Primal–Dual Tests for Safety and Reachability

Stephen Prajna[1] and Anders Rantzer[2]

[1] Control and Dynamical Systems,
California Institute of Technology,
Pasadena, CA 91125 – USA
prajna@cds.caltech.edu

[2] Department of Automatic Control,
Lund Institute of Technology,
SE 221 00 Lund – Sweden
rantzer@control.lth.se

Abstract. A methodology for safety verification using barrier certificates has been proposed recently. Conditions that must be satisfied by a barrier certificate can be formulated as a convex program, and the feasibility of the program implies system safety, in the sense that there is no trajectory starting from a given set of initial states that reaches a given unsafe region. The dual of this problem, i.e., the reachability problem, concerns proving the existence of a trajectory starting from the initial set that reaches another given set. Using insights from convex duality and the concept of density functions, in this paper we show that reachability can also be verified through convex programming. Several convex programs for verifying safety, reachability, and other properties such as eventuality are formulated. Some examples are provided to illustrate their applications.

1 Introduction

Safety verification or reachability analysis addresses the question whether an unsafe region in the state space is reachable by some system trajectories starting from a set of initial states. The need for safety verification arises as the complexity of the system increases, and is also underscored by the safety critical nature of the system. This is particularly important for modern engineering systems, many of which have hybrid (i.e., a mixture of discrete and continuous) dynamics.

Various methods have been proposed for safety verification. For verification of discrete (finite state) systems, model checking techniques [1] have been very successful and have garnered a popularity that prompts the development of analogous approaches for verification of continuous systems, which mostly require computing the propagation of initial states (see e.g. [2, 3, 4, 5, 6, 7, 8, 9, 10, 11]). Unfortunately, while these techniques allow us to compute an exact or near exact approximation of reachable sets, it is difficult to perform such a computation due to the infinite number of states. Not only that, the complexity is worse when the system is nonlinear and uncertain.

M. Morari and L. Thiele (Eds.): HSCC 2005, LNCS 3414, pp. 542–556, 2005.

Recently, we proposed a method for safety verification that does not require propagating the initial set, based on what we term barrier certificates [12]. Our conditions for safety can be stated as follows. Given a system $\dot{x} = f(x)$ with the state x taking its value in $\mathcal{X}$, a set of initial states $\mathcal{X}_0 \subset \mathcal{X}$, and an unsafe set $\mathcal{X}_u \subset \mathcal{X}$, suppose there exists a continuously differentiable function $B : \mathcal{X} \to \mathbb{R}$ such that the inequalities

$$B(x) \leq 0 \qquad \forall x \in \mathcal{X}_0, \tag{1}$$

$$B(x) > 0 \qquad \forall x \in \mathcal{X}_u, \tag{2}$$

$$\frac{\partial B}{\partial x} f(x) \leq 0 \qquad \forall x \in \mathcal{X}. \tag{3}$$

are satisfied. Then the *safety* of the system is verified, namely, there is 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 $x(t) \in \mathcal{X}$ for all $t \in [0, T]$. In this case, the function $B(x)$ is called a barrier certificate.

The above method is analogous to the Lyapunov method for stability analysis [13], and is closely related to the viability theory [14] and invariant sets [15] approaches to safety verification. We would like to note that ideas parallel to ours also appear in [16, 17]. When the vector field $f(x)$ is polynomial and the sets $\mathcal{X}$, $\mathcal{X}_0$, $\mathcal{X}_u$ are semialgebraic, a polynomial barrier certificate $B(x)$ can be searched using sum of squares techniques [18, 19] in conjunction with semidefinite programming [20]. The method can also be extended to handle hybrid, uncertain, and stochastic systems [12, 21] and successful application to a hybrid system with 6 locations and 10 continuous state variables has been reported [22].

For hybrid systems, safety verification can also be performed by first constructing a discrete abstraction of the system [23, 2, 6, 7, 8, 9] and then performing verification on the resulting abstraction. This approach provides a nice hierarchical way 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 the abstracting process is still the continuous reachability analysis, which is required to determine whether or not a transition between two discrete states in the abstraction is possible.

