Dynamic Mode Decomposition of Control-Affine Nonlinear Systems using Discrete Control Liouville Operators

Zachary Morrison, Moad Abudia, Joel A. Rosenfeld, and Rushikesh Kamalapurkar

***Abstract*— The representation of a nonlinear dynamical system as infinite-dimensional linear operator over a Hilbert space of functions enables the study of the nonlinear system via pseudo-spectral analysis of the corresponding operator. In this paper, we develop a novel operator representation of discrete-time, control-affine nonlinear dynamical systems. We also demonstrate that this representation can be used to predict the behavior of the closed-loop system in response to a given feedback law. The representation is learned using recorded snapshots of the system state resulting from arbitrary, potentially open-loop control inputs. We thereby extend the predictive capabilities of dynamic mode decomposition to discrete-time nonlinear systems that are affine in control. We validate the method using two numerical experiments by predicting the response of a controlled Duffing oscillator to a known feedback law. The advantages of the developed method relative to existing techniques in the literature are also demonstrated.**

I. Introduction

In this paper, a novel representation of discrete-time control-affine nonlinear systems as infinite-dimensional linear operators over reproducing kernel Hilbert spaces (RKHSs) is introduced. This effort is inspired by the method first developed in [1], which introduced similar operator representations for *continuous-time* dynamical systems. The idea of representing a nonlinear system as an infinite-dimensional linear operator in Hilbert space was first put forth by B.O. Koopman in [2] and the resulting composition operator is aptly known as the Koopman operator. This higher-dimensional space is typically referred to as the *feature space* or *lifted space* and the Koopman operator acts as a composition operator on the lifted space. In recent years, dynamic mode decomposition (DMD) and other data-driven methods have seen a resurgence due to the abundance of data and increased the availability of computational power [3]. An example of the application of DMD can be seen in the fluid mechanics community, where modal decomposition of fluid flows is accomplished [4], [5]. In a more general sense, DMD is intimately connected to the Koopman operator as DMD is one method used to approximate the Koopman operator associated with the dynamical system [5], [6].

The Koopman approach is amenable to spectral methods in linear operator theory in certain cases, e.g. see [7], but spectral convergence cannot be guaranteed in general; therefore, Koopman DMD methods are pseudo-spectral numerical methods. Despite this theoretical limitation, Koopman DMD and Liouville DMD methods in both continuous and discrete time have been shown to exhibit remarkable predictive accuracy over finite-time horizons [3]. Moreover, Koopman DMD allows one to study dynamical systems without direct knowledge of the dynamics, as Koopman DMD is strictly data driven and requires no knowledge of the dynamical system [3]. For measurements corrupted by noise or in the case of stochastic systems, robust approximations of the Koopman operator can be formulated [8]. The ultimate goal of DMD is to develop a data-driven model via an eigendecomposition of the Koopman operator, under the assumption that the full-state observable (the identity function) is in the span of the eigenfunctions [7]. The addition of control adds greater difficulty to data-driven methods like DMD, as the Koopman operator associated with the dynamical system depends upon the control input. Furthermore, in discrete time, the Koopman operator is generally not linear in its symbol, which makes separating the influence of the controller from the drift dynamics challenging. Despite the difficulty, there have been several successful methods for generalizing Koopman DMD for dynamical systems with control in results such as [9], [10]–[14].

The method presented in [11] yields a DMD routine to represent a general nonlinear system with control as a control-affine linear system. This idea is generalized in [10] with extended DMD (eDMD), providing greater predictive power. Furthermore, for a general discrete-time, nonlinear dynamical system with control, the authors in [10] utilize the shift operator to describe the time evolution of the control signal. Also, in discrete-time, separation of the control input from the state can be achieved via first order approximations [15]. For continuous-time dynamical systems, the Koopman canonical transform (see [16]) is used in [9] to leverage a formulation of the dynamical system in the lifted space as a control-affine, bilinear system, called the Koopman bilinear form (KBF). The KBF is then amenable to the design of feedback laws using techniques from optimal control.

The aforementioned methods demonstrate the ability to predict the response of both discrete-time [10] and continuous-time [9] dynamical systems to open-loop inputs. The algorithm developed in this paper offers an advantage over the methods from [9]–[15], since in addition to a predictive model, it also estimates eigenfunctions, and consequently, a Koopman invariant subspace of the *closed-loop system*. A key contribution here is the extension of the method presented in [1] to the discrete-time case. The operator representation presented in [1] relies on linearity of differential and multiplication operators to separate the influence of the controlled and the uncontrolled part of the system dynamics. In discrete time, the differential operators need to be replaced by composition operators, and composition operators are typically not linear in their symbol. Herein lies the difficulty of extending continuous-time DMD results

to discrete-time DMD results, as separation of the effect of the control input from the effect of the drift dynamics is nontrivial.

In this paper, we take an operator-theoretic approach to DMD with a novel operator definition that accounts for the effect of control and the discrete-time nature of the problem. The algorithm is referred to as discrete control Liouville DMD or DCLDMD for brevity. To accomplish DCLDMD, the discrete, nonlinear dynamical system is represented as a composition of two operators acting on a Hilbert space of functions. The first operator mimics the effects of a composition operator, which maps from a reproducing kernel Hilbert space (RKHS) to a vector-valued RKHS (vvRKHS). In order to account for the effect of control, we make use of a multiplication operator which maps functions in the vvRKHS back into the RKHS. In doing so, we obtain an approximate representation of the dynamical system as a composition of the aforementioned operators.

