Moment-sum-of-squares hierarchies for set approximation and optimal control

THÈSE N^{O} 7012 (2016)

PRÉSENTÉE LE 10 JUIN 2016

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE D'AUTOMATIQUE 3
PROGRAMME DOCTORAL EN SYSTÈMES DE PRODUCTION ET ROBOTIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Milan KORDA

acceptée sur proposition du jury:

Dr A. Karimi, président du jury
Prof. C. N. Jones, directeur de thèse
Prof. J. B. Lasserre, rapporteur
Prof. A. Rantzer, rapporteur
Prof. A. Billard, rapporteuse

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

Acknowledgements

First of all, my greatest thanks goes to Colin Jones for giving me a great deal of freedom in research and in organizing my personal time, for plethora of travel opportunities, as well as for his invaluable creative input during our long discussions in front of a whiteboard. I especially appreciate Colin's friendly, positive and unpretentious attitude during these discussion that always left me with new ideas and optimism.

A great thanks goes also to Didier Henrion whose visit of EPFL a few months after the start of my thesis put my research on a completely different path than originally intended by me or Colin. What had been a side-topic of interest for me became the core of my research and the subsequent collaboration with Didier proved to be incredibly fruitful and enjoyable.

LA is a special place; everybody says that. And it is true. It has attracted special, exceptionally colorful characters that made me laugh even on the rainiest of days. I will never forget those long (pseudo)-philosophical discussions and gossip-packed pub outings as well as those countless little pranks that spiced up the everyday lab life. A special thanks goes to the LA's rock climbing gang for showing me that there exist other things almost as cool as solving large SDPs.

Finally, a big thanks goes to my family for treating me to great Czech food during my visits and for finally giving up on asking about what it is that I actually do, as well as to all my friends for their willingness to always share a good laugh.

Abstract

This thesis uses the idea of lifting (or embedding) a nonlinear controlled dynamical system into an infinite-dimensional space of measures where this system is equivalently described by a linear equation. This equation and problems involving it are subsequently approximated using well-known moment-sum-of-squares hierarchies.

First, we address the problems of region of attraction, reachable set and maximum controlled invariant set computation, where we provide a characterization of these sets in terms of an infinite-dimensional linear program in the cone of nonnegative measures and we describe a hierarchy of finite-dimensional semidefinite-programming (SDP) hierarchies providing a converging sequence of outer approximations to these sets.

Next, we treat the problem of optimal feedback controller design under state and input constraints. We provide a hierarchy of SDPs yielding an asymptotically optimal sequence of rational feedback controllers. In addition, we describe hierarchies of SDPs yielding approximations to the value function attained by any given rational controller, from below and from above, as well as a hierarchy of SDPs providing approximations from below to the optimal value function, hence obtaining performance certificates for the designed controllers as well as for any given rational controller.

Finally, we describe a method to verify properties of a closed loop interconnection of a nonlinear dynamical system and an optimization-based controller (e.g., a model predictive controller) for deterministic and stochastic nonlinear dynamical systems. Properties such as global stability, the ℓ_2 gain or performance with respect to a given infinite-horizon cost function can be certified.

The methods presented are easy to implement using freely available software packages and are documented by a number of numerical examples.

Key words: region of attraction, reachable set, maximum controlled invariant set, optimal control, moment hierarchy, sum-of-squares, semidefinite programming, controller verification, lifting, embedding

Résumé

Cette thèse utilise l'idée de lifting (ou embedding) d'un système dynamique non linéaire contrôlé dans un espace de dimension infinie de mesures où ce système est décrit de façon équivalente par une équation linéaire. Cette équation et les problèmes l'impliquant sont ensuite approximés à l'aide des hiérarchies bien connues moment-somme de carrés.

Tout d'abord, nous abordons les problèmes de calcul de la région d'attraction, de l'ensemble atteignable et de l'ensemble invariant contrôlé maximal, où nous fournissons une caractérisation de ces ensembles par un problème d'optimisation linéaire de dimension infinie dans le cône de mesures non négatifs et nous décrivons une hiérarchie de hierarchies de problèmes d'optimisation semi-définis (SDP) en dimension finie fournissant une séquence d'approximations extérieures convergentes de ces ensembles.

Ensuite, nous traitons le problème de la conception de contrôleur optimal sous contraintes d'état et d'entrée. Nous fournissons une hiérarchie de SDP qui donne une séquence asymptotiquement optimale de contrôleurs rationnels. En outre, nous décrivons des hiérarchie de SDP produisant des approximations de la fonction objectif atteinte par un contrôleur rationnel donné, lune par valeur supérieure, l'autre inférieure, ainsi qu'une hiérarchie de SDP donnant une approximation inférieure à la fonction objectif optimale, en obtenant donc des certificats de performance pour la contrôleurs conçus ainsi que pour tout contrôleur rationnel donné.

Enfin, nous décrivons une méthode pour vérifier les propriétés d'une boucle fermée formée par l'interconnexion d'un système dynamique non linéaire et un contrôleur résolvant un problème d'optimisation (par exemple, un contrôleur prédictif) pour les systèmes dynamiques non linéaires déterministes et stochastiques. Les propriétés telles que la stabilité globale, le gain ℓ_2 ou la performance par rapport à une fonction de coût à horizon infini peuvent être certifiés.

Les méthodes présentées sont faciles à mettre en œuvre en utilisant des logiciels disponibles gratuitement et sont documentées par un certain nombre d'exemples numériques.

Mots clefs : région d'attraction, ensemble atteignable, ensemble invariant contrôlé maximal, commande optimale, hiérarchie des moment, somme des carrés, programmation semi-définie, vérification de contrôleur, lifting, embedding

Contents

List of Figures

List of Tables

Nomenclature

$\overline{\mathbf{X}}$	The closure of a set $\mathbf{X}$
$\deg p$	Total degree of a multivariate polynomial
$\mathbf{X}^\circ$	The interior of a set $\mathbf{X}$
μ^{leb}	The Lebesgue measure, i.e., $\int f(x)\, d\mu^{\text{leb}}(x) = \int f(x)\, dx$ and $\mu^{\text{leb}}(A) = \text{vol}(A)$
$\\|f\\|_{C(\mathbf{X})}$	The supremum norm of f over $\mathbf{X}$, i.e., $\\|f\\|_{C(\mathbf{X})} = \sup_{x\in\mathbf{X}} \|f(x)\|$
$\\|f\\|_{C^0}$	The same as $\\|f\\|_{C(\mathbf{X})}$ when the set $\mathbf{X}$ is clear from the context
$\\|f\\|_{C^k}$	The same as $\\|f\\|_{C^k(\mathbf{X})}$ when the set $\mathbf{X}$ is clear from the context
$\\|f\\|_{C^1(\mathbf{X})}$	The C^1 norm of f over $\mathbf{X}$, i.e., $\\|f\\|_{C^1(\mathbf{X})} = \sup_{x\in\mathbf{X}} \|f(x)\| + \sup_{x\in\mathbf{X}} \\|\nabla f\\|_2$
$\text{dist}_{\mathbf{X}}(x)$	The Euclidean distance of a point x to the set $\mathbf{X}$
∇f	The spacial gradient of f. If f is a function of (t, x), then $\nabla f = \frac{\partial f}{\partial x}$
$\mathbb{R}[x]$	The ring of all polynomials
$\mathbb{R}[x]_d$	The vector space of all polynomials of total degree at most d
Σ	The cone of all sum-of-squares polynomials
Σ_d	The cone of all sum-of-squares polynomials of total degree at most d
$\text{spt}\,\mu$	Support of a measure μ, i.e, the smallest closed set whose complement has a zero measure. Equivalently, $x \in \text{spt}\,\mu$ if an only if $\mu(U) > 0$ for every open neighborhood U of x
$\mathbf{X}^\star$	The space of all bounded linear functionals on $\mathbf{X}$ (= topological dual of $\mathbf{X}$)

$A^\star$	The adjoint of a linear operator A
$AC([a,b];\mathbf{Y})$	The set of all absolutely continuous functions $f:[a,b]\to\mathbf{Y}$
$C(\mathbf{X};\mathbf{Y})$	The space of all continuous functions $f:\mathbf{X}\to\mathbf{Y}$
$C(\mathbf{X})$	The space of all continuous functions $f:\mathbf{X}\to\mathbb{R}$
$C(\mathbf{X})_+$	The cone of all nonnegative continuous functions $f:\mathbf{X}\to\mathbb{R}$
$C^k(\mathbf{X})$	The space of all functions $f:\mathbf{X}\to\mathbb{R}$ which are continuous on $\mathbf{X}$ and k-times continuously differentiable on $\mathbf{X}^\circ$
$C^k_b(\mathbf{X})$	The space of all bounded functions $f:\mathbf{X}\to\mathbb{R}$ which are continuous on $\mathbf{X}$ and k-times continuously differentiable on $\mathbf{X}^\circ$
$C_b(\mathbf{X})$	The space of all bounded continuous functions $f:\mathbf{X}\to\mathbb{R}$
$I_\mathbf{X}$	The indicator function of the set $\mathbf{X}$, i.e., $I_\mathbf{X}=1$ on $\mathbf{X}$ and $I_\mathbf{X}=0$ otherwise
$l(\mathbb{N};\mathbf{X})$	The set of all sequences indexed with natural numbers taking values in $\mathbf{X}$
$L(\mathbf{X};\mathbf{Y})$	The space of all Borel measurable functions $f:\mathbf{X}\to\mathbf{Y}$
$L(\mathbf{X})$	The space of all Borel measurable functions $f:\mathbf{X}\to\mathbb{R}$
$L_k(\mathbf{X})$	The space of all functions $f:\mathbf{X}\to\mathbb{R}^n$ such that $\int_\mathbf{X}\lvert f\rvert^k\,dx<\infty$ for $k<\infty$ and such that $\operatorname{ess\,sup}_\mathbf{X}\lvert f\rvert<\infty$ for $k=\infty$
$M(\mathbf{X})$	The space of all signed Borel measures on $\mathbf{X}$
$M(\mathbf{X})_+$	The cone of all nonnegative Borel measures on $\mathbf{X}$

Chapter 1

Introduction

This thesis builds on the idea of *lifting* (or embedding) a *nonlinear* problem into a larger (possibly infinite-dimensional) space where this problem is equivalently represented by a *linear* problem. This linear problem is subsequently approximated by a tractable finite-dimensional problem of a predefined complexity controlling the quality of the approximation. This abstract idea is depicted in Figure 1.1.

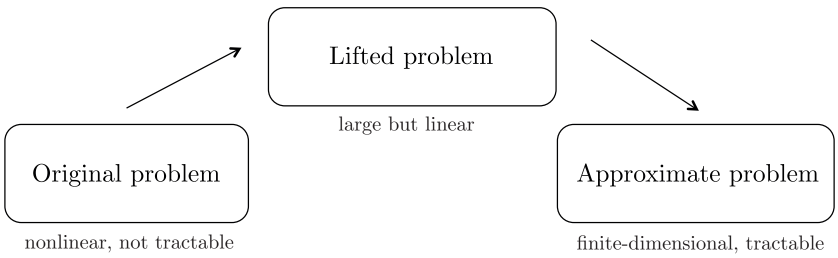


Figure 1.1 – Lifting idea

This idea permeates many areas of science under different names and guises including optimization, machine learning, graph theory, computer science, quantum physics and many others. In this thesis we use this idea to address several problems from the field of nonlinear controlled dynamical systems. In this field, the idea of lifting goes back to the work of L. C. Young [You69] (with first ideas laid out already in [You33]), which introduced the notions of generalized arcs and generalized flows. Building on this work, the work of [VL78a, VL78b] introduced a lifting of a nonlinear optimal control problem into a linear program in the space of measures (the so called weak formulation) and used this lifting along with convex duality theory to establish necessary and sufficient conditions of optimality for this problem. Similar results were obtained for controlled stochastic differential equation in [FV89] with much simplified arguments with the help of mollification techniques; these techniques were later adopted in [Vin93], giving simpler proofs under weaker assumptions of the results of [VL78a, VL78b]. Weak formulations of optimal control problems were also studied independently in many other works, for example in [Rub85] in a deterministic setting, in [Sto90, BB96] in a continuous time and in [HLL96] in a discrete time in a stochastic setting. Later, this lifting was used in [Ran01] for analysis of global stability of nonlinear dynamical systems.

To the best of our knowledge this lifting was not exploited *computationally* until the early 2000's starting with the work [HS00] dealing with stochastic optimal control. In a deterministic setting, the first applications appear to be [PPR04] for stabilizing controller design and [LHPT08] for optimal control.

In all these works, the nonlinear dynamics is replaced by a linear equation on measures or densities. In this thesis we term any such equation a *Liouville's equation* as a reference to the Liouville's equation known from classical mechanics which governs the time evolution of a measure (or a density) transported by the flow of a nonlinear dynamical system, although the use in this thesis is much broader. In particular we work with controlled dynamical systems and we do not use time-dependent measures but rather measures with time as a variable or measures with no time dependence at all capturing averaged-over-time properties of the trajectories starting from a given initial distribution. These measures are called *occupation measures* and have been studied in the stochastic systems literature at least since the work [DK57].

The approximation step in the lift-plus-approximate procedure seeks to approximate, from the inside or from the outside, the cone of nonnegative measures (which are the decision variables in the lifted problem) by easy-to-optimize-over finite-dimensional cones and by replacing the infinite-dimensional Liouville's equation by a suitable finite-dimensional approximation. By approximating the cone of nonnegative measures from the outside by a finite-dimensional cone and by replacing the Liouville's equation by a finite-dimensional truncation, one obtains a finite-dimensional *relaxation* of the original problem, i.e., a problem which has a constraint set looser than the original problem. On the other hand, by approximating the cone of nonnegative measures from the inside by a finite-dimensional cone and by replacing the Liouville's equation by a finite-dimensional linear equation whose satisfaction implies the satisfaction of the Liouville's equation, we obtain a *tightening* of the original problem, i.e., a problem whose set of constraints is tighter than that of the original problem.

There is a range of finite-dimensional approximations to the cone of nonnegative measures and, on the dual side, to the cone of nonnegative functions allowing one to trade off the approximation quality and the complexity of this approximation. In this thesis we use the so-called truncated moment cone and its dual, the truncated quadratic module, which are both semidefinite programming (SDP) representable cones and have been studied extensively in the literature with strong theoretical results available. In particular, these cones are the building block of the hierarchy of semidefinite programming relaxations for static polynomial optimization problems of [Las01].

Contribution and organization

In Chapter 3, Section 3.1, we address the problem of region of attraction (ROA) and reachable set computation; we treat the continuous time version of the problem on a finite time interval. The main contribution is a characterization of the ROA and the reachable set as an infinite-dimensional linear program in the cone of nonnegative measures whose finite-dimensional SDP relaxations provide a converging sequence of outer approximations to the ROA or to the reachable set. To the best of our knowledge, this is the first characterization of these sets as the solution to a convex optimization problem which can be systematically and tractably approximated with convergence guarantees. This chapter is based on the results of [HK14]. A converging sequence of inner approximations to these sets can be obtained by characterizing the complements of these sets rather than the sets itself; this is detailed in [KHJ13] for uncontrolled systems. In addition, a modified formulation which allows for controller extraction from the solutions to the finite-dimensional SDP relaxations is described in [KHJ14a]. These additional results are omitted from the thesis for brevity.

In Chapter 3, Section 3.2, we address the problem of maximum controlled invariant (MCI) set computation computation, where the MCI set is the set of all initial conditions that can be kept forever in the state constraint set using admissible controls. The main contribution is a formulation of the problem of MCI set computation as an infinite-dimensional linear program in the cone of nonnegative measure whose finite-dimensional SDP relaxations provide a converging sequence of outer approximations to the MCI set. Again, to the best of our knowledge, this is the first convex characterization of the MCI set which provides a systematic way of approximating it with convergence guarantees. The added complexity compared to the ROA problem is in the infinite-time nature which requires a different form of the Liouville's equation to be used and analyzed. In addition, we treat both discrete and continuous time systems. The results of this chapter are based on [KHJ14b].

In Chapter 4, we address the problem of optimal feedback controller design. The main contribution is a hierarchy of SDP *tightenings* of a lifted version of this problem whose solutions provide a sequence of rational feedback controllers, which is proven to be asymptotically optimal under certain technical conditions. The main challenge is in tightening the Liouville's equation in such a way that the tightening is feasible irrespective of how coarse it is and at the same time asymptotically equivalent to the Liouville's equation. In addition, as a means of certifying performance of the designed controllers, we describe converging hierarchies of approximations, from above and from below, to the value function attained in an optimal control problem by a given rational feedback controller as well as a hierarchy providing lower bounds on the optimal value function. To the best of our knowledge this is the first SDP hierarchy providing controllers with provable convergence as well as the first convex-optimization-based method for computing value function estimates, both from below and from above, valid globally over the whole state constraint set that is not required to be control invariant. The results of this chapter are based on [KHJ16]. The results of this and the previous chapters were obtained jointly with Didier Henrion.

In Chapter 5, we address the problem of verifying closed-loop properties of a nonlinear (possibly stochastic) discrete-time dynamical system controlled by an optimization-based controller. The main idea is the observation that the Karush-Kuhn-Tucker (KKT) system associated to an optimization problem with polynomial data is a basic semialgebraic set. This allows one to use polynomial optimization techniques to analyze the interconnection of a polynomial dynamical system and such a controller. Properties such as global stability, the ℓ_2 gain or performance with respect to a given infinite-horizon discounted cost function can be analyzed. To the best of our knowledge, this is the first computational method that allows for analysis of optimization-based controllers at this level of generality. This chapter is more practically oriented and no convergence results are given. The results of this chapter are based on [KJ13, KJ15].

In Chapter 6, we discuss computational aspects of the proposed methods and their software implementation.

Other work not included in the thesis manuscript

Besides the work included in the thesis manuscript, the candidate contributed to several other areas of control and optimization.

First, we contributed to the area of stochastic model predictive control (SMPC), where we developed a least-restrictive (in the sense of the size of the feasible set) SMPC algorithm for constrained linear discrete-time systems with additive disturbance [KGJO12, KGOJ14]. This algorithm guarantees the satisfaction of a stochastic constraint on the amount of state-constraint violation averaged over time, where the amount is quantified by a loss function and the averaging can be weighted. The freedom in the choice of the loss function makes this formulation highly flexible – for instance, probabilistic constraints, or integrated chance constraints, can be enforced by an appropriate choice of the loss function. The algorithm exploits the averaged-over-time nature by explicitly taking into account the amount of past constraint violations when determining the current control input, which leads to a significant reduction in conservatism compared to other SMPC schemes. The algorithm enjoys computational complexity, both offline and online, comparable to conventional (nominal or robust) MPC algorithms. This is a joint work with Ravi Gondhalekar and Frauke Oldewurtel.

Second, we developed an algorithmic scheme for solving the infinite-time *constrained* linear quadratic regulation (LQR) problem [SKJ14, SKJ15]. The basic idea is to apply first-order splitting methods (in our case the Alternating Minimization Algorithm or Forward-Backward Splitting) to a suitable reformulation of the problem, where in each iteration we solve an *unconstrained* infinite-horizon LQR problem (whose solution is known analytically) and a simple constrained infinite-horizon problem. We show that each iteration of the algorithm can be carried out using finite amount of memory and computation time and that the algorithm converges to the optimal infinite-horizon LQR solution with the worst-case rates of $O(1/k^2)$ for function values and $O(1/k)$ for the iterates, which are rates optimal for first-order methods. The algorithm requires no invariant sets or terminal weights to be computed and can be efficiently warm started. This is in contrast to most existing MPC schemes which solve a finite-horizon truncation of the problem in a receding horizon fashion and hence are sub-optimal and have to rely on invariant sets or on a difficult-to-estimate horizon length to ensure stability. This is a joint work with Georgios Stathopoulos.

Finally, we investigated turnpike properties of nonlinear optimal control problems. An optimal control problem is said to have a turnpike property on a given set if all optimal solutions starting from initial conditions from that set approach and stay close to a steady state point. Detection of this behavior justifies the use of simple near-optimal controllers which steer the state to this steady-state and stay on it from then on, hence replacing the problem of optimal control by that of stabilization. In addition, existence of this property (and the so-called storage function certifying it) can be used to design a stabilizing economic MPC controller ensuring convergence to the optimal steady state. In [FKJB14, FKJB15], we investigated relationships between the turnpike property, the property of dissipativity with respect to a steady state and the property of optimal steady state operation. In particular we showed that dissipativity with respect to a certain supply rate implies both the existence of a turnpike and optimal operation at a certain steady-state. This in particular implies that dissipativity, which is numerically checkable using sum-of-squares optimization, can be used to establish the existence of the turnpike property. In addition we established converse results under various technical assumptions. This is a joint work with Timm Faulwasser and Dominique Bonvin.

Chapter 2

Preliminaries

Many of the results of this thesis are derived by lifting an optimization problem involving a nonlinear dynamical system into an infinite-dimensional optimization problem on measures. This lifted problem is then approximated by a finite-dimensional optimization problem and tightness of this approximation as its size grows to infinity is theoretically analyzed.

The infinite-dimensional considerations are based on duality between continuous functions and measures on a compact set, which is treated in Section 2.1. The finite-dimensional approximations rely on identifying suitable sub- and super-cones of the cones of nonnegative functions and nonnegative measures; these sub- and super-cones are described in Section 2.2. The lifting of the nonlinear problem into a problem on measures relies on various forms of the Liouville's equation which are derived in Section 2.3.

2.1 Duality between $C(\mathbf{K})$ and $M(\mathbf{K})$

Let $C(\mathbf{K})$ denote the space of all continuous functions defined on a compact set $\mathbf{K} \subset \mathbb{R}^n$ and let $M(\mathbf{K})$ denote the set of all finite signed Borel measures on $\mathbf{K}$. By the Riesz representation theorem (e.g., [Rud86, Theorem 2.14]) we have $C(\mathbf{K})^\star = M(\mathbf{K})$, where $C(\mathbf{K})^\star$ denotes the topological dual of $C(\mathbf{K})$, i.e., the set of all bounded linear functionals on $C(\mathbf{K})$ with the duality pairing between any

$f \in C(\mathbf{K})$ and $\mu \in M(\mathbf{K}) = C(\mathbf{K})^\star$ defined by

$$\langle f, \mu \rangle := \int_{\mathbf{K}} f \, d\mu. \tag{2.1}$$

The same duality holds for the associated positive cones in $C(\mathbf{K})$ and $M(\mathbf{K})$, i.e.,

$$C(\mathbf{K})_+^\star := \left\{ \mu \in M(\mathbf{K}) \;\Big|\; \int_{\mathbf{K}} f \, d\mu \geq 0 \quad \forall\, f \in C(\mathbf{K})_+ \right\} = M(\mathbf{K})_+, \tag{2.2}$$

where $C(\mathbf{K})_+$ denotes the closed convex cone of nonnegative continuous functions on $\mathbf{K}$, $C(\mathbf{K})_+^\star$ is the dual cone of $C(\mathbf{K})_+$ and $M(\mathbf{K})_+$ denotes the closed convex cone of nonnegative Borel measures on $\mathbf{K}$. Similarly we have

$$M(\mathbf{K})_+^\star := \left\{ f \in C(\mathbf{K}) \;\Big|\; \int_{\mathbf{K}} f \, d\mu \geq 0 \quad \forall\, \mu \in M(\mathbf{K})_+ \right\} = C(\mathbf{K})_+, \tag{2.3}$$

where $M(\mathbf{K})_+^\star$ is the dual cone[1] of $M(\mathbf{K})_+$.

2.1.1 Weak-* topology

The duality pairing (2.1) induces the so-called weak-* topology on $M(\mathbf{K})$, which is the coarsest topology on $M(\mathbf{K})$ for which the linear functionals $\langle f\,, \cdot\rangle$ are continuous on $M(\mathbf{K})$ for all $f \in C(\mathbf{K})$, i.e., it is the topology generated by the collection of all sets of the form $\{\mu \in M(\mathbf{K}) \mid \int_{\mathbf{K}} f d\mu \in G\}$ for some open set $G \subset \mathbb{R}$ and some $f \in C(\mathbf{K})$.

A sequence of measures $\mu_k \in M(\mathbf{K})$ converges to a measure $\mu \in M(\mathbf{K})$ in the weak-* topology if and only if

$$\lim_{k\to\infty} \int_{\mathbf{K}} f \, d\mu_k = \int_{\mathbf{K}} f \, d\mu \quad \forall\, f \in C(\mathbf{K}). \tag{2.4}$$

In our setting, i.e., with the set $\mathbf{K}$ compact, this convergence is equivalent to the weak or narrow convergence used in probability theory. Here, however, we do not restrict ourselves to probability measures (i.e., nonnegative measures with unit mass), but work with arbitrary signed Borel measures.

[1] More precisely, $M(\mathbf{K})_+^\star$ should be called the predual cone of $M(\mathbf{K})_+$ since the topological dual of $M(\mathbf{K})$ is much larger than $C(\mathbf{K})$; in particular the topological dual of $M(\mathbf{K})$ contains all bounded measurable functions on $\mathbf{K}$.

2.2 Sub- and super-cones of $C(\mathbf{K})_+$ and $M(\mathbf{K})_+$

Much of the work of this thesis relies on identifying well-structured and easy-to-optimize-over sub-cones or super-cones of $C(\mathbf{K})_+$ and $M(\mathbf{K})_+$ since, being infinite-dimensional, the cones $C(\mathbf{K})_+$ and $M(\mathbf{K})_+$ are not suitable for numerical optimization.

Recalling that for any pair of cones C_1, C_2 we have

$$C_1 \subset C_2 \implies C_1^\star \supset C_2^\star, \tag{2.5}$$

we see that finding sub-cones of $C(\mathbf{K})_+$ yields super-cones of $M(\mathbf{K})_+$ and vice versa.

In order to find easy-to-handle sub- or super-cones we need to impose additional structure on the set $\mathbf{K}$. From now on, we assume that the set $\mathbf{K}$ is given by

$$\mathbf{K} := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0,\ i \in \{1, \dots, n_g\}\} \tag{2.6}$$

for some $g_i \in C(\mathbf{K})$, $i \in \{1, \dots, n_g\}$.

2.2.1 Sub-cones of $C(\mathbf{K})_+$

Given a set $\mathbf{K}$ defined by (2.6), there is a plethora of ways to construct subcones of $C(\mathbf{K})_+$. In particular, any cone of the form

$$\Big\{\sigma_0 + \sum_{i=1}^{n_{\tilde g}} \sigma_i \tilde g_i \mid \sigma_i \in C(\mathbb{R}^n)_+\Big\}, \tag{2.7}$$

where $\tilde g_i \in C(\mathbf{K})$ is an arbitrary finite product of g_i's (or any other function nonnegative on $\mathbf{K}$) is a (still infinite-dimensional) subcone of $C(\mathbf{K})_+$. This representation opens up a number of ways to obtain finite-dimensional sub-cones by restricting the multiplying functions σ_i to a finite-dimensional sub-cone of $C(\mathbb{R}^n)_+$ and by selecting a particular set of functions $\{\tilde g_i\}_{i=1}^{\tilde n_g}$. We will not survey here all the subcones proposed in the literature that stem from the general construction (2.7) but rather focus on the most classical one, which enjoys a good tradeoff between ease of optimization and richness of the set of functions belonging to this cone and in addition has interesting theoretical properties. This cone can be defined for an arbitrary algebra of functions g_i, but from now on we restrict our attention to multivariate polynomials for which strong theoretical results exist and are easily handled by a computer. This cone is the so-called *truncated quadratic module*[2] and is defined by

$$Q_d(\mathbf{K}) := \left\{ \sigma_0 + \sum_{i=1}^{n_g} \sigma_i g_i \mid \sigma_i \in \Sigma,\ \deg(\sigma_i g_i) \le d,\ i \in \{0,\dots,n_g\} \right\}, \tag{2.8}$$

where we have set $g_0 = 1$ and where Σ denotes the set of all sum-of-squares (SOS) polynomials, i.e., polynomials σ of the form $\sigma = \sum_i h_i(x)^2$ for some finitely many polynomials h_i. It is immediate to see that a polynomial $p \in \mathbb{R}[x]_{2k}$ is SOS *if and only if*

$$p(x) = r_k(x)^\top W r_k(x), \quad W \succeq 0, \tag{2.9}$$

where $r_k(x)$ denotes the vector of all monomials in the variable x up to a total degree k (and hence the size of W is $\binom{n+k}{k} \times \binom{n+k}{k}$). The representation (2.9) is sometimes referred to as a Gram matrix representation of an SOS polynomial and we remark that such representation is *not unique*.

From this we immediately conclude that $Q_d(\mathbf{K})$ is a finite-dimensional subcone of $C(\mathbf{K})_+$ of dimension $\binom{n+d}{d}$ (= dimension of $\mathbb{R}[x]_d$) and, importantly, is *semidefinite programming representable*[3] (SDP representable). Indeed, using the Gram matrix representation of SOS polynomials, we deduce that a polynomial $p \in \mathbb{R}[x]_d$ belongs to $Q_d(\mathbf{K})$ if and only if there exist matrices $W_0 \succeq 0, \dots, W_{n_g} \succeq 0$ such that

$$p(x) = r_{d_0}(x)^\top W_0 r_{d_0}(x) + \sum_{i=1}^{n_g} g_i(x) r_{d_i}(x)^\top W_i r_{d_i}(x), \tag{2.10}$$

where $d_0 = \lfloor d/2 \rfloor$ and $d_i = \lfloor (d - \deg g_i)/2 \rfloor$, $i \in \{1,\dots,n_g\}$ (hence the size of each matrix W_i, $i \in \{0,\dots,n_g\}$, is equal to $\binom{n+d_i}{d_i} \times \binom{n+d_i}{d_i}$). By comparing coefficients in equation (2.10) we obtain $\binom{n+d}{d}$ linear equalities between the coefficients of p and the matrices $W_i \succeq 0$ and therefore we conclude that $Q(\mathbf{K})_d$ is indeed SDP representable.

To the best of our knowledge, the SDP representability of $Q(\mathbf{K})$ was first noticed in [CLR95] but was not computationally exploited until the early 2000's in [Nes00,

[2] The symbol $Q_d(\mathbf{K})$ is a slight abuse of notation since the truncated quadratic module depends on the *algebraic description* of the set $\mathbf{K}$ (i.e., on the particular set of functions g_i used to describe it) rather than on the geometry of $\mathbf{K}$. Nevertheless, throughout this thesis we use this notation with the understanding that $Q_d(\mathbf{K})$ refers to the truncated quadratic module generated by those polynomials g_i that were used in the definition of $\mathbf{K}$.

[3] A set in $\mathbb{R}^n$ is semidefinite programming representable if it is a projection of the feasible set of a linear matrix inequality (i.e., a projection of a spectrahedron).

Las01, Par03].

One may ask why using the truncated quadratic module $Q_d(\mathbf{K})$ and not simply $P_d(\mathbf{K})_+$, the cone of all polynomials nonnegative on $\mathbf{K}$ of total degree no more than d. The answer simple – except for a few special cases (e.g., the scalar case), the cone $P_d(\mathbf{K})_+$ has no computationally tractable representation, despite being finite-dimensional. The fundamental question of quantifying the discrepancy between $Q_d(\mathbf{K}) \subset P_d(\mathbf{K})_+$ and $P_d(\mathbf{K})_+$ is not well understood at present, even in simple cases (e.g., $\mathbf{K}$ being a box or a ball). A notable exception is the case $\mathbf{K} = \mathbb{R}^n$, where [Ble06] proved that this discrepancy becomes large (in the sense of normalized volume in the space of coefficients) as the dimension n tends to infinity while the degree d is held fixed.

Despite the lack of quantitative understanding of this discrepancy, a fundamental asymptotic result holds provided that the functions defining the set $\mathbf{K}$ satisfy the Archimedianity condition:

Definition 1 *A set* $\mathbf{K}$ *defined by (2.6) satisfies the Archimedianity condition if there exists an* $N \geq 0$ *and a* $d \geq 0$ *such that* $N - \|x\|_2^2 \in Q_d(\mathbf{K})$.

We remark that the assumption of $\mathbf{K}$ satisfying the Archimedianity condition is made without loss of generality since $\mathbf{K}$ is assumed compact and hence a redundant ball constraint can be added to the description of $\mathbf{K}$, making the condition hold trivially.

Theorem 2.2.1 ([Put93]) *If* $p \in \mathbb{R}[x]$ *is strictly positive on a basic semialgebraic set*[4] $\mathbf{K}$ *satisfying the Archimedianity condition, then* $p \in Q_d(\mathbf{K})$ *for some* $d \geq 0$.

We remark that in Theorem 2.2.1 the degree d for which p belongs to $\in Q_d(\mathbf{K})$ can be far greater than $\deg(p)$ and in fact the discrepancy between the two can be arbitrarily large.

An immediate but crucial corollary for this work is the following:

Corollary 2.2.1 *Let a basic semialgebraic set* $\mathbf{K}$ *satisfy the Archimedianity condition. If* $f \in C(\mathbf{K})_+$*, then for any* $\varepsilon > 0$ *there exists a* $d \geq 0$ *and a polynomial* $p \in Q_d(\mathbf{K})$ *such that* $\|f - p\|_{C(\mathbf{K})} \leq \varepsilon$.

[4] A basic semialgebraic set is a set of the form (2.6) with g_i polynomial, i.e., it is the intersection of finitely many polynomial sub- or super-level sets.

Corollary 2.2.1 says that $Q_d(\mathbf{K})$ is dense in $C(\mathbf{K})_+$ as d tends to infinity, i.e.,

$$\overline{\bigcup_{d\geq 0} Q_d(\mathbf{K})} = C(\mathbf{K})_+,$$

where the closure is taken with respect to the standard uniform topology on $C(\mathbf{K})$.

2.2.2 Super-cones of $M(\mathbf{K})_+$

Using the results of the previous section and the cone duality relation (2.5) we immediately obtain finite-dimensional super-cones of $M(\mathbf{K})_+$ as $Q_d(\mathbf{K})^\star$. Henceforth we set $M_d^{\sup}(\mathbf{K})_+ := Q_d(\mathbf{K})^\star$ and call $M_d^{\sup}(\mathbf{K})_+$ the *truncated moment cone.*

The duality pairing between the two finite-dimensional cones $Q_d(\mathbf{K})$ and $M_d^{\sup}(\mathbf{K})_+$ is inherited from the infinite-dimensional case (2.1). Given a polynomial $p \in \mathbb{R}[x]_d \supset Q_d(\mathbf{K})$ expressed in a multivariate monomial basis as $p = \sum_{|\alpha|\leq d} \boldsymbol{p}_\alpha x^\alpha$, where $\boldsymbol{p} \in \mathbb{R}^{\binom{n+d}{d}}$ is the vector of coefficients of p and given a vector $\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \supset M_d^{\sup}(\mathbf{K})_+$, this duality pairing is given by

$$\langle p, \boldsymbol{y}\rangle = \sum_{|\alpha|\leq d} \boldsymbol{p}_\alpha \boldsymbol{y}_\alpha = \boldsymbol{p}^\top \boldsymbol{y},$$

which boils down to (2.1), i.e., to integration of p with respect to a measure μ, if the vector $\boldsymbol{y}$ is the truncated moment vector of some measure $\mu \in M(\mathbf{K})$, i.e., if $\boldsymbol{y}_\alpha = \int_{\mathbf{K}} x^\alpha \, d\mu$ for all multiindices $\alpha \in \mathbb{N}^n$ such that $|\alpha| := \sum_{i=1}^n \alpha_i \leq d$.

With these considerations, the cone $M_d^{\sup}(\mathbf{K})_+$ is given explicitly as

$$M_d^{\sup}(\mathbf{K})_+ = \left\{ \boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \langle p, \boldsymbol{y}\rangle \geq 0 \ \ \forall\, p \in Q_d(\mathbf{K}) \right\}$$

and we see that indeed $M_d^{\sup}(\mathbf{K})_+ \supset M(\mathbf{K})_+$ for all $d \geq 0$ in the sense that any measure $\mu \in M(\mathbf{K})_+$ gives rise to an element $\boldsymbol{y} \in M_d^{\sup}(\mathbf{K})_+$ defined by $\boldsymbol{y}_\alpha = \int_{\mathbf{K}} x^\alpha \, d\mu(x)$ for all α such that $|\alpha| \leq d$. Therefore

$$M_d^{\sup}(\mathbf{K})_+ \supset M_{d+1}^{\sup}(\mathbf{K})_+ \supset M(\mathbf{K})_+ \quad \forall\, d \geq 0,$$

where the first inclusion is to be understood in the sense that the vector $\tilde{\boldsymbol{y}}$ defined by the first $\binom{n+d}{d}$ elements of any vector $\boldsymbol{y} \in M_{d+1}^{\sup}(\mathbf{K})_+$ satisfies $\tilde{\boldsymbol{y}} \in M_d^{\sup}(\mathbf{K})_+$.

To derive an explicit SDP representation for $M_d^{\sup}(\mathbf{K})_+$ we recall that for any two cones we have $(C_1 + C_2)^\star = C_1^\star \cap C_2^\star$, where + denotes the elmentwise (or Minkowski) addition. Since

$$Q_d(\mathbf{K}) = \Sigma_{2d_0} + g_1\Sigma_{2d_1} + \ldots + g_{n_g}\Sigma_{2d_{n_g}},$$

where Σ_k denotes the set of all SOS polynomials of degree no more than k and d_i's are defined below Eq. (2.10), we conclude that

$$M_d^{\sup}(\mathbf{K})_+ = Q_d(\mathbf{K})^\star = \Sigma_{2d_0}^\star \cap (g_1\Sigma_{2d_1})^\star \cap \ldots \cap (g_{n_g}\Sigma_{2d_{n_g}})^\star \tag{2.11}$$

and hence it suffices to understand the dual cone of $g\Sigma_{2\bar{d}}$ for a polynomial g, where $\bar{d} = \lfloor (d - \deg g)/2 \rfloor$. A direct computation gives

$$\begin{aligned}
(g\Sigma_{2\bar{d}})^\star &= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \langle p, \boldsymbol{y}\rangle \geq 0 \;\; \forall\, p \in g\Sigma_{2\bar{d}}\} \\
&= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \langle gh^2, \boldsymbol{y}\rangle \geq 0 \;\; \forall\, h \in \mathbb{R}[x]_{\bar{d}}\} \\
&= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \langle g\boldsymbol{h}^\top r_{\bar{d}} r_{\bar{d}}^\top \boldsymbol{h}, \boldsymbol{y}\rangle \geq 0 \;\; \forall\, \boldsymbol{h} \in \mathbb{R}^{\binom{n+\bar{d}}{\bar{d}}}\} \\
&= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \boldsymbol{h}^\top \langle g r_{\bar{d}} r_{\bar{d}}^\top, \boldsymbol{y}\rangle \boldsymbol{h} \geq 0 \;\; \forall\, \boldsymbol{h} \in \mathbb{R}^{\binom{n+\bar{d}}{\bar{d}}}\} \\
&= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid \langle g r_{\bar{d}} r_{\bar{d}}^\top, \boldsymbol{y}\rangle \succeq 0\} \\
&= \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid M_d(\boldsymbol{y}, g) \succeq 0\},
\end{aligned} \tag{2.12}$$

where $\langle \cdot\,, \boldsymbol{y}\rangle$ acts componentwise on matrix arguments, $r_{\bar{d}}$ is the vector of all multivariate monomials up to degree $\bar{d}$ and the matrix

$$M_d(\boldsymbol{y}, g) := \langle g\, r_{\bar{d}} r_{\bar{d}}^\top, \boldsymbol{y}\rangle$$

is called the *(truncated) localizing matrix*. The matrix

$$M_d(\boldsymbol{y}) := M_d(\boldsymbol{y}, 1) = \langle r_{\bar{d}} r_{\bar{d}}^\top, \boldsymbol{y}\rangle$$

is called the *(truncated) moment matrix*. The size of $M_d(\boldsymbol{y}, g)$ is $\binom{n+\bar{d}}{\bar{d}} \times \binom{n+\bar{d}}{\bar{d}}$ (notice the dependence of the size both on d and $\deg(g)$). More importantly, notice that the matrix $M_d(\boldsymbol{y}, g)$ depends *linearly* on the truncated moment vector $\boldsymbol{y}$.

The characterization of $(g\Sigma_{2\bar{d}})^\star$ in (2.12) and (2.11) yield immediately a characterization of the truncated moment cone $M_d^{\sup}(\mathbf{K})_+$ as

$$M_d^{\sup}(\mathbf{K})_+ = \{\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}} \mid M_d(\boldsymbol{y}) \succeq 0,\; M_d(\boldsymbol{y}, g_i) \succeq 0,\; i \in \{1, \ldots, n_g\}\} \tag{2.13}$$

which is the desired explicit SDP representation of $M_d^{\sup}(\mathbf{K})_+$.

The discrepancy between $M_d^{\sup}(\mathbf{K})_+ \supset M(\mathbf{K})_+$ and $M(\mathbf{K})_+$ is quantitatively not understood at present (even less than for the case of $Q_d(\mathbf{K})$ and $C(\mathbf{K})_+$) but again an important asymptotic result holds:

Theorem 2.2.2 ([Put93]) *If a basic semialgebraic set* $\mathbf{K}$ *satisfies the Archimedianity condition and if a sequence of real numbers* $\boldsymbol{y}$ *satisfies* $M_d(\boldsymbol{y}) \succeq 0$, $M_d(\boldsymbol{y}, g_i) \succeq 0$ *for all* $i \in \{1, \ldots, n_g\}$ *and for all* $d \geq 0$, *then there exists a unique measure* $\mu \in M(\mathbf{K})_+$ *such that*

$$\boldsymbol{y}_\alpha = \int_{\mathbf{K}} x^\alpha \, d\mu \tag{2.14}$$

for all multiindices $\alpha \in \mathbb{N}^n$, *i.e.,*

$$\bigcap_{d \geq 0} M_d^{\sup}(\mathbf{K})_+ = M(\mathbf{K})_+ \, .$$

The measure μ from Theorem 2.2.2 is called the *representing measure* of the moment sequence $\boldsymbol{y}$ and we remark that the statement (2.14) is equivalent to $\langle p, \boldsymbol{y} \rangle = \int_{\mathbf{K}} p \, d\mu$ for all polynomials p.

2.2.3 Sub-cones of $M(\mathbf{K})_+$

A natural way of obtaining easy-to-optimize-over sub-cones of $M(\mathbf{K})_+$ is to restrict the measures in $M(\mathbf{K})_+$ to measures that have a density $\rho \in Q_d(\mathbf{K})$ with respect to some reference measure $\bar{\mu}$ with known moments over $\mathbf{K}$ (e.g., the Lebesgue measure). That is, we define

$$M_d^{\mathrm{sub}}(\mathbf{K})_+ := \left\{ \mu \in M(\mathbf{K})_+ \;\middle|\; \exists \, \rho \in Q_d(\mathbf{K}) \text{ s.t. } \int_{\mathbf{K}} f \, d\mu = \int_{\mathbf{K}} f \rho \, d\bar{\mu}, \; \forall \, f \in C(\mathbf{K}) \right\}. \tag{2.15}$$

The cones $Q_d(\mathbf{K})$ and $M_d^{\mathrm{sub}}(\mathbf{K})_+$ are isomorphic and optimizing over $M_d^{\mathrm{sub}}(\mathbf{K})_+$ amounts to optimizing over $Q_d(\mathbf{K})$, i.e., to semidefinite programming. Since $Q_d(\mathbf{K}) \subset Q_{d+1}(\mathbf{K})$, the cones $M_d^{\mathrm{sub}}(\mathbf{K})_+$ satisfy

$$M_d^{\mathrm{sub}}(\mathbf{K})_+ \subset M_{d+1}^{\mathrm{sub}}(\mathbf{K})_+ \subset M(\mathbf{K})_+ \quad \forall \, d \geq 0.$$

Combining Corollary 2.2.1 with the fact that measures with densities in $C(\mathbf{K})_+$ are dense in the weak-$\star$ topology in $M(\mathbf{K})_+$ provided that the support[5] of the reference measure $\operatorname{spt}\bar{\mu}$ is equal to $\mathbf{K}$, we obtain the following:

Corollary 2.2.2 *If* $\operatorname{spt}\bar{\mu} = \mathbf{K}$ *and the basic semialgebraic set* $\mathbf{K}$ *satisfies the Archimedianity condition, then*

$$\overline{\bigcup_{d\geq 0} M_d^{\text{sub}}(\mathbf{K})_+} = M(\mathbf{K})_+, \tag{2.16}$$

where the closure is with respect to the weak-$\star$ *topology on* $M(\mathbf{K})$.

The statement (2.16) has the following interpretation: For every $\mu \in M(\mathbf{K})_+$ there exists a sequence $\Big(\rho_d \in Q_d(\mathbf{K})\Big)_{d\geq 0}$ such that

$$\lim_{d\to\infty} \int_{\mathbf{K}} f(x)\rho_d(x)d\bar{\mu}(x) = \int_{\mathbf{K}} f(x)\,d\mu(x) \quad \forall f \in C(\mathbf{K}).$$

2.2.4 Super-cones of $C(\mathbf{K})_+$

Super-cones of $C(\mathbf{K})_+$, denoted by $C_d^{\text{sup}}(\mathbf{K})_+$, are obtained by dualizing $M_d^{\text{sub}}(\mathbf{K})_+$ using the duality relationship (2.5). A direct computation gives

$$\begin{aligned} C_d^{\text{sup}}(\mathbf{K})_+ &:= \Big(M_d^{\text{sub}}(\mathbf{K})_+\Big)^\star = \Big\{ f \in C(\mathbf{K}) \mid \int_{\mathbf{K}} f d\mu \geq 0 \; \forall \mu \in M_d^{\text{sub}}(\mathbf{K})_+ \Big\} \\ &= \Big\{ f \in C(\mathbf{K}) \mid \int_{\mathbf{K}} fp\, d\bar{\mu} \geq 0 \; \forall p \in Q_d(\mathbf{K}) \Big\} \\ &= \Big\{ f \in C(\mathbf{K}) \mid M_d(\boldsymbol{y}^{f\bar{\mu}}) \succeq 0,\; M_d(\boldsymbol{y}^{f\bar{\mu}}, g_i) \succeq 0,\; i \in \{1,\ldots,n_g\} \Big\}, \end{aligned}$$

where $\bar{\mu}$ is a given reference measure and $\boldsymbol{y}^{f\bar{\mu}}$ is the moment sequence of the measure $f d\bar{\mu}$, i.e.,

$$\boldsymbol{y}_\alpha^{f\bar{\mu}} = \int_{\mathbf{K}} x^\alpha f(x) d\bar{\mu}(x) \tag{2.17}$$

for all multiindices $\alpha \in \mathbb{N}^n$. The cones $C_d^{\text{sup}}(\mathbf{K})_+$ satisfy

$$C_d^{\text{sup}}(\mathbf{K})_+ \supset C_{d+1}^{\text{sup}}(\mathbf{K})_+ \supset C(\mathbf{K})_+ \quad \forall\, d \geq 0.$$

[5] The support of a nonnegative measure μ is the smallest closed set whose complement has a zero measure. Equivalently, $x \in \operatorname{spt}\mu$ if an only if $\mu(U) > 0$ for every open neighborhood U of x.

Given that the moments $\boldsymbol{y}^{f\bar{\mu}}$ are available, testing whether a given $f \in C(\mathbf{K})$ belongs to $C_d^{\mathrm{sup}}(\mathbf{K})_+$ amounts to solving a semidefinite programming feasibility problem.

Noticing that the cone $C_d^{\mathrm{sup}}(\mathbf{K})_+$ is isomorphic to $M_d^{\mathrm{sup}}(\mathbf{K})_+$, we can also optimize over $C_d^{\mathrm{sup}}(\mathbf{K})_+$ by optimizing directly over the truncated moment sequences $\boldsymbol{y}^{f\bar{\mu}} \in \mathbb{R}^{\binom{n+d}{d}}$ and then finding an $f \in C(\mathbf{K})$ such that (2.17) holds for all α such that $|\alpha| \leq d$. Finding such f is easily done whenever the moment matrix $M_{2d}(\boldsymbol{y}^{\bar{\mu}})$ associated to the moment sequence $\boldsymbol{y}^{\bar{\mu}}$ of the measure $\bar{\mu}$ is non-singular (a sufficient condition for this is, e.g., $\mathbf{K}$ having a nonempty interior and the Lebesgue measure being absolutely continuous with respect to $\bar{\mu}$). Then a polynomial $f \in \mathbb{R}[x]_d$ suffices with coefficients given by $\boldsymbol{f} = M_{2d}(\boldsymbol{y}^{\bar{\mu}})^{-1}\boldsymbol{y}^{f\bar{\mu}}$. This follows from the definition of $M_d(\cdot)$ which implies that $M_{2d}(\boldsymbol{y}^{\bar{\mu}})\boldsymbol{f} = \boldsymbol{y}^{f\bar{\mu}} \in \mathbb{R}^{\binom{n+d}{d}}$, with $\boldsymbol{y}^{f\bar{\mu}}$ defined by (2.17), for any polynomial $f \in \mathbb{R}[x]_d$ with coefficient vector $\boldsymbol{f}$ and any vector of moments $\boldsymbol{y}^{\bar{\mu}}$ of degree no more than d of a measure $\bar{\mu}$.

Theorem 2.2.1 leads immediately to the following corollary:

Corollary 2.2.3 *Let* $\mathbf{K}$ *be a basic semialgebraic set satisfying the Archimedianity condition and let the reference measure* $\bar{\mu}$ *satisfy* $\operatorname{spt}\bar{\mu} = \mathbf{K}$. *Then the following holds: if* $f \in C_d^{\mathrm{sup}}(\mathbf{K})_+$ *for all* $d \geq 0$, *then* $f \in C(\mathbf{K})_+$, *i.e.,*

$$\bigcap_{d\geq 0} C_d^{\mathrm{sup}}(\mathbf{K})_+ = C(\mathbf{K})_+.$$

This Corollary states that the approximations $C_d^{\mathrm{sup}}(\mathbf{K})_+$ to $C(\mathbf{K})_+$ are asymptotically tight provided that the support of the reference measure equals $\mathbf{K}$ and $\mathbf{K}$ satisfies the Archimedianity condition. In other words the obvious necessary condition for nonnegativity of f on $\mathbf{K}$, $f \in C_d^{\mathrm{sup}}(\mathbf{K})_+$ for *some* $d \geq 0$, is also sufficient if we require it to hold for *all* $d \geq 0$, under the above-stated assumptions.

