control, Nice, France" (2019)

- Matina Baradaran, Jorge I. Poveda, Andrew R. Teel, "Global Optimization on the Sphere: A Stochastic Hybrid Systems Approach" presented in "Proc. NOLCOS, Vienna, Austria" (2019)
- Matina Baradaran, Jorge I. Poveda, Andrew R. Teel, "Achieving Global Optimization on Smooth Manifolds: A Stochastic Hybrid Systems Approach" presented in "IEEE Conference on Decision and Control, Miami, FL, USA" (2018)

Awards

UCSB ECE Dissertation Fellowship	(Spring quarter 2022)
UCSB ECE Outstanding Teacher Assistant Award	(Winter quarter 2022)
Student Travel Award for "IEEE Conference on Decision and Control"	(Dec 2018)

Abstract

Design and Analysis of Hybrid and Hybrid-Inspired Control Systems in Stochastic and Non-Stochastic Settings

by

Matina Baradaran Hosseini

This dissertation is divided into three parts. The first part presents three chapters on a class of stochastic dynamical systems designed to solve non-convex optimization problems on smooth manifolds. The first chapter develops the stochastic, hybrid optimization algorithm. In this chapter, we show that the proposed dynamics combine continuous-time flows, characterized by a differential equation, and discrete-time jumps, characterized by a stochastic difference inclusion in order to guarantee convergence with probability one to the set of global minimizers of the cost function. By using the framework of stochastic hybrid inclusions, a detailed stability characterization of the dynamics, as well as a simple extension to address learning problems in games defined on manifolds is provided. In the second chapter, we cast a stochastic, hybrid algorithm for global optimization on the unit sphere using the framework of stochastic hybrid inclusions. The algorithm includes hysteresis switching between two coordinate charts in order to be able to fully explore the sphere by flowing without encountering singularities in the flow vector field. It also combines gradient flow with jumps that aim to escape singular points of the function to minimize, other than those singular points corresponding to global minima. For this case, the algorithm is stochastic because the jumps involve random probing on the sphere. Solutions are not unique because the jumps are governed by a set-valued mapping, i.e., an inclusion. Regarding the coordinate charts employed, we discuss both the use of spherical coordinates as well as stereographic projection. By using the framework of stochastic hybrid inclusions, we establish uniform global asymptotic stability in probability

for the set of global minimizers for arbitrary continuously differentiable (C^1) functions defined on the sphere. Lastly in the third chapter, we develop a stochastic, hybrid optimization algorithm for globally minimizing an arbitrary (C^1) function on the unit sphere intersected with an arbitrary half-space in $\mathbb{R}^3$. Hysteresis switching between coordinate charts is used to enable the algorithm to fully explore the sphere by flowing. During flows, the optimization algorithm uses (projected) gradient descent when near the boundary of the half-space. It may use an update rule inspired by accelerated gradient methods away from the boundary of the half-space. It uses hysteresis switching between the two continuous-time update methods. Periodically, stochastic probing on the sphere is used to attempt to improve the value of the cost function. A stability analysis of the algorithm is provided and the algorithm is demonstrated on a numerical example.

The second part of this dissertation consists of two chapters. The first chapter characterizes the asymptotic behavior that results from switching among asymptotically stable systems with distinct equilibria when the switching frequency satisfies an average dwell-time constraint with a small average rate. The asymptotic characterization is in terms of the Omega-limit set of an associated ideal hybrid system containing an average dwell-time automaton with the rate parameter set equal to zero. This set is globally asymptotically stable for the ideal system. The actual switched system, including small disturbances, constitutes a small perturbation of this ideal system, resulting in semi-global, practical asymptotic stability. In the second chapter, we consider some of convex optimization engineering challenges, such as those involving multi-agent systems and resource allocation, where the objective function can persistently switch during the execution of an optimization algorithm. Motivated by such applications, in Chapter 6 we analyze the effect of persistently switching objectives in continuous-time optimization algorithms. In particular, we take advantage of the robust stability results from Chapter 5 for switched systems with distinct equilibria and extend these results to systems described by differential inclusions, making the results applicable to recent optimization algorithms that em-

ploy differential inclusions for improving efficiency and/or robustness. Within the framework of hybrid systems theory, we provide an accurate characterization, in terms of Omega-limit sets, of the set to which the optimization dynamics converge. Finally, by considering the switching signal to be constrained in its average dwell time, we establish semi-global practical asymptotic stability of these sets with respect to the dwell-time parameter.

In the third part of this dissertation, Input-to-state stability (ISS) is considered for a nonlinear "soft-reset" system with inputs. The latter is a system that approximates a hard-reset system, which is modeled as a hybrid system with inputs. In contrast, a soft-reset system is modeled as a differential inclusion with inputs. Lyapunov conditions on the hard-reset system are given that guarantee ISS for the soft-reset system. In turn, it is shown when global asymptotic stability for the origin of the zero-input reset system guarantees ISS for nonzero inputs. Examples are given to demonstrate the theory.

Contents

Chapter 1

Introduction

Hybrid dynamical systems refer to a class of dynamical systems including both discrete and continuous time dynamics interacting and governing a certain system's evolution in time. Many classic dynamical systems are either continuous or discrete, but the need for hybrid dynamical systems arises when there is an interaction between digital discrete systems and continuous physical processes. Due to the increasing interactions between time-driven continuous dynamics and event-based discrete dynamics observed in many of today's applications and the tight interaction between them, one might not be able to study the discrete and continuous behavior separately and ignore the coupling. Many of these couplings between the continuous and discrete dynamics can be observed in advanced control systems, where multiple control laws are combined in order to address broader scenarios and with better performance i.e., control switch to avoid collision [1], control of multi-robot systems [2], control of legged robots [3] to name a few. The importance of hybrid systems in practice has attracted attention from the control system community during the past decades [4],[5], [6]. The multidisciplinary nature of hybrid dynamical systems can be challenging to study and analyze especially when stochasticity and randomness interplay with these systems. Stochastic hybrid systems are dynamical systems that combine continuous change and instantaneous change and that include random effects.

We can see such hybrid systems with stochasticity in many subclasses like Markov jump systems [7], stochastic switched systems [8], and stochastic hybrid inclusions [9]. Understanding the design and modeling of hybrid and stochastic hybrid systems has become a necessity. The model of a hybrid dynamical system can be represented in the following form using ordinary differential and difference equations

$$\dot{x} = f(x), \qquad x^+ = g(x). \tag{1.1}$$

To extend the above hybrid system into a more general setting, we can use differential and difference inclusions and have

$$\dot{x} \in F(x), \qquad x^+ \in G(x), \tag{1.2}$$

where F and G are set-valued mappings describing the continuous flow map and the discrete jump map respectively. The hybrid system (1.1) can be now considered as a particular case of (1.2). In the next step, we specify the space in which the continuous and discrete dynamics are allowed to evolve, namely the flow set C and the jump set D:

$$\begin{aligned} x \in C, \qquad & \dot{x} \in F(x), \\ x \in D, \qquad & x^+ \in G(x). \end{aligned}$$

The precise mathematical characterization of the properties of this system, the definition of solutions, stability notions, and generic robustness principles are given in [4]. This modeling framework is the fundamental hybrid system model, which will be used throughout this dissertation for stability analysis and modeling. We consider this fundamental model as the pillar for various examples that we study in this dissertation. Depending on the problem addressed in each chapter, there would be variations of this system with randomness or systems inspired

by this model introduced. Due to the hybrid system's theoretical and practical importance, it is important to obtain a comprehensive view of this field by designing and modeling hybrid and hybrid-inspired control systems and optimization algorithms implemented and tailored for different scenarios and problem settings. The goal of this dissertation is to design and analyze different systems in the framework of hybrid, stochastic hybrid, and hybrid-inspired systems through various examples and applications.

Part I

A Stochastic hybrid systems approach for global optimization on general manifolds

Chapter 2

Stochastic hybrid inclusions applied to global almost sure optimization on manifolds

There has been a renewed interest in developing dynamical systems to solve optimization problems [10], [11], [12], [13], [14]. These developments have led to novel algorithms for power systems [15], resource allocation [16], control of autonomous robots [17], stabilization of rigid bodies evolving on manifolds [18], and adaptive control problems [19], for example.

Designing such dynamical systems whose solutions from every possible initial condition converge to the solution of the optimization problem has been studied under convexity-like assumptions. However, the development of efficient algorithms that guarantee the global convergence property in non-convex settings is still an open research question. In fact, it is well-known that standard gradient-based dynamical systems cannot escape local minima or saddle points of the cost function [20, pp. 2]. Moreover, it is also well-known [21] that robust convergence to the set of global minimizers of an optimization problem with saddles and local minima cannot be achieved using a differential equation with a continuous right-hand side. On

the other hand, it has been shown in [22], [18], and [14], that deterministic *hybrid* optimization algorithms can overcome these limitations, provided the location of the critical points is known a priori. This restrictive assumption has motivated the development of systematic analytical frameworks for the design of optimization algorithms with randomization. For instance, random initialization and gradient descent is used in [23] to establish almost sure convergence to the set of local minimizers of a cost function. Stochastic gradient descent is used in [24] to escape saddles, under the assumption of having a Hessian matrix at the saddle point with a strictly negative eigenvalue. A finite-time algorithm based on gradient descent with small random perturbations is used in [20] to escape saddle points, and a similar stochastically perturbed version of Nesterov's accelerated gradient descent is studied in [25].

Motivated by this recent research line, we present in this chapter a family of optimization algorithms modeled as dynamical systems, designed to solve non-convex optimization problems on general smooth manifolds. In contrast to the existing results, the proposed method relies on asymptotic analytical tools for stochastic dynamical systems, such as Lyapunov functions and invariance principles. This fact allows us to establish a *uniform global asymptotic stability in probability* result for the system. Since the existence of degenerate saddle points and local minima prevent the implementation of purely continuous-time gradient-based optimization algorithms, we propose a class of optimization dynamics that implement continuous-time *and* discrete-time updates of the optimization variable, injecting stochasticity during the discrete-time updates of the system. Given that the dynamics of the algorithm are characterized by differential and difference inclusions rather than difference and differential equations, the optimization dynamics are modeled as a stochastic hybrid inclusion (SHI), a class of systems recently studied in [9]. Using Lyapunov-based tools and invariance principles for SHI, we characterize a class of non-convex optimization problems defined on general smooth manifolds for which global asymptotic stability in probability (GASp) can be established for the set of global minimizers of the cost function. Subsequently, we show that the hybrid stochastic

optimization dynamics can be adapted to solve distributed optimization and learning problems in a class of multi-agent systems (MAS) having a global potential function and modeled as a weighted potential game. For this type of problem it is shown that, provided a global intermittent communication protocol is implemented, GASp can be achieved for the set of global minimizers of the potential function.

2.1 Notation

We denote by $\mathbb{R}$ ($\mathbb{R}_{\geq 0}$) the set of real numbers (resp. non-negative real numbers), and by $\mathbb{Z}$ ($\mathbb{Z}_{\geq j}$) the set of integers (resp. set of all integers greater than or equal to j). A set-valued mapping $M : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ is outer semicontinuous (OSC) if, for each $(x_i, y_i) \mapsto (x, y) \in \mathbb{R}^p \times \mathbb{R}^n$ satisfying $y_i \in M(x_i)$ for all $i \in \mathbb{Z}_{\geq 0}$, we have $y \in M(x)$. A mapping M is locally bounded (LB) if, for each bounded set $K \in \mathbb{R}^p$, $M(K) := \cup_{x \in K} M(x)$ is bounded. $\mathbf{B}(\mathbb{R}^m)$ denotes the Borel field, i.e., the subsets of $\mathbb{R}^m$ generated from open subsets of $\mathbb{R}^m$ through complements and finite and countable unions. A set $F \subset \mathbb{R}^m$ is measurable if $F \in \mathbf{B}(\mathbb{R}^m)$. A mapping $M : \mathbb{R}^p \rightrightarrows \mathbb{R}^n$ is measurable if for each open $\mathcal{O} \subset \mathbb{R}^n$ the set $M^{-1}(\mathcal{O}) := \{v \in \mathbb{R}^p : M(v) \cap \mathcal{O} \neq \emptyset\}$ is measurable. The set $\mathbb{B}$ is defined as $\mathbb{B} := \{x \in \mathbb{R}^n : |x|_\infty \leq 1\}$ where $|x|_\infty := \max_{i \in (1,..,n)} |x_i|$.

2.2 Stochastic hybrid inclusions

2.2.1 Models

We consider optimization algorithms modeled by stochastic hybrid inclusions [9], which have the form

$$x \in C \quad \dot{x} \in F(x) \tag{2.1a}$$

$$x \in D \quad x^+ \in G(x, v^+) \qquad v \sim \mu(\cdot) \tag{2.1b}$$

where $x \in \mathbb{R}^n$ is the state of the system and v^+ is a placeholder for a sequence of i.i.d. random variables $\{\mathbf{v}_i\}_{i=1}^{\infty}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The distribution of each $\mathbf{v}_i : \Omega \to \mathbb{R}^m$ is given by $\mu(A) = \mathbb{P}(\omega \in \Omega : \mathbf{v}_i(\omega) \in A)$ for each $i \in \mathbb{Z}_{\geq 1}$ and each $A \in \mathbf{B}(\mathbb{R}^m)$. A distinguishing feature of these models compared to most other stochastic hybrid systems models in the literature (for a survey, see, [26]) is the allowance of non-unique solutions. Continuous evolution of the state x is allowed in the flow set $C \subset \mathbb{R}^n$, while jumps are allowed in the jump set $D \subset \mathbb{R}^n$. These sets may overlap. Flows are governed by the differential inclusion determined by the set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, called the flow map. Jumps are governed by the difference inclusion determined by the set-valued $G : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$. Set-valued mappings may be appropriate to capture ensembles of solutions in a single model or may be the result of a regularization required to ensure robustness or the validity of certain relaxed Lyapunov conditions for recurrence or asymptotic stability in probability. Indeed, such conditions often require the following stochastic hybrid basic conditions:

Assumption 1 *The data* (C, F, D, G) *satisfies:*

1. *The sets* $C, D \subset \mathbb{R}^n$ *are closed.*

2. *The mapping* $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *is locally bounded and outer semicontinuous (i.e., has a*

closed graph) and $F(x)$ is nonempty and convex for each $x \in C$.

3. *The mapping $G : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is locally bounded and the graphical mapping $v \mapsto \text{graph}(G(\cdot, v)) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in G(x, v)\}$ is measurable with respect to $\mathbf{B}(\mathbb{R}^{2n})$ with closed values.* ■

2.2.2 Solutions

The solution concept for (2.1) is adapted from the solution concept for non-stochastic hybrid systems in [27]. Following this, a *compact hybrid time domain* is a set of the form $\bigcup_{i=0}^{J}([t_i, t_{i+1}] \times \{i\})$ where $0 = t_0 \leq t_1 \leq \cdots t_{J+1} < \infty$. A *hybrid time domain* is a set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ such that, for each each $(T, J) \in E$, the set $E \cap ([0, T] \times \{0, \ldots, J\})$ is a compact hybrid time domain. A *hybrid arc* on $\mathbb{R}^n$ is a mapping $\phi : \text{dom}(\phi) \to \mathbb{R}^n$ such that $\text{dom}(\phi)$ is a hybrid time domain and, for each $i \in \mathbb{Z}_{\geq 0}$, $\phi(\cdot, i)$ is locally absolutely continuous on $\text{dom}(\phi)_i := \{t \in \mathbb{R}^n : (t, i) \in \text{dom}(\phi)\}$. The graph of a hybrid arc x is the set $\text{graph}(x) := \left\{(t, j, z) \in \mathbb{R}^{n+2} : (t, j) \in \text{dom}(x), z = x(t, j)\right\}$.

Given an initial condition $x \in C \cup D$, a solution $\mathbf{x}$ to the stochastic hybrid inclusion (2.1) is a mapping $\mathbf{x} : \Omega \to \left\{S : \mathbb{R}^2 \rightrightarrows \mathbb{R}^n : S \text{ osc, } \text{dom}(S) \neq \varnothing\right\}$ such that:

1. almost every sample path is a hybrid arc satisfying the constraints imposed by the data, i.e., for almost every $\omega \in \Omega$ and defining $\phi_\omega := \mathbf{x}(\omega)$,

   (a) $\phi_\omega(0, 0) = x$.

   (b) if $(t_1, j), (t_2, j) \in \text{dom}(\phi_\omega)$ with $t_1 < t_2$ then, for almost all $t \in [t_1, t_2]$, $\phi_\omega(t, j) \in C$ and $\frac{d}{dt}\phi_\omega(t, j) \in F(\phi_\omega(t, j))$.

   (c) if $(t, j), (t, j + 1) \in \text{dom}(\phi)$ then $\phi_\omega(t, j) \in D$ and $\phi_\omega(t, j + 1) \in G(\phi_\omega(t, j), \mathbf{v}_{j+1}(\omega))$.

2. the set-valued mapping

   $\omega \mapsto \text{graph}(\mathbf{x}(\omega)) \cap (\mathbb{R}_{\geq 0} \times \{0, \ldots, i\} \times \mathbb{R}^n)$

is $\mathcal{F}_i$-measurable for each $i \in \mathbb{Z}_{\geq 0}$ where $\{\mathcal{F}_i\}_{i=0}^{\infty}$ is the natural filtration associated with $\{\mathbf{v}_i\}_{i=1}^{\infty}$.

A solution is *almost surely complete* if almost every sample path has an unbounded time domain.

2.2.3 Stability concepts and sufficient conditions

We will use the classical concept of *uniform global asymptotic stability in probability* and a sufficient condition for this property given in [28].

We start with definitions culminating in the definition of uniform global asymptotic stability in probability.

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is *Lyapunov stable in probability* for (2.1) if, for each $\varepsilon > 0$ and $\rho > 0$, $\exists\, \delta > 0$ such that, if $\mathbf{x}$ is a solution with initial condition in $\mathcal{A} + \delta\mathbb{B}$ then

$$\mathbb{P}\Big(\text{graph}(\mathbf{x}) \subset \Big(\mathbb{R}^2 \times (\mathcal{A} + \varepsilon\mathbb{B})\Big)\Big) \geq 1 - \rho. \tag{2.2}$$

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is *Lagrange stable in probability* for (2.1) if, for each $\delta > 0$ and $\rho > 0$, $\exists\, \varepsilon > 0$ such that, if $\mathbf{x}$ is a solution with initial condition in $\mathcal{A} + \delta\mathbb{B}$ then (2.2) holds.

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is *globally stable in probability* for (2.1) if it is Lyapunov stable in probability and Lagrange stable in probability.

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is *uniformly globally attractive in probability* for (2.1) if, for each $\varepsilon > 0$, $\rho > 0$ and $R > 0$, $\exists\, \tau \geq 0$ such that, if $\mathbf{x}$ is a solution with initial condition in $\mathcal{A} + R\mathbb{B}$ with the set of hybrid times (t, j) defined as $\Gamma_{\geq \tau} := \{(t, j) \in \mathbb{R}^2 : t + j \geq \tau\}$ then

$\mathbb{P}\Big(\text{graph}(\mathbf{x}) \cap (\Gamma_{\geq \tau} \times \mathbb{R}^n) \subset \Big(\mathbb{R}^2 \times (\mathcal{A} + \varepsilon\mathbb{B})\Big)\Big) \geq 1 - \rho.$

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is *uniformly globally asymptotically stable in probability* for (2.1) if it is globally stable in probability and uniformly globally attractive in probability.

The result in [28, Theorem 8] provides Lyapunov function-based sufficient conditions for UGASp.

A continuous function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a *Lyapunov function* relative to the compact set $\mathcal{A} \subset \mathbb{R}^n$ for the system (C, F, D, G, μ) if $V(x) = 0$ if and only if $x \in \mathcal{A}$, V is positive definite with respect to $\mathcal{A}$ and radially unbounded, both relative to $C \cup D$, and satisfies

$$V(\phi(t)) \leq V(x) \qquad \forall \begin{cases} x \in C \\ t \in \operatorname{dom}(\phi) \\ \phi \in \mathcal{S}_C^F(x) \end{cases} \tag{2.3a}$$

$$\int_{\mathbb{R}^m} \max_{g \in G(x,v)} V(g)\mu(dv) \leq V(x) \qquad \forall x \in D, \tag{2.3b}$$

where $\mathcal{S}_C^F(x)$ denotes the solutions of $x \in C$, $\dot{x} \in F(x)$ starting at the initial condition x.

Theorem 1 *[28, Thm. 8] Let V be a Lyapunov function relative to the compact set $\mathcal{A} \subset \mathbb{R}^n$ for the system with the data (C, F, D, G, μ) from (2.1). Then $\mathcal{A}$ is uniformly globally asymptotically stable in probability if and only if there does not exist an almost surely complete solution that remains in a non-zero level set of the Lyapunov function almost surely.*

2.3 A stochastic hybrid optimization algorithm

In this section, we cast a stochastic hybrid optimization algorithm based on the framework of stochastic hybrid inclusions. We are especially interested in optimization problems that involve saddle points, local minima, and similar phenomena. Consequently, we will consider optimization problems defined on manifolds. In this section, the algorithm is not necessarily a distributed optimization algorithm. The latter is the topic of the next section.

2.3.1 Manifold structure and exploratory dynamics

We start with our assumptions about the manifold.

Assumption 2 *The closed set $\mathcal{M} \subset \mathbb{R}^n$, the countable set $\mathcal{Q}$, the sets $\left\{\widehat{C}_q, C_q\right\}_{q\in\mathcal{Q}}$, the integer $d \in \{1,\dots,n\}$, the matrix-valued functions $B_q : \overline{C_q} \cap \mathcal{M} \to \mathbb{R}^{n\times d}$ and the functions $h_q : \left(\overline{C_q} \cap \mathcal{M}\right) \times \mathbb{R}^d \to \mathcal{M}$, $q \in \mathcal{Q}$, and the positive real number r are such that:*

1. *$\forall\, q \in \mathcal{Q}$, $\widehat{C}_q \subset C_q \subset \mathbb{R}^n$, $\widehat{C}_q$ is closed,C_q is open;*
2. *$\cup_{q\in\mathcal{Q}} \widehat{C}_q \cap \mathcal{M} = \mathcal{M}$;*
3. *The collection of sets $\left\{C_q\right\}_{q\in\mathcal{Q}}$ is locally finite on $\mathcal{M}$, i.e., for each $z \in \mathcal{M}$ there exists an open set U_z containing z such that the cardinality of the set $\left\{q \in \mathcal{Q} : U_z \cap C_q \neq \emptyset\right\}$ is finite.*
4. *For each $q \in \mathcal{Q}$, the tangent space to $\mathcal{M}$ on $\overline{C_q} \cap \mathcal{M}$ is parametrized by the matrix-valued function B_q that is continuous and full column rank on its domain;*
5. *For each $q \in \mathcal{Q}$, for each $z \in \overline{C_q} \cap \mathcal{M}$ and each open set $U \subset \mathbb{R}^n$ contained in a ball of radius r around z there exists an open set $\mathcal{O} \subset r\mathbb{B} \subset \mathbb{R}^d$ such that $h_q(z, \mathcal{O}) \subseteq U \cap \mathcal{M}$.* ■

Continuous exploration

With Assumption 2 in place, we can create a hybrid dynamical control system with input $u \in \mathbb{R}^d$ that is able to fully explore the manifold $\mathcal{M}$ continuously. It has the form

$$z \in \overline{C_q} \cap \mathcal{M} \quad \dot{z} \in B_q(z)u \tag{2.4a}$$

$$z \in (\mathbb{R}^n \backslash C_q) \cap \mathcal{M} \quad q^+ \in \left\{p \in \mathcal{Q} : z \in \widehat{C}_p\right\} \tag{2.4b}$$

Let G_c denote the jump map in (2.4b), which is nonempty due to item 2 of Assumption 2. It is outer semicontinuous since the collection $\left\{\widehat{C}_q\right\}_{q\in\mathcal{Q}}$ is subordinate to $\left\{C_q\right\}_{q\in\mathcal{Q}}$, which is locally

finite, and because $\widehat{C}_q$ is closed for each $q \in Q$. Indeed, suppose that $p_i \in G_c(z_i)$ and that (p_i, z_i) is convergent to (p, z). By the local finiteness property, the sequence p_i eventually takes the value p, so that $p \in G_c(z_i)$, i.e., $z_i \in \widehat{C}_p$. Since $\widehat{C}_p$ is closed, it follows that $z \in \widehat{C}_p$, i.e., $p \in G_c(z)$. The jumps defined in the continuous exploration dynamics (2.4) do not change the value of the state z that evolves on the manifold $\mathcal{M}$. Rather, those jumps change the vectors that are used to parametrize the tangent space to $\mathcal{M}$. For some manifolds, this step may not be necessary, like for the unit circle. For other manifolds, like the 2-sphere, this step is important to avoid singularities, in light of the so-called "hairy ball theorem", which states that there is no non-vanishing continuous tangent vector field on even-dimensional n-spheres.

Next, we consider jumps of the variable z, which may also be a useful way in which to explore the manifold $\mathcal{M}$.

Exploration by jumps

Here we augment the dynamics in (2.4) with timed jumps that are governed by

$$(\tau, z) \in [0, T_{\max}] \times \left(\overline{C}_q \cap \mathcal{M}\right) \quad \dot{\tau} = 1 \tag{2.5a}$$

$$(\tau, z) \in [T_{\min}, T_{\max}] \times \left(\overline{C}_q \cap \mathcal{M}\right) \begin{cases} z^+ = \omega_a \\ q^+ = \omega_b \\ \tau^+ = 0 \end{cases} \tag{2.5b}$$

where $0 < T_{\min} \leq T_{\max} < \infty$, $\omega_a \in \mathcal{M}$, and $\omega_b \in Q$.

Combined exploration

The combination of (2.4)-(2.5) creates a composite hybrid control system with state $x := (z^T, q, \tau)^T$, controls (u, ω), where $\omega := (\omega_a^T, \omega_b^T)^T$, and flow set and flow map

$$C := \cup_{q \in Q} \left((\overline{C}_q \cap \mathcal{M}) \times \{q\} \right) \times [0, T_{\max}] \tag{2.6a}$$

$$H(x, u) := \begin{bmatrix} B_q(z)u \\ 0 \\ 1 \end{bmatrix}. \tag{2.6b}$$

The jump set is defined by

$$D_c := \cup_{q \in Q} \left((\mathbb{R}^n \backslash C_q) \cap \mathcal{M} \times \{q\} \right) \times [0, T_{\max}] \tag{2.7a}$$

$$D_d := \cup_{q \in Q} \left((\overline{C}_q \cap \mathcal{M}) \times \{q\} \right) \times [T_{\min}, T_{\max}] \tag{2.7b}$$

$$D := D_c \cup D_d. \tag{2.7c}$$

Finally, the jump map G is defined by

$$\widehat{G}_c(x,\omega) := \begin{bmatrix} \{z\} \\ G_c(z) \\ \{\tau\} \end{bmatrix} \quad \forall x \in D_c \tag{2.8a}$$

$$\widehat{G}_d(x,\omega) := \begin{bmatrix} \{w_a\} \\ \{w_b\} \\ \{0\} \end{bmatrix} \quad \forall x \in D_d \tag{2.8b}$$

$$G(x,\omega) := \begin{cases} \widehat{G}_c(x,\omega) & \forall x \in D_c \backslash D_d \\ \widehat{G}_d(x,\omega) & \forall x \in D_d \backslash D_d \\ \widehat{G}_c(x,\omega) \cup \widehat{G}_d(x,w) & \forall x \in D_c \cap D_d. \end{cases} \tag{2.8c}$$

2.3.2 Using exploratory inputs to solve an optimization problem

The optimization will take place on the manifold $\mathcal{M}$ and will be for a function f with the following properties:

Assumption 3 *The function $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, and the restriction of f to $\mathcal{M}$ is radially unbounded. Defining the set of global minimizers*

$$\mathcal{A} := \{z^* \in \mathcal{M} : f(z^*) \leq f(z), \quad \forall z \in \mathcal{M}\}, \tag{2.9}$$

for each $z_1 \in \mathcal{M} \backslash \mathcal{A}$ there exists $z_2 \in \mathcal{M} \cap (\{z_1\} + r\mathbb{B})$ such that $f(z_2) < f(z_1)$, where r comes from Assumption 2. ■

It is not difficult to relax the C^1 assumption to a locally Lipschitz regular assumption. We assume the former to make a clearer connection to the multi-agent case in the next section. Since f is continuous and radially unbounded, the set $\mathcal{A}$ defined in Assumption 3 is compact.

We define

$$\mathcal{Q}_0 := \left\{ q \in \mathcal{Q} : \mathcal{A} \cap \overline{C}_q \neq \emptyset \right\}. \tag{2.10}$$

Due to Assumption 2, $\mathcal{Q}_0$ is finite, and thus compact. We use the following stochastic control algorithm to drive x to the set $\mathcal{A} \times \mathcal{Q}_0 \times [0, T_{\max}]$:

$$u = -\Lambda B_q^T(z) \nabla f(z) \tag{2.11a}$$

$$\omega_a \in \operatorname{argmin}_{s \in \{z, h_q(z,v)\}} f(s) \tag{2.11b}$$

$$\omega_b \in \left\{ p \in \mathcal{Q} : \omega_a \in \widehat{C}_p \right\}. \tag{2.11c}$$

We make the following assumption about Λ and v.

Assumption 4 *The square matrix $\Lambda \in \mathbb{R}^{d \times d}$ is diagonal and positive definite, and v is a placeholder for an iid sequence of random variables that are uniformly distributed on the ball of radius r in $\mathbb{R}^d$.* ■

Theorem 2 *Suppose Assumptions 2-4 hold. For the hybrid control system (2.6)-(2.8), (2.11), the compact set $\mathcal{A} \times \mathcal{Q}_0 \times [0, T_{\max}]$, with $\mathcal{A}$ and $\mathcal{Q}_0$ defined in (2.9) and (2.10), is UGASp.*

Proof. We start by defining, for each $\gamma \in \mathbb{R}_{\geq 0}$,

$$\mathcal{Q}_\gamma := \left\{ q \in \mathcal{Q} : \zeta \in \overline{C}_q \cap \mathcal{M}, f(\zeta) - f(\mathcal{A}) \leq \gamma \right\}. \tag{2.12}$$

We note that, for each $\gamma \in \mathbb{R}_{\geq 0}$, $\mathcal{Q}_\gamma$ is compact due to item 3 of Assumption 2 and the assumption that f is radially unbounded with respect to $\mathcal{M}$. Also note that, conveniently, $\mathcal{Q}_0$ coincides with $\mathcal{Q}_0$ defined in (2.10). Now consider the Lyapunov function candidate

$$V(x) := f(z) - f(\mathcal{A}) + |q|_{\mathcal{Q}_{(f(z)-f(\mathcal{A}))}} + |\tau|^2_{[0,T_{\max}]}. \tag{2.13}$$

This function is continuous, positive definite with respect to to $\mathcal{A} \times Q_0 \times [0, T_{\max}]$ and radially unbounded. Moreover, except perhaps before a first jump at time $(0, 0)$, the state evolves in a set where

$$|q|_{Q_{(f(z)-f(\mathcal{A}))}} + |\tau|^2_{[0,T_{\max}]} = 0.$$

Indeed, τ never leaves the interval $[0, T_{\max}]$, and, perhaps after a first jump at time $(0, 0)$, the variables z and q are always related by $z \in \overline{C}_q$, which implies that $q \in Q_{(f(z)-f(\mathcal{A}))}$. Hence, the evolution of the value $V(x)$ is determined by the evolution of the value $f(z)$.

We note that, with λ_i, $i \in \{1, \ldots, d\}$ denoting the diagonal entries of Λ, which are positive by assumption,

$$\begin{aligned}\langle \nabla f(z), -B_q(z)\Lambda B_q^T(z)\nabla f(z)\rangle = \\ -\sum_{i=1}^{d} \lambda_i |(B_q^T(z)\nabla f(z))_i|^2 \leq 0.\end{aligned} \tag{2.14}$$

Therefore, condition (2.3a) is satisfied.

Next, we consider

$$\begin{aligned}\int_{\mathbb{R}^m} \max_{g \in \operatorname{argmin}_{s\in\{z, h_q(z,v)\}} f(s)} f(g)\mu(dv) = \\ \int_{\mathbb{R}^m} \min\left\{f(z), f(h_q(z, v))\right\}\mu(dv) \leq f(z).\end{aligned} \tag{2.15}$$

Therefore, condition (2.3b) is satisfied.

Finally, suppose $z \in \mathcal{M}\backslash\mathcal{A}$. Then, according to Assumption 3, there exists $z_2 \in \mathcal{M} \cap (\{z\} + r\mathbb{B})$ such that $f(z_2) < f(z)$. It follows that there exists an open set $U \subset (\{z\} + r\mathbb{B})$ and $\varepsilon > 0$ such that $f(\zeta) \leq f(z) - \varepsilon$ for all $\zeta \in U \cap \mathcal{M}$. Then, according to item 5 of Assumption 2,

there exists an open set $O \subset r\mathbb{B}$ such that $h_q(z, O) = U \cap \mathcal{M}$. In that case,

$$\int_{\mathbb{R}^m} \min\left\{f(z), f(h_q(z, v))\right\}\mu(dv) \leq$$
$$\mu(O)(f(z) - \varepsilon) + \mu(\mathbb{R}^m \backslash O) f(z) = f(z) - \mu(O)\varepsilon. \tag{2.16}$$

Since μ corresponds to a distribution uniformly distributed on a ball of radius r and $O \subset r\mathbb{B}$, it follows that $\mu(O) > 0$. Finally, since every solution jumps at least every $T_{\max}$ seconds, it follows that there is no almost surely complete solution that remains in a non-zero level set of the Lyapunov function almost surely. Hence, the claim of Theorem 2 follows from Theorem 1.