The paper is organized into the following sections. Section II establishes the mathematical background for dynamic mode decomposition with discrete control Liouville operators. Section III contains the problem description. Section IV provides the derivation for discrete control Liouville dynamic mode decomposition, as well as outlining the DCLDMD algorithm. Section VII contains the numerical experiments involving the Duffing oscillator. Lastly, section VIII concludes the paper.

II. Background

In this section, we provide a brief overview of RKHSs and vvRKHSs and their role in DCLDMD.

Definition 1: An RKHS $\tilde{H}$ over a set $X \subset \mathbb{R}^n$ is a Hilbert space of functions $f : X \to \mathbb{C}$ such that for all $x \in X$ the evaluation functional $E_x f \coloneqq f(x)$ is bounded. By the Riesz representation theorem, there exists a function $\tilde{K}_x \in \tilde{H}$ such that $f(x) = \langle f, \tilde{K}_x \rangle_{\tilde{H}}$ for all $f \in \tilde{H}$.

The snapshots of a dynamical system are embedded into an RKHS via a kernel map $x \mapsto \tilde{K}(\cdot, x) \coloneqq \tilde{K}_x$. Moreover, the span of the set $\{\tilde{K}_x : x \in X\}$ is dense in $\tilde{H}$.

Proposition 1: If $A \coloneqq \{\tilde{K}_x : x \in X\}$, then span $A = \tilde{H}$.

Proof: To show that the span of the set $\{\tilde{K}_x : x \in X\}$ is dense in $\tilde{H}$ amounts to showing that $(A^\perp)^\perp = \tilde{H}$. Let $h \in A^\perp$, then $\langle h, \tilde{K}_x \rangle = h(x) = 0$. Hence $h \equiv 0$ on X. Thus $A^\perp = \{0\}$. ∎

In order to account for the effect of control, we make use of a vvRKHS.

Definition 2: Let $\mathcal{Y}$ be a Hilbert space, and let H be a Hilbert space of functions from a set X to $\mathcal{Y}$. The Hilbert space H is a *vvRKHS* if for every $\bar{u} \in \mathcal{Y}$ and $x \in X$, the functional $f \mapsto \langle f(x), \bar{u} \rangle_{\mathcal{Y}}$ is bounded.

To each $x \in X$ and $\bar{u} \in \mathcal{Y}$, we can associate a linear operator over a vvRKHS given by $(x, \bar{u}) \mapsto K_{x,\bar{u}}$, following [1]. The function $K_{x,\bar{u}}$ is known as the kernel operator and the span of these functions constitutes a dense set in the respective vvRKHS [1, Proposition 1]. Given a function $f \in H$, the reproducing property of $K_{x,\bar{u}}$ implies $\langle f, K_{x,\bar{u}} \rangle_H = \langle f(x), \bar{u} \rangle_{\mathcal{Y}}$. For more discussion on vvRKHSs see [17].

III. Problem Statement

Consider a control-affine, discrete-time dynamical system of the form

$$x_{k+1} = F(x_k) + G(x_k)u_k, \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $F : \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are functions corresponding to the drift dynamics and the control effectiveness, respectively. We refer to the individual functions which comprise the columns of G by $G_j : \mathbb{R}^n \to \mathbb{R}^n$, for $1 \leq j \leq m$. Given a feedback law $\mu : \mathbb{R}^n \to \mathbb{R}^m$ and a set of data points $\{(x_k, x_{k+1}, u_k)\}_{k=1}^n$, where u_k are arbitrary (potentially open-loop) control inputs, the goal is to predict the response of the system in (1) to the feedback law μ.

In this paper, the set X is selected to be a compact subset of $\mathbb{R}^n$, the set $\mathcal{Y}$ is selected to be $\mathbb{C}^{1\times(m+1)}$, $\tilde{H}$ denotes an RKHS of continuous functions from X to $\mathbb{C}$, and H denotes a vvRKHS of continuous functions from X to $\mathcal{Y}$.

IV. Operator representation of controlled discrete-time systems

A linear operator can be associated with the dynamical system in (1), as a composition of two operators: a composition-like discrete Liouville operator and a multiplication operator. This operator representation is derived in this section.

A. A Composition-like Kernel Propagation Operator

The technical lemma below and the proposition that follows are needed for the kernel propagation operator to be well-defined.

Lemma 1: The set $\Omega \subset H$, defined as $\Omega \coloneqq \{K_{x,\bar{u}} : x \in X, u \in \mathbb{R}^m,$ and $\bar{u} \coloneqq \begin{pmatrix}1 & u^\top\end{pmatrix} \in \mathcal{Y}\}$, satisfies $\Omega^\perp = \{0\}$.

Proof: Let $h \in \Omega^\perp$. The reproducing property of $K_{x,\bar{u}}$ implies that for all $u \in \mathbb{R}^m$ and $x \in X$, $\left\langle h(x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = \langle h, K_{x,\bar{u}} \rangle_H$. Since $h \in \Omega^\perp$ and $K_{x,\bar{u}} \in \Omega$, we conclude that $\langle h, K_{x,\bar{u}} \rangle_H = 0$. As a result, for each fixed $x \in X$ and for all $u \in \mathbb{R}^m$, we have $\left\langle h(x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = 0$. That is, $\left\langle \Re\left(h(x)\right), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} + \mathrm{i} \left\langle \Im\left(h(x)\right), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = 0$, where $\Re\left(h(x)\right)$ denotes the real part of $h(x)$ and $\Im\left(h(x)\right)$ denotes the imaginary part of $h(x)$. Since the only such $h(x) \in \mathcal{Y}$ is the zero vector, we conclude that $h = 0$. ∎

Proposition 2: Let $L_z \in H$ be a function such that for all tuples (x, u, y) satisfying $y = F(x) + G(x)u$, we have

$$\left\langle [L_z](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = \tilde{K}_z(y). \tag{2}$$

For all $z \in X$, the map $\tilde{K}_z \mapsto L_z$ is a well-defined operator.

