Learning Regions of Attraction in Unknown Dynamical Systems via Zubov–Koopman Lifting: Regularities and Convergence

Yiming Meng, Ruikun Zhou, and Jun Liu, *Senior Member, IEEE*

***Abstract*—The estimation for the region of attraction of an asymptotically stable equilibrium point is crucial in the analysis of nonlinear systems. There has been a recent surge of interest in estimating the solution to Zubov's equation, whose nontrivial sublevel sets form the exact ROA. In this article, we propose a lifting approach to map observable data into an infinite-dimensional function space, which generates a flow governed by the proposed *Zubov–Koopman* operators. By learning a Zubov–Koopman operator over a fixed time interval, we can indirectly approximate the solution to Zubov's equation by iteratively applying the learned operator on certain functions. We also demonstrate that a transformation of such an approximator can be readily utilized as a near-maximal Lyapunov function. We approach our goal through a comprehensive investigation of the regularities of Zubov–Koopman operators and their associated quantities. Based on these findings, we present an algorithm for learning Zubov–Koopman operators that exhibit strong convergence to the true operator. We show that this approach reduces the amount of required data and can yield desirable estimation results, as demonstrated through numerical examples.**

***Index Terms*—Region of attraction (ROA), regularity analysis, unknown nonlinear systems, viscosity solution, Zubov's equation, Zubov–Koopman operators.**

I. Introduction

AN IMPORTANT aspect in dynamical systems is the determination of the region of attraction (ROA) for an equilibrium point. This is of significant importance in safety-critical industries, such as aviation, robotics, and power systems, where comprehending operational boundaries is crucial.

Estimating the ROA for nonlinear systems can be rigorously resolved using formal methods [20]. This is achieved through space discretization, formulating an "inclusion" of system transitions that serves as a symbolic abstraction, and using this surrogate abstraction to conservatively identify those states that will always reach an equilibrium point (which is the definition of the ROA). Formal methods require rigorous analysis to ensure the precision of the estimation, and face severe computational complexity arising from state space discretization. They are also not prepared for predicting the ROA of systems with limited knowledge.

This issue can be resolved by Lyapunov methods. Lyapunov functions qualitatively characterize stability properties for various nonlinear systems, and a forward-invariant sublevel set within their domains serves as an estimate of the ROA. The existence of Lyapunov functions is guaranteed by converse Lyapunov theorems [4], [21], [40]. The primary challenge lies in the construction of Lyapunov functions that possibly enhance the conservative estimation of ROAs for more general nonlinear systems. On the other hand, Zubov's theorem characterizes the maximal Lyapunov function [42] defined on the domain of attraction. However, it requires solving a partial differential equation (PDE) [13].

In real-world applications, limited knowledge of the system dynamics can make the identification of ROAs using Lyapunov methods even more challenging. Inspired by recent advances in Koopman operator-based system identification for unknown dynamical systems, we propose a *Zubov–Koopman* lifting approach to estimate the ROA by approximating the solutions to Zubov's equations, whose nontrivial sublevel sets form the exact ROA. We aim to: 1) study the regularities of Zubov's equation and introduce a linear Zubov–Koopman operator to characterize the solution; 2) configure the trajectory data and learn the Zubov-Koopman operator; 3) approximate the entire ROA for an asymptotically stable equilibrium point using the learned Zubov–Koopman operator; 4) approximate near-maximal Lyapunov functions and provide formal verification as a byproduct.

We review some crucial results from the literature that are pertinent to the work presented in this article.

A. Related Work

The computation of Lyapunov functions has a long history [11] and has gained increased attention with respect to data-driven methods [6]. In particular, the development of the Koopman operator theory provides a promising alternative learning approach for nonlinear system identification and Lyapunov function constructions [17], [35], [36].

In essence, Koopman operators simplify the nonlinear analysis by lifting the states of the system into the space of observable

Received 26 June 2024; revised 10 February 2025; accepted 10 April 2025. Date of publication 15 April 2025; date of current version 29 September 2025. This work was supported in part by the NSERC Discover Grant, an Ontario Early Researcher Award, and the Canada Research Chairs program, and in part by the Digital Research Alliance of Canada (alliance.ca). Recommended by Associate Editor L. Bako. *(Corresponding author: Yiming Meng.)*

Yiming Meng is with the Coordinated Science Laboratory, University of Illinois Urbana-Champaign, Champaign, IL 61820 USA (e-mail: ymmeng@illinois.edu, yiming.meng@uwaterloo.ca).

Ruikun Zhou and Jun Liu are with the Department of Applied Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: ruikun.zhou@uwaterloo.ca; j.liu@uwaterloo.ca).

Digital Object Identifier 10.1109/TAC.2025.3560653

functions, which evolve linearly governed by Koopman operators. The spectral properties of linear Koopman operators can facilitate Koopman mode decomposition for nonlinear systems using a specific set of observable functions. Several established techniques, such as dynamic mode decomposition [39] and extended dynamic mode decomposition (EDMD) [43], can be employed to acquire this linear representation. In addition, the autoencoder architecture can be considered as suitable neural-network (NN) observable functions, although it requires more training efforts to achieve an optimized linear representation [7].

Taking advantage of the spectral representation, in [25] and [26], the authors showed that a set of Lyapunov functions for nonlinear systems with global stability can be constructed based on the eigenfunctions of the learned Koopman operator. Deka et al. [7] improved the multistep trajectory prediction accuracy by a modified autoencoder architecture, with the corresponding Koopman eigenfunctions parameterizing a set of Lyapunov function candidates. In this framework, compared to the non-Koopman neural Lyapunov framework [47], it becomes possible to achieve a more desirable ROA estimation by exploring various valid combinations within the set of resulting Lyapunov functions.

The Koopman method in [7], [25], [26], [44], and [46] requires working on a forward-invariant compact subset within the state space for stability analysis. This allows the Koopman operators to preserve the associated function space on the invariant set. However, it sacrifices the ability to construct Lyapunov functions on a larger scale, given their unbounded nature near the boundary of the open ROA.

Considering this limitation, Zubov's construction of a Lyapunov function seems to address the issue by ensuring that the solution is always bounded. This approach can offer potential advantages in numerical approximations, particularly when solving a PDE to find a Lyapunov function. This property enables the extension of the function domain to the entire state space or any desired set where computations occur [22].

In this regard, recent investigations have focused on numerical solutions to Zubov's equation. In cases where system dynamics are known, the approach in [12] utilized local exponential stability conditions to train NN, similar in nature to physics-informed neural networks (PINNs) [38]. The most recent work [22], [23] introduced PINN algorithms for computing Lyapunov functions capable of approximating the entire ROA for an asymptotically stable compact set with high accuracy. The Lyapunov function candidate generated by the NN was also formally verified. The work in [15] employed a purely data-driven approach for estimating the solution to Zubov's PDE. This method, however, requires long-term observation of trajectory data, and may have limited predictability of ROAs when the observation time is restricted.

B. Contributions

Considering the pros and cons of the Koopman approach and Zubov's PDE, we propose a Zubov-Koopman operator-based data-driven technique, leveraging trajectory data within a fixed observation span to predict the ROA of a known asymptotically stable compact set or equilibrium point for unknown dynamics. In brief, we have following.

1) We conduct a regularity analysis for Zubov's equation, where we relax the assumption from continuously differentiable vector fields to local Lipschitz continuity. In addition, we explore how the solution to Zubov's equation connects with a more general notion of Lyapunov functions.
2) We introduce a Zubov–Koopman operator and establish its connections with Zubov's equation. Building upon the regularity analysis mentioned earlier, we provide a detailed proof demonstrating how a Zubov–Koopman operator can lead to convergence to the solution of Zubov's equation in a more general sense.
3) While it is often assumed that finite-rank approximations apply to Koopman-like operators, we rigorously investigate the theoretical feasibility of finite-dimensional approximations for Zubov–Koopman operators and demonstrate how they can guide the selection of observable functions. Notably, the existing theorems on the existence of Koopman eigenfunctions are based either on classical linearization theorems [19], [25, Sect. IV], which may not be suitable for locally Lipschitz continuous vector fields, or are complicated to verify [18, Prop. 6]. The spectral analysis in this article will focus on the specific goal of approximating the solution to Zubov's equation under mild conditions.
4) We introduce a learning algorithm for the Zubov–Koopman operator and employ the learned operator to approximate the solution to Zubov's equation, thereby estimating the ROA. We show that extra efforts are required to obtain a near-maximal Lyapunov function.
5) We provide numerical experiments and demonstrate the effectiveness of the proposed approach.

The rest of this article is organized as follows. Section II presents some preliminaries on the Koopman operator, Zubov's equation, and concepts for analyzing its solution regularity. Section III focuses on conducting regularity analysis under mild system conditions. Building on this, in Section IV, we introduce the Zubov–Koopman operator and demonstrate how the Zubov–Koopman operator can effectively characterize the solution to Zubov's equation. Section V establishes the theoretical feasibility of a finite-dimensional representation of the Zubov–Koopman operator and demonstrates its interplay with observable data. Section VI introduces the algorithms, with numerical experiments conducted in Section VII. Finally, Section VIII concludes this article.

C. Notations

We denote by $\mathbb{R}^n$ the Euclidean space of dimension $n > 1$, and by $\mathbb{R}$ the set of real numbers. For $x \in \mathbb{R}^n$ and $r \geq 0$, we denote the ball of radius r centered at x by $\mathcal{B}(x, r) = \{y \in \mathbb{R}^n : |y - x| \leq r\}$, where $|\cdot|$ is the Euclidean norm. For a closed set $A \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, we denote the distance from x to A by $|x|_A = \inf_{y \in A} |x - y|$ and r-neighborhood of A by $\mathcal{B}(A, r) = \cup_{x \in A} \mathcal{B}(x, r) = \{x \in \mathbb{R}^n : |x|_A \leq r\}$. For a set $A \subseteq \mathbb{R}^n$, $\overline{A}$ denotes its closure, $\mathrm{int}(A)$ denotes its interior, and ∂A denotes its boundary. For two sets $A, B \subseteq \mathbb{R}^n$, the set difference is defined by $A \setminus B = \{x : x \in A, x \notin B\}$. For finite-dimensional matrices, we use the Frobenius norm $\|\cdot\|_F$ as the metric. Given $a, b \in \mathbb{R}$, we define $a \wedge b := \min(a, b)$. Let $C(\Omega)$ be the set of continuous functions and $C_b(\Omega)$ be the set of bounded continuous functions with domain Ω. We denote the set of ith continuously differentiable functions by $C^i(\Omega)$, and similarly, bounded continuously differentiable functions

by $C_b^i(\Omega)$. We denote the set of locally and uniformly Lipschitz continuous functions by $\boldsymbol{LocLip}(\Omega)$ and $\boldsymbol{Lip}(\Omega)$. When making general statements for $v \in C^1(\mathbb{R}^n)$ with $n \geq 1$, we denote ∇ as its gradient (or as the derivative when $n = 1$).

II. Preliminaries

A. Dynamical Systems

Given a state space $\mathcal{X} \subseteq \mathbb{R}^n$, we consider a continuous-time nonlinear dynamical system of the form

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t)), \quad \mathbf{x}(0) = x \in \mathcal{X}, \ t \in [0, \infty) \tag{1}$$

where x denotes the initial condition, and the vector field $f : \mathcal{X} \to \mathcal{X}$ is assumed to be locally Lipschitz.

On the maximal interval of existence $\mathcal{I} \subseteq [0, \infty)$, the forward flow map (solution map) $\phi : \mathcal{I} \times \mathcal{X} \to \mathcal{X}$ should satisfy 1) $\partial_t(\phi(t, x)) = f(\phi(t, x))$, 2) $\phi(0, x) = x$, and 3) $\phi(s, \phi(t, x)) = \phi(t + s, x)$ for all $t, s \in \mathcal{I}$.

Without loss of generality, throughout this article, we will assume that the maximal interval of existence of the (unique) flow map to the initial value problem (1) is $\mathcal{I} = [0, \infty)$. We also generally consider that the state space $\mathcal{X} = \mathbb{R}^n$, i.e., the flow is not necessarily assumed to be invariant within a strict subset of $\mathbb{R}^n$.

To define Koopman operators, which govern the evolution of observable functions, let us now consider a complete function space $\mathcal{F}$ of the observable real-valued functions $h : \mathcal{X} \to \mathbb{R}$.

Definition 1 (Koopman operator): The Koopman operator family $\{\mathcal{K}_t\}_{t \geq 0}$ of system (1) is a collection of maps $\mathcal{K}_t : \mathcal{F} \to \mathcal{F}$ defined by

$$\mathcal{K}_t h = h \circ \phi(t, \cdot), \quad h \in \mathcal{F} \tag{2}$$

for each $t \geq 0$, where $\circ$ is the composition operator. The (infinitesimal) generator $\mathcal{L}_f$ of $\{\mathcal{K}_t\}_{t \geq 0}$ is defined by $\mathcal{L}_f h(x) := \lim_{t \to 0} \frac{\mathcal{K}_t h(x) - h(x)}{t}$, where the observable functions should be within the domain of $\mathcal{L}_f$, i.e., $\mathrm{dom}(\mathcal{L}_f) = \{h \in \mathcal{F} : \lim_{t \to 0} \frac{\mathcal{K}_t h(x) - h(x)}{t} \text{ exists}\}$.

Suppose that the observable functions are bounded and continuously differentiable, the generator is such that $\mathcal{L}_f h = \nabla h \cdot f$ for $h \in C_b^1(\mathcal{X})$.

Koopman operators form a linear C_0-semigroup that satisfies the following criteria. They allow us to study the nonlinear dynamics through the infinite-dimensional lifted space of observable functions with linear dynamics.

Definition 2 (Semigroup): A one-parameter family $\{\mathcal{S}_t\}_{t \geq 0}$, of bounded linear operators from $\mathcal{F}$ into $\mathcal{F}$ is a semigroup of bounded linear operators on $\mathcal{F}$ if

1) $\mathcal{S}_0 =$ identity operator.
2) $\mathcal{S}_t \circ \mathcal{S}_s = \mathcal{S}_{t+s}$ for every $t, s \geq 0$.

In addition, a semigroup $\{\mathcal{S}_t\}_{t \geq 0}$ is a strongly continuous semigroup, or C_0-semigroup, if $\lim_{t \downarrow 0} \mathcal{S}_t h = h$ for all $h \in \mathcal{F}$.

B. Concept of Stability

We are interested in systems of the form (1) with an intrinsic asymptotically stable set $\mathcal{A} \subseteq \mathcal{X}$. We define the set stability as follows.

Definition 3 (Set stability): A closed invariant set $\mathcal{A} \subseteq \mathcal{X}$ is said to be asymptotically stable for (1) if

1) for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|x|_{\mathcal{A}} < \delta$ implies $|\phi(t, x)|_{\mathcal{A}} < \varepsilon$ for all $t \geq 0$, and
2) there exists a $\delta > 0$ such that $|x|_{\mathcal{A}} < \delta$ implies $\lim_{t \to \infty} |\phi(t, x)|_{\mathcal{A}} = 0$.

Furthermore, $\mathcal{A}$ is said to be locally exponentially stable, if there exists a $\delta > 0, M > 0$ and $c > 0$ such that $|\phi(t, x)|_{\mathcal{A}} \leq M|x|_{\mathcal{A}} e^{-ct}$, for all $t \geq 0$ and $x \in \{x \in \mathcal{X} : |x|_{\mathcal{A}} \leq \delta\}$.

We further define the ROA of $\mathcal{A}$ given its asymptotic stability, which quantifies a region of the state space from which each starts and eventually converges to the attractor itself.

Definition 4 (ROA): Suppose that $\mathcal{A}$ is asymptotically stable, the ROA of $\mathcal{A}$ is a set defined as $\mathcal{D}(\mathcal{A}) := \{x \in \mathcal{X} : \lim_{t \to \infty} |\phi(t, x)|_{\mathcal{A}} = 0\}$.

Remark 5: It is a well-known result that the ROA is an open and forward invariant set. ◇

To better convey the idea of this article, we propose the following hypothesis for unknown systems in the form of (1).

