Controller design and value function approximation for nonlinear dynamical systems

Milan Korda[1], Didier Henrion[2,3,4], Colin N. Jones[1]

Draft of March 27, 2015

Abstract

This work considers the infinite-time discounted optimal control problem for continuous time input-affine polynomial dynamical systems subject to polynomial state and box input constraints. We propose a sequence of sum-of-squares (SOS) approximations of this problem obtained by first lifting the original problem into the space of measures with continuous densities and then restricting these densities to polynomials. These approximations are tightenings, rather than relaxations, of the original problem and provide a sequence of rational controllers with value functions associated to these controllers converging (under some technical assumptions) to the value function of the original problem. In addition, we describe a method to obtain polynomial approximations from above and from below to the value function of the extracted rational controllers, and a method to obtain approximations from below to the optimal value function of the original problem, thereby obtaining a sequence of asymptotically optimal rational controllers with explicit estimates of suboptimality. Numerical examples demonstrate the approach.

Keywords: Optimal control, nonlinear control, sum-of-squares, semidefinite programming, occupation measures, value function approximation

1 Introduction

This paper considers the infinite-time discounted optimal control problem for continuous-time input-affine polynomial dynamical systems subject to polynomial state constraints and box input constraints. This problem has a long history in both control and economics

[1]Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, Station 9, CH-1015, Lausanne, Switzerland. {`milan.korda,colin.jones`}`@epfl.ch`

[2]CNRS; LAAS; 7 avenue du colonel Roche, F-31400 Toulouse; France. `henrion@laas.fr`

[3]Université de Toulouse; LAAS; F-31400 Toulouse; France.

[4]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic.

literature. Various methods to tackle this problem have been developed, often based on the analysis of the associated Hamilton-Jacobi-Bellman equation.

In this work we take a different approach: We first lift the problem into an infinite-dimensional space of measures with continuous densities where this problem becomes convex; in fact a linear program (LP). This lifting is a *tightening*, i.e., its optimal value is greater than or equal to the optimal value of the original problem, and under suitable technical conditions the two optimal values coincide. This infinite-dimensional LP is then further tightened by restricting the class of functions to polynomials of a prescribed degree and replacing nonnegativity constraints by sufficient sum-of-squares (SOS) constraints. This leads to a hierarchy of semidefinite programming (SDP) tightenings of the original problem indexed by the degree of the polynomials. The solutions to the SDPs yield immediately a sequence of *rational* controllers, and we prove that, under suitable technical assumptions, the value functions associated to these controllers converge *from above* to the value function of the original problem.

We also describe how to obtain a sequence of polynomial approximations converging from above and from below to the value function associated to each rational controller. Combined with existing techniques to obtain polynomial under approximations of the value function of the original problem (adapted to our setting), this method can be viewed as a design tool providing a sequence of rational controllers asymptotically optimal in the original problem with explicit estimates of suboptimality in each step.

The idea of lifting a nonlinear problem to an infinite-dimensional space dates back at least to the work of L. C. Young [25] and subsequent works of Warga [26], Vinter and Lewis [24], Rubio [22] and many others, both in deterministic and stochastic settings. These works typically lift the original problem into the space of measures and this lifting is a *relaxation* (i.e., its optimal value is less than or equal to the optimal value of the original problem) and under suitable conditions the two values coincide.

More recently, this infinite-dimensional lifting was utilized numerically by *relaxing* the infinite-dimensional LP into a finite-dimensional SDP [13] or finite-dimensional LP [4]. Whereas the LP relaxations are obtained by classical state- and control-space gridding, the SDP relaxations are obtained by optimizing over truncated moment sequences (i.e., involving only finitely many moments) of the measures and imposing conditions *necessary* for these truncated moment sequences to be feasible in the infinite-dimensional lifted problem. These finite-dimensional relaxations provide *lower bounds* on the value function of the optimal control problem and seem to be difficult to use for control design with strong convergence guarantees; a controller extraction from the relaxations is possible although no convergence (e.g., [5, 4]) or only very weak convergence can be established (e.g., [8, 15] in the related context of region of attraction approximation).

Contrary to these works, in this approach we tighten the infinite-dimensional LP by optimizing over polynomial densities of the measures and imposing conditions *sufficient* for these densities to be feasible in the infinite-dimensional lifted problem, thereby obtaining upper bounds as opposed to lower bounds. Crucially, to ensure that polynomial densities of arbitrarily low degrees exist for our problem (and therefore the resulting SDP tightenings are feasible), we work with free initial and final measures and set up the cost function and constraints such that this additional freedom does not affect optimality. Importantly, we

2

do *not* assume that the state constraint set is control invariant, a requirement that is often imposed in the existing literature (e.g., [19]) but rarely met in practice.

The presented approach bears some similarity with the density approach of [21] for global stabilization later extended to optimal control (in a purely theoretical setting) in [19] and recently generalized to optimal stabilization of a given invariant set in [20] (providing both theoretical results and a practical computation method). However, contrary to [21] we consider the problem of optimal control, not stabilization and moreover we work under state constraints. Contrary to [20] we work in continuous time, consider a more general problem (optimal control, not optimal stabilization of a given set) and our approach of finite-dimensional approximation is completely different in the sense that it is based purely on convex optimization and it does not rely on state-space discretization. Moreover, and importantly, our approach comes with convergence guarantees.

Finally, let us mention that this work is inspired by [12], where a converging sequence of upper bounds on *static* polynomial optimization problems was proposed, as opposed to a converging sequence of lower bounds as originally developed in [11].

2 Preliminaries

2.1 Notation

We use $L(X;Y)$ to denote the space of all Lebesgue measurable functions defined on a set $X \subset \mathbb{R}^n$ and taking values in the set $Y \subset \mathbb{R}^m$. If the space Y is not specified it is understood to be $\mathbb{R}$. The spaces of integrable functions and essentially bounded functions are denoted by $L^1(X;Y)$ and $L^\infty(X;Y)$, respectively. The spaces of continuous respectively k-times continuously differentiable functions are denoted by $C(X;Y)$ respectively $C^k(X;Y)$. By a (Borel) measure we understand a countably additive mapping from (Borel) sets to nonnegative real numbers. Integration of a continuous function v with respect to a measure μ on a set X is denoted by $\int_X v(x)\,d\mu(x)$ or also $\int v\,d\mu$ when the variable and domain of integration are clear from the context. A probability measure is a measure with unit mass (i.e., $\int 1d\mu = 1$). The support of a measure μ, defined as the smallest closed set whose complement has zero measure, is denoted by spt μ. The ring of all multivariate polynomials in a variable x is denoted by $\mathbb{R}[x]$, the vector space of all polynomials of degree no more than d is denoted by $\mathbb{R}[x]_d$, and the vector space of m-dimensional polynomial vectors is denoted by $\mathbb{R}[x]^m$. The boundary of a set X is denoted by ∂X, the interior by X° and the closure by $\bar{X}$. The Euclidean distance of a point x from a set X is denoted by $\mathrm{dist}_X(x)$. For a possibly matrix-valued function $f \in C(X;\mathbb{R}^{n\times m})$ we define $\|f\|_{C^0(X)} := \sup_{x\in X} \max_{i,j} |f_{i,j}(x)|$ and for a vector-valued function $g \in C^1(X;\mathbb{R}^n)$ we define $\|g\|_{C^1(X)} := \|g\|_{C^0(X)} + \|\frac{\partial g}{\partial x}\|_{C^0(X)}$, where $\frac{\partial g}{\partial x}$ denotes the Jacobian of g. If clear from the context we write $\|\cdot\|_{C^0}$ for $\|\cdot\|_{C^0(X)}$ and similarly for the C^1 norm.

2.2 SOS programming

Crucial to the material presented in the paper is the ability to decide whether a polynomial $p \in \mathbb{R}[x]$ is nonnegative on a set

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, \ i = 1, \ldots, n_g\},$$

with $g_i \in \mathbb{R}[x]$. A *sufficient* condition for p to be nonnegative on X is that it belongs to the truncated quadratic module of degree d associated to X,

$$Q_d(X) := \left\{ s_0 + \sum_{i=1}^{n_a} g_i(x) s_i(x) \mid s_0 \in \Sigma_{2\lfloor \frac{d}{2} \rfloor}, s_i \in \Sigma_{2 \left\lfloor \frac{(d - \deg g_i)}{2} \right\rfloor} \right\},$$

where Σ_{2k} is the set of all polynomial sum-of-squares (SOS) of degree at most $2k$. Note in particular that $Q_{d+1}(X) \supset Q_d(X)$. If $p \in Q_d(X)$ for some $d \geq 0$ then clearly p is nonnegative on X, and the following fundamental result shows that a certain converse result holds.

Proposition 1 (Putinar [16]) *Let $N - \|x\|^2 \in Q_d(X)$ for some $d > 0$ and $N \geq 0$ and let $p \in \mathbb{R}[x]$ be strictly positive on X. Then $p \in Q_d(X)$ for some $d \geq 0$.*

Combining with the Stone-Weierstrass Theorem, as an immediate corollary we get:

Corollary 1 *Let $f \in C(X)$ be nonnegative on X and let $N - \|x\|^2 \in Q_d(X)$ for some $d > 0$ and $N \geq 0$. Then for every $\epsilon \geq 0$ there exists $d \geq 0$ and $p_d \in Q_d(X)$ such that $\|f - p_d\|_{C^0} < \epsilon$.*

Corollary 1 says that polynomials in $Q_d(X)$ are dense (with respect to the C^0 norm) in the space of continuous functions nonnegative on X when we let d tend to infinity.

In the rest of the text we use standard algebraic operations on sets. For instance if we write that $p \in gQ_d(X) + h\mathbb{R}[x]_d$, then it means that $p = gq + hr$ with $q \in Q_d(X)$ and $r \in \mathbb{R}[x]_d$.

The inclusion of $p \in Q_d(X)$ for a fixed d is equivalent to the existence of a positive semidefinite matrix W such that $p(x) = b(x)^\top W b(x)$, where $b(x)$ is a basis of $\mathbb{R}[x]_{d/2}$, the vector space of polynomials of degree at most $d/2$. Comparing coefficients leads to a set of affine constraints on the coefficients of p and the entries of W. Deciding whether $p \in Q_d(X)$ therefore translates to the feasibility of a semidefinite programming problem with the coefficients of p entering affinely. As a result, optimization of a linear function of the coefficients of p subject to the constraint $p \in Q_d(X)$ translates to a semidefinite programming problem (SDP) and hence to a well-understood and widely studied class of convex optimization problems for which powerful algorithms and off-the-shelf software are available. See, e.g., [10] and the references therein for more details.