■

2.4 A partially distributed stochastic hybrid optimization algorithm

We now show how the results of the previous section can be extended to partially address distributed optimization problems in multi-agent systems (MAS) where the actions of the agents are constrained to evolve on manifolds $\mathcal{M} \subset \mathbb{R}^n$ satisfying Assumption 2. In particular, we consider a MAS with N agents, where each agent has a state $z_i \in \mathcal{M}$, and a cost function $J_i : \mathbb{R}^{nN} \to \mathbb{R}$, which also depends on the actions of the other agents. The overall state of the system is defined as $z := (z_1, z_2, \cdots, z_N)$, such that $z \in \mathcal{M}^N$, where

$$\mathcal{M}^N := \underbrace{\mathcal{M} \times \ldots \times \mathcal{M}}_{\text{N times}} \subset \mathbb{R}^{nN}. \tag{2.17}$$

The manifold $\mathcal{M}^N$ plays $\mathcal{M}$'s role from the previous section.

Lemma 1 *If the manifold $\mathcal{M} \subset \mathbb{R}^n$, the countable set $\mathcal{Q}$, the sets $\left\{\widehat{C}_q, C_q\right\}_{q\in\mathcal{Q}}$, the integer $d \in$*

$\{1, \ldots, n\}$, the matrix-valued functions $B_q : \overline{C_q} \cap \mathcal{M} \to \mathbb{R}^{n\times d}$ and the functions $h_q : \left(\overline{C_q} \cap \mathcal{M}\right) \times \mathbb{R}^d \to \mathcal{M}$, $q \in \mathcal{Q}$, and the positive real number r satisfy Assumption 2 then the manifold $\mathcal{M}^N \subset \mathbb{R}^{nM}$, the countable set

$$\mathcal{Q}^N := \underbrace{\mathcal{Q} \times \ldots \times \mathcal{Q}}_{N \text{ times}},$$

the sets

$$\widehat{C}_q = \left\{\widehat{C}_{q_1} \times \cdots \times \widehat{C}_{q_N}\right\}_{(q_1,\ldots,q_N)\in\mathcal{Q}^N} \tag{2.18}$$

$$C_{q_1} = \left\{C_{q_1} \times \cdots \times C_{q_N}\right\}_{(q_1,\ldots,q_N)\in\mathcal{Q}^N} \tag{2.19}$$

the integer Nd, the matrix-valued functions

$$B_q^N(z) := \operatorname{diag}\left(B_{q_1}(z_1), \cdots, B_{q_N}(z_N)\right) \qquad q \in \mathcal{Q}^N, \tag{2.20}$$

and the functions

$$h_q^N(z, v) := \begin{bmatrix} h_{q_1}(z_1, v_1) \\ \vdots \\ h_{q_N}(z_N, v_N) \end{bmatrix} \quad q \in \mathcal{Q}^N,$$

and the positive real number r satisfy Assumption 2, where the balls of radius r have the form $\underbrace{r\mathbb{B} \times \cdots \times r\mathbb{B}}_{N \text{ times}}$. ■

Proof. In order to prove lemma 1, we need to show that all Assumption 2 conditions are satisfied for the cartesian product of multiple copies of the manifold $\mathcal{M}$, i.e.,

$$\mathcal{M}^N := \underbrace{\mathcal{M} \times \ldots \times \mathcal{M}}_{\text{N times}} \subset \mathbb{R}^{nN}. \tag{2.21}$$

1. Since $\widehat{C}_q$ is a closed set, the cartesian product of multiple closed sets i.e. equation 2.18 is also a closed set. Moreover, since C_q is an open set, the cartesian product of multiple open sets i.e. equation 2.19 is also an open set. So the first condition of Assumption 2 for multiple copies of the manifold is satisfied.

2. In this case we need to prove

$$\cup_{q\in Q^N}\widehat{C}_{q_1}\times\cdots\times\widehat{C}_{q_N}\cap\mathcal{M}^N=\mathcal{M}^N; \tag{2.22}$$

and we already know that $\cup_{q\in Q}\widehat{C}_q\cap\mathcal{M}=\mathcal{M}$.

$$\begin{gathered}\cup_{q\in Q^N}\widehat{C}_{q_1}\times\cdots\times\widehat{C}_{q_N}\cap\mathcal{M}^N=\\ \cup_{q\in Q^N}(\widehat{C}_{q_1}\cap\mathcal{M})\times(\widehat{C}_{q_2}\cap\mathcal{M})\cdots\times(\widehat{C}_{q_N}\cap\mathcal{M})=\\ \underbrace{\mathcal{M}\times\mathcal{M}\times\cdots\times\mathcal{M}}_{\text{N times}}=\mathcal{M}^N\end{gathered}$$

3. We need to show that the set A, defined as followed for multiple copies of the manifold has a finite cardinality.

$$A:=\left\{q\in Q_N:(U_1\times\cdots\times U_N)\cap(C_{q_1}\times\cdots\times C_{q_N})\neq\varnothing\right\};$$

which can be also written as

$$A:=\left\{q\in Q_N:(U_1\cap C_{q_1})\times\cdots\times(U_N\cap C_{q_N})\neq\varnothing\right\};$$

For each element of $q=(q_1,...,q_N)\in\mathcal{A}$, there exists a corresponding set

$$q_i\in A_i:=\left\{q_i\in Q_i:(U_i\cap C_{q_i})\neq\varnothing\right\}$$

with a finite cardinality according to the third condition of the Assumption 2. Since the cartesian product of finite sets is also finite, the cardinality of
$A := \prod_{i=1}^{N} \left\{ q_i \in \mathcal{Q}_i : U_i \cap C_{q_i} \neq \emptyset \right\}$ is finite.

4. Fourth claim of the Assumption 2 asks for each $q \in \mathcal{Q}$, the tangent space to $\mathcal{M}$ on $\overline{C_q} \cap \mathcal{M}$ is parametrized by the matrix-valued function B_q that is continuous and full column rank on its domain; B_q for one manifold is defined as $B_q : \overline{C_q} \cap \mathcal{M} \to \mathbb{R}^{n \times d}$ which is continuous and full rank and what we want to prove for multiple manifolds is that

$$B_q^N(z) := \operatorname{diag}\Big(B_{q_1}(z_1), \cdots, B_{q_N}(z_N) \Big) \qquad q \in \mathcal{Q}^N$$

is also continuous and full rank on its domain. We know that $n \geq d$, which means each $B_q^N(z)$ according to the Assumption 2's definitions is column linearly independant. Stacking each $B_{q_i}^N(z)$ matrix diagonally will give us a continuous and column linearly independant matrix belonging to $\mathbb{R}^{nN \times dN}$.

5. Fifth claim of the assumption says for each $q \in \mathcal{Q}$, for each $z \in \overline{C_q} \cap \mathcal{M}$ and each open set $U \subset \mathbb{R}^n$ contained in a ball of radius r around z there exists an open set $\mathcal{O} \subset r\mathbb{B} \subset \mathbb{R}^d$ such that $h_q(z, \mathcal{O}) = U \cap \mathcal{M}$. .
Firstly, we write this claim for multiple manifolds in the following way. For each $q \in \mathcal{Q}^N$ meaning for each element of $(q_1, q_2, \cdots, q_N) \in \mathcal{Q}^N$ there will be

$$(z_1, z_2, \cdots, z_N) \in \prod_{i=1}^{N} \overline{C_{q_i}} \cap \mathcal{M}^i$$

and open sets

$(U_{z_1} \times U_{z_2} \times \cdots \times U_{z_N})$ containing $(z_1, z_2, \cdots, z_N)$

there exists an open set $\mathcal{O} \subset \underbrace{r\mathbb{B} \times \cdots \times r\mathbb{B}}_{\text{N times}}$ such that

$h_q((z_1, \cdots, z_N), \mathcal{O}) = (U_{z_1} \times \cdots \times U_{z_N}) \cap \mathcal{M}^N$.

Take an arbitrary q_i and its corresponding U_{z_i} containing z_i. Then we have

$\exists \mathcal{O}_i$ such that $h_{q_i}(z_i, \mathcal{O}_i) = U_{z_i} \cap \mathcal{M}^i$. The $\mathcal{O}$ is defined as $\mathcal{O} = \mathcal{O}_1 \times \mathcal{O}_2 \times \cdots \times \mathcal{O}_N$ and so we have the new $h_q^N(z, \mathcal{O})$ given as following for multiple copies of the manifold.

$$h_q^N(z, \mathcal{O}) = \prod_{i=1}^{N} h_{q_i}^N(z_i, \mathcal{O}_i) = \prod_{i=1}^{N} U_i \cap \mathcal{M} = U \cap \mathcal{M}^N.$$

This will be the new projection in $\mathbb{R}^{nN}$. ■

Each agent controls only its own action $z_i \in \mathcal{M}$ aiming to minimize its own cost function J_i. Since for each agent $i \in \{1, 2, \ldots, N\}$, the cost function J_i also depends on the actions of the other agents, the multi-agent optimization defines a non-cooperative game on the manifold $\mathcal{M}^N$. The structure of the cost functions J_i is given by the following assumption.

Assumption 5 *The non-cooperative game characterized by the continuously differentiable cost functions $J_i : \mathbb{R}^{nN} \to \mathbb{R}$ is a* weighted potential game*, i.e., there exists a continuously differentiable function $J : \mathbb{R}^{nN} \to \mathbb{R}$ and a vector $w \in \mathbb{R}^N$ such that the following holds:*

1. *For each $i \in \{1, 2, \ldots, N\}$ and each pair $z_a, z_b \in \mathcal{M}^N$ satisfying $z_a - z_b = e_i(z_{a,i} - z_{b,i})$ we have that*

$$J(z_a) - J(z_b) = w_i(J_i(z_a) - J_i(z_b)), \tag{2.23}$$

2. *The partial derivatives are related as below for all $i \in \{1, 2, \ldots, N\}$.*

$$\frac{\partial J(z)}{\partial z_i} = w_i \frac{\partial J_i(z)}{\partial z_i}, \tag{2.24}$$

■

In terms of the optimization problem of the previous section, J plays the role of f. Hence, we also impose the following assumption:

Assumption 6 *The potential function J of Assumption 5 satisfies Assumption 3.* ■

In the control algorithm (2.11) of the previous section, now used to minimize the potential function J, the continuous-time input u can be written as $u := [u_1^\top, \ldots, u_N^\top]^\top$ and, with the definition (2.20), satisfies

$$u_i = -\Lambda_i B_{q_i}^T(z_i) \left(\frac{\partial J(z)}{\partial z_i} \right)^T \tag{2.25}$$

where Λ_i is a diagonal, positive definite matrix. According to the second item of Assumption 5, it follows that

$$u_i = -\Lambda_i B_{q_i}^T(z_i) w_i \left(\frac{\partial J_i(z)}{\partial z_i} \right)^T \tag{2.26}$$

for some real number w_i. Hence, during flows, if the ith agent implements (2.26), which is distributed because it uses only the ith state z_i and information about the derivative of the ith cost function J_i, the effect will be to implement the centralized control (2.25).

Unfortunately, it is difficult to make the jumps of the optimization algorithm completely distributed while guaranteeing convergence to the set of minimizers of the potential function. Indeed, there may be points in the state space where it is possible to decrease the potential function by changing the values of multiple agents but where it is impossible to decrease any of the individual cost functions through changes in just the state of the corresponding individual.

We can add additional, distributed jumps that do not harm, and usually help, the algorithm. But, for now, we are unable to remove the centralized jumps completely.

The additional jumps that we include are governed by

$$\tau_i \in [0, T_{\max,i}] \quad \dot{\tau}_i = 1$$

$$\tau_i \in [T_{\min,i}, T_{\max,i}] \begin{cases} z_i^+ \in \underset{s \in \{z_i, h_{q_i}(z_i, v_i)\}}{\operatorname{argmin}} w_i J_i(s, z_{-i}) \\ q_i^+ \in \left\{ p \in \mathcal{Q} : z_i^+ \in \widehat{C}_p \right\} \\ \tau_i^+ = 0 \end{cases}$$

where z_{-i} refers to the components of z other than z_i.

These additional jumps do not increase the potential function since, by construction, $w_i J_i(z_i^+, z_{-i}) \leq w_i J_i(z_i, z_{-1})$ so that, using the first condition of Assumption 5,

$$J(z_i^+, z_{-i}) - J(z_i, z_{-i}) = w_i(J_i(z_i^+, z_{-i}) - J_i(z_i, z_{-i})) \leq 0.$$

In the next section, we illustrate the efficacy of the proposed algorithm in the context of a weighted potential game.

2.5 Numerical Example

In this section, we illustrate the previous algorithm in a numerical example in the context of a location game, similar to the one considered in [29]. We consider two players, where each player controls its own individual action, which is constrained to the unit circle S^1, i.e., $\mathcal{M} := S^1$. Each player aims to selfishly minimize its own cost function $J_i : \mathcal{M}^2 \to \mathbb{R}$, where $\mathcal{M}^2 := S^1 \times S^1$, and which also depends on the actions of the other player. In polar coordinates,

the cost functions of the players are given by [29]

$$J_1(\theta_1, \theta_2) = -\cos\theta_1 + \alpha_1 \cos(\theta_1 - \theta_2) \tag{2.27}$$

$$J_2(\theta_1, \theta_2) = -\cos\theta_2 + \alpha_2 \cos(\theta_2 - \theta_1), \tag{2.28}$$

where $\alpha_1, \alpha_2 \in \mathbb{R}$ satisfy $\alpha_1, \alpha_2 \neq 0$. We considered $\alpha_1 = \alpha_2 = 1$. The above individual cost functions are transformed to the cartesian coordinate system, obtaining

$$J_1(z_2, z_2) = -ez_1 + \alpha_1 z_1^T z_2 \tag{2.29}$$

$$J_2(z_1, z_2) = -ez_2 + \alpha_2 z_1^T z_2, \tag{2.30}$$

where e is defined as $e := \begin{bmatrix} 1 & 0 \end{bmatrix}$.

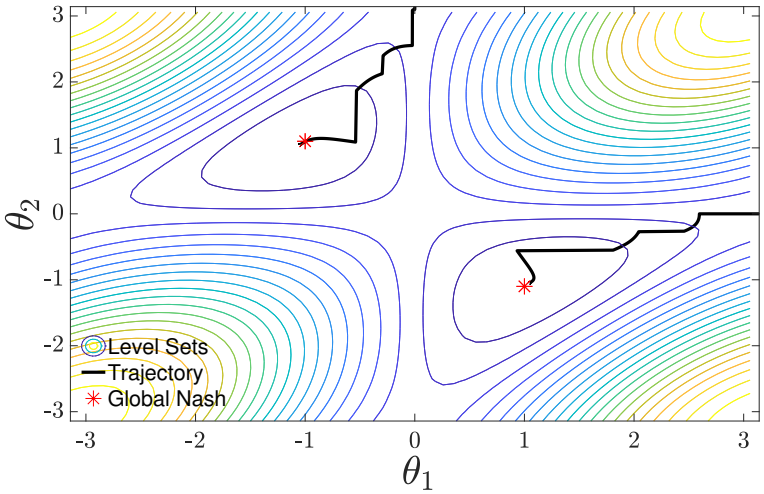

Figure 2.1: Evolution of the states θ from two different initial conditions, over the level sets of the potential function.

It can be shown that this location game is a potential game with a potential function

$$J(\theta_1, \theta_2) = -\frac{1}{\alpha_1}\cos(\theta_1) - \frac{1}{\alpha_2}\cos(\theta_2) + \cos(\theta_1 - \theta_2),$$

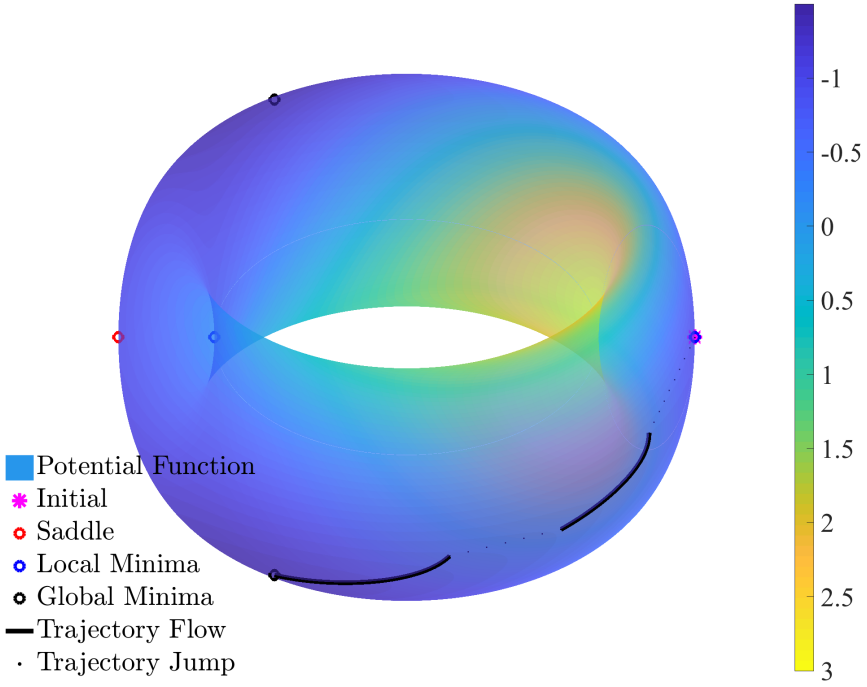

Figure 2.2: Evolution in the torus of the state z.

which in Cartesian coordinates is given by

$$J(z_1, z_2) = -\frac{1}{\alpha_1} e z_1 - \frac{1}{\alpha_2} e z_2 + z_1^T z_2.$$

The critical points of the potential function are solutions to

$$\frac{\partial J_1(\theta_1, \theta_2)}{\partial \theta_1} = \frac{\partial J_2(\theta_1, \theta_2)}{\partial \theta_2} = \frac{\partial J(\theta_1, \theta_2)}{\partial (\theta_1, \theta_2)} = 0.$$

Seven critical points are given as $\theta := [\theta_1, \theta_2]^\top$ with their corresponding angles. The critical points with angles $\theta_1 := [0, 0]^\top$, $\theta_2 := [\pi, 0]^\top$, $\theta_3 := [0, \pi]^\top$, $\theta_4 := [\frac{\pi}{3}, -\frac{\pi}{3}]^\top$ and $\theta_5 := [-\frac{\pi}{3}, \frac{\pi}{3}]^\top$ are all Nash equilibria [29] of the location game, however, not all these points are strict Nash equilibria. The critical point $\theta_1 := [0, 0]^\top$, with its corresponding value of $J = -1$, is a saddle point of the potential function. The points $\theta_2 := [\pi, 0]^\top$ and $\theta_3 := [0, \pi]^\top$ are both local

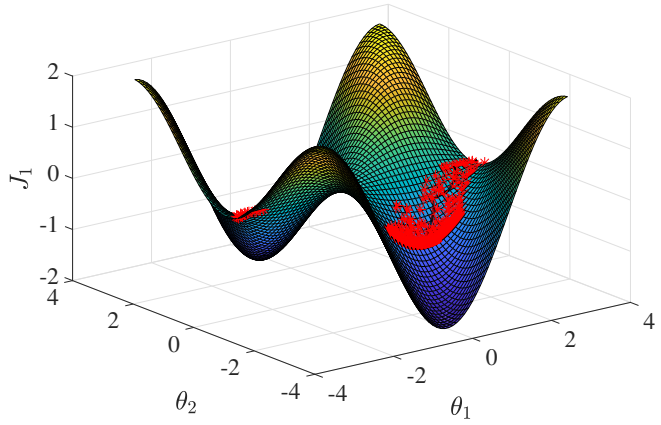

Figure 2.3: The cost function of player 1

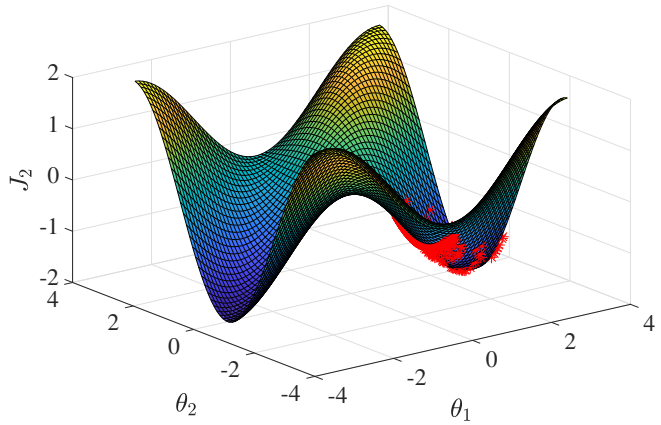

Figure 2.4: The cost function of player 2

minima, and only the two critical points $\theta_4 := [\frac{\pi}{3}, -\frac{\pi}{3}]^\top$ and $\theta_5 := [-\frac{\pi}{3}, \frac{\pi}{3}]^\top$ with both having their corresponding value of J as $J = -1.5$, are strict Nash equilibria of the game. These strict Nash equilibria are also global minimizers of the potential function. In polar coordinates, we denote by $\mathcal{A} := \{[\pi/3, -\pi/3]^\top\} \cup \{[-\pi/3, \pi/3]^\top\}$, the set of strict Nash equilibria of the game, i.e., the global minimizers of the potential function. In order to obtain convergence to $\mathcal{A}$ the stochastic hybrid algorithm from Section 2.4 is implemented. In this case, we have that $Q := \{1\}$, $\hat{C}_1 = C_1 = \mathbb{R}^2$, $d = 1$, $B_1(z) = [z_2, -z_1]^\top$, $h_1(z, v) = R(\varepsilon v)z$, where R is a rotation matrix with angle εv, ε is a tunable gain, and v is a random number that is uniformly selected

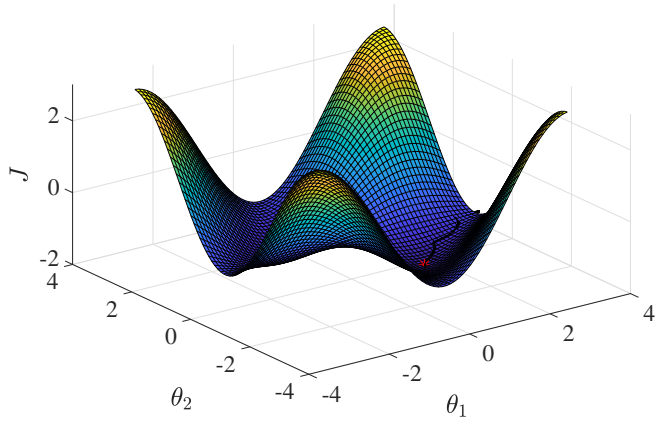

Figure 2.5: The potential function of the location game

from the set $[-\pi, \pi]$. Figure 2.1 shows two trajectories of the states $[\theta_1, \theta_2]$ converging to the set of strict Nash equilibria of the game, which correspond to the global minimizer of the potential function. The initial conditions are selected as the two local minima of the potential function.The parameter ε is selected as 0.6. The stochastic jumps are used to escape these minima in a finite time, after which the flows and jumps dominate the convergence of the states towards the set of global minimizers of the potential function.

Figure 2.2 shows the evolution of the state θ along the torus. Figure 2.5 shows a 3D plot of the potential function, and Figures 2.3 and 2.4 show 3D plots for the cost functions J_1 and J_2. Figure 2.6 shows the evolution in time of the potential function. It can be observed in the inset that at the beginning of the simulation, the states spend an initial amount of time in the local minimum of the potential function, until a stochastic jump moves the state out of the minimum.

2.6 Concluding Remarks

By using the framework of stochastic hybrid inclusions, we presented a class of dynamical systems that solve a class of non-convex optimization problems defined on manifolds. The global asymptotic stability in probability for the set of global minimizers of the cost function

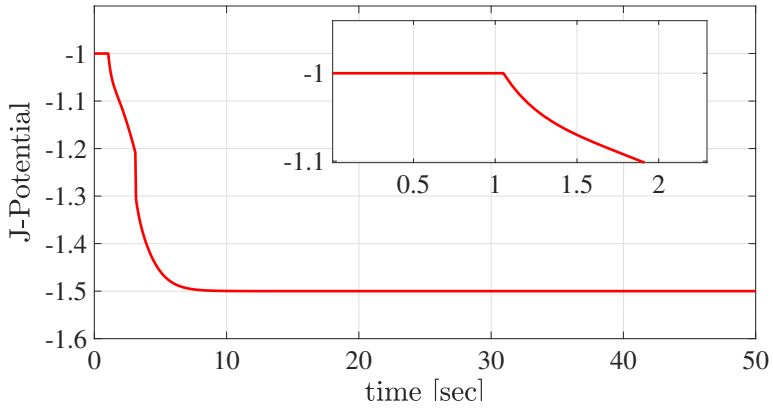

Figure 2.6: Evolution in time of the potential function.

is established, and the results extended to the multi-agent setting. A numerical example in a non-cooperative game defined on the torus was presented. Future research directions include the study of the computational complexities and the implementation of control architectures without centralized communication to solve distributed optimization problems in large-scale multi-agent systems.

Chapter 3

Global optimization on the sphere: A stochastic hybrid systems approach

In this chapter, we capture the stochastic hybrid optimization algorithm for minimizing any C^1 (continuously differentiable) function defined on the unit 2-sphere in the setting of stochastic hybrid inclusions (SHIs) presented in [9]. The observation that stochastic hybrid optimization algorithms fit into the SHI framework was already made in the previous chapter for general smooth manifolds (Also see [30]). This chapter, in particular, studies the special case of optimizing on the sphere as part of the general framework, which has been discussed in Chapter 2. Nevertheless, optimizing on the sphere is especially interesting since it may involve functions with saddle points, local minima, and similar phenomena among its critical points. Moreover, it is a compact manifold requiring more than one chart, in contrast to unit circle or torus, which is the setting of the example considered in Section 2.5 of Chapter 2.

In general, optimization problems defined on manifolds have been a developing paradigm of nonlinear optimization. Such problems emerge in engineering applications including phase synchronization of cyclic processes in [31] and tracking control of attitude dynamics of rigid bodies on SO(3) manifolds in [32]. Using our algorithm, optimizing any C^1 function on ellip-

soids or toroids as some specific cases of our general framework from Chapter 2 is straightforward. Thus, our focus specifically in this chapter is on the optimization of functions defined on the unit sphere, as it is a simple compact manifold that requires more than one but *not* more than two charts to cover it, which leads to a simplified case relative to the general setting. The general setting of such algorithms in the setting of SHI involves a combination of continuous and stochastic discrete-time updates of the optimization variables. In [33] for example, a hybrid optimization technique is presented, where stochastic optimization techniques like genetic algorithm (GA) and gradient-based local search algorithms are combined to overcome the local optima. The hybrid combination including gradient-based flows and randomness helps us to escape from local minima and saddle points of the function. It is a known fact that convergence to a set of global minimizers of a function defined on smooth manifolds using a gradient-descent method occurs only when the function possesses only one critical point, at which the derivative vanishes. Simple gradient-based methods for optimization in a non-convex setting like a manifold have been studied by [34] and [35], for example. Also, Morse theory used in the works of [36], [37] and [38], indicates that a continuous gradient-descent algorithm does not achieve global convergence to a single critical point, while a function defined on a compact manifold will have at least two critical points. Robustness guarantees with respect to arbitrary, small measurement noise or calculation errors are also reasons why discontinuous control and optimization are not good options. Encouraged by the mentioned research line, stochastic hybrid optimization algorithms as presented in the previous Chapter 2 are especially relevant for overcoming the existence of local minima and saddle points of functions defined on compact manifolds. An application of this algorithm would be minimization of any C^1 function on the sphere as mentioned. For this application on the sphere, we can observe an additional hybrid dynamics for switching between coordinate charts which overcomes the topological obstructions of the sphere and enables probing all over the sphere and converging to the set of global minimizers of the function. Using spherical coordinates or stereographic projection methods,

we generate the coordinate functions. Lastly, using Lyapunov functions and the invariance principle for SHIs, we establish global asymptotic stability in probability for the set of global minimizers of any C^1 function defined on the sphere. Finally, we demonstrate the results by providing a numerical example of the unit 2-sphere.

3.1 Problem Statement

The general framework of stochastic hybrid dynamical systems was introduced in (2.1) of Chapter 2, where the evolution of the state $x \in \mathbb{R}^n$ is a combination of continuous and discrete time dynamics. In this chapter, the problem that we would like to solve is creating a hybrid dynamical system of the form (2.1) to dynamically solve the following particular optimization problem:

$$\min_{z \in \mathcal{M}} f(z) \tag{3.1}$$

where the function f is any arbitrary C^1 function defined on a manifold. In other words, the state x of the system (2.1) includes a component that evolves on the manifold and the idea is to have this component converging to the optimal value of the function defined on the manifold as the time goes to infinity. In particular, we are considering the function f to be defined on a *unit sphere* and we create the stochastic hybrid optimization algorithm for this case. What we mean by (unit) sphere is the set

$$\mathbb{S}^2 := \left\{ z \in \mathbb{R}^3 : \sum_{i=1}^{3} z_i^2 = 1 \right\} \subset \mathbb{R}^3. \tag{3.2}$$

For creating the dynamics of the state evolving on the unit sphere, we need to create coordinate charts. In the next section, we discuss the reasons and methods of creating these coordinate charts.

3.2 Preliminaries: Coordinate Charts on the Sphere

A global stochastic hybrid optimization algorithm on the sphere typically contains two hybrid features. In this section, we discuss one of these hybrid features: jumps between the coordinate charts of the manifold to overcome certain topological obstructions. In the case of the sphere, such obstructions arise from the so-called "hairy ball theorem" which states that there is no non-vanishing continuous tangent vector field on even-dimensional n-spheres. The implication is that the optimization algorithm requires multiple local charts that cover the sphere as well as nonsingular vector fields on those charts and coordinate functions covering each chart. The vector fields can be used to flow in any arbitrary direction on the coordinate chart, while the coordinate functions are used for jumping to any arbitrary point on the corresponding chart. Consequently, we create two coordinate charts and allow hysteresis switching between them. One method for constructing charts involves using spherical coordinates; another method involves using stereographic projection. Both methods are discussed below.

3.2.1 Spherical Coordinates

In this section, we use the standard trigonometric parametrization of $\mathbb{S}^2$ in terms of spherical coordinates in order to create the coordinate functions. We can consider first the following parametrization:

$$\Psi_1(\theta, \phi) := (\cos(\theta)\cos(\phi), \cos(\theta)\sin(\phi), \sin(\theta))^\top. \tag{3.3}$$

By interchanging the role of z_2 and z_3 axes, we get a second parametrization as follows:

$$\Psi_2(\theta, \phi) := (\cos(\theta)\cos(\phi), \sin(\theta), \cos(\theta)\sin(\phi))^\top. \tag{3.4}$$

For each $z \in \mathbb{S}^2$ and each $q \in \{1, 2\}$, there exist $(\theta, \phi) \in (-\frac{\pi}{2}, \frac{\pi}{2}) \times (-\pi, \pi)$ such that

$$z = \Psi_q(\theta, \phi). \tag{3.5}$$

The mapping Ψ_q, with angles restricted to their corresponding sets, is invertible. For q fixed, we can derive the rate of change of z using the chain rule:

$$\frac{dz}{dt} = \frac{d\Psi_q(\theta, \phi)}{d\theta}\dot{\theta} + \frac{d\Psi_q(\theta, \phi)}{d\phi}\dot{\phi} \tag{3.6a}$$

$$= \begin{bmatrix} \frac{d\Psi_q(\theta,\phi)}{d\theta} & \frac{d\Psi_q(\theta,\phi)}{d\phi} \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\phi} \end{bmatrix}. \tag{3.6b}$$

Changing the trigonometric parameterizations into Cartesian form, for $q = 1$ we have:

$$\frac{d\Psi_1(z)}{d\theta} = \begin{bmatrix} -\sin(\theta)\cos(\phi) \\ -\sin(\theta)\sin(\phi) \\ \cos(\theta) \end{bmatrix} = \begin{bmatrix} -\frac{z_3 z_1}{\sqrt{1-z_3^2}} \\ -\frac{z_3 z_2}{\sqrt{1-z_3^2}} \\ \sqrt{1-z_3^2} \end{bmatrix} \tag{3.7a}$$

$$\frac{d\Psi_1(z)}{d\phi} = \begin{bmatrix} -\cos(\theta)\sin(\phi) \\ \cos(\theta)\cos(\phi) \\ 0 \end{bmatrix} = \begin{bmatrix} -z_2 \\ z_1 \\ 0 \end{bmatrix}. \tag{3.7b}$$

The dynamics using the mapping Ψ_1 can therefore be written as:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{z_3 z_1}{\sqrt{1-z_3^2}} & -z_2 \\ -\frac{z_3 z_2}{\sqrt{1-z_3^2}} & z_1 \\ \sqrt{1-z_3^2} & 0 \end{bmatrix}}_{B_1(z)} u, \tag{3.8}$$

where $u := \begin{bmatrix} \dot{\theta} & \dot{\phi} \end{bmatrix}^\top$ is the input and $B_1(z)$ is the matrix-valued vector field. Following the exact same procedure for $q = 2$ gives us:

$$\frac{d\Psi_2(z)}{d\theta} = \begin{bmatrix} -\sin(\theta)\cos(\phi) \\ \cos(\theta) \\ -\sin(\theta)\sin(\phi) \end{bmatrix} = \begin{bmatrix} -\frac{z_2 z_1}{\sqrt{1-z_2^2}} \\ \sqrt{1-z_2^2} \\ -\frac{z_3 z_2}{\sqrt{1-z_2^2}} \end{bmatrix} \tag{3.9a}$$

$$\frac{d\Psi_2(z)}{d\phi} = \begin{bmatrix} -\cos(\theta)\sin(\phi) \\ 0 \\ \cos(\theta)\cos(\phi) \end{bmatrix} = \begin{bmatrix} -z_3 \\ 0 \\ z_1 \end{bmatrix}. \tag{3.9b}$$

Accordingly, the dynamics are as below:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{z_2 z_1}{\sqrt{1-z_2^2}} & -z_3 \\ \sqrt{1-z_2^2} & 0 \\ -\frac{z_3 z_2}{\sqrt{1-z_2^2}} & z_1 \end{bmatrix}}_{B_2(z)} u. \tag{3.10}$$

Each of these parametrizations covers the whole sphere except the singular points at $(0, 0, \pm 1)$ and $(0, \pm 1, 0)$. In order to create the charts, for $q \in \{1, 2\}$ and $\delta > 0$ we define

$$\mathcal{X}_q := \left(-\frac{\pi}{2} + \delta, \frac{\pi}{2} - \delta\right) \times (-\pi, \pi).$$

Each set $\mathcal{X}_q$ assures the full coverage of each chart excluding a tube around a semi-circle on an orthodome of the sphere and the variable δ is relevant to this tube around the semi-circle. The coordinate charts can be defined as

$$C_1 := \Psi_1(\mathcal{X}_1) \qquad C_2 := \Psi_2(\mathcal{X}_2). \tag{3.11}$$

Note that $C_1 \cup C_2 = \mathbb{S}^2$. We also define coordinate functions h_q, which are the mappings from a set that restricts angles into the charts and they cover the chart. These coordinate functions can be written as

$$h_1(v_1) := \Psi_1(v_1) \qquad h_2(v_2) := \Psi_2(v_2) \tag{3.12}$$

where $v_1 \in \mathcal{X}_1$ and $v_2 \in \mathcal{X}_2$.

