DATA-DRIVEN OPTIMAL CONTROL VIA LINEAR TRANSFER OPERATORS: A CONVEX APPROACH

A PREPRINT

Joseph Moyalan
Department of Mechanical Engineering
Clemson University
Clemson, SC, USA
jmoyala@clemson.edu

Hyungjin Choi
Sandia National Laboratories
Albuquerque, NM, USA
hchoi@sandia.gov

Yongxin Chen
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA, USA
yongchen@gatech.edu

Umesh Vaidya *
Department of Mechanical Engineering
Clemson University
Clemson, SC, USA
uvaidya@clemson.edu

February 7, 2022

ABSTRACT

This paper is concerned with data-driven optimal control of nonlinear systems. We present a convex formulation to the optimal control problem with a discounted cost function. We consider OCP with both positive and negative discount factor. The convex approach relies on lifting nonlinear system dynamics in the space of densities using the linear Perron-Frobenius (P-F) operator. This lifting leads to an infinite-dimensional convex optimization formulation of the optimal control problem. The data-driven approximation of the optimization problem relies on the approximation of the Koopman operator using the polynomial basis function. We write the approximate finite-dimensional optimization problem as a polynomial optimization which is then solved efficiently using a sum-of-squares-based optimization framework. Simulation results are presented to demonstrate the efficacy of the developed data-driven optimal control framework.

Keywords Optimal Control · Nonlinear Systems · Linear Transfer Operators

1   Introduction

Data-driven optimal control of a nonlinear system is a problem that has significant interest with applications ranging from vehicle autonomy, robotics to manufacturing and power systems. The traditional approach to optimal control problem (OCP) relies on solving the Hamilton Jacobi Bellman (HJB) equation [Fleming and Rishel, 2012]. The HJB equation is a nonlinear partial differential equation and challenging to solve. Existing algorithms for solving HJB equations rely on iterative scheme [Beard et al., 1997, Bertsekas, 2011]. This iterative scheme is also at the heart of the variety of reinforcement learning (RL) algorithms for the data-driven optimal control [Sutton and Barto, 2018]. In this paper, we present an alternate approach based on the dual formulation of the OCP. This dual approach leads to a convex optimization formulation of the OCP, which can be solved using a single-shot algorithm. This is in contrast to the iterative scheme used for solving the HJB equation. Furthermore, the iterative algorithm required for solving the HJB equation requires an initial stabilizing controller. Finding stabilizing controller for a nonlinear system is, in general, a nontrivial problem. However, the computational framework for solving the OCP problem in the dual space does not require an initial stabilizing controller.

The dual formulation to the OCP we present is based on the theory of linear operators, namely the P-F and Koopman operators [Lasota and Mackey, 1994] and is developed in [Huang and Vaidya, 2020]. However, there are differences between the results presented in this paper and [Huang and Vaidya, 2020] as discussed in our contributions. The convex formulation to the OCP in the dual space of densities and occupation measure has been extensively studied in [Henrion and Korda, 2013, Korda et al., 2017, Lasserre et al., 2008]. The computational framework in these works relies on moment-based relaxation of the infinite-dimensional optimization problem. In contrast, our proposed computational framework uses data and depends on the linear operator theory for the finite-dimensional approximation of the infinite-dimensional convex optimization problem. The convex formulation for optimal control is also extended to study stochastic OCP, control design with safety constraints, and data-driven stabilization problems [Yu et al., 2021, Choi et al., 2020]. There is a growing body of literature on the use of the Koopman operator for data-driven control, where the control dynamical system is lifted in the space of functions or observables using the Koopman operator [Kaiser et al., 2021, Huang et al., 2018, Arbabi et al., 2018, Ma et al., 2019, Korda and Mezić, 2020, Mauroy and Mezić, 2013, Huang et al., 2020]. However, in this paper, we lift the control system using the P-F operator, which is dual to the Koopman operator. Unlike Koopman-based lifting, P-F lifting of the control dynamical system leads to a convex formulation of the OCP [Raghunathan and Vaidya, 2013, Vaidya et al., 2010].

The main contributions of this paper are stated as follows. First, we provide a convex formulation to the infinite horizon OCP with discounted cost involving continuous-time dynamics. We consider OCP problems with both positive and negative discount. For the continuous-time OCP, the negative (positive) discount corresponds to the case where the cost function is exponentially decreasing (increasing) with time. There is extensive literature on the OCP with negative discount factor [Modares and Lewis, 2014, Modares et al., 2016, Ghosh et al., 1993]. One of the main contributions of this paper is to provide condition for the existence of optimal control problem with a positive discount. The condition arises in the form of a stronger notion of almost everywhere exponential stability [Vaidya and Mehta, 2008]. Unlike [Huang and Vaidya, 2020], the computation framework relies on the use of polynomial basis for the approximation of linear Koopman operator using generator Extended Dynamic Mode Decomposition (gEDMD) algorithm [Klus et al., 2020]. Hence, we employ sum-of-square (SOS) optimization methods for solving a finite-dimensional optimization problem. The finite-dimensional approximation of the infinite-dimensional optimization problem is written as a semi-definite program (SDP). SOS-based optimization toolbox is then used to solve the SDP in a numerically efficient manner. Existing rigorous results for the convergence analysis of Koopman operator in the limit of data and number of basis functions goes to infinity are leveraged to provide convergence analysis of data-driven optimization problem. Simulation results are presented to verify the efficacy of the developed framework. The results presented in this paper are an extension of our conference paper [Moyalan et al., 2021]. In particular, the data-driven framework for optimal control design and theorems involving OCP with discounted cost function are new to this paper.

The rest of the paper is structured as follows. In Section 2, we introduce preliminaries and notations used throughout the paper. The main results of the paper on the convex formulation to OCP are presented in Section 3. The SOS and Koopman-based computation framework for the data-driven approximation of the convex optimization problem is discussed in Section 4. Simulation results are presented in Section 5. Conclusions are presented in Section 6.

2 Preliminaries and Notations

Notation: $\mathbb{R}^n$ denotes the n dimensional Euclidean space and $\mathbb{R}^n_{\geq 0}$ is the positive orthant. Given $\mathbf{X} \subseteq \mathbb{R}^n$ and $\mathbf{Y} \subseteq \mathbb{R}^m$, let $\mathcal{L}_1(\mathbf{X}, \mathbf{Y}), \mathcal{L}_\infty(\mathbf{X}, \mathbf{Y})$, and $\mathcal{C}^k(\mathbf{X}, \mathbf{Y})$ denote the space of all real valued integrable functions, essentially bounded functions, and space of k times continuously differentiable functions mapping from $\mathbf{X}$ to $\mathbf{Y}$ respectively. If the space $\mathbf{Y}$ is not specified then it is understood that the underlying space is $\mathbb{R}$. $\mathcal{B}(\mathbf{X})$ denotes the Borel σ-algebra on $\mathbf{X}$ and $\mathcal{M}(\mathbf{X})$ is the vector space of real-valued measure on $\mathcal{B}(\mathbf{X})$. $\mathbf{s}_t(\mathbf{x})$ denotes the solution of dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ starting from initial condition $\mathbf{x}$. We will use the notation $\mathcal{N}_\delta$ to denote the δ neighborhood of the equilibrium point at the origin for some fixed $\delta > 0$ and $\mathbf{X}_1 := \mathbf{X} \setminus \mathcal{N}_\delta$.

2.1 Koopman and Perron-Frobenius Operators

Consider a dynamical system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathbf{X} \subseteq \mathbb{R}^n \tag{1}$$

where the vector field is assumed to be $\mathbf{f}(\mathbf{x}) \in \mathcal{C}^1(\mathbf{X}, \mathbb{R}^n)$. There are two different ways of linearly lifting the finite dimensional nonlinear dynamics from state space to infinite dimension space of functions, Koopman and Perron-Frobenius operators.

Koopman Operator: $\mathbb{K}_t : \mathcal{L}_\infty(\mathbf{X}) \to \mathcal{L}_\infty(\mathbf{X})$ for dynamical system (1) is defined as

$$[\mathbb{K}_t \varphi](\mathbf{x}) = \varphi(\mathbf{s}_t(\mathbf{x})), \quad \varphi \in \mathcal{L}_\infty.$$

The infinitesimal generator for the Koopman operator is

$$\lim_{t \to 0} \frac{\mathbb{K}_t \varphi - \varphi}{t} = \mathbf{f}(\mathbf{x}) \cdot \nabla \varphi(\mathbf{x}) =: \mathcal{K}_{\mathbf{f}} \varphi. \tag{2}$$

Perron-Frobenius Operator: $\mathbb{P}_t : \mathcal{L}_1(\mathbf{X}) \to \mathcal{L}_1(\mathbf{X})$ for system (1) is defined as

$$[\mathbb{P}_t \psi](\mathbf{x}) = \psi(\mathbf{s}_{-t}(\mathbf{x})) \left| \frac{\partial \mathbf{s}_{-t}(\mathbf{x})}{\partial \mathbf{x}} \right|, \quad \psi \in \mathcal{L}_1, \tag{3}$$

where $|\cdot|$ stands for the determinant. The infinitesimal generator for the P-F operator is given by

$$\lim_{t \to 0} \frac{\mathbb{P}_t \psi - \psi}{t} = -\nabla \cdot (\mathbf{f}(\mathbf{x}) \psi(\mathbf{x})) =: \mathcal{P}_{\mathbf{f}} \psi. \tag{4}$$

These two operators are dual to each other where the duality is expressed as

$$\int_{\mathbb{R}^n} [\mathbb{K}_t \varphi](\mathbf{x}) \psi(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} [\mathbb{P}_t \psi](\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x}.$$

2.2 Sum of squares

Sum of squares (SOS) optimization [Topcu et al., 2010, Parrilo, 2003, Parrilo and Sturmfels, 2003, Parrilo, 2000] is a relaxation of positive polynomial constraints appearing in polynomial optimization problems. SOS polynomials are in a set of polynomials that can be described as a finite linear combination of monomials, i.e., $p = \sum_{i=1}^{\ell} d_i p_i^2$ where p is a SOS polynomial, p_i are polynomials, and d_i are nonnegative coefficients. Hence, SOS is a sufficient condition for the nonnegativity of a polynomial. Thus SOS relaxation provides a lower bound on the minimization problems of polynomial optimizations. Using the SOS relaxation, a large class of polynomial optimization problems with positive constraints can be formulated as SOS optimization as

$$\min_{\mathbf{d}} \mathbf{w}^\top \mathbf{d} \text{ s.t. } p_s(\mathbf{x}, \mathbf{d}) \in \Sigma[\mathbf{x}], p_e(\mathbf{x}; \mathbf{d}) = 0, \tag{5}$$

where $\Sigma[\mathbf{x}]$ denotes SOS set, $\mathbf{w}$ is weighting coefficient, p_s and p_e are polynomials parametrized by coefficients $\mathbf{d}$. The problem in (5) can be translated into a Semidefinite Programming (SDP) [Parrilo, 2003, Laurent, 2009]. There are readily available SOS optimization packages such as SOSTOOLS [Papachristodoulou et al., 2013] and SOSOPT [Seiler, 2013] for solving (5).

2.3 Almost everywhere uniform stability and Stabilization

This section derives results on a stronger notion of stability used in formulating optimal control problem with discounted cost. We first present the notion of, a.e., uniform stability as introduced in [Rajaram et al., 2010]. In the rest of the paper, we will use the following notation.

Definition 1 *[a.e. uniform stability] The equilibrium point of* (1) *is said to be a.e. uniform stable w.r.t. measure* $\mu \in \mathcal{M}(\mathbf{X})$ *if, for every given* ϵ, *there exists a time* $T(\epsilon)$ *such that*

$$\int_{T(\epsilon)}^{\infty} \mu(B_t) dt < \epsilon \tag{6}$$

where $B_t := \{\mathbf{x} \in \mathbf{X} : \mathbf{s}_t(\mathbf{x}) \in B\}$ *for any set* $B \in \mathcal{B}(\mathbf{X}_1)$.