Developments of this and the preceding section were inspired by the work [Las11], where similar reasoning was used to define SDP representable outer-approximations to the cone of nonnegative polynomials of a given degree on a closed set.

2.3 Lifting nonlinear dynamics

In this section we describe a lifting of a nonlinear controlled dynamical system into the space of measures. The lifting is described both for continuous-time dynamics

$$\dot{x}(t) = f(x(t), u(t)) \tag{2.18}$$

and for discrete-time dynamics

$$x_{t+1} = f(x_t, u_t), \tag{2.19}$$

where $x \in \mathbb{R}^n$ is the state of the system and $u \in \mathbb{R}^m$ the control input, which is constrained to lie in a compact set $\mathbf{U}$. Unless specified otherwise, throughout this section we only assume that f is measurable and bounded on every compact set, both in continuous and discrete time.

The goal is to lift the *nonlinear* dynamics into a *linear* equation on measures. This lifting leads to a different equation depending on whether we work in continuous or discrete-time and whether we work on a finite or infinite time horizon.

First we deal with a technical issue arising in continuous time.

2.3.1 Relaxed controls

In continuous time, the lifting presented will in general not be equivalent to the original dynamics (2.18) but rather to the *relaxed* differential inclusion

$$\dot{x}(t) \in \overline{\mathrm{conv}}\, f(x(t), \mathbf{U}), \tag{2.20}$$

where $\overline{\mathrm{conv}}$ denotes the closed convex hull. Trajectories of this differential inclusion are in a one-to-one correspondence with the trajectories of

$$\dot{x}(t) = \int_{\mathbf{U}} f(x(t), u)\, d\gamma_t(u), \tag{2.21}$$

where $\gamma_t \in P(\mathbf{U})$ is called a *relaxed* or measure-valued control. Here $P(\mathbf{U})$ is the set of all probability measures, i.e., nonnegative measures with unit mass.

To see the equivalence, observe that any point $y \in \overline{\mathrm{conv}}\, f(x(t), \mathbf{U})$ can be obtained as a convex combination of finitely many points $f(x(t), u_i(t))$, $u_i(t) \in \mathbf{U}$, with weights $w_i(t)$; defining $\gamma_t = \sum_i w_i(t)\delta_{u_i(t)}$, where $\delta_{u_i(t)}$ denotes the Dirac measure at $u_i(t)$, we see that $y = \int_{\mathbf{U}} f(x(t), u)\, d\gamma_t(u) = \sum_i w_i(t) f(x(t), u_i(t))$ as desired. In the other direction, the equivalence follows by approximating an arbitrary probability measure γ_t by a finite convex combination of Dirac measures and taking the limit, using the closedness of $\overline{\mathrm{conv}}\, f(x(t), \mathbf{U})$ and the density of convex combinations of Dirac measures in the space of all probability measures (in the weak-$\star$ topology).

Clearly, any trajectory of the original system (2.18) is also a trajectory of the relaxed systems (2.20) and (2.21). Importantly, the Filippov-Ważewski Theorem [AF09, Theorem 10.4.4, Corollary 10.4.5] gives a partial converse. This theorem states that the trajectories of the relaxed system (2.20) and (2.21) can be arbitrarily closely approximated by the trajectories of the original system (2.18) in the sense that the trajectories of (2.18) are dense (with respect to the supremum norm on $C([0, T]; \mathbb{R}^n)$) in the set of trajectories of (2.20) and (2.21) on any finite interval $[0, T]$.

This theorem justifies the use of the relaxed dynamics instead of the original one in practical applications, whenever working with the relaxed dynamics is more convenient. Nevertheless, contrived examples can be constructed, where the solution to a problem (e.g., the optimal value of an optimization problem or the size of the region of attraction) differs depending on whether the relaxed or non-relaxed dynamics is used. A typical such problem would be one where the state of a single integrator ($\dot{x} = u$) is constrained to stay at zero ($\mathbf{X} = \{0\}$) and the control input is constrained to the discrete set $\mathbf{U} = \{-1, 1\}$. The relaxed dynamics (2.21) can fulfill such constraint with the relaxed control given by $\gamma_t = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$, which corresponds to chattering between the two control values infinitely quickly. The non-relaxed control cannot chatter infinitely quickly and therefore cannot fulfill this constraint. It is, in general, precisely this chattering phenomenon that lies behind the discrepancy between the trajectories of the relaxed and non-relaxed dynamics.

Obvious sufficient conditions for a full equivalence of trajectories[6], i.e., for trajectories of (2.18), (2.20) and (2.21) to coincide are

- $f(x, \mathbf{U})$ closed and convex for all x,
- input affine dynamics $\dot{x} = f(x) + g(x)u$ with $\mathbf{U}$ closed and convex,

[6] Note that the fact that the set of trajectories of the relaxed and non-relaxed dynamics coincides does not imply that the optimal value of an optimal control problem involving these dynamics is the same. Problems exhibiting this behavior are, however, highly contrived and have necessarily cost function non-convex in u (otherwise chattering would imply higher rather than lower cost function through Jensen's inequality).

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

The input-affine (sometimes called control-affine) dynamics offers, besides the equivalence between the trajectories of the relaxed and non-relaxed dynamics, also computational advantages and is discussed separately in Section 2.3.5.

Now we proceed to describe the liftings of the (relaxed) dynamics to an infinite-dimensional space of measures.

2.3.2 Continuous-time, finite-horizon

Consider any trajectory $x(\cdot\,|\,x_0)$ of the relaxed system (2.21) starting from an initial condition x_0 generated by a relaxed control $\gamma_t(\cdot \mid x_0) \in P(\mathbf{U})$, $t \in [0,T]$.

Then we define the *conditional occupation measure* $\mu(\cdot \mid x_0) \in M([0,T] \times \mathbb{R}^n \times \mathbf{U})_+$ by

$$\mu(A \times B \times C \mid x_0) := \int_0^T \int_{\mathbf{U}} I_{A\times B\times C}(t, x(t\,|\,x_0), u)\, d\gamma_t(u\,|\,x_0)\, dt \tag{2.22}$$

for all sets[7] $A \subset [0,T]$, $B \subset \mathbb{R}^n$ and $C \subset \mathbf{U}$.

In words, the quantity $\mu(A \times B \times C \mid x_0)$ equals to the amount of time spent by the relaxed state-control trajectory $(x(\cdot\,|\,x_0), \gamma_\cdot(\cdot\,|\,x_0))$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$ during $A \subset [0,T]$.

Now suppose that the initial condition is not a single point but an *initial measure*[8] $\mu_0 \in M(\mathbb{R}^n)_+$ and a state trajectory along with an admissible relaxed control input generating it is associated to each initial condition from the support of μ_0. Then we define the *occupation measure* $\mu \in M([0,T] \times \mathbb{R}^n \times \mathbf{U})_+$ by

$$\mu(A \times B \times C) := \int_{\mathbb{R}^n} \mu(A \times B \times C \mid x_0)\, d\mu_0(x_0) \tag{2.23}$$

for all $A \subset [0,T]$, $B \subset \mathbb{R}^n$ and $C \subset \mathbf{U}$. The quantity $\mu(A \times B \times C)$ equals to the average amount of time spent by the relaxed state-control trajectories

[7] Throughout the thesis we somewhat downplay the role of measurability. In particular whenever we write "for all sets" we mean "for all Borel measurable sets".

[8] The initial measure μ_0 can be thought of as the probability distribution of the initial state, although we do not require the mass of μ_0 to be normalized to one.

$(x(\cdot\,|\,x_0), \gamma_\cdot(\cdot\,|\,x_0))$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$ during $A \subset [0,T]$, where the averaging is over the distribution of the initial conditions μ_0.

Lastly, we define the *final measure* as

$$\mu_T(B) := \int_{\mathbb{R}^n} I_B(x(T\,|\,x_0))\, d\mu_0(x_0) \tag{2.24}$$

for all $B \subset \mathbb{R}^n$. This measures captures the distribution of the state at time T after it has been propagated by the relaxed control system (2.21) starting at time zero from the initial measure μ_0.

Now we derive an equation linking the measures μ_0, μ and μ_T, which will be the sought lifting, for the finite time interval $[0,T]$, of the nonlinear relaxed dynamics (2.21) into a *linear* equation on measures. This equation will play a key role in subsequent developments, in particular in Section 3.1 on region of attraction and reachable set approximation. To derive the equation, fix an initial condition $x_0 \in \mathbb{R}^n$ and a trajectory $x(\cdot\,|\,x_0)$ generated by a relaxed control $\gamma_\cdot(\cdot\,|\,x_0)$ satisfying $\gamma_t(\cdot\,|\,x_0) \in P(\mathbf{U})$ for all $t \in [0,T]$. Then for any $v \in C^1([0,T] \times \mathbb{R}^n)$ we have

$$\begin{aligned}\frac{d}{dt}v(t,x(t\,|\,x_0)) &= \frac{\partial v}{\partial t}(t,x(t\,|\,x_0)) + \nabla v(t,x(t\,|\,x_0)) \cdot \frac{d}{dt}x(t\,|\,x_0)\\ &= \frac{\partial v}{\partial t}(t,x(t\,|\,x_0)) + \nabla v(t,x(t\,|\,x_0)) \cdot \int_{\mathbf{U}} f(x(t\,|\,x_0),u)\, d\gamma_t(u\,|\,x_0).\end{aligned}$$

Integrating from 0 to T and using the definition of the conditional occupation measure we obtain

$$\begin{aligned}v(T,x(T\,|\,x_0)) - v(0,x_0) &= \int_0^T \int_{\mathbf{U}} \frac{\partial v}{\partial t} + \nabla v(t,x(t\,|\,x_0)) \cdot f(x(t\,|\,x_0),u)\, d\gamma_t(u\,|\,x_0)\\ &= \int_{[0,T]\times\mathbb{R}^n\times\mathbf{U}} \frac{\partial v}{\partial t} + \nabla v(t,x) \cdot f(x,u)\, d\mu(t,x,u\,|\,x_0).\end{aligned}$$

Integrating with respect to μ_0 and using the definition of the occupation and final measures we get for all $v \in C^1([0,T] \times \mathbb{R}^n)$

$$\int_{\mathbb{R}^n} v(T,x)\, d\mu_T(x) = \int_{\mathbb{R}^n} v(0,x)\, d\mu_0(x) + \int\limits_{[0,T]\times\mathbb{R}^n\times\mathbf{U}} \frac{\partial v}{\partial t} + \nabla v(t,x) \cdot f(x,u)\, d\mu(t,x,u). \tag{2.25}$$

Equation (2.25) is called Liouville's equation. It is a *linear* equation in the variables (μ, μ_0, μ_T) and it is a lifting, on the time interval $[0,T]$, of the nonlinear relaxed dynamics (2.20) and (2.21) into the space of measures.

Remark 2.3.1 *The function v appearing in Eq. (2.25) is called a* test function *and we stress that v is not a variable in the equation, but rather the equation is required to hold for all test functions v from a suitable class of functions, in the case of Eq. (2.25) for all $v \in C^1([0,T] \times \mathbb{R}^n)$.*

It follows from the above discussion that any family of relaxed state-control trajectories starting from a given initial measure μ_0 gives rise to an occupation measure μ and a final measure μ_T such that the triplet (μ, μ_0, μ_T) satisfies (2.25). Importantly, the converse is true as well in the sense that for any triplet of measures (μ, μ_0, μ_T) satisfying (2.25) there exists a family of trajectories of (2.21) starting from μ_0 that generates μ and μ_T through (2.22), (2.23), (2.24). This is formalized in the following theorem.

Theorem 2.3.1 *If a triplet of nonnegative compactly supported finite measures*

$$(\mu, \mu_0, \mu_T) \in M([0,T] \times \mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+$$

satisfies (2.25) for all $v \in C^1([0,T] \times \mathbb{R}^n)$, then there exists a measure

$$\eta \in M(C([0,T];\mathbb{R}^n))_+$$

supported on the absolutely continuous trajectories of (2.20) such that

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

$$\mu_0(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(0))\, d\eta(x(\cdot))$$

and

$$\mu_T(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(T))\, d\eta(x(\cdot))$$

for all $A \subset \mathbb{R}$ and $B \subset \mathbb{R}^n$.

The proof is given in Appendix B.1 and is based on Ambrosio's superposition principle [Amb08, Theorem 3.2]. A particular case of the theorem for $\mu_0 = \delta_{x_0}$ was know at least since the work [VL78a].

Remark 2.3.2 *Note that in Theorem 2.3.1 we allow for superposition of multiple trajectories starting from the same initial condition; there can be many such trajectories even if f is Lipschitz (which we do not assume) since we are working in a* *controlled setting. Therefore we refer to this and related theorems stated below as superposition theorems.*

Remark 2.3.3 *Note that the theorem is stated only in terms of the state trajectories of (2.20). This was done for simplicity of the statement; the family of relaxed control trajectories generating the state trajectories of (2.20) can be readily backed out from the state trajectories (see the discussion below Eq. (2.21)).*

2.3.3 Continuous-time infinite horizon

Consider any trajectory $x(\cdot\,|\,x_0)$ of the relaxed system (2.21) starting from an initial condition x_0 generated by a relaxed control $\gamma_t(\cdot \mid x_0) \in P(\mathbf{U})$, $t \in [0, \infty)$.

Then, given a discount factor $\beta > 0$, we define the *conditional discounted occupation measure* $\mu(\cdot \mid x_0) \in M(\mathbb{R}^n \times \mathbf{U})_+$ as

$$\mu(B \times C \mid x_0) := \int_0^\infty \int_{\mathbf{U}} e^{-\beta t} I_{A\times B}(x(t\,|\,x_0), u)\, d\gamma_t(u\,|\,x_0)\, dt \tag{2.26}$$

for all sets $B \subset \mathbb{R}^n$ and $C \subset \mathbf{U}$.

The quantity $\mu(B \times C \mid x_0)$ is equal to the (discounted) time spent by the relaxed state-control trajectory $(x(\cdot\,|\,x_0), \gamma_\cdot(\cdot\,|\,x_0))$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$. The discounting in the definition of the occupation measure ensures that $\mu(B \times C \mid x_0)$ is always finite; in fact we have $\mu(\mathbb{R}^n \times \mathbf{U} \mid x_0) = \beta^{-1}$.

Now suppose that the initial condition is not a single point but an *initial measure*[9] $\mu_0 \in M(\mathbb{R}^n)_+$ and a state trajectory along with an admissible relaxed control input generating it is associated to each initial condition from the support of μ_0. Then we define the *discounted occupation measure* $\mu \in M(\mathbb{R}^n \times \mathbf{U})_+$ as

$$\mu(B \times C) := \int_{\mathbb{R}^n} \mu(B \times C \mid x_0)\, d\mu_0(x_0) \tag{2.27}$$

for all $B \subset \mathbb{R}^n$, $C \subset \mathbf{U}$. The quantity $\mu(B\times C)$ is equal to the average (discounted) amount of time spent by the relaxed state-control trajectories $(x(\cdot\,|\,x_0), \gamma_\cdot(\cdot\,|\,x_0))$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$, where the averaging is over the distribution of the initial conditions μ_0.

[9] The initial measure μ_0 can be thought of as the probability distribution of the initial state, although we do not require the mass of μ_0 to be normalized to one.

Now we derive an equation linking the measures μ_0 and μ, which will be the sought lifting, for the time interval $[0,\infty)$, of the nonlinear relaxed dynamics (2.21) into a *linear* equation on measures. This equation will play a key role in subsequent developments, in particular in Section 3.2 on maximum controlled invariant set approximation. To derive the equation, fix an initial condition $x_0 \in \mathbb{R}^n$ and a trajectory $x(\cdot\,|\,x_0)$ generated by a relaxed control $\gamma_\cdot(\cdot\,|\,x_0)$ satisfying $\gamma_t(\cdot\,|\,x_0) \in P(\mathbf{U})$ for all $t \in [0,\infty)$.

Let $C_b^1(\mathbb{R}^n)$ denote the space of all bounded continuously differentiable functions on $\mathbb{R}^n$. Then for any $v \in C_b^1(\mathbb{R}^n)$ integration by parts yields

$$\begin{aligned}
\int_{\mathbb{R}^n\times\mathbf{U}} \nabla v\cdot f(x,u)\,d\mu(x,u\,|\,x_0) &= \int_0^\infty \int_{\mathbf{U}} e^{-\beta t}\nabla v\cdot f(x(t\mid x_0),u)\,d\gamma_t(u\,|\,x_0)\,dt \\
&= \int_0^\infty e^{-\beta t}\frac{d}{dt}v(x(t\,|\,x_0))\,dt \\
&= \beta\int_0^\infty e^{-\beta t}v(x(t\,|\,x_0))\,dt - v(x(0\,|\,x_0)) \\
&= \beta\int_{\mathbb{R}^n\times\mathbf{U}} v(x)\,d\mu(x,u\,|\,x_0) - v(x(0\,|\,x_0)),
\end{aligned}$$

where the boundary term at infinity vanishes due to discounting and the fact v is bounded. Integrating with respect to μ_0 then gives the sought equation

$$\beta\int_{\mathbb{R}^n\times\mathbf{U}} v(x)\,d\mu(x,u) = \int_{\mathbb{R}^n} v(x)\,d\mu_0(x) + \int_{\mathbb{R}^n\times\mathbf{U}} \nabla v\cdot f(x,u)\,d\mu(x,u) \quad \forall v\in C_b^1(\mathbb{R}^n). \tag{2.28}$$

Equation (2.28) is called the discounted Liouville's equation. It is a *linear* equation in the variables (μ,μ_0) and it is a lifting, on the time interval $[0,\infty)$, of the nonlinear relaxed dynamics (2.20) and (2.21) into the space of measures.

Remark 2.3.1 regarding the test function v applies to Eq. (2.28) as well.

It follows from the above discussion that any family of relaxed state-control trajectories starting from a given initial measure μ_0 gives rise to a discounted occupation measure μ such that the pair (μ,μ_0) satisfies (2.28). Importantly, the converse is true as well in the sense that for any pair of measures (μ,μ_0) satisfying (2.28) there exists a family of trajectories of (2.21) starting from μ_0 that generates μ through (2.26) and (2.27). This is formalized in the following theorem.

Theorem 2.3.2 *If a pair of nonnegative compactly supported finite measures*

$$(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$$

satisfies (2.28) for all $v \in C_b^1(\mathbb{R}^n)$*, then there exists a measure*

$$\eta \in M(C([0,\infty); \mathbb{R}^n))_+$$

supported on the absolutely continuous trajectories of (2.20) such that

$$\mu(B \times \mathbf{U}) = \int_{C([0,\infty);\mathbb{R}^n)} \int_0^\infty e^{-\beta t} I_B(x(\cdot))\, dt\, d\eta(x(\cdot)),$$

and

$$\mu_0(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(0))\, d\eta(x(\cdot))$$

for all $B \subset \mathbb{R}^n$*.*

The proof of the Theorem is in Appendix B.2 and is based on results from stochastic optimal control theory and on Ambrosio's superposition principle [Amb08, Theorem 3.2].

Remark 2.3.3 regarding control trajectories and Remark 2.3.2 regarding superposition of trajectories apply to Theorem 2.3.2 as well.

2.3.4 Continuous-time infinite-horizon with stopping

In this section we discuss a generalization of the infinite-time setting of Section 2.3.3, where now we allow a state-dependent stopping of the trajectories. Hence assume that an initial condition, a relaxed control input $\gamma_t(\cdot \mid x_0)$ as well as a *stopping function* $\tau(x_0) \in L(\mathbb{R}^n; [0,\infty])$ are given (note that the stopping function is allowed to be infinite). Then we define the *conditional stopped discounted occupation measure* as

$$\mu(B \times C \mid x_0) := \int_0^{\tau(x_0)} \int_{\mathbf{U}} e^{-\beta t} I_{A\times B}(x(t\,|\,x_0), u)\, d\gamma_t(u\,|\,x_0)\, dt \tag{2.29}$$

for all sets $B \subset \mathbb{R}^n$ and $C \subset \mathbf{U}$.

Integrating over a given distribution of initial conditions μ_0 we define the *discounted* *stopped occupation measure* as

$$\mu(B \times C) := \int_{\mathbb{R}^n} \int_0^{\tau(x_0)} \int_{\mathbf{U}} e^{-\beta t} I_{A\times B}(x(t\,|\,x_0), u)\, d\gamma_t(u\,|\,x_0)\, dt\, d\mu_0(x) \tag{2.30}$$

and the *stopped final measure*

$$\mu_T(B) := \int_{\mathbb{R}^n} e^{-\beta\tau(x_0)} I_B(x(\tau(x_0)\,|\,x_0)). \tag{2.31}$$

Note that trajectories starting from distinct initial conditions x_0 can be stopped at different times, some finite, some infinite, based on the values of $\tau(x_0)$. The discounted stopped occupation measure μ captures the average (discounted) time spent in subsets of $\mathbb{R}^n \times \mathbf{U}$ until stopping, where the averaging is over the distribution of the initial conditions μ_0. The stopped final measure μ_T captures the (discounted) spatial distribution at the end points of the stopped trajectories (note that whenever $\tau(x_0) = +\infty$ in (2.31) the integrand is zero).

A computation completely analogous to the one in Section 2.3.3 reveals that the three measures μ, μ_0 and μ_T satisfy the equation

$$\int_{\mathbb{R}^n} v(x)\, d\mu_T(x) + \beta \int_{\mathbb{R}^n\times\mathbf{U}} v(x)\, d\mu(x,u) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \int_{\mathbb{R}^n\times\mathbf{U}} \nabla v \cdot f(x,u)\, d\mu(x,u) \tag{2.32}$$

for all $v \in C_b^1(\mathbb{R}^n)$, which is a linear equation in the variables μ, μ_0 and μ_T.

This equation is called the *stopped discounted Liouville's equation* and we remark that the unstopped version (2.28) is obtained by setting $\mu_T = 0$ in (2.32).

At present a superposition theorem similar to Theorem 2.3.2 is not available in full generality. However, and that is what will be needed in Chapter 4, such result can be proven in the uncontrolled setting and under the assumption that f is locally Lipschitz. We, however, have to allow for multiple stoppings for a single initial condition. This brings us from stopping functions to *stopping measures* that capture the distribution of stopping times for a given initial condition and mathematically are simply probability measures on $[0, \infty]$.

Theorem 2.3.3 *Let $\bar{f} : \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz and let the triplet of non-negative compactly supported measures*

$$(\mu, \mu_0, \mu_T) \in M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+$$

satisfy

$$\int_{\mathbb{R}^n} v(x)\,d\mu_T(x) + \beta \int_{\mathbb{R}^n} v(x)\,d\mu(x) = \int_{\mathbb{R}^n} v(x)\,d\mu_0(x) + \int_{\mathbb{R}^n} \nabla v \cdot \bar{f}(x)\,d\mu(x) \quad (2.33)$$

for all $v \in C_b^1(\mathbb{R}^n)$. *Then there exists an ensemble of probability measures* $\{\tau_{x_0} \in P([0,\infty])\}_{x_0 \in \mathbb{R}^n}$ *(a stochastic kernel*[10]*) such that*

$$\int_{\mathbb{R}^n} v(x)\,d\mu_0(x) = \int_{\mathbb{R}^n} v(x(0\,|\,x_0))\,d\mu_0(x_0), \quad (2.34a)$$

$$\int_{\mathbb{R}^n} v(x)\,d\mu(x) = \int_{\mathbb{R}^n} \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t\,|\,x_0))\,dt\,d\tau_{x_0}(\tau)\,d\mu_0(x_0), \quad (2.34b)$$

$$\int_{\mathbb{R}^n} v(x)\,d\mu_T(x) = \int_{\mathbb{R}^n} \int_0^\infty e^{-\beta\tau} v(\tau(x_0))\,d\tau_{x_0}(\tau)\,d\mu_0(x_0), \quad (2.34c)$$

for all bounded measurable functions v, *where* $x(\cdot\,|\,x_0)$ *denotes the unique trajectory of* $\dot{x} = \bar{f}(x)$ *starting from* x_0, *which is defined at least on* $[0, \sup \operatorname{spt} \tau_{x_0}]$ *for all* $x_0 \in \operatorname{spt} \mu_0$.

The proof of the Theorem is in Appendix B.3.

Remark 2.3.4 *Theorem 2.3.3 says that any measures satisfying (2.33) are generated by a superposition of the trajectories of the dynamical system* $\dot{x} = \bar{f}(x)$, *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* $\bar{f}$ *is locally 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* τ_{x_0}, $x_0 \in \mathbb{R}^n$. *For example, if* 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.*

[10]See Section A.1 for the definition of a stochastic kernel.

2.3.5 Continuous-time input-affine systems – special case

In this section we treat the special case of input-affine dynamics

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} f_{u_i}(x(t))u_i(t), \tag{2.35}$$

where the control input $u = (u_1, \ldots, u_m)$ is constrained to lie in a box $\mathbf{U} = [0, \bar{u}]^m$, $\bar{u} > 0$ (note that any box can be affinely transformed to a box of this form). Notice first that in this case the trajectories of the relaxed and non-relaxed dynamics coincide. More importantly from a practical point of view the input-affine form of the dynamics allows us to derive a special form of the Liouville's equation which allows for a controller extraction from the solutions to the Liouville's equation and allows for computational savings when relaxed or tightened later on.

We demonstrate the derivation on the infinite-horizon discounted Liouville's equation (2.28) and only state the final results for the case of finite time interval and for the case of infinite time interval with stopping, the derivation being identical.

Given a pair of measures $(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$ solving (2.28), which in our case reads

$$\beta \int_{\mathbb{R}^n \times \mathbf{U}} v(x)\, d\mu(x,u) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \int_{\mathbb{R}^n \times \mathbf{U}} \nabla v(x) \cdot \Big(f(x) + \sum_{i=1}^{m} f_{u_i}(x) u_i \Big)\, d\mu(x,u), \tag{2.36}$$

we disintegrate (see Section A.1) μ as $d\mu(x,u) = d\nu(u \mid x)d\bar{\mu}(x)$, where $\nu(\cdot \mid x)$ is a stochastic kernel on $\mathbf{U}$ given $\mathbb{R}^n$ and $\bar{\mu}$ is the x-marginal[11] of μ. Defining the feedback control law

$$u_i(x) = \int_{\mathbf{U}} u_i\, d\nu(u \mid x), \tag{2.37}$$

the above equation becomes

$$\beta \int_{\mathbb{R}^n} v(x)\, d\bar{\mu}(x) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \int_{\mathbb{R}^n} \nabla v(x) \cdot \Big(f(x) + \sum_{i=1}^{m} f_{u_i}(x) u_i(x) \Big)\, d\bar{\mu}(x).$$

[11] Given a measure $\mu \in M(\mathbf{X} \times \mathbf{U})$, the x-marginal of μ is the unique measure $\bar{\mu} \in M(\mathbf{X})$ satisfying $\bar{\mu}(B) = \mu(B \times \mathbf{U})$ for all $B \subset \mathbf{X}$.

Defining further the control measures $d\nu_i(x) = u_i(x)d\bar{\mu}(x)$, we obtain

$$\beta \int_{\mathbb{R}^n} v\, d\bar{\mu}(x) = \int_{\mathbb{R}^n} v\, d\mu_0(x) + \int_{\mathbb{R}^n} \nabla v(x) \cdot f(x)\, d\bar{\mu}(x) + \sum_{i=1}^{m} \int_{\mathbb{R}^n} \nabla v \cdot f_{u_i}(x)\, d\nu_i(x) \tag{2.38}$$

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

Notice that since $\mathbf{U} = [0, \bar{u}]^m$ and $\mathrm{spt}\mu \subset \mathbb{R}^n \times \mathbf{U}$, the controller (2.37) satisfies $u_i(x) \in [0, \bar{u}]$ for all $x \in \mathrm{spt}\,\bar{\mu}$, i.e., the extracted controller is admissible (which is not guaranteed in more general settings, in particular if $\mathbf{U}$ is not convex, in which case one can only guarantee $u(x) \in \overline{\mathrm{conv}}\,\mathbf{U}$ even if the dynamics is input-affine). In particular each control measure ν_i is absolutely continuous with respect to the x-marginal of the occupation measure with density $u_i(x) \in [0, \bar{u}]$ for all $x \in \mathrm{spt}\,\bar{\mu}$.

Importantly, any solution

$$(\bar{\mu}, \mu_0, \nu) \in M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+^m$$

to (2.38) satisfying $\nu_i \leq \bar{u}\bar{\mu}$ gives rise to a solution $(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$ to (2.36) with μ defined by

$$\mu(B \times C) = \int_B I_C(u(x))\, d\bar{\mu}(x)$$

for all $B \subset \mathbb{R}^n$, $C \subset \mathbb{R}^m$, where $u = (u_1, \ldots, u_m)$ with each $u_i(\cdot)$ being the density (i.e., the Radon-Nikodým derivative) of ν_i with respect to $\bar{\mu}$. Since $0 \leq \nu_i \leq \bar{u}\bar{\mu}$ we have $u_i(x) \in [0, \bar{u}]$ for all $x \in \mathrm{spt}\,\bar{\mu}$ and hence $\mathrm{spt}\,\mu \subset \mathbb{R}^n \times \mathbf{U}$ as desired.

Hence there is a one-to-one correspondence between solutions to (2.36) and solutions to (2.38) satisfying the additional constraint $\nu_i \leq \bar{u}\bar{\mu}$.

The advantage of using (2.38) compared to (2.36) is twofold. First, it is the easy extraction of a feedback controller as a density of ν_i with respect to $\bar{\mu}$ (this will be even more apparent in Section 4 when we tighten Eq. (2.38) and work with densities in which case the extracted controller is simply the ratio of the densities of ν_i and $\bar{\mu}$). Second, the equation (2.38) involves only measures on $\mathbb{R}^n$, whereas Eq. (2.36) involves measures on $\mathbb{R}^n \times \mathbb{R}^m$, which, when relaxed or tightened later on, leads to significantly smaller optimization problems involving approximations of these measures.

For the finite-time case, the equivalent equation to (2.25) is

$$\int_{\mathbb{R}^n} v\, d\mu_T(x) = \int_{\mathbb{R}^n} v\, d\mu_0(x) + \int_{\mathbb{R}^n} \frac{\partial v}{\partial t} + \nabla v \cdot f\, d\bar{\mu}(t,x) + \sum_{i=1}^{m} \int_{\mathbb{R}^n} \nabla v \cdot f_{u_i}\, d\nu_i(t,x) \tag{2.39}$$

for all $v \in C^1([0,T] \times \mathbb{R}^n)$, where the variables are

$$(\bar{\mu}, \mu_0, \mu_T, \nu) \in M([0,T] \times \mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M([0,T] \times \mathbb{R}^n)_+^m$$

and again we have the additional constraint $\nu_i \leq \bar{u}\bar{\mu}$.

For the case of infinite-time with stopping, the equivalent equation to (2.32) is

$$\int_{\mathbb{R}^n} v\, d\mu_T(x) + \beta \int_{\mathbb{R}^n} v\, d\bar{\mu}(x) = \int_{\mathbb{R}^n} v\, d\mu_0(x) + \int_{\mathbb{R}^n} \nabla v \cdot f\, d\bar{\mu}(x) + \sum_{i=1}^{m} \int_{\mathbb{R}^n} \nabla v \cdot f_{u_i}\, d\nu_i(x) \tag{2.40}$$

for all $v \in C_b^1(\mathbb{R}^n)$, where the variables are

$$(\bar{\mu}, \mu_0, \mu_T, \nu) \in M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+^m$$

and again we have the additional constraint $\nu_i \leq \bar{u}\bar{\mu}$.

2.3.6 Discrete-time infinite-horizon

Consider any trajectory $\{x_{t|x_0}\}_{t=0}^{\infty}$ of the discrete-time system (2.19) starting form an initial condition x_0 generated by the control $\{u_{t|x_0}\}_{t=0}^{\infty}$.

Then, given a discount factor $\alpha \in (0,1)$, we define the *conditional discounted occupation measure* $\mu(\cdot \mid x_0) \in M(\mathbb{R}^n \times \mathbf{U})_+$ as

$$\mu(B \times C \mid x_0) := \sum_{t=0}^{\infty} \alpha^t I_{B\times C}(x_{t|x_0}, u_{t|x_0}) \tag{2.41}$$

for all sets $A \subset \mathbb{R}^n$ and $B \subset \mathbf{U}$.

The quantity $\mu(B \times C \mid x_0)$ equals to the (discounted) number of visits of the state-control trajectory $\left(\{x_{t|x_0}, u_{t|x_0}\}_{t=0}^{\infty}\right)$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$. The discounting in the definition of the occupation measure ensures that $\mu(B \times C \mid x_0)$ is always finite; in fact we have $\mu(\mathbb{R}^n \times \mathbf{U} \mid x_0) = (1-\alpha)^{-1}$.

Now suppose that the initial condition is not a single point but an *initial measure*[12] $\mu_0 \in M(\mathbf{X})_+$ and a state trajectory along with an admissible control input generating it is associated to each initial condition from the support of μ_0. Then we define the *discounted occupation measure* $\mu \in M(\mathbb{R}^n \times \mathbf{U})_+$ as

$$\mu(B \times C) := \int_{\mathbb{R}^n} \mu(B \times C \,|\, x_0)\, d\mu_0(x_0). \tag{2.42}$$

The quantity $\mu(B \times C \mid x_0)$ equals to the average (discounted) number of visits of the state-control trajectories $\Big(\{x_{t|x_0}, u_{t|x_0}\}_{t=0}^{\infty}\Big)$ in $B \times C \subset \mathbb{R}^n \times \mathbf{U}$, where the averaging is over the distribution of the initial conditions μ_0.

Now we derive an equation linking the measures μ_0 and μ, which will be the sought lifting, on the infinite horizon $\{0, 1, \ldots\}$, of the nonlinear discrete-time dynamics (2.19) into a *linear* equation on measures. This equation will play a key role for maximum controlled invariant set approximation in Section 3.2, in discrete time. To derive this equation fix an initial condition $x_0 \in \mathbb{R}^n$ and a trajectory $\{x_{t|x_0}\}_{t=0}^{\infty}$ generated by a control sequence $\{u_{t|x_0}\}_{t=0}^{\infty}$. Then for any $v \in C_b(\mathbb{R}^n)$, where $C_b(\mathbb{R}^n)$ denotes the set of all bounded continuous functions, we have

$$\begin{aligned}
\int_{\mathbb{R}^n \times \mathbf{U}} v(x)\, d\mu(x, u \,|\, x_0) &= \sum_{t=0}^{\infty} \alpha^t v(x_{t|x_0}) = v(x_{0|x_0}) + \alpha \sum_{t=0}^{\infty} \alpha^t v(x_{t+1|x_0}) \\
&= v(x_{0|x_0}) + \alpha \sum_{t=0}^{\infty} \alpha^t v(f(x_{t|x_0}, u_{t|x_0})) \\
&= v(x_{0|x_0}) + \alpha \int_{\mathbf{X} \times \mathbf{U}} v(f(x, u))\, d\mu(x, u \,|\, x_0).
\end{aligned}$$

Integrating with respect to μ_0 we arrive at the sought equation

$$\int_{\mathbb{R}^n \times \mathbf{U}} v(x)\, d\mu(x, u) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \alpha \int_{\mathbb{R}^n \times \mathbf{U}} v(f(x, u))\, d\mu(x, u) \quad \forall v \in C_b(\mathbb{R}^n). \tag{2.43}$$

Equation (2.43) is called discounted Liouville's equation. It is a *linear* equation in the variables (μ, μ_0) and it is a lifting, on the infinite horizon $\{0, 1, \ldots\}$, of the nonlinear discrete-time dynamics (2.19) into the space of measures.

Remark 2.3.1 regarding the test function v applies to Eq. (2.43) as well.

[12] The initial measure μ_0 can be thought of as the probability distribution of the initial state, although we do not require the mass of μ_0 to be normalized to one.

It follows from the above discussion that any family of state-control trajectories of (2.19) starting from a given initial measure μ_0 gives rise to a discounted occupation measure μ such that the pair (μ, μ_0) satisfies (2.43). Importantly, the converse is true as well in the sense that for any pair of measures (μ, μ_0) satisfying (2.43) there exists a family of trajectories of (2.19) starting from μ_0 that generates μ through (2.41) and (2.42). This is formalized in the following theorem, where $l(\mathbb{N};\mathbb{R}^n)$ denotes the set of all $\mathbb{R}^n$-valued sequences indexed with natural numbers .

Theorem 2.3.4 *If a pair of nonnegative compactly supported finite measures*

$$(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$$

satisfies (2.43) for all $v \in C_b(\mathbb{R}^n)$, then there exists a measure

$$\eta \in M(l(\mathbb{N};\mathbb{R}^n))_+$$

supported on the trajectories of (2.19) such that

$$\mu(B \times \mathbf{U}) = \int_{l(\mathbb{N};\mathbb{R}^n)} \sum_{t=0}^{\infty} \alpha^t I_B(x_t)\, d\eta((x_k)_{k=0}^{\infty}),$$

and

$$\mu_0(B) = \mu(B \times \mathbf{U}) = \int_{l(\mathbb{N};\mathbb{R}^n)} I_B(x_0)\;\; d\eta((x_k)_{k=0}^{\infty}),$$

for all $B \subset \mathbb{R}^n$.

The proof of the Theorem is in Appendix B.4 and is based on results from Markov chain theory.

Remark 2.3.3 regarding control trajectories and Remark 2.3.2 regarding superposition of trajectories apply to Theorem 2.3.4 as well.

Chapter 3

Set approximation

This chapter addresses the problems of region of attraction (ROA), reachable set and maximum controlled invariant (MCI) set computation. These sets are fundamental in non-linear control theory and their approximations have a range of practical applications, e.g., for reach-avoid or target-hitting problems (see, e.g., [ML11]), collision avoidance and viability analysis (where the ROA is called capture basin and the MCI set viability kernel) [ABSP11] or for model predictive control design [RM09].

In this chapter we characterize each of these sets as an infinite-dimensional linear program (LP) in the cone of nonnegative measures. Then, for each of the sets, we derive an LP on continuous functions dual to the corresponding primal LP on measures and show that any feasible solution to this dual LP provides an outer approximation to the set. After that we approximate this infinite-dimensional LP by relaxing the primal LP and tightening the dual LP using moment-sum-of-squares hierarchies as described in Sections 2.2.1 and 2.2.2, obtaining a hierarchy of finite-dimensional semidefinite programming (SDP) relaxations of the primal LP and tightenings of the dual LP. Importantly, the hierarchy of dual SDP tightenings provides a sequence of outer approximations to these sets which is proven to converge in the sense of volume discrepancy tending to zero.

One of the main virtues of the proposed method is its simplicity and generality. The outer approximations are obtained as the solution to a single, convex, semidefinite programming problem with *no initialization* required and no ad hoc tuning parameters. The outer approximations are obtained as a super-level set of a single multivariate polynomial and hence are easy to manipulate. The only parameter of the problem is the degree of this polynomial allowing one to trade off computational complexity and tightness of the approximation. The only assumption is that all data is polynomial, hence covering a broad class of systems without imposing any structural assumptions. In addition, the approach is highly flexible: for example, convexity of the outer approximations and other shape constraints can be easily enforced and the approach extends to other classes of systems (e.g., trigonometric, piecewise polynomial, hybrid).

This is in contrast with most existing approaches which typically rely on non-convex bilinear matrix inequalities (see, e.g., [Che11, KSK14]) with their inherent numerical difficulties, including complicated initialization of iterative schemes and only local convergence, or rely on careful analysis of the dynamics for a particular class of systems (see, e.g., [CCG09, Sta09]).

We believe that our approach is closer in spirit to level-set and Hamilton-Jacobi approaches [MT03] or transfer operator approaches [WV10], but we do not rely on spatial and/or temporal discretization but rather on moment-sum-of-squares hierarchies, and our approach comes with convergence guarantees.

3.1 Region of attraction & Reachable set

Consider the relaxed control system

$$\dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}), \quad x(t) \in \mathbf{X},\ t \in [0, T], \tag{3.1}$$

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

$$\begin{aligned} x(t) \in \mathbf{X} &:= \{x \in \mathbb{R}^n : g_i^X(x) \geq 0, i \in \{1, \ldots, n_X\}\},\ t \in [0, T], \\ u(t) \in \mathbf{U} &:= \{u \in \mathbb{R}^m : g_i^U(u) \geq 0, i \in \{1, \ldots, n_U\}\},\ t \in [0, T], \end{aligned} \tag{3.2}$$

with $g_i^X \in \mathbb{R}[x]$ and $g_i^U \in \mathbb{R}[u]$. The target set is defined by

$$\boldsymbol{X_T} := \{x \in \mathbb{R}^n \ : \ g_i^{X_T}(x) \geq 0, i \in \{1, \ldots, n_T\}\} \subset \mathbf{X},$$

[1]Since f is continuous and $\mathbf{U}$ compact, conv $f(x, \mathbf{U})$ is necessarily compact and hence closure is not required in (3.22) as opposed to Section 2.3.1 where f was not required to be continuous.

[2]Note that the infinite-dimensional results of Sections 3.1.3 and 3.1.4 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 3.1.5.

and the initial set by

$$\boldsymbol{X_I} := \{x \in \mathbb{R}^n \;:\; g_i^{X_I}(x) \geq 0, i \in \{1, \dots, n_T\}\} \subset \mathbf{X},$$

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

Throughout this chapter we make the following standing assumption:

Assumption 3.1.1 *The sets* $\mathbf{X}$, $\mathbf{U}$, $\boldsymbol{X_T}$ *and* $\boldsymbol{X_I}$ *are compact.*

The region of attraction (ROA) is defined as

$$\boldsymbol{X}_0 := \Big\{x_0 \in \mathbf{X} \mid \exists\, x(\cdot) \in AC([0,T];\mathbb{R}^n) \text{ s.t. } \dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}) \text{ a.e.}, \\ x(0) = x_0,\; x(T) \in \boldsymbol{X_T},\; x(t) \in \mathbf{X}\, \forall\, t \in [0,T]\Big\},$$

where $AC([0,T];\mathbb{R}^n)$ denotes the set of all absolutely continuous functions on $[0,T]$ taking values in $\mathbb{R}^n$ and a.e. stands for "almost everywhere" with respect to the Lebesgue measure on $[0,T]$. In words, ROA is the set of all initial conditions that can be steered to the target $\boldsymbol{X_T}$ without violating the state constraints using relaxed controls satisfying the input constraints. The set $\boldsymbol{X}_0$ is bounded (by Assumption 3.1.1) and unique.

The reachable set is defined as

$$\boldsymbol{X_R} := \Big\{x_T \in \mathbf{X} \mid \exists\, x(\cdot) \in AC([0,T];\mathbb{R}^n) \text{ s.t. } \dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}) \text{ a.e.}, \\ x(0) \in \boldsymbol{X_I},\; x(T) = x_T,\; x(t) \in \mathbf{X}\, \forall\, t \in [0,T]\Big\},$$

In words, the reachable set is the set of all states that can be reached from the given initial set $\boldsymbol{X_I}$ without violating the state constraint set using relaxed controls satisfying the input constraints. The set $\boldsymbol{X_R}$ is also bounded and unique.

Remark 3.1.1 *The sets* $\boldsymbol{X}_0$ *and* $\boldsymbol{X_R}$ *are related by a change of direction of time and hence being able to characterize and compute approximation of one implies the same for the other. Indeed, the ROA for the dynamics* $\dot{x} \in \operatorname{conv} f(x, \mathbf{U})$ *with a target set* $\boldsymbol{X_T}$ *is precisely the reachable set for the time-reversed dynamics* $\dot{x} \in \operatorname{conv} -f(x, \mathbf{U})$ *with the initial set* $\boldsymbol{X_I} = \boldsymbol{X_T}$ *and vice versa. Hence from now on we focus on the ROA only, all the results following for the reachable set by changing the sign of* f *and replacing* $\boldsymbol{X_T}$ *with* $\boldsymbol{X_I}$.

Next we derive a characterization of the ROA as an infinite-dimensional LP in the cone of nonnegative measures. We proceed in several steps. First, we pose the problem as a nonlinear optimization problem. Then we lift this problem into an infinite-dimensional problem, which, however, turns out to be non-convex. In the last and crucial step of the derivation, we describe an equivalent convex (in fact linear) formulation of this problem. After that we derive a dual (in the space of continuous functions) to the infinite-dimensional LP on measures and present a hierarchy of SDP relaxations of the primal, and tightenings of the dual, which provides a converging sequence of outer approximations to the ROA.

3.1.1 ROA via optimization

The following characterization of ROA is almost tautological:

$$\begin{array}{ll} \sup\limits_{X_0 \subset \mathbb{R}^n} & \mu^{\mathrm{leb}}(X_0) \\ \text{s.t.} & \forall\, x_0 \in X_0 \;\exists\, x(\cdot \mid x_0) \in AC([0,T];\mathbb{R}^n) \text{ s.t.} \\ & \dot{x}(t \mid x_0) \in \operatorname{conv} f(x(t \mid x_0), \mathbf{U}) \text{ a.e. on } [0,T] \\ & x(t \mid x_0) \in \mathbf{X} \;\; \forall\, t \in [0,T] \\ & x(T \mid x_0) \in \boldsymbol{X_T}, \end{array} \tag{3.3}$$

where μ^{leb} denotes the Lebesgue measure (hence $\mu^{\mathrm{leb}}(X_0)$ is the volume of X_0). Any set X_0 feasible in (3.3) satisfies $X_0 \subset \boldsymbol{X}_0$ and the supremum in (3.3) is equal to the volume of the ROA $\boldsymbol{X}_0$. Hence any minimizer $X_0^\star$ of (3.3) satisfies $X_0^\star \subset \boldsymbol{X}_0$ and $\mu^{\mathrm{leb}}(\boldsymbol{X}_0 \setminus X_0^\star) = 0$.

3.1.2 Lifting: first attempt

Now we lift the problem (3.3) into the space of measures. The key ingredient is the finite-time Liouville's equation (2.25) which serves as the lifting of the nonlinear relaxed dynamics (3.1) into a *linear* equation on measures. Constraints on the trajectories are imposed via support constraints on these measures.

The lifting reads

$$
\begin{array}{rcll}
q^* & = & \sup\limits_{\mu_0,\mu,\mu_T} & \mu^{\mathrm{leb}}(\operatorname{spt}\mu_0) \\
& & \text{s.t.} & \int_{\boldsymbol{X_T}} v(T,\cdot)\,d\mu_T = \int_{\mathbf{X}} v(0,\cdot)\,d\mu_0 + \int\limits_{[0,T]\times\mathbf{X}\times\mathbf{U}} \mathcal{L}v\,d\mu \quad \forall v \in C^1([0,T]\times\mathbf{X}) \\
& & & \mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+ \\
& & & \mu_0 \in M(\mathbf{X})_+ \\
& & & \mu_T \in M(\boldsymbol{X_T})_+,
\end{array}
\tag{3.4}
$$

where the operator $\mathcal{L} : C^1([0,T]\times\mathbf{X}) \to C([0,T]\times\mathbf{X})$ is defined by

$$
\mathcal{L}v := \frac{\partial v}{\partial t} + \nabla v \cdot f, \tag{3.5}
$$

the symbol $\operatorname{spt}\mu_0$ denotes the support of μ_0 and the first constraint is precisely the Liouville's equation (2.25). Problem (3.4) is an infinite-dimensional optimization problem in the cone of nonnegative measures. Note, however, that the objective function of (3.3) is *nonconvex* and hence this infinite-dimensional problem is not a conic optimization problem; therefore, a convex reformulation described in Section 3.1.3, is needed.

We emphasize that the test function v is *not a decision variable* in the problem, but rather the equality constraint is required to hold for all $v \in C^1([0,T]\times\mathbf{X})$ (see Remark 2.3.1).

The rationale behind problem (3.4) is as follows. The first constraint is the Liouville's equation (2.25) which ensures that any triplet of measures (μ_0, μ, μ_T) feasible in (3.4) corresponds to an initial, an occupation and a terminal measure generated by trajectories of (3.1) in the sense of Theorem 2.3.1. Constraints on the trajectories and controls are encoded via constraints on the supports of the measures implicitly encoded by the inclusion to an appropriate cone of nonnegative measures. In particular, the constraint $\mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+$ ensures that $\operatorname{spt}\mu \subset [0,T]\times\mathbf{X}\times\mathbf{U}$ and hence that the trajectories satisfy the state and control constraints over the time interval $[0,T]$; the constraint $\mu_T \in M(\boldsymbol{X_T})_+$ ensures that $\operatorname{spt}\mu_T \subset \boldsymbol{X_T}$ and hence 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[3]. This discussion is summarized in the following Lemma.

[3]Even though the support of the initial measure attaining the maximum in (3.4) can differ from the ROA on the set of zero volume, the outer approximations obtained in Section 3.1.6 are valid "everywhere", not "almost everywhere".

Lemma 3.1.1 *The optimal value of problem (3.4) is equal to the volume of the ROA* $\boldsymbol{X}_0$*, that is,* $q^* = \mu^{\text{leb}}(\boldsymbol{X}_0)$*.*

Proof: By definition of the ROA, for any initial condition $x_0 \in \boldsymbol{X}_0$ there is a trajectory of (3.1) (which is necessarily generated by an admissible relaxed control input) which satisfies the state constraints and ends in the target set. Therefore for any initial measure μ_0 with $\operatorname{spt}\mu_0 \subset \boldsymbol{X}_0$ there exist an occupation measure μ and a final measure μ_T such that the constraints of problem (3.4) are satisfied. Thus, $q^* \geq \mu^{\text{leb}}(\boldsymbol{X}_0)$.

