Convex computation of the region of attraction of polynomial control systems

Didier Henrion[1,2,3], Milan Korda[4]

Draft of July 15, 2013

Abstract

We address the long-standing problem of computing the region of attraction (ROA) of a target set (typically a neighborhood of an equilibrium point) of a controlled nonlinear system with polynomial dynamics and semialgebraic state and input constraints. We show that the ROA can be computed by solving an infinite-dimensional convex linear programming (LP) problem over the space of measures. In turn, this problem can be solved approximately via a classical converging hierarchy of convex finite-dimensional linear matrix inequalities (LMIs). Our approach is genuinely primal in the sense that convexity of the problem of computing the ROA is an outcome of optimizing directly over system trajectories. The dual infinite-dimensional LP on nonnegative continuous functions (approximated by polynomial sum-of-squares) allows us to generate a hierarchy of semialgebraic outer approximations of the ROA at the price of solving a sequence of LMI problems with asymptotically vanishing conservatism. This sharply contrasts with the existing literature which follows an exclusively dual Lyapunov approach yielding either nonconvex bilinear matrix inequalities or conservative LMI conditions. The approach is simple and readily applicable as the outer approximations are the outcome of a single semidefinite program with no additional data required besides the problem description.

Keywords: Region of attraction, polynomial control systems, occupation measures, linear matrix inequalities (LMIs), convex optimization, viability theory, reachable set, capture basin.

1 Introduction

Given a nonlinear control system, a state-constraint set and a target set (e.g. a neighborhood of an attracting orbit or an equilibrium point), the constrained controlled region of attraction

---

[1]CNRS; LAAS; 7 avenue du colonel Roche, F-31400 Toulouse; France. `henrion@laas.fr`
[2]Université de Toulouse; LAAS; F-31400 Toulouse; France.
[3]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic.
[4]Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, Station 9, CH-1015, Lausanne, Switzerland. `milan.korda@epfl.ch`

(ROA) is the set of all initial states that can be steered with an admissible control to the target set without leaving the state-constraint set. The target set can be required to be reached at a given time or at any time before a given time. The problem of computing the ROA (and variations thereof) lies at the heart of viability theory (see, e.g., [4]) and goes by many other names, e.g., the reach-avoid or target-hitting problem (see, e.g., [30]); the ROA itself is sometimes referred to as the backward reachable set [32] or capture basin [4].

There are many variations on the ROA computation problem addressed in this paper as well as a large number of related problems. For instance, one could consider asymptotic convergence instead of finite-time formulation and / or inner approximations instead of outer approximations. Among the related problems we can name the computation of (forward) reachable sets and maximum / minimum (robust) controlled invariant sets. Most of these variations are amenable to our approach, sometimes with different quality of the results obtained; see the Conclusion for a discussion.

We show that, in the case of polynomial dynamics, semialgebraic state-constraint, input-constraint and target sets, the computation of the ROA boils down to solving an infinite-dimensional linear programming (LP) problem in the cone of nonnegative Borel measures. Our approach is genuinely primal in the sense that we optimize over state-space system trajectories modeled with occupation measures [27, 15].

In turn, this infinite-dimensional LP can be solved approximately by a classical hierarchy of finite-dimensional convex linear matrix inequality (LMI) relaxations. The dual infinite-dimensional LP on nonnegative continuous functions and its LMI relaxations on polynomial sum-of-squares provide explicitly an asymptotically converging sequence of nested semialgebraic outer approximations of the ROA.

Most of the existing literature on ROA computation follows Zubov's approach [31, 47, 18] and uses a dual Lyapunov certificate; see [43], the survey [16], Section 3.4 in [17], and more recently [40, 8] and [9] and the references therein. These approaches either enforce convexity with conservative LMI conditions (whose conservatism is difficult if not impossible to evaluate systematically) or they rely on nonconvex bilinear matrix inequalities (BMIs), with all their inherent numerical difficulties. In contrast, we show in this paper that the problem of computing the ROA has actually a convex infinite-dimensional LP formulation, and that this LP can be solved with a hierarchy of convex finite-dimensional LMIs with asymptotically vanishing conservatism.

We believe that our approach is closer in spirit to set-oriented approaches [10], level-set and Hamilton-Jacobi approaches [25, 30, 32] or transfer operator approaches [44], even though we do not discretize w.r.t. time and/or space. In our approach, we model a measure with a finite number of its moments, which can be interpreted as a frequency-domain discretization (by analogy with Fourier coefficients which are moments w.r.t. the unit circle).

Another way to evaluate the contribution of our paper is to compare it with the recent works [22, 20] which deal with polynomial approximations of semialgebraic sets. In these references, the sets to be approximated are given a priori (as a polynomial sublevel set, or as a feasibility region of a polynomial matrix inequality). In contrast, in the current paper the set to be approximated (namely the ROA of a nonlinear dynamical system) is not known in advance, and our contribution can be understood as an application and extension of the

techniques of references [22, 20] to sets defined implicitly by differential equations.

For the special case of linear systems, the range of computational tools and theoretical results is wider; see, e.g., [14, 6]. Nevertheless, even for this simple class of systems, the problem of ROA computation is notoriously hard, at least in a controlled setting where, after time-discretization, polyhedral projections are required [6].

The benefits of our occupation measure approach are overall the convexity of the problem of finding the ROA, and the availability of publicly available software to implement and solve the hierarchy of LMI relaxations. The use of occupation measures and related concepts has a long history in the fields of Markov decision processes and stochastic control; see, e.g., [23, 12]. Applications to deterministic control problems were, to the best of our knowledge, first systematically treated[1] in [38] and enjoyed a resurgence of interest in the last decade; see, e.g., [37, 27, 42, 15] and references therein. However, to the best of our knowledge, this the first time these methods are applied to region of attraction computation.

Our primary focus in this paper is the computation of the constrained finite-time controlled region of attraction of a given set. This problem is formally stated in Section 2 and solved using occupation measures in Section 4; the occupation measures themselves are introduced in Section 3. A dual problem on the space of continuous functions is discussed in Section 5. The hierarchy of finite-dimensional LMI relaxations of the infinite dimensional LP is described in Section 6. Convergence results are presented in Section 7. An extension to the free final time case is sketched in Section 8. Numerical examples are presented in Section 9, and we conclude in Section 10.

A reader interested only in the finite-dimensional relaxations providing the converging sequence of outer approximations can consult directly the dual infinite-dimensional LP (16), its finite-dimensional LMI approximations (22) and the resulting outer approximations in Section 7 with convergence proven in Theorem 6.

Notation and preliminaries Throughout the paper the symbol $\mathbb{Z}_{[i,j]}$ denotes the set of consecutive integers $\{i, i+1, \ldots, j\}$. The symbol $\mathbb{R}[\cdot]$ denotes the ring of polynomials in variables given by the argument. The difference of two sets A and B is denoted by $A \setminus B$. All sets are automatically assumed Borel measurable; in particular if we write "for all sets", we mean "for all Borel measurable sets". By a measure μ defined on a set $X \subset \mathbb{R}^n$, we understand a signed Borel measure, i.e., a (not necessarily nonnegative) mapping from subsets of X to real numbers satisfying $\mu(\emptyset) = 0$ and $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for every countable disjoint collection of sets $A_i \subset X$. The space of (signed) measures on a set X is denoted by $M(X)$ and the space of continuous functions is denoted by $C(X)$; the dual space of all linear functionals on $C(X)$ is denoted by $C(X)'$. The symbols $C(X; \mathbb{R}^m)$ (resp. $C^1(X; \mathbb{R}^m)$) then denote the spaces of all continuous (resp. continuously differentiable) functions taking values in $\mathbb{R}^m$. Any measure $\mu \in M(X)$ can be viewed as an element of

---

[1]J. E. Rubio in [38] used the so called Young measures [46], not the occupation measures, although the idea of "linearizing" the nonlinear problem by studying whole enables of trajectories represented by measures is the same. Similar techniques were studied around the same time by others, for instance, A. F. Filippov, J. Warga and R. V. Gamkrelidze.

$C(X)'$ via the duality pairing induced by Lebesgue integration of functions in $C(X)$ w.r.t. μ:

$$\langle \mu, g \rangle := \int_X g(x)\, d\mu(x), \ \mu \in M(X), \ g \in C(X).$$

In addition, we use the following notation:

- $I_A(\cdot)$ denotes the indicator function of a set A, i.e., a function equal to 1 on A and 0 elsewhere;

- λ denotes the Lebesgue measure on $X \subset \mathbb{R}^n$, i.e.,

$$\lambda(A) = \int_X I_A(x)\, d\lambda(x) = \int_X I_A(x)dx = \int_A dx$$

  is the standard n-dimensional volume of a set $A \subset X$;

- $\operatorname{spt} \mu$ denotes the support of a measure μ, that is, the set of all points x such that $\mu(A) > 0$ for every open neighborhood A of x; by construction this set is closed.

2 Problem statement

Consider the control system

$$\dot{x}(t) = f(t, x(t), u(t)), \ t \in [0, T], \tag{1}$$

with the state $x(t) \in \mathbb{R}^n$, the control input $u(t) \in \mathbb{R}^m$ and a terminal time $T > 0$. Each entry of the vector field f is assumed to be polynomial[2], i.e., $f_i \in \mathbb{R}[t, x, u]$, $i \in \mathbb{Z}_{[1,n]}$. The state and the control input are subject to the basic semialgebraic constraints

$$\begin{aligned}
x(t) &\in X := \{x \in \mathbb{R}^n : g_i^X(x) \geq 0, i \in \mathbb{Z}_{[1,n_X]}\}, \ t \in [0, T], \\
u(t) &\in U := \{u \in \mathbb{R}^m : g_i^U(u) \geq 0, i \in \mathbb{Z}_{[1,n_U]}\}, \ t \in [0, T],
\end{aligned} \tag{2}$$

with $g_i^X \in \mathbb{R}[x]$ and $g_i^U \in \mathbb{R}[u]$; the terminal state $x(T)$ is constrained to lie in the basic semialgebraic set

$$X_T := \{x \in \mathbb{R}^n \ : \ g_i^{X_T}(x) \geq 0, i \in \mathbb{Z}_{[1,n_T]}\} \subset X,$$

with $g_i^{X_T} \in \mathbb{R}[x]$.

Assumption 1 *The sets X, U and X_T are compact.*

A measurable control function $u : [0, T] \to \mathbb{R}^m$ is called admissible if $u(t) \in U$ for all $t \in [0, T]$. The set of all admissible control functions is denoted by $\mathcal{U}$. The set of all

---

[2]Note that the infinite-dimensional results of Sections 4 and 5 hold with any Lipschitz f; the assumption of f being polynomial and the constraint sets semialgebraic is required only for finite-dimensional relaxations of Section 6.

admissible trajectories starting from the initial condition x_0 generated by admissible control functions is then

$$\mathcal{X}(x_0) := \Big\{ x(\cdot) \ : \ \exists\, u(\cdot) \in \mathcal{U} \text{ s.t. } \dot{x}(t) = f(t, x(t), u(t)) \text{ a.e.}$$

$$x(0) = x_0,\ x(t) \in X,\ x(T) \in X_T,\ \forall\, t \in [0,T] \Big\}, \tag{3}$$

where $x(\cdot)$ is required to be absolutely continuous. Here, a.e. stands for "almost everywhere" w.r.t. the Lebesgue measure on $[0,T]$.

The constrained controlled region of attraction (ROA) is then defined as

$$X_0 := \big\{ x_0 \in X \ : \ \mathcal{X}(x_0) \neq \emptyset \big\}. \tag{4}$$

In words, the ROA is the set of all initial conditions for which there exists an admissible trajectory, i.e., the set of all initial conditions that can be steered to the target set in an admissible way. By construction the set X_0 is bounded and unique.

In the sequel we pose the problem of computing the ROA as an infinite-dimensional linear program (LP) and show how this formulation can be approximated by a sequence of linear matrix inequality (LMI) problems converging to the solution to the LP.

3 Occupation measures

In this section we introduce the concept of occupation measures and discuss its connection to trajectories of the control system (1).

3.1 Liouville's equation

Given an initial condition x_0 and an admissible trajectory $x(\cdot \,|\, x_0) \in \mathcal{X}(x_0)$ with its corresponding control $u(\cdot \,|\, x_0) \in \mathcal{U}$, which we assume to be a measurable function of x_0, we define the *occupation measure* $\mu(\cdot \,|\, x_0)$ by

$$\mu(A \times B \times C \,|\, x_0) := \int_0^T I_{A \times B \times C}(t, x(t \,|\, x_0), u(t \,|\, x_0))\, dt$$

for all $A \times B \times C \subset [0,T] \times X \times U$. The occupation measure measures the time spent by the trajectory of the state and control $(x(\cdot \,|\, x_0), u(\cdot \,|\, x_0))$ in subsets of $X \times U$. The occupation measure has the following important property: for any measurable function $g(t, x, u)$ the equality

$$\int_0^T g(t, x(t \,|\, x_0), u(t \,|\, x_0))\, dt = \int_{[0,T] \times X \times U} g(t, x, u)\, d\mu(t, x, u \,|\, x_0) \tag{5}$$

holds. Therefore, loosely speaking, the occupation measure $\mu(\cdot \,|\, x_0)$ encodes all information about the trajectory of the state and control input $(x(\cdot \,|\, x_0), u(\cdot \,|\, x_0))$.