The above stability definition essentially means that given any arbitrary set B not containing the origin, the measure of the set of all initial conditions that stay inside B can be made arbitrarily small after a sufficiently long time. Note that the above definition of a.e. uniform stability is stronger than the almost everywhere stability notion as introduced in [Rantzer, 2001] (refer to [Rajaram et al., 2010] for the proof). The following definition of a.e. exponential stability is introduced here and is stronger than the above Definition 1. The following exponential stability definition is a continuous-time counterpart of the discrete-time definition studied in [Vaidya and Mehta, 2008].

Definition 2 *[a.e. exponential stability] The equilibrium point is said to be almost everywhere exponential stable w.r.t. measure* μ *with rate* $\gamma > 0$ *if there exists a constant* M *such that*

$$\mu(B_t) \leq M e^{-\gamma t} \tag{7}$$

where $B_t := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{s}_t(\mathbf{x}) \in B\}$ *for any set* $B \in \mathcal{B}(\mathbf{X}_1)$.

In the following, we state theorems providing necessary and sufficient condition for a.e. uniform and a.e uniform exponential stability. These results are proved under the following assumption on the equilibrium point of (1).

Assumption 1 *We assume that* $\mathbf{x} = 0$ *is locally stable equilibrium point for the system (1) with local domain of attraction denoted by* $\mathcal{D}$ *and let* $0 \in \mathcal{N}_\delta \subset \mathcal{D}$.

Theorem 2 *The equilibrium point* $\mathbf{x} = 0$ *for the system (1) satisfying Assumption 1 is almost everywhere uniform stable w.r.t. measure* μ *if and only if there exists a density function* $\rho(\mathbf{x}) \in \mathcal{C}^1(\mathbf{X} \setminus \{0\}, \mathbb{R}^+)$ *which is integrable on* $\mathbf{X}_1$ *and satisfies*

$$\nabla \cdot (\mathbf{f}\rho) = h_0 \tag{8}$$

where $h_0 \in \mathcal{C}^1(\mathbf{X})$ *is the density function corresponding to the measure* μ.

Refer to [Rajaram et al., 2010, Theorem 13] for the proof.

3   Convex Formulation of Optimal Control Problem

In this section we briefly summarize the main results from [Huang and Vaidya, 2020] on the convex formulation of the optimal control problem. Consider a control affine system of the form

$$\dot{\mathbf{x}} = \bar{\mathbf{f}}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\bar{\mathbf{u}}, \quad \mathbf{x} \in \mathbf{X} \subseteq \mathbb{R}^n \tag{9}$$

where, $\mathbf{x}$ is the state, $\bar{\mathbf{u}} = [\bar{u}_1, \ldots, \bar{u}_m]^\top \in \mathbb{R}^m$ is the control input and $\mathbf{g}(\mathbf{x}) = (\mathbf{g}_1(\mathbf{x}), \ldots, \mathbf{g}_m(\mathbf{x}))$. All the vector fields are assumed to belong to $\mathcal{C}^1(\mathbf{X}, \mathbb{R}^n)$.

Remark 1 *The affine control assumption for a dynamical control system is not restrictive as any non-affine dynamical control system can be converted to control-affine by extending the state space. In particular, consider the control dynamical system of the form*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

then we can define $\mathbf{u}$ *as a new state and introduce* $\tilde{\mathbf{u}}$ *as another control input to write the above system as the following affine in the input control system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad \dot{\mathbf{u}} = \tilde{\mathbf{u}}$$

The following assumption is made on (9).

Assumption 3 *We assume that the linearization of the nonlinear system dynamics (9) at the origin is stabilizable i.e., the pair* $(\frac{\partial \bar{\mathbf{f}}}{\partial \mathbf{x}}(0), \mathbf{g}(0))$ *is stabilizable.*

Using the above stabilizability assumption, we can design a local stable controller using data. The detailed procedure for the design of such controller is given in Section 4. Let $\mathbf{u}_\ell$ be the locally stable controller. Defining $\mathbf{f}(\mathbf{x}) := \bar{\mathbf{f}}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}_\ell$ and $\mathbf{u} = \bar{\mathbf{u}} - \mathbf{u}_\ell$, we can rewrite control system (9) as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}. \tag{10}$$

The following is valid for the above dynamical system. With the control input $\mathbf{u} = 0$, the origin of system (10) is almost sure asymptotically stable locally in small neighborhood $\mathcal{D}$ of the origin such that $B_\delta \subset \mathcal{D}$.

Consider the discounted cost OCP of the form

$$J^\star(\mu_0) = \inf_{\mathbf{u}} \int_{\mathbf{X}_1} \left[\int_0^\infty e^{\gamma t}(q(\mathbf{x}(t)) + \beta\mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t)) \, dt \right] d\mu_0$$
$$\text{subject to (10)} \tag{11}$$

where $\gamma \in \mathbb{R}$. The existing literature on OCP with discounted cost address the case where γ is negative, i.e., negative discount factor. In this paper, with the stronger notion of a.e. uniform stability with geometric decay, we can address the case of cost with a positive discount. Note that the cost function is a function of initial measure μ_0, and this dependency on μ_0 can be explained as follows. The cost function can be written as

$$J(\mu_0) = \int_{\mathbf{X}_1} V(\mathbf{x}) d\mu_0(\mathbf{x}) \tag{12}$$

where, $V(\mathbf{x})$ can be written as

$$V(\mathbf{x}) = \int_0^\infty e^{\gamma t}(q(\mathbf{x}(t)) + \beta \mathbf{u}^\top(t)\mathbf{R}\mathbf{u}(t))dt$$

with $\mathbf{x}(\cdot)$ being a trajectory with initial condition $\mathbf{x}(0) = \mathbf{x}$. While $V(\mathbf{x})$ can be recognized with the familiar cost function used in the formulation of OCP in primal domain, the cost function $J(\mu_0)$ is minimized w.r.t. set of initial condition distributed with initial measure μ_0. In the rest of the paper we assume that the initial measure μ_0 is equivalent to Lebesgue with density function $0 < h_0(\mathbf{x}) \in \mathcal{L}_1(\mathbb{R}^n, \mathbb{R}_{>0}) \cap \mathcal{C}^1(\mathbb{R}^n)$. We make the following assumption on the OCP.

Assumption 4 *We assume that the state cost function $q : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is zero at the origin and uniformly bounded away from zero outside the neighborhood $\mathcal{N}_\delta$ and $\mathbf{R} > 0$. Furthermore, there exists a feedback control for which the cost function in (11) is finite and that the optimal control is feedback in nature, i.e., $\mathbf{u}^\star = \mathbf{k}^\star(\mathbf{x})$ with the function $\mathbf{k}^\star$ being in $\mathcal{C}^1(\mathbf{X}, \mathbb{R}^m)$.*

With the assumed feedback form of the optimal control input, the OCP can be written as

$$\inf_{\mathbf{k} \in \mathcal{C}^1(\mathbf{X})} \quad \int_{\mathbf{X}_1} \left[\int_0^\infty e^{\gamma t}(q(\mathbf{x}(t)) + \beta \mathbf{k}(\mathbf{x}(t))^\top \mathbf{R}\mathbf{k}(\mathbf{x}(t))) \, dt \right] d\mu_0$$

$$\text{s.t.} \qquad\qquad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x}) \tag{13}$$

The following is the main theorem on the OCP with discounted cost function.

Theorem 5 *Consider the OCP (13) with discount factor $\gamma \leq 0$ and assume that the cost function and optimal control satisfy Assumption 4. Then the OCP (13) can be written as the following infinite dimensional convex optimization problem*

$$J^\star(\mu_0) = \inf_{\rho \in \mathcal{S}, \bar{\boldsymbol{\rho}} \in \mathcal{C}^1(\mathbf{X}_1)} \int_{\mathbf{X}_1} q(\mathbf{x})\rho(\mathbf{x}) + \beta \frac{\bar{\boldsymbol{\rho}}(\mathbf{x})^\top \mathbf{R}\bar{\boldsymbol{\rho}}(\mathbf{x})}{\rho} d\mathbf{x}$$

$$\text{s.t.} \qquad \nabla \cdot (\mathbf{f}\rho + \mathbf{g}\bar{\boldsymbol{\rho}}) = \gamma\rho + h_0 \tag{14}$$

where $\bar{\boldsymbol{\rho}} = (\bar{\rho}_1, \ldots, \bar{\rho}_m)$ and $\mathcal{S} := \mathcal{L}_1(\mathbf{X}_1) \cap \mathcal{C}^1(\mathbf{X}_1, \mathbb{R}_{\geq 0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as

$$\mathbf{k}^\star(\mathbf{x}) = \frac{\bar{\boldsymbol{\rho}}^\star(\mathbf{x})}{\rho^\star(\mathbf{x})}. \tag{15}$$

Furthermore, if $\gamma = 0$, then optimal control $\mathbf{k}^\star(\mathbf{x})$ is a.e. uniformly stabilizing w.r.t. measure μ_0.

Proof of theorem (5) is given in Appendix.

Next, we consider discounted cost OCP with $\mathcal{L}_1$ norm on control term

$$\inf_{\mathbf{k} \in \mathcal{C}^1(\mathbf{X})} \quad \int_{\mathbf{X}_1} \left[\int_0^\infty e^{\gamma t}(q(\mathbf{x}(t)) + \beta \|\mathbf{k}(\mathbf{x}(t))\|_1) \, dt \right] d\mu_0(\mathbf{x})$$

$$\text{s.t.} \qquad\qquad \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x}). \tag{16}$$

We make the following assumption on the nature of optimal control for the $\mathcal{L}_1$-norm OCP (16).

Assumption 6 *We assume that the state cost function $q : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is zero at the origin and uniformly bounded away from zero outside the neighborhood $\mathcal{N}_\delta$ and $\mathbf{R} > 0$. Furthermore, there exists a feedback control input for which the cost function in (16) is finite. Furthermore, the optimal control is feedback in nature, i.e., $\mathbf{u}^\star = \mathbf{k}^\star(\mathbf{x})$ with the function $\mathbf{k}^\star$ is assumed to be $\mathcal{C}^1(\mathbf{X}, \mathbb{R}^m)$.*

Theorem 7 *Consider the OCP (16) with discount factor $\gamma \leq 0$ and assume that the cost function and optimal control satisfy Assumption 6 respectively. Then the OCP (16) can be written as following infinite dimensional convex optimization problem*

$$J^\star(\mu_0) = \inf_{\rho \in \mathcal{S}, \bar{\boldsymbol{\rho}} \in \mathcal{C}^1(\mathbf{X}_1)} \int_{\mathbf{X}_1} q(\mathbf{x})\rho(\mathbf{x}) + \beta \|\bar{\boldsymbol{\rho}}(\mathbf{x})\|_1 d\mathbf{x}$$

$$\text{s.t.} \qquad \nabla \cdot (\mathbf{f}\rho + \mathbf{g}\bar{\boldsymbol{\rho}}) = \gamma\rho + h_0 \tag{17}$$

where $\bar{\boldsymbol{\rho}} = (\bar{\rho}_1, \ldots, \bar{\rho}_m)$ and $\mathcal{S} := \mathcal{L}_1(\mathbf{X}_1) \cap \mathcal{C}^1(\mathbf{X}_1, \mathbb{R}_{\geq 0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as

$$\mathbf{k}^\star(\mathbf{x}) = \frac{\bar{\boldsymbol{\rho}}^\star(\mathbf{x})}{\rho^\star(\mathbf{x})}. \tag{18}$$

Furthermore, if $\gamma = 0$, then optimal control $\mathbf{k}(\mathbf{x})$ is a.e. uniformly stabilizing w.r.t. measure μ_0.