In constructing discrete abstractions of hybrid systems, barrier-certificate-based analysis can be used for ruling out transitions between discrete states. What is still missing is a method for proving that other transitions are indeed possible. This is the problem of *reachability*, which for a system $\dot{x} = f(x)$, the state set $\mathcal{X}$, the initial set $\mathcal{X}_0 \subset \mathcal{X}$, and the target set $\mathcal{X}_r \subset \mathcal{X}$, amounts to proving that there exists a trajectory $x(t)$ of the system such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_r$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$. It is important to note that failure in computing a barrier certificate that proves the unreachability of the target set from the initial set does not by itself mean that the target set is reachable from the initial set. For example, when using polynomial parameterization for $B(x)$, it may be the case that we fail to find $B(x)$ because the degree of the polynomial is not high enough.

In the present paper, we use the ideas of duality and density functions [24, 25] to formulate a "dual" test for reachability, thus forming a primal and dual pair of safety and reachability tests. We show that reachability can be verified through convex optimization, e.g., sum of squares technique and semidefinite programming when the vector field is polynomial and the sets are semialgebraic. In addition, 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 abstractions of hybrid systems. We will also show that this convex programming approach can be used to prove properties such as eventuality or weak eventuality — whose definitions will be presented later, or even other simple combinations of reachability/eventuality and safety.

The outline of the paper is as follows. In Section 2, we give an intuitive illustration of the duality idea by addressing the verification of a simple discrete system. The main results of the paper are presented and proven in Section 3. In Section 4, some examples will be presented to illustrate the applications of the tests. Finally, some conclusions will be given in Section 5.

2 A Discrete Verification Example

To give an intuitive flavor of the duality ideas used in this paper, let us consider the verification of a simple discrete system, shown in Figure 1. The system has four states, labelled 1 through 4, and three transitions between states, represented by the directed edges in the graph. We assume that node 1 is the initial state and node 4 is the unsafe state.

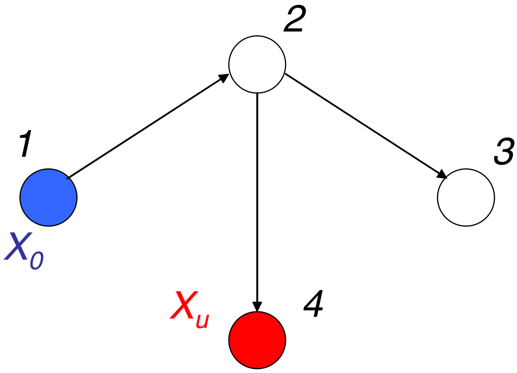


Fig. 1. A simple discrete system. The nodes represent the states of the system, while the directed edges represent transitions between states

For this system, conditions analogous to (1)–(3) that must be satisfied by a barrier certificate can be formulated. One way to find a barrier certificate is by solving the linear program

$$\begin{aligned} &\max\ B_4 - B_1 \\ &\text{subject to } B_2 - B_1 \leq 0, \\ &\qquad\qquad\quad B_3 - B_2 \leq 0, \\ &\qquad\qquad\quad B_4 - B_2 \leq 0, \end{aligned}$$

where the decision variables B_1, B_2, B_3, B_4 take values in the reals. This formulation is similar to the continuous case: analogous to (3), we ask that $B_i \leq B_j$ if there is a directed edge from node i to node j, whereas the objective function in this case is the difference between the values of B at the unsafe state and at the initial state. If there is a feasible solution of the above problem such that the objective function is strictly positive, then the value of B at the unsafe state is strictly greater than that at the initial state, i.e., there exists a barrier certificate for the system, and consequently we prove that there is no path going from node 1 to node 4.

The dual of the above linear program is as follows:

$$\begin{aligned} &\min\ 0 \\ &\text{subject to } \rho_{12} \geq 0, \rho_{23} \geq 0, \rho_{24} \geq 0, \\ &\qquad\qquad\quad \rho_{12} = 1, \\ &\qquad\qquad\quad \rho_{24} + \rho_{23} - \rho_{12} = 0, \\ &\qquad\qquad\quad \rho_{24} = 1, \\ &\qquad\qquad\quad \rho_{23} = 0. \end{aligned}$$

The dual decision variable ρ_{ij} 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 – the total flow into a node is equal to the total flow out. In addition, the first and third equality constraints indicate that there exist a unit source at node 1, i.e., the initial state, and a unit sink at node 4, i.e., the unsafe state. This duality interpretation has been studied extensively in the past, see e.g. [26] and references therein.