3 Problem statement

We consider the continuous-time input-affine[1] controlled dynamical system

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} f_{u_i}(x(t))u_i(t), \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, and the data are polynomial: $f \in \mathbb{R}[x]^n$, $f_{u_i} \in \mathbb{R}[x]^n$, $i = 1, \ldots, m$. The system is subject to semi-algebraic state and box[2] input constraints

$$x(t) \in X := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, \ i = 1, \ldots, n_g\}, \tag{2a}$$
$$u(t) \in U := [0, \bar{u}]^m, \tag{2b}$$

where $g \in \mathbb{R}[x]^{n_g}$ and $\bar{u} \geq 0$. The set X is assumed compact and the polynomials defining X are assumed to be such that

$$\bar{g}(x) := \prod_{i=1}^{n_g} g_i(x) > 0 \quad \forall x \in X^\circ. \tag{3}$$

Since X is assumed compact, we also assume, without loss of generality, that the inequalities defining the sets X contain the inequality $N - \|x\|^2 \geq 0$ for some $N \geq 0$.

The goal of the paper is to (approximately) solve the following optimal control problem (OCP):

$$
\begin{aligned}
V(x_0) \ := \ &\inf_{u(\cdot), \tau(\cdot)} \quad \int_0^{\tau(x_0)} e^{-\beta t}[l_x(x(t)) + \sum_{i=1}^m l_{u_i}(x(t))u_i(t)] \, dt + e^{-\beta \tau} M \\
&\text{s.t.} \quad x(t) = x_0 + \int_0^t f(x(s)) + \sum_{i=1}^m f_{u_i}(x(s))u_i(s) \, ds, \\
&\qquad (x(t), u(t)) \in X \times U \quad \forall t \in [0, \tau(x_0)] \\
&\qquad u \in L^\infty([0, \tau(x_0)]; U), \ \tau \in L(X; [0, \infty])
\end{aligned}
\tag{4}
$$

where $\beta > 0$ is a given discount factor and M is a constant chosen such that

$$M > \beta^{-1} \sup_{x \in X, u \in U} \{l(x, u)\}, \tag{5}$$

where the joint stage cost

$$l(x, u) := l_x(x) + \sum_{i=1}^{m} l_{u_i}(x)u_i \tag{6}$$

is, without loss of generality, assumed to be nonnegative on $X \times U$. The state and input stage cost functions l_x and l_{u_i}, $i = 1, \ldots, m$, are assumed to be polynomial. The function τ

---

[1]Any dynamical system $\dot{x} = f(x, u)$ depending nonlinearly on u can be transformed to the input-affine form by using state inflation $\begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} f(x, u) \\ v \end{bmatrix}$, where u is now a part of the state and v a new control input; constraints on v then correspond to rate constraints on u. Similarly, cost functions depending non-linearly on u in problem (4) can be handled using state inflation in exactly the same fashion.

[2]Any box can be affinely transformed to $[0, \bar{u}]$.

in OCP (4) is referred to as a *stopping function*; the optimization is therefore both over the control input u and over the final time $\tau(x_0)$, which can be finite or infinite and can depend on the initial condition x_0.

The function $x \mapsto V(x)$ in (4) is called the *value function*. The reason for choosing the slightly non-standard objective function in (4) is because with this objective function the value function V is bounded (by M) on X and it coincides with the standard[3] discounted infinite-horizon value function for all initial conditions $x_0 \in X$ for which the trajectories can be kept within the state constraint set X forever using admissible controls, i.e., for all x_0 in the maximum control invariant set associated to the dynamics (1) and the constraints (2). To see the first claim, set $\tau(x_0) = 0$ for all $x_0 \in X$. To see the second claim notice that with M chosen as in (5), it is always beneficial to continue the time evolution whenever possible and therefore $\tau(x_0) = +\infty$ for all x_0 in the maximum controlled invariant set associated to (1) and (2).

Remark 1 *A constant M satisfying (5) can be found either by analytically evaluating the supremum in (5) or by using the techniques of [11] to find an upper bound.*

Given a Lipschitz continuous *feedback* controller $u \in C(X; U)$ and a stopping function $\tau \in L(X; [0, \infty])$, the ODE (1) has a unique solution and we let $V_{u,\tau} \in L(X; [0, \infty])$ denote the value function attained by (u, τ) in (4), i.e., setting $u(t) = u(x(t))$. By V_u we denote the value function $V_{u,\tau_u^\star}$, where $\tau_u^\star \in L(X; [0, \infty])$ is the optimal stopping function associated to u. Note that, by the choice of M in (5), the optimal stopping function $\tau_u^\star$ is equal to the first hitting time of the complement of the constraint set X, i.e.,

$$\tau_u^\star(x_0) = \inf\{t \geq 0 \mid x(t \mid x_0) \notin X\},$$

where $x(t \mid x_0)$ is the trajectory of (1) with $u(t) = u(x(t))$ starting from x_0. Notice also that $V_{u,\tau}(x) \geq V(x)$ for all $x \in X$ and that for any pair (u, τ) feasible in (4) we have $V_{u,\tau}(x) \leq M$ for all $x \in X$.

Throughout the paper, we make the following technical assumption:

Assumption 1 *There exists a sequence of Lipschitz continuous feedback controllers $\{u^k \in C(X; U)\}_{k=1}^\infty$ and stopping functions $\{\tau^k \in L(X; [0, \infty])\}_{k=1}^\infty$ feasible in (4) such that*

$$\lim_{k \to \infty} \int_X (V_{u^k, \tau^k}(x) - V(x)) dx = 0 \tag{7}$$

and such that for every $k \geq 0$ there exist a function $\rho^k \in C^1(X)$ and a scalar $\gamma^k > 0$ such that $\rho^k(x) = 0$ if $\mathrm{dist}_{\partial X}(x) < \gamma_k$ and

$$\int_X \int_0^{\tau^k(x_0)} e^{-\beta t} v(x^k(t \mid x_0)) \, dt \, dx_0 = \int_X v(x) \rho^k(x) \, dx \quad \forall v \in C(X), \tag{8}$$

where $x^k(\cdot \mid x_0)$ denotes the solution to (1) controlled by u^k.

---

[3]By standard we mean a discounted optimal control problem with cost $\int_0^\infty e^{-\beta t}[l_x(x(t)) + \sum_{i=1}^m l_{u_i}(x(t)) u_i(t)] \, dt$ and no stopping function.

Remark 2 *Note that $V_{u^k,\tau^k} \geq V$ on X by construction and therefore (7) is equivalent to the L^1 convergence of V_{u^k,τ^k} to V.*

Assumption 1 says that the optimal control inputs and stopping functions for OCP (4) can be well approximated by Lipschitz continuous *feedback* controllers and measurable stopping functions such that the resulting densities of the discounted occupation measures are continuously differentiable and vanish near the boundary of X. Note that the existence of an optimal feedback controller, as well as whether it can be well approximated by Lipschitz controllers, are subtle issues. Similarly it is a subtle issue whether asymptotically optimal stopping functions can be found such that the associated densities ρ^k in (8) are continuously differentiable and vanish near the boundary of X (note, however, that the left hand side of (8) can always be represented as $\int_X v(x)d\mu^k(x)$ for some nonnegative measure μ^k). This problem is of rather technical nature and has been studied in the literature (e.g., [3, Section 1.4] or [18]), where affirmative results have been established in related settings. We do not undertake a study of this problem here and rely on Assumption 1, which is, for ease of reading, not stated in its most general form. For example, the functions ρ^k do not need to be C^1 but only weakly differentiable and the integration on the left-hand side of (8) can be weighted by a nonnegative function $\rho_0^k \in L_1(X)$ satisfying $\rho_0^k \geq 1$ on X and $\rho_0^k \to 1$ in $L_1(X)$. In addition, we conjecture that it is enough to require $\rho^k = 0$ on ∂X and not necessarily on some neighborhood of ∂X; this is in particular the case when X is a box or a ball but we expect all the results of the paper to hold with a general semialgebraic set for which the defining functions satisfy (3).

The main result of this paper is a hierarchy of sum-of-squares (SOS) problems providing an explicit sequence of *rational* feedback controllers $u^k \in C^\infty(X;U)$ such that, under Assumption 1, (7) holds with $\tau^k = \tau^\star_{u^k}$, i.e., a sequence of asymptotically optimal rational controllers in the sense of the L^1 convergence of the associated value functions (see Remark 2).

4 Converging hierarchy of solutions

In this section we present an infinite-dimensional linear program (LP) in the space of continuous functions whose sum-of-squares (SOS) approximations provide a sequence of *rational* controllers u^k satisfying (7). This infinite-dimensional LP is closely related (and in a weak sense equivalent) to OCP (4); the rationale behind the derivation of the LP and its relation to OCP (4) is detailed in Section 5.

The infinite-dimensional LP reads

$$
\begin{aligned}
&\inf_{\rho,\rho_0,\rho_T,\sigma} && \int_X l_x(x)\rho(x)\,dx + \sum_{i=1}^m \int_X l_{u_i}(x)\sigma_i(x)\,dx + M\int_X \rho_T(x)\,dx \\
&\text{s.t.} && \rho_T - \rho_0 + \beta\rho + \operatorname{div}(\rho f) + \sum_{i=1}^m \operatorname{div}(\sigma_i f_{u_i}) = 0 \\
& && \rho \leq 0 && \text{on } \partial X \\
& && \rho_0 \geq 1 && \text{on } X \\
& && \bar{u}\rho \geq \sigma_i && \text{on } X, \ i=1,\dots,m. \\
& && \rho_T \geq 0 && \text{on } X \\
& && \sigma_i \geq 0 && \text{on } X, \ i=1,\dots,m.
\end{aligned}
\tag{9}
$$

The optimization in (9) is over functions $(\rho, \rho_0, \rho_T, \sigma) \in C^1(X) \times C(X) \times C(X) \times C^1(X)^m$ with $\sigma = (\sigma_1, \ldots, \sigma_m)$.

The optimal value of (9) will be denoted by $p^\star$. The value attained in (9) by any tuple of densities $(\rho, \rho_0, \rho_T, \sigma)$ feasible in (9) will be denoted by $p(\rho, \rho_0, \rho_T, \sigma)$.