Proof 8 *The proof of Theorem 7 follows along similar lines to the proof of Theorem 5.*

Remark 2 *It is important to emphasize that the optimal feedback controller with discount factor $\gamma = 0$ is stabilizing in almost everywhere sense. This is analogous to the optimal control design in the primal formulation. The optimal cost function also serves as a Lyapunov function, thereby ensuring the stability of the feedback control system. In our proposed dual setting, the optimal density function serves as a.e. stability certificate for the case of discount factor $\gamma = 0$. However, due to the dual nature of the Lyapunov function and density function [Vaidya and Mehta, 2008, Rantzer, 2001], the density function has a singularity at the origin. Because of this singularity at the origin, the cost function is evaluated in $\mathbf{X}_1$ excluding the small region around the origin. Hence it may become necessary to design a local stabilizing or local optimal controller. The existence of such a local stabilizing controller is ensured following Assumption 3. The local controller, say $\mathbf{k}_\ell$, can be blended with global control $\mathbf{k}^\star$ using the following formula [Rantzer and Ceragioli, 2001].*

$$u(\mathbf{x}) = \frac{\rho_L}{\rho_L + \rho_N}\mathbf{k}_\ell(\mathbf{x}) + \frac{\rho_N}{\rho_L + \rho_N}\mathbf{k}^*(\mathbf{x})$$

$$\rho_L(\mathbf{x}) = max\{((\mathbf{x}^T P\mathbf{x}))^{-3} - \Delta, 0\}$$

where matrix $P > 0$ define a control Lyapunov function. The parameter Δ determines the region of operation for the local controller.

The optimal control results involving $\mathcal{L}_2$ and $\mathcal{L}_1$ norm on the control input with positive discount factor, i.e., $\gamma > 0$, are proved under the following assumption.

Assumption 9 *We assume that the state cost function $q : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is zero at the origin and uniformly bounded away from zero outside the neighborhood $\mathcal{N}_\delta$ and $\mathbf{R} > 0$. Furthermore, there exists a feedback control for which the cost function in (11) is finite and that the optimal control is feedback in nature, i.e., $\mathbf{u}^\star = \mathbf{k}^\star(\mathbf{x})$ with the function $\mathbf{k}^\star$ being in $\mathcal{C}^1(\mathbf{X}, \mathbb{R}^m)$. Furthermore, the feedback controller is assumed to be almost everywhere exponentially stabilizing (Definition 2) with decay rate $\gamma' > \gamma > 0$.*

Note that Assumption 9 is same as Assumption 4 except for the additional requirment that the feedback controller is a.e. exponentially stabilizing with decay rate strictly large than γ.

Theorem 10 *Consider the OCP (13) with discount factor $\gamma > 0$ and assume that the cost function and optimal control satisfy Assumption 9. Then the OCP (13) can be written as the following infinite dimensional convex optimization problem*

$$
\begin{aligned}
J^\star(\mu_0) &= \inf_{\rho \in \mathcal{S}, \bar{\rho} \in \mathcal{C}^1(\mathbf{X}_1)} \int_{\mathbf{X}_1} q(\mathbf{x})\rho(\mathbf{x}) + \beta\frac{\bar{\rho}(\mathbf{x})^\top \mathbf{R}\bar{\rho}(\mathbf{x})}{\rho}d\mathbf{x} \\
\text{s.t.} &\quad \nabla \cdot (\mathbf{f}\rho + \mathbf{g}\bar{\rho}) = \gamma\rho + h_0
\end{aligned}
\tag{19}
$$

where $\bar{\rho} = (\bar{\rho}_1, \ldots, \bar{\rho}_m)$ and $\mathcal{S} := \mathcal{L}_1(\mathbf{X}_1) \cap \mathcal{C}^1(\mathbf{X}_1, \mathbb{R}_{\geq 0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as

$$\mathbf{k}^\star(\mathbf{x}) = \frac{\bar{\rho}^\star(\mathbf{x})}{\rho^\star(\mathbf{x})}. \tag{20}$$

The proof of this theorem is provided in the Appendix. Theorem analogous to Theorem 7 can be stated and proved for the case involving $\mathcal{L}_1$ control norm and with positive discount factor $\gamma > 0$.

Remark 3 *In the above formulations of the OCPs, we did not explicitly impose constraints on the control input. Explicit constraints on the magnitude of the control input can be imposed in a convex manner as follows:*

$$\|\mathbf{u}\|_1 \leq M \iff |\bar{\rho}_1(\mathbf{x})|^2 + \ldots + |\bar{\rho}_m(\mathbf{x})|^2 \leq M\rho(\mathbf{x}), \tag{21}$$

for some positive constant M. To arrive at (21) we have used the formula for the optimal feedback control, i.e., (20) and the fact that $\rho(\mathbf{x}) > 0$. The above constraints are linear in the optimization variables $\bar{\rho}$ and ρ and hence can be implemented convexly. So the OCP involving explicit norm constraints on the control input can be implemented convexly by augmenting the optimization problem with linear constraints in (21).

Remark 4 *In the above formulations of the OCP problem, we have assumed that density function $h_0 > 0$, implying that the initial measure μ_0 is equivalent to Lebesgue. This assumption is necessary as $h_0 > 0$ guarantees that $\rho > 0$ (refer to Eq. (48)) and hence the the feedback control input $\mathbf{k} = \frac{\bar{\rho}}{\rho}$ is well defined. However, it is possible to relax this assumption and work with density function $h_0 \geq 0$. This will correspond to the case where the initial measure μ_0 is continuous w.r.t. Lebesgue measure and not equivalent to Lebesgue. In order to ensure that the feedback control input is well defined when μ_0 is continuous w.r.t. Lebesgue measure, we need to impose the following constraints on the control input*

$$|\bar{\rho}_k(\mathbf{x})|^2 \leq M\rho(\mathbf{x}), \quad k = 1, \ldots, m$$

for some large constant M. The above constraints will ensure that the for a.e. $\mathbf{x}$ if $\rho(\mathbf{x}) = 0 \implies \bar{\rho}_k(\mathbf{x}) = 0$ for $k = 1, \ldots, m$ thereby the feedback control input is well defined. Working with absolutely continuous initial measure μ_0 or equivalently $h_0 \geq 0$ will correspond to the case where optimality is guaranteed only from set of initial condition with support on μ_0.

In the following, we demonstrate how the results involving dual formulation to the OCP problem works out for the special case of scalar linear system. The main conclusion is that the optimal control obtained using dual formulation matches with the control obtained using a primal formulation of OCP, namely the linear quadratic regulator problem for a particular choice of $h_0(\mathbf{x})$. Note that the initial measure or the density function $h_0(\mathbf{x})$ is unique to our dual formulation with no parallel in the primal formulation.

4 Koopman and SOS-based Computation Framework for Optimal Control

This section provides Koopman and SOS-based computational framework for the finite-dimensional approximation of OCP involving $\mathcal{L}_1/\mathcal{L}_2$ costs. We begin with the following parameterization for the optimization variables ρ and $\bar{\rho}$.

$$\rho(\mathbf{x}) = a(\mathbf{x})/b(\mathbf{x})^{\alpha}, \quad \bar{\boldsymbol{\rho}}(\mathbf{x}) = \mathbf{c}(\mathbf{x})/b(\mathbf{x})^{\alpha}, \tag{22}$$

where $a(\mathbf{x}) \geq 0$ and $\mathbf{c}(\mathbf{x}) = [c_1(\mathbf{x}), \ldots, c_m(\mathbf{x})]^{\top}$. Here, $b(\mathbf{x})$ is a positive polynomial (positive at $\mathbf{x} \neq 0$), and α is a positive constant which is sufficiently large for integrability condition. In fact $b(\mathbf{x})$ is chosen to be control Lyapunov function based on the linearized control dynamics at the origin. The data-driven procedure for the identification of linear dynamics used in the construction of $b(\mathbf{x})$ is explained in Remark 7. The particular form for the parameterization of the optimization variable in (22) is chosen because of the fact that ρ has singularity at the origin (Remark 2). Using (22), we can write the constraints for the optimization problem as in (16) and (17) as

$$h = \nabla \cdot (\mathbf{f}\rho + \mathbf{g}\bar{\boldsymbol{\rho}}) - \gamma\rho = \nabla \cdot [(\mathbf{f}a + \mathbf{g}\mathbf{c})/b^{\alpha}] - \gamma\frac{a}{b^{\alpha}}$$

$$= \frac{1}{b^{\alpha+1}}[(1+\alpha)b\nabla \cdot (\mathbf{f}a + \mathbf{g}\mathbf{c}) - \alpha\nabla \cdot (b\mathbf{f}a + b\mathbf{g}\mathbf{c})$$

$$- \gamma ab].$$

With the above form of the constraints, we assume the following parameterization for $h = \frac{d}{b^{\alpha+1}}$, where d is an arbitrary positive polynomial. With the assumed form of h, we write the constraints in the optimization variable, a and $\mathbf{c}$ as

$$(1+\alpha)b\nabla \cdot (\mathbf{f}a + \mathbf{g}\mathbf{c}) - \alpha\nabla \cdot (b\mathbf{f}a + b\mathbf{g}\mathbf{c}) - \gamma ab = d \tag{23}$$

The above constraint in the optimization problem can be written in terms of the P-F generator as follows:

$$(1+\alpha)b\left(\mathcal{P}_{\mathbf{f}}a + \sum_{i=1}^{m}\mathcal{P}_{\mathbf{g_i}}c_i\right) - \alpha\left(\mathcal{P}_{\mathbf{f}}(ba) + \sum_{i=1}^{m}\mathcal{P}_{\mathbf{g_i}}(bc_i)\right)$$

$$- \gamma a = d \tag{24}$$

4.1 Data-driven Approximation of the Generators

From (24), it follows that the data-driven approximation of the constraints in the optimization problems (17) and (19) involves approximation of the P-F generators, $\mathcal{P}_{\mathbf{f}}$ and $\mathcal{P}_{\mathbf{g}_i}$. Furthermore, the P-F generator can be expressed in terms of the Koopman generator as

$$-\mathcal{P}_{\mathbf{f}}\psi = \nabla \cdot (\mathbf{f}\psi) = \mathbf{f} \cdot \nabla\psi + \nabla \cdot \mathbf{f}\psi = \mathcal{K}_{\mathbf{f}}\psi + \nabla \cdot \mathbf{f}\psi. \tag{25}$$

Expressing the P-F generator in terms of the Koopman generator allows us to use data-driven methods used to approximate the Koopman generator for the approximation of the P-F generator. In particular, we use generator Extended

Dynamic Mode Decomposition (gEDMD) algorithm from [Klus et al., 2020] for the approximation. To approximate Koopman generators, we first collect time-series data from the dynamical system in (9) by injecting different control inputs: i) zero control inputs, $\mathbf{u} = 0$, and ii) unit step control inputs, $\mathbf{u} = \mathbf{e}_j$[2] for $j = 1, \ldots, m$ for a finite time horizon with sampling step δt. Let,

$$\mathbf{X}_i = [\mathbf{x}_1, \ldots, \mathbf{x}_{T_i}], \ \dot{\mathbf{X}}_i = [\dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_{T_i}], \tag{26}$$

with $i = 0, 1, \ldots, m$ for zero and step control inputs where T_i are the number of data points for ith input case. The samples in $\mathbf{X}_i$ do not have to be from a single trajectory; it can be a concatenation of multiple experiment/simulation trajectories. Also, time derivatives of the states $\dot{\mathbf{x}}$ can be accurately estimated using numerical algorithms such as finite differences. Next, we construct a polynomial basis vector:

$$\boldsymbol{\Psi}(\mathbf{x}) = [\psi_1(\mathbf{x}), \ldots, \psi_Q(\mathbf{x})]^\top, \tag{27}$$

which can include monomials or Legendre/Hermite polynomials. The data-driven approximation of the generator will essentially involve the projection of the infinite-dimensional Koopman and P-F generators on the finite-dimensional space spanned by the basis function (27). Following [Klus et al., 2020], we define:

$$\dot{\boldsymbol{\Psi}}(\mathbf{x}, \dot{\mathbf{x}}) = [\dot{\psi}_1(\mathbf{x}, \dot{\mathbf{x}}), \ldots, \dot{\psi}_Q(\mathbf{x}, \dot{\mathbf{x}})]^\top, \tag{28}$$

$$\dot{\psi}_k(\mathbf{x}, \dot{\mathbf{x}}) = (\nabla_\mathbf{x} \psi_k)^\top \dot{\mathbf{x}} = \sum_{j=1}^n \frac{\partial \psi_k}{\partial x_j} \frac{dx_j}{dt} \tag{29}$$