The existence of a feasible solution to the dual linear program 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, showing that the dual linear program is feasible can be used for verifying reachability. As a matter of fact, if we also add the objective function $\sum \rho_{ij}$ to the dual linear program, we obtain a linear programming formulation of the shortest path problem. In this case, the nonzero entries corresponding to any optimal vertex solution of the linear program will indicate a shortest path from the initial node to the unsafe node.

The duality argument above 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 linear program being equal to zero. This objective value is attained by e.g., $B_i = 0$ for all i. The linear programming duality [20] implies that there exists a feasible solution to the dual linear program, 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 the continuous case, a converse theorem for barrier certificates is proven in [27].

For the above example, the optimal objective value of the primal linear program is equal to zero. The unique feasible solution to the dual linear program is given by $\rho_{12} = 1$, $\rho_{23} = 0$, $\rho_{24} = 1$, which shows the path from node 1 to node 4. Had the direction of the edge from node 2 to node 4 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.

3 Main Results

We denote the space of m-times continuously differentiable functions mapping $X \subseteq \mathbb{R}^n$ to $\mathbb{R}^p$ by $C^m(X, \mathbb{R}^p)$. The solution $x(t)$ of $\dot{x} = f(x)$ starting from $x(0) = x_0$ is denoted by $\phi_t(x_0)$. For a set Z, we define $\phi_t(Z) = \{\phi_t(x) : x \in Z\}$. The divergence of a vector field $f \in C^1(X, \mathbb{R}^n)$ is denoted by $\nabla \cdot f(x)$. Finally, let $\text{cl}(X)$ denote the closure of a set X, and ∂X denote the boundary of X.

The following version of Liouville's theorem (from [24]) will be used in the proofs of the main theorems.

Lemma 1. *Let $f \in C^1(D, \mathbb{R}^n)$ where $D \subseteq \mathbb{R}^n$ is open and let $\rho \in C^1(D, \mathbb{R})$ be integrable. Consider the system $\dot{x} = f(x)$. For a measurable set Z, assume that $\phi_\tau(Z)$ is a subset of D for all τ between 0 and t. Then*

$$\int_{\phi_t(Z)} \rho(x)dx - \int_Z \rho(z)dz = \int_0^t \int_{\phi_\tau(Z)} \left[\nabla \cdot (f\rho)\right](x)dxd\tau. \tag{4}$$

At this point, we are ready to state and prove the first pair of tests for safety and reachability.

Theorem 1. *Consider the differential equation $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{X}_0, \mathcal{X}_u, \mathcal{X}_r \subset \mathcal{X}$ be bounded open sets, and suppose that there exists a function $B \in C^1(\mathbb{R}^n, \mathbb{R})$ satisfying*

$$B(x) \le 0 \qquad \forall x \in \mathcal{X}_0, \tag{5}$$

$$B(x) > 0 \qquad \forall x \in \mathcal{X}_u, \tag{6}$$

$$\frac{\partial B}{\partial x} f(x) \le 0 \qquad \forall x \in \mathcal{X}. \tag{7}$$

Then the safety property holds, i.e., 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 $T \ge 0$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

On the other hand, if there exists a function $\rho \in C^1(\mathbb{R}^n, \mathbb{R})$ *satisfying*

$$\int_{\mathcal{X}_0} \rho(x)dx > 0, \tag{8}$$

$$\rho(x) < 0 \qquad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r), \tag{9}$$

$$\nabla \cdot (\rho f)(x) > 0 \qquad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{10}$$

then the reachability property holds, i.e., there exists a trajectory $x(t)$ *of the system such that* $x(0) \in \mathcal{X}_0$*,* $x(T) \in \mathcal{X}_r$ *for some* $T \geq 0$*, and* $x(t) \in \mathcal{X}$ *for all* $t \in [0, T]$*.*

Proof. For a proof of the first statement, assume that there exists a $B(x)$ satisfying (5)–(7), while at the same time there is an initial condition $x_0 \in \mathcal{X}_0$ such that the trajectory $x(t)$ of $\dot{x} = f(x)$ 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 (7) states that the Lie derivative of $B(x)$ along this flow is non-positive. A direct consequence of this is that $B(x(T))$ must be less than or equal to $B(x(0))$, which is contradictory to (5)–(6). Thus we conclude that the system is safe.