Now we show that $q^* \leq \mu^{\text{leb}}(\boldsymbol{X}_0)$. For contradiction, suppose that a triplet of measures (μ_0, μ, μ_T) is feasible in (3.4) and that $\mu^{\text{leb}}(\operatorname{spt}\mu_0 \setminus \boldsymbol{X}_0) > 0$. From Theorem 2.3.1 there is a family of admissible trajectories of the inclusion (3.1) 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 $\operatorname{spt}\mu_0 \setminus \boldsymbol{X}_0$ can be admissible. Thus, $\mu^{\text{leb}}(\operatorname{spt}\mu_0 \setminus \boldsymbol{X}_0) = 0$ and so $\mu^{\text{leb}}(\operatorname{spt}\mu_0) \leq \mu^{\text{leb}}(\boldsymbol{X}_0)$. Consequently, $q^* \leq \mu^{\text{leb}}(\boldsymbol{X}_0)$. □

3.1.3 Primal infinite-dimensional LP on measures

The key idea behind the results of this chapter consists in replacing the direct maximization of the support of the initial measure μ_0 by the maximization of the integral below the density (w.r.t. the Lebesgue measure) of μ_0 subject to the constraint that this density exists and is less than or equal one. This procedure is equivalent to maximizing the mass[4] of μ_0 under the constraint that μ_0 is dominated by the Lebesgue measure. This leads to the following infinite-dimensional LP:

$$
\begin{array}{rcll}
p^* & = & \sup\limits_{\mu,\mu_0,\mu_T} & \int_{\mathbf{X}} 1\, d\mu_0 \\
& & \text{s.t.} & \int_{\boldsymbol{X_T}} v(T,\cdot)\, d\mu_T = \int_{\mathbf{X}} v(0,\cdot)\, d\mu_0 + \int\limits_{[0,T]\times\mathbf{X}\times\mathbf{U}} \mathcal{L}v\, d\mu \quad \forall v \in C^1([0,T]\times\mathbf{X}) \\
& & & \mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+ \\
& & & \mu_0 \in M(\mathbf{X})_+ \\
& & & \mu_T \in M(\boldsymbol{X_T})_+ \\
& & & \mu_0 \leq \mu^{\text{leb}}
\end{array}
\tag{3.6}
$$

In problem (3.6) the constraint $\mu_0 \leq \mu^{\text{leb}}$ means that $\mu_0(A) \leq \mu^{\text{leb}}(A)$ for all $A \subset \mathbf{X}$. Note how the objective functions differ in problems (3.4) and (3.6).

[4] The mass of the measure μ_0 is defined as $\int_{\mathbf{X}} 1\, d\mu_0 = \mu_0(\mathbf{X})$.

The following theorem is then almost immediate.

Theorem 3.1.1 *The optimal value of the infinite-dimensional LP problem (3.6) is equal to the volume of the ROA $\boldsymbol{X}_0$, that is, $p^* = \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$. Moreover, the supremum is attained by a triplet of measure (μ, μ_0, μ_T) for which μ_0 is equal to the restriction of the Lebesgue measure to the ROA $\boldsymbol{X}_0$.*

Proof: Since the constraint set of problem (3.6) is tighter than that of problem (3.4), by Lemma 3.1.1 we have that $\mu^{\mathrm{leb}}(\mathrm{spt}\,\mu_0) \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$ for any feasible μ_0. From the constraint $\mu_0 \leq \mu^{\mathrm{leb}}$ we get $\int_{\mathbf{X}} 1\, d\mu_0 = \mu_0(\mathrm{spt}\,\mu_0) \leq \mu^{\mathrm{leb}}(\mathrm{spt}\,\mu_0) \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$ for any feasible μ_0. Therefore $p^* \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$. But by definition of the ROA $\boldsymbol{X}_0$, the restriction of the Lebesgue measure to $\boldsymbol{X}_0$ is feasible in (3.6), and so $p^* \geq \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$. Consequently $p^* = \mu^{\mathrm{leb}}(\boldsymbol{X}_0)$. □

Now we reformulate problem (3.6) to the so-called standard form of a linear program (see Appendix A.2 for a brief introduction to infinite-dimensional linear programming) which is more convenient for dualization and theoretical analysis. To this end, let us define the complementary measure (a slack variable) $\hat{\mu}_0 \in M(\mathbf{X})_+$ such that the inequality $\mu_0 \leq \mu^{\mathrm{leb}}$ in (3.6) can be written equivalently as the constraints $\mu_0 + \hat{\mu}_0 = \mu^{\mathrm{leb}}$, which is in turn equivalent to

$$\int_{\mathbf{X}} w(x)\, d\mu_0(x) + \int_{\mathbf{X}} w(x)\, d\hat{\mu}_0(x) = \int_{\mathbf{X}} w\, d\mu^{\mathrm{leb}}(x) \quad \forall\, w \in C(\mathbf{X}).$$

Then problem (3.6) is equivalent to the infinite-dimensional primal LP

$$\begin{array}{rcll} p^* & = & \sup\limits_{\mu,\mu_0,\mu_T,\hat{\mu}_0} & \int_{\mathbf{X}} 1\, d\mu_0 \\ & & \text{s.t.} & \int_{\boldsymbol{X_T}} v(T,\cdot)\, d\mu_T = \int_{\mathbf{X}} v(0,\cdot)\, d\mu_0 + \int\limits_{[0,T]\times\mathbf{X}\times\mathbf{U}} \mathcal{L}v\, d\mu \quad \forall\, v \in C^1([0,T]\times\mathbf{X}) \\ & & & \int_{\mathbf{X}} w\, d\mu_0 + \int_{\mathbf{X}} w\, d\hat{\mu}_0 = \int_{\mathbf{X}} w\, d\mu^{\mathrm{leb}} \qquad \forall\, w \in C(\mathbf{X}) \\ & & & \mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+ \\ & & & \mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+ \\ & & & \mu_0 \in M(\mathbf{X})_+ \\ & & & \mu_T \in M(\boldsymbol{X_T})_+ \\ & & & \hat{\mu}_0 \in M(\mathbf{X})_+, \end{array} \tag{3.7}$$

which is a standard form infinite-dimensional linear program.

3.1.4 Dual infinite-dimensional LP on functions

In this section we derive a linear program dual to (3.7) (and hence to (3.6)) on the space of continuous functions. A certain super-level set of one of the functions feasible in this dual LP will provide an outer approximation to the ROA $\boldsymbol{X}_0$.

The dual infinite-dimensional LP reads (see Appendix A.2 for a brief introduction to infinite-dimensional LP duality):

$$\begin{array}{rlll} d^* = \inf\limits_{v,w} & \displaystyle\int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x) & & \\ \text{s.t.} & -\mathcal{L}v(t,x,u) & \in & C([0,T]\times\mathbf{X}\times\mathbf{U})_+, \\ & w - v(0,\cdot) - 1 & \in & C(\mathbf{X})_+, \\ & v(T,\cdot) & \in & C(\boldsymbol{X_T})_+, \\ & w & \in & C(\mathbf{X})_+, \end{array} \tag{3.8}$$

where the infimum is over $(v,w) \in C^1([0,T]\times\mathbf{X})\times C(\mathbf{X})$. The dual has the following interpretation: the first constraint (equivalent to $\mathcal{L}v \leq 0$ on $[0,T]\times\mathbf{X}\times\mathbf{U}$) forces v to decrease along admissible trajectories and hence necessarily $v(0,x) \geq 0$ for all $x \in \boldsymbol{X}_0$ because of the third constraint (equivalent to $v(T,x) \geq 0$ on for all $x \in \boldsymbol{X_T}$). Consequently, $w \geq 1$ on $\boldsymbol{X}_0$ because of the second constraint equivalent to $w(x) \geq v(0,x)+1$ for all $x \in \mathbf{X}$. Therefore $w \geq I_{\boldsymbol{X}_0}$ on $\mathbf{X}$, where $I_{\boldsymbol{X}_0}$ denotes the indicator function[5] of the region of attraction. The objective function then minimizes the integral of w, trying to make w as small as possible under the implicit constraint that $w \geq I_{\boldsymbol{X}_0}$.

This instrumental observation is formalized in the following Lemma.

Lemma 3.1.2 *Let (v,w) be a pair of function feasible in (3.8). Then $v(0,\cdot) \geq 0$ on $\boldsymbol{X}_0$, $w \geq 1$ on $\boldsymbol{X}_0$ and $w \geq I_{\boldsymbol{X}_0}$ on $\mathbf{X}$.*

Proof: By definition of $\boldsymbol{X}_0$, given any $x_0 \in \boldsymbol{X}_0$ there exists $u(t)$ such that $x(t) \in \mathbf{X}$, $u(t) \in \mathbf{U}$ for all $t \in [0,T]$ and $x(T) \in \boldsymbol{X_T}$. Therefore, since $v(T,\cdot) \geq 0$ on $\boldsymbol{X_T}$ and

[5] The indicator function of a set is a function equal to one on that set and zero otherwise.

$\mathcal{L}v \leq 0$ on $[0,T] \times \mathbf{X} \times \mathbf{U}$,

$$\begin{aligned} 0 \leq v(T, x(T)) &= v(0, x_0) + \int_0^T \frac{d}{dt} v(t, x(t))\, dt \\ &= v(0, x_0) + \int_0^T \mathcal{L}v(t, x(t), u(t))\, dt \\ &\leq v(0, x_0) \leq w(x_0) - 1, \end{aligned}$$

where the last inequality follows from the second constraint of (3.8). The fact that $w \geq I_{\mathbf{X}_0}$ follows from the last constraint of (3.8) which requires w to be nonnegative on $\mathbf{X}$. □

Next, we have the following salient result:

Theorem 3.1.2 *There is no duality gap between the primal infinite-dimensional LP problems (3.6) and (3.7) and the infinite-dimensional dual LP problem (3.8) in the sense that $p^* = d^*$.*

Proof: To streamline the exposition, let

$$\begin{aligned} \mathcal{M} &:= M([0,T] \times \mathbf{X} \times \mathbf{U}) \times M(\mathbf{X}) \times M(\boldsymbol{X_T}) \times M(\mathbf{X}), \\ \mathcal{C} &:= C([0,T] \times \mathbf{X} \times \mathbf{U}) \times C(\mathbf{X}) \times C(\boldsymbol{X_T}) \times C(\mathbf{X}), \end{aligned}$$

and let $\mathcal{K}$ and $\mathcal{K}^\star$ denote the positive cones of $\mathcal{M}$ and $\mathcal{C}$ 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 $\mathcal{C}$ (see Section 2.1). The cone $\mathcal{K}$ is equipped with the weak-$\star$ topology (see Section 2.1.1). Then, the LP problem (3.7) can be rewritten as

$$\begin{array}{rcll} p^* & = & \sup_\gamma & \langle \gamma,\, c \rangle \\ & & \text{s.t.} & \mathcal{A}\gamma = \beta \\ & & & \gamma \in \mathcal{K}, \end{array} \tag{3.9}$$

where the infimum is over the vector $\gamma := (\mu, \mu_0, \mu_T, \hat{\mu}_0)$, the linear operator $\mathcal{A} : \mathcal{K} \to C^1([0,T] \times \mathbf{X})^\star \times M(\mathbf{X})$ is defined by

$$\mathcal{A}\gamma := \left(-\mathcal{L}^\star \mu - \delta_0 \otimes \mu_0 + \delta_T \otimes \mu_T \;,\;\; \mu_0 + \hat{\mu}_0\right),$$

where $\mathcal{L}$ is the adjoint of the operator $\mathcal{L}$ defined in (3.5), the right hand side of the equality constraint in (3.9) is the vector of measures $\beta := (0\,,\, \mu^{\mathrm{leb}}) \in M([0,T] \times \mathbf{X}) \times M(\mathbf{X})$, the vector function in the objective is $c := (0,1,0,0) \in \mathcal{C}$, so the objective function itself is

$$\langle \gamma,\, c\rangle = \int_{\mathbf{X}} 1\, d\mu_0 = \mu_0(\mathbf{X}).$$

The dual LP to (3.9) reads (see Section A.2)

$$\begin{array}{rcll} d^* & = & \inf_z & \langle \beta,\, z\rangle \\ & & \text{s.t.} & \mathcal{A}^\star(z) - c \in \mathcal{K}^\star, \end{array} \tag{3.10}$$

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

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

and satisfies the adjoint relation $\langle \mathcal{A}\gamma,\, z\rangle = \langle \gamma,\, \mathcal{A}^\star z\rangle$. The LP problem (3.10) is exactly the LP problem (3.8).

To conclude the proof we use an argument similar to that of [Las09, Section C.4]. From Theorem A.2.1 there is no duality gap between LPs (3.9) and (3.10) if the supremum p^* is finite and the set $S := \{(\mathcal{A}\gamma,\, \langle \gamma,\, c\rangle) : \gamma \in \mathcal{K}\}$ is closed in the weak-$\star$ topology of $\mathcal{K}$. The fact that p^* is finite follows readily from the constraint $\mu_0 + \hat{\mu}_0 = \mu^{\mathrm{leb}}$, $\hat{\mu}_0 \geq 0$, and from compactness of $\mathbf{X}$. To prove closedness, we first remark that $\mathcal{A}$ is weakly-$\star$ continuous[6] since $\mathcal{A}^\star(z) \in \mathcal{C}$ for all $z \in C^1([0, T] \times \mathbf{X}) \times C(\mathbf{X})$. Then we consider a sequence $\gamma_k = (\mu^k, \mu_0^k, \mu_T^k, \hat{\mu}_0^k) \in \mathcal{K}^\star$ and we want to show that $(\nu, a) := \lim_{k\to\infty}(\mathcal{A}\gamma_k,\, \langle \gamma_k,\, c\rangle)$ belongs to S, where $\nu \in C^1([0, T] \times \mathbf{X})^\star \times M(\mathbf{X})$ and $a \in \mathbb{R}$. To this end, consider first the test function $z_1 = (T - t, 1)$ which gives $\langle \mathcal{A}\gamma_k, z_1\rangle = \mu^k(0, T \times \mathbf{X} \times \mathbf{U}) + \mu_0^k(\mathbf{X}) + \hat{\mu}_0^k(\mathbf{X}) \to \langle \nu, z_1\rangle < \infty$; since the measures are nonnegative, this implies that the sequences of measures μ^k, μ_0^k and $\hat{\mu}_0^k$ are bounded. Next, taking the test function $z_2 = (1, 1)$ gives $\langle \mathcal{A}\gamma_k, z_2\rangle = \mu_T(\mathbf{X}) + \hat{\mu}_0(\mathbf{X}) \to \langle \nu, z_2\rangle < \infty$; this implies that the sequence μ_T^k is bounded as well. Thus, from the weak-$\star$ compactness of the unit ball (Alaoglu's Theorem [Lue69, Section 5.10, Theorem 1]) there is a subsequence γ_{k_i} that converges weakly-$\star$ to an element $\gamma \in \mathcal{K}^\star$ so that $\lim_{i\to\infty}(\mathcal{A}\gamma_{k_i},\, \langle \gamma_{k_i},\, c\rangle) \in S$ by continuity of $\mathcal{A}$. □

Note that, by Theorem 3.1.1, the supremum in the primal LPs (3.6) and (3.7) is attained (by the restriction of the Lebesgue measure to $\boldsymbol{X}_0$). In contrast, the infimum in the dual LP (3.8) is not attained in $C^1([0, T] \times \mathbf{X}) \times C(\mathbf{X})$, but there

[6] The weak-$\star$ topology on $C^1([0, T] \times \mathbf{X})^\star \times M(\mathbf{X})$ is induced by the standard topologies on C^1 and C – the topology of uniform convergence of the function and its derivative on C^1 and the topology of uniform convergence on C.

exists a sequence of feasible solutions to (3.8) whose w-component converges to the discontinuous indicator $I_{\boldsymbol{X}_0}$ as we show next.

Theorem 3.1.3 *There is a sequence of feasible solutions to the dual LP (3.8) such that its w-component converges from above to $I_{\boldsymbol{X}_0}$ in L^1 norm and almost uniformly*[7]*.*

Proof: By Theorem 3.1.1, the optimal solution to the primal is attained by the restriction of the Lebesgue measure to $\boldsymbol{X}_0$. Consequently,

$$p^* = \int_{\mathbf{X}} I_{\boldsymbol{X}_0}(x) d\mu^{\mathrm{leb}}(x). \tag{3.11}$$

By Theorem 3.1.2, there is no duality gap ($p^* = d^*$), and therefore there exists a sequence $(v_k, w_k) \in C^1([0,T] \times \mathbf{X}) \times C(\mathbf{X})$ feasible in (3.8) such that

$$p^* = d^* = \lim_{k\to\infty} \int_{\mathbf{X}} w_k(x)\, d\mu^{\mathrm{leb}}(x). \tag{3.12}$$

From Lemma 3.1.2 we have $w_k \geq I_{\boldsymbol{X}_0}$ on $\mathbf{X}$ for all k. Thus, subtracting (3.11) from (3.12) gives

$$\lim_{k\to\infty} \int_{\mathbf{X}} (w_k(x) - I_{\boldsymbol{X}_0}(x))\, d\mu^{\mathrm{leb}}(x) = 0,$$

where the integrand is nonnegative. Hence w_k converges to $I_{\boldsymbol{X}_0}$ in L^1 norm. From [Ash78, Theorems 2.5.2 and 2.5.3] there exists a subsequence converging almost uniformly. □

3.1.5 SDP approximations

In this section we formulate finite-dimensional SDP relaxations of the primal infinite-dimensional LP (3.7) and finite-dimensional SDP tightenings of the dual LP (3.10). These approximations are derived readily using the results of Section 2.2.2 and 2.2.1 on finite-dimensional SDP approximation of the cone of nonnegative measures from the outside and the cone of nonnegative functions from the inside. In addition,

[7] A sequence of functions w_k converges almost uniformly if $\forall\, \epsilon > 0, \exists\, B \subset \mathbf{X}$, $\mu^{\mathrm{leb}}(B) < \epsilon$, such that $w_k \to w$ uniformly on $\mathbf{X}\backslash B$. Note that almost uniform convergence implies convergence almost everywhere [Ash78, Theorem 2.5.2]

we truncate the equality constraints of (3.7)

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

$$\int_{\mathbf{X}} w(x)\,d\mu_0(x) + \int_{\mathbf{X}} w(x)\,d\hat{\mu}_0(x) = \int_{\mathbf{X}} w(x)\,d\mu^{\mathrm{leb}}(x)$$

by enforcing it only for the particular choice of test functions $v(t,x) = t^\alpha x^\beta$ and $w(x) = x^\gamma$ for all $\alpha \in \mathbb{N}$, $\beta \in \mathbb{N}^n$ and $\gamma \in \mathbb{N}^n$ such that $\alpha + |\beta| \le k_v$ and $\gamma \le k$, where

$$k_v := k - \deg f + 1.$$

The resulting finite-dimensional truncation of this linear system of equations is denoted by

$$A_k(\boldsymbol{y}, \boldsymbol{y}_0, \boldsymbol{y}_T, \hat{\boldsymbol{y}}_0) = b_k,$$

where $\boldsymbol{y}$, $\boldsymbol{y}_0$, $\boldsymbol{y}_T$ and $\hat{\boldsymbol{y}}_0$ represent the truncated moment sequences of the measures $(\mu, \mu_0, \mu_T, \hat{\mu}_0)$.

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

Assumption 3.1.2 *One of the polynomials defining the sets* $\mathbf{X}$, $\mathbf{U}$ *respectively* $\boldsymbol{X_T}$, *is equal to* $g_i^X(x) = R_X - \|x\|_2^2$, $g_i^U(u) = R_U - \|u\|_2^2$ *respectively* $g_i^{X_T}(x) = R_T - \|x\|_2^2$ *for some constants* $R_X \ge 0$, $R_U \ge 0$ *respectively* $R_T \ge 0$ *and the set* $[0,T]$ *is modeled as* $\{t \mid t(T-t) \ge 0\}$.

Assumption 3.1.2 is without loss of generality since the sets $\mathbf{X}$, $\mathbf{U}$ and $\boldsymbol{X_T}$ are bounded, and therefore redundant ball constraints of the form $R_X - \|x\|_2^2 \ge 0$, $R_U - \|u\|_2^2 \ge 0$ and $R_T - \|x\|_2^2 \ge 0$ can always be added to the description of the sets $\mathbf{X}$, $\mathbf{U}$ and $\boldsymbol{X_T}$ for sufficiently large R_X, R_U and R_T. Note also that this assumption implies the Archimedianity condition (see Definition 1) for the sets $[0,T] \times \mathbf{X} \times \mathbf{U}$, $\mathbf{X}$ and $\boldsymbol{X_T}$.

Using the results of Section (2.2.2), the finite-dimensional SDP relaxation of order

k reads

$$
\begin{array}{rclll}
p_k^* & = & \max\limits_{(\boldsymbol{y},\boldsymbol{y}_0,\boldsymbol{y}_T,\hat{\boldsymbol{y}}_0)} & (\boldsymbol{y}_0)_0 & \\
& & \text{s.t.} & A_k(\boldsymbol{y},\boldsymbol{y}_0,\boldsymbol{y}_T,\hat{\boldsymbol{y}}_0) = b_k & \\
& & & \boldsymbol{y} \ \in M_k^{\sup}([0,T]\times\mathbf{X}\times\mathbf{U})_+ & \\
& & & \boldsymbol{y}_0 \ \in M_k^{\sup}(\mathbf{X})_+, & \\
& & & \boldsymbol{y}_T \ \in M_k^{\sup}(\boldsymbol{X_T})_+, & \\
& & & \hat{\boldsymbol{y}}_0 \ \in M_k^{\sup}(\mathbf{X})_+, &
\end{array}
\tag{3.13}
$$

where the truncated moment cone is defined in (2.13) and the objective function is the first element (i.e., the mass) of the truncated moment sequence $\boldsymbol{y}_0$ corresponding to the initial measure. The constraint set of (3.13) is therefore looser than that of (3.7) in the sense the the truncated moment sequences associated to any tuple of measure $(\mu, \mu_0, \mu_T, \hat{\mu}_0)$ feasible in (3.7) are also feasible in (3.13) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the relaxation degree k tends to infinity (see Corollary 3.1.1 below).

In problem (3.13), a linear function is minimized subject to linear constraints and subject to inclusions in SDP representable cones. Therefore the problem (3.13) translates to and SDP.

Using the results of Section 2.2.1, the finite-dimensional SDP tightening of (3.10), which is also the SDP dual to (3.13), reads

$$
\begin{array}{rclll}
d_k^* & = & \inf\limits_{v\in\mathbb{R}[t,x]_{k_v},\, w\in\mathbb{R}[x]_k} & \boldsymbol{w}^\top \boldsymbol{l} & \\
& & \text{s.t.} & -\mathcal{L}v & \in Q_k([0,T]\times\mathbf{X}\times\mathbf{U}) \\
& & & w - v(0,\cdot) - 1 & \in Q_k(\mathbf{X}) \\
& & & v(T,\cdot) & \in Q_k(\boldsymbol{X_T}) \\
& & & w & \in Q_k(\mathbf{X}),
\end{array}
\tag{3.14}
$$

where $\boldsymbol{l}$ is the vector of the moments of the Lebesgue measure over $\mathbf{X}$ indexed in the same basis in which the polynomial $w(x)$ with coefficients $\boldsymbol{w}$ is expressed. In problem (3.14), a linear objective function is minimized subject to the inclusion into an SDP representable cone and hence this problem translates to an SDP. The constraint set of problem (3.14) is tighter than that of (3.8) in the sense that any pair of functions (v, w) feasible in (3.14) is also feasible in (3.8) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the degree k tends to infinity (see Corollary 3.1.1 below).

Theorem 3.1.4 *There is no duality gap between primal SDP problem (3.13) and* *dual SDP problem (3.14), i.e.,* $p_k^* = d_k^*$.

Proof: See the Appendix of this section. □

3.1.6 Outer approximations and convergence results

In this section we show how the dual SDP problem (3.14) gives rise to a sequence of outer approximations to the ROA $\boldsymbol{X}_0$ with a guaranteed convergence. In addition, we prove convergence of the primal and dual optimal values p_k^* and d_k^* to the volume of the ROA, and convergence of the w-components of optimal solutions to the dual SDP problem (3.14) to the indicator function of the ROA $I_{\boldsymbol{X}_0}$.

Let the polynomials $(w_k, v_k) \in \mathbb{R}[x]_{k_v} \times \mathbb{R}[x]_k$, denote an optimal solution to the k^{th} degree dual SDP approximation (3.14) and let $\bar{w}_k := \min_{i\le k} w_i$ and $\bar{v}_k := \min_{i\le k} v_i$ denote their running minima. Then, in view of Lemma 3.1.2 and the fact that any feasible solution to (3.14) is feasible in (3.8), the sets

$$\mathbf{X}_{0k} := \{x \in \mathbf{X} : v_k(0, x) \ge 0\} \tag{3.15}$$

and

$$\bar{\mathbf{X}}_{0k} := \{x \in \mathbf{X} : \bar{v}_k(0, x) \ge 0\} \tag{3.16}$$

provide outer approximations to the ROA; in fact, the inclusions $\mathbf{X}_{0k} \supset \bar{\mathbf{X}}_{0k} \supset \boldsymbol{X}_0$ hold for all $k \in \{1, 2, \ldots\}$.

Our first convergence result proves convergence of w_k and $\bar{w}_k$ to the indicator function of the ROA $I_{\boldsymbol{X}_0}$.

Theorem 3.1.5 *Let* $w_k \in \mathbb{R}_k[x]$ *denote the* w*-component of an optimal solution to the dual SDP (3.14) and let* $\bar{w}_k(x) = \min_{i\le k} w_i(x)$*. Then* w_k *converges from above to* $I_{\boldsymbol{X}_0}$ *in* L^1 *norm, i.e.,*

$$\lim_{k\to\infty} \int_{\mathbf{X}} w_k(x) - I_{\boldsymbol{X}_0}\, dx = 0.$$

In addition, $\bar{w}_k$ *converges to* $I_{\boldsymbol{X}_0}$ *from above both in* L^1 *norm and almost uniformly.*

Proof: From Lemma 3.1.2 and Theorem 3.1.3, for every $\epsilon > 0$ there exists a $(v, w) \in C^1([0, T] \times \mathbf{X}) \times C(\mathbf{X})$ feasible in (3.8) such that $w \ge I_{\boldsymbol{X}_0}$ and $\int_X (w - I_{\boldsymbol{X}_0})\, d\mu^{\text{leb}} < \epsilon$.

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 (3.8), 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 (3.8) with a margin at least ϵ. Since $[0,T] \times \mathbf{X}$ and $\mathbf{X}$ are compact, there exist[8] polynomials $\hat{v}$ and $\hat{w}$ of a sufficiently high degree such that $\sup_{[0,T]\times\mathbf{X}} |\tilde{v} - \hat{v}| < \epsilon$, $\sup_{[0,T]\times\mathbf{X}\times\mathbf{U}} |\mathcal{L}\tilde{v} - \mathcal{L}\hat{v}| < \epsilon$ and $\sup_{\mathbf{X}} |\tilde{w} - \hat{w}| < \epsilon$. The pair of polynomials $(\hat{v}, \hat{w})$ is therefore *strictly* feasible in (3.8) and as a result, under Assumption 3.1.2, feasible in (3.14) for a sufficiently large degree k (this follows from Theorem 2.2.1), and moreover $\hat{w} \geq w$. Consequently, $\int_{\mathbf{X}} |\tilde{w} - \hat{w}| \, d\mu^{\mathrm{leb}} \leq \epsilon \mu^{\mathrm{leb}}(\mathbf{X})$, and so $\int_X (\hat{w} - w) \, d\mu^{\mathrm{leb}} \leq \epsilon \mu^{\mathrm{leb}}(\mathbf{X})(T+4)$. Therefore

$$\int_{\mathbf{X}} (\hat{w} - I_{\boldsymbol{X}_0}) \, d\mu^{\mathrm{leb}} < \epsilon K, \quad \hat{w} \geq I_{\boldsymbol{X}_0},$$

where $K := [1 + (T+4)\mu^{\mathrm{leb}}(\mathbf{X})] < \infty$ is a constant. This proves the first statement since ϵ was arbitrary.

The second statement immediately follows since given a sequence $w_k \to I_{\boldsymbol{X}_0}$ in L_1 norm, there exists a subsequence w_{k_i} that converges almost uniformly to $I_{\boldsymbol{X}_0}$ by [Ash78, Theorems 2.5.2 and 2.5.3] and clearly $\bar{w}_k(x) \leq \min\{w_{k_i}(x) \; : \; k_i \leq k\}$. □

The following Corollary follows immediately from Theorem 3.1.5.

Corollary 3.1.1 *The sequence of infima of SDPs (3.14) converges monotonically from above to the supremum of the infinite-dimensional LP problem (3.8), 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 (3.13) converges monotonically from above to the maximum of the infinite-dimensional LP problem (3.6), 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 3.1.5 and from the fact that the higher the relaxation order k, the looser the constraint set of the minimization problem (3.14). To prove convergence of the primal maxima observe that from weak SDP duality we have $d^*_k \geq p^*_k$ and from Theorems 3.1.5 and 3.1.2 it follows that $d^*_k \to d^* = p^*$. In addition, clearly $p^*_k \geq p^*$

[8]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., [BBL02].

and $p^*_{k+1} \leq p^*_k$ since the higher the relaxation order k, the tighter the constraint set of the maximization problem (3.13). Therefore $p^*_k \to p^*$ monotonically from above. $\square$

Theorem 3.1.5 establishes a functional convergence of w_k to $I_{\boldsymbol{X}_0}$ and Corollary 3.1.1 a convergence of the primal and dual optima p^*_k and d^*_k to the volume of the ROA $\mu^{\mathrm{leb}}(\boldsymbol{X}_0) = p^* = d^*$. Finally, the following theorem establishes a set-wise convergence of the sets (3.15) and (3.16) to the ROA $\boldsymbol{X}_0$.

Theorem 3.1.6 *Let $(v_k, w_k) \in \mathbb{R}_{k_v}[t,x] \times \mathbb{R}_k[x]$ denote a solution to the dual SDP problems (3.14). Then the sets $\mathbf{X}_{0k}$ and $\bar{\mathbf{X}}_{0k}$ defined in (3.15) and (3.16) converge to the ROA $\boldsymbol{X}_0$ from the outside such that $\mathbf{X}_{0k} \supset \bar{\mathbf{X}}_{0k} \supset \boldsymbol{X}_0$ and*

$$\lim_{k\to\infty} \mu^{\mathrm{leb}}(\mathbf{X}_{0k} \setminus \boldsymbol{X}_0) = 0 \quad \textit{and} \quad \lim_{k\to\infty} \mu^{\mathrm{leb}}(\bar{\mathbf{X}}_{0k} \setminus \boldsymbol{X}_0) = 0.$$

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

Proof: The inclusion $\mathbf{X}_{0k} \supset \bar{\mathbf{X}}_{0k} \supset \boldsymbol{X}_0$ follows from Lemma 3.1.2 since any solution to (3.14) is feasible in (3.8) and since $\bar{\mathbf{X}}_{0k} = \cap_{i=1}^{k} \mathbf{X}_{0i}$. The latter fact also proves the monotonicity of the sequence $\bar{\mathbf{X}}_{0k}$. Next, from Lemma 3.1.2 we have $w_k \geq I_{\boldsymbol{X}_0}$ and therefore, since $w_k \geq v_k(0,\cdot) + 1$ on $\mathbf{X}$, we have $w_k \geq I_{\mathbf{X}_{0k}} \geq I_{\bar{\mathbf{X}}_{0k}} \geq I_{\boldsymbol{X}_0}$ on $\mathbf{X}$. In addition, from Theorem 3.1.5 we have $w_k \to I_{\boldsymbol{X}_0}$ in L^1 norm on $\mathbf{X}$; therefore

$$\begin{aligned} \mu^{\mathrm{leb}}(\boldsymbol{X}_0) = \int_{\mathbf{X}} I_{\boldsymbol{X}_0}\, d\mu^{\mathrm{leb}} &= \lim_{k\to\infty} \int_{\mathbf{X}} w_k\, d\mu^{\mathrm{leb}} \geq \lim_{k\to\infty} \int_{\mathbf{X}} I_{\mathbf{X}_{0k}}\, d\mu^{\mathrm{leb}} \\ &= \lim_{k\to\infty} \mu^{\mathrm{leb}}(\mathbf{X}_{0k}) \geq \lim_{k\to\infty} \mu^{\mathrm{leb}}(\bar{\mathbf{X}}_{0k}). \end{aligned}$$

But since $\boldsymbol{X}_0 \subset \bar{\mathbf{X}}_{0k} \subset \mathbf{X}_{0k}$ we must have $\mu^{\mathrm{leb}}(\boldsymbol{X}_0) \leq \mu^{\mathrm{leb}}(\bar{\mathbf{X}}_{0k}) \leq \mu^{\mathrm{leb}}(\mathbf{X}_{0k})$ and the theorem follows. $\square$

3.1.7 Free final time

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

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

$$
\begin{array}{rcll}
p^* & = & \sup\limits_{\mu_0,\mu,\mu_T} & \int_{\mathbf{X}} 1 \, d\mu_0 \\
& & \text{s.t.} & \int_{[0,T]\times \boldsymbol{X_T}} v \, d\mu_T = \int_{\mathbf{X}} v(0,\cdot)\, d\mu_0 + \int\limits_{[0,T]\times\mathbf{X}\times\mathbf{U}} \mathcal{L}v \, d\mu \quad \forall v \in C^1([0,T]\times\mathbf{X}) \\
& & & \int_{\mathbf{X}} w \, d\mu_0 + \int_{\mathbf{X}} w \, d\hat{\mu}_0 = \int_{\mathbf{X}} w \, d\mu^{\mathrm{leb}} \quad \forall w \in C(\mathbf{X}) \\
& & & \mu \in M([0,T]\times\mathbf{X}\times\mathbf{U})_+ \\
& & & \mu_0 \in M(\mathbf{X})_+ \\
& & & \mu_T \in M([0,T]\times\boldsymbol{X_T})_+ \\
& & & \hat{\mu}_0 \in M(\mathbf{X})_+
\end{array}
\tag{3.17}
$$

Note that the only difference to problem (3.6) is in the support constraint of the final measure μ_T.

The dual to this problem reads as

$$
\begin{array}{rlcl}
d^* = \inf\limits_{v,w} & \displaystyle\int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x) & & \\
\text{s.t.} & -\mathcal{L}v(t,x,u) & \in & C([0,T]\times\mathbf{X}\times\mathbf{U})_+, \\
& w - v(0,\cdot) - 1 & \in & C(\mathbf{X})_+, \\
& v(T,\cdot) & \in & C([0,T]\times\boldsymbol{X_T})_+, \\
& w & \in & C(\mathbf{X})_+,
\end{array}
\tag{3.18}
$$

The only difference to problem (3.8) is in the third constraint which now requires that $v(t,x)$ is nonnegative on $\boldsymbol{X_T}$ for *all* $t \in [0,T]$.

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

3.1.8 Numerical examples

In this section we present five examples of increasing complexity to illustrate our approach: a univariate uncontrolled cubic system, the Van der Pol oscillator, a double integrator, the Brockett integrator and an acrobot. For numerical implementation, one can either use Gloptipoly 3 [HLL09] 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 [Löf04] or SOSTOOLS [PAGV$^+$13]. As an SDP solver we used SeDuMi [PTZ07] 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 Chapter 6.

Whenever the approximations $\mathbf{X}_{0k}$ defined in (3.15) are monotonous (which is not guaranteed) we report these approximations (since then they are equal to the monotonous version $\bar{\mathbf{X}}_{0k}$ defined in (3.16)); otherwise we report $\bar{\mathbf{X}}_{0k}$.

Univariate cubic dynamics

Consider the system given by

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

the constraint set $\mathbf{X} = [-1, 1]$, the final time $T = 100$ and the target set $\boldsymbol{X_T} = [-0.01, 0.01]$. The ROA can in this case be determined analytically as $\boldsymbol{X}_0 = [-0.5, 0.5]$. Polynomial approximations to the ROA for degrees $d \in \{4, 8, 16, 32\}$ are shown in Figure 3.1. As expected the functional convergence of the polynomials to the discontinuous indicator function is rather slow; however, the set-wise convergence of the approximations $\mathbf{X}_{0k}$ is very fast as shown in Table 3.1. 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 $\mathbf{X}_{0k}$. Taking the monotonically decreasing approximations $\bar{\mathbf{X}}_{0k}$ defined in (3.16) 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 3.1 – Univariate cubic dynamics – relative volume error of the outer approximation to the ROA $\boldsymbol{X}_0 = [-0.5, 0.5]$ as a function of the approximating polynomial degree.

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

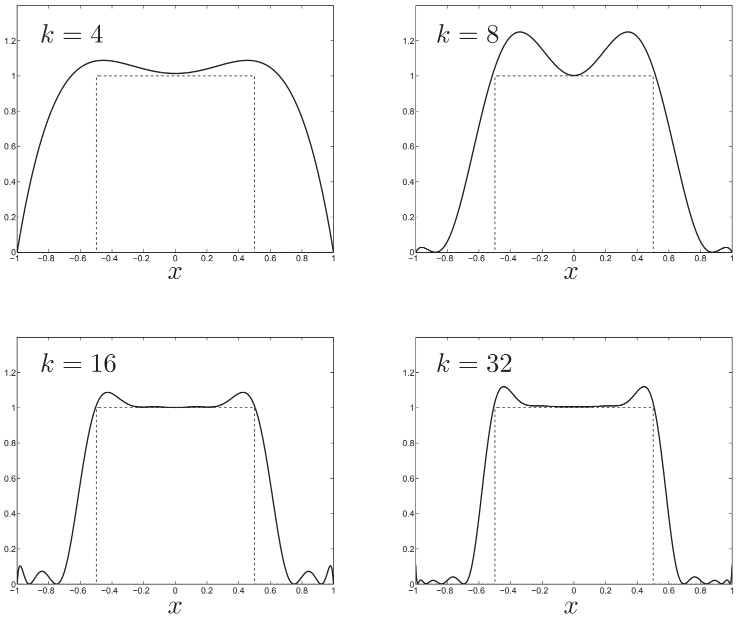


Figure 3.1 – Univariate cubic dynamics – polynomial approximations (solid line) to the ROA indicator function $I_{\boldsymbol{X}_0} = I_{[-0.5,0.5]}$ (dashed line) for degrees $k \in \{4, 8, 16, 32\}$.

Van der Pol oscillator

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

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

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

$$\boldsymbol{X}_0 \subset X := [-1.2,\ 1.2]^2.$$

In order to compute an outer approximation to this region we take $T = 100$ and $\boldsymbol{X_T} = \{x : \|x\|_2 \leq 0.01\}$. Plots of the ROA estimates $\mathbf{X}_{0k}$ for $k \in \{10, 12, 14, 16\}$ are shown in Figure 3.2. We observe a relatively-fast convergence of the super-level sets to the ROA – this is confirmed by the relative volume error[9] summarized in Table 3.2. Figure 3.3 then shows the approximating polynomial itself for degree $k = 18$. Here too, a better convergence is expected if instead of monomials, a more

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

appropriate polynomial basis is used.

Table 3.2 – Van der Pol oscillator – relative error of the outer approximation to the ROA $\boldsymbol{X}_0$ as a function of the approximating polynomial degree.

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

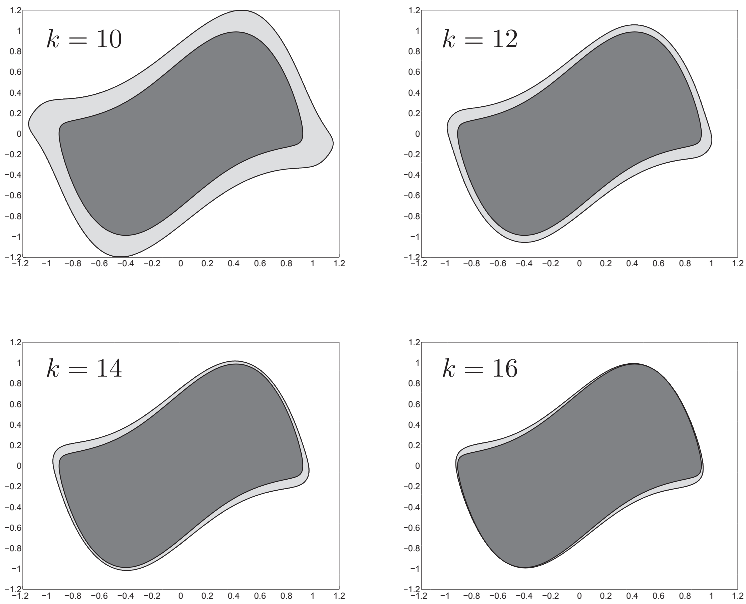


Figure 3.2 – Van der Pol oscillator – semialgebraic outer approximations (light gray) to the ROA (dark gray) for degrees $k \in \{10, 12, 14, 16\}$.

Double integrator

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

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

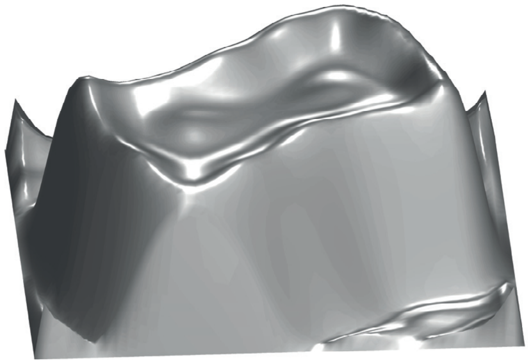


Figure 3.3 – Van der Pol oscillator – a polynomial approximation of degree 18 of the ROA indicator function $I_{\boldsymbol{X}_0}$.

The goal is to find an approximation to the set of all initial states $\boldsymbol{X}_0$ that can be steered to the origin at[10] the final time $T = 1$. Therefore we set $X_T = \{0\}$ and the constraint set such that $\boldsymbol{X}_0 \subset \mathbf{X}$, e.g., $\mathbf{X} = [-0.7, 0.7] \times [-1.2, 1.2]$. The solution to this problem can be computed analytically as

$$\boldsymbol{X}_0 = \{x \, : \, V(x) \leq 1\},$$

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 $\mathbf{X}_{0k}$ for $k \in \{6, 8, 10, 12\}$ are shown in Figure 3.4; again we observe a relatively fast convergence of the super-level set approximations, which is confirmed by the relative volume errors in Table 3.3.

Table 3.3 – Double integrator – relative error of the outer approximation to the ROA $\boldsymbol{X}_0$ as a function of the approximating polynomial degree.

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

[10]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 3.1.7.

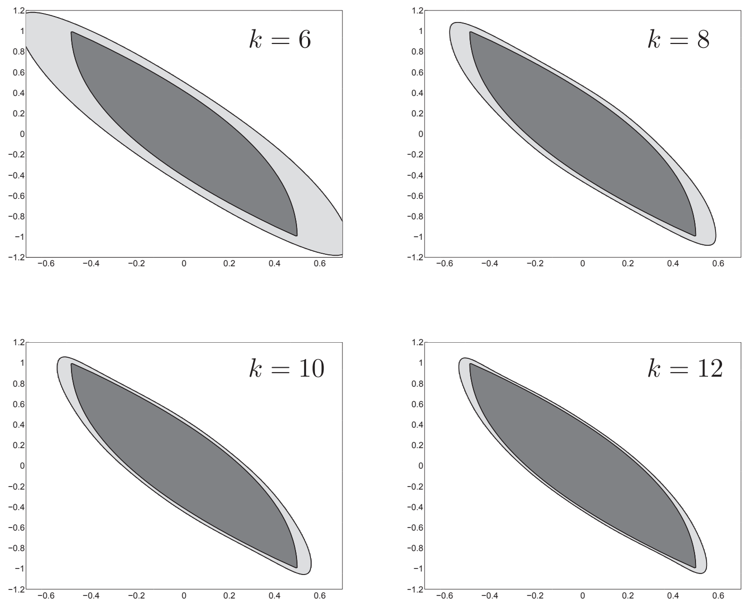


Figure 3.4 – Double integrator – semialgebraic outer approximations (light gray) to the ROA (dark gray) for degrees $k \in \{6, 8, 10, 12\}$.

Brockett integrator

Next, we consider the Brockett integrator

$$\begin{aligned}\dot{x}_1 &= u_1\\ \dot{x}_2 &= u_2\\ \dot{x}_3 &= u_1x_2 - u_2x_1\end{aligned}$$

with the constraint sets $\mathbf{X} = \{x \in \mathbb{R}^3 : ||x||_\infty \leq 1\}$ and $\mathbf{U} = \{u \in \mathbb{R}^2 : ||u||_2 \leq 1\}$, the target set $\boldsymbol{X_T} = \{0\}$ and the final time $T = 1$. The ROA can be computed analytically (see [LHPT08]) as $\boldsymbol{X}_0 = \{x \in \mathbb{R}^3 \,:\, \mathcal{T}(x) \leq 1\}$, where

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

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 $\mathbf{X}_{0k}$ are not monotonous in this case and therefore in Figure 3.5 we rather show the monotonous estimates $\bar{\mathbf{X}}_{0k}$ defined in (3.16) for degrees $k \in \{6, 10\}$. We observe fairly good tightness of the estimates.

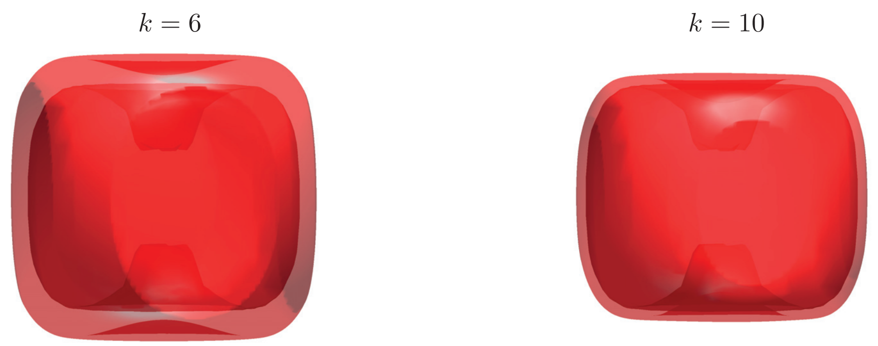


Figure 3.5 – Brockett integrator – semialgebraic outer approximations (light red, larger) to the ROA (dark red, smaller) for degrees $k \in \{6, 10\}$.

Acrobot

As our last example we consider the acrobot system adapted from [MH91], which is essentially a double pendulum with both joints actuated; see Figure 3.6. 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 $\mathbf{U} = [-10, 10] \times [-10, 10]$ and in the second case it is $\mathbf{U} = \{0\} \times [-10, 10]$. The state constraint set is for both cases $\mathbf{U} = [-\pi/2, \pi/2] \times [-\pi, \pi] \times [-5, 5] \times [-5, 5]$. 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 thesis and therefore we use the simpler (and non-exact) approach with Taylor expansion. Figure 3.7 shows the approximations $\mathbf{X}_{0k}$ of degree $k \in \{6, 8\}$; as expected disabling actuation of the first joint leads to a smaller ROA approximation. For this largest example presented in this section we also report computation times for two SDP solvers: the recently released MOSEK SDP solver and SeDuMi. Computation times[11] reported in Table 3.4 show that MOSEK outperforms SeDuMi in terms of speed by a large margin; this finding does not seem to be specific to this particular problem and holds for all ROA computation problems presented. Before solving, the problem data was scaled such that the constraint sets become unit boxes.

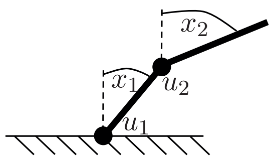


Figure 3.6 – Acrobot – sketch

Table 3.4 – 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	–

[11] Table 3.4 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.

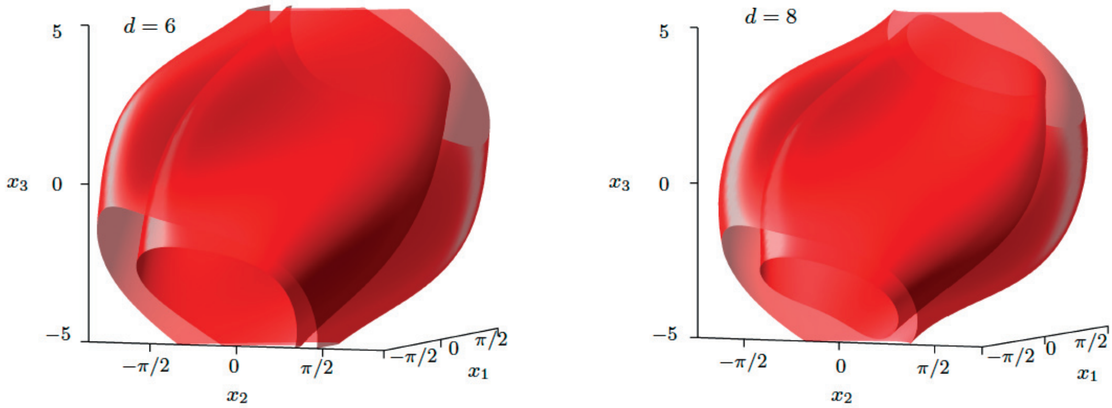


Figure 3.7 – Acrobot – section for $x_4 = 0$ of the semialgebraic outer approximations of degree $k \in \{6, 8\}$. Only the middle joint actuated – darker, smaller; both joints actuated – lighter, larger. The states displayed x_1, x_2 and x_3 are, respectively, the lower pendulum angle, the upper pendulum angle and the lower pendulum angular velocity.

3.1.9 Appendix

In this appendix we prove Theorem 3.2.5. In order to prove the theorem we rewrite primal LMI problem (3.13) in a vectorized form as follows

$$\begin{array}{rcll} p_k^* & = & \min & \mathbf{c}^\top \mathbf{y} \\ & & \text{s.t.} & \mathbf{A}\mathbf{y} = \mathbf{b} \\ & & & \mathbf{e} + \mathbf{D}\mathbf{y} \in \mathbf{K}, \end{array} \tag{3.19}$$

where $\mathbf{y} := [y^\top,\ y_0^\top,\ y_T^\top, \hat{y}_0^\top]^\top$ 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 (3.14) then becomes

$$\begin{array}{rcll} d_k^* & = & \max & \mathbf{b}^\top \mathbf{x} - \mathbf{e}^\top \mathbf{z} \\ & & \text{s.t.} & \mathbf{A}^\top \mathbf{x} + \mathbf{D}^\top \mathbf{z} = \mathbf{c} \\ & & & \mathbf{z} \in \mathbf{K}, \end{array} \tag{3.20}$$

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 (3.19) and dual LMI (3.20). The notation $\text{int}\,\mathbf{K}$ stands for the interior of $\mathbf{K}$.

