Safety Constraints In Control Systems

Dylan Janak

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Behçet Açıkmeşe, Chair

Dana Dabiri

Mehran Mesbahi

Lillian Ratliff

Program Authorized to Offer Degree:

William E. Boeing Department of Aeronautics & Astronautics

University of Washington

Abstract

Safety Constraints In Control Systems

Dylan Janak

Chair of the Supervisory Committee:
Behçet Açıkmeşe
William E. Boeing Department of Aeronautics & Astronautics

This dissertation highlights the research of myself and my numerous collaborators throughout my time at the University of Washington on topics such as: safety constraints, controlled invariant sets, Markov chains, Markov decision processes, and swarm guidance. As autonomous systems become more prevalent, great care must be taken to ensure that they remain in well-understood operating regimes, despite the unavoidable uncertainty of the real world. Safety constraints are therefore critical to consider for a wide variety of models, especially those which explicitly incorporate uncertainty. If safety cannot be guaranteed in some well-defined sense, then a model cannot confidently be applied to a practical problem—or may even guarantee an undesired outcome. Therefore, the analysis of control systems under safety constraints is integral to bridging the gap between mathematical models and real-world processes.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

"I cannot remember the books I've read any more than the meals I have eaten; even so, they have made me." —Ralph Waldo Emerson

It would be impossible to give the proper credit to everyone responsible for getting me to where I am today, as I write this. From family, friends, teachers, students, coworkers, and collaborators, to the countless strangers who have made my world a better place in one way or the other, I can only ever hope to thank a tiny fraction of these people. But I will try to list a few.

I first met my advisor, Dr. Behçet Açıkmeşe, at The University of Texas at Austin during my time as an undergrad, where he introduced me to control theory as my professor. Without him, I almost certainly would not have developed such a passion for guidance and control. He still loves recounting the conversation in which he urged me to pursue grad school and join his research group; I was nearing graduation and did not have much idea what I was going to do afterwards. (Of course, I gladly accepted!) Ever since, he has been a terrific mentor as head of the Autonomous Controls Laboratory (ACL), and has always been eager to teach and learn alike. He is also without a doubt my biggest academic inspiration.

Thanks also to everyone I met and occasionally worked with under Behçet at ACL: Mahmoud El Chamie, Danylo Malyuta, Sarah Li, Yue Yu, Daniel Dueri, Utku Eren, Nazlı Demir, Yuanqi Mao, Sean Rice, Dan Calderone, Miki Szmuk, Skye Mceowen, and many others. You all are always great company, and I appreciate all the fun and intellectually stimulating conversations we have had. Thanks to the RAIN lab as a whole too, for all the controls coffee hours our lab groups have shared, with a special shoutout to Mathias Hudoba de Badyn for being such a great coauthor and an overall positive presence.

Another special thanks to Pierre-Loïc Garoche, for being an especially wonderful

research partner and host during my stays in and around Toulouse. It has been an absolute pleasure attending conferences and working with him at ONERA and ENAC, and learning about lattices, formal verification, and positively invariant sets. I hope to see you in France again someday!

Lastly, I quite literally would not be here without my parents, Don and Mary Ann. Besides the obvious, they have set out to instill in me a desire to be the very best version of myself, and to guide me through life the best they could. They were—and still are—always there to talk with and to give all the advice I could ever ask for (and then some). Likewise, I wish to thank my grandma, as well as all my aunts, uncles, and cousins for all the care, motivation, and enjoyment they have provided throughout the years, with special thanks to Uncle David and Aunt Suzanne for assisting with my on-campus housing expenses and making sure that affordability did not keep me from getting the best possible education for myself.

DEDICATION

For C.U. and the countless others who have made me.

Nomenclature

$[N]$ $\{1, \ldots, N\}$, $N \in \mathbb{N}$

Δ_N $\{y \in \mathbb{R}^N_+ \mid \sum_i y_i = 1\}$. A simplex of dimension N.

$\mathbb{E}[\cdot]$ The expectation of a random variable.

$\mathbb{P}(\mathcal{X})$ The set of all *non-empty compact subsets* of $\mathcal{X}$.

$\mathcal{T}$ $\{0, \ldots, T\}$. Finite time horizon.

$\|\cdot\|$ Norm of a real vector space.

$\mathbb{1}_N$ A column vector of ones: $\mathbb{1}_N = [1, \ldots, 1]^\top \in \mathbb{R}^{N \times 1}$.

$\mathbb{R}(\mathbb{R}_+)$ The (positive) real number line.

$B_\epsilon(x^\star)$ $\{x \mid \|x - x^\star\| \leq \epsilon, x \in \mathbb{R}^n\}$. The set of all elements in $\mathbb{R}^n$ that is at most $\epsilon \in \mathbb{R}_+$ away from the input $x^\star \in \mathbb{R}^n$.

$C^r(\mathbb{R}^n, \mathbb{R})$ The set of continuously differentiable functions $f : \mathbb{R}^n \mapsto \mathbb{R}$, such that $\frac{d^j f(x)}{du^j}$ exists for all $x \in \mathbb{R}^n$ and $1 \leq j \leq r$.

I_N An identity matrix of size $N \times N$.

x^{-i} When $x = (x^i, x^{-i})$, x^{-i} is the strategy space of the opponents of player i.

Chapter 1

INTRODUCTION

1.1 Overview

The contributions of my research share the broad theme of swarm guidance, whether it be a literal swarm of vehicles moving between physical configurations, or a more abstract collection of possible trajectories in a stochastic model. A ubiquitous methodology for the theory and application of control theory is to build a mathematical model for the process being controlled, to design a controller which achieves certain design objectives for this idealized system, and then to experimentally verify the controller on a real system. Applying this methodology for controlling a swarm, each of these stages has its own challenges that require special attention; the model must capture the essential behavior of a swarm but not be too complex to rigorously analyze, there must be available algorithms for designing appropriate control laws across a wide range of multi-agent systems, and it must be possible to certify that a guidance algorithm is usable in practice and achieves the desired behavior. It is with these steps in mind that my research into the areas of Markov decision processes, optimal mass transport theory, certification of constraint satisfaction, and beyond will strive to further our knowledge of the theory and application of swarm guidance and control.

In the past decade alone, real-time optimization has matured into a proven technology in aerospace applications such as rocket landing. As our technological capabilities grow, so must our underlying confidence that our control algorithms will robustly and efficiently accomplish their objectives, and we must be able to examine how vulnerable our systems are to a broad set of unpredictable phenomena. An important challenge is to design complex systems which are certifiably safe while still achieving high performance. To this end, we must be able to translate the abstract idea of safety into concrete constraints on our systems, and then to optimize with respect to these constraints with the knowledge that our system model is fundamentally imprecise under real-world operating conditions.

1.1.1 Markov Decision Processes

A Markov Decision Process (MDP) is a stochastic control problem over a finite set of states and control actions, with associated costs. Since no model of a real-world system is perfect, each control action results in a probability distribution over possible outcomes. Under this framework, we can design a policy—a method of choosing which action to take from the current information—that optimizes a chosen objective while enforcing constraints on the probability distribution over states. The Sequentially Ordered MDP (SO-MDP) model is a variation on MDP where the action is chosen based on the current state and measured result of each future action in real time. My work with El Chamie and Açıkmeşe [27] shows that these extra measurement data result in improved performance and safety over a policy that chooses an action from only the current state information. We observed that benefit was the most dramatic in or near high-risk regions with large penalties, and therefore this kind of model is particularly useful to guaranteeing the satisfaction of probabilistic state constraints in the presence of significant uncertainty.

1.1.2 Safety Constraint Certification

Safety constraints are conditions requiring the state of a system to remain in some prescribed set $\mathcal{S}$ for all future times. For example, the constraint set for a swarm of vehicles may be the set of all configurations where no vehicle is colliding with an obstacle or another vehicle. Any example of a trajectory where the state leaves the constraint set proves the entire safety condition false, however proving that such a constraint is always satisfied cannot be proven true by brute force. One key result for dynamical systems of the form $x(t+1) = f(x(t))$ is that an initial condition satisfies the safety constraint if and only if it is within the maximal positively invariant subset of the constraint set. The standard algorithm to compute this set presented by Kolmanovsky [49] is a special case of a more general Kleene iteration algorithm, as discussed in §3.5.

1.1.3 Optimal Mass Transport for Swarms

Optimal mass transport (OMT) is the idealized problem of moving an initial configuration of "mass" into a desired final configuration in the most efficient way possible. This problem is of clear importance to swarm guidance. In our paper "Discrete-Time Linear-Quadratic Regulation via Optimal Transport, my coauthors and I demonstrated how the textbook

discrete-time LQR problem may be analyzed within the theoretical framework of OMT. By modeling the initial and final uncertainties in the state of an agent as probability distributions, the LQR problem for a stochastic system can be converted into an OMT problem with an explicit form. This idea also extends to a broader family of problems with constraints on the spatial density of agents as they move toward a desired configuration, and similar state constraints.

1.2 Notation

Arbitrary sets are denoted by capital script Latin letters.

$\mathbb{N}$ is the set of natural numbers $\{0, 1, 2, \dots\}$

$\mathbb{R}$ is the set of real numbers

$\mathbb{R}_+$ is the set of nonnegative real numbers

$\mathbb{R}^n$ is the set of $n$dimensional real vectors

$\mathbb{R}^{m\times n}$ is the set of real matrices of size $m \times n$

$\mathbf{1}$ is the column vector of ones of appropriate dimension

$\leq$ indicates entrywise inequality for vectors and matrices

$\overline{\mathcal{S}}$ is the closure of set $\mathcal{S}$

$$A \prec B \iff x^\intercal Ax < x^\intercal Bx,\ \forall x \in \mathbb{R}^n \backslash \{0\}$$

$$A \preceq B \iff x^\intercal Ax \leq x^\intercal Bx,\ \forall x \in \mathbb{R}^n$$

$$2^{\mathcal{A}} := \{\mathcal{B} \mid \mathcal{B} \subseteq \mathcal{A}\}$$

$$\mathcal{A} \oplus \mathcal{B} = \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$$

$$\|x\|_Q := \sqrt{x^\intercal Qx}$$

$$\Delta(n) = \{x \in \mathbb{R}^n \mid \mathbf{1}^\intercal x = 1,\ x \geq 0\}$$

$f^k(x)$ is defined recursively for $k \geq 0$ as $f^0(x) := x,\ f^{k+1}(x) = f(f^k(x))$

$$f^{-k}(\mathcal{A}) := \{x \mid f^k(x) \in \mathcal{A}\},\ k \in \mathbb{N}$$

$$\mathcal{O}_\ell(\mathcal{S}, f) := \bigcap_{k=0}^{\ell} f^{-k}(\mathcal{S})$$

$$\omega(x, f) := \bigcap_{k\in\mathbb{N}} \overline{\{f^l(x) \mid l \geq k\}}$$

$$\Omega(\mathcal{S}, f) := \{x \in \mathcal{S} \mid \omega(x, f) \subseteq \mathcal{S}\}$$

$$d(x, \mathcal{Y}) := \inf_{y\in\mathcal{Y}} \|x - y\|$$

$$\mathcal{B}_r := \{x \mid \|x\| < r\}$$

$$\kappa_p(V) := \left|\left|\left|V\right|\right|\right|_p \left|\left|\left|V^{-1}\right|\right|\right|_p$$

1.3 Major Contributions

Accepted:

- Mahmoud El Chamie, Dylan Janak, and Behçet Açıkmeşe. Markov decision processes with sequential sensor measurements. *Automatica*, 103:450–460, 2019. [27]

- Dylan Janak and Behçet Açıkmeşe. Maximal invariant set computation and design for Markov chains. In *2019 American Control Conference (ACC)*, pages 1244–1249. IEEE, 2019. [45]

- Mathias Hudoba de Badyn, Erik Miehling, Dylan Janak, Behçet Açıkmeşe, Mehran Mesbahi, Tamer Başar, John Lygeros, and Roy S. Smith. Discrete-time linear-quadratic regulation via optimal transport. [23]

In progress:

- Dylan Janak, Pierre-Loïc Garoche, and Behçet Açıkmeşe. Exact Computation of Maximal Invariant Sets for Safe Markov Chains. Pending approval.

Chapter 2

MARKOV DECISION PROCESSES

2.1 Markov Chains

A Markov process is defined as a stochastic process with the Markov property: for each $t > \tau$, $\Pr[X(t) \in \mathcal{S}_i \mid \{X(\sigma)\}_{\sigma \leq \tau}] = \Pr[X(t) \in \mathcal{S}_i \mid \{X(\tau)\}]$. In other words, the future trajectory can only depend on the current state, and not on the past history of states. If two trajectories have different histories but have the same state at time τ, then they will behave statistically identically from that time onward. Although the future trajectory often cannot be known with certainty, Markov processes correspond to linear systems over probabilities from the definition of conditional probability:

$$\Pr[X(t) \in \mathcal{S}_i] = (\Pr[X(t) \in \mathcal{S}_i \mid X(\tau) \in \mathcal{S}_j]) \times \Pr[X(\tau) \in \mathcal{S}_j]. \tag{2.1}$$

A *Markov chain* is a Markov process with a discrete, totally-ordered set of times indexed by $t \in \mathbb{N}$, and a finite set of states $\mathcal{S} = \{s_1, \ldots, s_n\}$. The conditional probabilities of Equation (2.1) are arranged into a sequence of state transition matrices $M(t)$ so that $M_{ij}(t) = \Pr[X(t+1) = s_i \mid X(t) = s_j]$. From the unitarity and nonnegativity of probabilities, these matrices have the properties

$$\sum_{i=1}^{n} M_{ij}(t) = 1, \ \forall j \in [1, n], \ \forall t \in \mathbb{N}, \tag{2.2}$$

$$M_{ij}(t) \geq 0, \ \forall (i, j) \in [1, n]^2, \ \forall t \in \mathbb{N}. \tag{2.3}$$

Properties (2.2) and (2.3) are written more compactly as the following:

$$\mathbf{1}^\intercal M(t) = \mathbf{1}^\intercal, \ \forall t \in \mathbb{N}, \tag{2.4}$$

$$M(t) \geq 0, \ \forall t. \tag{2.5}$$

In other words, each $M(t)$ is a square, column-stochastic matrix (also called a *Markov ma-*

trix).

Although Markov chains are fundamentally nondeterministic in general, they can be formulated as a linear system whose state $x(t)$ describes a probability distribution over $\mathcal{S}$ at time t:

$$x(t) = \begin{bmatrix} \Pr[X(t) = s_1] \\ \vdots \\ \Pr[X(t) = s_n] \end{bmatrix}. \tag{2.6}$$

This new state x is a vector in $\Delta(n)$ which satisfies the linear dynamics

$$x(t+1) = M(t)x(t). \tag{2.7}$$

If $x(\tau)$ is given, then $x(t)$ can be found for all $t > \tau$ in terms of $x(\tau)$ as

$$x(t) = M(t-1)M(t-2)\dots M(\tau)x(\tau). \tag{2.8}$$

The product of Markov matrices from times $t-1$ to τ is also a Markov matrix, and its entries are the conditional probabilities of $X(t)$, given $X(\tau)$.

2.2 Markov Decision Processes

A Markov decision process (MDP) is a controlled Markov chain with a discrete time index $t \in \mathbb{N}$, a finite set of states $\mathcal{S} = \{s_1, \dots, s_n\}$, a finite set of actions $\mathcal{A} = \{a_1, \dots, a_m\}$, a *transition kernel* G mapping state-action pairs to probability distributions on $\mathcal{S}$, and a cost (or alternatively, reward) functional on the feasible sequences of state-action pairs $(s(t), a(t))_{t\in\mathbb{N}}$. We also attach a cost functional over trajectories of state-action pairs; e.g., total cost $\sum_{t=0}^{\infty} r(s(t), a(t))$, discounted cost $\sum_{t=0}^{\infty} \gamma^t r(s(t), a(t))$ with a discount factor $\gamma \in [0, 1)$, or average cost $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} r(s(t), a(t))$. The current state is observed, and the controller chooses an action according a policy π.

We seek an optimal policy π^* for choosing an action based on knowledge of the state which minimizes the expected cost. A policy is said to be deterministic if the action is specified by

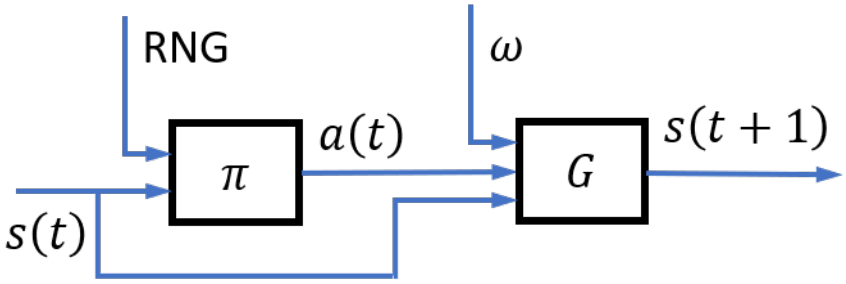

Figure 2.1: One time step of a Markov decision process. The policy π chooses a random action $a(t)$ using a random number generator. The state transitions to $s(t+1)$ stochastically from the state-action pair $(s(t), a(t))$ and a random variable ω.

the current state, and mixed if the action follows a distribution which is a function of state.

$$\Pr[s(t+1) = s_i] = \sum_{j=1}^{n} \sum_{k=1}^{m} \Pr[s(t+1) = s_i \mid s_j, a_k] \Pr[a_k \mid s_j] \Pr[s(t) = s_j]. \tag{2.9}$$

As with Markov chains, MDPs are formulated as a linear system on the space of probability distributions over $\mathcal{S}$, using the state vector $x(t)$ defined as $x_i(t) = \Pr[s(t) = s_i]$ similarly to Equation (2.6). This allows Equation (2.9) to be written in matrix form:

$$x(t+1) = M^{\pi}(t)x(t), \tag{2.10}$$

$$M^{\pi}_{ij}(t) = \sum_{k=1}^{m} G_{ijk}(t)\pi_{kj}(t), \tag{2.11}$$

$$G_{ikj}(t) = \Pr[s(t+1) = s_i \mid s(t) = s_j, a(t) = a_k], \tag{2.12}$$

$$\pi_{kj}(t) = \Pr[a(t) = a_k \mid s(t) = s_j]. \tag{2.13}$$

The state transition matrix $M^{\pi}(t)$ depends linearly on the policy π, which gives the probabilities of choosing each action upon observing the current state.

Rather than treating the set of actions as independent of state, we may instead consider the state-action pairs accessible from any particular state. There are several advantages of doing this, such as to express $M^{\pi}(t)$ as a product of a transition kernel matrix and a policy matrix, to conveniently frame the MDP dynamics over state-action pairs rather than just states or actions separately, and to allow for a state-dependent set of available actions $\mathcal{A}(s)$. Although the latter may appear to be more general, one could always choose $\mathcal{A}$ so

that $|\mathcal{A}| = \max_{s\in\mathcal{S}} |\mathcal{A}(s)|$, and introduce redundant actions for states with fewer options of action.

For convenience, the transition kernel G and policy π can be recast as $n \times mn$ and $mn \times n$ matrices, respectively. Let $P_{i\ell}$ be the probability that $s(t+1) = s_i$ given the ℓ^{th} state-action pair is realized at time t. Also let $\Pi_{\ell j}$ be the probability for being in the ℓ^{th} state-action pair at time t, conditioned upon $s(t) = s_j$. The MDP dynamics then simplify to

$$x(t+1) = P(t)\Pi(t)x(t). \tag{2.14}$$

2.2.1 Constrained MDP Synthesis

For any particular policy, the dynamics of the MDP $(\mathcal{S}, \mathcal{A}, P, R)$ reduces to the Markov chain shown in Equation (2.10). The transition matrix $M^\pi(t) = P(t)\Pi(t)$ separates the uncontrolled probabilities in transition kernel P from the policy matrix Π. To understand the set of feasible Π, the indicator matrix $I_o \in \{0,1\}^{n\times mn}$ is defined so that its $(i,k)^{\text{th}}$ entry is 1 iff the k^{th} state-action pair contains state s_i. If the set of actions is uniform across all states, then

$$I_o = I_{n\times n} \otimes \mathbf{1}_{1\times n} \tag{2.15}$$

$$= \begin{bmatrix} \mathbf{1}_{1\times n} & 0_{1\times n} & \cdots & 0_{1\times n} \\ 0_{1\times n} & \mathbf{1}_{1\times n} & \cdots & 0_{1\times n} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{1\times n} & 0_{1\times n} & \cdots & \mathbf{1}_{1\times n} \end{bmatrix}. \tag{2.16}$$

Because $\Pi_{kj} > 0$ only if the k^{th} state-action pair contains state s_j, and each column of Π is a probability vector, $\Pi(t)$ always satisfies the following linear constraints:

$$I_o\Pi(t) = I_{n\times n}, \tag{2.17a}$$

$$\Pi(t) \geq 0. \tag{2.17b}$$

Furthermore, any $\Pi(t)$ satisfying conditions (2.17) is compatible with the MDP. Since Π must have the sparsity structure of $I_o^\intercal$, the policy can also be represented by the vector $\pi \in \mathbb{R}_+^{mn}$

such that $\Pi(t) = \text{Diag}(\pi(t))I_o^\intercal$, and $\pi(t)$ satisfies the linear constraints

$$I_o\pi(t) = \mathbf{1}_{n\times 1}, \tag{2.18a}$$

$$\pi(t) \geq 0. \tag{2.18b}$$

The dynamics of the MDP can then be parameterized either by the policy matrix $\Pi(t)$ or the policy vector $\pi(t)$:

$$M^\pi(t) = P\Pi(t), \tag{2.19}$$

$$\Pi(t) = \text{Diag}(\pi(t))I_o^\intercal. \tag{2.20}$$

Since P and I_o are treated as fixed, the identities (2.19) and (2.20) are linear. Thus, any convex program in the state transition matrices $M^\pi(t)$ is also convex with respect to the policy vector formulation.

One can also consider the distribution over state-action pairs $y(t) \in \mathbb{R}_+^{mn}$, so that $y_k(t)$ is the probability of the t^{th} state-action pair occurring at time t. $y(t)$ captures both $x(t)$ and $\pi(t)$ simultaneously through the following identities:

$$y(t) = \Pi(t)x(t), \tag{2.21a}$$

$$x(t+1) = Py(t), \tag{2.21b}$$

$$x(t) = I_o y(t), \tag{2.21c}$$

Unlike in Equations (2.19) and (2.20), $y(t)$ is bilinear with respect to $\Pi(t)$ because $y(t) = \Pi(t)I_o y(t)$. This lets constraints which are linear in $y(t)$ but not in $\pi(t)$ be formulated as convex constraints. I.e., the constraint $x(t) \leq d$ is linear in y but not in the sequence of policy matrices, since $x(t) = I_o y(t) = P\Pi(t)P\Pi(t-1)\cdots P\Pi(0)x(0)$. Furthermore, the policy vector $\pi(t)$—and therefore $M^\pi(t)$—can be computed from $y(t)$ directly if $\text{Diag}(I_o^\intercal I_o y(t))$ is invertible:

$$\pi(t) = \text{Diag}(I_o^\intercal I_o y(t))^{-1} y(t). \tag{2.22}$$

This identity follows from the observation that each entry of $y(t)$ is a joint probability of being in the corresponding state $x(t)$ and taking the action with conditional probability given by $\pi(t)$. $I_o y(t)$ is simply $x(t)$, so $I_o^\intercal I_o y(t)$ is a vector whose i^{th} entry is the probability of the

state contained in the i^{th} state-action pair. With this choice of variables, the time-varying MDP takes the form

$$\begin{aligned} \min_{x(t),y(t)} \quad & \sum_{t=0}^{T} r(t)^\intercal y(t) \\ \text{s.t.} \quad & x(t) = I_o y(t), \\ & x(t+1) = P y(t), \\ & \mathbf{1}^\intercal y(t) = 1, \\ & y(t) \geq 0, \ \forall t \in 1, \ldots, T. \end{aligned} \tag{2.23}$$

The solution to this linear program gives the optimal sequence of $y(t)$. Additional convex constraints in $x(t)$ or $y(t)$ can be added without affecting the convexity of this problem. Then, we recover the time-varying policy from $y(t)$ using Equation (2.22).

2.3 Sequentially Observed MDP

A Sequentially Observed Markov Decision Process (SO-MDP) is a generalization of the usual MDP model. For a traditional MDP, the action is based only on the current state, while a SO-MDP determines the outcome of each action in sequence, either accepting or rejecting the action. At the start of each decision epoch, the outcome of action a_1 is measured in phase 1, and the action is either accepted or rejected. If accepted, the new state will take the observed value. Otherwise, action a_2 is sampled in phase 2, and the process repeats for all a_k in sequence up to a_{m-1}. If no actions are accepted in phases 1 through $m-1$, then a_m is forced to be accepted.

The policy for a SO-MDP takes the current state, phase, and measured transition, and then returns the probability of accepting the action. This contrasts with the normal MDP model in which the action is chosen at the start of the epoch without measuring the outcome of any action. Since the normal MDP policies are a subset of SO-MDP policies, the expected reward is generally increased with the additional measurement data. Furthermore, the optimal SO-MDP policy can be efficiently computed via backward induction. Since it can also be formulated as a linear program, off-the-shelf LP (or convex program) solvers can find the optimal policy when there are linear (convex) constraints on the state frequencies.

Numerical simulations show that the optimal SO-MDP policy typically outperforms the optimal MDP policy the most wherever there is a large spread in the distribution of rewards, such as when the state is near a very high-cost state such as an obstacle. Figure 2.4 shows the increase in expected utility from traditional MDP to sequantially ordered MDP in two similar

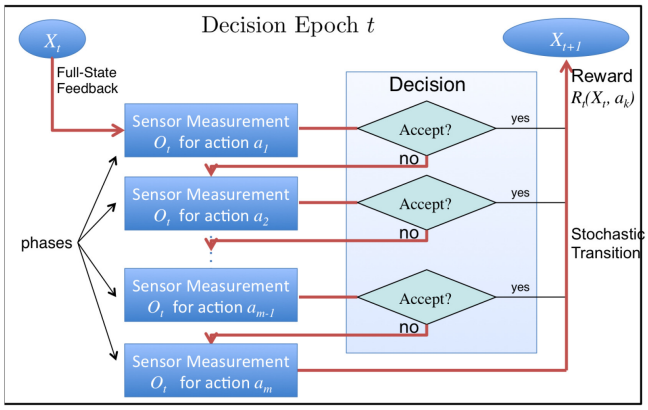

Figure 2.2: The SO-MDP model measures the outcome of each action, and either accepts or rejects each observed transition.

Algorithm 2 Backward Induction: Sequentially-Observed MDP Optimal Policy

1: Start with $V_N^*(i) = \bar{r}_N(i)$ for all $i \in S$
2: for $t = N-1, \ldots, 1$ given V_{t+1}^* and for $i = 1, \ldots, n$ calculate the optimal value

$$V_t^*(i) = \max_{P_i \in \mathcal{C}} \left\{ \bar{r}_t(i) + \sum_{j \in S} M_t(j,i) V_{t+1}^*(j) \right\}$$

and the optimal policy $P_i^*(t)$ given by:

$$P_i^*(t) = \operatorname{argmax}_{P_i \in \mathcal{C}} \left\{ \bar{r}_t(i) + \sum_{j \in S} M_t(j,i) V_{t+1}^*(j) \right\}$$

3: **Result:** $v_N^* = \mathbf{x}_1^T V_1^*$ where $\mathbf{x}_1$ is the initial state probability distribution.

Figure 2.3: Backward induction algorithm for SO-MDP. $V_t^*(i)$ is the optimal cost to go for measuring a transition to state i during phase t.

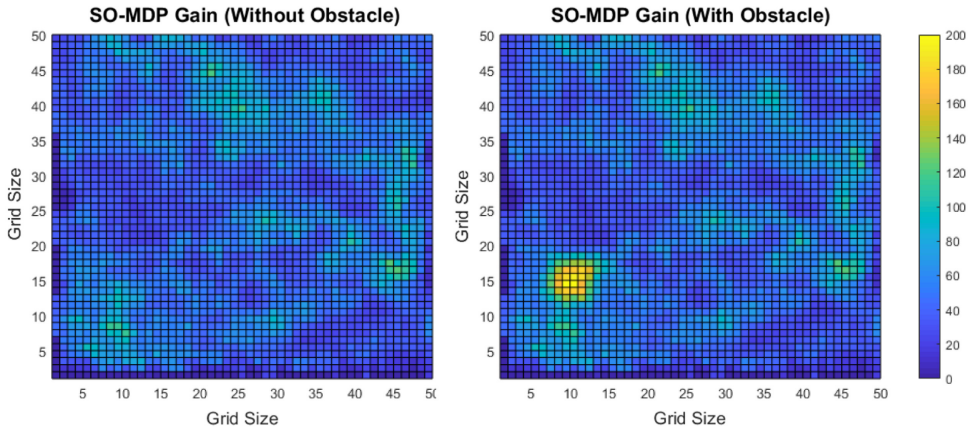

Figure 2.4: The SO-MDP experiences a higher reward across all states than the traditional MDP. This is most pronounced near the high-cost square of states added in the bottom-left of the rightmost figure.

square grids. The agent can either choose to move in each of the four cardinal directions or stay still, with a 60% chance to execute the desired motion, and a 10% chance for each of the others. Each state-action pair is assigned a uniform random reward from the start. In the figure, the right plot gives artificially low reward to a 5×5 square of cells near the bottom left, providing empirical evidence that SOMDP is particularly useful if there is a moderately high risk of incurring a large cost. For safety critical applications, the greater certainty from the outcomes of an agent's actions can lead a significant increase in performance and safety.

Chapter 3

SAFETY CONSTRAINTS

An important part of designing a controller is to be able to guarantee that certain requirements are practically always met. Formally, a safety condition is a logical statement that a particular property is true for all times. For example, an autonomous vehicle may be in danger of colliding with an obstacle only if they are ever within some known distance D. Then, collisions can be avoided if the following safety condition is met: that the distance between the vehicle and obstacle shall be greater than D always.

3.1 Formulation of Safety Constraints

First consider the time-invariant system with $f : \mathcal{X} \times \mathcal{U} \times \mathcal{W} \to \mathcal{X}$ and $h : \mathcal{X} \times \mathcal{W}$

$$x(t+1) = f(x(t), u(t), w(t)), \tag{3.1}$$

$$y(t) = h(x(t), w(t))x(0) \qquad \in \mathcal{X}_0. \tag{3.2}$$

x is the system state, y is the observation, u is a feedback control signal, and $w(t)$ is an uncontrolled disturbance signal. A safety constraint for this system would then take the form

$$x(t) \in \mathcal{S}, \ \forall t \in \mathbb{N}. \tag{3.3}$$

One natural question to ask is is there exists a feedback control law for which the safety condition (3.3) is satisfied for all sequences of disturbances.

3.2 Positively Invariant Sets

For each set $\mathcal{S} \subseteq \mathcal{X}$, there is an associated safety constraint $x(t) \in \mathcal{S}$, $\forall t \in \mathbb{N}$. This constraint is violated if and only if there exists some $t \in \mathbb{N}$ for which $x(t) \notin \mathcal{S}$. Therefore in principle, any safety constraint can be proven false by searching for a single counterexample, however infinitely many conditions $x(0) \in \mathcal{S}, x(1) \in \mathcal{S}, \ldots$ must be checked to prove it true.

3.2.1 Safety Certification Through Positive Invariance

Consider the idealized dynamical system without disturbance:

$$x(t+1) = f(x(t), u(t)), \tag{3.4}$$

$$y(t) = h(x(t)). \tag{3.5}$$

The goal of this section is to assess whether a particular feedback control law of the form $u(t) = \kappa(y(t))$ satisfies the safety constraint $x(t) \in \mathcal{S},\ \forall t \in \mathbb{N}$. By fixing the control as a function of the state, this system can be expressed without loss of generality as the dynamical process

$$x(t+1) = f_u(x(t)). \tag{3.6}$$

To certify that the system satisfies the safety constraint for a particular controller, it suffices to find a positively invariant subset of $\mathcal{S}$ with respect to the closed-loop system (3.6). A set $\mathcal{A}$ is said to be *positively invariant* with respect to a dynamic system with solution $x(t)$ if $x(\tau) \in \mathcal{A} \Rightarrow x(t) \in \mathcal{A},\ \forall t > \tau$. For a discrete-time system, positive invariance is equivalent to the condition $x(t) \in \mathcal{A} \Rightarrow x(t+1) \in \mathcal{A}$ by induction. If there exists some positively invariant set $\mathcal{A} \subseteq \mathcal{S}$, and if $x(\tau) \in \mathcal{A}$ for some time τ, then $x(t) \in \mathcal{S}$ for all times $t \geq \tau$.