Define further the linear operator $\mathcal{L} : C^1([0,T] \times X) \to C([0,T] \times X \times U)$ by

$$v \mapsto \mathcal{L}v := \frac{\partial v}{\partial t} + \sum_{i=1}^n \frac{\partial v}{\partial x_i} f_i(t, x, u) = \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot f$$

and its adjoint operator $\mathcal{L}' : C([0,T] \times X \times U)' \to C^1([0,T] \times X)'$ by the adjoint relation

$$\langle \mathcal{L}'\nu, v \rangle := \langle \nu, \mathcal{L}v \rangle = \int_{[0,T] \times X \times U} \mathcal{L}v(t,x,u)\, d\nu(t,x,u)$$

for all $\nu \in M([0,T] \times X \times U) = C([0,T] \times X \times U)'$ and $v \in C^1([0,T] \times X)$. Given a test function $v \in C^1([0,T] \times X)$, it follows from (5) that

$$
\begin{aligned}
v(T, x(T\,|\,x_0)) &= v(0, x_0) + \int_0^T \frac{d}{dt} v(t, x(t\,|\,x_0))\, dt \\
&= v(0, x_0) + \int_0^T \mathcal{L}v(t, x(t\,|\,x_0), u(t\,|\,x_0))\, dt \\
&= v(0, x_0) + \int_{[0,T] \times X \times U} \mathcal{L}v(t,x,u)\, d\mu(t,x,u\,|\,x_0) \\
&= v(0, x_0) + \langle \mathcal{L}'\mu(\cdot\,|\,x_0), v \rangle,
\end{aligned}
\tag{6}
$$

where we have used the adjoint relation in the last equality.

Remark 1 *The adjoint operator $\mathcal{L}'$ is sometimes expressed symbolically as*

$$\nu \mapsto \mathcal{L}'\nu = -\frac{\partial \nu}{\partial t} - \sum_{i=1}^n \frac{\partial (f_i \nu)}{\partial x_i} = -\frac{\partial \nu}{\partial t} - \operatorname{div} f\nu,$$

where the derivatives of measures are understood in the weak sense, or in the sense of distributions (i.e., via their action on suitable test functions), and the change of sign comes from the integration by parts formula. For more details the interested reader is referred to any textbook on functional analysis and partial differential equation, e.g., [11]. The concept of weak derivatives of measures is not essential for the remainder of the paper and only highlights the important connections of our approach to PDE literature.

Now consider that the initial state is not a single point but that its distribution in space is modeled by an *initial measure*[3] $\mu_0 \in M(X)$, and that for each initial state $x_0 \in \operatorname{spt} \mu_0$ there exists an admissible trajectory $x(\cdot\,|\,x_0) \in \mathcal{X}(x_0)$ with an admissible control function $u(\cdot\,|\,x_0) \in \mathcal{U}$. Then we can define the average occupation measure $\mu \in M([0,T] \times X \times U)$ by

$$\mu(A \times B \times C) := \int_X \mu(A \times B \times C\,|\,x_0)\, d\mu_0(x_0),
\tag{7}$$

and the *final measure* $\mu_T \in M(X_T)$ by

$$\mu_T(B) := \int_X I_B(x(T\,|\,x_0))\, d\mu_0(x_0).
\tag{8}$$

The average occupation occupation measure μ measures the average time spent by the state and control trajectories in subsets of $X \times U$, where the averaging is over the distribution

---

[3]The measure μ_0 can be thought of as the probability distribution of x_0 although we do not require that its mass be normalized to one.

of the initial state given by μ_0; the final measure μ_T represents the distribution of the state at the final time T after it has been transported along system trajectories from the initial distribution μ_0.

It follows by integrating (6) with respect to μ_0 that

$$\int_{X_T} v(T, x)\, d\mu_T(x) \;=\; \int_X v(0, x)\, d\mu_0(x)$$
$$+ \int_{[0,T] \times X \times U} \mathcal{L}v(t, x, u)\, d\mu(t, x, u),$$

or more concisely

$$\langle \mu_T,\, v(T, \cdot) \rangle = \langle \mu_0,\, v(0, \cdot) \rangle \;+\; \langle \mu,\, \mathcal{L}v \rangle \quad \forall\, v \in C^1([0, T] \times X), \tag{9}$$

which is a linear equation linking the nonnegative measures μ_T, μ_0 and μ. Denoting δ_t the Dirac measure at a point t and $\otimes$ the product of measures, we can write $\langle \mu_0,\, v(0, \cdot) \rangle = \langle \delta_0 \otimes \mu_0,\, v \rangle$ and $\langle \mu_T,\, v(T, \cdot) \rangle = \langle \delta_T \otimes \mu_T,\, v \rangle$. Then, Eq. (9) can be rewritten equivalently using the adjoint relation as

$$\langle \mathcal{L}'\mu,\, v \rangle = \langle \delta_T \otimes \mu_T, v \rangle \;-\; \langle \delta_0 \otimes \mu_0,\, v \rangle \quad \forall\, v \in C^1([0, T] \times X),$$

and since this equation is required to hold for all test functions v, we obtain the linear operator equation

$$\delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}'\mu. \tag{10}$$

This equation is classical in fluid mechanics and statistical physics, where $\mathcal{L}'$ is usually written using distributional derivatives of measures as remarked above; then the equation is referred to as Liouville's partial differential equation.

Each family of admissible trajectories starting from a given initial distribution $\mu_0 \in M(X)$ satisfies Liouville's equation (10). The converse may not hold in general although for the computation of the ROA the two formulations can be considered equivalent, at least from a practical viewpoint. Let us briefly elaborate more on this subtle point now.

3.2   Relaxed ROA

The control system $\dot{x}(t) = f(t, x(t), u(t))$, $u(t) \in U$, can be viewed as a differential inclusion

$$\dot{x}(t) \in f(t, x(t), U) := \{f(t, x(t), u)\,:\, u \in U\}. \tag{11}$$

We show in Lemma 3 in Appendix A that any triplet of measures satisfying Liouville's equation (10) corresponds to a family of trajectories of the convexified inclusion

$$\dot{x}(t) \in \operatorname{conv} f(t, x(t), U) \tag{12}$$

starting from the initial distribution μ_0, where conv denotes the convex hull[4]. Let us denote the set of absolutely continuous admissible trajectories of (12) by

$$\bar{\mathcal{X}}(x_0) := \{x(\cdot) : \dot{x}(t) \in \operatorname{conv} f(t, x(t), U) \text{ a.e., } x(0) = x_0,$$
$$x(T) \in X_T,\, x(t) \in X \,\forall\, t \in [0, T]\}.$$

---

[4]Note that the set conv $f(t, x(t), U)$ is closed for every t since f is continuous and U compact.

Given a family[5] of admissible trajectories of the convexified inclusion starting from an initial distribution μ_0, the occupation and final measures can be defined in a complete analogy via (7) and (8), but now there are only the time and space arguments in the occupation measure, not the control argument. In Appendix A, we state and prove the correspondence between the convexified inclusion (12) and the measures satisfying the Liouville equation (10).

Define now the relaxed region of attraction as

$$\bar{X}_0 := \left\{ x_0 \in X \; : \; \bar{\mathcal{X}}(x_0) \neq \emptyset \right\}.$$

Clearly $X_0 \subset \bar{X}_0$ and the inclusion can be strict; see Appendix C for concrete examples. However, by the Filippov-Ważewski relaxation Theorem [5, Theorem 10.4.4, Corollary 10.4.5], the trajectories of the original inclusion (11) are dense (w.r.t. the metric of uniform convergence of absolutely continuous functions of time) in the set of trajectories of the convexified inclusion (12). This implies that the relaxed region of attraction $\bar{X}_0$ corresponds to the region of attraction of the original system but with infinitesimally dilated constraint sets X and X_T; see Appendix B for more details. Therefore, we argue that there is little difference between the two ROAs from a practical point of view. Nevertheless, because of this subtle distinction we make the following standing assumption in the remaining part of the paper.

Assumption 2 *Control system (1) is such that $\lambda(X_0) = \lambda(\bar{X}_0)$.*

In other words, the volume of the classical ROA X_0 is assumed to be equal to the volume of the relaxed ROA $\bar{X}_0$. Obviously, this is satisfied if $X_0 = \bar{X}_0$, but otherwise these sets may differ by a set of zero Lebesgue measure. Any of the following conditions on control system (1) is *sufficient* for Assumption 2 to hold:

- $\dot{x}(t) \in f(t, x(t), U)$ with $f(t, x, U)$ convex for all t, x,

- $\dot{x}(t) = f(t, x(t)) + g(t, x(t))u(t)$, $u(t) \in U$ with U convex,

- uncontrolled dynamics $\dot{x}(t) = f(t, x(t))$,

as well as all controllability assumptions allowing the application of the constrained Filippov-Ważewski Theorem; see, e.g., [13, Corollary 3.2] and the discussion around Assumption I in [15].

3.3 ROA via optimization

The problem of computing ROA X_0 can be reformulated as follows:

$$
\begin{aligned}
q^* \;=\; \sup \quad & \lambda(\operatorname{spt} \mu_0) \\
\text{s.t.} \quad & \delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}'\mu \\
& \mu \geq 0, \; \mu_0 \geq 0, \; \mu_T \geq 0 \\
& \operatorname{spt} \mu \subset [0, T] \times X \times U \\
& \operatorname{spt} \mu_0 \subset X, \; \operatorname{spt} \mu_T \subset X_T,
\end{aligned}
\tag{13}
$$

---

[5]Each such family can be described by a measure on $C([0,T]; \mathbb{R}^n)$ which is supported on the absolutely continuous solutions to (12). Note that there may be more than one trajectory corresponding to a single initial condition since the inclusion (12) may admit multiple solutions.

where the given data are f, X, X_T, U, and the supremum is over a vector of nonnegative measures $(\mu, \mu_0, \mu_T) \in M([0, T] \times X \times U) \times M(X) \times M(X_T)$. Problem (13) is an infinite-dimensional optimization problem on the cone of nonnegative measures.

The rationale behind problem (13) is as follows. The first constraint is the Liouville's equation (10) which, along with the nonnegativity constraints, ensures that any triplet of measures (μ_0, μ, μ_T) feasible in (13) corresponds to an initial, an occupation and a terminal measure generated by trajectories of the controlled ODE (1) (or more precisely of the convexified differential inclusion (11)). The support constraint on the occupation measure μ ensures that these trajectories satisfy the state and control constraints; the support constraint on μ_T ensures that the trajectories end in the target set. Maximizing the volume of the support of the initial measure then yields an initial measure with the support equal to the ROA up to a set of zero volume[6] (in view of Assumption 2). This discussion is summarized in the following Lemma.

Lemma 1 *The optimal value of problem (13) is equal to the volume of the ROA X_0, that is, $q^* = \lambda(X_0)$.*

Proof: By definition of the ROA, for any initial condition $x_0 \in X_0$ there is an admissible trajectory in $\mathcal{X}(x_0)$. Therefore for any initial measure μ_0 with $\operatorname{spt} \mu_0 \subset X_0$ there exist an occupation measure μ and a final measure μ_T such that the constraints of problem (13) are satisfied. Thus, $q^* \geq \lambda(X_0) = \lambda(\bar{X}_0)$, where the equality follows from Assumption 2.

Now we show that $q^* \leq \lambda(X_0) = \lambda(\bar{X}_0)$. For contradiction, suppose that a triplet of measures (μ_0, μ, μ_T) is feasible in (13) and that $\lambda(\operatorname{spt} \mu_0 \setminus \bar{X}_0) > 0$. From Lemma 3 in Appendix A there is a family of admissible trajectories of the inclusion (12) starting from μ_0 generating the (t, x)-marginal of the occupation measure μ and the final measure μ_T. However, this is a contradiction since no trajectory starting from $\operatorname{spt} \mu_0 \setminus \bar{X}_0$ can be admissible. Thus, $\lambda(\operatorname{spt} \mu_0 \setminus \bar{X}_0) = 0$ and so $\lambda(\operatorname{spt} \mu_0) \leq \lambda(\bar{X}_0)$. Consequently, $q^* \leq \lambda(\bar{X}_0) = \lambda(X_0)$. $\qquad\square$

4 Primal infinite-dimensional LP on measures

The key idea behind the presented approach consists in replacing the optimization over the support of the measure μ_0 by the maximization of its mass[7] under the constraint that μ_0 is dominated by the Lebesgue measure. This leads to the following infinite-dimensional LP:

$$
\begin{aligned}
p^* = \sup \quad & \mu_0(X) \\
\text{s.t.} \quad & \delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}' \mu \\
& \mu \geq 0, \ \lambda \geq \mu_0 \geq 0, \ \mu_T \geq 0 \\
& \operatorname{spt} \mu \subset [0, T] \times X \times U \\
& \operatorname{spt} \mu_0 \subset X, \ \operatorname{spt} \mu_T \subset X_T,
\end{aligned}
\tag{14}
$$

---

[6]Even though the support of the initial measure attaining the maximum in (13) can differ from the ROA on the set of zero volume, the outer approximations obtained in Section 7 are valid "everywhere", not "almost everywhere".

[7]The mass of the measure μ_0 is defined as $\int_X 1 \, d\mu_0 = \mu_0(X)$.

where the supremum is over a vector of nonnegative measures $(\mu, \mu_0, \mu_T) \in M([0,T] \times X \times U) \times M(X) \times M(X_T)$. In problem (14) the constraint $\lambda \geq \mu_0$ means that $\lambda(A) \geq \mu_0(A)$ for all sets $A \subset X$. Note how the objective functions differ in problems (13) and (14).

The following theorem is then almost immediate.

