Optimization-Based Methods for Nonlinear and Hybrid Systems Verification

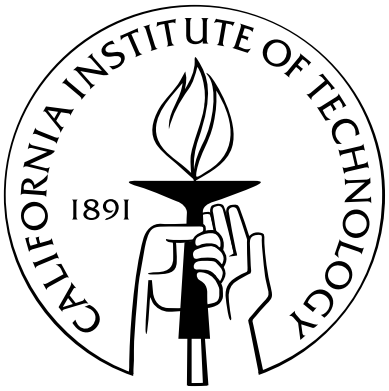
Thesis by

Stephen Prajna

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2005

(Defended January 14, 2005)

Acknowledgements

I would like to express my gratitude to my advisor, Professor John Doyle, for admitting me to Caltech and for providing his guidance and advice during my years at Caltech. Working with John has opened a whole new world to me. I really appreciate the intellectual freedom and the "think outside the box" spirit that he has given to me and his other students.

My acknowledgements also go to Professors Anders Rantzer, Pablo Parrilo, Ali Jadbabaie, and George Pappas for being my mentors. It has been an honor for me to work with them. In addition, I thank them for inviting me to visit Lund, ETH, and Penn at various times during my studies. Back at Caltech, thanks to Professors Richard Murray, Jerrold Marsden, Mani Chandy, and Hideo Mabuchi for serving on my committees and providing constructive feedbacks on my work.

I would like to thank Professors Manfred Morari, Arjan van der Schaft, Torkel Glad, Karl Henrik Johansson, Bjarne Foss, and Raffaello d'Andrea for their hospitality when I visited their institutions in the spring and summer of 2004. I also thank Professor Yaser Abu-Mostafa for supporting one of my conference trips in the same year.

To my friends Antonis Papachristodoulou, Domitilla del Vecchio, Xin Liu, Harish Bhat, Melvin Flores, Melvin Leok, and Shreesh Mysore, and the rest of CDS, thank you for the camaraderie and the good time that we had.

Last but definitely not least, I thank my family for the long distance support that they have given to me.

Abstract

Complex behaviors that can be exhibited by hybrid systems make the verification of such systems both important and challenging. Due to the infinite number of possibilities taken by the continuous state and the uncertainties in the system, exhaustive simulation is impossible, and also computing the set of reachable states is generally intractable. Nevertheless, the ever-increasing presence of hybrid systems in safety critical applications makes it evident that verification is an issue that has to be addressed.

In this thesis, we develop a unified methodology for verifying temporal properties of continuous and hybrid systems. Our framework does not require explicit computation of reachable states. Instead, functions of state termed barrier certificates and density functions are used in conjunction with deductive inference to prove properties such as safety, reachability, eventuality, and their combinations. As a consequence, the proposed methods are directly applicable to systems with nonlinearity, uncertainty, and constraints. Moreover, it is possible to treat safety verification of stochastic systems in a similar fashion, by computing an upper-bound on the probability of reaching the unsafe states.

We formulate verification using barrier certificates and density functions as convex programming problems. For systems with polynomial descriptions, sum of squares optimization can be used to construct polynomial barrier certificates and density functions in a computationally scalable manner. Some examples are presented to illustrate the use of the methods. At the end, the convexity of the problem formulation is also exploited to prove a converse theorem in safety verification using barrier certificates.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Much research effort has been devoted to the development of hybrid systems theory in the recent years. Hybrid systems [48, 92] are systems whose dynamics involve both continuous and discrete processes in interactions. Research on hybrid systems (see, e.g., [5, 7, 8, 25, 45, 51, 52, 89]) is partly motivated by the ubiquity of engineering and physical systems that are best modelled as such systems. One important example is the class of embedded and software-based control systems, which consist of discrete controllers, typically logical and event-based, interconnected with analog and often nonlinear actuators, sensors, and plants. Embedded and software-based systems have become increasingly ubiquitous in our everyday life. In fact, the trend shows that next generation control systems will be mostly of this type [53, 56].

Hybrid systems can exhibit very complex behaviors, which make their analysis both critical and challenging. Simulation is of limited use for analysis, due to the infinite number of possibilities taken by the continuous state and also the uncertainties of the system. Verifying by simulation that a hybrid system works correctly in all cases is never exact, simply because it is impossible to test all system behaviors. In fact, simulation alone may fail to uncover the existence of bad behaviors. Verification of hybrid systems is an area where deductive formal methods, relying on mathematical inferences and proofs to produce exact statements about the system, are indispensable. Formal methods are also needed in system synthesis, particularly

when correctness, robustness, and optimality are of paramount importance, which renders design by informal reasoning combined with trial and error ineffective.

Besides the more traditional properties such as stability and input-output performance, properties of interest in hybrid systems also include safety, reachability, and eventuality. In principle, safety verification aims to show that starting from any initial condition in some prescribed set, a system cannot evolve to some unsafe region in the state space. On the other hand, reachability verification aims to show that for *some* — and eventuality verification for *all* — initial conditions in some prescribed set, the system will evolve to some target region in the state space. The above properties are the most relevant when the system specifications are given in temporal logic formulas [36, 46] such as

> (from a multi-vehicle coordination scenario): "if Agent 1 starts at zone A and Agent 2 starts at zone B, then under the given control strategy,
>
> - Agent 1 will reach zone C in finite time,
> - Agent 2 will not reach zone D before Agent 1 reaches C,
> - both Agent 1 and Agent 2 will never enter a forbidden zone E at any time,"

which is the kind of specifications that seem likely to dominate next generation control systems. These verification questions are by no mean easy to answer, as for very simple classes of hybrid systems they are known to be undecidable already [30].

Scalable automated methods for verification of hybrid systems are definitely in demand. From computer science, there exist comprehensive bodies of techniques for verifying temporal logic formulas for discrete systems; they fall into two mainstream approaches: *model checking* [23] and *deductive verification* [47]. Model checking is applicable to finite state systems, and basically performs an exhaustive exploration of all possible system behaviors in a fully automated way. The drawback of model checking is the state explosion problem, i.e., the number of system trajectories that need to be explored grows very rapidly as the number of states increases, although

the use of an efficient data structure called ordered binary decision diagrams [18] has allowed model checking of systems with an astronomical number of states. Still, when the number of possible states is infinite, such as when the state space is continuous, model checking is no longer applicable. Indeed, the difficulty of applying model checking to hybrid systems is caused by the continuous part of their state space. Deductive verification, on the other hand, verifies system properties through formal deduction based on a set of inference rules. Deductive verification is applicable to infinite state systems, but has a drawback in the sense that guidance from a user is almost always needed in the process.

From control theory, there exist also comprehensive bodies of techniques for verifying properties of continuous systems such as stability, performance, robust stability, robust performance, and so on (see e.g., [40, 98]). These techniques are deductive in nature, since the systems considered have an infinite number of states. If the systems have a special structure (e.g., linear), then the verification can be automated. Unfortunately, the techniques are geared to verify properties that are expressed in terms of Lyapunov stability or signal/system norms, and as such are not directly applicable to verification of properties such as safety, reachability, and eventuality, let alone more general temporal logic formulas.

Naturally, there have been efforts to combine the results from computer science and control theory, to develop methodologies for verifying temporal properties of continuous and hybrid systems. Relevant references will be provided later in this thesis. In our view, however, what is still missing is a unified framework (although such a framework may not necessarily be the only one that can be proposed) that can directly handle systems with hybrid dynamics, nonlinearity, uncertainty, constraints, stochasticity, and so on. Moreover, many of the currently available techniques suffer from computational scalability issues: their computational cost grows exponentially with respect to the system size. Needless to say, the area of hybrid systems verification is still in its infancy, and we expect to see many more developments in the upcoming years.

1.2 Contributions and Outline

The objective of this thesis is to develop unified theoretical and computational frameworks that will facilitate automated verification of properties such as safety, reachability, and eventuality for continuous and hybrid systems. In doing so, we have used theoretical concepts called *barrier certificates* and *density functions*, in addition to a computational relaxation framework called *sum of squares optimization*, which involves sum of squares decompositions of multivariate polynomials, semidefinite programming, and real algebraic geometry. The contributions and outline of the thesis are as follows.

In Chapter 2, we introduce the concept of barrier certificates and propose using them for safety verification of continuous and hybrid systems in the worst-case setting. A barrier certificate is a function (or a set of functions) of state satisfying some inequalities on both the function itself and its derivative along the flow of the system. In this setting, a barrier certificate proves that all possible system trajectories starting from a given initial set cannot reach a given unsafe region. The use of barrier certificates for verifying safety is analogous to the use of Lyapunov functions for proving stability, and eliminates the need to propagate sets of states. As a consequence, our approach is directly applicable to systems with nonlinearity, uncertainty, constraints, and hybrid dynamics. We also propose using a class of convex relaxation, i.e., sum of squares optimization, to compute barrier certificates for systems whose descriptions are in terms of polynomials. Sum of squares optimization provides a hierarchical way to search for barrier certificates, where at each level the computational cost grows polynomially with respect to the system size. Because of this, our methodology seems to be more scalable than many other existing methods that can handle nonlinear continuous and hybrid systems. This chapter is based on the papers [65, 66].

When stochastic disturbance input, or stochastic discrete transition, or both are present in the system, answering the safety verification question in the worst-case setting usually leads to a very conservative answer, i.e., to the conclusion that the

system is not safe. Indeed, it is more natural to consider safety verification with probabilistic interpretation, e.g., to prove that the probability of reaching the unsafe set is lower than some safety margin. This is the subject of Chapter 3. Our method uses supermartingales as barrier certificates and upper-bounds the reach probability using a certain supermartingale inequality. The method is applicable to a large class of stochastic continuous and hybrid systems with polynomial descriptions, and is the first proposed computational method that can provide a verifiable upper bound on the reach probability. This chapter is based on the paper [67].

In Chapter 4, we consider the duality relation between proving safety and reachability. Using insights from the linear programming duality appearing in the discrete shortest path problem and the concept of density functions, we show that proving reachability or eventuality in continuous systems can also be performed by solving a convex optimization problem. Convex programs involving barrier certificates and density functions for verifying safety, reachability, eventuality, and some other temporal specifications are formulated. The chapter is based on the paper [76].

In Chapter 5, the duality relation between safety and reachability is used to prove a converse theorem for safety verification using barrier certificates. Under reasonable technical conditions, we prove that there exists a barrier certificate for a nonlinear continuous system if and only if the safety property holds. The chapter is based on the paper [75].

We end the thesis in Chapter 6 by presenting some conclusions and suggestions for future research.

1.3 Notations

We denote the set of real numbers by $\mathbb{R}$ and the Euclidean n-space by $\mathbb{R}^n$. The trace of an $n \times n$ matrix M, i.e., the sum of its diagonal elements, is denoted by $\mathrm{Tr}(M)$. In addition, we use $\mathrm{int}(X)$, $\mathrm{cl}(X)$, and ∂X to denote the interior, the closure, and the boundary of a set $X \subseteq \mathbb{R}^n$.

By $f : X \to Y$ we mean a function f mapping $X \subseteq \mathbb{R}^n$ to $Y \subseteq \mathbb{R}^m$. We denote

the spaces of k-times continuously differentiable functions mapping $X \subseteq \mathbb{R}^n$ to $\mathbb{R}^m$ by $C^k(X, \mathbb{R}^m)$, and when $m = 1$, we will write $C^k(X)$. Correspondingly, the spaces of continuous functions on X are denoted by $C(X, \mathbb{R}^m)$ and $C(X)$, equipped with the supremum norm if necessary. The zero subscript as in $C_0^1(\mathbb{R}^n)$ indicates that the functions have compact supports. The dual space of a normed linear space $\mathcal{K}$, i.e., the space of all continuous linear functionals on $\mathcal{K}$, is denoted by $\mathcal{K}^*$. By $\langle k^*, k \rangle$ we mean the value of a continuous linear functional $k^* \in \mathcal{K}^*$ applied to $k \in \mathcal{K}$.

For a differentiable function $F : \mathbb{R}^n \to \mathbb{R}$, we define

$$\frac{\partial F}{\partial x}(x) \triangleq \left[\frac{\partial F}{\partial x_1}(x) \quad \cdots \quad \frac{\partial F}{\partial x_n}(x) \right],$$

and

$$\frac{\partial^2 F}{\partial x^2}(x) \triangleq \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_1 \partial x_n}(x) & \cdots & \frac{\partial^2 F}{\partial x_n^2}(x) \end{bmatrix}.$$

The divergence of a differentiable vector field $f : \mathbb{R}^n \to \mathbb{R}^n$,

$$\frac{\partial f_1}{\partial x_1}(x) + ... + \frac{\partial f_n}{\partial x_n}(x),$$

is denoted by $\nabla \cdot f(x)$. The flow of $\dot{x} = f(x)$ starting at x_0 is denoted by $\phi_t(x_0)$. For a set $Z \subseteq \mathbb{R}^n$, we define $\phi_t(Z) \triangleq \{\phi_t(x_0) : x_0 \in Z\}$.

Finally, $P\{ \cdot \}$ and $P\{ \cdot \mid \cdot \}$ denote the total and conditional probability, respectively, whereas $E[\cdot]$ and $E[\cdot \mid \cdot]$ denote the total and conditional expectation.

Chapter 2

Worst-Case Safety Verification

In this chapter, we consider safety verification of nonlinear continuous and hybrid systems in the worst-case setting. Some disturbance signal and model uncertainty may also be included in the system description. We want to verify that under any circumstances, there is no trajectory of the system that starts from a given set of possible initial states and goes to an unsafe region in the state space. Such analysis is particularly important for safety critical systems like air traffic control [90], autonomous vehicle systems [33], and life support systems [28].

For safety verification of continuous and hybrid systems, several methods have been proposed. Explicit computation of either exact or approximate reachable sets corresponding to the continuous dynamics is crucial for most of these methods. For linear continuous systems with certain eigenvalue structures and semialgebraic initial sets, *exact* reachable set calculation using quantifier elimination has been proposed in [6, 43]. Unfortunately, their approach requires knowing the exact solution of the differential equations, and hence does not seem extendable to the nonlinear case. In another vein, several techniques have also been developed for *approximate* reachable set calculation. For linear systems, there are results based on quantifier elimination [86], ellipsoidal calculus [16, 41], polygonal approximation [10, 13], geometric programming [96], and real algebraic geometry [97]. Other techniques have been proposed for nonlinear systems, for example, based on the Hamilton Jacobi equations [91], polygonal approximations [22], and approximating the system as a piecewise linear system [9]. In the case of hybrid systems, most of the techniques are based on con-

structing abstractions (i.e., discrete quotients) of the systems, and then performing model checking on the resulting discrete systems. See for instance [2,4,10,13,22,87,91].

We will present a method for safety verification that is different from the above approaches as it does not require computation of reachable sets, but instead relies on what we term barrier certificates. For a continuous system, a barrier certificate is a function of state satisfying a set of inequalities on both the function itself and its Lie derivative along the flow of the system. In the state space, the zero level set of a barrier certificate separates an unsafe region from all system trajectories starting from a set of possible initial states. Therefore, the existence of such a function provides an exact certificate/proof of system safety.

Similar to the Lyapunov approach for proving stability, the main idea here is to study properties of the system without the need to compute the flow explicitly. Although an over-approximation of the reachable set may also be used as a proof for safety, a barrier certificate can be much easier to compute when the system is nonlinear and uncertain. Moreover, barrier certificates can be easily used to verify safety in infinite time horizon. Note also that there are some connections between our method and viability theory [11], invariant set theory [11,14], and also the verification approaches in [37, 83, 88]. We will discuss these connections later as we progress.

Our method can be easily extended to handle hybrid systems. In the hybrid case, a barrier certificate is constructed from a set of functions of continuous state indexed by the system location[1]. Instead of satisfying the aforementioned inequalities in the whole continuous state space, each function needs to satisfy the inequalities only within the invariant of the location. Functions corresponding to different locations are linked via appropriate conditions that must be satisfied during discrete transitions between the locations. The idea is analogous to using multiple Lyapunov-like functions [38] for stability analysis of hybrid systems.

With this methodology, it is possible to treat a large class of hybrid systems, including those with nonlinear continuous dynamics, uncertainty, and constraints. When the vector fields of the system are polynomials and the sets in the system de-

[1]The term "location" here means discrete state; cf. Section 2.2.1.

scription are semialgebraic (i.e., described by polynomial equalities and inequalities), a tractable computational method called sum of squares optimization [61, 62, 69, 72] can be utilized for constructing a polynomial barrier certificate, e.g., using the software SOSTOOLS [69,72]. While the computational cost of this construction depends on the degrees of the vector fields and the barrier certificate in addition to the number of discrete locations and the continuous state dimension, for fixed polynomial degrees the complexity grows polynomially with respect to the other quantities. Hence, we expect our method to be more scalable than many other existing methods. Successful application of our method to a NASA life support system, which is a nonlinear hybrid system with six discrete modes and ten continuous state variables, has been reported in [28].

This chapter is organized as follows. In Section 2.1, safety verification of continuous systems is addressed. We present some conditions for barrier certificates which guarantee the safety of the system. Later in the same section, we incorporate constraints into the framework. Safety verification of hybrid systems is then addressed in Section 2.2. Section 2.3 is devoted to computation of barrier certificates. Finally, Section 2.4 contains some examples illustrating the use of the methodology.

2.1 Continuous Systems

2.1.1 Convex Conditions

In this section, we address safety verification of continuous systems, to establish a foundation for the subsequent results. Consider a continuous system described by a set of ordinary differential equations in the state space form:

$$\dot{x}(t) = f(x(t), d(t)), \tag{2.1}$$

with the state $x(t)$ taking its value in $\mathbb{R}^n$ and the disturbance input $d(t)$ taking its value in $\mathcal{D} \subseteq \mathbb{R}^m$. Here, the signal $d(t)$ is assumed to be piecewise continuous and bounded on any finite time interval. Some smoothness conditions will be imposed on

the vector field $f(x, d)$. At the least it will be continuous, which makes $x(t)$ piecewise continuously differentiable.

In safety verification, only parts of trajectories that are contained in a given set $\mathcal{X} \subseteq \mathbb{R}^n$ and that start from a given set of possible initial states $\mathcal{X}_0 \subseteq \mathcal{X}$ are considered. We denote the unsafe region of the system by $\mathcal{X}_u$, with $\mathcal{X}_u \subseteq \mathcal{X}$. With these notations, the safety property in the worst-case setting can be defined as follows. The definition can be directly extended for other classes of systems as needed.

Definition 2.1 (Safety) *Given the system (2.1), the state set $\mathcal{X} \subseteq \mathbb{R}^n$, the initial set $\mathcal{X}_0 \subseteq \mathcal{X}$, the unsafe set $\mathcal{X}_u \subseteq \mathcal{X}$, and the disturbance set $\mathcal{D} \subseteq \mathbb{R}^m$, we say that the* safety *property holds if there exist no time instant $T \geq 0$ and a piecewise continuous and bounded disturbance $d : [0, T] \to \mathcal{D}$ that gives rise to an unsafe system trajectory, i.e., a trajectory $x : [0, T] \to \mathbb{R}^n$ satisfying $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$, and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$.*

Our method for verifying safety relies on the existence of what we will call barrier certificate. For continuous systems, the following proposition states the conditions that are satisfied by a barrier certificate.

Proposition 2.2 *Let the system $\dot{x} = f(x, d)$ and the sets $\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{D} \subseteq \mathbb{R}^m$ be given, with $f \in C(\mathbb{R}^{n+m}, \mathbb{R}^n)$. Suppose there exists a differentiable function $B : \mathbb{R}^n \to \mathbb{R}$ such that*

$$B(x) \leq 0 \qquad \forall x \in \mathcal{X}_0, \tag{2.2}$$

$$B(x) > 0 \qquad \forall x \in \mathcal{X}_u, \tag{2.3}$$

$$\frac{\partial B}{\partial x}(x) f(x, d) \leq 0 \qquad \forall (x, d) \in \mathcal{X} \times \mathcal{D}, \tag{2.4}$$

then the safety of the system in the sense of Definition 2.1 is guaranteed.

Proof. Our proof is by contradiction. Assume that there exists a barrier certificate $B(x)$ satisfying conditions (2.2)–(2.4), while at the same time the system is not safe, i.e., there exist a time instance $T \geq 0$, a disturbance signal $d : [0, T] \to \mathcal{D}$, and

an initial condition $x_0 \in \mathcal{X}_0$ such that a trajectory $x(t)$ of the system starting at $x(0) = x_0$ satisfies $x(t) \in \mathcal{X}$ for all $t \in [0, T]$ and $x(T) \in \mathcal{X}_u$. Condition (2.4) implies that the derivative of $B(x(t))$ with respect to time is non-positive on the time interval $[0, T]$. A direct consequence of this (which for example can be shown using the mean value theorem) is that $B(x(T))$ must be less than or equal to $B(x(0))$, which is contradictory to (2.2)–(2.3). Thus the initial hypothesis is not correct: the system must be safe. ■

A function $B(x)$ satisfying the conditions in Proposition 2.2 is termed a barrier certificate. The zero level set of a barrier certificate "provides a barrier" between possible system trajectories and the given unsafe region, in the sense that no trajectory of the system starting from the initial set can cross this level set to reach the unsafe region (cf. Section 2.4.1 for a visual illustration). In proving that the system is safe, no explicit computation of system trajectories nor reachable sets is required.

In the above proposition, we have assumed that the unknown disturbance input can vary arbitrarily fast. If the variation of the disturbance is bounded (for example, when there are uncertain parameters, which can be regarded as time-invariant disturbance), then a less conservative verification can be performed by considering a barrier certificate $B(x, d)$ that also depends on the instantaneous value of the disturbance and modifying (2.2)–(2.4) accordingly. For example, in (2.4) we need to take into account the extra derivative term $\frac{\partial B}{\partial d}(x, d)\dot{d}$, with the disturbance variation $\dot{d}$ taking its value in some bounded set.

Note that the set of barrier certificates satisfying the conditions in Proposition 2.2 is convex. This can be established by taking arbitrary $B_1(x)$ and $B_2(x)$ satisfying the above conditions and showing that for all $\alpha \in [0, 1]$, $B(x) = \alpha B_1(x) + (1 - \alpha)B_2(x)$ satisfies the conditions as well. The convexity property is very beneficial for the computation of $B(x)$. As we will see later in Section 2.3, a barrier certificate $B(x)$ in this convex set can be searched directly using convex optimization.

Since the set $\{x \in \mathcal{X} : B(x) \leq 0\}$ is actually an invariant set, the method presented above is closely related to the smallest invariant set approach for safety verification (see, e.g., [37]). The latter approach differs from ours in that it tries to

compute the smallest invariant set that contains $\mathcal{X}_0$, and then show that this set does not intersect $\mathcal{X}_u$. However, among invariant sets whose descriptions have *bounded complexity* (e.g., sets described using finite degree polynomials), the smallest set may not be one that does not intersect $\mathcal{X}_u$. Not only that, such smallest invariant set may be very difficult to find and may not be unique. Our approach, on the other hand, uses an arbitrary invariant set containing $\mathcal{X}_0$ that does not intersect $\mathcal{X}_u$. As such, our method is computationally much easier than the smallest invariant set approach.

We would like to remark that other approaches similar to ours are also presented in [83, 88]. These papers address the verification problem from a computer science point of view, and proposes methods for constructing *invariants* of the system. An invariant here is a property that holds for every reachable state of the system. Thus, in the barrier certificate framework, for example, $B(x) \leq 0$ is an invariant of the system. The difference is that their conditions for the invariants are more restrictive than ours, and the invariants are not computed using convex optimization, but instead using Gröbner basis method followed by solving a system of linear equations.

2.1.2 Non-Convex Conditions

Although the conditions in Proposition 2.2 are good for computation since they define a convex set of barrier certificates, the conditions seem rather conservative (i.e., within a class of barrier certificates with bounded complexity) as the derivative inequality (2.4) needs to be satisfied on the whole state set $\mathcal{X}$. It is natural to expect that the conditions can be relaxed by requiring a similar derivative inequality to hold only on and near the set of $x \in \mathcal{X}$ for which $B(x) = 0$. This kind of condition is used in Proposition 2.3 below. Unfortunately, the set of barrier certificates will no longer be convex, hence a direct computation of a barrier certificate using convex optimization is not possible, although we can still search for a barrier certificate in the non-convex set using an iterative method, as we will see in Section 2.3.3.

Proposition 2.3 *Let the system $\dot{x} = f(x, d)$ and the sets $\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{D} \subseteq \mathbb{R}^m$ be given, with $f \in C(\mathbb{R}^{n+m}, \mathbb{R}^n)$. If there exists a function $B \in C^1(\mathbb{R}^n)$ that*

satisfies the following conditions:

$$B(x) \leq 0 \qquad \forall x \in \mathcal{X}_0, \tag{2.5}$$

$$B(x) > 0 \qquad \forall x \in \mathcal{X}_u, \tag{2.6}$$

$$\frac{\partial B}{\partial x}(x) f(x,d) < 0 \qquad \forall (x,d) \in \mathcal{X} \times \mathcal{D} \text{ such that } B(x) = 0, \tag{2.7}$$

then the safety of the system in the sense of Definition 2.1 is guaranteed.

Proof. Consider $T > 0$, as the case where $T = 0$ is trivial. Suppose that a disturbance signal $d : [0,T] \to \mathcal{D}$ and a corresponding unsafe trajectory $x : [0,T] \to \mathcal{X}$ exist. Let t_1 and t_2 be two time instants such that $0 \leq t_1 < t_2 \leq T$, $B(x(t_1)) \leq 0$, $B(x(t_2)) \geq 0$, and

$$\frac{\partial B}{\partial x}(x(t)) f(x(t), d(t)) < 0 \quad \forall t \in [t_1, t_2].$$

Now integrate $\frac{\partial B}{\partial x}(x(t)) f(x(t), d(t))$ over the time interval $[t_1, t_2]$ to obtain a contradiction, thus proving that the system is safe. ■

The above proposition is sufficient for our purposes and its proof is also straightforward. However, it is interesting to note that other (non-convex) conditions can be derived using viability theory [11]. Interested readers are referred to the appendix in Section 2.5.