3.2.2 Lattice Theory and the Banach-Tarski Theorem

It is useful to consider the positively invariant subsets of a set $\mathcal{S} \subseteq \mathcal{X}$ with respect to a function $f_u : \mathcal{X} \to \mathcal{X}$ within the framework of lattice theory. A partially ordered set $(\mathcal{P}, \leq)$ consists of a set $\mathcal{P}$, and a partial order $\leq$ over $\mathcal{P}$. Namely, $\leq$ is a binary relation over $\mathcal{P}$ which is reflexive ($a \leq a$), antisymmetric ($a \leq b,\ b \leq a \Rightarrow a = b$), and transitive ($a \leq b,\ b \leq c \Rightarrow a \leq c$). In this context, a *lattice* is defined to be a partially ordered set $(\mathcal{P}, \leq)$ for which any pair of elements of $\mathcal{P}$ have a unique greatest lower bound (also called "meet" or "infimum") and a unique least upper bound (also called "join" or "supremum") in $\mathcal{P}$. A lattice $(\mathcal{P}, \leq)$ is said to be complete if every subset of $\mathcal{P}$ has a greatest lower bound and least upper bound. For example, $(\mathbb{R}, \leq)$ is a lattice because the infimum/supremum of any finite set of real numbers is their minimum/maximum, while no element of $\mathbb{R}$ is the infimum/supremum of $\mathbb{R}$. However, the extended real line $\mathbb{R} \cup \{+\infty, -\infty\}$ does form a com-

plete lattice. Another example of a complete lattice is $(2^{\mathbb{R}^n}, \subseteq)$, whose infimum/supremum represents intersection/union, respectively..

A fundamental result in lattice theory is the Banach-Tarski theorem, which concerns the fixed points of monotone functions in a complete lattice. A function g is monotone if $a \leq b \Rightarrow g(a) \leq g(b)$, and x is a fixed point of g if $f(x) = x$.

Theorem 3.1 (Banach-Tarski Theorem [72]) *Consider the complete lattice $(\mathcal{L}, \leq)$, and let $g : \mathcal{L} \to \mathcal{L}$ be a monotone function. Then, the set of fixed points of g is a complete lattice with respect to $\leq$.*

3.3 Complete Lattice of Positively Invariant Sets

Positively invariant sets satisfy certain interesting properties, in particular, both the union and intersection of any collection of positively invariant sets is also positively invariant. This property can be seen as a consequence of the Banach-Tarski Theorem, and can be used to define a complete lattice over the positively invariant subsets of the constraint set $\mathcal{S}$.

Firstly, a set is positively invariant with respect to f iff it is a fixed point of the mapping $g(\mathcal{A}) := \mathcal{A} \cap f^{-1}(A)$. If $\mathcal{A}$ is positively invariant, then the set of points mapping onto $\mathcal{A}$ must contain $\mathcal{A}$ itself, else there would be some point $x \in \mathcal{A}$ such that $f(x) \notin \mathcal{A}$, and $\mathcal{A}$ would not be positively invariant. Next, notice that $(2^{\mathcal{S}}, \subseteq)$ is a complete lattice. The Banach-Tarski theorem then implies that the set of fixed points of g forms a complete lattice under the partial order $\subseteq$, which is the same as the set of positively invariant subsets of $\mathcal{S}$. This guarantees that there is a maximal positively invariant subset which is the supremum of all the fixed points of g among subsets of $\mathcal{S}$. Therefore if $x(t)$ is in some positively invariant subset of $\mathcal{S}$, then it is in the maximal positively invariant subset $\mathcal{O}_\infty(\mathcal{S}, f)$.

Since the positively invariant subsets of $\mathcal{S}$ are a complete lattice by the Banach-Tarski theorem, one only needs to check if $x(\tau) \in \mathcal{O}_\infty(\mathcal{S}, f)$ instead of searching over all possible positively invariant sets to prove the safety condition $x(t) \in \mathcal{S},\ \forall t \geq \tau$. Conversely, if the safety condition is satisfied, then $x(\tau)$ is in the positively invariant set $\{x(\tau), x(\tau+1), x(\tau+2), \dots\}$, which is a subset of $\mathcal{O}_\infty(\mathcal{S}, f)$ by maximality.

3.3.1 Kleene Iteration

Kleene's Fixed Point Theorem provides an algorithm for computing the extremal elements of the complete lattice of positively invariant sets.

Theorem 3.2 (Kleene's Fixed Point Theorem (Special Case) [47]) *Let $(\mathcal{L}, \subseteq)$ be a complete lattice, and let $g : \mathcal{L} \to \mathcal{L}$ be continuous. If $g(x) \leq x$, then $\bigcap_{k\in\mathbb{N}} g^k(x)$ is the greatest fixed point of g on $\{y \in \mathcal{S} \mid y \leq x\}$.*

The procedure for computing the maximal positively invariant set presented by Kolmanovsky and Gilbert [49] is essentially Kleene iteration starting from the set $\mathcal{S}$, and achieving a sequence of tighter upper bounds for $\mathcal{O}_\infty(\mathcal{S}, f)$. By defining $\mathcal{O}_t(\mathcal{S}, f) = \bigcap_{k=0}^{t} f^k(\mathcal{S})$, it is apparent that the sequence $\mathcal{O}_t(\mathcal{S}, f)$ is nonincreasing in t, and approaches $\mathcal{O}_\infty(\mathcal{S}, f)$ from above. This sequence also has the recursive definition $\mathcal{O}_0(\mathcal{S}, f) = \mathcal{S}$ $\mathcal{O}_{t+1}(\mathcal{S}, f) = \mathcal{S} \cap f(\mathcal{O}_t(\mathcal{S}, f))$. This form makes it clear that $\mathcal{O}_{t^*}(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{S}, f)$ iff $\mathcal{O}_{t^*}(\mathcal{S}, f) = \mathcal{O}_{t^*+1}(\mathcal{S}, f)$, leading to Algorithm 3.1 for computing $\mathcal{O}_\infty(\mathcal{S}, f)$ exactly.

$\mathcal{K}_0 \leftarrow \mathcal{S}$
$\mathcal{K}_1 \leftarrow \mathcal{S} \cap f^{-1}(\mathcal{K}_0)$
$k = 0$
while $\mathcal{K}_0 \not\subseteq \mathcal{K}_1$ **do**
  $\mathcal{K}_0 \leftarrow \mathcal{K}_1$
  $\mathcal{K}_1 \leftarrow \mathcal{S} \cap f^{-1}(\mathcal{K}_1)$
  $k \leftarrow k + 1$
end while
$\mathcal{O}_\infty(\mathcal{S}, f) \leftarrow \mathcal{K}_0$

Figure 3.1: Kleene iteration with initial set $\mathcal{S}$.

If $\mathcal{O}_{t+1}(\mathcal{S}, f) \neq \mathcal{O}_t(\mathcal{S}, f)$ for all $t \in \mathbb{N}$, then this algorithm does not terminate, and $\mathcal{O}_\infty(\mathcal{S}, f)$ is not finitely determined from Algorithm 3.1. However, if there is a $t^* \in \mathbb{N}$ such that $\mathcal{O}_{t^*+1}(\mathcal{S}, f) = \mathcal{O}_{t^*}(\mathcal{S}, f)$, then Algorithm 3.1 finds $\mathcal{O}_\infty(\mathcal{S}, f)$ exactly as

$$\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_{t^*}(\mathcal{S}, f) \tag{3.7}$$

$$= \bigcap_{t=0}^{t^*} f^{-t}(\mathcal{S}). \tag{3.8}$$

Algorithm 3.1 does not need to start with the set $\mathcal{S}$ to compute $\mathcal{O}_\infty(\mathcal{S}, f)$ exactly. Notice that if the algorithm were called to compute $\mathcal{O}_\infty(\mathcal{O}_\infty(\mathcal{S}, f), f)$, then the algorithm immediately terminates and returns $\mathcal{O}_\infty(\mathcal{S}, f)$. It can be shown that $\mathcal{O}_\infty(\mathcal{R}, f) = \mathcal{O}_\infty(\mathcal{S}, f)$ for any

$\mathcal{R}$ in the interval $[\mathcal{O}_\infty(\mathcal{S}, f), \mathcal{S}]$, and furthermore, the number of iterations is maximum for $\mathcal{R} = \mathcal{S}$ and may be strictly smaller if $\mathcal{R}$ is a tighter upper bound for $\mathcal{O}_\infty(\mathcal{S}, f)$.

3.4 Maximal Invariant Sets for MDPs

Due to the inherent stochasticity of MDPs, it is natural to express safety conditions on the probability distribution rather than the physical state. For example, a swarm may be guided according to MDP dynamics. Suppose each agent is following the same stochastic policy to move among a set of locations to be continually surveyed. A lower-level collision avoidance subroutine may fail if there are too many agents near each other, so a safety constraint may impose a limit on the expected proportion of agents at each location. This constrains the overall distribution $x(t)$ to satisfy the linear inequality $x(t) \leq d$, so that the constraint set takes the form $\mathcal{S} = \{x \in \Delta(n) \mid Gx \leq g\}$.

For a time-invariant MDP with fixed policy, the distribution represented by $x(t)$ evolves as an LTI system $x(t+1) = Mx(t)$ so that $f(x) = Mx$. The safety constraint $x(t) \in \mathcal{S},\ \forall t \in \mathbb{N}$ is then equivalent to $x(0) \in \mathcal{O}_\infty(\mathcal{S}, f)$. The maximal invariant set $\mathcal{O}_\infty(\mathcal{S}, f)$ can be exactly computed if the following two assumptions are made:

Assumption 3.1 *There exists a $v \in \Delta(n)$ such that $\rho(M - v\mathbf{1}^\intercal) < 1$.*

Assumption 3.2 *The v from Assumption 3.1 satisfies $Gv + \mathbf{1}\epsilon \leq g$ for some $\epsilon > 0$.*

The safety constraint $x(t) \in \mathcal{S}$ can be certified if the closed-loop Markov matrix M satisfies Assumptions 3.1 and 3.2. One systematic way to satisfy these assumptions is to first choose a vector v in the relative interior of $\mathcal{S}$. Next, define r as the elementwise square root of v, and let $Q = \mathrm{Diag}(r)$. Then enforce the following LMI constraints on M and the policy vector π:

$$-\gamma I \preceq Q^{-1}MQ - rr^\intercal \preceq \gamma I, \tag{3.9}$$

$$M = P\ \mathrm{Diag}(\pi) I_o^\intercal, \tag{3.10}$$

$$I_o \pi = \mathbf{1}, \tag{3.11}$$

$$\pi \geq 0. \tag{3.12}$$

The parameter γ must be strictly less than 1, and can be found by minimizing γ subject to the other constraints.

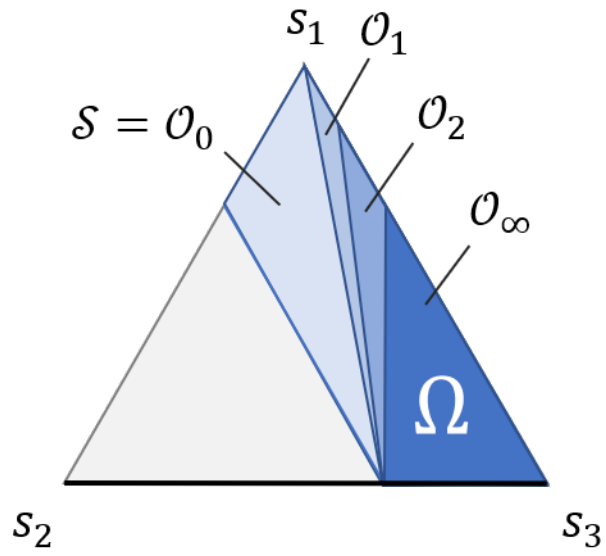

Figure 3.2: Sequence of sets $\mathcal{O}_t(\mathcal{S}, f)$ in the simplex with vertices s_1, s_2, s_3. The set Ω is an example of a set in the interval $[\mathcal{O}_\infty(\mathcal{S}, f), \mathcal{S}]$.

3.5 Safety Constraints

3.5.1 Introduction

Safety conditions are an important consideration for many dynamical systems. For example, an autonomous vehicle may minimize risk of failure by remaining in a safe operating regime, or a low-level control algorithm may be required to keep the state of the system sufficiently close to the nominal state used by a higher level motion planner. For these and other situations, safety constraints play a crucial role in simplifying the design and analysis of complex systems.

For this section, we developed a lattice-theoretic approach for verifying safety constraints, similar to what is used in formal verification of code. Specifically, we show that Kleene iteration explicitly computes the set of states which forever remain in the constraint set (often called the maximal invariant set because of the strong connection between safety constraints and positively invariant sets). The traditional method of computing this set is a special case of Kleene iteration. We propose a similar specialized algorithm for discrete-time Markov chains with polyhedral constraint sets based on an exact characterization of their omega-limit sets, and we show that for this application, the new algorithm exactly computes

the maximal invariant set over a strictly larger class of dynamics than the previous algorithm.

3.5.2 Related Literature

Many authors have described the applications and computation of invariant sets in the context of set-theoretic control [11] [12] [13] [49] [68] [26]. In this paper, we focus only on discrete-time, time-invariant dynamic systems with no explicit control or disturbance. Hennet showed that in the control-free LTI case, positive invariance of a polyhedron is equivalent to solving a linear feasibility problem [39], and Kolmanovsky and Gilbert considered the maximal and minimal positively invariant sets for LTI systems with bounded disturbances [49]. Various authors such as Hirata and Raković have considered polynomial systems with semi-algebraic constraints [42] [69]. Raković was able to approximate the maximal invariant set from below with positively invariant sets [68]. All of these approaches can be understood in a lattice-theoretic framework as well. The Banach-Tarski theorem [72] is a crucial result for complete lattices, guaranteeing that the fixed points of any monotone function form another complete lattice. Kleene's fixed point theorem [47] provides an algorithm for computing these fixed points, in this case the greatest fixed point of a mapping whose fixed points are precisely the positively invariant sets for the dynamical system.

3.5.3 Definitions

A partially ordered set (or poset) is a tuple $(\mathcal{X}, \leq)$ consisting of a set $\mathcal{X}$ and a partial order $\leq$. A complete lattice is a poset where all subsets admit a least upper bound and a greatest lower bound, hence a supremum and an infimum. The greatest fixed point of self-map $g : \mathcal{X} \to \mathcal{X}$ with respect to $(\mathcal{X}, \leq)$ is denoted $\mathrm{gfp}_{\mathcal{X}}^{\leq} g$. Given a dynamic system $x^+ = f(x)$, the sequence $(x_k)_{k\in\mathbb{N}}$ is a *trajectory* if $x_0 \in \mathcal{X}$, and $x_{k+1} = f(x_k)$ for all $k \in \mathbb{N}$. For the purpose of this paper, a *safety condition* has the form $x_k \in \mathcal{S}\ \forall k \in \mathbb{N}$, where $\mathcal{S}$ is the *constraint set*.

3.6 Invariant Sets and Safety

Throughout this section, we consider a general dynamic system

$$x^+ = f(x), \tag{3.13}$$

where $f : \mathcal{X} \to \mathcal{X}$. A set $\mathcal{A}$ is *positively invariant* with respect to (3.13) if $x \in \mathcal{A} \Rightarrow f(x) \in \mathcal{A}$. Equivalently, $\mathcal{A}$ is positively invariant if $\mathcal{A} \subseteq f^{-1}(\mathcal{A})$, i.e., $\mathcal{A} \subseteq \{x \mid f(x) \in \mathcal{A}\}$. This property

ensures that if $x \in \mathcal{A}$, then $f(x) \in \mathcal{A}$, $f(f(x)) \in \mathcal{A}$, and so on such that if $x \in \mathcal{A}$, then $f^k(x) \in \mathcal{A}$, $\forall k \in \mathbb{N}$. The safety condition $f^k(x) \in \mathcal{S}$, $\forall k \in \mathbb{N}$ therefore can be verified by proving that x is in some particular positively invariant subset of $\mathcal{S}$.

Our goal is to characterize the subset of $\mathcal{X}$ for which $f^k(x) \in \mathcal{S}$, $\forall k \in \mathbb{N}$, by showing that the poset of positively invariant subsets of $\mathcal{X}$ satisfy the assumptions of Kleene's Fixed Point Theorem [47]. Then, Kleene iteration can compute a nonincreasing sequence of outer approximations to the set of x satisfying the safety condition, denoted $\mathcal{O}_\infty(\mathcal{S}, f)$, which is also the largest positively invariant subset of $\mathcal{S}$. Hence our main objective is to compute $\mathcal{O}_\infty(\mathcal{S}, f)$, which will be accomplished by Kleene iteration.

3.6.1 Kleene Iteration from $\mathcal{S}$ - Existing algorithm for maximal invariant sets

Here we establish the lattice theoretic basis of our methodology by using two important theorems. We first use Tarski's fixed point theorem to show that the invariant subsets of $\mathcal{S}$ form a complete lattice. Then this result is used to establish the algorithm to compute the largest invariant set by using Kleene's fixed point theorem.

Theorem 3.3 (Tarski's Fixed Point Theorem [72]) *Consider the complete lattice $(\mathcal{L}, \leq)$, and let $g : \mathcal{L} \to \mathcal{L}$ be a monotone function, i.e., $f(x) \leq f(y)$, $\forall x \leq y$. Then, the set of fixed points of g is a complete lattice with respect to $\leq$.*

Applying Tarski's Fixed Point Theorem to the complete lattice $(2^{\mathcal{X}}, \subseteq)$, the fixed points of any monotone self-map over the subsets of $\mathcal{X}$ will also form a complete lattice with the same partial order $\subseteq$. What remains to be shown is that there exists a self-map whose fixed points are precisely the positively invariant subsets of $\mathcal{X}$. This is indeed the case, as proven by the following lemma with the monotone self-map $g(\mathcal{A}) = \mathcal{A} \cap f^{-1}(\mathcal{A})$.

Lemma 3.1 *Let $g(\mathcal{A}) = \mathcal{A} \cap f^{-1}(\mathcal{A})$. $\mathcal{A}$ is positively invariant with respect to (3.13) iff $\mathcal{A}$ is a fixed point of g.*

Proof: Suppose $\mathcal{A}$ is positively invariant. If $x \in \mathcal{A}$, then $f(x) \in \mathcal{A}$, and so $x \in f^{-1}(\mathcal{A})$ by definition. This implies that $\mathcal{A} \subseteq f^{-1}(\mathcal{A})$, and therefore $\mathcal{A} = \mathcal{A} \cap f^{-1}(\mathcal{A}) = g(\mathcal{A})$.

The converse is also true because if $\mathcal{A}$ is a fixed point of g, then $g(\mathcal{A}) = \mathcal{A} \cap f^{-1}(\mathcal{A}) = \mathcal{A}$, which is true only if $\mathcal{A} \subseteq f^{-1}(\mathcal{A})$, or equivalently, $x \in \mathcal{A} \Rightarrow f(x) \in \mathcal{A}$. This shows that $\mathcal{A}$ is positively invariant. ■

Corollary 3.1 *The positively invariant subsets of $\mathcal{X}$ for system (3.13) are a complete lattice with respect to the partial order $\subseteq$.*

Proof: Consider the poset $(2^{\mathcal{X}}, \subseteq)$. The least upper bound and greatest lower bound for any nonempty subset of $2^{\mathcal{X}}$ are respectively the union and intersection operations, forming a complete lattice. f^{-1} is monotone because if $\mathcal{A} \subseteq \mathcal{B}$, then $x \in f^{-1}(\mathcal{A}) \Rightarrow f(x) \in \mathcal{A} \Rightarrow f(x) \in \mathcal{B} \Rightarrow x \in f^{-1}(\mathcal{B})$. The function $g(\mathcal{A}) = \mathcal{A} \cap f^{-1}\mathcal{A}$ with domain $2^{\mathcal{X}}$ therefore is also monotone, since if $\mathcal{A} \subseteq \mathcal{B}$, then $g(\mathcal{A}) = \mathcal{A} \cap f^{-1}(\mathcal{A}) \subseteq \mathcal{B} \cap f^{-1}(\mathcal{A}) \subseteq \mathcal{B} \cap f^{-1}(\mathcal{B}) = g(\mathcal{B})$. Tarski's Fixed Point Theorem guarantees that the fixed points of g, i.e., the positively invariant subsets of $\mathcal{X}$, form a complete lattice with partial order $\subseteq$. ■

Kleene's fixed point theorem provides an algorithm for computing the least and greatest fixed points of a function over a complete lattice. Hence it provides an algorithm that computes the smallest and largest invariant sets for (3.13) in $\mathcal{S}$. The relevant special case of Kleene's fixed point theorem is given next.

Theorem 3.4 (Kleene Fixed Point Iteration Theorem [9][47]) *Let $(\mathcal{L}, \leq)$ be a complete lattice, and let $g : \mathcal{L} \to \mathcal{L}$ be a continuous monotone set-valued map. Then $\bigcap_{k \in \mathbb{N}} g^k(x)$ is the greatest fixed point of g on $\{y \in \mathcal{L} \mid y \leq x\}$.*

Following the principles of Abstract Interpretation [20], we can compute this invariant set $\mathcal{O}_\infty(\mathcal{S}, f) = \mathrm{gfp}^{\leqslant}_{\mathcal{S}} f^{-1}$ as the limit of the decreasing sequence $(\mathcal{S},\ \mathcal{S} \cap f^{-1}\mathcal{S},\ S \cap f^{-1}\mathcal{S} \cap f^{-2}\mathcal{S},\ \dots)$. Indeed, if $\mathcal{S} \subseteq \mathcal{S} \cap f^{-1}\mathcal{S}$, then the greatest element of $\{\mathcal{R} \in 2^{\mathcal{S}} \mid \mathcal{R} \subseteq \mathcal{R} \cap f^{-1}\mathcal{R}\}$ is $\mathcal{S}$, and $\mathcal{S} = \mathcal{O}_\infty(\mathcal{S}, f)$. Otherwise, assuming that $\mathcal{S}$ is strictly greater than $\mathcal{S} \cap f^{-1}\mathcal{S}$, one obtains $f^{-1}\mathcal{S} \supseteq f^{-1}\mathcal{S} \cap f^{-2}\mathcal{S}$ by monotonicity of $\cap$ and of f^{-1}, and we can build the decreasing sequence. If one works in a complete lattice—such as $(2^{\mathcal{S}}, \subseteq)$—then the limit of that sequence exists.

Since the maximal invariant set coincides with the set of safe initial conditions, it can be formally constructed by considering the sequence of sets for which the state remains in $\mathcal{S}$ for all times $k = 0, \dots, m$ and then letting m approach infinity. Let $\mathcal{O}_m(\mathcal{S}, f)$ be defined as follows:

$$\mathcal{O}_m(\mathcal{S}, f) = \{x_0 \in \mathcal{X} \mid f^k(x_0) \in \mathcal{S},\ \forall k \in \mathbb{N}_{\leq m}\}, \tag{3.14}$$

$$= \bigcap_{k=0}^{m} f^{-k}\mathcal{S}. \tag{3.15}$$

Since this is a nonincreasing sequence, the limit as $m \to \infty$ is simply the intersection over all $\mathcal{O}_m(\mathcal{S}, f)$. Letting $g(\mathcal{O}) = \mathcal{O} \cap f^{-1}\mathcal{O}$, this sequence can be expressed recursively:

$$\mathcal{O}_0(\mathcal{S}, f) = \mathcal{S}, \tag{3.16}$$

$$\mathcal{O}_{m+1}(\mathcal{S}, f) = g(\mathcal{O}_m(\mathcal{S}, f)), \ \forall m \in \mathbb{N}, \tag{3.17}$$

$$\mathcal{O}_\infty(\mathcal{S}, f) = \bigcap_{m \in \mathbb{N}} \mathcal{O}_m(\mathcal{S}, f). \tag{3.18}$$

Since a set $\mathcal{O}$ is positively invariant iff $\mathcal{O} = g(\mathcal{O})$, the previous equations show that if $\mathcal{O}_{m^*}(\mathcal{S}, f)$ is positively invariant for some $m^* \in \mathbb{N}$, then $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_{m^*}(\mathcal{S}, f)$. This fact is used as the stopping criterion for the following algorithm to compute the maximal invariant set [42].

$\mathcal{K}_0 \leftarrow \mathcal{S}$
$\mathcal{K}_1 \leftarrow \mathcal{S} \cap f^{-1}(\mathcal{K}_0)$
$k = 0$
while $\mathcal{K}_0 \not\subseteq \mathcal{K}_1$ **do**
  $\mathcal{K}_0 \leftarrow \mathcal{K}_1$
  $\mathcal{K}_1 \leftarrow \mathcal{S} \cap f^{-1}(\mathcal{K}_1)$
  $k \leftarrow k + 1$
end while
$\mathcal{O}_\infty(\mathcal{S}, f) \leftarrow \mathcal{K}_0$

Figure 3.3: Kleene iteration with initial set $\mathcal{S}$.

Algorithm 3.3 recursively computes $\mathcal{O}_m(\mathcal{S}, f)$, stored as the variable $\mathcal{K}_0$, until $\mathcal{K}_0$ is positively invariant. Positive invariance is determined by $\mathcal{K}_0$ being a fixed point of the mapping $g(\mathcal{K}) = \mathcal{K} \cap f^{-1}(\mathcal{K})$. Since g is non-increasing, $\mathcal{K}_0$ is a fixed point of g if and only if $\mathcal{K}_0 \subseteq g(\mathcal{K}_0) = \mathcal{K}_1$. This algorithm terminates only if the condition for positive invariance $\mathcal{K}_0 \subseteq \mathcal{K}_1$ is satisfied, in which case $\mathcal{O}_\infty(\mathcal{S}, f)$ is said to be *finitely determined* by Algorithm 3.3. Otherwise, the sequence of $\mathcal{O}_m(\mathcal{S}, f)$ approaches $\mathcal{O}_\infty(\mathcal{S}, f)$ asymptotically from above, and this algorithm must be terminated to give an outer approximation for the maximal invariant set. Finite determination is not guaranteed even in the case of linear dynamics and polyhedral $\mathcal{S}$, but sufficient conditions exist for special cases [36][49]. Algorithm 3.3 is therefore not sufficient for computing the maximal invariant set exactly, i.e., finite termination is

not always guaranteed.

3.6.2 Kleene Iteration from $\Omega(\mathcal{S}, f)$ - Proposed algorithm for maximal invariant sets

By initializing Kleene iteration with a tighter outer approximation of $\mathcal{O}_\infty(\mathcal{S}, f)$, Kleene iteration can potentially compute $\mathcal{O}_\infty(\mathcal{S}, f)$ more efficiently. For example, if $\mathcal{K}_0$ was initialized with $\mathcal{O}_\infty(\mathcal{S}, f)$ rather than $\mathcal{S}$, we would immediately detect that $\mathcal{K}_0 \subseteq \mathcal{K}_1$, and therefore that K_0 was positively invariant. Kleene's Fixed Point Theorem guarantees that we can initialize the iteration with any set $\mathcal{R}$ such that $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{R} \subseteq \mathcal{S}$, and the limit of Kleene iteration will still be the maximal invariant subset of $\mathcal{S}$. We can therefore generalize Algorithm 3.3 the following results.

Lemma 3.2 *Let $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{R} \subseteq \mathcal{S}$. Then, $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f)$, $\forall m \in \mathbb{N}$.*

Proof: The rightmost inequality $\mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f)$, $\forall m \in \mathbb{N}$ is shown by induction. The base case $m = 0$ is satisfied since $\mathcal{O}_0(\mathcal{R}, f) = \mathcal{R} \subseteq \mathcal{S} = \mathcal{O}_0(\mathcal{S}, f)$. Assuming the induction hypothesis $\mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f)$ is true for some $m \in \mathbb{N}$, then since the mapping $\mathcal{O} \to \mathcal{O} \cap f^{-1}\mathcal{O}$ is monotone, $\mathcal{O}_{m+1}(\mathcal{R}, f) = \mathcal{O}_m(\mathcal{R}, f) \cap f^{-1}\mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f) \cap f^{-1}\mathcal{O}_m(\mathcal{S}, f) = \mathcal{O}_{m+1}(\mathcal{S}, f)$, completing the proof by induction.

The inequality $\mathcal{O}_\infty(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{R}, f)$, $\forall m \in \mathbb{N}$ follows immediately from Equation (3.18). Using Equation (3.18) and the fact that $\mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f)$ for all $m \in \mathbb{N}$, $\mathcal{O}_\infty(\mathcal{R}, f) = \bigcap_{m\in\mathbb{N}} \mathcal{O}_m(\mathcal{R}, f) \subseteq \bigcap_{m\in\mathbb{N}} \mathcal{O}_m(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{S}, f)$. Therefore, $\mathcal{O}_\infty(\mathcal{R}, f) \subseteq \mathcal{O}_\infty(\mathcal{S}, f)$.

To show that $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{O}_\infty(\mathcal{R}, f)$, consider any $x \in \mathcal{O}_\infty(\mathcal{S}, f)$. Since we had assumed that $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{R}$, $f^m(x) \in \mathcal{O}_\infty(\mathcal{S}, f)$, $\forall m \in \mathbb{N} \Rightarrow f^m(x) \in \mathcal{R}$, $\forall m \in \mathbb{N} \Rightarrow x \in \mathcal{O}_\infty(\mathcal{R}, f)$. This shows that also $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{O}_\infty(\mathcal{R}, f)$, proving the equality $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{R}, f)$. ■

Theorem 3.5 shows that by initializing Algorithm 3.3 with a tighter outer bound of $\mathcal{O}_\infty(\mathcal{S}, f)$ than $\mathcal{S}$, the number of iterations to compute the maximal invariant set is either lower or the same.

Theorem 3.5 *Let $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{R} \subseteq \mathcal{S}$, and let l^* and m^* be the smallest elements of $\mathbb{N} \cup \{\infty\}$ for which $\mathcal{O}_{l^*}(\mathcal{R}, f) \subseteq f^{-1}\mathcal{O}_{l^*}(\mathcal{R}, f)$ and $\mathcal{O}_{m^*}(\mathcal{S}, f) \subseteq f^{-1}\mathcal{O}_{m^*}(\mathcal{S}, f)$, respectively. Then, $l^* \leq m^*$, and $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_{l^*}(\mathcal{R}, f)$.*

Proof: If $\mathcal{O}_\infty(\mathcal{S}, f)$ is finitely determined, then m^* is finite. Otherwise, $m^* = \infty$. By Lemma 3.2, $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{R}, f) \subseteq \mathcal{O}_m(\mathcal{S}, f)$ for all $m \in \mathbb{N} \cup \{\infty\}$. Because $\mathcal{O}_{m^*}(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{S}, f)$, setting $m = m^*$ gives $\mathcal{O}_\infty(\mathcal{S}, f) = \mathcal{O}_\infty(\mathcal{R}, f) = \mathcal{O}_{m^*}(\mathcal{R}, f)$. This also shows that m^* is an upper bound for l^*, since l^* is the minimum l for which $\mathcal{O}_\infty(\mathcal{R}, f) = \mathcal{O}_l(\mathcal{R}, f)$. ■

One outer approximation of $\mathcal{O}_\infty(\mathcal{S}, f)$ is obtained by considering the points in $\mathcal{S}$ whose omega-limit sets under f are entirely within $\mathcal{S}$. The omega-limit set $\omega(x, f)$ is defined as

$$\omega(x, f) = \bigcap_{k \in \mathbb{N}} \overline{\{f^l(x) \mid l \geq k,\ l \in \mathbb{N}\}}, \tag{3.19}$$

where $f : \mathcal{X} \to \mathcal{X}$, and the over bar denotes the set closure with respect to the topological space containing $\mathcal{X}$. We also define

$$\Omega(\mathcal{S}, f) = \{x \in \mathcal{S} \mid \omega(x, f) \subseteq \mathcal{S}\}. \tag{3.20}$$

As a simple case, suppose the system $x^+ = f(x)$ has an equilibrium point x_e with basin of attraction $\mathcal{A}$. For each initial condition $x_0 \in \mathcal{A}$, the state asymptotically approaches x_e, so $\omega(x_0, f) = \{x_e\}$. If $x_e \in \mathcal{S}$, then $\Omega(\mathcal{S}, f) = \{x \in \mathcal{S} \mid \omega(x, f) \subseteq \mathcal{S}\} \supseteq \{x \in \mathcal{S} \mid \omega(x, f) = \{x_e\}\} = \mathcal{S} \cap \mathcal{A}$. However if instead $x_e \notin \mathcal{S}$, then $\mathcal{A}$ and $\Omega(\mathcal{S}, f)$ are disjoint because the omega-limit sets of all states in $\mathcal{A}$ are $\{x_e\}$, which is not in $\mathcal{S}$. More generally, certain initial conditions for the system $x^+ = f(x)$ may approach a limit cycle or some other complex omega-limit set such as a strange attractor, rather than a single equilibrium point. However, computing the set $\Omega(\mathcal{S}, f)$ for nonlinear systems is beyond the scope of this paper, and we will focus on computing this set for Markov chains in §3.7.

The main motivation for the definition of $\Omega(\mathcal{S}, f)$ is to characterize a tighter outer approximation for $\mathcal{O}_\infty(\mathcal{S}, f)$ than $\mathcal{S}$, and which can be computed exactly for Markov chains with safety constraints on the probability distribution. Lemma 3.3 states that $\Omega(\mathcal{S}, f)$ is an upper bound for the maximal invariant set, and therefore meets the requirements for the set "$\mathcal{R}$" of Theorem 3.5.