Theorem 1 *The optimal value of the infinite-dimensional LP problem (14) is equal to the volume of the ROA X_0, that is, $p^* = \lambda(X_0)$. Moreover, the supremum is attained with the μ_0-component of the optimal solution equal to the restriction of the Lebesgue measure to the ROA X_0.*

Proof: Since the constraint set of problem (14) is tighter than that of problem (13), by Lemma 1 we have that $\lambda(\operatorname{spt} \mu_0) \leq \lambda(X_0)$ for any feasible μ_0. From the constraint $\mu_0 \leq \lambda$ we get $\mu_0(X) = \mu_0(\operatorname{spt} \mu_0) \leq \lambda(\operatorname{spt} \mu_0) \leq \lambda(X_0)$ for any feasible μ_0. Therefore $p^* \leq \lambda(X_0)$. But by definition of the ROA X_0, the restriction of the Lebesgue measure to X_0 is feasible in (14), and so $p^* \geq \lambda(X_0)$. Consequently $p^* = \lambda(X_0)$. $\qquad\square$

Now we reformulate problem (14) to an equivalent form more convenient for dualization and subsequent theoretical analysis. To this end, let us define the complementary measure (a slack variable) $\hat{\mu}_0 \in M(X)$ such that the inequality $\lambda \geq \mu_0 \geq 0$ in (14) can be written equivalently as the constraints $\mu_0 + \hat{\mu}_0 = \lambda$, $\mu_0 \geq 0$, $\hat{\mu}_0 \geq 0$. Then problem (14) is equivalent to

$$
\begin{aligned}
p^* = \quad & \sup \quad \mu_0(X) \\
& \text{s.t.} \quad \delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}'\mu \\
& \qquad \mu_0 + \hat{\mu}_0 = \lambda \\
& \qquad \mu \geq 0, \ \mu_0 \geq 0, \ \mu_T \geq 0, \ \hat{\mu}_0 \geq 0 \\
& \qquad \operatorname{spt} \mu \subset [0,T] \times X \times U \\
& \qquad \operatorname{spt} \mu_0 \subset X, \ \operatorname{spt} \mu_T \subset X_T \\
& \qquad \operatorname{spt} \hat{\mu}_0 \subset X.
\end{aligned}
\tag{15}
$$

5 Dual infinite-dimensional LP on functions

In this section, we derive a dual formulation of problem (15) (and hence (14)) on the space of continuous functions. A certain super-level set of any feasible solution to the dual problem yields an outer approximation to the ROA X_0.

Consider the infinite-dimensional LP problem

$$
\begin{aligned}
d^* = \quad & \inf \int_X w(x) \, d\lambda(x) \\
& \text{s.t.} \ \mathcal{L}v(t,x,u) \leq 0, \qquad \forall \, (t,x,u) \in [0,T] \times X \times U \\
& \qquad w(x) \geq v(0,x) + 1, \ \forall \, x \in X \\
& \qquad v(T,x) \geq 0, \qquad \quad \forall \, x \in X_T \\
& \qquad w(x) \geq 0, \qquad \qquad \forall \, x \in X,
\end{aligned}
\tag{16}
$$

where the infimum is over $(v, w) \in C^1([0,T] \times X) \times C(X)$. The dual has the following interpretation: the constraint $\mathcal{L}v \leq 0$ forces v to decrease along trajectories and hence

necessarily $v(0, x) \geq 0$ on X_0 because of the constraint $v(T, x) \geq 0$ on X_T. Consequently, $w(x) \geq 1$ on X_0. This instrumental observation is formalized in the following Lemma.

Lemma 2 *Let (v, w) be a pair of function feasible in (16). Then $v(0, \cdot) \geq 0$ on X_0 and $w \geq 1$ on X_0.*

Proof: By definition of X_0, given any $x_0 \in X_0$ there exists $u(t)$ such that $x(t) \in X$, $u(t) \in U$ for all $t \in [0, T]$ and $x(T) \in X_T$. Therefore, since $\mathcal{L}v \leq 0$ on $[0, T] \times X \times U$ and $v(T, \cdot) \geq 0$ on X_T,

$$0 \leq v(T, x(T)) = v(0, x_0) + \int_0^T \mathcal{L}v(t, x(t), u(t)) \, dt$$
$$\leq v(0, x_0) \leq w(x_0) - 1.$$

$\square$

Next, we have the following salient result:

Theorem 2 *There is no duality gap between the primal infinite-dimensional LP problems (14) and (15) and the infinite-dimensional dual LP problem (16) in the sense that $p^* = d^*$.*

Proof: To streamline the exposition, let

$$\mathcal{C} := C([0, T] \times X \times U) \times C(X) \times C(X_T) \times C(X),$$
$$\mathcal{M} := M([0, T] \times X \times U) \times M(X) \times M(X_T) \times M(X),$$

and let $\mathcal{K}$ and $\mathcal{K}'$ denote the positive cones of $\mathcal{C}$ and $\mathcal{M}$ respectively. Note that the cone $\mathcal{K}'$ of nonnegative measures of $\mathcal{M}$ can be identified with the topological dual of the cone $\mathcal{K}$ of nonnegative continuous functions of $\mathcal{C}$. The cone $\mathcal{K}'$ is equipped with the weak-* topology; see [29, Section 5.10]. Then, the LP problem (15) can be rewritten as

$$
\begin{aligned}
p^* \ = \ \sup \ & \langle \gamma, \ c \rangle \\
\text{s.t.} \ & \mathcal{A}'\gamma = \beta \\
& \gamma \in \mathcal{K}',
\end{aligned}
\tag{17}
$$

where the infimum is over the vector $\gamma := (\mu, \mu_0, \mu_T, \hat{\mu}_0)$, the linear operator $\mathcal{A}' : \mathcal{K}' \to C^1([0, T] \times X)' \times M(X)$ is defined by $\mathcal{A}'\gamma := (-\mathcal{L}'\mu - \delta_0 \otimes \mu_0 + \delta_T \otimes \mu_T \ , \ \mu_0 + \hat{\mu}_0)$, the right hand side of the equality constraint in (17) is the vector of measures $\beta := (0 \ , \ \lambda) \in M([0, T] \times X) \times M(X)$, the vector function in the objective is $c := (0, 1, 0, 0) \in \mathcal{C}$, so the objective function itself is

$$\langle \gamma, \ c \rangle = \int_X d\mu_0 = \mu_0(X).$$

The LP problem (17) can be interpreted as a dual to the LP problem

$$
\begin{aligned}
d^* \ = \ \inf \ & \langle \beta, \ z \rangle \\
\text{s.t.} \ & \mathcal{A}(z) - c \in \mathcal{K},
\end{aligned}
\tag{18}
$$

11

where the infimum is over $z := (v, w) \in C^1([0, T] \times X) \times C(X)$, and the linear operator $\mathcal{A} : C^1([0, T] \times X) \times C(X) \to \mathcal{C}$ is defined by

$$\mathcal{A}z := (-\mathcal{L}v, \ w - v(0, \cdot), \ v(T, \cdot), \ w)$$

and satisfies the adjoint relation $\langle \mathcal{A}'\gamma, z \rangle = \langle \gamma, \mathcal{A}z \rangle$. The LP problem (18) is exactly the LP problem (16).

To conclude the proof we use an argument similar to that of [26, Section C.4]. From [2, Theorem 3.10] there is no duality gap between LPs (17) and (18) if the supremum p^* is finite and the set $P := \{(\mathcal{A}'\gamma, \langle \gamma, c \rangle) : \gamma \in \mathcal{K}'\}$ is closed in the weak-* topology of $\mathcal{K}'$. The fact that p^* is finite follows readily from the constraint $\mu_0 + \hat{\mu}_0 = \lambda$, $\hat{\mu}_0 \geq 0$, and from compactness of X. To prove closedness, we first remark that $\mathcal{A}'$ is weakly-* continuous[8] since $\mathcal{A}(z) \in \mathcal{C}$ for all $z \in C^1([0, T] \times X) \times C(X)$. Then we consider a sequence $\gamma_k \in \mathcal{K}'$ and we want to show that its accumulation point $\lim_{k \to \infty}(\mathcal{A}'\gamma_k, \langle \gamma_k, c \rangle)$ belongs to P. Since the supports of the measures are compact and p^* is finite (hence μ_0, μ_T and $\hat{\mu}_0$ are bounded) and since $T < \infty$ (hence μ is bounded), the sequence γ_k is bounded and from the weak-* compactness of the unit ball (Alaoglu's Theorem, see [29, Section 5.10, Theorem 1]) there is a subsequence γ_{k_i} that converges weakly-* to an element $\gamma \in \mathcal{K}'$ so that $\lim_{i \to \infty}(\mathcal{A}'\gamma_{k_i}, \langle \gamma_{k_i}, c \rangle) \in P$. $\qquad\square$

In the above proof we observed that the equivalent infinite-dimensional LPs (14) and (15) are formulated in the dual of a Banach space equipped with the appropriate topology. It follows that the supremum in LPs (14) and (15) is attained (by the restriction of the Lebesgue measure to X_0), a statement already proved by a different means in Theorem 1. In contrast, the infimum in problem (16) is not attained in $C^1([0, T] \times X) \times C(X)$, but the w-component of feasible solutions of (16) converges to the discontinuous indicator I_{X_0} as we show next.

Before we state our convergence results, we recall the following types of convergence of a sequence of functions $w_k : X \to \mathbb{R}$ to a function $w : X \to \mathbb{R}$ on a compact set $X \subset \mathbb{R}^n$. As $k \to \infty$, the functions w_k converge to w:

- in L^1 norm if $\int_X |w_k - w| \, d\lambda \to 0$,

- in Lebesgue measure if $\lambda(\{x \ : \ |w_k(x) - w(x)| \geq \epsilon\}) \to 0 \ \forall \epsilon > 0$,

- almost everywhere if $\exists B \subset X$, $\lambda(B) = 0$, such that $w_k \to w$ pointwise on $X \backslash B$,

- almost uniformly if $\forall \epsilon > 0, \exists B \subset X, \lambda(B) < \epsilon$, such that $w_k \to w$ uniformly on $X \backslash B$.

We also recall that convergence in L^1 norm implies convergence in Lebesgue measure and that almost uniform convergence implies convergence almost everywhere (see [3, Theorems 2.5.2 and 2.5.3]). Therefore we will state our results in terms of the stronger notions of L^1 norm and almost uniform convergence.

Theorem 3 *There is a sequence of feasible solutions to problem (16) such that its w-component converges from above to I_{X_0} in L^1 norm and almost uniformly.*

---

[8]The weak-* topology on $C^1([0, T] \times X)' \times M(X)$ is induced by the standard topologies on C^1 and $\mathcal{C}$ – the topology of uniform convergence of the function and its derivative on C^1 and the topology of uniform convergence on $\mathcal{C}$.

Proof: By Theorem 1, the optimal solution to the primal is attained by the restriction of the Lebesgue measure to X_0. Consequently,

$$p^* = \int_X I_{X_0}(x) d\lambda(x). \tag{19}$$

By Theorem 2, there is no duality gap ($p^* = d^*$), and therefore there exists a sequence $(v_k, w_k) \in C^1([0,T] \times X) \times C(X)$ feasible in (16) such that

$$p^* = d^* = \lim_{k \to \infty} \int_X w_k(x) \, d\lambda(x). \tag{20}$$

From Lemma 2 we have $w_k \geq 1$ on X_0 and since $w_k \geq 0$ on X, we have $w_k \geq I_{X_0}$ on X for all k. Thus, subtracting (19) from (20) gives

$$\lim_{k \to \infty} \int_X (w_k(x) - I_{X_0}(x)) \, d\lambda(x) = 0,$$

where the integrand is nonnegative. Hence w_k converges to I_{X_0} in L^1 norm. From [3, Theorems 2.5.2 and 2.5.3] there exists a subsequence converging almost uniformly. $\qquad \square$

6 LMI relaxations and SOS approximations

In this section we show how the infinite-dimensional LP problem (15) can be approximated by a hierarchy of LMI problems with the approximation error vanishing as the relaxation order tends to infinity. The dual LMI problem is a sum-of-squares (SOS) problem and yields a converging sequence of outer approximations to the ROA.

The measures in equation (9) (or (10)) are fully determined by their values on a family of functions whose span is dense in $C^1([0,T] \times X)$. Hence, since all sets are assumed compact, the family of test functions in (9) can be restricted to any polynomial basis (since polynomials are dense in the space of continuous functions on compact sets equipped with the supremum norm). The basis of our choice is the set of all monomials. This basis is convenient for subsequent exposition and is employed by existing software (e.g., [21, 28]). Nevertheless, another polynomial basis may be more appropriate from a numerical point of view (see the Conclusion for a discussion).

Let $\mathbb{R}_k[x]$ denote the vector space of real multivariate polynomials of total degree less than or equal to k. Each polynomial $p(x) \in \mathbb{R}_k[x]$ can be expressed in the monomial basis as

$$p(x) = \sum_{|\alpha| \leq k} p_\alpha x^\alpha,$$

where α runs over the multi-indices (vectors of nonnegative integers) such that $|\alpha| = \sum_i \alpha_i \leq k$, and $x^\alpha = \prod_i x^{\alpha_i}$. A polynomial $p(x)$ is identified with its vector of coefficients $p := (p_\alpha)$ whose entries are indexed by α. Given a vector of real numbers $y := (y_\alpha)$ indexed by α, we define the linear functional $L_y : \mathbb{R}_k[x] \to \mathbb{R}$ such that

$$L_y(p) := p'y := \sum_\alpha p_\alpha y_\alpha,$$

where the prime denotes transposition[9]. When entries of y are moments of a measure ν, i.e.,

$$y_\alpha = \int x^\alpha d\nu(x),$$

the linear functional models integration of a polynomial w.r.t. ν, i.e.,

$$L_y(p) = \langle \nu,\, p \rangle = \int p(x)\, d\nu(x) = \sum_\alpha p_\alpha \int x^\alpha d\nu(x) = p'y.$$

When this linear functional acts on the square of a polynomial p of degree k, it becomes a quadratic form in the polynomial coefficients space, and we denote by $M_k(y)$ and call the *moment matrix* of order k the matrix of this quadratic form, which is symmetric and linear in y:

$$L_y(p^2) = p' M_k(y) p.$$

Finally, given a polynomial $g(x) \in \mathbb{R}[x]$ we define the *localizing matrix* $M_k(g,\, y)$ by the equality

$$L_y(gp^2) = p' M_k(g,\, y) p.$$

The matrix $M_k(g,\, y)$ is also symmetric and linear in y. For a detailed exposition and examples of moment and localizing matrices see [26, Section 3.2.1].