2.1.3 Incorporating Constraints

The method we have proposed in Section 2.1.1 can be extended to accommodate a larger class of systems. Consider the following system:

$$\dot{x}(t) = f(x(t), v(t)), \tag{2.8}$$

$$0 = g(x(t), v(t)), \tag{2.9}$$

$$0 \leq h(x(t), v(t)), \tag{2.10}$$

$$0 \leq \int_0^t \sigma(x(\tau), v(\tau)) d\tau \quad \forall t \geq 0, \tag{2.11}$$

where $x(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ is the state vector, and $v(t) \in \mathcal{V} \subseteq \mathbb{R}^m$ is a vector of auxiliary variables, which may include disturbance inputs. We assume that $v(t)$ is piecewise continuous along time, and that $f(x, v)$, $g(x, v)$, $h(x, v)$, and $\sigma(x, v)$ are continuous in their arguments. In general they will be vector-valued functions, for which the equality and inequality in (2.9)–(2.11) are interpreted entry-wise.

Note that the above formulation includes a very large class of systems, for example:

- Systems described by differential-algebraic equations (DAEs) can be accommodated by including the equality constraints (2.9) in the formulation.
- Memoryless uncertainties [40] relating some signals in the system can be taken into account by the inequality constraints (2.10).
- Uncertain time-varying inputs can be characterized using (2.10) for inputs with bounded magnitude, or (2.11) for inputs with bounded energy.
- Some classes of dynamic uncertainties can be described using hard[2] *integral quadratic constraints* (IQCs) [49], which is a special case of (2.11).

More importantly, their combinations clearly can still be described by (2.8)–(2.11).

First, we need to specify what is considered as a valid trajectory of the system. A trajectory $x : [0, T] \to \mathcal{X}$ is a valid trajectory of the system (2.8)–(2.11) on the time interval $[0, T]$ if there exists a piecewise continuous and bounded $v : [0, T] \to \mathcal{V}$ such that $x(t)$ is a solution of the differential equations (2.8), and the constraints (2.9)–(2.11) are satisfied by $x(t)$ and $v(t)$ for all $t \in [0, T]$. Since the vector field $f(x, v)$ is continuous, $x(t)$ will be piecewise continuously differentiable.

Similar to before, in the safety verification we will denote the initial set by $\mathcal{X}_0$ and the unsafe set by $\mathcal{X}_u$. The safety property for this system is defined as follows.

Definition 2.4 (Safety – Constrained Systems) *Given the system (2.8)–(2.11), the state set $\mathcal{X} \subseteq \mathbb{R}^n$, the initial set $\mathcal{X}_0 \subseteq \mathcal{X}$, the unsafe set $\mathcal{X}_u \subseteq \mathcal{X}$, and the set*

[2]The notion "hard" here means that the constraint must be satisfied for all $t \geq 0$; a "soft" integral constraint has the form $\int_0^\infty \sigma(x(\tau), d(\tau))d\tau \geq 0$. Some important integral constraints for robustness analysis of uncertain systems [49] are soft constraints.

$\mathcal{V} \subseteq \mathbb{R}^m$, *we say that the* safety *property holds if there exist no time instant* $T \geq 0$ *and a piecewise continuous and bounded signal* $v : [0, T] \to \mathcal{V}$ *that gives rise to an unsafe system trajectory, i.e., a trajectory* $x : [0, T] \to \mathbb{R}^n$ *such that (2.9)–(2.11) are satisfied by* $(x(t), v(t))$ *for all* $t \in [0, T]$*, and also* $x(0) \in \mathcal{X}_0$*,* $x(T) \in \mathcal{X}_u$*, and* $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$*.*

For handling this class of systems, we will multiply $g(x, v)$, $h(x, v)$, and $\sigma(x, v)$ given in (2.9)–(2.11) by some function multipliers satisfying certain positivity criteria, and add the products to the derivative condition that must be satisfied by the barrier certificate. This can be regarded as a generalization of the so-called *S-procedure* (in which the multipliers are constants; see [95]), and has been proposed in [58] for constructing Lyapunov functions for systems described by (2.8)–(2.11).

Proposition 2.5 *Let the system (2.8)–(2.11) and the sets* $\mathcal{X} \subseteq \mathbb{R}^n$*,* $\mathcal{X}_0 \subseteq \mathcal{X}$*,* $\mathcal{X}_u \subseteq \mathcal{X}$*,* $\mathcal{V} \subseteq \mathbb{R}^m$ *be given, with* $f \in C(\mathbb{R}^{n+m}, \mathbb{R}^n)$*,* $g \in C(\mathbb{R}^{n+m}, \mathbb{R}^p)$*,* $h \in C(\mathbb{R}^{n+m}, \mathbb{R}^q)$*,* $\sigma \in C(\mathbb{R}^{n+m}, \mathbb{R}^r)$*. Suppose there exist a function* $B \in C^1(\mathbb{R}^n)$*, function multipliers* $\lambda_1 \in C(\mathbb{R}^{n+m}, \mathbb{R}^p)$*,* $\lambda_2 \in C(\mathbb{R}^{n+m}, \mathbb{R}^q)$*, and constant multiplier* $\lambda_3 \in \mathbb{R}^r$ *such that*[3]

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{2.12}$$

$$B(x) > 0 \quad \forall x \in \mathcal{X}_u, \tag{2.13}$$

$$\frac{\partial B}{\partial x}(x) f(x, v) + \lambda_1^T(x, v) g(x, v) + \lambda_2^T(x, v) h(x, v) + \lambda_3^T \sigma(x, v) \leq 0 \quad \forall (x, v) \in \mathcal{X} \times \mathcal{V}, \tag{2.14}$$

$$\lambda_2(x, v) \geq 0 \quad \forall (x, v) \in \mathcal{X} \times \mathcal{V}, \tag{2.15}$$

$$\lambda_3 \geq 0. \tag{2.16}$$

Then the safety of the system in the sense of Definition 2.4 is guaranteed.

Proof. The proof is similar to the proof of Proposition 2.2, except that here the conditions (2.14)–(2.16) will be used to show that $B(x(t))$ is non-increasing along time. That can be shown directly by integrating the left hand side of (2.14) with

[3]Note that the inequalities (2.15)–(2.16) are interpreted entry-wise.

respect to time and using the fact that

$$\int_0^t \left[\lambda_1^T(x(\tau),v(\tau))g(x(\tau),v(\tau)) + \lambda_2^T(x(\tau),v(\tau))h(x(\tau),v(\tau)) + \lambda_3^T\sigma(x(\tau),v(\tau))\right] d\tau$$

is non-negative for $t \in [0,T]$, which follows from (2.9)–(2.11) and (2.15)–(2.16). ■

2.2 Hybrid Systems

2.2.1 Modelling Framework

Throughout this section, we adopt the hybrid modelling framework that was first proposed in [1]; see also [3] for a more detailed explanation and example. A hybrid system is a tuple $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$ with the following components:

- $\mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state space.
- L is a finite set of locations. The overall state space of the system is $X = L \times \mathcal{X}$, and a state of the system is denoted by $(l, x) \in L \times \mathcal{X}$.
- $X_0 \subseteq X$ is the set of initial states.
- $I : L \to 2^{\mathcal{X}}$ is the invariant, which assigns to each location l a set $I(l) \subseteq \mathcal{X}$ that contains all possible continuous states while at location l.
- $F : X \to 2^{\mathbb{R}^n}$ is a set of vector fields. F assigns to each $(l, x) \in X$ a set $F(l, x) \subseteq \mathbb{R}^n$ which constrains the evolution of the continuous state according to the differential inclusion $\dot{x}(t) \in F(l(t), x(t))$.
- $\mathcal{T} \subseteq X \times X$ is a relation capturing discrete transitions between two locations. A transition $((l, x), (l', x')) \in \mathcal{T}$ indicates that from the state (l, x) the system can undergo a discrete jump to the state (l', x').

Valid trajectories of the hybrid system H start at some initial state $(l_0, x_0) \in X_0$ and are concatenations of a sequence of continuous flows and discrete transitions. During a continuous flow, the discrete location l is maintained and the continuous

state evolves according to the differential inclusion $\dot{x}(t) \in F(l(t), x(t))$, with $x(t)$ remains inside the invariant set $I(l(t))$. For our purpose, we will model the uncertainty in the continuous flow by some disturbance inputs in the following manner:

$$F(l, x) = \{\dot{x} \in \mathbb{R}^n : \dot{x} = f_l(x, d) \text{ for some } d \in \mathcal{D}(l)\},$$

where $f_l(x, d)$ is a vector field that governs the flow of the system at location l, and $d(t)$ is a vector of disturbance inputs that takes value in the set $\mathcal{D}(l(t)) \subseteq \mathbb{R}^m$. We assume that $d(t)$ is piecewise continuous and bounded on any finite time interval, and that $f_l \in C(\mathbb{R}^{n+m}, \mathbb{R}^n)$ for all $l \in L$. Finally, at a state (l_1, x_1), a discrete transition to (l_2, x_2) can occur if $((l_1, x_1), (l_2, x_2)) \in \mathcal{T}$. We assume non-determinism in the discrete transition, i.e., the transition may or may not occur, but no stochastic characterization is used or given.

Given a hybrid system H and a set of unsafe states $X_u \subseteq X$, the safety verification problem is concerned with proving that all valid trajectories of the hybrid system H cannot enter the unsafe region X_u. More specifically, the safety property is defined as follows.

Definition 2.6 (Safety – Hybrid Systems) *Given a hybrid system H and an unsafe set $X_u \subseteq X$, the safety property holds if there exist no time instant $T \geq 0$, a piecewise continuous and bounded disturbance input $d : [0, T] \to \mathbb{R}^m$, and a finite sequence of transition times $0 \leq t_1 \leq t_2 \leq \ldots \leq t_N \leq T$ that give rise to an unsafe system trajectory, i.e., a trajectory $(l, x) : [0, T] \to X$ satisfying $(l(0), x(0)) \in X_0$, $x(t) \in I(l(t))$ for $t \in [0, T]$, and $(l(T), x(T)) \in X_u$. (Note that the disturbance input here must also satisfy $d(t) \in \mathcal{D}(l(t))$ for all $t \in [0, T]$.)*

In our analysis conditions, we will also need the following definitions. For each location $l \in L$, the sets of initial and unsafe continuous states are defined as, respec-

tively,

$$\text{Init}(l) = \{x \in \mathcal{X} : (l, x) \in X_0\},$$
$$\text{Unsafe}(l) = \{x \in \mathcal{X} : (l, x) \in X_u\},$$

both of which can be empty. To each tuple $(l, l') \in L^2$ with $l \neq l'$, we associate a guard set

$$\text{Guard}(l, l') = \{x \in \mathcal{X} : ((l, x), (l', x')) \in \mathcal{T} \text{ for some } x' \in \mathcal{X}\},$$

which is the set of continuous states from which the system can undergo a transition from location l to location l', and a (possibly set valued) reset map

$$\text{Reset}(l, l') : x \mapsto \{x' \in \mathcal{X} : ((l, x), (l', x')) \in \mathcal{T}\},$$

whose domain is $\text{Guard}(l, l')$. Obviously, if no discrete transition from location l to location l' is possible, then $\text{Guard}(l, l')$ will be regarded as empty, and the associated reset map needs not be defined.

2.2.2 Conditions for Safety

Verification of hybrid systems should use a barrier certificate that not only is a function of the continuous state, but also depends on the discrete location. For this purpose, we construct a barrier certificate from a set of functions of continuous state, where each function corresponds to a discrete location of the system. Since in each location the continuous state can only take value within the invariant of the location, each function only needs to satisfy inequalities similar to (2.2)–(2.4) or (2.5)–(2.7) in the invariant associated to its location. Functions corresponding to different locations are linked via appropriate conditions that take care of possible discrete transitions between the locations. An analogous idea was used in stability analysis of affine hybrid systems using piecewise quadratic Lyapunov functions [38, 63], and analysis of

polynomial hybrid systems using piecewise polynomial Lyapunov functions [68].

We state the conditions that must be satisfied by the barrier certificate in the following theorem. The notations and assumptions imposed on the system are as described in Section 2.2.1.

Theorem 2.7 *Let the hybrid system $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$ and the unsafe set $X_u \subseteq X$ be given. Suppose there exists a collection $\{B_l(x) : l \in L\}$ of functions $B_l \in C^1(\mathbb{R}^n)$ which, for all $l \in L$ and $(l, l') \in L^2$, $l \neq l'$, satisfy*

$$B_l(x) \leq 0 \quad \forall x \in \text{Init}(l), \tag{2.17}$$

$$B_l(x) > 0 \quad \forall x \in \text{Unsafe}(l), \tag{2.18}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x, d) < 0 \quad \forall (x, d) \in I(l) \times \mathcal{D}(l) \text{ such that } B_l(x) = 0, \tag{2.19}$$

$$B_{l'}(x') \leq 0 \quad \forall x' \in \text{Reset}(l, l')(x), \text{ for all } x \in \text{Guard}(l, l') \text{ s.t. } B_l(x) \leq 0. \tag{2.20}$$

Then the safety of the system in the sense of Definition 2.6 is guaranteed.

Proof. Assume that a barrier certificate $\{B_l(x) : l \in L\}$ satisfying the above conditions can be found. Take any trajectory of the hybrid system that starts at arbitrary $(l_0, x_0) \in X_0$, and consider the evolution of $B_{l(t)}(x(t))$ along this trajectory. Condition (2.17) asserts that $B_{l_0}(x_0) \leq 0$. Next, (2.19) implies that during a segment of continuous flow $B_{l(t)}(x(t))$ cannot become positive, which can be shown using Proposition 2.3. On the other hand, (2.20) guarantees that $B_{l(t)}(x(t))$ cannot jump to a positive value during a discrete transition. Consequently, any such trajectory can never reach an unsafe state $(l_u, x_u) \in X_u$, whose $B_{l_u}(x_u)$ is positive according to (2.17). We conclude that the safety of the system is guaranteed. ■

Similar to what we encounter in the continuous case, conditions (2.19)–(2.20) in the above theorem define a non-convex set of barrier certificates. Conditions defining a convex set of barrier certificates are given in the following theorem.

Theorem 2.8 *Let the hybrid system $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$, the unsafe set $X_u \subseteq X$, and a collection of nonnegative constants $\{\lambda_{l,l'} \in \mathbb{R} : (l, l') \in L^2,\ l \neq l'\}$ be given.*

Suppose there exists a collection $\{B_l(x) : l \in L\}$ *of differentiable functions* $B_l : \mathbb{R}^n \to \mathbb{R}$ *which, for all* $l \in L$ *and* $(l, l') \in L^2$, $l \neq l'$, *satisfy*

$$B_l(x) \leq 0 \quad \forall x \in \text{Init}(l), \tag{2.21}$$

$$B_l(x) > 0 \quad \forall x \in \text{Unsafe}(l), \tag{2.22}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x, d) \leq 0 \quad \forall (x, d) \in I(l) \times \mathcal{D}(l), \tag{2.23}$$

$$B_{l'}(x') - \lambda_{l,l'} B_l(x) \leq 0 \quad \forall x' \in \text{Reset}(l, l')(x), \text{ for all } x \in \text{Guard}(l, l'). \tag{2.24}$$

Then the safety of the system in the sense of Definition 2.6 is guaranteed.

Proof. Analogous to the proof of Theorem 2.7, but with Proposition 2.2 now being used to show that $B_{l(t)}(x(t))$ cannot become positive during a segment of continuous flow. ■

Remark 2.9 *The convexity of the set of barrier certificates in Theorem 2.8 can be established by taking two arbitrary collections* $\{B_l^1(x) : l \in L\}$ *and* $\{B_l^2(x) : l \in L\}$ *satisfying the conditions in the theorem and showing that for all* $\alpha \in [0, 1]$ *the collection* $\{\alpha B_l^1(x) + (1 - \alpha) B_l^2(x) : l \in L\}$ *satisfies the conditions as well. Note that for this convexity, it is crucial that the multipliers* $\lambda_{l,l'}$ *are fixed in advance.*

Remark 2.10 *Two possible choices for* $\lambda_{l,l'}$ *are 0 and 1. The choice* $\lambda_{l,l'} = 0$ *corresponds to modifying (2.20) to*

$$B_{l'}(x') \leq 0 \quad \forall x' \in \mathit{Reset}(l, l')(x), \text{ for some } l \in L \text{ and } x \in \mathit{Guard}(l, l'),$$

and in this case, a successful verification will actually prove that the system is safe even if during a transition from location l *to* l' *the continuous state is allowed to jump to any continuous state* x' *in the image of the reset map. On the other hand, choosing* $\lambda_{l,l'} = 1$ *is useful for handling integral constraints, as we will shortly see.*

2.2.3 Hybrid Systems with Constraints

In the remainder of this section, we will briefly discuss how constraints can be incorporated in verification of hybrid systems. Similar to the continuous case (cf. Section 2.1.3), there are three kinds of constraints that can be handled: algebraic equality, algebraic inequality, and integral constraints. Here we will focus on integral constraints, as verification by explicit calculation of reachable sets is the most difficult when such constraints exist. To the best of our knowledge, the only existing literature addressing this problem is [39], in which a method for bounding an image of the flow map between two affine switching surfaces for affine hybrid systems with integral quadratic constraints is presented.

Instead of assuming that the disturbance $d(t)$ is contained in $\mathcal{D}(l(t))$, suppose now that $d(t)$ and the continuous state $x(t)$ is constrained via a hard integral constraint:

$$\int_0^t \sigma(x(\tau), d(\tau))d\tau \geq 0 \quad \forall t > 0, \tag{2.25}$$

where $d(t)$ is again assumed to be piecewise continuous and bounded on any finite time interval. Apart from this change, valid trajectories of the system are generated in the same manner as in Section 2.2.1. Conditions guaranteeing safety when an integral constraint is present are given in the following theorem.

Theorem 2.11 *Let the hybrid system $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$, the unsafe set $X_u \subseteq X$, and the constraint (2.25) be given, with $\sigma \in C(\mathbb{R}^{n+m}, \mathbb{R}^r)$. Suppose there exist a collection $\{B_l(x) : l \in L\}$ of functions $B_l \in C^1(\mathbb{R}^n)$ and a constant multiplier $\lambda \in \mathbb{R}^r$ that satisfy*

$$B_l(x) \leq 0 \quad \forall x \in \text{Init}(l), \tag{2.26}$$

$$B_l(x) > 0 \quad \forall x \in \text{Unsafe}(l), \tag{2.27}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x, d) + \lambda^T \sigma(x, d) \leq 0 \quad \forall (x, d) \in I(l) \times \mathbb{R}^m, \tag{2.28}$$

$$B_{l'}(x') - B_l(x) \leq 0 \quad \forall x' \in \text{Reset}(l, l')(x), \text{ for all } x \in \text{Guard}(l, l'), \tag{2.29}$$

$$\lambda \geq 0, \tag{2.30}$$

for all $l \in L$ and $(l, l') \in L^2$, $l' \neq l$. Then the safety of the system is guaranteed in the sense of Definition 2.6 (except that $d(t)$ is not contained in $\mathcal{D}(l(t))$, but instead must satisfy (2.25)).

Proof. Assume that a barrier certificate satisfying the above conditions can be found, but at the same time there exists a $T \geq 0$ and a valid trajectory of the hybrid system on the time interval $[0, T]$ such that $(l(T), x(T)) \in X_u$. Assume that discrete transitions for this trajectory occur at time t_1, t_2, ..., t_N where the system switches to location l_1, l_2, ..., l_N. Denote the continuous states before and after the i-th transition by x_i^- and x_i^+, respectively. Then, from (2.28) and (2.30) we obtain

$$\begin{aligned}
&B_{l_0}(x_1^-) - B_{l_0}(x_0) + B_{l_1}(x_2^-) - B_{l_1}(x_1^+) + ... + B_{l_N}(x(T)) - B_{l_N}(x_N^+) \\
&= \int_0^{t_1^-} \frac{\partial B_{l_0}}{\partial x}(x(\tau)) f_{l_0}(x(\tau), d(\tau)) d\tau + ... + \int_{t_N^+}^{T} \frac{\partial B_{l_N}}{\partial x}(x(\tau)) f_{l_N}(x(\tau), d(\tau)) d\tau \\
&\leq -\lambda^T \int_0^T \sigma(x(\tau), d(\tau)) d\tau \leq 0.
\end{aligned}$$

Now, (2.29) guarantees that $B_{l_i}(x_i^+) - B_{l_{i-1}}(x_i^-) \leq 0$ for $i = 1, ..., N$, and hence, it follows from the above inequality that $B_{l_N}(x(T)) \leq B_{l_0}(x_0)$. Using (2.26)–(2.27), we obtain a contradiction, thus proving the theorem. ■

Remark 2.12 *The set of $\{B_l(x) : l \in L\}$ and λ satisfying the conditions in Theorem 2.11 is convex.*

2.3 Computational Method

Computation of barrier certificates is in general not easy, as is the case with computation of Lyapunov functions for nonlinear or hybrid systems. In fact, even verifying that a given barrier certificate satisfies the required conditions is hard. However, for systems whose vector fields are polynomial and whose set descriptions are *semialgebraic* (i.e., described by polynomial equalities and inequalities), a tractable computational method for verifying or constructing a barrier certificate exists, if we also

postulate the barrier certificate to be polynomial. The method uses sum of squares optimization [61,62,69,72] — a convex relaxation framework based on sum of squares decompositions of multivariate polynomials [80] and semidefinite programming [93].

2.3.1 Sum of Squares Optimization

In this subsection, we give a brief review on sum of squares optimization. Some parts of the subsection are based on [72]. See also [61, 62] for more detailed expositions.

Let the indeterminate x take its value in $\mathbb{R}^n$. From this point onward, we will consider polynomials in x with real coefficients. We say that a polynomial $p(x)$ is a sum of squares (SOS), if there exist polynomials $f_1(x), \ldots, f_m(x)$ such that

$$p(x) = \sum_{i=1}^{m} f_i^2(x). \tag{2.31}$$

It follows from the definition that the set of sums of squares polynomials in n variables is a convex cone. The existence of an SOS decomposition (2.31) can be shown equivalent to the existence of a real positive semidefinite matrix Q such that

$$p(x) = Z^T(x) Q Z(x), \tag{2.32}$$

where $Z(x)$ is the vector of monomials of degree less than or equal to degree$(p(x))/2$. By *monomial*, we mean a polynomial of the form $x_1^{\alpha_1} \ldots x_n^{\alpha_n}$, where the α_i's are nonnegative integers, and in this case, the degree of the monomial is $\alpha_1 + \ldots + \alpha_n$.

Expressing an SOS polynomial as a quadratic form in (2.32) has also been referred to as the Gram matrix method [21]. The decomposition (2.31) can be easily converted into (2.32) and vice versa. This equivalence makes an SOS decomposition computable using semidefinite programming, since finding a symmetric positive semidefinite matrix Q subject to the affine constraint (2.32) is nothing but a semidefinite programming problem [93]. Computation of SOS decompositions using semidefinite programming was first suggested in [61].

It is clear that a SOS polynomial is *globally nonnegative*. This is a property of

SOS polynomials that is crucial in many control applications, where we can obtain a tractable computational relaxation by replacing various polynomial inequalities with SOS conditions. However, it should be noted that not all nonnegative polynomials are sums of squares. The equivalence between nonnegativity and sum of squares is only guaranteed in three cases: univariate polynomials of any even degree, quadratic polynomials in any number of indeterminates, and quartic polynomials in three variables [80]. Indeed, nonnegativity is NP-hard to test [54], whereas the SOS condition is polynomial time verifiable through solving appropriate semidefinite programs. Despite this, in many cases we are able to obtain solutions to computational problems that are otherwise at the moment unsolvable, simply by replacing the nonnegativity conditions with SOS conditions.

A sum of squares program is a convex optimization problem of the following form:

$$\text{Minimize } \sum_{j=1}^{m} w_j c_j$$

$$\text{subject to}$$

$$a_{i,0}(x) + \sum_{j=1}^{m} a_{i,j}(x) c_j \text{ is SOS, for } i = 1, ..., p,$$

where the c_j's are scalar real decision variables, the w_j's are given real numbers, and the $a_{i,j}(x)$'s are given polynomials (with fixed coefficients). Note that equality constraint $a_{i,0}(x) + \sum_{j=1}^{m} a_{i,j}(x)c_j = 0$ can be included by asking both $(a_{i,0}(x) + \sum_{j=1}^{m} a_{i,j}(x)c_j)$ and $-(a_{i,0}(x) + \sum_{j=1}^{m} a_{i,j}(x)c_j)$ to be SOS. See also another equivalent canonical form of SOS programs in [69]. Sum of squares programs can still be solved via semidefinite programming using the Gram matrix method explained above. As a matter of fact, SOS programs and semidefinite programs are equivalent, since semidefinite programs can also be viewed as SOS programs with the polynomials $a_{i,j}(x)$ being quadratic. The software SOSTOOLS [69–72], in conjunction with a semidefinite programming solver such as SeDuMi [85], can be used to efficiently solve SOS programs.

It is notable that SOS programs can be used to prove emptiness of (basic) semi-

algebraic sets, i.e., sets of the form

$$\{x \in \mathbb{R}^n : f_j(x) \geq 0,\ g_k(x) \neq 0,\ h_\ell(x) = 0 \quad \forall j, k, \ell\},$$

where $f_j(x)$'s, $g_k(x)$'s, and $h_\ell(x)$'s are polynomials. The main tool used for this purpose is *Positivstellensatz* [84] (see also [15]), a theorem in real algebraic geometry characterizing *certificates for infeasibility* of the above system of polynomial equalities and inequalities. Computation of such infeasibility certificates using hierarchies of SOS programs has been proposed in [62]. The idea is to choose a degree bound for the certificates, then affinely parameterize a set of candidate certificates and find the proper ones in this set by solving a SOS program. If the semialgebraic set is empty and the degree bound is chosen to be large enough, then the SOS program will be feasible.