Lemma 3.3 *Let $f : \mathcal{X} \to \mathcal{X}$ and let $\mathcal{S}$ be a closed subset of $\mathcal{X}$. Then $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \Omega(\mathcal{S}, f) \subseteq \mathcal{S}$.*

Proof: The relation $\Omega(\mathcal{S}, f) \subseteq \mathcal{S}$ is apparent from the definition of $\Omega(\mathcal{S}, f)$. For any $x \in \mathcal{O}_\infty(\mathcal{S}, f)$, $f^l(x) \in \mathcal{O}_\infty(\mathcal{S}, f)$ for all $l \in \mathbb{N}$. Since $\mathcal{O}_\infty(\mathcal{S}, f) \subseteq \mathcal{S}$, and $\mathcal{S}$ is closed,

$\omega(x,f) \subseteq \overline{\{f^l(x) \mid l \in \mathbb{N}\}} \subseteq \overline{\mathcal{O}_\infty(\mathcal{S},f)} \subseteq \overline{\mathcal{S}} = \mathcal{S}$. Therefore, $x \in \Omega(\mathcal{S},f)$. ■
We then suggest the following algorithm for determining $\mathcal{O}_\infty(\mathcal{S},f)$ exactly.

$\mathcal{K}_0 \leftarrow \Omega(\mathcal{S},f)$
$\mathcal{K}_1 \leftarrow \Omega(\mathcal{S},f) \cap f^{-1}(\mathcal{K}_0)$
while $\mathcal{K}_0 \nsubseteq \mathcal{K}_1$ **do**
  $\mathcal{K}_0 \leftarrow \mathcal{K}_1$
  $\mathcal{K}_1 \leftarrow \Omega(\mathcal{S},f) \cap f^{-1}(\mathcal{K}_1)$
end while
$\mathcal{O}_\infty(\mathcal{S},f) \leftarrow \mathcal{K}_0$

Figure 3.4: Kleene iteration with improved initial set $\Omega(\mathcal{S},f)$.

Algorithm 3.4 computes $\mathcal{O}_\infty(\mathcal{S},f)$ in no more iterations than Algorithm 3.3 does, and even has finite termination in some cases where Algorithm 3.3 does not. The set $\Omega(\mathcal{S},f)$ must be pre-computed from $\mathcal{S}$ and f. To the authors' knowledge, there is no known method to compute this set efficiently for general systems, which may exhibit chaotic behavior. However, as we will show in §3.7, $\Omega(\mathcal{S},f)$ can be computed exactly for a Markov chain and for $\mathcal{S}$ that can be described by semi-algebraic constraints.

3.6.3 A Bound on the Number of Iterations

We make the following assumptions to bound the analytical complexity of Algorithm 3.4.

Assumption 3.3 (Uniform Exponential Decay of Transient Behavior) *There exist a* $C > 0$ *and a* $\gamma \in (0,1)$ *such that* $d(f^k(x), \omega(x,f)) \leq C\gamma^k$, $\forall x \in \Omega(\mathcal{S},f)$, $\forall k \in \mathbb{N}$.

Assumption 3.4 (Uniform Separation Between $\omega(x,f)$ and $\mathcal{S}^c$) *There exists an* $r > 0$ *such that* $\omega(x,f) \oplus \mathcal{B}_r \subseteq \mathcal{S}$, *for all* $x \in \Omega(\mathcal{S},f)$.

Assumptions 3.3 and 3.4 are generalizations of those in [45]. For example, all ergodic Markov chains have a unique asymptotically stable stationary distribution v, so $\omega(x,f) = \{v\}$ for all $x \in \Delta(n)$. The ball $\mathcal{B}_r$ is understood to be a subset of the domain of f. The exact value of r may be determined after the omega-limit sets are characterized. In the ergodic case, r is the distance from the stationary distribution to the boundary of the set $\mathcal{S}$.

For Lyapunov stable linear systems of the form $x^+ = Ax$, the eigenvalues of A can always be separated into those with magnitude 1, and those with magnitude strictly less than γ, for some $\gamma < 1$. Likewise, the sequence of A^k can be expressed as the sum of a steady-state component, and an exponentially decaying transient component. This property is evident from the Jordan decomposition of A^k:

$$A^k = \begin{bmatrix} V_1 & \cdots & V_p \end{bmatrix} \begin{bmatrix} J_1^k & & \\ & \ddots & \\ & & J_p^k \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_p \end{bmatrix}, \tag{3.21}$$

$$= \bar{V}_1 \bar{J}_1^k \bar{W}_1 + \bar{V}_2 \bar{J}_2^k \bar{W}_2, \tag{3.22}$$

where $\bar{J}_1$ is the direct sum of the Jordan blocks with eigenvalues on the unit circle, and the eigenvalues of $\bar{J}_2$ have magnitude less than γ. Since $\lim_{k\to\infty} A^k x - A^k \bar{V}_1 \bar{W}_1 x = \lim_{k\to\infty} \bar{V}_2 \bar{J}_2^k \bar{W}_2 x = 0$, the ω-limit set for the original system is the same as that of the projected system $\omega(x, A) = \omega(\bar{V}_1 \bar{W}_1 x, A)$, which has the property $A^k \bar{V}_1 \bar{W}_1 x \in \omega(\bar{V}_1 \bar{W}_1 x, A),\ \forall k \in \mathbb{N}$. Furthermore, if $\mathcal{S}$ is bounded, the transient part satisfies Assumption (3.3) by the existence of a $\bar{C}$ such that $\left|\left|\left| \bar{J}_2^k \right|\right|\right| \leq \bar{C}\gamma^k$ for some induced norm $|||\cdot|||$, because $\|\bar{V}_2 \bar{J}_2^k \bar{W}_2 x\| \leq \left|\left|\left| \bar{V}_2 \right|\right|\right| \left|\left|\left| \bar{W}_2 \right|\right|\right| \|x\| \bar{C}\gamma^k \leq C\gamma^k$.

Lemma 3.4 *Let function f and set $\mathcal{S}$ satisfy Assumptions 3.3 and 3.4. Then, $f^k(x) \in \mathcal{S}, \forall k \geq \log_\gamma(r/C),\ \forall x \in \Omega(\mathcal{S}, f)$.*

Proof: Let $k \geq \log_\gamma(r/C)$. Since $C\gamma^k$ is a decreasing function of k, $C\gamma^k \leq C\gamma^{\log_\gamma(r/C)} = r$. For all $x \in \Omega(\mathcal{S}, f)$, $d(f^k(x), \omega(x, f)) \leq C\gamma^k \leq r$. This implies that $f^k(x) \in \omega(x) \oplus \mathcal{B}_r$, and therefore that $f^k(x) \in \mathcal{S}$. ■

Theorem 3.6 *Consider the system $x^+ = f(x)$ and the set $\mathcal{S}$ satisfying Assumptions 3.3 and 3.4. Then, Algorithm 3.4 computes $\mathcal{O}_\infty(\mathcal{S}, f)$ in no more than* $\max\{0, \lceil \log_\gamma(r/C) \rceil\}$ *iterations.*

Proof: Let C, γ, and r be the constants described in Assumptions 3.3 and 3.4, and let $\bar{m} = \lceil \log_\gamma(r/C) \rceil$. For any $x \in \Omega(\mathcal{S}, f)$, Lemma 3.4 guarantees $f^{\bar{m}+k}(x) \in \mathcal{S}$ for all $k \in \mathbb{N}$, i.e., $f^{\bar{m}}(x) \in \mathcal{O}_\infty(\mathcal{S}, f)$, or equivalently $x \in f^{-\bar{m}}(\mathcal{O}_\infty(\mathcal{S}, f))$. Therefore, $\Omega(\mathcal{S}, f) \subseteq f^{-\bar{m}}\mathcal{O}_\infty(\mathcal{S}, f)$. The $(\bar{m}-1)^{\text{st}}$ iteration of $\mathcal{K}_0$ has the explicit form $\mathcal{K}_0^{[\bar{m}-1]} = \Omega(\mathcal{S}, f) \cap \cdots \cap f^{-(\bar{m}-1)}\Omega(\mathcal{S}, f)$, so that $\mathcal{K}_1^{[\bar{m}-1]} = \Omega(\mathcal{S}, f) \cap \mathcal{K}_0^{[\bar{m}-1]} = \Omega(\mathcal{S}, f) \cap \cdots \cap f^{-\bar{m}}\Omega(\mathcal{S}, f) = \Omega(\mathcal{S}, f) \cap \cdots \cap f^{-(\bar{m}-1)}\Omega(\mathcal{S}, f) =$

$\mathcal{K}_0^{[\bar{m}-1]}$, because $\Omega(\mathcal{S},f) \subseteq f^{-\bar{m}}\mathcal{O}_\infty(\mathcal{S},f) \subseteq f^{-\bar{m}}\Omega(\mathcal{S},f)$. Since Lemma 3.2 guarantees that $\mathcal{O}_\infty(\mathcal{S},f) \subseteq \mathcal{K}_0^{[\bar{m}-1]}$, and since $\mathcal{K}_0^{[\bar{m}-1]} = \mathcal{K}_1^{[\bar{m}-1]}$ proves that $\mathcal{K}_0^{[\bar{m}-1]}$ is a positively invariant subset of $\mathcal{S}$, we conclude that $\mathcal{K}_0^{[\bar{m}-1]} = \mathcal{O}_\infty(\mathcal{S},f)$. ■

Note the Assumptions 3.3 and 3.4 are not necessary to compute $\mathcal{O}(\mathcal{S},f)$ in finite time with Algorithm 3.4. For example, if $f(x) = x$, then any set $\mathcal{S}$ is positively invariant, and so $\mathcal{O}_\infty(\mathcal{S},f) = \mathcal{S}$. However, for any x on the boundary of $\mathcal{S}$, $\omega(x,f) = \{x\}$ will also be on the boundary of $\mathcal{O}_\infty(\mathcal{S},f)$, and so there is no $r > 0$ satisfying Assumption 3.4. Nevertheless, $d(f^k(x),\{x\}) = d(x,\{x\}) = 0$. The following theorem provides a more general result which applies to this case.

Theorem 3.7 *Let $\mathcal{S}$ be closed. If $\Omega(\mathcal{S},f) \subseteq f^{-m}\Omega(\mathcal{S},f)$ for some $m \in \mathbb{N}_{\geq 1}$, then $\mathcal{O}_\infty(\mathcal{S},f) = \bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f)$.*

Proof: First, we show that $\mathcal{O}_\infty(\mathcal{S},f) \subseteq \bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f)$. By Lemma 3.3, $\mathcal{O}_\infty(\mathcal{S},f) \subseteq \Omega(\mathcal{S},f)$. Since $\mathcal{O}_\infty(\mathcal{S},f)$ is positively invariant, $\mathcal{O}_\infty(\mathcal{S},f) \subseteq f^{-k}\mathcal{O}_\infty(\mathcal{S},f)$, $\forall k \in \mathbb{N}$. Therefore, since f^{-k} is a monotone mapping of sets, $\mathcal{O}_\infty(\mathcal{S},f) = \bigcap_{k=0}^{m-1} f^{-k}\mathcal{O}_\infty(\mathcal{S},f) \subseteq \bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f)$.

We now prove the reverse relation $\bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f) \subseteq \mathcal{O}_\infty(\mathcal{S},f)$. Recall that a set $\mathcal{A}$ is positively invariant if $\mathcal{A} \subseteq f^{-1}\mathcal{A}$. Let $\mathcal{A} = \bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f)$. $\mathcal{A} = \left(\bigcap_{k=1}^{m-1} f^{-k}\Omega(\mathcal{S},f)\right) \cap \Omega(\mathcal{S},f) \subseteq \left(\bigcap_{k=1}^{m-1} f^{-k}\Omega(\mathcal{S},f)\right) \cap f^{-m}\Omega(\mathcal{S},f) = \bigcap_{k=1}^{m} f^{-k}\Omega(\mathcal{S},f) = f^{-1}\left(\bigcap_{k=0}^{m-1} f^{-k}\Omega(\mathcal{S},f)\right) = f^{-1}\mathcal{A}$. Since $\mathcal{A} \subseteq f^{-1}\mathcal{A}$, $\mathcal{A}$ is positively invariant. Since $\mathcal{A}$ is a positively invariant subset of $\mathcal{S}$, it must be a subset of the maximal invariant subset of $\mathcal{S}$ due to Corollary (3.1), and therefore $\mathcal{A} \subseteq \mathcal{O}_\infty(\mathcal{S},f)$. ■

3.7 Markov Chains with Polyhedral Constraints

To implement Algorithm 3.4, we must compute $\Omega(\mathcal{S},f)$, compute $f^{-1}\mathcal{K}$, take the intersection of $\Omega(\mathcal{S},f)$ with $f^{-1}\mathcal{K}$, and determine whether $\mathcal{K}_0 \subseteq \mathcal{K}_1$. These computations can be costly in general, particularly if $\mathcal{S}$ is not a polyhedron. For example, the relatively simple problem of checking if an ellipse contains a polyhedron expressed in facet form is NP-complete [32]. However, each individual computation of Algorithm 3.4 can be performed in polynomial time for Markov chain dynamics and polyhedral $\mathcal{S}$, which is the focus of this section.

Any time-invariant, discrete-time, finite-dimensional Markov chain can be expressed as the linear system

$$x(k+1) = Mx(k), \tag{3.23}$$

where $x(0) \in \Delta(n)$, $f(x) = Mx$, and M is an $n \times n$ *Markov matrix*, i.e., $M \geq 0$ and $\mathbf{1}^\intercal M = \mathbf{1}^\intercal$. Perron-Frobenius theory ensures the eigenvalues of all Markov matrices are in the closed unit disk. We write the Jordan decomposition of M as

$$M = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}, \tag{3.24}$$

$$= V_1 J_1 W_1 + V_2 J_2 W_2, \tag{3.25}$$

where the block matrices J_1 and J_2 share the eigenvalues of M with modulus equal to 1, and strictly less than 1, respectively. We also represent M as the sum $M_1 + M_2$, defined as

$$M_1 = V_1 J_1 W_1, \tag{3.26}$$

$$M_2 = V_2 J_2 W_2. \tag{3.27}$$

Since $W = V^{-1}$, $W_1 V_2 = 0$ and $W_2 V_1 = 0$. This observation implies that M_1 and M_2 have the property

$$M_1 M_2 = M_2 M_1 = 0. \tag{3.28}$$

In the case where all eigenvalues of M have modulus 1, we define $M_1 = M$ and $M_2 = 0$. Every eigenvalue of a Markov matrix on the unit circle is semisimple and a root of unity [60]. Therefore, $J_1^p = I$ for some positive integer p, and so $M_1^{k+p} = V_1 J_1^{k+p} W_1 = V_1 J_1^k W_1 = M_1^k$ for all $k \geq 1$.

Theorem 3.8 *If M is a Markov matrix with M_1 given by Equation (3.26), and if p is the smallest positive integer such that $M_1 = M_1^{1+p}$ then $\omega(x, M) = \{M_1^k x \mid k \in \{1, \ldots, p\}\}$.*

Proof: A state y is in the ω-limit set $\omega(x, M)$ iff there exists an infinite subsequence $(M^{\sigma(1)}x,\ M^{\sigma(2)}x, \ldots)$ that converges to y. If $y = M_1^k x$ for some positive integer k no greater than p, then the subsequence $(M^k x, M^{k+p}x, M^{k+2p}x, \ldots)$ converges to $M_1^k x$, since $M^{k+ip}x = V_1 J_1^{k+ip} W_1 x + V_2 J_2^{k+ip} W_2 x = V_1 J_1^k W_1 x + V_2 J_2^{k+ip} W_2 x = M_1^k x + M_2^{k+ip} x$ from the Jordan decomposition (3.24). By construction, $\lim_{i\to\infty} M^{k+ip}x = M_1^k x + \left(\lim_{i\to\infty} M_2^{k+ip}\right) x = M_1^k x = y$. Therefore, $y \in \omega(x, M)$.

Conversely, assume that $y \notin \{M_1^k x \mid k \in \{1, \ldots, p\}\}$. Since this set is closed, $d(y, \{M_1^k x \mid k \in \{1, \ldots, p\}\}) > 0$. From Lyapunov theory for discrete LTI systems, $\rho(M_2) < 1$ if and only

if there exists a positive-definite matrix P and scalar $\gamma \in (0,1)$ such that $\gamma^2 P - M_2^\intercal P M_2 \succeq 0$. Equivalently, the norm $\|x\|_P = \sqrt{x^\intercal P x}$ then has the property $\|M_2 x\|_P \leq \gamma \|x\|_P$, which implies the inequality $\|M_2^t x\|_P \leq \gamma^t \|x\|_P, \ \forall t \in \mathbb{N}$. Let $V(x) = \|x - M_1^p x\|_P$. We show by induction $V(z(t)) \leq \gamma^t V(z(0)), \ \forall t \in \mathbb{N}$, for the system $z^+ = Mz$, $z(0) = x$. The base case $t = 0$ is trivial. If $V(z(t)) \leq \gamma^t V(x)$ for some $t \in \mathbb{N}$, then $V(z(t+1)) = V(Mz(t)) = \|Mz(t) - M_1^p Mz(t)\|_P = \|M_1 z(t) + M_2 z(t) - M_1^{p+1} z(t)\|_P = \|M_2 z(t)\|_P = \|M_2 z(t) - M_1^p M_2 z(t)\|_P = V(M_2 z(t)) \leq \gamma V(z(t)) \leq \gamma^{t+1} V(x)$, which proves the induction hypothesis for the $t+1$ case. $V(z(t))$ is nonnegative and bounded from above by the exponentially decreasing function $\gamma^t V(x)$, and $V(z(t))$ is the P-norm distance between $z(t)$ and some element of the set $\{M_1^k x \mid k \in \{1, \dots, p\}\}$. Therefore $\lim_{t\to\infty} d(z(t), \{M_1^k x \mid k \in \{1, \dots p\}\}) = 0$. Let $r = \min_{k \in \{1,\dots,p\}} \|y - M^k x\|_P$ Since $r > 0$, and $z(t)$ approaches $\{M_1^k x \mid k \in \{1, \dots p\}\}$, there cannot be infinitely many t for which $\|z(t) - y\|_P < \frac{r}{2}$, and therefore $y \notin \omega(x, M)$.

■

Markov chains necessarily satisfy Assumption 3.3 because for $k \geq 1$, $d(f^k(x), \omega(x, f)) \leq d(f^k(x), \{M_1^k x\}) = \|M_2^k x\| \leq \left|\left|\left|M_2^k\right|\right|\right| \|x\|$, which exponentially approaches 0 as $k \to \infty$. For $\mathcal{S} = \{x \in \Delta(n) \mid Gg \leq g\}$, if the Markov chain converges to a unique stationary distribution v for all initial conditions, e.g. an ergodic Markov chain, and if $Gv < g$, then Assumption 3.4 is satisfied with $r = \frac{1}{\|\|G\|\|_\infty} \min_i \{e_i^\intercal (g - Gv)\}$.

The omega-limit set $\omega(x, M)$ for System (3.23) has the explicit form

$$\omega(x, M) = \left\{ M_1^k x \mid k \in \{1, \dots, p\} \right\}. \tag{3.29}$$

For $\mathcal{S}$ defined as the sub-level set $\mathcal{S} = \{x \in \Delta(n) \mid H(x) \leq 0\}$,

$$\Omega(\mathcal{S}, M) = \left\{ x \in \mathcal{S} \mid H\left(M_1^k x\right) \leq 0, \ \forall k \in \{1, \dots, p\} \right\}. \tag{3.30}$$

Specifically, if $\mathcal{S}$ is the polyhedron $\{x \in \Delta(n) \mid Gx \leq h\}$, then $\Omega(\mathcal{S}, M)$ is also a polyhedron, namely $\{x \in \mathbb{R}^n \mid \mathbf{1}^\intercal x = 1, \ x \geq 0, \ GM_1^k x \leq h, \ \forall k \in \{1, \dots, p\}\}$.

For system (3.23) with polyhedral $\mathcal{S}$, Algorithm 3.4 initializes $\mathcal{K}_0$ with a polytope of the form $\mathcal{K}_0 = \{x \in \Delta(n) \mid Gx \leq h\}$. $\mathcal{K}_1$ is also a polytope, since $\mathcal{K}_1 = \{x \in \Delta(n) \mid Gx \leq h, \ GMx \leq h\}$. It can be shown by induction that all subsequent $\mathcal{K}_0$ and $\mathcal{K}_1$ inside the While loop are polytopes. Therefore $\mathcal{K}_0 \subseteq \mathcal{K}_1$ is a condition that one polytope is within another, which by the Extended Farkas' Lemma is equivalent to at least one of the following two conditions being true [40]:

1. $\mathcal{K}_0 = \emptyset$, or
2. There exist nonnegative matrix Y and real vector z such that $GM \leq YG + z\mathbf{1}^\intercal$, and $Yg + z \leq g$.

Each of these conditions is a linear feasibility problem which can be verified by solving a linear program (LP).

We provide a bound on the number of iterations as a set of LMI conditions.

Theorem 3.9 *Let M be an $n \times n$ Markov matrix. Also let $\mathcal{S}$ be a subset of $\Delta(n)$ such that for some $r > 0$, $\omega(x, M) \oplus \mathcal{B}_r \subseteq \mathcal{S}$, $\forall x \in \Omega(\mathcal{S}, M)$. Let $M_2 = V_2 J_2 W_2$ as defined in Equation (3.27). If there exist $\gamma \in (0, 1)$, $\kappa \in [1, \infty)$, and positive-definite matrix P such that $I \preceq P \preceq \kappa I$ and $M_2^\intercal P M_2 \preceq \gamma^2 P$, then $M^k x \in \mathcal{S}$, $\forall x \in \Delta(n)$, $\forall k > \max\{0, \log_\gamma(r/\sqrt{\kappa})\}$.*

Proof: Let k be a positive integer greater than $\log_\gamma(r/\sqrt{\kappa})$, and let $x \in \Delta(n)$. Let $M = M_1 + M_2$, with $M_1 = V_1 J_1 W_1$ and $M_2 = V_2 J_2 W_2$ as defined in Equation (3.24). By equation (3.29), $\omega(x, M) = \{M_1^k x \mid k \in \mathbb{N}_{>0}\}$, and so $d(M^k x, \omega(x, M)) \leq d(M^k x, \{M_1^k x\}) = \|M^k x - M_1^k x\|_2 = \|M_2^k x\|_2$. Since $I \preceq P$, $\|M_2^k x\|_2^2 = (M_2^k x)^\intercal (M_2^k x) \leq (M_2^k x)^\intercal P (M_2^k x) = \|M_2^k x\|_P^2$. Therefore, $d(M^k x, \omega(x, M)) \leq \|M_2^k x\|_2 \leq \|M_2^k x\|_P$.

Next, the LMI $M_2^\intercal P M_2 \preceq \gamma^2 P$ implies $\|M_2 x\|_P^2 \leq \gamma^2 \|x\|_P^2$, or equivalently $\|M_2 x\|_P \leq \gamma \|x\|_P$, for all $x \in \mathbb{R}^n$. The latter inequality is applied recursively to $\|M_2^k x\|_P$ to obtain $\|M_2^k x\|_P \leq \gamma \|M_2^{k-1} x\|_P \leq \cdots \leq \gamma^k \|x\|_P$. Similarly, because $P \preceq \kappa I$, $\|x\|_P^2 \leq \kappa \|x\|_2^2$, and $\|x\|_P \leq \sqrt{\kappa}\|x\|_2$. Since the 2-norm of x is maximized at the vertices of $\Delta(n)$, $\|x\|_2 \leq \max_i \|e_i\|_2 = 1$. As a result, $\|M_2^k\|_P \leq \gamma^k \|x\|_P \leq \sqrt{\kappa}\gamma^k \|x\|_2 \leq \sqrt{\kappa}\gamma^k$, and therefore $d(M^k x, \omega(x, M)) \leq \sqrt{\kappa}\gamma^k$.

Finally, $\sqrt{\kappa}\gamma^k$ is a strictly decreasing function of k. Since we required $k > \log_\gamma(r/\sqrt{\kappa})$, we have $\sqrt{\kappa}\gamma^k < \sqrt{\kappa}\gamma^{\log_\gamma(r/\sqrt{\kappa})} = r$. Combining this with the inequality at the end of the previous paragraph, $d(M^k x, \omega(x, M)) \leq \sqrt{\kappa}\gamma^k < r$. Thus, $M^k x \in \omega(x, M) \oplus \mathcal{B}_r$, and so $M^k x \in \mathcal{S}$ by the assumption $\omega(x, M) \oplus \mathcal{B}_r \subseteq \mathcal{S}$. This proves that $M^k x \in \mathcal{S}$ for all $x \in \Delta(n)$ and for all positive integers $k > \log_\gamma(r/\sqrt{\kappa})$. ■

Corollary 3.2 *Let $\mathcal{S} = \{x \in \Delta(n) \mid Gx \leq h\}$, and let M be a Markov matrix such that $\lim_{k\to\infty} M^k = v\mathbf{1}^\intercal$ and $Gv < h$. If $(M - v\mathbf{1}^\intercal)^\intercal P (M - v\mathbf{1}^\intercal) \preceq \gamma^2 P$ for some $\gamma \in (0, 1)$ and $P \succ 0$, then $\mathcal{O}_\infty(\mathcal{S}, M) = \mathcal{O}_{\bar{t}}(\mathcal{S}, M)$, where $\bar{t} = \lceil \log_\gamma \left(\frac{a_2}{a_1}\right) \rceil - 1$, $a_1 = \max_i \|e_i - v\|_P$, $a_2 = \min_i \frac{(h - Gv)_i}{\sqrt{(GP^{-1}G^\intercal)_{ii}}}$.*

Proof: We apply Theorem 3.7 to the Markov chain $f(x) = Mx$. From the LMI $(M-v\mathbf{1}^\intercal)^\intercal P(M-v\mathbf{1}^\intercal) \preceq \gamma^2 P$, we obtain the inequality $\|Mx-v\|_P \leq \gamma\|x-v\|_P,\ \forall x \in \Delta(n)$. By induction, it follows that $\|M^k x - v\|_P \leq \gamma^k \|x - v\|_P,\ \forall k \in \mathbb{N},\ \forall x \in \Delta(n)$.

For all $x \in \Delta(n)$, $\|x - v\|_P \leq \max_i \|e_i - v\|_P = a_1$, since this a_1 is the radius of the smallest P-norm ball centered at v, which contains $\Delta(n)$. Similarly, the largest P-norm ball centered at v which is a subset of the polytope $\{x \in \mathbb{R}^n \mid Gx \leq g\}$ has radius $a_2 = \min_i \frac{(h-Gv)_i}{\sqrt{(GP^{-1}G^\intercal)_{ii}}}$. Thus, $\{x \in \Delta(n) \mid \|x - v\|_P \leq a_2\} \subseteq \mathcal{S}$. If $t \geq \log_\gamma(\frac{a_2}{a_1})$, then $\|M^t x - v\|_P \leq \gamma^t \|x - v\|_P \leq \gamma^t a_1 \leq \gamma^{\log_\gamma(a_2/a_1)} a_1 = a_2,\ \forall x \in \mathcal{S}$. Since $M^t x \in \Delta(n)$ and $\|M^t x - v\|_P \leq a_2, M^t x \in \mathcal{S}, \forall x \in \mathcal{S}$. We notice that $\Omega(\mathcal{S}, f) = \mathcal{S}$ because $\lim_{t\to\infty} M^t x = v \in \mathcal{S},\ \forall x \in \Delta(n)$. We can then claim $\Omega(\mathcal{S}, f) \subseteq f^{-\lceil \log_\gamma(a_2/a_1)\rceil}\Omega(\mathcal{S}, f)$, which by Theorem 3.7 implies that $\mathcal{O}_\infty(\mathcal{S}, f) = \bigcap_{k=0}^{\lceil \log_\gamma(a_2/a_1)\rceil - 1} f^{-k}\mathcal{S} = \mathcal{O}_{\lceil \log_\gamma(a_2/a_1)\rceil - 1}(\mathcal{S}, f)$. ■

This result suggests that among ergodic Markov chains, those with fast mixing rate $\rho(M-v\mathbf{1}^\intercal)$ tend to let $\mathcal{O}_\infty(\mathcal{S}, f)$ be computed in fewest iterations, because all points in $\Delta(n)$ rapidly converge to v. The value $a_2 = \min_i \frac{(h-Gv)_i}{\sqrt{(GP^{-1}G^\intercal)_{ii}}}$ is the P-norm distance from v to the boundary of $\mathcal{S}$. Since $\gamma \in (0, 1)$, $\log_\gamma(a_2/a_1)$ is a decreasing function of a_2, and so $\bar{t}$ is smallest when a_2 is maximized. Note also that the representation of $\mathcal{S}$ as $\{x \in \Delta(n) \mid Gx \leq h\}$ is not unique, since if $\mathbf{1}^\intercal x = 1$, then $Gx \leq h$ is equivalent to $(G + b\mathbf{1})x \leq h + b$ for all real vectors b of appropriate dimension. Therefore to maximize a_2, one should choose b to minimize the square of the denominator $e_i^\intercal(G + b\mathbf{1}^\intercal)P^{-1}(G + b\mathbf{1}^\intercal)^\intercal e_i$ for all i, specifically $b = -\frac{GP^{-1}\mathbf{1}}{\mathbf{1}^\intercal P^{-1}\mathbf{1}}$. Therefore, a_2 in the formula in Corollary 3.2 from the optimal G becomes

$$a_2 = \min_i \frac{(h - Gv)_i}{\sqrt{(GP^{-1}G^\intercal - \frac{1}{\mathbf{1}^\intercal P^{-1}\mathbf{1}} GP^{-1}\mathbf{1}\mathbf{1}^\intercal P^{-1}G^\intercal)_{ii}}}. \tag{3.31}$$

3.8 Examples

3.8.1 Explicit Example

We present a low-dimensional example to show that Algorithm 3.4 may compute $\mathcal{O}_\infty(\mathcal{S}, f)$ in finitely many steps, while Algorithm 3.3 never terminates. Consider the system $x^+ =$

$f(x) = Mx$ and set be $\mathcal{S}$ given as:

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \\ 3/4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{3.32}$$

$$\mathcal{S} = \left\{ x \in \Delta(4) \mid \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} x \leq 1/2 \right\}. \tag{3.33}$$

Because the dynamics are linear, and the constraint set is a polyhedron, we check if $\mathcal{K}_0 \nsubseteq \mathcal{K}_1$ by solving a linear feasibility problem. Each iteration of Algorithm 3.3 computes the next set in the sequence of $\mathcal{O}_k(\mathcal{S}, f)$:

$$\mathcal{O}_k(\mathcal{S}, f) = \left\{ x \in \Delta(4) \mid \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} M^k x \leq 1/2 \right\}. \tag{3.34}$$

Observe that

$$x = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{3} - \frac{1}{6(-2)^k} & \frac{1}{6} + \frac{1}{6(-2)^k} \end{bmatrix}^\mathsf{T} \tag{3.35}$$

is in $\mathcal{O}_k(\mathcal{S}, f)$ but not $\mathcal{O}_{k+1}(\mathcal{S}, f)$. For $t \leq k$, $GM^t x = \frac{1}{2} - \frac{(-1)^t}{6(-2)^k} - \frac{1}{6(2^t)} \leq \frac{1}{2} + \frac{1}{6(2^k)} - \frac{1}{6(2^t)} \leq \frac{1}{2}$, while $GM^{k+1}x = \frac{1}{2} - \frac{(-1)^{k+1}}{6(-2)^k} - \frac{1}{6(2^{k+1})} = \frac{1}{2} + \frac{1}{12(2^k)} > \frac{1}{2}$. Since there is no $k \in \mathbb{N}$ for which $\mathcal{O}_k(\mathcal{S}, f) \subseteq \mathcal{O}_{k+1}(\mathcal{S}, f)$, Algorithm 3.3 does not terminate.

We use Algorithm 3.4 instead to finitely characterize $\mathcal{O}_\infty(\mathcal{S}, f)$. We first compute $\Omega(\mathcal{S}, f)$. The Jordan decomposition of M is computed symbolically, then M_1 is found by setting all transient Jordan blocks to zero:

$$M_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}. \tag{3.36}$$

Since the eigenvalues of M on the unit circle are 1 and -1, the Markov chain has period 2.

Therefore,

$$\omega(x, f) = \{M_1 x\} \cup \left\{M_1^2 x\right\}, \tag{3.37}$$

$$\Omega(\mathcal{S}, f) = \left\{x \in \Delta(4) \mid x \in \mathcal{S},\ M_1 x \in \mathcal{S},\ M_1^2 x \in \mathcal{S}\right\} \tag{3.38}$$

$$= \left\{ x \in \Delta(4) \mid \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} x \leq \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \right\}. \tag{3.39}$$

In this case, $\Omega(\mathcal{S}, f)$ is positively invariant, so Algorithm 3.4 immediately finds that $\mathcal{O}_\infty(\mathcal{S}, f) = \Omega(\mathcal{S}, f)$.

3.8.2 Numerical Example

Our second example shows how this algorithm can be applied to a larger example. A swarm of agents is placed on a 13-by-7 grid. At each time step, every agent moves to a uniformly random adjacent bin unless it is in the A, C, or L shaped region, where the agents move according to the Metropolis-Hastings algorithm with a uniform stationary distribution [1][6]. The eigenvalues of M in order of decreasing magnitude are $(1, 1, 1, -0.9397, 0.9397, 0.9211, ...)$. The eigenvalues with unit magnitude are all exactly 1, so M_1 has period $p = 1$.

