A Convex Approach to Data-Driven Optimal Control via Perron–Frobenius and Koopman Operators

Bowen Huang and Umesh Vaidya

Abstract—This article is about the data-driven computation of optimal control for a class of control affine deterministic nonlinear systems. We assume that the control dynamical system model is not available, and the only information about the system dynamics is available in the form of time-series data. We provide a convex formulation for the optimal control problem (OCP) of the nonlinear system. The convex formulation relies on the duality result in the dynamical system’s stability theory involving density function and Perron–Frobenius operator. We formulate the OCP as an infinite-dimensional convex optimization program. The finite-dimensional approximation of the optimization problem relies on the recent advances made in the Koopman operator’s data-driven computation, which is dual to the Perron–Frobenius operator. Simulation results are presented to demonstrate the application of the developed framework.

Index Terms—Convex optimization, data-driven control, linear operator approach.

I. INTRODUCTION

The development of data-driven approaches for solving optimal control problems (OCP) for the dynamical system has attracted tremendous interest from various research communities. The solution to OCP involves solving an infinite-dimensional nonlinear partial differential equation, namely Hamilton Jacobi Bellman (HJB) equation. The non-linear and infinite-dimensional nature of the HJB equation makes the OCP challenging. Alternate formulations of OCP have been sought, which are more amicable to data-driven computation.

Progress is made for a class of linearly solvable OCP using alternate Kullback–Leibler (KL) based formulation of OCP for a stochastic dynamical system, and path integral-based numerical scheme [1]–[3]. The convex approach to the OCP we propose for data-driven control draws a parallel with the following literature on this topic. A convex formulation to OCP as an infinite-dimensional linear program is proposed in [4] and [5]. Duality relationship is used for convex formulation of OCP in [6], [7]. The density function is introduced as dual to Lyapunov function for verifying weaker notion of almost everywhere (a.e.) stability. The convex formulation relies on the duality result in the dual-density space and the positivity property is intimately connected to the optimal control, and we show that the optimal control ensures, a.e. uniform stability as introduced in [14] and [15]. Another main advantage of the linear operator viewpoint is that it allows us to use recent advances in the data-driven approximation of linear operators to provide a data-driven solution to the OCP. The operator theoretic viewpoint also enables us to extend the convex formulation to optimal control from deterministic to stochastic setting in a straightforward manner [16].

The linear P–F and Koopman operators are used to lift nonlinear dynamics from state space to linear, albeit infinite-dimensional, dynamics in the space of functions. There is a growing body of literature on using these operators, especially the Koopman operator, for dynamical analysis and control. We refer the readers to the following references for more details on this topic [17]–[23]. Our approach differs from the existing methods involving the use of the Koopman operator for data-driven control. In particular, we perform bilinear lifting of the control system using the P–F operator, which is dual to the Koopman operator. We address the control challenges associated with bilinear lifting of the control system using the convexity property in the dual-density space and the positivity property of the linear operators [15], [24], [25]. While the theoretical formulation of OCP is based on the P–F operator, the numerical framework exploits the advantage of the Koopman operator’s data-driven computation.

The main contributions of this article are as follows. We present a systematic framework based on the linear operator theory for the data-driven optimal control of a class of continuous-time nonlinear systems. The lifting involving the P–F operator is instrumental in the convex formulation of the OCP in the dual space of density. The stability property is intimately connected to the optimal control, and we show that the optimal control ensures, a.e. uniform stability of closed-loop dynamics. The proposed data-driven computational framework exploits the recent advances in approximating the Koopman operator and the duality between the Koopman and P–F operator. In particular, the computational framework makes use of the naturally structured dynamic mode decomposition (NSDMD) [26] algorithm for the approximation, preserving positivity, and Markov properties of the linear operators. Time-series data from single or multiple trajectories corresponding to a system with zero input and unit step input are used in the training process.

This article is organized as follows. In Section II, we present some preliminaries on the linear operator theory and NSDMD algorithm for the finite-dimensional approximation of the Koopman and P–F operators. The main results on the formulation of the convex optimization problem for optimal control are presented in Section III. The computational framework for the finite-dimensional approximation of...
the OCP is presented in Section IV. Simulation results are presented in Section V. Finally Section VI concludes this article.

II. PRELIMINARIES AND NOTATIONS

Following notations will be used throughout this article. \( \mathbb{R}^n \) denotes the \( n \) dimensional Euclidean space and \( \mathbb{R}_{\geq 0}^n \) is the positive orthant. Let, \( X \subseteq \mathbb{R}^n \) and \( Y \subseteq \mathbb{R}^m \). \( \mathcal{L}_1(X, Y) \), \( \mathcal{L}_c(X, Y) \), and \( C^k(X, Y) \) denote the space of all real valued integrable functions, essentially bounded functions, and space of \( k \) times continuously differentiable functions mapping from \( X \) to \( Y \), respectively. If the space \( Y \) is not specified then it is understood that the underlying space if \( \mathbb{R} \). \( B(X) \) denotes the Borel \( \sigma \)-algebra on \( X \) and \( \mathcal{M}(X) \) is the vector space of real-valued measure on \( B(X) \). \( s_i(x) \) denotes the solution of dynamical system \( x = F(x) \) starting from initial condition \( x \).

A. Perron–Frobenius and Koopman Operator

Consider a dynamical system of the form

\[
\dot{x} = F(x), \quad x \in X \subseteq \mathbb{R}^n.
\]  

(1)

where the vector field is assumed to be \( F(x) \in C^1(X, \mathbb{R}^n) \). There are two different ways of lifting the finite dimensional nonlinear dynamics from state space to infinite dimension space of functions namely using Koopman and Perron–Frobenius operators. The definitions of these operators along with the infinitesimal generators of these operators are defined as follows [27].