Proof: For a given $z \in X$, suppose there are two functions, L_z^1 and L_z^2, each of which satisfy (2) given above. Then, for any tuple (x, y, u) which satisfies $y = F(x) + G(x)u$,

$$\left\langle [L_z^1](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = \left\langle [L_z^2](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}},$$

and therefore, $\left\langle [L_z^1](x) - [L_z^2](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = 0$.

Using the reproducing property,

$$\left\langle [L_z^1] - [L_z^2], K_{x,\bar{u}} \right\rangle_H = 0 \tag{3}$$

for all vectors in the set $\Omega \coloneqq \{K_{x,\bar{u}} : x \in X$ and $\bar{u} \in \mathcal{Y} \mid \bar{u} = \begin{pmatrix}1 & u^\top\end{pmatrix}, u \in \mathbb{R}^m\}$.

As a result, $[L_z^1] - [L_z^2] \in \Omega^\perp$, where $\perp$ denotes the orthogonal complement of $\Omega \subset \mathcal{Y}$.

Since $\Omega^\perp = \{0\}$ according to lemma 1, we conclude that for all $z \in X$, $[L_z^1] = [L_z^2]$. That is, the operator $\tilde{K}_z \mapsto L_z$ is well defined on the set $\{\tilde{K}_z\}_{z\in X}$. Linearity of the operator then implies that it is also well-defined on $D \coloneqq \{h + \mathrm{i}g \in \tilde{H} \mid h, g \in \operatorname{span}\{\tilde{K}_z\}_{z\in X}\}$. ∎

Definition 3: Let $A_{F,G} : \mathcal{D}(A_{F,G}) \to H$ be the operator with domain $\mathcal{D}(A_{F,G}) \coloneqq D$ that maps, for each $z \in X$, the function $\tilde{K}_z$ to a function $[A_{F,G}\tilde{K}_z] \in H$ such that for all tuples (x, u, y) satisfying $y = F(x) + G(x)u$, we have

$$\left\langle \left[A_{F,G}\tilde{K}_z\right](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = \tilde{K}_z(y). \tag{4}$$

A few remarks regarding definition 3 are in order. The kernel propagation operator $A_{F,G}$ is composition-like in the sense that if a linear kernel $\tilde{K}_z(x) = z^\top x$ is used, one could define $A_{F,G}$ explicitly as $A_{F,G}\tilde{K}_z = [\tilde{K}_z(F(\cdot)), \tilde{K}_z(G_1(\cdot)), \cdots, \tilde{K}_z(G_m(\cdot))]$. In that case, due to linearity of the kernel, $\left\langle \left[A_{F,G}\tilde{K}_z\right](x), \begin{pmatrix}1 & u^\top\end{pmatrix} \right\rangle_{\mathcal{Y}} = \tilde{K}_z(F(x)) + \sum_{j=1}^m \tilde{K}_z(G_j(x))u_j = \tilde{K}_z(y)$. That is, similar to the Koopman operator for autonomous systems, the operator $A_{F,G}$, when composed with the inner product operation in a RKHS with a linear kernel, propagates the observable $\tilde{K}_z$ one step forward in time.

In the case of nonlinear kernels, an explicit expression for the operator $A_{F,G}$ cannot be derived. However, the implicit definition above, which achieves one-step propagation of the kernels by definition, is still useful for DMD.

Since the set D is dense in H, the kernel propagation operator $A_{F,G}$ is densely defined. As such, the adjoint $A_{F,G}^*$ exists and can be defined through its domain.

Definition 4: The domain of the adjoint $A_{F,G}^*$ of $A_{F,G}$ is defined as $\mathcal{D}(A_{F,G}^*) \coloneqq \{f \in H : h \mapsto \langle A_{F,G}h, f\rangle_H$ is bounded on $\mathcal{D}(A_{F,G})\}$.

Note that for all $x \in X$ and $\bar{u} \in \mathcal{Y}$, the kernel functions $K_{x,\bar{u}}$ of H are in the domain of the adjoint $A_{F,G}^*$. Indeed, if $A_{F,G}h \in H$, $\langle A_{F,G}h, K_{x,\bar{u}}\rangle$ is bounded by definition 2 and hence $K_{x,\bar{u}} \in \mathcal{D}(A_{F,G}^*)$.

B. Multiplication Operators

Let $\nu : X \to \mathbb{R}^{1\times(m+1)} \subset \mathcal{Y}$ be a continuous function. The multiplication operator with symbol ν is denoted as $M_\nu : \mathcal{D}(M_\nu) \to \tilde{H}$. For a function $h \in \mathcal{D}(M_\nu)$, we define the action of the multiplication operator on h as

$$[M_\nu h](\cdot) = \langle h(\cdot), \nu(\cdot)\rangle_{\mathcal{Y}},$$

where the domain of the multiplication operator is given as

$$\mathcal{D}(M_\nu) \coloneqq \{h \in H \mid x \mapsto \langle h(x), \nu(x)\rangle_{\mathcal{Y}} \in \tilde{H}\}.$$

For completeness, we recall the interaction between multiplication operators and kernel operators from [1]. The interaction is used to calculate the finite-rank representation of the composition of the multiplication operator with the kernel propagation operator from Definition 3.