H1) We assume that there exists an equilibrium point x_{eq} of (1), i.e., a point such that $f(x_{\mathrm{eq}}) = \mathbf{0}$.
H2) We assume full knowledge of x_{eq}, and that $\{x_{\mathrm{eq}}\}$ is locally exponentially stable.

Based on Hypotheses (H1) and (H2), the purpose of this article is to employ a data-driven approach to estimate the ROA of an equilibrium point of unknown systems.

C. Zubov's Theorem

The ROA can be characterized by a maximal Lyapunov function as described in the following theorem [42].

Theorem 6: Let $D \subseteq \mathbb{R}^n$ be an open set. Suppose that there exists a function $V \in C^1(D)$ such that $\mathcal{A} \subseteq D$ and the following conditions hold: 1) V is positive definite on D with respect to $\mathcal{A}$, i.e., $V(x) = 0$ for all $x \in \mathcal{A}$ and $V(x) > 0$ for all $x \in D \setminus \mathcal{A}$; 2) the derivative of V along solutions of (1) is well-defined for all $x \in D$ and satisfies

$$\nabla V(x) \cdot f(x) = -q(x) \tag{3}$$

where the function $q : \mathbb{R}^n \to \mathbb{R}$ is continuous and positive definite with respect to $\mathcal{A}$; 3) $V(x) \to \infty$ as $x \to \partial D$ or $|x| \to \infty$. Then, $D = \mathcal{D}(\mathcal{A})$.

Theorem 6 is equivalent to Zubov's theorem [49], stated below, where $\eta(x)$ plays a similar role as the $q(x)$ in (3) above, and W represents the solution to the Zubov's PDE given the choice of η.

Theorem 7: Let $D \subset \mathbb{R}^n$ be an open set containing $\mathcal{A}$. Then, $D = \mathcal{D}(\mathcal{A})$ if and only if there exists two continuous functions $W : D \to \mathbb{R}$ and $\eta : \mathbb{R}^n \to \mathbb{R}$ such that the following conditions hold: 1) $0 < W(x) < 1$ for all $x \in D \setminus \mathcal{A}$ and $W(x) = 0$ for all $x \in \mathcal{A}$; 2) η is positive definite with respect to $\mathcal{A}$; 3) for any sufficiently small $c_3 > 0$, there exist $c_1, c_2 > 0$ such that $|x|_{\mathcal{A}} \geq c_3$ implies $W(x) > c_1$ and $\eta(x) > c_2$; 4) $W(x) \to 1$ as $x \to y$ for any $y \in \partial D$; 5) W and η satisfy (in the conventional differentiable sense)

$$\nabla W(x) \cdot f(x) + \eta(x)(1 - W(x)) = 0. \tag{4}$$

Remark 8: Based on Hypothesis (H1) and (H2), a typical choice for η is given by $\eta(x) = \alpha|x - x_{\mathrm{eq}}|$ for some $\alpha > 0$. In addition, Theorems 6 and 7 can be related by the following equation:

$$W(x) = 1 - \exp(-\alpha V(x)), \quad x \in \mathcal{D}(\mathcal{A}) \tag{5}$$

for some constant $\alpha > 0$ [22, Remark 1]. It is easy to verify that V satisfying (3) implies $\nabla W(x) \cdot f(x) = -\alpha(1 - W(x))q(x)$, which verifies (4) with $\eta(x) = \alpha q(x)$. ◇

Note that W remains bounded, with its value approaching 1 as x approaches the boundary. In contrast, the function V in Theorem 6 approaches infinity as x approaches the boundary. The boundedness can offer significant advantages in data-driven approximations. In particular, it allows us to focus on a desired region of interest where computations occur. However, this modification is not as straightforward as it might appear. To better understand how Zubov's construction can guide the data-driven approach to learning the ROA, we introduce an extension of Zubov's theorem in Section III and rigorously examine the regularity of the solutions.

D. Differentiability and Viscosity Solution

As we will see in Section III, the Zubov equation may not always possess differentiable solutions. One may need to consider a more general sense of solution to resolve.

Definition 9 (Viscosity solution): Define the superdifferential and the subdifferential sets of $v \in C(\mathbb{R}^n)$ at x, respectively, as

$$\partial^+ v(x) = \left\{ p \in \mathbb{R}^n : \limsup_{y \to x} \frac{v(y) - v(x) - p \cdot (y - x)}{|y - x|} \leq 0 \right\} \tag{6a}$$

$$\partial^- v(x) = \left\{ q \in \mathbb{R}^n : \liminf_{y \to x} \frac{v(y) - v(x) - q \cdot (y - x)}{|y - x|} \geq 0 \right\}. \tag{6b}$$

A continuous function v of a PDE of the form $F(x, v(x), \nabla v(x)) = 0$ (possibly encoded with boundary conditions) is a viscosity solution if the following conditions are satisfied.

1) (Viscosity subsolution) $F(x, v(x), p) \leq 0$ for all $x \in \mathbb{R}^n$ and for all $p \in \partial^+ v(x)$.
2) (Viscosity supersolution) $F(x, v(x), q) \geq 0$ for all $x \in \mathbb{R}^n$ and for all $q \in \partial^- v(x)$.

Remark 10: More details of viscosity solutions are provided in Appendix A. The concept of viscosity solution relaxes C^1 solutions. Note that, at differentiable points, $\nabla v(x)$ exists and $\{\nabla v(x)\} = \partial^+ v(x) = \partial^- v(x)$. In this case, to justify a viscosity solution, we can simply substitute $\nabla v(x)$ and check if $F(x, v(x), \nabla v(x)) = 0$. If v is not differentiable at a given point, then we have to go through the comparisons (1) and (2) in Definition 9 to verify that v is a viscosity solution. ◇

III. Extension of Zubov's Theorem and Regularity Analysis

For the rest of this article, we will assume that Hypothesis (H1) and (H2) hold and denote $\mathcal{A} := \{x_{\text{eq}}\}$.

Inspired by (4) and its connection with V through (5), to facilitate data-driven approximation, we explore the following dual form of Zubov's (4), namely Zubov's dual equation

$$H(x, U(x), \nabla U(x)) = 0, \; U(x_{\text{eq}}) = 1 \tag{7}$$

where

$$H(x, U(x), p) := -p \cdot f(x) + \eta(x) U(x), \quad p \in \mathbb{R}^n. \tag{8}$$

For now, we intentionally omit specifying the domain of U and would like to discuss regularity analysis later in this section.

Note that, on $\mathcal{D}(\mathcal{A})$, we always have $U(x) = 1 - W(x)$. We prefer this dual form of Zubov's equation because it allows us to potentially use a time series to approximate the solution [as in (10) below] through iterations. This approach avoids the need to evaluate (10) using trajectory data accumulated up to an extremely large time [15], thus reducing the amount of required data, as demonstrated in Section IV.

A. Solution to Zubov's Dual Equation

In this section, we construct the solution to the dual (7) of Zubov's equation and demonstrate that, in a general context, it satisfies (7) in a viscosity sense.

Let η be positive definite w.r.t. $\mathcal{A}$ and satisfy that for any $\delta > 0$, there exists $c > 0$ such that $\eta(x) > c$ for all $|x|_{\mathcal{A}} > \delta$. Define

$$V(x) = \int_0^\infty \eta(\phi(t, x)) dt, \quad x \in \mathbb{R}^n \tag{9}$$

where if the integral diverges, we let $V(x) = \infty$. Then, we have the following property [22, Prop. 1].

Lemma 11: The function $V : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ defined by (9) satisfies the following.

1) $V(x) < \infty$ if and only if $x \in \mathcal{D}(\mathcal{A})$.
2) $V(x) \to \infty$ as $x \to \partial\mathcal{D}(\mathcal{A})$.
3) V is positive definite with respect to $\mathcal{A}$.

Definition 12: Let $h \in C_b(\mathbb{R}^n)$. We further define

$$U_h(x) = \begin{cases} \exp\{-V(x)\} h(\phi_\infty(x)), & \text{if } V(x) < \infty \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where $\phi_\infty(x) := \lim_{t \to \infty} \phi(t, x)$. For test function $h(x) = \mathbb{1}(x)$, where $\mathbb{1}(x) \equiv 1$ for all x, we denote $U := U_{\mathbb{1}}$ for simplicity.

Remark 13: The notion of $h(\phi_\infty(x))$ seems redundant in that, given $V(x) < \infty$, by 1) of Lemma 11, we have $x \in \mathcal{D}(\mathcal{A})$ and $h(\phi_\infty(x)) = h \circ \lim_{t \to \infty} \phi(t, x) = h(x_{\text{eq}})$. However, we still keep it in (10) for consistency when using a time series to approximate in Section IV. ◇

We would like to show that the construction in (10) is a solution to (7). Getting the conventional continuously differentiable solutions to (7) or (4) depends on the differentiability of f (and hence the differentiability of $\phi(t, x)$ w.r.t. x) and η. However, this may not always be the case given the general assumptions on f and η. In view of numerical approximations, it is not appropriate to construct a convergent sequence with respect to the C^1-uniform norm (e.g., [9, Ch. 3]). Based on this viscosity property analysis, this article provides guidance for constructing numerical solutions through a Zubov–Koopman operator learning framework in subsequent sections. We provide a simple example below to demonstrate that we necessarily need to consider viscosity solutions to (7).

Example 14: Consider a simple dynamical system $\dot{\mathbf{x}}(t) = -\mathbf{x}(t) + \mathbf{x}^3(t)$, $\mathbf{x}(0) = x$, and a convex, positive definite function $\eta(x) = |x|$. Then, it can be verified that $V(x) = \frac{1}{2}\log(\frac{1+|x|}{1-|x|})$ for $x \in (-1, 1)$, $U(x) = \sqrt{\frac{1-|x|}{1+|x|}}$ for $x \in (-1, 1)$, and $U(x) = 0$ elsewhere. Clearly, $(x - x^3)U'(x) + |x|U(x) = 0$ for all $x \in \mathbb{R} \setminus \{0, \pm 1\}$ and U is not differentiable at 0 and ± 1. Given the fact that U is concave near 0, we also have that $\partial^- U(0) = \emptyset$ by [1, Chapter II, Prop. 4.7]. On the other hand, $\partial^+ U(0) = [-1, 1]$. For all $p \in \partial^+ U(0)$, we have that

$H(0,U(0),p)\le 0$, which show that U satisfies (7) at 0 in a viscosity sense. Similar proofs can be done at $x=\pm 1$. ◇

Lemma 15: Let $F(x,p):=-p\cdot f(x)-\eta(x)$. Then, on $\mathcal{D}(\mathcal{A})$, the function $V(x)$ in (9) is a viscosity solution to $F(x,\nabla V(x))=0$ with $V(x_{\rm eq})=0$.

Proof: It is clear that $V\in C(\mathbb{R}^n)$. On $\mathcal{D}(\mathcal{A})$, by Lemma 11 and [22, Prop. 1], we have $V(x)<\infty$ as well as the dynamic programming principle $V(x)=\int_0^t \eta(\phi(s,x))ds+V(\phi(t,x))$ for all $x\in\mathcal{D}(\mathcal{A})$ and $t>0$. We now show that V is a viscosity solution on $\mathcal{D}(\mathcal{A})$ using the equivalent conditions introduced in Appendix A. Let $\nu\in C^1$ and x be a local maximum point of $V-\nu$. Then, $V(x)-V(z)\ge \nu(x)-\nu(z)$ for all $z\in\mathcal{B}(x,r)$. For t sufficiently small, we have $\phi(t,x)\in\mathcal{B}(x,r)$. Therefore

$$\begin{aligned}&\nu(x)-\nu(\phi(t,x))\le V(x)-V(\phi(t,x))\\ &=\int_0^t \eta(\phi(s,x))ds+V(\phi(t,x))-V(\phi(t,x)).\end{aligned}\tag{11}$$

Considering the infinitesimal behavior on both sides of (11), i.e., dividing both sides by t and taking the limit as $t\downarrow 0$, we have $-\nabla\nu(x)\cdot f(x)\le\eta(x)$. The case when x is a local minimum of $V-\nu$ can be proved in the same manner. Then, V is a viscosity solution to $F(x,\nabla V(x))=0$ on $\mathcal{D}(\mathcal{A})$. ■

Remark 16: From the above proof, one may notice that the sign of $F(x,p)$ [or similarly, $H(x,U(x),p)$ as in (12)] matters. In other words, V is not a viscosity solution to $-F(x,\nabla V(x))=0$. ◇

Theorem 17 (Viscosity solution): For any $h\in C_b(\mathbb{R}^n)$, the U_h defined in (10) is a viscosity solution to

$$H(x,U_h(x),\nabla U_h(x))=0,\quad U_h(x_{\rm eq})=h(x_{\rm eq})\tag{12}$$

on $\mathbb{R}^n$, where H is defined in (8).

Proof: It is clear that $U_h\in C(\mathbb{R}^n)$ and $U_h(x_{\rm eq})=\exp\{-0\}h(x_{\rm eq})=h(x_{\rm eq})$. For h with $h(x_{\rm eq})=0$, $U_h\equiv 0$ is automatically a differentiable and, hence, a viscosity solution to (12). We verify the case when $h(x_{\rm eq})\neq 0$.

On $\mathcal{D}(\mathcal{A})$, since $V(x)<\infty$ by Lemma 11, the inverse function $(-\log(U_h(x)/h(x_{\rm eq})))$ of $U_h(x)=\exp(-V(x))h(x_{\rm eq})$ is well defined. Since $-V$ is a viscosity solution to $-F(x,\nabla V)=0$, where F is defined in Lemma 15, by [1, Ch. 2, Prop. 2.5], we immediately have that U_h is a viscosity solution of $-F(x,-\log'(U_h(x)/h(x_{\rm eq}))\nabla U_h(x))=-\frac{1}{U_h(x)}\nabla U_h(x)\cdot f(x)+\eta(x)=0$. This implies that U_h is a viscosity solution of (12) in $\mathcal{D}(\mathcal{A})$. For $x\in\mathbb{R}^n\setminus\overline{\mathcal{D}(\mathcal{A})}$, $u(x)\equiv 0$ is always a viscosity solution. It suffices to show the viscosity property on $\partial\mathcal{D}(\mathcal{A})$. Note that U_h may not be differentiable on $\partial\mathcal{D}(\mathcal{A})$. However, by the famous Rademacher's Theorem [8], U_h is differentiable a.e. on each subdomain $\Omega\subseteq\mathbb{R}^n\setminus\mathcal{D}(\mathcal{A})$. Equivalently, U_h satisfies

$$\eta(x)U_h(x)+H(x,\nabla U_h(x))=0,\quad \text{a.e. in } \mathbb{R}^n\setminus\mathcal{D}(\mathcal{A})\tag{13}$$

in the conventional sense, where $H(x,p)=-p\cdot f(x)$. Now we introduce a smooth mollifier $\rho\in C_0^\infty(\mathbb{R}^n)$ with compact support $\mathcal{B}(\mathbf{0},1)$ such that $\int_{\mathbb{R}^n}\rho(y)dy=1$, as well as a kernel $\rho_\varepsilon(x)=\frac{1}{\varepsilon^n}\rho(\frac{x}{\varepsilon})$ for $\varepsilon>0$. Then, we define the convolution for u_h by $u_h^\varepsilon(x):=(U_h*\rho_\varepsilon)(x):=\int_{\mathbb{R}^n}U_h(y)\rho_\varepsilon(x-y)dy$, and similarly for $\nabla u_h^\varepsilon(x)=(\nabla U_h*\rho_\varepsilon)(x)$. It is clear that $u_h^\varepsilon\to U_h$ locally uniformly in $\mathbb{R}^n$. Convolve both side of (13) with ρ_ε, then

$$\begin{aligned}&\eta(x)u_h^\varepsilon(x)+\int_{\mathbb{R}^n}H(y,\nabla U_h(y)\rho_\varepsilon(x-y)dy\\ &=\int_{\mathbb{R}^n}U_h(y)(\eta(x)-\eta(y))\rho_\varepsilon(x-y)dy\quad\forall x\in\mathbb{R}^n\setminus\mathcal{D}(\mathcal{A}).\end{aligned}\tag{14}$$

Since H is also linear in p for any fixed x, it follows that, for all $x\in\mathbb{R}^n\setminus\mathcal{D}(\mathcal{A})$

$$\begin{aligned}&\eta(x)u_h^\varepsilon(x)+H(x,\nabla u_h^\varepsilon(x))\\ &=\int_{\mathbb{R}^n}U_h(y)(\eta(x)-\eta(y))\rho_\varepsilon(x-y)dy\\ &\quad+\int_{\mathbb{R}^n}(H(x,\nabla U_h(y))-H(y,\nabla U_h(y)))\rho_\varepsilon(x-y)dy.\end{aligned}\tag{15}$$

Note that u_h^ε is differentiable, which is automatically a viscosity solution to the above equation. Since the right-hand side of (15) converges to 0 uniformly, by [1, Ch. II, Prop. 2.2.], it follows that U_h is a viscosity solution to (12) on $\mathbb{R}^n\setminus\mathcal{D}(\mathcal{A})$, which completes the proof. ■

The following corollary shows the connection between (7) and (4) in the viscosity sense. The proof follows the same procedure as outlined above. We hence do not repeat.