Definition 1 (Koopman Operator): \( U_t : \mathcal{L}_c(X) \rightarrow \mathcal{L}_c(X) \) for dynamical system (1) is defined as

\[
[U_t \phi](x) = \phi(s_t(x)).
\]

(2)

The infinitesimal generator for the Koopman operator is given by

\[
\lim_{t \to 0} \frac{(U_t - I)\phi}{t} = F(x) \cdot \nabla \phi(x) =: K_F \phi, \quad t \geq 0.
\]

(3)

Definition 2 (Perron–Frobenius Operator): \( \mathbb{P}_t : \mathcal{L}_1(X) \rightarrow \mathcal{L}_1(X) \) for dynamical system (1) is defined as

\[
[\mathbb{P}_t \psi](x) = \psi(s_t(x)) \left| \frac{\partial s_{-t}(x)}{\partial x} \right|
\]

(4)

where \( \cdot \) stands for the determinant. The infinitesimal generator for the P–F operator is given by

\[
\lim_{t \to 0} \frac{[\mathbb{P}_t - I] \psi}{t} = -\nabla \cdot (F(x) \psi(x)) =: Pf \psi, \quad t \geq 0.
\]

(5)

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

\[
\int_X [U_t \phi](x) \psi(x) dx = \int_X [\mathbb{P}_t \psi](x) \phi(x) dx.
\]

Property 1: These two operators enjoy positivity and Markov properties which are used in the finite dimension approximation of these operators.

1) Positivity: The P–F and Koopman operators are positive operators i.e., for any \( 0 \leq \phi(x) \in \mathcal{L}_c(X) \) and \( 0 \leq \psi(x) \in \mathcal{L}_1(X) \), we have

\[
[\mathbb{P}_t \psi](x) \geq 0, \quad [U_t \phi](x) \geq 0, \quad \forall t \geq 0.
\]

(6)

2) Markov Property: The P–F operator satisfies Markov property i.e.,

\[
\int_X [\mathbb{P}_t \psi](x) dx = \int_X \psi(x) dx.
\]

(7)

Definition 3 (Equivalent Measures): Two measures \( \mu_1 \in \mathcal{M}(X) \) and \( \mu_2 \in \mathcal{M}(X) \) are said to be equivalent i.e., \( \mu_1 = \mu_2 \) provided \( \mu_1(B) = 0 \) if and only if \( \mu_2(B) = 0 \) for every set \( B \in \mathcal{B}(X) \).

B. Almost Everywhere Uniform Stability and Stabilization

The formulation for the OCP, we present is intimately connected to density function and Lyapunov measure introduced for verifying following stronger notion of almost everywhere (a.e.) uniform stability as introduced in [14] and [28]. We first make following assumption on the system (1).

Assumption 1: We assume that \( x = 0 \) is locally stable equilibrium point for the system (1) with local domain of attraction denoted by \( \mathcal{N} \). Let \( B_\delta \) be the neighborhood of the origin for any given fixed \( \delta > 0 \) such that \( 0 \in B_\delta \subseteq \mathcal{N} \). We let \( X_1 := X \setminus B_\delta \)

Definition 4 (Almost everywhere (a.e.) uniform stability): The equilibrium point is said to be a.e. uniform stable w.r.t. measure \( \mu \in \mathcal{M}(X) \) if for any given \( \epsilon \), there exists a time \( T(\epsilon) \) such that

\[
\int_{T(\epsilon)}^\infty \mu(B_\delta) dt < \epsilon
\]

(8)

where \( B_\delta := \{ x \in X_1 : s_i(x) \in B \} \) for every set \( B \in \mathcal{B}(X_1) \).

The above stability definition implies that the system trajectories starting from a.e. initial conditions w.r.t. measure \( \mu \) of staying in any set \( B \in \mathcal{B}(X_1) \) can be made arbitrarily small after sufficiently large time. The main results of this article prove that the optimal control ensures a.e uniform stability of the equilibrium point. It is proved that a.e. uniform stability is stronger than a.e. stability [14, Lemma 7]. The a.e. stability as introduced in [7] is defined as follows.

Definition 5: The equilibrium point at \( x = 0 \) is said to be a.e. stable w.r.t. measure, \( \mu \), if

\[
\mu \{ x \in X_1 : \lim s_i(x) \neq 0 \} = 0.
\]

In the rest of this article we are going to make following assumption on measure \( \mu \).

Assumption 2: We assume that the measure \( \mu \in \mathcal{M}(X) \) is equivalent to Lebesgue with Radon–Nikodym derivative \( h \), i.e.,

\[
\frac{\mu}{\mu_X} = h(x) > 0 \quad \text{and} \quad h \in \mathcal{L}_1(X, \mathbb{R}_{>0}) \cap C^1(X).
\]

The following theorem is from [14, Th. 13] providing necessary and sufficient condition for a.e. uniform stability.