Let y, y_0, y_T and $\hat{y}_0$ respectively denote the sequences of moments of measures μ, μ_0, μ_T and $\hat{\mu}_0$ such that the constraints in problem (15) are satisfied. Then it follows that these sequences satisfy an infinite-dimensional linear system of equations corresponding to the equality constraints of problem (15) written explicitly as

$$\int_{X_T} v(T, x)\, d\mu_T(x) - \int_X v(0, x)\, d\mu_0(x)$$

$$- \int_{[0,T] \times X \times U} \mathcal{L}v(t, x, u)\, d\mu(t, x, u) = 0,$$

$$\int_X w(x)\, d\mu_0(x) + \int_X w(x)\, d\hat{\mu}_0(x) = \int_X w(x)\, d\lambda(x).$$

For the particular choice of test functions $v(t, x) = t^\alpha x^\beta$ and $w(x) = x^\beta$ for all $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}^n$, let us denote by

$$A_k(y, y_0, y_T, \hat{y}_0) = b_k$$

the finite-dimensional truncation of this system obtained by considering only the test functions of total degree less than or equal to $2k$. Let further

$$d_{Xi} := \left\lceil \frac{\deg g_i^X}{2} \right\rceil, \ \ d_{Ui} := \left\lceil \frac{\deg g_i^U}{2} \right\rceil, \ \ d_{Ti} := \left\lceil \frac{\deg g_i^{X_T}}{2} \right\rceil,$$

---

[9]In this paper, the operator prime is used to refer to the adjoint of a linear operator. If the linear operator is real-valued and finite-dimensional, the adjoint operator is the transposition operator.

where deg denotes the degree of a polynomial. The primal LMI relaxation of order k then reads

$$
\begin{aligned}
p_k^* \;=\; &\max \quad (y_0)_0 \\
&\text{s.t.} \quad A_k(y, y_0, y_T, \hat{y}_0) = b_k \\
M_k(y) \succeq 0, \quad & M_{k-d_{X\,i}}(g_i^X, y) \succeq 0, i \in \mathbb{Z}_{[1,n_X]} \\
M_{k-1}(t(T-t), y) \succeq 0, \quad & M_{k-d_{U\,i}}(g_i^U, y) \succeq 0, i \in \mathbb{Z}_{[1,n_U]} \\
M_k(y_0) \succeq 0, \quad & M_{k-d_{X\,i}}(g_i^X, y_0) \succeq 0, i \in \mathbb{Z}_{[1,n_X]} \\
M_k(y_T) \succeq 0, \quad & M_{k-d_{T\,i}}(g_i^{X_T}, y_T) \succeq 0, i \in \mathbb{Z}_{[1,n_T]} \\
M_k(\hat{y}_0) \succeq 0, \quad & M_{k-d_{X\,i}}(g_i^X, \hat{y}_0) \succeq 0, i \in \mathbb{Z}_{[1,n_X]},
\end{aligned}
\tag{21}
$$

where the notation $\succeq 0$ stands for positive semidefinite and the minimum is over sequences $(y, y_0, y_T, \hat{y}_0)$ truncated to degree $2k$. The objective function is the first element (i.e., the mass) of the truncated moment sequence y_0 corresponding to the initial measure; the equality constraint captures the two equality constraints of problem (15) evaluated on monomials of degree up to $2k$; and the LMI constraints involving the moment and localizing matrices capture the support constraints on the measures. Note that both the equality constraint and the LMI constraints are necessarily satisfied by the moment sequence of any vector of measures feasible in (15). The constraint set of (21) is therefore looser than that of (15); however, the discrepancy between the two constraint sets monotonically vanishes as the relaxation order k tends to infinity (see Corollary 1 below).

Problem (21) is a semidefinite program (SDP), where a linear function is minimized subject to convex LMI constraints, or equivalently a finite-dimensional LP in the cone of positive semidefinite matrices.

Without loss of generality we make the following standard assumption for the reminder of this section.

Assumption 3 *One of the polynomials defining the sets X, U resp. X_T, is equal to $g_i^X(x) = R_X - \|x\|_2^2$, $g_i^U(u) = R_U - \|u\|_2^2$ resp. $g_i^{X_T}(x) = R_T - \|x\|_2^2$ for some constants $R_X \geq 0$, $R_U \geq 0$ resp. $R_T \geq 0$.*

Assumption 3 is without loss of generality since the sets X, U and X_T are bounded, and therefore redundant ball constraints of the form $R_X - \|x\|_2^2 \geq 0$, $R_U - \|u\|_2^2 \geq 0$ and $R_T - \|x\|_2^2 \geq 0$ can always be added to the description of the sets X, U and X_T for sufficiently large R_X, R_U and R_T.

The dual to the SDP problem (21) is given by

$$
\begin{aligned}
d_k^* = &\inf \quad \mathbf{w}'l \\
&\text{s.t.} \; -\mathcal{L}v(t,x,u) = p(t,x,u) + q_0(t,x,u)t(T-t) \\
&\qquad + \sum_{i=1}^{n_X} q_i(t,x,u)g_i^X(x) + \sum_{i=1}^{n_U} r_i(t,x,u)g_i^U(x) \\
&\quad w(x) - v(0,x) - 1 = p_0(x) + \sum_{i=1}^{n_X} q_{0i}(x)g_i^X(x) \\
&\quad v(T,x) = p_T(x) + \sum_{i=1}^{n_T} q_{Ti}(x)g_i^{X_T}(x) \\
&\quad w(x) = s_0(x) + \sum_{i=1}^{n_X} s_{0i}(x)g_i^X(x),
\end{aligned}
\tag{22}
$$

where l is the vector of Lebesgue moments over X indexed in the same basis in which the polynomial $w(x)$ with coefficients $\mathbf{w}$ is expressed. The minimum is over polynomials $v(t,x) \in \mathbb{R}_{2k}[t,x]$ and $w \in \mathbb{R}_{2k}[x]$, and polynomial sum-of-squares $p(t,x,u)$, $q_i(t,x,u)$,

$i \in \mathbb{Z}_{[0,n_X]}$, $r_i(t,x,u)$, $i \in \mathbb{Z}_{[1,n_U]}$, $p_0(x)$, $p_T(x)$, $q_{0i}(x)$, $q_{Ti}(x)$, $s_0(x)$, $s_{0i}(x)$, $i \in \mathbb{Z}_{[1,n_X]}$ of appropriate degrees. The constraints that polynomials are sum-of-squares can be written explicitly as LMI constraints (see, e.g., [26]), and the objective is linear in the coefficients of the polynomial $w(x)$; therefore problem (22) can be formulated as an SDP.

Theorem 4 *There is no duality gap between primal LMI problem (21) and dual LMI problem (22), i.e., $p_k^* = d_k^*$.*

Proof: See Appendix D. □

7 Outer approximations and convergence results

In this section we show how the dual LMI problem (22) gives rise to a sequence of outer approximations to the ROA X_0 with a guaranteed convergence. In addition, we prove the convergence of the primal and dual optimal values p_k^* and d_k^* to the volume of the ROA, and the convergence of the w-component of an optimal solution to the dual LMI problem (22) to the indicator function of the ROA I_{X_0}.

Let the polynomials (w_k, v_k), each of total degree at most $2k$, denote an optimal solution to the problem k^{th} order dual SDP approximation (22) and let $\bar{w}_k := \min_{i \leq k} w_i$ and $\bar{w}_k := \min_{i \leq k} v_i$ denote their running minima. Then, in view of Lemma 2 using the fact that any feasible solution the (22) is feasible in (16), the sets

$$X_{0k} := \{x \in X : v_k(0,x) \geq 0\} \tag{23}$$

and

$$\bar{X}_{0k} := \{x \in X : \bar{v}_k(0,x) \geq 0\} \tag{24}$$

provide outer approximations to the ROA; in fact clearly $X_{0k} \supset \bar{X}_{0k} \supset X_0$ for all $k \in \{1, 2, \dots\}$.

Our first convergence results proves the convergence of w_k and $\bar{w}_k$ to the indicator function of the ROA I_{X_0}.

Theorem 5 *Let $w_k \in \mathbb{R}_{2k}[x]$ denote the w-component of an optimal solution to the dual LMI problem (22) and let $\bar{w}_k(x) = \min_{i \leq k} w_i(x)$. Then w_k converges from above to I_{X_0} in L^1 norm and $\bar{w}_k$ converges from above to I_{X_0} in L^1 norm and almost uniformly.*

Proof: From Lemma 2 and Theorem 3, for every $\epsilon > 0$ there exists a $(v, w) \in C^1([0, T] \times X) \times C(X)$ feasible in (16) such that $w \geq I_{X_0}$ and $\int_X (w - I_{X_0}) \, d\lambda < \epsilon$. Set

$$\tilde{v}(t,x) := v(t,x) - \epsilon t + (T+1)\epsilon,$$
$$\tilde{w}(x) := w(x) + (T+3)\epsilon.$$

Since v is feasible in (16), we have $\mathcal{L}\tilde{v} = \mathcal{L}v - \epsilon$, and $\tilde{v}(T,x) = v(T,x) + \epsilon$. Since also $\tilde{w}(x) - \tilde{v}(0,x) \geq 1 + 2\epsilon$, it follows that $(\tilde{v}, \tilde{w})$ is *strictly* feasible in (16) with a margin at least

ϵ. Since $[0, T] \times X$ and X are compact, there exist[10] polynomials $\hat{v}$ and $\hat{w}$ of a sufficiently high degree such that $\sup_{[0,T] \times X} |\tilde{v} - \hat{v}| < \epsilon$, $\sup_{[0,T] \times X \times U} |\mathcal{L}\tilde{v} - \mathcal{L}\hat{v}| < \epsilon$ and $\sup_X |\tilde{w} - \hat{w}| < \epsilon$. The pair of polynomials $(\hat{v}, \hat{w})$ is therefore *strictly* feasible in (16) and as a result, under Assumption 3, feasible in (22) for a sufficiently large relaxation order k (this follows from the classical Positivstellensatz by Putinar; see, e.g., [26] or [36]), and moreover $\hat{w} \geq w$. Consequently, $\int_X |\tilde{w} - \hat{w}| \, d\lambda \leq \epsilon\lambda(X)$, and so $\int_X (\hat{w} - w) \, d\lambda \leq \epsilon\lambda(X)(T + 4)$. Therefore

$$\int_X (\hat{w} - I_{X_0}) \, d\lambda < \epsilon K, \quad \hat{w} \geq I_{X_0},$$

where $K := [1 + (T + 4)\lambda(X)] < \infty$ is a constant. This proves the first statement since ϵ was arbitrary.

The second statement immediately follows since given a sequence $w_k \to I_{X_0}$ in L_1 norm, there exists a subsequence w_{k_i} that converges almost uniformly to I_{X_0} by [3, Theorems 2.5.2 and 2.5.3] and clearly $\bar{w}_k(x) \leq \min\{w_{k_i}(x) : k_i \leq k\}$. $\qquad\square$

The following Corollary follows immediately from Theorem 5.

Corollary 1 *The sequence of infima of LMI problems (22) converges monotonically from above to the supremum of the infinite-dimensional LP problem (16), i.e., $d^* \leq d^*_{k+1} \leq d^*_k$ and $\lim_{k\to\infty} d^*_k = d^*$. Similarly, the sequence of maxima of LMI problems (21) converges monotonically from above to the maximum of the infinite-dimensional LP problem (14), i.e., $p^* \leq p^*_{k+1} \leq p^*_k$ and $\lim_{k\to\infty} p^*_k = p^*$.*

Proof: Monotone convergence of the dual optima d^*_k follows immediately from Theorem 5 and from the fact that the higher the relaxation order k, the looser the constraint set of the minimization problem (22). To prove convergence of the primal maxima observe that from weak SDP duality we have $d^*_k \geq p^*_k$ and from Theorems 5 and 2 it follows that $d^*_k \to d^* = p^*$. In addition, clearly $p^*_k \geq p^*$ and $p^*_{k+1} \leq p^*_k$ since the higher the relaxation order k, the tighter the constraint set of the maximization problem (21). Therefore $p^*_k \to p^*$ monotonically from above. $\qquad\square$

Theorem 5 establishes a functional convergence of w_k to I_{X_0} and Corollary 1 a convergence of the primal and dual optima p^*_k and d^*_k to the volume of the ROA $\lambda(X_0) = p^* = d^*$. Finally, the following theorem establishes a set-wise convergence of the zero super-level sets of $v(0, \cdot)$ to X_0.

Theorem 6 *Let $(v_k, w_k) \in \mathbb{R}_{2k}[t, x] \times \mathbb{R}_{2k}[x]$ denote a solution to the dual LMI problem (22). Then the sets X_{0k} and $\bar{X}_{0k}$ defined in (23) and (24) converge to the ROA X_0 from the outside such that $X_{0k} \supset \bar{X}_{0k} \supset X_0$ and*

$$\lim_{k\to\infty} \lambda(X_{0k} \setminus X_0) = 0 \quad and \quad \lim_{k\to\infty} \lambda(\bar{X}_{0k} \setminus X_0) = 0.$$

Moreover the convergence of $\bar{X}_{0k}$ is monotonous, i.e., $X_{0i} \subset X_{0j}$ whenever $i \geq j$.

---

[10]This follows from an extension of the Stone-Weierstrass theorem that allows for a simultaneous uniform approximation of a function and its derivatives by a polynomial on a compact set; see, e.g., [24].