Corollary 18: For any $h\in C_b(\mathbb{R}^n)$, let $W_h=1-U_h$, where U_h is defined in (10). Define

$$Z(x,W_h(x),p):=-p\cdot f(x)+\eta(x)(1-W_h(x)),\quad p\in\mathbb{R}^n.\tag{16}$$

Then, W_h is a viscosity solution to

$$Z(x,W_h(x),\nabla W_h(x))=0,\quad W_h(x_{\rm eq})=1-h(x_{\rm eq})$$

on $\mathbb{R}^n$ if and only if U_h is a viscosity solution to (12). This particularly holds for $h=\mathbb{1}$.

Theorem 19: Suppose that $f\in\boldsymbol{Lip}(\mathbb{R}^n)$ and $\eta,h\in\boldsymbol{LocLip}(\mathbb{R}^n)$, and the set $\{x\in\mathbb{R}^n:U_h(x)>0\}$ is either $\mathbb{R}^n$ or bounded. Then, U_h is the unique bounded viscosity solution to (12).

The proof for more general perturbed systems is provided in [3, Thm. 3.8]. A succinct proof of Theorem 19 is offered in Appendix B, based on assumptions specific to our setting. By the proof of Theorem 17, and under the same assumptions as in the above theorem, it also implies the uniqueness of the viscosity solution V for $-\nabla V(x)\cdot f(x)-\eta(x)=0$. The Lipschitz continuity of V has also been proven in a recent work [23].

Remark 20: Note that this V may be a Lyapunov function on $\mathcal{D}(\mathcal{A})$ given that $\partial^- V(x)\neq\emptyset$ for all $x\in\mathcal{D}(\mathcal{A})$. This follows from Definition 9 and [48], since we have $-p\cdot f(x)-\eta(x)\ge 0$ for all $p\in\partial^- V(x)$ and for all $x\in\mathcal{D}(\mathcal{A})$. In other words, a differentiable V is already a Lyapunov function, while a nondifferentiable V may not be. However, theoretically, there always exists a smooth version of V that can be as a Lyapunov function [3, Sect. 5]. The subsequent sections of this article will provide a perspective on how to use data-driven Koopman-based methods to construct a Lyapunov function, serving as an alternative to the sum-of-squares (SOS) approach in [14]. ◇

B. Zubov's Dual Equation on a Compact Subdomain

To facilitate data-driven techniques and prevent significant underapproximation of the ROA, we directly choose a sufficiently large compact region of interest $\mathcal{R}\subseteq\mathbb{R}^n$. This region can either contain the entire ROA, assuming it is bounded, or cover a

significant portion of the ROA if it is unbounded. Our proposed method uses observable data to recover the ROA relative to $\mathcal{R}$.

To incorporate Zubov's dual equation, we need to recast the dynamics in $\mathcal{R}$. We first consider a first-hitting time of $\partial\mathcal{R}$ defined as follows:

$$\tau := \tau(x) = \inf\{t \geq 0 : \phi(t,x) \in \partial\mathcal{R}\}, \quad x \in \mathcal{R}. \tag{17}$$

We further define stopped-flow maps so that one can observe the trajectories in $\mathcal{R}$.

Definition 21: Given the compact region of interest $\mathcal{R}$, for each $x \in \mathcal{R}$, we define the stopped-flow maps $\hat{\phi} : [0,\infty) \times \mathcal{R} \to \mathcal{R}$ as $\hat{\phi}(t,x) := \phi(t \wedge \tau, x)$.

Note that in the above definition, the stopping time τ implicitly encodes the information of the starting position. It can also be verified that

1) $\hat{\phi}(0,x) = x$ and $\hat{\phi}(s,\phi(t,x)) = \hat{\phi}(t+s,x)$ for all $x \in \mathcal{R}$, and
2) $\partial_t(\hat{\phi}(t,x)) = f(\hat{\phi}(t,x))$ for all $x \in \operatorname{int}(\mathcal{R})$.

We then consider a recast version of functions V and U_h [defined in (9) and (10)] accordingly. Let $\hat{V}(x) = \int_0^\infty \eta(\hat{\phi}(t,x))dt$ for all $x \in \mathcal{R}$. For any $h \in C(\mathcal{R})$,[1] we define

$$\hat{U}_h(x) = \begin{cases} \exp\left\{-\hat{V}(x)\right\} h(\hat{\phi}_\infty(x)), & \text{if } \hat{V}(x) < \infty \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

where $\hat{\phi}_\infty(x) := \lim_{t\to\infty} \hat{\phi}(t,x)$. For test function $h(x) = \mathbb{1}(x)$, we denote $\hat{U} := \hat{U}_{\mathbb{1}}$ for simplicity. In this notion, it can be verified that $\hat{V}(x) = \infty$ if and only if $\tau < \infty$ and $x \notin \mathcal{D}(\mathcal{A})$. Working on the stopped flow, we seek to recover the bounded set $\{x \in \mathcal{R} : \hat{U}_h > 0\} \subseteq \mathcal{D}(\mathcal{A})$.

The following theorem collects nice properties of $\hat{V}$ and $\hat{U}$ on the refined region $\mathcal{R}$.

Theorem 22: For any $h \in C(\mathcal{R})$

1) $\hat{U}_h$ is a viscosity solution to
$$H(x, \hat{U}_h(x), \nabla\hat{U}_h(x)) = 0, \quad \hat{U}_h(x_{\text{eq}}) = h(x_{\text{eq}}) \tag{19}$$
on $\operatorname{int}(\mathcal{R})$, where H is defined in (8). Suppose that η and h are also locally Lipschitz continuous, then $\hat{U}_h$ is the unique bounded viscosity solution.
2) $\hat{W}_h = 1 - \hat{U}_h$ is a viscosity solution to $Z(x, \hat{W}_h(x), \nabla\hat{W}_h(x)) = 0$, $\hat{W}_h(x_{\text{eq}}) = 1 - h(x_{\text{eq}})$, on $\operatorname{int}(\mathcal{R})$ if and only if $\hat{U}_h$ is a viscosity solution to (19), where Z is defined in (16).
3) Let $F(x,p) := -p \cdot f(x) - \eta(x)$. On any invariant set $\mathcal{I} \subseteq \mathcal{D}(\mathcal{A}) \cap \operatorname{int}(\mathcal{R})$, the function $\hat{V}(x) = -\log(\hat{U}(x))$ is a viscosity solution to $F(x, \nabla\hat{V}(x)) = 0$ with $\hat{V}(x_{\text{eq}}) = 0$.

Proof: The proof follows a similar procedure as the proofs in Section III-A. Indeed, the proof should be the same for any $x \in \operatorname{int}(\mathcal{R})$ such that $\tau = \infty$. Particularly, the dynamic programming as in the proof of Lemma 15 still holds given the flow map property of $\hat{\phi}$. For any $x \in \operatorname{int}(\mathcal{R})$ such that $\tau < \infty$, the solution is trivial considering that the quantity $\hat{V}(x) = \int_0^\tau \eta(\hat{\phi}(s,x))ds + \eta(\hat{\phi}(\tau,x)) \int_\tau^\infty ds$ diverges. ■

[1]Note that we have subtly changed the space of test functions from $C_b(\mathbb{R}^n)$ [in (10)] to $C(\mathcal{R})$.

Remark 23: By Theorem 22, suppose that $\mathcal{D}(\mathcal{A}) \cap \operatorname{int}(\mathcal{R}) \neq \mathcal{D}(\mathcal{A})$, one can only recover a portion of $\mathcal{D}(\mathcal{A})$ that is not absorbed by the boundary by solving (19). This portion should be a sublevel set (relative to $\mathcal{D}(\mathcal{A}) \cap \operatorname{int}(\mathcal{R})$) of the $\hat{V}$. In view of [28], [29], and [30], this sublevel set is also a subset of the refined open and invariant subregion of ROA, from which trajectories will satisfy the reach-avoid-stay property. ◇

IV. Zubov–Koopman Operators and Semigroup Property

Addressing the problem of estimating the ROA, we have introduced Zubov's dual equation as well as its refined form on $\mathcal{R}$. The solution involves an improper integral up to ∞, which requires nearly the full knowledge of the trajectory [15]. To reduce the substantial amount of observation data, in this section, we derive an approximation approach using a time series. Specifically, this time series is governed by a convergent, time-homogeneous, and Feynman–Kac like semigroup, which allows us to approximate the long-term behavior through a simple iterative process.

We first work on $\mathbb{R}^n$ and then on the refined region $\mathcal{R}$.

A. Introducing Zubov–Koopman Operators

Consider

$$v_t(x) := \int_0^t \eta(\phi(r,x))dr \tag{20}$$

and, for any $h \in C_b(\mathbb{R}^n)$ and $t > 0$, we define $\mathcal{T}_t : C_b(\mathbb{R}^n) \to C_b(\mathbb{R}^n)$ as

$$\mathcal{T}_t h(x) := \exp\{-v_t(x)\} h(\phi(t,x)). \tag{21}$$

The following proposition shows the basic properties of $\{\mathcal{T}_t\}_{t\geq 0}$. We complete the proof in Appendix C.

Proposition 24: $\{\mathcal{T}_t\}_{t\geq 0}$ is a C_0-semigroup. In addition, for each $t \geq 0$ and for any $h \in C_b(\mathbb{R}^n)$, $\mathcal{T}_t U_h(x) = U_h(x)$ for all $x \in \mathbb{R}^n$.

The stochastic version of $\{\mathcal{T}_t\}_{t\geq 0}$ is the famous Feynman–Kac semigroup [37, Ch. 8]. While this is generally not true for stochastic systems, for deterministic systems, we observe that $\mathcal{T}_t$ depicts a form of separation of variables and can be written as a multiplication of a contraction operator with the Koopman operator, i.e., $\mathcal{T}_t = \exp\{-v_t\}\mathcal{K}_t$.

The following theorem also shows a close connection with the Zubov's dual equation. Particularly, we will introduce (22) below, which is a time-varying version of the Zubov's dual equation, and captures a steady point in the function space when $t \to \infty$. For the purpose of using the flow of h governed by $\mathcal{T}_t$ to approximate the solution of Zubov's dual equation, for any fixed t, we name $\mathcal{T}_t$ as the *Zubov–Koopman Operator*.

Theorem 25: Let the test function be $h \in C_b^1(\mathbb{R}^n)$. Suppose that $\eta \in C(\mathbb{R}^n)$ is nonnegative. Then, the following holds.

1) $u_h(t,x) = \mathcal{T}_t h(x)$ solves the following Cauchy problem, for all $t > 0$ and all $x \in \mathbb{R}^n$
$$\begin{cases} \partial_t u_h(t,x) = \nabla_x u_h(t,x) \cdot f(x) - \eta(x) u_h(t,x) \\ u_h(0,x) = h(x). \end{cases} \tag{22}$$

2) Conversely, for any $u_h \in C^{1,1}([0,\infty),\mathbb{R}^n)$ that satisfies (22), the solution should be of the form $u_h(t,x) = \mathcal{T}_t h(x)$.

Proof: Note that the (infinitesimal) generator of the Koopman semigroup $\{\mathcal{K}_t\}$ acting on u_h is such that $\mathcal{L}_f u_h(t,x) = \nabla_x u_h(t,x)\cdot f(x)$. To prove (1), we first look at how u_h evolves according to $\mathcal{K}_s$ for some small $s>0$. By the definition of $\mathcal{K}_s$, for any fixed $t>0$, we have

$$\begin{aligned}
&\frac{\mathcal{K}_s u_h(t,x) - u_h(t,x)}{s}\\
=&\frac{1}{s}\left[\exp\left\{-\int_0^t \eta(\phi(r,\phi(s,x)))dr\right\} h(\phi(t,\phi(s,x))) - u_h(t,x)\right]\\
=&\frac{1}{s}\left[\exp\left\{-\int_0^t \eta(\phi(r+s,x))dr\right\} h(\phi(t+s,x)) - u_h(t,x)\right]\\
=&\frac{1}{s}\left[\exp\left\{-\int_s^{t+s} \eta(\phi(\sigma,x))d\sigma\right\} h(\phi(t+s,x)) - u_h(t,x)\right].
\end{aligned} \tag{23}$$

Note that

$$\begin{aligned}
&\exp\left\{-\int_s^{t+s}\eta(\phi(\sigma,x))d\sigma\right\}\\
=&\exp\left\{-\int_0^{t+s}\eta(\phi(\sigma,x))d\sigma\right\}\exp\left\{\int_0^{s}\eta(\phi(\sigma,x))d\sigma\right\}.
\end{aligned} \tag{24}$$

Combining (23) and (24), we have

$$\begin{aligned}
&\frac{\mathcal{K}_s u_h(t,x) - u_h(t,x)}{s}\\
&= u_h(t+s,x)\cdot\frac{1}{s}\left[\exp\left\{\int_0^s \eta(\phi(r,x))dr\right\} - 1\right]\\
&\quad + \frac{1}{s}\left[u_h(t+s,x) - u_h(t,x)\right].
\end{aligned} \tag{25}$$

Sending $s\downarrow 0$ on both sides, it follows that

$$\begin{aligned}
\nabla_x u_h(t,x)\cdot f(x) &= u_h(t,x)\eta(\phi_0(x)) + \partial_t u_h(t,x)\\
&= \partial_t u_h(t,x) + \eta(x)u_h(t,x)
\end{aligned}$$

which completes the first part of the proof.

To prove (2), we suppose that $u_h \in C^{1,1}([0,\infty),\mathbb{R}^n)$ solves (22), then $\partial_t u_h - \nabla_x u_h\cdot f + \eta u_h = 0$, $\forall t>0$, $x\in\mathbb{R}^n$ and $u_h(0,x) = h(x)$ for all $x\in\mathbb{R}^n$. Introduce an auxiliary function $\Psi(t,x,v) = \exp\{-v\}u_h(t,x)$. Then, it is clear that $u_h(s,x) = \Psi(s,x,0)$. However

$$\begin{aligned}
&d\psi(s-t,\phi(t,x),v_t(x))\\
&= -\partial_t\psi(s-t,\phi(t,x),v_t(x)) + \mathcal{L}_f\Psi(s-t,\phi(t,x),v_t(x))\\
&\quad + \eta(\phi(t,x))\partial_v\Psi(s-t,\phi(t,x),v_t(x))\\
&= \exp\{-v_t(x)\}\left[-\partial_t u_h(s-t,\phi(t,x)) + \mathcal{L}_f u_h(s-t,\phi(t,x))\right.\\
&\quad \left.+\eta(\phi(t,x))u_h(s-t,\phi(t,x))\right] = 0
\end{aligned}$$

where the quantity $v_t(x)$ is defined as in (20). Therefore, the quantity $\Psi(s-t,\phi(t,x),v_t(x))$ is a constant for all t and x, which implies that

$$\begin{aligned}
u_h(s,x) &= \Psi(s,x,0)\\
&= \Psi(s,\phi(0,x),v_0(x)) = \Psi(0,\phi(s,x),v_s(x))\\
&= \exp\{-v_s(x)\}u_h(0,\phi(s,x)) = \exp\{-v_s(x)\}h(\phi(s,x)).
\end{aligned}$$

The proof is completed. ■

B. Time-Series Approximation

Clearly, by Theorem 25 and by the definition of U_h in (10), for any $h\in C_b^1(\mathbb{R}^n)$, $\mathcal{T}_t h$ is the unique solution to (22), and $\lim_{t\to\infty} u_h(t,x) = \lim_{t\to\infty}\mathcal{T}_t h(x) = U_h(x)$ for all $x\in\mathbb{R}^n$. In particular, $\lim_{t\to\infty}\mathcal{T}_t\mathbb{1} = U$ uniformly, and U is also the unique, up to multiplicative constants, fixed point of $\{\mathcal{T}_t\}_{t\geq 0}$. To approximate U, one can pick a fixed time interval Δt, and define

$$\mathcal{T}_\Delta := \mathcal{T}_{\Delta t} \tag{26}$$

as well as

$$\mathcal{T}_{k\Delta} := \underbrace{\mathcal{T}_\Delta\circ\mathcal{T}_\Delta\cdots\circ\mathcal{T}_\Delta}_{k \text{ iterations}}. \tag{27}$$

Then, by (10) and the uniqueness (up to multiplicative constants) of the fixed point, the composed operator $\mathcal{T}_{k\Delta}: C_b(\mathbb{R}^n)\to C_b(\mathbb{R}^n)$ for any $k\geq 1$, and $\lim_{k\to\infty}\mathcal{T}_{k\Delta}h = U$ for any h such that $h(x_{\text{eq}}) = 1$. Suppose that one can approximate $\mathcal{T}_\Delta$ properly, then $\mathcal{T}_{k\Delta}h$ for some large k should be a reasonably good approximate for U.