The partial derivatives of the basis function are computed analytically which is required for $\dot{\psi}_k(\mathbf{x}, \dot{\mathbf{x}})$. Note that we also need $\frac{dx_j}{dt}$ which is simply denoted by $\dot{x}_j$. The value of $\dot{x}_j$ is approximated using finite differences:

$$\dot{x}_j \approx \frac{x_j - x_{j-1}}{\Delta t} \tag{30}$$

where x_{j-1} and x_j are the $j-1^{th}$ and j^{th} data point in the system trajectory and Δt is the time difference between two consecutive data points. A more sophisticated finite-difference method can be used, e.g., total variation regularization [Chartrand, 2011] for noisy data and discontinuous derivatives, and also second-order central difference for better accuracy. Then, the Koopman generator approximate $\mathbf{L}_i$ for each input case can be approximated as:

$$\mathbf{L}_i = \operatorname*{argmin}_{\mathbf{L}_i} ||\mathbf{B}_i - \mathbf{A}_i \mathbf{L}_i||_F,$$
$$\mathbf{A}_i = \frac{1}{T_i} \sum_{\ell=1}^{T_i} \boldsymbol{\Psi}(\mathbf{X}_{i,\ell}) \boldsymbol{\Psi}(\mathbf{X}_{i,\ell})^\top, \tag{31}$$
$$\mathbf{B}_i = \frac{1}{T_i} \sum_{\ell=1}^{T_i} \boldsymbol{\Psi}(\mathbf{X}_{i,\ell}) \dot{\boldsymbol{\Psi}}(\mathbf{X}_{i,\ell}, \dot{\mathbf{X}}_{i,\ell})^\top,$$

and $\mathbf{X}_{i,\ell}$ and $\dot{\mathbf{X}}_{i,\ell}$ denote ℓth column of $\mathbf{X}_i$ and $\dot{\mathbf{X}}$, respectively. The solution of (31) is explicitly known, $\mathbf{K}_i = \mathbf{A}_i^\dagger \mathbf{B}_i$, where $\dagger$ stands for pseudo-inverse. Given the Koopman generator approximates for $\mathbf{f}$, $\mathcal{K}_\mathbf{f} \approx \mathbf{L}_0$, using the linearity of the generators,

$$\mathcal{K}_{\mathbf{g}_j} \approx \mathbf{L}_j - \mathbf{L}_0, \quad j = 1, \ldots, m. \tag{32}$$

The above is one method to estimate $\mathcal{K}_\mathbf{f}$ and $\mathcal{K}_{\mathbf{g}_j}$. They can also be approximated jointly by using trajectories subject to arbitrary inputs. Next, we approximate the divergence of vector field $\mathbf{f}$ as

$$\nabla \cdot \mathbf{f} = \nabla \cdot [\mathcal{K}_0 x_1, \ldots, \mathcal{K}_0 x_n]^\top \approx \nabla \cdot (\boldsymbol{\mathcal{C}}_x^\top \mathbf{L}_0 \boldsymbol{\Psi}) \tag{33}$$

where $\boldsymbol{\mathcal{C}}_x$ is a coefficient vector for $\mathbf{x}$, i.e., $\mathbf{x} = \boldsymbol{\mathcal{C}}_x^\top \boldsymbol{\Psi}$, which can be found easily if $\boldsymbol{\Psi}$ includes 1st-order monomials (i.e., $\mathbf{x}$). Similarly, the divergence of vector fields $\mathbf{g}_j$ are approximated as

$$\nabla \cdot (\mathbf{g}_j) \approx \nabla \cdot (\boldsymbol{\mathcal{C}}_x^\top \mathbf{L}_j \boldsymbol{\Psi}), \ \ j = 1, \ldots, m. \tag{34}$$

from (25), (32)–(34), P-F generators are approximated by

$$\mathbf{P}_j = \mathbf{L}_j + \nabla \cdot (\boldsymbol{\mathcal{C}}_x^\top \mathbf{L}_j \boldsymbol{\Psi}) \mathbf{I} \tag{35}$$

for $j = 0, 1, \ldots, m$.

---

[2]$\mathbf{e}_j \in \mathbb{R}^m$ denotes unit vectors, i.e., jth entry of $\mathbf{e}_j$ is 1, otherwise 0.

Remark 5 *While the above procedure describes an approach for the approximation of the Koopman generators corresponding to the drift vector field, $\mathbf{f}(\mathbf{x})$, and control vector fields, $\mathbf{g}_j(\mathbf{x})$, for $j = 1, \ldots, m$ using zero input and step input, it is also possible to identify these vector field using random inputs. The problem of data-driven identification of the system dynamics using random or arbitrary control input will involve identifying bilinear vector fields. It can again be reduced to a least-square optimization problem similar to (31).*

To parameterize the optimization variables, we express polynomial functions $a(\mathbf{x})$, $b(\mathbf{x})$, and $c_j(\mathbf{x})$ with respect to the basis $\boldsymbol{\Psi}(\mathbf{x})$. Let $\boldsymbol{\mathcal{C}}_a$, $\boldsymbol{\mathcal{C}}_b$, and $\boldsymbol{\mathcal{C}}_{c_j}$ be the coefficient vectors used in the expansion of $a(\mathbf{x})$, $b(\mathbf{x})$, and $c_j(\mathbf{x})$:

$$a(\mathbf{x}) = \boldsymbol{\mathcal{C}}_a^\top \boldsymbol{\Psi}(\mathbf{x}), b(\mathbf{x}) = \boldsymbol{\mathcal{C}}_b^\top \boldsymbol{\Psi}(\mathbf{x}), c_j(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{c_j}^\top \boldsymbol{\Psi}(\mathbf{x}). \tag{36}$$

Note that $\boldsymbol{\mathcal{C}}_b$ contains constant coefficients of $b(\mathbf{x})$ since $b(\mathbf{x})$ is a known polynomial function. Similarly, let $\boldsymbol{\mathcal{C}}_{ab}$, $\boldsymbol{\mathcal{C}}_{bc_1}$, $\ldots$, $\boldsymbol{\mathcal{C}}_{bc_m}$ denote coefficient vectors of polynomials, $a(\mathbf{x})b(\mathbf{x})$, and $b(\mathbf{x})c_j(\mathbf{x})$, for $j = 1, \ldots, m$, namely,

$$a(\mathbf{x})b(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{ab}^\top \boldsymbol{\Psi}(\mathbf{x}), b(\mathbf{x})c_j(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{bc_j}^\top \boldsymbol{\Psi}(\mathbf{x}). \tag{37}$$

Constructing the coefficient vectors $\boldsymbol{\mathcal{C}}_{ab}$, and $\boldsymbol{\mathcal{C}}_{bc_j}$ from the coefficient vectors $\boldsymbol{\mathcal{C}}_a$, $\boldsymbol{\mathcal{C}}_b$, and $\boldsymbol{\mathcal{C}}_{c_j}$ requires trivial numerical procedures. In case that $\boldsymbol{\Psi}(\mathbf{x})$ is a monomial, finding these coefficient vectors are straightforward. If $\boldsymbol{\Psi}(\mathbf{x})$ is not a monomial vector, it involves more complicated numerical steps. One approach is to express the basis with respect to a common monomial vector $\mathcal{M}(\mathbf{x})$ such that $\boldsymbol{\Psi}(\mathbf{x}) = \boldsymbol{\mathcal{C}}_\Psi^\top \mathcal{M}(\mathbf{x})$ where $\boldsymbol{\mathcal{C}}_\Psi$ is a coefficient matrix that connects $\boldsymbol{\Psi}(\mathbf{x})$ and $\mathcal{M}(\mathbf{x})$. Then, we can find coefficient vectors in terms of the monomial vector $\mathcal{M}(\mathbf{x})$, i.e., $a(\mathbf{x})b(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{ab}^{'}\mathcal{M}(\mathbf{x})$ and $b(\mathbf{x})c_j(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{bc_j}^{'}\mathcal{M}(\mathbf{x})$, and convert these coefficient vectors back to the original basis $\boldsymbol{\Psi}(\mathbf{x})$ by multiplying pseudo-inverse of $\boldsymbol{\mathcal{C}}_\Psi$, i.e., $\boldsymbol{\mathcal{C}}_{ab} = \boldsymbol{\mathcal{C}}_\Psi^\dagger \boldsymbol{\mathcal{C}}_{ab}^{'}$ and $\boldsymbol{\mathcal{C}}_{bc_j} = \boldsymbol{\mathcal{C}}_\Psi^\dagger \boldsymbol{\mathcal{C}}_{bc_j}^{'}$. The implementation of this procedure can be done easily using polynomial toolbox provided in SOSOPT in Matlab [Seiler, 2013].

Now, using approximated infinitesimal PF generators in (35), we restate the LHS of (23) in (19) and (17) as:

$$\begin{aligned}
&(1 + \alpha)b(\mathbf{x}) \left(\boldsymbol{\mathcal{C}}_a^\top \mathbf{P}_0 \boldsymbol{\Psi}(\mathbf{x}) + \sum_{j=1}^m \boldsymbol{\mathcal{C}}_c^\top \mathbf{P}_j \boldsymbol{\Psi}(\mathbf{x}) \right) \\
&- \alpha \left(\boldsymbol{\mathcal{C}}_{ab}^\top \mathbf{P}_0 \boldsymbol{\Psi}(\mathbf{x}) + \sum_{j=1}^m \boldsymbol{\mathcal{C}}_{bc_j}^\top \mathbf{P}_j \boldsymbol{\Psi}(\mathbf{x}) \right) - \gamma a(\mathbf{x})b(\mathbf{x}).
\end{aligned} \tag{38}$$

Remark 6 *In [Korda and Mezić, 2018, Klus et al., 2020], convergence results for the finite-dimensional approximation of the Koopman operator and generators in the limit as the number of basis functions and data points goes to infinity is studied. These convergence results combined with the finite-dimensional approximation of the cost function for the optimization problem can be used to provide theoretical justification for the solution obtained using the finite-dimensional approximation of the infinite-dimensional optimization problem.*

In the following sections, we discuss the finite dimensional approximation of the cost function leading up to the finite dimensional approximation of the infinite dimensional optimization problems (17) and (19) involving $\mathcal{L}_1$ and $\mathcal{L}_2$ norms on control input respectively.

4.2   Optimal Control with $\mathcal{L}_1$ norm of feedback control

Using the assumed paramaterization for the ρ and $\bar{\rho}$ from (22) the cost function for the $\mathcal{L}_1$ OCP problem in (17) can be written as

$$\inf_{a \geq 0, \mathbf{c}} \int_{\mathbf{X}_1} \frac{q(\mathbf{x})a(\mathbf{x})}{b(\mathbf{x})^\alpha} + \frac{\beta \|\mathbf{c}(\mathbf{x})\|_1}{b(\mathbf{x})^\alpha} d\mathbf{x} \tag{39}$$

where a small neighborhood of the origin, $\mathcal{N} = \{\mathbf{x} \in \mathbf{X} : |\mathbf{x}| \leq \epsilon, \ \epsilon > 0\}$, is chosen as a polytope and excluded from the integration of the cost function to remove singularity of the density at the origin (refer to Remark 2).