Proposition 3: Suppose that $\nu : X \to \mathbb{R}^{1\times(m+1)}$ corresponds to a densely defined multiplication operator $M_\nu : \mathcal{D}(M_\nu) \to \tilde{H}$ and $\tilde{K} : X \times X \to \mathbb{R}$ is the kernel function of the RKHS $\tilde{H}$. Then, for all $x \in X$, $\tilde{K}_x \in \mathcal{D}(M_\nu^*)$, where M_ν^* is the adjoint of M_ν, and $M_\nu^*\tilde{K}_x = K_{x,\nu(x)}$.

The composition of the kernel propagation operator from Definition 3 and the multiplication operator can be used to define the discrete control Liouville operator.

C. The Discrete Control Liouville Operator

Taking the composition of $A_{F,G}$ and $M_{(1,\mu^\top)}$, for a known feedback law $\mu : \mathbb{R}^n \to \mathbb{R}^m$, the evolution of an observable along trajectories of the dynamical system can be described in terms of an infinite-dimensional linear operator.

Definition 5: Let $\nu \coloneqq \begin{pmatrix}1 & \mu^\top\end{pmatrix}$. The discrete control Liouville operator corresponding to the closed-loop system

$$x_{k+1} = F(x_k) + G(x_k)\mu(x_k)$$

is defined as the composition $M_\nu A_{F,G} : \mathcal{D}(M_\nu A_{F,G}) \to \tilde{H}$. The discrete control Liouville operator governs the flow of observables in $\mathcal{D}(M_\nu A_{F,G}) \subseteq \tilde{H}$ along trajectories of the discrete-time dynamical system as $[M_\nu A_{F,G}h](x_k) = \left\langle [A_{F,G}h](x_k), \begin{pmatrix}1 & \mu(x_k)^\top\end{pmatrix}\right\rangle_{\mathcal{Y}} = h(x_{k+1})$. Furthermore, the composition $M_\nu A_{F,G}$ is a linear operator by linearity of the inner product and by Definition 3.

V. Discrete-time control Liouville DMD

In order to represent the infinite-dimensional discrete control Liouville operator as a finite-dimensional operator, we select bases $\alpha = \left\{\tilde{K}_{x_i}\right\}_{i=1}^n \subset \tilde{H}$ and $\beta = \{K_{x_i,\bar{u}_i}\}_{i=1}^n \subset H$, where $\bar{u}_i \coloneqq \begin{pmatrix}1 & u_i^\top\end{pmatrix} \in \mathcal{Y}$. DMD is then performed via an eigendecomposition of the finite-dimensional representation.

Given an observable $h \in \tilde{H}$, let $\tilde{h} \coloneqq P_\alpha h = \sum_{i=1}^n \tilde{a}_i\tilde{K}_{x_i}$ be the projection of h onto $\operatorname{span}\alpha$. One can recover a finite rank proxy of the discrete control Liovuille operator by observing its action restricted to $\operatorname{span}\alpha \subset \tilde{H}$ and projecting the output $M_\nu A_{F,G}\tilde{h}$ back onto $\operatorname{span}\alpha$. That is, recovering the finite-rank proxy amounts to writing $P_\alpha M_\nu A_{F,G}\tilde{h}$ as $\sum_{i=1}^n \tilde{b}_i\tilde{K}_{x_i}$ and finding a matrix that relates the coefficients $\{\tilde{a}_i\}_{i=1}^n$ and $\{\tilde{b}_i\}_{i=1}^n$. For brevity of notation, let $\tilde{a} \coloneqq \begin{pmatrix}\tilde{a}_1 & \dots & \tilde{a}_n\end{pmatrix}^\top$ and $\tilde{b} \coloneqq \begin{pmatrix}\tilde{b}_1 & \dots & \tilde{b}_n\end{pmatrix}^\top$. The coefficients can be computed by solving the linear system of equations (see [1] and [7])

$$\tilde{G}\begin{pmatrix}\tilde{b}_1\\ \vdots \\ \tilde{b}_n\end{pmatrix} = \begin{pmatrix}\langle M_\nu P_\beta A_{F,G}\tilde{h}, \tilde{K}_{x_1}\rangle_{\tilde{H}}\\ \vdots \\ \langle M_\nu P_\beta A_{F,G}\tilde{h}, \tilde{K}_{x_n}\rangle_{\tilde{H}}\end{pmatrix}, \tag{5}$$

where $\tilde{G} = \{\tilde{K}(x_i, x_j)\}_{i,j=1}^n$ is the kernel gram matrix for α. Since the kernel functions in $\alpha \subset \tilde{H}$ are in the domain of the adjoint of the multiplication operator (see proposition 3), for all j, $\langle M_\nu P_\beta A_{F,G}\tilde{h}, \tilde{K}_{x_j}\rangle_{\tilde{H}} = \langle A_{F,G}\tilde{h}, P_\beta M_\nu^*\tilde{K}_{x_j}\rangle_H$. Furthermore, by linearity of $A_{F,G}$,

$$\begin{aligned}\langle A_{F,G}\tilde{h}, P_\beta M_\nu^*\tilde{K}_{x_j}\rangle_H &= \sum_{i=1}^n \tilde{a}_i\langle A_{F,G}\tilde{K}_{x_i}, P_\beta M_\nu^*\tilde{K}_{x_j}\rangle_H\\ &= \sum_{i=1}^n \tilde{a}_i\langle A_{F,G}\tilde{K}_{x_i}, \sum_{k=1}^n w_{k,j}K_{x_k,\bar{u}_k}\rangle_H,\end{aligned}$$

where $\{w_{k,j}\}_{k=1}^n$ are weights in the projection of $M_\nu^* \tilde{K}_{x_j}$ onto $\operatorname{span}\beta$ and $w_j \coloneqq \begin{pmatrix} w_{1,j} & \dots & w_{n,j} \end{pmatrix}^\top$. Thus,