2.3.2 Direct Computation

The setting of Section 2.2.2 is used in this and the next subsections; other settings can be treated analogously. Consider a hybrid system $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$ whose vector fields $f_l(x, d)$ are polynomial for all $l \in L$. Furthermore, assume that for all $l \in L$, the invariant region $I(l)$ is given by

$$I(l) = \{x \in \mathbb{R}^n : g_{I(l)}(x) \geq 0\}.$$

In these set descriptions, the $g_{I(l)}$'s are vectors of polynomials, and the inequalities are satisfied entry-wise. For example, when $I(l)$ is the n-dimensional hypercube $[\underline{x_1}, \overline{x_1}] \times ... \times [\underline{x_n}, \overline{x_n}]$, we may define

$$g_{I(l)}(x) = \begin{bmatrix} (x_1 - \underline{x_1})(\overline{x_1} - x_1) \\ \vdots \\ (x_n - \underline{x_n})(\overline{x_n} - x_n) \end{bmatrix}.$$

Similarly, define the sets $\mathcal{D}(l)$, Init(l), Unsafe(l), and Guard(l, l') by the inequalities $g_{\mathcal{D}(l)}(d) \geq 0$, $g_{\text{Init}(l)}(x) \geq 0$, $g_{\text{Unsafe}(l)}(x) \geq 0$, and $g_{\text{Guard}(l,l')}(x) \geq 0$. Finally, assume

$$\text{Reset}(l, l')(x) = \{x' \in \mathbb{R}^n : g_{\text{Reset}(l,l')}(x, x') \geq 0\}$$

to be the value of the reset map Reset(l, l') evaluated at $x \in$ Guard(l, l').

When the $B_l(x)$'s are polynomials, verifying that a *given* barrier certificate $\{B_l(x) : l \in L\}$ satisfies the conditions in Theorems 2.7 or 2.8 is equivalent to proving that some basic semialgebraic sets are empty. Consider for example the condition (2.17) for a particular $l \in L$. The condition is satisfied if and only if the basic semialgebraic set

$$\{x \in \mathbb{R}^n : B_l(x) \geq 0,\ B_l(x) \neq 0,\ g_{\text{Init}(l)}(x) \geq 0\}$$

is empty. Proving that the above set is empty can be done using Positivstellensatz, with the help of SOS optimization as mentioned in the previous subsection.

What is more important, however, is the computation of barrier certificates. Sum of squares optimization has been exploited for algorithmically constructing Lyapunov functions for nonlinear systems [58, 61]. A similar approach can be used in the computation of barrier certificates. In this case, real coefficients $c_{1,l}, ..., c_{m,l}$ are used to parameterize sets of candidates for the functions $B_l(x)$, $\forall l \in L$, in the following way:

$$B_l(x) = \sum_j c_{j,l} b_{j,l}(x), \tag{2.33}$$

where the $b_{j,l}(x)$'s are elements of some finite polynomial basis; for example, they could be monomials of degree less than or equal to some pre-chosen bound. Then the search for a barrier certificate $\{B_l(x) : l \in L\}$ — or equivalently, the values of $c_{j,l}$'s, such that the convex conditions in Theorems 2.8 are satisfied — can be directly performed by solving a SOS program, as stated in the following algorithm.

Algorithm 2.13 (Direct Method) *Let the hybrid system H and the descriptions*

of $I(l)$, $\mathcal{D}(l)$, $\mathrm{Init}(l)$, $\mathrm{Unsafe}(l)$, $\mathrm{Guard}(l,l')$, *and* $\mathrm{Reset}(l,l')(x)$ *be given, along with some nonnegative constants* $\lambda_{l,l'}$, *for each* $l \in L$ *and* $(l,l') \in L^2$, $l \neq l'$.

1. ***Parameterize*** $B_l(x)$***'s:*** *Fix a degree bound for the barrier certificate, and parameterize* $B_l(x)$ $\forall l \in L$ *in terms of some unknown coefficients* $c_{j,l}$*'s as in (2.33), by having all monomials whose degrees are less than the degree bound as the* $b_{j,l}(x)$*'s.*

2. ***Parameterize the multipliers:*** *In a similar way, fix some degree bounds and use some other unknown coefficients to parameterize polynomial vectors* $\lambda_{\mathrm{Init}(l)}(x)$, $\lambda_{\mathrm{Unsafe}(l)}(x)$, $\lambda_{I(l)}(x,d)$, $\lambda_{\mathcal{D}(l)}(x,d)$, $\lambda_{\mathrm{Guard}(l,l')}(x,x')$, $\lambda_{\mathrm{Reset}(l,l')}(x,x')$ *of the same dimensions as the corresponding* $g_*(\cdot)$*'s.*

3. ***Compute the coefficients:*** *Choose a small positive number* ϵ. *Use SOS optimization to find values of the coefficients which make the expressions*

$$- B_l(x) - \lambda^T_{\mathrm{Init}(l)}(x) g_{\mathrm{Init}(l)}(x), \tag{2.34}$$

$$+ B_l(x) - \epsilon - \lambda^T_{\mathrm{Unsafe}(l)}(x) g_{\mathrm{Unsafe}(l)}(x), \tag{2.35}$$

$$- \frac{\partial B_l}{\partial x}(x) f_l(x,d) - \lambda^T_{I(l)}(x,d) g_{I(l)}(x) - \lambda^T_{D(l)}(x,d) g_{D(l)}(d), \tag{2.36}$$

$$- B_{l'}(x') + \lambda_{l,l'} B_l(x) - \lambda^T_{\mathrm{Guard}(l,l')}(x,x') g_{\mathrm{Guard}(l,l')}(x) - \lambda^T_{\mathrm{Reset}(l,l')}(x,x') g_{\mathrm{Reset}(l,l')}(x,x') \tag{2.37}$$

and the entries of $\lambda_{\mathrm{Init}(l)}(x)$, $\lambda_{\mathrm{Unsafe}(l)}(x)$, $\lambda_{I(l)}(x,d)$, $\lambda_{\mathcal{D}(l)}(x,d)$, $\lambda_{\mathrm{Guard}(l,l')}(x,x')$, $\lambda_{\mathrm{Reset}(l,l')}(x,x')$ *sums of squares, for each* $l \in L$ *and* $(l,l') \in L^2$, $l \neq l'$.

Proposition 2.14 *If the sum of squares optimization problem given in Algorithm 2.13 is feasible, then the polynomials* $\{B_l(x) : l \in L\}$ *obtained by substituting the corresponding values of* $c_{j,l}$*'s to their polynomial parameterization satisfy the conditions of Theorem 2.8, and therefore* $\{B_l(x) : l \in L\}$ *is a barrier certificate.*

Proof. We show that the entries of $\lambda_{\mathrm{Init}(l)}(x)$ and (2.34) being SOS implies (2.17) as follows. Notice that $-B_l(x) - \lambda^T_{\mathrm{Init}(l)}(x) g_{\mathrm{Init}(l)}(x)$ is globally nonnegative since it

is a SOS, and also that for any $x \in \text{Init}(l)$, the second term is nonnegative. Thus, $-B_l(x) \geq \lambda^T_{\text{Init}(l)}(x) g_{\text{Init}(l)}(x) \geq 0 \quad \forall x \in \text{Init}(l)$, i.e., condition (2.17) holds. Similar arguments can be used for the other conditions. ■

Remark 2.15 *If the reset map* $\text{Reset}(l, l')$ *actually maps* $x \in \text{Guard}(l, l')$ *to a singleton, e.g., if* $\text{Reset}(l, l') : x \mapsto g_{\text{Reset}(l,l')}(x)$ *for some polynomial vector* $g_{\text{Reset}(l,l')}(x)$*, then expression (2.37) can be simplified to*

$$-B_{l'}(g_{\text{Reset}(l,l')}(x)) + \lambda_{l,l'} B_l(x) - \lambda^T_{\text{Guard}(l,l')}(x) g_{\text{Guard}(l,l')}(x).$$

The computational cost of Algorithm 2.13 depends on three factors: the degrees of (2.34)–(2.37), the cardinality of L, and the dimension of (x, d). For fixed degrees, however, the required computations grow polynomially with respect to the cardinality of L and/or the dimension of (x, d). A hierarchy of computations can then be proposed, where we start with a low degree for the barrier certificate and increase it as needed. In many cases, a low degree barrier certificate can be used to verify safety if the system is "sufficiently" safe (in the sense that a small perturbation will not make the system unsafe).

We would also like to remark that although the computational approach discussed in this section assumes that the descriptions of the system and sets are polynomial, non-polynomial descriptions can be handled (although possibly with some conservatism) in at least two different ways:

- First, a non-polynomial vector field can be approximated by a polynomial vector field and the approximation error can be "covered" by including some uncertainty description, which has been treated in Section 2.1.3. In a similar way, we can cover sets with non-polynomial descriptions with those that are described using polynomials.

- Second, for some non-polynomial systems, algebraic recasting of variables can be used to transform the system to a polynomial system, possibly plus some algebraic constraints. Consider for example the system $\dot{x} = e^x$. By introducing

a new variable $\tilde{x} = e^x$, we can obtain an equivalent polynomial description in the new state variable $\dot{\tilde{x}} = e^x \dot{x} = \tilde{x}^2$, with inequality constraint $\tilde{x} \geq 0$. Then a polynomial barrier certificate can be constructed for the new system, which will correspond to a non-polynomial barrier certificate in the original system. The details of the recasting algorithm are outside the scope of this section, but we refer interested readers to [59].

2.3.3 Iterative Computation

The SOS optimization approach described in the previous subsection can be used to find a barrier certificate that lies in the *convex* set defined by the conditions in Theorem 2.8. The conditions in Theorem 2.7, however, define a *non-convex* set of barrier certificates. As a consequence, the search for a barrier certificate in this set cannot be performed through direct SOS optimization, although conditions for the barrier certificate can still be formulated as sum of squares conditions as follows.

Proposition 2.16 *Let the hybrid system H and the descriptions of $I(l)$, $\mathcal{D}(l)$, $\text{Init}(l)$, $\text{Unsafe}(l)$, $\text{Guard}(l,l')$, and $\text{Reset}(l,l')(x)$ be given. Suppose there exist polynomials $B_l(x)$ and $\lambda_{B_l}(x,d)$; positive numbers ϵ_1 and ϵ_2; and vectors of sums of squares $\lambda_{\text{Unsafe}(l)}(x)$, $\lambda_{\text{Init}(l)}(x)$, $\lambda_{I(l)}(x,d)$, $\lambda_{\mathcal{D}(l)}(x,d)$, $\lambda_{\text{Guard}(l,l')}(x,x')$, $\lambda_{\text{Reset}(l,l')}(x,x')$, and $\lambda_{l,l'}(x,x')$; such that the following expressions:*

$$-B_l(x) - \lambda^T_{\text{Init}(l)}(x) g_{\text{Init}(l)}(x), \tag{2.38}$$

$$+B_l(x) - \epsilon_1 - \lambda^T_{\text{Unsafe}(l)}(x) g_{\text{Unsafe}(l)}(x), \tag{2.39}$$

$$-\frac{\partial B_l}{\partial x}(x) f_l(x,d) - \epsilon_2 - \lambda^T_{D(l)}(x,d) g_{D(l)}(d) - \lambda^T_{I(l)}(x,d) g_{I(l)}(x) - \lambda_{B_l}(x,d) B_l(x), \tag{2.40}$$

$$-B_{l'}(x') + \lambda_{l,l'}(x,x') B_l(x) - \lambda^T_{\text{Guard}(l,l')}(x,x') g_{\text{Guard}(l,l')}(x) - \lambda^T_{\text{Reset}(l,l')}(x,x') g_{\text{Reset}(l,l')}(x,x') \tag{2.41}$$

are sums of squares for all $l \in L$ and $(l,l') \in L^2$, $l \neq l'$. Then the collection $\{B_l(x) : l \in L\}$ satisfies the conditions in Theorem 2.7, and therefore the safety property holds.

Proof. Analogous to the proof of Proposition 2.14. ■

In this case, direct computation of $\{B_l(x) : l \in L\}$ via SOS optimization is not possible due to the multiplication of the unknown coefficients of $B_l(x)$'s with those of $\lambda_{B_l}(x,d)$'s and $\lambda_{l,l'}(x,x')$'s in (2.40)–(2.41). By fixing either of them, all the unknown coefficients will be constrained in an affine manner, which reduces the problem[4] to a SOS program. For example, fixing the multipliers will convexify the set of $\{B_l(x) : l \in L\}$'s satisfying the conditions (2.38)–(2.41), resulting in a smaller convex set contained in the original non-convex set.

The motivation to search for barrier certificates in the non-convex set is the fact that when we put a bound on their complexity (e.g., by bounding the polynomial degrees), such barrier certificates are generally less conservative than barrier certificates in the convex set (cf. the comment at the beginning of Section 2.1.2). For instance, the former may prove safety for larger disturbance sets, guard sets, unsafe sets, etc. We will now present a simple iterative method to search for a barrier certificate in the non-convex set. In the iteration, we start with some sufficiently small sets, and increase their sizes as the iteration progresses.

Algorithm 2.17 (Iterative Method)

1. ***Initialization:*** *Start with sufficiently small $\mathcal{D}(l)$, Guard(l,l'), etc. Specify $\lambda_{B_l}(x,d)$ and $\sigma_{l,l'}(x,x')$ in advance, e.g., by choosing $\lambda_{B_l}(x) = 0$ and $\sigma_{l,l'}(x,x') = 0$ or 1. Search for $B_l(x)$'s and the remaining multipliers using SOS optimization as described in Algorithm 2.13.*

2. ***Fix the barrier certificate:*** *Fix the $B_l(x)$'s obtained from the previous step. Enlarge $\mathcal{D}(l)$, Guard(l,l'), etc. Search for $\lambda_{B_l}(x,d)$'s, $\sigma_{l,l'}(x,x')$'s, and the remaining multipliers.*

3. ***Fix the multipliers:*** *Fix the $\lambda_{B_l}(x,d)$'s and $\sigma_{l,l'}(x,x')$'s obtained from the previous step. Enlarge $\mathcal{D}(l)$, Guard(l,l'), etc. Search for $B_l(x)$'s and the remaining multipliers. Repeat to Step 2.*

[4]Note that the original problem is actually equivalent to a bilinear matrix inequality (BMI) problem [50].

For an example illustrating the benefit of using this method, we refer the reader to Section 2.4.2. It should be noted, however, that solving a non-convex optimization problem by an iteration like the above is not guaranteed to yield a globally optimal solution, as the iteration may actually converge to a local optimum. In our case, the barrier certificate we obtain at the end of our iteration may not be a barrier certificate that is able to prove safety for the maximum possible disturbance sets, etc.

2.4 Examples

2.4.1 Continuous System

Consider the two-dimensional system (taken from [40, page 180])

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \frac{1}{3}x_1^3 - x_2 \end{bmatrix},$$

with $\mathcal{X} = \mathbb{R}^2$. We want to verify that all trajectories of the system starting from the initial set $\mathcal{X}_0 = \{x \in \mathbb{R}^2 : (x_1 - 1.5)^2 + x_2^2 \leq 0.25\}$ will never reach the unsafe set $\mathcal{X}_u = \{x \in \mathbb{R}^2 : (x_1 + 1)^2 + (x_2 + 1)^2 \leq 0.16\}$. Note that the system has a stable focus at the origin and two saddle points at $(\pm\sqrt{3}, 0)$. Since $\mathcal{X}_0$ contains a part of the unstable manifold corresponding to the equilibrium $(\sqrt{3}, 0)$, the safety of this system cannot be verified exactly by computation of forward reachable sets in a finite time horizon.

For example, a polynomial barrier certificate $B(x)$ that satisfies (2.2)–(2.4) is given by

$$B(x) = -13 + 7x_1^2 + 16x_2^2 - 6x_1^2x_2^2 - \frac{7}{6}x_1^4 - 3x_1x_2^3 + 12x_1x_2 - \frac{12}{3}x_1^3x_2.$$

That the Lie derivative $\frac{\partial B}{\partial x}(x)f(x)$ is less than or equal to zero can be shown by

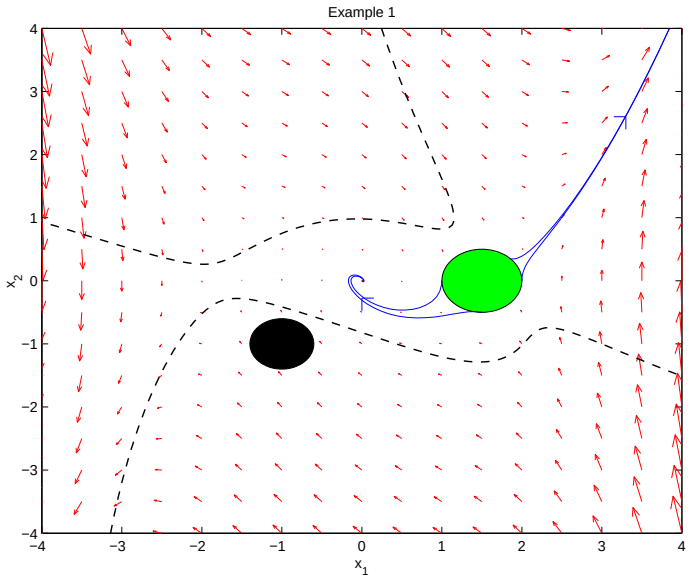


Figure 2.1: Phase portrait of the system in Section 2.4.1. Solid patches are (from left to right) $\mathcal{X}_u$ and $\mathcal{X}_0$, respectively. Dashed curves are the zero level set of $B(x)$, whereas solid curves are some trajectories of the system. The function $B(x)$ is strictly greater than zero for all $x \in \mathcal{X}_u$ and strictly less than zero for all $x \in \mathcal{X}_0$.

exhibiting the quadratic form $-\frac{\partial B}{\partial x}(x)f(x) = Z(x)^T Q Z(x)$, with

$$Q = \begin{bmatrix} 20 & 0 & 15 & 0 & -15/2 & -5 \\ 0 & 3 & 0 & 3/2 & 0 & 0 \\ 15 & 0 & 12 & 0 & -6 & -4 \\ 0 & 3/2 & 0 & 6 & 0 & 0 \\ -15/2 & 0 & -6 & 0 & 3 & 2 \\ -5 & 0 & -4 & 0 & 2 & 4/3 \end{bmatrix}, \quad Z(x) = \begin{bmatrix} x_2 \\ x_2^2 \\ x_1 \\ x_1 x_2 \\ x_1^2 x_2 \\ x_1^3 \end{bmatrix}.$$

In this case, the matrix Q is positive semidefinite, which implies the existence of a sum of squares decomposition for $-\frac{\partial B}{\partial x}(x)f(x)$ (and hence its nonnegativity). That (2.2)–(2.3) are satisfied can be shown by sum of squares arguments as well, and is also depicted pictorially in Figure 2.1. The zero level set of the barrier certificate separates $\mathcal{X}_u$ from all trajectories starting from $\mathcal{X}_0$. Hence, the safety of the system is verified.

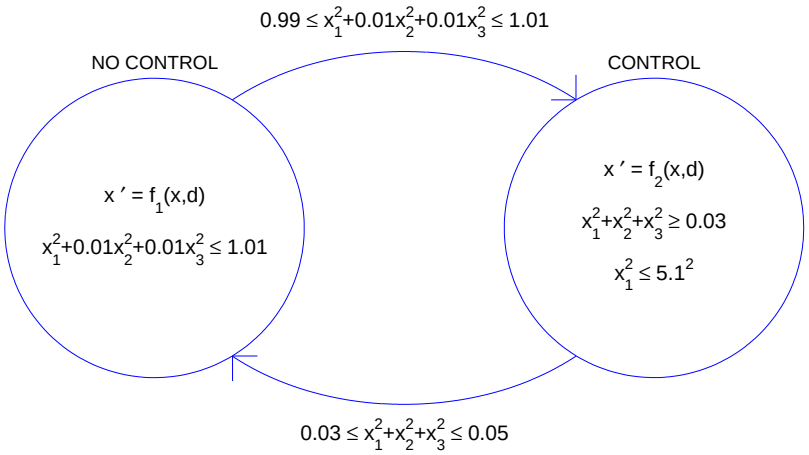


Figure 2.2: Discrete transition diagram of the system in Section 2.4.2. This system has two discrete locations: NO CONTROL and CONTROL, with the vector field and the invariant of each location depicted inside the corresponding circle. The texts labelling the transitions between locations describe the guard sets.

2.4.2 Hybrid System

Consider a hybrid system whose discrete transition diagram is depicted in Figure 2.2. The system starts in location 1 (NO CONTROL mode), with its continuous state initialized at $\text{Init}(1) = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 \leq 0.01\}$. In this location, the continuous state evolves according to

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + x_3 \\ x_1 + (2x_2 + 3x_3)(1 + x_3^2) + d \end{bmatrix} \triangleq f_1(x, d),$$

until it reaches some point in the guard set $\text{Guard}(1, 2) = \{x \in \mathbb{R}^3 : 0.99 \leq x_1^2 + 0.01x_2^2 + 0.01x_3^2 \leq 1.01\}$, at which instance a controller whose objective is to prevent $|x_1|$ from getting too big will be turned on, and the system jumps to location 2 (CONTROL mode). In location 2, the continuous dynamics is described by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + x_3 \\ -x_1 - 2x_2 - 3x_3 + d \end{bmatrix} \triangleq f_2(x, d).$$

Iteration	Description	Verified
1	Set $\lambda_{B_l}(x,d)=0$, find $B_l(x)$.	$-0.005 \le d \le 0.005$
2	Fix $B_l(x)$, find $\lambda_{B_l}(x,d)$.	$-0.625 \le d \le 0.625$
3	Fix $\lambda_{B_l}(x,d)$, find $B_l(x)$.	$-1 \le d \le 1$

Table 2.1: Description and results of the iterative method in Section 2.4.2. The third column indicates the disturbance range for which safety is verified.

The system will remain in this location until the continuous state enters the second guard set $\text{Guard}(2,1) = \{x \in \mathbb{R}^3 : 0.03 \le x_1^2 + x_2^2 + x_3^2 \le 0.05\}$, where the controller will be turned off and the system jumps to location 1. We assume nondeterminism in the jump from location 1 to location 2 and vice versa. For this system, the invariant of the discrete locations are given by $I(1) = \{x \in \mathbb{R}^3 : x_1^2 + 0.01x_2^2 + 0.01x_3^2 \le 1.01\}$ and $I(2) = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 \ge 0.03, x_1^2 \le 5.1^2\}$.

Our task in this example is to verify that $|x_1|$ never gets bigger than 5, if the instantaneous magnitude of the disturbance d is bounded by 1. We define our unsafe sets as $\text{Unsafe}(1) = \emptyset$ and $\text{Unsafe}(2) = \{x \in \mathbb{R}^3 : 5 \le x_1 \le 5.1\} \cup \{x \in \mathbb{R}^3 : -5.1 \le x_1 \le -5\}$, and compute a quartic barrier certificate satisfying the conditions in Theorem 2.7. Using the iterative method described in Section 2.3.3 to enlarge the verifiable disturbance set, we obtain the results shown in Table 2.1. At the third iteration, we are able to prove the safety of the system.

2.4.3 Limit of Design

.

In this example, we analyze the reachability of a linear system in feedback interconnection with a relay. The block diagram of the system is shown in Figure 2.3, with the matrices A, B, C, and D given by

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.2 & -0.3 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0.1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 \end{bmatrix},$$

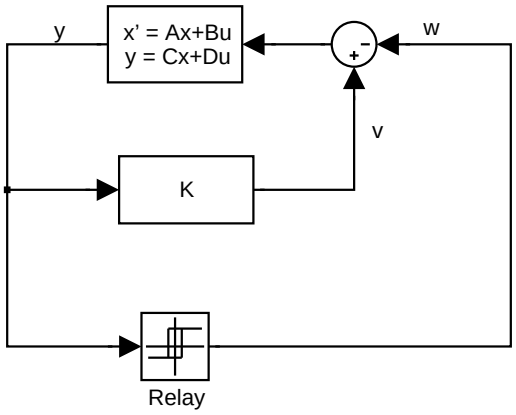


Figure 2.3: Block diagram of the system in Section 2.4.3. We ask if it is possible to design a controller K that steers the system from an initial set $\mathcal{X}_0$ to a destination set $\mathcal{X}_u$, subject to some other specifications.

and the relay element having the following characteristic:

$$w = \begin{cases} 10, & \text{if } y \geq 0, \\ -10, & \text{if } y < 0. \end{cases}$$

For the sets $\mathcal{X} = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 \leq 4^2\}$, $\mathcal{X}_0 = \{x \in \mathbb{R}^3 : (x_1+2)^2 + x_2^2 + x_3^2 \leq 0.1^2\}$, and $\mathcal{X}_u = \{x \in \mathbb{R}^3 : (x_1-2)^2 + x_2^2 + x_3^2 \leq 0.1^2\}$, we pose the following question: Is it possible to design a controller K (possibly nonlinear and time-varying) with the L_2-gain not greater than one, which is connected to the system in the way shown in Figure 2.3, such that the system can be steered from $\mathcal{X}_0$ to $\mathcal{X}_u$ while maintaining the state in $\mathcal{X}$?

The requirement that the L_2-gain of the controller is not greater than one can be equivalently formulated as an integral quadratic constraint (IQC) [49]

$$\int_0^T [y^2(t) - v^2(t)]dt \geq 0 \qquad \forall T \geq 0.$$

This specification introduces dynamic uncertainty to the problem. Nevertheless, we can perform reachability analysis by adjoining the above IQC using a nonnegative constant multiplier to the conditions on the time derivative of barrier certificates (cf. Theorem 2.11). For this example, a quartic barrier certificate that satisfies the

required conditions can be found. Hence we conclude that the given specification is impossible to meet.

2.5 Appendix: Non-Convex Conditions

In Section 2.1.2, it is mentioned that other non-convex conditions guaranteeing safety can be derived using viability theory. For example, using a viability theorem by Nagumo [55] and a characterization of contingent cone for a set described by inequalities [12], the following proposition can be obtained.

Proposition 2.18 *Let the system $\dot{x} = f(x)$ and the sets $\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$ be given, with the vector field $f : \mathbb{R}^n \to \mathbb{R}^n$ being locally Lipschitz continuous and $\mathcal{X}$ being open. Suppose there exists a function $B \in C^1(\mathbb{R}^n)$ that satisfies the following conditions:*

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{2.42}$$

$$B(x) > 0 \quad \forall x \in \mathcal{X}_u, \tag{2.43}$$

$$\frac{\partial B}{\partial x}(x) \neq 0 \quad \forall x \in \mathcal{X} \text{ such that } B(x) = 0, \tag{2.44}$$

$$\frac{\partial B}{\partial x}(x) f(x) \leq 0 \quad \forall x \in \mathcal{X} \text{ such that } B(x) = 0, \tag{2.45}$$

then the safety of the system in the sense of Definition 2.1 is guaranteed.

Notice in particular that if there is no disturbance input, the vector field $f(x)$ is locally Lipschitz continuous, and $\mathcal{X}$ is open, then the statement in Proposition 2.3 follows as a corollary of Proposition 2.18. We will now state some definitions and results needed to prove Proposition 2.18.