To prove the second statement, let $X \subset \mathcal{X}_0$ be an open set on which $\rho(x) > 0$. 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 1 as follows. Since $\phi_t(Z) \subset Y$, Y is bounded, and $\rho(x)$ is continuous, the left hand side of (4) is bounded. For (4) to hold, we must have $\int_{\phi_\tau(Z)} [\nabla \cdot (f\rho)](x)dx \to 0$ 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 $\lim_{t\to\infty} \int_Z \rho(x)dx + \int_0^t \int_{\phi_\tau(Z)} [\nabla \cdot (f\rho)](x)dxd\tau$ is strictly positive. Thus Z must have zero measure. 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 whose flow leaves $\mathcal{X} \setminus \mathcal{X}_r$ in finite time, and assume 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 $\mathcal{X}$. That is, let $\phi_t(x_0) \in \mathcal{X} \setminus \mathcal{X}_r$ for all $t \in [0, T)$ and $\phi_T(x_0) \notin \mathcal{X}$. Choose a neighborhood Z of x_0 such that

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

Now apply Lemma 1 again with $t = T$ to obtain a contradiction. According to the above, the left hand side of (4) 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 [28].

Theorem 2. *Consider the differential equation $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{X}_0, \mathcal{X}_u, \mathcal{X}_r \subset \mathcal{X}$ be bounded open sets, and suppose that there exists a function $B \in C^1(\mathbb{R}^n, \mathbb{R})$ satisfying*

$$\int_{\mathcal{X}_0} B(x) dx < 0, \tag{11}$$

$$B(x) > 0 \qquad \forall x \in \partial\mathcal{X} \setminus \partial\mathcal{X}_r, \tag{12}$$

$$\frac{\partial B}{\partial x} f(x) < 0 \qquad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r). \tag{13}$$

Then the reachability property holds, i.e., there exists a trajectory $x(t)$ of the system such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_r$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

On the other hand, if there exists a function $\rho \in C^1(\mathbb{R}^n, \mathbb{R})$ satisfying

$$\rho(x) \geq 0 \qquad \forall x \in \mathcal{X}_0, \tag{14}$$

$$\rho(x) < 0 \qquad \forall x \in \mathcal{X}_u, \tag{15}$$

$$\nabla \cdot (\rho f)(x) \geq 0 \qquad \forall x \in \mathcal{X}, \tag{16}$$

then the safety property holds, i.e., 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 $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

Proof. To prove the first statement, consider a point $x_0 \in \mathcal{X}_0$ such that $B(x_0) < 0$. The flow $\phi_t(x_0)$ must leave $\mathcal{X} \setminus \mathcal{X}_r$ in finite time, since the Lie derivative inequality (13) holds and $B(x)$ is bounded below on $\mathcal{X}$. Now assume 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. From (13), it follows that $B(\phi_T(x_0))$ is strictly less than zero, which is contradictory to (12). 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 the second statement. 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 $t \in [0, T]$. Let $Z \subset \mathcal{X}_0$ be a ball surrounding x_0 such that also $\phi_T(Z) \subset \mathcal{X}_u$ and $\phi_t(Z) \subset \mathcal{X}$ for $t \in [0, T]$. Now apply Lemma 1 with $t = T$ to obtain a contradiction. According to the assumptions of the theorem, the left hand side of (4) is negative and the right hand side is non-negative. Hence there is a contradiction and the proof is complete.

Remark 1. Modulo the following modifications on the assertions of the theorems, the conclusions will still hold even when the sets are not bounded. In particular, for the second part of Theorem 1, we need to add the condition that $\rho(x)$ is integrable on $\mathcal{X}$ and replace (10) by

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

for a positive number ϵ. In the first part of Theorem 2, we need to add the condition that $B(x)$ is bounded below on $\mathcal{X}$ and replace (13) by

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

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 (16) 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 (13) 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 (i.e., described by polynomial equalities and inequalities), a computational method called sum of squares optimization is available if we use polynomial parameterizations for $B(x)$ or $\rho(x)$. The method is based on the sum of squares decomposition of multivariate polynomials [18] and semidefinite programming [20]. Software tools [19] are helpful for this purpose. See [12] for details.