3.2.2 Stereographic Projection

An alternative to the standard spherical coordinates can be achieved by using the so-called stereographic projection. In this approach, we define a projection from all points on the sphere except the poles onto a plane. Using basic linear algebra, we write the equations of each projection onto the plane from north pole $N := (0, 0, 1)^\top$ and south pole $S := (0, 0, -1)^\top$. The mapping that corresponds to the projection of points on the sphere (except north pole) by projecting from the north pole $\Phi_1 : \mathbb{S}^2 \setminus N \to \mathbb{R}^2$ can be written as:

$$\Phi_1\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}\right) := \begin{bmatrix} \frac{z_1}{1-z_3} \\ \frac{z_2}{1-z_3} \end{bmatrix}. \tag{3.13}$$

The projection mapping from south pole to a plane $\Phi_2 : \mathbb{S}^2 \setminus N \to \mathbb{R}^2$ is given as:

$$\Phi_2\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}\right) := \begin{bmatrix} \frac{z_1}{1+z_3} \\ \frac{z_2}{1+z_3} \end{bmatrix}. \tag{3.14}$$

We use $X \in \mathbb{R}^2$ for the planar coordinate, while we have $z \in \mathbb{S}^2$. We can write $X = \Phi_q(z)$, for

any $q \in \{1, 2\}$. Note that Φ_q is a one-to-one projection and as a result, this mapping is invertible. Since we are interested in the rate of change of z on the sphere, we write the derivative of the above equation

$$\dot{X}_1 = u = \frac{\partial \Phi_1(z)}{\partial z} \dot{z}. \tag{3.15}$$

We can further expand (3.15) to:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{1-z_3} & 0 & \frac{z_1}{(1-z_3)^2} \\ 0 & \frac{1}{1-z_3} & \frac{z_2}{(1-z_3)^2} \end{bmatrix} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix}. \tag{3.16}$$

Since we have the constraint of staying on the sphere, we observe that

$$z^T \dot{z} = 0. \tag{3.17}$$

Using (3.16) and (3.17), we solve for $\dot{z}$ to get:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \underbrace{\begin{bmatrix} -(z_1^2 + z_3 - 1) & z_1 z_2 \\ z_1 z_2 & -(z_2^2 + z_3 - 1) \\ z_1(1 - z_3) & z_2(1 - z_3) \end{bmatrix}}_{B_1(z)} u. \tag{3.18}$$

Similarly, using the equations for projection from south pole we can write:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 + z_3 - z_1^2 & -z_1 z_2 \\ -z_1 z_2 & 1 + z_3 - z_2^2 \\ -z_1(1 + z_3) & -z_2(1 + z_3) \end{bmatrix}}_{B_2(z)} u. \tag{3.19}$$

Using stereographic projection, we cover the sphere with two subsets:

$$\mathcal{X}_q := \left\{(z_1, z_2) \in \mathbb{R}^2 : z^T z < r\right\},$$

where each subset covers the entire sphere except for a neighborhood of the north, respectively south, pole. The radius $r > 1$ is relevant to the size of the neighborhood around each singular point. The charts can be written as:

$$C_1 := \Phi_1^{-1}(\mathcal{X}_1) \qquad\qquad C_2 := \Phi_2^{-1}(\mathcal{X}_2). \tag{3.20}$$

We have $C_1 \cup C_2 = \mathbb{S}^2$. Similar to the previous method, by using the new mappings we can create the coordinate functions h_q:

$$h_1(v_1) := \Phi_1^{-1}(v_1) \qquad h_2(v_2) := \Phi_2^{-1}(v_2) \tag{3.21}$$

where $v_1 \in \mathcal{X}_1$ and $v_2 \in \mathcal{X}_2$.

3.3 Stochastic Hybrid Optimization Algorithm on the Unit Sphere

3.3.1 Development

In this section, we detail our dynamic stochastic hybrid optimization algorithm for an arbitrary C^1 function on the unit sphere. Using the data generated either by spherical coordinates as in Section 3.1 or by stereographic projection as in Section 3.2, we specify the optimization algorithm as follows:

1. We take the two relatively open charts from the preliminary section

$$C_i \subset \mathbb{S}^2, \quad i \in \{1, 2\} \tag{3.22}$$

where i is the number of coordinate charts.

Recall that we have the below property:

$$C_1 \cup C_2 = \mathbb{S}^2. \tag{3.23}$$

The tangent space to each coordinate chart is parametrized by a continuous and full rank matrix-valued function $B_i : \overline{C_i} \to \mathbb{R}^{3\times 2}$.

2. In this step, we build a dynamical control system with input $u \in \mathbb{R}^2$. We consider $z \in \mathbb{S}^2 \subset \mathbb{R}^3$ and two previously built coordinate charts, with each having a number assigned to them indicated by $q \in \{1, 2\}$. The dynamical control system evolves fully on the sphere using the continuous parameterizations B_q of the two charts as follows:

$$z \in \mathbb{S}^2, \qquad \dot{z} = B_q(z)u, \tag{3.24}$$

The input u is related to the function f. The function f is any C^1 function, that we would like to minimize and is defined on the sphere. We choose

$$u := -\Lambda B_q^T(z)\nabla f(z), \tag{3.25}$$

where Λ is a diagonal positive definite matrix.

3. Next, we build hysteresis switching dynamics between the coordinate charts. We define:

$$\begin{aligned} C &:= \left(\overline{C}_1 \times \{1\}\right) \cup \left(\overline{C}_2 \times \{2\}\right) \\ \mathcal{D} &:= \left(\left(\mathbb{S}^2 \setminus C_1\right) \times \{1\}\right) \cup \left(\left(\mathbb{S}^2 \setminus C_2\right) \times \{2\}\right). \end{aligned}$$

The flows on a chart and the jumps between charts are governed by:

$$(z, q) \in C \qquad \begin{cases} \dot{z} = B_q(z)u \\ \dot{q} = 0 \end{cases} \tag{3.26}$$

$$(z, q) \in \mathcal{D} \qquad \begin{cases} z^+ = z \\ q^+ = 3 - q. \end{cases} \tag{3.27}$$

4. As the last step, we inject stochastic jumps into the state and randomly probe the sphere in order to fulfill the task of optimization and escape local minima and saddle points. These stochastic jumps are triggered by a timer $\tau \in [0, T] \subset \mathbb{R}$, which controls the time expiration and resets to zero after each random jump. Here we have the variable T as a constant; however, it is straightforward to have T dependent on the size of the gradient in (3.25).

$$(z, q, \tau) \in C \times [0, T] \qquad \begin{cases} \dot{z} = B_q(z)u \\ \dot{q} = 0 \\ \dot{\tau} = 1 \end{cases} \tag{3.28}$$

$$(z, q, \tau) \in C \times \{T\} \qquad \begin{cases} z^+ = \omega_a \\ q^+ = \omega_b \\ \tau^+ = 0 \end{cases} \tag{3.29}$$

$$(z, q, \tau) \in \mathcal{D} \times [0, T] \qquad \begin{cases} z^+ = z \\ q^+ = 3 - q \\ \tau^+ = \tau. \end{cases} \tag{3.30}$$

In equation (3.29), the variable ω_a is given by:

$$\omega_a \in \text{argmin}_{s \in \{z, h_1(v_1), h_2(v_2)\}} f(s), \tag{3.31}$$

where v_i is the placeholder for an i.i.d. sequence of random variables. For simplicity, we take the distribution of each v_i to be uniform on $\mathcal{X}_i$. Also in equation (3.29), the variable ω_b is written as

$$\omega_b \in \left\{ q \in \{1, 2\} : \omega_a \in \overline{C}_q \right\}. \tag{3.32}$$

3.3.2 Overall Algorithm

Finally considering the state as $x := (z^\top, q, \tau)^\top$, controls as (u, ω), with $\omega := (\omega_a^\top, \omega_b^\top)^\top$, we can write the overall hybrid control algorithm:

$$\text{flow set} \qquad C := C \times [0, T] \tag{3.33a}$$

$$\text{flow map} \qquad F(x) := \begin{bmatrix} -B_q(z)\Lambda B_q^T(z)\nabla f(z) \\ 0 \\ 1 \end{bmatrix}. \tag{3.33b}$$

The jump set is defined as:

$$D_c := \mathcal{D} \times [0, T] \tag{3.34a}$$

$$D_d := \mathcal{D} \times \{T\} \tag{3.34b}$$

$$D := D_c \cup D_d. \tag{3.34c}$$

Finally, the jump map G is defined as:

$$\widehat{G}_c(x) := \begin{bmatrix} z \\ 3 - q \\ \tau \end{bmatrix} \quad \forall x \in D_c \tag{3.35a}$$

$$\widehat{G}_d(x, v) := \begin{bmatrix} w_a(z, v) \\ w_b(z, v) \\ \{0\} \end{bmatrix} \quad \forall x \in D_d \tag{3.35b}$$

$$G(x, v) := \begin{cases} \widehat{G}_c(x) & \forall x \in D_c \backslash D_d \\ \widehat{G}_d(x, v) & \forall x \in D_d \backslash D_c \\ \widehat{G}_c(x) \cup \widehat{G}_d(x, v) & \forall x \in D_c \cap D_d. \end{cases} \tag{3.35c}$$

It can be verified that the data (C, F, D, G) satisfies the stochastic hybrid basic conditions of

[28].

3.4 Algorithm Stability Result

Relying on [28, Theorem 8], we certify the behavior of the above minimization algorithm by establishing uniform global asymptotic stability in probability (UGASp) of the set of the function's global minimizers.

Given a C^1 function $f : \mathbb{S}^2 \to \mathbb{R}$, we define its set of global minimizers as

$$\mathcal{A} := \left\{ z \in \mathbb{S}^2 : f(z) = \min_{s \in \mathbb{S}^2} f(s) \right\}. \tag{3.36}$$

Theorem 3 *Given any C^1 function $f : \mathbb{S}^2 \to \mathbb{R}$, the proposed minimization algorithm in Section 3.3 renders the set of global minimizers of f globally asymptotically stable in probability.*

Proof. Let the function $f : \mathbb{S}^2 \to \mathbb{R}$ be C^1. Define the Lyapunov function candidate $V : \mathbb{S}^2 \times \{1, 2\} \times [0, T] \to \mathbb{R} \geq 0$ as

$$V(x) := V(z^T, q, \tau) = f(z) - \min_{s \in \mathbb{S}^2} f(s). \tag{3.37}$$

This function is positive definite with respect to $\mathcal{A}$ and its sublevel sets are compact since it is C^1 and its domain is compact. We note that

$$\langle \nabla V(x), F(x) \rangle = -\sum_{i=1}^{2} \lambda_i |(B_q^T(z) \nabla f(z))_i|^2 \leq 0 \tag{3.38}$$

where $\lambda_i > 0, i \in \{1, 2\}$, denote the diagonal entries of the positive definite matrix Λ. According to (3.38), the Lyapunov function candidate does not increase during flows.

Next, we consider the effect of jumps on the Lyapunov function candidate V. First, we observe that, due to the definition of G, for all $(v_1, v_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ from (3.30) and all $g \in G(x, v)$

from (3.29), we have $V(g) \leq V(x)$. It follows that

$$\int_{\mathcal{X}_1 \times \mathcal{X}_2} \max_{g \in G(x,v)} V(g) \mu(dv) \leq V(x) \tag{3.39}$$

regardless of the distribution μ. It follows that V is a Lypuanov function for $\mathcal{A}$ in the sense of [28, Theorem 8].

It remains to establish that there does not exist an almost surely complete solution that remains in a non-zero level set of the Lyapunov function almost surely. First, we note that there is no sample path with a purely discrete-time domain. This fact follows from the observation that, after at most two jumps, the sample path is no longer in the jump set. This is because if z belongs to the relative boundary of C_i then it belongs to the relative interior of C_j, $j \neq i$, by virtue of the fact that C_1 and C_2 are relatively open and cover $\mathbb{S}^2$. Consequently, each sample path is uniformly non-Zeno in the sense of [27, Proposition 6.35]. Hence each complete sample path has a time domain that is unbounded in the ordinary time direction. Since jumps due to $\widehat{G}_d$ happen every T units of ordinary time, it is enough to show that those jumps result in a decrease in the expected value of V from points outside the set of global minimizers. Let $z \in \mathbb{S}^2 \backslash \mathcal{A}$. Let $z^* \in \mathcal{A}$. Let $i \in \{1, 2\}$ and $v_i^* \in \mathcal{X}_i$ be such that $z^* = h_i(v_i^*) \in C_i$. In addition let $\mathbb{B} := \{x \in \mathbb{R}^n : |x| \leq 1\}$. By the continuity of h_i, there exists $\varepsilon > 0$ such that $\left\{v_i^*\right\} + \varepsilon\mathbb{B} \subset \mathcal{X}_i$ and $f(h_i(v_i)) \leq f(z) - \varepsilon$ for all $v_i \in \left\{v_i^*\right\} + \varepsilon\mathbb{B}$. Without loss of generality, assume that $i = 1$. Then

$$\int_{\mathcal{X}_1 \times \mathcal{X}_2} \max_{g \in \widehat{G}_d(x,v)} V(g) \mu(dv) \leq V(x) - \mu\left(\left(\{v_i^*\} + \varepsilon\mathbb{B}\right) \times \mathcal{X}_2\right) \varepsilon. \tag{3.40}$$

Since $\mu\left(\left(\left\{v_i^*\right\} + \varepsilon\mathbb{B}\right) \times \mathcal{X}_2\right) > 0$, by virtue of the distribution being uniform over $\mathcal{X}_1 \times \mathcal{X}_2$ (though any distribution for which each open set has positive measure would be sufficient), it follows that there does not exist an almost surely complete solution that remains in a non-zero level set

of the Lyapunov function almost surely. ■

3.5 Numerical Example

We illustrate the proposed algorithm in a numerical optimization example for a given function defined on a unit sphere with coordinate charts created by spherical coordinates. The proposed function to minimize is given by:

$$f(z) := z_2 z_3 + z_1^3. \tag{3.41}$$

The reason for this particular choice of function is the existence of both saddle points and local

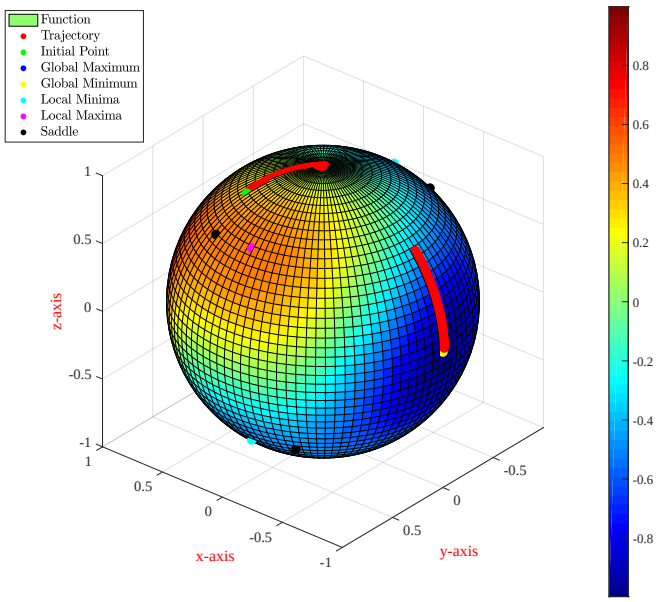

Figure 3.1: Trajectory on the Sphere

minima among its critical points while defined on the unit sphere. For the purpose of validating this example, the critical points of the given function are calculated using Lagrange multipliers

and are listed in Table 3.1.

Critical points	(x_1, x_2, x_3)	Value	Status
1	$(1, 0, 0)$	$+1$	Global Max
2	$(-1, 0, 0)$	-1	Global Min
3	$(0, 1/\sqrt{2}, 1/\sqrt{2})$	$1/2$	Local Max
4	$(0, -1/\sqrt{2}, -1/\sqrt{2})$	$1/2$	Local Max
5	$(1/3, 2/3, 2/3)$	$13/27$	Saddle
6	$(1/3, -2/3, -2/3)$	$13/27$	Saddle
7	$(0, 1/\sqrt{2}, -1/\sqrt{2})$	$-1/2$	Local Min
8	$(0, -1/\sqrt{2}, 1/\sqrt{2})$	$-1/2$	Local Min
9	$(-1/3, 2/3, -2/3)$	$-13/27$	Saddle
10	$(-1/3, -2/3, 2/3)$	$-13/27$	Saddle

Table 3.1: Critical Points

We create the dynamics $\dot{z} = B_q(z)u$ according to the second step of the optimization algorithm. We require the function's gradient to create the input u from (3.25):

$$\nabla f(z) = (3z_1^2, z_3, z_2)^\top. \tag{3.42}$$

The parameter δ is selected to be 0.01. Figure 3.1 illustrates a sample path starting from an initial point close to both a saddle point and a local maximum converging to the global minimizer of the function. The stochastic jumps are able to escape local minima and saddle points. Figure 3.2 and Figure 3.3 display the two coordinate charts of the sphere. The sample path, starting from an initial point located on the initial chart in Figure 3.2, starts decreasing the value of the function until it ends up converging to the global minimizer of f in the final chart in Figure 3.3. We observe the hysteresis switches between charts, which facilitates the convergence to a global minimizer of the function.

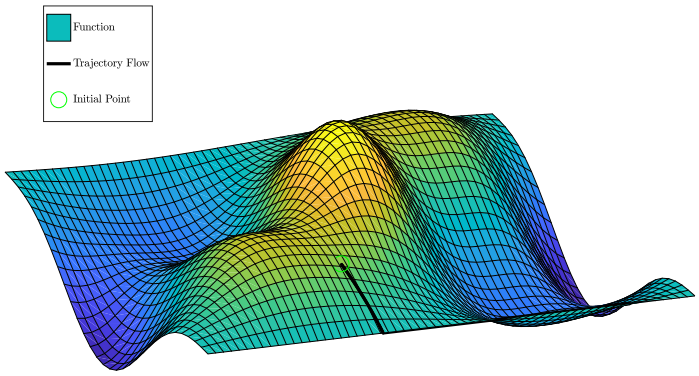

Figure 3.2: Initial Chart

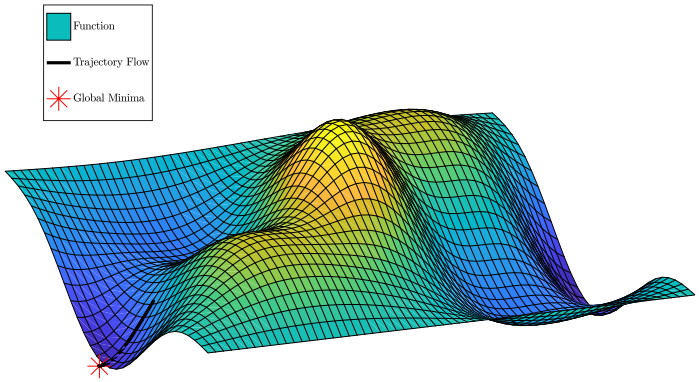

Figure 3.3: Final Chart

3.6 Concluding Remarks

We have captured a stochastic hybrid optimization algorithm in the framework of stochastic hybrid inclusions to minimize a function defined on the sphere. The idea of casting stochastic

hybrid optimization algorithms for general smooth manifolds was developed in [30], however optimization on the sphere, as a specific case of the general framework, results in several simplifications both in the algorithm and the stability analysis. In order to avoid singular points on the unit sphere in light of the hairy ball theorem, we have employed two coordinate charts, which are created using spherical coordinates or stereographic projection. Using the nature of stochastic hybrid optimization algorithm, we have gradient flows combined with random probing on the sphere which allows the escape from saddle points and local minima. We presented the stability characterization of this algorithm particularly for the unit sphere. Lastly, a numerical example was provided.

Chapter 4

Global Optimization on the Sphere with Half-space Constraints: A Stochastic Hybrid Systems Approach

In Chapter 2, we showed how the framework of stochastic hybrid inclusions from [9] can be applied to model *hybrid* stochastic algorithms for optimization on general manifolds. Hybrid algorithms have the advantage of being able to combine classical gradient descent, or any similar gradient-based approach, which can be very effective near global minima, with discrete-time stochastic probing, which is helpful for moving away from local minima, saddle points, etc. In Chapter 3 we specialized the results of Chapter 2 to the unit sphere, where the analysis simplifies somewhat (See [39]). However, the setting of manifolds without boundary limits the scope of applications since many engineering problems require a restriction on the feasible set, often appearing in the form of inequality constraints. For instance, optimization with inequality constraints on Riemannian manifolds appears in power systems [40]. We also encounter constraints on manifolds for applications involving integration [41].

In this chapter, we extend the stochastic, hybrid algorithm for global optimization of any

continuously differentiable function defined on the unit sphere [39] to the case where the optimization problem is subject to a half-space constraint. The main technicality involves dealing with the projection of a gradient descent law onto the tangent cone of the intersection of the sphere and the half space. The projection can induce additional singularity points of the gradient flow field but, like for other singular points that do not correspond to global minima, these singularities are overcome via periodic, stochastic probing.

As an additional feature, we allow the continuous-time flows to employ, at least away from the boundary of the constraint set, a minimization algorithm that is inspired by recent results in continuous-time accelerated gradient descent, as in [42], [43] and [44]. Near the boundary of the constraint set, we insist on using simple gradient descent, to minimize the number of induced weak equilibria, and we use hysteresis switching to switch between the two continuous-time algorithms.

There has been substantial interest in the study of Nesterov's accelerated gradient descent (AGD) [45] during the past years [46], [47] and [48]. The AGD provides a type of foresight via the momentum terms in the dynamics [49]. In specific settings, this feature helps better peeking of the actual objective values in the candidate search direction, which results in an improved Recurrent Neural Network (RNN) performance on a number of tasks [50]. In [51], [52] and [53], it was shown how the worst case convergence rate of conjugate gradient on a quadratic is the same as AGD's convergence rate. The efficacy of AGD-type algorithms in the non-convex case, like on manifolds, is still being investigated [54]. We illustrate one particular version here.

4.1 Algorithm

In this section, we develop a stochastic hybrid optimization algorithm for an arbitrary C^1 function on the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$ subject to a half-space constraint. The algorithm will

make use of simple gradient descent, especially near the boundary of the constraint set, and also updates inspired by accelerated gradient descent [55, 48] away from the boundary of the constraint set.

We employ coordinate charts on the unit sphere in order to be able to fully explore and flow on the sphere. In Chapter 3, we described a hysteresis switching mechanism for switching between coordinate charts created using either spherical coordinates or stereographic projection. We take advantage of these coordinate charts and develop our algorithm for the half-space constrained sphere. In case the feasible set contained a part of both coordinate charts, switching between these two coordinate charts enables the full exploration on the feasible half-space subset of the sphere.

1. *State variables*:

   The algorithm employs state variables

$$\begin{aligned} x &:= (z, \zeta, q, \alpha, \tau) \\ &\in (\mathbb{S}^2 \cap \mathcal{H}) \times \mathbb{R}^2 \times \{1, 2\} \times \{0, 1\} \times [0, T] \subset \mathbb{R}^8 \end{aligned} \tag{4.1}$$

   where $T > 0$ and $\mathcal{H}$ is a half-space in $\mathbb{R}^3$ defined subsequently. The variable z makes progress on the sphere and half-space toward a value that achieves the minimum of the cost function. The variable ζ is an auxiliary variable that plays a role in the accelerated gradient optimization algorithm. The variable q enables hysteresis switching between coordinate charts on the sphere. The variable α enables hysteresis switching between simple gradient descent and accelerated gradient decent. The variable τ is a timer that orchestrates periodic random jumps of the variable z to avoid getting stuck at a singular point of the gradient of the cost function.

2. *Coordinate charts on the sphere*:

The first step is adapted from Chapter 3, where we construct, for $i \in \{1, 2\}$, sets $C_i \subset \mathbb{S}^2$, vector fields $B_i : \overline{C}_i \to \mathbb{R}^{3\times 2}$, sets $\mathcal{X}_i \subset \mathbb{R}^2$, and coordinate functions $h_i : \mathcal{X}_i \to \mathbb{S}^2$ such that the following properties hold:

(a) C_1 and C_2 are open relative to $\mathbb{S}^2$;

(b) $C_1 \cup C_2 = \mathbb{S}^2$;

(c) $B_i(z)$ has full column rank for each $z \in \overline{C}_i$;

(d) $h_i(\mathcal{X}_i) = \overline{C}_i$.

These objects can be created using the stereographic projection method or the spherical coordinates method as described in the previous Chapter 3. Note that the vector fields B_i and coordinate functions h_i are continuously differentiable.

3. *Half-Space Constraint*:

The half space over which we optimize our cost function is expressed in terms of a normal vector $n \in \mathbb{R}^3$ and a point $z_0 \in \mathbb{R}^3$ and is given by

$$\mathcal{H} := \{z \in \mathbb{R}^3 : n^T(z - z_0) \geq 0\}. \tag{4.2}$$

The boundary of this half space is

$$\partial\mathcal{H} := \{z \in \mathbb{R}^3 : n^T(z - z_0) = 0\} \tag{4.3}$$

and the interior of the half space is

$$\text{Int}\mathcal{H} := \{z \in \mathbb{R}^3 : n^T(z - z_0) > 0\}. \tag{4.4}$$

4. *Flow dynamics:*

i. *Input projection:* We build a dynamical control system with input $u \in \mathbb{R}^2$ that is able to fully explore the sphere by flowing while forever remaining in the prescribed half space. The dynamics on $\overline{C}_q \cap \mathcal{H}$ are given by

$$z \in \overline{C}_q \cap \mathcal{H} \qquad \begin{cases} \dot{z} \in B_q(z)\overline{\text{co}}(\text{cl}(p_q(z,u))) \\ \dot{q} = 0 \end{cases} \tag{4.5}$$

where

$$T_q(z) := \begin{cases} \mathbb{R}^2 & z \in \text{Int}\mathcal{H} \\ \left\{u \in \mathbb{R}^2 : n^T B_q(z) u \geq 0\right\} & z \in \partial\mathcal{H} \end{cases} \tag{4.6a}$$

$$p_q(z,u) := \text{argmin}_{v \in T_q(z)} |v-u|^2. \tag{4.6b}$$

The value $p_q(z,u)$ is well-defined since the mapping $v \mapsto |v-u|^2$ is convex and the set $T_q(z)$ is convex for each (q,z) such that $z \in \overline{C}_q \cap \mathcal{H}$. However, the mapping $p_q(\cdot,u)$ is discontinuous in z. Here the $\overline{\text{co}}$ is the closed convex hull of the set $\text{cl}(p_q)$. The set-valued mapping $\text{cl}(p_q)$ is the outer-semicontinuous mapping whose graph is equal to the closure of the graph of p_q; that is, it is the outer semicontinuous hull of p_q. For more details, see [56, Section 5B].

ii. *Hysteresis switching between gradient descent and accelerated gradient descent:* In addition to the logic variable q, the algorithm employs another logic variable $\alpha \in \{0,1\}$ that allows the optimization algorithm to switch, via hysteresis, between a simple gradient descent law and a law inspired by accelerated gradient descent

laws. We use the former when "close" to the boundary of the constraint set and use the latter when "far" from the boundary of the constraint set. Here, "far" could mean far away enough that we never actually use or need to use the accelerated gradient descent. The law inspired by accelerated gradient descent may permit faster convergence for optimal values away from the boundary of the constraint set. On the other hand, its implementation on the boundary of the constraint set is problematic and can induce spurious equilibria, which we wish to avoid. In particular, we use

$$u = -(1-\alpha)B_q^T(z)\nabla f(z) + \alpha\zeta \tag{4.7a}$$

$$\dot{\zeta} \in -[\delta, \Delta]\zeta - \alpha B_q^T(z)\nabla f(z) \tag{4.7b}$$

$$\dot{\alpha} = 0 \tag{4.7c}$$

where $0 < \delta \leq \Delta$. The flows are additionally constrained by the condition

$$(z, \zeta) \in H_\alpha \times \mathbb{R}^2 \tag{4.8}$$

where

$$H_0 := \left\{ z \in \mathbb{R}^3 : n^T(z - z_0) \leq \beta_0 \right\} \tag{4.9a}$$

$$H_1 := \left\{ z \in \mathbb{R}^3 : n^T(z - z_0) \geq \beta_1 \right\} \tag{4.9b}$$

with $\beta_0 > \beta_1 > 0$.

iii. *Timer variable:* The algorithm employs a timer variable τ that flows according to

$$\tau \in [0, T] \qquad \dot{\tau} = 1. \tag{4.10}$$

iv. *System Summary:* The conditions (4.5)-(4.10) define the flow dynamics

$$x \in C \quad \dot{x} \in F(x) \tag{4.11}$$

where the flow map F is outer semicontinuous and locally bounded with nonempty convex values on the closed flow set C, which is the set of points in the state space where all of the constraints in (4.5)-(4.10) are satisfied. The flow set has the form

$$C = C \times [0, T] \tag{4.12}$$

where

$$C := \tag{4.13}$$
$$\left\{(z, \zeta, q, \alpha) \in \left(\mathbb{S}^2 \cap \mathcal{H}\right) \times \mathbb{R}^2 \times \{1, 2\} \times \{0, 1\} : z \in \overline{C}_q \cap H_\alpha\right\}.$$

5. *Jump Creation*:

Our algorithm has jumps from three different sources:

i. Hysteresis switching between charts, through the variable q: These jumps toggle q while leaving all other variables unchanged. These are the only jumps that can occur when $\tau \in [0, T)$, and $z \in \text{int}(H_\alpha)$. They can occur when $z \in (\mathbb{S}^2 \setminus C_q) \cap \mathcal{H}$. Thus, defining

$$\mathcal{D}_1 := \tag{4.14}$$
$$\left\{(z, \zeta, q, \alpha) : z \in \text{int}(H_\alpha) \cap ((\mathbb{S}^2 \setminus C_q) \cap \mathcal{H})\right\}$$

we can write

$$(z, \zeta, q, \alpha, \tau) \in \mathcal{D}_1 \times [0, T) \qquad \begin{cases} z^+ = z \\ \zeta^+ = \zeta \\ q^+ = 3 - q \\ \alpha^+ = \alpha \\ \tau^+ = \tau \end{cases} . \tag{4.15}$$

ii. Random probing of cost function to avoid getting stuck at singular points of the function's gradient, triggered by a timer variable τ and possibly inducing jumps in the variable z: These are the only jumps that can occur when $z \in \text{int}(H_\alpha)$ and $z \in C_q$. They can occur when $\tau = T$. Thus, defining

$$\mathcal{D}_2 := \left\{(z, \zeta, q, \alpha) : z \in \text{int}(H_\alpha) \cap C_q \cap \mathcal{H})\right\} \tag{4.16}$$

we can write

$$(z, \zeta, q, \alpha, \tau) \in \mathcal{D}_2 \times \{T\} \qquad \begin{cases} z^+ = \omega_a \\ \zeta^+ = \zeta \\ q^+ = \omega_b \\ \alpha^+ = \alpha \\ \tau^+ = 0 \end{cases} \tag{4.17}$$

where the variables ω_a and ω_b are given by:

$$\omega_a \in \text{argmin}_{s\in\{z,h_1(\upsilon_1),h_2(\upsilon_2)\}\cap\mathcal{H}} f(s) \tag{4.18a}$$

$$\omega_b \in \left\{q \in \{1,2\} : \omega_a \in (\overline{C}_q \cap \mathcal{H})\right\}. \tag{4.18b}$$

In (4.18a), the υ is the placeholder of an iid sequence of random variables and its distribution is uniform on the set $\mathcal{X}_1 \times \mathcal{X}_2$. The coordinate function h_q maps the random variable υ_q from $\mathcal{X}_q$ to $\overline{C}_q$. Since we are restricting our attention to jumps located on the feasible set, the algorithm jumps only if the jump lands on the intersection of the manifold and the half space. By using random jumps of state, we probe the feasible set and escape local minima and saddle points. In (4.18a), the set used for possible values of z will never be empty, as z will always lay on the feasible set. The condition (4.18b) adjusts the coordinate charts appropriately.

iii. Hysteresis switching between simple gradient descent and accelerated gradient descent, via the variable α: These are the only jumps that occur when $\tau \in [0, T)$ and $z \in C_q$. They can occur when $z \in \mathbb{S}^2 \cap \overline{\mathcal{H}\backslash H_\alpha}$. Thus, defining

$$\mathcal{D}_3 := \left\{(z, \zeta, q, \alpha) : z \in C_q \cap \overline{\mathcal{H}\backslash H_\alpha}\right\} \tag{4.19}$$

we can write

$$(z, \zeta, q, \alpha, \tau) \in \mathcal{D}_3 \times [0, T) \qquad \begin{cases} z^+ = z \\ \zeta^+ = 0 \\ q^+ = q \\ \alpha^+ = 1 - \alpha \\ \tau^+ = \tau. \end{cases} \tag{4.20}$$

The union of the closure of these jump sets provides the overall jump set D for the system. Moreover, the outer semicontinuous hull of the mapping defined by all of the jumps will define the overall jump map G.