Definition 2.19 (Contingent Cone) *Let X be a normed space, K be a non-empty subset of X, and x belong to K. The (Bouligand) contingent cone to K at x is*

$$T_K(x) = \left\{ v \in X : \liminf_{h \to 0^+} \frac{d_K(x + hv)}{h} = 0 \right\},$$

where $d_K(y)$ is the distance of y to K, i.e., $d_K(y) = \inf_{z \in K} \|y - z\|$.

In proving Proposition 2.18, we will use $X = \mathbb{R}^n$ and $K = \{x \in X : B(x) \leq 0\}$. The contingent cone to K is characterized in the following lemma.

Lemma 2.20 (See, e.g., [12]) *Let $X = \mathbb{R}^n$ and $K = \{x \in X : B(x) \leq 0\}$ for a continuously differentiable $B(x)$. Then $T_K(x) = X$ if x is in the interior of K, and*

$$T_K(x) = \left\{ v \in X : \frac{\partial B}{\partial x}(x) v \leq 0 \right\}$$

for any x such that $B(x) = 0$, under the condition that $\frac{\partial B}{\partial x}(x) \neq 0$.

Theorem 2.21 (Nagumo[5]) *Let X be a finite dimensional vector space, $K \subseteq X$ be locally compact and $f(x)$ be continuous from K to X. Then K is* locally viable *under $f(x)$, i.e., for any initial state $x_0 \in K$ there exist $\tau > 0$ such that at least one solution $x(t)$ of the differential equations $\dot{x} = f(x)$ starting at x_0 stays in K on $[0, \tau]$, if and only if*

$$f(x) \in T_K(x) \quad \forall x \in K.$$

Proof of Proposition 2.18. Let K be as defined in Lemma 2.20, and consider any initial condition $x_0 \in \partial K \cap \mathcal{X}$, where ∂K here denotes the boundary of K. Since $f(x) \in T_K(x)$ for all $x \in K \cap \mathcal{X}$, by Theorem 2.21 there is at least a trajectory of the system starting at x_0 that on a small enough time interval is contained in $K \cap \mathcal{X}$. But in fact there is only one such trajectory, since in the proposition we assert that $f(x)$ is locally Lipschitz continuous, which guarantees uniqueness of solutions to the differential equations. It follows that there is no trajectory $x : [0, T] \to \mathcal{X}$ starting from $\mathcal{X}_0$ that can intersect $\partial K \cap \mathcal{X}$ to reach $\mathcal{X}_u$, thus proving the proposition. ■

To see what can go wrong if condition (2.44) or the local Lipschitz continuity of the vector field is not fulfilled, consider the following examples.

[5] We use the version in [11].

Example 2.22 *Consider the system* $\dot{x} = 1$, *with* $\mathcal{X} = \mathbb{R}$, $\mathcal{X}_0 = (-\infty, -1]$, *and* $\mathcal{X}_u = [1, \infty)$. *Let* $B(x) = x^3$. *Then all the conditions in Proposition 2.18 are satisfied except (2.44). In fact, the system is not safe.*

Example 2.23 *Consider the system* $\dot{x} = x^{1/3}$, *with* $\mathcal{X} = \mathbb{R}$, $\mathcal{X}_0 = \{0\}$, *and* $\mathcal{X}_u = [1, \infty)$. *Let* $B(x) = x$. *Then all the conditions in Proposition 2.18 are satisfied except the local Lipschitz continuity of* $f(x)$. *The system is not safe, as there is a trajectory* $x(t) = (2t/3)^{3/2}$ *that connects* $\mathcal{X}_0$ *to* $\mathcal{X}_u$. *However, as guaranteed by Theorem 2.21, there is at least one trajectory, in this case* $x(t) = 0$, *that starts from* $\mathcal{X}_0$ *and stays in* $\{x \in \mathbb{R} : B(x) \leq 0\}$

Chapter 3

Stochastic Safety Verification

In this chapter, we consider safety verification of stochastic continuous and hybrid systems. The stochasticity of a continuous system may originate from random inputs to the dynamics, which can be taken into account by considering stochastic differential equations. In the case of stochastic hybrid systems, stochasticity may also be induced by randomness in the discrete transitions. Study of systems modelled by stochastic differential equations has a long history and readers can find relevant references, e.g., in [57]. On the other hand, only quite recently have people started to consider stochastic hybrid systems. See for instance [24, 27, 31, 34], and also [64] for an overview. Stochastic hybrid systems have been used as a modelling framework in various applications, such as air traffic management [29], manufacturing systems [27], communication networks [31], and stochastic modelling of chemical reactions [32].

When the system is stochastic, answering the safety verification question in a worst-case non-stochastic manner (i.e., to verify whether or not a trajectory of the system can reach the unsafe set) as presented in the previous chapter will usually lead to a very conservative and restrictive answer, since in most cases there is no hard bound on the value of stochastic input. Indeed, it is more natural to formulate and consider a safety verification problem that has a probabilistic interpretation. For example, it may be of interest to prove that the *probability* that a system trajectory reaches the unsafe region is *lower* than a certain safety margin. For some references on safety verification of stochastic continuous and hybrid systems, readers are referred to [19, 20, 35, 94]. Note also that there have been results on probabilistic model

checking (see [82] and references therein), but they are only applicable to systems with finite state.

The approach that we take to solve the stochastic safety verification problem still relies on barrier certificates. However, instead of using a barrier certificate whose zero level set separates the unsafe region from all possible system trajectories, we will use a barrier certificate that yields a *supermartingale* (loosely speaking, its expected value is non-increasing along time) under the given system dynamics. In addition, we ask that the value of the barrier certificate at the initial state be lower than its value at the unsafe region. The probability of reaching the unsafe region can then be bounded from above using a Chebyshev-like inequality for supermartingales. We derive conditions that must be satisfied by barrier certificates for stochastic continuous systems and various classes of stochastic hybrid systems. Similar to their non-stochastic counterpart, polynomial barrier certificates can be computed using sum of squares optimization when the description of the system is polynomial and the sets are semialgebraic.

For the above classes of systems, our method can be used to efficiently compute an *exactly guaranteed* upper bound on the probability that a system trajectory reaches the unsafe set. The references [19, 20], for example, suggest (theoretical) ways to calculate such a probability, yet they have not provided a computational technique for that. The reference [35] does provide a computational method to approximate the reach probability for stochastic differential equations, but since their method is based on discretizing the state space, there are still some unresolved issues with guaranteeing the accuracy of the computed probability and the scalability of the method. Similarly, the work in [94] approximates the reach probability for stochastic discrete time systems using randomized simulations; hence there is no accuracy guarantee either.

This chapter is organized as follows. In Section 3.1 we will consider safety verification of stochastic continuous systems. Safety verification of stochastic hybrid systems will be addressed in Section 3.2. Finally, the chapter will end with some examples in Section 3.3.

3.1 Continuous Systems

Consider a complete probability space $(\Omega, \mathcal{F}, P)$ and a standard $\mathbb{R}^m$-valued Wiener process $w(t)$ defined on this space. In this section, we will be dealing with stochastic differential equations of the form

$$dx(t) = f(x(t))dt + g(x(t))dw(t), \tag{3.1}$$

where $x(t) \in \mathbb{R}^n$, and $f(x)$, $g(x)$ are of appropriate dimensions. We denote the state space, the initial set, and the unsafe set, respectively by $\mathcal{X}$, $\mathcal{X}_0$, and $\mathcal{X}_u$, all of which are subsets of $\mathbb{R}^n$, with $\mathcal{X}$ assumed to be bounded and $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$. To guarantee the existence and uniqueness of solution, we will also assume that both $f(x)$ and $g(x)$ satisfy the local Lipschitz continuity and the linear growth condition on $\mathcal{X}$. Since $\mathcal{X}$ is bounded, the last condition can be replaced by the boundedness of $f(x)$ and $g(x)$ on $\mathcal{X}$.

It can be shown that the process $x(t)$ described above is right continuous and a strong Markov process [57]. The generator A of the process $x(t)$ is defined as follows.

Definition 3.1 (Generator) *The (infinitesimal) generator A of the process $x(t)$ is defined by*

$$AB(x_0) = \lim_{t \downarrow 0} \frac{E[B(x(t)) \mid x(0) = x_0] - B(x_0)}{t},$$

and the domain of the generator is the set of all functions $B : \mathbb{R}^n \to \mathbb{R}$ such that the above limit exists for all x_0.

The generator can be considered as the stochastic analog of the Lie derivative, and characterizes the evolution of the expectation of $B(x(t))$ via the so-called Dynkin's formula (see, e.g., [81]):

$$E[B(\tilde{x}(t_2))|\tilde{x}(t_1)] = B(\tilde{x}(t_1)) + E[\int_{t_1}^{t_2} AB(\tilde{x}(t))dt|\tilde{x}(t_1)] \tag{3.2}$$

for $t_2 \geq t_1$ and for any function $B(x)$ in the domain of the generator.

Since in general the process $x(t)$ is not guaranteed to always lie inside the set $\mathcal{X}$, we define the stopped process corresponding to $x(t)$ and $\mathcal{X}$ as follows.

Definition 3.2 (Stopped Process) *Suppose that τ is the first time of exit of $x(t)$ from the open set $\mathrm{int}(\mathcal{X})$. The stopped process $\tilde{x}(t)$ is defined by*

$$\tilde{x}(t) = \begin{cases} x(t) & \text{for } t < \tau, \\ x(\tau) & \text{for } t \geq \tau. \end{cases}$$

The stopped process $\tilde{x}(t)$ satisfies various properties. For example, it inherits the right continuity and strong Markovian property of $x(t)$. Furthermore, in most cases the generator corresponding to $\tilde{x}(t)$ is identical to the one corresponding to $x(t)$ on the set $\mathrm{int}(\mathcal{X})$, and is equal to zero outside of the set [42]. This will be implicitly assumed throughout the chapter. Having defined the system and the stopped process $\tilde{x}(t)$, we can now formulate the safety verification problem for stochastic differential equations in the probabilistic setting as follows.

Problem 3.3 *Given the system (3.1) and the bounded sets $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, compute an upper bound for the probability of the process $\tilde{x}(t)$ to reach $\mathcal{X}_u$. In other words, find $\gamma \in [0, 1]$ such that*

$$P\{\tilde{x}(t) \in \mathcal{X}_u \text{ for some } t \geq 0 \mid \tilde{x}(0) = x_0\} \leq \gamma \qquad \forall x_0 \in \mathcal{X}_0, \tag{3.3}$$

or

$$P\{\tilde{x}(t) \in \mathcal{X}_u \text{ for some } t \geq 0\} \leq \gamma, \tag{3.4}$$

if a probability distribution μ_0 whose support is in $\mathcal{X}_0$ is also given for $\tilde{x}(0)$.

Obviously, the ultimate objective of safety verification is to show that the above probability is small enough, for example, less than some safety margin. Hence, it is of interest to obtain an upper bound γ that is as tight as possible.

In this chapter, our approach to solve the above problem is based on finding an

appropriate barrier certificate $B(x)$ from which we can deduce an upper bound γ. As in the non-stochastic case, the approach is again analogous to using Lyapunov functions for proving stability[1]. However, instead of requiring the value of $B(\tilde{x}(t))$ to decrease along the trajectory of the system, we ask that the *expected* value of $B(\tilde{x}(t))$ decreases or stays constant as time increases. A process satisfying such a property is called a *supermartingale* (see, e.g., [81] for a technical definition). In our setting, a process $B(\tilde{x}(t))$ is a supermartingale with respect to the filtration $\{\mathcal{M}_t : t \geq 0\}$ generated by the process $\tilde{x}(t)$, if $B(\tilde{x}(t))$ is $\mathcal{M}_t$-measurable for all $t \geq 0$, $E[|B(\tilde{x}(t))|] < \infty$ for all $t \geq 0$, and

$$E[B(\tilde{x}(t_2))|\tilde{x}(t_1)] \leq B(\tilde{x}(t_1))$$

for all $t_2 \geq t_1$. Since we will use $B(x)$ that is twice continuously differentiable and $\tilde{x}(t)$ takes its value in a bounded set $\mathcal{X}$, the first and second conditions are always fulfilled. For nonnegative supermartingales, there exists the following result, which will be used several times in this chapter.

Lemma 3.4 ([42]; see [26] for the discrete version) *Let $B(\tilde{x}(t))$ be a supermartingale with respect to the process $\tilde{x}(t)$ and $B(x)$ be nonnegative on $\mathcal{X}$. Then for a positive λ and any initial condition $x_0 \in \mathcal{X}$,*

$$P\left\{ \sup_{0 \leq t < \infty} B(\tilde{x}(t)) \geq \lambda \;\middle|\; \tilde{x}(0) = x_0 \right\} \leq \frac{B(x_0)}{\lambda}. \tag{3.5}$$

At this point, we are ready to state and prove our first main result.

Theorem 3.5 *Let the stochastic differential equation (3.1) and the bounded sets $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$ be given, with $f(x)$, $g(x)$ being locally Lipschitz continuous and bounded on $\mathcal{X}$. Consider the stopped process $\tilde{x}(t)$. Suppose there exists a function*

[1] See, e.g., [42] for some notions of stochastic stability and stochastic Lyapunov functions.

$B \in C^2(\mathbb{R}^n)$ such that

$$B(x) \leq \gamma \quad \forall x \in \mathcal{X}_0, \tag{3.6}$$

$$B(x) \geq 1 \quad \forall x \in \mathcal{X}_u, \tag{3.7}$$

$$B(x) \geq 0 \quad \forall x \in \mathcal{X}, \tag{3.8}$$

$$\frac{\partial B}{\partial x}(x)f(x) + \frac{1}{2}\mathrm{Tr}\left(g^T(x)\frac{\partial^2 B}{\partial x^2}(x)g(x)\right) \leq 0 \quad \forall x \in \mathcal{X}, \tag{3.9}$$

then the probability bound (3.3) holds. If an initial probability distribution μ_0 is given, then (3.7)–(3.9) and

$$\int_{\mathcal{X}_0} B(x)d\mu_0(x) \leq \gamma \tag{3.10}$$

imply that the probability bound (3.4) holds.

Proof. For the stochastic differential equation (3.1), the generator of the process is given by (see, e.g., [57])

$$AB(x) = \frac{\partial B}{\partial x}(x)f(x) + \frac{1}{2}\mathrm{Tr}\left(g^T(x)\frac{\partial^2 B}{\partial x^2}(x)g(x)\right),$$

where the domain of the generator is the set of twice continuously differentiable functions with compact support. Since $\mathcal{X}$ is bounded, we can use any $B \in C^2(\mathbb{R}^n)$. Next, using Dynkin's formula, we have for $0 \leq t_1 \leq t_2 < \infty$

$$\begin{aligned} E[B(\tilde{x}(t_2))|\tilde{x}(t_1)] &= B(\tilde{x}(t_1)) + E[\int_{t_1}^{t_2} AB(\tilde{x}(t))dt|\tilde{x}(t_1)] \\ &\leq B(\tilde{x}(t_1)), \end{aligned}$$

and therefore (3.9) will imply that $B(\tilde{x}(t))$ is a supermartingale. By (3.8) and Lemma 3.4 we conclude that (3.5) holds. Now use (3.6) and the fact that $\mathcal{X}_u \subseteq \{x \in \mathcal{X} : B(x) \geq 1\}$, which follows from (3.7), to obtain the following series of

inequalities:

$$P\{\tilde{x}(t) \in \mathcal{X}_u \text{ for some } t \geq 0 \mid \tilde{x}(0) = x_0\}$$
$$\leq P\left\{ \sup_{0 \leq t < \infty} B(\tilde{x}(t)) \geq 1 \;\middle|\; \tilde{x}(0) = x_0 \right\} \leq B(x_0) \leq \gamma \quad \forall x_0 \in \mathcal{X}_0.$$

Thus, the probability bound (3.3) is proven.

Finally, if an initial probability distribution μ_0 is given, then the above derivation can be combined with the law of total probability and (3.10) to obtain

$$P\{\tilde{x}(t) \in \mathcal{X}_u \text{ for some } t \geq 0\} \leq \int_{\mathcal{X}_0} B(x) d\mu_0(x) \leq \gamma,$$

hence finishing the proof. ■

Note that it is possible to choose γ to be at most equal to one, since when $\gamma = 1$ the function $B(x) = 1$ will satisfy (3.6)–(3.9) and (3.10). The intuitive idea behind the theorem is clear. The process $B(\tilde{x}(t))$ is a supermartingale, and therefore its value is likely to stay constant or decrease as time increases. When we start from a lower initial value of $B(x)$ (i.e., as γ gets smaller), it becomes less likely for the trajectory to reach the unsafe set, on which the value of $B(x)$ is greater than or equal to one. This is quantified by Lemma 3.4, which provides a Chebyshev-like inequality for bounding the probability of the distribution tail.

An upper bound γ and a barrier certificate $B(x)$ which certifies the upper bound can be computed by formulating conditions (3.6)–(3.9) or (3.7)–(3.9) and (3.10) as a sum of squares optimization problem, similar to what we describe in Section 2.3. Furthermore, γ can be chosen as the objective function of the SOS program, whose value is to be minimized. The minimum value of γ obtained from the optimization will be the tightest upper bound for a given polynomial and sum of squares parameterization. Obviously, we may get a better bound as we expand the parameterization, for example, when we use higher degree barrier certificates. However, there is a trade-off between using a larger set of candidate barrier certificates and the computational complexity of finding a true certificate within it.

3.2 Hybrid Systems

In this section, we will consider several classes of stochastic hybrid systems, namely:

- Piecewise deterministic Markov processes [24],
- Switching diffusion processes [27],
- Stochastic hybrid systems of Hu et al. [34].

See also [64] for an overview. The method proposed in Section 3.1 can be extended to handle the above classes of systems. The main idea is similar to before, i.e., use the appropriate generator for the process, find a barrier certificate from the domain of the generator that yields a nonnegative supermartingale, and then bound the reach probability using the barrier certificate.

3.2.1 Piecewise Deterministic Markov Processes

In this section, we consider a class of stochastic hybrid systems called the piecewise deterministic Markov processes. Systems in this class have both continuous and discrete states, where the continuous state evolves according to an *ordinary* differential equation that depends on the discrete state. A discrete transition occurs either when the continuous state hits the boundary of the invariant, or in the interior of the invariant according to a generalized Poisson process with a state-dependent rate. In addition, during a transition the hybrid state is reset according to a probability distribution that is determined by the hybrid state before transition.

A piecewise deterministic Markov process is defined as $H = (\mathcal{X}, L, I, \mu_0, f_l, \lambda_l, R_l)$, with the following components[2]:

- $\mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state space.
- L is a finite set of locations. The overall state space of the system is $X = L \times \mathcal{X}$, and the state is denoted by $(l, x) \in L \times \mathcal{X}$.

[2]For notational simplicity and without loss of generality, we assume that the continuous state space has the same dimension for all $l \in L$; this is unlike in [20, 24].

- $I : L \to 2^{\mathcal{X}}$ is the invariant, which assigns to each location l an open set $I(l) \subseteq \mathcal{X}$ that contains all possible continuous states while at location l. For our purpose, it will also be assumed that $I(l)$ is bounded.

- μ_0 is a probability measure for the initial state, with its support contained in $X_0 \subseteq \bigcup_{l \in L}(\{l\} \times I(l))$.

- $f_l : I(l) \to \mathbb{R}^n$, $l \in L$, is a set of vector fields, where the subscript indicates the corresponding discrete location.

- $\lambda_l : I(l) \to [0, \infty)$, $l \in L$, is a set of state-dependent transition rates.

- $R_l(x)$, where $l \in L$, $x \in \text{cl}(I(l))$, is a set of reset probability measures for the hybrid state, with its support contained in $\bigcup_{l \in L}(\{l\} \times I(l))$.

In addition, we denote the unsafe region by $X_u = \cup_{l \in L}\{l\} \times \text{Unsafe}(l)$, where $\text{Unsafe}(l) \subseteq I(l)$ and can be empty for some l's.

A trajectory of the system starts with an initial condition (l_0, x_0) drawn from the initial probability measure μ_0. At this location, the continuous part of the state evolves according to the differential equation $\dot{x}(t) = f_{l_0}(x(t))$. Let $\hat{T}$ be the first time $\phi_t(x_0)$ exits $I(l)$, where $\phi_t(.)$ is the flow corresponding to the vector field $f_{l_0}(x)$. The time to transition T is governed by

$$P\{T > t\} = \begin{cases} \exp(-\int_0^t \lambda_l(\phi_\tau(x_0))d\tau) & \text{if } t < \hat{T}, \\ 0 & \text{if } t \geq \hat{T}. \end{cases}$$

Right when T elapses, the system undergoes a transition, and the hybrid state is reset to a new state (l_1, x_1) that is drawn from the reset probability measure $R_{l_0}(\phi_T(x_0))$. The above process is then repeated. Under some technical assumptions [24], the process will be right continuous and have the strong Markovian property. For our purpose, we will assume that $f_l(x)$ is globally Lipschitz continuous, $\lambda_l(x)$ is continuous, and the expected number of transitions is finite on any finite time interval. Since we reset the state when $x(t)$ goes out of $I(l)$, there is no need to use a stopped process here.

For this class of systems, the barrier certificate $B(l,x)$ will be constructed from several functions $B_l(x)$, where each $B_l(x)$ corresponds to a discrete location and we define $B(l,x) = B_l(x)$. The conditions that are satisfied by the barrier certificate are stated in the following theorem.

Theorem 3.6 *Let the piecewise deterministic Markov process $H = (\mathcal{X}, L, I, \mu_0, f_l, \lambda_l, R_l)$ with bounded $I(l)$'s, globally Lipschitz continuous $f_l(x)$'s, continuous $\lambda_l(x)$'s, and the unsafe set $X_u \subseteq \bigcup_{l\in L}(\{l\} \times I(l))$ be given. Suppose there exists a collection $\{B_l(x) : l \in L\}$ of functions $B_l \in C^1(\mathbb{R}^n)$, which satisfy*

$$B_l(x) \geq 1 \quad \forall x \in \text{Unsafe}(l), \tag{3.11}$$

$$B_l(x) \geq 0 \quad \forall x \in I(l), \tag{3.12}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x) + \lambda_l(x) \sum_{l'\in L} \int_{I(l')} (B_{l'}(x') - B_l(x)) dR_l(x)(x') \leq 0 \quad \forall x \in \text{cl}(I(l)), \tag{3.13}$$

$$B_l(x) - \sum_{l'\in L} \int_{I(l')} B_{l'}(x') dR_l(x)(x') = 0 \quad \forall x \in \partial I(l), \tag{3.14}$$

for all $l \in L$, and

$$\sum_{l\in L} \int_{I(l)} B_l(x) d\mu_0(l,x) \leq \gamma. \tag{3.15}$$

Then $P\{(l(t), x(t)) \in X_u \text{ for some } t \geq 0\} \leq \gamma$.

Proof. Define $B(l(t), x(t)) = B_{l(t)}(x(t))$. In this case,

$$AB(l,x) = \frac{\partial B_l}{\partial x}(x) f_l(x) + \lambda_l(x) \sum_{l'\in L} \int_{I(l')} (B_{l'}(x') - B_l(x)) dR_l(x)(x')$$

is the generator of the process, and $B(l,x)$ is in the domain of the generator if $B_l \in C^1(\mathbb{R}^n)$ and (3.14) holds (see [24]). Condition (3.13) implies that $B(l(t), x(t))$ is a supermartingale, which can be shown using Dynkin's formula. Since $B(l(t), x(t))$ is also nonnegative (as implied by (3.12)), Lemma 3.4 can be applied. The rest of the proof is similar to the proof of Theorem 3.5. ■

3.2.2 Switching Diffusion Processes

The continuous state of a switching diffusion process evolves according to a stochastic differential equation that depends on the discrete state, and the discrete trajectory itself is a Markov chain whose transition matrix depends on the continuous state. As implied by the name, these systems are switching systems, meaning that the value of the continuous state does not change during a discrete transition.

Formally, a switching diffusion process is a tuple $H = (\mathcal{X}, L, \mu_0, f_l, g_l, \lambda_{ll'})$ with the following components:

- $\mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state space, assumed to be bounded.
- L is a finite set of locations. The overall state space of the system is $X = L \times \mathcal{X}$, and the state is denoted by $(l, x) \in L \times \mathcal{X}$.
- μ_0 is an initial probability measure, with its support in $X_0 \subseteq X$.
- $f_l : \mathcal{X} \to \mathbb{R}^n$, $l \in L$, is a set of drift vector fields.
- $g_l : \mathcal{X} \to \mathbb{R}^{n \times m}$, $l \in L$, is a set of diffusion coefficients, where the i-th column of g_l corresponds to the i-th component of the $\mathbb{R}^m$-valued Wiener process $w(t)$.
- $\lambda_{ll'} : \mathcal{X} \to \mathbb{R}$, $(l, l') \in L^2$, is a set of x-dependent transition rates, with $\lambda_{ll'}(x) \geq 0$ for all x if $l \neq l'$, and $\sum_{l' \in L} \lambda_{ll'}(x) = 0$ for all $l \in L$.

Here we denote the unsafe set by $\mathcal{X}_u$, with $\mathcal{X}_u \subseteq \mathcal{X}$.

A trajectory of the system starts with an initial condition drawn from the initial probability measure μ_0. As mentioned above, the continuous part of the state evolves according to a stochastic differential equation, which at location l is given by

$$dx(t) = f_l(x(t))dt + g_l(x(t))dw(t).$$

On the other hand, the dynamics of the discrete state is described by the following

transition probability:

$$P\{l(t+\Delta)=j \mid l(t)=i\} = \begin{cases} \lambda_{ij}(x(t))\Delta + o(\Delta), \text{ if } i \neq j, \\ 1+\lambda_{ii}(x(t))\Delta + o(\Delta), \text{ if } i = j, \end{cases} \tag{3.16}$$

with $\Delta > 0$. See [27] for more details on how the discrete transitions are generated. During a discrete transition, the value of the continuous state is held constant. It is assumed that the discrete transition is independent from the Wiener process $w(t)$. In addition, we assume that $f_l(x)$, $g_l(x)$, and $\lambda_{ll'}(x)$ are bounded and locally Lipschitz continuous. Under these assumptions, the solution to the stochastic differential equation at each location exists and is unique, and also that $(l(t), x(t))$ is a Markov process and almost every sample path of it is a right continuous function [27]. Similar to the continuous case, we stop the process when $x(t)$ goes out from $\text{int}(\mathcal{X})$.