Remark 2. Strictly speaking, it should be noted that the tests in the above theorems are not pairs of *Lagrange dual* problems [20] 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 1 will require $\nabla \cdot (\rho f)(x)$ to be zero on $\mathcal{X} \setminus (\mathcal{X}_0 \cup \mathcal{X}_u)$ (see [27]). 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.

In the reachability test of Theorem 2, the set of states $\{x \in \mathcal{X}_0 : B(x) < 0\}$ is said to satisfy the *eventuality*[1] property: all trajectories starting from this set will eventually reach $\mathcal{X}_r$ in a finite time. Analogously, in Theorem 1, the set of states $\{x \in \mathcal{X}_0 : \rho(x) > 0\}$ is said to satisfy the *weak eventuality* property: almost all trajectories starting from this set will eventually reach $\mathcal{X}_r$ in a finite time. 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 or weak eventuality tests for the whole initial set $\mathcal{X}_0$ can be performed simply by replacing (11) and (8) by $B(x) < 0\ \forall x \in \mathcal{X}_0$ and $\rho(x) > 0\ \forall x \in \mathcal{X}_0$, respectively.

Example 1. 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 $x(t) = 0$.

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 1 and 2 is twofold. First, a solution to the convex programs for reachability will automatically indicate which state x_0 can be chosen as the initial state (or a set of states from which almost all points can be chosen as the initial state). Second, the use of these convex programs allows us to also consider the worst-case analysis of systems with disturbance or 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 take its value in a set D. Then solving (11)–(13) with the Lie derivative inequality replaced by

$$\frac{\partial B}{\partial x} f(x, d) < 0 \qquad\qquad \forall (x, d) \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r) \times D$$

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. On the other hand, the density function $\rho(x)$ is more appropriate for controller design, as pointed out in [24]. For a system $\dot{x} = f(x) + g(x)u$ where u is the control input, the inequalities (8)–(9) and

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

(and similarly for (14)–(16)) 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

[1] This property is also termed the *liveness* property in temporal logics [29]. We use "eventuality" to avoid possible confusion with "liveness" in the sense of viability theory.

optimization to find a feasible pair (ρ, ψ), then recover the control law as $u(x) = \psi(x)/\rho(x)$; see [28].

It is clear that the above tests 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 $t \in [0, T]$,

or the eventuality – safety (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 corollary.

Corollary 1. *Consider the differential equation $\dot{x} = f(x)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Let $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{X}_0, \mathcal{X}_u, \mathcal{X}_r \subset \mathcal{X}$ be bounded open sets, and suppose that there exists a function $B \in C^1(\mathbb{R}^n, \mathbb{R})$ satisfying*

$$B(x) < 0 \qquad \forall x \in \mathcal{X}_0, \tag{17}$$

$$B(x) > 0 \qquad \forall x \in (\partial\mathcal{X} \setminus \partial\mathcal{X}_r) \cup \mathcal{X}_u, \tag{18}$$

$$\frac{\partial B}{\partial x} f(x) < 0 \qquad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r). \tag{19}$$

Then the eventuality – safety property holds. Similarly, if there exists a function $\rho \in C^1(\mathbb{R}^n, \mathbb{R})$ satisfying

$$\rho(x) > 0 \qquad \forall x \in \mathcal{X}_0, \tag{20}$$

$$\rho(x) < 0 \qquad \forall x \in \mathrm{cl}(\partial\mathcal{X} \setminus \partial\mathcal{X}_r) \cup \mathcal{X}_u, \tag{21}$$

$$\nabla \cdot (\rho f)(x) > 0 \qquad \forall x \in \mathrm{cl}(\mathcal{X} \setminus \mathcal{X}_r), \tag{22}$$

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

4 Examples

4.1 Successive Primal–Dual Refinement

Consider the system

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

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

$$\mathcal{X}_0 = (-3.4, 3.4) \times (3.35, 3.45), \qquad \mathcal{X}_2 = (-3.5, 3.5) \times (-3.5, -3.45),$$
$$\mathcal{X}_1 = (3.45, 3.5) \times (-3.5, 3.5), \qquad \mathcal{X}_3 = (-3.5, -3.45) \times (-3.5, 3.5).$$

In this example, we will investigate the reachability of $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{X}_3$ from $\mathcal{X}_0$ (cf. Figure 3). 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.