For safety, no bin may have more than 3.6% of the agents at the same time. That is, the state vector $x(t)$ must satisfy the constraints $x(t) \in \mathcal{S} = \{x \in \Delta(91) \mid x \leq \frac{9}{250}\mathbf{1}\},\ \forall t \geq 0$. Since M_1 has period 1, $\Omega(\mathcal{S}, M) = \{x \in \Delta(91) \mid x \leq \frac{9}{250}\mathbf{1},\ M_1 x \leq \frac{9}{250}\mathbf{1}\}$. Rather than use Algorithm 3.3 to compute the minimum t^* for which $\mathcal{O}_{t^*}(\mathcal{S}, M) = \mathcal{O}_\infty(\mathcal{S}, M)$, we observe that M is diagonalizable as $M = VJV^{-1}$ with all real eigenvalues. Thus, the LMI condition $M_2^\intercal P M_2 \preceq \gamma^2 I$ is satisfied for $\gamma = \rho(M_2) \approx 0.9397$ and $P = (VV^\intercal)^{-1}$. This property satisfies Assumption 3.3 with respect to the P-norm $\|x\|_P = \sqrt{x^\intercal P x}$.

We choose an initial condition $x(0) \in \mathcal{O}_\infty(\mathcal{S}, M)$ that converges to the uniform distribution over the A, C, and L regions. We let v be this particular steady-state distribution. This v satisfies Assumption 3.4 within in the simplex $\mathcal{X} = \Delta(91)$, since $v \leq \frac{1}{28}\mathbf{1} < \frac{9}{250}\mathbf{1}$. Using Theorem 3.7 and Corollary 3.2, we find that $\mathcal{O}_\infty(\mathcal{S}, M) = \{x \in \Delta(91) \mid M^k x \leq \frac{9}{250}\mathbf{1},\ \forall k \in \mathbb{N}_{\leq 206}\}$. The safety constraint is then satisfied if and only if the initial condition is in the polytope $\mathcal{O}_\infty(\mathcal{S}, M)$. The initial condition is then optimized as part of a convex program with the linear constraints $x(0) \in \{x_0 \in \mathbb{R}^n \mid \mathbf{1}^\intercal x_0 = 1,\ x_0 \geq 0,\ M_1 x_0 = v,\ M^k x \leq \frac{9}{250}\mathbf{1},\ \forall k \in \mathbb{N}_{\leq 205}\}$.

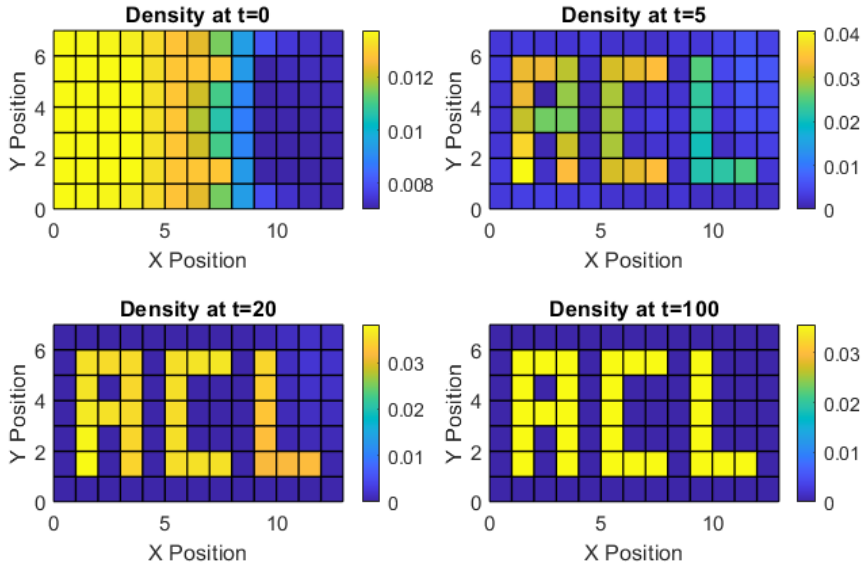

Figure 3.5: Evolution of the maximal entropy initial condition without the safety constraint.

To illustrate, we maximize the entropy of $x(0)$ to encourage a nearly uniform initial distribution.

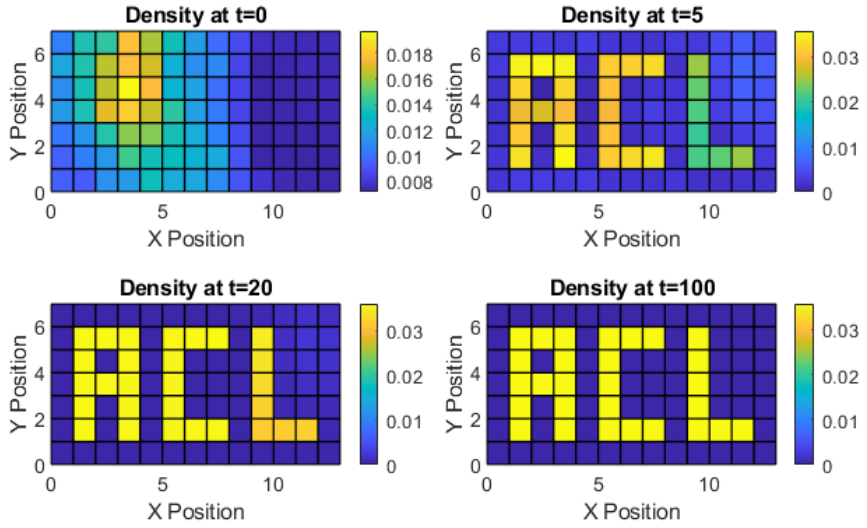

Figure 3.6: The initial condition is now constrained to be in $\mathcal{O}_\infty(\mathcal{S}, M)$, which ensures the safety constraint $x(t) \in \mathcal{S},\ \forall t \geq 0$.

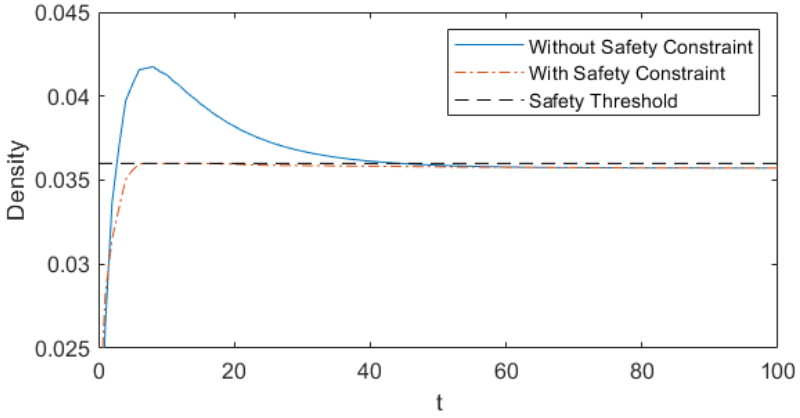

Figure 3.7: Comparison of the maximum density over time for the two trajectories.

Chapter 4

DISCRETE-TIME LINEAR-QUADRATIC REGULATION VIA OPTIMAL TRANSPORT

In this chapter, we consider a discrete-time stochastic control problem with uncertain initial and target states. We first discuss the connection between optimal transport and stochastic control problems of this form. Next, we formulate a linear-quadratic regulator problem where the initial and terminal states are distributed according to specified probability densities. A closed-form solution for the optimal transport map in the case of linear-time varying systems is derived, along with an algorithm for computing the optimal map. Two numerical examples pertaining to swarm deployment demonstrate the practical applicability of the model, and performance of the numerical method.

4.1 Introduction

The problem of steering the states of a linear system from an initial distribution to a terminal distribution has attracted much interest in recent years [16, 17]. Applications of such controllers include the density control of swarms [30, 25] and networked dynamical systems [22], as well as opinion dynamics [3].

Optimal mass transport is a mathematical framework for deriving mass-preserving maps between specified distributions that minimize a cost of transport. The optimal transport cost, in some specific contexts called the *Wasserstein metric*, provides a useful metric on the space of probability distributions. This has been employed in a wide variety of fields, such as economics [35], machine learning [33, 70], computer vision [67], and image processing [66]. The Wasserstein metric also allows one to tractably compute worst-case distributions in optimization problems [61], which have been applied in areas such as state estimation [71], and machine learning [53]. The computation of the Wasserstein metric and corresponding transport map has also attracted much attention, in particular techniques allowing for computational speedup such as entropic regularization and Sinkhorn scaling [21, 63].

The connection of optimal transport to continous-time control began with the seminal reformulation of optimal transport as a PDE-based fluid dynamics optimization problem [10].

In this approach, a velocity field is computed that minimizes the average kinetic energy of a fluid moving from one density to another. Equivalently, this approach can be thought of as a single-integrator particle moving from an initial state with uncertainty described by an initial distribution, to a final state with an uncertainty described by a final distribution. The cases of general linear time-varying (LTV) systems, and general LTV systems driven by noise (so-called *Schrödinger bridges*) were developed by [18].

The latter paper [18] employs a Lagrangian-based cost function, where the static quadratic cost is replaced with a time-varying cost with dynamical constraints. Such techniques were developed in [2], which dealt with optimal transport with nonholonomic constraints. In a similar problem configuration, the existence and uniqueness of transport maps were determined for linear–quadratic costs by [41]. Other works include distributed optimal transport for swarms of single-integrators [51, 52], Perron-Frobenius operator methods for computing optimal transport over nonlinear systems [29], and a related problem regarding the steering of an LTV systems to a terminal state with specified expected value and covariance [37, 7, 8].

While much attention has been paid to optimal transport of dynamical systems in continuous-time, there has been a marked lack of works discussing the implementation of such controllers in discrete time, which is a gap in the literature that needs to be addressed before optimal transport techniques can be implemented on digital controllers. One contribution of this paper is to provide a rigorous analysis of the optimal transport problem for linear-quadratic regulation of LTV systems in discrete time.

In the present work, we discuss the theory and implementation of optimal transport for discrete-time linear-quadratic regulation for LTV systems. Our contributions are as follows. We formalize a previously-developed method of applying optimal transport methods to control by converting a class of optimal control problems to an optimal transport problem where the cost function is the optimal cost-to-go from an initial state to a terminal state. This formalism is then applied to derive the closed-form solution of the discrete-time LQR problem with state-density constraints. This problem is solved numerically, and the solution is then implemented on an example involving swarm deployment.

The paper is organized as follows. We outline the notation and preliminaries on optimal transport in §4.2. Our problem statement is outlined in §4.3, where we discuss formulating optimal transport problems for control systems in terms of value functions. Our results concerning optimal transport for LQR and its numerical computation are in §4.4. We present numerical examples and an application to swarm deployment in §4.5, and conclude the paper

in §4.6.

4.2 Mathematical Preliminaries

In this section, we outline the notation used in the paper, as well as the necessary preliminaries on optimal transport.

4.2.1 Notation

The n-dimensional space of real numbers, non-negative real numbers, and positive real numbers are respectively denoted by $\mathbb{R}^n, \mathbb{R}^n_+$, and $\mathbb{R}^n_{++}$. We denote vectors in lower-case $x, y, z \in \mathbb{R}^n$, and matrices in capital-case $A, B, C \in \mathbb{R}^{n\times m}$. Inequalities are interpreted component-wise. Symmetric positive-definite and positive semi-definite cones of matrices are respectively denoted as $\mathcal{S}^n_{++}$ and $\mathcal{S}^n_+$. For $Q \in \mathcal{S}^n_+$, we let $x^TQx = \|x\|^2_Q$. The $n \times n$ identity matrix is denoted by I_n, or just I if comformable dimensions are assumed. $\mathbf{1}_n$ denotes the length-n vector of all ones, and $\mathbf{0}$ denotes a matrix of zeros. The identity map is denoted by $\mathrm{Id}(x) = x$. The direct sum of n $m \times m$ square matrices matrices $\{A_i\}_{i=1}^n$ is the $nm \times nm$ matrix formed by placing $A_1, \ldots, A_n$ on the block diagonal. It is denoted by $\bigoplus_{i=1}^n A_i$. The vectorization operation $U = \mathrm{vec}(\{u_k\}_{k=1}^n)$ denotes the vector $U \in \mathbb{R}^{nm}$ consisting of the stacked vectors $u_k \in \mathbb{R}^m$.

A measure space is a triple $(\mathcal{X}, \mathcal{A}, \mu)$ where $\mathcal{X}$ is a set, $\mathcal{A}$ is a σ-algebra on $\mathcal{X}$, and μ is measure on $(\mathcal{X}, \mathcal{A})$. We write a *probability space* as $(\mathcal{X}, \mu)$, where $\mu(\mathcal{X}) = 1$ is a non-negative measure, and we assume that $(\mathcal{X}, \mu)$ is equipped with the Borel σ-algebra.

For probability spaces $(\mathcal{X}, \mu_0)$, $(\mathcal{Y}, \mu_1)$, the *pushforward map*, denoted by $\mu_1 := T_\# \mu_0$, is defined by the relation

$$\mu_1(B) = \mu_0(T^{-1}(B)) \tag{4.1}$$

for each $B \in \mathcal{A}(Y)$. If a random variable x is distributed according to a probability density function $\rho(x)$, then we write $x \sim \rho(x)dx$.

4.2.2 Optimal Transport

In this section, we summarize four seminal forms of the optimal transport problem, and then specify the form of the optimal transport for our present work. One may consult excellent texts by Villani for a more in-depth discussion of the theory [74, 73].

Consider two probability spaces $(\mathcal{X}, \mu_0)$ and $(\mathcal{Y}, \mu_1)$. A *transport map* $T : \mathcal{X} \to \mathcal{Y}$ is said to *transport* μ_0 *to* μ_1 if $T_{\#}\mu_0 = \mu_1$. The Monge optimal transport problem seeks to find an optimal map T that minimizes some cost of transport $c(x, T(x))$,

$$\begin{array}{ll} \inf_T & \int_{\mathcal{X}} c(x, T(x)) d\mu_0(x) \\ \text{s.t.} & T_{\#}\mu_0 = \mu_1. \end{array} \tag{OT1}$$

In general, if one of the measures μ_0, μ_1 has infinite second moment, then the cost of (OT1) may be infinite. Furthermore, the pushforward constraint of (OT1) makes this problem computationally intractable. Kantorovich formulated a relaxation of (OT1) that obtains the same minimizer under quadratic costs[1], i.e., $c(x, y) = \frac{1}{2}\|x - y\|_2^2$. The problem considers the set of joint probability distributions $\pi(x, y)$ on $\mathcal{X} \times \mathcal{Y}$ whose marginals are the initial and target measures,

$$\pi(A, \mathcal{Y}) = \mu_0(A), \;\; \pi(\mathcal{X}, B) = \mu_1(B), \tag{4.2}$$

for all Borel sets $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$. With some abuse of notation, to make variables of operators (e.g., optimization, integration) we may write the above as

$$\pi(x, \cdot) = \mu_0(x), \;\; \pi(\cdot, y) = \mu_1(y). \tag{4.3}$$

The Kantorovich optimal transport is then given by,

$$\begin{array}{ll} \inf_\pi & \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) d\pi(x, y) \\ \text{s.t.} & \pi(x, \cdot) = \mu_0(x), \;\; \pi(\cdot, y) = \mu_1(y). \end{array} \tag{OT2}$$

For the case of quadratic costs, (OT2) obtains the same minimum as (OT1), and the optimal coupling satisfies $\pi^* = (\mathrm{Id} \times T^*)_{\#}\mu_0$, where $T^*(x)$ is the optimal map from (OT1).

The dual of (OT2) has an explicit interpretation in economic theory of transport pricing [35], but perhaps more importantly, it offers insight into the structure of the optimal map T in the case of quadratic costs. For ϕ, ψ in the dual space of probability measures, the

[1]The Kantorovich and Monge problems have corresponding minimizers under more general choices of $c(x, y)$, but we only consider the quadratic cost $c(x, y) = x^T Q_x x + y^T Q_y y + 2x^T Q_{xy} y$ in this paper.

dual is given by,

$$\begin{array}{ll} \sup_{\phi,\psi} & \int_{\mathcal{X}} \phi(x) d\mu_0(x) - \int_{\mathcal{Y}} \psi(y) d\mu_1(y) \\ \text{s.t.} & \phi(x) - \psi(y) \leq c(x,y), \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}. \end{array} \tag{OT3}$$

When $c(x,y) = \frac{1}{2}\|x-y\|_2^2$, the optimal map $T^*(x)$ of (OT1) can be written in terms of ψ^* from (OT3) as [73],

$$T^*(x) = \nabla \left(\frac{1}{2} x^T x + \psi^*(x) \right), \tag{4.4}$$

and in particular it can be shown that $(\frac{1}{2}x^T x + \psi^*(x))$ is a convex function [14]. Note that in our notation, $\psi^*(x)$ refers to the *optimal* ψ, and not its Fenchel conjugate.

One final formulation of optimal transport we describe here is given by Brenier and Benamou in the form of an optimal control problem in a fluid dynamics setting. Given initial and terminal densities ρ_0, ρ_1, one seeks to find a smooth, time-dependent velocity field $v(x,t)$ taking ρ_0 to ρ_1 in unit time, while satisfying the continuity equation. The velocity field minimizes the average kinetic energy of the fluid. The problem is explicitly defined as [10],

$$\begin{array}{ll} \sup_{\rho,v} & \int_0^1 \int_{\mathbb{R}^n} \|v(x,t)\|_2^2 \rho(x,t) dx dt \\ \text{s.t.} & \partial_t \rho(x,t) + \nabla \cdot (\rho(x,t) v(x,t)) = 0 \\ & \rho(x,0) = \rho_0(x), \ \rho(x,1) = \rho_1(x). \end{array} \tag{OT4}$$

In Lagrangian coordinates $X(x,t)$ with $X(x,0) := x$, $\partial_t X(x,t) = v(X(x,t),t)$, the solution to (OT4) is given by a linear interpolation with the optimal map,

$$X(x,t) = x + t(T(x) - x) =: T_t(x), \tag{4.5}$$

and so the densities at time t satisfy

$$\rho(x,t) := \rho_t(x) = (T_t)_{\#} \rho_0(x). \tag{4.6}$$

4.3 Stochastic Optimal Control with State-Density Constraints

In this section, we consider an optimal transport approach for the discrete-time linear-quadratic regulator. We present a formal discretization of the continuous-time controllers presented in [18], and extend this to the more general framework of LQR control.

We consider systems with a state $z_k \in \mathbb{R}^n$ of the form

$$\begin{aligned} z_{k+1} &= A_k z_k + B_k u_k \\ z_0 &\sim \rho_0(z)dz, \end{aligned} \tag{4.7}$$

where the initial condition z_0 has some uncertainty described by a probability density $\rho_0(x)$ and $u_k \in \mathbb{R}^m$ is the control. Our goal is to translate the system (4.7) to a terminal state $z_{t_f} \sim \rho_1$ over a time horizon $0 \leq k \leq t_f$, where ρ_1 captures some desired uncertainty in the terminal state[2]. The control should satisfy some optimality principle under an appropriate cost, and so an optimization problem with dynamics (4.7) is,

$$\begin{array}{ll} \min_{u,\pi} & \mathbb{E}_\pi \left[\sum_{k=0}^{t_f-1} c(z_k, u_k) \right] \\ \text{s.t.} & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 \sim \rho_0(z)dz, \;\; z_{t_f} \sim \rho_1(z)dz, \end{array} \tag{4.8}$$

where the expectation is with respect to a joint distribution $\pi(z_0, z_{t_f})$, as defined in (OT2). The remark below formalizes a solution technique for problems of the form (4.8) which was used by [18] to solve continuous-time optimal control problems with control costs.

Remark 4.1 *A general method to solve problems of the form (4.8) is to first solve the deterministic problem*

$$\left.\begin{array}{ll} \min_{u} & \sum_{k=0}^{t_f-1} c(z_k, u_k) \\ \text{s.t.} & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 = x, \;\; z_{t_f} = y \end{array}\right\} = C(x, y), \tag{4.9}$$

to determine a formula $C(x, y)$ for the optimal cost-to-go from x to y. Thus, (4.8) can be

[2] As a technical assumption, we let $t_f \geq n$. This is not a constricting assumption, because OT problems do not in general scale well with n, and so we expect that in a real-world setting the OT methods in this paper would be applied to a reduced-order model (and hence small n) to compute references that would be tracked by a local, higher-fidelity controller.

re-written as

$$\begin{array}{ll} \min_{\pi} & \int_{\mathbb{R}^n \times \mathbb{R}^n} C(x,y) d\pi(x,y) \\ \text{s.t.} & \pi(x,\cdot) = \rho_0(x)dx, \ \pi(\cdot,y) = \rho_1(y)dy, \end{array} \tag{4.10}$$

where the marginal constraints on π encode the relevant distributions on the initial state x and terminal state y. Problem (4.10) is clearly a Kantorovich optimal transport problem of the form (OT2), where the cost function is now the deterministic value function encoding the cost-to-go from x to y, and the optimal coupling π^* encodes a mapping between initial and terminal conditions x and y.

The solution to (4.10), under appropriate assumptions on the cost $C(x,y)$, yields a coupling of the form

$$\pi^*(x,y) = (\mathrm{Id} \times T^*)_{\#}\mu_0(x), \tag{4.11}$$

where $y = T^*(x)$. Finally, we note that $\{u_k^*(x, T(x))\}_{k=1}^{t_f-1}$ solves (4.8). △

When $c(z_k, u_k) = (z_k - y)^T Q_k (z_k - y) + u_k^T R_k u_k$ for $y \sim \rho_1(z)dz$, we have the following LQR problem with stochastic initial and terminal constraints,

$$\begin{array}{ll} \min_{u} & \mathbb{E}\left[\sum_{k=0}^{t_f-1} \left\{ \|z_k - y\|_{Q_k}^2 + \|u_k\|_{R_k}^2 \right\}\right] \\ \text{s.t.} & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 \sim \rho_0(z)dz, \ z_{t_f} = y \sim \rho_1(z)dz. \end{array} \tag{4.12}$$

We solve this problem in the following section.

4.4 Derivation of the Optimal Map

In this section, we outline the solutions to Problem (4.12), beginning with the simplified case of a cost on the control only. Our main contribution in this section is the more-general LQR problem, outlined in 4.4.2.

4.4.1 Discrete-Time Optimal Transport – Control Cost Case

Consider the task of transporting over a time horizon of $0 \leq k \leq t_f$ a linear time-varying system with uncertain initial state z_0 characterized by $\rho_0(z)$, to a final state z_{t_f} with an

uncertainty characterized by $\rho_1(z)$, with minimal control cost. We assume that the dynamics $z_{k+1} = A_k z_k + B_k u_k$ are controllable over the interval $0 \leq k \leq t_f$. The problem is posed as

$$\begin{aligned} \min \quad & \mathbb{E}\left[\sum_{k=0}^{t_f-1} \|u(k)\|_2^2\right] \\ \text{s.t.} \quad & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 \sim \rho_0(z)dz,\ z_{t_f} \sim \rho_1(z)dz. \end{aligned} \tag{P1}$$

Following a similar derivation as the continuous-time case studied in [18], one can consider first solving the deterministic problem,

$$\begin{aligned} \min \quad & \sum_{k=0}^{t_f-1} \|u(k)\|_2^2 \\ \text{s.t.} \quad & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 = x,\ z_{t_f} = y. \end{aligned} \tag{P2}$$

The solution to (P2) is given in closed form as

$$\begin{aligned} C^*(x,y) = &(y - \Phi(t_f,0)x)^T \\ &\cdot W_c(t_f,0)^{-1}\left(y - \Phi(t_f,0)x\right), \end{aligned} \tag{4.13}$$

$$u^*(k) = B_k^T \Phi(t_f, k+1)^T W_c(t_f,0)^{-1}\left(y - \Phi(t_f,0)x\right), \tag{4.14}$$

where

$$\Phi(t_f,k) = A_{t_f-1}A_{t_f-2}\cdots A_k, \tag{4.15}$$

and

$$W_c(t_f,0) = \sum_{k=0}^{t_f-1} \Phi(t_f,k+1)B_k B_k^T \Phi(t_f,k+1)^T. \tag{4.16}$$

Substituting the optimal cost (4.13) into (P1) and applying the coordinate transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} W_c(t_f,0)^{-1/2}\Phi(t_f,0)x \\ W_c(t_f,0)^{-1/2}y \end{bmatrix} \triangleq \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}, \tag{4.17}$$

transforms (4.13) into $\|\hat{x} - \hat{y}\|_2^2$, and so we arrive at the Kantorovich optimal transport problem

$$\begin{array}{ll} \min_{\hat{\pi}} & \int_{\mathbb{R}^n \times \mathbb{R}^n} \|\hat{x} - \hat{y}\|_2^2 d\hat{\pi}(\hat{x}, \hat{y}) \\ \text{s.t.} & \hat{\pi}(\hat{x}, \cdot) = \hat{\rho}_0(\hat{x}) d\hat{x}, \ \hat{\pi}(\cdot, \hat{y}) = \hat{\rho}_1(\hat{y}) d\hat{y}, \end{array} \tag{P3}$$

with the distributions in the $(\hat{x}, \hat{y})$ coordinates given by

$$\hat{\rho}_0(\hat{x}) = |W_c(t_f, 0)|^{1/2} |\Phi(t_f, 0)|^{-1} \rho_0 \left(\Phi(t_f, 0)^{-1} W_c(t_f, 0)^{1/2} \hat{x} \right) \tag{4.18}$$

$$\hat{\rho}_1(\hat{y}) = |W_c(t_f, 0)|^{1/2} \rho_0 \left(W_c(t_f, 0)^{1/2} \hat{y} \right). \tag{4.19}$$

Now, suppose that $\hat{T}$ is the Monge map that corresponds to the solution of (P3). Then, by using (4.17), the solution to the original problem (P1) can be written in terms of its Monge map

$$y = T(x) = W_c(t_f, 0)^{1/2} \hat{T} \left(W_c(t_f, 0)^{-1/2} \Phi(t_f, 0) x \right). \tag{4.20}$$

The optimal controls are thus given by

$$u(k, x(k)) = B^T \Phi(t_f, k+1)^T W_c(t_f, 0)^{-1} \left[T(x) - \Phi(t_f, 0) x \right]. \tag{4.21}$$

4.4.2 Discrete-Time Optimal Transport – Linear-Quadratic Case

In this subsection, we consider the more general case of a linear-quadratic cost function. The problem is formulated as

$$\begin{array}{ll} \min & \mathbb{E} \left[\sum_{k=0}^{t_f - 1} \left\{ \|u_k\|_{R_k}^2 + \|z_k - y\|_{Q_k} \right\}^2 \right] \\ \text{s.t.} & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 \sim \rho_0(z) dz, \ z_{t_f} = y \sim \rho_1(z) dz. \end{array} \tag{P3}$$

We proceed using the methodology in Remark 4.1 by considering the solution to the deterministic problem,

$$\begin{array}{ll} \min & \sum_{k=0}^{t_f - 1} \left\{ \|u_k\|_{R_k}^2 + \|z_k - y\|_{Q_k}^2 \right\} \\ \text{s.t.} & z_{k+1} = A_k z_k + B_k u_k \\ & z_0 = x, \ z_{t_f} = y. \end{array} \tag{P4}$$

We summarize the cost-to-go of (P4) in the following lemma. Note that we utilize a pseudoinverse, present in (4.47). While at first glance it may seem that this pseudoinverse severely limits the applicability of this lemma, this is not the case. We discuss in Remark 4.3 (after the proof of the lemma) why the pseudoinverse is well-defined for all controllable LTV systems, and we highlight an example in §4.5 that shows that the pseudoinverse is indeed well-behaved, even in pathological cases.

Lemma 4.1 *The optimal cost-to-go of* (P4) *is quadratic in* x, y*, in that*

$$C^*(x,y) = x^T Q_x x + y^T Q_y y + 2x^T Q_{xy} y, \tag{4.22}$$

where Q_{xy} *is invertible. The optimal control of* (P4) *is given by* $U^* := \text{vec}(\{u_k^*\}_{k=0}^{t_f-1})$*,*

$$\begin{aligned} U^* &= K^*(y - \Phi(t_f,0)x) \\ &\quad - \Gamma_{U_1} P^{-1} A_{U_1}^T \tilde{Q} \left(\Omega x - (\mathbf{1}_{t_f} \otimes I_n) y\right), \end{aligned} \tag{4.23}$$

$$K^* = \left(I - \Gamma_{U_1} P^{-1} \Gamma_{U_1}^T \tilde{R} - \Gamma_{U_1} P^{-1} A_{U_1}^T \tilde{Q} \Psi\right) \Gamma_y, \tag{4.24}$$

where the matrices comprising the optimal cost and control are given by

$$Q_x = K_1^T \tilde{Q} K_1 + K_3^T \tilde{R} K_3 \tag{4.25}$$

$$Q_y = K_2^T \tilde{Q} K_2 + K_4^T \tilde{R} K_4 \tag{4.26}$$

$$Q_{xy} = K_1^T \tilde{Q} K_2 + K_3^T \tilde{R} K_4 \tag{4.27}$$

$$K_1 = (I - A_{U_1} P^{-1} A_{U_1}^T \tilde{Q}) A_x - A_{U_1} P^{-1} \Gamma_{U_1}^T \tilde{R} \Gamma_x \tag{4.28}$$

$$K_2 = (I - A_{U_1} P^{-1} A_{U_1}^T \tilde{Q}) A_y - A_{U_1} P^{-1} \Gamma_{U_1}^T \tilde{R} \Gamma_y \tag{4.29}$$

$$K_3 = (I - \Gamma_{U_1} P^{-1} \Gamma_{U_1}^T \tilde{R}) \Gamma_x - \Gamma_{U_1} P^{-1} A_{U_1}^T \tilde{Q} A_x \tag{4.30}$$

$$K_4 = (I - \Gamma_{U_1} P^{-1} \Gamma_{U_1}^T \tilde{R}) \Gamma_y - \Gamma_{U_1} P^{-1} A_{U_1}^T \tilde{Q} A_y \tag{4.31}$$

$$P = A_{U_1}^T \tilde{Q} A_{U_1} + \Gamma_{U_1}^T \tilde{R} \Gamma_{U_1} \tag{4.32}$$

$$A_x = \Omega + \Psi \Gamma_x, \;\; A_y = \Psi \Gamma_y - \mathbf{1}_{t_f} \otimes I_n, \;\; A_{U_1} = \Psi \Gamma_{U_1} \tag{4.33}$$

$$\Gamma_x = \begin{bmatrix} \mathbf{0} \\ -S_2^\dagger \Phi(t_f,0) \end{bmatrix}, \;\; \Gamma_y = \begin{bmatrix} \mathbf{0} \\ S_2^\dagger \end{bmatrix}, \;\; \Gamma_{U_1} = \begin{bmatrix} I_{(t_f-n)m} \\ -S_2^\dagger S_1 \end{bmatrix}, \tag{4.34}$$

$$\tilde{Q} = \bigoplus_{k=0}^{t_f-1} Q_k, \;\; \tilde{R} = \bigoplus_{k=0}^{t_f-1} R_k, \tag{4.35}$$

where S_1, and S_2 are

$$S_1 = \tag{4.36}$$

$$\begin{bmatrix} \Phi(t_f,1)B_0 & \Phi(t_f,2)B_1 & \cdots & \Phi(t_f,t_f-n)B_{t_f-n-1} \end{bmatrix}, \tag{4.37}$$

$$S_2 = \begin{bmatrix} \Phi(t_f,t_f-n+1)B_{t_f-n} & \cdots & B_{t_f-1} \end{bmatrix}, \tag{4.38}$$

and the matrices defined by the dynamics are given by

$$\Psi = \begin{bmatrix} \bar{\Upsilon}(0) \\ \bar{\Upsilon}(1) \\ \vdots \\ \bar{\Upsilon}(t_f-1) \end{bmatrix}, \ \Omega = \begin{bmatrix} \Phi(1,0) \\ \Phi(2,0) \\ \vdots \\ \Phi(t_f,0) \end{bmatrix} \tag{4.39}$$

$$\Upsilon(l,0) = \begin{bmatrix} \Phi(l,1)B_0 & \Phi(l,2)B_1 & \cdots & B_{l-1} \end{bmatrix} \tag{4.40}$$

$$\bar{\Upsilon}(l) := \begin{bmatrix} \Upsilon(l,0) & | & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n \times mt_f}. \tag{4.41}$$

Remark 4.2 *For an LTI system, these matrices are simply*

$$\Psi = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & & \mathbf{0} \\ B & \mathbf{0} & & & \\ AB & B & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \\ A^{t_f-2}B & A^{t_f-3}B & \cdots & B & \mathbf{0} \end{bmatrix}, \ \Omega = \begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^{t_f-1} \end{bmatrix} \tag{4.42}$$

△

We now prove the lemma.