4.2 Behavior of algorithm

4.2.1 Solution Existence

Using the structure of the data created above and results from [27], local existence of solutions follows if we can show that $F(x) \cap T_C(x) \neq \emptyset$ for each $C \backslash D$, which boils down to showing that the projected flow map for the z dynamics intersects the tangent cone to $\mathbb{S}^2 \cap \mathcal{H}$ for each $z \in \partial\mathcal{H}$. Using results from [57], and the definition of $\overline{\text{co}}(\text{cl}(p_q(z, u))$, it is enough to establish that $B_q(z)p_q(z, u) \in T_{\mathbb{S}^2}(z) \cap T_{\mathcal{H}}(z)$. But this fact holds from the definition of $T_q(z)$ in the definition of $p_q(z, u)$ in (4.6a).

4.2.2 Stability Analysis

Theorem 4 *For the hybrid dynamical system defined above, the set*

$$\mathcal{A} := \mathcal{A}_0 \times \{0\} \times \{1, 2\} \times \{0, 1\} \times [0, 1] \tag{4.21}$$

where

$$\mathcal{A}_0 := \left\{ z \in \mathbb{S}^2 \cap \mathcal{H} : f(z) = \min_{s \in \mathbb{S}^2 \cap \mathcal{H}} f(s) \right\} \tag{4.22}$$

is uniformly globally asymptotically stable in probability (UGASp).

Proof. We use [28, Theorem 8]. Let $f^* := \min_{s \in \mathbb{S}^2 \cap \mathcal{H}} f(s)$. Define the Lyapunov function candidate $V : \mathbb{S}^2 \to \mathbb{R}_{\geq 0}$ as

$$V(z) := f(z) - f^* + 0.5\zeta^T \zeta.$$

This function is positive definite with respect to $\mathcal{A}$, when restricted to the union of the flow and jump sets. Moreover, its sublevel sets are compact, when again restricted to the union of the flow and jump sets. We first analyze the derivative of the Lyapunov function candidate using the projection function p_q. Subsequently, we show that a nonpositive derivative with the function p_q implies a nonpositive derivative when p_q is replaced by $\overline{\text{co}}(\text{cl}(p_q))$. We first observe that

$$u^T p_q(z, u) = |p_q(z, u)|^2. \tag{4.23}$$

Indeed, when $u \in T_q(z)$ we have $p_q(z, u) = u$ so that the above equality holds. Otherwise, $p_q(z, u)$ is a projection onto the plane and thus there exists a symmetric matrix $P_q(z)$ satisfying

$P_q(z)^2 = P_q(z)$ such that $p_q(z,u) = P_q(z)u$. Thus

$$u^T p_q(z,u) = u^T P_q(z)u = u^T P_q(z)^2 u = |p_q(z,u)|^2. \tag{4.24}$$

This calculation is relevant only for the case where $\alpha = 0$, i.e., simple gradient descent is being used. In that case,using (4.23) we note that

$$\langle \nabla f(z), B_q(z) p_q(z, -B_q^T(z)\nabla f(z))\rangle = \\ -|p_q(z, -B_q^T(z)\nabla f(z))|^2. \tag{4.25}$$

Hence, we have

$$\langle \nabla f(z), B_q(z) p_q(z, -B_q^T(z)\nabla f(z))\rangle \leq 0. \tag{4.26}$$

The inequality (4.26) holds for $u \in \text{cl}(p_q(z, -B_q^T(z)\nabla f(z)))$ since ∇f and B_q are continuous. We also have $u \in \overline{\text{co}}(\text{cl}(p_q(z, -B_q^T(z)\nabla f(z))))$ since u enters in an affine manner. Thus for $u \in \overline{\text{co}}(\text{cl}(p_q(z, -B_q^T(z)\nabla f(z))))$ equation (4.26) holds and the derivative of the Lyapunov function is negative semidefinite and it does not increase during flows. Now consider the case $\alpha = 1$. This case is used only for $z \in H_1$ defined in (4.9a). We have:

$$\langle \nabla f(z), -B_q(z)\zeta\rangle + \langle \zeta, -[\delta, \Delta]\zeta - B_q^T(z)\nabla f(z)\rangle \\ = -\zeta^T[\delta, \Delta]\zeta \leq 0. \tag{4.27}$$

Since $0 < \delta \leq \Delta$, the Lyapunov function does not increase. In this step, we check to see how the jumps affect the Lyapunov function candidate V. First, due to the definition of G, for all

$(v_1, v_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ and all $g \in G(x, v)$, we have $V(g) \leq V(x)$. Thus, we have:

$$\int_{\mathcal{X}_1 \times \mathcal{X}_2} \max_{g \in G(x,v)} V(g)\mu(dv) \leq V(x) \tag{4.28}$$

regardless of the distribution μ. As a result, V is a Lypuanov function for $\mathcal{A}$ in the sense of [28, Theorem 8]. We still have to establish that there does not exist an almost surely complete solution that remains in a non-zero level set of the Lyapunov function almost surely. We note that there is no sample path with a purely discrete-time domain. This fact follows from the hysteresis used to create each type of jump. For example, if z belongs to the relative boundary of $C_i \cap \mathcal{H}$ then it belongs to the relative interior of $C_j \cap \mathcal{H}$, $j \neq i$, by virtue of the fact that both $C_1 \cap \mathcal{H}$ and $C_2 \cap \mathcal{H}$ are relatively open and cover $\mathbb{S}^2 \cap \mathcal{H}$. Consequently, each sample path is uniformly non-Zeno and hence each complete sample path has a time domain that is unbounded in the ordinary time direction. Since jumps due to G happen every T units of ordinary time, it is enough to show that those jumps result in a decrease in the expected value of V from points outside the set of global minimizers. Let $z \in (\mathbb{S}^2 \cap \check{\mathcal{H}}) \backslash \mathcal{A}$ and let $z^* \in \mathcal{A}$. Let $i \in \{1, 2\}$ and $v_i^* \in \mathcal{X}_i$ be such that $z^* = h_i(v_i^*) \in (C_i \cap \mathcal{H})$. Let $\mathbb{B} := \{x \in \mathbb{R}^n : |x| \leq 1\}$. By the continuity of f and h_i, there exists $\varepsilon > 0$ such that $\left\{v_i^*\right\} + \varepsilon\mathbb{B} \subset \mathcal{X}_i$ and $f(h_i(v_i)) \leq f(z) - \varepsilon$ for all $v_i \in \left\{v_i^*\right\} + \varepsilon\mathbb{B}$. Without loss of generality, assume that $i = 1$. Then

$$\begin{aligned}\int_{\mathcal{X}_1 \times \mathcal{X}_2} \max_{g \in G(x,v)} V(g)\mu(dv) \leq & V(z) \\ & -\mu\left(\left(\{v_i^*\} + \varepsilon\mathbb{B}\right) \times \mathcal{X}_2\right)\varepsilon.\end{aligned}$$

Since $\mu\left(\left(\left\{v_i^*\right\} + \varepsilon\mathbb{B}\right) \times \mathcal{X}_2\right) > 0$, by virtue of the distribution being uniform over $\mathcal{X}_1 \times \mathcal{X}_2$ (though any distribution for which each open set has positive measure would be sufficient), it follows that there does not exist an almost surely complete solution that remains in a non-zero level set of the Lyapunov function almost surely. ■

4.3 Numerical Example

In this section, we provide examples of the unit sphere with half-space constraint created by two different plane equations. In order to simulate the algorithm, we need to make a (possibly time-varying) selection from the interval $[\delta, \Delta]$ to serve as the dissipation gain of the accelerated gradient algorithm. We make the selection $g(t) := \frac{3}{t+\varepsilon} + \varepsilon$ with ε being 0.01 to emulate the time-varying gain deduced in [55]. Thus, we are using $\delta = \varepsilon$ and $\Delta = \varepsilon + 3\varepsilon^{-1}$. The function defined on the sphere is given by:

$$f(z) := z_2 z_3 + z_1^3. \tag{4.29}$$

The given function has two local minima, two local maxima, four saddle points, one global maximum and one global minimum. The value of these critical points is analytically calculated and given in Chapter 3. We choose a small value of β_0 and β_1, which define H_0 and H_1 close to zero so that AGD is used except very close to the boundary. Consider a boundary plane according to (4.3), which passes through the sphere with the normal vector $n = (0, 1, 1)^T$

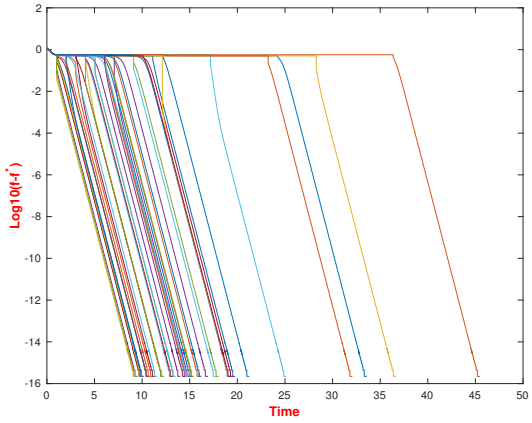

Figure 4.1: Standard Gradient Descent Sample Paths

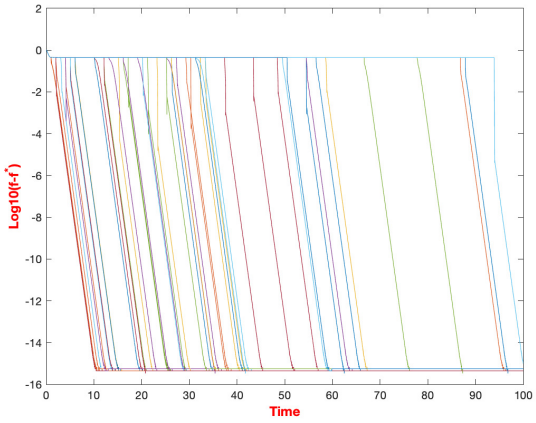

Figure 4.2: Standard Gradient Descent Sample Paths

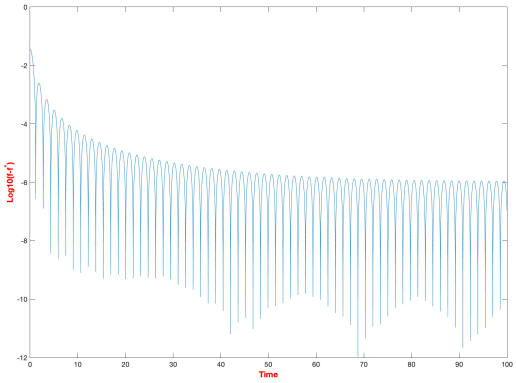

Figure 4.3: Accelerated Gradient Descent Sample Paths

and the point $(0.1343, 0.8543, -0.5021)$. In this case, the global minimum of the function is not located on the upper half-space subset of the sphere. In Figure 4.4, we demonstrate the projection of the sample paths on the boundary. Since the global minimum of the function is not located within the half space constrained set, the trajectory converges to the least value of the function in the feasible half space. The trajectory starts from an initial point (green point) in Figure 4.4 and it potentially goes towards the local minimum (light blue) or the saddle point (black), which are both located under the boundary constraint. However, the trajectory gets projected on the boundary when it reaches the plane. Finally, it converges to the least value of

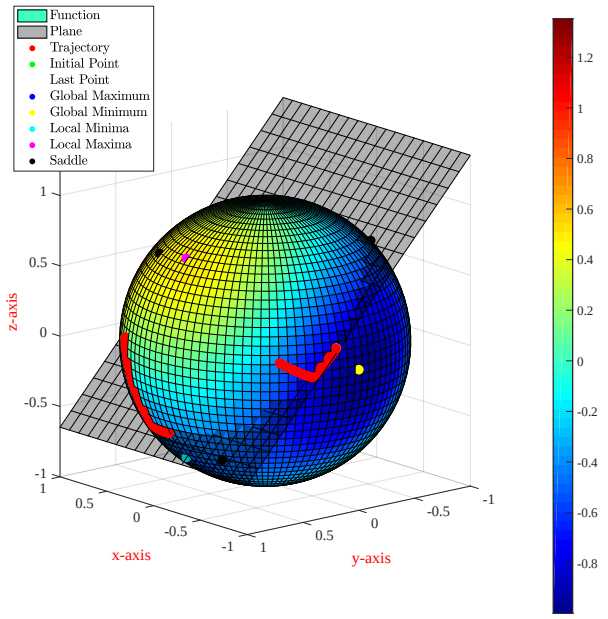

Figure 4.4: Sample Paths on the Half-Space Subset of the Sphere (Standard Gradient Descent-Optimal value on boundary)

the function within the half-space constraint created by the intersection of the given plane and the unit sphere. The analytically calculated value of the optimal point is -0.8774 and in Figure 4.4 the projected gradient descent has converged to this value.

Since the projected flows are combined with stochastic jumps, we have plotted $\log_{10}(f - f^*)$ with f^* being the optimal value in 4.1, for 50 different sample paths to demonstrate the variance of the algorithm.

In Figure 4.5, we have a different normal vector $n = (0, 1, 3)^T$. In this case, we observe that the global minimizer of the function with the value -1 is located in the interior of the feasible subset. The trajectories jump twice after getting stuck at a local minimum and a saddle point. Finally, it converges towards the least value of the function (white point) within the constrained half space.

In Figure 4.6, we also demonstrate a scenario in which the global minimum is located on the

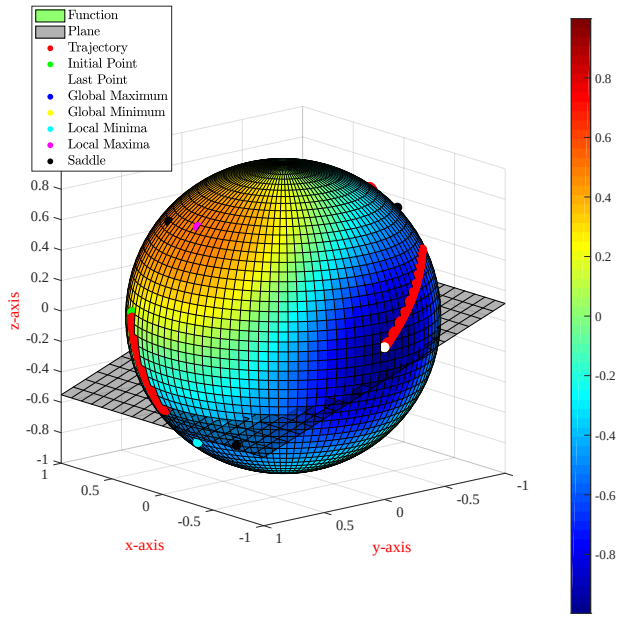

Figure 4.5: Sample Paths on the Half-Space Subset of the Sphere (Standard Gradient Descent-Optimal value in the interior)

interior of the feasible set and there is no need for boundary projection of the sample paths. In this case, the accelerated gradient descent flows for β_1 close enough to zero converge to the global minimum.

4.4 Concluding Remarks

We have captured stochastic hybrid optimization algorithms in the framework of stochastic hybrid inclusions (SHI). The SHI optimization algorithm is designed to minimize a function defined on the sphere with a half-space constraint. The algorithm has the ability of switching between accelerated gradient descent flows in the interior of the half-space and standard gradient descent flows for both interior and boundary. The idea of casting stochastic hybrid optimization algorithms for smooth manifolds was developed in Chapter 2 and applied to a unit sphere in Chapter 3. However, none of the previous chapters considered constraints for the

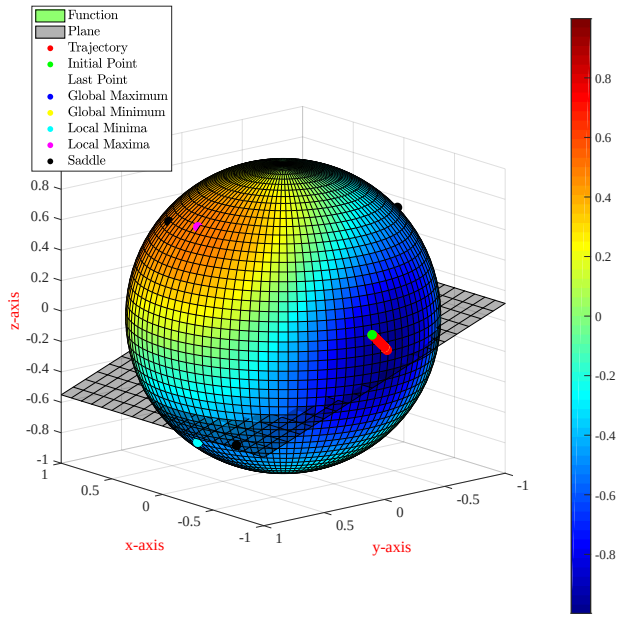

Figure 4.6: Sample Paths on the Half-Space Subset of the Sphere (Accelerated Gradient Descent- Optimal value in the interior)

optimization on manifolds. In order to avoid having dynamics and jumps outside of the half-space feasible set, we developed a projection method that keeps the standard gradient dynamics within the interior and boundary of the feasible set. The combination of the gradient flows with random probing on the sphere allows for the escape from saddle points and local minima. We presented a detailed stability characterization of this algorithm, particularly for the unit sphere. Lastly, a numerical example of optimization on the unit 2-sphere with half-space constraints is provided.

Part II

Robust stability analysis for switched systems with distinct equilibria using Omega-limit sets

Chapter 5

Omega-limit sets and robust stability for switched systems with distinct equilibria

In the following two chapters, we utilize hybrid dynamical systems for a different area of study and in conjunction with switched systems. In a series of recent papers, various authors have studied the robust stability properties of switched systems with multiple equilibria. See, for example, the pioneering results from [58] and the recent studies of [59], [60] and [61]. Studying this class of switched systems is motivated by many applications from game theory in [62], where the system switches between multiple games of different Nash equilibria, to robotics in [63],[64] for motion estimation of legged robots. The control of multicell wireless networks with mobile switching between cells (subsystems) [65] and modeling of non-spiking neurons in neurophysiology (see [66]) are other interesting applications of switched systems with multiple equilibrium points. Typically in the existing literature, an assumption about the existence of Lyapunov functions with certain properties is made and boundedness of solutions is established under a sufficiently small average dwell-time switching constraint. In this work, we eschew a Lyapunov-function-based approach and we aim to give a more precise characterization of the set to which trajectories converge. We approach the analysis problem using a hybrid systems

modeling framework [27]. We employ the notion of an Ω-limit set from a compact set of initial conditions, as considered in [67] for a hybrid system. We characterize this Ω-limit set for an associated, ideal hybrid system that employs an average dwell-time switching automaton coming from [68] or [67] with the switching rate set to zero so that only a finite number of switches is allowed. In turn, we draw conclusions for the switched system under small average dwell-time switching by using results developed for hybrid results on robust (semi-global, practical) asymptotic stability for a compact set.

5.1 Preliminaries

5.1.1 General notation

We use $\mathbb{R}^n$ to denote n-dimensional Euclidean space. We use $\mathbb{R}_{\geq 0}$ to denote the nonnegative real numbers and $\mathbb{Z}_{\geq 0}$ to denote the nonnnegative integers. We use $|x|$ to denote the Euclidean norm of the vector $x \in \mathbb{R}^n$. For a closed set $K \subset \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, the symbol $|x|_K$ denotes the distance of x to K, i.e., $|x|_K := \inf_{y \in K} |x - y|$. Given $r > 0$, we use $r\mathbb{B}$ to denote the set $\{x \in \mathbb{R}^n : |x| \leq r\}$ and $r\mathbb{B}^\circ$ to $\{x \in \mathbb{R}^n : |x| < r\}$. For a set $S \subset \mathbb{R}^n$, the symbol $\overline{S}$ denotes its closure. The closure of the convex hull of the set S is written as $\overline{\text{co}}S$. We say that $\alpha \in \mathcal{K}^+$ if $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is continuous and strictly increasing. Given $\alpha_1, \alpha_2 \in \mathcal{K}^+$, the symbol $\alpha_1 \circ \alpha_2$ denotes their composition, i.e., $\alpha_1 \circ \alpha_2(s) = \alpha_1(\alpha_2(s))$. We say that $\alpha \in \mathcal{K}$ if $\alpha \in \mathcal{K}^+$ and $\alpha(0) = 0$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if $\beta(., t)$ is of class $\mathcal{K}$ for each fixed $t \geq 0$ and $t \mapsto \beta(r, t)$ is nonincreasing and decreases to zero as $t \to \infty$ for each fixed $r \geq 0$.

5.1.2 Hybrid systems

We use the framework for hybrid systems described in [27]. The model of a hybrid system is written formally as

$$x \in C, \qquad \dot{x} \in F(x) \tag{5.1a}$$

$$x \in D, \qquad x^+ \in G(x) \tag{5.1b}$$

where $x \in \mathbb{R}^n$ is the state, $C \subset \mathbb{R}^n$ is the flow set, $D \subset \mathbb{R}^n$ is the jump set, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the flow map, and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the jump map. The data (C, F, D, G) is said to satisfy the *hybrid basic conditions* if C and D are closed, the graphs of F and G are closed, F and G are locally bounded, the values of F are nonempty and convex on C and the values of G are nonempty on D. A solution of the hybrid system (5.1) is a hybrid arc satisfying the constraints in (5.1); a hybrid arc is defined through the following concepts. A *compact hybrid time domain* is a set of the form

$$\cup_{j=0}^{J} \Big([t_j, t_{j+1}] \times \{j\} \Big) \subset \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$$

for some real numbers $0 = t_0 \leq t_1 \leq \cdots \leq t_{J+1}$. A *hybrid time domain* is a set $E \subset \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ having the property that, for each $(T, J) \in E$, the set $E \cap ([0, T] \times \{0, \ldots, J\})$ is a compact hybrid time domain. A *hybrid arc* is a function $x : E \to \mathbb{R}^n$ where E is a hybrid time domain and $x(\cdot, j)$ is locally absolutely continuous for each nonnegative integer j. We typically use $\mathrm{dom}(x)$ to denote the domain of the hybrid arc x. A hybrid arc is a *solution of (5.1)* if it satisfies the constraints implicit in (5.1), i.e.,

1. If $(t_1, j), (t_2, j) \in \mathrm{dom}(x)$ and $t_1 < t_2$ then, for almost all $t \in [t_1, t_2]$,

$$x(t, j) \in C, \quad \dot{x}(t, j) \in F(x(t, j))$$

2. If $(t, j), (t, j+1) \in \text{dom}(x)$ then

$$x(t, j) \in D, \qquad x(t, j+1) \in G(x(t, j)).$$

Given $K \subset \mathbb{R}^n$, we use $\mathcal{S}(K)$ to denote the *set of solutions to (5.1) that start in K*. Given $K \subset \mathbb{R}^n$, we use $R(K)$ to denote the *reachable set from K*, i.e.,

$$R(K) := \{z \in \mathbb{R}^n : z = x(t, j),\ x \in \mathcal{S}(K), (t, j) \in \text{dom}(x)\}.$$

Given $K \subset \mathbb{R}^n$, we use $\Omega(K)$ to denote the Ω-*limit set from K*, i.e.,

$$\Omega(K) := \Big\{ z \in \mathbb{R}^n : z = \lim_{i\to\infty} x_i(t_i, j_i),\ x_i \in \mathcal{S}(K), \\ (t_i, j_i) \in \text{dom}(x_i), \lim_{i\to\infty} t_i + j_i = \infty \Big\}$$

A sequence of hybrid arcs $\{x_i\}_{i=1}^{\infty}$ is said to be *locally eventually bounded* if for any $m > 0$, there exists $i_0 > 0$ and a compact set $K \subset \mathbb{R}^n$ such that for all $i > i_0$, all $(t, j) \in \text{dom}\,\phi_i$ with $t + j < m$, $x_i(t, j) \in K$.

Fundamental to our analysis are results derived from the properties of the solutions to (5.1) for the solutions of the inflated system

$$x \in C_\delta, \qquad \dot{x} \in F_\delta(x) \tag{5.2a}$$

$$x \in D_\delta, \qquad x^+ \in G_\delta(x) \tag{5.2b}$$

where $\delta > 0$ and

$$C_\delta := \{x \in \mathbb{R}^n : (x + \delta\mathbb{B}) \cap C \neq \emptyset\} \tag{5.3a}$$

$$F_\delta := \overline{\text{co}} F((x + \delta\mathbb{B}) \cap C) + \delta\mathbb{B} \tag{5.3b}$$

$$D_\delta := \{x \in \mathbb{R}^n : (x + \delta\mathbb{B}) \cap D \neq \emptyset\} \tag{5.3c}$$

$$G_\delta := G((x + \delta\mathbb{B}) \cap D) + \delta\mathbb{B}. \tag{5.3d}$$

We use $\mathcal{S}_\delta(K)$ to denote the solutions of (5.2) from K.

5.1.3 Stability concepts

We state several stability concepts for hybrid systems. They apply just as well to ordinary differential equations. The hybrid system (5.1) is said to be *Lagrange stable* if there exists $\alpha \in \mathcal{K}^+$ such that, for each $x_\circ \in \mathbb{R}^n$, each $x \in \mathcal{S}(x_\circ)$, and $(t, j) \in \text{dom}(x)$, we have

$$|x(t, j)| \leq \alpha(|x_\circ|).$$

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *(Lyapunov) stable* for the hybrid system (5.1) if, for each $\varepsilon > 0$ there exists $\delta > 0$ such that $|x_\circ|_\mathcal{A} \leq \delta$, $x \in \mathcal{S}(x_\circ)$ and $(t, j) \in \text{dom}(x)$ imply that $|x(t, j)|_\mathcal{A} \leq \varepsilon$.

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *attractive* for the hybrid system (5.1) if there exists $\delta > 0$ such that each solution $x \in \mathcal{S}(\mathcal{A} + \delta\mathbb{B})$ is bounded and, if complete, satisifes $x(t, j) \to \mathcal{A}$ as $t + j \to \infty$. The *basin of attraction* for an attractive set $\mathcal{A}$ is the set of initial conditions from which each solution is bounded and, if complete, converges to $\mathcal{A}$ as $t + j \to \infty$.

A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *asymptotically stable* for the hybrid system (5.1) if it is stable and attractive. It is said to be *globally asymptotically stable* for the hybrid system (5.1) if it is asymptotically stable with stable $\mathbb{R}^n$ as its basin of attraction.

The set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *semiglobally practically asymptotically stable in the parameter* $\delta > 0$ for the system (5.2) if there exists $\beta \in \mathcal{KL}$ and for each $\varepsilon > 0$ and $\Delta > 0$ there exists $\delta > 0$ such that each $x \in \mathcal{S}_\delta(\mathcal{A} + \Delta\mathbb{B})$ satisfies

$$|x(t, j)|_{\mathcal{A}} \leq \beta(|x(0, 0)|_{\mathcal{A}}, t + j) \qquad \forall (t, j) \in \mathrm{dom}(x). \tag{5.4}$$

5.1.4 Some useful preliminary results

The first preliminary result is contained in Exercise 4.3(b) of [56].

Lemma 2 *If the convergent sequence* $\{z_i\}_{i=1}^{\infty}$ *satisfies* $z_i \in S_i$ *for all i, where* $\{S_i\}_{i=1}^{\infty}$ *is a decreasing sequence of closed subsets of* $\mathbb{R}^n$*, i.e.,* $S_{i+1} \subset S_i \subset \mathbb{R}^n$ *for all i, then* $\lim_{i\to\infty} z_i \in \bigcap_i S_i$.

The next result is Corollary 7.7 from [27].

Lemma 3 *Suppose* (C, F, D, G) *satisfy the hybrid basic conditions. Let* K *be compact and suppose that* $R(K)$ *is bounded and* $\Omega(K)$ *is nonempty and contained in the interior of* K*. Then* $\Omega(K)$ *is asymptotically stable with basin of attraction containing* K*.*

The next result is Lemma 7.20 from [27].

Lemma 4 *Suppose* (C, F, D, G) *satisfy the hybrid basic conditions. If the compact set* $\mathcal{A}$ *is globally asymptotically stable for (5.1) then it is semiglobally practically asymptotically stable in the parameter* $\delta > 0$ *for the system (5.2).*

5.2 Problem setting

Let M be a positive integer and define

$$Q := \{1, \ldots, M\}. \tag{5.5}$$

For each $q \in \mathcal{Q}$, let $f_q : \mathbb{R}^n \to \mathbb{R}^n$. Let $\delta > 0$. We analyze the asymptotic behavior of the solutions of the differential inclusion

$$\dot{z} \in \overline{\text{co}} f_q(z + \delta\mathbb{B}) + \delta\mathbb{B} \tag{5.6}$$

where $q : \mathbb{R}_{\geq 0} \to \mathcal{Q}$ is any switching signal that satisfies an average dwell-time switching constraint paramterized by δ. In particular, letting N_0 be a positive integer, and letting $N(s,t)$ denote the number of switches of q in the interval $[s,t]$, we assume that

$$N(s,t) \leq \delta(t-s) + N_0 \qquad \forall\ 0 \leq s \leq t. \tag{5.7}$$

Our other assumption pertains to the family of differential equations

$$\dot{z} = f_q(z) \tag{5.8}$$

where $q \in \mathcal{Q}$ is constant.

Assumption 7 *For each $q \in \mathcal{Q}$, f_q is continuous and the point $z_q^* \in \mathbb{R}^n$ is globally asymptotically stable for (5.8).* ■

At times, we may also impose the following assumption, a sufficient condition for which is that the continuity in Assumption 7 is strengthened to local Lipschitz continuity.

Assumption 8 *For each $q \in \mathcal{Q}$, the solution of the initial value problem*

$$\dot{z} = -f_q(z)\,, \quad z(0) = z_q^* \tag{5.9}$$

is unique. ■

Our goal is to characterize the asymptotic behavior of (5.6) under the switching signal con-

straint (5.7) when $\delta > 0$ is small. To make progress toward this goal, we cast the combination of (5.1) and (5.7) as an equivalent hybrid system that employs an automaton to capture the average dwell-time switching constraint. That is, we consider the behavior of the hybrid system $\mathcal{H}_\delta$ given by

$$\mathcal{H}_\delta \begin{cases} (z,q,\tau) \in \mathbb{R}^n \times \mathcal{Q} \times [0, N_0] & \begin{cases} \dot{z} \in \overline{\text{co}} f_q(z + \delta\mathbb{B}) + \delta\mathbb{B} \\ \dot{q} = 0 \\ \dot{\tau} \in [0, \delta] \end{cases} \\ (z,q,\tau) \in \mathbb{R}^n \times \mathcal{Q} \times [1, N_0] & \begin{cases} z^+ = z \\ q^+ \in \mathcal{Q} \setminus \{q\} \\ \tau^+ = \tau - 1. \end{cases} \end{cases} \tag{5.10}$$

According to [69], the solutions to (5.10) are in a one-to-one correspondence with the solutions of (5.1) under the switching constraint (5.7). We also note that, under Assumption 7, the data of the hybrid system (5.10) satisfies the hybrid basic conditions spelled out in [27, Assumption 6.5].

5.3 Analysis of an ideal system

5.3.1 The model

To characterize the asymptotic behavior of the solutions of $\mathcal{H}_\delta$ in (5.10), we first characterize the asymptotic behavior of the the ideal system $\mathcal{H}_0$ that results from setting $\delta = 0$ in (5.10),

i.e.,

$$\mathcal{H}_0 \begin{cases} (z,q,\tau) \in \mathbb{R}^n \times Q \times [0, N_0] & \begin{cases} \dot{z} = f_q(z) \\ \dot{q} = 0 \\ \dot{\tau} = 0 \end{cases} \\ (z,q,\tau) \in \mathbb{R}^n \times Q \times [1, N_0] & \begin{cases} z^+ = z \\ q^+ \in Q \setminus \{q\} \\ \tau^+ = \tau - 1. \end{cases} \end{cases} \tag{5.11}$$

We will see that the asymptotic behavior of the solutions of this system will give an indication of the asymptotic behavior of the solutions to the system (5.10).

5.3.2 Boundedness

In this section, we establish a boundedness property for the solutions of $\mathcal{H}_0$ in (5.11) under Assumption 7. We start with such a boundedness result under a relaxation of Assumption 7.