The tests in Theorem 1 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/reachability for the whole $\mathcal{X}_0$. If neither $B(x)$ nor $\rho(x)$ can be found, we divide the interval of x_1 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.

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 2 (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 2 (b).

For proofs of the corresponding reachability and safety, see Figure 3.

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.

4.2 Proving Eventuality

For the second example, consider the four dimensional system

$$\begin{aligned} \dot{x}_1 &= x_2, & \dot{x}_3 &= x_4, \\ \dot{x}_2 &= -x_3, & \dot{x}_4 &= x_1^2 - x_4 + 2 + d, \end{aligned}$$

where the time-varying disturbance input $d(t)$ is assumed to take value in the interval $[-1, 1]$. Let the set of states and the initial set be

$$\begin{aligned} \mathcal{X} &= \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 < 7^2\}, \\ \mathcal{X}_0 &= \{x \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 < 0.1^2\}. \end{aligned}$$

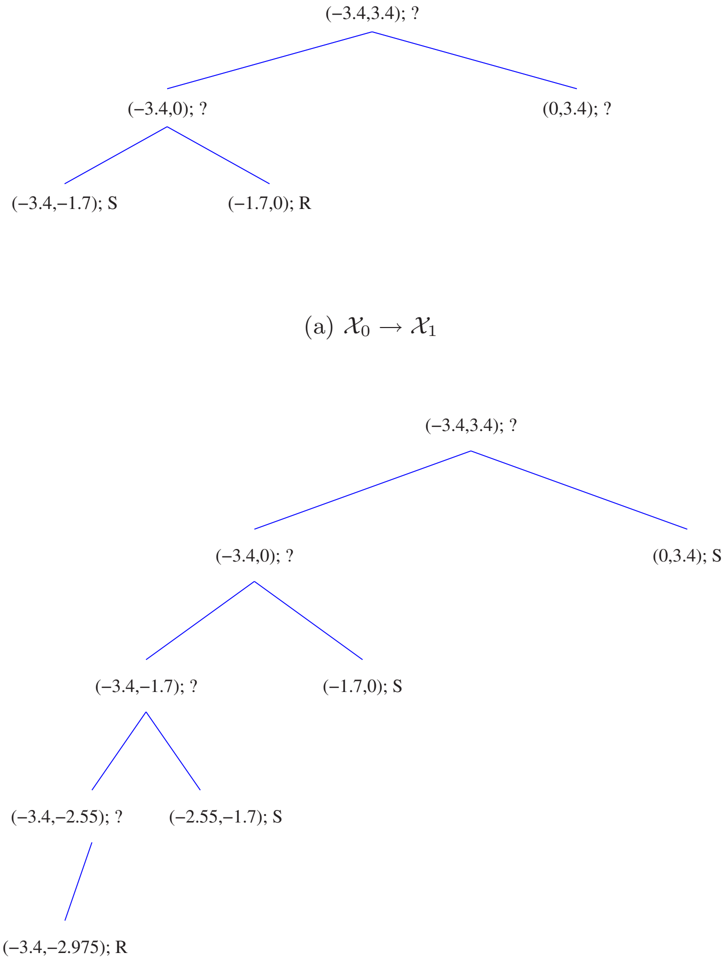
When there is no disturbance input, simulation indicates that the trajectory of the system starting from the origin reaches the set

$$\mathcal{X}_{r,1} = \{x \in \mathbb{R}^4 : (x_1 + 1)^2 + (x_2 + 1.75)^2 + (x_3 - 2.25)^2 + (x_4 - 2)^2 < 0.2^2\}$$

in finite time. As we introduce the disturbance input and also the uncertainty in the initial condition, some trajectories of the system will no longer reach the above set. However, it is expected that these trajectories will still reach a larger ball with the same center as $\mathcal{X}_{r,1}$. Using $B(x)$ of degree 4, it can be verified that the set

$$\mathcal{X}_{r,2} = \{x \in \mathbb{R}^4 : (x_1 + 1)^2 + (x_2 + 1.75)^2 + (x_3 - 2.25)^2 + (x_4 - 2)^2 < 3^2\}$$