The conditions for a barrier certificate are stated in the following theorem.

Theorem 3.7 *Let the switching diffusion process $H = (\mathcal{X}, L, \mu_0, f_l, g_l, \lambda_{ll'})$ be given, with bounded $\mathcal{X}$ and bounded, locally Lipschitz continuous $f_l(x)$'s, $g_l(x)$'s, and $\lambda_{ll'}(x)$'s. Suppose there exists a collection $\{B_l(x) : l \in L\}$ of functions $B_l \in C^2(\mathbb{R}^n)$, which satisfy*

$$B_l(x) \geq 1 \quad \forall x \in \mathcal{X}_u, \tag{3.17}$$

$$B_l(x) \geq 0 \quad \forall x \in \mathcal{X}, \tag{3.18}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x) + \frac{1}{2}\text{Tr}\left(g_l^T(x)\frac{\partial^2 B_l}{\partial x^2}(x) g_l(x)\right) + \sum_{l' \in L} \lambda_{ll'}(x) B_{l'}(x) \leq 0 \quad \forall x \in \mathcal{X}, \tag{3.19}$$

for all $l \in L$, and

$$\sum_{l \in L} \int_{\mathcal{X}} B_l(x) d\mu_0(l, x) \leq \gamma. \tag{3.20}$$

Then $P\{\tilde{x}(t) \in \mathcal{X}_u \text{ for some } t \geq 0\} \leq \gamma$.

Proof. Define $B(l(t), x(t)) = B_{l(t)}(x(t))$. In this case,

$$AB(l,x) = \frac{\partial B_l}{\partial x}(x) f_l(x) + \frac{1}{2}\mathrm{Tr}\left(g_l^T(x) \frac{\partial^2 B_l}{\partial x^2}(x) g_l(x) \right) + \sum_{l' \in L} \lambda_{ll'}(x) B_{l'}(x)$$

is the generator of the process, and $B(l,x)$ is in the domain of the generator if $B_l \in C^2(\mathbb{R}^n)\ \forall l \in L$ (see [27]). Condition (3.19) implies that $B(l(t), x(t))$ is a supermartingale, which can be shown using Dynkin's formula. Since $B(l(t), x(t))$ is also nonnegative (as implied by (3.18)), Lemma 3.4 can be applied. The rest of the proof is similar to the proof of Theorem 3.5. ■

3.2.3 Stochastic Hybrid Systems

In the class of stochastic hybrid systems proposed by Hu et al. [34], the continuous state of the system evolves according to a stochastic differential equation that depends on the discrete state. When the continuous state reaches a guard set, a discrete transition occurs, where the discrete state after the transition is chosen deterministically, but the continuous state is reset according to a probability distribution that is dependent on the hybrid state before the transition.

Formally, a stochastic hybrid system is $H = (\mathcal{X}, L, I, \mu_0, f_l, g_l, G, R_{ll'})$ with the following components[3]:

- $\mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state space.
- L is a finite set of locations. The overall state space of the system is $X = L \times \mathcal{X}$, and the state is denoted by $(l, x) \in L \times \mathcal{X}$.
- $I : L \to 2^{\mathcal{X}}$ is the invariant, which assigns to each location l an open set $I(l) \subseteq \mathcal{X}$ that contains all possible continuous states while at location l. For our purpose, it will also be assumed that $I(l)$ is bounded.
- μ_0 is a probability measure for the initial state, with its support in $X_0 \subseteq \bigcup_{l \in L}(\{l\} \times I(l))$.

[3]Following [64], we assume that the initial state is drawn according to an initial probability distribution.

- $f_l : I(l) \to \mathbb{R}^n$, $l \in L$, is a set of vector fields, where the subscript indicates the corresponding discrete location.

- $g_l : \mathcal{X} \to \mathbb{R}^n$, $l \in L$, is a set of diffusion coefficients corresponding to the 1-dimensional Wiener process $w(t)$.

- $G : L^2 \to 2^{\mathcal{X}}$ is guard, which assigns to each pair $(l, l') \in L^2$ a set $G(l, l')$ that is a measurable subset of $\partial I(l)$ (note that $G(l, l')$ is possibly empty for some (l, l')'s), and for each $l \in L$ the collection $\{G(l, l') : l' \in L\}$ forms a disjoint partition of $\partial I(l)$.

- $R_{ll'}(x)$, where $(l, l') \in L^2$, $x \in G(l, l')$, is a set of reset probability measures for the continuous state, with its support contained in $I(l')$.

We denote the unsafe region by $X_u = \cup_{l \in L}\{l\} \times \text{Unsafe}(l)$, where $\text{Unsafe}(l) \subseteq I(l)$ and can be empty for some l's.

A trajectory of the system starts with an initial condition drawn from the initial probability measure μ_0. The continuous part of the state evolves according to a stochastic differential equation, which at location l is given by

$$dx(t) = f_l(x(t))dt + g_l(x(t))dw(t).$$

When the continuous state reaches a guard set $G(l, l')$, a transition from location l to location l' occurs. In this transition, the continuous state is reset according to the probability measure $R_{ll'}(x)$, where x is the value of the continuous state before the transition. It is assumed that $f_l(x)$ and $g_l(x)$ are locally Lipschitz continuous and bounded, and also that $R_{ll'}(x)(A)$ is a measurable function in x for each measurable set $A \subset I(l')$. Under this assumption, the solution $(l(t), x(t))$ exists, is unique, and satisfies the right continuity and the Markovian property [64].

For these systems, the conditions that are satisfied by a barrier certificate are stated as follows.

Theorem 3.8 *Let the stochastic hybrid system $H = (\mathcal{X}, L, I, \mu_0, f_l, g_l, G, R_{ll'})$ with bounded $I(l)$'s, and bounded, locally Lipschitz continuous $f_l(x)$'s and $g_l(x)$'s; and the unsafe set $X_u \subseteq \cup_{l\in L}(\{l\} \times I(l))$ be given. Let $\mu_{ll'}(x) \triangleq R_{ll'}(x)$ if $x \in G(l, l')$, and zero otherwise. Suppose there exists a collection of functions $B_l \in C^2(\mathbb{R}^n)$, which satisfy*

$$B_l(x) \geq 1 \quad \forall x \in \text{Unsafe}(l), \tag{3.21}$$

$$B_l(x) \geq 0 \quad \forall x \in I(l), \tag{3.22}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x) + \frac{1}{2}\text{Tr}(g_l^T(x)\frac{\partial^2 B_l}{\partial x^2}(x) g_l(x)) + \int_{I(l')} (B_{l'}(x') - B_l(x)) d\mu_{ll'}(x)(x') \leq 0 \quad \forall x \in \text{cl}(I(l)), \tag{3.23}$$

$$B_l(x) - \int_{I(l')} B_{l'}(x') d\mu_{ll'}(x)(x') = 0 \quad \forall x \in \partial I(l), \tag{3.24}$$

for all $l \in L$ and $(l, l') \in L^2$, and

$$\sum_{l\in L} \int_{I(l)} B_l(x) d\mu_0(l, x) \leq \gamma. \tag{3.25}$$

Then $P\{(l(t), x(t)) \in X_u \text{ for some } t \geq 0\} \leq \gamma$.

Proof. Similar to the proof of Theorem 3.6, but in this case

$$AB(l, x) = \frac{\partial B_l}{\partial x}(x) f_l(x) + \frac{1}{2}\text{Tr}(g_l^T(x)\frac{\partial^2 B_l}{\partial x^2}(x) g_l(x)) + \int_{I(l')} (B_{l'}(x') - B_l(x)) d\mu_{ll'}(x)(x')$$

is the generator of the process, and $B(l, x)$ is in the domain of the generator if $B_l \in C^2(\mathbb{R}^n)$ and the boundary condition (3.24) holds [64]. ■

3.3 Examples

3.3.1 Stochastic Differential Equation

Consider the nonlinear stochastic differential equation

$$\begin{aligned} dx_1(t) &= x_2(t)dt, \\ dx_2(t) &= (-x_1(t) - x_2(t) - 0.5x_1^3(t))dt + \sigma dw(t), \end{aligned}$$

where the diffusion coefficient σ is assumed to be a constant. In this case, the deterministic system corresponding to $\sigma = 0$ has a globally asymptotically stable equilibrium at the origin, as can be proven by a quartic polynomial Lyapunov function. Because of the asymptotic stability of the deterministic system, we expect that for small enough diffusion coefficient σ, the trajectories of the stochastic system will also evolve to a region around the origin.

We use $\mathcal{X} = \{x \in \mathbb{R}^2 : -3 \leq x_1 \leq 3, -3 \leq x_2 \leq 3, x_1^2 + x_2^2 \geq 0.5^2\}$ as the set of states and $\mathcal{X}_0 = \{x \in \mathbb{R}^2 : (x_1 + 2)^2 + x_2^2 \leq 0.1^2\}$ as the initial set. Finally, the set $\mathcal{X}_u = \{x \in \mathcal{X} : x_2 \geq 2.25\}$ will be regarded as the unsafe set. Some realizations of the process $\tilde{x}(t)$ starting from $\mathcal{X}_0$ are depicted in Figure 3.1.

We will compute an upper bound γ on the probability that a stopped process starting from $\mathcal{X}_0$ intersects $\mathcal{X}_u$, as the state evolves toward the origin. For example, this may correspond to the control objective of keeping the value of x_2 lower than the given threshold. Using the theory described in Section 3.1 and the computational method described in Section 2.3, we are able to compute upper bounds as well as polynomial barrier certificates that prove these upper bounds. The verification results for various degrees of barrier certificates and various values of σ are given in Table 3.1. As we include more candidates in the set of candidate barrier certificates to be searched (i.e., as we increase the degree of the barrier certificate), we are able to obtain a better upper bound. However, the computational complexity of solving the sum of squares problem also increases. When we decrease σ, the bound on the reach probability decreases as well. This agrees with our intuition, as the system is

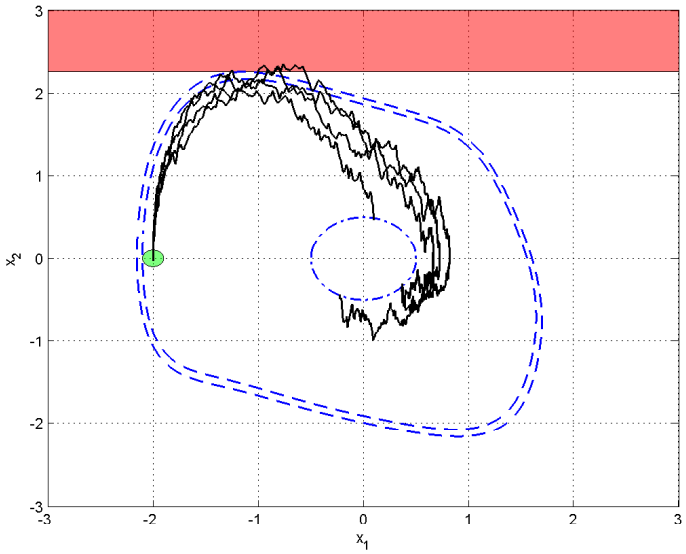


Figure 3.1: Phase portrait of the system in Section 3.3.1. Black curves are some realizations of the stopped process $\tilde{x}(t)$ for $\sigma = 0.5$, all starting at $\tilde{x}(0) = (-2, 0)$. We stop the process when it enters the region whose boundary is depicted by the dash-dotted curve. The shaded region at the top is the unsafe set. Shown as dashed curves are the level sets $B(x) = 1$ (outer) and $B(x) = 0.792$ (inner) of the barrier certificate that proves the upper bound $\gamma = 0.792$ (cf. Table 3.1).

	Degree= 4	Degree= 6	Degree= 8	Degree= 10
$\sigma = 0.5$	$\gamma = 1$	$\gamma = 0.847$	$\gamma = 0.792$	$\gamma = 0.771$
$\sigma = 0.25$	$\gamma = 0.848$	$\gamma = 0.616$	$\gamma = 0.472$	$\gamma = 0.412$
$\sigma = 0.1$	$\gamma = 0.824$	$\gamma = 0.450$	$\gamma = 0.257$	$\gamma = 0.157$

Table 3.1: Results of the stochastic safety verification in Section 3.3.1.

safe when there is no stochastic input.

3.3.2 Switching Diffusion Process

In this example, we consider the system

$$dx(t) = A_{l(t)}x(t) + \sigma(x(t))dw(t),$$

where $l(t) \in \{1, 2\}$ and

$$A_1 = \begin{bmatrix} -5 & -4 \\ -1 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -2 & -4 \\ 20 & -2 \end{bmatrix}, \quad \sigma(x) = \begin{bmatrix} 0 \\ 0.5x_2 \end{bmatrix}.$$

It can be shown using a common polynomial Lyapunov function of degree six that the deterministic system corresponding to $\sigma(x) = 0$ is globally asymptotically stable under arbitrary switching.

We assume that the initial condition is given by $l(0) = 1$ or 2, with equal probability for both locations, and $x(0) = (0, 3)$. For the initial continuous condition $x(0) = (0, 3)$, trajectories of the deterministic system corresponding to the first and second locations are shown in Figure 3.2. We choose $\mathcal{X} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 \leq 4^2, -1.5 \leq x_2 \leq 4\}$ as the set of continuous states, and the unsafe set is given by $\mathcal{X}_u = \{(x_1, x_2) \in \mathcal{X} : x_2 \leq -1\}$. The safety of the stochastic system with transition rates

$$\lambda_{11} = -0.5, \qquad \lambda_{12} = 0.5,$$
$$\lambda_{21} = \lambda, \qquad \lambda_{22} = -\lambda,$$

is to be verified, where the nonnegative parameter λ will be varied. Larger λ means that from location 2 the system tends to switch to location 1 faster.

This problem can be given the following interpretation. Although in both locations the system will evolve toward the origin, location 2 is different from location 1 in the sense that it has an oscillatory response which tends to bring the system to the unsafe region, whereas the trajectory corresponding to location 1 will evolve directly to the origin without going through the unsafe region. In the verification, we will show that by using a large λ – i.e., making the system be in location 1 for most of the time – the probability of reaching the unsafe set can be kept small.

Using polynomial barrier certificates of degree 10, we can prove that the probability of reaching the unsafe region is bounded by $\gamma = 0.346$ for $\lambda = 10$, $\gamma = 0.145$ for $\lambda = 20$, and $\gamma = 0.069$ for $\lambda = 30$. As expected, the probability bound decreases

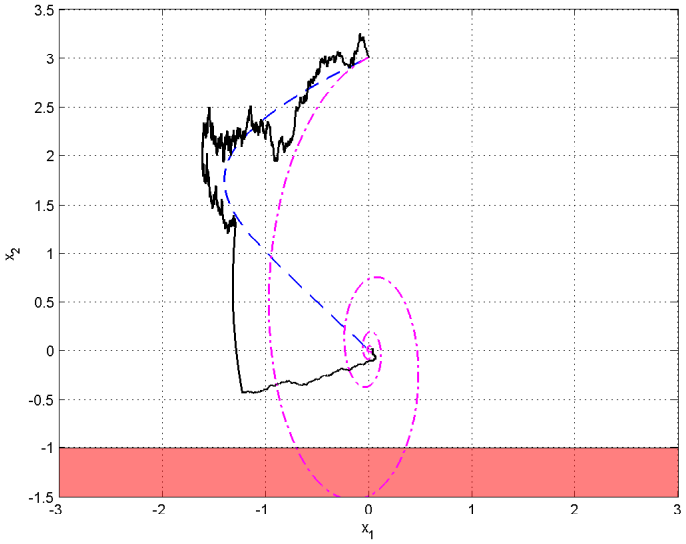


Figure 3.2: Phase portrait of the system in Section 3.3.2. Trajectories of the systems $\dot{x} = A_1 x$ and $\dot{x} = A_2 x$ starting at $x(0) = (0, 3)$ are shown by the dashed and dash-dotted curves, respectively. A realization of the switching diffusion process for $\lambda = 10$ is depicted by the solid curve. Shaded region at the bottom of the figure is the unsafe set.

when we increase λ.

Chapter 4

Reachability and Eventuality Verification

In Chapter 2, a method for safety verification in the worst-case setting using barrier certificates has been proposed. For continuous or hybrid systems, safety verification can also be performed by first constructing a discrete abstraction of the system [2, 4, 10, 87] and then performing verification on the resulting abstraction. This approach provides another hierarchical way (i.e., besides simply increasing the degree of the barrier certificate) for managing the complexity of verification: start with a coarse abstraction and successively refine it until safety is verified or a non-spurious counter-example is found. However, a crucial and computationally demanding component of this approach is still the continuous reachability analysis, which is required to determine whether or not transitions connecting two discrete states in the abstraction is possible (see Figure 4.1).

In constructing discrete abstractions, barrier-certificate-based analysis can be used for ruling out transitions between discrete states. Up to this point, what is still missing is a method for proving that other transitions are indeed possible. This is the problem of *reachability verification*, which can be regarded as the "dual" of safety verification, and concerns with proving that at least one trajectory of the system starting from a set of initial states will reach another given set of states in a finite time. It is important to note that a failure in computing a barrier certificate which proves the unreachability of the target set from the initial set does not, by itself, mean that

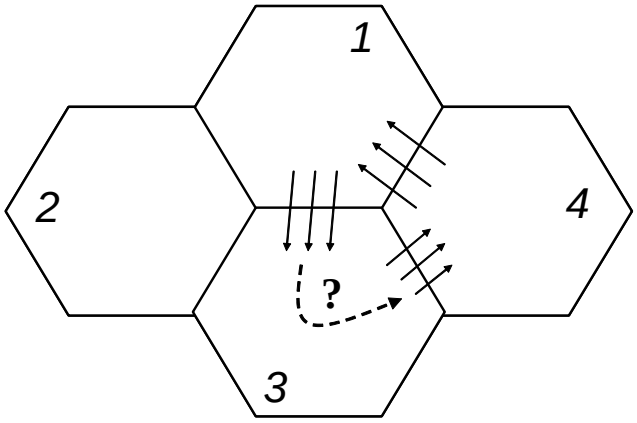


Figure 4.1: System analysis by abstraction. The continuous state space is partitioned into cells, four of which are shown in the figure above. Vector field analysis at the boundaries of the cells indicates that a direct transition from 1 to 4 is not possible, but transitions from 1 to 3, as well as 3 to 4, are possible. The question now is whether the system can evolve from 1 to 4 via 3.

the target set is reachable from the initial set. For example, when using polynomial candidates for $B(x)$, it may be the case that we fail to find $B(x)$ because the degree of the polynomials is not high enough.

It should be noted that besides its usage for constructing discrete abstraction, reachability verification has a purpose on its own right, as properties of interest of the system can often be specified in terms of reachability. For example, it may be of interest to prove that a "good" set of states can be reached by the system, something which can be conveniently expressed as a reachability property.

In the present chapter, we use the ideas of duality and density functions [77,78] to formulate a "dual" test for reachability, thus forming a primal-dual pair of safety and reachability tests, each of which can be solved using convex optimization. This opens the possibility of proving reachability using sum of squares optimization, when the vector field of the system is polynomial and the sets are semialgebraic. Another pair of convex programs for safety and reachability tests will also be formulated, where the primal test now proves reachability and the dual test proves safety. Either of these pairs can be used to rule out or establish transitions between discrete states when

creating and analyzing abstractions of hybrid systems. In addition, we will show that this approach can be used to prove properties such as eventuality and weak eventuality, whose definitions will be presented later, or even other simple combinations of reachability and safety, or eventuality and safety.

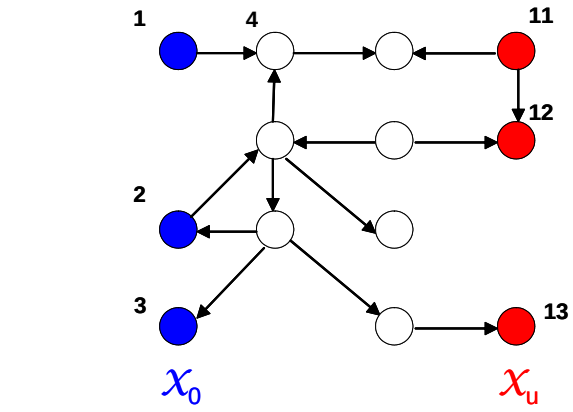
The outline of the chapter is as follows. In Section 4.1, we give an intuitive illustration of the duality idea by addressing the verification of a simple discrete system. Various primal and dual tests for safety, reachability, eventuality, and other simple temporal specifications are presented and proven in Section 4.2. The tests for hybrid systems will be discussed in Section 4.3. Two examples will then be given in Section 4.4. In the first example, a pair of primal-dual tests is used in a successive manner to prove safety and reachability, whereas in the second example, some temporal properties of a Van der Pol oscillator with a disturbance input will be verified.

4.1 Discrete Example

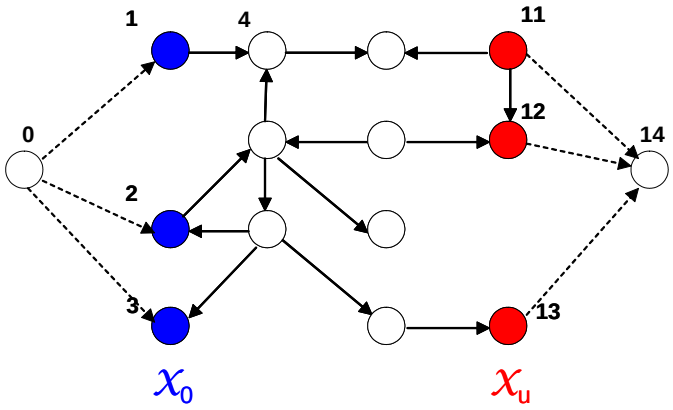
To give an intuitive flavor of the duality ideas used in this chapter, let us consider the verification of a simple discrete system shown in Figure 4.2(a). The system has thirteen states, labelled 1 through 13, and fourteen transitions between states, represented by the directed edges in the graph. We assume that nodes 1–3 are the possible initial states and nodes 11–13 are the bad/unsafe states. The safety verification then amounts to verifying that there is no path that connects any of the initial states to any of the unsafe states.

An equivalent formulation of this problem, but whose conditions for safety are easier to write, is shown in Figure 4.2(b). This graph is obtained by augmenting an extra "source" node (i.e., node 0) and edges that connect it to all initial states, as well as an extra "sink" node (i.e., node 14) and edges that connect all unsafe states to it. It is obvious that the safety property holds for the original transition system, if and only if there is no path connecting node 0 to node 14.

For verifying the safety property, conditions analogous to (2.2)–(2.4) that must



(a)



(b)

Figure 4.2: Verification of a simple discrete transition system. The nodes represent the states of the system, while the directed edges represent transitions between states. In (a), nodes 1–3 are the initial states and nodes 11–13 are the unsafe states. In (b), an extra "source" node (i.e., node 0) and an extra "sink" node (i.e., node 14) are augmented to the graph. It is clear that there is no path that connects any of nodes 1–3 to any of nodes 11–13, if and only if there is no path that connects node 0 to node 14.

be satisfied by a barrier certificate can be formulated. One way to find a barrier certificate which proves safety is by solving the linear program (LP)

$$\begin{aligned} &\max\ s^T B \\ &\text{subject to } A^T B \leq 0 \end{aligned}$$

where $B \triangleq \text{col}(B_0, B_1, B_2, \ldots, B_{14}) \in \mathbb{R}^{15}$ is the decision variable of the LP (i.e., the

barrier certificate), A is the incidence matrix of the graph, which in this case is a 15×20 matrix

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}^T,$$

and s is a 15×1 column vector whose entries are equal to 1 at the first position, -1 at the last position, and zero otherwise. This formulation is similar to the continuous case. Analogous to (2.4), we ask that $B_i \leq B_j$ if there is a directed edge from node i to node j. The objective function of the LP is just the difference between the values of B at the unsafe state and at the initial state. If there is a feasible solution to the above LP such that the objective function is strictly positive, then the safety property can be inferred, i.e., we prove that there is no path going from node 0 to node 14.

The dual of the above LP is as follows:

$$\begin{aligned} &\min\ 0 \\ &\text{subject to } A\rho = s, \\ &\qquad \rho \geq 0, \end{aligned}$$

where $\rho \triangleq \text{col}(\rho_{0,1}, \rho_{0,2}, \rho_{0,3}, \rho_{1,4}, \dots \rho_{13,14}) \in \mathbb{R}^{20}$ is the dual decision variables, whose entries correspond to the edges in the graph. The dual decision variable $\rho_{i,j}$ can be interpreted as the transportation density from node i to node j. The equality constraints basically state that conservation of flows holds at each node, namely that the total flow into a node is equal to the total flow out. In addition, the first and last equality constraints indicate that there exist a unit source at node 0, and a unit sink at node 14. This duality interpretation has been studied extensively in the past; see,

e.g., [60] and references therein.

The existence of a feasible solution to the dual LP implies the existence of a path from the initial state to the unsafe state. This can be shown using the facts that the flows are conserved and that there are a unit source and a unit sink at the initial state and unsafe state, respectively. Hence, solving the dual LP can be used for verifying reachability. As a matter of fact, we obtain a linear programming formulation of the shortest path problem if we also add the objective function $\sum \rho_{i,j}$ to the dual LP. In this case, the nonzero entries corresponding to any optimal vertex solution to the LP will indicate a shortest path from the initial node to the unsafe node [60].

This duality argument can also be used to prove that the existence of a barrier certificate is both sufficient and necessary for safety. For this, suppose that there exists no barrier certificate for the system, which is equivalent to the maximum objective value of the primal LP being equal to zero. This objective value is attained, e.g., by $B_i = 0$ for all i. The linear programming duality [17] implies that there exists a feasible solution to the dual LP, from which we can further conclude the existence of a path from the initial state to the unsafe state, as explained in the previous paragraph. In Chapter 5, a strong duality argument will also be used to prove a converse theorem for barrier certificates later in the continuous case.

For the above example, the optimal objective value of the primal linear program is equal to zero, and hence the safety property does not hold. A feasible solution to the dual linear program is given by $\rho_{0,2} = 1$, $\rho_{2,5} = 1$, $\rho_{5,6} = 1$, $\rho_{6,10} = 1$, $\rho_{10,13} = 1$, $\rho_{13,14} = 1$, and all the other ρ's equal to zero. This solution shows a path from node 0 to node 14. Had the direction of the edge from node 2 to node 5 been reversed, for example, the optimal objective value of the corresponding primal linear program will be ∞, and there will be no feasible solution to the dual linear program.