Proposition 1 *If, for each $q \in Q$, the system (5.8) is Lagrange stable then the hybrid system $\mathcal{H}_0$ in (5.11) is Lagrange stable.*

Proof: According to the assumption of the proposition, there exists a family of functions $\left\{\alpha_q\right\}_{q\in Q}$ with $\alpha_q \in \mathcal{K}^+$ for each $q \in Q$, such that each solution $x = (z, q, \tau)$ of the flow dynamics in (5.11), i.e., of

$$(z,q,\tau) \in \mathbb{R}^n \times Q \times [0, N_0] \quad \begin{cases} \dot{z} = f_q(z) \\ \dot{q} = 0 \\ \dot{\tau} = 0 \end{cases} \tag{5.12}$$

satisfies

$$|x(t)| \leq \alpha_{q(0)}(|x(0)|) \qquad \forall t \in \mathrm{dom}(x). \tag{5.13}$$

Let $\mathcal{N}$ denote the family of functions obtained from k compositions of the functions α_q for any $k \in \{1, \ldots, N_0\}$ without composing the same function with itself, i.e.,

$$\begin{aligned} \mathcal{N} := \Big\{ \alpha : \alpha = \alpha_{q_k} \circ \cdots \circ \alpha_{q_1},\ k \in \{1, \ldots, N_0\}, \\ q_j \in \mathcal{Q} \quad \forall j \in \{1, \ldots, k\}, \\ q_{j+1} \neq q_j \quad \forall j \in \{1, \ldots, k-1\} \Big\}. \end{aligned} \tag{5.14}$$

Since the composition of continuous, nondecreasing functions is continuous and nondecreasing, it follows that $\mathcal{N} \subset \mathcal{K}^+$. Note that the number of functions in the set $\mathcal{N}$ is finite. Thus we can define

$$\widetilde{\alpha}(s) := \max_{\alpha \in \mathcal{N}} \alpha(s) \qquad \forall s \geq 0, \tag{5.15}$$

yielding $\widetilde{\alpha} \in \mathcal{K}^+$ since the pointwise maximum of continuous, nondecreasing functions is continuous and nondecreasing.

Let $x = (z, q, \tau)$ be a complete solution of $\mathcal{H}_0$ in (5.11) and define $J := \max_{(t,j) \in \mathrm{dom}(x)} j$. Note that J is well-defined and satisfies $J \in \{0, \ldots, N_0\}$; it denotes the number of switches experienced by the solution x. Using this definition, we can write $\mathrm{dom}(x)$ as

$$\mathrm{dom}(x) = \left(\bigcup_{j=0}^{J-1} ([t_j, t_{j+1}] \times \{j\}) \right) \cup \Big([t_J, \infty) \times \{J\} \Big) \tag{5.16}$$

where $0 = t_0 \leq t_1 \leq \cdots t_J < \infty$. For notational convenience, we let $t_{J+1} > t_J$ denote an arbitrarily large positive number.

It follows from the assertion in (5.13) for the solutions of the system (5.12) that, for each $j \in \{0, \ldots, J\}$ and each $t \in [t_j, t_{j+1}]$,

$$|x(t, j)| \leq \alpha_{q(t_j,j)}(|x(t_j, j)|). \tag{5.17}$$

By concatenating these bounds, it follows that, for all $k \in \{0, \ldots, J\}$ and each $t \in [t_k, t_{k+1}]$,

$$|x(t, k)| \leq \alpha_{q(t_k,k)} \circ \alpha_{q(t_{k-1},k-1)} \circ \ldots \circ \alpha_{q(0,0)}(|x(0, 0)|). \tag{5.18}$$

By the definition of the flow map and jump map in (5.11), it follows that $q(t_k, k) \neq q(t_{k-1}, k-1)$ for each $k \in \{1, \ldots, J\}$. Hence, for each $k \in \{0, \ldots, J\}$, we have

$$\alpha_{q(t_k,k)} \circ \alpha_{q(t_{k-1},k-1)} \circ \ldots \circ \alpha_{q(0,0)} \in \mathcal{N}. \tag{5.19}$$

It follows from the definition of $\widetilde{\alpha}$ in (5.15) that, for each $k \in \{0, \ldots, J\}$ and each $t \in [t_k, t_{k+1}]$,

$$|x(t, k)| \leq \widetilde{\alpha}(|x(0, 0)|). \tag{5.20}$$

Since t_{J+1} is arbitrary, it follows that

$$|x(t, k)| \leq \widetilde{\alpha}(|x(0, 0)|) \qquad \forall (t, k) \in \mathrm{dom}(x). \tag{5.21}$$

Thus, the hybrid system (5.11) is Lagrange stable. ■

Since global asymptotic stability of a compact set implies Lagrange stability, the following corollary is a consequence of Proposition 1.

Corollary 1 *Under Assumption 7, the hybrid system $\mathcal{H}_0$ defined in (5.11) is Lagrange stable.*

5.3.3 The Ω-limit set for $\mathcal{H}_0$

Let $K \subset \mathbb{R}^{n+2}$. For the system $\mathcal{H}_0$ defined in (5.11), we use $\Omega_0(K)$ to denote the Ω-limit set from K and we use $R_0(K)$ to denote the reachable set from K. We define

$$S_q := \bigcap_{j \in \mathbb{Z}_{\geq 0}} \overline{R_0\left(\left(\left\{z_q^*\right\} + \frac{1}{j+1}\mathbb{B}\right) \times \{q\} \times [0, N_0]\right)} \tag{5.22a}$$

$$S := \bigcup_{q \in Q} S_q. \tag{5.22b}$$

The next lemma is a result of Corollary 1 and the construction of S in (5.22).

Lemma 5 *Under Assumption 7, the set S defined in (5.22) is compact.*

The rest of this section is devoted to establishing that $\Omega_0(K) = S$ for sufficiently large compact sets K.

Proposition 2 *If Assumptions 7 and 8 hold then, for each compact set $K \subset \mathbb{R}^{n+2}$ containing the set*

$$\left(\bigcup_{q \in Q} \left\{z_q^*\right\} \times \{q\}\right) \times [0, N_0]$$

in its interior, $\Omega_0(K) = S$.

Proposition 2 follows from the subsequent two lemmas.

Lemma 6 *If Assumption 7 holds then, for each compact set $K \subset \mathbb{R}^{n+2}$, $\Omega_0(K) \subset S$.*

Proof: Let $p \in \Omega_0(K)$ and let the sequence of solutions $\phi_i \in \mathcal{S}(K)$ and times $(t_i, j_i) \in \mathrm{dom}(\phi_i)$ satisfy

$$\lim_{i \to \infty} t_i + j_i = \infty \tag{5.23a}$$

$$\lim_{i \to \infty} \phi_i(t_i, j_i) = p. \tag{5.23b}$$

Since K is compact and the system $\mathcal{H}_0$ is Lagrange stable (due to Corollary 1) the sequence $\{\phi_i\}_{i=1}^{\infty}$ is locally eventually bounded. Consequently, it contains a subsequence converging to a complete solution $\phi \in \mathcal{S}(K)$ [27, Theorem 6.1]. Henceforth, we use $\{\phi_i\}_{i=1}^{\infty}$ for the converging subsequence. Define $J := \max_{(t,j)\in \mathrm{dom}(\phi)} j$. Note that J is well-defined and satisfies $J \in \{0, \ldots, N_0\}$; it denotes the number of switches experienced by the solution ϕ. Using this definition, we can write $\mathrm{dom}(\phi)$ as

$$\mathrm{dom}(\phi) = \left(\bigcup_{j=0}^{J-1} ([t_j, t_{j+1}] \times \{j\}) \right) \cup \left([t_J, \infty) \times \{J\} \right) \tag{5.24}$$

where $0 = t_0 \leq t_1 \leq \cdots t_J < \infty$. Moreover, with $(z, q, \tau) = \phi$, since q and τ are constant during flows, there exists $(q^*, \tau^*) \in \mathcal{Q} \times [0, N_0]$ such that $(q(t, J), \tau(t, J)) = (q^*, \tau^*)$ for all $t \in [t_J, \infty)$. Also, due to Assumption 7,

$$\lim_{t\to\infty} |z(t, J) - z^*_{q^*}| = 0. \tag{5.25}$$

Thus, there exists an increasing, unbounded sequence of times $\left\{s_j\right\}_{j\in\mathbb{Z}_{\geq 0}}$ such that, for each $j \in \mathbb{Z}_{\geq 0}$,

$$t_J \leq s_j\,, \quad |z(s_j, J) - z^*_{q^*}| \leq \frac{1}{2(j+1)}. \tag{5.26}$$

For each $j \in \mathbb{Z}_{\geq 0}$, let $i^*(j) \in \mathbb{Z}_{\geq 0}$ be such that, for all $i \geq i^*(j)$, we have

$$t_i + j_i \geq s_j + J + 0.5 \tag{5.27}$$

and there exists $\hat{t}_i$ such that $(\hat{t}_i, J) \in \mathrm{dom}(\phi_i)$ satisfying

$$|\hat{t}_i - s_j| \leq \frac{1}{2(j+1)} \tag{5.28a}$$

$$|\phi_i(\hat{t}_i, J) - \phi(s_j, J)| \leq \frac{1}{2(j+1)}. \tag{5.28b}$$

By combining (5.26)-(5.28), it follows that

$$t_i + j_i \geq \hat{t}_i + J \tag{5.29a}$$

$$|z_i(\hat{t}_i, J) - z^*_{q^*}| \leq \frac{1}{j+1}. \tag{5.29b}$$

It follows that

$$\phi_{i^*(j)}(t_{i^*(j)}, j_{i^*(j)}) \in \overline{R_0\left(\left(\left\{z^*_{q^*}\right\} + \frac{1}{j+1}\mathbb{B}\right) \times \{q^*\} \times \{\tau^*\}\right)}. \tag{5.30}$$

Without loss of generality, we may assume that

$$i^*(j+1) \geq i^*(j) + 1 \tag{5.31}$$

so that $i^*(j)$ grows unbounded in j, and hence, using (5.23),

$$\lim_{j\to\infty} t_{i^*(j)} + j_{i^*(j)} = \infty \tag{5.32a}$$

$$\lim_{j\to\infty} \phi_{i^*(j)}(t_{i^*(j)}, j_{i^*(j)}) = p. \tag{5.32b}$$

It follows from Lemma 2 that

$$p \in \bigcap_{j\in\mathbb{Z}_{\geq 0}} \overline{R_0\left(\left(\left\{z^*_{q^*}\right\} + \frac{1}{j+1}\mathbb{B}\right) \times \{q^*\} \times \{\tau^*\}\right)}$$
$$\subset S_{q^*} \subset \bigcup_{q\in Q} S_q = S. \tag{5.33}$$

This containment establishes the result. ■

Lemma 7 *If Assumptions 7 and 8 hold then, for each compact set $K \subset \mathbb{R}^{n+2}$ containing the set*

$$K_0 := \left(\bigcup_{q\in Q}\left\{z^*_q\right\} \times \{q\}\right) \times [0, N_0] \tag{5.34}$$

in its interior, $S \subset \Omega_0(K)$.

Proof: Since K_0 belongs to the interior of K, there exists $\varepsilon > 0$ such that $K_0 + \varepsilon\mathbb{B} \subset K$. Let $p \in S$. According to (5.22b), we have

$$p \in \bigcup_{q\in Q} S_q\,. \tag{5.35}$$

Thus, there exists a $q^* \in Q$ such that

$$p \in S_{q^*} \tag{5.36}$$
$$= \bigcap_{j\in\mathbb{Z}_{\geq 0}} \overline{R_0\left(\left(\left\{z^*_{q^*}\right\} + \frac{1}{j+1}\mathbb{B}\right) \times \{q^*\} \times [0, N_0]\right)}.$$

As a result we have that, for all $j \in \mathbb{Z}_{\geq 0}$,

$$p \in \overline{R_0\left(\left(\left\{z^*_{q^*}\right\} + \frac{1}{j+1}\mathbb{B}\right) \times \{q^*\} \times [0, N_0]\right)}. \tag{5.37}$$

It follows that there exist a solution ϕ_j^* and $(t_j^*, l_j^*) \in \operatorname{dom} \phi_j^*$ such that

$$\phi_j^*(0,0) \in \left(\left(\left\{ z_{q^*}^* \right\} + \frac{1}{j+1} \mathbb{B} \right) \times \{q^*\} \times [0, N_0] \right) \tag{5.38}$$

and

$$|\phi_j^*(t_j^*, l_j^*) - p| \leq \frac{1}{j+1}. \tag{5.39}$$

Let z_j^* and τ^* be such that $\phi_j^*(0,0) = (z_j^*, q^*, \tau_j^*)$. Let z_j be a solution to the system $\dot{z} = -f_{q^*}(z)$ with the initial condition

$$z_j(0) = z_j^*. \tag{5.40}$$

Define

$$h(j) := \min\{j, \inf\{t \in \operatorname{dom} z_j : |z_j(t) - z_q^*| = \varepsilon\}\}. \tag{5.41}$$

It follows from Assumptions 7 and 8 that

$$\lim_{j \to \infty} h(j) = \infty. \tag{5.42}$$

Next, we define a hybrid arc $\overline{\phi}_j$ with the domain

$$\operatorname{dom} \overline{\phi}_j := ([0, h(j)] \times \{0\}) \bigcup (\operatorname{dom} \phi_j^* + (\{h(j)\} \times \{0\})), \tag{5.43}$$

given by

$$\overline{\phi}_j(t,k) := \tag{5.44}$$
$$\begin{cases} (z_j(h(j)-t), q^*, \tau_j^*) & \forall (t,k) \in [0, h(j)] \times \{0\} \\ \phi_j^*(t-h(j),k) & \forall (t,k) \in \operatorname{dom} \phi_j^* + (\{h(j)\} \times \{0\}). \end{cases}$$

It can be verified that $\overline{\phi}_j$ is a solution of system (5.11) starting at $(z_j(h(j)), q^*, \tau_j^*)$. This point belongs to K, due to (5.41) and the definition of ε.

Next, we define

$$t_j := t_j^* + h(j), \qquad l_j := l_j^* \tag{5.45}$$

so that, due to (5.42),

$$\lim_{j\to\infty} t_j + l_j = \infty. \tag{5.46}$$

It follows from (5.45), (5.44) and (5.39) that

$$\begin{aligned} |\overline{\phi}_j(t_j, l_j) - p| &= |\overline{\phi}_j(t_{j^*} + h(j), l_{j^*}) - p| \\ &= |\phi_j^*(t_j^*, l_j^*) - p| \leq \frac{1}{(j+1)}. \end{aligned} \tag{5.47}$$

As a result, we have that

$$\lim_{j\to\infty} \overline{\phi}_j(t_j, l_j) = p. \tag{5.48}$$

Now follows from (5.46) and (5.48) that $p \in \Omega_0(K)$. ■

5.4 Main result

We are now ready to state our main results.

Theorem 5 *Under Assumptions 7 and 8, the set S defined in (5.22) is semiglobally, practically asymptotically stable in the parameter $\delta > 0$ for the system $\mathcal{H}_\delta$ defined in (5.10).*

Proof: Let the compact set $K \subset \mathbb{R}^{n+2}$ be such that the set S defined in (5.22), which is compact according to Lemma 5, is contained in the interior of K. According to Proposition 2, $\Omega(K)$ is contained in the interior of K. According to Lemma 3, the compact set $\Omega(K)$ is asymptotically stable with basin of attraction containing K for the system $\mathcal{H}_0$ defined in (5.11). Since K can be taken to be arbitrarily large, it follows that S is globally asymptotically stable for the system $\mathcal{H}_0$ defined in (5.11). It then follows from Lemma 4 that the set S is semi-globally practically asymptotically stable in $\delta > 0$ for the system $\mathcal{H}_\delta$ defined in (5.10). ■

In the case where Assumption 8 does not hold, we still have the following result:

Theorem 6 *Let Assumption 7 hold and let $r > 0$ be such that $S \subset r\mathbb{B}^0$. Then the set $\Omega_0(r\mathbb{B})$ is compact, contained in S, and semi-globally, practically asymptotically stable in the parameter $\delta > 0$ for the system $\mathcal{H}_\delta$ defined in (5.10).*

Proof: According to Lemma 6, $\Omega_0(r\mathbb{B}) \subset S$. Then, due to the assumption that $S \subset r\mathbb{B}^\circ$, we have that $\Omega_0(r\mathbb{B}) \subset r\mathbb{B}^\circ$. It follows from Lemma 3 that the set $\Omega_0(r\mathbb{B})$ is asymptotically stable with basin of attraction containing $r\mathbb{B}$. We claim that the basin of attraction is $\mathbb{R}^{n+2}$. Indeed, for any value $r' > r$, we again have that $\Omega_0(r'\mathbb{B}) \subset S$ is asymptotically stable with basin of attraction containing $r'\mathbb{B}$. It follows from the containment $S \subset r\mathbb{B}^\circ$ that each complete solution from $r'\mathbb{B}$ reaches $r\mathbb{B}$ in finite time, and thus each point in $r'\mathbb{B}$ belongs to the basin of attraction of the asymptotically stable set $\Omega_0(r\mathbb{B})$. [1] Since $r' > r$ was arbitrary, this observation establishes that the set $\Omega_0(r\mathbb{B})$ is globally asymptotically stable for the system $\mathcal{H}_0$ defined in

[1] In fact, it can be shown that $\Omega(r'\mathbb{B}) = \Omega(r\mathbb{B})$ for each $r' > r$, though this is not needed for the proof.

(5.11). It then follows from Lemma 4 that the set S is semi-globally practically asymptotically stable in $\delta > 0$ for the system $\mathcal{H}_\delta$ defined in (5.10). ■

5.5 Numerical Example

In this section, we consider an example with $N_0 = 1$, to ease the visualization of the ideal Ω-limit set. Consider the following linear time-invariant systems

$$\dot{x} = A_1 x + b_1 \tag{5.49a}$$

$$\dot{x} = A_2 x + b_2, \tag{5.49b}$$

where $x \in \mathbb{R}^2$ and

$$A_1 = \begin{bmatrix} 0 & 1 \\ -10 & -1 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 0 & 10 \\ -1 & -1 \end{bmatrix}$$
$$b_1 = \begin{bmatrix} -20 \\ 4 \end{bmatrix}; \qquad b_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}.$$

The matrices A_1 and A_2 are invertible, yielding the unique equilibrium points for (5.49a) and (5.49b), respectively, at $x_1^* = [-1.6, 20]^T$ and $x_2^* = [3.9, 0.1]^T$. Each equilibrium is exponentially stable since A_1 and A_2 are Hurwitz.

In Figure 5.1, we have generated the Ω-limit set for the ideal hybrid system.

Figure 5.2 shows behavior under a particular, persistently switching signal that satisfies a dwell-time constraint with $\delta = 0.1$. The initial condition starts near the ideal Ω-limit set and so corresponds to a type of "steady-state" behavior. Figure 5.3 shows behavior from an initial condition farther from the ideal Ω-limit set.

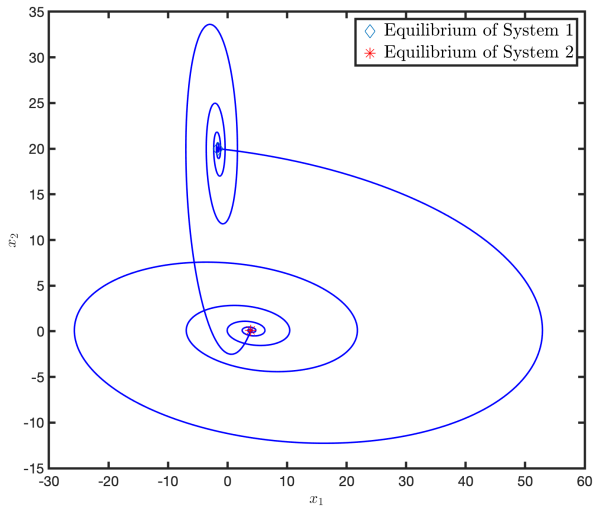

Figure 5.1: The Ω-limit set corresponding to the ideal system (5.11) for the example (5.49)

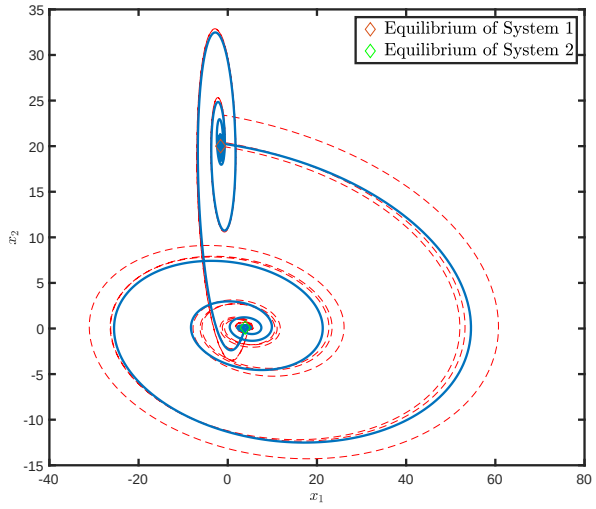

Figure 5.2: "Steady-state" behavior near the ideal Ω-limit set for dwell-time switching with $\delta = \frac{1}{10}$ and disturbances

5.6 Concluding Remarks

This chapter provides a characterization of the asymptotic behavior of a perturbed, switched system with distinct equilibria under average dwell-time switching with a small rate parameter. The asymptotic behavior of an ideal hybrid system without disturbances and without persistent

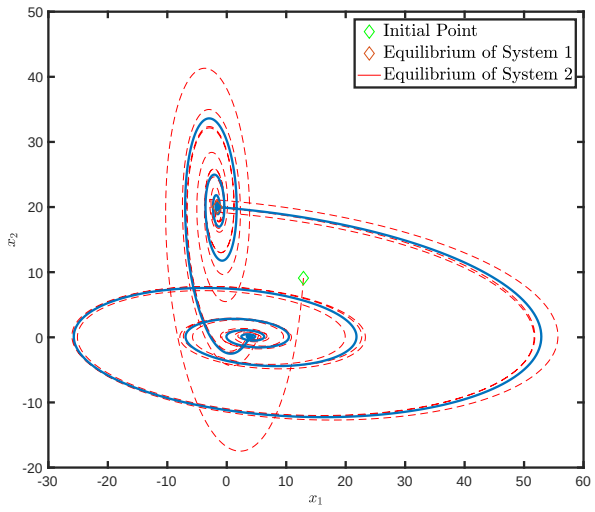

Figure 5.3: Transient response from a non-equilibrium initial condition

switching was analyzed first. It was shown that the solutions of such a system are bounded if each subsystem is Lagrange stable. Subsequently, the Ω-limit for the ideal hybrid system was characterized and was shown to be semiglobally practically asymptotically stable in the average dwell-time parameter for the switched system. Finally, an example for a system with two equilibria was provided.

Chapter 6

Analyzing the Effect of Persistent Asset Switches on a Class of Hybrid-Inspired Optimization Algorithms

In this chapter, we exploit the previously developed theory for an optimization problem. Convex optimization challenges are pervasive across many current science and technology fields. Such optimization problems are often solved using iterative algorithms such as first-order gradient-based methods that can be naturally represented and analyzed as dynamical systems. However, most studies of these algorithms do not account for applications in which the objective function to be optimized can instantaneously change at discrete moments in time during the algorithm's execution. Switching objectives arise in increasingly many real-world applications, such as multi-agent systems in which the agents must be replaced according to a real-time mission constraint, as well as resource allocation problems in which the allocated assets can experience persistent changes in the reward they generate. Other examples can be found in branches of human science, such as sociology, psychology, and organization science, which study the group interactions and performance of teams with multiple individuals engaged in a

common task [70], [71]. In [72], for instance, an algorithm is proposed and studied for optimizing the sum of the team members' performance measures, and each member's performance measure is determined by variables such as the member's specific skill level. Thus, the objective will switch whenever a team member is replaced during execution of the algorithm. Persistent switches of assets could also play an important role in the realm of intelligent control of unmanned aerial vehicles (UAVs), also referred to as drones. In recent years, there has been a surge in the use of UAVs for surveillance and security, parcel shipment, traffic monitoring, disaster recovery, and military reconnaissance [73],[74]. Cooperative and collaborative optimization of UAV performance is essential in such applications. In [75], applications and engineering constraints for UAV-based mobile ad hoc networking are surveyed. In all these settings, there are few studies of how optimization dynamics are impacted by objectives that persistently switch due to failures or replacements of UAVs in a network/relay. In this chapter, we analyze how the presence of a persistently switching objective impacts the asymptotic stability properties of optimization dynamics. Our analysis aims to be applicable to systems such as the aforementioned UAV ad hoc networks, where the UAVs that suffer from low battery, potential damages, or other disabling aspects may be replaced during the execution of optimization dynamics. Our analysis targets applications where a system must instantaneously replace assets (UAVs, team members, etc.) at discrete moments in time, while it optimizes performance continuously in time. To prevent instability and be able to characterize the set to which the optimization dynamics converge, such switches should satisfy an average dwell-time constraint. In Chapter 5 (see also [76]), stability is studied for systems involving switches between multiple differential equations with distinct equilibria, with switches satisfying an average dwell-time condition. We extend the results from [76] and apply the provided asymptotic characterization method to consider switching between differential inclusions with distinct equilibria. The differential inclusions each take the form of the Hybrid-inspired Heavy Ball System from [77], which takes advantage of the differential inclusion to achieve efficiency comparable to acceler-

ated gradient methods while also retaining certain robustness properties. We characterize the set to which the resulting switched system converges, in terms of the Omega-limit set of an associated ideal hybrid system. This ideal hybrid system involves an automaton with solutions that are in one-to-one correspondence with time domains satisfying the average dwell-time constraint and with the rate parameter set to zero. Finally, we show that the system switching between differential inclusions with a small disturbance is a perturbed version of the mentioned ideal hybrid system with a globally asymptotically stable Omega-limit set. Thus, we can establish semi-global, practical asymptotic stability for the perturbed system. The robust stability of switched systems with distinct equilibria is a well-studied subject as discussed in the previous Chapter 5 (see [58],[61]). For example, [78] studies the steady-state optimization of switched systems with time-varying cost functions. We take advantage of the results from Chapter 5 and demonstrate our stability results on an application involving a data relay formed by UAVs, where the objective model's various performance measures such as communication quality or battery consumption. Sections 6.2, 6.3, 6.3.1, and 6.3.2 review some relevant ideas from Chapter 5.

6.1 Notations

In this work, $\mathbb{R}^n$ is used to demonstrate the n-dimensional Euclidean space. The $\mathbb{R}_{\geq 0}$ is used to show the nonnegative real numbers. We use $|x|$ to denote the Euclidean norm of the verctor $x \in \mathbb{R}^n$. Given $r > 0$, we use $r\mathbb{B}$ for the set $\{x \in \mathbb{R}^n : |x| \leq r\}$. For a function α, we say $\alpha \in \mathcal{K}^+$ if $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is continuous and strictly increasing.

6.2 Hybrid Systems

The hybrid systems framework that we use is described in (5.1) Chapter 5. We use $\mathcal{S}(K)$ to denote the *set of solutions* to (5.1) that start in K.

6.3 Preliminaries

In the first part of this section, we introduce some notions later used in this work and state some stability concepts for hybrid systems. In the second part of this section, we review the concept of average dwell-time constraint [79]. Next, we consider a perturbed hybrid system, in which the perturbation is parameterized by a dwell-time parameter satisfying the average dwell-time condition from [69]. Finally, we discuss the advantages of optimization methods employing differential inclusions for applications with persistent asset switches, introducing an efficient differential inclusion-based optimization algorithm inspired by the hybrid algorithms studied in [77]. This optimization algorithm has been shown in [77] to have desirable robustness properties and will be shown to be especially suitable for online optimization problems with persistent asset switches.

6.3.1 Stability concept for hybrid systems

We state some stability analysis concepts particularly for the hybrid system framework (5.1). The hybrid system (5.1) is said to be Lagrange stable if there exists $\alpha \in \mathcal{K}^+$ such that, for each $z_0 \in \mathbb{R}^n$, each $x \in \mathcal{S}(x_0)$ and $(t, j) \in \mathrm{dom}\,(x)$, we have $|x(t, j)| \leq \alpha(|x_0|)$. A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *stable* for the hybrid system (5.1) if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that $|x_\circ|_{\mathcal{A}} \leq \delta$, $x \in \mathcal{S}(x_\circ)$ and $(t, j) \in \mathrm{dom}(x)$ imply that $|x(t, j)|_{\mathcal{A}} \leq \varepsilon$. A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *attractive* for (5.1) if there exists $\delta > 0$ such that each solution $x \in \mathcal{S}(\mathcal{A} + \delta\mathbb{B})$ is bounded and, if complete, satisfies $\lim_{t+j\to\infty} |x(t, j)|_{\mathcal{A}} = 0$. The *basin of attraction* for an

attractive set $\mathcal{A}$ is the set of initial conditions from which each solution is bounded and, if complete, satisfies $\lim_{t+j\to\infty} |x(t,j)|_{\mathcal{A}} = 0$. A compact set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *asymptotically stable* for (5.1) if it is stable and attractive. It is said to be *globally asymptotically stable* for (5.1) if it is asymptotically stable with $\mathbb{R}^n$ as its basin of attraction. The set $\mathcal{A} \subset \mathbb{R}^n$ is said to be *semiglobally practically asymptotically stable in the parameter* $\delta > 0$ for the perturbed hybrid system if there exists $\beta \in \mathcal{KL}$ and, for each $\varepsilon > 0$ and $\Delta > 0$, there exists $\delta > 0$ such that each $x \in \mathcal{S}_\delta(\mathcal{A} + \Delta\mathbb{B})$ satisfies $|x(t,j)|_{\mathcal{A}} \leq \beta(|x(0,0)|_{\mathcal{A}}, t+j) + \varepsilon$ for all $(t,j) \in \mathrm{dom}(x)$.

6.3.2 Average dwell time switching and its automaton

Let a family of differential inclusions be given

$$\dot{z} \in \overline{\mathrm{co}} F_\sigma(z + \delta\mathbb{B}) + \delta\mathbb{B}, \qquad \sigma \in \Sigma. \tag{6.1}$$

Let $\Sigma := \{1, \cdots, M\}$ with M being a positive integer. The switching signal is denoted by $\sigma : \mathbb{R}_{\geq 0} \to \Sigma$ which satisfies the average dwell-time constraint parameterized by a small $\delta > 0$. We formalize the concept of average dwell time for the switching signal σ. Let $N_\sigma[s,t]$ denote the number of switches σ within the interval $[s,t]$ and N_0 being a positive integer. We then assume

$$N_\sigma(s,t) \leq \delta(t-s) + N_0 \qquad \forall\ 0 \leq s \leq t. \tag{6.2}$$

We cast a hybrid system, which associates the dynamics from differential inclusions given in (6.1) and hybrid time domains satisfying the constraint from (6.2). This hybrid system employs an automaton to capture the average dwell-time condition. The model of this hybrid system $\hat{\mathcal{H}}_\delta$ is given

$$\hat{\mathcal{H}}_\delta \begin{cases} (z,\sigma,\tau) \in C\times\Sigma\times[0,N_0] & \begin{cases} \dot{z} \in \overline{\text{co}} F_\sigma(z+\delta\mathbb{B})+\delta\mathbb{B} \\ \dot{\sigma} = 0 \\ \dot{\tau} \in [0,\delta] \end{cases} \\ (z,\sigma,\tau) \in C\times\Sigma\times[1,N_0] & \begin{cases} z^+ = z \\ \sigma^+ \in \Sigma\setminus\{\sigma\} \\ \tau^+ = \tau - 1. \end{cases} \end{cases} \tag{6.3}$$

According to [69, Proposition 1.1], the solutions to $\hat{\mathcal{H}}_\delta$ are in a one-to-one correspondence with the solutions of (6.1) under the average dwell-time switching constraint (6.2) and $\hat{\mathcal{H}}_\delta$ satisfies the hybrid basic conditions from [27, Assumption 6.5]. Note that, for convenience, the parameter $\delta > 0$ describes both the maximum flow rate of the automaton τ and the perturbed differential inclusion.

6.3.3 Ideal system analysis

The goal is to extend the results from Chapter 5 and [76] to differential inclusions and characterize the asymptotic behavior of (6.1) under average dwell-time constraint parametrized by a small δ. In order to realize this goal, we characterize the asymptotic behavior of the $\hat{\mathcal{H}}_\delta$ without the perturbation, which results from setting $\delta = 0$ in (6.3). This new ideal hybrid

system obtained by setting $\delta = 0$ is denoted by $\hat{\mathcal{H}}_0$ and corresponds to

$$\hat{\mathcal{H}}_0 \begin{cases} (z,\sigma,\tau) \in C\times\Sigma\times[0,N_0] & \begin{cases} \dot{z} \in F_\sigma(z) \\ \dot{\sigma} = 0 \\ \dot{\tau} = 0 \end{cases} \\ (z,\sigma,\tau) \in C\times\Sigma\times[1,N_0] & \begin{cases} z^+ = z \\ \sigma^+ \in \Sigma\setminus\{\sigma\} \\ \tau^+ = \tau - 1. \end{cases} \end{cases} \tag{6.4}$$

As we will see, the asymptotic behavior of (6.4) approximates the asymptotic behavior of the (6.3).

6.3.4 A differential inclusion-based optimization algorithm

Differential inclusions have been useful for studying stability of steepest descent/ascent dynamics in convex optimization [80]. Differential inclusions are also discussed in [81], in which a continuous-time analogue of the Alternating Direction Method-of-Multipliers is given. In [82] and [83], differential inclusions are used to approximate a high-gain anti-windup strategy for handling input constraints in feedback-based optimization. In this section, we focus on an algorithm, represented as a differential inclusion, whose trajectories seek a solution to the problem

$$\min_{q\in\mathbb{R}^n} \left\{ \phi(q) \coloneqq \sum_{i=1}^{n} \phi_i(q_i) \right\} \tag{6.5a}$$

$$\text{s.t.} \quad \mathbb{1}^T q = d, \quad d \in \mathbb{R}, \tag{6.5b}$$

under the following assumptions.