Theorem 2: The equilibrium point \( x = 0 \) for system (1) satisfying Assumption 1 is a.e. uniformly stable w.r.t. measure \( \mu \) if and only if there exists a function \( \rho(x) \in C^1(X \setminus \{0\}, \mathbb{R}_{>0}) \cap \mathcal{L}_1(X) \) and satisfies

\[
\nabla \cdot (F \rho) = h.
\]

(9)

C. DATA-DRIVEN APPROXIMATION: NSDMD

NSDMD is a modification of extended dynamic mode decomposition (EDMD) algorithm [29], one of the popular algorithms for Koopman approximation from data. The modifications are introduced to incorporate the natural properties of these operators namely positivity and Markov. For the continuous-time dynamical system (1), consider snapshots of dataset obtained as time-series data from single or multiple trajectories

\[
X = [x_1, x_2, \ldots, x_M], \quad Y = [y_1, y_2, \ldots, y_M]
\]

(10)

where \( x_i \in X \) and \( y_i \in Y \). The pair of datasets are assumed to be two consecutive snapshots i.e., \( y_i = s_{\Delta t}(x_i) \), where \( s_{\Delta t} \) is solution of (1). Let \( \Psi = [\psi_1, \psi_\alpha] \) be the choice of basis functions. The popular EDMD algorithm provides the finite-dimensional approximation of the
Koopman operator as the solution of the following least square problem:

$$\min_K \| GK - A \|_F$$  \hspace{1cm} (11)

where

$$G = \frac{1}{M} \sum_{m=1}^{M} \Psi(x_m)\Psi(x_m)^\top, \quad A = \frac{1}{M} \sum_{m=1}^{M} \Psi(x_m)\Psi(y_m)^\top$$  \hspace{1cm} (12)

with $K, G, A \in \mathbb{R}^{N \times N}$, $\| \cdot \|_F$ stands for Frobenius norm. The above least square problem admits an analytical solution

$$K_{EDMD} = G^\top A.$$  \hspace{1cm} (13)

Convergence results for EDMD algorithms in the limit as the number of data points and basis functions go to infinity are provided in [30] and [31]. In this article, we work with Gaussian Radial Basis Function (RBF) for the finite-dimensional approximation of the linear operators. Under the assumption that the basis functions are positive, like the Gaussian RBF, the NSDMD algorithm propose following convex optimization problem for the approximation of the Koopman operator that preserves positivity and Markov property in Property 1:

$$\min_P \| GP - \hat{A} \|_F$$

s.t. $[P]_{ij} \geq 0, \quad \hat{P} = \mathbb{I}$  \hspace{1cm} (14)

where

$$\hat{G} = GA^{-1}, \quad \hat{A} = AA^{-1}, \quad \hat{\Lambda} = \int_{\mathbb{X}} \Psi \Psi^\top dx$$  \hspace{1cm} (15)

with $G$ and $A$ are as defined in 12 and $\mathbb{I}$ is a vector of all ones. All the matrices in (15) are precomputed from the data. In fact, since the basis functions are assumed to be Gaussian RBF, the constant $A$ matrix can be computed explicitly as

$$\Lambda_{i,j} = \left( \frac{\pi \sigma^2}{2} \right)^{n/2} \exp \left( -\frac{||x_i-x_j||^2}{2\sigma^2} \right), \quad i, j = 1, 2, \ldots, N$$

where $x_i, x_j$ are the centers of the $\psi_i$ and $\psi_j$ Gaussian RBFs, respectively.

The approximation for the P–F operator and its generator are obtained as the solution of the optimization problem (14) as

$$P_{\Delta t} \approx \hat{P}^\top =: P, \quad \hat{P} = \frac{\hat{P}^\top - I}{\Delta t} =: M.$$  \hspace{1cm} (16)

III. MAIN RESULTS

We consider OCP for control affine system of the form

$$\dot{x} = \bar{f}(x) + g(x)\bar{u}, \quad x \in \mathbb{X} \subseteq \mathbb{R}^n$$  \hspace{1cm} (17)

where, the vector field $\bar{f}, g \in C^1(\mathbb{X}, \mathbb{R}^n)$ and $\bar{u} \in \mathbb{R}$ is the control input. For the simplicity of presentation we present the results for the case of single input, the results generalize to the multi-input case in straight-forward manner. We make following assumption on the control system (17).

Assumption 3: We assume that for the OCP (19) for a given $\mu_0$ there exists a feedback control input, $u = k(x)$, such that the cost function corresponding to this input is finite. Furthermore, the optimal control is feedback in nature i.e., $u^* = k^*(x)$ with the function $k^*(x)$ is assumed to be $C^1(X)$.

Remark 3: The Assumption 5 ensures the existence of the solution to the OCP and serves as a necessary condition for solving the OCP. The presence of measure $\mu_0$ and the associated density function $h_0$ is unique to our formulation. For systems evolving on a compact state space, integrability of $h_0$ and stabilizability property of the feedback controller is sufficient to ensure that the optimal cost, $J^*(\mu_0)$, is finite. However, on unbounded state-space, additional assumptions on function $h_0$ will be necessary. For example, with quadratic cost function i.e., $q(x) = \|x\|^2$ it is necessary that the function $h_0$ decay at the rate $\|x\|^{-\alpha}$ with $\alpha \geq 4$ as $x \to \infty$ to ensure optimal cost is finite. Given the

$$u := \bar{u} - u_\epsilon,$$ we can write system equation (17) as

$$\dot{x} = f(x) + g(x)u.$$  \hspace{1cm} (18)

Following is true for the control system (18). The origin of system (18) is locally stable with $u = 0$. We denote the local domain of attraction around the origin by $\mathcal{N}$. Let $B_\delta \subset \mathcal{N}$ is a $\delta$ neighborhood of equilibrium point for any fixed $\delta > 0$. We denote $X_1 := X \setminus B_\delta$. The objective is to design optimal controller which is active outside $B_\delta$. In Section III-B, we outline procedure for combining the local and global optimal controller using a procedure of smooth blending controller discussed in [32]. In the following, we assume that the measure $\mu_0 \in M(X)$ and is equivalent to Lebesgue with Radon–Nikodým derivative $h$ i.e., $\frac{d\mu_0}{dx} = h_0(x) > 0$ and $h_0 \in L_1(X, \mathbb{R}_{>0}) \cap C^1(X)$.