Proof: An analysis in the simpler case of LTI systems with $y = 0$ may be found in [31]. By setting $U(l) = \text{vec}(\{u_k\})_{k=0}^{l}$, $U = U(t_f)$, and $\tilde{R} = \bigoplus_{k=0}^{t_f-1} R_k$, the control cost term can be written as

$$\sum_{k=0}^{t_f-1} \|u_k\|_{R_k}^2 = U^T \tilde{R} U. \tag{4.43}$$

Similarly, by writing $z_k = \Phi(k,0)x + \Upsilon(k,0)U(k)$ and $\tilde{Q} = \bigoplus_{k=0}^{t_f-1} Q_k$, the state cost term

can be written as

$$\sum_{k=0}^{t_f-1} \|z_k - y\|_{Q_k}^2 = \tag{4.44}$$

$$\left(\Omega x + \Psi U - (\mathbf{1}_{t_f} \otimes I_n)y\right)^T \tilde{Q} \left(\Omega x + \Psi U - (\mathbf{1}_{t_f} \otimes I_n)y\right). \tag{4.45}$$

To eliminate the equality constraints, we can write,

$$y = \Phi(t_f, 0)x + \begin{bmatrix} S_1 & S_2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \tag{4.46}$$

and we can thus parameterize U as follows:

$$U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} U_1 \\ S_2^\dagger \left(y - \Phi(t_f, 0)x - S_1 U_1\right) \end{bmatrix} \tag{4.47}$$

$$= \begin{bmatrix} I_{(t_f-n)m} \\ -S_2^\dagger S_1 \end{bmatrix} U_1 + \begin{bmatrix} \mathbf{0} \\ -S_2^\dagger \Phi(t_f, 0) \end{bmatrix} x + \begin{bmatrix} \mathbf{0} \\ S_2^\dagger \end{bmatrix} y \tag{4.48}$$

$$= \Gamma_{U_1} U_1 + \Gamma_x x + \Gamma_y y. \tag{4.49}$$

Substituting (4.49) into (4.43) and (4.45) yields the total cost as

$$\begin{aligned} J(x, y, U_1) = &(N_Q + A_{U_1} U_1)^T \tilde{Q} (N_Q + A_{U_1} U_1) \\ &+ (N_R + \Gamma_{U_1} U_1)^T \tilde{R} (N_R + \Gamma_{U_1} U_1) \end{aligned} \tag{4.50}$$

$$N_Q = A_x x + A_y y \tag{4.51}$$

$$N_R = \Gamma_x x + \Gamma_y y. \tag{4.52}$$

Taking the gradient of (4.50) with respect to U_1, setting it to zero, and solving for U_1^* yields,

$$U_1^* = -P^{-1}(A_{U_1}^T \tilde{Q} N_Q + \Gamma_{U_1}^T \tilde{R} N_R). \tag{4.53}$$

Substituting this form of U_1^* into (4.49) yields the optimal control as in (4.23), and substituting U_1^* into (4.50) yields the optimal cost as in (4.22). It can be checked that Q_{xy} is positive definite. ■

Remark 4.3 *Our technical assumption in the above lemma is that the LTV system is con-*

trollable, in the sense that the controllability Gramian is positive definite. The matrix S_2 *in* (4.47) *may be zero, even in cases when the underlying system is controllable. The pseudoinverse is well-defined in this case due to the elementary property* $\mathbf{0}^{\dagger}_{m\times n} = \mathbf{0}_{n\times m}$. *See the example in §4.5 for more details.* △

We can now state and prove the main theorem regarding the solution of Problem (P3).

Theorem 4.1 *Consider the setting of Problem* (P3)*, and the Kantorovich optimal transport problem,*

$$\begin{array}{ll} \min_{\pi} & \int_{\mathbb{R}^n\times\mathbb{R}^n} C^*(x,y)d\pi(x,y) \\ \text{s.t.} & x\sim\rho_0(x)dx,\ y\sim\rho_1(y)dy \end{array}, \tag{4.54}$$

where $C^*(x,y)$ *is given by* (4.22)*. Then, the optimal coupling* π^* *of* (4.54) *is given by*

$$\pi^*(x,y) = (\mathrm{Id}\times T^*)_{\#}\rho_0(x), \tag{4.55}$$

where $T^*_{\#}\rho_0 = \rho_1$*. Furthermore, the control inputs optimizing Problem* (P3) *is given by,*

$$U = K^*(T^*(x) - \Phi(t_f,0)x) \tag{4.56}$$

$$-\Gamma_{U_1}P^{-1}A_{U_1}^T\tilde{Q}\left(\Omega x - (\mathbf{1}_{t_f}\otimes I_n)T^*(x)\right), \tag{4.57}$$

where the relevant matrices are defined in Lemma 4.1.

Proof: We can see that the Kantorovich problem (4.54) corresponds to Problem (P3). By Theorem 2.2 in [41], the solution to Problem (4.54) under the cost $C^*(x,y)$ from (4.22) exists, and is of the form

$$\pi^*(x,y) = (\mathrm{Id}\times T^*)_{\#}\rho_0(x), \tag{4.58}$$

with

$$T^*(x) = -\frac{1}{2}Q_{xy}^{-1}\nabla\varphi(x), \tag{4.59}$$

where $\varphi(x)$ is a convex function. From the calculation in Lemma 4.1, the optimal control in (4.57) follows. ■

Remark 4.4 *The solution to Problem (4.54) yields a Monge map T^* that transports $x \sim \rho_0(x)dx$ to $y := T^*(x) \sim \rho_1(y)dy$, minimizing the expected cost-to-go from x to y. Another interpretation of this map is that it pairs initial and terminal states (x, y) in such a manner that it minimizes the LQR cost averaged over the distribution of initial states. We exploit this interpretation in §4.5, where we discuss an application to swarm deployment.* △

First, we examine the numerical computation of $T^*(x)$.

4.4.3 Numerical Computation of the Monge Map

In general, the Monge map $T(x)$ is difficult to compute numerically [34, 64]. In fact, the optimal control formulation (OT4) was devised by Brenier and Benamou precisely to numerically compute $T(x)$, and devising fast solvers for this problem is an area of active research [62]. In one dimension, a classical result (used in [18]) determines the Monge map in terms of the cumulative distribution functions of the initial and terminal densities as

$$\int_{-\infty}^{x} \rho_0(x)dx = \int_{-\infty}^{T(x)} \rho_1(y)dy. \tag{4.60}$$

This can readily be solved to high precision with a bisection algorithm.

For systems with $n > 1$ states, the situation is more complicated. For example, in the single-integrator system $x_{k+1} = x_k + u_k$, the one-timestep Monge map exists explicitly when the initial and termimal distributions are Gaussian. Suppose ρ_0, ρ_1 are, $\rho_0(x) \sim \mathcal{N}(m_0, \Sigma_0)$, $\rho_1(x) \sim \mathcal{N}(m_1, \Sigma_1)$. Then, the optimal Monge map is a shift and scaling [48], $T(x) = \tilde{A}(x - m_0) + m_1$, with $\tilde{A} = \Sigma_0^{-1/2}\left(\Sigma_0^{1/2}\Sigma_1\Sigma_0^{1/2}\right)^{1/2}\Sigma_0^{-1/2}$.

For general distributions, we outline a discretization-based method for computing π^* from (OT2), and then generating the image of $T^*(x)$ from this approximate π^*. Suppose we discretize $\mathcal{X} = \mathcal{Y} := \mathbb{R}^n$ into cells $\{X_i\}_{i=1}^{n_x}$, $\{Y_j\}_{j=1}^{n_y}$, and then define probability mass vectors $\boldsymbol{\rho}_0 \in \mathbb{R}^{n_x}$, $\boldsymbol{\rho}_1 \in \mathbb{R}^{n_y}$ representing ρ_0, ρ_1, as

$$\boldsymbol{\rho}_i = \int_{X_i} \rho_0(x)dx, \;\; \boldsymbol{\rho}_j = \int_{Y_j} \rho_1(y)dy. \tag{4.61}$$

The cost in (OT2) can be written over this discrete space as

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} C(x, y)d\pi(x, y) \longrightarrow \sum_{i,j} C(x_i, y_j)\boldsymbol{\pi}(x_i, y_j), \tag{4.62}$$

where x_i, y_i are representative coordinates of the cell, say their centroids. The marginal constraints can be imposed on $\boldsymbol{\pi}$ as

$$\boldsymbol{\pi}\mathbf{1}_{n_y} = \boldsymbol{\rho}_0, \ \boldsymbol{\pi}^T\mathbf{1}_{n_x} = \boldsymbol{\rho}_1. \tag{4.63}$$

Letting $C(x_i, y_j) := C_{ij}$, and $\boldsymbol{\pi}(x_i, y_j) := \boldsymbol{\pi}_{ij}$, we arrive at the linear program

$$\begin{array}{ll} \min_{\boldsymbol{\pi}} & \sum_{i,j} C_{ij}\boldsymbol{\pi}_{ij} \\ \text{s.t.} & \boldsymbol{\pi}\mathbf{1}_{n_y} = \boldsymbol{\rho}_0, \ \boldsymbol{\pi}^T\mathbf{1}_{n_x} = \boldsymbol{\rho}_1 \\ & \boldsymbol{\pi}_{ij} \geq 0, \ \forall i, j. \end{array} \tag{4.64}$$

To recover a discrete image of the map T, one has to numerically 'un-do' the pushforward operation that $\boldsymbol{\pi}$ represents. This is done by the element-wise division of $\boldsymbol{\pi}(x_i, y_j)$ by $\boldsymbol{\rho}_0(x_i)$:

$$\boldsymbol{\tau}(x_i, y_j) := \frac{\boldsymbol{\pi}(x_i, y_j)}{\boldsymbol{\rho}_0(x_i)}. \tag{4.65}$$

Note that this definition requires that $\boldsymbol{\rho}_0$ must be strictly positive over the discrete domain; alternately if $\boldsymbol{\rho}_0(x_i) = 0$, then the corresponding row of $\boldsymbol{\pi}$ must also be 0 from the constraints in (4.64). In this case, we can define $\boldsymbol{\tau}(x_i, y_j)$ arbitrarily, since if there is no mass to move from x_i, it is irrelevant where that mass should move to. Note that Problem (4.64) suffers from the 'curse of dimensionality' due to the discretization of $\mathbb{R}^n$. Fast approximations of optimal transport are an ongoing area of research, and in the near future one may expect that Problem (4.64), or approximations of it, may be computationally tractable for large state-spaces [4, 21, 5].

The graph of T over $\{x_i\}_{i=1}^k$ can then be determined by applying the map $\boldsymbol{\tau}$ to the domain $\{x_i\}_{i=1}^k$. Suppose $[X_1, \ldots, X_n]$ are appropriately-vectorized coordinates in each of the n directions of the discretized domain in $\mathbb{R}^n$. Then, the matrix $\boldsymbol{\tau}$ generates the image of T as follows:

$$\boldsymbol{\tau}[X_1, \ldots, X_n] = [T_1, \ldots, T_n], \tag{4.66}$$

where T_i is the vectorized map over the domain in the ith direction of $\mathbb{R}^n$.

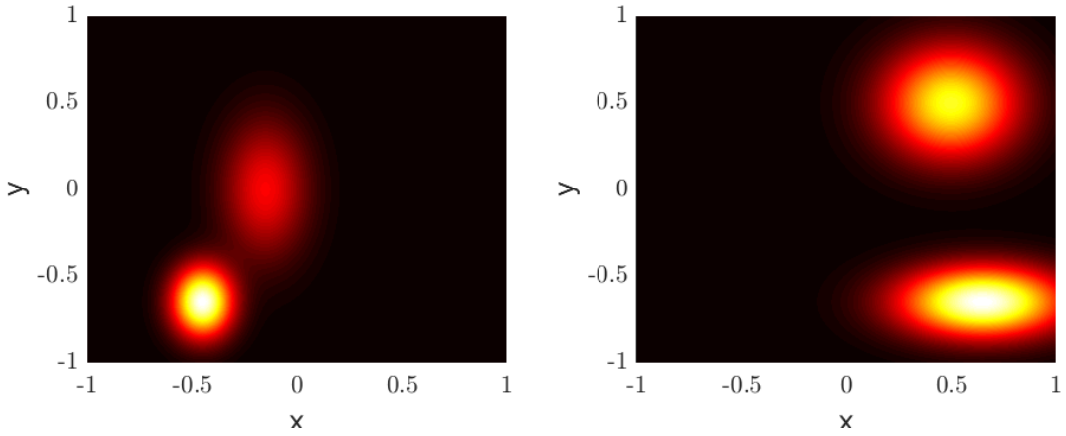

Figure 4.1: True distribitions of the initial (left: $\rho_0(x)$) and target (right: $\rho_1(x)$) states.

4.5 Examples

In this section, we provide numerical experiments of the results in §4.4. Code (and parameters) for the examples can be accessed at [43]. The runtimes for the computation of the optimal transport maps are 12s and 9s, for examples 4.5.1 and 4.5.2 respectively, on an Intel Core i7-9700K CPU (3.60GHz).

4.5.1 2D LQR Example on LTI System

First, we provide an example of the numerical implementation of Theorem 4.1. We implemented the optimal transport method for LQR on a 2-state, 1-input system with matrices

$$A = \begin{bmatrix} 0.9 & -0.1 \\ -0.1 & 0.8 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.67}$$

$$Q_k = I_2, \; R_k = 1, \; 0 \leq k \leq 10. \tag{4.68}$$

Let the states be denoted by $z_k := (z_k^{(1)}, z_k^{(2)})$. Our initial states were distributed according to $\rho_0(x)$ depicted in Fig. 4.1, and we sought to steer the system to the distribution $\rho_1(y)$, also depicted in Fig. 4.1.

The distributions ρ_0, ρ_1 are supported on a discrete grid on the cube $[-1, 1]^2$ with a discretization length $\Delta x = \Delta y = 0.0588$. The optimal transport map T^* was computed by solving (4.64) discretized on this grid, with a cost matrix C computed using (4.22) from

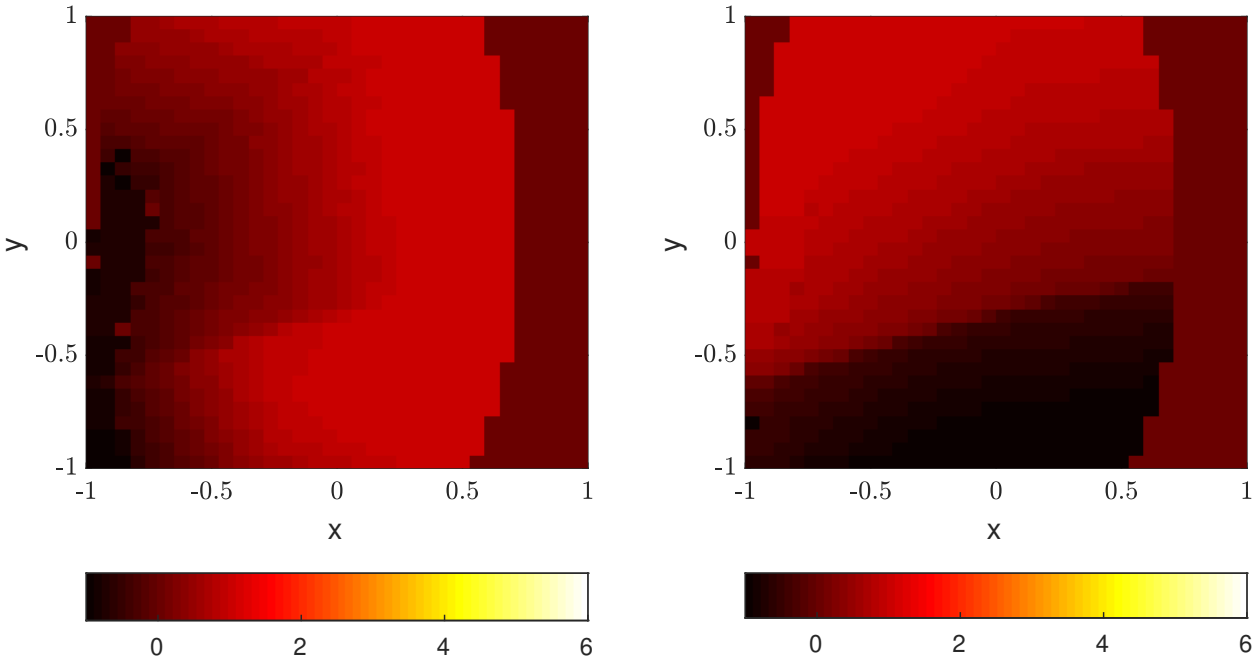

Figure 4.2: Images of the optimal map $T^*(x)$ in the $y^{(1)}$ (left) and $y^{(2)}$ (right) coordinates of the target domain.

Lemma 4.1. In Fig. 4.2, we show a color map of the image of T^*. On the left is the coordinate in the first dimension as a function of the initial $(x^{(1)}, x^{(2)}) = (z_0^{(1)}, z_0^{(2)})$, and on the right is the coordinate in the second dimension as a function of the initial $(z_0^{(1)}, z_0^{(2)})$. One can note that the numerical approximation contains outliers in regions where $\boldsymbol{\rho}_0$ has little mass.

We ran an experiment with 10,000 i.i.d. random initial conditions sampled from ρ_0. For each initial condition z_0, the optimal map computed the corresponding final condition as $z_{t_f} = T^*(z_0)$. The simulation then used the optimal control inputs (4.57) to guide the system to $T^*(z_0)$ over a time horizon of $0 \leq k \leq 10$. Plots of the empirical distributions are shown in Fig. 4.3 at times $t = 0, 3, 7, 10$.

4.5.2 Swarm Deployment - 2D LQR on an LTV System

Consider the task of assigning target positions to n agents whose initial states have an empirical distribution approximating $\rho_0(x)dx$, but not randomly instantiated. For example, consider n agents spaced at constant intervals in the cube $[-1, 1]^2$, as depicted in the middle-left subfigure of Fig. 4.5. Clearly, this is an approximation of a uniform distribution. Our

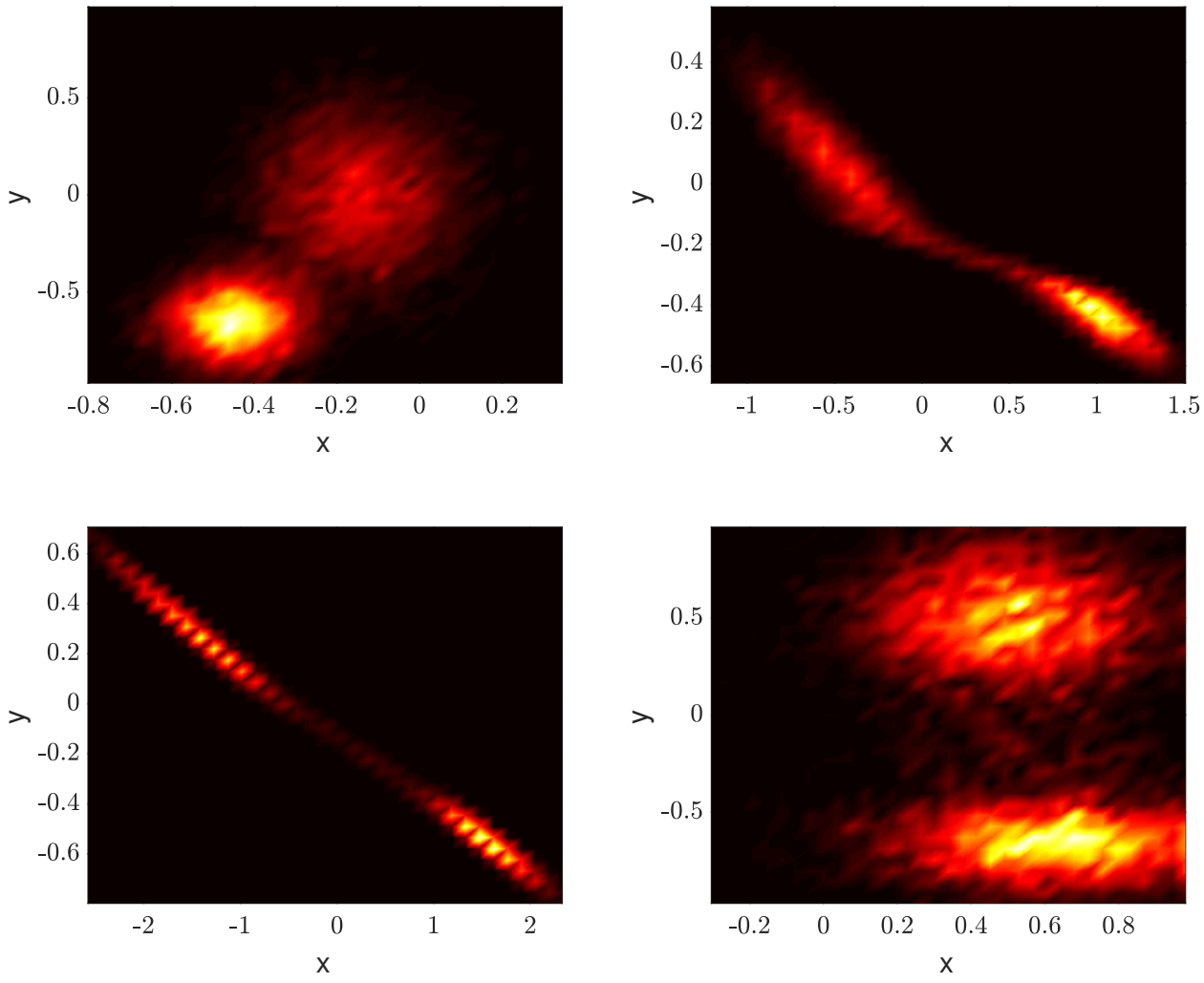

Figure 4.3: Empirical distributions of the states of the system over time. Top left: $t = 0$. Top right: $t = 3$. Bottom left: $t = 7$. Bottom right: $t = 10$.

target distribution is the logo of the Swiss Federal Institute of Technology, Zürich, discretized over a 35×35 pixel domain. A target application could be a swarm of UAVs providing a background performance act during a university event.

Using LTV discrete single-integrator dynamics

$$\begin{aligned} &A_k = Q_k = R_k = I_2,\ 0 \leq k \leq 10, \\ &B_k = I_2,\ 0 \leq k \leq 5,\ B_k = \mathbf{0}_{2\times 2},\ 5 < k \leq 10, \end{aligned} \tag{4.69}$$

we compute the optimal map using (4.64) with the cost matrix (4.22) from Lemma 4.1, depicted in Figure 4.6. The simulation was again produced over a time horizon of $0 \leq k \leq 10$. This time, we plot the explicit mapping between points in a grid and their target states as generated by the map T^*, as shown in the bottom-right of Figure 4.6.

The dynamics (4.69) are controllable, in the sense that the controllability Gramian (4.16) $W_c(t_f, 0)$ is positive-definite, however the matrix S_2 in (4.47) is $\mathbf{0}_{4\times 2}$. Since $\mathbf{0}_{4\times 2}^\dagger = \mathbf{0}_{2\times 4}$, by (4.47), this simply means that the control is zero for $6 \leq k \leq 10$. As the system is controllable, it is steered to the final position by timestep $k = 5$, as evident in Figure 4.4.

4.6 Conclusion

We studied the discrete-time linear-quadratic regulator with uncertainties in the initial state, and how optimal transport can be used to guide the system to a final state with an uncertainty specified by a target probability density. We derived the form of the optimal control from the optimal transport map, and discussed numerical implementations of this map. Finally, we provided numerical examples with an application to swarm deployment.

Future work may include online computation of the optimal transport map corresponding to the LQ cost, and studying systems where additional uncertainty comes from being driven by noise of arbitrary distributions.

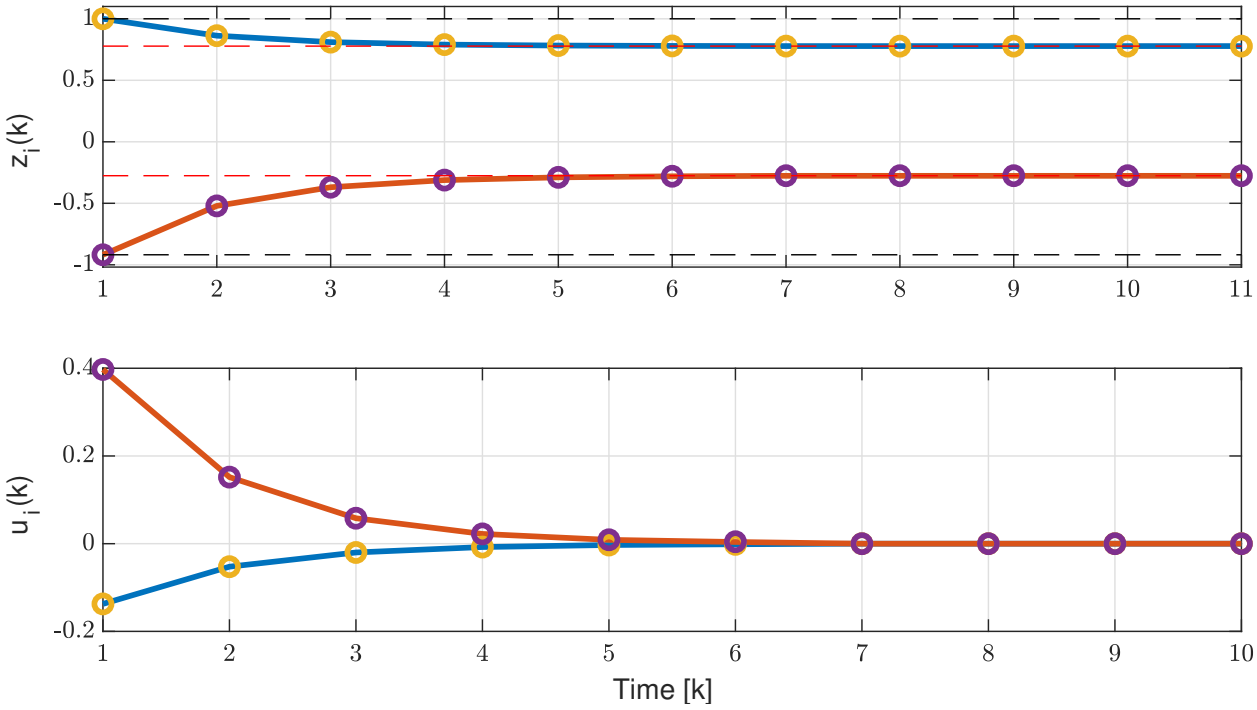

Figure 4.4: Top: Trajectory of dynamics (4.69). Bottom: Control computed by solving (P3) using `cvx`[38] (solid line) and via (4.23) (markers). Red (black) dashed lines indicate target (initial) states.

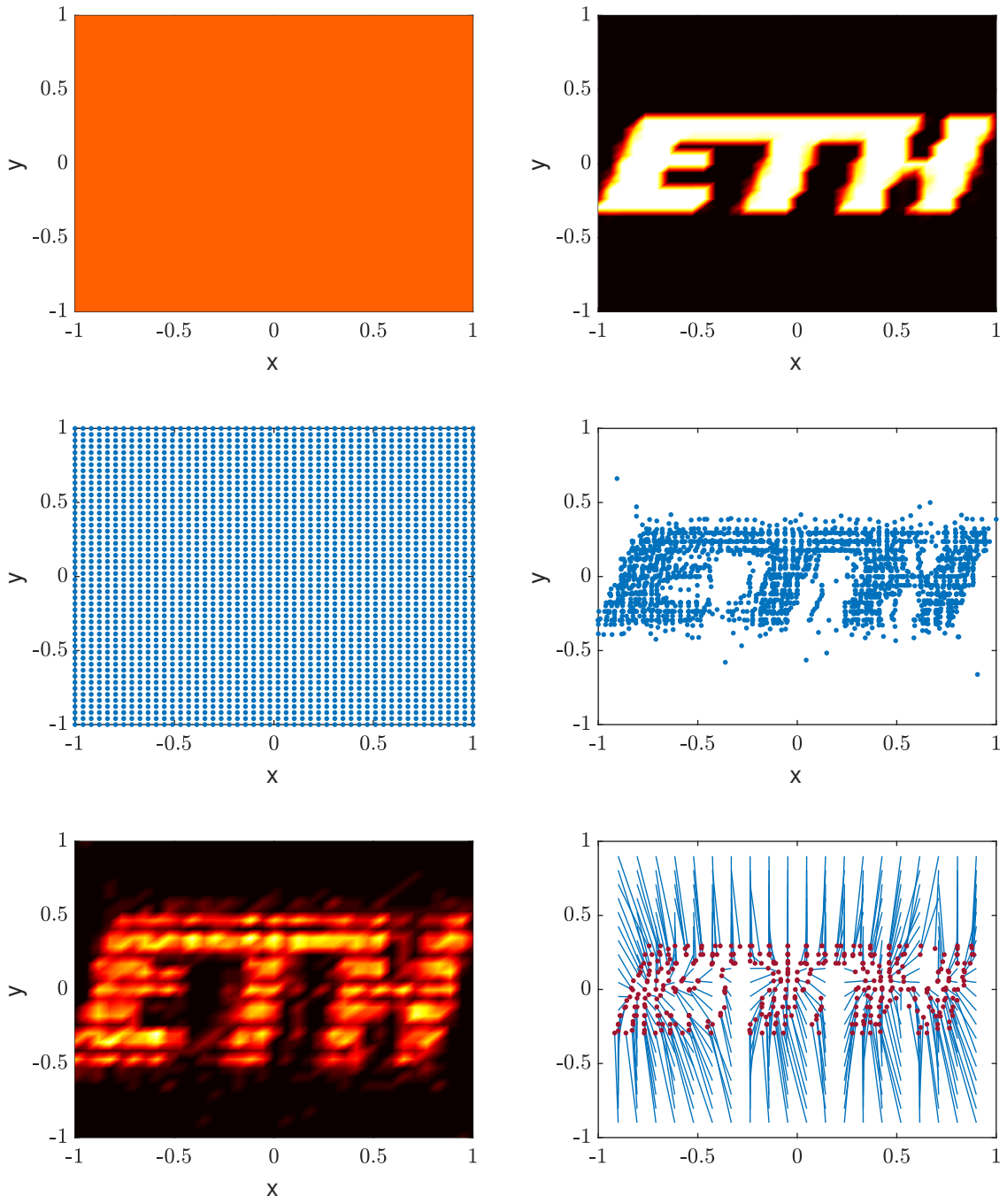

Figure 4.5: Top: plots of the initial uniform distribution (left), and target distribution representing the ETH logo (right). Middle: Initial conditions (x, y) uniformly spaced in $[-1, 1]^2$ (left), and their corresponding terminal conditions $T^*(x, y)$ (right). Bottom: empirical distribution of the terminal states from the middle-right, and a plot showing the corresponding final states interpolated from an initial state.

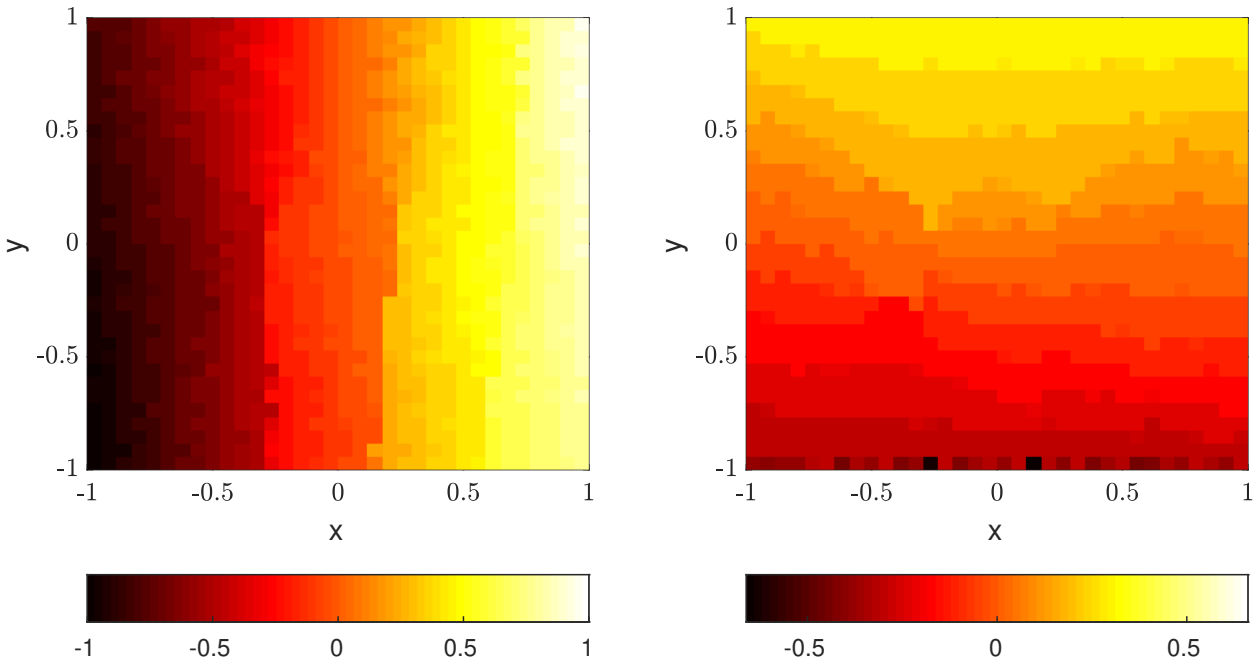

Figure 4.6: Images of the optimal map $T^*(x)$ in the x (left) and y (right) coordinates of the target domain.