Assumption 9 *The objective $\phi : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, has compact sub-level sets, has an L-Lipschitz gradient $\nabla\phi$, and is convex.*

Under Assumption 9, [84, Prop. 5.3.7] implies that the set Q^* of solutions to (6.5) is non-empty and compact. Furthermore, [84, Prop. 5.3.3] implies that $q^* \in Q^*$ if and only if there exists $\mu^* \in \mathbb{R}$ such that

$$\nabla\phi(q^*) + \mu^*\mathbb{1} = 0, \quad \mathbb{1}^T q^* = d. \tag{6.6}$$

Defining $\mathcal{L}$ as the Laplacian of a connected undirected graph, we have that the nullspace of $\mathcal{L}$ is span($\mathbb{1}$) [85, Sec. II], and thus, the conditions (6.6) can be equivalently expressed as

$$\mathcal{L}\nabla\phi(q^*) = 0, \quad \mathbb{1}^T q^* = d. \tag{6.7}$$

It is convenient to express the above conditions in terms of $\mathcal{L}$ because $\mathcal{L}$ will play an important role in the analysis and application of our proposed algorithm. Our algorithm is based on the following differential equation, which we refer to as the *Laplacian-Gradient Heavy Ball Method* (HBM) and is a continuous-time analogue of the algorithm in [86]. The HBM is given by

$$\dot{x} = \begin{bmatrix} p \\ -Kp - \mathcal{L}\nabla\phi(q) \end{bmatrix}. \tag{6.8}$$

To achieve fast convergence without oscillations, first-order convex optimizations methods often require knowledge of problem parameters for a precise algorithmic parameter tuning. Theoretically, convergence of iterative optimization algorithms for convex problems can be achieved through momentum in the sense of Nesterov's method [87]. However, often the algorithmic parameters such as momentum and stepsize should be accurately specified depending on the

problem parameters. When such precise parameters are not known, momentum methods such as Nesterov's method suffer from oscillations in their trajectories. Such oscillations exhibited by Nesterov's method often restrict its application to real-world systems [88]. Therefore, we consider the differential inclusion, based on HBM and inspired by the hybrid algorithm in [89], which can greatly reduce oscillations without precise algorithmic parameter tuning. This algorithm is referred to as the *Laplacian-Gradient Hybrid-inspired Heavy Ball Method* (HiHBM). This approach using a differential inclusion can also improve transient performance in settings where the objective is persistently switching during the algorithm execution. In this system, the state is denoted as $x := (q, p)$. The parameters $\{\underline{K}, \overline{K}\} \in \mathbb{R}^2$ satisfy $0 < \underline{K} \leq \overline{K}$. The system is defined as

$$\dot{x} \in F(x) := \begin{bmatrix} p \\ -\kappa(x)p - \mathcal{L}\nabla\phi(q) \end{bmatrix}, \tag{6.9a}$$

$$\begin{aligned} \kappa(x) &:= \kappa(x; \underline{K}, \overline{K}) \\ &:= \begin{cases} \overline{K} & \text{if } \langle \mathcal{L}\nabla\phi(q), p\rangle > 0, \\ \underline{K} & \text{if } \langle \mathcal{L}\nabla\phi(q), p\rangle < 0, \\ \left[\underline{K}, \overline{K}\right] & \text{if } \langle \mathcal{L}\nabla\phi(q), p\rangle = 0, \end{cases} \end{aligned} \tag{6.9b}$$

$$C := \{(q, p) \in \mathbb{R}^{2n} : \ \mathbb{1}^T q = d \ \& \ \mathbb{1}^T p = 0\}. \tag{6.9c}$$

To establish the following global asymptotic stability property of HiHBM, it will be convenient to write $\mathcal{F}_d$ to denote the feasible set defined by (6.5b). That is, we have

$$\mathcal{F}_d := \{q \in \mathbb{R}^n : \mathbb{1}^T q = d\}, \tag{6.10}$$

and the set defined in (6.9c) can be written $C = \mathcal{F}_d \times \mathcal{F}_0$.

Theorem 7 *Under Assumption* (9)*, the set* $\mathcal{A} := \mathcal{Q}^* \times \{0\}$ *is GAS for the system* (6.9).

Proof: Let $\mathcal{L}^\dagger$ denote the generalized inverse Laplacian [90], which can be shown to be a symmetric positive semi-definite matrix of rank $n-1$ satisfying

$$\mathcal{L}^\dagger \mathbb{1} = 0, \qquad \mathcal{L}\mathcal{L}^\dagger = I_n - \frac{1}{n}\mathbb{1}\mathbb{1}^T. \tag{6.11}$$

We aim to apply [27, Thm. 8.2]. Toward this goal, consider

$$V(q,p) \coloneqq \phi(q) - \phi^* + \frac{1}{2}p^T\mathcal{L}^\dagger p, \tag{6.12}$$

$$\phi^* \coloneqq \min_{w\in\mathcal{F}_d} \phi(w). \tag{6.13}$$

The map $q \mapsto \phi(q) - \phi^*$ is positive definite on C with respect to $\mathcal{Q}^*$ because $q \in \mathcal{F}_d$ for all points in C. The map $p \mapsto p^T\mathcal{L}^\dagger p$ is positive definite on C with respect to $\{0\}$, which follows from (6.11) and the fact that $p \in \mathcal{F}_0$ for all points in C. By Assumption (9), V is radially unbounded with respect to $\mathcal{A}$ relative to C. The Lie derivative of V with respect to (6.9a) satisfies

$$\begin{aligned}
&\langle \nabla\phi(q),\ p\rangle - \left\langle \mathcal{L}\nabla\phi(q),\ \mathcal{L}^\dagger p\right\rangle - \kappa(x)p^T\mathcal{L}^\dagger p \\
&= p^T\left(\frac{1}{n}\mathbb{1}\mathbb{1}^T\right)\nabla\phi(q) - \kappa(x)p^T\mathcal{L}^\dagger p = -\kappa(x)p^T\mathcal{L}^\dagger p \le 0,
\end{aligned}$$

where the first equality follows from (6.11), and the second equality follows from the fact that $\mathbb{1}^T p = 0$. Stability of $\mathcal{A}$ has now been shown. To show attractivity, first recall that $p \mapsto p^T\mathcal{L}^\dagger p$ is positive definite when restricted to $\mathcal{F}_0$, and thus, $\kappa(x)p^T\mathcal{L}^\dagger p$ can remain at zero for all time only if p remains at zero. However, if p remains at zero, then (6.9a) implies that both $\dot{q}$ and $\mathcal{L}\nabla\phi(q)$ must remain at zero, which can happen only if q remains in $\mathcal{Q}^*$, due to (6.7). In summary, $\kappa(x)p^T\mathcal{L}^\dagger p$ remains at zero only if (q,p) remains in $\mathcal{A}$, and the result then follows from [91, Thm. 2.11]. ■

6.4 Online optimization with persistent switches

In this section, we discuss an optimization problem that models a team of drones collectively executing a task. The drones are the assets that persistently switch in this scenario, due to each drone's limited battery life, potential physical damages, or other disabling aspects. Consider a team of n drones with the collective task of forming a relay to transmit data from a source to a destination. The drones form a straight path of length $d \in \mathbb{R}_{>0}$ from the first to n-th drone. The network-wide state vector is denoted by $q \in \mathbb{R}^n$, where q_i is the relative distance from drone i to the node that precedes it, while the data source is considered to precede the first drone. We assume that the drones 1 to n never cross over their neighboring drones, and their ordering remains the same during their movement. In practice, enforcement of a box constraint on each q_i is needed, which can be done using the penalty-based approach in [92], but the details are beyond the scope of this work. The problem can be modeled as a variant of the problem (6.5), in which the objective switches according to a switching signal $\sigma : \mathbb{R}_{\geq 0} \to \Sigma$:

$$\min_{q \in \mathbb{R}^n} \left\{ \phi_{\sigma(t)}(q) \coloneqq \sum_{i=1}^{n} \phi_{i,\sigma(t)}(q_i) \right\} \tag{6.14a}$$

$$\text{s.t.} \quad \mathbb{1}^T q = d. \tag{6.14b}$$

In a similar setting to (6.1), the switching signal satisfies the average dwell-time constraint parametrized by a small δ.

Assumption 10 *For each constant $\sigma(.) \in \Sigma$, the function $q \mapsto \phi_\sigma(q)$ satisfies the conditions of Assumption 9 and there exists a unique solution q^* to* (6.14).

Theorem 8 *If Assumptions 9 and 10 hold then the Ω-limit set associated with the ideal hybrid*

system $\hat{\mathcal{H}}_0$ *given in (6.4) with* $z = (q, p)$ *and with*

$$F_\sigma(z) := \begin{bmatrix} p \\ -\kappa(z)p - \mathcal{L}\nabla\phi_\sigma(q) \end{bmatrix},$$

$$:= \begin{cases} \overline{K} & \text{if } \langle \mathcal{L}\nabla\phi_\sigma(q), p\rangle > 0, \\ \underline{K} & \text{if } \langle \mathcal{L}\nabla\phi_\sigma(q), p\rangle < 0, \\ \left[\underline{K}, \overline{K}\right] & \text{if } \langle \mathcal{L}\nabla\phi_\sigma(q), p\rangle = 0, \end{cases} \tag{6.15}$$

$$C := \{(q, p) \in \mathbb{R}^{2n} : \ \mathbb{1}^T q = d \ \& \ \mathbb{1}^T p = 0\}, \tag{6.16}$$

is semiglobally, practically asymptotically stable in the parameter $\delta > 0$ *for the perturbed system* $\hat{\mathcal{H}}_\delta$ *given in (6.3).*

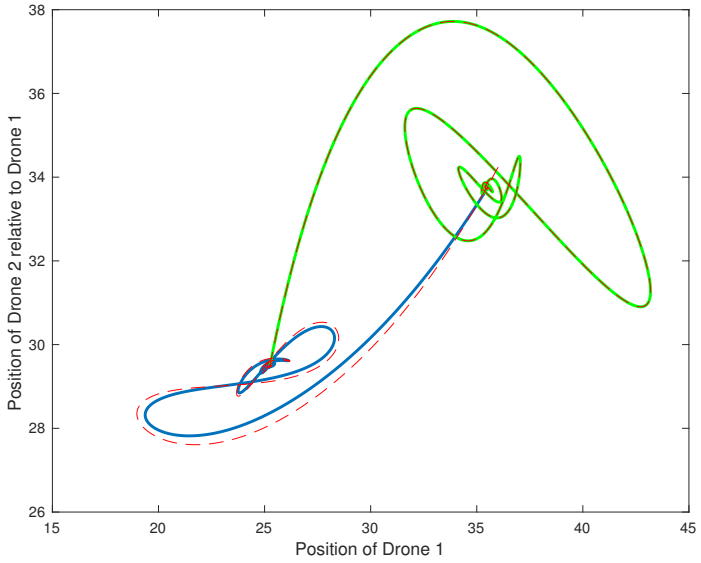

Figure 6.1: Steady State Behavior near Omega-limit set

The proof of Theorem 8 is given in Section 6.6.

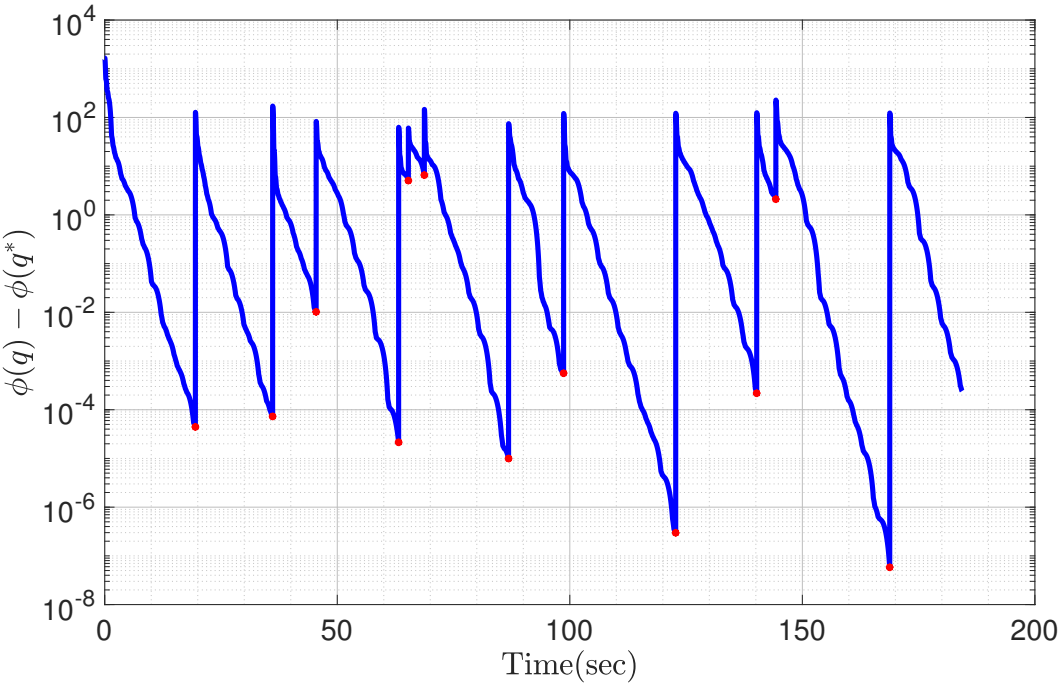

Figure 6.2: Optimization with persistent switches

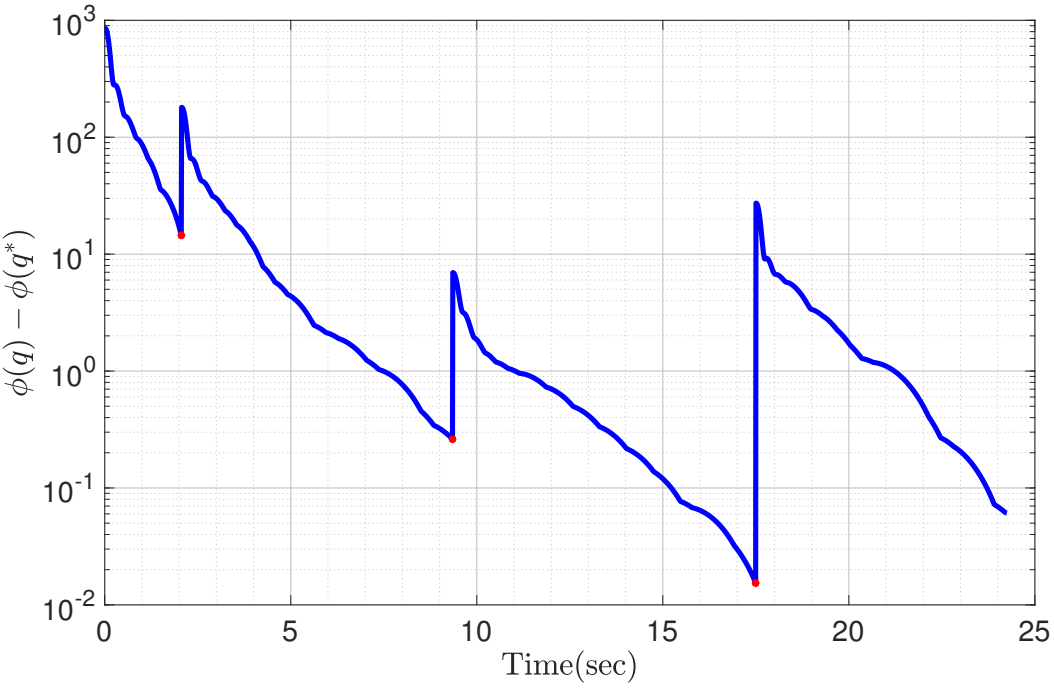

Figure 6.3: Optimization with fewer asset switches

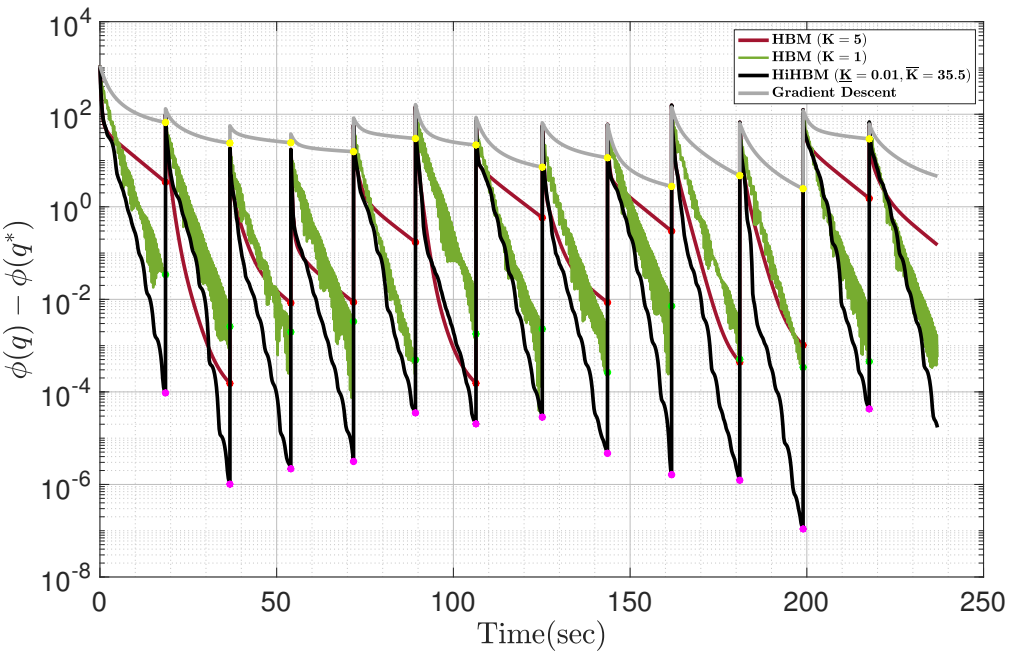

Figure 6.4: Comparison of HBM (6.8) and HiHBM (6.9)

6.5 Numerical Example

In this section, we give numerical examples for the application and the optimization method discussed in Sections 6.4 and 6.3.4 respectively. In Figure 6.1, we consider a team of 2 drones in a relay of distance $d = 100$, and we consider $N_0 = 1$ in order to demonstrate the Ω-limit set. The blue line represents the position of the first drone (initial condition: distance of 35.5071 units from the data source), and the green line represents the position of a second drone relative to the first drone in the relay (initial condition: distance of 33.7398 units from the first drone). These two lines depict the Ω-limit set of the ideal hybrid system given in (6.4). The red lines are created by allowing for the small dwell-time parameter $\delta = 0.0338$. The switches satisfy the average dwell-time condition. Considering the switching behavior to be a small disturbance as in Theorem 8, we observe that the solutions converge to a small neighborhood of the Ω-limit set of the ideal hybrid system. These observations agree with the results from Theorem 8. Figure

6.2 shows the decrease in the value of the objective function under the dynamics of the HiHBM algorithm applied to the problem (6.14) for 20 drones in a relay of length $d = 100$, with the network-wide objective being $\phi_\sigma(q) = \frac{1}{2}q^T P_\sigma q + b_\sigma^T q$ for $\sigma \in \{1, 2\}$. Each P_σ is diagonal with eigenvalues i.i.d. uniform on $[10, 20]$, and each b_σ has its entries i.i.d. uniform on $[-10, 10]$. The objective at each drone is $\phi_{i,\sigma}(q_i) = \frac{1}{2}P_{\sigma,ii}q_i^2 + b_{\sigma,i}q_i$, where q_i is the distance from drone i to drone $i-1$, as described in Sec. 6.4. We allow persistent switches satisfying average dwell time with $\delta = 0.06$. Figures 6.2 shows that HiHBM converges efficiently even as the objective switches persistently. Figure 6.3 shows the same example with fewer switches. In Figure 6.4, the gray line displays the gradient descent optimization. The red line demonstrates the convergence of the HBM optimization algorithm with $K = 5$ and the green line demonstrates the same optimization algorithm with $K = 1$. The black line in Figure 6.4 shows the HiHBM optimization algorithm with a faster, more efficient convergence. The lower and upper-bound for HiHBM are 0.01 and 35.5 respectively. In Figure 6.4, we see the efficiency of the HiHBM algorithm that generates smaller errors in comparison with the simple gradient descent and the HBM method.

6.6 The proof of theorem 8

6.6.1 A new result on switching between constrained inclusions

Motivated by applications requiring persistent asset switches while optimizing, in this subsection we provide results on characterization of the asymptotic behavior that results from persistent switches among asymptotically stable differential inclusions with distinct equilibria. The switching signal is considered to satisfy an average dwell-time constraint as mentioned in the preliminaries 6.3.2.

Assumptions

We have two assumptions regarding the following family of differential inclusions parametrized by $\sigma \in \Sigma$ given as

$$\dot{z} \in F_\sigma(z) \quad z \in C. \tag{6.17}$$

Assumption 11 *For each $\sigma \in \Sigma$, F_σ is outer semi-continuous, locally bounded relative to $C \subset \text{dom}\, F_q$, and, for each z, $F_\sigma(z)$ is non-empty and convex for all values of $z \in C$. Furthermore, the point z^*_σ is globally asymptotically stable for (6.17).*

Assumption 12 *For each $\sigma \in \Sigma$, the only solution to*

$$\begin{cases} z \in C, \quad \dot{z} \in -F_\sigma(z) \end{cases} \tag{6.18}$$

with an initial value of $z(0) = z^*_\sigma$ is $z(t) = z^*_\sigma$ for all $t \geq 0$.

Assumption 12 extends the Assumption 8 from Chapter 5 (see [76, Assumption 2]) from differential equations to differential inclusions.

6.6.2 Extended Main Result

In this section, we adapt the results from Chapter 5 and [76] to the hybrid system $\hat{\mathcal{H}}_\delta$. First, we establish boundedness for solutions to $\hat{\mathcal{H}}_0$ by claiming the following proposition and corollary.

Proposition 3 *If, for each $\sigma \in \Sigma$, the system (6.17) is Lagrange stable, the hybrid system $\hat{\mathcal{H}}_0$ is Lagrange stable.*

The proof of Proposition 3 follows the boundedness results from [76, Section 4.2] for the ideal hybrid system $\hat{\mathcal{H}}_0$. As a consequence of Proposition 3 we have the following corollary.

Corollary 2 *If Assumption 11 holds then the hybrid system $\hat{\mathcal{H}}_0$ is Lagrange stable.*

Using the definition of Ω-limit set from [76], the rest of this section is devoted to the characterization of the Ω-limit set of $\hat{\mathcal{H}}_0$ from a compact set $K \subset \mathbb{R}^{n+2}$ denoted by $\Omega_0(K)$. We use the definition of reachable sets from [76] and demonstrate the reachable set from K with $R_0(K)$. Following the setting from [76], we define

$$S_\sigma := \bigcap_{j \in \mathbb{Z}_{\geq 0}} \overline{R_0\left(\left(\{z_\sigma^*\} + \frac{1}{j+1}\mathbb{B}\right) \times \{\sigma\} \times [0, N_0]\right)} \tag{6.19a}$$

$$S := \bigcup_{\sigma \in \Sigma} S_\sigma. \tag{6.19b}$$

Lemma 8 *Under Assumption 11 and as a result of Corollary 2, the set S is compact.*

Proposition 4 *Under Assumptions 11 and 12, for each compact set K with $(\bigcup_{\sigma\in\Sigma}\{z_\sigma^*\} \times \{\sigma\}) \times [0, N_0]$ in its interior, $\Omega_0(K) = S$.*

This result follows from the following Lemmas.

Lemma 9 *If Assumption 11 holds then, for each compact set $K \subset \mathbb{R}^{n+2}$, $\Omega_0(K) \subset S$.*

Lemma 10 *If Assumptions 11 and 12 hold then, for each compact set $K \subset \mathbb{R}^{n+2}$ containing the set*

$$K_0 := \left(\bigcup_{\sigma\in\Sigma}\{z_\sigma^*\} \times \{\sigma\}\right) \times [0, N_0] \tag{6.20}$$

in its interior, $S \subset \Omega_0(K)$.

The proof of Lemmas 9 and 10 follows the same lines as the proof for unconstrained differential equations given in Chapter 5 and [76, Section 4.3]. Since the required changes to the proof are minimal the details are omitted. We can finally state our extension of the main results according to [76].

Theorem 9 *Under Assumptions 11 and 12, the set S defined in (6.19b) is semi-globally, practically asymptotically stable in the parameter $\delta > 0$ for the system $\hat{\mathcal{H}}_\delta$.*

6.6.3 Verifying that the systems in Section III satisfy the assumptions of Theorem 9

Theorem 10 *Under Assumptions 9 and 10, and for $\sigma \in \Sigma$, HiHBM with objective function ϕ_σ satisfies Assumption 12.*

Proof: Under the given assumptions, for each constant σ, let $z_\sigma^* := (q_\sigma^*, 0)$, where q_σ^* is the unique solution of (6.14) be the initial value to $\dot{z} \in -F_\sigma(z)$. Let z be also a solution to (6.18). There exist $L \geq 0$ and $r > 0$ such that $|F_\sigma(z) - F_\sigma(z_\sigma^*)| \leq L|z - z_\sigma^*|$ for all $z \in \{z_\sigma^*\} + r\mathbb{B}$. Then, with $e := z - z_\sigma^*$ and noting that $e(0) = 0$ and $\langle e, \dot{e} \rangle \leq L|e|^2$, it follows from standard comparison theorems that $e(t) = 0$ for all $t \geq 0$. ■

Theorem 11 *Under Assumptions 9 and 10, and for $\sigma \in \Sigma$, HiHBM with objective function ϕ_σ satisfies Assumption 11.*

Proof: For each constant $\sigma \in \Sigma$, let q_σ^* denote the solution to problem (6.14). Then, global asymptotic stability of $z_\sigma^* := (q_\sigma^*, 0)$ for HiHBM with objective ϕ_σ follows from Theorem 7 by setting $\mathcal{Q} = \{q_\sigma^*\} \times \{0\}$ for each σ. ■

Proof: We now prove Theorem 8. Suppose Assumptions 9 and 10 hold. It follows from Theorems 10 and 11 that Assumptions 11 and 12 hold. From Theorem 9, it follows that the

set S in (6.19b) is semi-globally practically asymptotically stable for the system (6.3)-(6.5) embedded with an average dwell-time automaton as in (6.3) with respect to the parameter δ. Finally, according to Proposition 4, the set S is the Ω-limit set indicated in Theorem 8. ■

6.7 Concluding Remarks

We discussed the importance of enabling persistent switches of objective functions during an online execution of an optimization algorithm. The reason for these persistent switches is motivated by many engineering applications of convex optimization. We further extend the existing results of Chapter 5 (see also [76]) on differential equations with multiple equilibria to differential inclusions with distinct multiple equilibria. We present the asset switches occurring in an optimization problem consisting of drones in a relay aiming for maximizing the signal strength. The asset switches are analyzed through the use of average dwell-time parameter which determines the rate of objective's switching during the online HiHBM optimization. Hence, we establish semi-global practical asymptotic stability of a certain set with respect to this parameter. We characterize this set via Omega-limit sets.

Part III

Soft-reset systems

Chapter 7

Input-to-state stability of soft-reset systems with nonlinear data

In this chapter, we investigate the ISS property for a class of reset control systems. The concept of input-to-state stability (ISS) property was introduced in the 1989 paper [93]. This property became a bedrock for analyzing nonlinear dynamical systems with inputs, including to establish closed-loop stability results based on the ISS nonlinear small-gain theorem [94]. Though most nonlinear control algorithms are based on either continuous-time systems or discrete-time systems, reset control systems employ hard state resets that lead to a mixture of continuous-time and discrete-time dynamics. Reset systems reset their states or a subset of their states to zero based on a determined condition. Typically, in a closed-loop control system, resets occur in the controller state rather than the plant state. Reset systems have various applications in control systems. A reset integrator, also referred to as a Clegg integrator (CI) [95], is one of the earliest examples of a reset system. This integrator circuit has a describing function similar to the frequency response of a linear integrator but with a different phase lag. Later, in the 1970's, Horowitz and co-authors attempted to provide a systematic approach for designing reset controllers in order to add flexibility in linear controller designs and eliminate the limitations of

Clegg's integrator [96],[97]. In the 1990's, Hollot presented performance and stability analysis for general linear reset control systems [98],[99],[100]. Later, Beker and co-authors achieved controller design specifications that demonstrated some advantages of reset controllers over linear controllers in [101], [102], and [103]. In the last two decades, researchers have begun to approach reset systems from a hybrid dynamical systems point of view [104]. Due to to the nature of continuous/discrete interplay, reset systems can be modeled within the framework of hybrid dynamical systems in the sense of [105], [106], and [4]. Looking at reset controllers from the hybrid systems point of view has led to a more rigorous analysis on stability and robustness of reset systems; see, for example, [107] and [108]. In many practical instances it has been shown that embedding reset-like behavior in an otherwise continuous controller results in desirable performance; see, for example, [109] and [110]. However, due to the difficulty in providing rigorous analytical tools that can handle instantaneous changes of system solutions and a lack of performance analysis in nonlinear settings, reset systems have not been applied extensively.

Recently, attempts have been made to broaden the applicability of reset systems by introducing an alternative "soft-reset" implementation, which does not require the framework of hybrid systems but instead is modeled by a differential inclusion. The idea was introduced in [111] for linear reset systems without inputs. Follow-up work related to numerical verification of stability for these soft-reset systems is given in [112]. Soft resets were studied from a passivity point of view for nonlinear reset systems with inputs in [113].

In this work, we study the soft-reset implementation of hard-reset systems having nonlinear data, establishing conditions for ISS of the soft-reset system. We begin in Section 7.2 by exploiting strong convexity of a Lyapunov function for the hard-reset system, as done in [111]. However, in contrast with [111], which studies global asymptotic stability given linear data, we study ISS given nonlinear data that satisfy certain local sector growth conditions. Then, in Section 7.3, under a different sector growth condition, we infer ISS of the soft-reset system

using global asymptotic stability of the origin for the zero-input system. We do this using two approaches, one involving strong Lyapunov conditions and the other involving homogeneity of the system data. In Section 7.4, by assuming a strongly convex weak Lyapunov function for the zero-input hard-reset system, we establish global exponential stability of the origin of the zero-input soft-reset system, using homogeneity of both the data and the Lyapunov function. We then use this result to claim that the conditions for ISS in Section 7.3 are satisfied for the soft-reset system with nonzero inputs.

Lastly, some numerical examples on asymptotic stability and ISS of a soft-reset system in a closed loop are provided.

7.1 Notation

The set of (nonnegative) real numbers is denoted by ($\mathbb{R}_{\geq 0}$) $\mathbb{R}$. The set of (nonnegative) integers is denoted by ($\mathbb{Z}_{\geq 0}$) $\mathbb{Z}$. For any two vectors $u, v \in \mathbb{R}^n$, we use $\langle u, v\rangle \coloneqq u^T v$. For $x \in \mathbb{R}^n$, we use $|x| \coloneqq \sqrt{\langle x, x\rangle}$. We denote by $\mathcal{G}$ the set of functions from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}_{\geq 0}$ that are continuous, nondecreasing, and zero at zero. The subset of strictly increasing functions in $\mathcal{G}$ is denoted by $\mathcal{K}$. The subset of unbounded functions in $\mathcal{K}$ is denoted by $\mathcal{K}_\infty$. Moreover, $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is said to belong to class $\mathcal{KL}$ if $\beta(\cdot, s)$ belongs to class $\mathcal{K}$ for each $s \geq 0$, and for each fixed $r \geq 0$, the mapping $\beta(r, \cdot)$ is decreasing to zero. A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *sector bounded near the origin* if there exist $\delta > 0$ and $L > 0$ such that $|f| \leq L|z|$ for all $z \in \mathbb{R}^n$ satisfying $|z| \leq \delta$ and all $f \in F(z)$. It is said to be *quadratically bounded near the origin* if there exist $\delta > 0$ and $L > 0$ such that $|f| \leq L|z|^2$ for all $z \in \mathbb{R}^n$ satisfying $|z| \leq \delta$ and all $f \in F(z)$. It is said to be homogeneous of degree $k \in \mathbb{Z}_{\geq 0}$ if $F(\lambda x) = \lambda^k F(x)$ for all $x \in \mathbb{R}^n$ and $\lambda > 0$. We use C^1 for any function that is continuously differentiable. A C^1 function

$V : \mathbb{R}^n \to \mathbb{R}$ is called *strongly convex* if there exists a $\mu > 0$ such that, for all $x, y \in \mathbb{R}^n$, we have

$$V(y) \geq V(x) + \langle \nabla V(x), y - x \rangle + \mu |x - y|^2. \tag{7.1}$$

Given $x \in \mathbb{R}^n$ and a nonempty set $\mathcal{A} \subset \mathbb{R}^n$, the distance of x to $\mathcal{A}$ is denoted $|x|_{\mathcal{A}}$ and is defined by $|x|_{\mathcal{A}} := \inf_{y \in \mathcal{A}} |x - y|$. The origin of the system $\dot{x} \in F(x)$ is said to be *(Lyapunov) stable* if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that $|x(0)| \leq \delta$ implies $|x(t)| \leq \varepsilon$ for all $t \geq 0$. It is said to be *globally attractive* if every solution x satisfies $\lim_{t \to \infty} |x(t)| = 0$. It is said to be *globally asymptotically stable (GAS)* if it is both stable and globally attractive. It is said to be *globally exponentially stable (GES)* if there exist positive constants c_0 and c_1 such that every solution x satisfies $|x(t)| \leq c_0 |x(0)| \exp(-c_1 t)$ for all $t \geq 0$. The system $\dot{x} \in F(x, d)$ is said to be *input-to-state stable (ISS)* if, for each locally essentially bounded input d, maximal solutions are defined on $[0, \infty)$ and there exist a class $\mathcal{KL}$ function β and a class $\mathcal{G}$ function γ such that every solution x satisfies

$$|x(t)| \leq \beta(|x(0)|, t) + \gamma(\|d\|_\infty) \qquad \forall t \geq 0. \tag{7.2}$$

7.2 ISS for nonlinear soft-reset systems

A hard-reset system with input is a hybrid dynamical system in the modeling framework of [27] with state $x \in \mathbb{R}^n$ and an external disturbance $d \in \mathbb{R}^m$ given as:

$$x \in C := \{x \in \mathbb{R}^n : \varphi(x) \leq 0\}, \qquad \dot{x} \in \widehat{F}(x, d), \tag{7.3a}$$

$$x \in D := \{x \in \mathbb{R}^n : \varphi(x) \geq 0\}, \qquad x^+ = g(x). \tag{7.3b}$$

The set C indicates where the continuous change of the state is allowed, and the set D indicates where the instantaneous change of the state is allowed. Continuous change is governed by the input-driven differential inclusion in (7.3a) while instantaneous change is governed by the difference equation in (7.3b). The mappings that define these entities have the following properties:

Assumption 13

1. $\widehat{F} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is outer semi-continuous (that is, its graph is closed) and locally bounded with nonempty, convex values.