Proof: The inclusion $X_{0k} \supset \bar{X}_{0k} \supset X_0$ follows from Lemma 2 since any solution to (22) is feasible in (16) and since $\bar{X}_{0k} = \cap_{i=1}^k X_{0i}$. The latter fact also proves the monotonicity of the sequence $\bar{X}_{0k}$. Next, from Lemma 2 we have $w_k \geq I_{X_0}$ and therefore, since $w_k \geq v_k(0,\cdot) + 1$ on X, we have $w_k \geq I_{X_{0k}} \geq I_{\bar{X}_{0k}} \geq I_{X_0}$ on X. In addition, from Theorem 5 we have $w_k \to I_{X_0}$ in L^1 norm on X; therefore

$$\lambda(X_0) = \int_X I_{X_0}\, d\lambda = \lim_{k \to \infty} \int_X w_k\, d\lambda \geq \lim_{k \to \infty} \int_X I_{X_{0k}}\, d\lambda$$
$$= \lim_{k \to \infty} \lambda(X_{0k}) \geq \lim_{k \to \infty} \lambda(\bar{X}_{0k}).$$

But since $X_0 \subset \bar{X}_{0k} \subset X_{0k}$ we must have $\lambda(X_0) \leq \lambda(\bar{X}_0) \leq \lambda(X_{0k})$ and the theorem follows. $\square$

8 Free final time

In this section we outline a straightforward extension of our approach to the problem of reaching the target set X_T at *any* time before $T < \infty$ (and not necessarily staying in X_T afterwards).

It turns out that the set of all initial states x_0 from which it is possible to reach X_T at a time $t \leq T$ can be obtained as the support of an optimal solution μ_0^* to the problem

$$
\begin{aligned}
\sup \quad & \mu_0(X) \\
\text{s.t.} \quad & \mu_T = \mu_0 \otimes \delta_0 + \mathcal{L}'\mu \\
& \mu \geq 0,\ \lambda \geq \mu_0 \geq 0,\ \mu_T \geq 0 \\
& \operatorname{spt}\mu \subset [0,T] \times X \times U \\
& \operatorname{spt}\mu_0 \subset X,\ \operatorname{spt}\mu_T \subset [0,T] \times X_T,
\end{aligned}
\tag{25}
$$

where the supremum is over a vector of nonnegative measures $(\mu, \mu_0, \mu_T) \in M([0,T] \times X \times U) \times M(X) \times M([0,T] \times X_T)$. Note that the only difference to problem (14) is in the support constraints of the final measure μ_T.

The dual to this problem reads as

$$
\begin{aligned}
\inf \quad & \int_X w(x)\, d\lambda(x) \\
\text{s.t.} \quad & \mathcal{L}v(t,x,u) \leq 0, && \forall\, (t,x,u) \in [0,T] \times X \times U \\
& w(x) \geq v(0,x) + 1, && \forall\, x \in X \\
& v(t,x) \geq 0, && \forall\, (t,x) \in [0,T] \times X_T \\
& w(x) \geq 0, && \forall\, x \in X.
\end{aligned}
\tag{26}
$$

The only difference to problem (16) is in the third constraint which now requires that $v(t,x)$ is nonnegative on X_T for *all* $t \in [0,T]$.

All results from the previous sections hold with proofs being almost verbatim copies.

9 Numerical examples

In this section we present four examples of increasing complexity to illustrate our approach: a univariate uncontrolled cubic system, an uncontrolled Van der Pol oscillator, a double integrator and an acrobot. For numerical implementation, one can either use Gloptipoly 3 [21] to formulate the primal problem on measures and then extract the dual solution provided by a primal-dual SDP solver or formulate directly the dual SOS problem using, e.g., YALMIP [28] or SOSTOOLS [35]. As an SDP solver we used SeDuMi [34] for the first three examples and MOSEK for the last two examples. For computational purposes the problem data should be scaled such that the constraint sets are contained in, e.g., unit boxes or unit balls; in particular the time interval $[0, T]$ should be scaled to $[0, 1]$ by multiplying the vector field f by T. Computational aspects are further discussed in the Conclusion.

Whenever the approximations X_{0k} defined in (23) are monotonous (which is not guaranteed) we report these approximations (since then they are equal to the monotonous version $\bar{X}_{0k}$ defined in (24)); otherwise we report $\bar{X}_{0k}$.

9.1 Univariate cubic dynamics

Consider the system given by

$$\dot{x} = x(x - 0.5)(x + 0.5),$$

the constraint set $X = [-1, 1]$, the final time $T = 100$ and the target set $X_T = [-0.01, 0.01]$. The ROA can in this case be determined analytically as $X_0 = [-0.5, 0.5]$. Polynomial approximations to the ROA for degrees $d \in \{4, 8, 16, 32\}$ are shown in Figure 1. As expected the functional convergence of the polynomials to the discontinuous indicator function is rather slow; however, the set-wise convergence of the approximations X_{0k} (where $k = d/2$) is very fast as shown in Table 1. Note that the volume error is not monotonically decreasing – indeed what is guaranteed to decrease is the integral of the approximating polynomial $w(x)$, not the volume of X_{0k}. Taking the monotonically decreasing approximations $\bar{X}_{0k}$ defined in (24) would prevent the volume increase. Numerically, a better behavior is expected when using alternative polynomial bases (e.g., Chebyshev polynomials) instead of the monomials; see the conclusion for a discussion.

Table 1: Univariate cubic dynamics – relative volume error of the outer approximation to the ROA $X_0 = [-0.5, 0.5]$ as a function of the approximating polynomial degree.

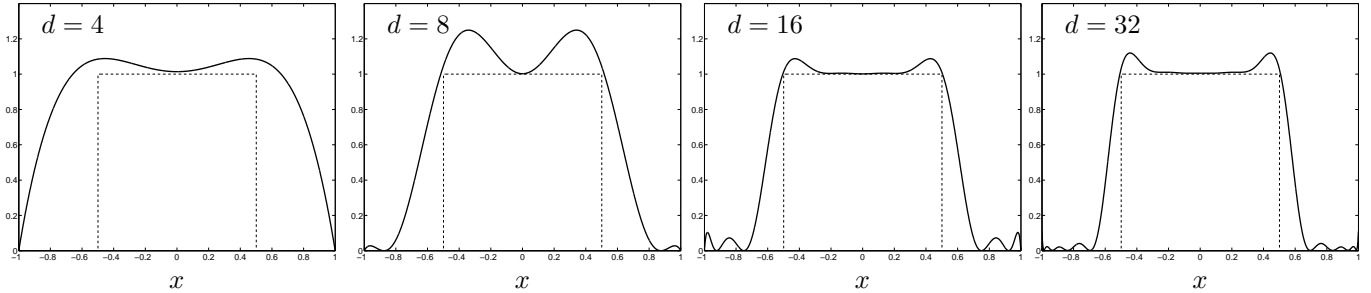| degree	4	8	16	32
error	31.60 %	3.31 %	0.92 %	1.49 %

Figure 1: Univariate cubic dynamics – polynomial approximations (solid line) to the ROA indicator function $I_{X_0} = I_{[-0.5,0.5]}$ (dashed line) for degrees $d \in \{4, 8, 16, 32\}$.

9.2 Van der Pol oscillator

As a second example consider a scaled version of the uncontrolled reversed-time Van der Pol oscillator given by

$$\dot{x}_1 = -2x_2,$$
$$\dot{x}_2 = 0.8x_1 + 10(x_1^2 - 0.21)x_2.$$

The system has one stable equilibrium at the origin with a bounded region of attraction

$$X_0 \subset X := [-1.2,\ 1.2]^2.$$

In order to compute an outer approximation to this region we take $T = 100$ and $X_T = \{x : \|x\|_2 \leq 0.01\}$. Plots of the ROA estimates X_{0k} for $d = 2k \in \{10, 12, 14, 16\}$ are shown in Figure 2. We observe a relatively-fast convergence of the super-level sets to the ROA – this is confirmed by the relative volume error[11] summarized in Table 2. Figure 3 then shows the approximating polynomial itself for degree $d = 18$. Here too, a better convergence is expected if instead of monomials, a more appropriate polynomial basis is used.

Table 2: Van der Pol oscillator – relative error of the outer approximation to the ROA X_0 as a function of the approximating polynomial degree.

degree	10	12	14	16
error	49.3 %	19.7 %	11.1 %	5.7 %

9.3 Double integrator

To demonstrate our approach in a controlled setting we first consider a double integrator

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = u.$$

The goal is to find an approximation to the set of all initial states X_0 that can be steered to the origin at[12] the final time $T = 1$. Therefore we set $X_T = \{0\}$ and the constraint set

---

[11]The relative volume error was computed approximately by Monte Carlo integration.

[12]In this case, the sets of all initial states that can be steered to the origin *at* time T and at any time *before* T are the same. Therefore we could also use the free-final-time approach of Section 8.

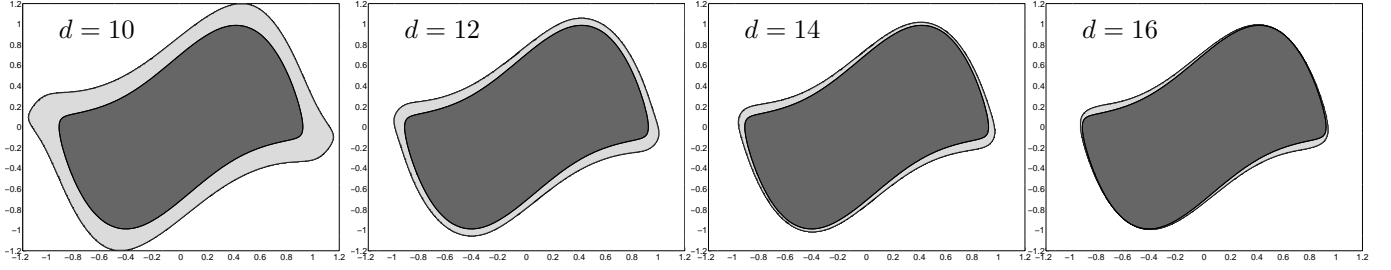Figure 2: Van der Pol oscillator – semialgebraic outer approximations (light gray) to the ROA (dark gray) for degrees $d \in \{10, 12, 14, 16\}$.
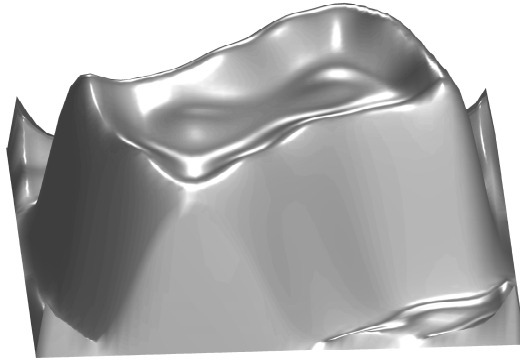

Figure 3: Van der Pol oscillator – a polynomial approximation of degree 18 of the ROA indicator function I_{X_0}.

such that $X_0 \subset X$, e.g., $X = [-0.7, 0.7] \times [-1.2, 1.2]$. The solution to this problem can be computed analytically as

$$X_0 = \{x \,:\, V(x) \leq 1\},$$

where

$$V(x) = \begin{cases} x_2 + 2\sqrt{x_1 + \frac{1}{2}x_2^2} & \text{if} \quad x_1 + \frac{1}{2}x_2|x_2| > 0, \\ -x_2 + 2\sqrt{-x_1 + \frac{1}{2}x_2^2} & \text{otherwise.} \end{cases}$$

The ROA estimates X_{0k} for $d = 2k \in \{6, 8, 10, 12\}$ are shown in Figure 4; again we observe a relatively fast convergence of the super-level set approximations, which is confirmed by the relative volume errors in Table 3.

Table 3: Double integrator – relative error of the outer approximation to the ROA X_0 as a function of the approximating polynomial degree.

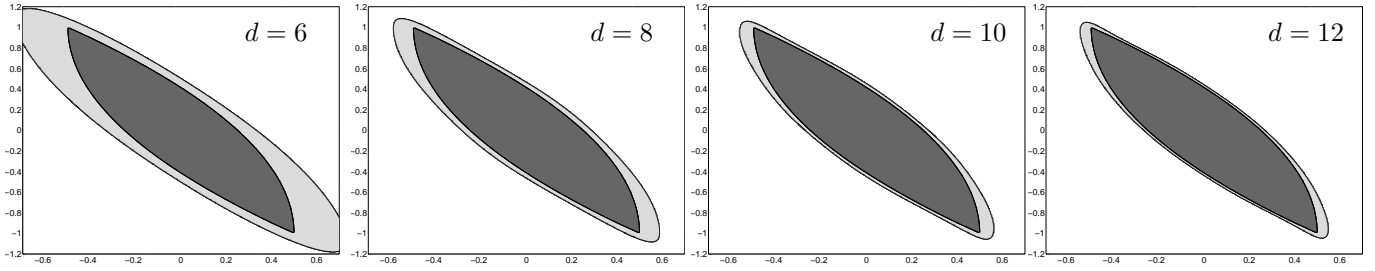| degree	6	8	10	12
error	75.7 %	32.6 %	21.2 %	16.0 %

Figure 4: Double integrator – semialgebraic outer approximations (light gray) to the ROA (dark gray) for degrees $d \in \{6, 8, 10, 12\}$.