A. Convex Formulation of Optimal Control

Consider the OCP with cost function of the form

$$J^*(\mu_0) = \inf_u \int_{X_1} \int_0^\infty (q(x(t)) + ru^2(t)) \; dt d\mu_0(x)$$

subject to (18).  \hspace{1cm} (19)

We make following assumptions on state cost function $q(x)$ and optimal control.

Assumption 4: We assume that the state cost function $q : \mathbb{R}^n \to \mathbb{R}_{>0}$ is zero at the origin and uniformly bounded away from zero outside the neighborhood $B_\delta$ and $r > 0$.

Some comments on the nature of the assumed cost function are necessary. Let

$$V(x) = \int_0^\infty \left( q(x(t)) + ru(t)^2 \right) dt$$  \hspace{1cm} (20)

which is the cost function for the usual optimization problem formulated in the state space. Using (20) the cost function (19) can be written as

$$J(\mu_0) = \int_{X_1} V(x) d\mu_0(x) = \int_{X_1} V(x) h_0(x) dx$$

where we used the fact that $d\mu_0(x) = h_0(x) dx$. Note that the new cost function in the density space is weighted with respect to the given density function $h_0$ associating different weightage to different initial conditions or sets.

Another distinguishing feature of the cost function is that the cost function is evaluated and minimized only on set $X_1$. Using Assumption 3, we can design a local stabilizing controller and restrict the construction of optimal control outside the set $B_\delta$ around the origin. In Section III-B, we outline a procedure for blending the local stabilizing controller with the global optimal controller smoothly.

Assumption 5: We assume that for the OCP (19) for a given $\mu_0$ there exists a feedback control input, $u = k(x)$, such that the cost function corresponding to this input is finite. Furthermore, the optimal control is feedback in nature i.e., $u^* = k^*(x)$ with the function $k^*(x)$ is assumed to be $C^1(X)$.
computational focus of this article, we can restrict to a compact state space where integrability of $h_0$ combined with stabilizability property of the feedback controller is sufficient to ensure a finite value of optimal cost. The stabilizability property of the control system can be verified using data (refer to Section IV-C on data-driven stabilization).

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

$$J^*(\mu_0) = \inf_{\rho \in C^1(X_1)} \int_{X_1} \left[ \int_0^\infty q(s^e_t(x)) + r k^2(s^e_t(x)) \; dt \right] d\mu_0(x)$$

subject to $\dot{x} = f(x) + g(x)k(x)$

(21)

where $s^e_t(x)$ is the solution of the closed loop system (21) from initial condition $x$. We now state the main theorem on the convex formulation of the OCP.

**Theorem 4:** Consider the OCP (21), with the cost function and optimal control satisfying Assumptions 4, and 5, respectively. The OCP (21) can be written as following infinite dimensional convex problem

$$J^*(\mu_0) = \inf_{\rho \in S} \int_{X_1} q(x)\rho(x) + \frac{\rho(x)^2}{\rho} \; dx$$

subject to $\nabla \cdot (\mathbf{f}(x) + \mathbf{g}(x)\bar{k}(x)) = 0$

(22)

where $S := L_1(X_1) \cap C^1(X_1, \mathbb{R}_{>0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as

$$k^*(x) = \frac{\bar{\rho}^*(x)}{\rho^*(x)}$$

(23)

where $(\rho^*, \bar{\rho}^*)$ are solution of (22). Furthermore, the optimal control $k^*(x)$ is a.e. uniformly stabilizing the origin.

**Proof:** Consider the feedback control system

$$\dot{x} = f(x) + g(x)k(x)$$

(24)

where $u = k(x)$ be the feedback controller satisfying Assumption 5. The cost in (21) can be written as

$$J(\mu_0) = \int_{X_1} \left[ \int_0^\infty (q + r k^2)(s^e_t(x)) \; dt \right] d\mu_0(x)$$

where $s^e_t(x)$ is the solution of (24) and we use the fact that $d\mu_0 = h_0 \; dx$. Performing change of coordinates $y = s^e_t(x)$, $x = s^e_t(y)$ and $dx = \frac{\partial s^e_t}{\partial y} \; dy$. With the change of coordinates and after relabeling the cost function above can be written as

$$J(\mu_0) = \int_{X_1} \left[ \int_0^\infty (q + r k^2)(s^e_t(x)) \left| \frac{\partial s^e_t}{\partial y}(y) \right| \; dt \right] \; dx$$

$$= \int_{X_1} \left( q + r k^2 \right)(x) \left( \int_0^\infty h_0(s^e_t(x)) \left| \frac{\partial s^e_t}{\partial y}(y) \right| \; dt \right) \; dx$$

$$= \int_{X_1} \left( q + r k^2 \right)(x) \left( \int_0^\infty [\bar{P}^*_t h_0(x)](x) \; dt \right) \; dx$$

(25)

where $\bar{P}^*_t$ is the P-F operator for the closed loop system (24).