Remark 3 (Non-uniform weighting) *Note that we could have imposed $\rho_0 \geq \bar{\rho}_0$ for any polynomial $\bar{\rho}_0$ nonnegative on X. Choosing a different $\bar{\rho}_0$ has no impact on the asymptotic convergence of the value functions established in the rest of the paper as long as $\bar{\rho}_0$ is strictly positive on X. It may, however, influence the speed of convergence in different subsets of X. In general we expect faster convergence where $\bar{\rho}_0$ is large and slower convergence where it is small. Choosing a non-constant $\bar{\rho}_0$ therefore allows to assign a different importance to different subsets of X.*

The infinite-dimensional LP (9) is then approximated by a hierarchy of sum-of-squares (SOS) problems, which immediately translate to finite-dimensional semidefinite programs (SDPs).

The SOS approximation of degree d of (9) reads

$$
\begin{aligned}
&\inf_{(\rho, \rho_0, \rho_T, \sigma) \in \mathbb{R}[x]_d^{3+m}} && \int_X l_x(x)\rho(x)\,dx + \sum_{i=1}^m \int_X l_{u_i}(x)\sigma_i(x)\,dx + M \int_X \rho_T(x)\,dx \\
&\text{s.t.} && \rho_T - \rho_0 + \beta\rho + \mathrm{div}(\rho f) + \sum_{i=1}^m \mathrm{div}(\sigma_i f_{u_i}) = 0 \\
& && -\rho \in Q_d(X) + g_i \mathbb{R}[x]_{d-\deg g_i} + \bar{g}\mathbb{R}[x]_{d-\deg \bar{g}} && i = 1, \ldots, n_g \\
& && \rho_0 - 1 \in Q_d(X) \\
& && \bar{u}\rho - \sigma_i \in Q_d(X) + \bar{g}Q_{d-\deg \bar{g}}(X) && i = 1, \ldots, m \\
& && \rho_T \in Q_d(X) \\
& && \sigma_i \in Q_d(X) + \bar{g}Q_{d-\deg \bar{g}}(X), && i = 1, \ldots, m.
\end{aligned}
\tag{10}
$$

Once a basis for $\mathbb{R}[x]_d$ is fixed (e.g., the standard monomial basis), the objective becomes linear in the coefficients of polynomials ρ, σ and ρ_T, and the equality constraint is imposed by equating the coefficients. The inclusions in the quadratic modules translate to semidefinite constraints and affine equality constraints; see Section 2.2. Optimization problem (10) therefore immediately translates to an SDP.

Remark 4 (Feasibility) *Trivially, any feasible solution to (10) is feasible in (9). Also, problem (10) is feasible for any $d \geq 0$. Indeed $(\rho, \rho_0, \rho_T, \sigma) = (0, 1, 1, 0)$ is always feasible in (10). See also Remark 6 below.*

If non-uniform weighting of initial conditions (see Remark 3) was required, the constraint $\rho_0 - 1 \in Q_d(X)$ would be replaced by $\rho_0 - \bar{\rho}_0 \in Q_d(X)$ for a polynomial weighting function $\bar{\rho}_0$ nonnegative on X.

Given an optimal solution $(\rho^d, \rho_0^d, \rho_T^d, \sigma^d)$ to (10), we define a rational control law u^d by

$$
u_i^d(x) := \frac{\sigma_i^d(x)}{\rho^d(x)} \quad \forall x \in X, \ i = 1, \ldots, m.
\tag{11}
$$

The main result of the paper is the following theorem stating that the controllers u^d are asymptotically optimal:

Theorem 1 *For all $d \geq 0$ we have $u^d(x) \in U$ for all $x \in X$ and if Assumption 1 holds, then*

$$\lim_{d \to \infty} \int_X (V_{u^d}(x) - V(x)) \, dx = 0, \tag{12}$$

that is, $V_{u^d} \to V$ in $L_1(X)$ (note that $V_{u^d} \geq V$ on X).

5 Rationale behind the LP formulation (9) and proof of the main Theorem 1

This section explains the rationale behind the LP problem (9) and its relation to the OCP (4) and gives the proof of Theorem 1. First, we lift the original problem (4) into the space of measures with nonnegative densities in $C(X)$; this lifting is problem (9). Next we tighten the problem by considering only polynomials of prescribed degree and with nonnegativity constraints enforced via SOS conditions; this is problem (10). Importantly, the lifting (9) is a *tightening* of the original problem (4) as show in Theorem 3 below. This is in contrast with [13] where the original problem was lifted into the space of measures and this lifting was a *relaxation.*

To be more concrete, observe that any initial measure μ_0, stopping function $\tau \in L(X; [0, \infty])$, and family of trajectories $\{x(\cdot \mid x_0)\}_{x_0 \in X}$ of (1) generated by a Lipschitz controller $u \in C(X; U)$ give rise to a triplet of measures defined by

$$\int_X v(x) d\mu(x) = \int_X \int_0^{\tau(x_0)} e^{-\beta t} v(x(t \mid x_0)) \, dt \, d\mu_0(x_0), \tag{13a}$$

$$\int_X v(x) d\mu_T(x) = \int_X e^{-\beta \tau(x_0)} v(x(\tau(x_0) \mid x_0)) \, d\mu_0(x_0), \tag{13b}$$

$$\int_X v(x) d\nu_i(x) = \int_X \int_0^{\tau(x_0)} e^{-\beta t} v(x(t \mid x_0)) u_i(x(t \mid x_0)) \, dt \, d\mu_0(x_0). \tag{13c}$$

The measure μ is called *discounted occupation measure*, the measure μ_T *terminal measure* and the measures ν_i, $i = 1, \ldots, m$, *control measures*. These measures satisfy the *discounted Liouville equation*

$$\int_X v \, d\mu_T(x) = \int_X v \, d\mu_0(x) + \int_X (\nabla v \cdot f - \beta v) \, d\mu(x) + \sum_{i=1}^m \int_X \nabla v \cdot f_{u_i} \, d\nu_i(x) \tag{14}$$

for all $v \in C^1(X)$. This follows by direct computation; see, e.g., [8]. Notice also that $d\nu_i(x) = u_i(x) d\mu(x)$, i.e., ν_i is absolutely continuous with respect to μ with Radon-Nikodým derivative equal to u_i.

Crucially, the converse statement is also true, although we have to go from stopping functions to stopping measures:

Theorem 2 (Superposition) *If measures μ, μ_0, μ_T and ν_i, $i = 1, \ldots, m$, satisfy (14) with $\text{spt} \, \mu_0 \subset X$, $\text{spt} \, \mu \subset X$ and $\text{spt} \, \mu_T \subset X$ and $d\nu_i = u_i d\mu$ for some Lipschitz $u \in C(X, U)$, then*

there exists an ensemble of probability measures (i.e., measures with unit mass) $\{\tau_{x_0}\}_{x_0 \in X}$ and an ensemble of trajectories $\{x(\cdot \mid x_0)\}_{x_0 \in X}$ of the system (1) controlled with $u(t) = u(x(t))$ such that $x(t \mid x_0) \in X$ for all $t \in \mathrm{spt}\,\tau_{x_0}$ and

$$\int_X v(x)\,d\mu_0(x) = \int_X v(x(0 \mid x_0))\,d\mu_0(x_0), \tag{15a}$$

$$\int_X v(x)\,d\mu(x) = \int_X \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t \mid x_0))\,dt\,d\tau_{x_0}(\tau)\,d\mu_0(x_0), \tag{15b}$$

$$\int_X v(x)\,d\mu_T(x) = \int_X \int_0^\infty e^{-\beta \tau} v(\tau(x_0))\,d\tau_{x_0}(\tau)\,d\mu_0(x_0), \tag{15c}$$

$$\int_X v(x)\,d\nu_i(x) = \int_X \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t \mid x_0)) u_i(x(t \mid x_0))\,dt\,d\tau_{x_0}(\tau)\,d\mu_0(x_0) \tag{15d}$$

for all $v \in C^1(X)$.

Proof: See Appendix B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 5 (Interpretation of Theorem 2) *Theorem 2 says that any measures satisfying (14) are generated by a superposition of the trajectories of the dynamical system $\dot{x} = f(x) + \sum_{i=1}^m f_{u_i}(x) u_i(x)$, where the superposition is over the final time of the trajectories. Note that there is a unique trajectory corresponding to each initial condition (since the vector field $f(x) + \sum_{i=1}^m f_{u_i}(x) u_i(x)$ is Lipschitz) but this unique trajectory can be stopped at multiple times (in fact at a whole continuum of times) allowing for superposition; this superposition is captured by the stopping measures $\{\tau_{x_0}\}_{x_0 \in X}$. For example, if the τ_{x_0} is a Dirac measure at a given time, then there is no superposition; if τ_{x_0} has a discrete distribution, then there is a superposition of finitely or countably many overlapping trajectories starting at x_0 stopped at different time instances; if τ_{x_0} has a continuous distribution then there is a superposition of a continuum of overlapping trajectories starting from x_0 stopped at different time instances.*

If in addition the measures μ_0, μ, μ_T satisfying the discounted Liouville equation (14) are absolutely continuous with respect to the Lebesgue measure with densities $\rho_0 \in C(X)$, $\rho \in C^1(X)$, $\rho_T \in C(X)$ such that $\rho = 0$ on ∂X, then these densities satisfy

$$\rho_T - \rho_0 + \beta\rho + \mathrm{div}(f\rho) + \sum_{i=1}^m \mathrm{div}(f_{u_i}\sigma_i) = 0 \tag{16}$$

with $\sigma_i = u_i\rho$, $i = 1, \ldots, m$. This follows directly by substituting $d\mu_0 = \rho_0 dx$, $d\mu = \rho dx$, $d\mu_T = \rho_T dx$ and $d\nu_i = u_i d\mu = u_i \rho dx = \sigma_i dx$ in (14) and using integration by parts. Equation (16) holds almost everywhere in X with a Lipschitz controller u, since Lipschitz functions are differentiable almost everywhere and the integration by parts formula applies to them, and everywhere with $u \in C^1(X; U)$.

Remark 6 (Role of the terminal measure) *An important feature of the SOS tightenings (10) is that they are feasible for arbitrarily low degrees (see Remark 4), which is crucial from a practical point of view and is not satisfied with other, more obvious, formulations*

(e.g., those not involving a stopping function in (4)); the reason for this is that, in the absence of a terminal measure (i.e., $\rho_T = 0$), the discounted Liouville equation (16) may not have a solution with a polynomial ρ even though ρ_0 and the dynamics are polynomial. Indeed, for example with $f = -x$, $f_{u_i} = 0$, $\beta = 1$, $\rho_0 = 1$ on $X = [-1, 1]$ and zero elsewhere, the only solution to (16) with $\rho_T = 0$ is $\rho(x) = -\ln(|x|)$.