Other properties of this discrete transition system such as eventuality can also be verified by solving some appropriate linear programs. We will not state them here, but instead we will now proceed to present the corresponding convex programs for continuous systems.

4.2 Continuous Systems

Several convex programs for verifying safety, reachability, eventuality, and some other specifications will be derived in this section. Our notations are as described in Section 1.3. The following version of Liouville's theorem will be used several times in the proofs of the main theorems.

Lemma 4.1 ([77]) *Let $f \in C^1(D, \mathbb{R}^n)$ where $D \subseteq \mathbb{R}^n$ is open and let $\rho \in C^1(D)$ be integrable. Consider the system $\dot{x} = f(x)$. For a measurable set Z, the relation*

$$\int_{\phi_T(Z)} \rho(x)dx - \int_Z \rho(x)dx = \int_0^T \int_{\phi_t(Z)} \left[\nabla \cdot (f\rho)\right](x)dxdt \tag{4.1}$$

holds, provided that $\phi_t(Z)$ is a subset of D for all $t \in [0, T]$.

4.2.1 Safety and Reachability Verification

We define the reachability property of a continuous system as follows.

Definition 4.2 (Reachability) *Given a system $\dot{x} = f(x)$, the state set $\mathcal{X} \subseteq \mathbb{R}^n$, the initial set $\mathcal{X}_0 \subseteq \mathcal{X}$, and the target set $\mathcal{X}_r \subseteq \mathcal{X}$, we say that the reachability property holds if there exist a finite $T \geq 0$ and a trajectory $x(t)$ of the system such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_r$, and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$. (Note that there is no need for $x(t)$ to stay in $\mathcal{X}_r$ for all $t \geq T$).*

At this point, we are ready to state and prove the first pair of convex programs that verify safety and reachability for continuous systems. As the reader may have noticed from Definition 4.2, for now we will assume that there is no disturbance input in the system. In addition, we will also assume that the sets of states are bounded. Some remarks on how to relax these assumptions will be given later.

Theorem 4.3 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded.*

(a) If there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{4.2}$$

$$B(x) > 0 \quad \forall x \in \mathcal{X}_u, \tag{4.3}$$

$$\frac{\partial B}{\partial x}(x) f(x) \leq 0 \quad \forall x \in \mathcal{X}. \tag{4.4}$$

Then the safety property in the sense of Definition 2.1 holds.

(b) If $\mathcal{X}_0$ has a non-empty interior and if there exists a function $\rho \in C^1(\mathbb{R}^n)$ satisfying

$$\int_{\mathcal{X}_0} \rho(x) dx \geq 0, \tag{4.5}$$

$$\rho(x) < 0 \quad \forall x \in \mathrm{cl}(\partial \mathcal{X} \setminus \partial \mathcal{X}_r), \tag{4.6}$$

$$\nabla \cdot (\rho f)(x) > 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{4.7}$$

then the reachability property in the sense of Definition 4.2 holds.

Proof. The statement (a) is a special case of Proposition 2.2, and has been proven in Chapter 2.

To prove the statement (b), let $X \subseteq \mathcal{X}_0$ be an open set on which $\rho(x) \geq 0$. We will first prove that there must be an initial condition $x_0 \in X$ whose flow $\phi_t(x_0)$ leaves $\mathcal{X} \setminus \mathcal{X}_r$ in finite time. In fact, the set of all initial conditions in X whose flows do not leave $\mathcal{X} \setminus \mathcal{X}_r$ in finite time is a set of measure zero. To show this, let Y be an open neighborhood of $\mathcal{X} \setminus \mathcal{X}_r$ such that $\nabla \cdot (\rho f)(x) > 0$ on $\mathrm{cl}(Y)$. Now define

$$Z = \bigcap_{i=1,2,\ldots} \{x_0 \in X : \phi_t(x_0) \in Y \quad \forall t \in [0, i]\}.$$

The set Z is an intersection of countable open sets and hence is measurable. It contains all initial conditions in X for which the trajectories stay in Y for all $t \geq 0$. That Z is a set of measure zero can be shown using Lemma 4.1 as follows. Since

$\phi_t(Z) \subset Y$, Y is bounded, and $\rho(x)$ is continuous, the left hand side of

$$\int_{\phi_t(Z)} \rho(x)dx - \int_Z \rho(x)dx = \int_0^t \int_{\phi_\tau(Z)} \left[\nabla \cdot (f\rho)\right](x)dxd\tau$$

is bounded for all $t \geq 0$. Therefore, for the above equation to hold, we must have $\int_{\phi_\tau(Z)} \left[\nabla \cdot (f\rho)\right](x)dx \to 0$ as $\tau \to \infty$, or equivalently, the measure of $\phi_\tau(Z)$ converges to zero as $\tau \to \infty$. Suppose now that Z has non-zero measure. We have a contradiction since $\lim_{t\to\infty} \int_{\phi_t(Z)} \rho(x)dx = 0$, whereas $\lim_{t\to\infty} \int_0^t \int_{\phi_\tau(Z)} \left[\nabla \cdot (f\rho)\right](x)dxd\tau + \int_Z \rho(x)dx$ is strictly positive, as implied by (4.5) and (4.7). Using this argument, we conclude that Z has measure zero. Since $\mathcal{X} \setminus \mathcal{X}_r \subset Y$, it follows immediately that the set of all initial conditions in X whose flows stay in $\mathcal{X} \setminus \mathcal{X}_r$ for all time is a set of measure zero.

Now take any $x_0 \in X$ whose flow leaves $\mathcal{X} \setminus \mathcal{X}_r$ in finite time; we will show that such a flow must enter $\mathcal{X}_r$ before leaving $\mathcal{X}$. Suppose to the contrary that the flow $\phi_t(x_0)$ leaves $\mathcal{X}$ without entering $\mathcal{X}_r$ first. Let $T > 0$ be the "first" time instant $\phi_t(x_0)$ leaves $\mathcal{X}$. By this, we mean that either $\phi_t(x_0) \in \mathcal{X} \setminus \mathcal{X}_r$ for all $t \in [0, T)$ and $\phi_T(x_0) \notin \mathcal{X}$; or $\phi_t(x_0) \in \mathcal{X} \setminus \mathcal{X}_r$ for all $t \in [0, T]$ and $\phi_{T+\epsilon}(x_0) \notin \mathcal{X}$ for any $\epsilon > 0$. From conditions (4.6)–(4.7), it follows that for a sufficiently small neighborhood U of x_0, we have

$$\begin{aligned} \rho(x) &\geq 0 \quad \forall x \in U, \\ \rho(x) &< 0 \quad \forall x \in \phi_T(U), \\ \nabla \cdot (\rho f)(x) &> 0 \quad \forall x \in \phi_t(U), t \in [0, T]. \end{aligned}$$

Apply Lemma 4.1 again to obtain a contradiction. According to the above, the left hand side of

$$\int_{\phi_T(U)} \rho(x)dx - \int_U \rho(x)dx = \int_0^T \int_{\phi_\tau(U)} \left[\nabla \cdot (f\rho)\right](x)dxd\tau$$

is negative while the right hand side is positive. Thus, there is a contradiction, and we conclude that for $x(0) = x_0$, there must exist $T \geq 0$ such that $x(T) \in \mathcal{X}_r$ and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$. ■

It is interesting to see that the roles of $B(x)$ and $\rho(x)$ in proving safety and reachability can be interchanged, as in the second pair of tests stated in the next theorem. The possibility of using the density function $\rho(x)$ to prove safety was first suggested in [79].

Theorem 4.4 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded.*

(a) If $\mathcal{X}_0$ has a non-empty interior and if there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying

$$\int_{\mathcal{X}_0} B(x)dx \leq 0, \tag{4.8}$$

$$B(x) > 0 \quad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r), \tag{4.9}$$

$$\frac{\partial B}{\partial x}(x)f(x) < 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{4.10}$$

then the reachability property in the sense of Definition 4.2 holds.

(b) If there exist open sets $\tilde{\mathcal{X}}_0$, $\tilde{\mathcal{X}}$ and a function $\rho \in C^1(\mathbb{R}^n)$ such that $\mathcal{X}_0 \subseteq \tilde{\mathcal{X}}_0$, $\mathcal{X} \subseteq \tilde{\mathcal{X}}$, and

$$\rho(x) \geq 0 \quad \forall x \in \tilde{\mathcal{X}}_0, \tag{4.11}$$

$$\rho(x) < 0 \quad \forall x \in \mathcal{X}_u, \tag{4.12}$$

$$\nabla \cdot (\rho f)(x) \geq 0 \quad \forall x \in \tilde{\mathcal{X}}, \tag{4.13}$$

then the safety property in the sense of Definition 2.1 holds.

Proof. To prove (a), consider a point $x_0 \in \mathcal{X}_0$ such that $B(x_0) \leq 0$. The flow $\phi_t(x_0)$ must leave $\mathcal{X} \setminus \mathcal{X}_r$ in finite time, since the derivative inequality (4.10) holds and $B(x)$ is bounded below on $\mathcal{X}$. Now, suppose that $\phi_t(x_0)$ leaves $\mathcal{X}$ without entering $\mathcal{X}_r$

first, and consider the "first" time instant $t = T$ at which it happens. Similar to the proof of Theorem 4.3, by this we mean that either $\phi_t(x_0) \in \mathcal{X} \setminus \mathcal{X}_r$ for all $t \in [0, T)$ and $\phi_T(x_0) \notin \mathcal{X}$; or $\phi_t(x_0) \in \mathcal{X} \setminus \mathcal{X}_r$ for all $t \in [0, T]$ and $\phi_{T+\epsilon}(x_0) \notin \mathcal{X}$ for any $\epsilon > 0$. From (4.10) and $B(x_0) \leq 0$, it follows that $B(\phi_T(x_0))$ is non-positive, which is contradictory to (4.9). Thus, we conclude that for $x(0) = x_0$, there must exist $T \geq 0$ such that $x(T) \in \mathcal{X}_r$ and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

We proceed to proving (b). Assume that there is a $\rho(x)$ satisfying the conditions of the theorem, while at the same time there exists an $x_0 \in \mathcal{X}_0$ such that $\phi_T(x_0) \in \mathcal{X}_u$ for some $T \geq 0$ and $\phi_t(x_0) \in \mathcal{X}$ for $t \in [0, T]$. Then it follows from (4.11)–(4.13) that for a sufficiently small neighborhood Z of x_0, we have

$$\begin{aligned} \rho(x) &\geq 0 \quad \forall x \in Z, \\ \rho(x) &< 0 \quad \forall x \in \phi_T(Z), \\ \nabla \cdot (\rho f)(x) &\geq 0 \quad \forall x \in \phi_t(Z), t \in [0, T]. \end{aligned}$$

Now apply Lemma 4.1 to obtain a contradiction. According to the above, the left hand side of

$$\int_{\phi_T(Z)} \rho(x)dx - \int_Z \rho(x)dx = \int_0^T \int_{\phi_\tau(Z)} \left[\nabla \cdot (f\rho)\right](x)dxd\tau.$$

is negative and the right hand side is non-negative. Hence there is a contradiction and the proof is complete. ■

Remark 4.5 *Modulo the following modifications on the assertions of the theorems, the conclusions of Theorems 4.3 and 4.4 will still hold even when the sets are not bounded. In particular, for the second statement of Theorem 4.3, we need to add the condition that $\rho(x)$ is integrable on $\mathcal{X}$ and replace (4.7) by*

$$\nabla \cdot (\rho f)(x) \geq \epsilon \qquad\qquad \forall x \in (\mathcal{X} \setminus \mathcal{X}_r)$$

for a positive number ϵ. In the first statement of Theorem 4.4, we need to add the

condition that $B(x)$ is bounded below on $\mathcal{X}$ and replace (4.10) by

$$\frac{\partial B}{\partial x}(x)f(x) \leq -\epsilon \qquad\qquad \forall x \in (\mathcal{X} \setminus \mathcal{X}_r)$$

for a positive number ϵ.

In applications where the system has stable equilibrium points, it is often convenient to exclude a neighborhood of the equilibria from the region where the divergence inequality (4.13) must be satisfied, since the inequality is otherwise impossible to satisfy without a singularity in $\rho(x)$. This does not make the conclusion of the theorem weaker, as long as the excluded set does not intersect $\mathcal{X}_u$ and is entirely surrounded by a region of positive $\rho(x)$.

Similarly, the Lie derivative inequality (4.10) is impossible to satisfy when the system has equilibrium points in $\mathcal{X} \setminus \mathcal{X}_r$. In this case, a neighborhood of the equilibria should also be excluded from the region where the inequality is to be satisfied. The conclusion of the theorem is still valid as long as the excluded set is entirely surrounded by a region of positive $B(x)$.

Notice in particular that all the tests presented above are convex programming problems. This opens the possibility of computing $B(x)$ and $\rho(x)$ using convex optimization. For systems whose vector fields are polynomial and whose set descriptions are semialgebraic, a computational method based on sum of squares optimization is available, if we use polynomial parameterizations for $B(x)$ or $\rho(x)$. This computational technique has been described in Section 2.3 of this thesis.

When we set $\mathcal{X}_u = \mathcal{X}_r$, each pair of the convex programs in Theorems 4.3 and 4.4 form a pair of *weak alternatives*: at most one of them can be feasible. Nevertheless, strictly speaking it should be noted that the tests in the above theorems are not pairs of *Lagrange dual* problems [17] in the sense of convex optimization. We deliberately do not use Lagrange dual problems to avoid computational problems when we postulate $B(x)$ or $\rho(x)$ as polynomials. For example, the Lagrange dual problem of the safety test in Theorem 4.3 will require $\nabla \cdot (\rho f)(x)$ to be zero on $\mathcal{X} \setminus (\mathcal{X}_0 \cup \mathcal{X}_u)$ (cf. Section 5.1). Although useful for theoretical purposes, this will hinder the computation

of $\rho(x)$ through polynomial parameterization and sum of squares optimization. In this regard, some interesting future directions would be to see if a pair of Lagrange dual problems can be formulated so that both problems can be solved using sum of squares optimization, or more importantly, to see if the dual infeasibility certificate of one convex program can be interpreted directly as a feasible solution to the dual convex program.

4.2.2 Eventuality Verification

In the reachability test of Theorem 4.4, the set of states $\{x \in \mathcal{X}_0 : B(x) \leq 0\}$ is said to satisfy the *eventuality* property, which we define as follows.

Definition 4.6 (Eventuality) *Given a system $\dot{x} = f(x)$, the state set $\mathcal{X} \subseteq \mathbb{R}^n$, the initial set $\mathcal{X}_0 \subseteq \mathcal{X}$, and the target set $\mathcal{X}_r \subseteq \mathcal{X}$, we say that the eventuality property holds if for all initial conditions $x_0 \in \mathcal{X}_0$, any trajectory $x(t)$ of the system starting at $x(0) = x_0$ satisfies $x(T) \in \mathcal{X}_r$ and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$ for some $T \geq 0$.*

Analogously, in Theorem 4.3, the set of states $\{x \in \mathcal{X}_0 : \rho(x) \geq 0\}$ is said to satisfy the *weak eventuality* property, defined as follows.

Definition 4.7 (Weak Eventuality) *Given a system $\dot{x} = f(x)$, the state set $\mathcal{X} \subseteq \mathbb{R}^n$, the initial set $\mathcal{X}_0 \subseteq \mathcal{X}$, and the target set $\mathcal{X}_r \subseteq \mathcal{X}$, we say that the weak eventuality property holds if for almost all[1] initial conditions $x_0 \in \mathcal{X}_0$, any trajectory $x(t)$ of the system starting at $x(0) = x_0$ satisfies $x(T) \in \mathcal{X}_r$ and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$ for some $T \geq 0$.*

These facts are evident from the proofs of the theorems. In many applications, it is of paramount importance to prove eventuality (or even weak eventuality), e.g., to prove that something "good" will happen. The eventuality and weak eventuality tests for the whole initial set $\mathcal{X}_0$ can be performed simply by replacing (4.8) by $B(x) \leq 0$ $\forall x \in \mathcal{X}_0$, and (4.5) by $\rho(x) \geq 0$ $\forall x \in \mathcal{X}_0$, where in the latter we also require that $\mathcal{X}_0$ has a non-empty interior.

[1] This is in the sense that the set of all initial conditions that do not satisfy the criteria has measure zero.

Proposition 4.8 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded. Suppose that there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying*

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{4.14}$$

$$B(x) > 0 \quad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r), \tag{4.15}$$

$$\frac{\partial B}{\partial x}(x) f(x) < 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r). \tag{4.16}$$

Then the eventuality property in the sense of Definition 4.6 holds.

Proof. Analogous to the proof of the first statement of Theorem 4.4. ■

Proposition 4.9 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded. If $\mathcal{X}_0$ has a non-empty interior and if there exists a function $\rho \in C^1(\mathbb{R}^n)$ satisfying*

$$\rho(x) \geq 0 \quad \forall x \in \mathcal{X}_0, \tag{4.17}$$

$$\rho(x) < 0 \quad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r), \tag{4.18}$$

$$\nabla \cdot (\rho f)(x) > 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{4.19}$$

then the weak eventuality property in the sense of Definition 4.7 holds.

Proof. Analogous to the proof of the second statement of Theorem 4.3. ■

Example 4.10 *To show that the weak eventuality property mentioned above cannot in general be strengthened to eventuality, consider the system $\dot{x} = x$, with $\mathcal{X} = [-5, 5] \subset \mathbb{R}$, $\mathcal{X}_0 = [-1, 1]$, $\mathcal{X}_r = [-5, -4] \cup [4, 5]$. The function $\rho(x) = 1$ satisfies all the conditions that guarantee weak eventuality. Hence, almost all trajectories starting from $\mathcal{X}_0$ will reach $\mathcal{X}_r$ in finite time. The only exception in this case is the trajectory $x(t) = 0$.*

4.2.3 Other Verification

While one may argue that the reachability property can be shown by running a numerical simulation of $\dot{x} = f(x)$ starting from a properly chosen $x_0 \in \mathcal{X}_0$, the merit of the tests in Theorems 4.3 and 4.4 is twofold. First, a solution to the convex programs for reachability will automatically indicate a set from which all (or almost all) points can be chosen as the initial state. Second, the use of these convex programs allows us to consider also the worst-case analysis of systems with disturbance, or even the controller design problem. For example, consider a system $\dot{x} = f(x, d)$, where the disturbance signal $d(t)$ is assumed to be piecewise continuous, bounded, and takes its value in a set $\mathcal{D}$. Then solving (4.8)–(4.10) with the Lie derivative inequality replaced by

$$\frac{\partial B}{\partial x}(x) f(x, d) \leq -\epsilon \quad \forall (x, d) \in (\mathcal{X} \setminus \mathcal{X}_r) \times \mathcal{D} \tag{4.20}$$

will prove reachability under all possible disturbance $d(t)$, which obviously *cannot be proven using simulation*. The same remark applies to eventuality, which cannot be proven using simulation even when there exists no disturbance.

At the moment, it is unclear how a similar worst-case analysis for systems with time-varying disturbance can be formulated using $\rho(x)$. However, as pointed out in [77], the density function $\rho(x)$ seems to have a better convexity property that is more beneficial for controller design. For a system $\dot{x} = f(x) + g(x)u(x)$ where $u(x)$ is the control input (assumed to be in a state feedback form), the inequalities (4.5)–(4.6) and

$$\nabla \cdot [\rho(f + ug)](x) > 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r),$$

(and similarly for (4.11)–(4.13)) are certainly convex conditions on the pair $(\rho, \rho u)$. It is therefore natural to introduce $\psi = \rho u$ as a search variable and use convex optimization to find a feasible pair (ρ, ψ), then recover the control law as $u(x) = \psi(x)/\rho(x)$. Some results on this are available in [79].

It is clear that the tests in the previous subsections can be combined to prove the reachability – safety property:

> There exists a trajectory $x(t)$ such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_r$ for some $T \geq 0$, and $x(t) \notin \mathcal{X}_u$, $x(t) \in \mathcal{X}$ for all $t \in [0, T]$,

and the eventuality – safety[2] (or weak eventuality – safety) property:

> For all (or almost all) initial states $x_0 \in \mathcal{X}_0$, the trajectory $x(t)$ starting at $x(0) = x_0$ will satisfy $x(T) \in \mathcal{X}_r$ for some $T \geq 0$ and $x(t) \notin \mathcal{X}_u$, $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

For instance, the tests for eventuality – safety and weak eventuality – safety properties are stated in the following corollaries.

Corollary 4.11 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded. Suppose that there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying*

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{4.21}$$

$$B(x) > 0 \quad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r) \cup \mathcal{X}_u, \tag{4.22}$$

$$\frac{\partial B}{\partial x}(x) f(x) < 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r). \tag{4.23}$$

Then the eventuality – safety property holds.

Corollary 4.12 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$, $\mathcal{X}_r \subseteq \mathcal{X}$ be bounded. If $\mathcal{X}_0$ has a non-empty interior and if there exists a function $\rho \in C^1(\mathbb{R}^n)$ satisfying*

$$\rho(x) \geq 0 \quad \forall x \in \mathcal{X}_0, \tag{4.24}$$

$$\rho(x) < 0 \quad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r) \cup \mathcal{X}_u, \tag{4.25}$$

$$\nabla \cdot (\rho f)(x) > 0 \quad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{4.26}$$

[2] In linear temporal logic (LTL), for example, this property corresponds to the "until" operator [46].

then the weak eventuality – safety property holds. In this case, the safety property holds also for trajectories that do not reach $\mathcal{X}_r$ in finite time.

4.3 Hybrid Systems

Some of the tests proposed in the last section, namely those that are based on $B(x)$, can be directly extended to handle hybrid systems. We will illustrate this by presenting a test for the eventuality property of a hybrid system, although similar extension can be derived for other temporal properties. For this, the idea of using multiple $B(x)$'s, similar to the one in Chapter 2, will be adopted. The model of hybrid system that we use is the same as in Section 2.2.1. First, we define the eventuality property as follows.

Definition 4.13 (Eventuality – Hybrid Systems) *Given a hybrid system H and a target set $X_r \subseteq X$, the eventuality property holds if for all initial conditions $(l_0, x_0) \in X_0$, any valid trajectory (i.e., trajectory that corresponds to a piecewise continuous and bounded disturbance $d(t) \in \mathcal{D}(l(t))$) $(l(t), x(t))$ of the system starting at $(l(0), x(0))$ satisfies $(l(T), x(T)) \in X_r$ and $x(t) \in I(l(t))$ $\forall t \in [0, T]$ for some $T \geq 0$.*

In the proposition below, we define the sets of continuous target states as $\text{Reach}(l) = \{x \in \mathcal{X} : (l, x) \in X_r\}$, for $l \in L$.

Proposition 4.14 *Let the hybrid system $H = (\mathcal{X}, L, X_0, I, F, \mathcal{T})$ with bounded $\mathcal{X} \subset \mathbb{R}^n$ and the target set $X_r \subseteq X$ be given. Suppose there exists a collection $\{B_l(x) : l \in L\}$ of functions $B_l \in C^1(\mathbb{R}^n, \mathbb{R})$ which, for all $l \in L$ and $(l, l') \in L^2$, $l \neq l'$ satisfy*

$$B_l(x) \leq 0 \quad \forall x \in \text{Init}(l), \tag{4.27}$$

$$B_l(x) > 0 \quad \forall x \in \text{cl}(\partial I(l) \setminus \partial \text{Reach}(l)), \tag{4.28}$$

$$\frac{\partial B_l}{\partial x}(x) f_l(x, d) \leq -\epsilon \quad \forall (x, d) \in (I(l) \setminus \text{Reach}(l)) \times \mathcal{D}(l), \tag{4.29}$$

$$B_{l'}(x') - B_l(x) \leq -\epsilon \quad \forall x' \in \text{Reset}(l, l')(x) \setminus \text{Reach}(l'), \text{ for all } x \in \text{Guard}(l, l') \setminus \text{Reach}(l) \tag{4.30}$$

for some positive number ϵ. Then the eventuality property in the sense of Definition 4.13 is guaranteed.

Proof. Consider a valid trajectory of the system starting at $(l_0, x_0) \in X_0$ and the evolution of $B_l(t)(x(t))$ along this trajectory. Since the inequalities (4.29)–(4.30) hold and each $B_l(x)$ is bounded below on the bounded set $\mathcal{X}$, the trajectory must leave $\cup_{l \in L}\{l\} \times (I(l) \setminus \mathrm{Reach}(l))$ in a finite time and after a finite number of discrete transitions. Using an argument similar to the proof of (a) in Theorem 4.3, it can be shown that the above trajectory does not leave $\cup_{l \in L}\{l\} \times I(l)$ without entering X_r first. Thus the statement of the proposition follows. ■

Although the currently available tests for hybrid systems are based on $B(x)$, we expect that the tests based on $\rho(x)$ can also be extended to handle hybrid systems directly, using an approach that is analogous to the one presented here.

4.4 Examples

4.4.1 Successive Safety and Reachability Refinements

Consider the system

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 + \frac{1}{3}x_1^3 - x_2, \end{aligned}$$

and let the set of states be $\mathcal{X} = [-3.5, 3.5] \times [-3.5, 3.5] \subset \mathbb{R}^2$. Furthermore, define

$$\begin{aligned} \mathcal{X}_0 &= [-3.4, 3.4] \times [3.35, 3.45], & \mathcal{X}_2 &= [-3.5, 3.5] \times \{-3.5\}, \\ \mathcal{X}_1 &= \{3.5\} \times [-3.5, 3.5], & \mathcal{X}_3 &= \{-3.5\} \times [-3.5, 3.5]. \end{aligned}$$

In this example, we will investigate the reachability of $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{X}_3$ from $\mathcal{X}_0$. This kind of analysis is encountered when constructing a discrete abstraction of continuous or hybrid systems, or when analyzing a counter-example found during the verification

of such an abstraction (cf. Figure 4.1).