Lemma 3.1.3 *If matrix* $\mathbf{D}$ *has full column rank, exactly one of these statements is true:*

- *there exists* $\mathbf{x}$ *and* $\mathbf{z} \in \text{int}\,\mathbf{K}$ *such that* $\mathbf{A}^\top \mathbf{x} + \mathbf{D}^\top \mathbf{z} = \mathbf{c}$
- *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 3.1.3: A classical lemma of alternatives states that if matrix $\bar{\mathbf{D}}$ has full column rank, then either there exists $\mathbf{z} \in \text{int}\mathbf{K}$ such that $\bar{\mathbf{D}}^\top \mathbf{z} = \bar{\mathbf{c}}$ or there exists $\bar{\mathbf{y}}$ such that $\bar{\mathbf{D}}\bar{\mathbf{y}} \in \mathbf{K}$ and $\bar{\mathbf{c}}^\top \bar{\mathbf{y}} \leq 0$, but not both, see e.g. [Trn05, 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 $\mathbf{A}^\top \mathbf{x} + \mathbf{D}^\top \mathbf{z} = \mathbf{c}$ if and only if $\mathbf{z}$ is such that $\bar{\mathbf{D}}^\top \mathbf{z} = \bar{\mathbf{c}}$ with $\bar{\mathbf{D}} = \mathbf{D}\mathbf{F}$, $\bar{\mathbf{c}} = \mathbf{F}^\top \mathbf{c}$ 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 3.2.5: First notice that the feasibility set of LMI problem (3.19) is nonempty and bounded. Indeed, a triplet of zero measures is a trivial feasible point for (3.6) and hence $(0, 0, 0, \mu^{\text{leb}})$ is feasible in (3.7); consequently a concatenation of truncated moment sequences corresponding to the quadruplet of measures $(0, 0, 0, \mu^{\text{leb}})$ is feasible in (3.19) 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 3.1.2 and from the fact that the masses (zero-th moments) of the measures are bounded because of the constraint $\mu_0 + \hat{\mu}_0 = \mu^{\text{leb}}$ 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 [Trn05, Theorem 4] and show that boundedness of the feasibility set of LMI problem (3.19) implies existence of an interior point for LMI problem (3.20), and then from standard SDP duality it follows readily that $p^* = d^*$ since D has a full column rank; see, e.g., [Trn05, Theorem 5] and references therein.

Let $\bar{\mathbf{y}}$ denote a point in the feasibility set of LMI problem (3.19), 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 (3.20), i.e. there is no $\mathbf{x}$ and $\mathbf{z} \in \operatorname{int} \mathbf{K}$ such that $\mathbf{A}'\mathbf{x} + \mathbf{D}'\mathbf{z} = \mathbf{c}$. Then from Lemma 3.1.3 there exists $\mathbf{y} \neq 0$ such that $\mathbf{A}\mathbf{y} = 0$, $\mathbf{D}\mathbf{y} \in \mathbf{K}$ and $\mathbf{c}'\mathbf{y} \leq 0$. It follows that there exists a ray $\bar{\mathbf{y}} + t\mathbf{y}$, $t \geq 0$ of feasible points for LMI problem (3.19), hence implying that the feasibility set is not bounded. ■

3.2 Maximum controlled invariant set

The approach is developed in parallel for discrete and continuous time.

Discrete time

Consider the discrete-time controlled system

$$x_{t+1} = f(x_t, u_t), \quad x_t \in \mathbf{X}, \quad u_t \in \mathbf{U}, \quad t \in \{0, 1, \ldots\} \tag{3.21}$$

with a given polynomial transition mapping f with entries $f_i \in \mathbb{R}[x, u]$, $i = 1, \ldots, n$, and given compact[12] basic semialgebraic state and input constraints

$$\begin{aligned} x_t \in \mathbf{X} &:= \{x \in \mathbb{R}^n \; : \; g_{X\,i}(x) \geq 0, i = 1, 2, \ldots, n_X\}, \\ u_t \in \mathbf{U} &:= \{u \in \mathbb{R}^m \; : \; g_{U\,i}(u) \geq 0, i = 1, 2, \ldots, n_U\} \end{aligned}$$

with $g_{X\,i} \in \mathbb{R}[x]$, $g_{U\,i} \in \mathbb{R}[u]$.

The maximum controlled invariant (MCI) set is defined as

$$\begin{aligned} \boldsymbol{X}_\infty := \Big\{ x_0 \in \mathbf{X} \; : \exists \left(\{x_t\}_{t=1}^\infty, \{u_t\}_{t=1}^\infty \right) \text{ s.t. } x_{t+1} = f(x_t, u_t), \\ u_t \in \mathbf{U},\, x_t \in \mathbf{X},\, \forall t \in \{0, 1, \ldots\} \Big\}. \end{aligned}$$

A control sequence $\{u_t\}_{t=0}^\infty$ is called *admissible* if $u_t \in \mathbf{U}$ for all $t \in \{0, 1, \ldots\}$.

[12]The assumption of compactness is crucial for the subsequent theoretical developments. On the other hand, the assumptions that the constraint sets are basic semialgebraic and the mapping f polynomial plays no role in the infinite-dimensional considerations and only facilitates the convergence results of the finite-dimensional relaxations. The above is valid for both discrete and continuous time.

In words, the MCI set is the set of all initial states which can be kept inside the constraint set $\mathbf{X}$ forever using admissible control inputs.

Continuous time

Consider the relaxed continuous-time controlled system

$$\dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}), \quad x(t) \in \mathbf{X}, \quad t \in [0, \infty), \tag{3.22}$$

where conv denotes the convex hull[13], f is a polynomial vector field with entries $f_i \in \mathbb{R}[x, u]$, $i = 1, \ldots, n$, and compact basic semialgebraic state and input constraint sets are defined by

$$\begin{aligned} \mathbf{X} &:= \{x \in \mathbb{R}^n \ : \ g_{X_i}(x) \geq 0, i = 1, 2, \ldots, n_X\}, \\ \mathbf{U} &:= \{u \in \mathbb{R}^m \ : \ g_{U_i}(u) \geq 0, i = 1, 2, \ldots, n_U\} \end{aligned}$$

with $g_{X_i} \in \mathbb{R}[x]$, $g_{U_i} \in \mathbb{R}[u]$.

The maximum controlled invariant (MCI) set is defined as

$$\boldsymbol{X}_\infty := \Big\{ x_0 \in X \ : \ \exists\, x(\cdot) \ \text{s.t.} \ \dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}) \text{ a.e.}, \ x(t) \in \mathbf{X}\, \forall\, t \in [0, \infty) \Big\},$$

where $x(\cdot)$ is required to be absolutely continuous and a.e. stands for "almost everywhere" with respect to the Lebesgue measure on $[0, \infty)$.

In words, the MCI set is the set of all initial states for which there exists a trajectory of the convexified inclusion (3.22) which remains in $\mathbf{X}$ forever.

3.2.1 Lifting: Primal LP

In this section we show how the MCI set computation problem can be cast as an infinite-dimensional LP in the cone of nonnegative measures. The procedure is completely analogous to Section 3.1.3 and hence we only outline it. Lifting of the system dynamics (3.21) and (3.22) in continuous and discrete time, respectively, are the corresponding *discounted Liouville's* equations (2.43) and (2.28). Analogously to Section 3.1.3 the problem of MCI set computation can be cast

[13] Since f is continuous and $\mathbf{U}$ compact, conv $f(x, \mathbf{U})$ is necessarily compact and hence closure is not required in (3.22) as opposed to Section 2.3.1 where f was not required to be continuous.

as an infinite-dimensional optimization problem with the decision variables being the measures (μ, μ_0) appearing in the Liouville's equation and the objective being the maximization of the support of the initial measure μ_0. The non-convexity of support maximization is than circumvented by maximizing the mass of the initial measure μ_0 subject to the constraint that it is dominated by the Lebesgue measure, that is, $\mu_0 \leq \mu^{\mathrm{leb}}$. State and input constraints are expressed through constraints on the supports of the initial and occupation measure. The constraint that $\mu_0 \leq \mu^{\mathrm{leb}}$ can be equivalently rewritten as $\mu_0 + \hat{\mu}_0 = \mu^{\mathrm{leb}}$ for some non-negative slack measure $\hat{\mu}_0 \in M(\mathbf{X})_+$. This constraint is in turn equivalent to $\int_{\mathbf{X}} w(x)\, d\mu_0(x) + \int_{\mathbf{X}} w(x)\, d\hat{\mu}_0(x) = \int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x)$ for all $w \in C(\mathbf{X})$. These considerations lead to the following primal LPs.

Discrete time

The primal LP in discrete time reads

$$
\begin{array}{rclll}
p^* & = & \sup\limits_{\mu,\mu_0,\hat{\mu}_0} & \int_{\mathbf{X}} 1\, d\mu_0 & \\
& & \text{s.t.} & \int_{\mathbf{X}\times\mathbf{U}} v\, d\mu = \int_{\mathbf{X}} v\, d\mu_0 + \alpha \int_{\mathbf{X}\times\mathbf{U}} v \circ f\, d\mu & \forall\, v \in C(\mathbf{X}) \\
& & & \int_{\mathbf{X}} w\, d\mu_0 + \int_{\mathbf{X}} w\, d\hat{\mu}_0 = \int_{\mathbf{X}} w\, d\mu^{\mathrm{leb}} & \forall\, w \in C(\mathbf{X}) \\
& & & \mu \in M(\mathbf{X}\times\mathbf{U})_+, & \\
& & & \mu_0 \in M(\mathbf{X})_+, & \\
& & & \hat{\mu}_0 \in M(\mathbf{X})_+, &
\end{array}
\tag{3.23}
$$

where μ^{leb} denotes the Lebesgue measure on $\mathbf{X}$ and the first equality constraint is precisely the discrete-time discounted Liouville's equation (2.43). This is an infinite-dimensional LP in the cone of nonnegative measures. The following crucial theorem relates an optimal solution of this LP to the MCI set $\boldsymbol{X}_\infty$.

Theorem 3.2.1 *The optimal value of LP problem (3.23) is equal to the volume of the MCI set $\boldsymbol{X}_\infty$, that is, $p^* = \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty)$. Moreover, the supremum is attained by the restriction of the Lebesgue measure to the MCI set $\boldsymbol{X}_\infty$.*

Proof: By definition of the MCI set $\boldsymbol{X}_\infty$, for any initial condition $x_0 \in \boldsymbol{X}_\infty$ there exists an admissible control sequence such that the associated state sequence remains in $\mathbf{X}$. Therefore for any initial measure $\mu_0 \leq \mu^{\mathrm{leb}}$ with $\operatorname{spt}\mu_0 \subset \boldsymbol{X}_\infty$ there exist a discounted occupation measure μ with $\operatorname{spt}\mu \subset \mathbf{X}\times\mathbf{U}$ and a slack measure $\hat{\mu}_0$ with $\operatorname{spt}\hat{\mu}_0 \subset \mathbf{X}$ such that the constraints of problem (3.23) are satisfied. One such measure μ_0 is the restriction of the Lebesgue measure to $\boldsymbol{X}_\infty$, and therefore $p^* \geq \mu^{\rm leb}(\boldsymbol{X}_\infty)$. The fact $p^* \leq \mu^{\rm leb}(\boldsymbol{X}_\infty)$ follows from Lemma 3.2.1 below since the first equality constraint of (3.23) is the discounted Liouville's equation (2.43). □

The following lemma is a direct consequence of the superposition Theorem 2.3.4.

Lemma 3.2.1 *For any pair of measures* (μ_0, μ) *satisfying equation (2.43) with* $\mathrm{spt}\,\mu_0 \subset \mathbf{X}$ *and* $\mathrm{spt}\,\mu \subset \mathbf{U} \times \mathbf{X}$ *we have* $\mu^{\rm leb}(\mathrm{spt}\,\mu_0) \leq \mu^{\rm leb}(\boldsymbol{X}_\infty)$.

Proof: Suppose that a pair of measures (μ_0, μ) satisfies (2.43) and that $\mu^{\rm leb}(\mathrm{spt}\,\mu_0 \setminus \boldsymbol{X}_\infty) > 0$. From Theorem 2.3.4 there is a family of trajectories of (3.21) starting from μ_0 with discounted occupation measure whose x-marginal coincides with the x-marginal of μ. However, this is a contradiction since no trajectory starting from $\mathrm{spt}\,\mu_0 \setminus \boldsymbol{X}_\infty$ remains in $\mathbf{X}$ for all times and $\mathrm{spt}\,\mu \subset \mathbf{X}$. Thus, $\mu^{\rm leb}(\mathrm{spt}\,\mu_0 \setminus \boldsymbol{X}_\infty) = 0$ and so $\mu^{\rm leb}(\mathrm{spt}\,\mu_0) \leq \mu^{\rm leb}(\boldsymbol{X}_\infty)$. □

Continuous time

The primal LP in continuous time reads

$$
\begin{array}{rcll}
p^* & = & \sup\limits_{\mu,\mu_0,\hat\mu_0} & \int_{\mathbf{X}} 1 \, d\mu_0 \\
 & & \text{s.t.} & \beta \int_{\mathbf{X}\times\mathbf{U}} v \, d\mu = \int_{\mathbf{X}} v \, d\mu_0 + \int_{\mathbf{X}\times\mathbf{U}} \nabla v \cdot f \, d\mu \qquad \forall\, v \in C^1(\mathbf{X}) \\
 & & & \int_{\mathbf{X}} w \, d\mu_0 + \int_{\mathbf{X}} w \, d\hat\mu_0 = \int_{\mathbf{X}} w \, d\mu^{\rm leb} \qquad \forall\, w \in C(\mathbf{X}) \\
 & & & \mu \in M(\mathbf{X}\times\mathbf{U})_+, \\
 & & & \mu_0 \in M(\mathbf{X})_+, \\
 & & & \hat\mu_0 \in M(\mathbf{X})_+,
\end{array}
\tag{3.24}
$$

where $\mu^{\rm leb}$ denotes the Lebesgue measure on $\mathbf{X}$ and the first equality constraint is precisely the continuous-time discounted Liouville's equation (2.28). This is an infinite-dimensional LP in the cone of nonnegative measures. The following crucial theorem relates an optimal solution of this LP to the MCI set $\boldsymbol{X}_\infty$.

Theorem 3.2.2 *The optimal value of LP problem (3.24) is equal to the volume of the MCI set* $\boldsymbol{X}_\infty$*, that is,* $p^* = \mu^{\rm leb}(\boldsymbol{X}_\infty)$*. Moreover, the supremum is attained by the restriction of the Lebesgue measure to the MCI set* $\boldsymbol{X}_\infty$.

Proof: The fact that μ_0 equal to the restriction of the Lebesgue measure to $\boldsymbol{X}_\infty$ is feasible in (3.24) (and therefore $p^* \geq \mu^{\rm leb}(\boldsymbol{X}_\infty)$) follows by the same arguments as in discrete time. The fact that $p^* \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty)$ follows from Lemma 3.2.2 below since the first equality constraint of (3.24) is the discounted Liouville's equation (2.28). □

The following lemma is a direct consequence of the superposition Theorem 2.3.2.

Lemma 3.2.2 *For any pair of measures* (μ_0, μ) *satisfying equation (2.28) with* $\operatorname{spt} \mu_0 \subset \mathbf{X}$ *and* $\operatorname{spt} \mu \subset \mathbf{U} \times \mathbf{X}$ *we have* $\mu^{\mathrm{leb}}(\operatorname{spt} \mu_0) \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty)$.

Proof: Suppose that a pair of measures (μ_0, μ) satisfies (2.28) and that $\mu^{\mathrm{leb}}(\operatorname{spt} \mu_0 \setminus \boldsymbol{X}_\infty) > 0$. From Theorem 2.3.2 there is a family of trajectories of (3.22) starting from μ_0 with discounted occupation measure whose x-marginal coincides with the x-marginal of μ. However, this is a contradiction since no trajectory starting from $\operatorname{spt} \mu_0 \setminus \boldsymbol{X}_\infty$ remains in $\mathbf{X}$ for all times and $\operatorname{spt} \mu \subset \mathbf{X}$. Thus, $\mu^{\mathrm{leb}}(\operatorname{spt} \mu_0 \setminus \boldsymbol{X}_\infty) = 0$ and so $\mu^{\mathrm{leb}}(\operatorname{spt} \mu_0) \leq \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty)$. □

3.2.2 Lifting: Dual LP

In this section we derive LPs dual to the primal LPs (3.23) and (3.24). Since the primal LPs are in the space of measures, the dual LPs will be on the space of continuous functions. Super-level sets of feasible solutions to these LPs then provide outer approximations to the MCI sets, both in discrete and in continuous time. As in Section 3.1, both duals are derived by standard infinite-dimensional LP duality theory (see Appendix A.2 for a brief introduction).

Discrete time

The dual LP in discrete time reads

$$\begin{array}{rclcl} d^* & = & \inf\limits_{v,w} & \displaystyle\int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x) & \\ & & \text{s.t.} & v - \alpha \cdot v \circ f & \in \; C(\mathbf{X} \times \mathbf{U})_+ \\ & & & w - v - 1 & \in \; C(\mathbf{X})_+ \\ & & & w & \in \; C(\mathbf{X})_+ \end{array} \tag{3.25}$$

where the infimum is over the pair of functions $(v, w) \in C(\mathbf{X}) \times C(\mathbf{X})$.

The following key observation shows that the unit super-level set of any function w feasible in (3.25) provides an outer-approximation to $\boldsymbol{X}_\infty$.

Lemma 3.2.3 *Any feasible solution to problem (3.25) satisfies* $v \geq 0$*,* $w \geq 1$ *on* $\boldsymbol{X}_\infty$ *and* $w \geq I_{\boldsymbol{X}_\infty}$ *on* $\mathbf{X}$*.*

Proof: Given any $x_0 \in \boldsymbol{X}_\infty$ there exists a sequence $\{u_t\}_{t=0}^{\infty}$, $u_t \in \mathbf{U}$, such that $x_t \in \mathbf{X}$ for all t. The first constraint of problem (3.25) is equivalent to $\alpha v(x_{t+1}) \leq v(x_t)$, $t \in \{0, 1, \ldots\}$. By iterating this inequality we get

$$v(x_0) \geq \alpha^t v(x_t) \to 0 \quad \text{as} \quad t \to \infty$$

since $x_t \in \mathbf{X}$ and $\mathbf{X}$ is bounded. Therefore $v(x_0) \geq 0$ and $w(x_0) \geq 1$ for all $x_0 \in \boldsymbol{X}_\infty$. The fact that $w \geq I_{\boldsymbol{X}_\infty}$ on $\mathbf{X}$ then follows from the last constraint of (3.25). □

The following theorem is instrumental in proving the convergence results of Section 3.2.3.

Theorem 3.2.3 *There is no duality gap between primal LP problems (3.23) on measures and dual LP problem (3.25) on functions in the sense that* $p^* = d^*$*.*

Proof: Follows by the same arguments as Theorem 3.1.2 using standard infinite-dimensional LP duality theory (see Appendix A.2) and the fact that the feasible set of the primal LP is nonempty and bounded (in the total variation norm on $M(\mathbf{X} \times \mathbf{U}) \times M(\mathbf{X}) \times M(\mathbf{X})$). To see non-emptiness, notice that the vector of measures $(\mu_0, \mu, \hat{\mu}_0) = (0, 0, \mu^{\mathrm{leb}})$ is trivially feasible. To see the boundedness, it suffices to evaluate the equality constraints of (3.23) for $v(x) = w(x) = 1$. This gives $\mu_0(\mathbf{X}) + \hat{\mu}_0(\mathbf{X}) = \mu^{\mathrm{leb}}(\mathbf{X}) < \infty$ and $\mu(\mathbf{X}) = \mu_0(\mathbf{X})/(1 - \alpha)$, which, since $\alpha \in (0, 1)$ and all measures are nonnegative, proves the assertion. □

Continuous time

The dual LP in continuous time reads

$$
\begin{array}{rcll}
d^* & = & \inf\limits_{v,w} & \displaystyle\int_{\mathbf{X}} w(x)\, d\mu^{\text{leb}}(x) \\
& & \text{s.t.} & \beta v - \nabla v \cdot f \quad \in \; C(\mathbf{X} \times \mathbf{U})_+ \\
& & & w - v - 1 \quad \in \; C(\mathbf{X})_+ \\
& & & w \quad \in \; C(\mathbf{X})_+
\end{array}
\tag{3.26}
$$

where the infimum is over the pair of functions $(v, w) \in C^1(\mathbf{X}) \times C(\mathbf{X})$.

The following key observation shows that the unit super-level set of any function w feasible in (3.26) provides an outer-approximation to $\boldsymbol{X}_\infty$.

Lemma 3.2.4 *Any feasible solution to problem (3.26) satisfies $v \geq 0$, $w \geq 1$ on $\boldsymbol{X}_\infty$ and $w \geq I_{\boldsymbol{X}_\infty}$ on $\mathbf{X}$.*

Proof: Given any $x_0 \in \boldsymbol{X}_\infty$ there exists an admissible relaxed control function $\gamma_t(\cdot)$, $\gamma_t(\mathbf{U}) = 1$, such that $x(t) \in X$ for all t. For that $x(t)$ we have $\frac{d}{dt}v(x(t)) = \int_{\mathbf{U}} \nabla v \cdot f(x(t), u)\, d\gamma_t(u) \leq \int_{\mathbf{U}} \beta v(x(t))\, d\gamma_t(u) = \gamma_t(\mathbf{U})\beta v(x(t)) = \beta v(x(t))$. Then by Gronwall's inequality $v(x(t)) \leq e^{\beta t} v(x_0)$, and consequently

$$v(x_0) \geq e^{-\beta t} v(x(t)) \to 0 \quad \text{as} \quad t \to \infty$$

since $x(t) \in \mathbf{X}$ and $\mathbf{X}$ is bounded. Therefore $v(x_0) \geq 0$ and $w(x_0) \geq 1$ for all $x_0 \in \boldsymbol{X}_\infty$. The fact that $w \geq I_{\boldsymbol{X}_\infty}$ on $\mathbf{X}$ then follows from the last constraint of (3.26). □

The following theorem is instrumental in proving the convergence results of Section 3.2.3.

Theorem 3.2.4 *There is no duality gap between primal LP problems (3.24) on measures and dual LP problem (3.26) on functions in the sense that $p^* = d^*$.*

Proof: Follows by the same arguments as Theorem 3.1.2 using standard infinite-dimensional LP duality theory (see Appendix A.2) and the fact that the feasible set of the primal LP is nonempty and bounded (in the total variation norm on $M(\mathbf{X} \times \mathbf{U}) \times M(\mathbf{X}) \times M(\mathbf{X})$). To see non-emptiness, notice that the vector of measures $(\mu_0, \mu, \hat{\mu}_0) = (0, 0, \mu^{\mathrm{leb}})$ is trivially feasible. To see the boundedness, it suffices to evaluate the equality constraints of (3.24) for $v(x) = w(x) = 1$. This gives $\mu_0(\mathbf{X}) + \hat{\mu}_0(\mathbf{X}) = \mu^{\mathrm{leb}}(\mathbf{X}) < \infty$ and $\mu(\mathbf{X}) = \mu_0(\mathbf{X})/\beta$, which, since $\beta > 0$ and all measures are nonnegative, proves the assertion. □

3.2.3 SDP approximations

In this section we formulate finite-dimensional SDP relaxations of the primal infinite-dimensional LPs (3.23) and (3.24) and finite-dimensional SDP tightenings of the dual LPs (3.25) and (3.26). In complete analogy to Section 3.1.5, these approximations are derived readily using the results of Section 2.2.2 and 2.2.1 on finite-dimensional SDP approximation of the cone of nonnegative measures from the outside and the cone of nonnegative functions from the inside in conjunction with truncating the infinite-dimensional linear eqauality constraints of (3.23) and (3.24) by enforcing it only for polynomial test functions up to a prescribed degree.

Discrete time

The equality constraint of (3.23)

$$\int_{\mathbf{X}\times\mathbf{U}} v(x)\, d\mu(x,u) = \int_{\mathbf{X}} v(x)\, d\mu_0(x) + \alpha \int_{\mathbf{X}\times\mathbf{U}} v(f(x,u))\, d\mu(x,u) \quad \forall\, v \in C(\mathbf{X}),$$

$$\int_{\mathbf{X}} w(x)\, d\mu_0(x) + \int_{\mathbf{X}} w(x)\, d\hat{\mu}_0(x) = \int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x) \quad \forall\, w \in C(\mathbf{X})$$

is truncated by enforcing it only for the particular choice of test functions $v(x) = x^\alpha$ and $w(x) = x^\beta$ for all $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le k_v$ and $|\beta| \le k$, where

$$k_v := \left\lfloor \frac{k}{\deg f} \right\rfloor.$$

The resulting finite-dimensional truncation of this linear system of equations is denoted by

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

where $\boldsymbol{y}$, $\boldsymbol{y}_0$ and $\hat{\boldsymbol{y}}_0$ represent the truncated moment sequences of the measures $(\mu, \mu_0, \hat{\mu}_0)$.

Combining this truncation with the results of Section 2.2.2, the primal relaxation of degree k in discrete time reads

$$
\begin{array}{rcll}
p_k^* & = & \max\limits_{\boldsymbol{y},\boldsymbol{y}_0,\hat{\boldsymbol{y}}_0} & (\boldsymbol{y}_0)_0 \\
& & \text{s.t.} & A_k(\boldsymbol{y},\boldsymbol{y}_0,\hat{\boldsymbol{y}}_0) = b_k \\
& & & \boldsymbol{y} \in M_k^{\sup}(\mathbf{X}\times\mathbf{U})_+ \\
& & & \boldsymbol{y}_0 \in M_k^{\sup}(\mathbf{X})_+, \\
& & & \hat{\boldsymbol{y}}_0 \in M_k^{\sup}(\mathbf{X})_+,
\end{array}
\tag{3.27}
$$

where the truncated moment $M_k^{\sup}(\cdot)_+$ cone is defined in (2.13) and the objective function is the first element (i.e., the mass) of the truncated moment sequence $\boldsymbol{y}_0$ corresponding to the initial measure. In problem (3.27), a linear objective is minimized subject to SDP representable conic inclusions and therefore problem (3.27) translates to a semidefinite program (SDP). The constraint set of (3.27) is looser than that of (3.23) in the sense the the truncated moment sequences associated to any triple of measure $(\mu,\mu_0,\hat{\mu}_0)$ feasible in (3.23) are also feasible in (3.27) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the relaxation degree k tends to infinity (see Theorem 3.2.5).

Using the results of Section 2.2.1, we immediately arrive at a degree-k SDP tightening of (3.25), which is also the SDP dual to (3.27):

$$
\begin{array}{rclll}
d_k^* & = & \inf\limits_{v\in\mathbb{R}[x]_{k_v},\, w\in\mathbb{R}[x]_k} & \boldsymbol{w}^\top \boldsymbol{l} & \\
& & \text{s.t.} & v - \alpha\cdot v\circ f & \in Q_k(\mathbf{X}\times\mathbf{U}) \\
& & & w - v - 1 & \in Q_k(\mathbf{X}) \\
& & & w & \in Q_k(\mathbf{X}),
\end{array}
\tag{3.28}
$$

where $\boldsymbol{l}$ is the vector of Lebesgue moments over $\mathbf{X}$ indexed in the same basis in which the polynomial $w(x)$ with coefficients $\boldsymbol{w}$ is expressed and the quadratic module $Q_k(\cdot)$ is defined in (2.8). In problem (3.28), a linear objective function is minimized subject to the inclusion into an SDP representable cone and hence this problem translates to an SDP. The constraint set of problem (3.28) is tighter than that of (3.25) in the sense that any pair of functions (v,w) feasible in (3.28) is also feasible in (3.25) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the degree k tends to infinity (see Theorem 3.2.5 below).

Continuous time

The equality constraint of (3.24)

$$\beta \int_{\mathbf{X}\times\mathbf{U}} v(x)\, d\mu(x,u) = \int_{\mathbf{X}} v(x)\, d\mu_0(x) + \int_{\mathbf{X}\times\mathbf{U}} \nabla v \cdot f(x,u)\, d\mu(x,u) \quad \forall\, v \in C^1(\mathbf{X}).$$

$$\int_{\mathbf{X}} w(x)\, d\mu_0(x) + \int_{\mathbf{X}} w(x)\, d\hat{\mu}_0(x) = \int_{\mathbf{X}} w(x)\, d\mu^{\mathrm{leb}}(x) \quad \forall\, w \in C(\mathbf{X})$$

is truncated by enforcing it only for the particular choice of test functions $v(x) = x^\alpha$ and $w(x) = x^\beta$ for all $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le k_v$ and $|\beta| \le k$, where

$$k_v := k - \deg f + 1.$$

The resulting finite-dimensional truncation of this linear system of equations is denoted by

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

where $\boldsymbol{y}$, $\boldsymbol{y}_0$ and $\hat{\boldsymbol{y}}_0$ represent the truncated moment sequences of the measures $(\mu, \mu_0, \hat{\mu}_0)$.

Combining this truncation with the results of Section 2.2.2, the primal relaxation of order k in discrete time reads

$$\begin{array}{rcll} p_k^* & = & \max\limits_{\boldsymbol{y},\boldsymbol{y}_0,\hat{\boldsymbol{y}}_0} & (\boldsymbol{y}_0)_0 \\ & & \text{s.t.} & A_k(\boldsymbol{y}, \boldsymbol{y}_0, \hat{\boldsymbol{y}}_0) = b_k \\ & & & \boldsymbol{y} \in M_k^{\mathrm{sup}}(\mathbf{X}\times\mathbf{U})_+ \\ & & & \boldsymbol{y}_0 \in M_k^{\mathrm{sup}}(\mathbf{X})_+, \\ & & & \hat{\boldsymbol{y}}_0 \in M_k^{\mathrm{sup}}(\mathbf{X})_+, \end{array} \tag{3.29}$$

where the truncated moment $M_k^{\mathrm{sup}}(\cdot)_+$ cone is defined in (2.13) and the objective function is the first element (i.e., the mass) of the truncated moment sequence $\boldsymbol{y}_0$ corresponding to the initial measure. In problem (3.27), a linear objective is minimized subject to SDP representable conic inclusions and therefore problem (3.27) translates to a semidefinite program (SDP). The constraint set of (3.29) is looser than that of (3.24) in the sense the the truncated moment sequences associated to any triple of measure $(\mu, \mu_0, \hat{\mu}_0)$ feasible in (3.24) are also feasible in (3.29) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the relaxation degree k tends to infinity (see Theorem 3.2.5).

Using the results of Section 2.2.1, we immediately arrive at a degree-k SDP tightening of (3.26), which is also the SDP dual to (3.29):

$$\begin{array}{rclll} d_k^* & = & \inf\limits_{v\in\mathbb{R}[x]_{k_v}, w\in\mathbb{R}[x]_k} & \boldsymbol{w}^\top \boldsymbol{l} & \\ & & \text{s.t.} & \beta v - \nabla v\cdot f & \in Q_k(\mathbf{X}\times\mathbf{U}) \\ & & & w - v - 1 & \in Q_k(\mathbf{X}) \\ & & & w & \in Q_k(\mathbf{X}), \end{array} \tag{3.30}$$

where $\boldsymbol{l}$ is the vector of Lebesgue moments over $\mathbf{X}$ indexed in the same basis in which the polynomial $w(x)$ with coefficients $\boldsymbol{w}$ is expressed and the quadratic module $Q_k(\cdot)$ is defined in (2.8). In problem (3.30), a linear objective function is minimized subject to the inclusion into an SDP representable cone and hence this problem translates to an SDP. The constraint set of problem (3.30) is tighter than that of (3.26) in the sense that any pair of functions (v, w) feasible in (3.30) is also feasible in (3.26) but the converse is not true in general. However, the discrepancy between the two constraint sets monotonically vanishes as the degree k tends to infinity (see Theorem 3.2.5 below).

3.2.4 Convergence results

In this section we state several convergence results for the finite dimensional relaxations resp. tightenings (3.27), (3.29) resp. (3.28), (3.30). Let w_k and v_k denote an optimal solution to the degree k dual SDP approximation (3.28) or (3.30), and define

$$\boldsymbol{X}_{\infty,k} := \{x \in X \ : \ v_k(x) \geq 0\}.$$

Then, in view of Lemmas 3.2.3 and 3.2.4, we know that w_k over-approximates the indicator function of the MCI set $\boldsymbol{X}_\infty$ on $\mathbf{X}$, i.e., $w_k \geq I_{\boldsymbol{X}_\infty}$ on $\mathbf{X}$, and that the sets $\boldsymbol{X}_{\infty,k}$ approximate from the outside the MCI set $\boldsymbol{X}_\infty$, i.e., $\boldsymbol{X}_{\infty,k} \supset \boldsymbol{X}_\infty$. In the sequel we prove the following:

- The optimal values of the finite-dimensional primal and dual problems p_k^* and d_k^* coincide and converge to the optimal values of the infinite dimensional primal and dual LPs p^* and d^* which also coincide (in view of Theorems 3.2.3 and 3.2.4) and are equal to the volume of the MCI set.

- The sequence of functions w_k converges on $\mathbf{X}$ from above to the indicator function of the MCI set in L_1 norm. In addition, the running minimum

$\min_{i\le k} w_i$ converges on $\mathbf{X}$ from above to the indicator function of the MCI set set in L_1 norm and almost uniformly.

- The sequence of sets $\boldsymbol{X}_{\infty,k}$ converges to the MCI set $\boldsymbol{X}_\infty$ in the sense that the volume discrepancy tends to zero, i.e., $\lim_{k\to\infty}\mu^{\mathrm{leb}}(\boldsymbol{X}_{\infty,k}\setminus\boldsymbol{X}_\infty)=0$.

The proofs of the results follow very similar reasoning as the results on region of attraction approximations in Section 3.1.

Lemma 3.2.5 *There is no duality gap between primal SDP problems (3.27 and 3.29) and dual SDP problems (3.28 and 3.30), i.e.* $p_k^* = d_k^*$.

Proof: The proof is similar to the proof of Theorem 3.2.5 and therefore we only outline the key points. To prove the absence of duality gap, it is sufficient to show that the feasible sets of the primal SDPs (3.27) and (3.29) are non-empty and compact. The result then follows by standard SDP duality theory (see the proof of Theorem 3.2.5 for a detailed argument). The non-emptiness follows trivially since the vector of measures $(\mu_0,\mu,\hat\mu)=(0,0,\mu^{\mathrm{leb}})$ is feasible in the primal infinite-dimensional LPs (3.23) and (3.24) and therefore the truncated moment sequences corresponding to these measures are feasible in the primal SDP relaxations (3.27) and (3.29). To see the compactness observe that the first components (i.e., masses) of the truncated moment vectors $\boldsymbol{y}_0$, $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}_0$ are bounded. This follows by evaluating the equality constraints of (3.23) and (3.24) for $w(x)=v(x)=1$. Indeed, in discrete-time we get $(y)_0=(y_0)_0/(1-\alpha)$ and in continuous-time we get $(y)_0=(y_0)_0/\beta$; in addition, in both cases we have $(y_0)_0+(\hat{\boldsymbol{y}}_0)_0=\mu^{\mathrm{leb}}(\mathbf{X})<\infty$ and therefore the first components are indeed bounded (since they are trivially bounded from below, in fact nonnegative, due to the constraints on moment matrices). Boundedness of the even components of each truncated moment vector then follows from the structure of the localizing matrices corresponding to the functions from Assumption 3.1.2. Boundedness of the entire truncated moment vectors then follows since the even moments appear on the diagonal of the positive semidefinite moment matrices. □

The following result shows the convergence of the optimal values of the finite-dimensional relaxations / approximations to the optimal values of the infinite-dimensional LPs.

Theorem 3.2.5 *The sequence of infima of SDP problems (3.28) and (3.30) converges monotonically from above to the supremum of the LP problems (3.25) and*

(3.26), i.e., $d^* \le d^*_{k+1} \le d^*_k$ *and* $\lim_{k\to\infty} d^*_k = d^* = p^*$*. Similarly, the sequence of maxima of SDP problems (3.27) and (3.29) converges monotonically from above to the maximum of the LP problems (3.23) and (3.24), i.e.,* $p^* \le p^*_{k+1} \le p^*_k$ *and* $\lim_{k\to\infty} p^*_k = p^* = d^*$*.*

Proof: The monotonicity of the optimal values of the relaxations p^*_k resp. approximations d^*_k is evident form the structure of the feasible sets of the corresponding SDPs. The convergence of the primal relaxations p_k to p^* follows from the compactness of the feasible sets of the primal SDPs (3.27) and (3.29) (shown in the proof of Lemma 3.2.5) by standard arguments on the convergence of Lasserre's SDP hierarchy (see, e.g., [Las09]). The converge of the optimal value of the dual approximations d^*_k to d^* then follows from Lemma 3.2.5. The equality between p^* and d^* is the statement of Theorems 3.2.3 and 3.2.4. □

The next theorem shows functional convergence from above to the indicator function of the MCI set.

Theorem 3.2.6 *Let* $w_k \in \mathbb{R}_k[x]$ *denote the* w*-component of a solution to the dual SDP problems (3.28) or (3.30) and let* $\bar{w}_k(x) = \min_{i\le k} w_i(x)$*. Then* w_k *converges from above to* $I_{\boldsymbol{X}_\infty}$ *in* L^1 *norm, i.e.,*

$$\lim_{k\to\infty} \int_{\mathbf{X}} w_k(x) - I_{\boldsymbol{X}_\infty}\, dx = 0.$$

In addition $\bar{w}_k$ *converges to* $I_{\boldsymbol{X}_\infty}$ *from above both in* L^1 *norm and almost uniformly*[14]*.*

Proof: The convergence in L_1 norm follows immediately from Theorem 3.2.5 and from the fact that $w_k \ge I_{\boldsymbol{X}_\infty}$ by Lemmas 3.2.3 and 3.2.4. The convergence of the running minima follows from the fact that there exists a subsequence of $\{w_k\}_{k=0}^\infty$ which converges almost uniformly (by, e.g., [Ash78, Theorems 2.5.2 and 2.5.3]). □

Our last theorem shows a set-wise convergence of the outer-approximations to the MCI set.

Theorem 3.2.7 *Let* $(v_k, w_k) \in \mathbb{R}_{k_v}[x] \times \mathbb{R}_k[x]$ *denote an optimal solution to the dual SDP problem (3.28) or (3.30) and let* $\boldsymbol{X}_{\infty,k} := \{x \in \mathbb{R}^n : v_k(x) \ge 0\}$*. Then*

[14] A sequence of functions w_k converges almost uniformly if $\forall\,\epsilon > 0, \exists\, B \subset \mathbf{X}$, $\mu^{\text{leb}}(B) < \epsilon$, such that $w_k \to w$ uniformly on $\mathbf{X}\backslash B$. Note that almost uniform convergence implies convergence almost everywhere [Ash78, Theorem 2.5.2]

$\boldsymbol{X}_\infty \subset \boldsymbol{X}_{\infty,k}$,

$$\lim_{k\to\infty} \mu^{\mathrm{leb}}(\boldsymbol{X}_{\infty,k} \setminus \boldsymbol{X}_\infty) = 0 \quad \textit{and} \quad \mu^{\mathrm{leb}}(\cap_{k=1}^\infty \boldsymbol{X}_{\infty,k} \setminus \boldsymbol{X}_\infty) = 0.$$

Proof: From Lemmas 3.2.3 or 3.2.4 we have $\boldsymbol{X}_{\infty,k} \supset \boldsymbol{X}_\infty$ and $w_k \geq I_{\boldsymbol{X}_\infty}$; therefore, since $w \geq v + 1$ and $w \geq 0$ on $\mathbf{X}$, we have $w_k \geq I_{\boldsymbol{X}_{\infty,k}} \geq I_{\boldsymbol{X}_\infty}$ and $\{x : w_k(x) \geq 1\} \supset \boldsymbol{X}_{\infty,k} \supset \boldsymbol{X}_\infty$. From Theorem 3.2.6, we have $w_k \to I_{\boldsymbol{X}_\infty}$ in L^1 norm on $\mathbf{X}$. Consequently,

$$\begin{aligned}\mu^{\mathrm{leb}}(\boldsymbol{X}_\infty) = \int_{\mathbf{X}} I_{\boldsymbol{X}_\infty}\, d\mu^{\mathrm{leb}} &= \lim_{k\to\infty} \int_{\mathbf{X}} w_k\, d\mu^{\mathrm{leb}} \geq \lim_{k\to\infty} \int_{\mathbf{X}} I_{\boldsymbol{X}_{\infty,k}}\, d\mu^{\mathrm{leb}} \\ &= \lim_{k\to\infty} \mu^{\mathrm{leb}}(\boldsymbol{X}_{\infty,k}) \geq \lim_{k\to\infty} \mu^{\mathrm{leb}}(\cap_{i=1}^k \boldsymbol{X}_{\infty,i}) = \mu^{\mathrm{leb}}(\cap_{k=1}^\infty \boldsymbol{X}_{\infty,k}).\end{aligned}$$

But since $\boldsymbol{X}_\infty \subset \boldsymbol{X}_{\infty,k}$ for all k, we must have

$$\lim_{k\to\infty} \mu^{\mathrm{leb}}(\boldsymbol{X}_{\infty,k}) = \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty) \quad \text{and} \quad \mu^{\mathrm{leb}}(\cap_{k=1}^\infty \boldsymbol{X}_{\infty,k}) = \mu^{\mathrm{leb}}(\boldsymbol{X}_\infty),$$

and the theorem follows. □

3.2.5 Numerical examples

In this section we present numerical examples that illustrate our results. The primal SDP relaxations were modeled using Gloptipoly 3 [HLL09] and the dual SOS problems using Yalmip [Löf04]. The resulting SDP problems were solved using SeDuMi [PTZ07] (which, in the case of primal relaxations, also returns the dual solution providing the outer approximations). For numerical computation (especially for higher relaxation orders), the problem data should be scaled such that the constraint sets are (within) unit boxes or unit balls; for ease of reproduction, most of the numerical problems shown are already scaled. On our problem class we observed only marginal sensitivity to the values of the discrete- and continuous-time discount factors α and β and report results with $\alpha = 0.9$ and $\beta = 1$ for all examples presented.

For a discussion on the scalability of our approach and the performance of alternative SDP solvers see the acrobot-on-a-cart example below and Chapter 6.

Discrete time

Double integrator Consider the discrete-time double integrator:

$$x_1^+ = x_1 + 0.1x_2$$
$$x_2^+ = x_2 + 0.05u$$

with the state constraint set $\mathbf{X} = [-1, 1]^2$ and input constraint set $\mathbf{U} = [-0.5, 0.5]$. The resulting of MCI set outer approximations of degree 8 and 12 are shown in Figure 3.8; the approximation is fairly tight for modest degrees. The true MCI set was computed using the standard algorithm based on polyhedral projections [Bla99].

Cathala system Consider the Cathala system borrowed from [KK11]:

$$x_1^+ = x_1 + x_2$$
$$x_2^+ = -0.5952 + x_2 + x_1^2.$$

The chaotic attractor of this system is contained in the set $\mathbf{X} = [-1.6, 1.6]^2$. MCI set outer approximations are shown in Figure 3.9; again, the approximations are relatively tight for small relaxation orders. The true MCI set was (approximately) computed by gridding.

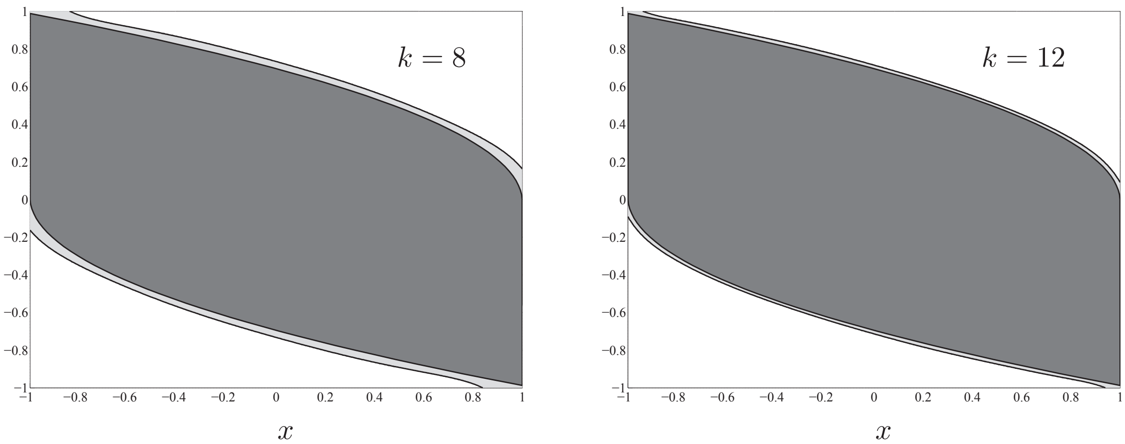


Figure 3.8 – Discrete time double integrator – polynomial outer approximations (light gray) to the MCI set (dark gray) for degrees $k \in \{8, 12\}$.

Julia sets Consider over $z \in \mathbb{C}$, or equivalently over $x \in \mathbb{R}^2$ with $z := x_1 + ix_2$, the quadratic recurrence

$$z^+ = z^2 + c$$

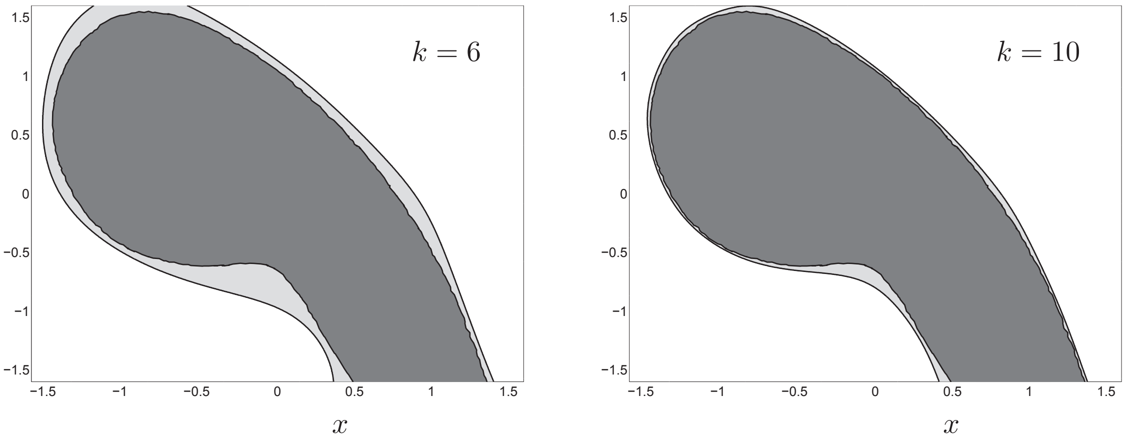
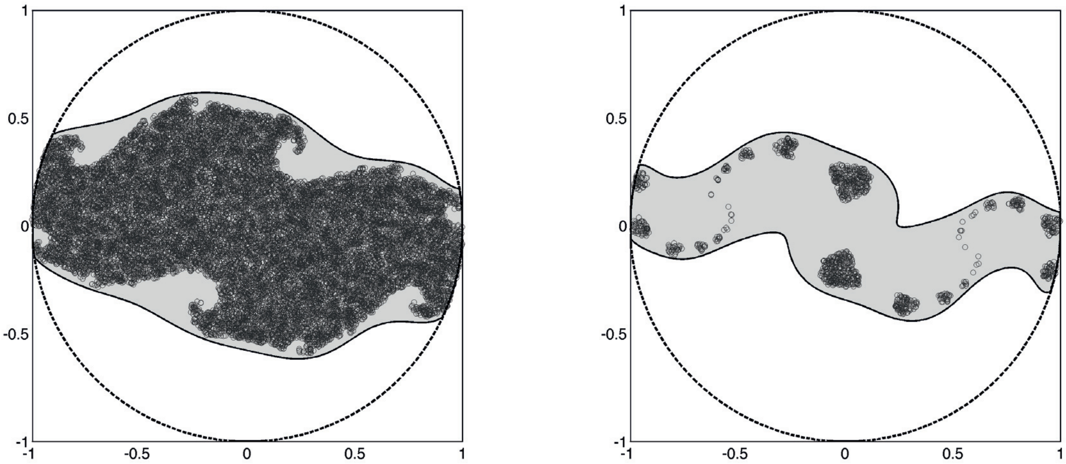


Figure 3.9 – Cathala system – polynomial outer approximations (light gray) to the MCI set (dark gray) for degrees $k \in \{6, 10\}$.

with $c \in \mathbb{C}$ a given complex number and i the imaginary unit. The filled Julia set is the set of all initial conditions of the above recurrence for which the trajectories remain bounded. The shape of the Julia set depends strongly on the parameter c. If c lies inside the Mandelbrot set, then the Julia set is connected; otherwise the set is disconnected. In both cases the boundary of the set has a very complicated (in fact fractal) structure. Here we shall compute outer approximations of the filled Julia set intersected with the unit ball. To this end we set $\mathbf{X} = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$. Figure 3.10 shows outer approximations of degree 12 for parameter values $c = -0.7 + i0.2$ (inside the Mandelbrot set) and $c = -0.9 + i0.2$ (outside the Mandelbrot set). The "true" filled Julia set was (approximately) obtained by randomly sampling initial conditions within the unit ball and iterating the recurrence for one hundred steps. Taking higher degree of the approximating polynomials does not give significant improvements due to our choice of the monomial basis to represent polynomials. An alternative basis (e.g. Chebyshev polynomials – see the related discussions in [HLS09] in Section 3.1) would allow us to improve further the outer estimates and better capture the intricate structure of the filled Julia set's boundary.

Hénon map Consider the modified controlled Hénon map

$$
\begin{aligned}
x_1^+ &= 0.44 - 0.1x_3 - 4x_2^2 + 0.25u,\\
x_2^+ &= x_1 - 4x_1x_2,\\
x_3^+ &= x_2,
\end{aligned}
$$



$c = -0.7 + i0.2$ $\quad$ $c = -0.9 + i0.2$

Figure 3.10 – Filled Julia set – polynomial outer approximation of degree 12 (light gray) and (an approximation of) the "true" set (dark grey) represented as an ensemble of initial conditions randomly sampled within the state-constraint set. The dashed line shows the boundary of the unit-ball state-constraint set.

adapted from [LZFQ12] with $\mathbf{X} = [-1, 1]^3$ and $\mathbf{U} = [-u_{\max}, u_{\max}]$. We investigate two cases: uncontrolled (i.e., $u_{\max} = 0$) and controlled with $u_{\max} = 1$. Figure 3.11 shows outer approximations to the MCI set of degree eight for both settings and the "true" MCI set in the uncontrolled setting (approximately) obtained by random sampling of initial conditions inside the constraint set $\mathbf{X}$. The outer approximations suggest that, as expected, allowing for control leads to a larger MCI set.