Using the Assumption 5 on the existence of optimal control for which the optimal cost is finite, we obtain

$$\kappa \int_{X_1} \left[ \int_0^\infty [\bar{P}^*_t h_0(x)](x) \; dt \right] \leq J(\mu_0) \leq M \times \infty$$

(26)

for some constant $M > 0$, where $\kappa$ is assumed lower bound on the state cost $g(x)$ following Assumption 4. We define

$$\rho(x) := \int_0^\infty [\bar{P}^*_t h_0(x)](x) \; dt$$

(27)

It follows from (26) that $\rho(x)$ is well defined for a.e. $x$ and is an integrable function. From the definition of the P-F operator and the assumption made on function $h_0$ it follows that the function $[\bar{P}^*_t h_0(x)](x)$ is uniformly continuous function of time. Hence, using Barbalat Lemma [33, pg. 269], we have

$$\lim_{t \to \infty} [\bar{P}^*_t h_0(x)](x) = 0$$

(28)

for a.e. $x$

Now, using the definition of $\rho$ from (27) we write (25) as

$$J(\mu_0) = \int_{X_1} (q + r k^2)(x)\rho(x) \; dx$$

(29)

Defining $\bar{\rho}(x) := \rho(x)k(x)$ it follows that (29) can be written in the form (22). We next show that $\rho(x)$ and $\bar{\rho}$ satisfies the constraints in (22). Substituting (27) in the constraint of (22), we obtain

$$\nabla \cdot (f(x)\rho(x)) = \int_0^\infty \nabla \cdot (f(x)[\bar{P}^*_t h_0(x)](x)) \; dt$$

$$= \int_0^\infty \frac{d}{dt} [\bar{P}^*_t h_0(x)](x) \; dt = -[\bar{P}^*_t h_0(x)](x)^\infty_{t=0} = h_0(x)$$

(30)

where, $f(x) := f(x) + g(x)k(x)$. In deriving (30), we have used the infinitesimal generator property of P-F operator (5) and the fact that $\lim_{t \to \infty} [\bar{P}^*_t h_0(x)](x) = 0$ following (28). Furthermore, since $h_0 > 0$, it follows that $\rho > 0$ from the positivity property of the P-F operator. Combining (25) and (30) along with the definition of $\bar{\rho}(x) := \rho(x)k(x)$, it follows that the OCP problem can be written as convex optimization problem (22). The optimal control $k^*(x)$ from (23) is a.e. uniformly stabilizing w.r.t. measure $\mu_0$ follows from the results of Theorem 2 and using the fact that closed-loop system satisfies (30), $\bar{\rho}(x) \in L_1(X_1), \rho(x) \in L_1(X_1) \cap C^1(X_1, \mathbb{R}_{>0})$ follows from the fact that $h_0 \in C^1(X)$, the integral formula for $\rho$ in (27), and the definition of P-F operator (4).

We next consider the optimization problem involving 1-norm on the control input.

$$\inf_{\mu \in C^1(X)} \int_{X_1} \left[ \int_0^\infty q(x(t)) + r|k(x(t))| \; dt \right] d\mu_0(x)$$

subject to $\dot{x} = f(x) + g(x)k(x)$

(31)

We make following assumption

**Assumption 6:** We assume that for the OCP (31) there exists a feedback control input such that the cost function corresponding to the input is finite. Furthermore, the optimal control is feedback in nature i.e., $u^* = k^*(x)$ with the function $k^*(x)$ is assumed to be $C^1(X)$.

**Theorem 6:** Consider the OCP (31), with the cost function, and optimal control satisfying Assumptions 4 and 6, respectively. The OCP (21) is written as following infinite dimensional optimization problem

$$J^*(\mu_0) = \inf_{\rho \in S} \int_{X_1} q(x)\rho(x) + r|\bar{\rho}(x)| \; dx$$

subject to $\nabla \cdot (\mathbf{f}(x) + \mathbf{g}(x)\bar{k}(x)) = 0$

(32)

where $S := L_1(X_1) \cap C^1(X_1, \mathbb{R}_{>0})$. The optimal feedback control input is recovered from the solution of the above optimization problem as $k^*(x) = \frac{\bar{\rho}^*(x)}{\rho^*(x)}$. Furthermore, the optimal control $k^*(x)$ is a.e. uniformly stabilizing the origin.

**Proof:** The proof of this theorem follows along the lines of proof of Theorem 4.
### B. Local Optimal Control and Controller Blending

The density function $\rho$ for the solution of optimization problem satisfy

$$\rho(x) = \int_0^\infty [P(t)h_0](x)dt$$

where $P(t)$ is the P–F operator for the closed-loop system $\dot{x} = f(x) + g(x)k(x)$ and hence $\rho$ serves as an occupancy measure i.e.,

$$\int_A \rho(x)dx = \langle \int_0^\infty [U\chi_A]dt, h \rangle$$

for any set $A \in B(\mathbb{R}^n)$ signifies the amount of time closed-loop system trajectories spend in the set $A$ with initial condition supported on measure $\mu_0$, where $\chi_A$ is the indicator function of set $A$. Because of this, $\rho(x)$ has a singularity at the equilibrium point stabilized by the closed-loop system. We exclude the small neighborhood around the origin for the proper parameterization of the density function $\rho$ in the computation of optimal control due to singularity at the origin. In particular, the optimization problem (22) is solved excluding the small neighborhood around the origin. For the small region around the origin, we either design stabilizing controller or optimal control depending upon the linearized system’s stabilizability or controllability property around the origin. For the design of the local controller we use time-series for the identification of linear dynamics. Let $k(x)$ and $\rho(x)$ be the global optimal controller obtained as the solution of optimization problem (22). Let $k_0$ be the local feedback controller obtained by solving linear quadratic regular (LQR) optimal control or Lyapunov based stabilization control using the identified linearized dynamics and $\rho_0(x) = \max[1/(x^TPx)^{-3} - \gamma, 0]$, where $P$ is a solution of Riccati equation or controlled Lyapunov equation obtained based on linearized dynamics. The local and global controllers are combined using the blending procedure from [32] as follows:

$$\bar{u} = \frac{\rho_0(x)}{\rho(x) + \rho_0(x)}k_0(x) + \frac{\rho(x)}{\rho(x) + \rho_0(x)}k(x)$$

The region where local controller is active is given by $x^TPx \leq (\frac{1}{2})^2$.