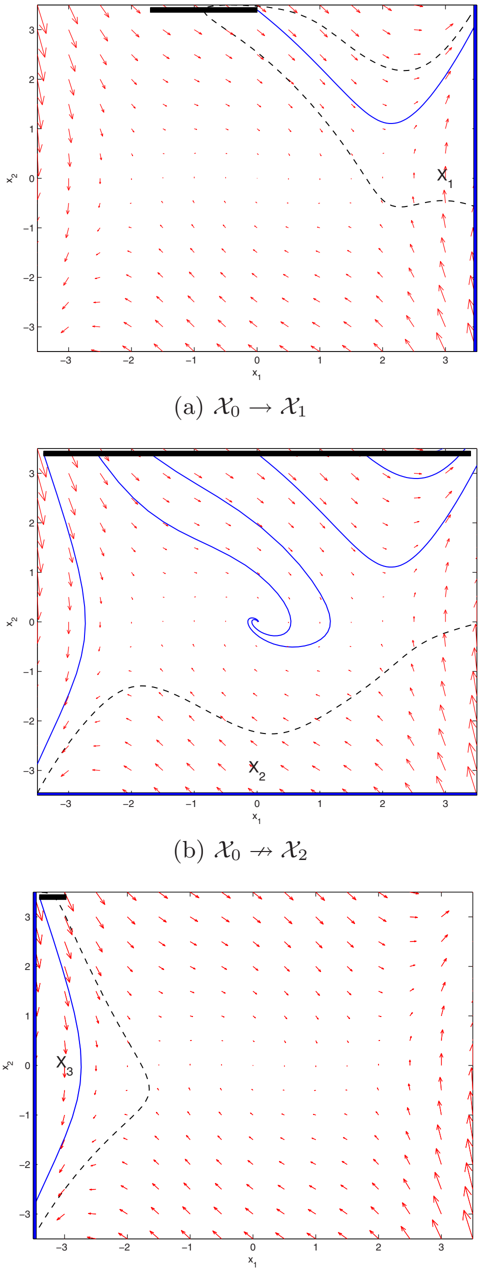
is reached in finite time by all trajectories of the system starting from $\mathcal{X}_0$.



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

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

Fig. 2. Proving the reachability of $\mathcal{X}_1$ and $\mathcal{X}_3$ from $\mathcal{X}_0$ in the example of Section 4.1. At each node we indicate the range of x_1 on $\mathcal{X}_0$ 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. When the degree of $B(x)$ or $\rho(x)$ is chosen equal to 8, the semidefinite program for each safety or reachability test at any node can be solved in less than 4 seconds on a Pentium III 600 MHz laptop. 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) $\mathcal{X}_0 \rightarrow \mathcal{X}_3$

Fig. 3. Possible transitions from $\mathcal{X}_0$ to $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{X}_3$ for the example in Section 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

5 Conclusions

In the previous sections, we use the insight from convex duality and the concept of density functions to formulate a test for reachability, which together with safety analysis using barrier certificates form a pair of convex programs for safety and reachability tests. We have additionally presented another pair of safety and reachability tests, also in the form of convex programs. This opens the possibility to perform these tests using convex optimization. In particular, sum of squares optimization can be used for this purpose when the vector field of the system is polynomial and the sets are semialgebraic.

We have further commented on the use of this methodology for worst-case reachability analysis or controller synthesis. It is pointed out that similar tests can be derived for proving eventuality or weak eventuality, and the tests can be combined to verify properties such as reachability–safety and eventuality–safety. Some examples have been presented for illustration. While the present tests are aimed for continuous reachability or safety analysis and hence are useful for constructing abstractions of hybrid systems, we expect that all of them can also be extended to handle hybrid systems directly, using an approach similar to the one presented in [12].