Similar to the approximation of Koopman operators, to obtain a discrete version $\mathbf{T}$ of the bounded linear operator $\mathcal{T}_\Delta$, it usually relies on the choice of a (discrete) dictionary of observable test functions, denoted by

$$\mathcal{Z}_N(x) := [\mathfrak{z}_0(x),\mathfrak{z}_1(x),\ldots,\mathfrak{z}_{N-1}(x)],\quad N\in\mathbb{N}\cup\{\infty\}. \tag{28}$$

Then, the approximation $\widetilde{\mathcal{T}}: \text{span}\{\mathfrak{z}_i\}_{i=0}^{N-1}\to\text{span}\{\mathfrak{z}_i\}_{i=0}^{N-1}$ is valid in the sense that, for each $h\in C_b(\mathbb{R}^n)$, there exists an $\mathfrak{h}\in\text{span}\{\mathfrak{z}_i\}_{i=0}^{N-1}$ and a uniformly continuous residual term $\vartheta\in C_b(\mathbb{R}^n)$ such that

$$\mathcal{T}_\Delta h = \widetilde{\mathcal{T}}\mathfrak{h} + \vartheta. \tag{29}$$

Possible choices of the dictionary $\mathcal{Z}_N$ have been discussed in [7] and [43], including polynomials, Fourier basis, spectral elements, and NN-based functions. These choices are generally locally Lipschitz continuous, but there may be cases where differentiability is not exhibited, as seen in NN-based functions with rectified linear unit (ReLU) as the activation function. In line with Theorem 25 and viscosity regularity of Zubov's dual equation, to make the time series approximation of U robust, we extend Theorem 25 for test functions $h\in C_b(\mathbb{R}^n)\cap\boldsymbol{LocLip}(\mathbb{R}^n)$ and verify the regularity.

Theorem 26: Let the test function be $h\in C_b(\mathbb{R}^n)\cap\boldsymbol{LocLip}(\mathbb{R}^n)$. Suppose that $\eta\in C(\mathbb{R}^n)$ is nonnegative and locally Lipschitz. Then, for each $t>0$, $u_h := \mathcal{T}_t h$ is the unique viscosity solution to (22).

Furthermore, for each fixed $t>0$, suppose that there exists a family of functions $\{u_k^t\in C_b(\mathbb{R}^n)\cap\boldsymbol{LocLip}(\mathbb{R}^n)\}_{k=0}^\infty$ and, accordingly, a family of uniformly bounded continuous residuals $\{\vartheta_k^t\}_{k=0}^\infty$ with Lipschtiz constant L_k, such that

$$u_h(t,x) := \mathcal{T}_t h(x) = u_k^t(x) + \vartheta_k^t(x),\quad x\in\mathbb{R}^n$$

and $\vartheta_k^t(x_{\text{eq}}) = 0$. Assume $\sup_{x\in\Omega}|\vartheta_k^t(x)|\to 0$ on each bounded subdomain $\Omega\subseteq\mathbb{R}^n$. Then, as $t,k\to\infty$, u_k^t converges uniformly to U_h.

In addition, suppose we also have $L_k \to 0$, then the function $\mathfrak{v}_k^t = -\log(\mathfrak{u}_k^t)$ is the unique viscosity solution to $-\nabla \mathfrak{v}_k^t(x) \cdot f(x) - \eta(x) + \mathcal{O}_k(x) = 0$ with $\mathfrak{v}_k^t(x_{\text{eq}}) = 0$, where $\sup_{x\in\Omega} |\mathcal{O}_k(x)| \to 0$ (as $k \to \infty$) on each bounded subdomain $\Omega \subseteq \mathbb{R}^n$.

Proof: Since the proof of viscosity property is similar to the proof for U_h, we omit the first part of the proof.

Now, as $t \to \infty$, by the uniform convergence of $\mathcal{T}_t h$ for any $h \in C_b(\mathbb{R}^n)$, one has $|\partial_t u_h(t,x)|$ uniformly converges to 0 on each subdomain $\Omega \subseteq \mathbb{R}^n$. By the construction of $\mathfrak{u}_k^t$, as $k \to \infty$, $\mathfrak{u}_k^t$ also uniformly converges to $u_h(t,\cdot)$. Therefore, for each h, $\mathfrak{u}_k^t \to U_h$ uniformly, which completes the second part of the proof.

By the construction of $\mathfrak{u}_k^t$, and by the first part of Theorem 26, we can immediately have that, for each t and for each k, $\mathfrak{u}_k^t$ is the unique viscosity solution to

$$\begin{aligned}&\nabla \mathfrak{u}_k^t(x) \cdot f(x)\\ &= \partial_t u_h(t,x) + \eta(x)\mathfrak{u}_k^t(x) + \eta(x)\vartheta_k^t(x) - \nabla \vartheta_k^t(x) \cdot f(x).\end{aligned} \tag{30}$$

Since ϑ_k^t is necessarily locally Lipschitz continuous, $\nabla \vartheta_k^t(x)$ exists a.e. At a differentiable point, we have $|\nabla \vartheta_k^t(x)| \le L_k \to 0$. This implies that $|\nabla \vartheta_k^t|$ converges uniformly to 0 a.e. For sufficiently large t and k, by a similar argument as the first part,[2] one can verify that there exists an $\widetilde{\mathcal{O}}_k$ such that $\sup_{x\in\Omega} |\widetilde{\mathcal{O}}_k(x)| \to 0$ and $\mathfrak{u}_k^t$ is the viscosity solution to

$$-\nabla \mathfrak{u}_k^t(x) \cdot f(x) = -\eta(x)\mathfrak{u}_k^t(x) + \widetilde{\mathcal{O}}_k(x). \tag{31}$$

By a similar argument as in Proposition 15, given the smoothness of $-\log(\mathfrak{u}_k^t)$, there exists an $\mathcal{O}_k$ with $\sup_{x\in\Omega} |\mathcal{O}_k(x)| \to 0$ such that $\mathfrak{v}_k^t$ is the unique viscosity solution to $-\nabla \mathfrak{v}_k^t(x) \cdot f(x) - \eta(x) + \mathcal{O}_k(x) = 0$ with $\mathfrak{v}_k^t(x_{\text{eq}}) = 0$. ■

Remark 27: The first two parts of the above theorem state that, even though we may only use a family of uniformly convergent locally Lipschitz functions to approximate, the limit still satisfies (22) in a viscosity sense. In addition, we can find a proper approximation $\mathfrak{u}_k^t$ for U_h. For the purpose of approximating $\{x \in \mathbb{R}^n : U(x) \ge 0\}$ w.r.t. the Hausdorff metric,[3] this approximation is satisfactory.

On the other hand, without the assumption that $L_k \to 0$, the approximation $\mathfrak{u}_k^t$ may not solve (30) with vanishing $\nabla \vartheta_k^t \cdot f$ in a proper sense. This may eventually cause the third part of the statement to fail to hold, meaning that the approximation $\mathfrak{v}_k^t$ cannot be readily used as a Lyapunov function. We will also demonstrate this effect in Section VII via examples. ◇

At the end of this section, we make a quick extension of the aforementioned results, further refining our observations on the compact region of interest $\mathcal{R}$. Due to the similarity with previous results, we omit the proof for the following Corollary.

[2]One can take a convolution on both side of (30) with a mollifier ρ_ε and let $\widetilde{\mathcal{O}}_k = (\eta \vartheta_k^t - \nabla \vartheta_k^t \cdot f) * \rho_\varepsilon$. Then, $\widetilde{\mathcal{O}}_k$ satisfies the requirement.

[3]For any approximator H of U such that $\|U - H\| \le \varepsilon$, one can show that the set $\{x : H(x) \ge \varepsilon\}$ is a tight inner approximation of the ROA. By definition, the Hausdorff distance between the ROA and $\{x : H(x) \ge \varepsilon\}$ is bounded by a small value determined by ε, reflecting the proximity of the two sets in terms of their level set definitions. However, there is a possibility that the approximated boundary 'crosses' the $\partial\mathcal{D}(\mathcal{A})$. To determine whether it provides a proper inner approximation of the ROA, one must undergo formal verifications by checking the Lie derivative properties, as stated in Section VI-C and detailed in [22] and [23].

Corollary 28: Recall the stopped-flow map $\hat{\phi}$ defined in Definition 21. Let $\hat{v}_t(x) := \int_0^t \eta(\hat{\phi}(r,x))dr$. For any $h \in C(\mathcal{R})$ and $t > 0$, we redefine $\mathcal{T}_t : C(\mathcal{R}) \to C(\mathcal{R})$ as

$$\mathcal{T}_t h(x) := \exp\{-\hat{v}_t(x)\}\, h(\hat{\phi}(t,x)). \tag{32}$$

Then, the following holds.

1) $\{\mathcal{T}_t\}_{t\ge 0}$ is a C_0-semigroup.
2) For each each $h \in C(\mathcal{R})$ and for each t, $\hat{U}_h$ is an eigenfunction such that $\mathcal{T}_t \hat{U}_h = \hat{U}_h$.
3) For any test function $h \in \boldsymbol{Lip}(\mathcal{R})$, given that $\eta \in \boldsymbol{Lip}(\mathcal{R})$ is nonnegative, then $\hat{u}_h(t,x) := \mathcal{T}_t h(x)$ is the unique viscosity solution to (22) for all $t > 0$ and $x \in \text{int}(\mathcal{R})$.
4) Suppose there exists a family of functions $\{\mathfrak{u}_k^t \in \boldsymbol{Lip}(\mathcal{R})\}_{k=0}^\infty$ that uniformly converges to $\hat{u}_h(t,\cdot)$ for any t, then, as $t, k \to \infty$, $\mathfrak{u}_k^t$ converges uniformly to $\hat{U}_h$, where $\hat{U}_h$ is defined in (18).
5) Suppose we also have $L_k \to 0$, which is the Lipschitz constant for $\hat{u}_h(t,\cdot) - \mathfrak{u}_k^t$ for each t. Then, for sufficiently large t and k, the function $\mathfrak{v}_k^t = -\log(u_k^t)$ is the unique viscosity solution to $-\nabla \mathfrak{v}_k^t(x) \cdot f(x) - \eta(x) + \mathcal{O}(x) = 0$ with $\mathfrak{v}_k^t(x_{\text{eq}}) = 0$ on any invariant set $\mathcal{I} \subseteq \mathcal{D}(\mathcal{A}) \cap \text{int}(\mathcal{R})$, where $\sup_{x\in\mathcal{I}} |\mathcal{O}(x)|$ is arbitrarily small.

Remark 29: We have chosen not to introduce the "hat" notation for the redefined $\mathcal{T}_t$ in (32). The difference between the two versions [(21) and (32)] lies in their domains. We encourage readers to verify the domain of $\mathcal{T}_t$ in the context before using it. ◇

V. Finite-Dimensional Approximation of Zubov–Koopman Operators

As we have seen in (29), we expect to find an approximation for the Zubov–Koopman operators such that the image functions converge uniformly. In practice, we would like to see if the training discrete dictionary $\mathcal{Z}_N$ [as in (28)] of observable test functions can be reduced to finite, such that the approximation behaves like a finite-rank operator, and preserves $\text{span}\{\mathfrak{z}_i\}_{i=0}^{N-1}$ for some $N < \infty$.

In this section, we rigorously investigate basic properties and a finite-dimensional approximation of Zubov–Koopman operators. We will work on the compact region of interest $\mathcal{R} \subseteq \mathbb{R}^n$ for the rest of this article. Since our purpose is to learn Zubov–Koopman operators based on training data, we propose a three-step intermediate approximation for $\{\mathcal{T}_t\}$, such that for any $t > 0$, an approximation of the form (29) holds.

A. Compact Approximation of Zubov–Koopman

$\{\mathcal{T}_t\}_{t\ge 0}$ defined in (32) is clearly a family of bounded linear operators. One may attempt to show that $\mathcal{T}_t$ is also compact for each t. It suffices to show that $\mathcal{T}_t(\mathcal{B}_r) \subseteq C(\mathcal{R})$ is relatively compact, where $\mathcal{B}_r = \{h \in C(\mathcal{R}) : \|h\|_\infty \le r\}$ for some $r > 0$. However, equicontinuity within $\mathcal{T}_t(\mathcal{B}_r)$ is not guaranteed. To see this, we set $h_n(x) = \sin(nx) \in \mathcal{B}_1$ (or similarly, the Fourier basis), and let $\hat{\phi}(t,x) = x \cdot e^{-t}$ for all $x \in \text{int}(\mathcal{R})$ and $\hat{\phi}(t,x) = x$ elsewhere. Then, the sequence $\{h_n \circ \hat{\phi}(t,\cdot)\}_n$ for each t does not possess equicontinuity due to the rapid oscillation as n increases. Nonetheless, the following proposition states that

one can use compact operators to strongly approximate $\mathcal{T}_t$ for each t.

Proposition 30: For each $t > 0$, there exists a family of compact linear operator $\{\mathcal{T}_t^\varepsilon\}_{\varepsilon>0}$, such that for all $h \in C(\mathcal{R})$, we have $\|\mathcal{T}_t^\varepsilon h - \mathcal{T}_t h\|_\infty \to 0$ as $\varepsilon \to 0$.

Proof: For each t, we rewrite the Koopman operator as $\mathcal{T}_t h(x) = \int_{\mathcal{R}} \delta(x-y)\mathcal{T}_t h(y)dy$, where δ is the Dirac delta. Now we use a family of integral operators with smooth kernels to approximate the above distribution. The idea is to approximate the Dirac delta using a smooth mollifier $\rho \in C_0^\infty(\mathcal{R})$ as we have seen in the proof of Theorem 26. For each $\varepsilon > 0$, let $\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\zeta(\frac{x}{\varepsilon})$ and $\int_{\mathcal{R}} \rho(y)dy = 1$. It can be verified that $\rho_\varepsilon \subset C_0^\infty(\mathcal{R})$ with a compact support in $\mathcal{B}_\varepsilon$. We define the approximation as $\mathcal{T}_t^\varepsilon h(x) = \int_{\mathcal{R}} \rho_\varepsilon(x-y)\mathcal{T}_t h(y)dy$ for all $x \in \mathcal{R}$. It is a well-known result that, for each t and $\varepsilon > 0$, the operator $\mathcal{T}_t^\varepsilon : C(\mathcal{R}) \to C(\mathcal{R})$ is compact given its smooth kernel [5, Ch. II.6].