2. $g : D \to \mathbb{R}^n$ is continuous.

In this chapter, we are interested in the ISS property for reset, or reset-inspired, systems with inputs. If we were to consider ISS for the hard-reset system (7.3), a potential Lyapunov condition would be the existence of a C^1 function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that admits a continuous, positive definite function $\widehat{W} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, functions $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, and a function $\gamma \in \mathcal{K}_\infty$ such that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|), \qquad \forall x \in \mathbb{R}^n, \tag{7.4}$$

and

$$x \in C,\ |x| \geq \gamma(|d|),\ \widehat{f} \in \widehat{F}(x,d) \quad \Longrightarrow \quad \left\langle \nabla V(x), \widehat{f} \right\rangle \leq -\widehat{W}(x), \tag{7.5a}$$

$$x \in D \quad \Longrightarrow \quad V(g(x)) \leq V(x). \tag{7.5b}$$

However, there is a problem with this condition. In particular, (7.5b) does not rule out the possibility of solutions that exclusively jump without decreasing the Lyapunov function V. To avoid having to worry about such solutions, it is of interest to recast the hard-reset system

(7.3) via its "soft-reset" implemention, first introduced in [111] for linear reset systems without inputs. This implementation corresponds to the differential inclusion

$$\dot{x} \in \widehat{F}(x,d) + \kappa(x)\Big(\mathrm{SGN}\left(\varphi(x)\right) + 1\Big)(g(x) - x) =: F(x,d), \tag{7.6}$$

where $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ and the set-valued mapping SGN is defined as

$$\mathrm{SGN(s)} := \begin{cases} \frac{s}{|s|} & s \neq 0 \\ [-1,1] & s = 0. \end{cases} \tag{7.7}$$

Lemma 11 *If Assumption 13 holds and $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ is continuous then the set-valued mapping F defined in (7.6) is outer semicontinous and locally bounded with nonempty, convex values.*

Proof: (Sketch) See [114, Prop. 2.23(a), Prop. 5.51(a),(b)]. ■

As we will show below, the Lyapunov conditions (7.4)-(7.5) guarantee ISS of (7.6) at least for $\kappa(\cdot)$ sufficiently large if the following additional conditions hold:

Assumption 14

1. The function V in (7.4)-(7.5) is strongly convex.
2. There exists a $M = M^T$ such that $\varphi(x) = x^T M x$ for all $x \in \mathbb{R}^n$.
3. $\widehat{F} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ sector bounded near the origin, $g : D \to \mathbb{R}^n$ is sector bounded near the origin, and $\widehat{W}$ is quadratically bounded near the origin.
4. Jumps starting in the jump set $x \in D$ land in the flow set i.e., $g(x) \in C$ for all $x \in D$.

5. The function γ in (7.5) belongs to $\mathcal{K}_\infty$ and satisfies

$$\limsup_{s\to 0^+} \frac{s}{\gamma(s)} < \infty. \tag{7.8}$$

6. The condition (7.5a) holds with C replaced by the inflated set $C_\varepsilon := \left\{x \in \mathbb{R}^n : \varphi(x) \leq \varepsilon x^T x\right\}$ for some $\varepsilon > 0$.

The condition (7.8) can always be satisfied by, for example, adding a linear term to γ near the origin. By assuming Item 4 of Assumption 14, we allow for solutions of the hard-reset system (7.3) to flow without immediately jumping after a prior jump, but this does not remove the purely discrete-time solutions to (7.3) that do not decrease the function V. The strong convexity of V in Item 1 of Assumption 14 enables establishing ISS of the soft-reset system (7.6), as in the following theorem:

Theorem 12 *The conditions (7.4)-(7.5), augmented with the conditions in Assumptions 13 and 14, guarantee that there exists a continuous function $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ with sufficiently large values and a continuous, positive definite function $W : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that $|x| \geq \gamma(|d|)$ implies $\langle \nabla V(x), f\rangle \leq -W(x)$ for all $f \in F(x, d)$.*

Proof: Since the function g is assumed to be continuous and sector bounded near the origin, there exists a continuous function $\sigma_0 : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that is positive definite and sector bounded near the origin satisfying

$$|M(g(x) + x)| \leq \sigma_0(x), \quad \forall x \in D. \tag{7.9}$$

Using the inequality from (7.9) and the Cauchy-Schwarz inequality, $M = M^T$, and Item 4 of

Assumption 14, giving $g(x)^T M g(x) \leq 0$ when $x^T M x \geq 0$, it follows that

$$\begin{aligned} x \neq 0,\ x^T M x \geq 0 \quad \implies \\ |g(x) - x| \geq \frac{-(g(x)-x)^T M (g(x)+x)}{\sigma_0(x)} = \frac{x^T M x - g(x)^T M g(x)}{\sigma_0(z)} \\ \geq \frac{x^T M x}{\sigma_0(x)}. \end{aligned} \tag{7.10}$$

By the strong convexity of V, as in (7.1), from Item 1 of Assumption 1, we can write

$$V(g(x)) \geq V(x) + \langle \nabla V(x), g(x) - x \rangle + \mu |x - g(x)|^2. \tag{7.11}$$

From (7.11) and (7.5b), we have:

$$x \in D \quad \implies \quad \langle \nabla V(x), x - g(x) \rangle \geq \mu |x - g(x)|^2. \tag{7.12}$$

It then follows from the definition of the set D in (7.3b) and the definition of SGN below (7.6) that (7.12) can be rewritten as

$$\begin{aligned} s \in \mathrm{SGN}(x^T M x) \quad \implies \\ \langle \nabla V(x), (s+1)(g(x) - x) \rangle \leq -(s+1)\mu |x - g(x)|^2. \end{aligned} \tag{7.13}$$

Combining (7.13), (7.10), and Assumption 1 gives

$$\begin{aligned} x \neq 0,\ s \in \mathrm{SGN}(x^T M x) \quad \implies \\ \langle \nabla V(x), (s+1)(g(x) - x) \rangle \leq -2\mu \max\{0, x^T M x\} \frac{x^T M x}{\sigma_0(x)^2}. \end{aligned} \tag{7.14}$$

Due to Item 5 of Assumption 1, there exist $\varepsilon, m > 0$ such that $s/\gamma(s) \leq m$ for all $s \in (0, \varepsilon]$. We will show that there exists a continuous function $\sigma : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that is quadratically bounded

near the origin such that, for all $x \in \mathbb{R}^n$ and $|d| \leq \varepsilon$, we have

$$\widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|) \geq \frac{|d|}{m} \quad \Longrightarrow \quad \left\langle \nabla V(x), \widehat{f} \right\rangle + \widehat{W}(x) \leq \sigma(x). \tag{7.15}$$

Due to Item 6 of Assumption 14, $\sigma(x)$ may take any nonnegative values for $x \in C_\varepsilon$. If $\varepsilon > \overline{\sigma}(M)$, the latter denoting the maximum singular value of M, then $C_\varepsilon = \mathbb{R}^n$. Since V is a C^1 function, ∇V is sector bounded near the origin, and, due to Assumption 14, $\widehat{W}$ is quadratically bounded near the origin. Hence, using the Cauchy-Schwarz inequality on the left-hand side of the inequality (7.15), and the local sector boundedness of $\widehat{F}$ and ∇V, as well as W being quadratically bounded near the origin, we have that there exist positive constants $L_{\nabla V}$, L_f, and L_W such that, for values of $x \in \mathbb{R}^n$ near the origin,

$$\begin{aligned}
\widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|) \geq \frac{|d|}{m} \quad \Longrightarrow& \\
|\nabla V(x)||\widehat{f}| + \widehat{W}(x) &\leq L_{\nabla V}|x| L_F\,(|x| + |d|) + L_W|x|^2 \\
&\leq L_{\nabla V} L_f |x|^2 + m L_{\nabla V} L_f |x|^2 + L_W |x|^2 \\
&\leq (L_{\nabla V} L_f (1+m) + L_W)|x|^2.
\end{aligned}$$

Let us define $\sigma_1(x) := (L_{\nabla V} L_f (1+m) + L_W)|x|^2$. We have now shown that (7.15) holds for small values of x and d. For large values of x and d, due to the fact that $\widehat{F}$, ∇V, and $\widehat{W}$ are each locally bounded, there exist $h \in \mathcal{G}$ and a constant $c > 0$ such that

$$\begin{aligned}
\widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|) \quad \Longrightarrow& \\
\left\langle \nabla V(x), \widehat{f} \right\rangle + \widehat{W}(x) &\leq h(|x| + |d| + c) \\
&\leq h(|x| + \gamma^{-1}(|x|) + c) =: \sigma_2(x).
\end{aligned} \tag{7.16}$$

Let $\xi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a smooth function satisfying $\xi(r) = 0$ for $r \leq \gamma(0.5\varepsilon)$ and satisfying

$\xi(r) = 1$ for $r \geq \gamma(\varepsilon)$. Then, letting $\sigma(x) \coloneqq \sigma_1(x) + \xi(|x|)\sigma_2(x)$, and noting that (7.15) holds for $\gamma^{-1}(|x|) \leq \varepsilon$, we have from (7.16) that

$$\widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|) \quad \Longrightarrow \quad \left\langle \nabla V(x), \widehat{f} \right\rangle + \widehat{W}(x) \leq \sigma(x). \tag{7.17}$$

It follows from (7.17) that

$$\begin{aligned} x \neq 0,\ \widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|),\ x^T M x \geq \varepsilon |x|^2 \quad &\Longrightarrow \\ \left\langle \nabla V(x), \widehat{f} \right\rangle + \widehat{W}(x) &\leq \sigma(x) \\ &\leq \frac{\sigma_0(x)^2 \sigma(x)}{\varepsilon^2 |x|^4} \max\{0, x^T M x\} \frac{x^T M x}{\sigma_0(x)^2}. \end{aligned} \tag{7.18}$$

Due to σ and σ_0 being quadratically bounded and sector bounded near the origin, respectively, we can establish

$$\limsup_{x \to 0, x \in D \backslash \{0\}} \frac{\sigma_0(x)^2 \sigma(x)}{\varepsilon^2 |x|^4} < \infty.$$

Pick the continuous function $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ such that

$$\frac{1}{2\mu} \left(\kappa_0(x) + \frac{\sigma_0(x)^2 \sigma(x)}{\varepsilon^2 |x|^4} \right) \leq \kappa(x), \qquad \forall x \in D \backslash \{0\}, \tag{7.19}$$

where $\kappa_0 : \mathbb{R}^n \to \mathbb{R}_{>0}$ is continuous. Here, we can combine our so far derived bounds from (7.14), (7.18) with (7.19) and write

$$\begin{aligned} x \neq 0,\ \widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|),\ s \in \mathrm{SGN}(x^T M x) \quad &\Longrightarrow \\ \left\langle \nabla V(x), \widehat{f} - \kappa(x)(s+1)(x - g(x)) \right\rangle &\\ \leq -\widehat{W}(x) - \kappa_0(x) \max\{0, x^T M x\} \frac{x^T M x}{\sigma_0(x)^2}. \end{aligned} \tag{7.20}$$

Let us define

$$W(x) := \widehat{W}(x) + \kappa_0(x)\max\{0, x^T M x\}\frac{x^T M x}{\sigma_0(x)^2}.$$

Due to Item 3 of Assumption 1, the positivity and continuity of $\kappa_0(\cdot)$, and the fact that σ_0 is positive definite and sector bounded near the origin, we have that $W(\cdot)$ is continuous and positive definite. From (7.20), we have

$$\begin{aligned}\widehat{f} \in \widehat{F}(x,d),\ |x| \geq \gamma(|d|),\ s \in \mathrm{SGN}(x^T M x) \quad &\Longrightarrow \\ \left\langle \nabla V(x), \widehat{f} - \kappa(x)(s+1)(x - g(x)) \right\rangle &\leq -W(x)\end{aligned}$$

It follows that the soft-reset system (7.6) is ISS for large enough κ (see [115, Theorem 1]), particularly for κ satisfying (7.19). ■

Remark 1 *Note that the gain function γ is preserved in the input-to-state stability property from the hard-reset system (7.3) to the soft-reset system (7.6).*

7.3 From Global Asymptotic Stability to ISS

In this section, we demonstrate two methods for inferring ISS from a globally asymptotically stable (GAS) soft-reset system (7.6) together with the following sector growth condition:

Assumption 15 *There exists a positive constant L such that*

$$|\widehat{f}|_{\widehat{F}(x,0)} \leq L|d|, \qquad \forall x \in \mathbb{R}^n,\ d \in \mathbb{R}^m,\ \widehat{f} \in \widehat{F}(x,d).$$

The first method takes advantage of the strong Lyapunov conditions for GAS of the origin for soft-reset system (7.6) with zero disturbance to achieve ISS. The second method establishes

ISS by exploiting a homogeneity assumption on the data of (7.6) and a GAS assumption for the origin of (7.6) with zero disturbance. The latter can be establish under weakened Lyapunov conditions with a strong convexity assumption, as in Section 7.4.

7.3.1 ISS from Strong Lyapunov Conditions

ISS of (7.6) is established under the following assumptions involving strong Lyapunov conditions on the system having zero disturbance.

Assumption 16 There exists a continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that

1. the below inequality holds

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|), \qquad \forall x \in \mathbb{R}^n,\ \alpha_1, \alpha_2 \in \mathcal{K}_\infty \tag{7.21}$$

2. there exist positive constants c and $\hat{c}$ such that, recalling the definition of F in (7.6),

$$\langle \nabla V(x), f_0 \rangle \leq -c|x|^2, \qquad \forall x \in \mathbb{R}^n,\ f_0 \in F(x, 0) \tag{7.22a}$$

$$|\nabla V(x)| \leq \hat{c}|x|, \qquad \forall x \in \mathbb{R}^n, \tag{7.22b}$$

Theorem 13 *If Assumptions 15 and 16 hold then the system (7.6) is ISS.*

Proof: Let $x \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, and $f \in F(x, d)$. Let $\widehat{f} \in \widehat{F}(x, d)$ and $h \in \kappa(x)\left(\mathrm{SGN}\left(\varphi(x)\right) + 1\right)(g(x) - x)$ be such that $f = \widehat{f} + h$. Let $\widehat{f_0^*} \in \widehat{F}(x, 0)$ be such that $|\widehat{f}|_{\widehat{F}(x,0)} = |\widehat{f} - \widehat{f_0^*}|$. Note that

$\widehat{f_0^*} + h \in F(x, 0)$. Then, using (7.22), we have

$$\begin{aligned}\langle \nabla V(x), f\rangle &= \left\langle \nabla V(x), \widehat{f} + h\right\rangle \\ &= \left\langle \nabla V(x), \widehat{f_0^*} + h\right\rangle + \left\langle \nabla V(x), \widehat{f} - \widehat{f_0^*}\right\rangle \\ &\leq -c|x|^2 + |\nabla V(x)|\,|\widehat{f} - \widehat{f_0^*}| \\ &= -c|x|^2 + |\nabla V(x)|\,|\widehat{f}|_{\widehat{F}(x,0)} \\ &\leq -c|x|^2 + \hat{c}L|x||d|.\end{aligned} \tag{7.23}$$

To dominate the term $\hat{c}L|x||d|$ for large $|x|$, we introduce $0 < \theta < 1$ in the previous inequality as follows:

$$\begin{aligned}\frac{\partial V}{\partial x} f &\leq -c(1-\theta)|x|^2 - c\theta|x|^2 + \hat{c}L|x||d| \\ &\leq -c(1-\theta)|x|^2, \qquad \forall |x| \geq \frac{\hat{c}L|d|}{c\theta}.\end{aligned}$$

Hence, the conditions for input-to-state stability from [115, Theorem 1] are satisfied with $\rho(r) = (\hat{c}L/c\theta)r$, and we conclude that system (7.6) is input-to-state stable with $\gamma(r) := \alpha_1^{-1} \circ \alpha_2 \circ \rho(r)$. ■

7.3.2 ISS from Homogeneity Conditions and GAS

In this section, we emulate the results from the previous section and obtain ISS by using GAS of the origin and homogeneity for the soft-reset system (7.6) with $d \equiv 0$.

Assumption 17

1. The origin of the soft-reset system (7.6) is GAS when $d \equiv 0$.
2. The function g and the mapping $x \mapsto \widehat{F}(x, 0)$ are homogeneous of degree 1, and the function κ is homogeneous of degree 0.

Theorem 14 *If Assumptions 13 and 17 hold and there exists a $M = M^T$ such that $\varphi(x) = x^T M x$ for all $x \in \mathbb{R}^n$ then the system (7.6) is ISS.*

Proof: Due to Item 2 of Assumption 17 and the assumption that $\varphi(x) = x^T M x$, we have that F is homogeneous of degree 1. Indeed,

$$\begin{aligned} F(\lambda x, 0) &= \widehat{F}(\lambda x, 0) + \kappa(\lambda x)\Big(\mathrm{SGN}\left(\varphi(\lambda x)\right) + 1\Big)(g(\lambda x) - \lambda x), \\ &= \lambda \widehat{F}(x, 0) + \kappa(x)\Big(\mathrm{SGN}\left(\varphi(x)\right) + 1\Big)\lambda(g(x) - x), \\ &= \lambda\Big[\widehat{F}(x, 0) + \kappa(x)\Big(\mathrm{SGN}\left(\varphi(x)\right) + 1\Big)(g(x) - x)\Big], \\ &= \lambda F(x, 0). \end{aligned} \tag{7.24}$$

Next, due to Item 1 of Assumption 13 and Lemma 11, $x \mapsto F(x, 0)$ is outer semi-continuous and locally bounded. Therefore $F(x, 0)$ is compact for each $x \in \mathbb{R}^n$; see [114, Theorem 5.19]. It follows from Item 1 of Assumption 17 and [116, Theorem 1.2] that there exists a C^∞ function V and a positive definite, continuous function $W : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that (7.21) holds and, for all $x \in \mathbb{R}^n$ and $f_0 \in F(x, 0)$, $\langle \nabla V(x), f_0 \rangle \leq -W(x)$. Then, from [117, Prop. 8], there exists a C^1 function $\overline{V} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that is homogeneous of degree 2 and a positive definite, continuous function $\overline{W} : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ such that

$$\overline{\alpha}_1(|x|) \leq \overline{V}(x) \leq \overline{\alpha}_2(|x|), \qquad \forall x \in \mathbb{R}^n,\ \overline{\alpha}_1, \overline{\alpha}_2 \in \mathcal{K}_\infty, \tag{7.25a}$$

$$\left\langle \nabla \overline{V}(x), f_0 \right\rangle \leq -\overline{W}(x), \qquad \forall x \in \mathbb{R}^n,\ f_0 \in F(x, 0). \tag{7.25b}$$

Note that $\overline{V}$ can be constructed from V using [117, Eq. 36] and [117, Eq. 37]. We will show that Assumption 16 holds with V replaced by $\overline{V}$ and thereby invoke Theorem 13 to conclude that (7.6) is ISS.

Due to the fact that $\overline{V}$ is homogeneous of degree 2, we have that, for all $x \in \mathbb{R}^n$ such that

$x \neq 0$, with $w \coloneqq x/|x|$,

$$\overline{V}(x) = \overline{V}(|x|w) = |x|^2 \overline{V}(w) \geq |x|^2 \min_{v:\ |v|=1} \overline{V}(v). \tag{7.26}$$

Letting $a_1 \coloneqq \min_{v:\ |v|=1} \overline{V}(v)$, we have from (7.25a) that $a_1 > 0$. It follows from (7.26) that

$$\overline{V}(x) \geq a_1 |x|^2, \quad \forall x \in \mathbb{R}^n. \tag{7.27}$$

Due to the fact that $\overline{V}$ is homogeneous of degree 2, the Euler homogeneous function theorem ensures that $\nabla\overline{V}$ is homogeneous of degree 1. Then, due to the homogeneity of degree 1 of $\nabla\overline{V}$ and $x \mapsto F(x, 0)$, along with the homogeneity of degree 2 of $\overline{V}$, it can be shown that, for all $x \in \mathbb{R}^n$ and $f_0 \in F(x, 0)$,

$$\begin{aligned} c_1 &\coloneqq \sup_{v:\ |v|=1} \frac{\langle \nabla\overline{V}(v), f_0 \rangle}{\overline{V}(v)}, \\ \left\langle \nabla\overline{V}(x), f_0 \right\rangle &\leq c_1 \overline{V}(x). \end{aligned} \tag{7.28}$$

From (7.25), we have that c_1 must be negative. Hence, with $\lambda \coloneqq -c_1 > 0$, we have

$$\left\langle \nabla\overline{V}(x), f_0 \right\rangle \leq -\lambda \overline{V}(x), \quad \forall x \in \mathbb{R}^n,\ f_0 \in F(x, 0). \tag{7.29}$$

Combining (7.29) with (7.27), we have that $\overline{V}$ satisfies (7.22a) with $c \coloneqq a_1\lambda$. Due to the fact that $\nabla\overline{V}$ is homogeneous of degree 1, $\overline{V}$ satisfies (7.22b) with $\hat{c} \coloneqq \max_{v:\ |v|=1} \left|\nabla\overline{V}(v)\right|$. We conclude that Assumption 16 holds with V replaced by $\overline{V}$, and it follows from Theorem 13 that (7.6) is ISS. ■

7.4 GAS for homogeneous soft-reset systems

The goal of this section is to give weak Lyapunov conditions on the hard-reset system with $d \equiv 0$ that guarantee the assumptions of the previous sections 7.3.1 and 7.3.2 on the soft-reset system with $d \equiv 0$. We consider the continuous-time implementation (7.6) of the hybrid, reset control system (7.3) with $d \equiv 0$ and show that the origin of the soft implementation is globally exponentially stable (GES) if the following assumption holds.

Assumption 18

1. The function g and the mapping $x \mapsto \widehat{F}(x, 0)$ are homogeneous of degree 1, and the function κ is homogeneous of degree 0.

2. Let $M = M^T$. There exist $\varepsilon > 0$ and a C^1 and strongly convex, homogeneous of degree 2, positive definite function $V : \mathbb{R}^n \to \mathbb{R}_{>0}$ such that, with the definition $C_\varepsilon := \{x \in \mathbb{R}^n : x^T M x \leq \varepsilon x^T x\}$, the following inequalities hold:

$$\left\langle \nabla V(x), \widehat{f_0} \right\rangle \leq 0, \qquad \forall x \in C_\varepsilon,\ \widehat{f_0} \in \widehat{F}(x, 0), \tag{7.30a}$$

$$V(g(x)) \leq V(x), \qquad \forall x \in D. \tag{7.30b}$$

3. Jumps starting in the jump set $x \in D$ land in the flow set i.e., $g(x) \in C$ for all $x \in D$.

4. There is no solution to (7.3a) in C with an unbounded time domain that keeps the function V equal to a nonzero constant.

According to Assumption 18, we have stability of the origin for system (7.3) with $d \equiv 0$. However, these assumptions are not strong enough to establish GES of the origin. It is possible for hard-reset systems to have discrete-time solutions which do not converge to zero. This can be addressed by considering the corresponding continuous-time soft-reset implementation

(7.6) of the hard-reset system (7.3), for which GES of the origin can be obtained with $d \equiv 0$ as in the following result.

Theorem 15 *Under Assumptions 13 and Assumption 18, the origin of (7.6) with $d \equiv 0$ is globally exponentially stable for $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ taking sufficiently large values.*

Proof: For the origin of (7.6) with $d \equiv 0$, consider the Lyapunov candidate V, which is positive definite and radially unbounded due to Item 2 of Assumption 18. First, we bound the inner product $\langle \nabla V(x), (\mathrm{SGN}(x^T M x) + 1)(g(x) - x)\rangle$. By using (7.3*b*), strong convexity of V as in (7.1), and (7.30b), we have

$$x^T M x \geq 0 \quad \Longrightarrow \quad \langle \nabla V(x), (g(x) - x)\rangle \leq -\mu |g(x) - x|^2. \tag{7.31}$$

We can rewrite (7.31) as follows:

$$\begin{aligned} &s \in \mathrm{SGN}(x^T M x) \quad \Longrightarrow \\ &\qquad \langle \nabla V(x), (s+1)(g(x) - x)\rangle \leq -(s+1)\mu |g(x) - x|^2. \end{aligned} \tag{7.32}$$

Let $\delta_0 > 0$ be such that, for all $x \in \mathbb{R}^n$, we have $|M(g(x) + x)|_2 \leq \delta_0 |x|_2$, and then, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} x \neq 0,\ x^T M x \geq 0 \quad \Longrightarrow \quad |g(x) - x| &\geq \frac{-(g(x) - x)^T M (g(x) + x)}{\delta_0 |x|} \\ &= \frac{x^T M x - g(x)^T M g(x)}{\delta_0 |x|} \\ &\geq \frac{x^T M x}{\delta_0 |x|}. \end{aligned} \tag{7.33}$$

Combining (7.32) and (7.33) results in

$$x \neq 0,\ s \in \mathrm{SGN}(x^T M x) \implies$$
$$\langle \nabla V(x), (s+1)(g(x)-x)\rangle \leq -2\mu \max\{0, x^T M x\}\frac{x^T M x}{\delta_0^2 |x|^2}. \tag{7.34}$$

The next step is to bound $\left\langle \nabla V(x), \widehat{f_0}\right\rangle$ for $\widehat{f_0} \in \widehat{F}(x,0)$ and for all $x \in \mathbb{R}^n$. Due to Assumption 18, this quantity is negative when the inequality $x^T M x \leq \varepsilon |x|_2^2$ holds. Using homogeneity of degree 2 for V and thus, homogeneity of degree 1 of ∇V due to Euler's homogeneous function theorem, along with homogeneity of degree 1 of $x \mapsto \widehat{F}(x,0)$, it follows for $x^T M x \geq \varepsilon |x|_2^2$ that there exists a $\Gamma > 0$ such that, for all $\widehat{f_0} \in \widehat{F}(x,0)$,

$$\begin{aligned} x^T M x \geq \varepsilon |x|_2^2 > 0 \implies \left\langle \nabla V(x), \widehat{f_0}\right\rangle &\leq \Gamma |x|^2 \\ &\leq \frac{\Gamma}{\varepsilon} x^T M x \\ &\leq \frac{\Gamma \delta_0^2}{\varepsilon^2} \max\{0, x^T M x\}\frac{x^T M x}{\delta_0^2 |x|^2}. \end{aligned} \tag{7.35}$$

It follows from (7.34), (7.35), and Assumption 18 that, for each constant $v > 0$, there exists κ taking sufficiently large values such that, for $\widehat{f_0} \in \widehat{F}(x,0)$,

$$s \in \mathrm{SGN}(x^T M x) \implies$$
$$\left\langle \nabla V(x), \widehat{f_0} - \kappa(x)(s+1)(x-g(x)\right\rangle \leq -v \max\{0, x^T M x\}\frac{x^T M x}{|x|^2} \leq 0.$$

Due to the invariance principle for differential inclusion [118, Theorem 11] and properties of F, the origin is globally asymptotically stable if and only if there is no $c > 0$ and solution x such that $V(x(t)) = c$ for all $t \geq 0$. Let us assume that $V(x(t))$ is equal to a nonzero constant.

Let $\widehat{f_0}(t) \in \widehat{F}(x(t), 0)$ and $s(t) \in \mathrm{SGN}\left(x(t)^T M x(t)\right)$ satisfy, for almost all $t \geq 0$,

$$\dot{x}(t) = \widehat{f_0}(t) - \kappa(x(t))(s(t) + 1)(x(t) - g(x(t))). \tag{7.36}$$

Then we have

$$\left\langle \nabla V(x(t)), \widehat{f_0}(t) - \kappa(x(t))(s(t) + 1)(x(t) - g(x(t)) \right\rangle = 0.$$

According to (7.35), such solutions require $x^T(t) M x(t) \leq 0$ for all $t \geq 0$. As a result, it follows from (7.30a), (7.32), and the positivity of κ that, for almost all $t \geq 0$,

$$\begin{aligned} &\left\langle \nabla V(x(t)), \widehat{f_0}(t) \right\rangle = 0, \\ &\langle \nabla V(x(t)), -\kappa(x(t))(s(t) + 1)(g(x(t) - x(t)) \rangle = 0. \end{aligned}$$

With (7.32) and the positivity of $\kappa(\cdot)$ and μ, it follows that for almost all $t \geq 0$

$$(s(t) + 1)|g(x(t)) - x(t)| = 0.$$

It follows from (7.36) and the definition of $\widehat{f}(t)$ that $x(\cdot)$ is also a solution of (7.3a). Since we have assumed in Item 4 of Assumption 18 that solutions of (7.3a) do not keep V equal to a nonzero constant, it follows that no solution keeps V equal to a nonzero constant and thus we have estalished GAS of the origin of $\dot{x} \in F(x, 0)$. That is, GAS of the origin of (7.6) with $d \equiv 0$ can be concluded. With this conclusion, along with Assumption 13, it can be shown using [116, Theorem 1.2] and [117, Prop. 8], as done in the proof of Theorem 14, that there exists a C^1 function $\overline{V}$ that is homogeneous of degree 2 and a constant $\lambda > 0$ such that (7.29) holds. Furthermore, due to the fact that $\overline{V}$ is homogeneous of degree 2, it can be shown using steps similar to those in the proof of Theorem 14 that, with $a_1 := \min_{v:\, |v|=1} \overline{V}(v)$ and

$a_2 \coloneqq \max_{v:\,|v|=1} \overline{V}(v)$, we have $a_1|x|^2 \leq \overline{V}(x) \leq a_2|x|^2$ for all $x \in \mathbb{R}^n$. It follows that the origin of (7.6) with $d \equiv 0$ is GES. ■

The following corollary establishes that the soft-reset system (7.6) is ISS by combining Theorems 14 and 15.

Corollary 3 *If Assumptions 15 and 18 hold then the system (7.6) is ISS for $\kappa : \mathbb{R}^n \to \mathbb{R}_{>0}$ taking sufficiently large values.*

Proof: Given that Assumption 18 holds, Theorem 15 ensures that Item 1 of Assumption 17 holds for $\kappa(\cdot)$ taking sufficiently large values. The remaining items of Assumption 17 follow immediately from the assumptions stated here. Hence, Theorem 14 can be invoked to conclude ISS of (7.6). ■

In the next section, we present a numerical example.

7.5 Numerical Example

7.5.1 GAS example

Consider the soft-reset system (7.6) with state x comprising a plant state $x_p \in \mathbb{R}$ and a compensator state $x_c \in \mathbb{R}$, i.e., $x \coloneqq (x_p, x_c)^T \in \mathbb{R}^2$, with the system data given by

$$\widehat{F}(x,0) = \begin{bmatrix} x_c \\ -K\,\mathrm{SGN}(x_c)|x| - x_p \end{bmatrix}, \tag{7.38a}$$

$$g(x) = \begin{bmatrix} x_p \\ 0 \end{bmatrix}, \qquad \varphi(x) = x^T \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x, \tag{7.38b}$$

for some $K \in \mathbb{R}_{>0}$. We choose $\kappa(\cdot)$ in (7.6) to be constant, namely with $\kappa(x) = \kappa \in \mathbb{R}_{\geq 0}$ for all $x \in \mathbb{R}^2$. Consequently, $\kappa(\cdot)$ is homogeneous of degree 0, and, noting that $x \mapsto \widehat{F}(x,0)$ and g

in (7.38) are homogeneous of degree 1, we have that Item 1 of Assumption 18 holds. For all $x \in \mathbb{R}^2$, consider $V(x) \coloneqq x^T x$, which is a C^1, strongly convex, and positive definite function that is homogeneous of degree 2 and satisfies

$$\left\langle \nabla V(x), \widehat{f_0} \right\rangle = -2K|x|x_c \leq 0, \qquad \forall x \in \mathbb{R}^2,\ \widehat{f_0} \in \widehat{F}(x, 0), \tag{7.39a}$$

$$V(g(x)) = x_p^T x_p \leq V(x), \qquad \forall x \in \mathbb{R}^2. \tag{7.39b}$$

Hence, V satisfies Item 2 of Assumption 18 with $M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $C_\varepsilon = \mathbb{R}^n$. For all $x \in \mathbb{R}^2$, $g(x)^T M g(x) = 0$, and thus Item 3 of Assumption 18 holds. Supposing that a solution x of system (7.3a) keeps $V(x(t))$ at a nonzero constant for all $t \geq 0$, it follows from (7.39a) that $x_c \equiv 0$ and, therefore, that $\dot{x}_c \equiv 0$. It then follows from (7.38a) that solutions x of system (7.3a) satisfy $\dot{x}_p = x_c \equiv 0$ and $\dot{x}_c = -x_p \equiv 0$. It follows that $x \equiv 0$ and $V(x) = x^T x \equiv 0$, contradicting the premise that $V(x)$ is a nonzero constant. We have thus shown by contradiction that Item 4 of Assumption 18 holds. It is now verified that Assumption 18 holds, and Theorem 15 may be applied to conclude GES of the origin of (7.6) with $d \equiv 0$ and with the system data given by (7.38).