### IV. DATA DRIVEN APPROXIMATION

For the data-driven computation of optimal control, we need to provide finite dimensional approximation of the infinite dimensional linear program (22) and (32). Toward this goal, we need the data-driven approximation of the generator corresponding to vector field $f$ and $g$ i.e., $\nabla \cdot (f\rho)$ and $\nabla \cdot (g\dot{\rho})$.

**Remark 7:** In this article, we use Gaussian RBF to obtain all the simulation results i.e., $\psi_k(x) = \exp\left(-\frac{|x-e_k|^2}{2\sigma^2}\right)$. where $e_k$ is the center of the $k$th Gaussian RBF.

### A. Approximation of Convex Optimization Problem

Let $P_0 \in \mathbb{R}^{N \times N}$ be the finite-dimensional approximation of the P–F operator corresponding to uncontrolled dynamical system $\dot{x} = f(x)$. Similarly, let $P_1$ be the P–F operator of the system obtained with unit step input i.e., $\dot{x} = f(x) + g(x)$. These operators are obtained using NSDMD algorithm from Section II-C with time series data generated from the dynamical system with discretization time-step of $\Delta t$. The approximation of the P–F generator for the vector field $f$ is

$$P_{\ell} \approx \frac{1}{\Delta t}(P_0 - I) = : M_0.$$  \hspace{1cm} (33)

Using linearity property of the generator it follows that:

$$P_{g} = P_{\ell+g} - P_{\ell} \approx \frac{P_1 - P_0}{\Delta t} = : M_1.$$  \hspace{1cm} (34)

Let $h(x)$, $\rho(x)$, and $\rho_0(x)$ be expressed in terms of the basis

$$h(x) \approx \Psi^\top m, \quad \rho(x) \approx \Psi^\top v, \quad \rho_0(x) \approx \Psi^\top w.$$  \hspace{1cm} (35)

With the above approximation of the generators $P_{\ell}$ and $P_{g}$, and $\rho, \rho_0$, we can approximate the equality constraints in the optimization problem (22) as finite dimensional equality constraints

$$-\langle \Psi(x)^\top \begin{pmatrix} M_0 v + M_1 w \end{pmatrix} \rangle = \Psi(x)^\top m.$$

We now proceed with the approximation of the cost function

$$\int_{x_1} q(x)\rho(x)dx \approx \int_{x_1} q(x)\Psi^\top d\Psi = d^\top v$$

where the vector $d := \int_{X} q(x)\Psi dx$ can be precomputed. We use following approximation for the term:

$$\frac{\rho_0(x)}{\rho(x)} \approx \Psi(x)^\top \begin{pmatrix} v \\ w \end{pmatrix}$$

where the division is assumed element-wise. The above approximation is justified as we use Gaussian RBFs as basis functions in our simulation and apply following thump rule to select parameters for Gaussian RBFs, $d \leq 3\sigma \leq 1.5 d$, where $d$ is the distance between the Gaussian RBFs centers and $\sigma$ is the standard deviation. The choice of parameters ensure that in the region of intersection of two Gaussian RBFs, the function takes smaller values. With the above approximation, we have

$$\frac{\rho_0^2}{\rho} \approx w^\top \Psi^\top \begin{pmatrix} v \\ w \end{pmatrix}, \quad \int_{x_1} \frac{\rho_0^2}{\rho}dx = w^\top D w$$

where, $D = \int_{X} \Psi \Psi^\top dx$. We have the following approximation to the optimization problem (22):

$$\min_{\Psi \geq 0, w} \quad \begin{pmatrix} d^\top v + rw^\top D w \end{pmatrix}$$

s.t. $-\langle \Psi(x)^\top \begin{pmatrix} M_0 v + M_1 w \end{pmatrix} \rangle = \Psi(x)^\top m.$  \hspace{1cm} (36)

Since the basis functions are taken to be positive, (see Remark 7), the approximation for the $\rho$ and $\rho_0$ in (35) can be obtained by solving following finite-dimensional convex problem:

$$\min_{\Psi \geq 0, w} \quad \begin{pmatrix} d^\top v + rw^\top D w \end{pmatrix}$$

s.t. $-\langle \Psi(x)^\top \begin{pmatrix} M_0 v + M_1 w \end{pmatrix} \rangle = \Psi(x)^\top m.$  \hspace{1cm} (37)

The optimization problem is convex as the cost function is quadratic over linear and the constraints are linear in the decision variables. The optimal control is then approximated as $u = \Psi^\top \frac{x}{\bar{u}}$, where $\bar{u}$ is the decision variable. Similarly, the finite dimensional approximation of the OCP in Theorem 6 corresponding to $L_1$ norm on control is given by

$$\min_{\Psi \geq 0, w} \quad \begin{pmatrix} d^\top v + rc|w| \end{pmatrix}$$

s.t. $-\langle \Psi(x)^\top \begin{pmatrix} M_0 v + M_1 w \end{pmatrix} \rangle = \Psi(x)^\top m.$  \hspace{1cm} (37)

where $c = \int_{X} \psi(x)dx = \int_{X} \psi(x)dx$ is a positive constant.