Continuous time

Double integrator Consider the continuous-time double integrator

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

with state constraint set $\mathbf{X} = [-1, 1]^2$ and input constraint set $\mathbf{U} = [-1, 1]$. The resulting MCI set outer approximations for degrees 8 and 12 are in Figure 3.12. The approximations are fairly tight even for relatively low relaxation orders. The true MCI set was (approximately) computed as in Section 3.2.5 by methods of [Bla99] after dense time discretization.

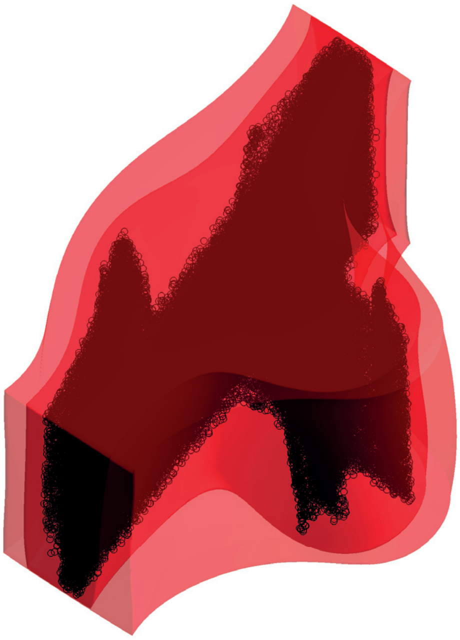


Figure 3.11 – Controlled Hénon map – polynomial outer approximation of degree eight in the uncontrolled setting (darker red, smaller) and in the controlled setting (lighter red, larger). The (approximation of) the "true" set (black) in the uncontrolled setting is represented as an ensemble of initial conditions randomly sampled within the state-constraint set.

Spider-web system As our second example we take the spider-web system from [Ahm08] given by equations

$$\dot{x}_1 = -0.15x_1^7 + 200x_1^6x_2 - 10.5x_1^5x_2^2 - 807x_1^4x_2^3 + 14x_1^3x_2^4 + 600x_1^2x_2^5 - 3.5x_1x_2^6 + 9x_2^7$$
$$\dot{x}_2 = -9x_1^7 - 3.5x_1^6x_2 - 600x_1^5x_2^2 + 14x_1^4x_2^3 + 807x_1^3x_2^4 - 10.5x_1^2x_2^5 - 200x_1x_2^6 - 0.15x_2^7$$

with the constraint set $\mathbf{X} = [-1, 1]^2$. Here we exploit the fact that the system dynamics are captured by constraints on v only whereas w is merely over approximating $v + 1$, and the fact that outer approximations to the MCI set are given not only by $\{x : v(x) \geq 0\}$ but also by $\{x : w(x) \geq 1\}$. Therefore, if low-complexity outer approximations are desired, it is reasonable to choose different degrees of v and w in (3.30) – high for v and lower for w – and use the set $\{x : w(x) \geq 1\}$ as the outer approximation. That way, we expect to obtain relatively tight low-order approximations. This is confirmed by numerical results shown in Figure 3.13. The degree of v is equal to 16 for both figures, whereas $\deg w = 8$ for the left figure and $\deg w = 16$ for the right figure. We observe no significant loss in tightness by choosing a smaller degree of w. The true MCI set was (approximately) computed

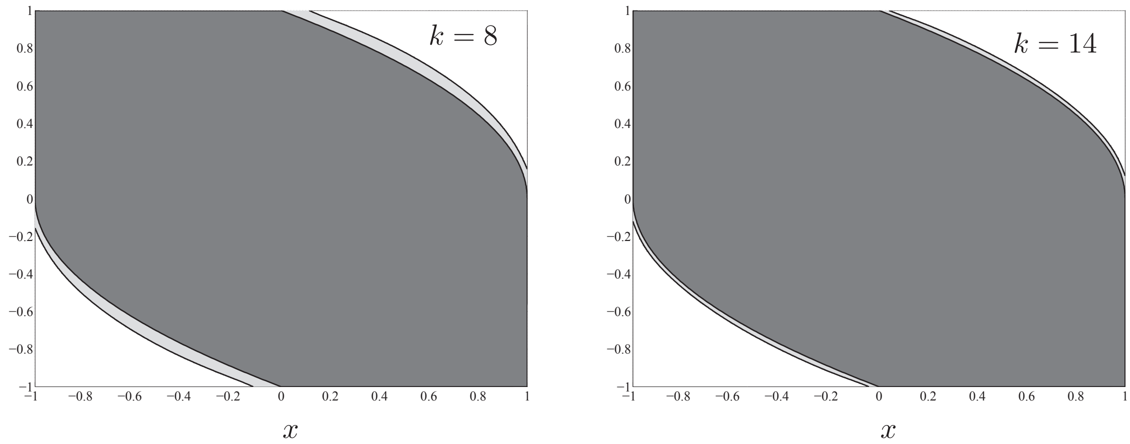


Figure 3.12 – Continuous-time double integrator – polynomial outer approximations (light gray) to the MCI set (dark gray) for degrees $k \in \{8, 14\}$.

by gridding.

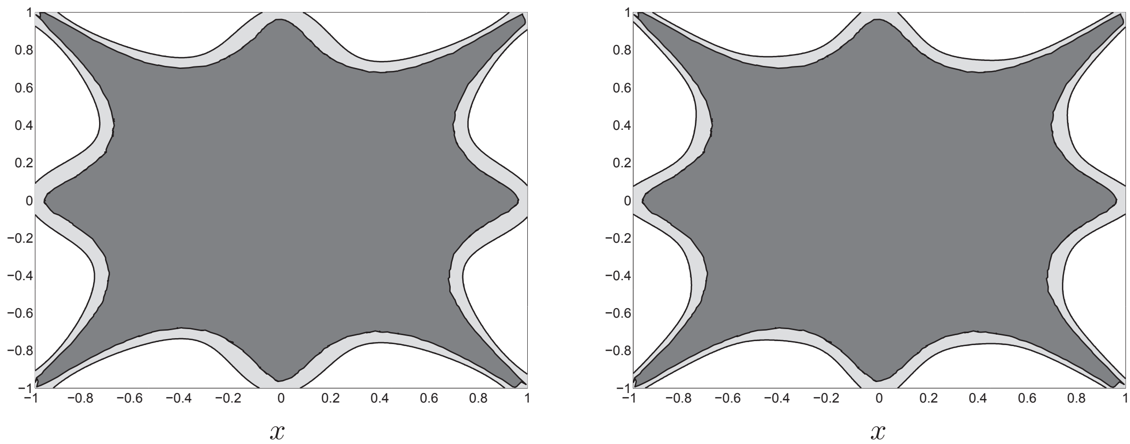


Figure 3.13 – Spider-web system – polynomial outer approximations (light gray) to the MCI set (dark gray) for degrees $\deg v = 16$ and $\deg w = 8$ on the left and $\deg w = 16$ on the right.

Acrobot on a cart As our last example we consider the acrobot on a cart system adapted from [JJ10], which is essentially a double pendulum on a cart where the inputs are the force acting on the cart and the torque in the middle joint of the double pendulum. The system is sketched in Figure 3.14. It is a sixth order system with with two control inputs; the dynamic equation is given by

$$\dot{x} = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ M(x)^{-1}N(x,u) \end{bmatrix} \in \mathbb{R}^6$$

where

$$M(x) = \begin{bmatrix} a_1 & a_2 \cos x_2 & a_3 \cos x_3 \\ a_2 \cos x_2 & a_4 & a_5 \cos(x_2 - x_3) \\ a_3 \cos x_3 & a_5 \cos(x_2 - x_3) & a_6 \end{bmatrix}$$

and

$$N(x,u) = \begin{bmatrix} u_1 + a_2 x_5^2 \sin x_2 + a_3 x_6^2 \sin x_3 - \delta_0 x_4 \\ -a_5 x_6^2 \sin(x_2 - x_3) + \delta_2 x_6 + a_7 \sin x_2 - x_5(\delta_1 + \delta_2) \\ u_2 + a_5 \sin(x_2 - x_3) x_5^2 + \delta_2 x_5 - \delta_2 x_6 + a_8 \sin x_3 \end{bmatrix}.$$

The states x_1, x_2, x_3 represent, respectively, the position of the cart (in meters), the angle of the lower rod and the angle of the upper rod of the double pendulum (both in radians); the states x_4, x_5 and x_6 are then the corresponding velocities in meters per second for the cart and radians per second for the pendulum rods. The constants are given by $a_1 = 0.85$, $a_2 = 0.2063$, $a_3 = 0.0688$, $a_4 = 0.0917$, $a_5 = 0.0344$, $a_6 = 0.0229$, $a_7 = 2.0233$, $a_8 = 0.6744$, $\delta_0 = 0.3$, $\delta_1 = 0.1$, $\delta_2 = 0.1$. We are interested in computing the maximum controlled invariant subset of the state constraint set

$$\mathbf{X} = [-1,1] \times [-\pi/3, \pi/3] \times [-\pi/3, \pi/3] \times [-0.5, 0.5] \times [-5,5] \times [-5,5].$$

We investigate two cases. First, we consider the situation where only the middle joint is actuated and there is no force on the cart; therefore we impose the constraint $(u_1, u_2) \in \mathbf{U} = \{0\} \times [-1,1]$. Second, we consider the situation where we can also exert a force on the cart; in this case we impose $(u_1, u_2) \in \mathbf{U} = [-1,1] \times [-1,1]$. Naturally, the MCI set for the second case is larger (or at least the same) as for the first case. This is confirmed[15] by outer approximations of degree four whose section for $x_1 = 0$, $x_4 = 0$, $x_5 = 0$ is shown in Figure 3.15. In order to compute the outer approximations we took a third order Taylor expansion of the non-polynomial dynamics even though exact treatment would be possible via a coordinate transformation leading to rational dynamics to which our methods can be readily extended; this extension is, however, not treated in this thesis and therefore we opted for the simpler (and non-exact) approach using Taylor expansion.

[15]There is no a priori guarantee on set-wise ordering of the outer approximations; what is guaranteed is the ordering of optimal values of the optimization problems (3.29) or (3.30).

Before solving the problem we made a linear coordinate transform so that the state constraint set becomes the unit hypercube $[-1, 1]^6$.

This example, which is the largest of those considered in this section, took 110 seconds to solve[16] with SeDuMi for $k = 4$; the corresponding time with the MOSEK SDP solver was 10 seconds. Using MOSEK we could also solve this example for $k = 6$ (in 420 seconds) although there the solver converged to a solution with a rather poor accuracy[17] and therefore we do not report the results.

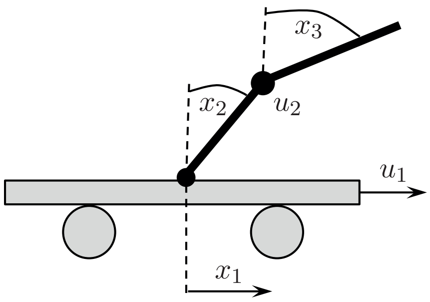


Figure 3.14 – Acrobot on a cart – sketch

[16]All examples were run on an Apple iMac with 3.4 GHz Intel Core i7, 8 GB RAM and Mac OS X 10.8.2. The time reported is the pure solver time, not including the Yalmip preprocessing time.

[17]Note that the MOSEK SDP solver is still being developed and its accuracy is likely to improve in the future.

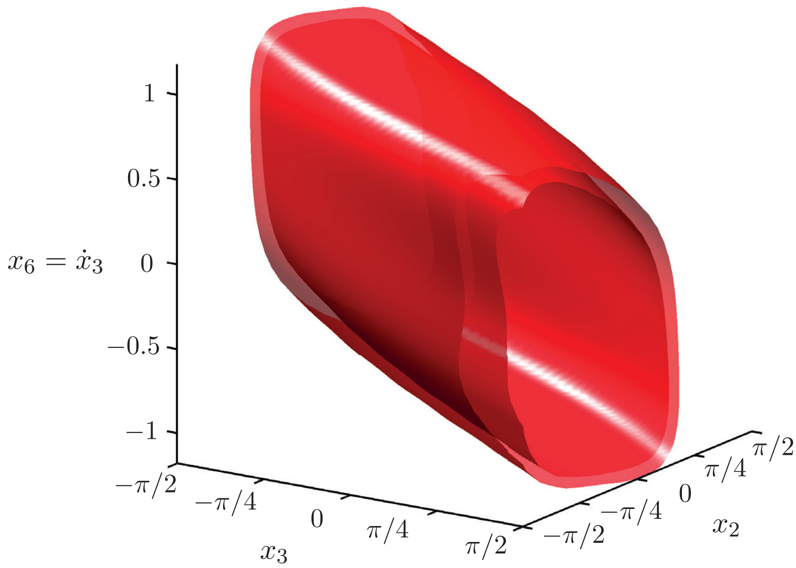


Figure 3.15 – Acrobot on a cart – section of the polynomial outer approximations of degree four for $(x_1, x_4, x_5) = (0, 0, 0)$. Only the middle joint actuated – darker, smaller; middle joint and the cart actuated – lighter, larger. The states displayed x_2, x_3 and x_6 are, respectively, the lower pendulum angle, the upper pendulum angle and the upper pendulum angular velocity.

Chapter 4

Optimal control

In this chapter we use the lift-plus-approximate scheme for infinite-horizon discounted optimal control problems. The main contribution is a hierarchy of SDP *tightenings* to the lifted problem whose solutions provide a sequence of rational *feedback controllers* which is proven to be asymptotically optimal, under certain technical conditions. In addition, we describe converging SDP hierarchies of approximations, from above and from below, to the value function attained in an optimal control problem by any given rational feedback controller as well as an SDP hierarchy providing lower bounds on the optimal value function. This value function approximations can serve as performance certificates of the designed controllers or of any given rational controller.

The novelty is in tightening the Liouville's equation (by optimizing over measures having polynomial densities) rather than relaxing it (as was done, e.g., in [LHPT08]), which facilitates theoretical analysis of the extracted controllers, but also brings about additional technical difficulties. In particular, the discounted Liouville's equation (2.28) is not suitable for this purpose since measures with polynomial densities (of arbitrarily large degrees) satisfying this equation may not exist. Hence, we resort to an alternative formulation of the optimal control problem by introducing stopping, which leads to the stopped discounted Liouville's equation (2.32) for which solutions with polynomial densities exist. Importantly, this alternative formulation is in a well-defined sense equivalent to the original one.

The results of this chapter can be seen as generalization of the stabilizing controller design procedure of [PPR04] to the optimal control setting (with convergence analysis) or as a generalization of the results of [Las11] from the setting of static polynomial optimization to the setting of optimal control.

The method presented here is also related to [RV14], where, however, the approximation step of the lift-plus-approximate procedure is carried out using spacial discretization rather than using moment-sum-of-squares hierarchies, no state constraints are considered and no convergence guarantees are given.

4.1 Problem statement

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

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} f_{u_i}(x(t))u_i(t), \tag{4.1}$$

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

$$x(t) \in \mathbf{X} := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0,\ i = 1, \dots, n_g\}, \tag{4.2a}$$

$$u(t) \in \mathbf{U} := [0, \bar{u}]^m, \tag{4.2b}$$

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

$$\bar{g}(x) := \prod_{i=1}^{n_g} g_i(x) > 0 \quad \forall x \in \mathbf{X}^\circ, \tag{4.3}$$

where $\mathbf{X}^\circ$ denotes the (topological) interior of $\mathbf{X}$.

Since $\mathbf{X}$ is assumed compact, we also assume, without loss of generality, that the inequalities defining the sets $\mathbf{X}$ contain the inequality $N - \|x\|^2 \geq 0$ for some $N \geq 0$, which implies the Archimedianity condition (see Definition 1).

[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 $\begin{bmatrix} \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} f(x,u) \\ \tilde{u} \end{bmatrix}$, where u is now a part of the state and $\tilde{u}$ a new control input; constraints on $\tilde{u}$ then correspond to rate constraints on u. Similarly, cost functions depending non-linearly on u in problem (4.4) can be handled using state inflation in exactly the same fashion.

[2] Any box can be affinely transformed to $[0, \bar{u}]^m$. For an axis-aligned box given by a vector of lower bounds $\mathrm{lb} \in \mathbb{R}^m$ and a vector of upper bounds $\mathrm{ub} \in \mathbb{R}^m$ one such transformation is $u = \bar{u}^{-1}\mathrm{diag}(\mathrm{ub} - \mathrm{lb})\tilde{u} + \mathrm{lb}$ with $\tilde{u} \in [0, \bar{u}]^m$. The general, non-axis-aligned, case follows by composing with appropriate rotation matrices.

The goal of this chapter is to (approximately) solve for all $x_0 \in \mathbf{X}$ the following optimal control problem (OCP):

$$\begin{array}{rcll} V(x_0) & := & \inf\limits_{u(\cdot),\tau} & \int_0^\tau e^{-\beta t}[l_x(x(t)) + \sum_{i=1}^m l_{u_i}(x(t))u_i(t)]\,dt + e^{-\beta\tau}M \\ & & \text{s.t.} & x(t) = x_0 + \int_0^t f(x(s)) + \sum_{i=1}^m f_{u_i}(x(s))u_i(s)\,ds, \\ & & & (x(t),u(t)) \in \mathbf{X}\times\mathbf{U} \;\; \forall t\in[0,\tau] \\ & & & u \in L([0,\tau];\mathbb{R}^m),\;\; \tau\in[0,\infty], \end{array} \tag{4.4}$$

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

$$M > \beta^{-1} \sup_{x\in\mathbf{X},u\in\mathbf{U}} \{l(x,u)\}, \tag{4.5}$$

where the joint stage cost

$$l(x,u) := l_x(x) + \sum_{i=1}^m l_{u_i}(x)u_i \tag{4.6}$$

is, without loss of generality, assumed to be nonnegative on $\mathbf{X}\times\mathbf{U}$. The state and input stage cost functions l_x and l_{u_i}, $i = 1,\dots,m$, are assumed to be polynomial. The function τ in OCP (4.4) is referred to as a *stopping time*; the optimization is therefore both over the control input u and over the final time τ, which can be finite or infinite and can depend on the initial condition x_0.

The function $x \mapsto V(x)$ in (4.4) is called the *value function*. The reason for choosing the slightly non-standard objective function in (4.4) is twofold. First, with this objective function the value function V is bounded (by M from above and by zero from below) on $\mathbf{X}$. Second, the value function V coincides with the value function associated to the more traditional[3] discounted infinite-horizon optimal control problem for all initial conditions $x_0 \in \mathbf{X}$ for which the trajectories can be kept within the state constraint set $\mathbf{X}$ forever using admissible controls, i.e., for all x_0 in the maximum control invariant set associated to the dynamics (4.1) and the constraints (4.2). To see the first claim, set $\tau = 0$ for all $x_0 \in \mathbf{X}$. To see the second claim notice that with M chosen as in (4.5), it is always beneficial to continue the time evolution whenever possible and therefore $\tau = +\infty$ for all x_0 in the maximum controlled invariant set associated to (4.1) and (4.2).

[3]By more traditional we mean a discounted optimal control problem with cost $\int_0^\infty e^{-\beta t} l(x(t),u(t))\,dt$ and no stopping.

The use of stopping times is very common in stochastic systems literature (see, e.g., [PS06]). Here, however, we use stopping times mostly for mathematical reasons[4].

Remark 4.1.1 *A constant M satisfying (4.5) can be found either by analytically evaluating the supremum in (4.5) or by using the SDP relaxation techniques of [Las01] to find an upper bound.*

Given a Lipschitz continuous *feedback* controller $u \in C(\mathbf{X};\mathbf{U})$ and a *stopping function* $\tau \in L(\mathbf{X};[0,\infty])$ such that the unique solution to (4.1) is contained in $\mathbf{X}$ for all $t \in [0,\tau(x_0)]$ and all $x_0 \in \mathbf{X}$, we let

$$V_{u,\tau} \in L(\mathbf{X};[0,\infty])$$

denote the value function attained by (u,τ) in (4.4), i.e., setting $u(t) = u(x(t))$ and $\tau = \tau(x_0)$. By V_u we denote the value function $V_{u,\tau_u^\star}$, where $\tau_u^\star \in L(\mathbf{X};[0,\infty])$ is the optimal stopping function associated to u. Note that, by the choice of M in (4.5), the optimal stopping function $\tau_u^\star$ is equal to the first hitting time of the complement of the constraint set $\mathbf{X}$, i.e.,

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

where $x(t\,|\,x_0)$ is the trajectory of (4.1) with $u(t) = u(x(t))$ starting from x_0. Notice also that $V_{u,\tau}(x) \geq V(x)$ for all $x \in \mathbf{X}$ and that for any pair (u,τ) feasible in (4.4) we have $V_{u,\tau}(x) \leq M$ for all $x \in \mathbf{X}$.

Throughout this chapter, we make the following technical assumption:

Assumption 4.1.1 *There exists a sequence of Lipschitz continuous feedback controllers $\{u^k \in C(\mathbf{X};\mathbf{U})\}_{k=1}^\infty$ and stopping functions $\{\tau^k \in L(\mathbf{X};[0,\infty])\}_{k=1}^\infty$ feasible in (4.4) such that*

$$\lim_{k\to\infty} \int_{\mathbf{X}} (V_{u^k,\tau^k}(x) - V(x))dx = 0 \tag{4.7}$$

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

$$\int_{\mathbf{X}} \int_0^{\tau^k(x_0)} e^{-\beta t} v(x^k(t\,|\,x_0))\,dt\,dx_0 = \int_{\mathbf{X}} v(x)\rho^k(x)\,dx \quad \forall v \in C(\mathbf{X}), \tag{4.8}$$

[4]One reason is the boundedness of the value function V. The second, less obvious, reason is the existence of solutions with polynomial densities to the stopped discounted Liouville's equation (2.32), which is not guaranteed without allowing for stopping (i.e., for the standard discounted Liouville's equation (2.28) without the terminal measure); see Remark 4.3.1 below.

where $x^k(\cdot\,|\,x_0)$ denotes the solution to (4.1) controlled by u^k.

Remark 4.1.2 *Note that $V_{u^k,\tau^k} \geq V$ on $\mathbf{X}$ by construction and therefore (4.7) is equivalent to the L^1 convergence of V_{u^k,τ^k} to V.*

Assumption 4.1.1 says that the optimal control inputs and stopping times for OCP (4.4) 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 $\mathbf{X}$. 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 (4.8) are continuously differentiable and vanish near the boundary of $\mathbf{X}$ (note, however, that the left hand side of (4.8) can always be represented as $\int_{\mathbf{X}} v(x)d\mu^k(x)$ for some measure $\mu^k \in M(\mathbf{X})_+$). This problem is of rather technical nature and has been studied in the literature (e.g., [Cri07, Section 1.4], [Ran02] or [RH03]), where affirmative results have been established in related settings[5]. We do not undertake a study of this problem here and rely on Assumption 4.1.1, 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 (4.8) can be weighted by a nonnegative function $\rho_0^k \in L_1(\mathbf{X})$ satisfying $\rho_0^k \geq 1$ on $\mathbf{X}$ and $\rho_0^k \to 1$ in $L_1(\mathbf{X})$. In addition, we conjecture that it is enough to require $\rho^k = 0$ on $\partial\mathbf{X}$ and not necessarily on some neighborhood of $\partial\mathbf{X}$; this is in particular the case when $\mathbf{X}$ is a box or a ball but we expect all the results of the chapter to hold with a general semialgebraic set for which the defining functions satisfy (4.3). We also remark that the requirement of Assumption 4.1.1 that an asymptotically optimal sequence of Lipschitz controllers exists can be removed if the problem (4.4) is posed directly in a closed-loop form, i.e., by requiring that the control input in problem (4.4) is generated by a Lipschitz continuous feedback controller. The closed-loop form is slightly less standard and hence here we adhere to the open-loop formulation (4.4) and rely on Assumption 4.1.1.

The main result of this chapter is a hierarchy of semidefinite programming *tighten-*

[5]In particular [Cri07, Section 1.4] establishes the relationship between the regularity of the densities transported along the flow of a nonlinear vector field and the regularity of the vector field itself; [Ran02] establishes the existence of densities certifying almost global stability, whereas [RH03] establishes the existence of asymptotically optimal densities in a constrained optimal control setting where the state constrained set is assumed to be control invariant and to have a C^1 boundary.

ings of problem (4.4), which provides a sequence of *rational* controllers u^k such that, under Assumption 4.1.1, (4.7) 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 4.1.2).

4.2 Lifting

First we lift the problem (4.4) to an infinite-dimensional LP in the space of measures. As in Chapter 3, the lifting consists of replacing the nonlinear dynamics with a particular form of Liouville's equation. In our case, since we are working on infinite horizon with stopping and the system dynamics is input affine with box constraints, the best suited form is Eq. (2.40).

In order to lift the problem for *all* initial conditions $x_0 \in \mathbf{X}$ we impose the constraint that $\mu_0 \geq \mu^{\text{leb}}$, where μ^{leb} is the Lebesgue measure on $\mathbf{X}$. This will enable us to obtain a *feedback* controller from the subsequent SDP tightenings of the lifted problem rather than an open-loop control trajectory (see Section 2.3.5 and the discussion below Eq. (4.10)).

The infinite-dimensional LP reads

$$\begin{array}{ll} \inf\limits_{\mu,\,\mu_0,\,\mu_T,\,\nu} & \int_{\mathbf{X}} l_x(x)\,d\mu(x) + \sum_{i=1}^m \int_{\mathbf{X}} l_{u_i}(x)\,d\nu_i(x) + M\int_{\mathbf{X}} 1\,d\mu_T(x) \\ \text{s.t.} & \int_{\mathbf{X}} v\,d\mu_T = \int_{\mathbf{X}} v\,d\mu_0 + \int_{\mathbf{X}}(\nabla v\cdot f - \beta v)\,d\mu + \sum_{i=1}^m \int_{\mathbf{X}} \nabla v\cdot f_{u_i}\,d\nu_i \quad \forall v \in C^1 \\ & \mu \in M(\mathbf{X})_+,\ \mu_0 \in M(\mathbf{X})_+,\ \mu_T \in M(\mathbf{X}_+),\ \nu \in M(\mathbf{X})_+^m \\ & \mu_0 \geq \mu^{\text{leb}} \\ & \bar{u}\mu \geq \nu_i, \quad i \in \{1,\ldots,m\}, \end{array} \tag{4.9}$$

where the first constraint is precisely the stopped discounted Liouville's equation in the input-affine form (2.40), which we repeat here for convenience in full:

$$\int_{\mathbf{X}} v\,d\mu_T = \int_{\mathbf{X}} v\,d\mu_0 + \int_{\mathbf{X}}(\nabla v\cdot f - \beta v)\,d\mu + \sum_{i=1}^m \int_{\mathbf{X}} \nabla v\cdot f_{u_i}\,d\nu_i \quad \forall\, v \in C^1(\mathbf{X}). \tag{4.10}$$

The constraints $\bar{u}\mu \geq \nu_i$ ensure the satisfaction of input constraints (see Section 2.3.5). In particular $\bar{u}\mu \geq \nu_i$ implies that each control measure ν_i is absolutely continuous with respect to μ with density (i.e., Radon-Nikodým derivative) bounded by $\bar{u}$, i.e., there exists a $u_i \in L_\infty(\mathbf{X};[0,\bar{u}])$ such that $d\nu_i(x) = u_i(x)d\mu(x)$ for all $i \in \{1,\ldots,m\}$. The function $u(x) = (u_1(x),\ldots,u_m(x))$ can be viewed as a *feedback* *controller*; indeed, substituting $d\nu_i(x) = u_i(x)d\mu(x)$ in (4.10) yields

$$\int_{\mathbf{X}} v\, d\mu_T = \int_{\mathbf{X}} v\, d\mu_0 + \int_{\mathbf{X}} (\nabla v \cdot \bar{f}(x) - \beta v)\, d\mu \quad \forall\, v \in C^1(\mathbf{X}), \tag{4.11}$$

where the $\bar{f} = f(x) + \sum_{i=1}^{m} f_{u_i} u_i(x)$ denotes the vector field of the ODE (4.1) when the loop is closed with the feedback controller $u(x)$. The equation (4.11) is nothing but Eq. (2.33), i.e., a stopped discounted Liouville's equation for the closed-loop vector field $\bar{f}$. Note that, at this point, the controller $u(x)$ is not guaranteed to be Lipschitz and so is not the closed-loop vector field $\bar{f}$; later on, when considering polynomial tightenings of (4.9), this controller will be Lipschitz and so will be the closed-loop vector field $\bar{f}$.

Remark 4.2.1 (Non-uniform weighting) *Note that instead of $\mu_0 \geq \mu^{\mathrm{leb}}$, we could have imposed that μ_0 is greater than any measure having a polynomial density $\bar{\rho}_0$ with respect to the Lebesgue measure μ^{leb}. This has no impact on all theoretical results established in the rest of the chapter as long as $\bar{\rho}_0$ is strictly positive on $\mathbf{X}$. In particular, asymptotic convergence of the value functions associated to the designed controllers is preserved. The choice of non-uniform $\bar{\rho}_0$ may, however, influence the speed of convergence in different subsets of $\mathbf{X}$. 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 one to assign a different importance to different subsets of $\mathbf{X}$.*

4.3 Tightening

Now we use the ideas of Section 2.2.3 to tighten, rather than relax (as in Chapter 3 or in [LHPT08]), the lifted problem (4.9). The reason for using tightenings rather than relaxations is the fact that, contrary to relaxations, feasible solutions to the tightenings are also feasible in the original problem, which makes them more suited for controller design with strong theoretical guarantees (as in Theorem 4.3.1 below).

The idea is to restrict the measures in (4.9), to measures which possess a density with respect to the Lebesgue measure. We do that in two steps: first we restrict the densities to measures with continuous (resp. C^1) densities and then we restrict further to polynomial densities belonging to suitable SDP representable quadratic modules, as in Section 2.2.3.

4.3.1 Tightening with continuous densities

The first tightening reads:

$$\begin{array}{lll}\inf\limits_{\rho,\,\rho_0,\,\rho_T,\,\sigma} & \int_{\mathbf{X}} l_x(x)\rho(x)\,dx + \sum_{i=1}^m \int_{\mathbf{X}} l_{u_i}(x)\sigma_i(x)\,dx + M\int_{\mathbf{X}} \rho_T(x)\,dx & \\ \text{s.t.} & \rho_T - \rho_0 + \beta\rho + \operatorname{div}(\rho f) + \sum_{i=1}^m \operatorname{div}(\sigma_i f_{u_i}) = 0 & \\ & \rho \le 0 & \text{on } \partial\mathbf{X} \\ & \rho_0 \ge 1 & \text{on } \mathbf{X} \\ & \bar{u}\rho \ge \sigma_i & \text{on } \mathbf{X},\ \ i = 1,\ldots,m. \\ & \rho_T \ge 0 & \text{on } \mathbf{X} \\ & \sigma_i \ge 0 & \text{on } \mathbf{X},\ \ i = 1,\ldots,m. \end{array} \tag{4.12}$$

The optimization in (4.12) is over functions $(\rho, \rho_0, \rho_T, \sigma) \in C^1(\mathbf{X}) \times C(\mathbf{X}) \times C(\mathbf{X}) \times C^1(\mathbf{X})^m$ with $\sigma = (\sigma_1, \ldots, \sigma_m)$. The optimal value of (4.12) will be denoted by $p^\star$ and the value attained in (4.12) by any tuple of densities $(\rho, \rho_0, \rho_T, \sigma)$ feasible in (4.12) will be denoted by $p(\rho, \rho_0, \rho_T, \sigma)$.

The equality constraint of (4.12)

$$\rho_T - \rho_0 + \beta\rho + \operatorname{div}(f\rho) + \sum_{i=1}^m \operatorname{div}(f_{u_i}\sigma_i) = 0 \tag{4.13}$$

is termed the functional stopped discounted Liouville's equation. Satisfaction of (4.13) by

$$(\rho, \rho_0, \rho_T, \sigma) \in C^1(\mathbf{X}) \times C(\mathbf{X}) \times C(\mathbf{X}) \times C^1(\mathbf{X})^m$$

implies the satisfaction of the (measure) Liouville's equation (4.10) by the measures

$$(d\mu, d\mu_0, d\mu_T, d\nu) = (\rho dx, \rho_0 dx, \rho_T dx, \sigma dx) \in M(\mathbf{X})_+ \times M(\mathbf{X})_+ \times M(\mathbf{X})_+ \times M(\mathbf{X})_+^m \tag{4.14}$$

provided that $\rho = 0$ on $\partial\mathbf{X}$. This follows readily from integration by parts applied on the divergence terms in (4.13), where the constraint $\rho = 0$ on $\partial\mathbf{X}$ ensures that the boundary terms vanish. This is the reason for adding the constraint $\rho \le 0$ on $\partial\mathbf{X}$ to (4.12), which in conjunction with $\bar{u}\rho \ge \sigma_i \ge 0$, implies that indeed $\rho = 0$ on $\partial\mathbf{X}$. Noticing also that the constraint $\bar{u}\rho \ge \sigma$ and $\rho_0 \ge 1$ imply that $\bar{u}\mu \ge \nu$ and $\mu_0 \ge \mu^{\mathrm{leb}}$, we conclude that the problem (4.12) is indeed a tightening of (4.9) in the sense that any feasible solution to (4.12) gives rise to a feasible solution to (4.9) via (4.14).

4.3.2 Tightening with polynomial densities

Using the results of Sections 2.2.1 and Section 2.2.3, the infinite-dimensional LP (4.12) is then further tightened by optimizing over polynomials instead of continuous functions and by replacing nonnegativity constraints by inclusions in appropriate SDP representable quadratic modules.

Polynomial approximation of degree d of (4.12) reads

$$\begin{array}{lll}\inf\limits_{(\rho,\rho_0,\rho_T,\sigma)\in\mathbb{R}[x]_d^{3+m}} & \int_{\mathbf{X}} l_x(x)\rho(x)\,dx + \sum_{i=1}^m \int_{\mathbf{X}} l_{u_i}(x)\sigma_i(x)\,dx + M\int_{\mathbf{X}} \rho_T(x)\,dx & \\ \text{s.t.} & \rho_T - \rho_0 + \beta\rho + \mathrm{div}(\rho f) + \sum_{i=1}^m \mathrm{div}(\sigma_i f_{u_i}) = 0 & \\ & -\rho \in Q_d(\mathbf{X}) + g_i\mathbb{R}[x]_{d-\deg g_i} + \bar{g}\mathbb{R}[x]_{d-\deg \bar{g}} & i = 1,\ldots,n_g \\ & \rho_0 - 1 \in Q_d(\mathbf{X}) & \\ & \bar{u}\rho - \sigma_i \in Q_d(\mathbf{X}) + \bar{g}Q_{d-\deg\bar{g}}(\mathbf{X}) & i = 1,\ldots,m \\ & \rho_T \in Q_d(\mathbf{X}) & \\ & \sigma_i \in Q_d(\mathbf{X}) + \bar{g}Q_{d-\deg\bar{g}}(\mathbf{X}), & i = 1,\ldots,m,\end{array} \tag{4.15}$$

where $\bar{g}$ is the polynomial defined in (4.3). 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 $(\rho, \rho_0, \sigma, \rho_T)$, 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 2.2.1. Optimization problem (4.15) therefore immediately translates to an SDP.

Remark 4.3.1 (Feasibility & role of terminal measure) *Trivially, any feasible solution to (4.15) is feasible in (4.12). Also, problem (4.15) is feasible for any* $d \geq 0$ *as* $(\rho, \rho_0, \rho_T, \sigma) = (0, 1, 1, 0)$ *is always feasible in (4.15). This is crucial from a practical point of view and is not satisfied with other, more obvious, formulations (e.g., those not involving a stopping time in (4.4)); the reason for this is that, in the absence of a terminal measure (i.e.,* $\rho_T = 0$*), the functional stopped discounted Liouville's equation (4.13) 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* $\mathbf{X} = [-1, 1]$ *and zero elsewhere, the only solution to (4.13) with* $\rho_T = 0$ *is* $\rho(x) = -\ln(|x|)$*.*

Remark 4.3.2 (Why $\bar{g}$) *Note that the satisfaction of the constraints of (4.15) imply the satisfaction of the constraints of (4.12) even without the terms involving* $\bar{g}$ *in the second, fourth and the sixth constraint. The terms involving* $\bar{g}$ *are included in order to increase the algebraic strength of the nonnegativity certificates which* *enables us to prove our main result, Theorem 4.3.1, with the help of Lemma 4.4.3. Mere application of the Putinar's positivstellensatz (Proposition 2.2.1) and its immediate corrolary (Corrolary 2.2.1) is not sufficient since here we approximate with nonnegative polynomials vanishing on the boundary of $\partial\mathbf{X}$ (note that the constraints of (4.12) imply that $\rho \geq 0$ on $\mathbf{X}$ and $\rho = 0$ on $\partial\mathbf{X}$).*

If non-uniform weighting of initial conditions (see Remark 4.2.1) was required, the constraint $\rho_0 - 1 \in Q_d(\mathbf{X})$ would be replaced by $\rho_0 - \bar{\rho}_0 \in Q_d(\mathbf{X})$ for a polynomial weighting function $\bar{\rho}_0$ nonnegative on $\mathbf{X}$.

Rational feedback controller

Given an optimal solution $(\rho^d, \rho_0^d, \rho_T^d, \sigma^d)$ to (4.15), we define a *rational* feedback controller u^d by

$$u_i^d(x) := \frac{\sigma_i^d(x)}{\rho^d(x)} \;\; \forall x \in \mathbf{X},\; i = 1, \ldots, m, \tag{4.16}$$

where the control input is defined by the limit when $\rho^d(x) = 0$ (since $0 \leq \sigma_i^d(x) \leq \bar{u}\rho(x)$, this limit always exists and that the resulting controller is continuous on $\mathbf{X}$ and satisfies $u_i(x) \in [0, \bar{u}]$ for all $x \in \mathbf{X}$). The controller u^d is nothing but the density of the measure $d\nu^d(x) = \sigma^d(x)dx$ with respect to the measure $d\mu^d(x) = \rho^d(x)dx$ (see the discussion around Eq. (4.11)).

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

Theorem 4.3.1 *For all $d \geq 0$ we have $u^d(x) \in \mathbf{U}$ for all $x \in \mathbf{X}$ and if Assumption 4.1.1 holds, then*

$$\lim_{d\to\infty} \int_{\mathbf{X}} (V_{u^d}(x) - V(x))\, dx = 0, \tag{4.17}$$

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

4.4 Proof of Theorem 4.3.1

In this section we prove the main result of this chapter, Theorem 4.3.1. A crucial ingredient to the proof is Theorem 2.3.3, which immediately leads to the following specialization for our setting:

Theorem 4.4.1 (Superposition) *If measures μ, μ_0, μ_T and ν_i, $i = 1, \ldots, m$, satisfy (4.10) with* $\operatorname{spt}\mu_0 \subset \mathbf{X}$, $\operatorname{spt}\mu \subset \mathbf{X}$ *and* $\operatorname{spt}\mu_T \subset \mathbf{X}$ *and $d\nu_i = u_i d\mu$ for some Lipschitz $u \in C(\mathbf{X}; \mathbf{U})$, then there exists an ensemble of probability measures (i.e., measures with unit mass) $\{\tau_{x_0}\}_{x_0 \in \mathbf{X}}$ such that*

$$\int_{\mathbf{X}} v(x)\, d\mu_0(x) = \int_{\mathbf{X}} v(x(0\,|\,x_0))\, d\mu_0(x_0), \tag{4.18a}$$

$$\int_{\mathbf{X}} v(x)\, d\mu(x) = \int_{\mathbf{X}} \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t\,|\,x_0))\, dt\, d\tau_{x_0}(\tau)\, d\mu_0(x_0), \tag{4.18b}$$

$$\int_{\mathbf{X}} v(x)\, d\mu_T(x) = \int_{\mathbf{X}} \int_0^\infty e^{-\beta\tau} v(\tau(x_0))\, d\tau_{x_0}(\tau)\, d\mu_0(x_0), \tag{4.18c}$$

$$\int_{\mathbf{X}} v(x)\, d\nu_i(x) = \int_{\mathbf{X}} \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t\,|\,x_0)) u_i(x(t\,|\,x_0))\, dt\, d\tau_{x_0}(\tau)\, d\mu_0(x_0) \tag{4.18d}$$

for all bounded measurable functions $v : \mathbb{R}^n \to \mathbb{R}$, where $x(\cdot \mid x_0)$ denotes the unique trajectory of system (4.1) controlled with the Lipschitz controller $u(t) = u(x(t))$ which satisfies $x(t\,|\,x_0) \in \mathbf{X}$ for all $t \in$ $\operatorname{spt}\tau_{x_0}$.

Proof: Theorem 4.4.1 follows from Theorem 2.3.3 by setting $\bar{f} = f + \sum_{i=1}^m f_{u_i} u_i$. The conclusion that $x(t\,|\,x_0) \in \mathbf{X}$ for all $t \in \operatorname{spt}\tau_{x_0}$ follows by taking $v = I_{\mathbb{R}^n \setminus \mathbf{X}}$ in (4.18b). □

Theorem 4.4.1 immediately enables us to prove a representation of the cost of problem (4.12) in terms of trajectories of (4.1).

Lemma 4.4.1 *If $(\rho, \rho_0, \rho_T, \sigma)$ is feasible in (4.12) and $u = \sigma/\rho$, then*

$$\begin{aligned} p(\rho, \rho_0, \rho_T, \sigma) = &\int_{\mathbf{X}} \int_0^\infty \int_0^\tau e^{-\beta t} l_x(x(t\,|\,x_0)) dt\, d\tau_{x_0}(\tau) \rho_0(x_0)\, dx_0 \\ &+ \sum_{i=1}^m \int_{\mathbf{X}} \int_0^\infty \int_0^\tau e^{-\beta t} l_{u_i}(x(t\,|\,x_0)) u_i(x(t\,|\,x_0))\, dt\, d\tau_{x_0}(\tau) \rho_0(x_0) dx_0 \\ &+ M \int_{\mathbf{X}} \int_0^\infty e^{-\beta\tau}\, d\tau_{x_0}(\tau) \rho_0(x_0) dx_0, \end{aligned} \tag{4.19}$$

where $x(\cdot \mid x_0)$ are trajectories of (4.1) 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 (4.4) in the sense that $x(t\,|\,x_0) \in \mathbf{X}$ and $u(t\,|\,x_0) \in \mathbf{U}$ for all $t \in$* $\operatorname{spt}\tau_{x_0}$.

Proof: Let $(\rho, \rho_0, \rho_T, \sigma)$ be feasible in (4.12) and let $p(\rho, \rho_0, \rho_T, \sigma)$ denote the value attained by $(\rho, \rho_0, \rho_T, \sigma)$ in (4.12). The equality constraint of (4.12) is exactly (4.13). Since the constraint of (4.12) implies $\rho = 0$ on $\partial\mathbf{X}$, equation (4.10) 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 $u_i = \frac{\sigma_i}{\rho} \in C(\mathbf{X}; \mathbf{U})$, $i = 1, \ldots, m$. By Theorem 4.4.1 (setting $v(x) = l_x(x)$ in (4.18b), $v(x) = 1$ in (4.18c) and $v(x) = l_{u_i}(x)$ in (4.18d)) we obtain the result (noticing that the constraints of (4.12) imply that $u(x) \in \mathbf{U}$ for all $x \in \mathbf{X}$). □

Corollary 4.4.1 *If $(\rho, \rho_0, \rho_T, \sigma)$ is feasible in (4.12) and $u = \sigma/\rho$, then*

$$p(\rho, \rho_0, \rho_T, \sigma) \geq \int_{\mathbf{X}} V_u(x_0)\rho_0(x_0)dx_0. \tag{4.20}$$

If in addition the stopping measures $\{\tau_{x_0}\}_{x_0 \in \mathbf{X}}$ in (4.19) are equal to the Dirac measures $\{\delta_{\tau(x_0)}\}_{x_0 \in \mathbf{X}}$ for some stopping function $\tau \in L(\mathbf{X}; [0, \infty])$, then

$$p(\rho, \rho_0, \rho_T, \sigma) = \int_{\mathbf{X}} V_{u,\tau}(x_0)\rho_0(x_0)dx_0. \tag{4.21}$$

Proof: Let $(\rho, \rho_0, \rho_T, \sigma)$ be feasible in (4.12). Using Lemma 4.4.1, $p(\rho, \rho_0, \rho_T, \sigma)$ has representation (4.19), where the state-control trajectories in (4.19) are feasible in (4.4). Since the measures τ_{x_0} in (4.19) have unit mass for all $x_0 \in \mathbf{X}$, we obtain (4.20). If $\tau_{x_0} = \delta_{\tau(x_0)}$ for some stopping function $\tau \in L(\mathbf{X}; [0, \infty])$, then the integrals with respect to τ_{x_0} in (4.19) become evaluations at $\tau(x_0)$ and hence (4.21) holds. □

Corollary 4.4.1 immediately implies that the problem (4.12) (and hence problem (4.15)) is a tightening of the original problem (4.4):

Theorem 4.4.2 *The optimal value of (4.12) of $p^\star$ satisfies*

$$p^* \geq \int_{\mathbf{X}} V(x)\, dx. \tag{4.22}$$

Proof: Follows from Corollary 4.4.1 since $\rho_0 \geq 1$ and $V_u \geq V \geq 0$. □

Now we are in a position to prove the following crucial lemma linking problems (4.4) and (4.12).

Lemma 4.4.2 *If $\{u^k \in C(\mathbf{X}; \mathbf{U})\}_{k=1}^{\infty}$ and $\{\tau^k \in L(\mathbf{X}; [0, \infty])\}_{k=1}^{\infty}$ are respectively* *sequences of controllers and stopping functions satisfying the conditions of Assumption 4.1.1, 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 (4.12) and satisfy*

$$\lim_{k\to\infty} p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) = \int_{\mathbf{X}} V(x_0)dx_0. \tag{4.23}$$

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 (4.1.1) holds, then equation (4.7) 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 (4.1.1). Setting $\rho_0^k = 1$ and defining $\sigma_i^k := u_i^k\rho^k$ and $\rho_T^k := \rho_0^k - \beta\rho^k - \mathrm{div}(\rho^k f) - \sum_{i=1}^m \mathrm{div}(f_{u_i}\sigma_i^k)$ we see that $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ satisfy (4.13) with $\rho^k = 0$ on $\partial\mathbf{X}$. Therefore $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ are feasible in (4.12). In addition, in view of (4.8), the representation (4.19) holds with $\tau_{x_0} = \delta_{\tau^k(x_0)}$. Therefore by Lemma 4.4.1

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

and hence (4.23) holds since $\{V_{u^k,\tau^k}\}_{k=1}^\infty$ satisfies (4.7) and $\rho_0^k = 1$ for all $k \ge 0$.

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 (4.23) by Theorem 4.4.2 and by the first part of Lemma 4.4.2 just proven. Therefore (4.7) holds with $u^k := \sigma^k/\rho^k$ since

$$p(\rho^k, \rho_0^k, \rho_T^k, \sigma^k) \ge \int_{\mathbf{X}} V_{u^k}(x_0)dx_0$$

by Corollary 4.4.1. □

We will also need the following result showing that nonnegative C^1 functions vanishing on a neighborhood of $\partial\mathbf{X}$ can be approximated by polynomials in $\bar{g}Q_d(\mathbf{X})$ (which necessarily vanish on $\partial\mathbf{X}$).

Lemma 4.4.3 *Let* $\rho \in C^1(\mathbf{X})$ *such that* $\rho \ge 0$ *on* $\mathbf{X}$ *and* $\rho = 0$ *on* $\{x \in \mathbf{X} : \mathrm{dist}_{\partial\mathbf{X}}(x) < \zeta\}$ *for some* $\zeta > 0$. *Then for any* $\epsilon > 0$ *there exists* $d \ge 0$ *and a polynomial* $p_d \in \bar{g}Q_{d-\deg\bar{g}}(\mathbf{X})$ *such that*

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

and $p_d = 0$ *on* $\partial\mathbf{X}$.

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

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

Since polynomials are dense in C^1 (e.g., [BBL02]) there exists for every $\delta > 0$ a polynomial $\hat{h} > 0$ such that

$$\|\hat{h} - h\|_{C^1(\mathbf{X})} < \delta. \tag{4.24}$$

Applying Theorem 2.2.1 to $\hat{h}$ we see that there exists $\hat{p}_{\hat{d}} \in Q_{\hat{d}}(\mathbf{X})$ for some $\hat{d} \geq 0$ such that

$$\|\hat{h} - \hat{p}_{\hat{d}}\|_{C^1(\mathbf{X})} < \delta. \tag{4.25}$$

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

Now we are ready to prove our main result, Theorem 4.3.1.

Proof (of Theorem 4.3.1): Consider the sequences $\{u^k \in C(\mathbf{X};\mathbf{U})\}_{k=1}^\infty$, $\{\tau^k \in L(\mathbf{X};[0,\infty])\}_{k=1}^\infty$ from Assumption 4.1.1. By the first part of Lemma 4.4.2 the sequence of associated densities $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ generated by (u^k, τ^k) is feasible in (4.12) and satisfies (4.23). By Assumption 4.1.1, $\rho^k = 0$ and $\sigma^k = 0$ on $\{x \in \mathbf{X} : \operatorname{dist}_{\partial\mathbf{X}}(x) < \gamma^k\}$ (since $\sigma^k = u^k\rho^k$) with $\gamma^k > 0$.