Theorem 2 immediately enables us to prove a representation of the cost of problem (9) in terms of trajectories of (1).

Lemma 1 *If $(\rho, \rho_0, \rho_T, \sigma)$ is feasible in (9) and $u = \sigma/\rho$, then*

$$
p(\rho, \rho_0, \rho_T, \sigma) = \int_X \int_0^\infty \int_0^\tau e^{-\beta t} l_x(x(t\,|\,x_0))dt\, d\tau_{x_0}(\tau)\rho_0(x_0)\, dx_0
$$
$$
+ \sum_{i=1}^m \int_X \int_0^\infty \int_0^\tau e^{-\beta t} l_{u_i}(x(t\,|\,x_0))u_i(x(t\,|\,x_0))\, dt\, d\tau_{x_0}(\tau)\rho_0(x_0)dx_0
$$
$$
+ M \int_X \int_0^\infty e^{-\beta\tau}\, d\tau_{x_0}(\tau)\rho_0(x_0)dx_0, \tag{17}
$$

where $x(\cdot\,|\,x_0)$ are trajectories of (1) controlled by $u(t) = u(x(t))$ and τ_{x_0} are stopping probability measures with support $\operatorname{spt} \tau_{x_0}$ included in $[0, \infty]$. Moreover the state-control trajectories $x(\cdot\,|\,x_0)$ and $u(x(\cdot\,|\,x_0))$ are feasible in (4) in the sense that $x(t\,|\,x_0) \in X$ and $u(t\,|\,x_0) \in U$ for all $t \in \operatorname{spt} \tau_{x_0}$.

Proof: Let $(\rho, \rho_0, \rho_T, \sigma)$ be feasible in (9) and let $p(\rho, \rho_0, \rho_T, \sigma)$ denote the value attained by $(\rho, \rho_0, \rho_T, \sigma)$ in (9). The equality constraint of (9) is exactly (16). Since the constraint of (9) implies $\rho = 0$ on ∂X, equation (14) holds with $d\mu_0 = \rho_0 dx$, $d\mu = \rho dx$, $d\mu_T = \rho_T dx$ and $d\nu_i = u_i d\mu = u_i \rho dx = \sigma_i dx$, where $u_i = \frac{\sigma_i}{\rho} \in C^1(X; U)$, $i = 1, \ldots, m$. By Theorem 2 (setting $v(x) = l_x(x)$ in (15b), $v(x) = 1$ in (15c) and $v(x) = l_{u_i}(x)$ in (15d)) we obtain the result (noticing that the constraints of (9) imply that $u(x) \in U$ for all $x \in X$). $\qquad\square$

Corollary 2 *If $(\rho, \rho_0, \rho_T, \sigma)$ is feasible in (9) and $u = \sigma/\rho$, then*

$$
p(\rho, \rho_0, \rho_T, \sigma) \geq \int_X V_u(x_0)\rho_0(x_0)dx_0. \tag{18}
$$

If in addition the stopping measures $\{\tau_{x_0}\}_{x_0 \in X}$ in (17) are equal to the Dirac measures $\{\delta_{\tau(x_0)}\}_{x_0 \in X}$ for some stopping function $\tau \in L(X; [0, \infty])$, then

$$
p(\rho, \rho_0, \rho_T, \sigma) = \int_X V_{u,\tau}(x_0)\rho_0(x_0)dx_0. \tag{19}
$$

Proof: Let $(\rho, \rho_0, \rho_T, \sigma)$ be feasible in (9). Using Lemma 1, $p(\rho, \rho_0, \rho_T, \sigma)$ has representation (17), where the state-control trajectories in (17) are feasible in (4). Since the measures τ_{x_0} in (17) have unit mass for all $x_0 \in X$, we obtain (18). If $\tau_{x_0} = \delta_{\tau(x_0)}$ for some stopping function $\tau \in L(X; [0, \infty])$, then the integrals with respect to τ_{x_0} in (17) become evaluations at $\tau(x_0)$ and hence (19) holds. $\qquad\square$

Corollary 2 immediately implies that the problem (9) (and hence problem (10)) is a tightening of the original problem (4):

Theorem 3 *The optimal value of (9) of $p^\star$ satisfies*

$$p^* \geq \int_X V(x)\, dx. \tag{20}$$

Proof: Follows from Corollary 2 since $\rho_0 \geq 1$ and $V_u \geq V \geq 0$. $\qquad\square$

Now we are in a position to prove the following crucial lemma linking problems (4) and (9).

Lemma 2 *If $\{u^k \in C(X;U)\}_{k=1}^{\infty}$ and $\{\tau^k \in L(X;[0,\infty])\}_{k=1}^{\infty}$ are respectively sequences of controllers and stopping functions satisfying the conditions of Assumption 1, then the corresponding densities $\{\rho^k, \rho_0^k, \rho_T^k, \sigma^k\}_{k=1}^{\infty}$ with $\rho_0^k = 1$ are feasible in (9) and satisfy*

$$\lim_{k\to\infty} p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) = \int_X V(x_0)dx_0. \tag{21}$$

Conversely, if $\{\rho^k, \rho_0^k, \rho_T^k, \sigma^k\}_{k=1}^{\infty}$ is a sequence such that $\lim_{k\to\infty} p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) = p^\star$ and if Assumption (1) holds, then equation (7) holds with $u^k = \sigma^k/\rho^k$.

Proof: To prove the first part of the statement consider the controllers u^k, stopping functions τ^k and densities ρ^k from Assumption (1). Setting $\rho_0^k = 1$ and defining $\sigma_i^k := u_i^k \rho^k$ and $\rho_T^k := \rho_0^k - \beta\rho^k - \operatorname{div}(\rho^k f) - \sum_{i=1}^{m} \operatorname{div}(f_{u_i}\sigma_i^k) = 0$ we see that $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ satisfy (16) with $\rho^k = 0$ on ∂X. Therefore $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ are feasible in (9). In addition, in view of (8), the representation (17) holds with $\tau_{x_0} = \delta_{\tau^k(x_0)}$. Therefore by Lemma 1

$$p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) = \int_X V_{u^k,\tau^k}(x_0)\rho_0^k(x_0)dx_0$$

and hence (21) holds since $\{V_{u^k,\tau^k}\}_{k=1}^{\infty}$ satisfies (7) and $\rho_0^k = 1$ for all $k \geq 0$.

To prove the second part of the statement, let $\{\rho^k, \rho_0^k, \rho_T^k, \sigma^k\}_{k=1}^{\infty}$ be any sequence such that $\lim_{k\to\infty} p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) = p^\star$. Then this sequence satisfies (21) by Theorem 3 and by the first part of Lemma 2 just proven. Therefore (7) holds with $u^k := \sigma^k/\rho^k$ since

$$p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) \geq \int_X V_{u^k}(x_0)dx_0$$

by Corollary 2. $\qquad\square$

We will also need the following result showing that nonnegative C^1 functions vanishing on a neighborhood of ∂X can be approximated by polynomials in $Q_d(K)$ vanishing on ∂X.

Lemma 3 *Let $\rho \in C^1(X)$ such that $\rho \geq 0$ on X and $\rho = 0$ on $\{x \in X : \operatorname{dist}_{\partial X}(x) < \zeta\}$ for some $\zeta > 0$. Then for any $\epsilon > 0$ there exists $d \geq 0$ and a polynomial $p_d \in \bar{g}Q_{d-\deg\bar{g}}(X)$ such that*

$$\|\rho - p_d\|_{C^1(X)} < \epsilon$$

and $p_d = 0$ on ∂X.

Proof: Since $\bar{g} > 0$ on X°, we can factor $\rho = \bar{g}h$ with $h \in C^1(X)$ given by

$$h(x) := \begin{cases} \rho(x)/\bar{g}(x) & \text{if } \text{dist}_{\partial X}(x) \geq \zeta \\ 0 & \text{otherwise.} \end{cases}$$

Since polynomials are dense in C^1 there exists for every $\delta > 0$ a polynomial $\hat{h} > 0$ such that

$$\|\hat{h} - h\|_{C^1(X)} < \delta. \tag{22}$$

Applying Proposition 1 to $\hat{h}$ we see that there exists $\hat{p}_{\hat{d}} \in Q_{\hat{d}}(X)$ for some $\hat{d} \geq 0$ such that

$$\|\hat{h} - \hat{p}_{\hat{d}}\|_{C^1(X)} < \delta. \tag{23}$$

Defining $p_d := \hat{p}_{\hat{d}}\bar{g}$ we see that $p_d \in \bar{g}Q_{d-\deg\bar{g}}(X)$ with $d = \hat{d} + \deg(\bar{g})$ and that $p_d = 0$ on ∂X. Finally,

$$\|\rho - p_d\|_{C^0} = \|h\bar{g} - \hat{p}_{\hat{d}}\bar{g}\|_{C^0} \leq \|\bar{g}\|_{C^0}\|h - \hat{p}_{\hat{d}}\|_{C^0} < 2\delta\|\bar{g}\|_{C^0}$$

and

$$\begin{aligned} \|\nabla\rho - \nabla p_d\|_{C^0} &= \|\bar{g}\,\nabla h + h\,\nabla\bar{g} - \bar{g}\,\nabla\hat{p}_{\hat{d}} + \hat{p}_{\hat{d}}\,\nabla\bar{g}\|_{C^0} \\ &\leq \|\nabla\bar{g}\|_{C^0}\|h - \hat{p}_{\hat{d}}\|_{C^0} + \|\bar{g}\|_{C^0}\|\nabla h - \nabla\hat{p}_{\hat{d}}\|_{C^0} \\ &\leq 2\delta\big(\|\nabla\bar{g}\|_{C^0} + \|\bar{g}\|_{C^0}\big). \end{aligned}$$

Therefore choosing δ such that $2\delta\big(\|\nabla\bar{g}\|_{C^0} + 2\|\bar{g}\|_{C^0}\big) < \epsilon$ gives the desired result. $\qquad\square$

Now we are ready to prove our main result, Theorem 1.

Proof (of Theorem 1): Consider the sequences $\{u^k \in C^1(X;U)\}_{k=1}^\infty$, $\{\tau^k \in L(X;[0,\infty])\}_{k=1}^\infty$ from Assumption 1. By the first part of Lemma 2 the sequence of associated densities $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ generated by (u^k, τ^k) is feasible in (9) and satisfies (21). By Assumption 1, $\rho^k = 0$ and $\sigma^k = 0$ on $\{x \in X : \text{dist}_{\partial X}(x) < \gamma^k\}$ (since $\sigma^k = u^k\rho^k$) with $\gamma^k > 0$.