9.4 Brocket integrator

Next, we consider the Brocket integrator

$$\dot{x}_1 = u_1$$
$$\dot{x}_2 = u_2$$
$$\dot{x}_3 = u_1 x_2 - u_2 x_1$$

with the constraint sets $X = \{x \in \mathbb{R}^3 : ||x||_\infty \leq 1\}$ and $U = \{u \in \mathbb{R}^2 : ||u||_2 \leq 1\}$, the target set $X_T = \{0\}$ and the final time $T = 1$. The ROA can be computed analytically (see [27]) as $X_0 = \{x \in \mathbb{R}^3 : T(x) \leq 1\}$, where

$$T(x) = \frac{\theta \sqrt{x_1^2 + x_2^2 + 2|x_3|}}{\sqrt{\theta + \sin^2 \theta - \sin \theta \cos \theta}},$$

and $\theta = \theta(x)$ is the unique solution in $[0, \pi)$ to

$$\frac{\theta - \sin \theta \cos \theta}{\sin^2 \theta}(x_1^2 + x_2^2) = 2|x_3|.$$

The ROA estimates X_{0k} are not monotonous in this case and therefore in Figure 5 we rather show the monotonous estimates $\bar{X}_{0k}$ defined in (24) for degrees six $d = 2k \in \{6, 10\}$. We observe fairly good tightness of the estimates.
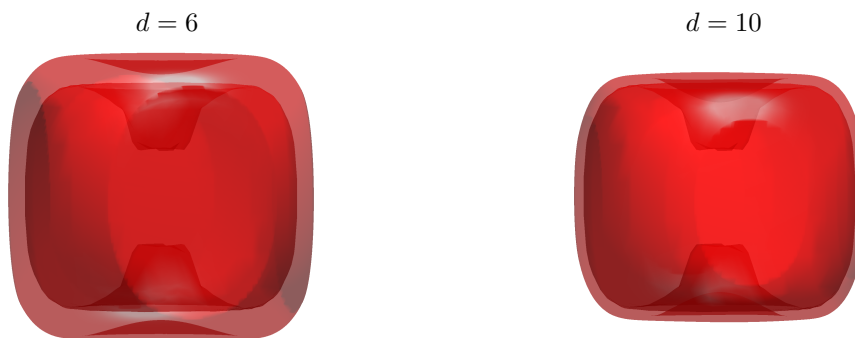

Figure 5: Brocket integrator – semialgebraic outer approximations (light red, larger) to the ROA (dark red, smaller) for degrees $d \in \{6, 10\}$.

9.5 Acrobot

As our last example we consider the acrobot system adapted from [33], which is essentially a double pendulum with both joints actuated; see Figure 6. The system equations are given by

$$\dot{x} = \begin{bmatrix} x_3 \\ x_4 \\ M(x)^{-1}N(x,u) \end{bmatrix} \in \mathbb{R}^4,$$

where

$$M(x) = \begin{bmatrix} 3 + \cos(x_2) & 1 + \cos(x_2) \\ 1 + \cos(x_2) & 1 \end{bmatrix}$$

and

$$N(x,u) = \begin{bmatrix} g\sin(x_1 + x_2) - a_1 x_3 + a_2 \sin(x_1) + x_4 \sin(x_2)(2x_3 + x_4) + u_1 \\ -\sin(x_2)x_3^2 - a_1 x_4 + g\sin(x_1 + x_2) + u_2 \end{bmatrix}$$

with $g = 9.8$, $a_1 = 0.1$ and $a_2 = 19.6$. The first two states are the joint angles (in radians) and the second two the corresponding angular velocities (in radians per second). The two control inputs are the torques in the two joints. Here, rather than comparing our approximations with the true ROA (which is not easily available), we study how the size of the ROA approximations is influenced by the actuation of the first joint. We consider two cases: with both joints actuated and with only the middle joint actuated. In the first case the input constraint set is $U = [-10, 10] \times [-10, 10]$ and in the second case it is $U = \{0\} \times [-10, 10]$. The state constraint set is for both cases $X = [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-5, 5] \times [-5, 5]$. Since this system is not polynomial we take a third order Taylor expansion of the vector field around the origin. An exact treatment would be possible via a coordinate transformation leading to rational dynamics to which our approach can be readily extended; this extension is, however, not treated in this paper and therefore we use the simpler (and non-exact) approach with Taylor expansion. Figure 7 shows the approximations X_{0k} of degree $d = 2k \in \{6, 8\}$; as expected disabling actuation of the first joint leads to a smaller ROA approximation. Contrary to previous examples, here we used the recently released MOSEK SDP solver, which on our problem class seems to perform much better than SeDuMi. This is documented by computation time[13] comparison in Table 4. Before solving, the problem data was scaled such that the constraint sets become unit boxes.
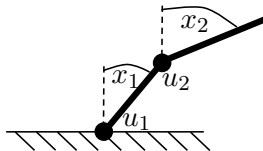

Figure 6: Acrobot – sketch

---

[13]Table 4 reports pure solver times, excluding the Yalmip parsing and preprocessing overhead, using Apple iMac with 3.4 GHz Intel Core i7, 8 GB RAM, Mac OS X 10.8.3 and Matlab 2012a.
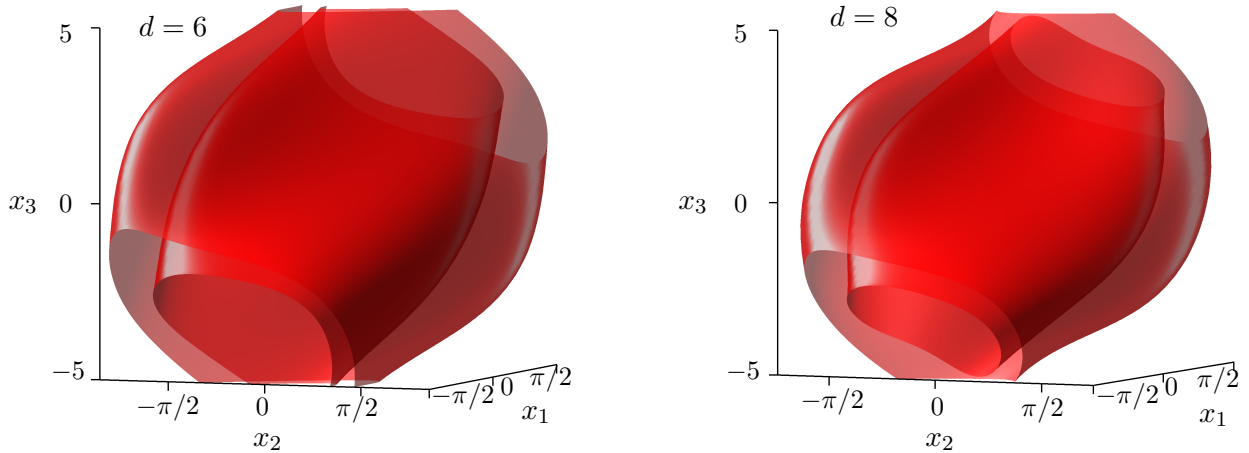
Figure 7: Acrobot – section for $x_4 = 0$ of the semialgebraic outer approximations of degree $d \in \{6, 8\}$. Only the middle joint actuated – darker, smaller; both joints actuated – lighter, larger. The states displayed x_1, x_2 and x_3 are, respectively, the lower pendulum angle, the upper pendulum angle and the lower pendulum angular velocity.

Table 4: Acrobot – comparison of computation time of MOSEK and SeDuMi for different degrees of the approximating polynomial. The "–" in the last cell signifies that SeDuMi could not solve the problem.

degree	4	6	8
MOSEK	0.93 s	23.5 s	2029 s
SeDuMi	7.1 s	2775 s	–

10 Conclusion

The main contributions of this paper can be summarized as follows:

- contrary to most of the existing systems control literature, we propose a convex formulation for the problem of computing the controlled region of attraction;

- our approach is constructive in the sense that we rely on standard hierarchies of finite-dimensional LMI relaxations whose convergence can be guaranteed theoretically and for which public-domain interfaces and solvers are available;

- we deal with polynomial dynamics and semialgebraic input and state constraints, therefore covering a broad class of nonlinear control systems;

- additional properties (e.g., convexity) of the approximations can be enforced by additional constraints on the polynomial (e.g., Hessian being negative definite).

- the approach extremely simple to use – the outer approximations are the outcome of a single semidefinite program with *no additional data* required besides the problem description.

24

The problem of computing the (forward) reachable set, i.e. the set of all states that can be reached from a given set of initial conditions under input and state constraints, can be addressed with the same techniques – the problem can be formulated as ROA computation using a time-reversal argument. Similar ideas can also be used to characterize and compute outer approximations of the maximum controlled invariant set (both for discrete and continuous time); this is a work in progress.

The hierarchy of LMI relaxations described in this paper generates a sequence of nested outer approximations of the ROA, but it is also possible, using a similar approach, to compute valid inner approximations. Results for uncontrolled systems will be reported elsewhere. The extension to the controlled case is far more involved (both theoretically and computationally) and is a subject of future research.

Furthermore, there is a straightforward extension to piecewise polynomial dynamics defined over a semialgebraic partition of the state and input spaces – one measure is then defined for each region of the partition. Our approach should also allow for extensions to stochastic systems (either discrete-time controlled Markov processes or controlled SDEs) and/or uncertain systems.

Since it is based on (an extension of) Lasserre's hierarchy of LMI relaxations (originally proposed for static polynomial optimization [26]), our approach scales similarly as Lasserre's approach. Namely, the number of moments (variables in the LMI relaxation) grows as $O(k^{n+m})$ when the problem dimension $n+m$ is kept constant and the relaxation order $k = d/2$ varies, and grows as $O((n+m)^k)$ when k is kept constant and $n+m$ varies. Therefore, at present, the approach is limited to systems of moderate size (say, $n+m \leq 6$) unless one is willing to compromise the accuracy of the approximations by taking a small relaxation order k. However, given the rapid progress of computing and optimization, the authors expect the approach to scale to larger dimensions in the future. One possible direction is sparsity exploitation; indeed, the recently released MOSEK SDP solver seems to have far superior performance to SeDuMi on our problem class, most likely due to more sophisticated sparsity exploitation. Another direction is parallelization; for instance, parallel interior point solvers (e.g., SDPARA [45]) should allow the approach to scale to larger dimensions. Alternative algorithms to standard primal-dual interior-point methods, such as first-order methods, are also worth considering. This is currently investigated by the authors.

Numerical examples indicate that the choice of monomials as a dense basis for the set of continuous functions on compact sets, while mathematically appropriate (and notationally convenient), is not always satisfactory regarding convergence and quality of the approximations. However, this is not peculiar to ROA computation problems – a similar behavior was already observed when computing the volume (and moments) of semialgebraic sets in [22]. To achieve better performance, we recommend the use of alternative polynomial bases such as Chebyshev polynomials; see [19, Section 4] for more details.

Appendix A

In this Appendix we state and prove the correspondence between the Liouville PDE on measures (10) and the convexified differential inclusion (12). We will need the notion of a

stochastic kernel. Let Y and Z be two Borel sets (in two Euclidean spaces of not necessarily the same dimension). The object $\nu(\cdot \,|\, \cdot)$ is called a stochastic kernel on Y given Z if, first, $\nu(\cdot \,|\, z)$ is a probability measure on Y for every fixed $z \in Z$, and, second, if $\nu(A \,|\, \cdot)$ is a measurable function on Z for every fixed $A \subset Y$. Let also $\bar{\mu}(t, x)$ denote the (t, x)-marginal of the occupation measure μ defined through (7), that is,

$$\bar{\mu}(A \times B) := \mu(A \times B \times U) \quad \forall\, A \subset [0, T],\ B \subset X.$$

Lemma 3 *Let (μ_0, μ, μ_T) be a triplet of measures satisfying the Liouville equation (10) such that* $\operatorname{spt} \mu_0 \subset X$, $\operatorname{spt} \mu \subset [0, T] \times X \times U$ *and* $\operatorname{spt} \mu_T \subset X_T$. *Then there exists a family of absolutely continuous admissible trajectories of (12) starting from μ_0 (i.e., trajectories in $\bar{\mathcal{X}}(x_0)$) such that the occupation measure and the terminal measure generated by this family of trajectories are equal to $\bar{\mu}$ and μ_T, respectively.*

Proof: Since the occupation measure μ is defined on a Euclidean space which is Polish and therefore Souslin, it can be, in view of [7, Corollary 10.4.13], disintegrated as

$$d\mu(t, x, u) = d\nu(u \,|\, t, x) d\bar{\mu}(t, x),$$

where $d\nu(u \,|\, t, x)$ is a stochastic kernel on U given $[0, T] \times X$. Then we can rewrite equation (9) as

$$
\begin{aligned}
&\int_{X_T} v(T, \cdot)\, d\mu_T - \int_X v(0, \cdot)\, d\mu_0 \\
&= \int_{[0,T] \times X} \int_U \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot f(t, x, u)\, d\nu(u \,|\, t, x)\, d\bar{\mu}(t, x) \\
&= \int_{[0,T] \times X} \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot \Big[\int_U f(t, x, u)\, d\nu(u \,|\, t, x) \Big] d\bar{\mu}(t, x) \\
&= \int_{[0,T] \times X} \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot \bar{f}(t, x)\, d\bar{\mu}(t, x),
\end{aligned}
\tag{27}
$$

where

$$\bar{f}(t, x) := \int_U f(t, x, u)\, d\nu(u \,|\, t, x) \in \operatorname{conv} f(t, x, U).$$

Therefore we will study the trajectories of the differential equation

$$\dot{x}(t) = \bar{f}(t, x(t)). \tag{28}$$

In the remainder of the proof we show that the measures μ_T and $\bar{\mu}$ are generated by a family of absolutely continuous trajectories of this differential equation (which is clearly a subset of trajectories of the convexified inclusion (12)) starting from μ_0. Note that the vector field $\bar{f}$ is only known to be measurable[14], so this equation may not admit a unique solution.

---

[14]Measurability of $\bar{f}(t, x)$ follows by first observing that for $f(t, x, u) = I_{A \times B \times C}(t, x, u)$ we have $\bar{f}(t, x) = I_A(t) I_B(x) \nu(C \,|\, t, x)$, which is a product of measurable functions, and then by approximating an arbitrary measurable $f(t, x, u)$ by simple functions (i.e., sums of indicator functions). This is a standard measure theoretic argument; details are omitted for brevity.