Chapter 5

MEASURE FORMULATION OF CONTROL SYSTEMS

5.1 Introduction

At the heart of control theory is the idea of a control system which represents some relationship between input and output signals. We will look at both discrete-time and continuous-time systems, and then develop an alternate way to characterize the dynamics in terms of measures, rather than simply the evolution of a state.

This formulation has some important advantages and disadvantages compared to a state-space model;

- It allows for stochasticity in control and dynamics,
- It can describe a swarm of identical agents which is independent of size (as long as the agents are indestinguishable and follow the same overall dynamics), and
- It losslessly converts a nonlinear, finite-dimensional system into an infinite-dimensional linear system.

5.2 Measures

5.2.1 Definitions and Preliminaries

A *measure* is fundamentally a type of mapping from sets to the nonnegative real numbers $\mathbb{R}_+$. Specifically, we define a *measure space* $(\mathcal{X}, \mathcal{A}, \mu)$ where $\mathcal{X}$ is some set, $\mathcal{A}$ is a σ-algebra on the set $\mathcal{X}$, and μ is a measure. Given the set $\mathcal{X}$, a σ-algebra is a nonempty collection of subsets of $\mathcal{X}$ which are closed under the complement with respect to $\mathcal{X}$, countable unions, and countable intersections. One commonly used σ-algebra is the collection of *Borel* subsets of $\mathbb{R}^n$—here denoted $\mathcal{B}(\mathbb{R}^n)$—defined as the σ-algebra generated by the closed subsets of $\mathbb{R}^n$ or equivalently by the collection of rectangles of the form $\{x \in \mathbb{R}^n \mid a < x \leq b\}$ for some vectors a and $b \in \mathbb{R}^n$.

The measure μ is a mapping with the following properties: if $E_1, E_2, \cdots \in \mathcal{A}$ are mutually disjoint sets, then

1. $\mu(\emptyset) = 0$,
2. $\mu(E_i) \geq 0, \ \forall i$, and
3. $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$.

A measure μ is said to be a *probability measure* if $\mu(\mathcal{X}) = 1$. Furthermore, the *support* of μ—denoted sptμ—is the smallest closed set whose complement has measure zero. Intuitively, the support indicates the set which contains all the "stuff" picked up by the measure.

The *push-forward operator* $f_{\#}$ acts on a measure μ such that:

$$\int_S \varphi(x) \ f_{\#}\mu(\mathrm{d}x) = \int_{f^{-1}(S)} \varphi(f(y)) \ \mu(\mathrm{d}y). \tag{5.1}$$

By abuse of notation, the set $f^{-1}(S)$ is the preimage of set S under mapping f, i.e., $f^{-1}(S) = \{y \mid f(y) \in S\}$. Notice that the domain and co-domain of f are not the same, so the operator $f_{\#}$ re-expresses an intergral in one measure space into that of another measure space, analogous to u-substitution.

[Define marginal and conditional measures]

5.2.2 Lebesgue Integration

The idea of an integral can be extended to measures in a very similar way to the Lebesgue integral, which relies on the Lebesgue measure space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ corresponding to the n-dimensional volume in Euclidean space. E.g., for a function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\int_{\mathcal{X}} f(x) \ \mathrm{d}x := \sup_{y_0 < \cdots < y_N} \sum_{i=0}^{N-1} y_i \lambda\left(X_i\right), \ \text{where} \tag{5.2}$$

$$X_i := \{x \in \mathcal{X} \mid y_i \leq f(x) < y_{i+1})\}. \tag{5.3}$$

This supremum is over all finite partitions of the range of f.

More generally, we can define integrals with respect to other measures, by replacing $\lambda(X_i)$ in Equation (5.2) with another measure $\mu(X_i)$. We would write this as either

$$\int_{\mathbb{R}} f(x)\ \mathrm{d}\mu(x),\ \text{or} \tag{5.4}$$

$$\int_{\mathbb{R}} f(x)\ \mu(\mathrm{d}x). \tag{5.5}$$

An important result connecting integrals with respect to different measures is the *Radon-Nikodym Theorem:*[cite Cohn book on Measure Theory]

Theorem 5.1 (Radon-Nikodym) *Let ν and μ be measures on the measurable space $(X, \mathcal{A})$. If $\nu(X) < \infty$ and $\mu(X) < \infty$, and if ν is absolutely continuous with respect to μ—i.e., $\mu(A) = 0 \Rightarrow \nu(A) = 0,\ \forall A \in \mathcal{A}$—then there is an $\mathcal{A}$-measureable nonnegative function g such that $\nu(A) = \int_A g(x)\ \mathrm{d}\mu(x)$. Furthermore, g is unique up to μ-almost-everywhere equality.*

If such a $g(x)$ in the statement of Theorem 5.1 exists, then g is said to be a *Radon-Nikodym derivative* of ν with respect to μ and is typically denoted $g = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$. The collection of all such g must agree almost everywhere, and so we can consider them all to form an equivalence class. (This type of object is closely tied to the idea of a *distribution*, i.e., a generalized function which cannot be evaluated at a specific point, but rather is defined by how it integrates with respect to a set of so-called test functions.)

One important implication of this theorem is that if the measure μ of any set with zero Lebesgue measure (e.g., any subset of R^3 with zero volume), then μ has an associated density function $\rho(x)$ such that $\mu(A) = \int_A \rho(x)\ \lambda(\mathrm{d}x)$. If μ is a probability measure, then ρ is its probability density function. Furthermore, we can simply evaluate integrals with respect to a measure ν by converting this to a more standard Lebesgue integral:

$$\int_A f(x)\ \nu(\mathrm{d}x) = \int_A f(x)\ \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x)\ \lambda(\mathrm{d}x).$$

To simplify things, we often treat the Radon-Nikodym derivative as a unique *function*, especially if there exists a function in this equivalence class which is continuous. For example, we might say that the probability density function for the standard normal distribution is $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, despite being able to arbitrarily change the value of this function on a set of (Lebesgue) measure zero.

5.2.3 Moments of Measures

One important practical challenge is how to quantify a measure. For example, assuming absoute continuity, we could represent μ as its Radon-Nikodym derivative with respect to the Lebesgue measure (the pdf being a special case for probability measures). One can also directly specify the measures of a family of sets which generate the σ-algebra of the measure space. E.g., probability measures over $\mathbb{R}$, have a cumulative distribution function, giving measures of sets of the form $(-\infty, x]$ explicitly.

Another approach is to give the information required to compute the integral $\int_{\mathcal{S}} \varphi(x)\, \mu(\mathrm{d}x)$, for some family of functions $\varphi(x)$. One way to do this is by specifying the integrals of a set of basis functions $v_i(x)$. Then, the hope is that for $\varphi(x)$ in the span of $v_i(x)$, the integral of phi will be represented as the appropriate linear combination of the integrals of basis functions. If $\varphi(x)$ is in the span of a finite subset of basis functions, then the integral can indeed be computed in this way. However, some caution is needed for functions requiring infinitely many basis functions to ensure convergence. There are many sufficient conditions to this end, but the following theorem is especially relevant.

Theorem 5.2 (Lebesgue's Dominated Convergence Theorem [19]) *Let $(X, \mathcal{A}, \mu)$ be a measure space, let g be a $[0, +\infty]$-valued integrable function on X, and let f and f_1, f_2, ... be $[-\infty, +\infty]$-valued $\mathcal{A}$-measurable functions on X such that $f(x) = \lim_n f_n(x)$ and $|f_n(x)| \leq g(x)$, $n = 1, 2, \ldots$ hold at μ-almost every x in X. Then f and f_1, f_2, ... are integrable, and $\int f(x)\mu(\mathrm{d}x) = \lim_n \int f_n(x)\mu(\mathrm{d}x)$.*

Consider the integral of some function $\varphi(x)$ with respect to a measure μ. By linearity, if we can decompose $\varphi(x)$ as the linear combination of countably many basis functions $v_i(x)$, then

$$\begin{aligned}\int_{\mathcal{X}} \varphi(x)\ \mu(\mathrm{d}x) &= \int_{\mathcal{X}} \left(\sum_{i \in \mathbb{N}} \alpha_i v_i(x) \right)\ \mu(\mathrm{d}x) \\ &= \sum_{i \in \mathbb{N}} \int_{\mathcal{X}} \alpha_i v_i(x)\ \mu(\mathrm{d}x) \\ &= \sum_{i \in \mathbb{N}} \alpha_i \int_{\mathcal{X}} v_i(x)\ \mu(\mathrm{d}x).\end{aligned}$$

Therefore if we know the integrals of a set of basis functions, then we can in principle integrate any suitably nice function in their span by taking the corresponding linear combination of

basis integrals.

To demonstrate the equivalence between measures and and their moments, let $\mathbb{I}_{\mathcal{S}}(x)$ denote the *indicator* function on set $\mathcal{S}$, which is a measurable subset of $\mathbb{X}$:

$$\mathbb{I}_{\mathcal{S}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{S} \\ 0 & \text{if } x \notin \mathcal{S}. \end{cases} \tag{5.6}$$

If $\mathbb{I}_{\mathcal{S}}(x)$ is represented as $\sum_{i \in \mathbb{N}} \alpha_i v_i(x)$, then

$$\mu(\mathcal{S}) = \int_{\mathcal{X}} \mathbb{I}_{\mathcal{S}}(x)\ \mu(\mathrm{d}x) = \sum_{i \in \mathbb{N}} \alpha_i \int_{\mathcal{X}} v_i(x)\ \mu(\mathrm{d}x),$$

assuming that the dominated convergence theorem allows this series to converge. Therefore, the measures of some class of sets $\mathcal{S}$ in this measure space can be computed as a particular linear combination of the moments of μ.

However, be cautioned that a candidate set of moments does not necessarily correspond to a measure. As an example, consider the derivative operator of a function $f : \mathcal{R} \to \mathcal{R}$ evaluated at zero:

$$\begin{aligned} \mathcal{L}(f) &:= f'(0), \\ \mathcal{L}(x^n) &= n. \end{aligned}$$

5.3 Measure Dynamics over State-Space

A typical state-space model tracks how a single state x (e.g., in $\mathbb{R}^n$) changes over time. If one wants to track the evolution of multiple agents at once, then one approach is to stack the states of each individual agent into a larger overall state, which can become intractible as the dimensionality grows exceedingly large. This procedure also struggles to capture an infinite number of agents, such as in an idealized fluid. Instead, measures can be used to characterize state dynamics of a complex dynamical system by instead considering how a measure derived from the distribution of agent states evolves.

5.3.1 Discrete Time

Consider the following time-invariant, discrete-time model with state x, control input u, and disturbance w:

$$x_{k+1} = f(x_k, u_k, w_k). \tag{5.7}$$

The state at time step $k+1$ entirely depends on the tuple of state, control input, and disturbance at time step k. For a single agent, we can recursively generate a sequence of states $(x_0, x_1, \dots)$ from the initial state x_0, and the control and disturbance sequences $(u_0, u_1, \dots)$ and $(w_0, w_1, \dots)$, respectively.

Another way to understand System (5.3.1) is as a sequence of three-phase processes. The first phase generates control input u_k which converts the state x_k into the state-action pair (x_k, u_k). The second phase generates the disturbance w_k to form the tuple (x_k, u_k, w_k). Finally, the third phase applies the function f, resulting in x_{k+1}. This perspective makes it straightforward to explicitly consider u_k as a function of x_k, and for w_k to be a function of both x_k and u_k.

Now consider the case in which x_0 is not precisely known, but rather follows some probability distribution represented by the measure $\mu_0(S) := \Pr[x_0 \in S]$. Then for any specified control sequence, we can hope to solve for a sequence of measures, rather than states: $(\mu_0, \mu_1, \dots)$ such that $\mu_k(S) = \Pr[x_k \in S]$. To make this idea rigorous, one can construct the measure spaces $(\mathcal{X}, \mathcal{A}, \rho_k)$, $(\mathcal{X} \times \mathcal{U}, \mathcal{B}, \mu_k)$, and $(\mathcal{X} \times \mathcal{U} \times \mathcal{W}, \mathcal{C}, \nu_k)$—which correspond to the distributions of states, state-action pairs, and state-action-disturbance tuples, respectively—such that ρ_k is a marginal measure of μ_k, which is in turn a marginal measure of ν_k, and ρ_{k+1} is a push-forward measure of ν_k:

$$\int_{\mathcal{U}} \mu_k(\mathcal{S}, \mathrm{d}u) = \rho_k(\mathcal{S}), \tag{5.8}$$

$$\int_{\mathcal{W}} \nu_k(\mathcal{S}, \mathcal{T}, \mathrm{d}w) = \mu_k(\mathcal{S}, \mathcal{T}), \tag{5.9}$$

$$\rho_{k+1}(\mathcal{S}, \mathcal{T}) = (f_{\#}\nu_k)(\mathcal{S}, \mathcal{T}), \quad \forall \mathcal{S} \in \mathcal{A},\ \forall \mathcal{T} \in \mathcal{B}. \tag{5.10}$$

A more useful form for these relations is obtained by integrating a continuous and bounded test function $\varphi(x)$ (or $\psi(x, u)$) with respect to the measures on both sides of the

equations:

$$\iint_{\mathcal{X}\times\mathcal{U}} \varphi(x)\ \mu_k(\mathrm{d}x,\mathrm{d}u) = \int_{\mathcal{X}} \varphi(x)\ \rho_k(\mathrm{d}x), \tag{5.11}$$

$$\iiint_{\mathcal{X}\times\mathcal{U}\times\mathcal{W}} \psi(x,u)\ \nu_k(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w) = \iint_{\mathcal{X}\times\mathcal{U}} \psi(x,u)\ \mu_k(\mathrm{d}x,\mathrm{d}u), \tag{5.12}$$

$$\int_{\mathcal{X}} \varphi(x)\ \rho_{k+1}(\mathrm{d}x) = \iiint_{\mathcal{X}\times\mathcal{U}\times\mathcal{W}} \varphi(f(x,u,w))\ \nu_k(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w). \tag{5.13}$$

5.3.2 Continuous Time

We can also consider the continuous-time system

$$\dot{x}(t) = f(x(t),u(t),w(t)), \tag{5.14}$$

which can be approximated for small Δt as

$$x(t+\Delta t) = x(t) + f(x(t),u(t),w(t))\Delta t + o(\Delta t). \tag{5.15}$$

Analogously to §5.3.1, we can consider the time-varying measure μ_t by letting the step size Δt in this discrete-time approximation approach zero. Let $g(x,u,w) = x + f(x,u,w)\Delta t$, and assume $\varphi(x)$ is smoothly differentiable:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi(x)\rho_t(\mathrm{d}x) &:= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\int \varphi(x)\ \rho_{t+\Delta t}(\mathrm{d}x) - \int \varphi(x)\ \rho_t(\mathrm{d}x)\right)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\int \varphi(x)\ g_{\#}\nu_t(\mathrm{d}x) - \int \varphi(x)\ \rho_t(\mathrm{d}x)\right)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\int \varphi(g(x,u,w))\ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w) - \int \varphi(x)\ \rho_t(\mathrm{d}x)\right)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left(\int \varphi(x + f(x,u,w)\Delta t + o(\Delta t))\ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w) - \int \varphi(x)\ \rho_t(\mathrm{d}x)\right)\\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int \nabla\varphi(x)\cdot[f(x,u,w)\Delta t + o(\Delta t)]\ \ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w)\\
&= \int \nabla\varphi(x)\cdot f(x,u,w)\ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w).
\end{aligned}$$

Analogously to the discrete-time case, $\rho_t(\mathcal{X})$ is a marginal measure of $\mu_t(\mathcal{X},\mathcal{U})$, which in turn is a marginal measure of $\nu_t(\mathcal{X},\mathcal{U},\mathcal{W})$. These are combined with the above dynamic

relationship for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t(\mathcal{X})$ and are expressed in the form of integrals of general test functions $\varphi(x)$ and $\psi(x,u)$:

$$\iint_{\mathcal{X}\times\mathcal{U}} \varphi(x)\ \mu_t(\mathrm{d}x,\mathrm{d}u) = \int_{\mathcal{X}} \varphi(x)\ \rho_t(\mathrm{d}x), \tag{5.16}$$

$$\iiint_{\mathcal{X}\times\mathcal{U}\times\mathcal{W}} \psi(x,u)\ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w) = \iint_{\mathcal{X}\times\mathcal{U}} \psi(x,u)\ \mu_t(\mathrm{d}x,\mathrm{d}u), \tag{5.17}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{X}} \varphi(x)\ \rho_t(\mathrm{d}x) = \iiint_{\mathcal{X}\times\mathcal{U}\times\mathcal{W}} \nabla\varphi(x)\cdot f(x,u,w)\ \nu_t(\mathrm{d}x,\mathrm{d}u,\mathrm{d}w). \tag{5.18}$$

Chapter 6

MAXIMAL CONTROL INVARIANT SETS FOR MARKOV DECISION PROCESSES

6.1 Introduction

Invariant sets play a crucial role in certifying safety constraints for control systems. Specifically, if there is a control law that keeps the state of a system in a desirable region, then we can keep the state in a safe region. Finding such controlled invariant sets—and their corresponding control laws—has been the subject of much study [12][13][46], specifically in the context of Markov decision processes (MDPs) [28][1].

The technique of transforming a nonlinear state-space model into an infinite-dimensional linear system is a powerful tool for the analysis of control systems. This extends the tools of linear system theory to more general nonlinear systems, and various quantities of interest in control theory such as reachable sets and positively invariant sets can be formulated as solutions to infinite-dimensional linear programs over the space of measures [57][50][59]. However, this comes at the cost of needing to truncate the dimensionality of the space of measures from uncountably infinite to finite dimensions in practice.

The method of Lasserre hierarchies [57] offers a systematic way to bound the solutions to these infinite-dimensional linear programs with solutions to finite-dimensional convex programs of fixed degree, e.g., by optimizing over finitely many moments of the measures, or by optimizing over nonnegative polynomials of fixed degree in the dual problem.

Here we apply the results of Korda et al. [50] to find the maximal controlled invariant (MCI) set for a Markov decision process (MDP) under polyhedral constraints to the state vector and policy. To the authors' knowledge, sum-of-squares (SoS) programming has not yet been analyzed in much detail, apart from the more general class of polynomial systems with basic semi-algebraic constraints. We exploit the particular structure of MDP dynamics to explicitly eliminate the equality constraints from the state and control to reduce the dimension of the problem. We then present the specific form of this SoS program for MDPs, using the moments of the so-called canonical simplex [56], which are of particular importance to MDP dynamics. Furthermore, since the SoS procedure necessarily produces an outer

approximation to the MCI set, we can compare this method with previous work on invariant sets for Markov chains to bound the MCI set from both the outside and inside [50][45][46].

6.1.1 Notation

We write the vector of ones of dimension n as 1_n, and the $m \times n$ matrix of zeros as $0_{m\times n}$ (or just 0 if the dimensions are apparent). If $x, y \in \mathbb{R}^n$, then $x \leq y$ means that $x_i \leq y_i$, for all $i = 1, \ldots, n$. R^n_+ is the set of nonnegative vectors in $\mathbb{R}^n$, and Δ^n denotes the so-called unit simplex embedded in $\mathbb{R}^n$, i.e., $\Delta^n := \{x \in \mathbb{R}^n_+ \mid 1^\intercal x = 1\}$. Note that Δ^n has dimension $n-1$ because of the affine constraint $1^\intercal x = 1$. The support of measure μ is denoted $\mathrm{spt}(\mu)$, and is loosely defined as the closure of the set on which μ is nonzero. When f is a function with domain $\mathcal{X}$ and codomain $\mathcal{Y}$, then we denote the image of set $\mathcal{A}$ as $f(\mathcal{A}) := \{f(x) \mid x \in \mathcal{A}\}$. We similarly denote the pre-image of set $\mathcal{B}$ as $f^{-1}(\mathcal{B}) := \{x \in \mathcal{X} \mid f(x) \in \mathcal{B}\}$. $A \bigotimes B$ denotes the Kronecker product of matrices A and B. $\mathrm{diag}(v)$ is the square diagonal matrix whose (i,i)-th entry is v_i. For continuous function f and measure μ, the *push-forward measure* $f_{\#}\mu$ has the defining property

$$\int_{\mathcal{A}} \varphi(y)\ (f_{\#}\mu)(\mathrm{d}y) := \int_{f^{-1}(\mathcal{A})} \varphi(f(x))\ \mu(\mathrm{d}x)$$

for all test functions $\varphi : \mathcal{A} \to \mathbb{R}$, where both integrals are well-defined (e.g., $\mathcal{A}$ is a Borel set). Informally, the push-forward measure results from applying the mapping f to every infinitesimal bit of mass which makes up measure μ. E.g., if μ is the Dirac delta concentrated at (x_k, u_k), then $f_{\#}\rho$ is the Dirac delta concentrated at $x_{k+1} = f(x_k, u_k)$.

6.2 MCI Set for a Discrete-Time System

For the sake of making this paper self-contained, we review the general theoretical framework outlined by Korda et al. [50] to formulate MCI sets as solutions to infinite-dimensional linear programs of positive measures, and we present the specific case of MDPs in §6.3. In particular, we consider the following finite-dimensional nonlinear time-invariant control system in discrete time:

$$x_{k+1} = f(x_k, u_k), \quad f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \tag{6.1}$$

$$x_k \in \mathcal{X} \subset \mathbb{R}^n, \quad u_k \in \mathcal{U} \subset \mathbb{R}^m, \quad k = 0, 1, 2, \ldots. \tag{6.2}$$

We call $\mathcal{X}$ the *constraint set* and $\mathcal{U}$ the set of feasible controls. In principle $\mathcal{U}$ may be a function of the state, but here we just consider uniform control constraints through careful formulation of the dynamics. A set $\mathcal{A} \subseteq \mathrm{R}^n$ is said to be *controlled invariant* (abbreviated CI) for System (6.1) if for all $x_0 \in \mathcal{A}$, there exists a sequence of $u_k \in \mathcal{U}$ such that $x_k \in \mathcal{A}$ for all $k > 0$. The empty set is therefore controlled invariant, as is $\mathbb{R}^n$.

A key property of controlled invariant sets is that they are closed under the union operation $\cup$. In other words, controlled invariant sets form a complete join-semilattice for the partial order $\subseteq$, where the "join" (i.e. least upper bound) of two CI sets $\mathcal{A}_1$ and $\mathcal{A}_2$ is simply $\mathcal{A}_1 \cup \mathcal{A}_2$. As a result, the *maximal controlled invariant (MCI)* subset of $\mathcal{X}$ is the union of all controlled invariant subsets of $\mathcal{X}$. The MCI set is maximal in the sense that it is itself controlled invariant, and all controlled invariant subsets of $\mathcal{X}$ are subsets of the MCI set.

Controlled invariant sets play a crucial role in verifying that a trajectory satisfies a safety constraint. If x_0 is in some controlled invariant subset of $\mathcal{X}$, then by definition there exists a control sequence which satisfies the safety constraint $x_k \in \mathcal{X}$, $\forall k \in \mathbb{N}$. Since this would also imply that x_0 is in the MCI set, we may equivalently define the MCI set as the set of initial states for System (6.1) that satisfy the safety constraint for some control sequence. As a result, if x_0 is not in the MCI subset of $\mathcal{X}$, then every feasible trajectory will always eventually leave the constraint set $\mathcal{X}$.

6.2.1 Linear Conditions on the Occupation Measure

To motivate the measure solutions of System (6.1), we could consider how a collection of agents following these dynamics evolves. The trajectory of state-action pairs is completely determined by the initial state vector x_0 and the sequence of inputs $\{u_k\}_{k\in\mathbb{N}}$. More generally, we allow for a randomized control law $\kappa_k(\mathcal{U} \mid x)$ which is a probability measure over the set of controls $\mathcal{U}$ at time k and state vector x. Now consider a distribution of initial states $\rho_0(x)$, with each state evolving according to these dynamics. We can then describe how this distribution evolves over time under control law $\kappa_k(\cdot \mid x)$—a probability measure on $\mathcal{U}$ which is a function of the state vector x—in a similar way:

$$\mu_k(\mathrm{d}x, \mathrm{d}u) = \kappa_k(\mathrm{d}u \mid x)\rho_k(\mathrm{d}x) \tag{6.3}$$

$$\rho_{k+1}(\mathrm{d}x) = (f_{\#}\mu_k)(\mathrm{d}x), \tag{6.4}$$

where $f_{\#}\mu$ is the push-forward measure of f acting upon measure μ. In other words, Equation (6.3) states that we decompose μ_k into the product of a conditional distribution κ_k and the corresponding marginal measure ρ_k, by the disintegration theorem (see: e.g., [15]).

We also can incorporate control into the dynamics of the measures more generally. Consider the joint measure $\mu = M(\mathcal{X} \times \mathcal{U})$. Then, $\mu(\mathcal{A} \times \mathcal{B})$ is the mass of the agents in the state region $\mathcal{A}$ and taking an action in set $\mathcal{B}$. The measure $\rho(\mathrm{d}x)$ over states is the marginal measure $\mu(\mathrm{d}x, \mathcal{U})$, and the (mixed) control law is a probability distribution over $\mathcal{U}$ conditioned upon state vector $x \in \mathcal{X}$.

Korda et al. [50] presented an infinite-dimensional LP whose solution implicitly yields the MCI set. Let $I_{\mathcal{P}}$ denote the indicator function on set $\mathcal{P}$; namely $I_{\mathcal{P}}(x, u) = 1$ when $(x, u) \in \mathcal{P}$, and $I_{\mathcal{P}}(x, u) = 0$ otherwise. The *discounted occupation measure* is defined for a discount factor $\alpha \in (0, 1)$, initial state x_0, and sequence of admissible state-action pairs $\{(x_{k|x_0}, u_{k|x_0})\}$ consistent with equation (6.1):

$$\mu\left(\mathcal{P} \mid \{(x_k, u_k)\}; \alpha\right) = \sum_{k=0}^{\infty} \alpha^k I_{\mathcal{P}}\left(x_k, u_k\right), \tag{6.5}$$

with $\mathcal{P}$ being any measurable subset of $\mathcal{X} \times \mathcal{U}$. The occupation measure is a discounted count of how many times the state-action pair enters $\mathcal{P}$, given the initial condition x_0 and control sequence $\{u_{k|x_0}\}$. This definition is extended to a sequence of measures μ_k over $\mathcal{X} \times \mathcal{U}$:

$$\tilde{\mu}(\mathcal{P} \mid \{\mu_k\}_{k \in \mathbb{N}}; \alpha) = \sum_{k=0}^{\infty} \alpha^k \mu_k(\mathcal{P}). \tag{6.6}$$

We will adopt the methodology for computing the MCI set for polynomial systems given by Korda et al. [50] to MDP dynamics. System (6.1) defines a trajectory with initial condition $x(0) = x_0$, with a time-varying input $u(t)$. We then extend these dynamics to an ensemble of trajectories described by the sequence of positive measures $\{\mu_t\}_{t \in \mathbb{N}}$ so that $\mu_t(\mathcal{A} \times \mathcal{B})$ is the total "mass" on the set $\mathcal{A}$ taking an action in $\mathcal{B}$ at time t.

The following lemma shows that any discounted occupation measure $\tilde{\mu}$ follows a linear relation with its initial measure over state space ρ_0.

Theorem 6.1 *Let ρ_0 be the initial mass distribution for the system $x_{k+1} = f(x_k, u_k)$, and let $\tilde{\mu}$ be the occupation measure with discount factor $\alpha \in (0, 1)$ generated by some (possibly*

time-varying) stochastic control policy, then $\tilde{\mu}$ satisfies the equation

$$\iint_{\mathcal{X}\times\mathcal{U}} (\varphi(x) - \alpha\varphi(f(x,u)))\,\tilde{\mu}(\mathrm{d}x,\mathrm{d}u) = \int_{\mathcal{X}} \varphi(x)\rho_0(\mathrm{d}x) \tag{6.7}$$

for all continuous, bounded functions $\varphi : \mathcal{X} \to \mathbb{R}$.

Proof: Now we derive a condition relating a discounted occupation measure to the initial distribution $\rho_0 \in M(\mathcal{X})$. We first give the equations relating the measure sequences $\rho_k(X)$ and $\mu_k(X,U)$:

$$\int_{\mathcal{X}} \psi(x)\rho_k(\mathrm{d}x) = \iint_{\mathcal{X}\times\mathcal{U}} \psi(x)\mu_k(\mathrm{d}x,\mathrm{d}u), \tag{6.8}$$

$$\int_{\mathcal{X}} \varphi(x)\rho_{k+1}(\mathrm{d}x) = \iint_{\mathcal{X}\times\mathcal{U}} \varphi(f(x,u))\mu_k(\mathrm{d}x,\mathrm{d}u), \tag{6.9}$$

for all continuous ψ and φ, and for all $k \in \mathbb{N}$. We then integrate a test function φ with respect to the occupation measure defined in equation (6.6).

$$\begin{aligned}
&\textstyle\iint_{\mathcal{X}\times\mathcal{U}} \varphi(x)\tilde{\mu}(\mathrm{d}x,\mathrm{d}u;\alpha) \\
&\quad = \textstyle\iint_{\mathcal{X}\times\mathcal{U}} \varphi(x) \sum_{k=0}^{\infty} \alpha^k \mu_k(\mathrm{d}x,\mathrm{d}u) && \text{using (6.6)} \\
&\quad = \textstyle\sum_{k=0}^{\infty} \alpha^k \iint_{\mathcal{X}\times\mathcal{U}} \varphi(x)\mu_k(\mathrm{d}x,\mathrm{d}u) \\
&\quad = \textstyle\sum_{k=0}^{\infty} \alpha^k \int_{\mathcal{X}} \varphi(x)\rho_k(\mathrm{d}x) && \text{using (6.8)} \\
&\quad = \textstyle\int_{\mathcal{X}} \varphi(x)\rho_0(\mathrm{d}x) \\
&\qquad \textstyle+\alpha \sum_{k=0}^{\infty} \alpha^k \int_{\mathcal{X}} \varphi(x)\rho_{k+1}(\mathrm{d}x) \\
&\quad = \textstyle\int_{\mathcal{X}} \varphi(x)\rho_0(\mathrm{d}x) && \text{using (6.9)} \\
&\qquad \textstyle+\alpha \sum_{k=0}^{\infty} \alpha^k \iint_{\mathcal{X}\times\mathcal{U}} \varphi(f(x,u))\mu_k(\mathrm{d}x,\mathrm{d}u)
\end{aligned}$$

$$\begin{aligned}
\textstyle\iint_{\mathcal{X}\times\mathcal{U}} \varphi(x)\tilde{\mu}(\mathrm{d}x,\mathrm{d}u;\alpha) &= \textstyle\int_{\mathcal{X}} \varphi(x)\rho_0(\mathrm{d}x) \\
\textstyle+\alpha \iint_{\mathcal{X}\times\mathcal{U}} \varphi(f(x,u))\tilde{\mu}(\mathrm{d}x,\mathrm{d}u;\alpha) && \text{using (6.6)}
\end{aligned}$$

Therefore, any occupation measure $\tilde{\mu}$ must satisfy the linear equality

$$\begin{aligned}
\textstyle\iint_{\mathcal{X}\times\mathcal{U}} (\varphi(x) - \alpha\varphi(f(x,u)))\,\tilde{\mu}(\mathrm{d}x,\mathrm{d}u;\alpha) \\
= \textstyle\int_{\mathcal{X}} \varphi(x)\rho_0(\mathrm{d}x), \quad \forall\varphi.
\end{aligned} \tag{6.10}$$

■

Notice that the specific policy κ does not appear in Equation (6.10), rather this is a necessary (and sufficient) condition that all occupation measures satisfy. We may also extract a stochastic time-invariant policy $\kappa(\mathrm{d}u \mid x)$ for any such $\tilde{\mu}$, as we show in the following theorem.

Theorem 6.2 *Let $\tilde{\mu}(\mathrm{d}x, \mathrm{d}u)$ be a positive measure satisfying Equation (6.7), and let $\tilde{\kappa}(\mathrm{d}u|x)$ be the corresponding conditional measure of $\tilde{\mu}$, i.e.,*

$$\tilde{\mu}(\mathrm{d}x, \mathrm{d}u) = \tilde{\kappa}(\mathrm{d}u \mid x)\left(\rho_0(\mathrm{d}x) + \alpha f_{\#}\tilde{\mu}(\mathrm{d}x)\right). \tag{6.11}$$

Then, the control system (6.3) has discounted occupation measure $\tilde{\mu}$ with discount factor $\alpha \in (0, 1)$.

Proof: First, we confirm that $\tilde{\kappa}$ is the conditional measure of $\tilde{\mu}$ as a function of x by noticing that Equation (6.7) may be rearranged and rewritten in brief form as $\int_{\mathcal{U}} \tilde{\mu} = \rho_0 + \alpha f_{\#}\tilde{\mu}$, and thus $\tilde{\kappa}$ is an appropriate conditional measure: $\tilde{\mu}(\mathrm{d}x, \mathrm{d}u) = \tilde{\kappa}(\mathrm{d}u|x \int_{\mathcal{U}} \tilde{\mu}(\mathrm{d}x, \mathrm{d}u)$.