Hence by Lemma 3 there exist polynomial densities $\rho^{k,\text{pol}} \in \bar{g}Q_{d_k-\deg\bar{g}}(X)$, $\sigma^{k,\text{pol}} \in \bar{g}Q_{d_k-\deg\bar{g}}(X)^m$ for some degrees $d^k \geq 0$ such that

$$\|\rho^k - \rho^{k,\text{pol}}\|_{C^1(X)} < 1/k \tag{24}$$

$$\|\sigma_i^k - \sigma_i^{k,\text{pol}}\|_{C^1(X)} < 1/k \tag{25}$$

$\bar{u}\rho^{k,\text{pol}} - \sigma_i^{k,\text{pol}} \in \bar{g}Q_{d^k-\deg\bar{g}}(X)$ for all $i = 1,\ldots,m$ (since $u^k(x) \in U = [0,\bar{u}]^m$ for all $x \in X$ and hence $\bar{u}\rho^k \geq \sigma_i^k$ on X). Notice also that since $\rho^{k,\text{pol}} \in \bar{g}Q_{d^k-\deg\bar{g}}(X)$, we have $-\rho^{k,\text{pol}} \in \bar{g}\mathbb{R}_{d^k-\deg\bar{g}}$. Next, since $\rho_0^k \geq 1$ and $\rho_T^k \geq 0$, we can find, by Corollary 1, polynomial densities $\hat{\rho}_0^{k,\text{pol}} \in 1 + Q_{d^k}(X)$ and $\hat{\rho}_T^{k,\text{pol}} \in Q_{d^k}(X)$ such that

$$\|\rho_0^k - \hat{\rho}_0^{k,\text{pol}}\|_{C^0(X)} < 1/k, \tag{26}$$

$$\|\rho_T^k - \hat{\rho}_T^{k,\text{pol}}\|_{C^0(X)} < 1/k. \tag{27}$$

Since $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ satisfy the equality constraint of (9) we have

$$\hat{\rho}_T^{k,\mathrm{pol}} + \beta\rho^{k,\mathrm{pol}} - \hat{\rho}_0^{k,\mathrm{pol}} + \mathrm{div}(\rho^{k,\mathrm{pol}}f) + \sum_{i=1}^m \mathrm{div}(\sigma_i^{k,\mathrm{pol}}f_{u_i}) = \omega^k$$

where

$$\omega^k := \hat{\rho}_T^{k,\mathrm{pol}} - \rho_T^k + \beta(\rho^{k,\mathrm{pol}} - \rho^k) - (\hat{\rho}_0^{k,\mathrm{pol}} - \rho_0^k) + \mathrm{div}[(\rho^{k,\mathrm{pol}} - \rho^k)f] + \sum_{i=1}^m \mathrm{div}[(\sigma_i^{k,\mathrm{pol}} - \sigma_i^k)f_{u_i}]$$

is a polynomial such that $\|\omega^k\|_{C^0} \to 0$ as $k \to \infty$ in view of (24)-(27). Defining the constants $\epsilon^k = 1/k + \|\omega^k\|_{C^0}$ and setting

$$\rho_T^{k,\mathrm{pol}} := \hat{\rho}_T^{k,\mathrm{pol}} + \epsilon^k$$
$$\rho_0^{k,\mathrm{pol}} := \hat{\rho}_0^{k,\mathrm{pol}} + \epsilon^k + \omega^k$$

we see that

$$\rho_T^{k,\mathrm{pol}} + \beta\rho^{k,\mathrm{pol}} - \rho_0^{k,\mathrm{pol}} + \mathrm{div}(\rho^{k,\mathrm{pol}}f) + \sum_{i=1}^m \mathrm{div}(\sigma_i^{k,\mathrm{pol}}f_{u_i}) = 0,$$

and $\rho_0^{k,\mathrm{pol}} - 1$ and $\rho_T^{k,\mathrm{pol}}$ are strictly positive on X and hence belong to $Q_{d_k}(X)$. The densities $(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}})$ are therefore feasible in (10) for some $d^k \geq 0$. In addition, by construction, $\|\rho_0^{k,\mathrm{pol}} - \rho_0^k\|_{C^0} \to 0$ and $\|\rho_T^{k,\mathrm{pol}} - \rho_T^k\|_{C^0} \to 0$ as $k \to \infty$. Therefore we have obtained a sequence of polynomial densities $(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}})$ that are feasible in (10) and such that

$$\|\rho_0^{k,\mathrm{pol}} - \rho_0^k\|_{C^0} \to 0, \quad \|\rho_T^{k,\mathrm{pol}} - \rho_T^k\|_{C^0} \to 0, \quad \|\rho^{k,\mathrm{pol}} - \rho^k\|_{C^1} \to 0, \quad \|\sigma^{k,\mathrm{pol}} - \sigma^k\|_{C^1} \to 0$$

as $k \to \infty$. This implies that

$$|p(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}}) - p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)| \to 0$$

and hence $(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}})$ satisfies (21) and so $p(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}}) \to p^\star$ by Theorem 3. Therefore (7) holds with the rational controllers $u^k := \sigma^{k,\mathrm{pol}}/\rho^{k,\mathrm{pol}}$ by the second part of Lemma 2. This finishes the proof. $\qquad\square$

6   Value function approximations

In this section we propose a converging hierarchy of approximations from below and from above to the value function V_u associated to a rational controller $u = \sigma/\rho$ with $\sigma \in \mathbb{R}[x]^m$ and $\rho \in \mathbb{R}[x]$ satisfying $0 \leq \sigma_i \leq \bar{u}\rho$ on X. In addition we describe a hierarchy of approximations from below to the optimal value function V. This is useful as a post-processing step, once a rational control law has been computed as described in Section 4, providing an explicit bound on the suboptimality of the controller.

Note that, trivially, approximations from above to V_u provide approximations from above to V. Defining $\hat{f} = \rho f + \sum_{i=1}^m f_{u_i} \sigma_i \in \mathbb{R}[x]^n$ and $\hat{l} = \rho l_x + \sum_{i=1}^m l_{u_i} \sigma_i \in \mathbb{R}[x]$, the degree d polynomial upper and lower bounds are given by

$$
\begin{aligned}
\min_{\overline{V_u} \in \mathbb{R}[x]_d} \quad & \int_X \overline{V_u}(x)\,dx \\
\text{s.t.} \quad & \beta\rho\overline{V_u} - \nabla\overline{V_u} \cdot \hat{f} - \hat{l} \in Q_d(X) \\
& \overline{V_u} - M \in Q_d(X) + \bar{g}\mathbb{R}_{d-\deg \bar{g}},
\end{aligned} \tag{28}
$$

and

$$
\begin{aligned}
\max_{\underline{V_u} \in \mathbb{R}[x]_d} \quad & \int_X \underline{V_u}(x)\,dx \\
\text{s.t.} \quad & -(\beta\rho\underline{V_u} - \nabla\underline{V_u} \cdot \hat{f} - \hat{l}) \in Q_d(X) \\
& M - \underline{V_u} \in Q_d(X) + \bar{g}\mathbb{R}_{d-\deg \bar{g}},
\end{aligned} \tag{29}
$$

respectively. Fixing a basis of $\mathbb{R}[x]_d$, the objective functions of (28) and (29) become linear in the coefficients of $\overline{V_u}$ respectively $\underline{V_u}$ in this basis. Problems (28) and (29) are therefore convex SOS problems and immediately translate to SDPs (see Section 2.2).

Theorem 4 *Let $\overline{V_u}^d$ and $\underline{V_u}^d$ denote solutions to (28) and (29) of degree d. Then $\overline{V_u}^d \geq V_u \geq \underline{V_u}^d$ on X and*

$$
\lim_{d \to \infty} \int_X \overline{V_u}^d(x)\,dx = \int_X V_u(x)\,dx = \lim_{d \to \infty} \int_X \underline{V_u}^d(x)\,dx. \tag{30}
$$

Proof: See Appendix A. □

As a simple corollary we obtain a converging sequence of polynomial over-approximations to V, the optimal value function of (4):

Theorem 5 *Let $\overline{V}_{u^{d_1}}^{d_2}$ denote the degree d_2 polynomial approximation from above to the value function associated to the rational controller u^{d_1} obtained from (10) using (11). Then $\overline{V}_{u^{d_1}}^{d_2} \geq V$ on X and*

$$
\lim_{d_1 \to \infty} \lim_{d_2 \to \infty} \int_X (\overline{V}_{u^{d_1}}^{d_2}(x) - V(x))\,dx = 0.
$$

Now we describe a hierarchy of lower bounds on V:

$$
\begin{aligned}
\max_{\underline{V} \in \mathbb{R}[x]_d,\, p \in \mathbb{R}[x]_d^m} \quad & \int_X \underline{V}(x)\,dx \\
\text{s.t.} \quad & l_x - \beta\underline{V} + \nabla\underline{V} \cdot f + \bar{u}\sum_{i=1}^m p_i \in Q_d(X) \\
& l_{u_i} + \nabla\underline{V} \cdot f_{u_i} - p_i \in Q_d(X) \\
& -p_i \in Q_d(X) \\
& M - \underline{V} \in Q_d(X) + \bar{g}\mathbb{R}_{d-\deg \bar{g}}.
\end{aligned} \tag{31}
$$

Theorem 6 *If $\underline{V} \in \mathbb{R}[x]_d$ is feasible in (31), then $\underline{V} \leq V$ on X.*

Proof: Follows by similar arguments based on Gronwall's Lemma as in the proof of Theorem 4. □

Remark 7 *The question whether $\underline{V}$ converges from below to V as degree d in (31) tends to infinity is open (although likely to hold). A proof would require an extension of the superposition Theorem 7 to non-Lipschitz vector fields (in the spirit of the finite-time superposition result of [1, Theorem 4.4]) or an extension of the argument of [4] to the case of $\mu_T \neq 0$, either of which is beyond the scope of this paper.*

Remark 8 *Besides closed-loop cost function with respect to the OCP (4), one can assess other aspects of the closed-loop behavior of the dynamical system (1) controlled by the rational controller $u = \sigma/\rho$. In particular, regions of attraction or maximum controlled invariant sets can be estimated by methods of [6, 9, 7], which extend readily to the case of rational systems.*

7   Numerical examples

This section demonstrates the approach on numerical examples. To improve the numerical conditioning of the SDPs solved, we use the Chebyshev basis to parametrize all polynomials. More specifically, we use tensor products of univariate Chebyshev polynomials of the first kind to obtain a multivariate Chebyshev basis. We note, however, that similar results, albeit slightly less accurate could be obtained with the standard multivariate monomial basis (in which case the SDPs can be readily formulated using high level modelling tools such as Yalmip [14] or SOSOPT [17]). The resulting SDPs were solved using MOSEK.