To verify the convergence property, the rest of the proof falls in a standard procedure. Now, for each t and all $x \in \mathcal{R}$, by change of variable, we have that $\mathcal{T}_t^\varepsilon h(x) = \int_{\mathcal{B}_1} \rho(y)\mathcal{T}_t h(x-\varepsilon y)dy$. It follows that

$$\begin{aligned}|\mathcal{T}_t^\varepsilon h(x) - \mathcal{T}_t h(x)| &\le \left|\int_{\mathcal{R}} \rho_\varepsilon(x-y)[\mathcal{T}_t h(y) - \mathcal{T}_t h(x)]dy\right| \\ &\le \int_{\mathcal{B}_1} \rho(y)|\mathcal{T}_t h(x-\varepsilon y) - \mathcal{T}_t h(x)|dy \\ &\le \sup_{y\in\mathcal{B}_1} |\mathcal{T}_t h(x-\varepsilon y) - \mathcal{T}_t h(x)|. \qquad (33)\end{aligned}$$

Note that, tor each t, given any $h \in C(\mathcal{R})$ and any Lipshitz continuous map $\hat{\phi}(t,\cdot) : \mathcal{R} \to \mathcal{R}$, the composition $h \circ \hat{\phi}$ is a uniform continuous function on $\mathcal{R}$. Therefore, by definition, $\mathcal{T}_t h$ is also uniformly continuous on $\mathcal{R}$. The last term above converges to 0 as $\varepsilon \downarrow 0$ for all $x \in \mathcal{R}$. Therefore, for each t, taking the supremum on both sides of (33) and sending ε to 0, we have $\|\mathcal{T}_t^\varepsilon h - \mathcal{T}_t h\|_\infty \to 0$. ■

Corollary 31: For each $t, s > 0$, there exists a family of compact linear operator $\{\mathcal{T}_t^\varepsilon\}_{\varepsilon>0}$, such that for all $h \in C(\mathcal{R})$, we have $\|(\mathcal{T}_t^\varepsilon \circ \mathcal{T}_t^\varepsilon)h - (\mathcal{T}_t \circ \mathcal{T}_s)h\|_\infty \to 0$ as $\varepsilon \to 0$.

Proof: We can use the same $\{\mathcal{T}_t^\varepsilon\}_{\varepsilon>0}$ as in Proposition 30. Then, the composition $(\mathcal{T}_t^\varepsilon \circ \mathcal{T}_t^\varepsilon)h = (\rho_\varepsilon * \rho_\varepsilon) * \mathcal{T}_{t+s}h$, where $(\rho_\varepsilon * \rho_\varepsilon)$ is a smoother mollifier than ρ_ε with the same convergence property relative to the Dirac Delta. The rest of the proof should be the same as in Proposition 30. ■

Remark 32: The convergence in Proposition 30 cannot be extended to the convergence w.r.t. the operator norm, i.e., $\|\mathcal{T}_t^\varepsilon - \mathcal{T}_t\| := \sup_{\|h\|_\infty=1} \|(\mathcal{T}_t^\varepsilon - \mathcal{T}_t)h\| \to 0$. One may revisit the example $h_n(x) = sin(nx)$ for $n \in \mathbb{N}$. It is clear that $\|h_n\|_\infty = 1$ for all n, but based on the inequality in (33), the uniform limit fails to exist due to the unbounded Lipschitz constants of $\{h_n\}$. However, for the purpose of this section, the convergence in Proposition 30 is already satisfactory. ◇

B. Finite Dimensional Representation

Since the kernels of $\{\mathcal{T}_t^\varepsilon\}$ are smooth and compactly supported, it is a well-known result that the operators within this family are Hilbert–Schmidt operators [45, Ch. 4]. We then investigate the spectral behavior of $\mathcal{T}_t^\varepsilon$ within a separable Hilbert space $\mathcal{H}$ with the inner product $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_{L_2}$. Let $\{\zeta_i\}_{i\in\mathbb{Z}} \subset \mathcal{H}$ be the eigenfunctions of $\mathcal{T}_t^\varepsilon$. Then, for all $h \in C(\mathcal{R}) \subset \mathcal{H}$, we have that $\mathcal{T}_t^\varepsilon h = \sum_{i=0}^\infty e^{\lambda_i^\varepsilon t}\langle h, \zeta_i\rangle\zeta_i$ as well as $\mathcal{T}_t^\varepsilon \zeta_j = \sum_{i=0}^\infty e^{\lambda_i^\varepsilon t}\langle\zeta_j,\zeta_i\rangle\zeta_i = e^{\lambda_j^\varepsilon t}\zeta_j$ where $\{\lambda_i^\varepsilon\}$ are the corresponding eigenvalues in the log-scale.

Remark 33: In the above infinite-sum representations, we have implicitly assumed that the eigenfunctions are real-valued. However, this is not always the case. For more general situations, $\langle\zeta_i,\bar{\zeta}_i\rangle = 1$ for all i, and $\mathcal{T}_t^\varepsilon h = \sum_{i\in\mathbb{Z}} e^{\lambda_i^\varepsilon t}\langle h,\bar{\zeta}_i\rangle\zeta_i$ as well as $\mathcal{T}_t^\varepsilon \zeta_j = \sum_{i\in\mathbb{Z}} e^{\lambda_i^\varepsilon t}\langle\zeta_j,\bar{\zeta}_i\rangle\zeta_i = e^{\lambda_j^\varepsilon t}\zeta_j$. For simplicity, we still use the above expressions without further indicating whether they are real or complex-valued. ◇

Proposition 34: For any fixed $t > 0$, for any arbitrarily small $\vartheta > 0$, there exists a sufficiently large N and a finite-dimensional approximation $\mathcal{T}_{t,N}^\varepsilon$ such that $\|\mathcal{T}_{t,N}^\varepsilon h - \mathcal{T}_t h\|_\infty < \vartheta, \quad h \in C(\mathcal{R})$.

Proof: The Hilbert–Schmidt operator $\mathcal{T}_t^\varepsilon$ is also necessarily a finite-rank operator, i.e., for any complete orthonormal basis $\{g_j\} \subset \mathcal{H}$, $\sum_{j=0}^\infty\langle\mathcal{T}_t^\varepsilon g_j, g_j\rangle = \sum_{j=0}^\infty\sum_{i=0}^\infty e^{\lambda_i^\varepsilon t}|\langle g_j,\zeta_i\rangle|^2 = \sum_{j=0}^\infty e^{\lambda_j^\varepsilon t} < \infty$. Denoting the finite truncation as $\mathcal{T}_{t,N}^\varepsilon := \sum_{i=0}^{N-1} e^{\lambda_i t}\langle\,\cdot\,,\zeta_i\rangle\zeta_i$, we have $\lim_{N\to\infty}\|\mathcal{T}_{t,N}^\varepsilon h - \mathcal{T}_t^\varepsilon h\|_\infty = 0$ for all $h \in C(\mathcal{R})$. Combining Proposition 30, we have the desired property. ■

Remark 35: In contrast to the approach described in [24, Sect. III, eq. (2)], wherein the finite-rank operator is limited to preserving only a finite-dimensional subspace of the initial function space, our proposed finite-rank approximate operator, as in Proposition 34, demonstrates better theoretical robustness. Notably, it retains a valid infinite-dimensional domain, i.e., $h \in C(\mathcal{R})$, identical to that of $\mathcal{T}_t$ for all t.

Particularly, suppose $\{\zeta_i\}$ are complex-valued and U is the solution to (7), by the eigenvalue problem in Proposition 24, we have that $U = \mathcal{T}_t U \approx \mathcal{T}_t^\varepsilon U = \sum_{i\in\mathbb{Z}} e^{\lambda_i^\varepsilon t}\langle U,\bar{\zeta}_i\rangle\zeta_i \approx \sum_{i=-N}^N e^{\lambda_i^\varepsilon t}\langle U,\bar{\zeta}_i\rangle\zeta_i$ and $\langle U,\bar{\zeta}_i\rangle = \langle\mathcal{T}_t U,\bar{\zeta}_i\rangle \approx e^{\lambda_i^\varepsilon t}\langle U,\bar{\zeta}_i\rangle$ for some i such that $\langle U,\bar{\zeta}_i\rangle \neq 0$. Then, the corresponding $e^{\lambda_i^\varepsilon t} \in \mathbb{R}$ is an approximation of the eigenvalue 1 of $\mathcal{T}_t$, but the eigenfunction ζ_i is not a direct approximation for U. We still need the expansion form to properly approximate U. ◇

C. Feasibility of Data-Driven Approximation

For any fixed Δt, based on Proposition 34, the original operator $\mathcal{T}_\Delta$ (recall Definition 26 for $\mathcal{T}_\Delta$) can be approximated (strongly) with an arbitrarily high precision by a finite-dimensional bounded linear operator $\mathcal{T}_{\Delta,N}^\varepsilon$, which preserves the nonignorable eigenmodes. We would like to further use a data-driven approach to learn the eigenvalues and the eigenfunctions for this finite-dimensional operator.

To do this, we consider a sufficiently dense N-dimensional subspace $\mathcal{F}_N \subset C(\mathcal{R})$ and choose observable test functions $\mathcal{Z}_N(x) = [\mathfrak{z}_0(x), \mathfrak{z}_1(x), \ldots, \mathfrak{z}_{N-1}(x)]$ from $\mathcal{F}_N$. Then, $\mathcal{T}_{\Delta,N}^\varepsilon \mathcal{Z}_N(\cdot) = \sum_{i=0}^{N-1} e^{\lambda_i^\varepsilon t}\zeta_i(\cdot)\mathfrak{p}_i \xrightarrow{\|\cdot\|_\infty} \sum_{i=0}^\infty e^{\lambda_i^\varepsilon t}\zeta_i(\cdot)\mathfrak{p}_i$, where $\mathfrak{p}_i = [\langle\mathfrak{z}_0,\zeta_i\rangle, \langle\mathfrak{z}_1,\zeta_i\rangle, \cdots\langle\mathfrak{z}_{N-1},\zeta_i\rangle]$ represents the vector of the projection inner products. If $\mathrm{span}\{\mathfrak{z}_i\}_{i=0}^\infty$ is also dense in $C(\mathcal{R})$, then this finite-dimensional approximation is of the desired form as in (29).

Suppose that $\mathcal{F}_N$ is given arbitrarily, then each $\mathfrak{z}_i$ may not be invariant within $\{\zeta_i\}_{i=0}^{N-1}$ under $\mathcal{T}_\Delta$. However, when restricted on the N-dimensional truncation, the approximate operator $\mathcal{T}_{\Delta,N}^\varepsilon$ preserves the space $\{\zeta_i\}_{i=0}^{N-1}$. We aim to express $\{\zeta_i\}_{i=0}^{N-1}$ by functions from $\mathrm{span}\{\mathfrak{z}_0,\mathfrak{z}_1,\ldots,\mathfrak{z}_{N-1}\}$.

Taking advantage of the linear transformation, for systems with unknown dynamics, data-driven methods will approximate the Zubov–Koopman eigenvalues and eigenfunctions by fitting the data $\mathcal{Z}_N(x^{(m)})$ and $\mathcal{T}_\Delta \mathcal{Z}_N(x^{(m)})$ for some $\{x^{(m)}\}_{m=0}^{M-1} \subset \mathcal{R}$. We also denote by $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{M\times N}$ the stack of input and output data. The best-fit matrix $\mathbf{T}$ is formulated as an optimization problem $\mathbf{T} = \operatorname{argmin}_{A\in\mathbb{C}^{N\times N}} \|\mathbf{Y} - \mathbf{X}A\|_F$.

The learned operator has the following sense of approximations for sufficiently large N and M.

1) Let $(\mu_i, \mathbf{e}_i)_{i=0}^N$ be the eigenvalues and eigenvectors of $\mathbf{T}$. Then, for each i, the learned eigenvalues $\tilde{\lambda}_i$ is such that $\lambda_i^\varepsilon \approx \tilde{\lambda}_i = \log(\mu_i)/\Delta t$, and the learned eigenfunction $\widetilde{\zeta}_i$ satisfies $\zeta_i(x) \approx \widetilde{\zeta}_i(x) = \mathcal{Z}_N(x)\mathbf{e}_i$. [4]
2) For any $\mathfrak{h} \in \operatorname{span}\{\mathfrak{z}_0, \mathfrak{z}_1, \ldots, \mathfrak{z}_{N-1}\}$ such that $\mathfrak{h}(x) = \mathcal{Z}_N(x)\mathbf{w}$ for some column vector $\mathbf{w}$, we have that

$$\mathcal{T}_t\mathfrak{h}(\cdot) \approx \mathcal{T}_{t,N}^\varepsilon \mathfrak{h}(\cdot) \approx \mathcal{Z}_N(\cdot)(\mathbf{T}\mathbf{w}). \tag{34}$$

Due to its substantial similarity to EDMD [43], we omit the proof for this part of the convergence. By the well-known universal approximation theorem, all of the function approximations from above should have uniform convergence.

VI. Data-Driven Algorithms

By the finite rank expansion (34), one can easily verify that

$$\mathcal{T}_{k\Delta}\mathfrak{h}(\cdot) \approx \mathcal{Z}_N(\cdot)(\mathbf{T}^k\mathbf{w}) = \sum_{i=0}^{N-1} e^{\tilde{\lambda}_i k\Delta t}\widetilde{\zeta}_i(\cdot)(\mathbf{e}_i \cdot \mathbf{w}) \tag{35}$$

where $\mathbf{w}$ is such that $h(\cdot) = \mathcal{Z}_N(\cdot)\mathbf{w}$. Recall the beginning part of Section IV-B, particularly (27) and $\lim_{k\to\infty} \mathcal{T}_{k\Delta}h = U$ for any h such that $h(x_{\text{eq}}) = 1$. Eq (35) is then a numerical version for approximating U using iterations. By virtue of Proposition 24 and Corollary 28, for all t, the real-valued eigenfunction of $\mathcal{T}_t$ with the corresponding eigenvalue 1 should be U, up to multiplicative constants. Similarly, the only nonvanishing mode(s) as $k \to \infty$ should be the ones with $\tilde{\lambda}_i = 0$, which will be readily a solution to the Zubov's dual equation suppose it is real-valued. If this happens, we can directly choose $\widetilde{\zeta}_i$ associated with $\tilde{\lambda}_i = 0$ as the approximation for U without any iteration. However, due to the numerical error and the possibly complex-valued eigenvectors of $\mathbf{T}$, one still needs to use the iteration form for some sufficiently large k as the final approximation.

Nonetheless, the key is to learn $\mathcal{T}_\Delta$ using observable data of trajectories within a relatively shorter time period, and obtain the matrix $\mathbf{T}$ so that the (35) can be utilized. To do this, we modify the existing Koopman learning techniques for Zubov–Koopman operators $\mathcal{T}_\Delta$, as defined in (26), with a fixed training time interval Δt.

A. Generating Training Data

As we have briefly discussed about the feasibility of using data-driven approaches to approximate the operator $\mathcal{T}_\Delta$ for each $\mathfrak{z} \in \mathcal{Z}_N$ and each $x \in \mathcal{R}$, we consider $\mathfrak{z}(x)$ as the features and $\mathcal{T}_{\Delta\mathfrak{z}}(x) := \exp(-\int_0^{\Delta t} \eta(\hat{\phi}(s,x))ds)\mathfrak{z}(\hat{\phi}(\Delta t, x))$ as the labels.

[4]The eigen-approximation is only for the compact operator $\mathcal{T}_t^\varepsilon$ introduced in Section V-A. The original operator $\mathcal{T}_t$ may not possess nontrivial point spectrum other than 1.

Algorithm 1: Training Data Preprocessing

Let the region of interest $\mathcal{R}$, the dictionary $\mathcal{Z}_N(\cdot) = [\mathfrak{z}_0(\cdot), \mathfrak{z}_1(\cdot), \ldots, \mathfrak{z}_{N-1}(\cdot)]^\mathsf{T}$, the function $\eta \in C(\mathbb{R}^n)$, and the time interval Δt be given. Then, for each $i \in \{0, 1, \cdots N-1\}$ and $x \in \mathcal{R}$, we compute $\mathcal{T}_{\Delta\mathfrak{z}_i}(x)$ as follows.