The tests in Theorem 4.3 will be used for our analysis. Since the vector field is polynomial and the sets are semialgebraic, we use polynomial parameterization for $B(x)$ and $\rho(x)$, and then apply the sum of squares method to compute them. Degree bound is imposed on $B(x)$ and $\rho(x)$. Because of this, we might not be able to find a single $B(x)$ or $\rho(x)$ that prove safety or reachability for the whole $\mathcal{X}_0$. If neither $B(x)$ nor $\rho(x)$ can be found, we divide the interval of x_1 in $\mathcal{X}_0$ into two parts and apply the tests again to the smaller sets. A set is pruned if $B(x)$ is found, and this process is repeated until a $\rho(x)$ is found or the whole $\mathcal{X}_0$ is proven safe. When the degree of $B(x)$ or $\rho(x)$ is chosen equal to eight, the semidefinite program for each safety or reachability test at any node can be solved in less than four seconds on a Pentium III 600 MHz machine.

The result is as follows.

1. We prove that the set $\mathcal{X}_1$ is reachable from $\mathcal{X}_0$. The verification progress is shown in Figure 4.3 (a).

2. It can be proven directly that $\mathcal{X}_2$ is not reachable from $\mathcal{X}_0$.

3. It is proven that the set $\mathcal{X}_3$ is reachable from $\mathcal{X}_0$ (see Figure 4.3 (b)).

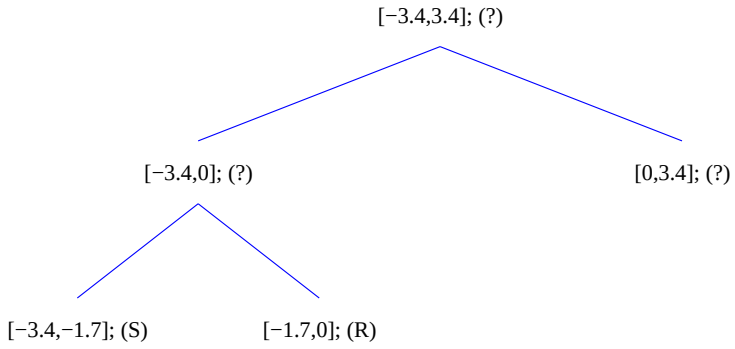
For visualizations of reachability and safety proofs, see Figure 4.4.

Obviously, the above bisection algorithm is just a simple, straightforward approach to refine and prune the initial set, and other algorithms that are more efficient can be proposed in the future.

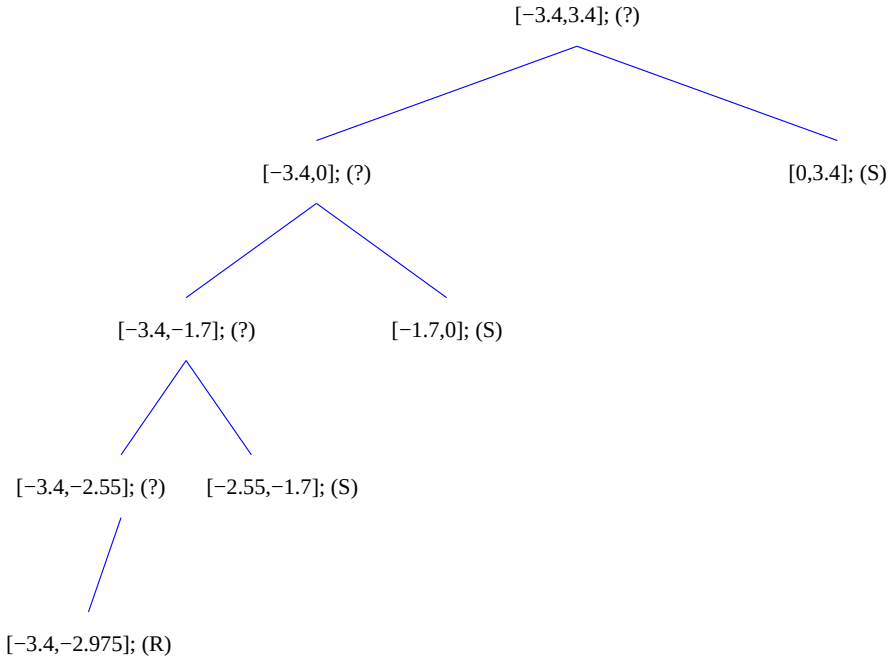
4.4.2 Eventuality and Eventuality – Safety Verification

Consider a Van der Pol oscillator with disturbance input:

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= x_2(1-x_1^2) - x_1 + d,\end{aligned}$$

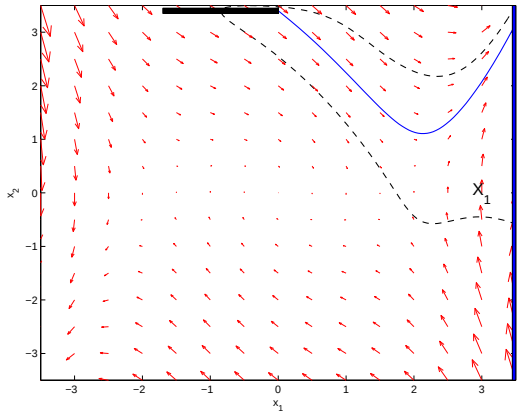


(a) $\mathcal{X}_0 \to \mathcal{X}_1$

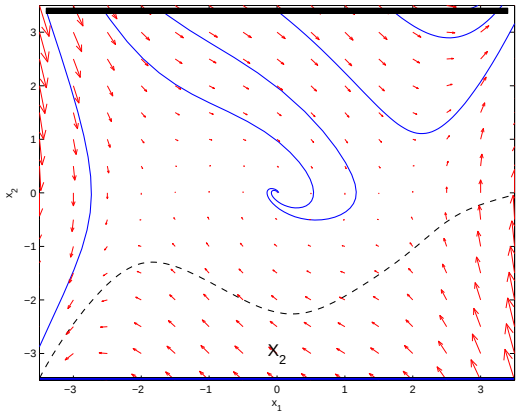


(b) $\mathcal{X}_0 \to \mathcal{X}_3$

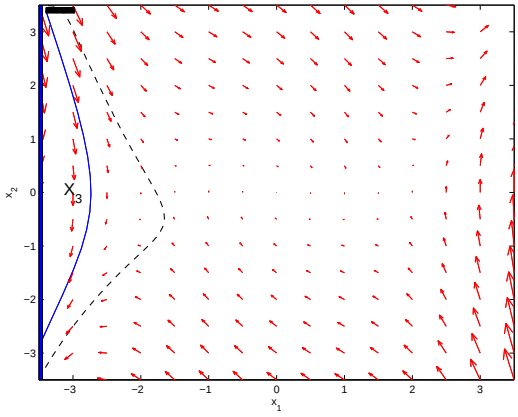
Figure 4.3: Proving the reachability of $\mathcal{X}_1$ and $\mathcal{X}_3$ from $\mathcal{X}_0$ in the example of Section 4.4.1. At each node, we indicate the range of x_1 in $\mathcal{X}_0$ for which safety and reachability are tested. If neither is verified (denoted by ?), then the x_1-interval is divided into two and the tests are applied to the smaller sets. The annotation S (respectively R) indicates that $B(x)$ (respectively $\rho(x)$) is found. Breadth-first search starting from the leftmost branch is used. The verification of $\mathcal{X}_0 \nrightarrow \mathcal{X}_2$ terminates at the top node, since a barrier certificate $B(x)$ can be found directly.



(a) $\mathcal{X}_0 \rightarrow \mathcal{X}_1$



(b) $\mathcal{X}_0 \nrightarrow \mathcal{X}_2$



(c) $\mathcal{X}_0 \rightarrow \mathcal{X}_3$

Figure 4.4: Possible transitions from $\mathcal{X}_0$ to $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ in the example of Section 4.4.1. In (a) and (c), dashed curves are the zero level sets of $\rho(x)$'s that certify reachability. In (b), dashed curve is the zero level set of $B(x)$ that certifies safety. Thick solid lines at the top of the figures are the initial sets for which the certificates are computed. Some trajectories of the system are depicted by solid curves.

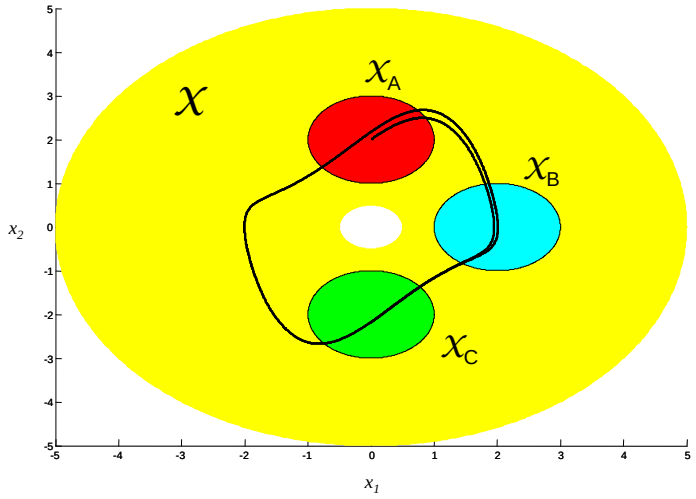


Figure 4.5: Verifying the temporal properties of a Van der Pol oscillator with disturbance. We want to verify that under all possible disturbance input, if the system starts in $\mathcal{X}_A$, then both $\mathcal{X}_B$ and $\mathcal{X}_C$ are reached in finite time, but $\mathcal{X}_C$ will not be reached before the system reaches $\mathcal{X}_B$. The nominal trajectory of the system (i.e., for $d = 0$) starting at $x = (0, 2)$ is depicted by the solid curve.

where d is the disturbance input, taking its value in $\mathcal{D} = [-0.25, 0.25] \subset \mathbb{R}$. Let $\mathcal{X} = \{x \in \mathbb{R}^2 : 0.5 \leq \|x\|_2 \leq 5\}$. In addition, let

$$\mathcal{X}_A = \{x \in \mathbb{R}^2 : (x_1)^2 + (x_2 - 2)^2 \leq 1\},$$
$$\mathcal{X}_B = \{x \in \mathbb{R}^2 : (x_1 - 2)^2 + (x_2)^2 \leq 1\},$$
$$\mathcal{X}_C = \{x \in \mathbb{R}^2 : (x_1)^2 + (x_2 + 2)^2 \leq 1\}.$$

These sets are depicted in Figure 4.5, where a nominal trajectory of the system starting at $x = (0, 2)$ is also shown. Our objective in this example is to verify that under all possible piecewise continuous and bounded disturbance $d(t)$, if the system starts in $\mathcal{X}_A$, then both $\mathcal{X}_B$ and $\mathcal{X}_C$ are reached in finite time, but $\mathcal{X}_C$ will not be reached before the system reaches $\mathcal{X}_B$.

To verify this temporal specification, we will search for two barrier certificates

$B_1(x)$ and $B_2(x)$ satisfying the following conditions:

$$\begin{cases} B_1(x) \leq 0 \quad \forall x \in \mathcal{X}_A, \\ B_1(x) > 0 \quad \forall x \in \partial\mathcal{X} \cup \mathcal{X}_C, \\ \frac{\partial B_1}{\partial x}(x) f(x,d) \leq -\epsilon \quad \forall (x,d) \in (\mathcal{X} \setminus \mathcal{X}_B) \times \mathcal{D}, \end{cases}$$

$$\begin{cases} B_2(x) \leq 0 \quad \forall x \in \mathcal{X}_A, \\ B_2(x) > 0 \quad \forall x \in \partial\mathcal{X}, \\ \frac{\partial B_2}{\partial x}(x) f(x,d) \leq -\epsilon \quad \forall x \in (\mathcal{X} \setminus \mathcal{X}_C) \times \mathcal{D}, \end{cases}$$

for some positive ϵ. Using sum of squares optimization, polynomial $B_1(x)$ and $B_2(x)$ of degree ten can be found, thus the temporal specification is verified.

Chapter 5

On the Necessity of Barrier Certificates

In this chapter, the idea of *strong duality* between convex programs (cf. Section 4.1) will be exploited to derive a converse theorem for safety verification of continuous systems using barrier certificates. Under some reasonable technical conditions, we will prove in Section 5.1 that the existence of a barrier certificate is both sufficient and necessary for safety. In Section 5.2, we will give comments on some cases in which these technical conditions are automatically satisfied.

5.1 A Converse Theorem

The main result of the section can be stated as follows.

Theorem 5.1 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$, and $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$ be compact sets, and suppose that there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x)f(x) < 0$ for all $x \in \mathcal{X}$. Then there exists a function $B \in C^1(\mathbb{R}^n)$ that satisfies*

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{5.1}$$

$$B(x) > 0 \quad \forall x \in \mathcal{X}_u, \tag{5.2}$$

$$\frac{\partial B}{\partial x}(x)f(x) \leq 0 \quad \forall x \in \mathcal{X} \tag{5.3}$$

if and only if the safety property holds.

Notice that in the theorem we have used a seemingly strong assumption that there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x)f(x) < 0 \ \forall x \in \mathcal{X}$. As mentioned at the beginning of the chapter, in the next section we will show that in many cases the existence of $\tilde{B}(x)$ is actually guaranteed.

Before proving the theorem, we will present and prove the following lemmas, which will be used in the proof of the main theorem.

Lemma 5.2 *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, and $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$ be compact sets. Suppose there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x)f(x) < 0$ for all $x \in \mathcal{X}$. Then there exists no $B \in C^1(\mathbb{R}^n)$ satisfying (5.1)–(5.3) only if there are measures of bounded variation ψ_0, ψ_u, ρ (each defined on $\mathbb{R}^n$) such that ψ_0, ψ_u, ρ are nonnegative on $\mathbb{R}^n$ and equal to zero outside $\mathcal{X}_0$, $\mathcal{X}_u$, and $\mathcal{X}$ respectively; and*

$$\begin{aligned} \int_{\mathcal{X}_0} d\psi_0 &= 1, \\ \int_{\mathcal{X}_u} d\psi_u &= 1, \\ \nabla \cdot (\rho f) &= \psi_0 - \psi_u, \end{aligned}$$

where $\nabla \cdot (\rho f)$ is interpreted as a distributional derivative.

Proof. Let us consider the convex optimization problem

$$\begin{aligned} &\sup B_u - B_0, \\ &\text{subject to } B(x) - B_0 \leq 0 \quad \forall x \in \mathcal{X}_0, \\ &\qquad\qquad B(x) - B_u \geq 0 \quad \forall x \in \mathcal{X}_u, \\ &\qquad\qquad \frac{\partial B}{\partial x}(x)f(x) \leq 0 \quad \forall x \in \mathcal{X}, \end{aligned}$$

with the supremum denoted by γ, and taken over all $B_0 \in \mathbb{R}$, $B_u \in \mathbb{R}$, and $B \in C^1(\mathbb{R}^n)$. Since $B_0 = 0$, $B_u = 0$, and $B(x) = 0$ satisfy the constraint, γ must be greater than or equal to zero. In addition, since the objective function and the

constraints are all linear, the value of γ is either 0 or ∞. There exists no $B \in C^1(\mathbb{R}^n)$ satisfying (5.1)–(5.3) if and only if the value of γ is equal to zero.

Now suppose that $\gamma = 0$. Let $\mathcal{K} = \mathbb{R} \times (C(\mathcal{X}))^3$, $\mathcal{B} = \mathbb{R}^2 \times C_0^1(\mathbb{R}^n)$, and define $\mathcal{K}_1$, $\mathcal{K}_2$ as follows:

$$\begin{aligned}\mathcal{K}_1 = \{(z, h_0, h_u, h) \in \mathcal{K} : h_0 = B_0 - B,\ h_u = B - B_u,\ h = -\frac{\partial B}{\partial x} f \text{ on } \mathcal{X};\\ z = B_u - B_0;\ \text{ and } (B_0, B_u, B) \in \mathcal{B}\},\\ \mathcal{K}_2 = \{(z, h_0, h_u, h) \in \mathcal{K} : z \geq 0,\ h_0 \geq 0 \text{ on } \mathcal{X}_0,\ h_u \geq 0 \text{ on } \mathcal{X}_u,\ h \geq 0 \text{ on } \mathcal{X}\}.\end{aligned}$$

Then both $\mathcal{K}_1$ and $\mathcal{K}_2$ are convex sets, and $\mathcal{K}_2$ has non-empty interior in $\mathcal{K}$. Furthermore, since $\gamma = 0$, it follows that the first component in $\mathcal{K}_1$ is less than or equal to zero when the second, third, and fourth components are greater than or equal to zero, and therefore $\mathcal{K}_1 \cap \text{int}(\mathcal{K}_2) = \emptyset$. Now, by the Hahn-Banach theorem [44], there exists a nonzero $k^* = (a, \tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho}) \in \mathcal{K}^* = \mathbb{R} \times (C(\mathcal{X})^*)^3$ such that

$$\sup_{k_1 \in \mathcal{K}_1} \langle k^*, k_1 \rangle \leq \inf_{k_2 \in \mathcal{K}_2} \langle k^*, k_2 \rangle, \tag{5.4}$$

where $C(\mathcal{X})^*$ in this case is the set of measures on $\mathcal{X}$ with bounded variation. The right-hand side of the inequality can be expanded as follows

$$\begin{aligned}\inf_{k_2 \in \mathcal{K}_2} \langle k^*, k_2 \rangle &= \inf_{(z, h_0, h_u, h) \in \mathcal{K}_2} az + \langle \tilde{\psi}_0, h_0 \rangle + \langle \tilde{\psi}_u, h_u \rangle + \langle \tilde{\rho}, h \rangle \\ &= \begin{cases} 0, & \text{if } a \geq 0;\ \tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho} \geq 0;\ \text{and} \\ & \quad \tilde{\psi}_0, \tilde{\psi}_u \text{ are zero outside } \mathcal{X}_0, \mathcal{X}_u \text{ respectively,} \\ -\infty, & \text{otherwise.} \end{cases}\end{aligned}$$

Now denote the extension of $\tilde{\psi}_0$, $\tilde{\psi}_u$, $\tilde{\rho}$ to the whole $\mathbb{R}^n$ by ψ_0, ψ_u, ρ, which are obtained by letting them equal to zero outside of $\mathcal{X}$. Then, for the left-hand side of

(5.4), we have the following equality:

$$\begin{aligned}
\sup_{k_1 \in \mathcal{K}_1} \langle k^*, k_1 \rangle &= \sup_{(B_0, B_u, B) \in \mathcal{B}} a(B_u - B_0) + \langle \psi_0, B_0 - B \rangle \\
&\quad + \langle \psi_u, B - B_u \rangle + \langle \rho, -\frac{\partial B}{\partial x} f \rangle \\
&= \sup_{(B_0, B_u, B) \in \mathcal{B}} (-a + \int d\psi_0) B_0 + (a - \int d\psi_u) B_u \\
&\quad + \langle -\psi_0 + \psi_u + \nabla \cdot (\rho f), B \rangle \\
&= \begin{cases} 0, & \text{if } \int_{\mathbb{R}^n} d\psi_0 = a, \int_{\mathbb{R}^n} d\psi_u = a, \text{ and} \\ & \quad -\psi_0 + \psi_u + \nabla \cdot (\rho f) = 0 \\ \infty, & \text{otherwise,} \end{cases}
\end{aligned}$$

where $\nabla \cdot (\rho f)$ is interpreted as a distributional derivative. Thus, for the supremum to be less than or equal to the infimum, we must have a nonzero $(a, \psi_0, \psi_u, \rho)$, where ψ_0, ψ_u, ρ are measures of bounded variation on $\mathbb{R}^n$, such that $a \geq 0$; ψ_0, ψ_u, ρ are nonnegative; ψ_0, ψ_u, ρ are equal to zero outside $\mathcal{X}_0$, $\mathcal{X}_u$, and $\mathcal{X}$ respectively; and

$$\begin{aligned}
\int_{\mathbb{R}^n} d\psi_0 &= a, \\
\int_{\mathbb{R}^n} d\psi_u &= a, \\
\nabla \cdot (\rho f) &= \psi_0 - \psi_u.
\end{aligned}$$

We will next show that because of the assumption that there exists a $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x) f(x) < 0$ for all $x \in \mathcal{X}$, we must have $a > 0$. For this, let $\mathcal{L} = (C(\mathcal{X}))^3$, and define

$$\begin{aligned}
\mathcal{L}_1 = \{(h_0, h_u, h) \in \mathcal{L} : h_0 = B_0 - B,\ h_u = B - B_u,\ h = -\frac{\partial B}{\partial x} f \text{ on } \mathcal{X}; \\
\text{and } (B_0, B_u, B) \in \mathcal{B}\}, \\
\mathcal{L}_2 = \{(h_0, h_u, h) \in \mathcal{L} : h_0 \geq 0 \text{ on } \mathcal{X}_0,\ h_u \geq 0 \text{ on } \mathcal{X}_u,\ h \geq 0 \text{ on } \mathcal{X}\}.
\end{aligned}$$

Note in particular that due to the above assumption and the compactness of $\mathcal{X}_0$, $\mathcal{X}_u$, $\mathcal{X}$, we have $\mathcal{L}_1 \cap \mathrm{int}(\mathcal{L}_2) \neq \emptyset$. Now consider $k^* = (a, \tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho})$ that we have before. Suppose that $a = 0$ and substitute this to (5.4). Then we have a nonzero $(\tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho}) \in (C(\mathcal{X})^*)^3$, such that

$$\sup_{\ell_1 \in \mathcal{L}_1} \langle (\tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho}), \ell_1 \rangle \leq \inf_{\ell_2 \in \mathcal{L}_2} \langle (\tilde{\psi}_0, \tilde{\psi}_u, \tilde{\rho}), \ell_2 \rangle.$$

This implies that $\mathcal{L}_1 \cap \mathrm{int}(\mathcal{L}_2) = \emptyset$, which is contradictory to the above. Thus a must be strictly positive. Without loss of generality, assume that k^* is scaled such that $a = 1$. This completes the proof of our lemma. ■

Next, we will show that the existence of ψ_0, ψ_u, ρ in the conclusion of Lemma 5.2 implies that there exists an unsafe trajectory of the system. Since in this case we have a density function ρ which is in fact a measure, we need a version of Liouville theorem which applies to measures.

Lemma 5.3 *Let $f \in C^1(D, \mathbb{R}^n)$ where $D \subseteq \mathbb{R}^n$ is open. For a measurable set Z, assume that $\phi_t(Z)$ is a subset of D for all t between 0 and T. If ρ is a measure of bounded variation on D such that ρ has a compact support and the distributional derivative $\nabla \cdot (\rho f)$ is also a measure of bounded variation with compact support, then*

$$\int_{\phi_T(Z)} d\rho - \int_Z d\rho = \int_0^T \int_{\phi_t(Z)} d(\nabla \cdot (\rho f)) dt.$$

Proof. Choose $\rho_1, \rho_2, \ldots \in C_0^\infty(D)$ such that $\rho_k \to \rho$ in the (weak) topology of distributions. Then also $\nabla \cdot (\rho_k f) \to \nabla \cdot (\rho f)$ in the sense of distributions. In particular,

$$\lim_{k \to \infty} \int_X d|\rho_k - \rho| = 0,$$
$$\lim_{k \to \infty} \int_X d|\nabla \cdot (\rho_k f) - \nabla \cdot (\rho f)| = 0$$

for every $X \subset D$. The lemma (cf. Lemma 4.1) was proven for the case of smooth ρ in [77], i.e.,

$$\int_{\phi_T(Z)} \rho_k(x)dx - \int_Z \rho_k(x)dx = \int_0^T \int_{\phi_t(Z)} [\nabla \cdot (\rho_k f)(x)]dxdt.$$

The desired equality is obtained in the limit as $k \to \infty$. ■

Lemma 5.4 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, and let $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{X}_0 \subseteq \mathcal{X}$, $\mathcal{X}_u \subseteq \mathcal{X}$ be compact sets. Suppose there exist measures of bounded variations ψ_0, ψ_u, ρ such that ψ_0, ψ_u, ρ are nonnegative on $\mathbb{R}^n$ and equal to zero outside $\mathcal{X}_0$, $\mathcal{X}_u$, and $\mathcal{X}$ respectively; and $\int_{\mathcal{X}_0} d\psi_0 = 1$, $\int_{\mathcal{X}_u} d\psi_u = 1$, $\nabla \cdot (\rho f) = \psi_0 - \psi_u$. Then there exists a $T \geq 0$ and a trajectory $x(t)$ of the system such that*

$$\begin{aligned} x(0) &\in \mathcal{X}_0, \\ x(T) &\in \mathcal{X}_u, \\ x(t) &\in \mathcal{X} \quad \forall t \in [0, T]. \end{aligned}$$

Proof. Let $X_1, X_2, ... \subseteq \mathbb{R}^n$ be a sequence of open sets such that $\mathcal{X}_0 \subseteq X_i$ for all i and $\lim_{i\to\infty} X_i = \mathcal{X}_0$. In addition, define the measurable sets

$$Z_i = \bigcup_{x_0 \in X_i} \{x \in \mathbb{R}^n : x = \phi_t(x_0) \text{ for some } t \geq 0\}, \text{ for } i = 1, 2, ...$$

By the assertions of the lemma, both ρ and $\nabla \cdot (\rho f)$ are measures with bounded variation and compact support, so we can use Lemma 5.3 and $\nabla \cdot (\rho f) = \psi_0 - \psi_u$ to obtain the relation

$$\int_{\phi_t(Z_i)} d\rho - \int_{Z_i} d\rho = \int_0^t \int_{\phi_\tau(Z_i)} d(\psi_0 - \psi_u)d\tau$$

for all $t \geq 0$. Since $\rho \geq 0$ and $\phi_t(Z_i) \subseteq Z_i$ for all $t \geq 0$, the left-hand side of the above expression is less than or equal to zero. It follows from $\int_{\mathcal{X}_0} d\psi_0 = 1$ and $\psi_0 \geq 0$ that $\mathcal{X}_u \cap Z_i \neq \emptyset$ for all $i = 1, 2, ...$, for otherwise the right-hand side of the

expression can be made strictly greater than zero by taking some $t > 0$ and we obtain a contradiction. Since the sets $\mathcal{X}_0$ and $\mathcal{X}_u$ are closed, we conclude that $\phi_T(x_0) \in \mathcal{X}_u$ for some $T \geq 0$ and $x_0 \in \mathcal{X}_0$. For our purpose, let T be the first time instance such that $\phi_T(x_0) \in \mathcal{X}_u$.