Hence by Lemma 4.4.3 there exist polynomial densities $\rho^{k,\mathrm{pol}} \in \bar{g}Q_{d_k-\deg\bar{g}}(\mathbf{X})$, $\sigma^{k,\mathrm{pol}} \in \bar{g}Q_{d_k-\deg\bar{g}}(\mathbf{X})^m$ for some degrees $d^k \geq 0$ such that

$$\|\rho^k - \rho^{k,\mathrm{pol}}\|_{C^1(\mathbf{X})} < 1/k \tag{4.26}$$

$$\|\sigma_i^k - \sigma_i^{k,\mathrm{pol}}\|_{C^1(\mathbf{X})} < 1/k \tag{4.27}$$

$\bar{u}\rho^{k,\mathrm{pol}} - \sigma_i^{k,\mathrm{pol}} \in \bar{g}Q_{d^k-\deg\bar{g}}(\mathbf{X})$ for all $i = 1,\dots,m$ (since $u^k(x) \in \mathbf{U} = [0,\bar{u}]^m$ for all $x \in \mathbf{X}$ and hence $\bar{u}\rho^k \geq \sigma_i^k$ on $\mathbf{X}$). Notice also that since $\rho^{k,\mathrm{pol}} \in \bar{g}Q_{d^k-\deg\bar{g}}(\mathbf{X})$, we have $-\rho^{k,\mathrm{pol}} \in \bar{g}\mathbb{R}_{d^k-\deg\bar{g}}$. Next, since $\rho_0^k \geq 1$ and $\rho_T^k \geq 0$, we can find, by Corollary 2.2.1, polynomial densities $\hat{\rho}_0^{k,\mathrm{pol}} \in 1 + Q_{d^k}(\mathbf{X})$ and $\hat{\rho}_T^{k,\mathrm{pol}} \in Q_{d^k}(\mathbf{X})$ such that

$$\|\rho_0^k - \hat{\rho}_0^{k,\mathrm{pol}}\|_{C^0(\mathbf{X})} < 1/k, \tag{4.28}$$

$$\|\rho_T^k - \hat{\rho}_T^{k,\mathrm{pol}}\|_{C^0(\mathbf{X})} < 1/k. \tag{4.29}$$

Since $(\rho^k, \rho_0^k, \rho_T^k, \sigma^k)$ satisfy the equality constraint of (4.12) 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 (4.26)-(4.29). 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 $\mathbf{X}$ and hence belong to $Q_{d_k}(\mathbf{X})$. The densities $(\rho^{k,\mathrm{pol}}, \rho_0^{k,\mathrm{pol}}, \rho_T^{k,\mathrm{pol}}, \sigma^{k,\mathrm{pol}})$ are therefore feasible in (4.15) 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 (4.15) 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 (4.23) 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 4.4.2. Therefore (4.7) holds with the rational controllers $u^k := \sigma^{k,\mathrm{pol}}/\rho^{k,\mathrm{pol}}$ by the second part of Lemma 4.4.2. This finishes the proof. □

4.5 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 $0 \le \sigma_i \le \bar{u}\rho$ on $\mathbf{X}$. In addition we describe a hierarchy of approximations from below to the optimal value function V. This is useful as a post-processing step that can be used to get explicit bounds on the suboptimality of the rational controller obtained from the solution to (4.15), although the results of this section apply to any rational 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

$$\begin{array}{cl} \min\limits_{\overline{V_u} \in \mathbb{R}[x]_d} & \int_X \overline{V_u}(x)\,dx \\ \text{s.t.} & \beta\rho\overline{V_u} - \nabla\overline{V_u}\cdot\hat{f} - \hat{l} \in Q_d(\mathbf{X}) \\ & \overline{V_u} - M \in Q_d(\mathbf{X}) + \bar{g}\mathbb{R}_{d-\deg\bar{g}}, \end{array} \tag{4.30}$$

and

$$\begin{array}{cl} \max\limits_{\underline{V_u} \in \mathbb{R}[x]_d} & \int_X \underline{V_u}(x)\,dx \\ \text{s.t.} & -(\beta\rho\underline{V_u} - \nabla\underline{V_u}\cdot\hat{f} - \hat{l}) \in Q_d(\mathbf{X}) \\ & M - \underline{V_u} \in Q_d(\mathbf{X}) + \bar{g}\mathbb{R}_{d-\deg\bar{g}}, \end{array} \tag{4.31}$$

respectively. Fixing a basis of $\mathbb{R}[x]_d$, the objective functions of (4.30) and (4.31) become linear in the coefficients of $\overline{V_u}$ respectively $\underline{V_u}$ in this basis. Problems (4.30) and (4.31) immediately translate to SDPs using the results of Section 2.2.1.

Theorem 4.5.1 *Let* $\overline{V_u}^d$ *and* $\underline{V_u}^d$ *denote solutions to (4.30) and (4.31) of degree* d. *Then* $\overline{V_u}^d \geq V_u \geq \underline{V_u}^d$ *on* $\mathbf{X}$ *and*

$$\lim_{d\to\infty} \int_{\mathbf{X}} \overline{V_u}^d(x)\,dx = \int_{\mathbf{X}} V_u(x)\,dx = \lim_{d\to\infty} \int_{\mathbf{X}} \underline{V_u}^d(x)\,dx. \tag{4.32}$$

Proof: See the Appendix of this Chapter. □

As a simple corollary we obtain a converging sequence of polynomial over-approximations to V, the optimal value function of (4.4):

Theorem 4.5.2 *Let* $\overline{V}^{d_2}_{u^{d_1}}$ *denote the degree* d_2 *polynomial approximation from above to the value function associated to the rational controller* u^{d_1} *obtained from (4.15) using (4.16). Then* $\overline{V}^{d_2}_{u^{d_1}} \geq V$ *on* $\mathbf{X}$ *and*

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

Now we describe a hierarchy of lower bounds on V:

$$\begin{array}{ll} \max\limits_{\underline{V}\in\mathbb{R}[x]_d,\, p\in\mathbb{R}[x]_d^m} & \int_{\mathbf{X}} \underline{V}(x)\,dx \\ \text{s.t.} & l_x - \beta\underline{V} + \nabla\underline{V}\cdot f + \bar{u}\sum_{i=1}^m p_i \in Q_d(\mathbf{X}) \\ & l_{u_i} + \nabla\underline{V}\cdot f_{u_i} - p_i \in Q_d(\mathbf{X}) \\ & -p_i \in Q_d(\mathbf{X}) \\ & M - \underline{V} \in Q_d(\mathbf{X}) + \bar{g}\mathbb{R}_{d-\deg\bar{g}}. \end{array} \tag{4.33}$$

Theorem 4.5.3 *If* $\underline{V} \in \mathbb{R}[x]_d$ *is feasible in (4.33), then* $\underline{V} \leq V$ *on* $\mathbf{X}$.

Proof: Follows by similar arguments based on Gronwall's Lemma as in the proof of Theorem 4.5.1. □

Remark 4.5.1 *The question whether* $\underline{V}$ *converges from below to* V *as degree* d *in (4.33) tends to infinity is open (although likely to hold). A proof would require* *an extension of the superposition Theorem 2.3.3 to non-Lipschitz vector fields or an extension of the argument of [GQ09] to the case of $\mu_T \neq 0$, either of which is beyond the scope of this thesis.*

Remark 4.5.2 *Besides closed-loop cost function with respect to the OCP (4.4), one can assess other aspects of the closed-loop behavior of the dynamical system (4.1) controlled by the rational controller $u = \sigma/\rho$. In particular, the region of attraction or the maximum controlled invariant set can be estimated by methods of Chapter 3, which extend readily to the case of rational systems.*

4.6 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 [Löf04] or SOSOPT [Sei10]). The resulting SDPs were solved using MOSEK.

4.6.1 Nonlinear double integrator

As our first example we consider the nonlinear double integrator

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

subject to the constraints $u \in [-1, 1]$ and $x \in \mathbf{X} := \{x : \|x\|_2 \leq 1\}$ 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 $1.01 > \sup_{x \in \mathbf{X}}\{x^\top x\}/\beta = 1$. The input constraint set $[-1, 1]$ is transformed to $[0, 1]$ using the transformation $u = 2\tilde{u} - 1$ with $\tilde{u} \in [0, 1]$. First we obtain a rational controller of degree six by solving (4.15) with $d = 6$. The graph of the controller is shown in Figure 4.1. Next we obtain a polynomial upper bound $\overline{V_u}$ of degree 14 on the value function associated to u by solving (4.30) 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 (4.4) obtained by solving (4.33) with $d = 14$. The graphs of the two value functions are plotted in Figure 4.2. 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; this is confirmed by looking at the sections of the value function approximations in Figure 4.4. Quantitatively, the average performance gap defined as $100 \int_{\mathbf{X}} (\overline{V_u} - \underline{V}) dx / \int_X \underline{V} dx$ is equal to 19.5%. Finally, Figure 4.3 shows the extracted densities; notice in particular the density of the terminal measure ρ_T which is, as expected, small everywhere except for those regions near the boundary of $\mathbf{X}$ through which the state trajectories are most likely to exit $\mathbf{X}$. Notice also that the initial density ρ_0 is not identically equal to one; this shows that, for this example and for this choice of the degree bounds on the densities, it is beneficial to have an initial density not identically equal to one which results in a higher cost in (4.12) for a *fixed* controller, but allows more freedom in shaping the other densities ρ, σ and ρ_T (and hence the controller $u = \sigma/\rho$) through Equation (4.13), and overall results in a lower cost in (4.15) (and hence in (4.12)).

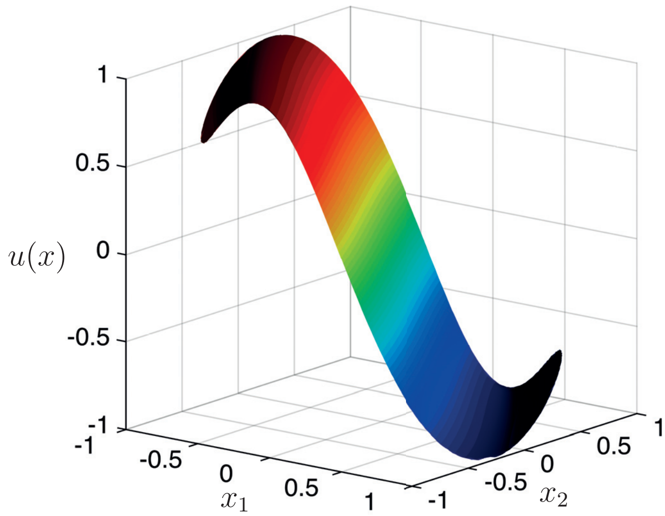


Figure 4.1 – Nonlinear double integrator – rational controller of degree six.

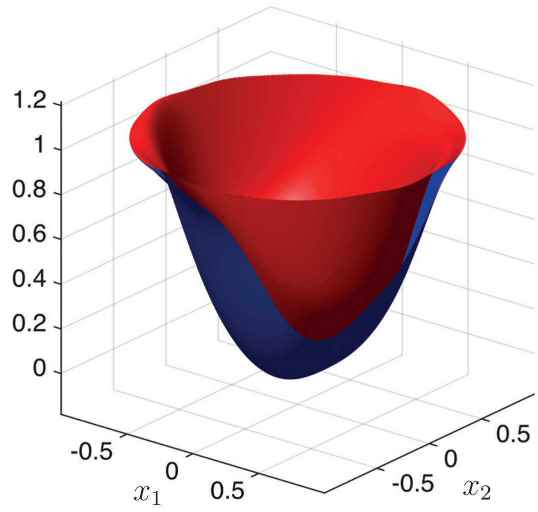


Figure 4.2 – Nonlinear double integrator – upper bound on the value function associated to the designed controller (red); lower bound on the optimal value function (blue).

4.6.2 Controlled Lotka-Volterra

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

$$\dot{x} = r \circ x \circ (\mathbb{1} - Ax) + u^+ - u^-,$$

where $\mathbb{1} \in \mathbb{R}^n$ 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

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

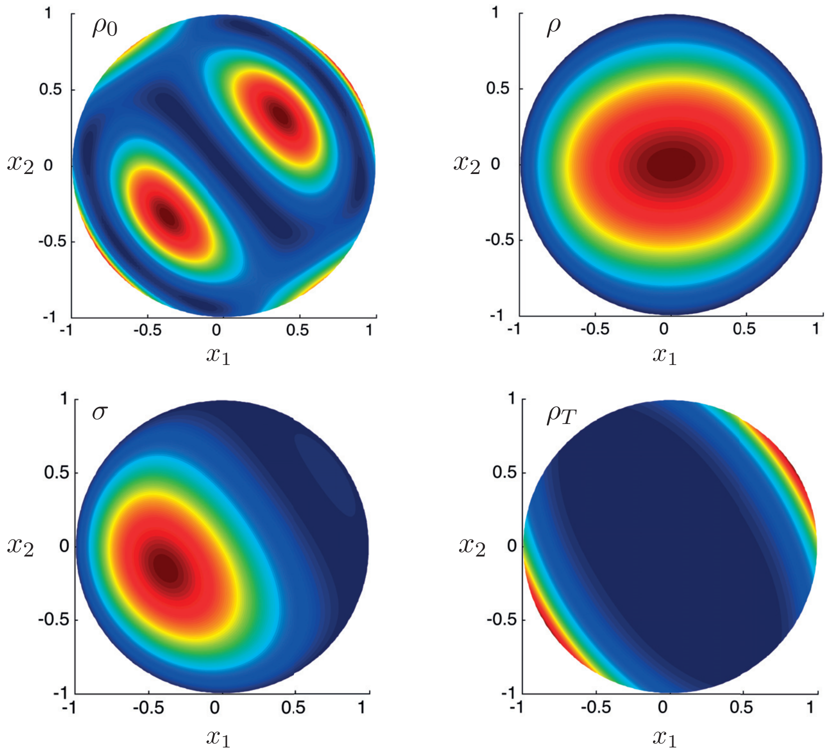


Figure 4.3 – Nonlinear double integrator – densities

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 $l(x, u)$ (4.6) 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 $Q = \mathrm{diag}(0.475 \cdot \mathbb{1})$ and $q = 0.525 \cdot \mathbb{1}$. We choose $\beta = 1$ and $M = 16.16 > \sup_{x\in\mathbf{X},u\in\mathbf{U}}\{l(u, x)\}/\beta = 16$. We apply the coordinate transformation $x = Q\hat{x} + q$ and obtain a rational controller of degree eight by solving (4.15). Figure 4.5 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 (4.33). Using Monte Carlo simulation with 1000 samples of initial conditions drawn from a uniform distribution over the

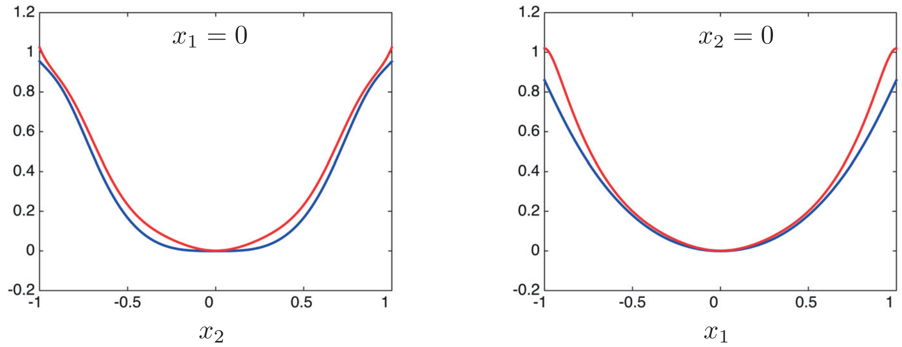


Figure 4.4 – Nonlinear double integrator – sections of the value function approximations for $x_1 = 0$ (left) and $x_2 = 0$ (right). The upper bound on the value function associated to the designed controller is in red; the lower bound on the optimal value function is in blue.

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 (4.30). In this case, however, the upper bound (4.30) is not informative. Nevertheless, the Monte Carlo simulation along with the lower bound (4.33) is a viable alternative in this case, since the extracted controller is simple and hence trajectories of the controlled system can be simulated rapidly.

4.7 Appendix

This Appendix contains the proof of Theorem 4.5.1; we use the same notation as in Section 4.5. 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 (4.30) and (4.31) imply that

$$\nabla \overline{V_u}^d \cdot (f + \sum_{i=1}^{m} f_{u_i} u_i) \leq \beta \overline{V_u}^d - (l_x + \sum_{i=1}^{m} l_{u_i} u_i), \tag{4.34}$$

$$\nabla \underline{V_u}^d \cdot (f + \sum_{i=1}^{m} f_{u_i} u_i) \geq \beta \underline{V_u}^d - (l_x + \sum_{i=1}^{m} l_{u_i} u_i) \tag{4.35}$$

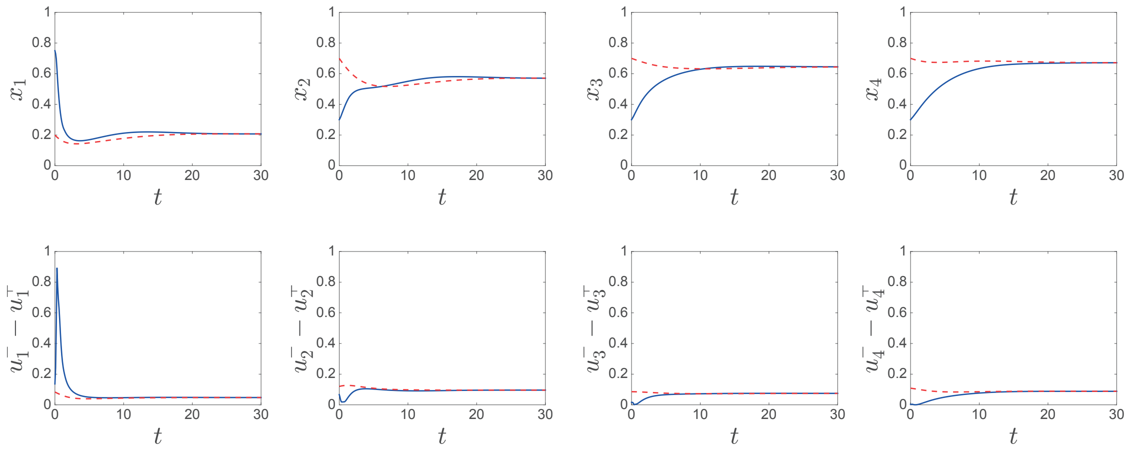


Figure 4.5 – Controlled Lotka-Volterra – (blue) trajectory starting from a high initial population of the first species and low initial population of the other species; (red) trajectory starting from low initial population of the first species and high initial population of the other species.

on X and $\overline{V_u} \geq M$, $\underline{V_u} \leq M$ on $\partial\mathbf{X}$. 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 $x_0 \in \mathbf{X}$ the inequality (4.34) 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

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

and hence

$$\overline{V_u}^d(x_0) \geq e^{-\beta t}\overline{V_u}^d(x(t\,|\,x_0)) + \int_0^t e^{-\beta s}\left[l_x(x(s\,|\,x_0)) + \sum_{i=1}^m l_{u_i}(x(s\,|\,x_0))u_i(x(s\,|\,x_0))\right] ds \tag{4.36}$$

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 $V_u(x_0)$, the value function associated to u, is equal to

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

In view of (4.36), 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 $\partial\mathbf{X}$.

Convergence of the upper and lower bounds (4.32) follows from Theorem 4.4.1 using infinite-dimensional LP duality and standard arguments for proving convergence of moment hierarchies in exactly the same fashion as in the proofs of Theorems 3.1.5 and 3.2.6; hence we only outline the proof. The hierarchy of SOS programming problems (4.30) and (4.31) 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 $C^1(\mathbf{X})$ and feasible solutions of this dual provide upper or lower bounds on V_u. Crucial to applying infinite-dimensional duality strong duality result of Theorem A.2.1 is the boundedness of measures satisfying the discounted Liouville's equation (4.10) with $\nu_i \leq \bar{u}\mu$ and $\mu_0 = \mu_X^{\mathrm{leb}}$, where $\mu_{\mathbf{X}}^{\mathrm{leb}}$ is the restriction of the Lebesgue measure to $\mathbf{X}$. Plugging $v = 1$ in (4.10) we have $\mu_T(\mathbf{X}) + \beta\mu(\mathbf{X}) = \mu_0(\mathbf{X})$. Since $\mu_0(\mathbf{X}) = \mu_{\mathbf{X}}^{\mathrm{leb}}(\mathbf{X}) = \mathrm{vol}\,\mathbf{X} < \infty$ 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 $\mathbf{X}$, which implies the boundedness of the truncated moment sequences feasible in the moment relaxations. The absence of duality gap then follows from [Trn05, Lemma 2]. □

Chapter 5

Verification of optimization-based controllers

This chapter presents a computational approach to analyze closed-loop properties of optimization-based controllers for constrained polynomial discrete-time dynamical systems. We assume that we are given an optimization-based controller that at each time instance generates a control input by solving an optimization problem parametrized by a function of the past measurements of the controlled system's output, and we ask about closed-loop properties of this interconnection. This setting encompasses a wide range of control problems including the control of a polynomial dynamical system by a linear controller (e.g., a PID) with an input saturation, output feedback model predictive control (MPC) with inexact model and soft constraints, or a general optimization-based controller where the underlying problem is solved approximately with a fixed number of iterations of a first-order optimization method. Importantly, the method verifies all KKT points; hence it can be used to verify closed-loop properties of optimization-based controllers where the underlying, possibly nonconvex, optimization problem is solved with a local method with guaranteed convergence to a KKT point only.

The closed-loop properties possible to analyze by the approach include: global stability and stability on a given subset, performance with respect to a discounted infinite-horizon cost (where we provide polynomial upper and lower bounds on the cost attained by the controller over a given set of initial conditions, both in a deterministic and a stochastic setting), the ℓ_2 gain from a given disturbance input to a given performance output (where we provide a numerical upper bound).

The main idea behind the presented approach is the observation that the KKT system associated to an optimization problem with polynomial data is a basic semi-algebraic set (i.e., a system of polynomial equalities and inequalities). Consequently, provided that suitable constraint qualification conditions hold, the solution of this optimization problem belongs to a projection of this set. Hence, the closed-loop evolution of a polynomial dynamical system controlled by an optimization-based controller solving at each time step an optimization problem with polynomial data can be seen as a difference inclusion where the successor state lies in a set defined by polynomial equalities and inequalities. This difference inclusion is then analyzed using sum-of-squares (SOS) techniques.

The approach is based on the observation of Primbs [Pri01] who noticed that the KKT system of a constrained linear quadratic optimization problem is a set of polynomial equalities and inequalities and used the S-procedure to derive sufficient linear matrix inequality (LMI) conditions for a given linear MPC controller to be stabilizing. This chapter significantly extends the approach in terms of both the range of closed-loop properties analyzed and the range of practical problems amenable to the method. Indeed, our approach is applicable to general polynomial dynamical systems, both deterministic and stochastic, and allows the analysis not only of stability but also of various performance measures. The approach is not only applicable to an MPC controller with linear dynamics and a quadratic cost function as in [Pri01] but also to a general optimization-based controller, where the optimization problem may not be solved exactly, encompassing all the above-mentioned control problems.

To the best of our knowledge, this is the first computational method that allows for analysis of optimization-based controllers at this level of generality with applications ranging from classical setups of a PID + saturation to modern embedded control applications where a model predictive controller (MPC) is deployed under tight constraints on computation time. The latter has been an active research topic with methods existing for the constrained linear quadratic case which, however, typically rely on a bound on the number of iterations of a given optimization method to achieve a given accuracy where these bounds are conservative and/or computationally difficult to obtain (e.g., requiring the solution to a mixed-integer optimization problem whose size grows quickly with the input data dimension) [BP12, RJM12].

5.1 Problem statement

We consider the nonlinear discrete-time dynamical system

$$x^+ = f_x(x, u), \tag{5.1a}$$
$$y = f_y(x), \tag{5.1b}$$

where $x \in \mathbb{R}^{n_x}$ is the state, $u \in \mathbb{R}^{n_u}$ the control input, $y \in \mathbb{R}^{n_y}$ the output, $x^+ \in \mathbb{R}^{n_x}$ the successor state, $f_x : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ a transition mapping and $f_y : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ an output mapping. We assume that each component of f_x and f_y is a multivariate polynomial in (x, u) and x, respectively.

We assume that the system is controlled by a given set-valued controller

$$u \in \kappa(\mathbf{K}_s), \tag{5.2}$$

where $\kappa : \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_u}$ is polynomial and

$$\mathbf{K}_s := \{\theta \in \mathbb{R}^{n_\theta} \mid \exists \lambda \in \mathbb{R}^{n_\lambda} \text{ s.t. } g(s, \theta, \lambda) \geq 0, h(s, \theta, \lambda) = 0\}, \tag{5.3}$$

where each component of the vector-valued functions $g : \mathbb{R}^{n_s} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_\lambda} \to \mathbb{R}^{n_g}$ and $h : \mathbb{R}^{n_s} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_\lambda} \to \mathbb{R}^{n_h}$ is a polynomial in the variables (s, θ, λ). The set $\mathbf{K}_s$ is parametrized by the output of a dynamical system

$$z^+ = f_z(z, y), \tag{5.4a}$$
$$s = f_s(z, y), \tag{5.4b}$$

where $f_z : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_z}$ and $f_s : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \to \mathbb{R}^{n_s}$ are polynomial. The problem setup is depicted in Figure 5.1. In the rest of this chapter we develop a method to analyze the closed-loop stability and performance of this interconnection. Before doing that we present several examples which fall into the presented framework.

5.2 Examples

The framework considered allows for the analysis of a large number of practical control scenarios. The common idea is to write the control input u as the output of an optimization problem with polynomial data parametrized by the state of the dynamical system (5.4). The control input u then belongs to the associated

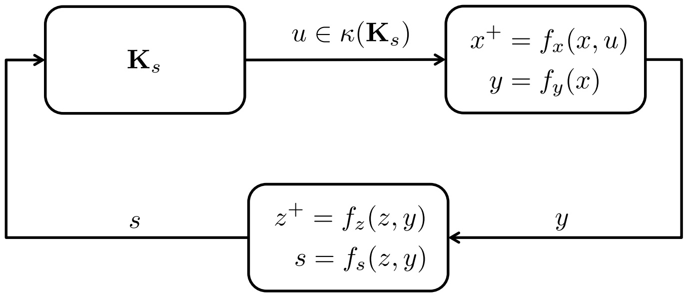


Figure 5.1 – Control scheme

KKT system (provided that mild regularity conditions are satisfied) which is of the form (5.3).

5.2.1 Polynomial dynamical controller + input saturation

Any polynomial dynamical controller (e.g., a PID controller) plus an input saturation can be written in the presented form provided that the input constraint set is defined by finitely many polynomial inequalities satisfying mild constraint qualification conditions (see, e.g., [Pet73]). Indeed, regarding z as the state and s as the output of the controller and generating u according to $u \in \mathrm{proj}_{\mathbf{U}}(s)$, where

$$\mathrm{proj}_{\mathbf{U}}(s) = \arg\min_{\theta \in \mathbf{U}} \frac{1}{2}\|\theta - s\|_2^2 \tag{5.5}$$

is the set of Euclidean projections of s on the constraint set $\mathbf{U}$ (there can be multiple solutions since $\mathbf{U}$ is in general nonconvex). Assuming that the input constraint set is of the form

$$\mathbf{U} = \{v \in \mathbb{R}^{n_u} \mid g_U(v) \geq 0\}, \tag{5.6}$$

where $g_U : \mathbb{R}^{n_u} \to \mathbb{R}^{n_{g_U}}$ has polynomial entries, the KKT conditions associated to the optimization problem (5.5) read

$$\theta - s - \nabla g_U(\theta)\lambda = 0 \tag{5.7a}$$
$$\lambda^\top g_U(\theta) = 0 \tag{5.7b}$$
$$\lambda \geq 0 \tag{5.7c}$$
$$g_U(\theta) \geq 0, \tag{5.7d}$$

where $\lambda \in \mathbb{R}^{n_u}$ is the vector of Lagrange multipliers associated with the constraints defining $\mathbf{U}$ and ∇g_U is the transpose of the Jacobian of g_U (i.e., $[\nabla g_U]_{i,j} = \frac{\partial [g_U]_j}{\partial x_i}$). Assuming that constraint qualification conditions hold such that any minimizer of (5.5) satisfies the KKT conditions (5.7) we conclude that

$$u \in \kappa(\mathbf{K}_s)$$

with κ being the identity (i.e., $\kappa(\theta) = \theta$),

$$h(s, \theta, \lambda) = \begin{bmatrix} \theta - s - \nabla g_U(\theta)\lambda \\ \lambda^\top g_U(\theta) \end{bmatrix}$$

and

$$g(s, \theta, \lambda) = \begin{bmatrix} \lambda \\ g_U(\theta) \end{bmatrix},$$

where h and g are polynomials in (s, θ, λ) as required.

Note that the description of the input constraint set (5.6) is not unique. For example, if the input constraint set is $[-1, 1]$, then the function g_U can be

$$g_U(\theta) = (1 - \theta)(1 + \theta) \tag{5.8}$$

or

$$g_U(\theta) = \begin{bmatrix} 1 - \theta \\ 1 + \theta \end{bmatrix} \tag{5.9}$$

or any odd powers of the above. Depending on the particular description, the constraint qualification conditions may or may not hold. It is therefore important to choose a suitable description of $\mathbf{U}$ which is both simple and such that the constraint qualification conditions hold. We remark that in the case of $\mathbf{U} = [-1, 1]$ both (5.8) and (5.9) satisfy these requirements.

Note also that the KKT system (5.7) may be satisfied by points which are not global minimizers of (5.5) if the set $\mathbf{U}$ is nonconvex; this is an artefact of the presented method and cannot be avoided within the presented framework. We note, however, that the input constraint set $\mathbf{U}$ is in most practical cases convex.

5.2.2 Output feedback nonlinear MPC with model mismatch and soft constraints

This example shows how to model within the presented framework a nonlinear MPC controller with state estimation, a model mismatch[1], soft constraints and no a priori stability guarantees (enforced, e.g., using a terminal penalty and/or terminal set) and possibly only locally optimal solutions delivered by the optimization algorithm. In this case the system (5.4) is an estimator of the state of the dynamical system (5.1) and in each time step the following optimization problem is solved

$$\begin{array}{ll}\underset{\hat{\boldsymbol{u}},\hat{\boldsymbol{x}},\boldsymbol{\varepsilon}}{\text{minimize}} & l_{\mathrm{s}}(\boldsymbol{\varepsilon}) + \sum_{i=0}^{N-1} l_i(s,\hat{x}_i,\hat{u}_i) + l_N(s,x_N) \\ \text{subject to} & \hat{x}_{i+1} = \hat{f}(\hat{x}_i,\hat{u}_i), \quad i = 0,\ldots,N-1 \\ & a(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon}) \geq 0 \\ & b(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}}) = 0,\end{array} \tag{5.10}$$

where $\hat{\boldsymbol{x}} = (\hat{x}_0,\ldots,\hat{x}_N)$, $\hat{\boldsymbol{u}} = (\hat{u}_0,\ldots,\hat{u}_{N-1})$, $\boldsymbol{\varepsilon} = (\varepsilon_1,\ldots,\varepsilon_{n_\varepsilon})$ are slack variables for the inequality constraints, $\hat{f}$ is a polynomial model of the true transition mapping f_x, l_{s} is a polynomial penalty for violations of the inequality constraints, l_i, $i = 0,\ldots,N$, are polynomial stage costs and a and b (vector) polynomial constraints parametrized by the state estimate s produced by (5.4). If the dimension of the state estimate s is equal to the dimension of the state of the model $\hat{x}$ (which we do not require), then most MPC formulations will impose $\hat{x}_0 = s$, which is encoded by making one of the components of b equal to $\hat{x}_0 - s$. The formulation (5.10) is, however, not restricted to this scenario and allows arbitrary dependence of the constraints (along the whole prediction horizon) on the state estimate s. The control input applied to the system is then $u = \hat{u}_0^\star$, where $\hat{u}_0^\star$ is the first component of any vector $\hat{\boldsymbol{u}}^\star$ optimal in (5.10).

[1] In this example we assume that we know the true model of the system but in the MPC controller we intentionally use a different model (e.g., we use a linearized or otherwise simplified model for the sake of computation speed); the true model is used only to *verify* closed-loop properties of the true model controlled by the MPC controller. See Section 5.3.5 for the case where the true model is not known exactly even for the verification purposes and the model mismatch is captured by an exogenous disturbance.

The KKT system associated to (5.10) reads

$$\nabla_{\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon}}\mathcal{L}(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon},\lambda) = 0 \tag{5.11a}$$
$$\lambda_a^\top a(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon}) = 0 \tag{5.11b}$$
$$b(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}}) = 0 \tag{5.11c}$$
$$\hat{x}_{i+1} - \hat{f}(\hat{x}_i,\hat{u}_i) = 0, \quad i = 0,\ldots,N-1 \tag{5.11d}$$
$$\lambda_a \geq 0 \tag{5.11e}$$
$$a(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon}) \geq 0, \tag{5.11f}$$

where $\lambda := (\lambda_a, \lambda_b, \lambda_{\hat{f}}^0, \ldots, \lambda_{\hat{f}}^{N-1})$ and

$$\begin{aligned}\mathcal{L}(s,\hat{x},\hat{u},\boldsymbol{\varepsilon},\lambda) := {} & l_{\mathrm{s}}(\boldsymbol{\varepsilon}) + \sum_{i=0}^{N-1} l_i(s,\hat{x}_i,\hat{u}_i) - \lambda_a^\top(a(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}},\boldsymbol{\varepsilon})) \\ & + l_N(s,x_N) + \lambda_b^\top b(s,\hat{\boldsymbol{x}},\hat{\boldsymbol{u}}) + \sum_{i=0}^{N-1} \lambda_{\hat{f}}^i(\hat{x}_{i+1} - \hat{f}(\hat{x}_i,\hat{u}_i))\end{aligned}$$

is the Lagrangian of (5.10). The KKT system (5.11) is a system of polynomial equalities and inequalities. Consequently, setting

$$\theta := (\hat{\boldsymbol{u}},\hat{\boldsymbol{x}},\boldsymbol{\varepsilon}), \quad \kappa(\theta) = \hat{u}_0$$

and assuming that constraint qualification conditions hold such that every optimal solution to (5.10) satisfies the KKT condition (5.11), there exist polynomial functions $h(s,\theta,\lambda)$ and $g(s,\theta,\lambda)$ such that $\hat{u}_0^\star \in \kappa(\mathbf{K}_s)$ for every $\hat{u}_0^\star$ optimal in (5.10).

Remark 5.2.1 *Let us mention that, provided suitable constraint qualification conditions hold, not only every globally optimal $\hat{u}_0^\star$ will satisfy the KKT system but also every* locally optimal *solution to (5.10) and every critical point of (5.10) will; hence the proposed method can be used to verify stability and performance properties even if only local solutions to (5.10) are delivered by the optimization algorithm.*

Remark 5.2.2 *Note that the situation where the optimization problem (5.10) is not solved exactly can be handled as well. One way to do so is to include an auxiliary variable δ capturing the inaccuracy in the solution, either in the satisfaction of the KKT system (5.11) or directly as an error on the delivered control action $\hat{u}_0$ (e.g., defining $\kappa(\theta) = \hat{u}_0(1+\delta)$ with $\theta = (\hat{\boldsymbol{u}},\hat{\boldsymbol{x}},\boldsymbol{\varepsilon},\delta)$) and imposing $|\delta| \leq \Delta$, where $\Delta > 0$*

is a known bound on the solution accuracy. If the solution inaccuracy is due to a premature termination of a first-order optimization method used to solve (5.10), a more refined analysis can be carried out within the presented framework; this is detailed in Section 5.2.4.

5.2.3 General optimization-based controller

Clearly, there was nothing specific about the MPC structure of the optimization problem solved in the previous example and therefore the presented framework can be used to analyze arbitrary optimization-based controllers which at each time step solve an optimization problem parametrized by the output of the dynamical system (5.4):

$$\begin{array}{ll} \underset{\theta\in\mathbb{R}^{n_\theta}}{\text{minimize}} & J(s,\theta) \\ \text{subject to} & a(s,\theta)\geq 0 \\ & b(s,\theta)=0, \end{array} \tag{5.12}$$

with J, a and b polynomial. The associated KKT system reads

$$\nabla_\theta J(s,\theta) - \nabla_\theta a(s,\theta)\lambda_a + \nabla_\theta b(s,\theta)\lambda_b = 0 \tag{5.13a}$$
$$\lambda_a^\top a(s,\theta) = 0 \tag{5.13b}$$
$$\lambda_a \geq 0 \tag{5.13c}$$
$$a(s,\theta) \geq 0 \tag{5.13d}$$
$$b(s,\theta) = 0, \tag{5.13e}$$

which is a system of polynomial equalities and inequalities in (s,θ,λ), where $\lambda = (\lambda_a,\lambda_b)$ and hence can be treated within the given framework. In particular the functions h and g defining the set $\mathbf{K}_s$ read

$$h(s,\theta,\lambda) = \begin{bmatrix} \nabla_\theta J(s,\theta) - \nabla_\theta a(s,\theta)\lambda_a + \nabla_\theta b(s,\theta)\lambda_b \\ \lambda_a^\top a(s,\theta) \\ b(s,\theta) \end{bmatrix}$$

and

$$g(s,\theta,\lambda) = \begin{bmatrix} \lambda_a \\ a(s,\theta) \end{bmatrix}.$$

See Remark 5.2.2 for the situation where the problem (5.12) is not solved exactly.

5.2.4 Optimization-based controller solved using a fixed number of iterations of a first order method

The presented framework can also handle the situation where the optimization problem (5.12) is solved using an iterative optimization method, each step of which is either an optimization problem or a polynomial mapping. This scenario was elaborated on in detail in [KJ13], where it was shown that the vast majority of first order optimization methods fall into this category. Here we present the basic idea on one of the simplest optimization algorithms, the projected gradient method. When applied to problem (5.12), the iterates of the projected gradient method are given by

$$\theta_{k+1} \in \mathrm{proj}_{\mathbf{S}}(\theta_k - \eta \nabla_\theta J(s,\theta_k)), \tag{5.14}$$

where $\mathrm{proj}_{\mathbf{S}}(\cdot)$ denotes the set of Euclidean projections on the constraint set

$$\mathbf{S} = \{\theta \in \mathbb{R}^{n_\theta} \mid a(s,\theta) \geq 0,\ b(s,\theta) = 0\}$$

and $\eta > 0$ is a step size. The update formula (5.14) can be decomposed into two steps: step in the direction of the negative gradient and projection on the constraint set. The first step is a polynomial mapping and the second step is an optimization problem. Indeed, equation (5.14) can be equivalently written as

$$\begin{aligned} \theta_{k+1} \in \ & \underset{\theta\in\mathbb{R}^{n_\theta}}{\arg\min} && \tfrac{1}{2}\|\theta - (\theta_k - \eta\nabla_\theta J(s,\theta_k))\|_2^2 \\ & \text{s.t.} && a(s,\theta) \geq 0 \\ & && b(s,\theta) = 0. \end{aligned} \tag{5.15}$$

For each $k \in \{0, 1, \ldots\}$, the KKT system associated to (5.15) reads

$$\begin{aligned} \theta_{k+1} - (\theta_k - \eta\nabla_\theta J(s,\theta_k)) - \nabla_\theta\, a(s,\theta_{k+1})\lambda_a^{k+1}& \\ +\nabla_\theta\, b(s,\theta_{k+1})\lambda_b^{k+1} &= 0 && (5.16a) \\ b(s,\theta_{k+1}) &= 0 && (5.16b) \\ a(s,\theta_{k+1})^\top \lambda_a^{k+1} &= 0 && (5.16c) \\ a(s,\theta_{k+1}) &\geq 0 && (5.16d) \\ \lambda_a^{k+1} &\geq 0, && (5.16e) \end{aligned}$$

which is a system of polynomial equalities and inequalities. Note in particular the coupling between θ_k and θ_{k+1} in equation (5.16b). Assuming we apply M steps of the projected gradient method, the last iterated θ_M is therefore characterized by M coupled KKT systems of the form (5.16), which is a system of polynomial equalities and inequalities as required by the proposed method.

Other optimization methods, in particular most of the first order methods (e.g., fast gradient method [Nes04], AMA [Tse91], ADMM [GM76] and their accelerated versions [GOSB14]), including local non-convex methods (e.g., [HJ14]), and some of the second order methods (e.g., the interior point method with exact line search) are readily formulated in this framework as well; see [KJ13] for more details on first-order methods.

5.3 Closed-loop analysis

In this section we describe a method to analyze closed-loop properties of the interconnection depicted in Figure 5.1 and described in Section 5.1. First, notice that the closed-loop evolution is governed by the difference inclusion

$$x^+ \in f_x(x, \kappa(\mathbf{K}_s)), \tag{5.17a}$$

$$z^+ = f_z(z, f_y(x)). \tag{5.17b}$$

Since all problem data is polynomial and the set $\mathbf{K}_s$ is basic semialgebraic it is possible to analyze stability and performance using sum-of-squares (SOS) programming. This is detailed next.

We will use the following notation:

$$\hat{h}(x, z, \theta, \lambda) := h(f_s(z, f_y(x)), \theta, \lambda) \tag{5.18a}$$

$$\hat{g}(x, z, \theta, \lambda) := g(f_s(z, f_y(x)), \theta, \lambda). \tag{5.18b}$$

For the rest of the chapter we impose the following standing assumption:

Assumption 5.3.1 *The set $\mathbf{K}_s$ is nonempty for all $s \in \mathbb{R}^{n_s}$.*

Assumption 5.3.1 implies that the control input (5.2) is well defined for all $s \in \mathbb{R}^{n_s}$.

5.3.1 Stability analysis – global

A sufficient condition for the state x of the difference inclusion (5.17) to be stable is the existence of a function V satisfying

$$V(x^+, z^+, \theta^+, \lambda^+) - V(x, z, \theta, \lambda) \leq -\|x\|_2^2 \tag{5.19a}$$
$$V(x, z, \theta, \lambda) \geq \|x\|_2^2 \tag{5.19b}$$

for all

$$(x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+) \in \mathbf{K},$$

where

$$\begin{aligned}&\mathbf{K} = \{(x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+) \mid x^+ = f_x(x, \kappa(\theta)),\\ &\hat{h}(x, z, \theta, \lambda) = 0,\ \hat{g}(x, z, \theta, \lambda) \geq 0,\ \hat{h}(x^+, z^+, \theta^+, \lambda^+) = 0,\\ &\hat{g}(x^+, z^+, \theta^+, \lambda^+) \geq 0,\ z^+ = f_z(z, f_y(x))\}.\end{aligned} \tag{5.20}$$

These equations require that a Lyapunov function V exists which decreases on the basic semialgebraic set $\mathbf{K}$ implicitly characterizing the closed-loop evolution (5.17). Therefore, we can seek a Lyapunov function for system (5.17) by restricting V to be a polynomial of a pre-defined degree and replacing the inequalities (5.19) by tractable sufficient conditions. For simplicity, we use here the inclusion to a quadratic module as the sufficiency condition, although from a practical point of view it is advisable to use simpler subsets of the quadratic modules in order for the resulting optimization problem to be tractable; see Section 6.1 for a discussion of computation aspects. Since we will refer to individual SOS multipliers appearing in the definition of a quadratic module later when discussing computational results, we write out the inclusion to a quadratic module explicitly when necessary.

Setting

$$\xi := (x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+),$$

a sufficient condition for (5.19) to hold is

$$\begin{aligned} &V(x,z,\theta,\lambda) - V(x^+,z^+,\theta^+,\lambda^+) - \|x\|_2^2 = \qquad (5.21a)\\ &\quad \sigma_0(\xi) + \sigma_1(\xi)^\top \hat{g}(x,z,\theta,\lambda) + \sigma_2(\xi)^\top \hat{g}(x^+,z^+,\theta^+,\lambda^+)\\ &\quad + p_1(\xi)^\top \hat{h}(x,z,\theta,\lambda) + p_2(\xi)^\top \hat{h}(x^+,z^+,\theta^+,\lambda^+)\\ &\quad + p_3(\xi)(x^+ - f_x(x,\kappa(\theta)) + p_4(\xi)(z^+ - f_z(z,f_y(x)),\\ &V(x,z,\theta,\lambda) - \|x\|_2^2 = \qquad (5.21b)\\ &\quad \bar{\sigma}_0(\xi) + \bar{\sigma}_1(\xi)^\top \hat{g}(x,z,\theta,\lambda) + \bar{p}_1(\xi)^\top \hat{h}(x,z,\theta,\lambda), \end{aligned}$$

where $\sigma_i(\xi)$ and $\bar{\sigma}_i(\xi)$ are SOS multipliers and $p_i(\xi)$ and $\bar{p}_i(\xi)$ polynomial multipliers of compatible dimensions and pre-specified degrees. The satisfaction of (5.21a) implies the satisfaction of (5.19a) and the satisfaction of (5.21b) implies the satisfaction of (5.19b) for all $\xi \in \mathbf{K}$; this follows readily from the results of Section 2.2.1 since the satisfaction of (5.21) implies that the left-hand sides of (5.21a) and (5.21b) belong to $Q_d(\mathbf{K})$ for some $d \geq 0$.

Remark 5.3.1 *Note that instead of including the equalities $x^+ - f_x(x,\kappa(\theta))$ and $z^+ - f_z(z,f_y(x))$ in the description of $\mathbf{K}$ we could also directly substitute for x^+ and z^+. In general, direct substitution is preferred if the mappings f_x, f_z and f_y are of low degree, especially linear, in which case there is no increase in the degree of the composition of V with f_x or with f_z and f_y. Otherwise, the formulation (5.21) is preferred.*

From the previous discussion we conclude that closed-loop stability of the state x of (5.17) is implied by the feasibility of the following SOS problem, which immediately translates to an SDP:

$$\begin{array}{ll} \text{find} & V,\sigma_0,\sigma_1,\sigma_2,p_1,p_2,p_3,p_4,\bar{\sigma}_0,\bar{\sigma}_1,\bar{p}_1 \\ \text{s.t.} & (5.21a),(5.21b) \\ & \sigma_0,\sigma_1,\sigma_2,\bar{\sigma}_0,\bar{\sigma}_1 \quad \text{SOS polynomials} \\ & V,p_1,p_2,p_3,p_4,\bar{p}_1 \quad \text{arbitrary polynomials,} \end{array} \qquad (5.22)$$

where the decision variables are the coefficients of the polynomials

$$(V,\sigma_0,\sigma_1,\sigma_2,p_1,p_2,p_3,p_4,\bar{\sigma}_0,\bar{\sigma}_1,\bar{p}_1).$$

The following theorem summarizes the results of this section.

Theorem 5.3.1 *If the problem (5.22) is feasible, then the state x of the closed-loop system (5.17) is globally asymptotically stable.*

5.3.2 Stability analysis – on a given subset

This section addresses the stability analysis on a given subset $\mathbf{X}$ of the state space $\mathbb{R}^{n_x} \times \mathbb{R}^{n_z}$ of the difference inclusion (5.17). The set $\mathbf{X}$ serves to restrict the search for a stability certificate to a given subset of the state space if global stability cannot be proven, as well as it can be used to encode physical constraints on the states of the original system x and/or known relationships between the states x and z (e.g., if z is an estimate of x and a bound on the estimation error is known).

We assume that the set $\mathbf{X}$ is defined as

$$\mathbf{X} := \{(x,z) \in \mathbb{R}^{n_x+n_z} \mid \psi_i(x,z) \geq 0,\ i = 1,\dots,n_\psi\}. \tag{5.23}$$

where $\psi_i(\cdot)$ are polynomials. The Lyapunov conditions (5.19) are then enforced on the intersection of $\mathbf{X}$ with the set $\mathbf{K}$ (defined in (5.20)), i.e., on the set

$$\begin{aligned}\bar{\mathbf{K}} := \{&(x,z,\theta,\lambda,x^+,z^+,\theta^+,\lambda^+) \mid \\ &(x,z,\theta,\lambda,x^+,z^+,\theta^+,\lambda^+) \in \mathbf{K},\ (x,z) \in \mathbf{X}\},\end{aligned}$$

which is a set defined by finitely many polynomial equalities and inequalities. Hence the problem of stability verification on $\mathbf{X}$ leads to an SOS problem completely analogous to (5.22).

However, the pitfall here is that the satisfaction of Lyapunov conditions (5.19) (with $\mathbf{K}$ replaced by $\bar{\mathbf{K}}$) does not ensure the invariance of the closed-loop evolution of (x,z) in the set $\mathbf{X}$. Asymptotic stability is guaranteed only on the largest sub-level set contained in $\mathbf{X}$ of the function

$$\bar{V}(x,z) = \sup_{\theta,\lambda}\{V(x,z,\theta,\lambda) \mid (x,z,\theta,\lambda) \in \hat{\mathbf{K}}\},$$

where

$$\begin{aligned}\hat{\mathbf{K}} := \{(x,z,\theta,\lambda) \mid \hat{g}(x,z,\theta,\lambda) \geq 0,\ \hat{h}(x,z,\theta,\lambda) = 0,\\ (x,z) \in \mathbf{X}\}.\end{aligned}$$

Finding this largest sub-level set or an inner approximation to it is hard in general.

Nevertheless, if we choose V as a function of (x, z) only, then trivially $\bar{V}(x, z) = V(x, z)$ and in this case finding an inner approximation is possible.

This can be done by solving the following optimization problem:

$$\begin{aligned} &\underset{\gamma\in\mathbb{R}_+,\{\sigma_{0,i}\},\{\sigma_{0,i}\}}{\text{maximize}} \quad \gamma \\ &\text{s.t. } \psi_i(x,z) = \sigma_{0,i}(x,z) + \sigma_{1,i}(x,z)(\gamma - V(x,z)), \\ &\quad \{\sigma_{0,i}\},\ \{\sigma_{1,i}\} \quad \text{SOS polynomials} \quad \forall\, i \in \{1,\dots,n_\psi\}. \end{aligned} \tag{5.24}$$

Satisfaction of the first constraint implies that $\psi_i(x, z) \geq 0$ for all (x, z) such that $V(x, z) \leq \gamma$ and all $i \in \{1, \dots, n_\psi\}$; therefore $\{(x, z) \mid V(x, z) \leq \gamma\} \subset \mathbf{X}$ for any γ feasible in (5.24). Maximizing γ then maximizes the size of the inner approximation.

Problem (5.24) is only quasi-convex because of the bilinearity between σ_1 and γ but can be efficiently solved using a bi-section on γ. Indeed, for a fixed value of γ problem (5.24) is an SDP, typically of much smaller size than (5.22).

This immediately leads to the following theorem.