7.1   Nonlinear double integrator

As our first example we consider the nonlinear double integrator

$$\dot{x}_1 = x_2 + 0.1x_1^3$$
$$\dot{x}_2 = 0.3u$$

subject to the constraints $u \in [-1, 1]$ and $x \in X := \{x : \|x\|_2 < 1\}$ and stage costs $l_x(x) = x^\top x$ and $l_u(x) = 0$. The discount factor β was set to 1; the constant M to $1.01 > \sup_{x \in X}\{x^\top x\}/\beta = 1$. First we obtain a rational controller of degree six by solving (9) with $d = 6$. The graph of the controller is shown in Figure 1. Next we obtain a polynomial upper bound $\overline{V_u}$ of degree 14 on the value function associated to u by solving (28) with $d = 14$. To assess suboptimality of the controller u we compare it with a lower bound $\underline{V}$ on the optimal value function of the problem (4) obtained by solving (31) with $d = 14$. The graphs of the two value functions are plotted in Figure 2. We see that the gap between the upper bound on V_u and lower bound on V is relatively small, verifying a good performance of the extracted controller. Quantitatively, the average performance gap defined as $100 \int_X (\overline{V_u} - \underline{V})dx / \int_X \underline{V}dx$ is equal to 19.5%.
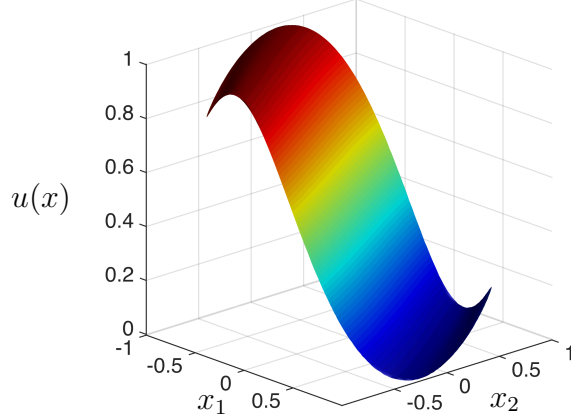
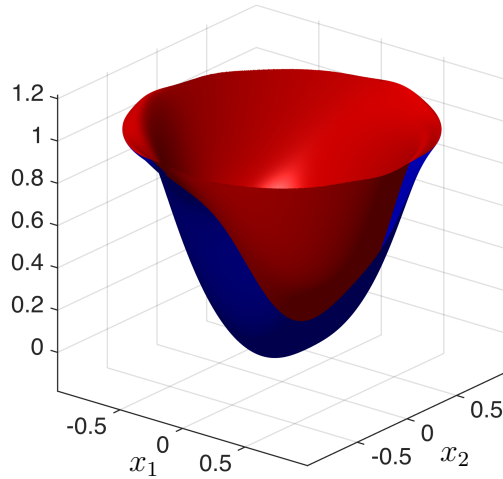Figure 1: Nonlinear double integrator – rational controller of degree six.

Figure 2: Nonlinear double integrator – upper bound on the value function $\overline{V_u}$ associated to the extracted controller (red); lower bound on the optimal value function $\underline{V}$ (blue).

7.2 Controlled Lotka-Volterra

In our second example we apply the proposed method to a population model governed by n-dimensional controlled Lotka-Volterra equations

$$\dot{x} = r \circ x \circ (e - Ax) + u^+ - u^-,$$

where $e \in \mathbb{R}^n$ is the vector of ones and $\circ$ denotes the componentwise (Hadamard) product. Each component x_i of the state $x \in \mathbb{R}^n$ represents the size of the population of species i. The vector $r \in \mathbb{R}^n$ contains the intrinsic growth rates of each species and the matrix $A \in \mathbb{R}^{n \times n}$ captures the interaction between the species. If $A_{i,j} > 0$, then species j is harmful to species

i (e.g., competes for resources) and if $A_{i,j} < 0$, then species j is helpful to species i (e.g., species i feeds on species j); the diagonal components $A_{i,i}$ are normalized to one. The control inputs $u^+ \in [0,1]^n$ and $u^- \in [0,1]^n$ represent, respectively, the inflow and outflow of new species from the outside. For our numerical example we select $n = 4$ and model parameters

$$r = \begin{bmatrix} 1 \\ 0.6 \\ 0.4 \\ 0.2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0.3 & 0.4 & 0.2 \\ -0.2 & 1 & 0.4 & -0.1 \\ -0.1 & -0.2 & 1 & 0.3 \\ -0.1 & -0.2 & -0.3 & 1 \end{bmatrix},$$

which results in a system with four states and eight control inputs. The economic objective is to harvest species number one while ensuring that no species goes extinct. More specifically the cost function is $l_u(x) = (-1.0, 0.5, 0.6, 0.8, 1.1, 2, 4, 6)$ and $l_x(x) = 1$, where the vector $l_u(x)$ is associated with the control input vector $u = (u^-, u^+)$. Therefore there is a reward for harvesting species number one and cost associated with both introduction and hunting of all other species, the cost of hunting being lower than the cost of introduction. The reason for choosing $l_x(x) = 1$ is in order to make the joint stage cost $l(x, u)$ (6) nonnegative; this choice does not affect optimality since $l_x(x(t)) = 1$ irrespective of the control input applied. The non-extinction constraint is expressed as $g(x) = 1 - (Q^{-1}x - q)^\top(Q^{-1}x - q) \geq 0$ with $Q = \mathrm{diag}(0.475 \cdot e)$ and $q = 0.525 \cdot e$. We choose $\beta = 1$ and $M = 16.16 > \sup_{x \in X, u \in U}\{l(u, x)\}/\beta = 16$. We apply the coordinate transformation $x = Q\hat{x} + q$ and solve solve obtain a rational controller of degree eight by solving (10). Figure 3 we shows plots for two different initial conditions, one with low population size of the first species and one with high. Finally, we evaluate the suboptimality of the extracted controller using the polynomial lower bound on the optimal value function of degree 11 obtained from (31). Using Monte Carlo simulation with 1000 samples of initial conditions drawn from a uniform distribution over the constraint set we obtain average cost of the extracted controller to be 0.89 whereas the lower bounds predicts average cost of 0.72; hence the extracted controller is no more than 23.6 % suboptimal (modulo the statistical estimation error). Note that we could also obtain a deterministic suboptimality estimate using the upper bound on the value function of the extracted controller obtained from (28). In this case, however, the upper bound (28) is not informative. Nevertheless, the Monte Carlo simulation along with the lower bound (31) is a viable alternative in this case, since the extracted controller is simple and hence trajectories of the controlled system can be simulated rapidly.

8 Conclusion

This paper presented a method to obtain a sequence of rational controllers asymptotically optimal (under suitable technical assumptions) in a discounted optimal control problem and a method to explicitly estimate suboptimality of each controller. The rational controller of a given degree is obtained by solving a single sum-of-squares problem with no extraction step. The SOS problem solved is feasible for any degree and therefore this method allows to trade off complexity of the controller against performance.

The approach is based on lifting the nonconvex optimal control problem into an infinite dimensional space of measures with continuous densities, where this problem becomes linear.
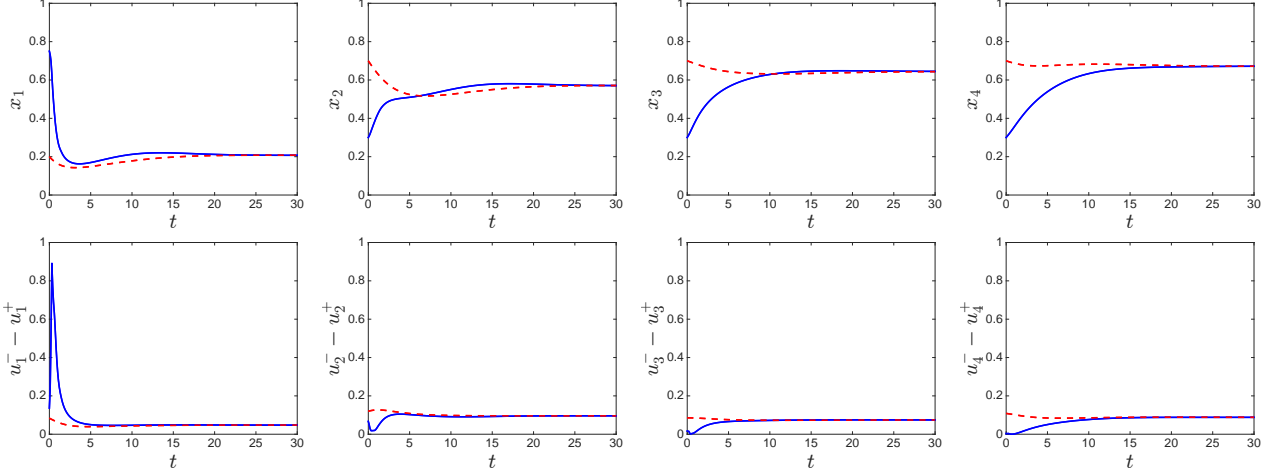
Figure 3: Controlled Lotka-Volterra – (blue) trajectory starting from a high initial population of the first species and low initial population of the other species; (red) trajectory starting from low initial population of the first species and high initial population of the other species.

Crucially, this problem is a tightening of the original problem, which follows immediately from the representation result for solutions of the discounted Liouville's equation with a terminal measure (Theorem 7). Asymptotic optimality of the extracted controllers then follows by approximating the asymptotically optimal continuous densities (guaranteed to exist by Assumption 1) with polynomial densities in such a way that these densities correspond to the densities of the dynamical system (this is the essence of the proof of Theorem 1).