$$\langle A_{F,G}\tilde{h}, P_\beta M_\nu^* \tilde{K}_{x_j}\rangle_H = \sum_{i,k=1}^n \tilde{a}_i w_{k,j}^* \langle A_{F,G}\tilde{K}_{x_i}, K_{x_k,\bar{u}_k}\rangle_H$$

$$= \sum_{i=1}^n \sum_{k=1}^n \tilde{a}_i w_{k,j}^* \langle [A_{F,G}\tilde{K}_{x_i}](x_k), \begin{pmatrix} 1 & u_k{}^\top \end{pmatrix}\rangle_{\mathcal{Y}} = \tilde{a}^\top \tilde{I} w_j^*,$$

where $\tilde{I} = \left(\left\langle [A_{F,G}\tilde{K}_{x_i}](x_k), \begin{pmatrix} 1 & u_k{}^\top \end{pmatrix} \right\rangle_{\mathcal{Y}} \right)_{i,k=1}^n$ is computed using the fact that $\left\langle [A_{F,G}\tilde{K}_{x_i}](x_k), \begin{pmatrix} 1 & u_k{}^\top \end{pmatrix} \right\rangle_{\mathcal{Y}} = \tilde{K}_{x_i}(x_{k+1})$ (here w^* denotes the complex conjugate of w).

Since M_ν^* maps $\tilde{K}_{x_j}$ to $K_{x_j,\nu(x_j)}$, the coefficients w_j in the projection of $K_{x_j,\nu(x_j)}$ onto $\operatorname{span}\beta \subset H$ are solutions of

$$G \begin{pmatrix} w_{1,j} \\ \vdots \\ w_{n,j} \end{pmatrix} = \begin{pmatrix} \langle K_{x_j,\nu(x_j)}, K_{x_1,\bar{u}_1}\rangle_H \\ \vdots \\ \langle K_{x_j,\nu(x_j)}, K_{x_n,\bar{u}_n}\rangle_H \end{pmatrix}, \tag{6}$$

where $G = \left(\left\langle K_{x_i,\bar{u}_i}, K_{x_j,\bar{u}_j} \right\rangle_H \right)_{i,j=1}^n$ and $\nu(x_j) = \begin{pmatrix} 1 & \mu(x_j)^\top \end{pmatrix}$. If a diagonal kernel operator $K_{x_i} \coloneqq \operatorname{diag}\begin{pmatrix} \tilde{K}_{x_1} & \dots & \tilde{K}_{x_{m+1}} \end{pmatrix}$ is used, with $\tilde{K}_{x_j} = \tilde{K}_{x_i}$ for $1 \le j \le m+1$, then the inner products in G can be computed as

$$\langle K_{x_i,\bar{u}_i}, K_{x_j,\bar{u}_j}\rangle_H = \langle K_{x_i,\bar{u}_i}(x_j), \begin{pmatrix} 1 & u_j^\top \end{pmatrix}\rangle_{\mathcal{Y}} = \begin{pmatrix} 1 & u_i^\top \end{pmatrix} \tilde{K}(x_j, x_i) \begin{pmatrix} 1 & u_j^\top \end{pmatrix}^\top. \tag{7}$$

Letting $I_j^\top$ denote the column vector on the right-hand side of (6), he jth row of I is given by

$$I_j = \left(\langle K_{x_j,\nu(x_j)}, K_{x_1,\bar{u}_1}\rangle_H, \dots, \langle K_{x_j,\nu(x_j)}, K_{x_n,\bar{u}_n}\rangle_H \right).$$

The complete finite-rank representation of the DCLDMD operator is then recovered as $[M_\nu P_\beta A_{F,G}]_\alpha^\alpha = \tilde{G}^\dagger I G^\dagger \tilde{I}^\top$, where the subscript α denotes the restriction of the operator to the $\operatorname{span}\alpha$; the superscript α denotes projection of the output onto $\operatorname{span}\alpha$; and $(\cdot)^\dagger$ denotes the Moore-Penrose pseudoinverse.

A. Discrete Control Liouville Dynamic Mode Decomposition

DMD can be accomplished via an eigendecomposition of the finite-rank proxy of discrete control Liovuille operator. Let $\{v_i, \lambda_i\}_{i=1}^n$ be the eigenvalue-eigenvector pairs of the matrix $[M_\nu P_\beta A_{F,G}]_\alpha^\alpha$. Following [7], if v_j is an eigenvector of the matrix $[M_\nu P_\beta A_{F,G}]_\alpha^\alpha$, then the function $\varphi_j = \sum_{i=1}^n (v_j)_i \tilde{K}_{x_i}$ is an eigenfunction of the operator $P_\alpha M_\nu P_\beta A_{F,G}|_\alpha$, where the subscript i in $(v_j)_i$ denotes the i-th component of the eigenvector v_j.

If φ_j is an eigenfunction of $P_\alpha M_\nu P_\beta A_{F,G}|_\alpha$ with eigenvalue λ_j, then

$$\varphi_j(x_{k+1}) = M_\nu A_{F,G}\varphi_j(x_k) = \lambda_j \varphi_j(x_k).$$

Hence, the eigenfunctions evolve linearly along the flow. The normalized eigenfunctions are defined as $\hat{\varphi}_j \coloneqq \frac{1}{\sqrt{(v_j^*)^\top \tilde{G} v_j}} \sum_{i=1}^n (v_j)_i \tilde{K}_{x_i}$.