To make (39) solvable, we introduce dummy polynomials $\mathbf{s}(\mathbf{x}) = [s_1(\mathbf{x}), \ldots, s_m(\mathbf{x})]^\top$, adding constraints:

$$\mathbf{s}(\mathbf{x}) - \mathbf{c}(\mathbf{x}) \geq 0, \ \mathbf{s}(\mathbf{x}) + \mathbf{c}(\mathbf{x}) \geq 0. \tag{40}$$

The polynomial $\mathbf{s}$ is expressed in terms of the basis function using the coefficient vector $\boldsymbol{\mathcal{C}}_{s_j}$ as:

$$s_j(\mathbf{x}) = \boldsymbol{\mathcal{C}}_{s_j}^\top \boldsymbol{\Psi}(\mathbf{x}), \tag{41}$$

for $j = 1, \ldots, m$. Substituting (37) in the integral cost (39), we obtain following finite dimensional approximation of the cost function.

$$\mathbf{d}_1^\top \boldsymbol{\mathcal{C}}_a + \beta \sum_{j=1}^m \mathbf{d}_2^\top \boldsymbol{\mathcal{C}}_{s_j} \tag{42}$$

where $\mathbf{d}_1$ and $\mathbf{d}_2$ are the coefficient vectors given by

$$\mathbf{d}_1 = \int_{\mathbf{X}_1} \frac{q(\mathbf{x})\boldsymbol{\Psi}(\mathbf{x})}{b(\mathbf{x})^\alpha} d\mathbf{x}, \ \mathbf{d}_2 = \int_{\mathbf{X}_1} \frac{\boldsymbol{\Psi}(\mathbf{x})}{b(\mathbf{x})^\alpha} d\mathbf{x}. \tag{43}$$

Using (40)–(43) and SOS positivity constraints denoted by $\Sigma[\mathbf{x}]$, (39) can be expressed as a SOS problem as:

$$\min_{\substack{\boldsymbol{\mathcal{C}}_a, \boldsymbol{\mathcal{C}}_{c_j}, \boldsymbol{\mathcal{C}}_{s_j} \\ j=1,\dots,m}} \ \mathbf{d}_1^\top \boldsymbol{\mathcal{C}}_a + \beta \textstyle\sum_{j=1}^m \mathbf{d}_2^\top \boldsymbol{\mathcal{C}}_{s_j}$$

$$\text{s.t. } (38) \in \Sigma[\mathbf{x}], \ a(\mathbf{x}) \in \Sigma[\mathbf{x}],$$
$$(\mathbf{s}(\mathbf{x}) - \mathbf{c}(\mathbf{x})) \in \Sigma[\mathbf{x}], \ (\mathbf{s}(\mathbf{x}) + \mathbf{c}(\mathbf{x})) \in \Sigma[\mathbf{x}]. \tag{44}$$

4.3 Optimal Control with $\mathcal{L}_2$ norm of feedback control

$\mathcal{L}_2$ OCP in (19) is restated as:

$$\min_{a,\mathbf{c}} \quad \int_{\mathbf{X}_1} \frac{q(\mathbf{x})a(\mathbf{x})}{b(\mathbf{x})^\alpha} + \beta \frac{\mathbf{c}(\mathbf{x})^\top \mathbf{R} \mathbf{c}(\mathbf{x})}{a(\mathbf{x})b(\mathbf{x})^\alpha} dx$$
$$\text{s.t. } \quad (38) \geq 0, \ a(\mathbf{x}) \geq 0. \tag{45}$$

by following the same parameterization in (22). Subsequently, we reformulate (45) as follows:

$$\min_{a,\mathbf{c},w} \quad \int_{\mathbf{X}_1} \frac{q(\mathbf{x})a(\mathbf{x})}{b(\mathbf{x})^\alpha} + \beta \frac{w(\mathbf{x})}{b(\mathbf{x})^\alpha} d\mathbf{x}$$
$$\text{s.t. } \quad (38) \geq 0, \ a(\mathbf{x}) \geq 0,$$
$$\mathbf{M}(\mathbf{x}) = \begin{bmatrix} w(\mathbf{x}) & \mathbf{c}(\mathbf{x})^\top \\ \mathbf{c}(\mathbf{x}) & a(\mathbf{x})\mathbf{R}^{-1} \end{bmatrix} \succcurlyeq 0, \tag{46}$$

where the positive semidefinite (PSD) of $\mathbf{M}(\mathbf{x})$ is a result of applying the Schur complement lemma on $\mathcal{L}_2$ cost bounded by $w(\mathbf{x})$, i.e., $\frac{\mathbf{c}(\mathbf{x})^\top \mathbf{R} \mathbf{c}(\mathbf{x})}{a(\mathbf{x})} \leq w(\mathbf{x})$. Now, to algebraically express $\mathbf{M}(\mathbf{x}) \succcurlyeq 0$, we introduce the lemma:

Lemma 11 (Positive semidefinite polynomial matrix) *[Scherer and Hol, 2006] A $p \times p$ matrix $\mathbf{H}(\mathbf{x})$ whose entries are polynomials is positive semidefinite with respect to the monomial vector $\mathbf{z}(\mathbf{x})$, if and only if, there exist $\mathbf{D} \succcurlyeq 0$ such that*

$$\mathbf{H}(\mathbf{x}) = \left(\mathbf{z}(\mathbf{x}) \otimes \mathbf{I}_p\right)^\top \mathbf{D} \left(\mathbf{z}(\mathbf{x}) \otimes \mathbf{I}_p\right),$$

where $\otimes$ denotes a Kronecker product (tensor product) and $\mathbf{I}_p$ is an identity matrix with dimension p.

Following Lemma 11, let $\mathbf{z}(\mathbf{x})$ be a monomial vector with the maximum degree equal to $\text{floor}(\deg(\boldsymbol{\Psi}(\mathbf{x}))/2) + 1$, then $\mathbf{M}(\mathbf{x})$ in (46) is PSD when there exists $\mathbf{D} \succcurlyeq 0$ such that $\mathbf{M}(\mathbf{x}) = \mathbf{H}(\mathbf{x})$. Using this result and (43), a SOS problem equivalent to (46) can be formulated as follows:

$$\min_{\substack{\boldsymbol{\mathcal{C}}_a, \boldsymbol{\mathcal{C}}_w, \boldsymbol{\mathcal{C}}_{c_j} \\ j=1,\dots,m}} \ \mathbf{d}_1^\top \boldsymbol{\mathcal{C}}_a + \beta \mathbf{d}_2^\top \boldsymbol{\mathcal{C}}_w$$

$$\text{s.t. } \quad (38) \in \Sigma[\mathbf{x}], \ a(\mathbf{x}) \in \Sigma[\mathbf{x}],$$
$$w(\mathbf{x}) - \mathbf{H}_{11}(\mathbf{x}) = 0, \ \mathbf{c}(\mathbf{x}) - \mathbf{H}_{12}(\mathbf{x}) = 0,$$
$$a(\mathbf{x})\mathbf{R}^{-1} - \mathbf{H}_{22}(\mathbf{x}) = 0, \ \mathbf{D} \succcurlyeq 0 \tag{47}$$

where $\mathbf{H}_{ij}(\mathbf{x})$ denotes the ijth entry of $\mathbf{H}(\mathbf{x})$; and $\boldsymbol{\mathcal{C}}_w$ is a coefficient vector of $w(\mathbf{x})$, i.e., $w(\mathbf{x}) = \boldsymbol{\mathcal{C}}_w^\top \boldsymbol{\Psi}(\mathbf{x})$.

Remark 7 *To obtain $b(\mathbf{x})$, the control Lyapunov function for the linearized dynamics, we first identify the linearized control system dynamics from time-series data collected near the origin. to identify the linear dynamics, we use the gEDMD algorithm discussed in Section 4.1 for the special case of identity basis functions i.e., $\boldsymbol{\Psi}(\mathbf{x}) = \mathbf{x}$. Once linearized system dynamics is identified we use linear quadratic regulator based optimal control for the construction of $b(\mathbf{x})$, namely $b(\mathbf{x}) = \mathbf{x}^\top P \mathbf{x}$, where P is the solution of algebraic Riccati equation (ARE). Following Assumption 3, we know that there exists a positive definite solution, P, to the ARE, which serves as control Lyapunov function for the linearized control system.*

5 Simulation Results

In this section, we present simulation results to illustrate the proposed data-driven control framework. All the simulation results are performed using MATLAB on a desktop computer with 64GB RAM. We have taken the value of $\alpha = 4$ and $\beta = 1$ for all our examples. The cost function and control matrix for each example is $q(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ and $R = 1$ respectively. Furthermore, the cost function is computed outside the region $\mathcal{N}$ of the neighborhood. However, we did not implement a local stabilizing controller around the origin and hence no blending controller (Remark 2). Also, the maximum degree of $a(\mathbf{x})$ is taken to be 1. We take the simulation time step $\Delta t = 0.01$ [s] for sampling time-series data, and also, Legendre polynomials are used as dictionary functions for all examples. We use SOSOPT [Seiler, 2013] toolbox to solve the formulated SOS optimization problems for the OCPs in (44) and (47). All other parameters used for each example are listed in Table 1.

Example 1: Consider the dynamics of controlled simple nonlinear numerical system:

$$\dot{x}_1 = -x_1 + x_2, \quad \dot{x}_2 = -0.5(x_1 + x_2) + 0.5x_1^2 x_2 + x_1 u.$$

With this example we use $\mathcal{L}_2$ cost on the control input. For this example, optimal control and optimal cost can be found by solving the HJB equation [Primbs and Nevistić, 1996] analytically and are given as below:

$$u^\star(\mathbf{x}) = -x_1 x_2, \quad V^\star(\mathbf{x}) = 0.5x_1^2 + x_2^2.$$

Next, using the proposed method, we get an optimal control with discount factor $\gamma = 0$ as follows.

$$k^\star(\mathbf{x}) = -1.38x_1 x_2 + 0.00005x_1 - 0.00004x_2 - 0.0063$$

By rounding off the coefficients of $k^\star(\mathbf{x})$, we see that $u^\star(\mathbf{x}) = k^\star(\mathbf{x})$. The small mismatch in decimal is due to the choice of $h_0(\mathbf{x})$, which is unique to our formulation but is absent from the primal formulation of OCP. The simulation results are obtained by solving inequality in the constraints corresponding to the case where $h_0 \geq 0$. In this example, we collected 2×10^4 time-series data points by simulating the system to estimate the Koopman generator. Figure 1 shows the comparison of trajectories simulated from the closed-loop system using the optimal control solutions obtained by the HJB approach (dotted red) and the proposed data-driven convex approach (solid black).

Example 2. Consider the dynamics of controlled Van der Pol oscillator as follows:

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = (1 - x_1^2)x_2 - x_1 + u.$$

In this example, we solve the $\mathcal{L}_2$ OCP for different values of the discount factor. Total 2×10^4 time-series data points are collected from repeated simulations to estimate the Koopman generator. Fig.2 and 4 show the trajectories of the closed-loop system starting from arbitrary initial points obtained from discount factor values, $\gamma = 0$ and $\gamma = 1$, respectively. We notice that the controller becomes more aggressive for larger γ and trajectories converge to the origin faster. This is expected as the OCPs achieve optimal control solutions at an exponential rate for $\gamma > 0$ for which the closed-loop system converges faster than uniform stability. On the other hand, we observe that, for negative discount factor $\gamma < 0$, the control solution is not guaranteed to stabilize the system to the origin, and the closed-loop dynamics converge to a limit cycle as shown in Fig. (4) resulting from $\gamma = -5$. This is again expected as the cost function is decreasing exponentially, and even without the stabilizing feedback controller, the optimal cost function is finite.