1) (Solving (36)) Choose a number of partition points N. Generate a uniform partition with N points from the interval $[0, \Delta t]$. Denote the size of each partition as δt. Solve (36) using an ODE solver with this partition. Denote the solution arrays as $\operatorname{vec}(\phi, x) := [\phi(j \cdot \delta t, x), \text{ for } j \in \{0, 1, \ldots, \mathsf{N}-1\}]$ and $\operatorname{vec}(I) := [I(j \cdot \delta t), \text{ for } j \in \{0, 1, \ldots, \mathsf{N}-1\}]$. Also, denote the j-th element of the vectors as $\operatorname{vec}_j()$.
2) (Identifying out-of-domain index) Identify the first element numbered as ι in $\operatorname{vec}(\phi, x)$ that is not within $\mathcal{R}$.
3) (Identifying the intersection point on the boundary) If $\iota \notin \{0, \ldots, \mathsf{N}-1\}$, continue to the next step. Otherwise, find the line segment connecting the $(\iota-1)$-th and ι-th elements (i.e. $\operatorname{vec}_{\iota-1}(\phi, x)$ and $\operatorname{vec}_\iota(\phi, x)$). Find the intersection point $\mathcal{R}_{\text{sect}}$ of the line segment and $\partial\mathcal{R}$.
4) (Modifying the trajectory and the integral) If $\iota \notin \{0, \ldots, \mathsf{N}-1\}$, continue to the next step. Otherwise, for each $j \in \{\iota, \cdots \mathsf{N}-1\}$, replace $\operatorname{vec}_j(\phi, x)$ by $\mathcal{R}_{\text{sect}}$, and replace $\operatorname{vec}_j(I)$ by $\operatorname{vec}_{\iota-1}(I) + (j-\iota+1)\mathcal{R}_{\text{sect}} \cdot \delta t$. Keep the notation for the modified $\operatorname{vec}(\phi, x)$ and $\operatorname{vec}(I)$.
5) $\mathcal{T}_{\Delta\mathfrak{z}_i}(x) = \exp\{-\operatorname{vec}_{\mathsf{N}-1}(I)\}\mathfrak{z}_i(\operatorname{vec}_{\mathsf{N}-1}(\phi, x))$.

We inevitably need to acquire the stopped-flow map $\hat{\phi}$ and compute the integral. Drawing inspiration from [15], for each termination time, we can assess both the trajectory (without stopping) and the integral, i.e., the pair $(\phi(t,x), \int_0^t \eta(\phi(s,x))ds)$, by solving a single augmented ODE system

$$\begin{aligned} \dot{\mathbf{x}}(t) &= f(\mathbf{x}(t)), \quad \mathbf{x}(0) = x \in \mathbb{R}^n \\ \dot{I}(t) &= \eta(\mathbf{x}), \quad I(0) = 0. \end{aligned} \tag{36}$$

We propose the following algorithm based on this technique for our purpose of evaluating $\mathcal{T}_{\Delta\mathfrak{z}}(x)$ for each $\mathfrak{z}$ and each x.

We summarize the algorithm for generating training data for one time period in Algorithm 2.

B. EDMD Algorithm for Operator Learning

After gathering training data, the rest of the training process should be the same as the existing Koopman operator learning techniques. We perform the EDMD algorithm to obtain the final approximation for $\mathcal{T}_\Delta$. EDMD algorithm provides an estimation of $\mathcal{T}_\Delta$ using one time period observation data. We use Algorithm 2 to obtain training data $(\mathbf{X}, \mathbf{Y})$. As we have briefly discussed in Section V-C, we need to find $\mathbf{T} = \operatorname{argmin}_{A\in\mathbb{C}^{N\times N}} \|\mathbf{Y} - \mathbf{X}A\|_F$ in order to obtain the approximation of the form (34). For EDMD, the $\mathbf{T}$ is given in closed-form as

$$\mathbf{T} = \left(\mathbf{X}^\mathsf{T}\mathbf{X}\right)^\dagger \mathbf{X}^\mathsf{T}\mathbf{Y} \tag{37}$$

where $\dagger$ is the pseudoinverse.

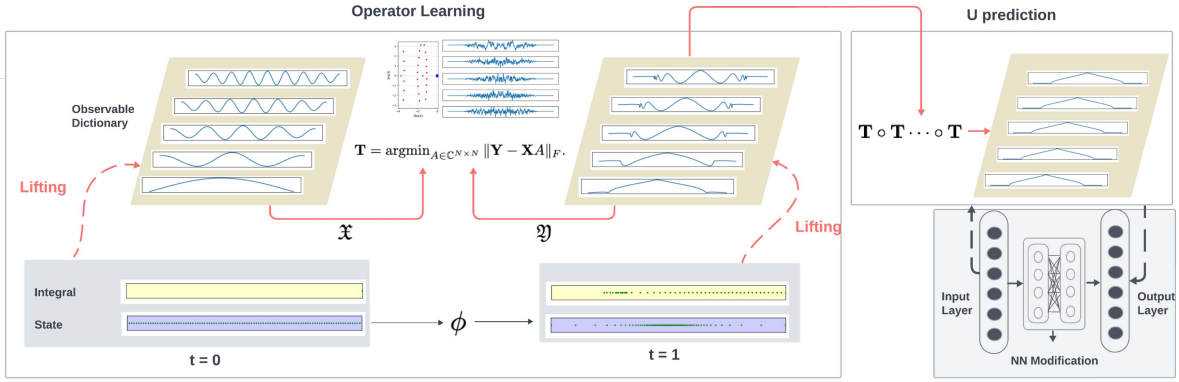

Fig. 1. Zubov–Koopman approach of ROA prediction and Lyapunov function construction.

Algorithm 2: Generating Training Data.

Require: f, η, $\mathcal{R}$, $\mathcal{Z}_N$, Δt, and a set $\{x^{(m)}\}_{m=0}^{M-1} \subset \mathcal{R}$.
1: **for** m **from** 0 **to** $M-1$ **do**
2: **for** i **from** 0 **to** $N-1$ **do**
3: Calculate $\mathfrak{z}_i(x^{(m)})$
4: Calculate $\mathcal{T}_{\Delta\mathfrak{z}_i}(x^{(m)})$ using Algorithm 1
5: **end for**
6: Stack

$$\mathcal{T}_\Delta\mathcal{Z}_N(x^{(m)}) = [\mathcal{T}_{\Delta\mathfrak{z}0}(x^{(m)}), \ldots, \mathcal{T}_{\Delta\mathfrak{z}N-1}(x^{(m)})]$$

7: **end for**
8: Stack $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{M\times N}$ such that
$\mathbf{X} = [\mathcal{Z}_N(x^{(0)}), \mathcal{Z}_N(x^{(1)}), \ldots, \mathcal{Z}_N(x^{(M-1)})]^\mathsf{T}$ and
$\mathbf{Y} = [\mathcal{T}_\Delta\mathcal{Z}_N(x^{(0)}), \mathcal{T}_\Delta\mathcal{Z}_N(x^{(1)}), \ldots, \mathcal{T}_\Delta\mathcal{Z}_N(x^{(M-1)})]^\mathsf{T}$

C. Predicting ROA and Constructing Lyapunov Function

To predict U, (35) is also ready to use. We simply pick a tolerance and the largest number of iteration K, then iterate $\mathbf{T}$ until $\|\mathbf{T}^k - \mathbf{T}^{k-1}\|_F$ reaches the threshold or $k = K$, whichever comes first. For simplicity, we can also choose the dictionary $\mathcal{Z}_N$ to include a real-valued $\mathfrak{z}_i$ such that $\mathfrak{z}_i(x_{\text{eq}}) = 1$. Then, we obtain a Zubov–Koopman approximation of U

$$U_{\text{ZK}}(\cdot) := \mathcal{Z}_N(\cdot)(\mathbf{T}^k\mathbf{w}) \tag{38}$$

where $\mathbf{w}$ is the ith standard basis vector. Since U_{ZK} and U are uniformly close, we can approximate the ROA (w.r.t. the Hausdorff metric) by the largest connected positive level sets of U_{ZK}.

As for constructing a Lyapunov function, in virtue of (30) and Remark 27, we may not directly use the result U_{ZK} just yet. Fortunately, by the viscosity regularity of U and the alternative comparison principles given in Lemma 37, one can seek a smooth function that uniformly approximates U_{ZK}, and hence uniformly approximates U, with its Lipschitz constant also converges to that of U.

We achieve this extra modification by NN using a new set of samples $\{x^{(m)}\}$ (could be different from the one used in Algorithm 2) and $\{U_{\text{ZK}}(x^{(m)})\}$. This NN approximation will be named U_{NN}. We omit this algorithm as it follows the standard procedure. The formal verification of U_{NN} utilizes satisfiability modulo theories (SMT) solvers and follows the exact procedures in [22, Sect. V].

Remark 36: Note that for any Lipschitz continuous function, the Lipschitz constant of its smooth neural approximations typically converges to the true Lipschitz constant. This phenomenon has been extensively studied by [16]. When verifying the Lipschitz constant of the residual is challenging, assuming this phenomenon holds is reasonable. ◇

We summarize the learning procedures in Fig. 1.

VII. Numerical Examples

In this section, we provide numerical examples to demonstrate the proposed Zubov–Koopman approach for predicting ROAs, as well as learning and verifying neural Lyapunov functions. For consistency, the reported computation times were recorded on a MacBook Pro equipped with an M1 Pro 10-Core CPU and 32 GB of memory. Multiprocessing was applied to all the examples, and the recorded calculation times for smaller data volumes can deviate from real values due to computational overhead. Code for the numerical examples can be found online.[5] An illustration of the overall computational approach using a 1-D example can be found in the appendix of the extended version [32] of this article. It is omitted here due to space limitations.

A. Reversed Van Der Pol Oscillator

Consider the reversed Van der Pol oscillator

$$\dot{\mathbf{x}}_1(t) = -\mathbf{x}_2(t), \quad \dot{\mathbf{x}}_2(t) = \mathbf{x}_1(t) - (1 - \mathbf{x}_1^2(t))\mathbf{x}_2(t)$$

with $\mathbf{x}(0) := [\mathbf{x}_1(0), \mathbf{x}_2(0)] = [x_1, x_2]$. The system has a stable equilibrium point at $\mathbf{0}$. We use this example to compare the data efficiency of predicting the ROA of $\{\mathbf{0}\}$ as well as finding a Lyapunov function between the Zubov–Koopman approach and the data-driven approach presented in [15].

In this example, we set $\eta(x) = (x_1^2 + x_2^2)/5$, and pick the region of interest as $\mathcal{R} = [-3, 3]^2$. A total of $M = 200^2$, 250^2, and 300^2 uniformly spaced samples $\{x^{(m)}\}$ in $\mathcal{R}$ are generated, respectively, for three parallel experiments.

[5][Online]. Available: https://github.com/Yiming-Meng/Zubov-Koopman-Operator-Learning

TABLE I
VERIFICATION OF NEURAL LYAPUNOV FUNCTIONS FOR REVERSED VAN DER POL OSCILLATOR

Approaches	Sample size	IVP solving time	Stacking time	Operator learning time	Function U learning time	Modification time (300^2 samples)	NN final loss	Verified volume
Data-driven	100×100	10.13(s)	0.81 (s)	-	54.24(s)	-	8.48×10^{-5}	84.71%
ZK + NN	100×100	7.68(s)	5.40 (s)	89.26(s)	52.91(s)	482.10 (s)	1.25×10^{-6}	91.68%
Data-driven	200×200	40.83(s)	0.20(s)	-	366.90(s)	-	1.35×10^{-6}	88.87%
ZK + NN	200×200	31.24(s)	12.59(s)	128.57 (s)	53.18(s)	482.10 (s)	1.25×10^{-6}	90.32%
Data-driven	300×300	87.06(s)	0.32(s)	-	545.02(s)	-	9.37×10^{-7}	87.09%
ZK + NN	300×300	61.69(s)	16.82(s)	178.90(s)	51.10(s)	486.67(s)	1.05×10^{-5}	89.70%

1) Zubov–Koopman Method: We simulate the trajectory up to $\Delta t = 1.5$. The dictionary $\mathcal{Z} := [\mathfrak{z}_{0,0}, \mathfrak{z}_{0,1}, \mathfrak{z}_{1,0}, \ldots, \mathfrak{z}_{i,j}, \ldots, \mathfrak{z}_{2N-1,2N-1}]$ is selected in a similar manner as in the 1-D example with adjustments for 2-D inputs, i.e., we set $\mathfrak{z}_{i,j}(x_1, x_2) = \cos(\frac{2\pi(ix_1+jx_2)}{3}) \exp\{-(x_1^2 + x_2^2)/4\}$ for each i and j. We let $N = 50$, which means a total of 9801 observable basis functions are used. To obtain U_{ZK} through iterations, we set the termination tolerance to 10^{-2} with a maximum of 8 iterations. The prediction of the ROA follows the same procedure as in of [32, Sect. VII.A]. For the construction and verification of a Lyapunov function, we need to incorporate an additional NN modification, as described in Section VI-C, due to inaccuracies in the derivatives of U_{ZK}. The training set should be independent of the one used for obtaining U_{ZK}. Therefore, for all the experiments at this stage, we use 300^2 uniformly spaced samples, along with the evaluations of U_{ZK} at those points, to train U_{NN}. We use a network with 2 hidden layers, each containing 15 neurons. We terminate training when the mean-square training loss was smaller than 10^{-8} or after 300 epochs.

2) Data-Driven Method: As for the data-driven method in [15], for each experiment, we use the same samples as above and generate trajectory data up to a termination time $t = 10$, or until $\int_0^t \eta(\phi(s, x))ds \geq 200$. We then use $\{x^{(m)}\}$ and $\{\exp\{-\int_0^t \eta(\phi(s, x^{(m)}))ds\}\}$ to train a NN U_{DD}. We use a network model with two hidden layers, each containing 15 neurons. We terminate training when the mean-square training loss is smaller than 10^{-8} or after 500 epochs. We predict the ROA as the largest connected positive level sets of U_{DD}. This neural solution U_{DD} is also ready to be verified as a Lyapunov function without extra modification.

3) Verification: Note that the approximations U_{NN} and U_{DD} are intended for Zubov's dual equation. This is in contrast with the procedure for verifying a neural Lyapunov function computing using Zubov's equation [22].

4) Experiment Results: The comparison results are reported in Table I. In the table, "IVP solving" corresponds to steps 1) to 4) of Algorithm 1, while "stacking" involves preparing the data for learning. For the data-driven method, "function learning" refers to the NN training, whereas for the Zubov–Koopman approach, it pertains to the iterative procedure to obtain U_{ZK} [see (38)]. We also recorded the NN final losses for both approaches whenever a NN was employed.

We show the predicted and verified boundaries of the ROA for $M = 100^2$ in Fig. 2. The prediction using the Zubov–Koopman approach shows overall better accuracy, while the verified regions are similar for all the methods. On the other hand, with $M = 100^2$ initial samples, the region verified by the data-driven method is slightly smaller. This is because the verification relies on NN approximation for both methods, but the Zubov–Koopman approach's NN stage is based on an already good approximation of U, which is not constrained by the initial sample size. These phenomena also reflect that the U_{ZK} and the data generated for the data-driven method have similar quality in cases where $M = 200^2$ and 300^2.

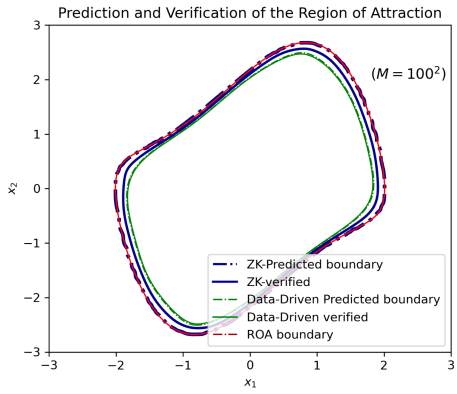

Fig. 2. Predictions and verification of the ROA for the reversed Van der Pol with Zubov–Koopman and data-driven method.