Next, notice that (6.3) implies $\mu_{k+1} = \kappa_k f_{\#}\mu_k$. We substitute $\kappa_k = \tilde{\kappa}$ and solve for μ_k in terms of ρ:

$$\mu_k = (\tilde{\kappa} f_{\#})^k \mu_0 = (\tilde{\kappa} f_{\#})^k \tilde{\kappa} \rho_0. \tag{6.12}$$

Rearranging Equation (6.11) as $\tilde{\kappa}\rho_0 = \tilde{\mu} - \alpha\tilde{\kappa} f_{\#}\tilde{\mu}$, we can then explicitly write μ_k in terms of $\tilde{\mu}$:

$$\begin{aligned} \mu_k = (\tilde{\kappa} f_{\#})^k \tilde{\kappa} \rho_0 &= (\tilde{\kappa} f_{\#})^k (\tilde{\mu} - \alpha \tilde{\kappa} f_{\#} \tilde{\mu}) \\ &= (\tilde{\kappa} f_{\#})^k \tilde{\mu} - \alpha (\tilde{\kappa} f_{\#})^{k+1} \tilde{\mu}. \end{aligned}$$

We can finally compute the discounted occupation measure:

$$\begin{aligned}\sum_{k=0}^{\infty}\alpha^k\mu_k &= \sum_{k=0}^{\infty}\alpha^k\left((\tilde{\kappa}f_{\#})^k\tilde{\mu}-\alpha(\tilde{\kappa}f_{\#})^{k+1}\tilde{\mu}\right)\\ &= \sum_{k=0}^{\infty}\left((\alpha\tilde{\kappa}f_{\#})^k\tilde{\mu}-(\alpha\tilde{\kappa}f_{\#})^{k+1}\tilde{\mu}\right)\\ &= \sum_{k=0}^{\infty}(\alpha\tilde{\kappa}f_{\#})^k\tilde{\mu}-\sum_{k=1}^{\infty}(\alpha\tilde{\kappa}f_{\#})^k\tilde{\mu}\\ &= \tilde{\mu}.\end{aligned}$$

■

6.2.2 Linear Program for MCI Set

As shown by Korda et al. [50], the MCI set for System (6.1) subject to the safety constraints $(x(t),u(t))\in\mathcal{X}\times\mathcal{U},\ \forall t\in\mathbb{N}$ is the support of an initial mass measure ρ_0 which is consistent with some occupation measure $\tilde{\mu}$ satisfying the safety constraints. We write the primal equations from Korda et al. [50] in a form consistent with our notation (eliminating the slack measure for clarity):

$$\begin{aligned}&p^* = \sup_{\tilde{\mu},\rho_0}\rho_0(\mathcal{X})\\ &\text{s.t.}\begin{cases}\iint[v(x)-\alpha v(f(x,u))]\tilde{\mu}(\mathrm{d}x,\mathrm{d}u)\\ \qquad=\int v(x)\rho_0(\mathrm{d}x), & \forall v\\ \int w(x)\rho_0(\mathrm{d}x)\le\int w(x)\lambda(\mathrm{d}x), & \forall w\ge 0\\ \tilde{\mu}\ge 0,\quad \rho_0\ge 0,\quad \text{spt }\tilde{\mu}\subseteq\mathcal{X}\times\mathcal{U},\end{cases}\end{aligned}\tag{6.13}$$

If $\mathcal{X}\times\mathcal{U}$ is bounded (as is the case for MDPs), we only need the first and second constraints to be satisfied for all polynomials v and w such that w is nonnegative over $\mathcal{X}$. Note that we have also removed the explicit constraint on the support of ρ_0, since spt $\tilde{\mu}\subseteq\mathcal{X}\times\mathcal{U}\ \Rightarrow$ spt $\rho_0\subseteq\mathcal{X}$ for all $\tilde{\mu}$ and ρ_0 satisfying the "v" constraint. The main idea is that we can find the MCI set by maximizing the mass of the initial measure ρ_0 with a finite density upper bound, which is consistent with some occupation measure contained entirely inside the constraint sets, thereby satisfying the state constraints for some state feedback control law. The optimal value p^* is therefore the volume of the MCI set. This primal problem may be solved as-is by optimizing over their moments subject to the support constraints [50], but this approach is

outside the scope of this paper. Rather, we instead focus on the dual problem.

The dual of this infinite-dimensional LP is as follows:

$$d^* = \inf_{v,w} \int_{\mathcal{X}} w(x)\lambda(\mathrm{d}x) \tag{6.14}$$

$$\text{s.t.} \begin{cases} v(x) - \alpha v(f(x,u)) \geq 0, & \forall (x,u) \in \mathcal{X} \times \mathcal{U} \\ w(x) - v(x) - 1 \geq 0, & \forall x \in \mathcal{X} \\ w(x) \geq 0, & \forall x \in \mathcal{X}. \end{cases} \tag{6.15}$$

This gives another characterization of the MCI set as $\{x \in \mathcal{X} \mid w(x) \geq 1\}$, for any optimal $w(x)$. Furthermore, for any $w_k(x)$ in the feasible set for the dual LP, the MCI set is a subset of $\{x \in \mathcal{X} \mid w_k(x) \geq 1\}$. The dual problem solves for the value function $v(x)$, and an auxilliary nonnegative function $w(x)$ which bounds $v(x)+1$ from above. By optimizing $v(x)$ and $w(x)$ over a sufficiently expressive family of functions satisfying the dual constraints, then we could obtain an arbitrarily good outer approximation of the MCI set, in principle.

Suppose $\mathcal{X}$ and $\mathcal{U}$ are compact basic semialgebraic sets, i.e.,

$$\mathcal{X} = \{x \in \mathbb{R}^n \mid \tilde{g}_i(x) \geq 0,\ i = 1, 2, \ldots, n_X\}, \tag{6.16}$$

$$\mathcal{U} = \{u \in \mathbb{R}^m \mid \tilde{h}_i(u) \geq 0,\ i = 1, 2, \ldots, n_U\}. \tag{6.17}$$

The constraints of the dual LP are satisfied if and only if there exist nonnegative functions $p_i(x)$, $q_i(x,u)$, $r_i(x,u)$, and $s_i(x)$ such that

$$\begin{aligned} v(x) - \alpha v(f(x,u)) &\geq \textstyle\sum_{i=1}^{n_X} q_i(x,u)\tilde{g}_i(x) \\ &\quad + \textstyle\sum_{i=1}^{n_U} r_i(x,u)\tilde{h}_i(u), \\ w(x) - v(x) - 1 &\geq \textstyle\sum_{i=1}^{n_X} p_i(x)\tilde{g}_i(x), \\ w(x) &\geq \textstyle\sum_{i=1}^{n_X} s_i(x)\tilde{g}_i(x), \end{aligned} \tag{6.18}$$

$\forall (x,u) \in \mathbb{R}^{n \times m}$. Notice that if $(x,u) \in \mathcal{X} \times \mathcal{U}$, then the left-hand sides are greater than or equal to sums of products of nonnegative functions, and are thus nonnegative on the domain of interest.

6.2.3 MCI Set for Polynomial System with Semi-algebraic Constraints

Although the problem (6.14) is infinite-dimensional, we can obtain an outer approximate solution by optimizing over a finite-dimensional subset of feasible v and w by replacing each

inequality with a stronger sum-of-squares polynomial constraint. Let $f(x,u)$, $\tilde{g}(x)$, and $\tilde{h}_i(u)$ be polynomials. Then, the inequality conditions (6.18) are satisfied for polynomials $v(x)$ and $w(x)$ such that:

$$v + \alpha v \circ f - \sum_i q_i \tilde{g}_i - \sum_i r_i \tilde{h}_i \in SoS, \tag{6.19}$$

$$w - v - 1 - \sum_i p_i \tilde{g}_i \in SoS, \tag{6.20}$$

$$w - \sum_i s_i \tilde{g}_i \in SoS. \tag{6.21}$$

One hopes that the SoS polynomials are expressive enough so that we can achieve an arbitrarily good approximation of the true dual constraints by choosing a high enough degree. This is indeed the case for a general MDP. We justify replacing the inequality constraints of (6.14) with polynomial SoS conditions with the aid of the following powerful result [58][65]:

Assumption 6.1 *Assume there exist SoS polynomials $p_i(x)$ such that the set $\{x \in \mathbb{R}^n \mid p_0(x) + \sum_i p_i(x)\tilde{g}_i(x) \geq 0\}$ is compact.*

Theorem 6.3 (Putinar's Positivstellensatz) *Let $\mathcal{X}$ be the simple semi-algebraic set $\{x \in \mathbb{R}^n \mid \tilde{g}_i(x) \geq 0,\ i = 1,\ldots,n_X\}$ such that the polynomials g_i satisfy Assumption 6.1. If the polynomial $f(x)$ is positive for all $x \in \mathcal{X}$, then there exist SoS polynomials $q_i(x)$ such that $f(x) = q_0(x) + \sum_i q_i(x)\tilde{g}_i(x)$.*

Assumption 6.1 is broad enough to cover the entire MDP context; it is satisfied whenever $\mathcal{X}$ is compact (since $\mathcal{X}$ is a subset of a simplex represented by linear constraints [44][54]). This result is the foundation of Lasserre hierarchies since it shows that if the Archimedian condition holds, then we may replace polynomial constraints of the form $f(x) > 0,\ \forall x\ s.t.\ \tilde{g}(x) \geq 0$ with the more tractable conditions $f(x) = q_0(x) + q(x)^\intercal \tilde{g}(x),\ q_i(x) \in SoS$, where the degrees of q_i are uniformly less than some integer k. By a limit argument, we can then replace the strict inequality with $\geq$ by taking the limit as the maximum degree k tends towards infinity. This justifies approximating the constraint $f(x) > 0,\ \forall x\ s.t.\ \tilde{g}(x) \geq 0$ with the corresponding SoS condition of some large (but finite) degree k.

6.3 MCI Set for Controlled Markov Chain

For completeness, we re-derive the Lasserre hierarchy presented by Korda et al. [50] first in the context of general polynomial dynamics, and then specialized to the specific case of

Markov decision processes.

6.3.1 MDP Dynamics

A *Markov chain* is a discrete-time stochastic process over a finite set of states $\mathcal{S} = \{s_1, \ldots, s_n\}$, not to be confused with the state vector x in the previous section. During each time step, the state transitions stochastically as a function of the current state:

$$\Pr[s(t+1) = s_j \mid s(t) = s_i] = M_{i,j}, \ \forall t \in \mathbb{Z}_{\geq 0}. \tag{6.22}$$

Although the state trajectory of a Markov chain is nondeterministic, the distribution over the states $\xi(t) \in \mathbb{R}^n$ evolves as a linear system:

$$\xi(t+1) = M\xi(t), \ \forall t \in \mathbb{N} \tag{6.23}$$

where $\xi_i(t) = \Pr[s(t) = s_i]$, and $M_{i,j} = \Pr[s(t+1) = s_i \mid s(t) = s_j]$. Since all probabilities conditioned on $s(t) = s_j$ are nonnegative and sum to 1, M must be a column stochastic matrix:

$$\begin{aligned} \textstyle\sum_{i=1:n} M_{i,j} = 1 \quad & \forall j, \\ M_{i,j} \geq 0 \quad & \forall i, j. \end{aligned} \tag{6.24}$$

In a *Markov decision process* (MDP), an action $a(t)$ is chosen from a finite set $\mathcal{A}$ at each time step t. Transition probabilities depend on the state-action pair $(s(t), a(t)) \in \mathcal{S} \times \mathcal{A}$, and there is a reward functional which associates real number to each infinite sequence of state-action pairs. One typically wishes to find an optimal policy π associating each state with a probability distribution over actions, to maximize expected reward.

Lemma 6.1 *The MDP dynamics may be represented in matrix form as* $\xi(t+1) = P \operatorname{Diag}(\pi(t)) \, I_o^\intercal \, \xi(t)$, *where:*

- $\xi_j(t) = \Pr[s(t) = s_j]$, s_j *is the* j^{th} *state,*
- $P_{i,k} = \Pr[s(t+1) = s_i \mid p(t) = p_k]$, p_k *is the* k^{th} *state-action pair,*
- $\pi_k = \Pr[p(t) = p_k \mid s(t) = s_{j_k}]$, j_k *is the index of the state in state-action pair* p_k, *and*

- $I_{o[j,k]} = \delta_{j,j_k} = \begin{cases} 1 & j = j_k \\ 0 & j \neq j_k \end{cases}$.

Proof: This lemma follows straightforwardly from the law of total probability and matrix manipulation. Specifically, for all i:

$$\begin{aligned}
\xi_i^+ &= \Pr[s^+ = s_i] \\
&= \sum_{k,j} \big(\Pr[s^+ = s_i \mid p = p_k] \dots \\
&\qquad \cdot \Pr[p = p_k \mid s = s_j] \cdot \Pr[s = s_j] \big) \\
&= \sum_{k,j} P_{i,k} \cdot \pi_k \ \delta_{j,j_k} \cdot \xi_j \\
&= \sum_{k,j} P_{i,k} \cdot \pi_k \ I_{o[j,k]} \cdot \xi_j \\
&= \sum_{k} P_{i,k} \cdot \pi_k \cdot \left[I_o^\intercal \cdot \xi \right]_k \\
&= \left[P \cdot \mathrm{Diag}(\pi) I_o \xi \right]_i .
\end{aligned}$$

Thus $\xi^+ = P \cdot \mathrm{Diag}(\pi) I_o \xi$, as hoped. ■

For any fixed policy π, the system simplifies to a Markov chain with a particular Markov matrix denoted M^π. If we consider mixed policies, the set of achievable M^π is a polytope in $\mathbb{R}^{n \times n}$. Specifically, we may characterize this polytope as

$$M^\pi \in \{ P \ \mathrm{Diag}(\pi) I_o^\intercal \mid \pi \in \mathbb{R}_+^{mn}, \ I_o \pi = 1_n \}, \tag{6.25}$$

with transition kernel $P \in \mathbb{R}^{n \times mn}$ whose k^{th} column is the distribution over states from the k^{th} state-action pair, and indicator matrix $I_o \in \{0,1\}^{n \times mn}$ such that $I_o(i,k) = 1$ if and only if the k^{th} state-action pair includes state s_i. We can represent the MDP dynamics as a polynomial control system:

$$\begin{aligned}
&\xi_{k+1} = P \ \mathrm{diag}(\pi_k) I_o^\intercal \xi_k, \\
&\xi_k \in \{ \xi \in \mathbb{R}_+^n \mid 1^\intercal \xi = 1 \}, \\
&\pi_k \in \{ \pi \in \mathbb{R}_+^{mn} \mid I_o \pi = 1_n \}.
\end{aligned} \tag{6.26}$$

Notice that the constraints on π require the entries associated with any particular state to

be nonnegative and sum to 1, since π is a vector of conditional probabilities. We explicitly include these constraints in the representation of the semi-algebraic sets $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{U}}$ defined in the form of (6.16):

$$\begin{aligned}\mathcal{X} = \{&\xi \in \mathbb{R}^n \mid 1_n^\intercal \xi = 1,\ \xi \geq 0,\ g_i(\xi) \geq 0,\\ &\forall i = 1, \ldots, n_X\}\,,\end{aligned} \tag{6.27}$$

$$\begin{aligned}\mathcal{U} = \{&\pi \in \mathbb{R}^{mn} \mid I_o \pi = 1_n,\ u \geq 0,\ h_i(\pi) \geq 0,\\ &\forall i = 1, \ldots, n_U\}\,.\end{aligned} \tag{6.28}$$

Although the set of available actions may not be uniform across all states in practice, we assume that the full set of m actions is accessible from any state for notational convenience *w.l.o.g.*, since we may pad out the number of actions with redundant ones. This ensures the policy vector π_k is in $\mathbb{R}^{mn}$ and lets us construct I_o as the Kronecker product of the $n \times n$ identity matrix and an m-dimensional row vector of ones, i.e.,

$$I_o = I_{n \times n} \otimes 1_m^\intercal, \tag{6.29}$$

since m actions may be chosen from each of the n states.

6.3.2 Dynamics in Unit Simplex

We now present the dual problem (6.14) for a state vector is constrained to a subset of the unit simplex.

Lemma 6.2 *The exact MCI dual problem for the dynamics $\xi_{k+1} = f(\xi_k, \pi_k)$ with constraint*

sets $\mathcal{X}$ *and* $\mathcal{U}$ *respectively defined by* (6.27) *and* (6.28) *is the following:*

$$d^* = \inf_{v,w,p,q,r,s} \int_{\Delta(n)} w(\xi)\ \mathrm{d}\lambda(x) \tag{6.30}$$

$$\text{s.t.} \begin{cases} v(\xi) - \alpha v(f(\xi,\pi)) - \mathbf{q}_1(\xi,\pi)^\intercal g(\xi) \\ \qquad -\mathbf{q}_2(\xi,\pi)^\intercal \xi + q_3(\xi,\pi)(1_n^\intercal \xi - 1) \\ \qquad -\mathbf{r}_1(\xi,\pi)^\intercal h(\pi) - \mathbf{r}_2(\xi,\pi)^\intercal \pi \\ \qquad +\mathbf{r}_3^\intercal(\xi,\pi)(I_o\pi - 1_n) \geq 0, \\ w(\xi) - v(\xi) - 1 - \mathbf{p}_1(\xi)^\intercal g(\xi) - \mathbf{p}_2(\xi)^\intercal \xi \\ \qquad +p_3(\xi)(1_n^\intercal \xi - 1) \geq 0, \\ w(\xi) - \mathbf{s}_1(\xi)^\intercal g(\xi) - \mathbf{s}_2(\xi)^\intercal \xi \\ \qquad +s_3(\xi)(1_n^\intercal \xi - 1) \geq 0, \\ \mathbf{q}_1(\xi,\pi),\ \mathbf{q}_2(\xi,\pi),\ \mathbf{r}_1(\xi,\pi),\ \mathbf{r}_2(\xi,\pi) \geq 0, \\ \qquad \mathbf{p}_1(\xi),\ \mathbf{p}_2(\xi),\ \mathbf{s}_1(\xi),\ \mathbf{s}_2(\xi) \geq 0, \\ \qquad \forall (\xi,\pi) \in \mathbb{R}^n \times \mathbb{R}^m \end{cases} \tag{6.31}$$

where the multiplier functions $\mathbf{q}_{1,2}$, $\mathbf{r}_{1,2,3}$, $\mathbf{p}_{1,2}$, *and* $\mathbf{s}_{1,2}$ *are vector-valued.*

Proof: Take the problem (6.14), with basic semi-algebraic sets $\mathcal{X} = \{\xi \in \mathbb{R}^n \mid 1^\intercal \xi = 1, \xi \geq 0, g_i(\xi) \geq 0,\ \forall i\}$ and $\mathcal{U} = \{\pi \in \mathbb{R}^m \mid I_o\pi = 1, \pi \geq 0, h_i(\pi) \geq 0,\ \forall i\}$. We do not require q_3, $\mathbf{r}_3$, p_3, or s_3 to be nonnegative because they are multipliers associated with equality constraints in the primal problem.

■

The infimum is taken over continuous functions—or equivalently polynomials, from the Stone-Weierstrass theorem [24]—$v(\xi)$, $w(\xi)$, $q_i(\xi,\pi)$, $r_i(\xi,\pi)$, $p_i(\xi)$, and $s_i(\xi)$ satisfying the above constraints. Since the coefficients of any nonnegative polynomial are approximated arbitrarily well on a bounded set by those of SoS polynomials [55], all nonnegativity constraints can be replaced with stronger SoS conditions of sufficiently high degree:

$$d^* = \inf_{v,w,p,q,r,s} \int_{\Delta(n)} w(\xi)\ \mathrm{d}\lambda(\xi) \tag{6.32}$$

$$\text{s.t.} \begin{cases} v(\xi) - \alpha v(f(\xi,\pi)) - q_1(\xi,\pi)^\intercal g(\xi) - q_2(\xi,\pi)^\intercal \xi \\ \qquad +q_3(\xi,\pi)(1_n^\intercal \xi - 1) - r_1(\xi,\pi)^\intercal h(\pi) \\ \qquad -r_2(\xi,\pi)^\intercal \pi + r_3^\intercal(\xi,\pi)(I_o \pi - 1_n) = q_0(\xi,\pi), \\ w(\xi) - v(\xi) - 1 - p_1(\xi)^\intercal g(\xi) - p_2(\xi)^\intercal \xi \\ \qquad +p_3(\xi)(1_n^\intercal \xi - 1) = p_0(\xi), \\ w(\xi) - s_1(\xi)\xi^\intercal g(\xi) - s_2(\xi)^\intercal \xi \\ \qquad +s_3(\xi)(1_n^\intercal \xi - 1) = s_0(\xi), \\ q_{0,1,2}(\xi,\pi),\ r_{1,2}(\xi,\pi) \in SoS, \\ p_{0,1,2}(\xi), s_{0,1,2}(\xi) \in SoS, \qquad \forall (\xi,\pi) \in \mathbb{R}^n \times \mathbb{R}^m \end{cases} \tag{6.33}$$

6.3.3 Transformed Dynamics

The MDP defined in §6.3 has polynomial dynamics in the probability vector $\xi \in \mathbb{R}^n$ and policy $u \in \mathbb{R}^{mn}$.

$$\xi_{k+1} = P\ \mathrm{diag}(\pi_k)\ \ (I_n \otimes 1_m)\ \ \xi_k \tag{6.34}$$

The vector ξ describes a probability distribution over the finite set of MDP states, so ξ satisfies the constraints $1^\intercal \xi = 1$ and $\xi \geq 0$. Likewise, the entries of π corresponding to each individual state should also be nonnegative and sum to 1. These equality constraints imply that the domains of ξ and π are polytopes of dimension $n-1$ and $(m-1)n$, respectively. For example:

$$\begin{aligned} &\xi_k = \begin{bmatrix} x_k \\ 1 - 1^\intercal x_k \end{bmatrix}, \quad x_k \in \left\{ x \in \mathbb{R}_+^{n-1} \mid 1^\intercal x \leq 1 \right\}, \\ &\pi_k = R u_k + c, \\ &u_k \in \left\{ u \in \mathbb{R}_+^{(m-1)n} \mid \left(I_n \otimes 1_{m-1}^\intercal \right) u \leq 1_n \right\}, \end{aligned} \tag{6.35}$$

where $R = I_n \otimes \begin{bmatrix} I_{m-1} \\ -1_{m-1}^\intercal \end{bmatrix}$, and $c = 1_n \otimes \begin{bmatrix} 0_{(m-1)\times 1} \\ 1 \end{bmatrix}$. This change of variables removes the equality constraints and reduces the dimensionality of the state and control vectors. Left-multiplying Equation (6.34) by $\begin{bmatrix} I_{n-1} & 0 \end{bmatrix}$ and making the appropriate substitutions, we have an equivalent system with no equality constraints:

$$x_{k+1} = \begin{bmatrix} I_{n-1} & 0 \end{bmatrix} f\left(\begin{bmatrix} x_k \\ 1 - 1^\intercal x_k \end{bmatrix}, Ru_k + c\right) \tag{6.36}$$

$$= \begin{bmatrix} I_{n-1} & 0 \end{bmatrix} P \operatorname{diag}(Ru_k + c)(I_n \otimes 1_m) \begin{bmatrix} x_k \\ 1 - 1^\intercal x_k \end{bmatrix}. \tag{6.37}$$

Likewise, this eliminates the equality constraints in the representations of (6.27), becoming the transformed constraint sets $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{U}}$:

$$\begin{aligned}
\tilde{\mathcal{X}} &= \{x \in \mathbb{R}^{n-1} \mid \begin{bmatrix} x_k \\ 1 - 1^\intercal x_k \end{bmatrix} \in \mathcal{X}\} \\
&= \Big\{x \in \mathbb{R}^{n-1} \mid x \geq 0,\ 1 - 1^\intercal x \geq 0, \\
&\qquad g_i\left(\begin{bmatrix} x_k^\intercal & 1 - 1^\intercal x_k \end{bmatrix}^\intercal\right) \geq 0,\ \forall i\Big\}, \\
\tilde{\mathcal{U}} &= \{u \in \mathbb{R}^{mn-n} \mid Ru + c \in \mathcal{U}\} \\
&= \{u \in \mathbb{R}^{mn-n} \mid Ru + c \geq 0,\ h_i(Ru + c) \geq 0,\ \forall i\}.
\end{aligned}$$

6.3.4 Specific Form for Controlled Markov Chain

So far, we were able to over approximate the exact solution to (6.14) as the solution to a finite-dimensional SDP by further constraining the functions v and w to be polynomials of degree d such that the nonnegativity constraints are replaced with the stronger constraint of being sum-of-squares polynomials of degree $2k$. Since w is a polynomial, we may represent it as a linear combination of monomials, noting that x_i denotes the i-th component of $x \in \mathbb{R}^n$:

$$w(x) = \sum_\tau c_\tau x^\tau := \sum_{\tau \in T} c_\tau \prod_{i=1}^n x_i^{\tau_i},$$

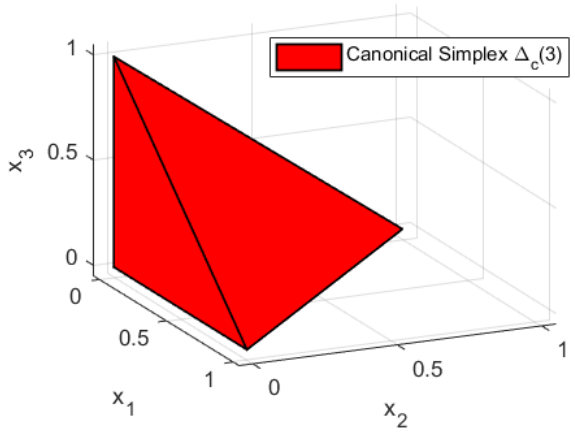

Figure 6.1: Canonical simplex in $\Delta_c(3)$. This is the projection of the unit simplex $\Delta(4)$ into $\mathbb{R}^3$.

where $T = \{\tau \in \mathbb{Z}_{\geq 0} \mid \tau_1 + \cdots + \tau_n \leq 2k\}$. By linearity, the objective function is a weighted sum of the coefficients c_τ:

$$\begin{aligned}\int_{\Delta(n)} w(x)\ \lambda(\mathrm{d}x) &= \int_{\Delta(n)} \sum_{\tau} c_\tau x^\tau\ \lambda(\mathrm{d}x) \\ &= \sum_{\alpha} c_\tau \left(\int_{\Delta(n)} x^\tau \lambda(\mathrm{d}x) \right).\end{aligned}$$

These integrals over the canonical simplex $\Delta_c(n-1) = \{x \in \mathbb{R}_+^{n-1} \mid 1^\intercal x \leq 1\}$ converge if $\tau_i > -1,\ \forall i$ (which is the case if w is a polynomial) and have the explicit form [56]

$$\begin{aligned}\int_{\Delta_c(n-1)} x^\tau\ \lambda(\mathrm{d}x) &:= \int_{\Delta_c(n-1)} \prod_{i=1}^{n} x_i^{\tau_i}\ \lambda(\mathrm{d}x) \\ &= \frac{\prod_{i=1}^n \tau_i!}{(n + \sum_{i=1}^n \tau_i)!} \\ &= B(\tau_1 + 1, \ldots, \tau_n + 1, 1)\end{aligned}$$

where B is the multivariate beta function. Note that we may scale the dual problem objective function by an arbitrary positive constant without changing the optimal solutions v and w, so it makes no difference whether the basis polynomials are integrated with respect to the Lebesgue measure or to the uniform probability measure on the simplex, since the former

simply scales the integral by the volume. Replacing the integral with an expected value w.r.t. the uniform distribution on the canonical simplex, the maximum moment is scaled to 1 rather than to the volume $\frac{1}{n!}$:

$$\mathbb{E}_X\left[X^\tau\right] = n!B(\tau_1+1,\ldots,\tau_n+1,1) = \frac{n!\prod_{i=1}^n \tau_i!}{(n+\sum_{i=1}^n \tau_i)!}.$$

We therefore find the final version of the SoS program by making the substitutions to Equation (6.32) as defined in Equations (6.35). The final version of our optimization problem is a $2k^{th}$ order SoS program in the coefficients of the polynomials $w(x)$, $v(x)$, $p(x)$, $q(x,u)$, $r(x,u)$, and $s(x)$.

$$\begin{aligned}
\min_{w,v,p,q,r,s} \quad & \sum_{\tau_1+\cdots+\tau_n\leq 2k} w_\tau \mathbb{E}_X[X^\tau] \\
s.t. \quad & v(x) - \alpha\, v(f(x,u)) - q(x,u)^\intercal \tilde{G}(x) \\
& \quad - r(x,u)^\intercal \tilde{H}(u) \in SoS, \\
& w(x) - v(x) - 1 - p(x)^\intercal g(x) \in SoS, \\
& w(x) - s(x)^\intercal g(x) \in SoS, \\
& q(x,u),\ r(x,u) \in SoS, \\
& p(x),\ s(x) \in SoS,
\end{aligned}$$

where

$$\tilde{G}(x) = \begin{bmatrix} x \\ 1-1^\intercal x \\ g\left(\begin{bmatrix} x \\ 1-1^\intercal x \end{bmatrix}\right) \end{bmatrix}, \tag{6.38}$$

$$\tilde{H}(u) = \begin{bmatrix} Ru+c \\ h(Ru+c) \end{bmatrix}. \tag{6.39}$$

6.4 Examples

To simplify the notation in the examples, we note the time-invariance of the dynamics and use $x^+ = f(x,u)$ to denote $x_{k+1} = f(x_k,u_k),\ \forall k = 0,1,2,....$

6.4.1 Example 1

In our first example, we compare the SoS approach to the exact maximal invariant set for the uncontrolled Markov chain

$$x^+ = \begin{bmatrix} 0.9 & 0 & 0.1 \\ 0.1 & 0.85 & 0 \\ 0 & 0.15 & 0.9 \end{bmatrix} x, \tag{6.40}$$

with safety constraint

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} x \leq \begin{bmatrix} -0.1 \\ -0.1 \\ -0.1 \end{bmatrix} \tag{6.41}$$

for all time steps $k \geq 0$. Since the unit simplex in $\mathbb{R}^3$ has dimension 2, we reduce the dimension with the substitution $x_3 = 1 - x_1 - x_2$:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^+ = \begin{bmatrix} 0.8 & -0.1 \\ 0.1 & 0.85 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0 \end{bmatrix},$$
$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} -0.1 \\ -0.1 \\ 0.9 \end{bmatrix}.$$

Since there is no control, the solution polynomials $v(x)$ and $w(x)$ are functions only of x (in particular, the first two components x_1 and x_2). Each non-negativity constraint $poly(x) \geq 0$

is expressed as the stronger SoS condition $A \succeq 0$, where

$$
\begin{aligned}
poly(x) &= \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \\ \vdots \\ x_2^k \end{bmatrix}^\intercal \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N_A} \\ a_{21} & a_{22} & \cdots & a_{2N_A} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_A1} & a_{N_A2} & \cdots & a_{N_AN_A} \end{bmatrix}}_{A \succeq 0} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \\ \vdots \\ x_2^k \end{bmatrix} \\
&= m(x)^\intercal \; A \; m(x).
\end{aligned}
$$

For concreteness, we chose $\alpha = 0.9$. Explicitly, the SoS problem of degree $2k$ is:

$$\min_{w,v,p,q} \sum_{\tau_1=0}^{2k} \sum_{\tau_2=0}^{2k-\tau_1} \frac{2!\tau_1!\tau_2!}{(2+\tau_1+\tau_2)!} w_{\tau_1,\tau_2} \tag{6.42}$$

$$s.t.\ v(x) - 0.9v\left(\begin{bmatrix} 0.1 + 0.8x_1 - 0.1x_2 \\ 0.1x_1 + 0.85x_2 \end{bmatrix}\right) \tag{6.43}$$

$$- q(x,u)^\intercal \begin{bmatrix} -0.1 + x_1 \\ -0.1 + x_2 \\ 0.9 - x_1 - x_2 \end{bmatrix} \in SoS, \tag{6.44}$$

$$w(x) - v(x) - 1 - p(x)^\intercal \begin{bmatrix} -0.1 + x_1 \\ -0.1 + x_2 \\ 0.9 - x_1 - x_2 \end{bmatrix} \in SoS, \tag{6.45}$$

$$w(x) \in SoS. \tag{6.46}$$

We illustrate the maximum invariant sets with polynomials of varying degrees in Figure 2.

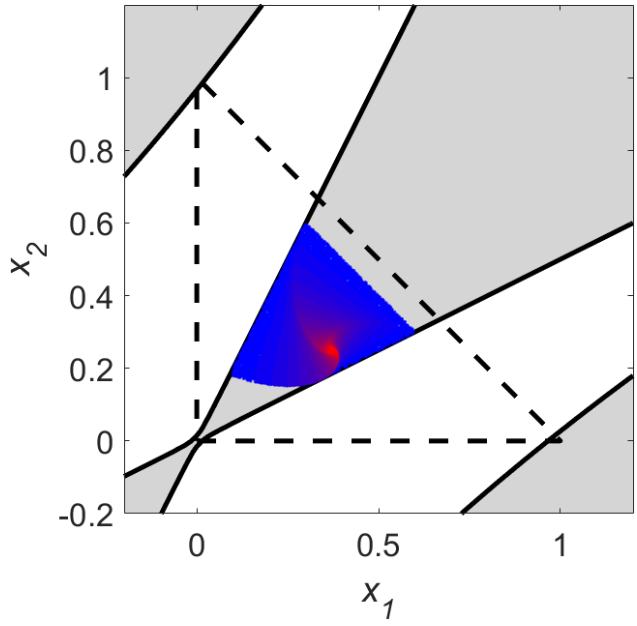

Figure 6.2: Degree 4 Polynomial MCI set approximation for Example 1. The computed set is the segment of the grey region contained within the canonical simplex (outlined in dashed lines).