Theorem 5.3.2 *If a polynomial $V \in \mathbb{R}[x, z]$ satisfies (5.19) for all*

$$(x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+) \in \bar{\mathbf{K}}$$

and $\gamma \in \mathbb{R}_+$ is feasible in (5.24), then all trajectories of the closed-loop system (5.17) starting from the set $\{(x, z) \mid V(x, z) \leq \gamma\}$ lie in the set $\mathbf{X}$ and $\lim_{t\to\infty} x_t = 0$.

Since $\bar{\mathbf{K}}$ is defined by finitely many polynomial equalities and inequalities, the search for V satisfying the conditions of Theorem 5.3.2 can be cast as an SOS problem completely analogous to (5.22).

5.3.3 Performance analysis – deterministic setting

In this section we analyze the performance of the controller (5.2) with respect to a given cost function. The performance is analyzed for all initial conditions belonging to a given set $\mathbf{X}$ defined in (5.23). In order to facilitate the performance analysis of the difference inclusion (5.17) we introduce a *selection oracle*:

Definition 2 (Selection oracle) *A selection oracle is any function* $\mathcal{O} : 2^{\mathbb{R}^{n_u}} \setminus \{\emptyset\} \to \mathbb{R}^{n_u}$ *satisfying* $\mathcal{O}(A) \in A$ *for all* $A \subset \mathbb{R}^{n_u}$*,* $A \neq \emptyset$*.*

In words, a selection oracle is a function which selects one point from any nonempty subset of $\mathbb{R}^{n_u}$ (note that at least one such function exists by the axiom of choice). The performance analysis of this section then pertains to the discrete-time recurrence

$$x_{t+1} = f_x(x_t, \mathcal{O}_t(\kappa(\mathbf{K}_{s_t}))), \tag{5.25a}$$

$$z_{t+1} = f_z(z_t, f_y(x_t)), \tag{5.25b}$$

and all results hold for an arbitrary and possibly time-dependent sequence of selection oracles $\mathcal{O}_t$; hence in what follows we suppress the dependence of all quantities on the selection oracle sequence. The cost function with respect to which we analyze performance is

$$\mathcal{C}(x_0, z_0) = L\alpha^{\tau(x_0,z_0)} + \sum_{t=0}^{\tau(x_0,z_0)-1} \alpha^t l(x_t, u_t), \tag{5.26}$$

$(x_t, z_t)_{t=0}^{\infty}$ is the solution to (5.17), $u_t = \mathcal{O}(\kappa(\mathbf{K}_{s_t}))$, $\alpha \in (0, 1)$ is a discount factor, l is a polynomial stage cost,

$$\tau(x, z) := \inf\{t \in \{1, 2, \ldots\} \mid (x_t, z_t) \notin \mathbf{X},\ (x_0, z_0) = (x, z)\} \tag{5.27}$$

is the first time that the state (x_t, z_t) leaves $\mathbf{X}$ (setting $\tau(x, z) = +\infty$ if $(x_t, u_t) \in \mathbf{X}$ for all t) and

$$L > \sup\{l(x, u) \mid (x, z) \in \mathbf{X}, u \in \kappa(\mathbf{K}_s)\}/(1 - \alpha) \tag{5.28}$$

is a constant upper bounding the stage cost l on $\mathbf{X}$ divided by $1 - \alpha$. We assume that $L < \infty$, which is fulfilled if the projection of $\mathbf{X}$ on $\mathbb{R}^{n_x}$ is bounded and the set $\mathbf{K}_s$ is bounded for all $s = f_s(z, f_y(x))$ with $(x, z) \in \mathbf{X}$. A constant L satisfying (5.28) is usually easily found since $\mathbf{X}$ is known and the controller (5.2) is usually set up in such a way that it satisfies the input constraints of system (5.1), which are typically a bounded set of a simple form.

The reason for choosing (5.26) is twofold, similarly to Chapter 4 where a similar cost function was employed. First, $\mathcal{C}(x_0, z_0) = \sum_{t=0}^{\infty} \alpha^t l(x_t, u_t)$ for all (x_0, z_0) such that $(x_t, z_t) \in \mathbf{X}$ for all t; that is, whenever (x_t, z_t) stays in the state constraint set $\mathbf{X}$ for all t, the cost (5.26) coincides with the standard infinite-horizon discounted cost. Second, $\mathcal{C}(x_0, z_0) \leq L$ for all $(x_0, z_0) \in \mathbf{X}$; that is, the cost function is bounded on

$\mathbf{X}$, which enables us to obtain polynomial upper and lower bounds on $\mathcal{C}$ (which is not possible if $\mathcal{C}$ is infinite outside the maximum positively invariant subset of (5.17) included in $\mathbf{X}$ as is the case for the standard infinite-horizon discounted cost).

In the rest of this section we derive polynomial upper and lower bounds on $\mathcal{C}(x, z)$. To this end define

$$\hat{\mathbf{K}}_c := \{(x, z, \theta, \lambda) \mid \hat{g}(x, z, \theta, \lambda) \geq 0,\ \hat{h}(x, z, \theta, \lambda) = 0,\\ (x, z) \notin \mathbf{X}\}.$$

The upper bound is based on the following lemma:

Lemma 5.3.1 *If*

$$V(x, z, \theta, \lambda) - \alpha V(x^+, z^+, \theta^+, \lambda^+) - l(x, \kappa(\theta)) \geq 0 \\ \forall\, (x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+) \in \bar{\mathbf{K}}, \tag{5.29}$$

$$V(x, z, \theta, \lambda) \geq L \quad \forall\, (x, z, \theta, \lambda) \in \hat{\mathbf{K}}_c \tag{5.30}$$

and

$$\overline{V}(x, z) \geq V(x, z, \theta, \lambda) \quad \forall\, (x, z, \theta, \lambda) \in \hat{\mathbf{K}}, \tag{5.31}$$

then $\overline{V}(x, z) \geq \mathcal{C}(x, z)$ *for all* $(x, z) \in \mathbf{X}$.

Proof: See the Appendix of this Chapter. □

The lower bound is based on the following lemma:

Lemma 5.3.2 *If*

$$V(x, z, \theta, \lambda) - \alpha V(x^+, z^+, \theta^+, \lambda^+) - l(x, \kappa(\theta)) \leq 0 \\ \forall\, (x, z, \theta, \lambda, x^+, z^+, \theta^+, \lambda^+) \in \bar{\mathbf{K}}, \tag{5.32}$$

$$V(x, z, \theta, \lambda) \leq L \quad \forall\, (x, z, \theta, \lambda) \in \hat{\mathbf{K}}_c \tag{5.33}$$

and

$$\underline{V}(x, z) \leq V(x, z, \theta, \lambda) \quad \forall\, (x, z, \theta, \lambda) \in \hat{\mathbf{K}}, \tag{5.34}$$

then $\underline{V}(x, z) \leq \mathcal{C}(x, z)$ *for all* $(x, z) \in \mathbf{X}$.

Proof: Analogous to the proof of Lemma 5.3.1. □

The previous two lemmas lead immediately to optimization problems providing upper and lower bounds on $\mathcal{C}(x,z)$.

An upper bound on $\mathcal{C}(x,z)$ is given by the following optimization problem:

$$\begin{array}{ll} \underset{V,\overline{V}}{\text{minimize}} & \int_{\mathbf{X}} \overline{V}(x,z)\rho(x,z)d(x,z) \\ \text{s.t.} & (5.29),(5.30),(5.31), \end{array} \tag{5.35}$$

where $\rho(x,z)$ is a user-defined nonnegative weighting function allowing one to put a different weight on different initial conditions. Typical examples are $\rho(x,z)=1$ or $\rho(x,z)$ equal to the indicator function of a certain subset of $\mathbf{X}$ (see Example 5.4.1).

A lower bound on $\mathcal{C}(x,z)$ is given by the following optimization problem:

$$\begin{array}{ll} \underset{V,\underline{V}}{\text{maximize}} & \int_{\mathbf{X}} \underline{V}(x,z)\rho(x,z)d(x,z) \\ \text{s.t.} & (5.32),(5.33),(5.34). \end{array} \tag{5.36}$$

In both optimization problems, the optimization is over continuous functions $(V,\overline{V})$ or $(V,\underline{V})$; in order to make the problems tractable we tighten the problems by restricting the class of functions to polynomials and by replacing the constraints by inclusions to SDP representable cones. For (5.29), (5.31) and (5.32), (5.34), the resulting conditions are completely analogous to (5.21) (i.e., they are inclusions to truncated quadratic modules as defined in Section 2.2.1). For (5.30) and (5.33) we have to deal with the condition $(x,z)\notin\mathbf{X}$. A sufficient condition for (5.30) is

$$\begin{aligned} V(x,z,\theta,\lambda)-L &= -\sigma_{\psi_i}(\zeta)\psi_i(x,z)+\sigma_0(\zeta) \\ &+\sigma_1(\zeta)^\top\hat{g}(x,z,\theta,\lambda)+\bar{p}_1(\zeta)^\top\hat{h}(x,z,\theta,\lambda)\;\forall\, i\in\{1,\ldots,n_\psi\}, \end{aligned} \tag{5.37}$$

where

$$\zeta := (x,z,\theta,\lambda),$$

σ_0, σ_1 and σ_{ψ_i}'s are SOS and $\bar{p}_1$ is a polynomial. For each $i\in\{1,\ldots,n_\psi\}$ this condition implies that $V(x,z,\theta,\lambda)-L\geq 0$ on

$$\begin{aligned} \mathbf{K}_{c,i} = \{(x,z,\theta,\lambda) \mid \hat{g}(x,z,\theta,\lambda)\geq 0,\ \hat{h}(x,z,\theta,\lambda)=0, \\ \psi_i(x,z)\leq 0\}. \end{aligned}$$

Since $\cup_{i=1}^{n_\psi} \mathbf{K}_{c,i} = \mathbf{K}_c$, the condition (5.37) indeed implies (5.30). A sufficient condition for (5.33) is obtained by replacing the left-hand side of (5.37) by $L - V(x, z, \theta, \lambda)$.

Therefore, all constraints of the optimization problems (5.35) and (5.36) can be enforced through sufficient SDP representable constraints. The objective function is linear in the coefficients of the polynomials $\overline{V}$ or $\underline{V}$ and can be evaluated in closed form as long as the moments of the Lebesgue measure on $\mathbf{X}$ are known.

In conclusion, by restricting the class of decision variables in (5.35) and (5.36) to polynomials of a prescribed degree and replacing the nonnegativity constraints by sufficient SDP representable constraints, we have tightened the problems (5.35) and (5.36) to an SDP.

5.3.4 Performance analysis – stochastic setting

A small modification of the developments from the previous section allows us to analyze the performance in a stochastic setting, where (5.17) is replaced by

$$x_{t+1} = f_x(x_t, \mathcal{O}_t(\kappa(\mathbf{K}_{s_t})), w_t), \tag{5.38a}$$

$$z_{t+1} = f_z(z_t, f_y(x_t, v_t)), \tag{5.38b}$$

where $(\mathcal{O}_t)_{t=0}^{\infty}$ is an arbitrary selection oracle sequence (see Definition 2) and $(w_t, v_t)_{t=0}^{\infty}$ is an iid (with respect to time) process and measurement noise with known joint probability distributions $P_{w,v}$, i.e.,

$$\mathbb{P}(w_t \in A, v_t \in B) = P_{w,v}(A \times B)$$

for all Borel sets $A \subset \mathbb{R}^{n_w}$ and $B \subset \mathbb{R}^{n_v}$. We analyze the performance with respect to the cost function

$$\mathcal{C}_{\mathrm{s}}(x_0, z_0) = \mathbb{E}\left\{ L\alpha^{\tau(x_0,z_0)} \} + \sum_{t=0}^{\tau(x_0,z_0)-1} \alpha^t l(x_t, u_t) \right\}, \tag{5.39}$$

where $\tau(x_0, z_0)$ defined in (5.27) is now a random variable and L satisfies (5.28). The expectation in (5.39) is over the realizations of the stochastic process $(w_t, v_t)_{t=0}^{\infty}$. The rationale behind (5.39) is the same as behind (5.26).

The stochastic counterpart to Lemma 5.3.1 reads

Lemma 5.3.3 *If*

$$\overline{V}(x,z) - \alpha\,\mathbb{E}\,\overline{V}(f_x(x,\kappa(\theta),w), f_z(z, f_y(x,v))) - l(x,\kappa(\theta)) \geq 0 \quad \forall\,(x,z,\theta,\lambda) \in \hat{\mathbf{K}}, \tag{5.40}$$

and

$$\overline{V}(x,z) \geq L \quad \forall\,(x,z) \in \mathbf{X}^c, \tag{5.41}$$

then $\overline{V}(x,z) \geq \mathcal{C}_{\mathrm{s}}(x,z)$ *for all* $(x,z) \in \mathbf{X}$,

where $\mathbf{X}^c$ is the complement of $\mathbf{X}$.

Proof: See the Appendix of this chapter. □

The stochastic counterpart to Lemma 5.3.2 reads

Lemma 5.3.4 *If*

$$\underline{V}(x,z) - \alpha\,\mathbb{E}\,\underline{V}(f_x(x,\kappa(\theta),w), f_z(z, f_y(x,v))) - l(x,\kappa(\theta)) \leq 0 \quad \forall\,(x,z,\theta,\lambda) \in \hat{\mathbf{K}}, \tag{5.42}$$

and

$$\underline{V}(x,z) \leq L \quad \forall\,(x,z) \in \mathbf{X}^c, \tag{5.43}$$

then $\underline{V}(x,z) \leq \mathcal{C}_{\mathrm{s}}(x,z)$ *for all* $(x,z) \in \mathbf{X}$.

Proof: Similar to the proof of Lemma 5.3.3. □

Upper and lower bounds on $\mathcal{C}_{\mathrm{s}}$ are then obtained by

$$\begin{array}{ll} \underset{\overline{V}}{\text{minimize}} & \int_{\mathbf{X}} \overline{V}(x,z)\rho(x,z)d(x,z) \\ \text{s.t.} & (5.40), (5.41), \end{array} \tag{5.44}$$

and

$$\begin{array}{ll} \underset{\underline{V}}{\text{maximize}} & \int_{\mathbf{X}} \underline{V}(x,z)\rho(x,z)d(x,z) \\ \text{s.t.} & (5.42), (5.43), \end{array} \tag{5.45}$$

where $\rho(x,z)$ is a given nonnegative weighting function. Polynomial upper and lower bounds on $\mathcal{C}_{\mathrm{s}}$ are obtained by restricting the functions $\overline{V}$ and $\underline{V}$ to polynomials and replacing the nonnegativity constraints by sufficient SDP representable constraints in exactly the same fashion as in the deterministic setting. The expectation in the constraints (5.40) and (5.42) is handled as follows: Given a polynomial $p(x, w, v) = \sum_{\alpha,\beta,\gamma} p_{(\alpha,\beta,\gamma)} x^\alpha w^\beta v^\gamma$ with coefficients $\{p_{(\alpha,\beta,\gamma)}\}$ indexed by multiindices (α, β, γ), we have

$$\begin{aligned}\mathbb{E}\; p(x, w, v) &= \int p(x, w, v) dP_{w,v}(w, v) \\ &= \sum_{\alpha,\beta,\gamma} p_{(\alpha,\beta,\gamma)} x^\alpha \int w^\beta v^\gamma dP_{w,v}(w, v),\end{aligned}$$

where the moments $\int w^\beta v^\gamma dP_{w,v}(w, v)$ are fixed numbers and can be precomputed offline. Hence, the expectation in (5.40) and (5.42) is linear in the decision variables, as required, and is available in closed form provided that the moments of $P_{w,v}$ are known.

Remark 5.3.2 *Note that in problems (5.44) and (5.45) we use only one function* $\overline{V}$ *and* $\underline{V}$ *instead of pairs of functions* $(V, \overline{V})$ *and* $(V, \underline{V})$ *in problems (5.35) and (5.36). Using a pair of functions gives more degrees of freedom and hence smaller conservatism of the upper and lower bounds, but is difficult to use in the stochastic setting because of the need to evaluate the expectation of a function of* (θ^+, λ^+) *which has an unknown dependence on* (w, v)*. In order to overcome this, one would either have to impose additional assumptions or resort to a worst-case approach.*

5.3.5 Robustness analysis – global ℓ_2 gain, ISS

In this section we describe how to analyze performance in a robust setting in terms of the ℓ_2 gains from w and v to a performance output

$$\hat{y} = f_{\hat{y}}(x), \tag{5.46}$$

where $f_{\hat{y}}$ is a polynomial. We assume the same dynamics (5.38) as in Section (5.3.4) but now all that is known about w and v is that they take values in a given (possibly state-dependent) set

$$\mathbf{W}(x, z) = \{(w, v) \in \mathbb{R}^{n_w} \times \mathbb{R}^{n_v} \mid \psi_w(x, z, w, v) \geq 0\}, \tag{5.47}$$

where each component of $\psi_w : \mathbb{R}^{n_x+n_z+n_w+n_v} \to \mathbb{R}^{n_{\psi_w}}$ is a polynomial in (x, z, w, v). Note that we do not a priori assume that the set $\mathbf{W}(x, z)$ is compact.

Defining

$$\begin{aligned}\mathbf{K}_w = \Big\{ & (x, z, \theta, \lambda, w, v, x^+ z^+, \theta^+, \lambda^+, w^+, v^+) \mid \\ & \hat{h}(x, z, \theta, \lambda) = 0,\ \hat{g}(x, z, \theta, \lambda) \geq 0, \\ & \hat{h}(x^+, z^+, \theta^+, \lambda^+) = 0,\ \hat{g}(x^+, z^+, \theta^+, \lambda^+) \geq 0, \\ & \psi_w(x, z, w, v) \geq 0,\ \psi_w(x^+, z^+, w^+, v^+) \geq 0, \\ & x^+ - f_x(x, \kappa(\theta), w) = 0,\ z^+ - f_z(z, f_y(x, v)) = 0 \Big\},\end{aligned} \tag{5.48}$$

and

$$\begin{aligned}\hat{\mathbf{K}}_w = \Big\{ & (x, z, \theta, \lambda, w, v) \mid \hat{h}(x, z, \theta, \lambda) = 0,\ \hat{g}(x, z, \theta, \lambda) \geq 0, \\ & \psi_w(x, z, w, v) \geq 0 \},\end{aligned} \tag{5.49}$$

we can seek a function V such that

$$\begin{aligned}& V(x^+, z^+, \theta^+, \lambda^+, w^+, v^+) - V(x, z, \theta, \lambda, w, v) \leq \\ & \qquad - \|f_{\hat{y}}(x)\|_2^2 + \alpha_w \|w\|_2^2 + \alpha_v \|v\|_2^2 \\ & \forall\ (x, z, \theta, \lambda, w, v, x^+ z^+, \theta^+, \lambda^+, w^+, v^+) \in \mathbf{K}_w,\end{aligned} \tag{5.50}$$

$$V(x, z, \theta, \lambda, w, v) \geq 0 \quad \forall\ (x, z, \theta, \lambda, w, v) \in \hat{\mathbf{K}}_w. \tag{5.51}$$

and

$$V(0, 0, \theta, \lambda, w, v) = 0 \quad \forall\ (0, 0, \theta, \lambda, w, v) \in \hat{\mathbf{K}}_w. \tag{5.52}$$

The following lemma and its immediate corollary links the satisfaction of (5.50), (5.51) and (5.52) to the ℓ_2 gain from w and v to $\hat{y}$.

Lemma 5.3.5 *If V satisfies (5.50), (5.51) for some $\alpha_w \geq 0$ and $\alpha_v \geq 0$, then*

$$\begin{aligned}\sum_{t=0}^{\infty} \|\hat{y}_t\|_2^2 \leq & V(x_0, z_0, \theta_0, \lambda_0, w_0, v_0) \\ & + \alpha_w \sum_{t=0}^{\infty} \|w_t\|_2^2 + \alpha_v \sum_{t=0}^{\infty} \|v_t\|_2^2.\end{aligned} \tag{5.53}$$

Proof: See the Appendix of this Chapter. □

Corollary 5.3.1 *If V satisfies (5.50), (5.51) and (5.52) for some $\alpha_w \geq 0$ and $\alpha_v \geq 0$, then the ℓ_2 gain from w to $\hat{y}$ respectively from v to $\hat{y}$ is bounded by α_w respectively α_v.*

Proof: Follows by setting $(x_0, z_0) = (0, 0)$ and using (5.52) which implies that $V(x_0, z_0, \theta_0, \lambda_0, w_0, v_0) = 0$ in (5.53). □

Minimization of an upper bound on the ℓ_2 gain from w and v to $\hat{y}$ is then achieved by the following optimization problem:

$$\begin{array}{ll} \underset{V,\alpha_w,\alpha_v}{\text{minimize}} & \alpha_w + \gamma\alpha_v \\ \text{s.t.} & (5.50), (5.51), (5.52), \end{array} \tag{5.54}$$

where the parameter $\gamma \geq 0$ trades off the minimization of the ℓ_2 gains from w to $\hat{y}$ and from v to $\hat{y}$.

Remark 5.3.3 *If instead of (5.51) we require $V(x, z, \theta, \lambda, w, v) \geq \|x\|^2$ for all $(x, z, \theta, \lambda, w, v) \in \hat{\mathbf{K}}_w$, then this along with (5.50) implies that the system (5.38) is input-to-state stable (ISS) with respect to the input $(w, v) \in \mathbf{W}(x, z)$.*

Since the sets $\mathbf{K}_w$ and $\hat{\mathbf{K}}_w$ are basic semialgebraic, we can find upper bounds on the ℓ_2 gains α_w and α_v by restricting V to be a polynomial of a prescribed degree and by replacing the nonnegaivity constraints of (5.54) by sufficient SDP representable constraints (e.g., by inclusions to truncated quadratic modules as in Section 2.2.1). By doing so we immediately obtain a tractable SDP tightening of (5.54).

5.4 Numerical examples

This section illustrates the approach on two numerical examples. The SOS problems were modeled using SOSOPT [Sei10] and solved using MOSEK. For both examples we report the parsing time of SOSOPT, the time to carry out monomial reduction by SOSOPT and the solve time of MOSEK. The bottleneck of the approach is the monomial reduction phase (which, however, is very effective in the sense of reducing the size of the problem significantly). It is expected that a more efficient implementation of the reduction phase and polynomial handling in general would allow the approach to scale much beyond what is presented here. The reduction is carried out in part automatically by SOSOPT and in part by selecting a given subset of variables whose monomials are then used in the basis of SOS multipliers (see Section 6.1)

5.4.1 Bilinear system + PI with saturation – performance analysis

First we demonstrate the approach on a bilinear dynamical system

$$\begin{aligned} x^+ &= f_x(x,u) := \begin{bmatrix} 0.9x_1 + u + 0.2ux_1 \\ 0.85x_2 + x_1 \end{bmatrix} \\ y &= f_y(x) := x_2 \end{aligned}$$

controlled by a PI controller with input saturation given by

$$\begin{aligned} z^+ &= f_z(z,y) := z - k_i y \\ s &= f_s(z,y) := k_p(z-y) \end{aligned}$$

with $k_p = 0.05$, $k_i = 0.02$. The control input is given by saturating u on the input constraint set $\mathbf{U} = [-0.5, 0.5]$, i.e, $u = \mathrm{proj}_{\mathbf{U}}(s)$. In addition the system is subject to the state constraints $\|x\|_\infty \le 10$. In view of Section 5.2.1, this set up can be analyzed using the presented method. The goal is to estimate the performance of this closed-loop system with respect to the cost function (5.26) with $l(x,u) = \|x\|^2 + u^2$, $\alpha = 0.95$ and $L = (2 \cdot 10^2 + 0.5^2)/(1-\alpha) = 4.05 \cdot 10^3$ chosen according to (5.28). We estimate the performance using the optimization problem (5.35), where we consider V as function of (x,z) only and therefore do not need the upper bounding function $\overline{V}$. Assume that we are interested only in closed-loop performance for initial conditions starting from $\mathbf{X}' = \{x \mid \|x\|_\infty \le 1\}$ and $z = 0$ (i.e., zero integral component at the beginning of the closed-loop evolution), which is a strict subset of the set $\mathbf{X} := \{(x,z) \mid \|x\|_\infty \le 10, z \in \mathbb{R}\}$. To this effect we minimize $\int_{\mathbf{X}'} V(x,0)\,dx$ as the objective of (5.35), which corresponds to setting $\rho(x,z) = I_{\mathbf{X}'}(x)\delta_0(z)$, where δ_0 is the Dirac distribution centered at zero. We compare the upper bound obtained by solving (5.35) with the exact cost function evaluated on a dense grid of initial conditions in $\mathbf{X}'$ by forward simulation of the closed-loop system. The comparison is in Figure (5.2); we see a relatively good fit over the whole region of interest $\mathbf{X}'$. The constraints (5.29) and (5.30) of (5.35) were replaced with sufficient SOS conditions with SOS multipliers of degree four containing only monomials in (x,z) and polynomial multipliers of degree three containing monomials in (x,z,θ,λ).

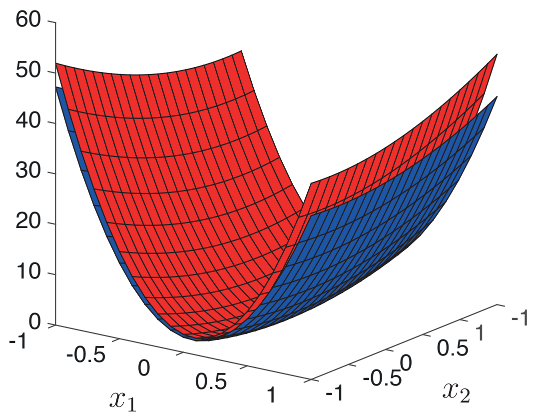


Figure 5.2 – Bilinear system performance bound – Red: upper bound $V(x,0)$ of degree 6. Blue: true closed-loop cost $J(x,0)$.

5.4.2 Uncertain linear system – global asymptotic stability

Consider the Quanser active suspension model in continuous-time $\dot{x} = A_c x + B_c u$ with

$$A_c = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -K_s/M_s & -B_s/M_s & 0 & B_s/M_s \\ 0 & 0 & 0 & 1 \\ K_s/M_{us} & B_s/M_{us} & -K_{us}/M_{us} & -(B_s+B_{us})/M_{us} \end{bmatrix},$$

$$B_c = \begin{bmatrix} 0 & 1/M_s & 0 & -1/M_{us} \end{bmatrix}^T,$$

where $K_s = 1205$, $K_{us} = 2737$, $M_{us} = 1.5$, $B_s = 20$, $B_{us} = 20$ and the mass M_s is unknown and possibly time-varying in the interval $[2.85, 4]$. After discretization[2] with sampling period 0.01, this model can be written as $x^+ = (A_0 + A_1 w)x + (B_0 + B_1 w)x$, where $w := 1/M_s \in [1/4, 1/2.85]$.

We consider an MPC controller (with perfect state measurement) with cost function given by matrices $Q = I$ and $R = 20$ minimized over prediction horizon N subject to input constraint $|u| \leq 250$ and nominal dynamics $x^+ = A_0 x + B_0 u$. This problem is expressed in a dense form (i.e., the state is eliminated using the dynamics equation) to which we apply M steps of the projected gradient method (5.14) (see Section 5.2.4) initialized with the LQ solution and seek a quadratic ISS Lyapunov function V (see Remark 5.3.3) while minimizing the ℓ_2 gain α_w using the optimization problem (5.54) (with $\alpha_v = 0$). The problem (5.54) is feasible (for all combinations of M and N tested) when we take the SOS multipliers σ_1,

[2]The matrices A_0, A_1, B_0, B_1 were found as a least-squares fit of the continuous-time dynamics discretized on a grid of values of $w \in [1/4, 1/2.85]$.

σ_2 in equation (5.21a) of degree two in (x, θ) and the polynomial multipliers p_1, p_2 of degree one in (x, θ, λ). The list of monomials $r(x, \theta, \lambda, w)$ constituting the multiplier σ_0 in the Gram matrix form $\sigma_0 = r(x, \theta, \lambda, w)^T W r(x, \theta, \lambda, w)$, $W \succeq 0$, is determined automatically by SOSOPT and contains monomials linear in λ, x, θ and w, and products $x \cdot w$ and $\theta \cdot w$. In Eq. (5.21b) we set all multipliers to zero except for $\bar{\sigma}_0$, monomials of which are again determined automatically by SOSOPT. Determining the smallest list of monomials $r(x, \theta, \lambda, w)$ takes the most time of the whole procedure; this is documented by Table 5.1 reporting the time breakdown for different values of N and M. The optimal ℓ_2 gain α_w is equal to zero, showing closed-loop global robust asymptotic stability (i.e., convergence $\|x_k\| \to 0$ for any sequence $\{w_k \in [1/4, 1/2.85]\}_{k=0}^{\infty}$). Figure 5.3 shows a sample trajectory of $\|x_k\|$, $V(x_k)$ (the Lyapunov function is a function of x only in this case) and u_k and w_k for $N = M = 4$.

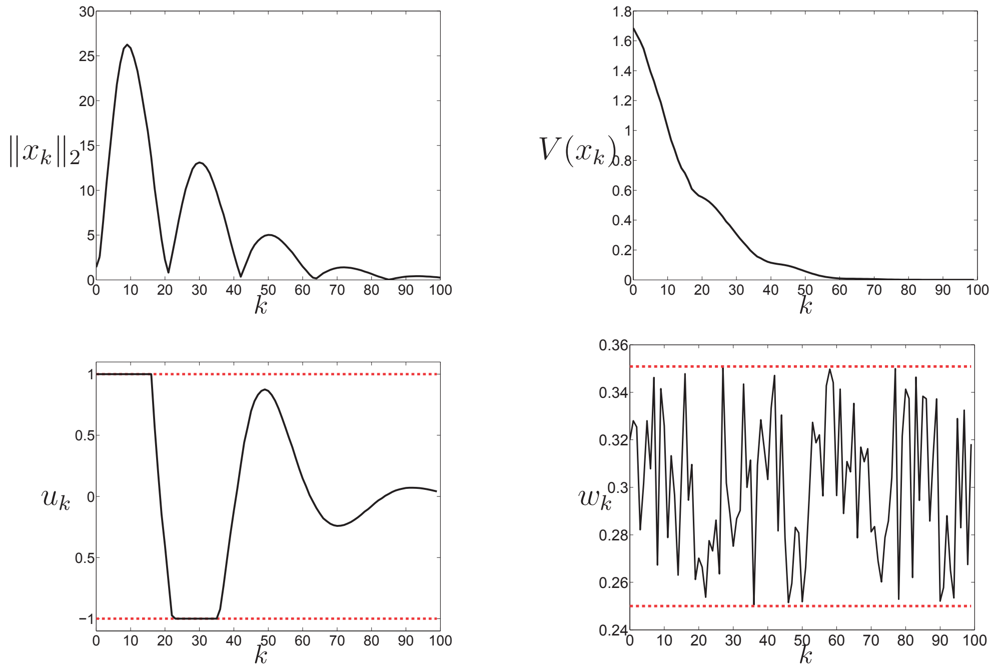


Figure 5.3 – Global asymptotic stability of an uncertain system – one trajectory starting from the initial condition $x_0 = [1\ 1\ 1\ 1]^T$ of the norm of the state $\|x_k\|$, the Lyapunov function $V(x_k)$, the control input u_k and the disturbance w_k.

5.4.3 Stability of a quadcopter on a given subset

This example investigates stability of a linearized attitude and vertical velocity model of a quadcopter. The system has seven states (Roll, Pitch and Yaw angles

Table 5.1 – Global asymptotic stability of an uncertain system – timing breakdown as a function of the number of iterations of the projected gradient method M and the horizon length N used in the cost function. The parsing and monomial reduction was carried out by SOSOPT; the SDP solve by MOSEK.

	parsing	monomial reduction	SDP solve
$M, N = 1, 1$	0.99 s	1.12 s	0.45 s
$M, N = 2, 2$	1.56 s	5.9 s	0.5 s
$M, N = 3, 3$	4.85 s	65.5 s	1.95 s
$M, N = 4, 4$	45 s	755 s	13.9 s

and angular velocities, and velocity in the vertical direction) and four control inputs (the thrusts of the four rotors). The system is controlled by a one-step MPC controller (with perfect state information) which at time k approximately minimizes the cost $x_k^T Q x_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}$, where $Q = I$, $R = 10I$ and P is the infinite-time LQ matrix associated to the cost matrices Q and R, using one step of the projected gradient method (5.14) subject to the input constraints $\|u\|_\infty \leq 1$. This model is open-loop unstable and therefore we investigate closed-loop stability in the region $\mathbf{X} = [-1, 1]^7$ as described in Section 5.3.2. The SOS problem (5.22) is feasible when seeking a quadratic Lyapunov function using SOS multipliers σ_1, σ_2 in equation (5.21a) of degree two in x and the polynomial multipliers p_1, p_2 of degree one in (x, θ, λ). The smallest set of monomials constituting σ_0 is chosen automatically by SOSOPT. In (5.21b), we chose all multipliers zero except for $\bar{\sigma}_0$ whose monomials are determined automatically by SOSOPT. Computing the largest γ such that $\{x \mid V(x) \leq \gamma\}$ is included in $\mathbf{X}$ yields $\gamma = 6.37$; this proves that all trajectories starting in $\{x \mid V(x) \leq \gamma\}$ stay there and converge to the origin. One closed-loop trajectory of $\|x\|_2$, $V(x)$ and u are depicted in Figure 5.4; note that this trajectory does not start in $\{x \mid V(x) \leq \gamma\}$ but still converges to the origin and the Lyapunov function decreases. The parsing time and monomial reduction carried out by SOSOPT took 2.7 s and 16.2 s, respectively; the MOSEK solve time was 0.55 s.

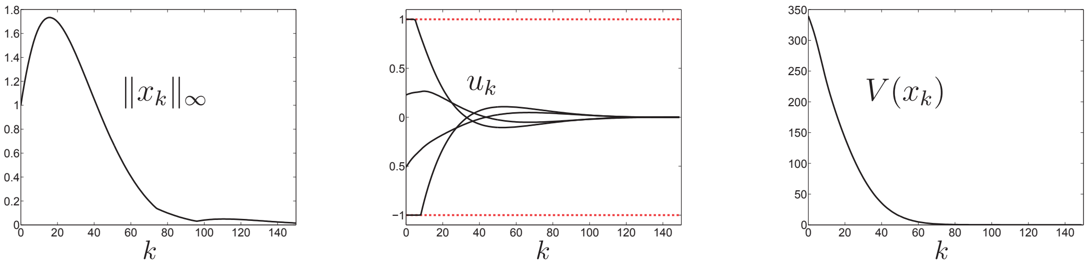


Figure 5.4 – Stability of a quadcopter – trajectories of the norm of the state $\|x_k\|$, the Lyapunov function $V(x_k)$, the control input u_k for initial condition $x_0 = [1\ 1\ 1\ 1\ 1\ 1\ 1]^\top$.

5.5 Appendix

Proof of Lemma 5.3.1

Let (x_t, z_t) be a solution to (5.17) for $t \in \{0, 1, \ldots\}$ and let $u_t \in \kappa(\mathbf{K}_s)$. Then there exist $\theta_t \in \mathbb{R}^{n_\theta}$ and $\lambda_t \in \mathbb{R}^{n_\lambda}$ such that

$$(x_t, z_t, \theta_t, \lambda_t, x_{t+1}, z_{t+1}, \theta_{t+1}, \lambda_{t+1}) \in \bar{\mathbf{K}} \tag{5.55}$$

and

$$(x_t, z_t, \theta_t, \lambda_t) \in \hat{\mathbf{K}} \tag{5.56}$$

for all $t \in \{0, 1, \ldots, \tau - 1\}$, where $\tau := \tau(x_0, z_0)$ is defined in (5.27) and

$$(x_\tau, z_\tau, \theta_\tau, \lambda_\tau) \in \hat{\mathbf{K}}_{\mathbf{c}}. \tag{5.57}$$

Using (5.29) and (5.55), we conclude that

$$V(x_t, z_t, \theta_t, \lambda_t) - \alpha V(x_{t+1}, z_{t+1}, \theta_{t+1}, \lambda_{t+1}) - l(x_t, u_t) \geq 0$$

for all $t \in \{0, 1, \ldots, \tau - 1\}$. This implies that

$$\alpha^\tau V(x_\tau, z_\tau, \theta_\tau, \lambda_\tau) + \sum_{t=0}^{\tau-1} \alpha^t l(x_t, u_t) \leq V(x_0, z_0, \theta_0, \lambda_0).$$

Using (5.30) and (5.57) we conclude that

$$\mathcal{C}(x_0, z_0) = \alpha^\tau L + \sum_{t=0}^{\tau-1} \alpha^t l(x_t, u_t) \leq V(x_0, z_0, \theta_0, \lambda_0).$$

Using (5.31) and (5.56) we have $V(x_0, z_0, \theta_0, \lambda_0) \leq \overline{V}(x_0, z_0)$ and hence $\overline{V}(x_0, z_0) \geq \mathcal{C}(x_0, z_0)$ as desired. ■

Proof of Lemma 5.3.3

The proof proceeds along similar lines as the deterministic version by decomposing the probability space according to the values of the stopping time τ. On the probability event $\{\tau = k\}$ we get, by iterating the inequality (5.40),

$$\alpha^k V(x_k, z_k) + \mathbb{E}\left\{\sum_{t=0}^{k-1} \alpha^t l(x_t, u_t) \mid \tau = k\right\} \leq V(x_0, z_0).$$

Since $(x_k, z_k) \notin \mathbf{X}$ on $\{\tau = k\}$ we have $V(x_k, z_k) \geq L$ by (5.41) and hence

$$\alpha^k L + \mathbb{E}\{\sum_{t=0}^{k-1} \alpha^t l(x_t, u_t) \mid \tau = k\} \leq V(x_0, z_0)$$

on $\{\tau = k\}$. Summing over k gives the result.

Proof of Lemma 5.3.5

Let (x_t, z_t) be a solution to (5.38) with $(w_t, v_t) \in \mathbf{W}(x_t, z_t)$ for all $t \in \{0, 1, \ldots\}$. Then there exist $\theta_t \in \mathbb{R}^{n_\theta}$ and $\lambda_t \in \mathbb{R}^{n_\lambda}$ such that

$$(x_t, z_t, \theta_t, \lambda_t, w_t, v_t, x_{t+1} z_{t+1}, \theta_{t+1}, \lambda_{t+1}, w_{t+1}, v_{t+1}) \in \mathbf{K}_w$$

and

$$(x_t, z_t, \theta_t, \lambda_t, w_t, v_t,) \in \hat{\mathbf{K}}_w$$

for all $t \in \{0, 1, \ldots\}$. Hence, by (5.50) and (5.51),

$$\begin{aligned} V(x_{t+1}, z_{t+1}, \theta_{t+1}, \lambda_{t+1}, w_{t+1}, v_{t+1}) - V(x_t, z_t, \theta_t, \lambda_t, w_t, v_t) \\ \leq -\|\hat{y}_t\|_2^2 + \alpha_w \|w_t\|_2^2 + \alpha_v \|v_t\|_2^2 \end{aligned} \tag{5.58}$$

and

$$V(x_t, z_t, \theta_t, \lambda_t) \geq 0$$

for all $t \in \{0, 1, \ldots\}$. Iterating (5.58) we obtain

$$\begin{aligned}
&\sum_{t=0}^{T-1} \|\hat{y}_t\|_2^2 \leq -V(x_T, z_T, \theta_T, \lambda_T, w_T, v_T) \\
&+ V(x_0, z_0, \theta_0, \lambda_0, w_0, v_0) + \sum_{t=0}^{T-1} \alpha_w \|w_t\|_2^2 + \alpha_v \|v_t\|_2^2 \\
&\leq V(x_0, z_0, \theta_0, \lambda_0, w_0, v_0) + \sum_{t=0}^{T-1} \alpha_w \|w_t\|_2^2 + \alpha_v \|v_t\|_2^2,
\end{aligned}$$

where we have used the fact that $V(x_T, z_T, \theta_T, \lambda_T) \geq 0$ in the second inequality. Letting T tend to infinity gives the result. ■

Chapter 6

Numerical aspects

In this chapter we briefly discuss computational aspects of the proposed methods. We discuss software implementation as well as computational complexity.

On the side of polynomials, the bulk of computational complexity and software implementation work boils down to imposing the inclusion of a polynomial $p \in \mathbb{R}[x]$, $x \in \mathbb{R}^n$, in a truncated quadratic module $Q_d(\mathbf{K})$ defined in (2.8). As discussed in Section 2.2.1, this is equivalent to the existence of matrices $W_0 \succeq 0, \ldots, W_{n_g} \succeq 0$ such that

$$p(x) = r_{d_0}(x)^\top W_0 r_{d_0}(x) + \sum_{i=1}^{n_g} g_i(x) r_{d_i}(x)^\top W_i r_{d_i}(x), \tag{6.1}$$

where g_i's are the polynomials defining the set $\mathbf{K}$ and $d_0 = \lfloor d/2 \rfloor$ and $d_i = \lfloor (d - \deg g_i)/2 \rfloor$, $i \in \{1, \ldots, n_g\}$.

On the side of moments, the main computational complexity and software implementation work comes down to imposing the inclusion of a vector $\boldsymbol{y} \in \mathbb{R}^{\binom{n+d}{d}}$ in the truncated moment cone $M_d^{\mathrm{sup}}(\mathbf{K})_+$ defined in (2.13). This boils down to constructing the moment and localizing matrices $M_d(\cdot)$ and $M_d(\cdot, g_i)$ defined in Section 2.2.2 and imposing the satisfaction of the constraints

$$M_d(\boldsymbol{y}) \succeq 0,\ M_d(\boldsymbol{y}, g_i) \succeq 0,\ \ i \in \{1, \ldots, n_g\}, \tag{6.2}$$

which is equivalent to semidefinite programming feasibility problem.

6.1 Computational complexity

For computational complexity discussion, we restrict our attention to the polynomial side as the moment side is dual to it and hence the same conclusions apply there.

By and large, the computational complexity is governed by the complexity of imposing the inclusion of a polynomial in a truncated quadratic module $Q_d(\mathbf{K})$ through Eq. (6.1). The size of each matrix W_i, $i \in \{0, \dots, n_g\}$, in (6.1) is equal to $\binom{n+d_i}{d_i} \times \binom{n+d_i}{d_i}$ and hence the size of the largest matrix, W_0, grows asymptotically as $O(n^{d/2})$ when d is held fixed and n varies and grows as $O((d/2)^n)$ when d varies and n is held fixed.

This relatively steep growth in the size of the SDP matrices precludes the use of these methods when the number of variables n or the degree d is large unless special structure is imposed on the matrices W_i. The simplest way to impose structure on W_i is to discard certain monomials from the monomial vector $r_{d_i}(x)$. The simplest way to do so is to restrict the vectors of monomials $r_{d_i}(x)$ to monomials of a certain subset of the variables $(x_1, \dots, x_n)$. This clearly leads to conservatism, i.e., to the shrinkage of the set of polynomials spanned by the right-hand side of (6.1) (i.e., the right-hand side of (6.1) now spans only a subset of $Q_d(\mathbf{K})$). Nevertheless, if the polynomial p and the polynomials g_i possess a certain kind of sparsity structure and this structure is taken into account in the variable reduction process, the asymptotic result of Theorem 2.2.1 still holds [Las06].

In addition, there exist automatic monomial reduction techniques (e.g., the Newton polytope [Stu98]) which discard those monomials in $r(x)$ which cannot appear in the decomposition of $p(x)$. It is also possible to directly reduce the size of the SDP having (6.1) as a constraint using a facial reduction algorithm, e.g., [PP14]. The benefit of facial reduction is that it works at the SDP level and is therefore not limited to the situation where $r(x)$ is a subset of the monomial basis, allowing one to use numerically better conditioned bases (e.g., the Chebyshev basis) while still benefiting from a possible size reduction. Importantly, either of these reduction techniques incurs no conservatism, i.e., the set of polynomials spanned by the right-hand side of (6.1) remains unaltered.

Besides monomial reduction, one can also replace the constraint $W_i \succeq 0$ by a sufficient and less computationally demanding constraint. This is the idea of [AM14], where the authors propose to replace the constraint $W_i \succeq 0$ either by the constraint that W_i is diagonally dominant or scaled diagonally dominant. The former replacement leads to a set of linear inequality constraints and the latter to a set of second order cone constraints. This is clearly a conservative replacement but the hope is that the simpler constraint and the higher degree of maturity of linear programming or second order cone solvers (compared to SDP solvers) can outweigh this conservatism by allowing one to solve the problem for much higher values of the degree d.

In the numerical examples accompanying Chapters 3 and 4, we observed little benefit in using any of the above mentioned complexity reduction techniques. This is due to the fact that in these cases, the polynomial p in (6.1) is itself a variable (e.g., a linear transformation of the polynomials w and v in (3.9)) and there is no apparent structure to exploit in the dynamical systems used (or their constraints) in those examples. On the other hand, the results of Chapter 5 benefited strongly from reduction techniques due the large amount of structure of the KKT systems involved.

6.2 Software implementation

Fortunately, user friendly software packages exist both for manipulating polynomials and imposing their inclusion in truncated quadratic modules as well as for manipulating truncated moment sequences and imposing their inclusion in truncated moment cones.

These software packages parse a high-level code, call an underlying SDP solver and parse its output. Hence, implementing any of the methods presented in this thesis is straightforward, necessitating only a few lines of code and a call to one of these software packages.

On the polynomial side, the list of available packages is very long. Let us mention Yalmip [Löf04], SOSOPT [Sei10] and SOSTOOLS [PAGV$^+$13], to name just a few. All of these packages provide similar functionality. On the moment side, the only available package, to our knowledge, is Gloptipoly 3 [HLL09].

As for the underlying SDP solver, there we used MOSEK or SeDuMi [PTZ07] throughout the thesis. These are both primal-dual interior point solvers suitable for medium-size problems. In order for the methods of this thesis to scale higher (without any structure exploitation), first-order or augmented Lagrangian methods (most notably SDPNAL [ZST10]) need to be employed. These methods, however, are more sensitive to numerical conditioning and have, so far, not shown particularly promising results on the class of problems studied here. One exception is the work [Nie12] where SDPNAL was applied iteratively, with solution of one iteration used to construct a preconditioner for the subsequent iteration and reported promising results for the case of static polynomial optimization. However, the question of preconditioning general moment-sum-of-squares problems (e.g., those encountered in this thesis) remains largely open.

Chapter 7

Conclusion and outlook

This thesis used the idea of lifting a nonlinear problem into an equivalent infinite-dimensional problem and subsequently approximating this problem by a sequence of linear (or conic) finite-dimensional problems. We have demonstrated the efficiency of this approach on several problems from the field of nonlinear control and dynamical systems, providing both novel theoretical results and easy-to-use numerical methods.

The main virtue of the proposed methods is in their simplicity and generality: The results are obtained as the solution to a single convex semidefinite programming problem with no initialization and no ad hoc tuning parameters and with the only assumption that the data defining the problem are polynomial.

The price to pay for this generality is scalability of the approach, where, currently, the proposed methods are applicable to small and moderate-size systems only, although applicability to larger systems is expected as numerical methods for semidefinite programming progress and/or as a structure of a particular class of problems is discovered and computationally exploited. First application of the methods of this thesis to larger systems from robotics were already reported [MVTT14].

The particular lifting (i.e., the Liouville's equation) and the particular way of approximating the infinite-dimensional problem (i.e., the moment-sum-of-squares hierarchy) are by no means the only way of carrying out the lift-plus-approximate procedure and it is entirely possible that different implementations of this procedure may be more suited for the class of problems studied in this thesis. This is a subject of future research.

Appendix A

Mathematical background

A.1 Stochastic kernels and disintegration

In this appendix we discuss stochastic kernels and disintegration, which play an important role in theoretical results of this thesis.

Let $\mathbf{Y} \subset \mathbb{R}^n$ and $\mathbf{Z} \subset \mathbb{R}^m$ be two Borel sets. The object $\nu(\cdot\,|\,\cdot)$ is called a stochastic kernel on $\mathbf{Y}$ given $\mathbf{Z}$ if, first, $\nu(\cdot\,|\,z)$ is a probability measure on $\mathbf{Y}$ for every fixed $z \in \mathbf{Z}$, and, second, if $\nu(A\,|\,\cdot)$ is a measurable function on $\mathbf{Z}$ for every Borel $A \subset \mathbf{Y}$.

Stochastic kernels appear when *disintegrating* a measure $\mu \in M(\mathbf{Y} \times \mathbf{Z})_+$ as $d\mu(y, z) = d\nu(y\,|\,z)d\bar{\mu}(z)$ by which we mean

$$\int_{\mathbf{Y}\times\mathbf{Z}} g(y,z)\,d\mu(y,z) = \int_{\mathbf{Z}}\int_{\mathbf{Y}} g(y,z)\,d\nu(y\,|\,z)\,d\bar{\mu}(z)$$

for all μ-integrable functions g, where $\nu(\cdot\,|\,\cdot)$ is a stochastic kernel on $\mathbf{Y}$ given $\mathbf{Z}$ and $\bar{\mu} \in M(\mathbf{Z})_+$ is the z-marginal of μ, i.e.,

$$\bar{\mu}(A) = \mu(\mathbf{Y} \times A).$$

for all Borel $A \subset \mathbf{Z}$. The existence of such disintegration is guaranteed by [Bog06, Corollary 10.4.13].

Intuitively, the probability measure $\nu(\cdot\,|\,z)$ can be thought of as the conditional distribution, given z, of a random variable defined on $\mathbf{Y} \times \mathbf{Z}$ and having the joint distribution μ, although we do not assume that the mass of μ is normalized to one.

A.2 Infinite-dimensional linear programming

In this appendix we give a brief overview of infinite-dimensional linear programming. The material is taken from [HLL96, Chapter 6].

Let $\mathbf{X}$ and $\mathbf{Y}$ be two real vector spaces and let a bilinear form $\langle \cdot\,, \cdot\rangle_{\mathbf{X},\mathbf{Y}} : \mathbf{X}\times\mathbf{Y} \to \mathbb{R}$ be given satisfying:

- For all $x \in \mathbf{X}$ there exists $y \in \mathbf{Y}$ such that $\langle x, y\rangle \neq 0$,
- For all $y \in \mathbf{Y}$ there exists $x \in \mathbf{X}$ such that $\langle x, y\rangle \neq 0$.