Setting $K = 0.5$, Figures 7.1 and 7.2 respectively show the evolution of the states x_p and x_c, highlighting the effect of setting κ equal to various constant values. Figure 7.3 shows the evolution of $V(x) = x^T x$.

7.5.2 ISS example

Figures 7.4-7.6 show the same example from Section 7.5.1 but with

$$\widehat{F}(x, d) = \begin{bmatrix} x_c + d \\ -K\,\mathrm{SGN}(x_c)|x| - x_p \end{bmatrix} \tag{7.40}$$

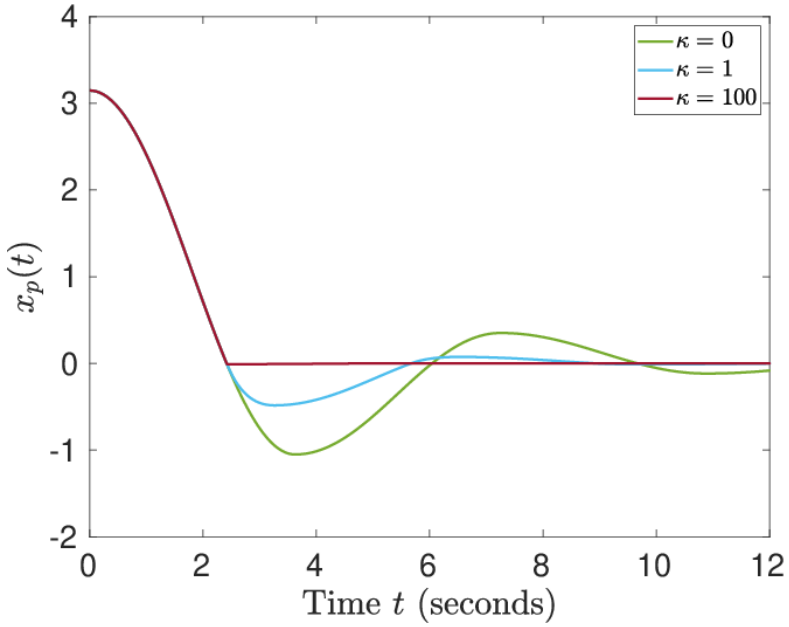

Figure 7.1: The value of $x_p(t)$ as a function of t using nonlinear homogeneous system data and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

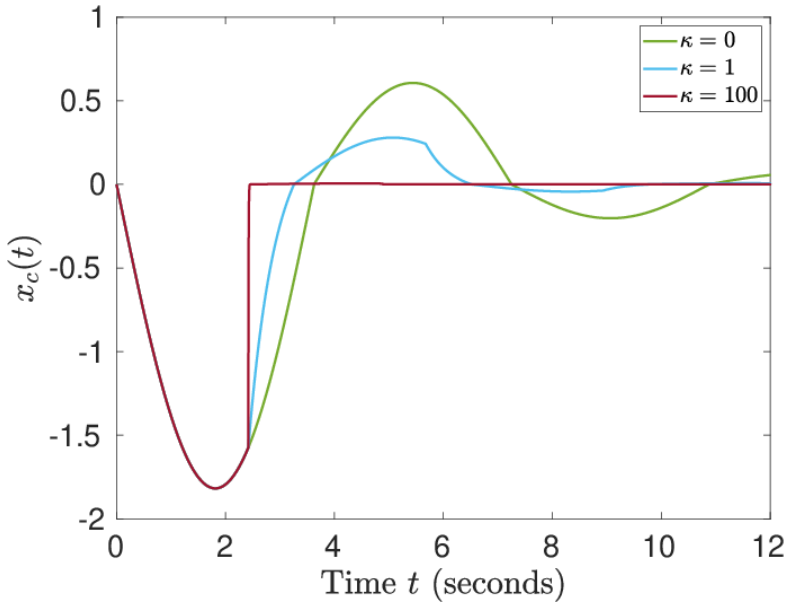

Figure 7.2: The value of $x_c(t)$ as a function of t using nonlinear homogeneous system data and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

and a disturbance of $d(t) = 0.1\sin(t)$ for all $t \geq 0$.

Setting $d \equiv 0$ makes (7.40) equivalent to (7.38a), and thus, Assumption 18 holds for the

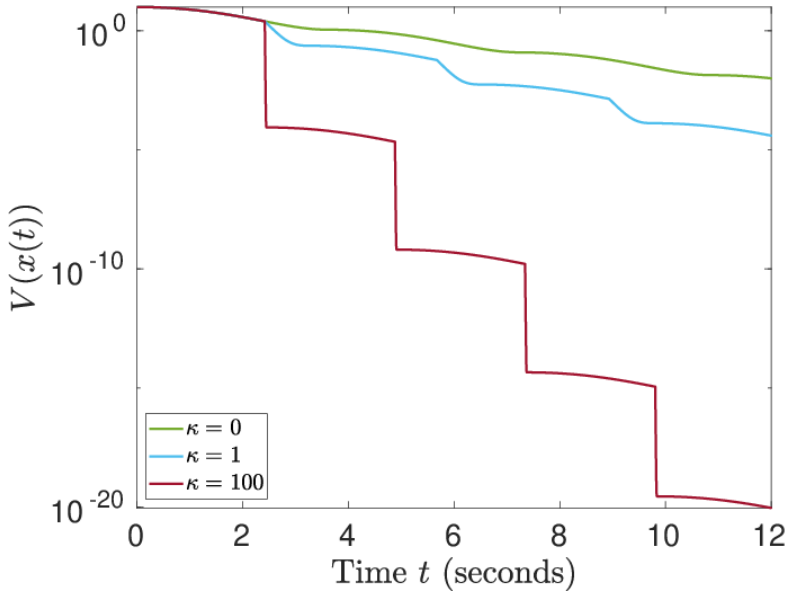

Figure 7.3: The value of $V(x(t))$ as a function of t using nonlinear homogeneous system data and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

same reasons given in Section 7.5.1. With F given by (7.6), (7.38b), and (7.40), for any $x \in \mathbb{R}^2$, $d \in \mathbb{R}$, and $\widehat{f} \in \widehat{F}(x,d)$, letting $\widehat{f}_0^*$ be such that $|\widehat{f}|_{\widehat{F}(x,0)} = |\widehat{f} - \widehat{f}_0^*|$, we have that $\widehat{f} - \widehat{f}_0^* = d$, and thus, Assumption 15 holds with $L = 1$. We conclude that the conditions of Corollary 3 hold for the soft-reset system (7.6) with system data given by (7.38b) and (7.40), ensuring that, for sufficiently large $\kappa(\cdot)$, (7.6) is ISS with this system data. Figures 7.4-7.6 show results for the various constant values of κ considered in Section 7.5.1.

7.6 Concluding Remarks

Conditions were provided for input-to-state stability of nonlinear soft-reset systems with inputs. Assuming strong convexity of a Lyapunov function for the hard-reset system, sufficient conditions are provided for input-to-state stability of the corresponding soft-reset system. Moreover, two methods are described for obtaining input-to-state stability for the soft-reset system via asymptotic stability for the origin of the zero-input soft-reset system. Lastly, exponen-

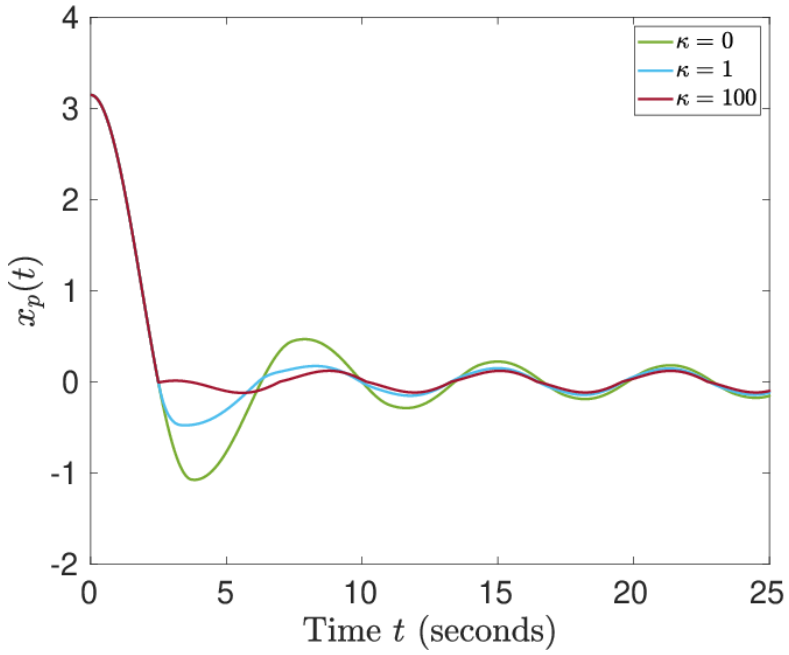

Figure 7.4: The value of $x_p(t)$ as a function of t using nonlinear homogeneous system data, a sinusoidal input disturbance in the plant, and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

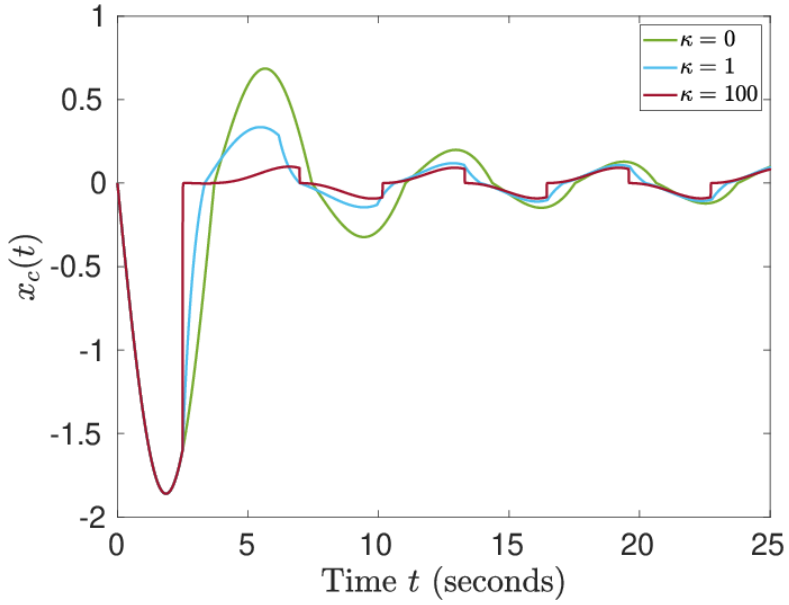

Figure 7.5: The value of $x_c(t)$ as a function of t using nonlinear homogeneous system data, a sinusoidal input disturbance in the plant, and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

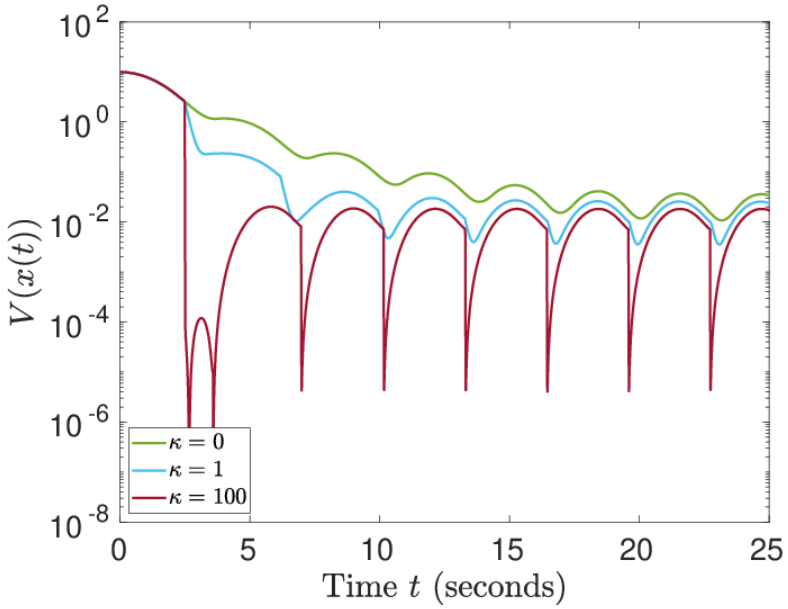

Figure 7.6: The value of $V(x(t))$ as a function of t using nonlinear homogeneous system data, a sinusoidal input disturbance in the plant, and initializing the system state at $(x_p^\circ, 0)^T$ with x_p° being randomly selected.

tial stability is established for a soft-reset system using strong convexity of a weak Lyapunov function for the hard-reset system, along with homogeneity of the data and the Lyapunov function. Numerical examples are given, illustrating global asymptotic stability and input-to-state stability of a soft-reset control system in a closed loop.

Bibliography

[1] S. Mastellone, D. M. Stipanović, C. R. Graunke, K. A. Intlekofer, and M. W. Spong, *Formation control and collision avoidance for multi-agent non-holonomic systems: Theory and experiments*, *The International Journal of Robotics Research* **27** (2008), no. 1 107–126.

[2] L. E. Parker, D. Rus, and G. S. Sukhatme, *Multiple mobile robot systems*, in *Springer Handbook of Robotics*, pp. 1335–1384. Springer, 2016.

[3] P. Holmes, R. J. Full, D. Koditschek, and J. Guckenheimer, *The dynamics of legged locomotion: Models, analyses, and challenges*, *SIAM review* **48** (2006), no. 2 207–304.

[4] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid dynamical systems*, *IEEE Control Systems Magazine* **29** (April, 2009) 28–93.

[5] A. V. Savkin and R. J. Evans, *Hybrid dynamical systems: controller and sensor switching problems*. Springer Science & Business Media, 2002.

[6] A. J. Van Der Schaft and H. Schumacher, *An introduction to hybrid dynamical systems*, vol. 251. springer, 2007.

[7] D. Chatterjee and D. Liberzon, *Stability analysis and stabilization of randomly switched systems*, in *Proceedings of the 45th IEEE Conference on Decision and Control*, pp. 2643–2648, IEEE, 2006.

[8] D. Chatterjee and D. Liberzon, *On stability of stochastic switched systems*, in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, vol. 4, pp. 4125–4127, IEEE, 2004.

[9] A. R. Teel, *Lyapunov conditions certifying stability and recurrence for a class of stochastic hybrid systems*, *Annual Reviews in Control* **37** (2013) 1–24.

[10] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*. Springer-Verlag London, 1994.

[11] S. S. Kia, J. Cortes, and S. Martinez, *Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication*, *Automatica* **55** (2015) 254–264.

[12] J. Cortes, *Finite-time convergent gradient flows with applications to network consensus*, *Automatica* **42** (2006), no. 11 1993–2000.

[13] R. Goebel, *Stability and robustness for saddle-point dynamics through monotone mappings*, *Systems & Control Lett.* **108** (2017) 16–22.

[14] T. Strizic, J. I. Poveda, and A. R. Teel, *Hybrid gradient descent for robust global optimization on the circle*, *IEEE 56th Annual Conference on Decision and Control* (2017) 2985–2990.

[15] S. Trip and C. D. Persis, *Distributed optimal load frequency control with non-passive dynamics*, *IEEE Trans. Autom. Control, DOI 10.1109/TCNS.2017.2698259* (2017).

[16] J. Barreiro-Gomez, N. Quijano, and C. Ocampo-Martinez, *Constrained distributed optimization: A population dynamics approach*, *Automatica* **69** (2016), no. 101-116.

[17] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, *IEEE Transactions on Automatic Control* **49** (2004), no. 9 1520–1533.

[18] C. G. Mayhew and A. R. Teel, *Synergistic hybrid feedback for global rigid-body attitude tracking on so(3)*, *IEEE Trans. Autom. Control.* **58** (2013), no. 11 2730–2742.

[19] J. I. Poveda and A. R. Teel, *A robust event-triggered approach for fast sampled-data extremization and learning*, *IEEE Transactions on Automatic Control* **62** (2017), no. 10 4949–4964.

[20] C. Jin, R. Ge, P. Netrapalli, and S. M. Kakade, *How to escape saddle points efficiently*, *arXiv:1703.00887* (2017).

[21] C. G. Mayhew and A. R. Teel, *On the topological structure of attraction basins for differential inclusions*, *Systems & Control Lett.* **60** (2011), no. 12 1045–1050.

[22] R. G. Sanfelice, M. J. Messina, S. E. Tuna, and A. R. Teel, *Robust hybrid controllers for continuous-time systems with applications to obstacle avoidance and regulation to disconnected set of points*, in *Proc. 26th American Control Conference*, pp. 3352–3357, July, 2006.

[23] J. D. Lee, M. Simchowitz, M. I. Jordan, and B. Recht, *Gradient descent converges to minimizers*, *arXiv:1602.04915* (2016).

[24] R. Ge, F. Huang, C. Jin, and Y. Yuan, *Escaping from saddle-points - online stochastic gradient for tensor decomposition-*, *In proc. of 28th Conference on Learning Theory* (2015) 797–842.

[25] C. Jin, P. Netrapalli, and M. I. Jordan, *Accelerated gradient descent escapes saddle points faster than gradient descent*, *arXiv:1711.10456* (2018).

[26] A. R. Teel, A. Subbaraman, and A. Sferlazza, *Stability analysis for stochastic hybrid systems: a survey*, *Automatica* **50** (2014), no. 10 2435–2456.

[27] R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid Dynamical Systems*. Princeton University Press, 2012.

[28] A. Subbaraman and A. Teel, *Recurrence principles and their application to stability theory for a class of stochastic hybrid systems*, *IEEE Transactions on Automatic Control* **61** (2016), no. 11 3477–3492.

[29] L. J. Ratliff, S. A. Burden, and S. S. Sastry, *Characterization and computation of local Nash equilibria in continuous games*, in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*, pp. 917–924, IEEE, 2013.

[30] M. Baradaran, J. I. Poveda, and A. R. Teel, *Achieving global optimization on smooth manifolds: A stochastic hybrid systems approach*, *Proc. 57th IEEE Conference on Decision and Control* (2018).

[31] A. S. Bandeira, N. Boumal, and A. Singer, *Tightness of the maximum likelihood semidefinite relaxation for angular synchronization*, *Mathematical Programming* **163** (2017), no. 1-2 145–167.

[32] T. Lee, *Geometric tracking control of the attitude dynamics of a rigid body on SO(3)*, in *American Control Conference (ACC), 2011*, pp. 1200–1205, IEEE, 2011.

[33] V. Kelner, F. Capitanescu, O. Léonard, and L. Wehenkel, *A hybrid optimization technique coupling an evolutionary and a local search algorithm*, *Journal of Computational and Applied Mathematics* **215** (2008), no. 2 448–456.

[34] P.-A. Absil and R. Sepulchre, *A hybrid control scheme for swing up acrobatics*, in *European Conference on Control ECC*, 2001.

[35] D. Luenberger, *Optimization by Vector Space Methods*. Wiley Interscience, 1969.

[36] C. G. Mayhew and A. R. Teel, *Hybrid control of planar rotations*, in *American Control Conference (ACC), 2010*, pp. 154–159, IEEE, 2010.

[37] S. P. Bhat and D. S. Bernstein, *A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon*, *Systems & Control Letters* **39** (2000), no. 1 63–70.

[38] C. G. Mayhew, *Hybrid control for topologically constrained systems*. University of California, Santa Barbara, 2010.

[39] M. Baradaran, J. I. Poveda, and A. R. Teel, *Global optimization on the sphere: A stochastic hybrid systems approach*, *to appear in Proc. NOLCOS* (2019).

[40] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, *Projected gradient descent on Riemannian manifolds with applications to online power system optimization*, in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 225–232, IEEE, 2016.

[41] J. M. Lee, *Introduction to smooth manifolds*. Springer, 2001.

[42] Y. Nesterov, *Gradient methods for minimizing composite functions*, *Mathematical Programming* **140** (2013), no. 1 125–161.

[43] G. Teschke and C. Borries, *Accelerated projected steepest descent method for nonlinear inverse problems with sparsity constraints*, *Inverse Problems* **26** (2010), no. 2.

[44] D. Ochoa, J. I. Poveda, C. A. Uribe, and N. Quijano, *Hybrid robust optimal resource allocation*, *to appear in Proc. 58th IEEE Conference on Decision and Control* (2019).

[45] Y. Nesterov, *A method for unconstrained convex minimization problem with the rate of convergence o (1/k^2)*, in *Doklady AN USSR*, vol. 269, pp. 543–547, 1983.

[46] Y. Liu, F. Shang, J. Cheng, H. Cheng, and L. Jiao, *Accelerated first-order methods for geodesically convex optimization on riemannian manifolds*, in *Advances in Neural Information Processing Systems*, pp. 4868–4877, 2017.

[47] W. Krichene, A. Bayen, and P. L. Bartlett, *Adaptive averaging in accelerated descent dynamics*, in *Advances in Neural Information Processing Systems*, pp. 2991–2999, 2016.

[48] W. Krichene, A. Bayen, and P. L. Bartlett, *Accelerated mirror descent in continuous and discrete time*, in *Advances in neural information processing systems*, pp. 2845–2853, 2015.

[49] S. Ruder, *An overview of gradient descent optimization algorithms*, *arXiv preprint arXiv:1609.04747* (2016).

[50] Y. Bengio, N. Boulanger-Lewandowski, and R. Pascanu, *Advances in optimizing recurrent networks*, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8624–8628, IEEE, 2013.

[51] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.

[52] J. R. Shewchuk, *An introduction to the conjugate gradient method without the agonizing pain*, 1994.

[53] B. Oadonoghue and E. Candes, *Adaptive restart for accelerated gradient schemes*, *Foundations of computational mathematics* (2015) 715–732.

[54] C. Jin, P. Netrapalli, and M. I. Jordan, *Accelerated gradient descent escapes saddle points faster than gradient descent*, *arXiv preprint arXiv:1711.10456* (2017).

[55] W. Su, S. Boyd, and E. Candes, *A differential equation for modeling nesterovs accelerated gradient method: Theory and insights*, in *Advances in Neural Information Processing Systems*, pp. 2510–2518, 2014.

[56] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer, 1998.

[57] Z. Wu, *Tangent cone and contingent cone to the intersection of two closed sets*, *Nonlinear Analysis: Theory, Methods & Applications* **73** (2010), no. 5 1203–1220.

[58] T. Alpcan and T. Basar, *A stability result for switched systems with multiple equilibria*, *Dynamics of Continuous, Discrete and Impulsive Systems Series A: Mathematical Analysis* **17** (2010), no. 4 949–958.

[59] S. Veer and I. Poulakakis, *Switched systems with multiple equilibria under disturbances: boundedness and practical stability*, *To appear in IEEE Transactions on Automatic Control* (2019) 4935–2307.

[60] S. Veer and I. Poulakakis, *Practical stability of switched systems with multiple equilibria under disturbances*, *Proceedings of the American Control Conference* (2019) 4935–4940.

[61] S. Veer, M. S. Motahar, and I. Poulakakis, *Generation of and switching among limit-cycle bipedal walking gaits*, in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 5827–5832, IEEE, 2017.

[62] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*, vol. 23. Siam, 1999.

[63] R. D. Gregg, A. K. Tilton, S. Candido, T. Bretl, and M. W. Spong, *Control and planning of 3D dynamic walking with asymptotically stable gait primitives*, *IEEE Transactions on Robotics* **28** (2012), no. 6 1415–1423.

[64] S. Veer and I. Poulakakis, *Safe adaptive switching among dynamical movement primitives: Application to 3d limit-cycle walkers*, in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 3719–3725, IEEE, 2019.

[65] T. Alpcan, *Noncooperative games for control of networked systems*. Citeseer, 2006.

[66] O. Makarenkov and A. Phung, *Dwell time for local stability of switched affine systems with application to non-spiking neuron models*, *Applied Mathematics Letters* **86** (2018) 89–94.

[67] C. Cai, R. Goebel, R. G. Sanfelice, and A. R. Teel, *Hybrid systems: limit sets and zero dynamics with a view toward output regulation*, in *Analysis and Design of Nonlinear control systems*, pp. 241–261. Springer, 2008.

[68] L. Vu, D. Chatterjee, and D. Liberzon, *Input-to-state stability of switched systems and switching adaptive control*, *Automatica* **43** (2007), no. 4 639–646.

[69] C. Cai, A. R. Teel, and R. Goebel, *Smooth Lyapunov functions for hybrid systems part ii:(pre) asymptotically stable compact sets*, *IEEE Transactions on Automatic Control* **53** (2008), no. 3 734–748.

[70] E. G. Anderson Jr and K. Lewis, *A dynamic model of individual and collective learning amid disruption*, *Organization Science* **25** (2014), no. 2 356–376.

[71] J. G. March, *Exploration and exploitation in organizational learning*, *Organization science* **2** (1991), no. 1 71–87.

[72] W. Mei, N. E. Friedkin, K. Lewis, and F. Bullo, *Dynamic models of appraisal networks explaining collective learning*, *IEEE Transactions on Automatic Control* **63** (2018), no. 9 2898–2912.

[73] H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluagac, A. Kadri, and A. Tuncer, *UAV-enabled intelligent transportation systems for the smart city: Applications and challenges*, *IEEE Communications Magazine* **55** (2017), no. 3 22–28.

[74] Y. Zhou, N. Cheng, N. Lu, and X. S. Shen, *Multi-UAV-aided networks: Aerial-ground cooperative vehicular networking architecture*, *ieee vehicular technology magazine* **10** (2015), no. 4 36–44.

[75] I. Bekmezci, O. K. Sahingoz, and Ş. Temel, *Flying ad-hoc networks (fanets): A survey*, *Ad Hoc Networks* **11** (2013) 1254–1270.

[76] M. Baradaran and A. R. Teel, *Omega-limit sets and robust stability for switched systems with distinct equilibria*, *IFAC Proceedings Volumes IFAC (International Federation of Automatic Control) World Congress in Berlin* (2020).

[77] J. H. Le and A. R. Teel, *Hybrid heavy-ball systems: Reset methods for optimization with uncertainty*, *in Proc. American Control Conference* (2021).

[78] G. Bianchin, J. Poveda, and E. Dall'Anese, *Online optimization of switched LTI systems using continuous-time and hybrid accelerated gradient flows*, *in Proc. American Control Conference* (2021).

[79] J. P. Hespanha and A. S. Morse, *Stability of switched systems with average dwell-time*, in *Proc. 38th IEEE Conference on Decision and Control*, (Sidney, Australia), pp. 2655–2660, 1999.

[80] R. Goebel, *A glimpse at pointwise asymptotic stability for continuous-time and discrete-time dynamics*, in *Splitting Algorithms, Modern Operator Theory, and Applications*, p. 243. Springer, 2019.

[81] G. França, D. Robinson, and R. Vidal, *A nonsmooth dynamical systems perspective on accelerated extensions of ADMM*, *arXiv:1808.04048* (2018).

[82] A. Hauswirth, F. Dörfler, and A. Teel, *On the robust implementation of projected dynamical systems with anti-windup controllers*, in *2020 American Control Conference (ACC)*, pp. 1286–1291, 2020.

[83] A. Hauswirth, F. Dörfler, and A. Teel, *On the differentiability of projected trajectories and the robust convergence of non-convex anti-windup gradient flows*, *IEEE Control Systems Letters* **4** (2020), no. 3 620–625.

[84] D. Bertsekas, *Convex Optimization Theory*. Athena Scientific optimization and computation series. Athena Scientific, 2009.

[85] R. Olfati-Saber, J. A. Fax, and R. M. Murray, *Consensus and cooperation in networked multi-agent systems*, *Proceedings of the IEEE* **95** (2007), no. 1 215–233.

[86] E. Ghadimi, I. Shames, and M. Johansson, *Multi-step gradient methods for networked optimization*, *IEEE Transactions on Signal Processing* **61** (2013), no. 21 5417–5429.

[87] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*, *Kluwer Academic Publishers* (2003).

[88] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, *Timescale separation in autonomous optimization*, *IEEE Transactions on Automatic Control* **66** (2021), no. 2 611–624.

[89] A. R. Teel, J. I. Poveda, and J. Le, *First-order optimization algorithms with resets and Hamiltonian flows*, in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 5838–5843, 2019.

[90] I. Gutman and W. Xiao, *Generalized inverse of the Laplacian matrix and some applications*, *Bulletin (Academie serbe des sciences et des arts. Classe des sciences mathematiques et naturelles. Sciences mathematiques)* (2004) 15–23.

[91] E. P. Ryan, *An integral invariance principle for differential inclusions with applications in adaptive control*, *SIAM J. Control Optim.* **36** (1998), no. 3 960–980.

[92] A. Cherukuri and J. CortÃ©s, *Distributed generator coordination for initialization and anytime optimization in economic dispatch*, *IEEE Transactions on Control of Network Systems* **2** (Sep., 2015) 226–237.

[93] E. D. Sontag, *Smooth stabilization implies coprime factorization*, *IEEE Trans. Autom. Control* **34** (1989) 435–443.

[94] Z. Jiang, A. Teel, and L. Praly, *Small-gain theorem for iss systems and applications*, *Math. Control Signal Systems* **7** (1994) 95–120.

[95] J. C. Clegg, *A nonlinear integrator for servomechanisms*, *Transactions of the American Institute of Electrical Engineers, Part II: Applications and Industry* **77** (1958), no. 1 41–42.

[96] I. Horowitz and P. Rosenbaum, *Non-linear design for cost of feedback reduction in systems with large parameter uncertainty*, *International Journal of Control* **21** (1975), no. 6 977–1001.

[97] K. Krishnan and I. Horowitz, *Synthesis of a non-linear feedback system with significant plant-ignorance for prescribed system tolerances*, *International Journal of Control* **19** (1974), no. 4 689–706.

[98] H. Hu, Y. Zheng, Y. Chait, and C. Hollot, *On the zero-input stability of control systems with clegg integrators*, in *Proceedings of the 1997 American Control Conference*, vol. 1, pp. 408–410, 1997.

[99] C. Hollot, Y. Zheng, and Y. Chait, *Stability analysis for control systems with reset integrators*, in *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 2, pp. 1717–1719, 1997.

[100] Y. Chait and C. Hollot, *On Horowitz's contributions to reset control*, *International Journal of Robust and Nonlinear Control* **12** (2002), no. 4 335–355.

[101] O. Beker, C. Hollot, Y. Chait, and H. Han, *Fundamental properties of reset control systems*, *Automatica* **40** (2004), no. 6 905–915.

[102] O. Beker, C. V. Hollot, and Y. Chait, *Plant with integrator: an example of reset control overcoming limitations of linear feedback*, *IEEE Transactions on Automatic Control* **46** (2001), no. 11 1797–1799.

[103] O. Beker, C. Hollot, and Y. Chait, *Stability of a MIMO reset control system under constant inputs*, in *Proceedings of the 38th IEEE Conference on Decision and Control*, vol. 3, pp. 2780–2781, 1999.

[104] W. M. Haddad, V. Chellabsina, and N. Kablar, *Active control of combustion instabilities via hybrid resetting controllers*, in *Proceedings of the 2000 American Control Conference. ACC*, vol. 4, pp. 2378–2382, 2000.

[105] R. Goebel, J. Hespanha, A. R. Teel, C. Cai, and R. Sanfelice, *Hybrid systems: generalized solutions and robust stability*, *IFAC Proceedings Volumes* **37** (2004), no. 13 1–12.

[106] R. Goebel and A. R. Teel, *Solutions to hybrid inclusions via set and graphical convergence with stability theory applications*, *Automatica* **42** (2006), no. 4 573–587.

[107] D. Nešić, L. Zaccarian, and A. R. Teel, *Stability properties of reset systems*, *Automatica* **44** (2008), no. 8 2019–2026.

[108] D. Nesic, A. R. Teel, and L. Zaccarian, *Stability and performance of SISO control systems with first-order reset elements*, *IEEE Transactions on Automatic Control* **56** (2011), no. 11 2567–2582.

[109] J. Zheng, Y. Guo, M. Fu, Y. Wang, and L. Xie, *Improved reset control design for a PZT positioning stage*, in *2007 IEEE International Conference on Control Applications*, pp. 1272–1277, 2007.

[110] A. Fernández, A. Barreiro, A. Banos, and J. Carrasco, *Reset control for passive teleoperation*, in *34th Annual Conference of IEEE Industrial Electronics*, pp. 2935–2940, 2008.

[111] A. R. Teel, *Continuous-time implementation of reset control systems*, in *Trends in Nonlinear and Adaptive Control – A tribute to Laurent Praly for his 65th birthday, Springer*, pp. 27–42, 2021.

[112] R. Bertollo, A. Teel, and L. Zaccarian, "Continuous-time reset control with max-of-quadratics Lyapunov certificates." Submitted.

[113] J. H. Le and A. R. Teel, *Passive soft-reset controllers for nonlinear systems*, in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 5320–5325, 2021.

[114] R. T. Rockafellar and R. J. Wets, *Variational Analysis*, vol. 317. Springer Berlin, Heidelberg, 1998.

[115] M. Dellarossa, A. Tanwani, and L. Zaccarian, *Non-pathological iss-Lyapunov functions for interconnected differential inclusions*, *IEEE Transactions on Automatic Control* (2021).

[116] F. H. Clarke, Y. S. Ledyaev, and R. J. Stern, *Asymptotic stability and smooth Lyapunov functions*, *Journal of differential Equations* **149** (1998), no. 1 69–114.

[117] H. Nakamura, Y. Yamashita, and H. Nishitani, *Smooth Lyapunov functions for homogeneous differential inclusions*, in *Proceedings of the 41st SICE Annual Conference. SICE 2002.*, vol. 3, pp. 1974–1979, 2002.

[118] E. Ryan, *A universal adaptive stabilizer for a class of nonlinear systems*, *Systems & Control Letters* **16** (1991), no. 3 209–218.