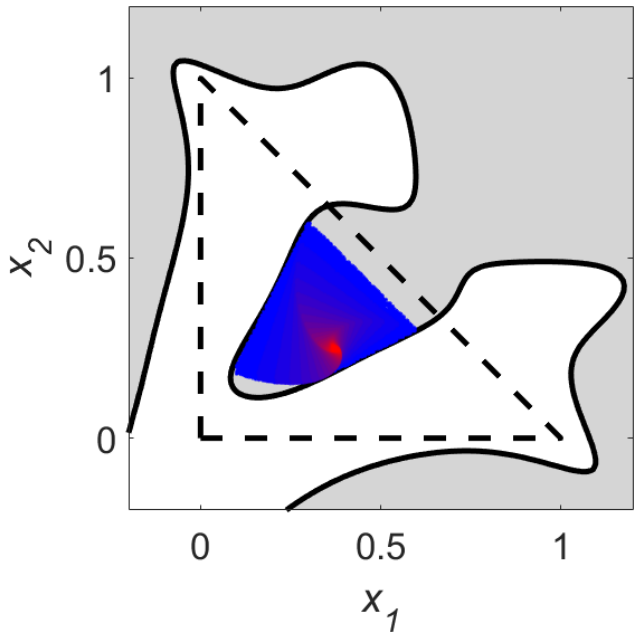

Figure 6.3: Degree 6 Polynomial MCI set approximation for Example 1.

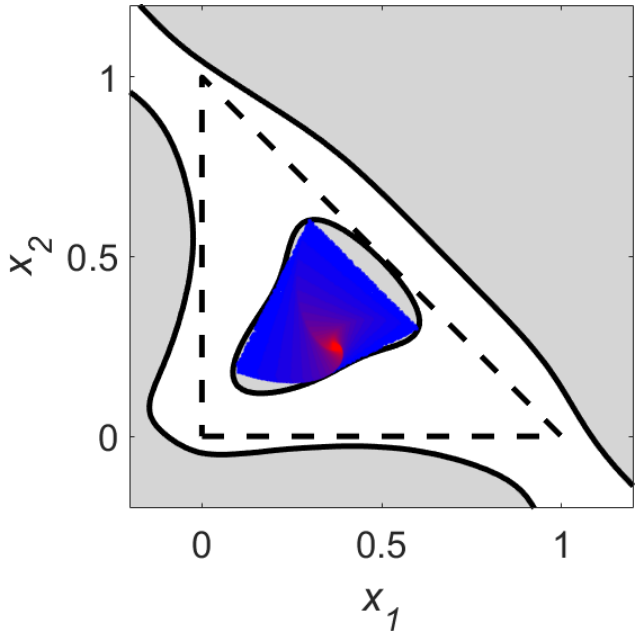

Figure 6.4: Degree 8 Polynomial MCI set approximation for Example 1.

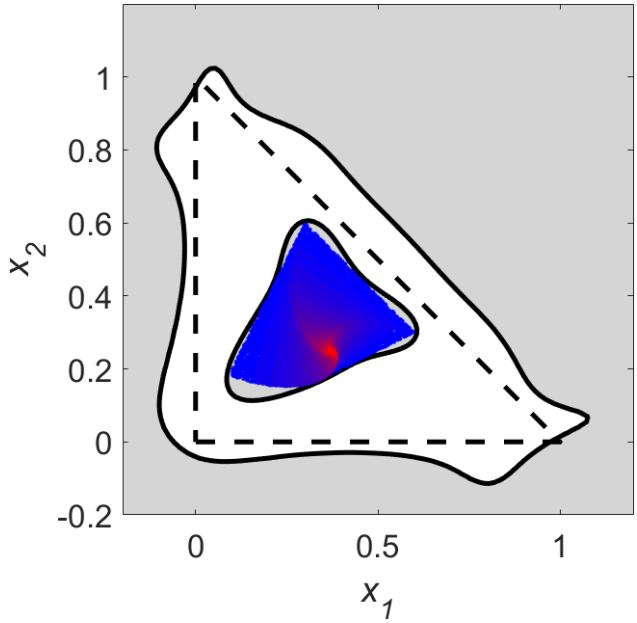

Figure 6.5: Degree 10 Polynomial MCI set approximation for Example 1.

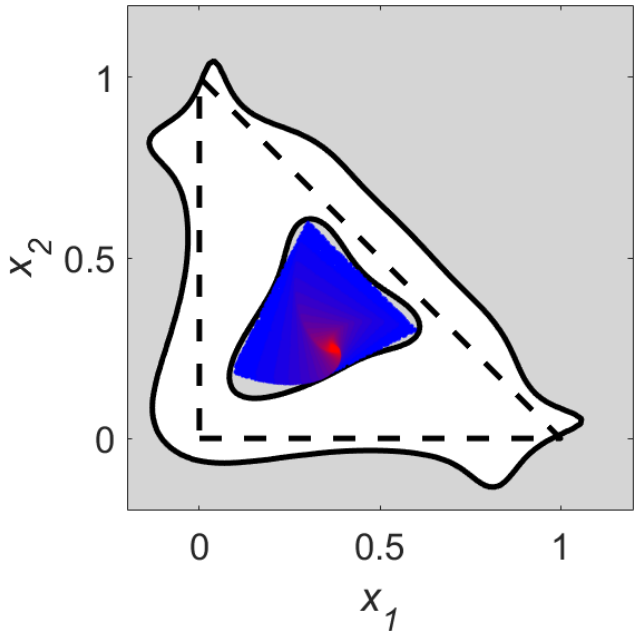

Figure 6.6: Degree 12 Polynomial MCI set approximation for Example 1.

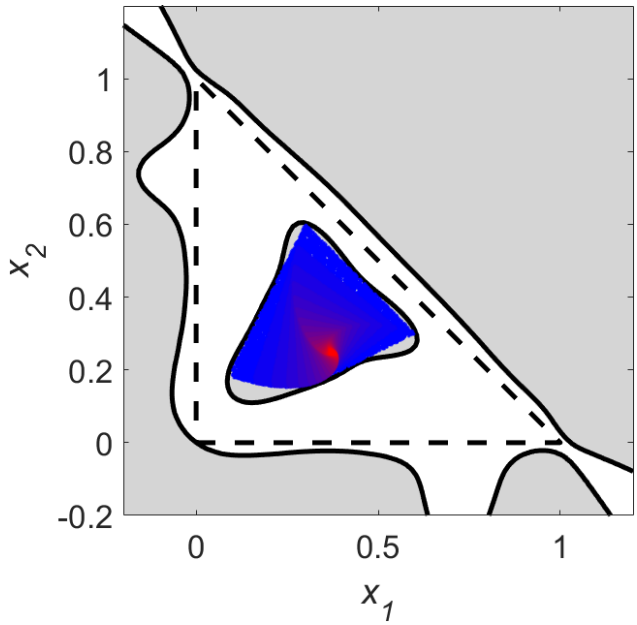

Figure 6.7: Degree 14 Polynomial MCI set approximation for Example 1.

6.4.2 Example 2

In the second example, we introduce control to the system from Example 1. Each state now has two actions: move to the next consecutive state, and remain in the current state. The dynamics have the form

$$
\begin{aligned}
x^+ &= P\ \mathrm{diag}(u)\ I_o\ x \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \ddots & & \\ & u & \\ & & \ddots \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} x \\
&= \begin{bmatrix} u_1 & 0 & u_6 \\ u_2 & u_3 & 0 \\ 0 & u_4 & u_5 \end{bmatrix} x = \begin{bmatrix} u_1 x_1 + u_6 x_3 \\ u_2 x_1 + u_3 x_2 \\ u_4 x_2 + u_5 x_3 \end{bmatrix}
\end{aligned}
$$

and constraints

$$
\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & -2 & 0 \\ -2 & 1 & 0 \end{bmatrix} x \leq \begin{bmatrix} -0.1 \\ -0.1 \\ -0.1 \\ 0 \\ 0 \end{bmatrix}, \quad 1^\intercal x = 1,\ x \geq 0
$$

$$
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} u \leq \begin{bmatrix} 0.9 \\ 0.85 \\ 0.9 \end{bmatrix}, \quad I_o u = 1,\ u \geq 0.
$$

Notice that these control constraints are a relaxation of Example 1. Therefore we would expect the MCI set to contain the one in the previous example. As expected, the semi-algebraic outer approximation also contains the previously computed maximal invariant set (in blue and red).

To see that this outer approximation represents a real increase in size of the MCI set, notice that the vertex $x_0 = \begin{bmatrix} 0.1 & 0.1 & 0.8 \end{bmatrix}^\intercal$ is sent to itself with the feasible control vector

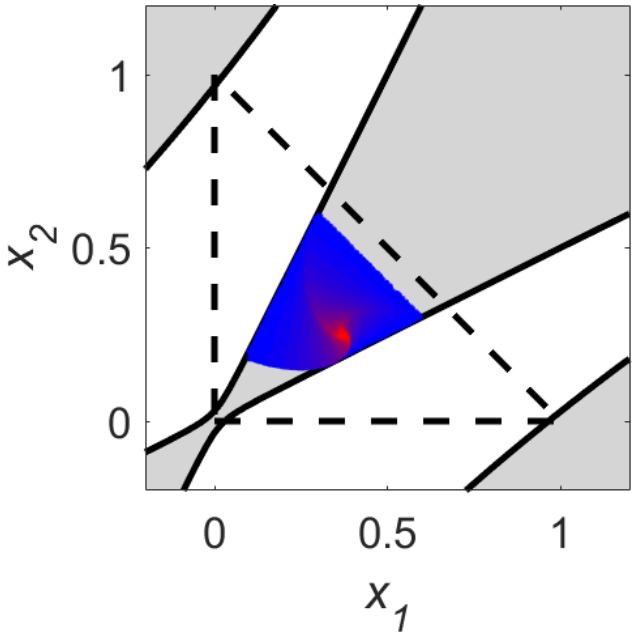

Figure 6.8: Degree 4 Polynomial MCI set approximations for Example 2.

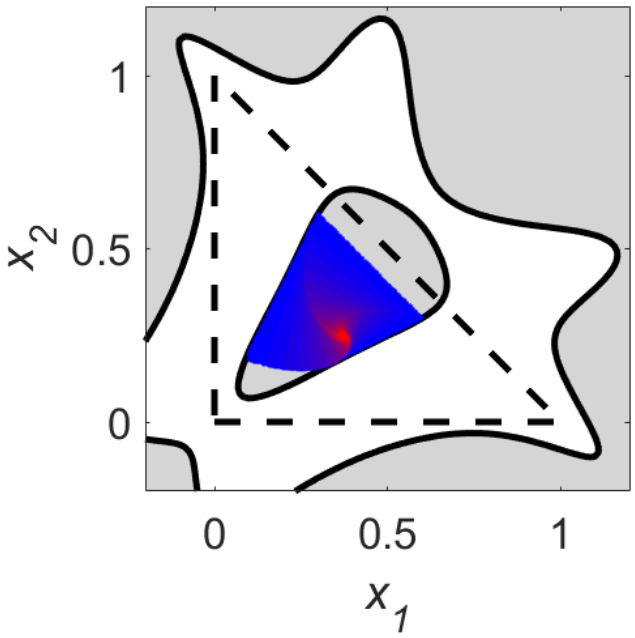

Figure 6.9: Degree 6 Polynomial MCI set approximations for Example 2.

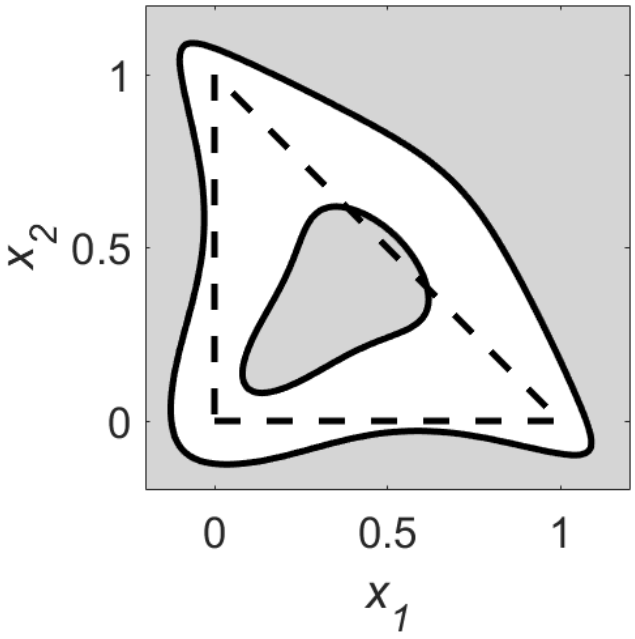

Figure 6.10: Degree 8 Polynomial MCI set approximations for Example 2.

$u_0 = \begin{bmatrix} 0.875 & 0.125 & 0 & 1 & 0 & 1 \end{bmatrix}^\intercal$, however x_0 violates the safety constraint from Example 1 since

$$f(x_0, u_0) = \begin{bmatrix} 0.9 & 0 & 0.1 \\ 0.1 & 0.85 & 0 \\ 0 & 0.15 & 0.9 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.1 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 0.17 \\ 0.095 \\ 0.735 \end{bmatrix},$$

which is outside the polyhedron of Equation (6.41).

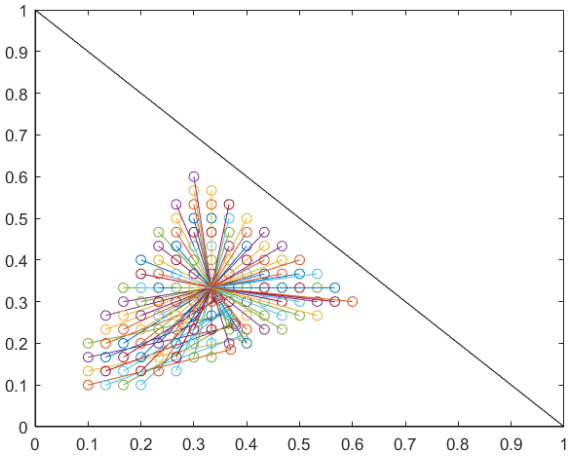

Figure 6.11: Example feasible trajectories of points within the true MCI set.

6.5 Conclusion

We presented a brief summary of the measure formulation of control system to provide background for the infinite-dimensional linear program to compute the MCI set under state and control constraints. We then applied the Lasserre hierarchy adapted from Korda et al. [50] to reduce this to a sequence of finite-dimensional SoS programs for specifically an MDP subject to basic semi-algebraic safety constraints in both the full-dimension"unit simplex" and reduced-dimension "canonical simplex" formulations. We then gave numerical examples showing the effectiveness of this method for various polynomial degrees.

Chapter 7

CONCLUDING REMARKS

7.1 Future Research Directions

Though the primary topic for this dissertation, the study of safety constraints is far from my only topic of interest. The techniques presented here have many potential applications apart from the specific scenarios we considered, and I hope that they will prove to be fruitful for future research by myself, my collaborators, and the broader controls community.

7.1.1 Velocity Fields for Swarm Guidance

The main insight is to treat a swarm as a continuous fluid instead of as a discrete set of agents. Every agent has a position and velocity, so we can think of each agent as a particle in a fluid. For very large swarms, this approach may be useful because we can think of a velocity field carrying the collection of agents rather than looking at each agent's individual state.

$$\dot{x}(t) = f(t, x(t)) \tag{7.1}$$

If $f(t,x)$ is continuous, then $\lim_{dx\to 0} f(t, x+dx) - f(t,x) = 0$, and so the relative velocity of two nearby agents will approach zero as the relative displacements approach zero. By requiring f to be continuous, we can ensure that all collisions will occur with zero relative velocity to minimize the damage done by each collision. However, this is only a heuristic whenever the vehicles have nonzero volume, and two agents can still have arbitrarily large relative velocity when they are within some nonzero distance R. Continuity alone is not sufficient to mitigate the danger of collisions. To quantify the relationship between distance and relative velocity, f needs to satisfy stronger continuity properties such as smoothness.

Assumption 7.1 *Assume that $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable, and that $\frac{\partial}{\partial x} f(t,x)$ is uniformly bounded for all t and x.*

Let $x_1(t)$ and $x_2(t)$ be solutions of Equation (7.1). Then the rate of change of the squared

P-distance between these two trajectories is

$$\frac{d}{dt}\|x_2(t) - x_1(t)\|_P^2 = \frac{d}{dt}\left[x_2(t) - x_1(t)\right]^\intercal P \left[x_2(t) - x_1(t)\right] \tag{7.2}$$

$$= 2\left[x_2(t) - x_1(t)\right]^\intercal P \left[f(t, x_2(t)) - f(t, x_1(t))\right]. \tag{7.3}$$

If we assume that $\frac{\partial}{\partial x} f(t, x_1(t)), \frac{\partial}{\partial x} f(t, x_2(t)) \in \Omega$, where Ω is a closed and convex subset of $\mathbb{R}^{n\times n}$, then a generalization of the Mean Value Theorem states that

$$\exists J \in \Omega \mid f(t, z_2) - f(t, z_1) = J \cdot (z_2 - z_1). \tag{7.4}$$

This lets us continue from Equation (7.3).

$$\frac{d}{dt}\|x_2(t) - x_1(t)\|_P^2 = 2\left[x_2(t) - x_1(t)\right]^\intercal P \left[f(t, x_2(t)) - f(t, x_1(t))\right], \tag{7.5}$$

$$= \left[x_2(t) - x_1(t)\right]^\intercal \left(PJ + J^\intercal P\right) \left[x_2(t) - x_1(t)\right], \tag{7.6}$$

for some $J \in \Omega$. If $\gamma_1 I \preceq PJ + J^\intercal P \preceq \gamma_2 I$ for all $J \in \Omega$, then the right-hand side of Equation (7.6) is in the interval $[\gamma_1\|x_2(t) - x_1(t)\|^2, \gamma_2\|x_2(t) - x_1(t)\|^2]$. It is then possible to bound the rate of change of $\|x_2(t) - x_1(t)\|^2$ as

$$\gamma_1\|x_2(t) - x_1(t)\|^2 \leq \frac{d}{dt}\|x_2(t) - x_1(t)\|_P^2 \leq \gamma_2\|x_2(t) - x_1(t)\|^2. \tag{7.7}$$

The change in separation distance between two agents with trajectories x_1 and x_2 can therefore be controlled by the parameter γ_1, which is a property of the gradient of f.

Consider the squared Euclidean distance $\|x_2(t) - x_1(t)\|_P,\ P = I$. The inequality (7.7) then can be converted into a bound on the distance itself using Grönwall's inequality.

$$e^{\gamma_1 t}\|x_2(0) - x_1(0)\|^2 \leq \|x_2(t) - x_1(t)\|^2 \leq e^{\gamma_2 t}\|x_2(0) - x_1(0)\|^2, \tag{7.8}$$

$$e^{(\gamma_1/2)t}\|x_2(0) - x_1(0)\| \leq \|x_2(t) - x_1(t)\| \leq e^{(\gamma_2/2)t}\|x_2(0) - x_1(0)\|. \tag{7.9}$$

7.1.2 Distributed Swarm Control

Another approach to efficiently control a large swarm is to utilize a decentralized control algorithm. Previously in this report, the curse of dimensionality was addressed by treating a swarm as a continuum so that swarms of any size can be modeled as a fluid. Rather than

a central decision-maker broadcasting a policy to all agents, each vehicle may be given the ability to communicate with and coordinate its policy with nearby agents.

I plan to investigate consensus over a swarm following MDP dynamics. Each agent will behave according to some initial policy. If two agents are in close proximity, then they will average their own policy with that of their neighbors. This forms a dynamic network represented by a time-varying graph whose vertices are the agents, and whose edges are pairs of agents in direct communication. The connectivity of this network depends on the states of all the agents.

For example, each agent may be given an initial mixed policy so that the entire swarm evolves nondeterministically. The entire swarm initially follows a mixture of the distributions of the agents. How this macroscopic distribution evolves according to various consensus algorithms is worth investigating. In one case I plan to consider, these policies may be parameterized by their stationary distribution vector so that each policy solves a constrained MDP problem. The consensus algorithm could be performed over these vectors, or by some other parameterization of the policies.

Some questions which remain to be answered are:

1. How does the distribution of the swarm evolve over time?
2. Under what conditions are convergence guaranteed, and what is the rate of convergence?
3. What is the performance of such an architecture, and how robust is it?

7.1.3 Verifiably Safe Control Synthesis

Our techniques to certify safety constraints as presented in Section 3.5 guarantee finite determination of the maximal positively invariant safe set, and under Assumptions 3.3 and 3.4 in §3.6.3, there is an easily computable upper bound on algorithmic complexity of the Kleene iteration based algorithm. We may want to design controllers for our systems so that this algorithm terminates, and $\mathcal{O}_\infty(\mathcal{S}, f)$ can be easily found for a large set of constraint sets which do not necessarily satisfy these assumptions. Further research may reveal new insights into which systems tend to have easily-verified safety constraints. A related challenge is to maximize the size of $\mathcal{O}_\infty(\mathcal{S}, f)$, for MDPs, since the maximal invariant set is a nonlinear function of the policy.

7.2 Conclusion

Safety constraints play a vital role in the control or real-world systems, and they must be analyzed for a variety of models, from a state-space model of a single vehicle, to Markov decision processes in a complex stochastic environment, to measure-based models of a swarm of autonomous agents. This dissertation touches on a variety of topics including positively invariant set computation, Markov decision processes, and swarm guidance, all with the overarching theme of the satisfaction of safety constraints. However, this is only a small sample of the kinds of processes for which safety must be considered. The increasing prevalence of novel approaches to modeling comes with the need for theoretical and practical guarantees that they will behave in a safe and desirable manner. For this reason, I am confident that the consideration of safety constraints will continue to be a fruitful topic of research for the foreseeable future.

BIBLIOGRAPHY

[1] Behçet Açıkmeşe, Nazlı Demir, and Matthew W Harris. Convex necessary and sufficient conditions for density safety constraints in markov chain synthesis. *IEEE Transactions on Automatic Control*, 60(10):2813–2818, 2015.

[2] Andrei Agrachev and Paul Lee. Optimal transportation under nonholonomic constraints. *Transactions of the American Mathematical Society*, 361(11):6019–6047, 2009.

[3] Giacomo Albi, Young-Pil Choi, Massimo Fornasier, and Dante Kalise. Mean field control hierarchy. *Applied Mathematics & Optimization*, 76:93–135, 2017.

[4] Jason Altschuler, Francis Bach, Alessandro Rudi, and Jonathan Weed. Approximating the quadratic transportation metric in near-linear time. *arXiv preprint arXiv:1810.10046*, 2018.

[5] Jason Altschuler, Jonathan Niles-Weed, and Philippe Rigollet. Near-linear time approximation algorithms for optimal transport via sinkhorn iteration. *Advances in neural information processing systems*, 30, 2017.

[6] Aristotle Arapostathis, Ratnesh Kumar, and Sekhar Tangirala. Controlled markov chains with safety upper bound. *IEEE Transactions on Automatic Control*, 48(7):1230–1234, 2003.

[7] Efstathios Bakolas. Optimal covariance control for discrete-time stochastic linear systems subject to constraints. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 1153–1158. IEEE, 2016.

[8] Efstathios Bakolas. Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints. *Automatica*, 91:61–68, 2018.

[9] Andrei Baranga. The contraction principle as a particular case of kleene's fixed point theorem. *Discrete Mathematics*, 98(1):75–79, 1991.

[10] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[11] Georges Bitsoris. Positively invariant polyhedral sets of discrete-time linear systems. *International Journal of Control*, 47(6):1713–1726, 1988.

[12] Franco Blanchini. Survey paper: Set invariance in control. *Automatica (Journal of IFAC)*, 35(11):1747–1767, 1999.

[13] Franco Blanchini and Stefano Miani. *Set-theoretic methods in control*. Birkhäuser, 2008.

[14] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

[15] Joseph T Chang and David Pollard. Conditioning as disintegration. *Statistica Neerlandica*, 51(3):287–317, 1997.

[16] Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Optimal steering of a linear stochastic system to a final probability distribution, part i. *IEEE Transactions on Automatic Control*, 61(5):1158–1169, 2015.

[17] Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Optimal steering of a linear stochastic system to a final probability distribution, part ii. *IEEE Transactions on Automatic Control*, 61(5):1170–1180, 2015.

[18] Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. Optimal transport over a linear dynamical system. *IEEE Transactions on Automatic Control*, 62(5):2137–2152, 2016.

[19] Donald L Cohn. *Measure theory*, volume 5. Springer, 2013.

[20] P. Cousot and R. Cousot. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Conference Record of the Fourth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 238–252, Los Angeles, California, 1977. ACM Press, New York, NY.

[21] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

[22] Mathias Hudoba De Badyn, Utku Eren, Behçet Açikmeşe, and Mehran Mesbahi. Optimal mass transport and kernel density estimation for state-dependent networked dynamic systems. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 1225–1230. IEEE, 2018.

[23] Mathias Hudoba De Badyn, Erik Miehling, Dylan Janak, Behçet Açıkmeşe, Mehran Mesbahi, Tamer Başar, John Lygeros, and Roy S Smith. Discrete-time linear-quadratic regulation via optimal transport. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 3060–3065. IEEE, 2021.

[24] Louis De Branges. The stone-weierstrass theorem. *Proceedings of the American Mathematical Society*, 10(5):822–824, 1959.

[25] Nazlı Demir, Utku Eren, and Behçet Açıkmeşe. Decentralized probabilistic density control of autonomous swarms with safety constraints. *Autonomous Robots*, 39:537–554, 2015.

[26] Mahmoud El Chamie and Behçet Açıkmeşe. Convex synthesis of optimal policies for markov decision processes with sequentially-observed transitions. In *American Control Conference (ACC), 2016*, pages 3862–3867. IEEE, 2016.

[27] Mahmoud El Chamie, Dylan Janak, and Behçet Açıkmeşe. Markov decision processes with sequential sensor measurements. *Automatica*, 103:450–460, 2019.

[28] Mahmoud El Chamie, Yue Yu, Behçet Açıkmeşe, and Masahiro Ono. Controlled markov processes with safety state constraints. *IEEE Transactions on Automatic Control*, 64(3):1003–1018, 2018.

[29] Karthik Elamvazhuthi, Piyush Grover, and Spring Berman. Optimal transport over deterministic discrete-time nonlinear systems using stochastic feedback laws. *IEEE control systems letters*, 3(1):168–173, 2018.

[30] Utku Eren and Behçet Açıkmeşe. Velocity field generation for density control of swarms using heat equation and smoothing kernels. *IFAC-PapersOnLine*, 50(1):9405–9411, 2017.

[31] Christian Feller. *Relaxed barrier function based model predictive control*. Logos Verlag Berlin GmbH, 2017.

[32] Robert M Freund and James B Orlin. On the complexity of four polyhedral set containment problems. *Mathematical programming*, 33(2):139–145, 1985.

[33] Charlie Frogner, Chiyuan Zhang, Hossein Mobahi, Mauricio Araya, and Tomaso A Poggio. Learning with a wasserstein loss. *Advances in neural information processing systems*, 28, 2015.

[34] Peyré Gabriel. The numerical tours of signal processing. *IEEE Comput. Sci. Eng.*, 13:94–97, 2011.

[35] Alfred Galichon. *Optimal transport methods in economics*. Princeton University Press, 2018.

[36] Elmer G Gilbert and K Tin Tan. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transactions on Automatic control*, 36(9):1008–1020, 1991.

[37] Maxim Goldshtein and Panagiotis Tsiotras. Finite-horizon covariance control of linear time-varying systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 3606–3611. IEEE, 2017.

[38] Michael Grant and Stephen Boyd. Cvx: Matlab software for disciplined convex programming, version 2.1, 2014.

[39] J. C. Hennet and C. E. T. Dórea. Invariant regulators for linear systems under combined input and state constraints. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, volume 2, pages 1030–1035 vol.2, Dec 1994.

[40] Jean-Claude Hennet. Discrete time constrained linear systems. *Control and Dynamic Systems*, Vol.71:157–213, 1995.

[41] Ahed Hindawi, J-B Pomet, and Ludovic Rifford. Mass transportation with lq cost functions. *Acta applicandae mathematicae*, 113:215–229, 2011.

[42] Kenji Hirata and Yoshito Ohta. Exact determinations of the maximal output admissible set for a class of nonlinear systems. *Automatica*, 44(2):526–533, 2008.

[43] Mathias Hudoba de Badyn. Supplementary software for" discrete-time linear quadratic regulation via optimal transport", 2021.

[44] Thomas Jacobi and Alexander Prestel. Distinguished representations of strictly positive polynomials. *Journal für die reine und angewandte Mathematik*, 2001(532):223–235, 2001.

[45] Dylan Janak and Behçet Açikrneşe. Maximal invariant set computation and design for markov chains. In *2019 American Control Conference (ACC)*, pages 1244–1249. IEEE, 2019.

[46] Dylan Janak, Pierre-Loïc Garoche, and Behçet Açıkmeşe. Exact computation of maximal invariant sets for safe markov chains—lattice theoretic approach. *IEEE Transactions on Automatic Control*, 67(12):6980–6986, 2022.

[47] Stephen Cole Kleene, NG De Bruijn, J de Groot, and Adriaan Cornelis Zaanen. *Introduction to metamathematics*, volume 483. van Nostrand New York, 1952.

[48] Martin Knott and Cyril S Smith. On the optimal mapping of distributions. *Journal of Optimization Theory and Applications*, 43:39–49, 1984.

[49] Ilya Kolmanovsky and Elmer G Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical problems in engineering*, 4(4):317–367, 1998.

[50] 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.

[51] Vishaal Krishnan and Sonia Martínez. Distributed online optimization for multi-agent optimal transport. *arXiv preprint arXiv:1804.01572*, 2018.

[52] Vishaal Krishnan and Sonia Martínez. Distributed optimal transport for the deployment of swarms. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 4583–4588. IEEE, 2018.

[53] Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations research & management science in the age of analytics*, pages 130–166. Informs, 2019.

[54] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.

[55] Jean B Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM review*, 49(4):651–669, 2007.

[56] Jean B Lasserre. Simple formula for integration of polynomials on a simplex. *BIT Numerical Mathematics*, 61(2):523–533, 2021.

[57] Jean B Lasserre and Eduardo S Zeron. Solving a class of multivariate integration problems via laplace techniques. *Applicationes Mathematicae*, 28:391–405, 2001.

[58] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2009.

[59] Victor Magron, Pierre-Loïc Garoche, Didier Henrion, and Xavier Thirioux. Semidefinite approximations of reachable sets for discrete-time polynomial systems. *SIAM Journal on Control and Optimization*, 57(4):2799–2820, 2019.

[60] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, USA, 2000.

[61] Peyman Mohajerin Esfahani and Daniel Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1):115–166, 2018.

[62] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences*, 7(1):212–238, 2014.

[63] Gabriel Peyré. Entropic approximation of wasserstein gradient flows. *SIAM Journal on Imaging Sciences*, 8(4):2323–2351, 2015.

[64] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[65] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[66] Julien Rabin and Gabriel Peyré. Wasserstein regularization of imaging problem. In *2011 18th IEEE international conference on image processing*, pages 1541–1544. IEEE, 2011.

[67] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot. Wasserstein barycenter and its application to texture mixing. In *Scale Space and Variational Methods in Computer Vision: Third International Conference, SSVM 2011, Ein-Gedi, Israel, May 29–June 2, 2011, Revised Selected Papers 3*, pages 435–446. Springer, 2012.

[68] Saša V. Raković and Mirko Fiacchini. Invariant approximations of the maximal invariant set or "encircling the square". *IFAC Proceedings Volumes*, 41(2):6377 – 6382, 2008. 17th IFAC World Congress.

[69] Saša V. Raković and Mario E. Villanueva. The maximal positively invariant set: Polynomial setting, 2017.

[70] Antoine Rolet, Marco Cuturi, and Gabriel Peyré. Fast dictionary learning with a smoothed wasserstein loss. In *Artificial intelligence and statistics*, pages 630–638. PMLR, 2016.

[71] Soroosh Shafieezadeh Abadeh, Viet Anh Nguyen, Daniel Kuhn, and Peyman M Mohajerin Esfahani. Wasserstein distributionally robust kalman filtering. *Advances in Neural Information Processing Systems*, 31, 2018.

[72] Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.*, 5(2):285–309, 1955.

[73] Cedric Villani. Topics in optimal transportation.(books). *OR/MS Today*, 30(3):66–67, 2003.

[74] Cédric Villani et al. *Optimal transport: old and new*, volume 338. Springer, 2009.