### B. Computation of Local Optimal Controller

For the computation of the local optimal controller, we identify local linearized dynamics from data. To identify the linearized dynamics, we use time-series data generated by initializing the system around the origin. For the approximation of local linear dynamics, we use the EMD algorithm instead of NSDMD with basis function $\Psi(x) = x$, i.e., identity function. In particular, let $A$ and $b$ are the matrices for the
Comparison of the state trajectories and optimal feedback control obtained using analytical formula and the proposed data-driven P–F approach. (a) $x_1$ versus $t$. (b) Comparison of feedback control. (Local linear approximation of the system dynamics in discrete-time is obtained by solving least square problem:

$$\min_{A, b} \| Y - A X - b U \|$$

$$Y := [y_1, \ldots, y_L, y_{L+1}, \ldots, y_M], X = [x_1, \ldots, x_L, x_{L+1}, \ldots, x_M]$$

$$U = [0, \ldots, 0, 1, \ldots, 1]$$

$y_k = s_{\Delta t}(x_k, \bar{u} \in \{0, 1\})$ is the solution of (17) with either zero input or step input and initial condition $\{x_k\}$ initialized around the origin. So, part of the data is generated using zero input and remaining using step input. Once we have the local approximation of the system matrices, the local controller, $u_k$ is designed using the Lyapunov-based approach for stabilization or using the LQR based optimal control.

C. Data-Driven Nonlinear Stabilization

The data-driven stabilization of a nonlinear system will be the particular case of our data-driven optimal control system. As the constraint in the optimization problem guarantees stability following the results of Theorem 2. In particular, using the optimal control formulation in (22), the stabilization problem can be posed as a feasibility problem. Our proposed data-driven approach for stabilization will stand in contrast to the model-based approach for stabilization using density function as presented in [8].

D. Data-Driven Nonlinear Stabilization

The data-driven stabilization of a nonlinear system will be the particular case of our data-driven optimal control system. As the constraint in the optimization problem guarantees stability following the results of Theorem 2. In particular, using the optimal control formulation in (22), the stabilization problem can be posed as a feasibility problem. Our proposed data-driven approach for stabilization will stand in contrast to the model-based approach for stabilization using density function as presented in [8].
approximation of P–F operator, we applied NSDMD algorithm using one-step time-series data with 1e5 initial conditions, with Δt = 0.01. The comparison between the analytical derived optimal control and P–F based optimal control for the state trajectories shows a close match.

### B. Duffing Oscillator

\[
\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 - x_3^3 - 0.5x_2 + u. \tag{39}
\]

In this example, we used 225 Gaussian RBF with σ = 0.21, and the centers of basis functions are distributed uniformly within the range of D = [−3, 3] × [−3, 3], γ = 1000. The cost function is chosen to be \( x^T x + u^2 \). For the approximation of P–F operator, we applied NSDMD algorithm using one-step time-series data with 5e4 initial conditions. The region where the blending controller is active is marked by dotted ellipsoid around the origin in Fig. 2(b). Simulation results show that the optimal control is stabilizing the origin.

### C. Controlled 3-D Vanderpol Oscillator

\[
\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 + x_2 - x_3 - x_1^2 x_2 \\
\dot{x}_3 = x_3 - x_2^2 + 0.5 u. \tag{40}
\]

For this example, we are using 512 Gaussian RBF with σ = 0.14, and the centers of basis functions are distributed uniformly within the range of D = [−1, 1] × [−1, 1] × [−1, 1]. The cost function is chosen to be \( x^T x + u^2 \). For the approximation of P–F operator, we applied NSDMD algorithm using one-step time-series data with 5e4 initial conditions, Δt = 0.01. Fig. 3(a) shows the results for the optimal control of 3-D oscillator system with 10 initial conditions.

## VI. CONCLUSION

In this article, we have provided a convex optimization-based formulation for the infinite horizon OCP in the dual density space. We provided a data-driven approach for the computation of optimal control. The data-driven approach is based on the approximation of the P–F and Koopman operator for the finite-dimensional approximation of the convex optimization problem. Future research efforts will focus on the development of a computationally efficient numerical scheme and the choice of appropriate basis function for the implementation of the developed algorithm to system with large dimensional state space.

## ACKNOWLEDGMENT

The authors would like to acknowledge useful discussion with Prof. Yongxin Chen from Georgia Tech, Duvan Tellez-Castro from Universidad Nacional de Colombia, Bogotá, Colombia, and the anonymous reviewers who helped improve the quality and presentation of this article.

## REFERENCES

[1] H. J. Kappen, “An introduction to stochastic control theory, path integrals and reinforcement learning,” in Proc. AIP Conf. Conf., 2007, pp. 149–181.

[2] E. Todorov, “Efficient computation of optimal actions,” Proc. Nat. Acad. Sci., vol. 106, pp. 11 478–11 483, 2009.

[3] E. Theodorou, J. Buchli, and S. Schaal, “A generalized path integral control approach to reinforcement learning,” J. Mach. Learn. Res., vol. 11, pp. 3137–3181, 2010.

[4] D. Hernández-Hernández, O. Hernández-Lerma, and M. Takasar, “The linear programming approach to deterministic optimal control problem,” Applicationes Mathematicae, vol. 24, no. 1, pp. 17–33, 1996.