B. Polynomial System

We consider the polynomial system in [26]

$$\dot{\mathbf{x}}_1(t) = \mathbf{x}_2(t), \quad \dot{\mathbf{x}}_2(t) = -2\mathbf{x}_1(t) + \frac{1}{3}\mathbf{x}_1^3(t) - \mathbf{x}_2(t).$$

The origin of this system is known to have an unbounded domain of attraction, delimited by the stable manifolds of the saddle equilibrium points at $(\pm\sqrt{6}, 0)$ [22]. We restrict our computations to the domain $\mathcal{R} = [-6, 6]^2$, and use the proposed Zubov–Koopman approach to predict the refined ROA within $\mathrm{int}(\mathcal{R})$. We also use NN modification to construct a Lyapunov function[6] to characterize both stability and safety.

We use three experiments to investigate the effect of sample size and the width of the NN: 1) $M = 300^2$, width $= 15$; 2) $M = 500^2$, width $= 15$; 3) $M = 500^2$, width $= 30$. In all of the experiments, we choose $\Delta t = 2$, $N = 50$, and $\mathfrak{z}_{i,j}(x_1, x_2) = \cos(\frac{2\pi(ix_1+jx_2)}{6}) \exp\{-(x_1^2 + x_2^2)/25\}$ for all $i, j \in \{0, 1, \ldots, 2N-1\}$. We use 2-layer NN model with 30 neurons each for the modification stage. The verified portions of the refined ROA are 96.41%, 98.12%, and 98.50%. Fig. 3

[6]In this case, this Lyapunov function is also a Lyapunov-barrier function. Previous work [28], [30] has shown that under mild conditions, Lyapunov and barrier functions can be united and behave as a single Lyapunov function.

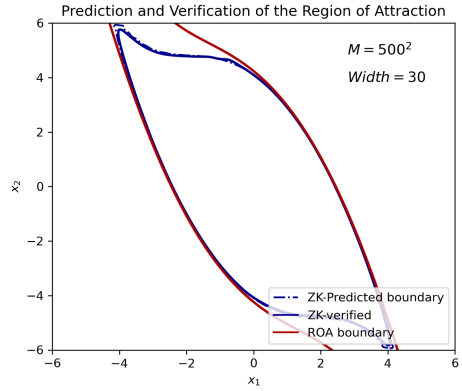

Fig. 3. Predictions and verification of the ROA with safety for the Polynomial System with Zubov–Koopman approach.

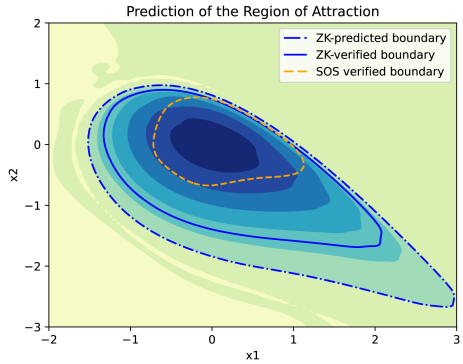

Fig. 4. Verification of the ROA with Safety for the Two-Machine Power System with Zubov–Koopman approach using $M = 300^2$ samples.

shows the predicted and verified ROA with safety guarantees for experiment (3). Although the actual ROA is unbounded, the proposed algorithm approximates the largest certifiable ROA using a neural Lyapunov function.

C. Two-Machine Power System

Consider the two-machine power system [42] modeled by

$$\dot{\mathbf{x}}_1(t) = \mathbf{x}_2(t),\ \dot{\mathbf{x}}_2(t) = -0.5\mathbf{x}_2(t) - (\sin(\mathbf{x}_1(t) + \delta) - \sin(\delta))$$

where $\delta = \frac{\pi}{3}$. We restrict our computations to the domain $\mathcal{R} = [-2, 3] \times [-3, 2]$ and use the proposed Zubov–Koopman approach to predict the refined ROA of the origin. Note that the system has an unstable equilibrium point at $(\pi/3, 0)$.

In the experiment, we choose $\Delta t = 2$, $N = 40$, $M = 300^2$, and $\mathfrak{z}_{k,l}(x_1, x_2) = \exp\{i\frac{\pi}{6}(kx_1 + lx_2)\}$ for all $k, l \in \{-(N-1), -(N-2), \ldots, N-1\}$, where i in this example is the imaginary unit, i.e., $i^2 = -1$. We use a NN with two hidden layers, 30 neurons each, for the modification stage.

The example (see Fig. 4) shows that the proposed method outperforms SOS Lyapunov functions [41] when dealing with nonpolynomial nonlinearities. Note that the verified ROA using dReal [10] may appear smaller because the verification using an SMT solver tends to be more conservative, i.e., it only returns the largest level set within which the flow is forward invariant. However, to guarantee a subset of ROA, forward invariance is not necessary. Readers can, based on their needs, use our approach as a data-driven numerical algorithm to predict (larger) ROA and roughly gauge where the boundary approximately lies.

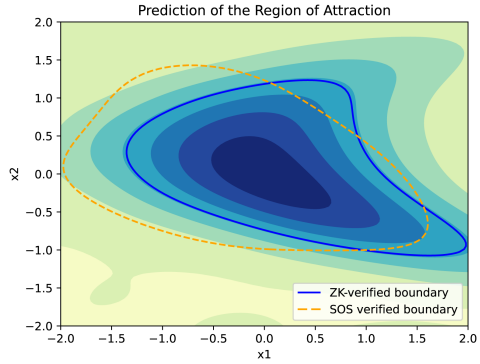

Fig. 5. Verification of the ROA with safety for the 3-D System with Zubov–Koopman approach using $M = 40^3$ samples.

D. 3-D System

We use this example to showcase the ability of the proposed method in predicting the ROA of a 3-D system. Consider

$$\begin{cases} \dot{\mathbf{x}}_1(t) = \mathbf{x}_2(t) + 2\mathbf{x}_2(t)\mathbf{x}_3(t), & \dot{\mathbf{x}}_2(t) = \mathbf{x}_3(t) \\ \dot{\mathbf{x}}_3(t) = -0.5\mathbf{x}_1(t) - 2\mathbf{x}_2(t) - \mathbf{x}_3(t). \end{cases}$$

We restrict our computations to the domain $\mathcal{R} = [-2, 2]^3$ and use the proposed Zubov–Koopman approach to predict the refined ROA of the origin. In the experiment, we choose $\Delta t = 2$, $N = 10$, $M = 40^3$, and $\mathfrak{z}_{i,j,k}(x_1, x_2, x_3) = \exp\{i\frac{\pi}{6}(ix_1 + jx_2 + kx_3)\}$ for all $i, j, k \in \{-(N-1), -(N-2), \ldots, N-1\}$. We use a NN with one hidden layer and 100 neurons for the modification stage. The verified ROA within the region of interest is $\{x \in \mathcal{R} : U_{\text{NN}} \geq 0.2\}$. We present the x_1-x_2 section by performing the projection with x_3 fixed at 0. Fig. 5 demonstrates that the verified ROA within $\mathcal{R}$ is almost the largest forward invariant set within $\mathcal{R}$. Again, it is worth mentioning that we restrict the ROA to $\mathcal{R}$, which is within the actual ROA of the system. Therefore, the resulting ROA estimates from the proposed Zubov–Koopman approach and the SOS approach are not comparable (i.e., one is not a subset of the other). However, we would like to emphasize that the proposed Zubov–Koopman approach is performed to estimate the ROA for unknown systems (using a much less dense set of samples compared to previous examples), while the SOS technique requires exact model information.

E. Stiff System I

In this example, we examine Van der Pol oscillators

$$\dot{\mathbf{x}}_1(t) = -\mathbf{x}_2(t), \quad \dot{\mathbf{x}}_2(t) = \mathbf{x}_1(t) - \mu(1 - \mathbf{x}_1^2(t))\mathbf{x}_2(t)$$

with $\mu = 4$ and 6. The value of U changes rapidly due to the high evolution speed of the state within the stiff systems. For both systems, we use $M = 300^2$ samples to determine U_{ZK} and employ a network model with two hidden layers and 15 neurons for modification. However, the rapid changes prevent the validation of a Lyapunov function. We present the color map of U_{ZK} for predicting the ROAs in Fig. 6. We can observe that as μ increases, the same number of samples becomes insufficiently dense to provide adequate information for operator learning. This results in substantial oscillations at the boundaries of sublevel regions.

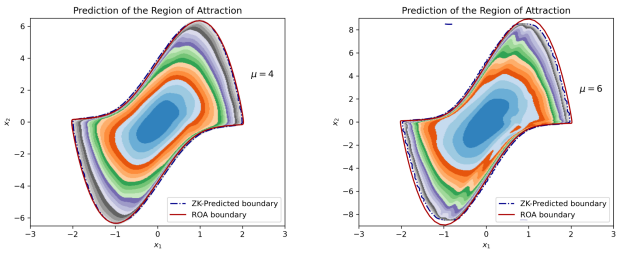

Fig. 6. Zubov–Koopman approximations of the ROAs for stiff Van der Pol systems.

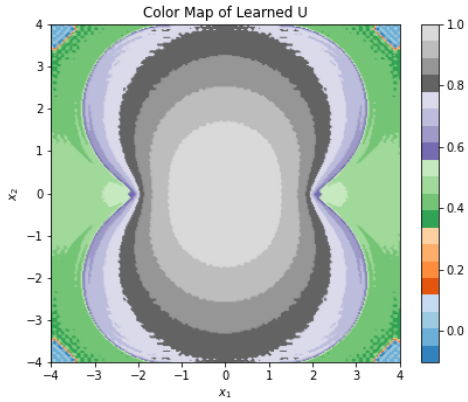

Fig. 7. Predictions and verification of the ROA with safety for the Stiff System II with Zubov–Koopman approach using $M = 300^2$ samples.

F. Stiff System II

We use this example to compare the predictability of ROA from [25, Example 3] for dynamics

$$\dot{\mathbf{x}}_1(t) = -2\mathbf{x}_1(t) + \mathbf{x}_1^2(t) - \mathbf{x}_2^2(t),\ \dot{\mathbf{x}}_2(t) = -2.5\mathbf{x}_2(t) + 2\mathbf{x}_1(t)\mathbf{x}_2(t).$$

The origin is globally stable on $\mathbb{R}^2 \setminus \{x \in \mathbb{R}^2 : x_1 \geq 2, x_2 = 0\}$. We choose $\mathcal{R} = [-4, 4]^2$ and use $M = 300^2$ samples to learn U_{ZK}. The nonzero color map represents the ROA relative to $\mathcal{R}$, as shown in Fig. 7. The learned U_{ZK} demonstrates better predictability than the local Lyapunov functions generated by conventional Koopman operators. We omit the NN-modification and formal verification in this example to make a fair comparison with [25].

VIII. Conclusion

In this article, we introduced a Zubov–Koopman approach to characterize the solution to Zubov's equation, based on regularity and convergence analysis for a broader range of nonlinear systems. The proposed operator exhibits a semigroup property, and the learning technique is similar to the conventional Koopman operator. In particular, we executed an EDMD-like algorithm using a Fourier-like dictionary of observable functions based on our understanding of the solution properties. We then devise an iterative approach that employs the learned Zubov–Koopman operator to approximate the solution to Zubov's equation, achieving a high level of precision. Compared to the existing data-driven methods, our proposed technique demonstrates a better predictability of the ROAs via numerical examples.

It is worth mentioning that, through numerical examples, we observed clear spectral gaps in the learned operators. The effectiveness of using semigroup properties with such spectral gaps to find low-dimensional approximations for characterizing the long-term fixed points of PDEs has been demonstrated in the literature [2], [31]. Similarly, this approach can potentially lead to dimension reduction for the Zubov–Koopman operators and effectively address the "curse of dimensionality" for unknown systems. Future directions will be on the spectral analysis with applications to high-dimensional physical systems.

One potential limitation of this work is its predictability of the flow in the state space. We have to rely on the knowledge of fixed points to predict the ROAs. However, recent works [27], [33], [34] have provided a way to use the same set of training data as this article for identifying system transitions, achieving improved accuracy compared to the benchmark Koopman-logarithm-based approach [24]. It would be interesting to investigate how to align this approach with dimension reduction analysis and the proposed method in this paper, so that a streamlined computational tool can be developed for situations with limited information about stable attractors. Another immediate area for future work with the Zubov–Koopman framework could involve systems with robust control or other types of measurable inputs. It would be intriguing to explore its potential in predicting controllable regions with stability and safety properties for unknown systems.

Appendix A
Fundamental Properties of Viscosity Solutions

We provide some fundamental properties of viscosity solutions in this appendix. The following lemma [1, Lemma 1.7, Lemma 1.8, Ch. I] provides some insights on $\partial^+ v(x)$ and $\partial^+ v(x)$ for some $v \in C(\mathcal{X})$.

Lemma 37 (Subdifferential and Superdifferential): Let $v \in C(\mathcal{X})$. Then, the following holds:

1) $p \in \partial^+ v(x)$ if and only if there exists $\nu \in C^1(\mathcal{X})$ such that $\nabla\nu(x) = p$ and $v - \nu$ has a local maximum at x;
2) $q \in \partial^- v(x)$ if and only if there exists $\nu \in C^1(\mathcal{X})$ such that $\nabla\nu(x) = q$ and $v - \nu$ has a local minimum at x;
3) if for some x both $\partial^+ v(x)$ and $\partial^- v(x)$ are nonempty, then $\partial^+ v(x) = \partial^- v(x) = \{\nabla v(x)\}$;
4) the sets $\{x \in \mathcal{X} : \partial^+ v(x) \neq \emptyset\}$ and $\{x \in \mathcal{X} : \partial^- v(x) \neq \emptyset\}$ are dense.

In view of (1) and (2) in Lemma 37, (1) and (2) in Definition 9 are equivalent to

1) for any $\nu \in C^1$, if x is a local maximum for $v - \nu$, then $F(x, \nu(x), \nabla\nu(x)) \leq 0$ and
2) for any $\nu \in C^1$, if x is a local minimum for $v - \nu$, then $F(x, \nu(x), \nabla\nu(x)) \geq 0$.

Appendix B
Unique Bounded Viscosity Solutions

Proof of Theorem 19: The uniqueness property within $\mathcal{D}(\mathcal{A})$ follows the result [3, Thm. 3.8] for zero perturbation of (1). Alternatively, we can suppose there exists another viscosity solution $\tilde{V}$ to $-\nabla V \cdot f(x) - \eta(x) = 0$ with $V(x_{\mathrm{eq}}) = 0$. Let $w = V - \tilde{V}$. Then, w solves $-\nabla w(x) \cdot f(x) = 0$ with $w(x_{\mathrm{eq}}) = 0$ in a viscosity sense. One can follows a similar proof as in

[1, Ch. II, Prop. 5.18] to show that $w(\phi(t,x)) = w(x)$ for all $t \geq 0$ and all $x \in \mathcal{D}(\mathcal{A})$. By the stability assumption, one has $w \equiv w(x_{\text{eq}}) = 0$, which shows that V is the unique viscosity solution on $\mathcal{D}(\mathcal{A})$. For any valid h, by the definition (10) and by [1, Ch. 2, Prop. 2.5], we immediately have that U_h is the unique viscosity solution to (12) on $\mathcal{D}(\mathcal{A})$.

Now we suppose $\mathcal{D}(\mathcal{A}) \neq \mathbb{R}^n$. Then, $\partial\mathcal{D}(\mathcal{A}) \neq \emptyset$ and is bounded by the assumption. By the above uniqueness argument within $\mathcal{D}(\mathcal{A})$, and by the construction of U_h, all viscosity solutions to (12) should be equal to 0 on $\partial\mathcal{D}(\mathcal{A})$. We then work on $\mathbb{R}^n \setminus \overline{\mathcal{D}(\mathcal{A})}$ and verify that 0 is the unique bounded viscosity solution to (12).