9 Appendix A

This Appendix contains the proof of Theorem 4; we use the same notation as in Section 6. The inequalities $\overline{V_u}^d \geq V_u \geq \underline{V_u}^d$ follow from Gronwall's Lemma by noticing that the constraints of (28) and (29) imply that

$$\nabla \overline{V_u}^d \cdot (f + \sum_{i=1}^m f_{u_i} u_i) \leq \beta \overline{V_u}^d - (l_x + \sum_{i=1}^m l_{u_i} u_i), \tag{32}$$

$$\nabla \underline{V_u}^d \cdot (f + \sum_{i=1}^m f_{u_i} u_i) \geq \beta \underline{V_u}^d - (l_x + \sum_{i=1}^m l_{u_i} u_i) \tag{33}$$

on X and $\overline{V_u} \geq M$, $\underline{V_u} \leq M$ on ∂X. We detail the argument for the inequality $\overline{V_u}^d \geq V_u$, the inequality $V_u \geq \underline{V_u}^d$ being similar. Given $x_0 \in X$ the inequality (32) implies that

$$\frac{d}{dt} \overline{V_u}^d(x(t \,|\, x_0)) \leq \beta \overline{V_u}^d(x(t \,|\, x_0)) - \left[l_x(x(t \,|\, x_0)) + \sum_{i=1}^m l_{u_i}(x(t \,|\, x_0)) u_i(x(t \,|\, x_0)) \right],$$

and therefore by Gronwall's Lemma

$$\overline{V_u}^d(x(t \,|\, x_0)) \leq e^{\beta t} \overline{V_u}^d(x_0) - \int_0^t e^{\beta(t-s)} \left[l_x(x(s \,|\, x_0)) + \sum_{i=1}^m l_{u_i}(x(s \,|\, x_0)) u_i(x(s \,|\, x_0)) \right] ds$$

19

and hence

$$\overline{V_u}^d(x_0) \geq e^{-\beta t}\overline{V_u}^d(x(t\,|\,x_0)) + \int_0^t e^{-\beta s}\Big[l_x(x(s\,|\,x_0)) + \sum_{i=1}^m l_{u_i}(x(s\,|\,x_0))u_i(x(s\,|\,x_0))\Big]\,ds \quad (34)$$

for all $t \in [0, \tau]$, where $\tau := \inf\{t \geq 0 \mid x(t\,|\,x_0) \notin X\} \in [0, \infty]$ is the first exit time of X. Next we observe that V_u, the value function associated to u, is equal to

$$V_u(x_0) = \begin{cases} \int_0^\infty e^{-\beta s}\Big[l_x(x(s\,|\,x_0)) + \sum_{i=1}^m l_{u_i}(x(s\,|\,x_0))u_i(x(s\,|\,x_0))\Big]ds, & \tau = \infty \\ Me^{-\beta\tau} + \int_0^\tau e^{-\beta s}\Big[l_x(x(s\,|\,x_0)) + \sum_{i=1}^m l_{u_i}(x(s\,|\,x_0))u_i(x(s\,|\,x_0))\Big]ds, & \tau < \infty. \end{cases}$$

In view of (34), we conclude that $\overline{V_u}^d(x_0) \geq V_u(x_0)$ if $\tau = \infty$ since $\overline{V_u}^d$ is polynomial and hence bounded on X (and hence $e^{-\beta t}\overline{V_u}^d(x(t\,|\,x_0)) \to 0$); and we conclude that $\overline{V_u}^d(x_0) \geq V_u(x_0)$ if $\tau < \infty$ since $x(\tau\,|\,x_0) \in \partial X$ and $\overline{V_u}^d \geq M$ on ∂X.

Convergence of the upper and lower bounds (30) follows from Theorem 7 using infinite-dimensional LP duality and standard results on the convergence of moment relaxations. The proof is similar to the proof of Theorem 5 in [7] or Theorem 3.6 in [13] and therefore we only outline it. The hierarchy of SOS programming problems (28) and (29) is dual to the hierarchy of moment relaxations of an infinite-dimensional LP in the cone of nonnegative measures whose dual is an infinite-dimensional LP in $C^1(X)$ and feasible solutions of this dual provide upper or lower bounds on V_u. Crucial to applying infinite-dimensional duality results (e.g., [2, Theorem 3.10]) is the boundedness of measures satisfying the discounted Liouville equation (14) with $\nu_i \leq \bar{u}\mu$ and $\mu_0 = \lambda_X$, where λ_X is the restriction of the Lebesgue measure to X. Plugging $v = 1$ in (14) we have $\mu_T(X) + \beta\mu(X) = \mu_0(X)$. Since $\mu_0(X) = \lambda_X(X) = \text{vol}\,X < \infty$ and $\beta > 0$ we conclude that μ_T and μ are indeed bounded, which implies that ν_i is also bounded for $i = 1, \ldots, m$. Equally important is the absence of duality gap between the finite-dimensional moment relaxations and SOS tightenings (which are both SDP problems); this follows immediately from the presence of the constraint $g_i = N - \|x\|_2^2$ among the constraints describing X, which implies the boundedness of the truncated moment sequences feasible in the moment relaxations. The absence of duality gap then follows from [23, Lemma 2]. $\qquad\square$

10 Appendix B

This appendix presents a proof of Theorem 2. We will prove a slightly more general version of the result from which Theorem 2 immediately follows:

Theorem 7 *Let $\bar{f} : \mathbb{R}^n \to \mathbb{R}^n$ be globally Lipschitz and let the nonnegative measures μ, μ_0, μ_T on $\mathbb{R}^n$ satisfy*

$$\int_{\mathbb{R}^n} v\,d\mu_T = \int_{\mathbb{R}^n} v\,d\mu_0 + \int_{\mathbb{R}^n} (\nabla v \cdot \bar{f} - \beta v)\,d\mu \quad (35)$$

for all $v \in C^1(\mathbb{R}^n)$. Then there exists an ensemble of probability measures $\{\tau_{x_0}\}_{x_0 \in X}$ with

spt $\tau_{x_0} \subset [0, \infty]$ *and an ensemble of trajectories* $\{x(\cdot \,|\, x_0)\}_{x_0 \in X}$ *of the ODE* $\dot{x} = \bar{f}(x)$ *and*

$$\int_{\mathbb{R}^n} v(x) \, d\mu_0(x) = \int_{\mathbb{R}^n} v(x(0 \,|\, x_0)) \, d\mu_0(x_0), \tag{36a}$$

$$\int_{\mathbb{R}^n} v(x) \, d\mu(x) = \int_{\mathbb{R}^n} \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t \,|\, x_0)) \, dt \, d\tau_{x_0}(\tau) \, d\mu_0(x_0), \tag{36b}$$

$$\int_{\mathbb{R}^n} v(x) \, d\mu_T(x) = \int_{\mathbb{R}^n} \int_0^\infty e^{-\beta \tau} v(\tau(x_0)) \, d\tau_{x_0}(\tau) \, d\mu_0(x_0), \tag{36c}$$

for all $v \in L^1(\mathbb{R}^n)$.

Theorem 2 follows from Theorem 7 by setting $\bar{f} = f + \sum_{i=1}^m f_{u_i} u_i$ and modifying f and f_{u_i} outside the compact set X such that $\bar{f}$ is globally Lipschitz[4]. The conclusion that $x(t \,|\, x_0) \in X$ for all $t \in \mathrm{spt}\,\tau_{x_0}$ follows by taking $v(x) = e^{-\|x\|^2} I_{\mathbb{R}^n \setminus X}(x)$ in (36), where I_A is the indicator function of a set A, i.e. $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise.

Suppose therefore that (36) holds. First we will prove a simple result. In the rest of this Appendix we will use the notation C_c^k for the space of all compactly supported k-times continuously differentiable functions.

Lemma 4 *For any* $w \in C_c^1(\mathbb{R}^n)$, *the equation*

$$\nabla v \cdot \bar{f} - \beta v = w \tag{37}$$

has a solution v *such that for all* $x_0 \in \mathbb{R}^n$ *it holds*

$$v(x_0) = -\int_0^\infty e^{-\beta t} w(x(t \,|\, x_0)) \, dt. \tag{38}$$

Proof: Since $\bar{f}$ is globally Lipschitz the solution $x(t \,|\, x_0)$ is defined for all $x_0 \in \mathbb{R}^n$ and all $t \geq 0$. Therefore (38) is well defined (notice that w is bounded and $\beta > 0$). Direct

---

[4]Such modification is always possible. For instance let $\bar{f}(x) = \min_{y \in X}\{f(y) + \sum_{i=1}^m f_{u_i}(y) u_i(y) + L\|x - y\|\}$, where L is the Lipschitz constant of $f + \sum_{i=1}^m f_{u_i} u_i$ on X.

computation then gives:

$$\nabla v \cdot \bar{f}(x(t \mid x_0)) = \frac{d}{dt} v(x(t \mid x_0))$$

$$= -\frac{d}{dt} \int_0^\infty e^{-\beta s} w(x(s \mid x(t \mid x_0))) \, ds$$

$$= -\frac{d}{dt} \int_0^\infty e^{-\beta s} w(x(t+s \mid x_0)) \, ds$$

$$= -\int_0^\infty e^{-\beta s} \nabla w(x(t+s \mid x_0)) \cdot \bar{f}(x(t+s \mid x_0)) \, ds$$

$$= -\int_0^\infty e^{-\beta s} \nabla w(t+s \mid x_0)) \cdot \bar{f}(x(t+s \mid x_0)) \, ds$$

$$= -\int_0^\infty e^{-\beta s} \frac{d}{ds} w(x(t+s \mid x_0)) \, ds$$

$$= -\beta \int_0^\infty e^{-\beta s} w(x(t+s \mid x_0)) \, ds - [e^{-\beta s} w(x(t+s \mid x_0))]_0^\infty$$

$$= -\beta \int_0^\infty e^{-\beta s} w(x(s \mid x(t \mid x_0))) \, ds + w(x(t \mid x_0))$$

$$= \beta v(x(t \mid x_0)) + w(x(t \mid x_0)).$$

Setting $t = 0$, we arrive at (37). $\qquad \square$

Proof (of Theorem 7) We will proceed in several steps.