Example 3: Consider the dynamics of controlled simple inverted pendulum involing nonpolynomial dynamics:

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -\sin x_1 - 0.2x_2 + u$$

The number of data points used in the estimation of the Koopman operator equals 2×10^4. The simulation results for the $\mathcal{L}_2$ optimal control for different discount factor γ are shown in Fig.5 - Fig.7. Similar to Example 2, we notice that, for zero and positive discount factor γ, the controller obtained by positive discount factor can stabilize the origin at a

Table 1: Parameters for different examples

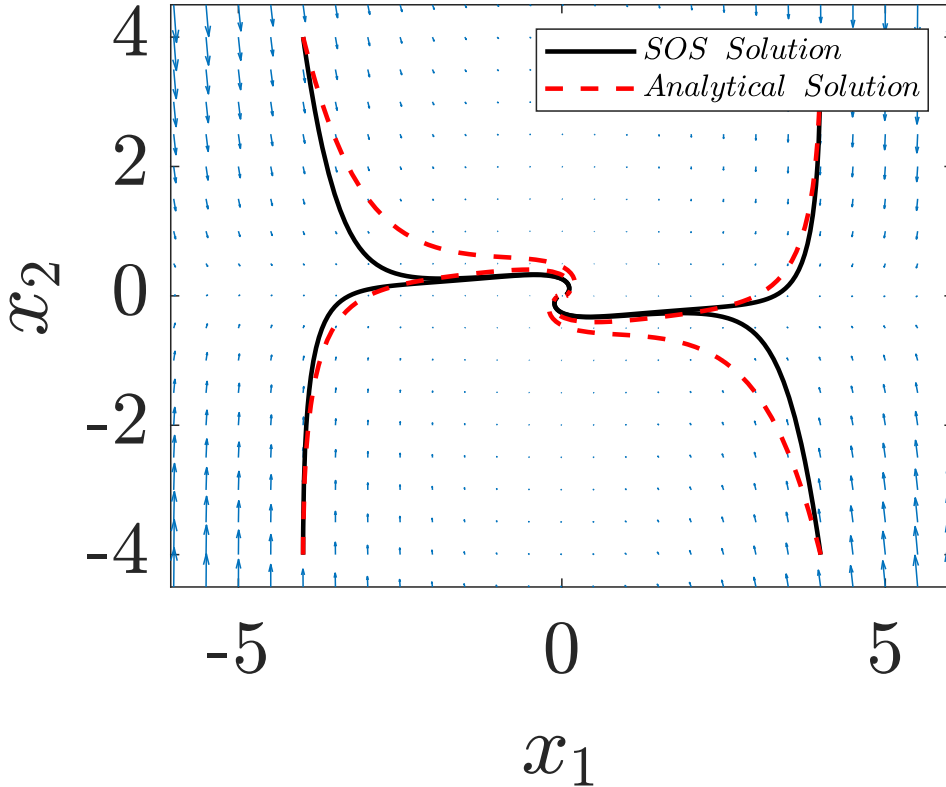|	Ex 1	Ex 2	Ex 3	Ex 4
$\mathbf{X}$	$[-5,5]^2$	$[-5,5]^2$	$[-5,5]^2$	$[-5,5]^3$
$\mathcal{N}$	$[-0.1,0.1]^2$	$[-0.1,0.1]^2$	$[-0.1,0.1]^2$	$[-0.1,0.1]^3$
$\deg(c(\mathbf{x}))$	2	6	3	6
$\deg(s(\mathbf{x}))$	2	7	7	6
$\boldsymbol{\Psi}(\mathbf{x})$	4th order	9th order	7th order	8th order

Figure 1: $\mathcal{L}_2$ OCP results of system in Example 1, showing converging state trajectories for $\gamma = 0$.

faster rate than the case of $\gamma = 0$ whereas, for negative discount factor, the origin is not stabilized for the closed-loop system.

Example 4: Consider the controlled Lorentz attractor:

$$\dot{x}_1 = \sigma(x_2 - x_1), \dot{x}_2 = x_1(\rho - x_3) - x_2 + u, \dot{x}_3 = x_1 x_2 - \eta x_3,$$

where $\sigma = 10$, $\rho = 28$, and $\eta = \frac{8}{3}$. The open-loop dynamics of the Lorentz system with the above parameter values are chaotic. For this example, we provide simulation results with $\mathcal{L}_1$ optimal control with $\gamma = 0$. We notice that the optimal control can stabilize the system.

6   Conclusion

A systematic convex optimization-based framework is provided for optimal control of nonlinear systems with a discounted cost function. In contrast to the existing literature of OCP with discounted cost, we consider the OCP problem with both positive and negative discount factor and provide a condition for the existence of optimal control. The OCP is formulated in the dual space of density function as an infinite-dimensional convex optimization problem. A new data-driven algorithm is provided for the computation of optimal control combining methods from Sum-of-Squares optimization and data-driven approximation of linear transfer operators. Simulation results are presented to verify the developed framework on data-driven optimal control.

Appendix

 Proof of Theorem 5
Consider the feedback control system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x})$$
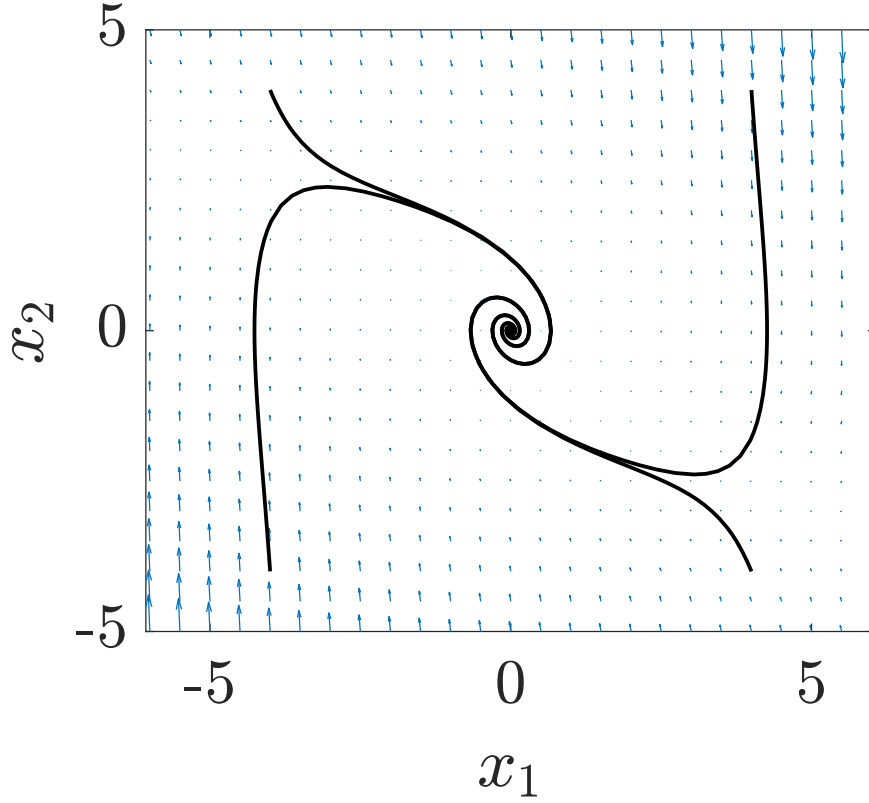
Figure 2: Trajectories of closed-loop system with $\mathcal{L}_2$ OCP solution for $\gamma = 0$ for Van der pol Oscillator.

and let $\mathbb{P}_t^c$ and $\mathbb{U}_t^c$ be the P-F and Koopman operator for the feedback control system. Using the definition of the Koopman operator, the cost in (13) can be written as

$$J(\mu_0) = \int_{\mathbf{X}_1} \int_0^\infty e^{\gamma t} [\mathbb{U}_t^c (q + \mathbf{k}^\top \mathbf{R} \mathbf{k})](\mathbf{x}) dt h_0(\mathbf{x}) d\mathbf{x}.$$

Using the duality and linearity of the Koopman and P-F operators, we obtain

$$J(\mu_0) = \int_{\mathbf{X}_1} \int_0^\infty (q + \mathbf{k}^\top \mathbf{R} \mathbf{k})(\mathbf{x}) e^{\gamma t} [\mathbb{P}_t^c h_0](\mathbf{x}) dt d\mathbf{x}.$$

Defining

$$\rho(\mathbf{x}) := \int_0^\infty e^{\gamma t} [\mathbb{P}_t^c h_0](\mathbf{x}) dt, \tag{48}$$

the $J(\mu_0)$ can be written as

$$J(\mu_0) = \int_{\mathbf{X}_1} (q(\mathbf{x}) + \mathbf{k}(\mathbf{x})^\top \mathbf{R} \mathbf{k}(\mathbf{x})) \rho(\mathbf{x}) d\mathbf{x}. \tag{49}$$

Defining $\bar{\rho}(\mathbf{x}) := \mathbf{k}(\mathbf{x}) \rho(\mathbf{x})$, the cost function can be written in the form given in (19). We next show that $\rho(\mathbf{x})$ and $\bar{\rho}$ satisfies the constraints in (19). Following Assumptions 4, we know that the state cost q is uniformly bounded away from zero in $\mathbf{X}_1$ and optimal cost function is finite and hence we have

$$\infty > J(\mu_0) \geq \int_{\mathbf{X}_1} q(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x} \geq c \int_{\mathbf{X}_1} \rho(\mathbf{x}) d\mathbf{x} \tag{50}$$

where c is the lower bound for the state cost function $q(\mathbf{x})$ on $\mathbf{X}_1$. The above proves that there exists a constant M such that

$$\int_{\mathbf{X}_1} \rho(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{X}_1} \int_0^\infty e^{\gamma t} [\mathbb{P}_t^c h_0](\mathbf{x}) dt d\mathbf{x} \leq M. \tag{51}$$
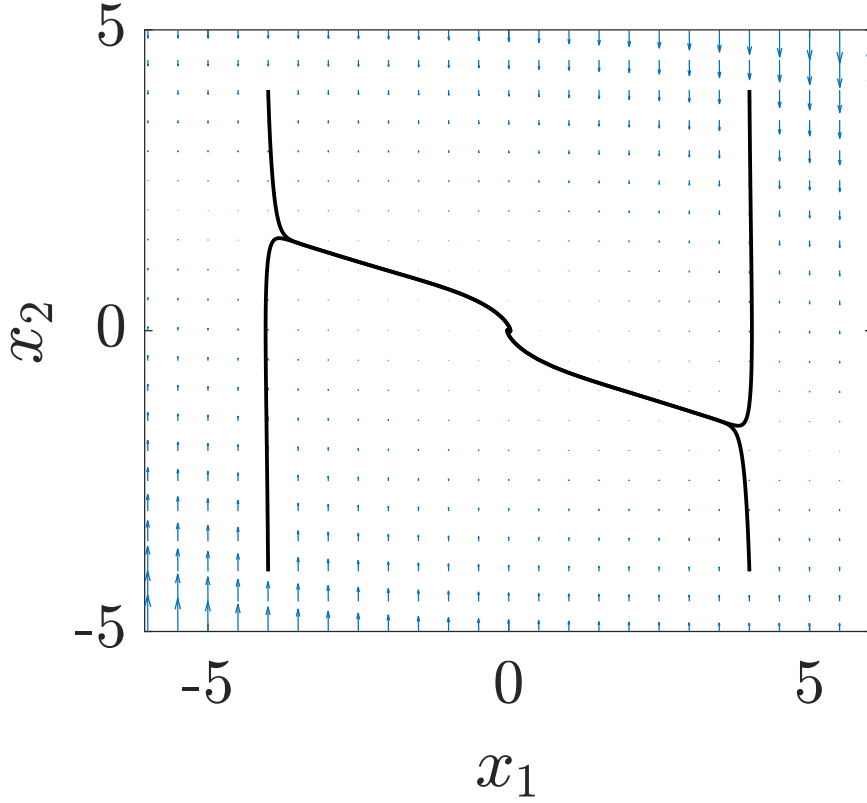
Figure 3: Trajectories of closed-loop system with $\mathcal{L}_2$ OCP solution for $\gamma = 1$ for Van der pol Oscillator.

We next claim that

$$\lim_{t \to \infty} e^{\gamma t}[\mathbb{P}_t h](\mathbf{x}) = 0 \tag{52}$$

for μ_0 almost all $\mathbf{x} \in \mathbf{X}_1$. Using Barbalat Lemma, we know that for $f(t) \in \mathcal{C}^1$, and $\lim_{t \to \infty} f(t) = \alpha$. If $f'(t)$ is uniformly continuous, then $\lim_{t \to \infty} f'(t) = 0$.