Let $\Omega = \mathbb{R}^n \setminus \overline{\mathcal{D}(\mathcal{A})}$, then, $\partial\Omega = \partial\mathcal{D}(\mathcal{A})$. Suppose u is another bounded viscosity solution to (12) and let $\Psi(x,y) = u(x) - |x-y|^2/(2\varepsilon)$ for any $\varepsilon > 0$. Then, $\lim_{|x|+|y|\to\infty} \Psi(x,y) = -\infty$. By the continuity of Ψ, there exists $x_\varepsilon, y_\varepsilon \in \overline{\Omega}$ such that $\Psi(x_\varepsilon, y_\varepsilon) = \max_{\overline{\Omega}\times\overline{\Omega}} \Psi(x,y)$. Then, for any $\varepsilon > 0$, $\max_{x\in\overline{\Omega}} u(x) = \max_{x\in\overline{\Omega}} \Psi(x,x) \leq \Psi(x_\varepsilon, y_\varepsilon) \leq u(x_\varepsilon)$. In addition, by the exact same argument as in [1, Ch. II, Thm. 3.1], one has that $|x_\varepsilon - y_\varepsilon| \to 0$ and $\frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \to 0$ as $\varepsilon \to 0$.

Now we argue that $u(x_\varepsilon) \leq 0$ as $\varepsilon \to 0$, which will imply that $u(x) \leq 0$ on Ω. If $x_\varepsilon \in \partial\Omega$, then $u(x_\varepsilon) = 0$. If $y_\varepsilon \in \partial\Omega$, then $u(x_\varepsilon) = u(x_\varepsilon) - u(y_\varepsilon) \to 0$ by the continuity of u. If $x_\varepsilon, y_\varepsilon \in \Omega$, set $\psi_1(y) := u(x_\varepsilon) - \frac{|x_\varepsilon - y|^2}{2\varepsilon}$ and $\psi_2(x) := \frac{|x-y_\varepsilon|^2}{2\varepsilon}$. Then, x_ε is a local maximum for $u - \psi_2$ and y_ε is a local minimum for $U_h - \psi_1$. Note that $\nabla\psi_1(y_\varepsilon) = \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} = \nabla\psi_2(x_\varepsilon)$. In addition, since u and U_h are viscosity (sub/supper) solutions on Ω, we have that $\eta(x_\varepsilon)u(x_\varepsilon) - \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} \cdot f(x_\varepsilon) \leq 0$ and $\frac{x_\varepsilon - y_\varepsilon}{\varepsilon} \cdot f(y_\varepsilon) \leq 0$. This implies $u(x_\varepsilon) \leq \frac{x_\varepsilon - y_\varepsilon}{\varepsilon\eta(x_\varepsilon)} \cdot (f(x_\varepsilon) - f(y_\varepsilon))$ since $\eta(x_\varepsilon)$ is bounded below by a positive number (recall that x_ε is bounded away from 0 on Ω). By the Lipschitz continuity of f, we have $u(x_\varepsilon) \leq 0$ as $\varepsilon \to 0$. The other side, i.e., $u(x) \geq 0$ on Ω, can be proved in the same manner. Therefore, $u(x) = 0$ on Ω. ■

APPENDIX C
FUNDAMENTAL PROPERTIES OF ZUBOV–KOOPMAN

In the appendix, we complete the proof for Proposition 24.

Proof: It is clear that $\mathcal{T}_0 h(x) = h(x)$. We also have the following identities:

$$
\begin{aligned}
&\mathcal{T}_s \circ \mathcal{T}_t h(x) \\
&= \exp\{-v_s(x)\}\, \mathcal{T}_t h(\phi(s,x)) \\
&= \exp\{-v_s(x)\} \exp\left\{-\int_0^t \eta(\phi(r,\phi(s,x)))dr\right\} h(\phi(t,\phi(s,x))) \\
&= \exp\{-v_s(x)\} \exp\left\{-\int_0^t \eta(\phi(r+s,x)))dr\right\} h(\phi(t+s,x)) \\
&= \exp\{-v_s(x)\} \exp\left\{-\int_s^{t+s} \eta(\phi(\sigma,x))d\sigma\right\} h(\phi(t+s,x)) \\
&= \exp\left\{-\int_0^{t+s} \eta(\phi(\sigma,x))dr\right\} h(\phi(t+s,x)) = \mathcal{T}_{s+t} h(x)
\end{aligned}
$$

which show the associativity of $\{\mathcal{T}_t\}_{t\geq 0}$. The strong continuity of $\{\mathcal{T}_t\}_{t\geq 0}$ can be verified straightforwardly by Definition 2. For the second part of the proof, we apply dynamic programming and use the similar technique as above. We omit the details due to the similarity. ■

REFERENCES

[1] M. Bardi and I. C. Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, vol. 12. Berlin, Germany: Springer, 1997.

[2] D. Blömker and M. Hairer, "Multiscale expansion of invariant measures for spdes," *Commun. Math. Phys.*, vol. 251, no. 3, pp. 515–555, 2004.

[3] F. Camilli, L. Grüne, and F. Wirth, "A generalization of Zubov's method to perturbed systems," *SIAM J. Control Optim.*, vol. 40, no. 2, pp. 496–515, 2001.

[4] F. H. Clarke, Y. S Ledyaev, and R. J. Stern, "Asymptotic stability and smooth Lyapunov functions," *J. Differ. Equ.*, vol. 149, no. 1, pp. 69–114, 1998.

[5] J. B. Conway, *A Course in Functional Analysis*, vol. 96. Berlin, Germany: Springer, 2019.

[6] C. Dawson, S. Gao, and C. Fan, "Safe control with learned certificates: A survey of neural Lyapunov, barrier, and contraction methods for robotics and control," *IEEE Trans. Robot.*, vol. 39, no. 3, pp. 1749–1767, Jun. 2023.

[7] S. A. Deka, A. M. Valle, and C. J. Tomlin, "Koopman-based neural Lyapunov functions for general attractors," in *Proc. IEEE Conf. Decis. Control*, 2022, pp. 5123–5128.

[8] L. C. Evans, *Partial Differential Equations*, vol. 19. Providence, RI, USA: American Mathematical Society, 2010.

[9] M. Farsi and J. Liu, *Model-Based Reinforcement Learning: From Data to Actions With Python-Based Toolbox*. Hoboken, NJ, USA: Wiley-Blackwell, 2023.

[10] S. Gao, S. Kong, and E. M. Clarke, "Dreal: An SMT solver for nonlinear theories over the reals," in *Proc. Int. Conf. Automated Deduction*, 2013, pp. 208–214.

[11] P. Giesl and S. Hafstein, "Review on computational methods for Lyapunov functions," *Discrete Continuous Dyn. Syst.-B*, vol. 20, no. 8, 2015, Art. no. 2291.

[12] L. Grüne, "Computing Lzyapunov functions using deep neural networks," *J. Comput. Dyn.*, vol. 8, no. 2, pp. 131–152, 2021.

[13] K. H. Khalil, *Nonlinear Systems*. Upper Saddle River, NJ, USA: Pearson, 2001.

[14] M. Jones and M. M Peet, "Converse Lyapunov functions and converging inner approximations to maximal regions of attraction of nonlinear systems," in *Proc. IEEE Conf. Decis. Control*, 2021, pp. 5312–5319.

[15] W. Kang, K. Sun, and L. Xu, "Data-driven computational methods for the domain of attraction and Zubov's equation," *IEEE Trans. Autom. Control*, vol. 69, no. 3, pp. 1600–1611, Mar. 2024.

[16] G. Khromov and S. P. Singh, "Some fundamental aspects about Lipschitz continuity of neural network functions," in *Proc. Int. Conf. Learn. Representations*, 2024.

[17] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proc. Nat. Acad. Sci.*, vol. 17, no. 5, pp. 315–318, 1931.

[18] M. D. Kvalheim and S. Revzen, "Existence and uniqueness of global Koopman eigenfunctions for stable fixed points and periodic orbits," *Physica D: Nonlinear Phenomena*, vol. 425, 2021, Art. no. 132959.

[19] Y. Lan and I. Mezić, "Linearization in the large of nonlinear systems and Koopman operator spectrum," *Physica D: Nonlinear Phenomena*, vol. 242, no. 1, pp. 42–53, 2013.

[20] Y. Li and J. Liu, "Robustly complete synthesis of memoryless controllers for nonlinear systems with reach-and-stay specifications," *IEEE Trans. Autom. Control*, vol. 66, no. 3, pp. 1199–1206, Mar. 2021.

[21] Y. Lin, E. D Sontag, and Y. Wang, "A smooth converse Lyapunov theorem for robust stability," *SIAM J. Control Optim.*, vol. 34, no. 1, pp. 124–160, 1996.

[22] J. Liu, Y. Meng, M. Fitzsimmons, and R. Zhou, "Towards learning and verifying maximal neural Lyapunov functions," in *Proc. IEEE Conf. Decis. Control*, 2023, pp. 8012–8019.

[23] J. Liu, Y. Meng, M. Fitzsimmons, and R. Zhou, "Physics-informed neural network Lyapunov functions: PDE characterization, learning, and verification," *Automatica*, vol. 175, 2025, Art. no. 112193.

[24] A. Mauroy and J. Goncalves, "Koopman-based lifting techniques for nonlinear systems identification," *IEEE Trans. Autom. Control*, vol. 65, no. 6, pp. 2550–2565, Jun. 2020.

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

[26] A. Mauroy and I. Mezić, "Global stability analysis using the eigenfunctions of the Koopman operator," *IEEE Trans. Autom. Control*, vol. 61, no. 11, pp. 3356–3369, Nov. 2016.
[27] Y. Meng, H. Li, M. Ornik, and X. Li, "Koopman-based data-driven techniques for adaptive cruise control system identification," in *Proc. 27th IEEE Int. Conf. Intell. Transp. Syst.*, 2024, pp. 849–855.
[28] Y. Meng, Y. Li, M. Fitzsimmons, and J. Liu, "Smooth converse Lyapunov-barrier theorems for asymptotic stability with safety constraints and reach-avoid-stay specifications," *Automatica*, vol. 144, 2022, Art. no. 110478.
[29] Y. Meng, Y. Li, and J. Liu, "Control of nonlinear systems with reach-avoid-stay specifications: A Lyapunov-barrier approach with an application to the Moore-Greizer model," in *Proc. Amer. Control Conf.*, 2021, pp. 2284–2291.
[30] Y. Meng and J. Liu, "Lyapunov-barrier characterization of robust reach–avoid–stay specifications for hybrid systems," *Nonlinear Anal.: Hybrid Syst.*, vol. 49, 2023, Art. no. 101340.
[31] Y. Meng, N. S. Namachchivaya, and N. Perkowski, "Hopf bifurcations of Moore-Greitzer pde model with additive noise," *J. Nonlinear Sci.*, vol. 33, no. 5, 2023, Art. no. 74.
[32] Y. Meng, R. Zhou, and J. Liu, "Learning regions of attraction in unknown dynamical systems via Zubov-Koopman lifting: Regularities and convergence," 2023, *arXiv:2311.15119*.
[33] Y. Meng, R. Zhou, M. Ornik, and J. Liu, "Koopman-based learning of infinitesimal generators without operator logarithm," in *Proc. IEEE Conf. Decis. Control*, 2024, pp. 8302–8307.
[34] Y. Meng, R. Zhou, M. Ornik, and J. Liu, "Resolvent-type data-driven learning of generators for unknown continuous-time dynamical systems," 2024, *arXiv:2411.00923*.
[35] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dyn.*, vol. 41, pp. 309–325, 2005.
[36] I. Mezić, "Koopman operator, geometry, and learning of dynamical systems," *Not. Am. Math. Soc.*, vol. 68, no. 7, pp. 1087–1105, 2021.
[37] B. Oksendal, *Stochastic Differential Equations: An Introduction With Applications*. Berlin, Germany: Springer, 2013.
[38] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.
[39] P. J. Schmid, "Application of the dynamic mode decomposition to experimental data," *Experiments Fluids*, vol. 50, pp. 1123–1130, 2011.
[40] A. R. Teel and L. Praly, "A smooth Lyapunov function from a class-estimate involving two positive semidefinite functions," *ESAIM: Control, Optimisation Calculus Variations*, vol. 5, pp. 313–367, 2000.
[41] U. Topcu, A. Packard, P. Seiler, and G. Balas, "Help on sos [ask the experts]," *IEEE Control Syst. Mag.*, vol. 30, no. 4, pp. 18–23, Aug. 2010.
[42] A. Vannelli and M. Vidyasagar, "Maximal Lyapunov functions and domains of attraction for autonomous nonlinear systems," *Automatica*, vol. 21, no. 1, pp. 69–80, 1985.
[43] 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, pp. 1307–1346, 2015.
[44] B. Yi and I. R. Manchester, "On the equivalence of contraction and Koopman approaches for nonlinear stability and control," *IEEE Trans. Autom. Control*, vol. 69, no. 7, pp. 4336–4351, Jul. 2024.
[45] E. Zeidler, *Applied Functional Analysis: Applications to Mathematical Physics*, vol. 108. Berlin, Germany: Springer, 2012.
[46] Z. Zeng, Z. Yue, A. Mauroy, J. Gonçalves, and Y. Yuan, "A sampling theorem for exact identification of continuous-time nonlinear dynamical systems," *IEEE Trans. Autom. Control*, vol. 69, no. 12, pp. 8402–8417, Dec. 2024.
[47] R. Zhou, T. Quartz, H. De Sterck, and J. Liu, "Neural Lyapunov control of unknown nonlinear systems with stability guarantees," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 29113–29125, 2022.
[48] Q. J. Zhu, "Lower semicontinuous Lyapunov functions and stability," *J. Nonlinear Convex Anal.*, vol. 4, 2003, Art. no. 325.
[49] V. I. Zubov, *Methods AM Lyapunov Their Application*, vol. 4439, Washington, D.C., USA: US Atomic Energy Commission, 1961.

Yiming Meng received the bachelor's degree in chemical engineering from Tianjin University, Tianjin, China, in 2013, and the M.Sc. degree in process system engineering from Chemical Engineering Department, Imperial College London, London, U.K., in 2017, and the Ph.D. degree in applied mathematics from the University of Waterloo, Waterloo, ON, Canada, in 2022.

He was a Postdoctoral Fellow with the Department of Applied Mathematics, University of Waterloo between 2022 and 2023. Currently, he is a Postdoctoral Research Associate with the Coordinated Science laboratory, University of Illinois Urbana-Champaign, Champaign, IL, USA. His main research interests include theory and applications of stochastic systems and control, including scientific computing and rigorous AI-enhanced computational methods for control design and autonomy in complex dynamical systems and other physical sciences.

Ruikun Zhou received the bachelor's degree in vehicle engineering from Chongqing University, Chongqing, China, in 2014, and the M.ASc degree in mechanical engineering from the University of Ottawa, Ottawa, ON, Canada, in 2020. He is currently working toward the Ph.D. degree in applied mathematics with the University of Waterloo, Waterloo, ON.

His research interests include intersection of machine learning and control systems, including learning-based stability analysis with formal guarantees, and learning-based control with applications in robotics.

Jun Liu (Senior Member, IEEE) received the B.S. degree in applied mathematics from Shanghai Jiao-Tong University, Shanghai, China, in 2002, the M.S. degree in mathematics from Peking University, Beijing, China, in 2005, and the Ph.D. degree in applied mathematics from the University of Waterloo, Waterloo, ON, Canada, in 2011.

He held an NSERC Postdoctoral Fellowship in control and dynamical systems with Caltech between 2011 and 2012. He was a Lecturer in control and systems engineering with the University of Sheffield, Sheffield, U.K., between 2012 and 2015. In 2015, he joined the Faculty of Mathematics with the University of Waterloo, where he is currently a Professor of Applied Mathematics and directs the Hybrid Systems Laboratory. His main research interests include theory and applications of hybrid systems and control, including rigorous computational methods for control design with applications in cyber-physical systems and robotics.

Dr. Liu was the recipient of the Marie-Curie Career Integration Grant from the European Commission in 2013, a Canada Research Chair from the Government of Canada in 2017, an Early Researcher Award from the Ontario Ministry of Research, Innovation and Science in 2018, and an Early Career Award from the Canadian Applied and Industrial Mathematics Society and Pacific Institute for the Mathematical Sciences (CAIMS/PIMS) in 2020, and the best paper awards include the Zhang Si-Ying Outstanding Youth Paper Award (2010, 2015) and the Nonlinear Analysis: Hybrid Systems Paper Prize (2017). He is a Member of SIAM and a Lifetime Member of CAIMS.