The case in which $T = 0$ is trivial since $\mathcal{X}_0 \subseteq \mathcal{X}$. Consider now the case in which $T > 0$. We will show that $\phi_t(x_0) \in \mathcal{X}$ for all $t \in [0, T]$ by a contradiction. Suppose to the contrary that there exists $\tilde{T} \in (0, T)$ such that $\phi_{\tilde{T}}(x_0) \notin \mathcal{X}$. Then, for a sufficiently small open neighborhood U of x_0, we have

$$\begin{aligned} &\phi_{\tilde{T}}(U) \subset \mathbb{R}^n \setminus (\mathcal{X}), \\ &\phi_t(U) \cap \mathcal{X}_u = \emptyset \quad \forall t \in [0, \tilde{T}]. \end{aligned}$$

Using Lemma 5.3 once again, we have

$$\int_{\phi_{\tilde{T}}(U)} d\rho - \int_U d\rho = \int_0^{\tilde{T}} \int_{\phi_\tau(U)} d(\psi_0 - \psi_u) d\tau.$$

Since $\rho = 0$ on $\mathbb{R}^n \setminus (\mathcal{X})$, the first term on the left is equal to zero, and therefore, the left-hand side is non-positive, which leads to a contradiction since the right-hand side is strictly greater than zero. This lets us conclude that $\phi_t(x_0) \in \mathcal{X}$ for all $t \in [0, T]$, thus finishing the proof of the lemma. ■

We are now ready to present the proof of the main theorem.

Proof of Theorem 5.1.

($\Rightarrow$): This is a special case of Proposition 2.2, and has been proven in Chapter 2.

($\Leftarrow$): Follows from Lemmas 5.2 and 5.4. ■

5.2 Some Remarks

The result stated in Theorem 5.1 uses the assumption that the following Slater-like condition [17] is fulfilled: that there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x) f(x) < 0$ for all $x \in \mathcal{X}$. While in the discrete case strong duality holds (and

hence the necessity of barrier certificates too) without such an assumption, its proof depends on a special property of polyhedral convex sets, which does not carry over to the continuous case. Eliminating this condition in the continuous case will presumably require a different proof technique than the one presented in this paper. Nevertheless, there are cases in which the condition is automatically fulfilled, for instance when the trajectories of the system starting from any $x_0 \in \mathcal{X}$ leave a neighborhood of $\mathcal{X}$ at least once, as shown in the following proposition.

Proposition 5.5 *Consider the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and let $\mathcal{X} \subset \mathbb{R}^n$ be a compact set. Suppose there exist an open neighborhood $\tilde{\mathcal{X}}$ of $\mathcal{X}$ and a time instant $T > 0$ such that for all initial conditions $x_0 \in \mathcal{X}$, we have the flow $\phi_t(x_0)$ outside of $\mathrm{cl}(\tilde{\mathcal{X}})$ for some $t \in [0, T]$. Then there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x) f(x) < 0$ for all $x \in \mathcal{X}$.*

Proof. Let $\mathcal{Y}$ be an open neighborhood of $\mathcal{X}$ such that its closure is contained in $\tilde{\mathcal{X}}$. In addition, let $\xi \in C^1(\mathbb{R}^n)$ be a nonnegative function such that $\xi(x) = 1$ for all $x \in \mathcal{Y}$ and $\xi(x) = 0$ for all $x \notin \tilde{\mathcal{X}}$; also let $\psi \in C^1(\mathbb{R}^n)$ be a function such that $\psi(x) > 0$ for all $x \in \mathcal{X}$ and $\psi(x) = 0$ for all $x \notin \mathcal{Y}$. Now consider the differential equation $\dot{x} = \xi(x) f(x)$. Denote the flow of $\dot{x} = \xi(x) f(x)$ starting at x_0 by $\tilde{\phi}_t(x_0)$. Modulo a time re-parameterization, the flows $\tilde{\phi}_t(x_0)$ and $\phi_t(x_0)$ are identical up to some finite time. Next define

$$\tilde{B}(x_0) = \int_0^\infty \psi(\tilde{\phi}_t(x_0)) dt.$$

For all x_0 in a neighborhood of $\mathcal{X}$, the flow $\tilde{\phi}_t(x_0)$ is outside of $\mathcal{Y}$ for large t, and thus by its construction $\psi(\tilde{\phi}_t(x_0))$ is equal to zero for large t and for all such x_0. It follows that $\tilde{B}(x)$ is well defined on a neighborhood of $\mathcal{X}$. The function $\tilde{B}(x)$ is continuously differentiable on $\mathcal{X}$ since both $\psi(x)$ and $\tilde{\phi}_t(x)$ are also continuously differentiable. Taking the total derivative of $\tilde{B}(x)$ with respect to time, we obtain

$$\frac{\partial \tilde{B}}{\partial x}(x) \xi(x) f(x) = -\psi(x),$$

which is strictly less than zero, on $\mathcal{X}$. Finally, recall that on $\mathcal{X}$ we have $\xi(x) = 1$. This completes the proof of the proposition. ■

While the above Slater-like condition excludes the possibility of applying Theorem 5.1 when there is, e.g., an equilibrium point in $\mathcal{X}$, analysis can still be performed by excluding a neighborhood of the equilibrium point from $\mathcal{X}$ in the condition (4.4). If the excluded region is either backward or forward invariant, and does not intersect $\mathcal{X}_0$ and $\mathcal{X}_u$, then the safety criterion (5.1)–(5.3) will still apply in terms of the original sets.

Finally, note also that when *all* the connected components of $\mathbb{R}^n \setminus \mathcal{X}$ are either forward or backward invariant, an even stronger safety criterion can be obtained, as in the following proposition.

Proposition 5.6 *Let the system $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and the compact sets $\mathcal{X}_0 \subset \mathbb{R}^n$, $\mathcal{X}_u \subset \mathbb{R}^n$ be given, with $0 \notin \mathcal{X}_0 \cup \mathcal{X}_u$. Suppose that the origin is a globally asymptotically stable equilibrium of the system with a global strict Lyapunov function $V(x)$ [1]. Let $\epsilon_1 = \min_{x \in \mathcal{X}_0 \cup \mathcal{X}_u} V(x)$ and $\epsilon_2 = \max_{x \in \mathcal{X}_0 \cup \mathcal{X}_u} V(x)$. Then there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying*

$$B(x) \leq 0 \quad \forall x \in \mathcal{X}_0, \tag{5.5}$$

$$B(x) > 0 \quad \forall x \in \mathcal{X}_u, \tag{5.6}$$

$$\frac{\partial B}{\partial x}(x) f(x) \leq 0 \quad \forall x \in \{x \in \mathbb{R}^n : \epsilon_1 \leq V(x) \leq \epsilon_2\}, \tag{5.7}$$

if and only if there exists no trajectory $x(t)$ of the system such that

$$x(0) \in \mathcal{X}_0, \tag{5.8}$$

$$x(T) \in \mathcal{X}_u \text{ for some } T \geq 0. \tag{5.9}$$

Proof. Define

$$\mathcal{X} = \{x \in \mathbb{R}^n : \epsilon_1 \leq V(x) \leq \epsilon_2\}.$$

[1] That is, $V \in C^1(\mathbb{R}^n)$ is radially unbounded, $V(x) > 0 \ \forall x \neq 0$, and $\frac{\partial V}{\partial x}(x) f(x) < 0 \ \forall x \neq 0$.

In this case, the existence of a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $\frac{\partial \tilde{B}}{\partial x}(x)f(x) < 0$ for all $x \in \mathcal{X}$ is guaranteed by Proposition 5.5, and even the Lyapunov function $V(x)$ can be used as $\tilde{B}(x)$. By Theorem 5.1, there exists a function $B \in C^1(\mathbb{R}^n)$ satisfying (5.5)–(5.7) if and only if there exists no trajectory $x(t)$ of the system such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$.

Since the connected components of $\mathbb{R}^n \backslash \mathcal{X}$ are either forward or backward invariant, however, there can be no trajectory $x(t)$ of the system and time instants T_1, T_2, T_3 such that $T_1 < T_2 < T_3$, $x(T_1) \in \mathcal{X}$, $x(T_2) \in \mathbb{R}^n \setminus \mathcal{X}$, and $x(T_3) \in \mathcal{X}$. This, combined with the fact that $\mathcal{X}_0, \mathcal{X}_u \subseteq \mathcal{X}$, implies that the set of trajectories satisfying $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ $\forall t \in [0, T]$ is the same as the set of trajectories satisfying (5.8)–(5.9), and therefore the statement of the proposition follows. ■

Chapter 6

Conclusions

We have presented in the previous chapters a methodology based on barrier certificates and density functions for verifying system properties such as safety, reachability, eventuality, and their combinations. Within our framework, such properties can be verified without explicitly computing the set of reachable states. This makes the methodology directly applicable to continuous and hybrid systems with nonlinear, uncertain, and constrained dynamics. In addition, by using barrier certificates that generate nonnegative supermartingales under the given system dynamics, we are able to handle safety verification of stochastic continuous and hybrid systems by computing certified upper bounds on the probability of reaching the unsafe region.

Most of the conditions satisfied by barrier certificates and density functions form convex programming problems. When the system is described in terms of polynomials, this provides the possibility to search for appropriate barrier certificates and density functions using a convex relaxation framework called sum of squares optimization. Moreover, a hierarchical search based on bounding the degrees of the polynomial expressions can be performed, such that at each level the complexity grows polynomially with respect to the system size. Some examples have been presented to illustrate the use of the proposed methods. In addition, the convexity of the problem has been exploited to derive a converse theorem for safety verification using barrier certificates.

The framework presented in this thesis opens several future research avenues, some of which we will attempt to outline here.

While the duality between safety and reachability verification is now understood

for deterministic continuous systems, much remains to be discovered when hybrid dynamics, uncertainty, and stochasticity are present in the system. In particular, we expect that the density-based approach can also be extended to handle systems with hybrid dynamics or time-varying uncertainty or both. For stochastic systems, developing a method for proving that the reach probability is higher than some margin and discovering schemes to obtain tighter probability bounds (other than by increasing the degree of the barrier certificate) are just a few directions worth pursuing.

We have shown that a combination of properties such as safety and reachability/eventuality can be verified in our framework. Related to this, it would be of interest to consider verification of more general temporal properties. It seems likely that our approach can be extended for this purpose. One possible research direction would be to develop temporal logics for continuous and hybrid systems utilizing barrier certificates and density functions as certificates of formulas.

While our results are stated for the general case, we believe that it is also beneficial to consider special problem classes, e.g., systems with linear continuous dynamics. Various questions can be asked, such as under what conditions it will be enough to consider barrier certificates of low degrees; whether it is possible to obtain a convex reformulation for the non-convex conditions if we consider restriction to special problem classes; and whether the structure of the problem can be exploited for more efficient numerical computation.

In a slightly different vein, it would be interesting to investigate the synthesis problem, i.e., to design a controller for control objectives expressed in terms of safety, reachability, and eventuality. Synthesis conditions are typically non-convex, which makes the synthesis problem harder than analysis. Preliminary results on synthesis of safe controllers using density functions can be found in [79]. Results on synthesis of controllers for stabilization of nonlinear systems based on density functions [74] and Lyapunov functions [73] might also be relevant here. Additionally, it may be interesting to consider alternative computational methods, such as randomized algorithms, to see if they can be used on their own or combined with the sum of squares optimization to solve the non-convex synthesis conditions.

Bibliography

[1] R. Alur, C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P.-H. Ho, X. Nicollin, A. Oliviero, J. Sifakis, and S. Yovine. The algorithmic analysis of hybrid systems. *Theoretical Computer Science*, 138:3–34, 1995.

[2] R. Alur, T. Dang, and F. Ivancic. Reachability analysis of hybrid systems via predicate abstraction. In *Hybrid Systems: Computation and Control, LNCS 2289*, pages 35–48. Springer-Verlag, Heidelberg, 2002.

[3] R. Alur, T. Dang, and F. Ivancic. Progress on reachability analysis of hybrid systems using predicate abstraction. In *Hybrid Systems: Computation and Control, LNCS 2623*, pages 4–19. Springer-Verlag, Heidelberg, 2003.

[4] R. Alur, T. Henzinger, G. Lafferriere, and G. J. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88(2):971–984, 2000.

[5] R. Alur and G. J. Pappas, editors. *Hybrid Systems: Computation and Control*, volume 2993 of *Lecture Notes in Computer Science*. Springer-Verlag, Heidelberg, 2004.

[6] H. Anai and V. Weispfenning. Reach set computations using real quantifier elimination. In *Hybrid Systems: Computation and Control, LNCS 2034*, pages 63–76. Springer-Verlag, Heidelberg, 2001.

[7] P. J. Antsaklis. *Proceedings of the IEEE*, 88(7), 2000. Special issue on Hybrid Systems: Theory and Applications.

[8] P. J. Antsaklis and A. Nerode. *IEEE Transactions on Automatic Control*, 43(4), 1998. Special issue on Hybrid Control Systems.

[9] E. Asarin, T. Dang, and A. Girard. Reachability analysis of nonlinear systems using conservative approximation. In *Hybrid Systems: Computation and Control, LNCS 2623*, pages 20–35. Springer-Verlag, Heidelberg, 2003.

[10] E. Asarin, T. Dang, and O. Maler. The d/dt tool for verification of hybrid systems. In *Computer Aided Verification, LNCS 2404*, pages 365–370. Springer-Verlag, Heidelber, 2002.

[11] J.-P. Aubin. *Viability Theory*. Birkhäuser, Boston, MA, 1991.

[12] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Birkhäuser, Boston, MA, 1990.

[13] A. Bemporad, F. D. Torrisi, and M. Morari. Optimization-based verification and stability characterization of piecewise affine and hybrid systems. In *Hybrid Systems: Computation and Control, LNCS 1790*, pages 45–58. Springer-Verlag, Heidelberg, 2000.

[14] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[15] J. Bochnak, M. Coste, and M.-F. Roy. *Real Algebraic Geometry*. Springer-Verlag, Berlin, 1998.

[16] O. Botchkarev and S. Tripakis. Verification of hybrid systems with linear differential inclusions using ellipsoidal approximations. In *Hybrid Systems: Computation and Control, LNCS 1790*, pages 73–88. Springer-Verlag, Heidelberg, 2000.

[17] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[18] R. E. Bryant. Graph-based algorithms for Boolean function manipulation. *IEEE Transactions on Computers*, 35(8):677–691, 1986.

[19] M. L. Bujorianu. Extended stochastic hybrid systems and their reachability problem. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 234–249. Springer-Verlag, Heidelberg, 2004.

[20] M. L. Bujorianu and J. Lygeros. Reachability questions in piecewise deterministic Markov processes. In *Hybrid Systems: Computation and Control, LNCS 2623*, pages 126–140. Springer-Verlag, Heidelberg, 2003.

[21] M. D. Choi, T. Y. Lam, and B. Reznick. Sum of squares of real polynomials. *Proceedings of Symposia in Pure Mathematics*, 58(2):103–126, 1995.

[22] A. Chutinan and B. H. Krogh. Computational techniques for hybrid system verification. *IEEE Transactions on Automatic Control*, 48(1):64–75, 2003.

[23] E. M. Clarke, Jr., O. Grumberg, and D. A. Peled. *Model Checking*. MIT Press, Cambridge, MA, 2000.

[24] M. H. A. Davis. *Markov Processes and Optimization*. Chapman-Hall, London, 1993.

[25] M. D. Di Benedetto and A. L. Sangiovanni-Vincentelli, editors. *Hybrid Systems: Computation and Control*, volume 1790 of *Lecture Notes in Computer Science*. Springer-Verlag, Heidelberg, 2001.

[26] G. A. Edgar and L. Sucheston. *Stopping Times and Directed Processes*. Cambridge University Press, Cambridge, 1992.

[27] M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM Journal on Control and Optimization*, 31(5):1183–1204, 1993.

[28] S. Glavaski, A. Papachristodoulou, and K. Ariyur. Safety verification of controlled advanced life support system using barrier certificates. In *Hybrid Systems: Computation and Control, LNCS 3414*, pages 306–321. Springer-Verlag, Heidelberg, 2005.

[29] W. Glover and J. Lygeros. A stochastic hybrid model for air traffic control simulation. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 372–386. Springer-Verlag, Heidelberg, 2004.

[30] T. A. Henzinger, P. W. Kopke, A. Puri, and P. Varaiya. What's decidable about hybrid automata? *Journal of Computer and System Sciences*, 57(1):94–124, 1998.

[31] J. P. Hespanha. Stochastic hybrid systems: Application to communication networks. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 387–401. Springer-Verlag, Heidelberg, 2004.

[32] J. P. Hespanha. Polynomial stochastic hybrid systems. In *Hybrid Systems: Computation and Control*. Springer-Verlag, Heidelberg, 2005. To appear.

[33] R. Horowitz and P. Varaiya. Control design of an automated highway system. *Proceedings of the IEEE*, 88(7):913–925, 2000.

[34] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In *Hybrid Systems: Computation and Control, LNCS 1790*, pages 160–173. Springer-Verlag, Heidelberg, 2000.

[35] J. Hu, M. Prandini, and S. Sastry. Probabilistic safety analysis in three dimensional aircraft flight. In *Proceedings of the IEEE Conference on Decision and Control*, 2003.

[36] M. Huth and M. Ryan. *Logic in Computer Science: Modelling and Reasoning about Systems*. Cambridge University Press, Cambridge, 2000.

[37] M. Jirstrand. Invariant sets for a class of hybrid systems. In *Proceedings of the IEEE Conference on Decision and Control*, 1998.

[38] M. Johansson and A. Rantzer. Computation of piecewise quadratic Lyapunov functions for hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):555–559, 1998.

[39] U. T. Jönsson. On reachability analysis of uncertain hybrid systems. In *Proceedings of the IEEE Conference on Decision and Control*, 2002.

[40] H. K. Khalil. *Nonlinear Systems*. Prentice-Hall, Inc., Upper Saddle River, NJ, second edition, 1996.

[41] A. Kurzhanski and P. Varaiya. Ellipsoidal techniques for reachability analysis. In *Hybrid Systems: Computation and Control, LNCS 1790*, pages 202–214. Springer-Verlag, Heidelberg, 2000.

[42] H. J. Kushner. *Stochastic Stability and Control*. Academic Press, New York, 1967.

[43] G. Lafferriere, G. J. Pappas, and S. Yovine. Symbolic reachability computations for families of linear vector fields. *Journal of Symbolic Computation*, 32(3):231–253, 2001.

[44] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, NY, 1969.

[45] O. Maler and A. Pnueli, editors. *Hybrid Systems: Computation and Control*, volume 2623 of *Lecture Notes in Computer Science*. Springer-Verlag, Heidelberg, 2003.

[46] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, New York, NY, 1992.

[47] Z. Manna and A. Pnueli. *Temporal Verification of Reactive Systems: Safety*. Springer-Verlag, New York, NY, 1995.

[48] A. S. Matveev and A. V. Savkin. *Qualitative Theory of Hybrid Dynamical Systems*. Birkhäuser, Boston, MA, 2000.

[49] A. Megretski and A. Rantzer. System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42(6):819–830, 1997.

[50] M. Mesbahi, M. G. Safonov, and G. P. Papavassilopoulos. Bilinearity and complementarity in robust control. In *Advances in Linear Matrix Inequality Methods in Control*, pages 269–292. SIAM, Philadelphia, PA, 2000.

[51] M. Morari and L. Thiele, editors. *Hybrid Systems: Computation and Control*. Springer-Verlag, Heidelberg, 2005. Forthcoming.

[52] A. S. Morse, C. C. Pantelides, S. S. Sastry, and J. M. Schumacher. *Automatica*, 35(3), 1999. Special issue on Hybrid Systems.

[53] R. M. Murray, editor. *Control in an Information Rich World: Report of the Panel on Future Directions in Control, Dynamics, and Systems*. SIAM, Philadelphia, PA, 2003. Available at `http://www.cds.caltech.edu/~murray/cdspanel`.

[54] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

[55] M. Nagumo. Über die lage der integralkurven gewöhnlicher differentialgleichungen. *Proceedings of the Physico–Mathematical Society of Japan*, 24:551–559, 1942.

[56] National Research Council. *Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers*. National Academy Press, Washington, DC, 2001.

[57] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, Berlin, 2000.

[58] A. Papachristodoulou and S. Prajna (equal contribution). On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings of the IEEE Conference on Decision and Control*, 2002. Journal version submitted to *IEEE Transactions on Automatic Control*.

[59] A. Papachristodoulou and S. Prajna (equal contribution). Analysis of non-polynomial systems using the sum of squares decomposition. In *Positive Polynomials in Control*. Springer-Verlag, 2005. To appear.

[60] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Dover Publications Inc., Mineola, NY, 1998.

[61] P. A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, Pasadena, CA, 2000.

[62] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Series B*, 96(2):293–320, 2003.

[63] S. Pettersson and B. Lennartson. Stability and robustness of hybrid systems. In *Proceedings of the IEEE Conference on Decision and Control*, 1996.

[64] G. Pola, M. L. Bujorianu, J. Lygeros, and M. D. Di Benedetto. Stochastic hybrid models: An overview. In *Proceedings IFAC Conference on Analysis and Design of Hybrid Systems*, 2003.

[65] S. Prajna. Barrier certificates for nonlinear model validation. In *Proceedings of the IEEE Conference on Decision and Control*, 2003. Journal version submitted to *Automatica*.

[66] S. Prajna and A. Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 477–492. Springer-Verlag, Heidelberg, 2004.

[67] S. Prajna, A. Jadbabaie, and G. J. Pappas. Stochastic safety verification using barrier certificates. In *Proceedings of the IEEE Conference on Decision and Control*, 2004.

[68] S. Prajna and A. Papachristodoulou. Analysis of switched and hybrid systems — Beyond piecewise quadratic methods. In *Proceedings of the American Control Conference*, 2003.

[69] S. Prajna, A. Papachristodoulou, and P. A. Parrilo. Introducing SOSTOOLS: A general purpose sum of squares programming solver. In *Proceedings of the IEEE Conference on Decision and Control*, 2002. Software available at `http://www.cds.caltech.edu/sostools` and `http://www.mit.edu/~parrilo/sostools`.

[70] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. New developments in sum of squares optimization and SOSTOOLS. In *Proceedings of the American Control Conference*, 2004.

[71] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. SOSTOOLS: Control applications and new developments. In *Proceedings of the IEEE Conference on Computer Aided Control Systems Design*, 2004.

[72] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. SOSTOOLS and its control applications. In *Positive Polynomials in Control*. Springer-Verlag, 2005. To appear.

[73] S. Prajna, A. Papachristodoulou, and F. Wu. Nonlinear control synthesis by sum of squares optimization: A Lyapunov-based approach. In *Proceedings of the Asian Control Conference*, 2004.

[74] S. Prajna, P. A. Parrilo, and A. Rantzer. Nonlinear control synthesis by convex optimization. *IEEE Transactions on Automatic Control*, 49(2):310–314, 2004.

[75] S. Prajna and A. Rantzer. On the necessity of barrier certificates. In *Proceedings of the IFAC World Congress*, 2005. To appear.

[76] S. Prajna and A. Rantzer. Primal-dual tests for safety and reachability. In *Hybrid Systems: Computation and Control, LNCS 3414*, pages 542–556. Springer-Verlag, 2005.

[77] A. Rantzer. A dual to Lyapunov's stability theorem. *Systems and Control Letters*, 42(3):161–168, 2001.

[78] A. Rantzer and S. Hedlund. Duality between cost and density in optimal control. In *Proceedings of the IEEE Conference on Decision and Control*, 2003.

[79] A. Rantzer and S. Prajna. On analysis and synthesis of safe control laws. In *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2004.

[80] B. Reznick. Some concrete aspects of Hilbert's 17th Problem. In *Real Algebraic Geometry and Ordered Structures*, pages 251–272. American Mathematical Society, Providence, RI, 2000.

[81] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales. Volume 1: Foundations*. Cambridge University Press, Cambridge, 2000.

[82] J. J. M. M. Rutten, M. Kwiatkowska, G. Norman, and D. Parker. *Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems*. American Mathematical Society, Providence, RI, 2004.

[83] S. Sankaranarayanan, H. Sipma, and Z. Manna. Constructing invariants for hybrid systems. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 539–554. Springer-Verlag, Heidelberg, 2004.

[84] G. Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207:87–97, 1974.

[85] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999. Software available at `http://fewcal.kub.nl/sturm/software/sedumi.html`.

[86] A. Tiwari. Approximate reachability for linear systems. In *Hybrid Systems: Computation and Control, LNCS 2623*, pages 514–525. Springer-Verlag, Heidelberg, 2003.

[87] A. Tiwari and G. Khanna. Series of abstractions for hybrid automata. In *Hybrid Systems: Computation and Control, LNCS 2289*, pages 465–478. Springer-Verlag, Heidelberg, 2002.

[88] A. Tiwari and G. Khanna. Nonlinear systems: Approximating reach sets. In *Hybrid Systems: Computation and Control, LNCS 2993*, pages 600–614. Springer-Verlag, Heidelberg, 2004.

[89] C. Tomlin and M. R. Greenstreet, editors. *Hybrid Systems: Computation and Control*, volume 2289 of *Lecture Notes in Computer Science*. Springer-Verlag, Heidelberg, 2002.

[90] C. Tomlin, I. Mitchell, and R. Ghosh. Safety verification of conflict resolution maneuvers. *IEEE Transactions on Intelligent Transportation Systems*, 2(2):110–120, 2001.

[91] C. J. Tomlin, I. Mitchell, A. M. Bayen, and M. Oishi. Computational techniques for the verification of hybrid systems. *Proceedings of the IEEE*, 91(7):986–1001, 2003.

[92] A. van der Schaft and H. Schumacher. *An Introduction to Hybrid Dynamical Systems*. Springer-Verlag, London, 2000.

[93] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[94] O. Watkins and J. Lygeros. Stochastic reachability for discrete time systems: An application to aircraft collision avoidance. In *Proceedings IEEE Conference on Decision and Control*, 2003.

[95] V. A. Yakubovich. S-procedure in nonlinear control theory. *Vestnik Leningrad University*, 4(1):73–93, 1977. English translation.

[96] H. Yazarel and G. Pappas. Geometric programming relaxations for linear systems reachability. In *Proceedings of the American Control Conference*, 2004.

[97] H. Yazarel, S. Prajna, and G. Pappas. SOS for safety. In *Proceedings of the IEEE Conference on Decision and Control*, 2004.

[98] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1996.