Letting $f(t) = \int_0^t e^{\gamma \tau} \int_{\mathbf{X}_1} [\mathbb{P}_\tau^c h_0](\mathbf{x}) d\mathbf{x} d\tau$ and hence $f'(t) = e^{\gamma t} \int_{\mathbf{X}_1} [\mathbb{P}_t h_0](\mathbf{x}) d\mathbf{x}$. By Barbalat Lemma, since $e^{\gamma t} \int_{\mathbf{X}_1} [\mathbb{P}_t^c h_0](\mathbf{x}) d\mathbf{x}$ is uniformly continuous w.r.t. time for $\gamma \leq 0$. The uniformly continuity follows from the definition of the P-F semi-group and the fact that solution of dynamical system is uniformly continuous w.r.t. time. We have

$$0 = \lim_{t \to \infty} e^{\gamma t} \int_{\mathbf{X}_1} [\mathbb{P}_t^c h_0](\mathbf{x}) d\mathbf{x} = \lim_{t \to \infty} \int_{\mathbf{X}_1} e^{\gamma t} [\mathbb{P}_t^c h_0](\mathbf{x}) d\mathbf{x}$$

which implies

$$\lim_{t \to \infty} e^{\gamma t}[\mathbb{P}_t^c h_0](\mathbf{x}) = 0 \tag{53}$$

for a.e. $\mathbf{x}$ w.r.t. μ_0 in $\mathbf{X}_1$. Next, we claim that $\rho(\mathbf{x})$ as defined in (48) can be obtained as a solution of the following equation

$$\nabla \cdot ((\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{k}(\mathbf{x}))\rho(\mathbf{x})) = \gamma \rho(\mathbf{x}) + h(\mathbf{x}), \tag{54}$$

for $x \in \mathbf{X}_1$. Substituting (48) in (54), we obtain

$$\nabla \cdot (\mathbf{f}_c \rho) = \int_0^\infty \nabla \cdot (\mathbf{f}_c(\mathbf{x}) e^{\gamma t}[\mathbb{P}_t^c h_0](\mathbf{x})) dt$$

$$= \int_0^\infty -e^{\gamma t} \frac{d}{dt}[\mathbb{P}_t^c h_0](\mathbf{x}) dt$$

$$-e^{\gamma t}[\mathbb{P}_t^c h_0](\mathbf{x})|_{t=0}^\infty + \int_0^\infty \gamma e^{\gamma t}[\mathbb{P}_t^c h_0](\mathbf{x}) dt$$

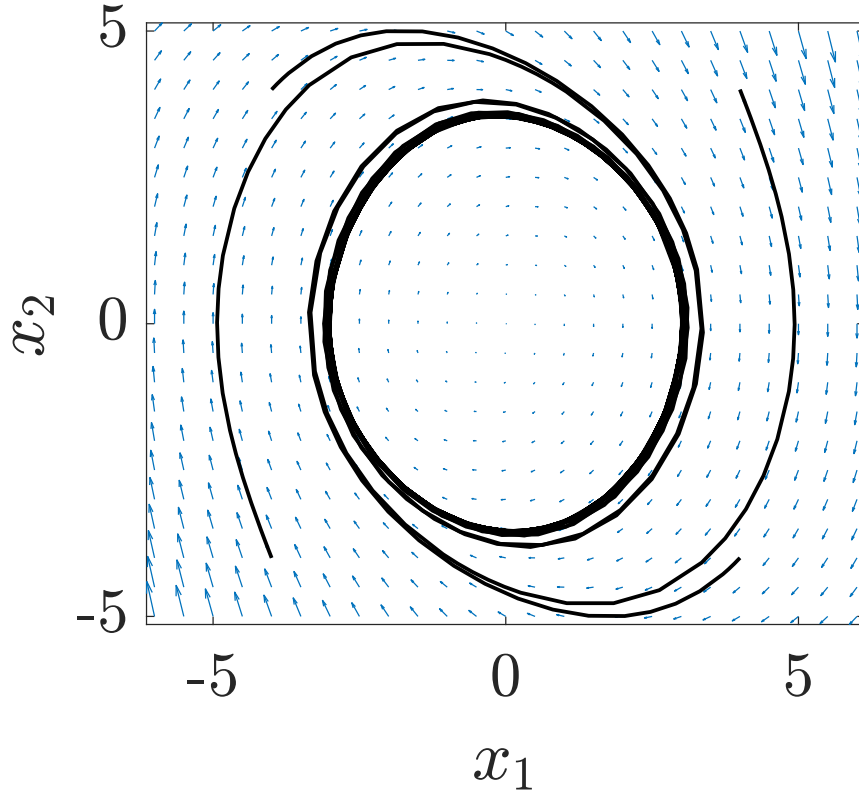$$= h_0(\mathbf{x}) + \gamma \rho(\mathbf{x}) \tag{55}$$

14

Figure 4: Trajectories of closed-loop system with $\mathcal{L}_2$ OCP solution for $\gamma = -5$ for Van der pol Oscillator.

In deriving (55) we have used the infinitesimal generator property of P-F operator Eq. (4) and the fact that $\lim_{t \to \infty} e^{\gamma t}[\mathbb{P}_t^c h_0](\mathbf{x}) = 0$ following (53). Furthermore, since $h_0(\mathbf{x}) > 0$, it follows that $\rho(\mathbf{x}) > 0$ from the positivity property of the P-F operator. Combining (49) and (55) along with the definition of $\bar{\rho}$, it follows that the OCP problem can be written as convex optimization problem (17). The optimal control $\mathbf{k}^\star(\mathbf{x})$ obtained as the solution of optimization problem (17) is a.e. uniform stable (for $\gamma = 0$) follows from the results of Theorem 2 using the fact that closed loop system satisfies (54) with ρ that is integrable. The optimal solution $\rho^\star(\mathbf{x}) \in \mathcal{L}_1(\mathbf{X}_1) \cap \mathcal{C}^1(\mathbf{X}_1, \mathbb{R}_{\geq 0})$ follows from the fact that $h_0 \in \mathcal{L}_1(\mathbb{R}^n, \mathbb{R}_{>0}) \cap \mathcal{C}^1(\mathbb{R}^n)$ and the definitions of ρ (48) and P-F operator (3).

Proof of Theorem 10 The proof of this theorem follows exactly along the lines of proof of Theorem 5 until equation (51). Unlike the proof of Theorem 5, where the claim (52) is proved using Barbalat Lemma, in this proof $\lim_{t \to \infty}[\mathbb{P}_t h_0](\mathbf{x}) = 0$ for μ_0 almost all $\mathbf{x} \in \mathbf{X}_1$ follows from Assumption 9. As $\mu_0(B_t) \leq M e^{-\gamma' t}$ implies $e^{\gamma t}[\mathbb{P}_t h_0](\mathbf{x}) \to 0$ for μ_0 a.e. $\mathbf{x} \in \mathbf{X}_1$. The rest of the proof then follows again along the lines of proof of Theorem 5.

References

[Arbabi et al., 2018] Arbabi, H., Korda, M., and Mezić, I. (2018). A data-driven koopman model predictive control framework for nonlinear partial differential equations. In *2018 IEEE Conference on Decision and Control*, pages 6409–6414. IEEE.

[Beard et al., 1997] Beard, R. W., Saridis, G. N., and Wen, J. T. (1997). Galerkin approximations of the generalized hamilton-jacobi-bellman equation. *Automatica*, 33(12):2159–2177.

[Bertsekas, 2011] Bertsekas, D. P. (2011). Approximate policy iteration: A survey and some new methods. *Journal of Control Theory and Applications*, 9(3):310–335.

[Chartrand, 2011] Chartrand, R. (2011). Numerical differentiation of noisy, nonsmooth data. *ISRN Applied Mathematics*, 2011:164564.
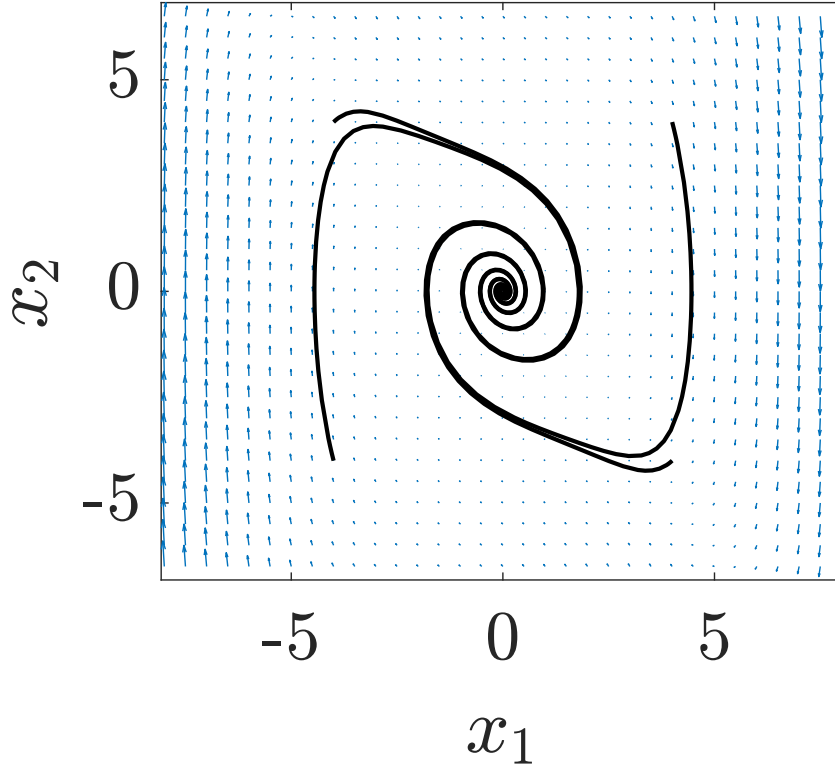
Figure 5: $\mathcal{L}_2$ OCP results of simple inverted pendulum, showing converging state trajectories for $\gamma = 0$.

[Choi et al., 2020] Choi, H., Vaidya, U., and Chen, Y. (2020). A convex data-driven approach for nonlinear control synthesis.

[Fleming and Rishel, 2012] Fleming, W. H. and Rishel, R. W. (2012). *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media.

[Ghosh et al., 1993] Ghosh, M. K., Arapostathis, A., and Marcus, S. I. (1993). Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM Journal on Control and Optimization*, 31(5):1183–1204.

[Henrion and Korda, 2013] Henrion, D. and Korda, M. (2013). Convex computation of the region of attraction of polynomial control systems. *IEEE Transactions on Automatic Control*, 59(2):297–312.

[Huang et al., 2018] Huang, B., Ma, X., and Vaidya, U. (2018). Feedback stabilization using koopman operator. In *2018 IEEE Conference on Decision and Control*, pages 6434–6439. IEEE.

[Huang et al., 2020] Huang, B., Ma, X., and Vaidya, U. (2020). Data-driven nonlinear stabilization using koopman operator. In *The Koopman Operator in Systems and Control*, pages 313–334. Springer.

[Huang and Vaidya, 2020] Huang, B. and Vaidya, U. (2020). A convex approach to data-driven optimal control via perron-frobenius and koopman operators. *Under review*.

[Kaiser et al., 2021] Kaiser, E., Kutz, J. N., and Brunton, S. L. (2021). Data-driven discovery of koopman eigenfunctions for control. *Machine Learning: Science and Technology*, 2(3):035023.

[Klus et al., 2020] Klus, S., Nüske, F., Peitz, S., Niemann, J.-H., Clementi, C., and Schütte, C. (2020). Data-driven approximation of the koopman generator: Model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416.

[Korda et al., 2017] Korda, M., Henrion, D., and Jones, C. N. (2017). Convergence rates of moment-sum-of-squares hierarchies for optimal control problems. *Systems & Control Letters*, 100:1–5.