Observe that the t-marginal of μ (and hence of $\bar{\mu}$) is equal to the Lebesgue measure restricted to $[0, T]$ scaled by $\rho := \mu_0(X)$ $(=\mu_T(X))$. Indeed, plugging $v(t, x) = t^k$, $k \in \mathbb{N}$, in (9), we obtain $\mu_T(X) = \int t^k d\mu_0 + \int kt^{k-1} d\mu$; taking $k = 0$ gives $\mu_T(X) = \mu_0(X)$ and $k \geq 1$ gives $\int t^{k-1} d\mu = \mu_T(X)T^k/k$, which is nothing but the Lebesgue moments on $[0, T]$ scaled by $\mu_T(X) = \mu_0(X)$. Therefore, using [7, Theorem 6.4], we can disintegrate $\bar{\mu}$ as

$$d\bar{\mu}(t, x) = d\mu_t(x)dt, \tag{29}$$

where $d\mu_t(x)$ is a stochastic kernel on X given t scaled by ρ and dt is the standard Lebesgue measure on $[0, T]$. The kernel μ_t can be thought of as the distribution[15] of the state at time t. The kernel μ_t is defined uniquely dt-almost everywhere, and we will show that there is a version such that the function $t \mapsto \int_X w(x) \, d\mu_t(x)$ is absolutely continuous for all $w \in C^1(X)$ and such that the continuity equation

$$\frac{d}{dt} \int_X w(x) \, d\mu_t(x) = \int_X \operatorname{grad} w(x) \cdot \bar{f}(t, x) \, d\mu_t(x) \ \forall w \in C^1(X) \tag{30}$$

with the initial condition μ_0 is satisfied almost everywhere w.r.t. the Lebesgue measure on $[0, T]$.

Fix $w \in C^1(X)$ and define the test function $v(t, x) := \psi(t)w(x)$, where $\psi \in C^1([0, T])$. Then from equation (27)

$$\psi(T) \int_{X_T} w \, d\mu_T - \psi(0) \int_X w \, d\mu_0$$

$$= \int_{[0,T] \times X} \frac{\partial(\psi w)}{\partial t} + \operatorname{grad}(\psi w) \cdot \bar{f}(t, x) \, d\bar{\mu}(t, x)$$

$$= \int_0^T \int_X \dot{\psi}(t)w(x) + \psi(t)\operatorname{grad} w(x) \cdot \bar{f}(t, x) \, d\mu_t(x)dt$$

$$= \int_0^T \left[\dot{\psi} \int_X w \, d\mu_t + \psi \int_X \operatorname{grad} w \cdot \bar{f} \, d\mu_t \right] dt,$$

which can be seen as an equation of the form

$$\psi(T)d - \psi(0)c = \int_0^T \dot{\psi}(t)a(t) + \psi(t)b(t) \, dt \ \forall \psi \in C^1([0, T]), \tag{31}$$

where $c := \int_X w(x) \, d\mu_0(x)$, $d := \int_{X_T} w(x) \, d\mu_T$ and $b(t) := \int_X \operatorname{grad} w \cdot \bar{f}(t, x) \, d\mu_t(x)$ are constants and $a(t)$ is an unknown function. One solution is clearly $a(t) = \int_X w \, d\mu_t$. Now we show that

$$\tilde{a}(t) := c + \int_0^t b(\tau) \, d\tau = \int_X w \, d\mu_0 + \int_0^t \int_X \operatorname{grad} w \cdot \bar{f} \, d\mu_\tau d\tau$$

also solves the equation. Indeed, since from (27) with v replaced by w we have $\tilde{a}(T) = \int_X w \, d\mu_T = d$, integration by parts gives

$$\int_0^T \dot{\psi}(t)\tilde{a}(t) \, dt = \psi(T)d - \psi(0)c - \int_0^T \psi(t)b(t) \, dt,$$

---

[15]It will become clear from the following discussion that for $t = 0$ and $t = T$ this kernel (or a version thereof) coincides with μ_0 and μ_T, respectively; hence there is no ambiguity in notation. Note also that the kernel μ_t, $t \in [0, T]$, is defined uniquely up to a subset of $[0, T]$ of Lebesgue measure zero; by a "version" we then mean a particular choice of the kernel.

so $\tilde{a}(t)$ indeed solves equation (31). Now we prove that this solution is unique. Since $\tilde{a}$ is a solution we have

$$\psi(T)d - \psi(0)c = \int_0^T \dot{\psi}(t)\tilde{a}(t) + \psi(t)b(t)\,dt,$$

and subtracting this from (31) we get

$$0 = \int_0^T \dot{\psi}(t)[a(t) - \tilde{a}(t)]\,dt \quad \forall\,\psi \in C^1([0,T]),$$

or equivalently

$$0 = \int_0^T \phi(t)[a(t) - \tilde{a}(t)]\,dt \quad \forall\,\phi \in C([0,T]).$$

Since $C([0,T])$ is dense in $L^1([0,T])$, this implies $a(t) = \tilde{a}(t)$ dt-almost everywhere. Consequently, since $C^1(X)$ is separable,

$$\int_X w\,d\mu_t = \int_X w(x)d\mu_0 + \int_0^t \int_X \operatorname{grad} w \cdot \bar{f}\,d\mu_\tau d\tau \quad \forall\,w \in C^1(X) \tag{32}$$

dt-almost everywhere. The right-hand side of this equality is an absolutely continuous function of time for each $w \in C^1(X)$ and the left-hand side is a bounded positive linear functional on $C(X)$ for all $t \in [0,T]$. By continuity of the right-hand-side of (32) with respect to time, this right-hand side is a bounded positive linear functional on $C^1(X)$ for all $t \in [0,T]$ and therefore can be uniquely extended to a bounded positive linear functional on $C(X)$ (since C^1 is dense in C). Therefore, for all $t \in [0,T]$ the right-hand side has a representing measure [39, Theorem 2.14] and hence there is a version of μ_t such that the equality (32) holds for *all* $t \in [0,T]$. With this version of μ_t the function $t \mapsto \int_X w(x)\,d\mu_t(x)$ is absolutely continuous and μ_t solves the continuity equation (30).

To finish the proof, we use [1, Theorem 3.2] which asserts the existence of a nonnegative measure σ on $C([0,T];\mathbb{R}^n)$ which corresponds to a family of absolutely continuous solutions to ODE (28) whose projection at each time $t \in [0,T]$ coincides with μ_t. More precisely, there is a nonnegative measure $\sigma \in M(C([0,T];\mathbb{R}^n))$ supported on a family of absolutely continuous solutions to ODE (28) such that for all measurable $w : \mathbb{R}^n \to \mathbb{R}$

$$\int_X w(x)\mu_t(x) = \int_{C([0,T];\mathbb{R}^n)} w(x(t))\,d\sigma(x(\cdot)) \quad \forall\,t \in [0,T]. \tag{33}$$

Using $I_{A\times B}(t,x) = I_A(t)I_B(x)$, it follows from (29) that

$$\bar{\mu}(A \times B) = \int_{[0,T]\times X} I_A(t)I_B(x)\,d\bar{\mu}(t,x) = \int_0^T I_A(t)\int_X I_B(x)\,d\mu_t(x)\,dt.$$

Therefore, using (33) with $w(x) = I_B(x)$ and Fubini's theorem [39, Theorem 8.8], we get

$$\bar{\mu}(A \times B) = \int_{C([0,T];\mathbb{R}^n)} \int_0^T I_{A\times B}(t,x(t))\,dt\,d\sigma(x(\cdot)),$$

and so the occupation measure of the family of trajectories coincides with $\bar{\mu}$. Clearly, the initial and the final measures of this family coincide with μ_0 and μ_T as well. As a result σ-almost all trajectories of this family are admissible. The proof is completed by discarding the null-set of trajectories that are not admissible, which does not change the measure σ and the generated measures $\bar{\mu}, \mu_0, \mu_T$. $\qquad\square$

Appendix B

In this Appendix we elaborate further on the discussion from Section 3.2 on the connection between the classical ROA and the relaxed ROA. Let us recall the definition of the classical ROA

$$X_0 := \big\{ x_0 \in X \ : \ \mathcal{X}(x_0) \neq \emptyset \big\},$$

where

$$\mathcal{X}(x_0) := \big\{ x(\cdot) \ : \dot{x}(t) \in f(t, x(t), U) \text{ a.e., } x(0) = x_0,$$
$$x(T) \in X_T, \ x(t) \in X \ \forall t \in [0, T] \big\}$$

and $x(\cdot)$ is required to be absolutely continuous. Similarly, recall the definition of the relaxed ROA

$$\bar{X}_0 := \big\{ x_0 \in X \ : \ \bar{\mathcal{X}}(x_0) \neq \emptyset \big\},$$

where

$$\bar{\mathcal{X}}(x_0) := \big\{ x(\cdot) \ : \dot{x}(t) \in \text{conv } f(t, x(t), U) \text{ a.e., } x(0) = x_0,$$
$$x(T) \in X_T, \ x(t) \in X \ \forall t \in [0, T] \big\}$$

with $x(\cdot)$ absolutely continuous. Obviously, it holds

$$X_0 \subset \bar{X}_0, \tag{34}$$

and the question is whether this inclusion is strict or not.

Denote $B_\epsilon(a) := \{ x \in \mathbb{R}^n \ : \ ||x - a||_2 < \epsilon \}$ and define the dilated constraint sets

$$X^\epsilon := X \oplus B_\epsilon(0) \quad \text{and} \quad X_T^\epsilon := X_T \oplus B_\epsilon(0),$$

where $\oplus$ denotes the Minkowski sum of two sets. Accordingly, the dilated ROA and the dilated relaxed ROA are

$$\begin{aligned} X_0^\epsilon &:= \big\{ x_0 \in X \ : \ \mathcal{X}^\epsilon(x_0) \neq \emptyset \big\}, \\ \bar{X}_0^\epsilon &:= \big\{ x_0 \in X \ : \ \bar{\mathcal{X}}^\epsilon(x_0) \neq \emptyset \big\}, \end{aligned}$$

where

$$\mathcal{X}^\epsilon(x_0) := \big\{ x(\cdot) \ : \dot{x}(t) \in f(t, x(t), U) \text{ a.e., } x(0) = x_0,$$
$$x(T) \in X_T^\epsilon, \ x(t) \in X^\epsilon \ \forall t \in [0, T] \big\},$$
$$\bar{\mathcal{X}}^\epsilon(x_0) := \big\{ x(\cdot) \ : \dot{x}(t) \in \text{conv } f(t, x(t), U) \text{ a.e., } x(0) = x_0,$$
$$x(T) \in X_T^\epsilon, \ x(t) \in X^\epsilon \ \forall t \in [0, T] \big\}$$

Since the constraint sets are compact and the vector field f Lipschitz, it follows from the equivalence between the trajectories of the convexified inclusion (12) and solutions to the Liouville equation (10), stated in Lemma 3 of Appendix A, and from Filippov-Ważewski's relaxation Theorem (see, e.g., [5]) that

$$\bar{X}_0 = \text{spt } \mu_0 \ \subset \ \bigcap_{\epsilon > 0} X_0^\epsilon.$$

In contrast, for all $\epsilon > 0$ it holds

$$X_0^\epsilon = \bar{X}_0^\epsilon.$$

In general inclusion (34) is strict. However, we argue that for most practical purposes the relaxed ROA $\bar{X}_0$ and the true ROA X_0 are the same. Indeed, for any $x_0 \in \bar{X}_0$ there exists a sequence of admissible control functions $u_k(\cdot)$ such that

$$\sup_{t \in [0,T]} \mathrm{dist}_X(x_k(t)) \to 0 \quad \text{and} \quad \mathrm{dist}_{X_T}(x_k(T))$$

as $k \to \infty$, where $x_k(\cdot)$ denotes the solution to the ODE (1) corresponding to the control function $u_k(\cdot)$, and $\mathrm{dist}_A(x) := \inf\{\|z - x\|_2 \, : \, z \in A\}$ denotes the distance to a set A.

Appendix C

In this Appendix we describe two contrived examples of control systems (1) for which the relaxed ROA $\bar{X}_0$ is strictly larger than the classical ROA X_0; see Appendix B for definitions.

Let $f(t, x, u) = u$, $U = \{-1, +1\}$, $X = X_T = \{0\}$ for, e.g., $T = 1$. Obviously there is no admissible trajectory in $\mathcal{X}(0)$, whereas there is a feasible triplet of measures satisfying (10) given by $\mu_0 = \delta_0$, $\mu_T = \delta_0$ and $\mu = \lambda_{[0,1]} \otimes \delta_0 \otimes \frac{1}{2}(\delta_{-1} + \delta_{+1})$, where $\lambda_{[0,1]}$ denotes the restriction of the Lebesgue measure to $[0, 1]$. Therefore in this case $X_0 = \emptyset \neq \bar{X}_0 = \{0\}$, but $\lambda(X_0) = \lambda(\bar{X}_0)$. Assumption 2 is therefore satisfied. Note that the relaxed solution corresponds to an infinitely fast chattering of the control input between -1 and $+1$ which can be arbitrarily closely approximated by chattering solutions of finite speed; the singleton constraint set X, however, renders such solutions infeasible.

Another example for which the gap (e.g., in volume) between $\bar{X}_0$ and X_0 can be as large as desired is the following. Consider $\dot{x} = u \in \mathbb{R}^2$ with

$$x \in X := B_r((-1 - r, 0)) \cup [-1, 1] \times \{0\} \cup B_1((+2, 0)) \subset \mathbb{R}^2$$

for a given $r > 0$ and with $u \in U := \{-1, 1\}^2$ and $X_T := B_1((+2, 0))$. That is, the constraint set consist of two balls (one of radius r and the other of radius 1) connected by a line; the target set is the ball of radius 1. Then $X_0 = X_T$ is strictly smaller than $\bar{X}_0 = X$, and $\lambda(X_0) = \pi$, whereas $\lambda(\bar{X}_0) = (1 + r^2)\pi$. Assumption 2 is therefore not satisfied for $r > 0$. In this example, regular solutions starting in the left ball cannot transverse the line to the right ball; this is, by contrast, possible for the relaxed solutions using an infinitely fast chattering.