Then we call $(\mathbf{X}, \mathbf{Y})$ a *dual pair* with the duality pairing $\langle \cdot\,, \cdot\rangle_{\mathbf{X},\mathbf{Y}}$. We endow $\mathbf{X}$ with the *weak topology*, which is the coarsest topology that makes $\langle \cdot, y\rangle_{\mathbf{X},\mathbf{Y}}$ continuous on $\mathbf{X}$ for all $y \in \mathbf{Y}$. Then $\mathbf{Y}$ is the topological dual of $\mathbf{X}$ w.r.t. the weak topology, i.e., the space of all linear functionals on $\mathbf{X}$ continuous w.r.t the weak topology.

Example *A typical example of these spaces encountered throughout the thesis is* $\mathbf{X} = M(\mathbf{K})$ *and* $\mathbf{Y} = C(\mathbf{K})$ *with the duality pairing given by*

$$\langle f, \mu\rangle = \int_{\mathbf{K}} f\, d\mu$$

for all $f \in C(\mathbf{K})$ *and* $\mu \in M(\mathbf{K})$*. The weak topology on* $\mathbf{X} = M(\mathbf{K})$ *as just defined then coincides with the weak-** *topology on* $M(\mathbf{K})$*.*

Let $(\mathbf{Z}, \mathbf{W})$ be another dual pair of vector spaces with duality pairing $\langle \cdot\,, \cdot\rangle_{\mathbf{Z},\mathbf{W}}$ and let a continuous (w.r.t. the weak topologies on $\mathbf{X}$ and $\mathbf{Z}$) operator $\mathcal{A} : \mathbf{X} \to \mathbf{Z}$ be given along with its adjoint $\mathcal{A}^\star : \mathbf{W} \to \mathbf{Y}$ satisfying

$$\langle \mathcal{A}x, w\rangle_{\mathbf{Z},\mathbf{W}} = \langle x, \mathcal{A}^\star w\rangle_{\mathbf{X},\mathbf{Y}}$$

for all $x \in \mathbf{X}$ and $w \in \mathbf{W}$.

Let also a convex cone $C \subset \mathbf{X}$ satisfying $0 \in C$ be given as well as two vectors $c \in \mathbf{Y}$ and $b \in \mathbf{Z}$. Then a standard-form primal infinite-dimensional LP reads

$$\begin{array}{rcll} p^\star & = & \inf\limits_{x\in\mathbf{X}} & \langle x, c\rangle_{\mathbf{X},\mathbf{Y}} \\ & & \text{s.t.} & \mathcal{A}x = b \\ & & & x \in C. \end{array} \tag{A.1}$$

The dual LP reads

$$\begin{array}{rcll} d^\star & = & \sup\limits_{w \in \mathbf{W}} & \langle b, w \rangle_{\mathbf{Z},\mathbf{W}} \\ & & \text{s.t.} & c - \mathcal{A}^\star w \in C^\star, \end{array} \tag{A.2}$$

where

$$C^\star := \{ y \in \mathbf{Y} \mid \langle x, y \rangle_{\mathbf{X},\mathbf{Y}} \geq 0 \; \forall \, x \in \mathbf{X} \}$$

is the dual cone to C.

We remark that this setting is very general and encompasses all of the finite-dimensional conic programs, e.g., linear programs, second-order cone programs, semidefinite programs, exponential cone programs etc. as well as all the infinite-dimensional LPs encountered in this thesis.

As is the case for finite-dimensional LPs, the following holds for the pair of infinite-dimensional LPs (A.1) and (A.2)

- (**Weak duality**) If (A.1) and (A.2) are feasible, then $p^\star \leq d^\star$.
- (**Complementarity**) If x is feasible for (A.1) and w is feasible for (A.2) and $\langle x, c - \mathcal{A}^\star w \rangle_{\mathbf{X},\mathbf{Y}} = 0$, then x is optimal in (A.1) and w is optimal in (A.2).

Importantly, the following strong duality theorem holds:

Theorem A.2.1 *If (A.1) is feasible and if the set*

$$\{ (Ax, \langle x, c \rangle_{\mathbf{X},\mathbf{Y}}) \mid x \in C \} \subset \mathbf{Z} \times \mathbb{R}$$

is closed (in the product topology of the weak topology on $\mathbf{Z}$ *and the standard topology on* $\mathbb{R}$*), then the optimal value in (A.1) is attained and* $p^\star = d^\star$*.*

Proof: Theorem 3.10 in [AN87]. □

Appendix B

Superposition theorems

The goal of this Appendix is to prove the superposition Theorems 2.3.1, 2.3.2. For convenience we restate each theorem here.

B.1 Proof of Theorem 2.3.1

Theorem 2.3.1 *If a triplet of nonnegative compactly supported finite measures*

$$(\mu, \mu_0, \mu_T) \in M([0,T] \times \mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+$$

satisfies

$$\int_{\mathbb{R}^n} v(T,x)\, d\mu_T(x) = \int_{\mathbb{R}^n} v(0,x)\, d\mu_0(x) + \int\limits_{[0,T]\times\mathbb{R}^n\times\mathbf{U}} \frac{\partial v}{\partial t} + \nabla v(t,x) \cdot f(x,u)\, d\mu(t,x,u) \tag{B.1}$$

for all $v \in C^1([0,T] \times \mathbb{R}^n)$*, then there exists a measure*

$$\eta \in M(C([0,T];\mathbb{R}^n))_+$$

supported on the absolutely continuous trajectories of

$$\dot{x}(t) \in \mathrm{conv}\, f(x(t), \mathbf{U}) \tag{B.2}$$

such that

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

$$\mu_0(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(0))\, d\eta(x(\cdot))$$

and

$$\mu_T(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(T))\, d\eta(x(\cdot))$$

for all $A \subset \mathbb{R}$ *and* $B \subset \mathbb{R}^n$.

Proof: Let $\mathbf{X}$ denote a compact set such that $\mathrm{spt}\mu \subset [0,T] \times \mathbf{X} \times \mathbf{U}$, $\mathrm{spt}\mu_0 \subset \mathbf{X}$ and $\mathrm{spt}\mu_T \subset \mathbf{X}$ and let $\bar{\mu}$ denote the (t,x) marginal of μ, i.e.,

$$\bar{\mu}(A \times B) = \mu(A \times B \times \mathbf{U})$$

for all $A \subset [0,T]$ and $B \subset \mathbf{X}$.

We start by disintegrating (see Section A.1) the measure μ 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 $\mathbf{U}$ given $[0,T] \times \mathbf{X}$. Then we can rewrite equation (B.1) as

$$\begin{aligned}
&\int_{\mathbf{X}} v(T,\cdot)\, d\mu_T - \int_X v(0,\cdot)\, d\mu_0 \\
&= \int_{[0,T]\times X} \int_{\mathbf{U}} \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot f(t,x,u)\, d\nu(u\,|\,t,x)\, d\bar{\mu}(t,x) \\
&= \int_{[0,T]\times \mathbf{X}} \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot \left[\int_U f(t,x,u)\, d\nu(u\,|\,t,x) \right] d\bar{\mu}(t,x) \\
&= \int_{[0,T]\times \mathbf{X}} \frac{\partial v}{\partial t} + \operatorname{grad} v \cdot \bar{f}(t,x)\, d\bar{\mu}(t,x), \qquad \text{(B.3)}
\end{aligned}$$

where

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

Therefore we will study the trajectories of the differential equation

$$\dot{x}(t) = \bar{f}(t,x(t)). \qquad \text{(B.4)}$$

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 (B.2)) starting from μ_0. Note that the vector field $\bar{f}$ is only known to be measurable[1], so this equation may not admit a unique solution.

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

$$d\bar{\mu}(t,x) = d\mu_t(x)dt, \tag{B.5}$$

where $d\mu_t(x)$ is a stochastic kernel on $\mathbf{X}$ 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[2] 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 $t \mapsto \int_{\mathbf{X}} w(x)\, d\mu_t(x)$ is absolutely continuous for all $w \in C^1(\mathbf{X})$ and such that the continuity equation

$$\frac{d}{dt}\int_{\mathbf{X}} w(x)\, d\mu_t(x) = \int_{\mathbf{X}} \operatorname{grad} w(x) \cdot \bar{f}(t,x)\, d\mu_t(x) \;\; \forall\, w \in C^1(\mathbf{X}) \tag{B.6}$$

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

Fix $w \in C^1(\mathbf{X})$ and define the test function $v(t,x) := \psi(t)w(x)$, where $\psi \in$

[1]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.

[2]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.

$C^1([0,T])$. Then from equation (B.3)

$$\begin{aligned}
&\psi(T)\int_{\mathbf{X}} w\,d\mu_T - \psi(0)\int_X w\,d\mu_0\\
&\quad= \int_{[0,T]\times\mathbf{X}} \frac{\partial(\psi w)}{\partial t} + \operatorname{grad}(\psi w)\cdot \bar{f}(t,x)\,d\bar{\mu}(t,x)\\
&\quad= \int_0^T\int_X \dot{\psi}(t)w(x) + \psi(t)\operatorname{grad} w(x)\cdot \bar{f}(t,x)\,d\mu_t(x)dt\\
&\quad= \int_0^T \left[\dot{\psi}\int_X w\,d\mu_t + \psi\int_{\mathbf{X}} \operatorname{grad} w\cdot \bar{f}\,d\mu_t\right]dt,
\end{aligned}$$

which can be seen as an equation of the form

$$\psi(T)d\ -\ \psi(0)c = \int_0^T \dot{\psi}(t)a(t) + \psi(t)b(t)\,dt\ \ \forall\psi\in C^1([0,T]), \tag{B.7}$$

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 $a(t) = \int_{\mathbf{X}} w\,d\mu_t$. 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 (B.3) with v replaced by w we have $\tilde{a}(T) = \int_X w\,d\mu_T = d$, integration by parts gives

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

so $\tilde{a}(t)$ indeed solves equation (B.7). 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 (B.7) 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 $C^1(\mathbf{X})$ is separable,

$$\int_{\mathbf{X}} w\,d\mu_t = \int_{\mathbf{X}} w(x)d\mu_0 + \int_0^t \int_X \operatorname{grad} w \cdot \bar{f}\,d\mu_\tau d\tau \;\; \forall\,w \in C^1(\mathbf{X}) \tag{B.8}$$

dt-almost everywhere. The right-hand side of this equality is an absolutely continuous function of time for each $w \in C^1(\mathbf{X})$ and the left-hand side is a bounded positive linear functional on $C(\mathbf{X})$ for all $t \in [0,T]$. By continuity of the right-hand-side of (B.8) with respect to time, this right-hand side is a bounded positive linear functional on $C^1(\mathbf{X})$ for all $t \in [0,T]$ and therefore can be uniquely extended to a bounded positive linear functional on $C(\mathbf{X})$ (since $C^1(\mathbf{X})$ is dense in $C(\mathbf{X})$). Therefore, for all $t \in [0,T]$ the right-hand side has a representing measure [Rud86, Theorem 2.14] and hence there is a version of μ_t such that the equality (B.8) holds for *all* $t \in [0,T]$. With this version of μ_t the function $t \mapsto \int_{\mathbf{X}} w(x)\,d\mu_t(x)$ is absolutely continuous and μ_t solves the continuity equation (B.6).

To finish the proof, we use [Amb08, Theorem 3.2] which asserts the existence of a nonnegative measure η on $C([0,T];\mathbb{R}^n)$ which corresponds to a family of absolutely continuous solutions to ODE (B.4) whose projection at each time $t \in [0,T]$ coincides with μ_t. More precisely, there is a nonnegative measure $\eta \in M(C([0,T];\mathbb{R}^n))_+$ supported on a family of absolutely continuous solutions to ODE (B.4) such that for all measurable $w : \mathbb{R}^n \to \mathbb{R}$

$$\int_{\mathbf{X}} w(x)\mu_t(x) = \int_{C([0,T];\,\mathbb{R}^n)} w(x(t))\,d\eta(x(\cdot)) \quad \forall\,t \in [0,T]. \tag{B.9}$$

Using $I_{A\times B}(t,x) = I_A(t)I_B(x)$, it follows from (B.5) that

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

Therefore, using (B.9) with $w(x) = I_B(x)$ and Fubini's theorem [Rud86, Theorem 8.8], we get

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

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.

□

B.2 Proof of Theorem 2.3.2

Theorem 2.3.2 *If a pair of nonnegative compactly supported finite measures*

$$(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$$

satisfies

$$\beta \int_{\mathbb{R}^n \times \mathbf{U}} v(x)\, d\mu(x,u) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \int_{\mathbb{R}^n \times \mathbf{U}} \nabla v(x) \cdot f(x,u)\, d\mu(x,u) \quad \text{(B.10)}$$

for all $v \in C_b^1(\mathbb{R}^n)$, then there exists a measure

$$\eta \in M(C([0,\infty); \mathbb{R}^n))_+$$

supported on the absolutely continuous trajectories of

$$\dot{x}(t) \in \operatorname{conv} f(x(t), \mathbf{U}) \quad \text{(B.11)}$$

such that

$$\mu(B \times \mathbf{U}) = \int_{C([0,\infty);\mathbb{R}^n)} \int_0^\infty e^{-\beta t} I_B(x(\cdot))\, dt\, d\eta(x(\cdot)),$$

and

$$\mu_0(B) = \int_{C([0,T];\mathbb{R}^n)} I_B(x(0))\, d\eta(x(\cdot))$$

for all $B \subset \mathbb{R}^n$.

Proof: The proof is based on fundamental results of [Amb08] and [BB96] and on the compactification procedure discussed in [Kur11].

Let $\mathbf{X}$ denote a compact set such that spt$\mu \subset \mathbf{X} \times \mathbf{U}$, spt$\mu_0 \subset \mathbf{X}$. We begin by embedding the problem in a stochastic setting. To this end, define the extended state space E as the one-point compactification of $\mathbb{R}^n$, i.e., $E = \mathbb{R}^n \cup \{\Delta\}$, where Δ is the point compactifying $\mathbb{R}^n$. Define also the linear operator $A : \mathcal{D}(A) \to C(E \times \mathbf{U})$

by

$$w \mapsto Aw := \mathrm{grad}\, w \cdot f,$$

where the domain of A, $\mathcal{D}(A)$, is defined as

$$\mathcal{D}(A) := \{w : E \to \mathbb{R} \mid w \in C^1(\mathbb{R}^n),\ w(\Delta) = 0, \lim_{x\to\Delta} w(x) = 0, \\ \lim_{x\to\Delta} \mathrm{grad}\, w \cdot f(x,u) = 0\ \forall\, u \in \mathbf{U}\}.$$

In words, $\mathcal{D}(A)$ is the space all continuously differentiable functions vanishing at infinity such that $\mathrm{grad}\, w \cdot f$ also vanishes at infinity for all $u \in \mathbf{U}$. Now consider the relaxed martingale problem [BB96]: find a stochastic process $Y : [0,\infty] \times \Omega \to E$ defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and a stochastic kernel $\nu(\cdot\,|\,\cdot)$ (stationary relaxed Markov control) on $\mathbf{U}$ given E such that

- $P(Y(0) \in A) = \mu_0(A) \quad \forall\, A \subset E$

- for all $w \in \mathcal{D}(A)$ the stochastic process

$$w(Y(t)) - \int_0^t \int_{\mathbf{U}} Aw(Y(s),u)\, \nu(du\,|\,Y(s))\, ds \tag{B.12}$$

  is an $\mathcal{F}_t$-martingale (see, e.g., [Kal10] for a definition).

Observe that there exists a countable subset of $\mathcal{D}(A)$ (e.g., all polynomials with rational coefficients attenuated near infinity) dense in $\mathcal{D}(A)$ in the supremum norm. Next, $\mathcal{D}(A)$ is clearly an algebra that separates points of E and $A1 = 0$. Finally, since $f(x,u)$ is polynomial and hence locally Lipschitz, the ODE $\dot{x} = f(x,u)$ has a solution on $[0,\infty)$ for any $x_0 \in E$ and any fixed $u \in \mathbf{U}$ in the sense that if there is a finite escape time t_e, then we define $x(t) = \Delta$ for all $t \geq t_e$. Each such solution satisfies the martingale relation (B.12) (with a trivial probability space). Therefore, A satisfies Conditions 1-3 of [BB96] and it follows from Theorem 2.2 and Corollary 2.2 therein that for any pair of measures satisfying the discounted Liouville's equation (B.10), there exists a solution to the above martingale problem whose discounted occupation measure is equal to μ, that is,

$$\mu(A \times B) = \mathbb{E}\left\{ \int_0^\infty e^{-\beta t} I_{A\times B}(Y(t),u)\, \nu(du\,|\,Y(t))\, dt \right\}, \quad P(Y(0) \in A) = \mu_0(A),$$

where $\mathbb{E}$ denotes the expectation w.r.t. the probability measure P. From the martingale property of (B.12) and the definition of A we get

$$\mathbb{E}\{w(Y(t))\} - \mathbb{E}\left\{\int_0^t \int_{\mathbf{U}} \operatorname{grad} w \cdot f(Y(s), u)\, \nu(du\,|\,Y(s))\, ds\right\} = \mathbb{E}\{Y(0)\}.$$

Now let μ_t denote the marginal distribution of $Y(t)$ at time t; that is,

$$\mu_t(A) := P(Y(t) \in A) = \mathbb{E}\{I_A(Y(t))\} \quad \forall\, A \subset \mathbf{X}.$$

Then the above relation becomes

$$\int_{\mathbf{X}} w(x)\, d\mu_t(x) - \int_0^t \int_{\mathbf{X}} \int_{\mathbf{U}} \operatorname{grad} w(x) \cdot f(x, u)\, \nu(du\,|\,x)\, d\mu_s(x)\, ds = \int_{\mathbf{X}} w(x)\, d\mu_0(x),$$

where we have used Fubini's thorem to interchange the expectation operator and integration w.r.t. time. Defining the relaxed vector field

$$\bar{f}(x) = \int_{\mathbf{U}} f(x, u)\, \nu(du\,|\,x) \in \operatorname{conv} f(x, \mathbf{U})$$

and rearranging we obtain

$$\int_{\mathbf{X}} w(x)\, d\mu_t(x) = \int w(x)\, d\mu_0(x) + \int_0^t \int_{\mathbf{X}} \operatorname{grad} w(x) \cdot \bar{f}(x)\, d\mu_s(x)\, ds, \qquad \text{(B.13)}$$

where the equation holds for all $w \in C^1(\mathbf{X})$ almost everywhere with respect to the Lebesgue measure on $[0, \infty)$. The Lemma then follows from Ambrosio's superposition principle [Amb08, Theorem 3.2] using the same arguments as in the proof of Theorem 2.3.1.

□

B.3 Proof of Theorem 2.3.3

Theorem 2.3.3 *Let* $\bar{f} : \mathbb{R}^n \to \mathbb{R}^n$ *be locally Lipschitz and let the triplet of nonnegative compactly supported measures*

$$(\mu, \mu_0, \mu_T) \in M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+ \times M(\mathbb{R}^n)_+$$

satisfy

$$\int_{\mathbb{R}^n} v(x)\, d\mu_T(x) + \beta \int_{\mathbb{R}^n} v(x)\, d\mu(x) = \int_{\mathbb{R}^n} v(x)\, d\mu_0(x) + \int_{\mathbb{R}^n} \nabla v \cdot \bar{f}(x)\, d\mu(x) \quad \text{(B.14)}$$

for all $v \in C_b^1(\mathbb{R}^n)$. *Then there exists an ensemble of probability measures* $\{\tau_{x_0} \in P([0,\infty])\}_{x_0 \in \mathbb{R}^n}$ *(a stochastic kernel) such that*

$$\int_{\mathbb{R}^n} v(x)\, d\mu_0(x) = \int_{\mathbb{R}^n} v(x(0\,|\,x_0))\, d\mu_0(x_0), \quad \text{(B.15a)}$$

$$\int_{\mathbb{R}^n} v(x)\, d\mu(x) = \int_{\mathbb{R}^n} \int_0^\infty \int_0^\tau e^{-\beta t} v(x(t\,|\,x_0))\, dt\, d\tau_{x_0}(\tau)\, d\mu_0(x_0), \quad \text{(B.15b)}$$

$$\int_{\mathbb{R}^n} v(x)\, d\mu_T(x) = \int_{\mathbb{R}^n} \int_0^\infty e^{-\beta\tau} v(\tau(x_0))\, d\tau_{x_0}(\tau)\, d\mu_0(x_0), \quad \text{(B.15c)}$$

for all bounded measurable functions v, *where* $x(\cdot\,|\,x_0)$ *denotes the unique trajectory of* $\dot{x} = \bar{f}(x)$ *starting from* x_0, *which is defined at least on* $[0, \sup \operatorname{spt} \tau_{x_0}]$ *for all* $x_0 \in \operatorname{spt} \mu_0$.

Proof: Let $\mathbf{X}$ denote a compact set such that $\operatorname{spt}\mu \subset \mathbf{X}$, $\operatorname{spt}\mu_0 \subset \mathbf{X}$ and $\operatorname{spt}\mu_T \subset \mathbf{X}$. Since $\mathbf{X}$ is compact we can modify $\bar{f}$ outside of $\mathbf{X}$ such that $\bar{f}$ is *globally* Lipschitz and (B.14) still holds.

Suppose therefore that (B.14) holds with a globally Lipschitz $\bar{f}$. First we will prove a simple result. In the rest of this Appendix we will use the notation C_b^k for the space of all bounded k-times continuously differentiable functions.

Lemma B.3.1 *If* $\beta > 0$, *then for any* $w \in C_b^1(\mathbb{R}^n)$, *the equation*

$$\nabla v \cdot \bar{f} - \beta v = w \quad \text{(B.16)}$$

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

$$v(x_0) = -\int_0^\infty e^{-\beta t} w(x(t\,|\,x_0))\, dt. \quad \text{(B.17)}$$

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 (B.17) is well defined (since w is bounded and $\beta > 0$).

Direct computation then gives:

$$
\begin{aligned}
\nabla v \cdot \bar{f}(x(t \mid x_0)) &= \frac{d}{dt} v(x(t|x_0)) \\
&= -\frac{d}{dt} \int_0^\infty e^{-\beta s} w(x(s|x(t|x_0)))\, ds \\
&= -\frac{d}{dt} \int_0^\infty e^{-\beta s} w(x(t+s|x_0))\, ds \\
&= -\int_0^\infty e^{-\beta s} \nabla w(x(t+s|x_0)) \cdot \bar{f}(x(t+s|x_0))\, ds \\
&= -\int_0^\infty e^{-\beta s} \frac{d}{ds} w(x(t+s|x_0))\, ds \\
&= -\beta \int_0^\infty e^{-\beta s} w(x(t+s|x_0))\, ds - [e^{-\beta s} w(x(t+s|x_0))]_0^\infty \\
&= -\beta \int_0^\infty e^{-\beta s} w(x(s|x(t|x_0)))\, ds + w(x(t|x_0)) \\
&= \beta v(x(t|x_0)) + w(x(t|x_0)).
\end{aligned}
$$

Setting $t = 0$, we arrive at (B.16). □

Proof (of Theorem 2.3.3) 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 (B.14) then there exists a time $\tau \geq 0$ such that $x(\tau|x_0) = x_T$. Consider now any $w \in C_b^1(\mathbb{R}^n)$, $w \geq 0$ and the associated $v \in C^1(\mathbb{R}^n)$ solving (B.16). Then we have

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

Therefore, by Lemma B.3.1,

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

Using (B.17) again on $v(x_T)$ we get

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

or

$$
a \int_0^\infty e^{-\beta t} w(x(t|x_T))\, dt \leq \int_0^\infty e^{-\beta t} w(x(t|x_0))\, dt. \tag{B.18}
$$

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

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

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

Assuming there is no $\tau \geq 0$ such that $x(\tau \,|\, x_0) = x_T$ we have $\mathcal{X}_0 \cap \mathcal{X}_T = \emptyset$ and since $\mathcal{X}_0$ and $\mathcal{X}_T$ are compact there exists (e.g., by Uryshon's Lemma with mollification) a function $w \in C_b^1(\mathbb{R}^n; [0,1])$ such that $w = 0$ on $\mathcal{X}_0$ and $w = 1$ on $\mathcal{X}_T$. Then the left hand side of (B.18) 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 (B.18) by picking a sufficiently large S. Therefore there exists a $\tau \geq 0$ such that $x(\tau \,|\, x_0) = x_T$ (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 $a \leq e^{-\beta\tau}$. Since $x_T = x(\tau \mid x_0)$ and x_0 are on the same trace we have

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

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

$$av(x_T) \geq e^{-\beta\tau} v(x_T) - \int_0^\tau w(x(t \,|\, x_0))\, dt, \;\text{ or}$$

$$(e^{-\beta\tau} - a) \int_0^\infty e^{-\beta t} w(x(t \,|\, x_T)) dt \geq - \int_0^\tau w(x(t \,|\, x_0))\, dt. \tag{B.19}$$

Consider the set

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

Since x_0 and x_T are on the same trace (and x_T follows x_0) there exists $w \in C_b^1(\mathbb{R}^n)$, $w \geq 0$, such that $w = 0$ on $\mathcal{X}_\tau$ and $w > 0$ elsewhere (such function always exists since every closed set in $\mathbb{R}^n$ is the zero set of a nonnegative bounded smooth function). With this choice of w the equation (B.19) gives

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

and therefore $a \leq e^{-\beta\tau}$ because the integral is strictly positive (since $x(t \mid x_T) \notin \mathcal{X}_\tau$ for for all $t \geq 0$). This proves the first two claims.

To finish we observe that any solution to (B.16) satisfies

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

Therefore

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

Using (B.17) we get

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

Since

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

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. □

B.4 Proof of Theorem 2.3.4

Theorem 2.3.4 *If a pair of nonnegative compactly supported finite measures*

$$(\mu, \mu_0) \in M(\mathbb{R}^n \times \mathbf{U})_+ \times M(\mathbb{R}^n)_+$$

satisfies

$$\int_{\mathbb{R}^n \times \mathbf{U}} v(x) \, d\mu(x, u) = \int_{\mathbb{R}^n} v(x) \, d\mu_0(x) + \alpha \int_{\mathbb{R}^n \times \mathbf{U}} v(f(x, u)) \, d\mu(x, u) \quad \forall v \in C_b(\mathbb{R}^n). \tag{B.20}$$

for all $v \in C_b(\mathbb{R}^n)$, then there exists a measure

$$\eta \in M(l(\mathbb{N}; \mathbb{R}^n))_+$$

supported on the trajectories of (2.19)

$$x_{t+1} = f(x_t, u_t), \quad u_t \in \mathbf{U}, \tag{B.21}$$

such that

$$\mu(B \times \mathbf{U}) = \int_{l(\mathbb{N};\mathbb{R}^n)} \sum_{t=0}^{\infty} \alpha^t I_B(x_t)\, d\eta((x_k)_{k=0}^{\infty}),$$

and

$$\mu_0(B) = \mu(B \times \mathbf{U}) = \int_{l(\mathbb{N};\mathbb{R}^n)} I_B(x_0)\;\, d\eta((x_k)_{k=0}^{\infty}),$$

for all $B \subset \mathbb{R}^n$*.*

Proof: Let $\mathbf{X}$ denote a compact set such that spt$\mu \subset \mathbf{X} \times \mathbf{U}$, spt$\mu_0 \subset \mathbf{X}$. We start by embedding our problem in the setting of discrete-time Markov control processes; terminology and notation is borrowed from the classical reference [HLL96]. Let us define a stochastic kernel on $\mathbf{U}$ given $\mathbf{X}$ as a map $\nu(\cdot\,|\,\cdot)$ such that $\nu(\cdot\,|\,x)$ is a probability measure on $\mathbf{U}$ for all $x \in \mathbf{X}$ and $\nu(B\,|\,\cdot)$ is a measurable function on $\mathbf{X}$ for all $B \subset \mathbf{U}$. Any such stochastic kernel gives rise to a discrete-time Markov process when applied to system (B.21) as a stationary randomized control policy (a policy which, given x, chooses the control action randomly based on the probability distribution $\nu(\cdot\,|\,x)$, i.e., $\mathrm{Prob}(u \in B\,|\,x) = \nu(B\,|\,x)$ for all $B \subset \mathbf{U}$). The transition kernel $\mathcal{R}_\nu(\cdot\,|\,\cdot)$ of this stationary Markov process is then given by

$$\mathcal{R}_\nu(A\,|\,x) = \int_{\mathbf{U}} I_A(f(x,u))\, d\nu(u\,|\,x) = \mathrm{Prob}(x^+ \in A\,|\,x) \quad \forall\, A \subset \mathbb{R}^n,$$

where x is the current state and x^+ the successor state. The t-step transition kernel is then defined by induction as

$$\mathcal{R}_\nu^t(A\,|\,x) := \int_{\mathbb{R}^n} \mathcal{R}(A\,|\,y)\, d\,\mathcal{R}_\nu^{t-1}(y\,|\,x), \quad t \in \{2, 3, \ldots\}$$

with $\mathcal{R}_\nu^1 := \mathcal{R}_\nu$. Given an initial distribution μ_0, the distribution of the Markov chain at time t, $\tilde{\mu}_t$, is given by

$$\tilde{\mu}_t(A) = \int_{\mathbf{X}} \mathcal{R}_\nu^t(A\,|\,x)\, d\mu_0(x) = \mathrm{Prob}(x_t \in A).$$

The joint distribution of state and control is then

$$\mu_t(A \times B) = \int_A \nu(B\,|\,x)\, d\tilde{\mu}_t(x).$$

The discounted occupation measure associated to the Markov process is defined by

$$\mu(A \times B) = \sum_{t=0}^{\infty} \alpha^t \mu_t(A \times B). \tag{B.22}$$

Note that this relation reduces to (2.41) when $\mu_t = \delta_{(x_t,u_t)}$

These considerations along with the following result from [HLL96] lead immediately to the desired conclusion.

Lemma B.4.1 *For any pair of measures (μ_0, μ) satisfying equation (B.20) there exists a stationary randomized control policy $\nu(\cdot\,|\,x)$ such the Markov chain $\{x_t\}_{t=0}^{\infty}$ with transition kernel $\mathcal{R}_\nu$ starting from initial distribution μ_0 (i.e., Markov chain obtained by applying the randomized control policy ν to the difference equation (B.21)) has the discounted occupation measure (B.22) equal to μ.*

The measure $\eta \in M(l(\mathbb{N}\ \mathbb{R}^n))_+$ from Theorem 2.3.4 is then obtained as the law (i.e., the pushforward measure of the underlying probablity space to $l(\mathbb{N};\mathbb{R}^n)$) of the Markov chain $\{x_t\}_{t=0}^{\infty}$ from Lemma B.4.1.

□

Proof of Lemma B.4.1 For completeness we give here a proof of Lemma B.4.1. We start by disintegrating (see Section A.1) μ as $d\mu(x,u) = d\nu(u\,|\,x)d\tilde{\mu}(x)$, where $\tilde{\mu}$ denotes the x-marginal of μ and ν is a stochastic kernel on $\mathbf{U}$ given $\mathbf{X}$. According to the discussion preceding Lemma B.4.1, applying ν to (B.21) gives rise to a stationary discrete-time Markov process with the transition kernel $\mathcal{R}_\nu$ starting form the initial distribution μ_0.

With this notation, equation (B.20) can be equivalently rewritten as

$$\int_{\mathbf{X}} v(x)\, d\tilde{\mu}(x) = \int_{\mathbf{X}} v(x)\, d\mu_0(x) + \alpha \int_{\mathbf{X}} \int_{\mathbb{R}^n} v(y)\, d\,\mathcal{R}_\nu(y\,|\,x)\, d\tilde{\mu}(x) \tag{B.23}$$

for any measurable $v(x)$ (derivation of equation (B.20) did not depend on the continuity of v). Taking $v(x) := I_A(x)$ we obtain

$$\tilde{\mu}(A) = \mu_0(A) + \alpha \int_{\mathbf{X}} \mathcal{R}_\nu(A\,|\,x)\, d\tilde{\mu}(x) \quad \forall\, A \subset \mathbf{X}. \tag{B.24}$$

Using relation (B.23) with $v(x) := \mathcal{R}_\nu(A\,|\,x)$ to evaluate the integral w.r.t. $\tilde{\mu}$ on the right hand side of (B.24) we get

$$\tilde{\mu}(A) = \mu_0(A) + \alpha \int_{\mathbf{X}} \mathcal{R}_\nu(A\,|\,x)\, d\mu_0(x) + \alpha^2 \int_{\mathbf{X}} \mathcal{R}_\nu^2(A\,|\,x)\, d\tilde{\mu}(x).$$

By iterating this procedure we obtain

$$\tilde{\mu}(A) = \mu_0(A) + \sum_{i=1}^{t} \alpha^i \underbrace{\int_{\mathbf{X}} \mathcal{R}_\nu^i(A\,|\,x)\, d\mu_0(x)}_{\mu_i(A)} \; + \; \underbrace{\alpha^{t+1} \int_{\mathbf{X}} \mathcal{R}_\nu^{t+1}(A\,|\,x)\, d\tilde{\mu}(x)}_{\to 0}, \quad \text{(B.25)}$$

and taking the limit as $t \to \infty$ gives

$$\tilde{\mu}(A) = \sum_{t=0}^{\infty} \alpha^t \tilde{\mu}_t(A),$$

where the third term in (B.25) converges to zero because $\alpha \in (0,1)$, $\mathcal{R}_\nu^{t+1}(A\,|\,x) \le 1$ and $\tilde{\mu}$ is a finite measure. Hence the x-marginal of the discounted occupation measure of the Markov chain coincides with the x-marginal of μ.

Finally, to establish equality of the whole measures observe that

$$\sum_{t=0}^{\infty} \alpha^t \mu_t(A \times B) = \sum_{t=0}^{\infty} \alpha^t \int_A \nu(B\,|\,x)\, d\tilde{\mu}_t(x) = \int_A \nu(B\,|\,x)\, d\tilde{\mu}(x) = \mu(A \times B).$$

□

Bibliography

[ABSP11] Jean-Pierre Aubin, Alexandre M Bayen, and Patrick Saint-Pierre. *Viability theory: New directions*. Springer, 2011.

[AF09] J. P. Aubin and H. Frankowska. *Set-valued analysis*. Springer, 2009.

[Ahm08] Amir Ali Ahmadi. Non-monotonic lyapunov functions for stability of nonlinear and switched systems: theory and computation. Master's thesis, Massachusetts Institute of Technology, 2008.

[AM14] Amir Ali Ahmadi and Angshul Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *48th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–5, 2014.

[Amb08] Luigi Ambrosio. Transport equation and cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, pages 1–41. 2008.

[AN87] Edward J Anderson and P. Nash. Linear programming in infinite-dimensional spaces: theory and applications. 1987.

[Ash78] Robert B. Ash. *Real analysis and Probability*. Academic Press, San Diego, 1978.

[BB96] Abhay G Bhatt and Vivek S Borkar. Occupation measures for controlled markov processes: Characterization and optimality. *The Annals of Probability*, pages 1531–1562, 1996.

[BBL02] Thomas Bagby, Len Bos, and Norman Levenberg. Multivariate simultaneous approximation. *Constructive approximation*, 18(4):569–577, 2002.

[Bla99] F. Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[Ble06] Grigoriy Blekherman. There are significantly more nonegative polynomials than sums of squares. *Israel Journal of Mathematics*, 153(1):355–380, 2006.

[Bog06] Vladimir I Bogachev. *Measure theory*, volume 2. Springer, 2006.

[BP12] Alberto Bemporad and Panagiotis Patrinos. Simple and certifiable quadratic programming algorithms for embedded linear model predictive control. In *IFAC Nonlinear Model Predictive Control Conference*, pages 14–20, 2012.

[CCG09] Maria Letizia Corradini, Andrea Cristofaro, and Fabio Giannoni. Sharp estimates on the region of attraction of planar linear systems with bounded controls. In *Conference on Decision and Control*, pages 5345–5350, 2009.

[Che11] Graziano Chesi. *Domain of attraction: analysis and control via SOS programming*, volume 415. Springer, 2011.

[CLR95] Man-Duen Choi, Tsit Yuen Lam, and Bruce Reznick. Sums of squares of real polynomials. *Proceedings of Symposia in Pure mathematics*, 58:103–126, 1995.

[Cri07] Gianluca Crippa. The flow associated to weakly differentiable vector fields. *Theses of Scuola Normale Superiore di Pisa*, 2007.

[DK57] D. A. Darling and M. Kac. On occupation times for Markoff processes. *Transactions of the American Mathematical Society*, pages 444–458, 1957.

[FKJB14] Timm Faulwasser, Milan Korda, Colin N Jones, and Dominique Bonvin. Turnpike and dissipativity properties in dynamic real-time optimization and economic mpc. In *IEEE Decision and Control*, pages 2734–2739, 2014.

[FKJB15] Timm Faulwasser, Milan Korda, Colin N Jones, and Dominique Bonvin. On turnpike and dissipativity properties of continuous-time optimal control problems. *arXiv preprint arXiv:1509.07315*, 2015.

[FV89] Wendell H Fleming and Domokos Vermes. Convex duality approach to the optimal control of diffusions. *SIAM journal on control and optimization*, 27(5):1136–1155, 1989.

[GM76] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[GOSB14] Tom Goldstein, Brendan O'Donoghue, Simon Setzer, and Richard Baraniuk. Fast alternating direction optimization methods. *SIAM Journal on Imaging Sciences*, 7(3):1588–1623, 2014.

[GQ09] Vladimir Gaitsgory and Marc Quincampoix. Linear programming approach to deterministic infinite horizon optimal control problems with discounting. *SIAM Journal on Control and Optimization*, 48(4):2480–2512, 2009.

[HJ14] J. H. Hours and Colin N. Jones. A parametric multi-convex splitting technique with application to real-time NMPC. In *53st IEEE Conference on Decision and Control (CDC 2014)*, 2014.

[HK14] Didier Henrion and Milan Korda. Convex computation of the region of attraction of polynomial control systems. *IEEE Transactions on Automatic Control*, 59:297–312, 2014.

[HLL96] O. Hernandez-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, first edition, 1996.

[HLL09] D. Henrion, J.B. Lasserre, and J. Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24:761–779, 2009.

[HLS09] Didier Henrion, Jean B Lasserre, and Carlo Savorgnan. Approximate volume and integration for basic semialgebraic sets. *SIAM review*, 51(4):722–743, 2009.

[HS00] Kurt Helmes and Richard H Stockbridge. Numerical comparison of controls and verification of optimality for stochastic control problems. *Journal of Optimization Theory and Applications*, 106(1):107–127, 2000.

[JJ10] S Jadlovská and A Jadlovská. Inverted pendula simulation and modeling–a generalized approach. In *International Conference on Process Control*, pages 7–10, 2010.

[Kal10] O. Kallenberg. *Foundations of Modern Probability*. Springer, second edition, 2010.

[KGJO12] M. Korda, R. Gondhalekar, C. N. Jones, and F. Oldewurtel. Stochastic model predictive control: Controlling the average number of constraint violations. In *Conference on Decision and Control*, 2012.

[KGOJ14] Milan Korda, Ravi Gondhalekar, Frauke Oldewurtel, and Colin N. Jones. Stochastic model predictive control: Controlling the average amount of constraint violation. *IEEE Transactions on Automatic Control*, 59:1706–1721, 2014.

[KHJ13] Milan Korda, Didier Henrion, and Colin N. Jones. Inner approximations of the region of attraction for polynomial dynamical systems. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS)*, 2013.

[KHJ14a] Milan Korda, Didier Henrion, and Colin N Jones. Controller design and region of attraction estimation for nonlinear dynamical systems. In *IFAC World Congress on Automatic Control*, pages 2310–2316, 2014.

[KHJ14b] Milan Korda, Didier Henrion, and Colin N Jones. Convex computation of the maximum controlled invariant set for polynomial control systems. *SIAM Journal on Control and Optimization*, 52(5):2944–2969, 2014.

[KHJ16] Milan Korda, Didier Henrion, and Colin N Jones. Controller design and value function approximation for nonlinear dynamical systems. *Automatica*, 67:54–66, 2016.

[KJ13] Milan Korda and Colin N. Jones. Certification of fixed computation time first-order optimization-based controllers for a class of nonlinear dynamical systems. In *American Control Conference*, 2013.

[KJ15] M. Korda and C. Jones. Stability and performance verification of optimization-based controllers. *arXiv preprint arXiv:1501.03919*, 2015.

[KK11] Alexander P Krishchenko and Anatoly N Kanatnikov. Maximal compact positively invariant sets of discrete-time nonlinear systems. In *Proceedings of the IFAC World Congress on Automatic Control, Milano, Italy*, 2011.

[KSK14] Larissa Khodadadi, Behzad Samadi, and Hamid Khaloozadeh. Estimation of region of attraction for polynomial nonlinear systems: A numerical method. *ISA transactions*, 53(1):25–32, 2014.

[Kur11] Thomas G Kurtz. Equivalence of stochastic equations and martingale problems. In *Stochastic Analysis*, pages 113–130. 2011.

[Las01] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[Las06] J. B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM J. Optim*, 17:822–843, 2006.

[Las09] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications,*. Imperial College Press, first edition, 2009.

[Las11] Jean B Lasserre. A new look at nonnegativity on closed sets and polynomial optimization. *SIAM Journal on Optimization*, 21(3):864–885, 2011.

[LHPT08] Jean B Lasserre, Didier Henrion, Christophe Prieur, and Emmanuel Trélat. Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM Journal on Control and Optimization*, 47(4):1643–1666, 2008.

[Löf04] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[Lue69] David G Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1969.

[LZFQ12] Meiqin Liu, Senlin Zhang, Zhen Fan, and Meikang Qiu. State estimation for discrete-time chaotic systems based on a unified model. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 42(4):1053–1063, 2012.

[MH91] Richard M Murray and John Edmond Hauser. *A case study in approximate linearization: The acrobat example*. Electronics Research Laboratory, College of Engineering, University of California, 1991.

[ML11] Kostas Margellos and John Lygeros. Hamilton–jacobi formulation for reach–avoid differential games. *IEEE Transactions on Automatic Control*, 56(8):1849–1861, 2011.

[MT03] Ian M Mitchell and Claire J Tomlin. Overapproximating reachable sets by Hamilton-Jacobi projections. *Journal of Scientific Computing*, 19(1-3):323–346, 2003.

[MVTT14] Anirudha Majumdar, Ram Vasudevan, Mark M Tobenkin, and Russ Tedrake. Convex optimization of nonlinear feedback controllers via occupation measures. *The International Journal of Robotics Research*, pages 1209–1230, 2014.

[Nes00] Yurii Nesterov. Squared functional systems and optimization problems. In *High performance optimization*, pages 405–440. Springer, 2000.

[Nes04] Yurii Nesterov. *Introductory lectures on convex optimization*, volume 87. Springer, 2004.

[Nie12] Jiawang Nie. Regularization methods for SDP relaxations in large-scale polynomial optimization. *SIAM Journal on Optimization*, 22(2):408–428, 2012.

[PAGV+13] A. Papachristodoulou, J. Anderson, S. Prajna G. Valmorbida, P. Seiler, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, 2013.

[Par03] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical programming*, 96(2):293–320, 2003.

[Pet73] David W. Peterson. A review of constraint qualifications in finite-dimensional spaces. *SIAM Review*, 15(3):639–654, 1973.

[PP14] F. Permenter and P. Parrilo. Partial facial reduction: Simplified, equivalent semidefinite programs via approximations of the positive semidefinite cone. *Arxiv preprint arXiv:0905.3447*, 2014.

[PPR04] Stephen Prajna, Pablo Parrilo, and Anders Rantzer. Nonlinear control synthesis by convex optimization. *Automatic Control, IEEE Transactions on*, 49(2):310–314, 2004.

[Pri01] J. A. Primbs. The analysis of optimization based controllers. *Automatica*, 37(6):933–938, 2001.

[PS06] Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Springer, 2006.

[PTZ07] I. Pólik, T. Terlaky, and Y. Zinchenko. Sedumi: a package for conic optimization. In *IMA workshop on Optimization and Control*, 2007.

[Put93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal,*, 42:969–984, 1993.

[Ran01] Anders Rantzer. A dual to lyapunov's stability theorem. *Systems & Control Letters*, 42(3):161–168, 2001.

[Ran02] Anders Rantzer. A converse theorem for density functions. In *IEEE Conference on Decision and Control*, pages 1890–1891, 2002.

[RH03] Anders Rantzer and Sven Hedlund. Duality between cost and density in optimal control. In *Proceedings. 42nd IEEE Conference on Decision and Control*, pages 1218–1221, 2003.

[RJM12] Stefan Richter, Colin Neil Jones, and Manfred Morari. Computational complexity certification for real-time mpc with input constraints based on the fast gradient method. *IEEE Transactions on Automatic Control*, 57(6):1391–1403, 2012.

[RM09] James B. Rawlings and David Q. Mayne. *Model Predictive Control: Theory and Design.* Nob Hill Publishing, first edition, 2009.

[Rub85] Julio E Rubio. *Control and Optimization: the linear treatment of nonlinear problems.* Manchester University Press, 1985.

[Rud86] Walter Rudin. *Real and complex analysis.* New York: McGraw-Hill Inc, third edition, 1986.

[RV14] Anand Raghunathan and Umesh Vaidya. Optimal stabilization using lyapunov measures. *IEEE Transactions on Automatic Control*, 59(5):1316–1321, 2014.

[Sei10] P. Seiler. SOSOPT: A toolbox for polynomial optimization. *University of Minnesota*, 2010.

[SKJ14] Georgios Stathopoulos, Milan Korda, and Colin Jones. Solving the infinite-horizon constrained lqr problem using splitting techniques. In *IFAC World Congress on Automatic Control*, pages 2285–2290, 2014.

[SKJ15] Giorgos Stathopoulos, Milan Korda, and Colin N Jones. Solving the infinite-horizon constrained lqr problem using accelerated dual proximal methods. *arXiv preprint arXiv:1501.04352*, 2015.

[Sta09] Konstantin E Starkov. Bounds for compact invariant sets of the system describing dynamics of the nuclear spin generator. *Communications in Nonlinear Science and Numerical Simulation*, 14(6):2565–2570, 2009.

[Sto90] Richard H Stockbridge. Time-average control of martingale problems: A linear programming formulation. *The Annals of Probability*, 18(1):206–217, 1990.

[Stu98] B. Sturmfels. Polynomial equations and convex polytopes. American Mathematical Monthly. *American Mathematical Monthly*, 105(10):907–922, 1998.

[Trn05] Mária Trnovská. Strong duality conditions in semidefinite programming. *Journal of Electrical Engineering*, 56(12), 2005.

[Tse91] Paul Tseng. Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM Journal on Control and Optimization*, 29(1):119–138, 1991.

[Vin93] Richard B Vinter. Convex duality and nonlinear optimal control. *SIAM journal on control and optimization*, 31(2):518–538, 1993.

[VL78a] Richard B Vinter and Richard M Lewis. The equivalence of strong and weak formulations for certain problems in optimal control. *SIAM Journal on Control and Optimization*, 16(4):546–570, 1978.

[VL78b] Richard B Vinter and Richard M Lewis. A necessary and sufficient condition for optimality of dynamic programming type, making no a priori assumptions on the controls. *SIAM Journal on Control and Optimization*, 16(4):571–583, 1978.

[WV10] Kai Wang and Umesh Vaidya. Transfer operator approach for computing domain of attraction. In *IEEE Conference on Decision and Control*, pages 5390–5395, 2010.

[You33] L. C. Young. On approximation by polygons in the calculus of variations. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, pages 325–341, 1933.

[You69] L. C. Young. Lectures on the calculus of variations and optimal control theory. *Sunders, Philadelphia*, 1969.

[ZST10] Xin-Yuan Zhao, Defeng Sun, and Kim-Chuan Toh. A Newton-CG augmented lagrangian method for semidefinite programming. *SIAM Journal on Optimization*, 20(4):1737–1765, 2010.

Milan Korda

Academic CV

Education

2012 – 2016 **PhD in Automatic Control**, *École Polytechnique Fédérale de Lausanne*.
Supervisor: Prof. Colin N. Jones
Topic: Polynomial optimization for dynamical systems
Student supervision: two master theses, five semester projects, three team projects
Teaching: assistant for the Control systems class and the Model predictive control class

2009–2011 **Master in Cybernetics and Robotics**, *Czech Technical University in Prague*.
Master thesis: "Stochastic Model Predictive control", conducted at ETH Zurich
First Class Honours; GPA 4.0 / 4.0

2006–2009 **Bachelor in Electrical Engineering**, *Czech Technical University in Prague*.
First Class Honours; GPA 4.0 / 4.0

Awards & Scholarships

2014 Best presentation in session, IEEE American Control Conference, Portland

2011 Czech Siemens Award – for the master thesis titled "Stochastic Model Predictive Control"

2011 Preciosa prize – for the master thesis titled "Stochastic Model Predictive Control"

2011 Dean's prize – for outstanding results during master studies

2011 Erasmus scholarship, ETH Zurich

2009 Dean's prize – for outstanding results during bachelor studies

Workshop organization

2015 Operator theoretical methods for dynamical systems control and optimization, European Control Conference (ECC), 2015 – co-organizer

2014 Mini-course on polynomial optimization and control, International Symposium on Mathematical Theory of Networks and Systems (MTNS), 2014 – co-organizer

Invited talks

2016 Mini-workshop on Applied Koopmanism, Mathematisches Forschungsinstitut Oberwolfach (MFO), talk title "Moment-sum-of-squares hierarchies for set approximation and optimal control"

2015 SPOT (Multidisciplinary Optimization Seminar in Toulouse) seminar series, talk title "Constrained LQR control with accelerated proximal gradient algorithm"

2015 SIAM conference on Applications of Dynamical Systems, minisymposium on Uncertainty Propagation in High Dimensional Systems, talk title "Optimal Control Design and Value Function Estimation for Nonlinear Dynamical Systems"

2014 VORACE (Verified Fast Optimization for Embedded Control) workshop, talk title "Certification of fixed computation time first-order optimization-based controllers for a class of nonlinear dynamical systems"

2013 GeoLMI - Conference on Geometry and Algebra of Linear Matrix Inequalities, talk title "Region of attraction approximation for polynomial dynamical systems"

2011 CTU Prague, Department of Automatic control – talk title "Non-quadratic stochastic model predictive control"