[Korda and Mezić, 2018] Korda, M. and Mezić, I. (2018). On convergence of extended dynamic mode decomposition to the koopman operator. *Journal of Nonlinear Science*, 28(2):687–710.
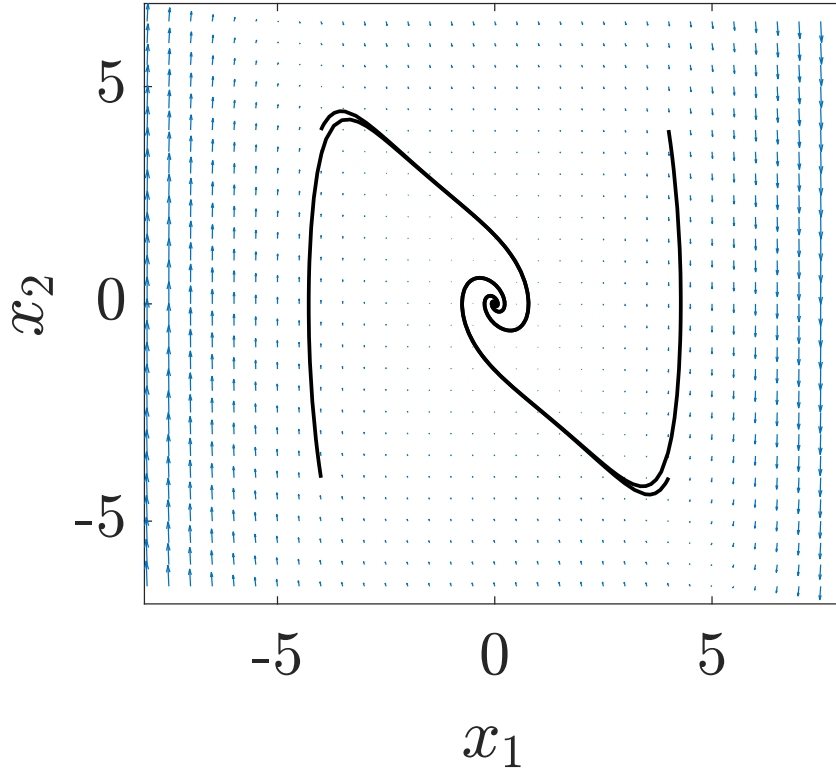
Figure 6: $\mathcal{L}_2$ OCP results of simple inverted pendulum, showing converging state trajectories for $\gamma = 2$.

[Korda and Mezić, 2020] Korda, M. and Mezić, I. (2020). Optimal construction of koopman eigenfunctions for prediction and control. *IEEE Transactions on Automatic Control*, 65(12):5114–5129.

[Lasota and Mackey, 1994] Lasota, A. and Mackey, M. C. (1994). *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, New York.

[Lasserre et al., 2008] Lasserre, J. B., Henrion, D., Prieur, C., and Trélat, E. (2008). Nonlinear optimal control via occupation measures and lmi-relaxations. *SIAM journal on control and optimization*, 47(4):1643–1666.

[Laurent, 2009] Laurent, M. (2009). *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, pages 157–270. Springer New York, New York, NY.

[Ma et al., 2019] Ma, X., Huang, B., and Vaidya, U. (2019). Optimal quadratic regulation of nonlinear system using koopman operator. In *2019 American Control Conference*, pages 4911–4916. IEEE.

[Mauroy and Mezić, 2013] Mauroy, A. and Mezić, I. (2013). A spectral operator-theoretic framework for global stability. In *52nd IEEE Conference on Decision and Control*, pages 5234–5239. IEEE.

[Modares and Lewis, 2014] Modares, H. and Lewis, F. L. (2014). Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning. *IEEE Transactions on Automatic control*, 59(11):3051–3056.

[Modares et al., 2016] Modares, H., Lewis, F. L., and Jiang, Z.-P. (2016). Optimal output-feedback control of unknown continuous-time linear systems using off-policy reinforcement learning. *IEEE Transactions on Cybernetics*, 46(11):2401–2410.

[Moyalan et al., 2021] Moyalan, J., Choi, H., Chen, Y., and Vaidya, U. (2021). Sum of squares based convex approach for optimal control synthesis. In *2021 29th Mediterranean Conference on Control and Automation (MED)*, pages 1270–1275. IEEE.

[Papachristodoulou et al., 2013] Papachristodoulou, A., Anderson, J., Valmorbida, G., Prajna, S., Seiler, P., and Parrilo, P. A. (2013). *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*. http://arxiv.org/abs/1310.4716.
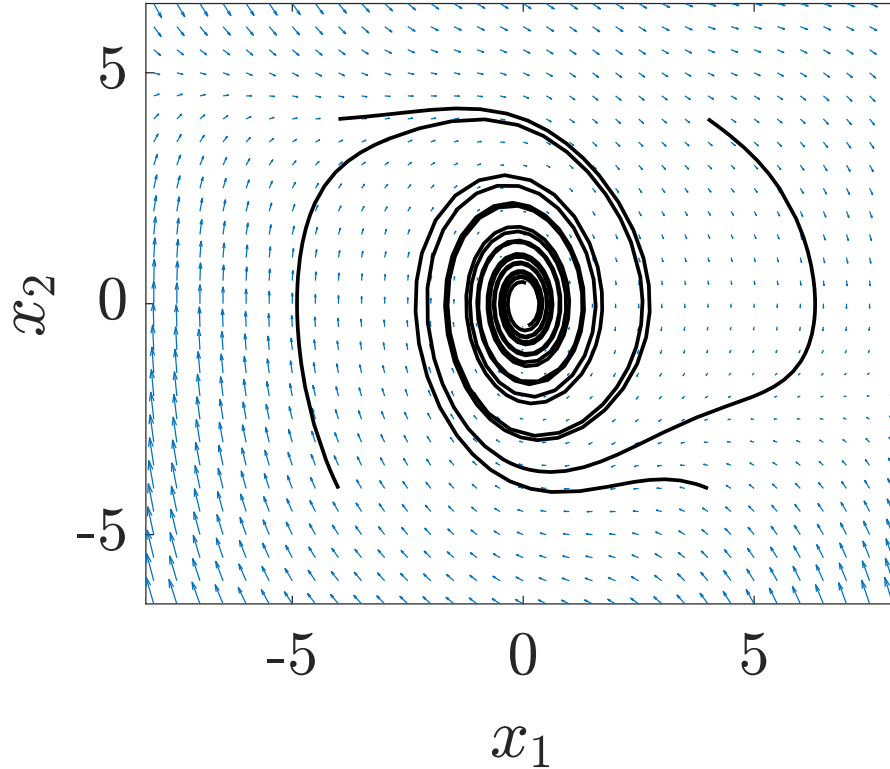
Figure 7: $\mathcal{L}_2$ OCP results of simple inverted pendulum, showing converging state trajectories for $\gamma = -5$.

[Parrilo, 2000] Parrilo, P. A. (2000). *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology.

[Parrilo, 2003] Parrilo, P. A. (2003). Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96:293–320.

[Parrilo and Sturmfels, 2003] Parrilo, P. A. and Sturmfels, B. (2003). Minimizing polynomial functions. *Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 60:83–99.

[Primbs and Nevistić, 1996] Primbs, J. A. and Nevistić, V. (1996). Optimality of nonlinear design techniques: A converse hjb approach.

[Raghunathan and Vaidya, 2013] Raghunathan, A. and Vaidya, U. (2013). Optimal stabilization using lyapunov measures. *IEEE Transactions on Automatic Control*, 59(5):1316–1321.

[Rajaram et al., 2010] Rajaram, R., Vaidya, U., Fardad, M., and Ganapathy-subramanian, B. (2010). Stability in the almost everywhere sense: A linear transfer operator approach. *Journal of Mathematical analysis and applications*, 368(1):144–156.

[Rantzer, 2001] Rantzer, A. (2001). A dual to Lyapunov's stability theorem. *Systems & Control Letters*, 42:161–168.

[Rantzer and Ceragioli, 2001] Rantzer, A. and Ceragioli, F. (2001). Smooth blending of nonlinear controllers using density functions. In *2001 European Control Conference (ECC)*, pages 2851–2853. IEEE.

[Scherer and Hol, 2006] Scherer, C. W. and Hol, C. W. J. (2006). Matrix sum-of-squares relaxations for robust semidefinite programs. *Mathematical Programming*, (6):189–211.

[Seiler, 2013] Seiler, P. (2013). Sosopt: A toolbox for polynomial optimization.

[Sutton and Barto, 2018] Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

[Topcu et al., 2010] Topcu, U., Packard, A., Seiler, P., and Balas, G. (2010). Help on sos [ask the experts]. *IEEE Control Systems Magazine*, 30(4):18–23.
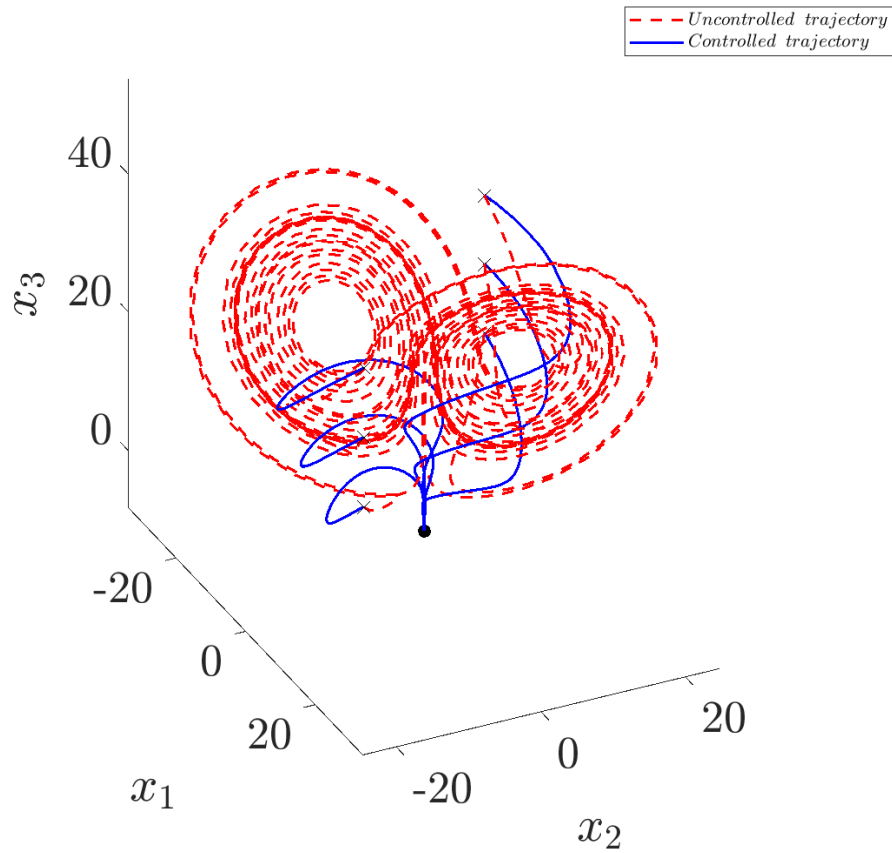
Figure 8: $\mathcal{L}_1$ OCP results of Lorentz attractor, showing converging state trajectories for $\gamma = 0$.

[Vaidya and Mehta, 2008] Vaidya, U. and Mehta, P. G. (2008). Lyapunov measure for almost everywhere stability. *IEEE Transactions on Automatic Control*, 53(1):307–323.

[Vaidya et al., 2010] Vaidya, U., Mehta, P. G., and Shanbhag, U. V. (2010). Nonlinear stabilization via control lyapunov measure. *IEEE Transactions on Automatic Control*, 55(6):1314–1328.

[Yu et al., 2021] Yu, H., Moyalan, J., Tellez-Castro, D., Vaidya, U., and Chen, Y. (2021). Convex optimal control synthesis under safety constraints. *Accepted for publication in IEEE CDC*.