References

1. Clarke, Jr., E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press, Cambridge, MA (2000)
2. Bemporad, A., Torrisi, F.D., Morari, M.: Optimization-based verification and stability characterization of piecewise affine and hybrid systems. In: Hybrid Systems: Computation and Control, LNCS 1790. Springer-Verlag (2000) 45–58
3. Kurzhanski, A., Varaiya, P.: Ellipsoidal techniques for reachability analysis. In: Hybrid Systems: Computation and Control, LNCS 1790. Springer-Verlag, Heidelberg (2000) 203–213
4. Lafferriere, G., Pappas, G.J., Yovine, S.: Symbolic reachability computations for families of linear vector fields. Journal of Symbolic Computation **32** (2001) 231–253
5. Anai, H., Weispfenning, V.: Reach set computations using real quantifier elimination. In: Hybrid Systems: Computation and Control, LNCS 2034. Springer-Verlag (2001) 63–76
6. Asarin, E., Dang, T., Maler, O.: The d/dt tool for verification of hybrid systems. In: Computer Aided Verification, LNCS 2404. Springer-Verlag (2002) 365–370
7. Alur, R., Dang, T., Ivancic, F.: Progress on reachability analysis of hybrid systems using predicate abstraction. In: Hybrid Systems: Computation and Control, LNCS 2623. Springer-Verlag, Heidelberg (2003) 4–19
8. Tomlin, C.J., Mitchell, I., Bayen, A.M., Oishi, M.: Computational techniques for the verification of hybrid systems. Proceedings of the IEEE **91** (2003) 986–1001
9. Chutinan, A., Krogh, B.H.: Computational techniques for hybrid system verification. IEEE Transactions on Automatic Control **48** (2003) 64–75
10. Tiwari, A.: Approximate reachability for linear systems. In: Hybrid Systems: Computation and Control, LNCS 2623. Springer-Verlag (2003) 514–525

11. Yazarel, H., Pappas, G.: Geometric programming relaxations for linear systems reachability. In: Proceedings of the American Control Conference. (2004)
12. Prajna, S., Jadbabaie, A.: Safety verification of hybrid systems using barrier certificates. In: Hybrid Systems: Computation and Control, LNCS 2993. Springer-Verlag, Heidelberg (2004) 477–492
13. Khalil, H.K.: Nonlinear Systems. Second edn. Prentice-Hall, Inc., Upper Saddle River, NJ (1996)
14. Aubin, J.P.: Viability Theory. Birkhäuser, Boston, MA (1991)
15. Jirstrand, M.: Invariant sets for a class of hybrid systems. In: Proceedings of the IEEE Conference on Decision and Control. (1998)
16. Sankaranarayanan, S., Sipma, H., Manna, Z.: Constructing invariants for hybrid systems. In: Hybrid Systems: Computation and Control, LNCS 2993. Springer-Verlag (2004) 539–554
17. Tiwari, A., Khanna, G.: Nonlinear systems: Approximating reach sets. In: Hybrid Systems: Computation and Control, LNCS 2993. Springer-Verlag (2004) 600–614
18. Parrilo, P.A.: Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. PhD thesis, California Institute of Technology, Pasadena, CA (2000)
19. Prajna, S., Papachristodoulou, A., Parrilo, P.A.: Introducing SOSTOOLS: A general purpose sum of squares programming solver. In: Proceedings of the IEEE Conference on Decision and Control. (2002) Available at http://www.cds.caltech.edu/sostools and http://www.aut.ee.ethz.ch/˜parrilo/sostools.
20. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
21. Prajna, S., Jadbabaie, A., Pappas, G.J.: Stochastic safety verification using barrier certificates. In: Proceedings of the IEEE Conference on Decision and Control. (2004)
22. Glavaski, S., Papachristodoulou, A., Ariyur, K.: Controlled hybrid system safety verification: Advanced life support system testbed. Submitted (2005)
23. Alur, R., Henzinger, T., Lafferriere, G., Pappas, G.J.: Discrete abstractions of hybrid systems. Proceedings of the IEEE **88** (2000) 971–984
24. Rantzer, A.: A dual to Lyapunov's stability theorem. Systems and Control Letters **42** (2001) 161–168
25. Rantzer, A., Hedlund, S.: Duality between cost and density in optimal control. In: Proceedings of the IEEE Conference on Decision and Control. (2003)
26. Papadimitriou, C.H., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover Publications Inc., Mineola, NY (1998)
27. Prajna, S., Rantzer, A.: On the necessity of barrier certificates. In: Proceedings of the IFAC World Congress. (2005) To appear.
28. Rantzer, A., Prajna, S.: On analysis and synthesis of safe control laws. In: Proceedings of the Allerton Conference on Communication, Control, and Computing. (2004)
29. Manna, Z., Pnueli, A.: The Temporal Logic of Reactive and Concurrent Systems: Specification. Springer-Verlag, New York, NY (1992)