Assuming that the j-th component identity function, g_{id}, defined as $g_{\mathrm{id},j}(x) \coloneqq x_j$ is in $\mathcal{D}(M_\nu A_{F,G}) \subset \tilde{H}$, for each $j = 1, 2, \dots, n$, we can describe the evolution of the full-state observable $g_{\mathrm{id}}(x) = x$ as a linear combination of eigenfunctions of $M_\nu A_{F,G}$. This approach yields a data-driven model of the closed-loop dynamical system as a linear combination of eigenfunctions of the operator $P_\alpha M_\nu P_\beta A_{F,G}|_\alpha$. That is, for a given $x_0 \in X$ we have a pointwise approximation of the flow of the closed-loop system

$$x_{k+1} = F(x_k) + G(x_k)\mu(x_k) \approx \sum_{i=1}^n \lambda_i^k \xi_i \hat{\varphi}_i(x_1). \tag{8}$$

We refer to the vectors ξ_i as the *Liouville Modes*, these are the coefficients required to represent the full-state observable as a linear combination of the eigenfunctions. We can calculate the modes by solving $g_{\mathrm{id}}(x) = x = \sum_{i=1}^n \xi_i \varphi_i$ for ξ_i, which yields $\xi \coloneqq \begin{pmatrix} \xi_1 & \cdots & \xi_n \end{pmatrix} = X(V^\top \tilde{G})^\dagger$, where V is the matrix of normalized eigenvectors of the finite-rank representation $[M_\nu P_\beta A_{F,G}]_\alpha^\alpha$ and $X \coloneqq \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ is the data matrix. We refer to this method as the *direct reconstruction* of the flow.

We can also formulate an *indirect reconstruction* of the flow by considering the function $F_\mu \coloneqq x \mapsto \sum_{i=1}^n \lambda_i \xi_i \hat{\varphi}_i(x)$ that approximates the closed loop dynamics under the feedback law μ as $x_{k+1} \approx F_\mu(x_k)$. The indirect method generally performs better for approximating the nonlinear dynamics; we hypothesize that the better performance is due to the fact that indirect reconstruction yields a nonlinear model of the flow, as opposed to the direct reconstruction, where the estimated model is linear. Due to its superior performance, we use the indirect reconstruction in the numerical experiments in Section VII. The DCLDMD algorithm is summarized in Algorithm 1.

VI. Convergence Properties of DCLDMD

Discrete control Liouville DMD enjoys convergence guarantees on par with current state-of-the-art Koopman methods. That is, the sequence of finite-rank operators $P_\alpha^n M_\nu P_\beta^n A_{F,G} P_\alpha^n$, where P^n denotes the projection onto the n-dimensional span of α and β, respectively, converges to the operator $M_\nu A_{F,G}$ in the *strong operator topology* (SOT). Underlying this fact is the assumption that as $n \to \infty$, the Gram matrices $\tilde{G}$ and G do not become rank deficient.

Theorem 1: If $A_{F,G} : \tilde{H} \to H$ and $M_\nu : \mathcal{D}(M_\nu) \to \tilde{H}$ are bounded, and $\alpha \coloneqq \{\tilde{K}_{x_n}\}_{n=1}^\infty \subset \tilde{H}$ and $\beta \coloneqq \{K_{x_n,\bar{u}_n}\}_{n=1}^\infty \subset H$ are two orthonormal sequences in $\tilde{H}$ and H, respectively, then for all $f \in \tilde{H}$, $\lim_{n\to\infty} \left\| P_\alpha^n M_\nu P_\beta^n A_{F,G} P_\alpha^n f - M_\nu A_{F,G} f \right\|_{\tilde{H}} = 0$.

Proof: Suppose $f \in \tilde{H}$, then

$$\left\| P_\alpha^n M_\nu P_\beta^n A_{F,G} P_\alpha^n f - M_\nu A_{F,G} f \right\|_{\tilde{H}} =$$
$$\left\| (P_\alpha^n - I) M_\nu P_\beta^n A_{F,G} P_\alpha^n f + M_\nu (P_\beta^n A_{F,G} P^n f - A_{F,G} f) \right\|_{\tilde{H}}$$
$$\le \left\| (P_\alpha^n - I)(M_\nu P_\beta^n A_{F,G} P_\alpha^n f - M_\nu A_{F,G} f) \right\|_{\tilde{H}} +$$
$$\| (P_\alpha^n - I) M_\nu A_{F,G} f \|_{\tilde{H}} + \left\| M_\nu (P_\beta^n A_{F,G} P_\alpha^n f - A_{F,G} f) \right\|_{\tilde{H}}$$
$$\le \| (P_\alpha^n - I) \|_{op} \left\| (M_\nu P_\beta^n A_{F,G} P_\alpha^n f - M_\nu A_{F,G} f) \right\|_{\tilde{H}} +$$
$$\| (P_\alpha^n - I) M_\nu A_{F,G} f \|_{\tilde{H}} + \left\| M_\nu (P_\beta^n A_{F,G} P_\alpha^n f - A_{F,G} f) \right\|_{\tilde{H}},$$

where $\|\cdot\|_{op}$ denotes the operator norm. Since M_ν is continuous and $\| (P_\alpha^n - I) \|_{op}$ is bounded (by Parseval's identity,

Algorithm 1 The DCLDMD algorithm

Input: Data points $\{(x_k, y_k, u_k)\}_{k=1}^n$ that satisfy $y_k = F(x_k) + G(x_k)u_k$, reproducing kernels $\tilde{K}_{x_j}$ and $K_{x_j,\bar{u}_j}$ for $\tilde{H}$ and H, respectively. A feedback law μ, kernel parameter σ, and a regularization parameter ϵ.