Two Diracs. We start with the simplest case of $\mu_0 = \delta_{x_0}$ and $\mu_T = a\delta_{x_T}$, $a > 0$, $x_T \in \mathbb{R}^n$, and some $\mu \geq 0$. First we will show that if (μ_T, μ_0, μ) solves (35) then there exists a time $\tau \geq 0$ such that $x(\tau \mid x_0) = x_T$. Consider now any $w \in C_c^1(\mathbb{R}^n)$, $w \geq 0$ and the associated $v \in C^1(\mathbb{R}^n)$ solving (37). Then we have

$$av(x_T) - v(x_0) = \int_{\mathbb{R}^n} (\nabla v \cdot f - \beta v) \, d\mu = \int_{\mathbb{R}^n} w \, d\mu \geq 0.$$

Therefore, by Lemma 4,

$$av(x_T) \geq v(x_0) = -\int_0^\infty e^{-\beta t} w(x(t \mid x_0)) \, dt.$$

Using (38) again on $v(x_T)$ we get

$$-a \int_0^\infty e^{-\beta t} w(x(t \mid x_T)) \, dt \geq -\int_0^\infty e^{-\beta t} w(x(t \mid x_0)) \, dt,$$

or

$$a \int_0^\infty e^{-\beta t} w(x(t \mid x_T)) \, dt \leq \int_0^\infty e^{-\beta t} w(x(t \mid x_0)) \, dt. \qquad (39)$$

Now pick $S \geq 0$ (to be specified later) and consider the traces

$$\mathcal{X}_0 = \{x(t \mid x_0) \mid 0 \leq t \leq S\}.$$

22

$$\mathcal{X}_T = \{x(t \mid x_T) \mid 0 \le t \le S\}.$$

Assuming there is no $\tau \ge 0$ such that $x(\tau \mid x_0) = x_T$ we have $\mathcal{X}_0 \cap \mathcal{X}_T = \emptyset$ and since $\mathcal{X}_0$ and $\mathcal{X}_T$ are compact there exist (by Uryshon's Lemma with mollification) a function $w \in C_c^1(\mathbb{R}^d; [0, 1])$ such that $w = 0$ on $\mathcal{X}_0$ and $w = 1$ on $\mathcal{X}_T$. Then the left hand side of (39) is greater than or equal to $a(1 - e^{-\beta S})/\beta$ whereas the right hand side is less than or equal to $e^{-\beta S}/\beta$. Since $a > 0$ and $\beta > 0$ we arrive at a contradiction with (39) by picking a sufficiently large S. Therefore there exists a $\tau \ge 0$ such that $x(\tau \mid x_0) = x_T$ (i.e., x_T and x_0 are on the same trace of the flow associated to $\dot{x} = f(x)$).

Now we prove that $a \le e^{-\beta\tau}$. Since $x_T = x(\tau)$ and x_0 are on the same trace we have

$$v(x_0) = e^{-\beta\tau} \underbrace{v(x_T)}_{v(x(\tau))} - \int_0^\tau w(x(t \mid x_0)) \, dt.$$

Using again $av(x_T) \ge v(x_0)$ if $w \ge 0$ we get

$$av(x_T) \ge e^{-\beta\tau} v(x_T) - \int_0^\tau w(x(t \mid x_0)) \, dt, \quad \text{or}$$

$$(e^{-\beta\tau} - a) \int_0^\infty e^{-\beta t} w(x(t \mid x_T)) dt \ge - \int_0^\tau w(x(t \mid x_0)) \, dt. \tag{40}$$

Consider the set

$$\mathcal{X}_\tau = \{x(t \mid x_0) \mid 0 \le t \le \tau\}.$$

Since x_0 and x_T are on the same trace (and x_T follows x_0) there exists $w \in C_c^1(X)$, $w \ge 0$, such that $w = 0$ on $\mathcal{X}_\tau$ and $w > 0$ elsewhere (e.g., let $w(x) = \min(\text{dist}(x, \mathcal{X}_\tau), 1)$ with appropriate mollification). With this choice of w the equation (40) gives

$$(e^{-\beta\tau} - a) \int_0^\infty e^{-\beta t} w(x(t \mid x_T)) dt \ge 0$$

and therefore $a \le e^{-\beta\tau}$ since the integral is strictly positive. This proves the first two claims.

To finish we observe that any solution to (37) satisfies

$$e^{-\beta\tau} v(x_T) = v(x_0) + \int_0^\tau e^{-\beta t} w(x(t \mid x_0)) \, dt.$$

Therefore

$$av(x_T) = v(x_0)ae^{\beta\tau} + ae^{\beta\tau} \int_0^\tau e^{-\beta t} w(x(t \mid x_0)) \, dt.$$

Using (38) we get

$$av(x_T) = v(x_0) + \underbrace{ae^{\beta\tau}}_{\ge 0} \int_0^\tau e^{-\beta t} w(x(t \mid x_0)) \, dt + \underbrace{(1 - ae^{\beta\tau})}_{\ge 0} \int_0^\infty e^{-\beta t} w(x(t \mid x_0)) \, dt.$$

Since

$$av(x_T) - v(x_0) = \int_{\mathbb{R}^n} w \, d\mu$$

23

we conclude that

$$\int_{\mathbb{R}^n} w \, d\mu = ae^{\beta\tau} \int_0^\tau e^{-\beta t} w(x(t \,|\, x_0)) \, dt + (1 - ae^{\beta\tau}) \int_0^\infty e^{-\beta t} w(x(t \,|\, x_0)) \, dt,$$

i.e., μ is indeed generated by trajectories of $\dot{x} = f(x)$ (in this case by two trajectories, both starting at x_0, one stopping at τ, the other one continuing to infinity with weights given by the ratio of masses of μ_0 and μ_T). That is the measure τ_{x_0} is given by

$$\tau_{x_0} = ae^{\beta\tau}\delta_\tau + (1 - ae^{\beta\tau})\delta_\infty$$

as expected.

Dirac at x_0, sum of Diracs for μ_T. Next we treat the case where $\mu_T = \sum_{i=1}^\infty a_i \delta_{x_i}$ for some $a_i \geq 0$ and $x_i \in \mathbb{R}^n$. Using the same argument as in the previous case we can show that

$$x_i \in \mathcal{X}_0 = \{x(t \mid x_0) \mid t \geq 0\}$$

for all i and that the condition

$$\sum_{i=1}^\infty a_i e^{\beta\tau_i} \leq 1,$$

holds with τ_i being the times to reach x_i from x_0. Then we have

$$\tau_{x_0} = \sum_{i=1}^\infty a_i e^{\beta\tau_i}\delta_{\tau_i} + (1 - \sum_{i=1}^\infty a_i e^{\beta\tau_i})\delta_\infty.$$

Dirac at x_0 arbitrary μ_T. In the same way as before we can show that the support of μ_T must be on the trace $\mathcal{X}_0$. Then we can define the measure $\hat{\tau}_{x_0}$ by

$$\hat{\tau}_{x_0}(A) := \mu_T(x(A \mid x_0)), \quad A \subset [0, \infty)$$

and show that it has to satisfy the condition $\int_0^\infty e^{\beta t} d\hat{\tau}_{x_0}(t) \leq 1$. Next, using the fact that the mapping $t \mapsto x(t \,|\, x_0)$ is invertible, we obtain

$$\int_{\mathbb{R}^n} v \, d\mu_T = \int_0^\infty v(x(t \,|\, x_0)) \, d\hat{\tau}_{x_0}(t).$$

The conclusion of the theorem then holds with τ_{x_0} defined by

$$\tau_{x_0}(A) = \int_0^\infty I_A(t)e^{\beta t} d\hat{\tau}_{x_0}(t) + \left[1 - \int_0^\infty e^{\beta t} d\hat{\tau}_{x_0}(t)\right] I_A(\infty), \quad A \subset [0, \infty].$$

Arbitrary μ_0, arbitrary μ_T. The general case follows by approximating μ_0 by a sum of Dirac measures, using the fact that any measure is the weak limit of a sequence of Dirac measures.

References

[1] L. Ambrosio, G. Crippa. Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields. Transport Equations and Multi-D Hyperbolic Conservation Laws, Lecture Notes of the Unione Matematica Italiana, Vol. 5, Springer, 2008.

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

[3] G. Crippa. The flow associated to weakly differentiable vector fields, Theses of Scuola Normale Superiore di Pisa, Edizioni della Normale, Pisa, 2009.

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

[5] D. Henrion, J. B. Lasserre, C. Savorgnan. Nonlinear optimal control synthesis via occupation measures. IEEE Conf. Decision and Control, Cancún, Mexico, 2008.

[6] D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. IEEE Trans. Autom. Control, 59:297-312, 2014.

[7] M. Korda, D. Henrion, C. N. Jones. Convex computation of the maximum controlled invariant set for polynomial control systems. SIAM J. Control Optim., 52:2944-2969, 2014.

[8] M. Korda, D. Henrion, C. N. Jones. Controller design and region of attraction estimation for nonlinear dynamical systems. Proc. IFAC World Congress Autom. Control, Cape Town, South Africa, 2014.

[9] M. Korda, D. Henrion, C. N. Jones. Controller design and region of attraction estimation for nonlinear dynamical systems. Proc. IFAC Symp. Nonlinear Control Systems, Toulouse, France, 2013. Extended version: `arXiv:1310.2213`, 2013.

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

[11] J. B. Lasserre. Global Optimization with Polynomials and the Problem of Moments, SIAM J. Optim. 11(3):796-817, 2001.

[12] J. B. Lasserre. A New Look at Nonnegativity on Closed Sets and Polynomial Optimization. SIAM J. Optim. 21(3):864-885, 2011.

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

[14] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In Proc. IEEE CCA/ISIC/CACSD, Taipei, Taiwan, 2004.

[15] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake. Convex optimization of nonlinear feedback controllers via occupation measures. Int. J. Robotics Research, 33:1209-1230, 2014.

[16] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. 42:969-984, 1993.

[17] P. Seiler. SOSOPT: A toolbox for polynomial optimization. University of Minnesota, 2010.

[18] A. Rantzer. An converse theorem for density functions. Proc. IEEE Conf. Decision and Control, Las Vegas, Nevada, 2002.

[19] A. Rantzer and S. Hedlund. Duality between cost and density in optimal control. Proc. IEEE Conf. Decision and Control, Maui, Haiwai, 2003.

[20] A. Raghunathan, U. Vaidya. Optimal stabilization using Lyapunov measures. IEEE Trans. Autom. Control, 59:1316-1321, 2014.

[21] S. Prajna. P. A. Parrilo. A. Rantzer. Nonlinear control synthesis by convex optimization. IEEE Trans. Autom. Control, 49:310-314, 2004.

[22] J. E. Rubio. Control and optimization: the linear treatment of nonlinear problems. Manchester University Press, Manchester, UK, 1985.

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

[24] R. B. Vinter, R. M. Lewis. The equivalence of strong and weak formulations for certain problems in optimal control, SIAM J. Control Optim., 16:546-570, 1978.

[25] L. C. Young, Lectures on the calculus of variations and optimal control theory, W. B. Saunders, Philadelphia, 1969.

[26] J. Warga. Optimal control of differential and functional equations, Academic Press, New York, 1972.