Appendix D

In this appendix we prove Theorem 4. In order to prove the theorem we rewrite primal LMI problem (21) in a vectorized form as follows

$$
\begin{aligned}
p_k^* \quad = \quad \min \quad & \mathbf{c}'\mathbf{y} \\
\text{s.t.} \quad & \mathbf{A}\mathbf{y} = \mathbf{b} \\
& \mathbf{e} + \mathbf{D}\mathbf{y} \in \mathbf{K},
\end{aligned}
\tag{35}
$$

where $\mathbf{y} := [y', y_0', y_T', \hat{y}_0']'$ and $\mathbf{K}$ is a direct product of cones of positive semidefinite matrices of appropriate dimensions, here corresponding to the moment matrix and localizing matrix constraints. The notation $\mathbf{e} + \mathbf{D}\mathbf{y} \in \mathbf{K}$ means that vector $\mathbf{e} + \mathbf{D}\mathbf{y}$ contains entries of positive semidefinite moment and localizing matrices, and by construction matrix $\mathbf{D}$ has full column rank (since a moment matrix is zero if and only if the corresponding moment vector is zero). Dual LMI problem (22) then becomes

$$
\begin{aligned}
d_k^* \ = \ \max \quad & \mathbf{b}'\mathbf{x} - \mathbf{e}'\mathbf{z} \\
\text{s.t.} \quad & \mathbf{A}'\mathbf{x} + \mathbf{D}'\mathbf{z} = \mathbf{c} \\
& \mathbf{z} \in \mathbf{K},
\end{aligned}
\tag{36}
$$

and we want to prove that $p_k^* = d_k^*$. The following instrumental result is a minor extension of a classical lemma of the alternatives for primal LMI (35) and dual LMI (36). The notation int $\mathbf{K}$ stands for the interior of $\mathbf{K}$.

Lemma 4 *If matrix* $\mathbf{D}$ *has full column rank, exactly one of these statements is true:*

- *there exists* $\mathbf{x}$ *and* $\mathbf{z} \in$ int $\mathbf{K}$ *such that* $\mathbf{A}'\mathbf{x} + \mathbf{D}'\mathbf{z} = \mathbf{c}$

- *there exists* $\mathbf{y} \neq 0$ *such that* $\mathbf{A}\mathbf{y} = 0$, $\mathbf{D}\mathbf{y} \in \mathbf{K}$ *and* $\mathbf{c}'\mathbf{y} \leq 0$.

Proof of Lemma 4: A classical lemma of alternatives states that if matrix $\bar{\mathbf{D}}$ has full column rank, then either there exists $\mathbf{z} \in$ int $\mathbf{K}$ such that $\bar{\mathbf{D}}'\mathbf{z} = \bar{\mathbf{c}}$ or there exists $\bar{\mathbf{y}}$ such that $\bar{\mathbf{D}}\bar{\mathbf{y}} \in \mathbf{K}$ and $\bar{\mathbf{c}}'\bar{\mathbf{y}} \leq 0$, but not both, see e.g. [41, Lemma 2] for a standard proof based on the geometric form of the Hahn-Banach separation theorem. Our proof then follows from restricting this lemma of alternatives to the null-space of matrix $\mathbf{A}$. More explicitly, there exists $\mathbf{x}$ and $\mathbf{z}$ such that $\mathbf{A}'\mathbf{x} + \mathbf{D}'\mathbf{z} = \mathbf{c}$ if and only if $\mathbf{z}$ is such that $\bar{\mathbf{D}}'\mathbf{z} = \bar{\mathbf{c}}$ with $\bar{\mathbf{D}} = \mathbf{D}\mathbf{F}$, $\bar{\mathbf{c}} = \mathbf{F}'\mathbf{c}$ for $\mathbf{F}$ a full-rank matrix such that $\mathbf{A}\mathbf{F} = 0$. Matrix $\bar{\mathbf{D}}$ has full column rank since it is the restriction of the full column rank matrix $\mathbf{D}$ to the null-space of $\mathbf{A}$. ∎

Proof of Theorem 4: First notice that the feasibility set of LMI problem (35) is nonempty and bounded. Indeed, a triplet of zero measures is a trivial feasible point for (14) and hence $(0, 0, 0, \lambda)$ is feasible in (15); consequently a concatenation of truncated moment sequences corresponding to the quadruplet of measures $(0, 0, 0, \lambda)$ is feasible in (35) for each relaxation order k. Boundedness of the even components of each moment vector follows from the structure of the localizing matrices corresponding to the functions from Assumption 3 and from the fact that the masses (zero-th moments) of the measures are bounded because of the constraint $\mu_0 + \hat{\mu}_0 = \lambda$ and because $T < \infty$. Boundedness of the whole moment vectors then follows since the even moments appear on the diagonal of the positive semidefinite moment matrices.

To complete the proof, we follow [41, Theorem 4] and show that boundedness of the feasibility set of LMI problem (35) implies existence of an interior point for LMI problem (36), and then from standard SDP duality it follows readily that $p^* = d^*$ since D has a full column rank; see, e.g., [41, Theorem 5] and references therein.

Let $\bar{\mathbf{y}}$ denote a point in the feasibility set of LMI problem (35), i.e. a vector satisfying $\mathbf{A}\bar{\mathbf{y}} = \mathbf{b}$ and $\mathbf{e} + \mathbf{D}\bar{\mathbf{y}} \in \mathbf{K}$. Suppose that there is no interior point for LMI problem (36),

31

i.e. there is no $\mathbf{x}$ and $\mathbf{z} \in \text{int } \mathbf{K}$ such that $\mathbf{A}'\mathbf{x} + \mathbf{D}'\mathbf{z} = \mathbf{c}$. Then from Lemma 4 there exists $\mathbf{y} \neq 0$ such that $\mathbf{Ay} = 0$, $\mathbf{Dy} \in \mathbf{K}$ and $\mathbf{c}'\mathbf{y} \leq 0$. It follows that there exists a ray $\bar{\mathbf{y}} + t\mathbf{y}$, $t \geq 0$ of feasible points for LMI problem (35), hence implying that the feasibility set is not bounded. $\blacksquare$

Acknowledgments

References

[1] L. Ambrosio. Transport equation and Cauchy problem for non-smooth vector fields. In L. Ambrosio et al. (eds.), Calculus of variations and nonlinear partial differential equations. Lecture Notes in Mathematics, Vol. 1927, Springer-Verlag, Berlin, 2008.

[2] E. J. Anderson, P. Nash. Linear programming in infinite-dimensional spaces: theory and applications. Wiley, New York, NY, 1987.

[3] R. B. Ash. Real analysis and probability. Academic Press, San Diego, CA, 1972.

[4] J. P. Aubin, A. M. Bayen, P. Saint-Pierre. Viability Theory: New Directions. Springer-Verlag, Berlin, 2011.

[5] J. P. Aubin, H. Frankowska. Set-valued analysis. Springer-Verlag, Berlin, 1990.

[6] F. Blanchini. Set invariance in control. Automatica, 35: 1747-1767, 1999.

[7] V. I. Bogachev. Measure Theory, volume 2. Springer, 2006.

[8] G. Chesi. Domain of attraction; analysis and control via SOS programming. Lecture Notes in Control and Information Sciences, Vol. 415, Springer-Verlag, Berlin, 2011.

[9] G. Chesi. LMI techniques for optimization over polynomials in control: a survey, IEEE Transactions on Automatic Control, 55: 2500-2510, 2010.

[10] M. Dellnitz, G. Froyland, O. Junge. The algorithms behind GAIO – Set oriented numerical methods for dynamical systems. In B. Fiedler (ed.), Ergodic theory, analysis, and efficient simulation of dynamical systems, Springer-Verlag, Berlin, 2001.

[11] L. C. Evans. Partial differential equations. 2nd edition, Amer. Math. Society, Providence, NJ, 2010.

[12] W. Fleming and D. Vermes. Convex duality approach to the optimal control of diffusions. SIAM J. Control Optimization, 36(2):1136-1155, 1989.

[13] H. Frankowska, F. Rampazzo. Filippov's and Filippov-Ważewski's theorems on closed domains. Journal of Differential Equations, 161:449-478, 2000.

[14] E. G. Gilbert and K. T. Tan. Linear systems with state and control constraints: the theory and application of maximal output admissible sets. IEEE Transactions on Automatic Control, 36: 1008-1020, 1991.

[15] V. Gaitsgory, M. Quincampoix. Linear programming approach to deterministic infinite horizon optimal control problems with discounting. SIAM J. on Control and Optimization, 48:2480-2512, 2009.

[16] R. Genesio, M. Tartaglia, A. Vicino. On the estimation of asymptotic stability regions: state of the art and new proposals. IEEE Transactions on Automatic Control, 30:747–755, 1985.

[17] W. M. Haddad, V. Chellaboina. Nonlinear dynamical systems and control, a Lyapunov-based approach. Princeton Univ. Press, Princeton, NJ, 2008.

[18] W. Hahn. Stability of motion. Springer-Verlag, Berlin, 1967.

[19] D. Henrion. Semidefinite characterisation of invariant measures for one-dimensional discrete dynamical systems. LAAS-CNRS Research Report 12052, February 2012. To appear in Kybernetika, 2012.

[20] D. Henrion, J. B. Lasserre. Inner approximations for polynomial matrix inequalities and robust stability regions. IEEE Transactions on Automatic Control, 57:1456-1467, 2012.

[21] D. Henrion, J. B. Lasserre, and J. Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. Optimization Methods and Software, 24:761–779, 2009. URL: `http://homepages.laas.fr/henrion/software/gloptipoly/`.

[22] D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. SIAM Review, 51:722-743, 2009.

[23] O. Hernandez-Lerma, J. B. Lasserre. Discretre-Time Markov Control Processes: Basic Optimality Criteria. Springer, 1996.

[24] M. W. Hirsch. Differential topology. Springer-Verlag, Berlin, 1976.

[25] A. B. Kurzhanski, P. Varaiya. Ellipsoidal techniques for reachability analysis. In Hybrid systems: computation and control (B. Krogh and N. Lynch, eds.), LNCIS 1790, Springer-Verlag, Berlin, 2000.

[26] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, London, UK, 2009.

[27] J. B. Lasserre, D. Henrion, C. Prieur, E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. SIAM Journal on Control and Optimization, 47:1643-1666, 2008.

[28] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MAT-LAB. In Proc. IEEE CCA/ISIC/CACSD Conference, Taipei, Taiwan, 2004. URL: `http://users.isy.liu.se/johanl/yalmip/`.

[29] D. G. Luenberger. Optimization by vector space methods. Wiley, New York, NY, 1969.

[30] K. Margellos, J. Lygeros. Hamilton-Jacobi formulation for reach-avoid differential games. IEEE Transactions on Automatic Control, 56:1849-1861, 2011.

[31] S. G. Margolis, W. G. Vogt. Control engineering applications of V. I. Zubov's construction procedure for Lyapunov functions. IEEE Transactions on Automatic Control, 8:101-113, 1963.

[32] I. Mitchell, C. Tomlin. Overapproximating reachable sets by Hamilton-Jacobi projections. Journal of Scientific Computing, 19:323-346, 2003.

[33] R. M. Murray, J. Hauser. A Case Study in Approximate Linearization: The Acrobot Example. EECS Department, University of California, Berkeley. Memorandum No. UCB/ERL M91/46, 1991.

[34] I. Pólik, T. Terlaky, and Y. Zinchenko. SeDuMi: a package for conic optimization. IMA workshop on Optimization and Control, Univ. Minnesota, Minneapolis, Jan. 2007. URL: `http://sedumi.ie.lehigh.edu/`.

[35] S. Prajna, A. Papachristodoulou, P. Seiler, P. A. Parrilo. SOSTOOLS: sum of squares optimization toolbox for MATLAB, 2004. URL: `http://www.cds.caltech.edu/sostools/`.

[36] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana University Mathematics Journal, 42:969-984, 1993.

[37] A. Rantzer. A dual to Lyapunov's stability theorem. Systems & Control Letters, 42:161-168, 2001.

[38] J. E. Rubio. Control and Optimization: The Linear Treatment of Nonlinear Problems. Manchester University Press, Manchester, UK, 1985.

[39] W. Rudin. Real and Complex Analysis. McGraw-Hill Science, 1986.

[40] U. Topcu, A. K. Packard, P. Seiler, G. J. Balas. Robust region-of-attraction estimation. IEEE Transactions on Automatic Control, 55:137-142, 2010.

[41] M. Trnovská. Strong duality conditions in semidefinite programming. Journal of Electrical Engineering (Bratislava), 56:1-5, 2005.

[42] U. Vaidya, P. G. Mehta. Lyapunov Measure for Almost Everywhere Stability. IEEE Transactions on Automatic Control, 53: 307-323, 2008.

[43] A. Vannelli, M. Vidyasagar. Maximal Lyapunov functions and domains of attraction for autonomous nonlinear systems. Automatica, 21:69-80, 1985.

[44] K. Wang, U. Vaidya. Transfer operator approach for computing domain of attraction. IEEE Conference on Decision and Control (CDC), Atlanta, GA, 2010.

[45] M. Yamashita, K. Fujisawa, M. Kojima. SDPARA: SemiDefinite Programming Algorithm paRAllel version. Parallel Computing 29:1053-1067, 2003.

[46] L. C. Young. Calculus of variations and optimal control theory. Sunders, Philadelphia,

[47] V. I. Zubov. Methods of A. M. Lyapunov and their application. Noordhoff, Groningen, 1964.