[5] V. Gaitsgory and M. Quincampoix, “Linear programming approach to deterministic infinite horizon optimal control problems with discounting,” SIAM J. Control Optim., vol. 48, no. 4, pp. 2480–2512, 2009.

[6] R. Vinter, “Convex duality and nonlinear optimal control,” SIAM J. Control Optim., vol. 31, no. 2, pp. 518–538, 1993.

[7] A. Rantzer, “A dual to Lyapunov’s stability theorem,” Syst. Control Lett., vol. 42, pp. 161–168, 2001.

[8] S. Prajna, P. A. Parrilo, and A. Rantzer, “Nonlinear control synthesis by convex optimization,” IEEE Trans. Autom. Control, vol. 49, no. 2, pp. 310–314, Feb. 2004.

[9] A. Rantzer and S. Hedlund, “Duality between cost and density in optimal control,” in Proc. 42nd IEEE Conf. Decis. Control, 2003, pp. 1218–1221.

[10] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” IEEE Trans. Autom. Control, vol. 59, no. 2, pp. 297–312, Feb. 2014.

[11] M. Korda, “Moment-sum-of-squares hierarchies for set approximation and optimal control,” IGM, Lausanne, Switzerland, Tech. Rep., 2016.

[12] M. Korda, D. Henrion, and C. N. Jones, “Convergence rates of moment-sum-of-squares hierarchies for optimal control problems,” Syst. Control Lett., vol. 100, pp. 1–5, 2017.

[13] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trelat, “Nonlinear optimal control via occupation measures and LMI-relaxations,” SIAM J. Control Optim., vol. 47, no. 4, pp. 1643–1666, 2008.

[14] R. Rajaram, U. Vaidya, M. Fardad, and B. Ganapathysubramanian, “Stability in the almost everywhere sense: A linear transfer operator approach,” J. Math. Anal. Appl., vol. 368, no. 1, pp. 144–156, 2010.

[15] U. Vaidya and P. G. Mehta, “Lyapunov measure for almost everywhere stability,” IEEE Trans. Autom. Control, vol. 53, no. 1, pp. 307–325, Feb. 2008.

[16] U. Vaidya and B. Huang, “A convex approach to stochastic optimal control using linear operator,” in Proc. Amer. Control Conf., 2021, pp. 1306–1311.

[17] E. Kaiser, J. N. Kutz, and S. L. Brunton, “Data-driven discovery of Koopman eigenfunctions for control,” Bull. Amer. Phys. Soc., vol. 62, pp. 1–31, 2017.

[18] B. Huang, X. Ma, and U. Vaidya, “Feedback stabilization using Koopman operator,” in Proc. IEEE Conf. Decis. Control, 2018, pp. 6434–6439.
[19] H. Arbabi, M. Korda, and I. Mezić, “A data-driven Koopman model predictive control framework for nonlinear partial differential equations,” in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 6409–6414.

[20] X. Ma, B. Huang, and U. Vaidya, “Optimal quadratic regulation of nonlinear system using Koopman operator,” in *Proc. Amer. Control Conf.*, 2019, pp. 4911–4916.

[21] M. Korda and I. Mezić, “Optimal construction of Koopman eigenfunctions for prediction and control,” *IEEE Trans. Autom. Control*, vol. 65, no. 12, pp. 5114–5129, Dec. 2020.

[22] A. Mauroy and I. Mezic, “A spectral operator-theoretic framework for global stability,” in *Proc. IEEE Conf. Decis. Control*, 2013, pp. 5234–5239.

[23] B. Huang, X. Ma, and U. Vaidya, “Data-driven nonlinear stabilization using Koopman operator,” in *The Koopman Operator in Systems and Control*. Berlin, Germany: Springer, 2020, pp. 313–334.

[24] U. Vaidya, P. Mehta, and U. Shanbhag, “Nonlinear stabilization via control Lyapunov measure,” *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1314–1328, Jun. 2010.

[25] A. Raghunathan and U. Vaidya, “Optimal stabilization using Lyapunov measures,” *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1316–1321, May 2014.

[26] B. Huang and U. Vaidya, “Data-driven approximation of transfer operators: Naturally structured dynamic mode decomposition,” in *Proc. Amer. Control Conf.*, 2018, pp. 5659–5664.

[27] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. New York, NY, USA: Springer, 1994.

[28] U. Vaidya, “Converse theorem for almost everywhere stability using Lyapunov measure,” in *Proc. Amer. Control Conf.*, 2007, pp. 4835–4840.

[29] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, “A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,” *J. Nonlinear Sci.*, vol. 25, no. 6, pp. 1307–1346, 2015.

[30] M. Korda and I. Mezić, “On convergence of extended dynamic mode decomposition to the Koopman operator,” *J. Nonlinear Sci.*, vol. 28, no. 2, pp. 687–710, 2018.

[31] S. Klus, I. Schuster, and K. Muandet, “Eigendecompositions of transfer operators in reproducing kernel Hilbert spaces,” *J. Nonlinear Sci.*, vol. 30, no. 1, pp. 283–315, 2020.

[32] A. Rantzer and F. Ceragioli, “Smooth blending of nonlinear controllers using density functions,” in *Proc. Eur. Control Conf.*, 2001, pp. 2851–2853.

[33] I. Barbalat, “Systemes d'équations différentielles d'oscillations non linéaires,” *Rev. Math. Pures Appl.*, vol. 4, no. 2, pp. 267–270, 1959.