Output: $\{\hat{\varphi}_j, \lambda_j, \xi_j\}_{j=1}^n$

$\tilde{G} \leftarrow \{\tilde{K}(x_i, x_j)\}_{i,j=1}^n$

$\tilde{I} \leftarrow \{\tilde{K}(x_{k+1}, x_i)\}_{k,i=1}^n$

$G \leftarrow \{\langle K_{x_i,\bar{u}_i}, K_{x_j,\bar{u}_j}\rangle_H\}_{i,j=1}^n$ (see (7))

$I \leftarrow \{\langle K_{x_j,\nu(x_j)}, K_{x_i,\bar{u}_i}\rangle_H\}_{i,j=1}^n$ (see (7))

Compute $[M_\nu P_\beta A_{F,G}]_\alpha^\alpha = \tilde{G}^\dagger I G^\dagger \tilde{I}^\top$

Eigendecomposition: $\{v_j, \lambda_j\}_{j=1}^n \leftarrow [M_\nu P_\beta A_{F,G}]_\alpha^\alpha$

Normalize the eigenfunctions: $\{\hat{\varphi}_j\}_{j=1}^n \leftarrow \hat{\varphi}_j = \frac{1}{\sqrt{(v_j^*)^\top \tilde{G} v_j}} \sum_{i=1}^n (v_j)_i \tilde{K}_{x_i}$

Liouville modes: $\xi \leftarrow X(V^\top \tilde{G})^\dagger$

return $\{\hat{\varphi}_j, \lambda_j, \xi_j\}_{j=1}^n$

see [18, Section 3.1.11]), and since $P_\beta^n A_{F,G} P_\alpha^n$ converges to $A_{F,G}$ in the SOT [18, Page 172]), the first and the third terms in the inequality above converge to 0 as $n \to \infty$. The fact that P_α^n converges to I in the SOT implies the convergence of the second term to zero. Therefore, the sequence of operators $P_\alpha^n M_\nu P_\beta^n A_{F,G} P_\alpha^n$ converges to $M_\nu A_{F,G}$ in the SOT. ■

Convergence in the SOT does not guarantee convergence of the spectrum, but by Theorem 4 in [19], it does guarantee that there is a subsequence of eigenvalue-eigenfunction pairs of the finite-rank representation which converges to an eigenvalue-eigenfunction pair of the true discrete control Liouville operator.

VII. Numerical Experiments

As a demonstration of the efficacy of the developed DCLDMD algorithm, we apply the method to the controlled Duffing oscillator and compare it with the linear predictor developed in [10].

Experiment 1: The controlled Duffing oscillator is a nonlinear dynamical system with state-space form

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = \begin{pmatrix}x_2\\ -\delta x_2 - \beta x_1 - \alpha x_1^3\end{pmatrix} + \begin{pmatrix}0\\ 2+\sin(x_1)\end{pmatrix} u \quad (9)$$

where α, β, δ are coefficients in $\mathbb{R}$, $[x_1, x_2]^\top \in \mathbb{R}^2$ is the state, and $u \in \mathbb{R}$ is the control input. For the experiments the parameters are selected to be: $\delta = 0$, $\alpha = 1$, and $\beta = -1$.

We descretize (9) using a time step of 0.01 seconds to yield a discrete-time, control-affine dynamical system of the form $x_{k+1} = F(x_k) + G(x_k)u_k$. Using the tuples $\{(x_k, x_{k+1}, u_k)\}_{k=1}^n$ generated by the dynamical system, we aim to predict the response of the system starting from the initial condition $x_0 = [2, -2]^\top$ to the feedback law $\mu(x_k) = -2x_{k,1}^3 - x_{k,2}$ for a total of 5 seconds.

In the implementation of DCLDMD, we generate 1225 data points from initial conditions sampled from a 35×35 grid within the set $[-5, 5] \times [-5, 5] \subset \mathbb{R}^2$ and the control inputs are sampled uniformly from the interval $[-8, 8] \subset \mathbb{R}$.

The Gaussian radial basis function kernel $\tilde{K}(x, y) = \mathrm{e}^{\frac{-\|x-y\|_2^2}{\sigma}}$ is used for $\tilde{H}$, with kernel width $\sigma = 20$. For H, we associate to each pair $\{(x_k, u_k)\}_{k=1}^n$ a kernel $K_{x_k,\bar{u}_k} := \begin{pmatrix}1 & u_k^\top\end{pmatrix} K_{x_k} \in H$. Here we use the kernel operator $K_{x_i} := \operatorname{diag}\begin{pmatrix}\tilde{K}_{x_1} & \cdots & \tilde{K}_{x_{m+1}}\end{pmatrix}$ where $\tilde{K}_{x_j}(y) = \mathrm{e}^{\frac{-\|x_j-y\|_2^2}{\sigma}}$ for $j = 1, \ldots, m+1$, with $\sigma = 20$. Lastly, we select $\varepsilon = 10^{-6}$ for regularization of the Gram matrices in order to ensure invertibility of both $\tilde{G}$ and G, instead of the pseudoinverse, in the finite-rank representation (see Algorithm 1).

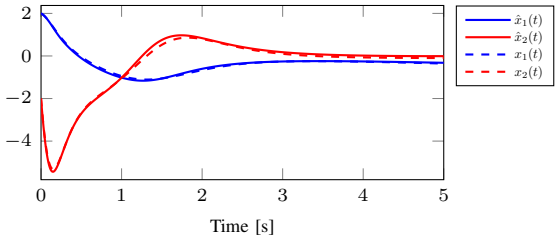

Fig. 1. A comparison of indirectly reconstructed trajectories $\hat{x}_1(t)$ and $\hat{x}_2(t)$ with the true trajectories $x_1(t)$ and $x_2(t)$ of the Duffing oscillator resulting from the nonlinear feedback law μ in experiment 1.

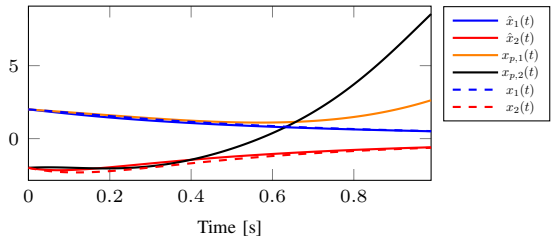

Fig. 2. A comparison between the linear predictor developed in [10] and the indirect reconstruction via DCLDMD in experiment 2. Here, $\hat{x}_i(t)$, $x_{p,i}(t)$, and $x_i(t)$ represent the indirect reconstruction, the linear predictor, and the actual trajectories, respectively, where i is a subscript denoting an element of the state.

A comparison between the true trajectories and the indirectly reconstructed trajectories corresponding to the feedback law μ can be seen in Figure 1.

Experiment 2: In this experiment, we compare the predictive capabilities of the indirect reconstruction via DCLDMD with the linear predictor derived in [10]. The linear predictor in [10] is of the form $z_{k+1} = Az_k + Bu_k$ with $x_k = Cz_k$ and z being the lifted state (see [10] for more details). For a given feedback law μ, we can estimate the response of the Duffing oscillator described by equation (9) to the feedback law μ by using the linear predictor $z_{k+1} = Az_k + B\mu(Cz_k)$.

For this experiment, we generate 1000 data points and DCLDMD is implemented using the same kernels as in Experiment 1, except the kernel widths are both set to $\sigma = 100$. For regularization we set $\varepsilon = 10^{-6}$. For the linear predictor, extended DMD (eDMD) is performed with the Gaussian radial basis functions as in [10]. For the initial condition $x_0 = [2, -2]^\top$ and the feedback law $\mu(x_k) = -2x_{k,1} - 2x_{k,2}$, we compare the predictions of the indirect DCLDMD method and the linear predictor with the true trajectories (see Figure 2).

A. Discussion

The experiments demonstrate the efficacy of DCLDMD in an academic setting with the Duffing oscillator. The experiments are done with no prior model knowledge, besides the system being affine in control. The novelty of the

representation can be seen in the separation of the control input and the state on the operator-theoretic level, while still preserving the nonlinearity of the dynamical system. This is opposed to the standard approach for discrete-time dynamical systems where the lifted state $z_k \in \mathbb{R}^N$ can be approximated as $z_{k+1} \approx Az_k + Bu_k$, with $A \in \mathbb{R}^{N\times N}$ and $B \in \mathbb{R}^{1\times N}$ found using extended DMD. Unless the original nonlinear system admits an exact lifting, which is not generally the case, the trajectories of the linear lifted systems are expected to diverge from the trajectories of the nonlinear system with increasing prediction horizons.

In Experiment 2, specifically, in Figure 2, we observe that as expected, the behavior of the linear predictor from [10] diverges from the behavior of the nonlinear Duffing oscillator under the given feedback law, while the indirect reconstruction approach developed in this paper accurately tracks the actual trajectory of the Duffing oscillator. We postulate that the improved predictive capability can be attributed to the fact that the indirect predictive model is a *nonlinear* predictor, as opposed to the model from [10], which is a linear predictor, albeit in a higher dimensional lifted state space.

In both experiments, indirect reconstruction is used to estimate the flow. The indirect reconstruction explicitly depends upon the eigenfunctions of $P_\alpha M_\nu P_\beta A_{F,G}|_\alpha$. Whether or not we can always represent the full-state observable (i.e. the flow) in terms of the eigenfunctions is not entirely clear, but this is a standard assumption in the DMD literature. With this assumption in mind, DCLDMD is termed a heuristic approach for estimation of the dynamics. Regardless, the numerical experiments in section VII demonstrate the capability of DCLDMD for prediction of the response of the control-affine system to given feedback laws.

VIII. Conclusion

In this paper, a novel operator representation of a control-affine nonlinear system is developed as a composition of a multiplication operator and a composition-like kernel propagation operator over an RKHS. The multiplication operator takes advantage of the affine nature of the system to capture the effect of control on the system behavior, while the kernel-propagation operator captures the effect of the system dynamics on the kernels of the underlying RKHS. The resulting DMD algorithm is entirely data driven and requires no model knowledge besides the dynamical system being affine in control. Furthermore, the DCLDMD formulation provides a novel way to separate the state from the control input on the operator-theoretic level. This separation leads to better prediction capabilities over existing methods, as evidenced by the results of Experiment 2. Moreover, since DCLDMD can be used to predict closed-loop trajectories of a nonlinear system under feedback laws, it could potentially be utilized for control synthesis, which is a topic for future research.

References

[1] J. A. Rosenfeld and R. Kamalapurkar, "Dynamic mode decomposition with control Liouville operators," in *IFAC-PapersOnLine*, vol. 54, no. 9, 2021, pp. 707–712.

[2] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 17, no. 5, p. 315, 1931.

[3] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynanic mode decomposition - data-driven modeling of complex systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2016.

[4] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.*, vol. 656, pp. 5–28, 2010.

[5] I. Mezić, "Analysis of fluid flows via spectral properties of the Koopman operator," *Annu. Rev. Fluid Mech.*, vol. 45, pp. 357–378, 2013.

[6] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *J. Comput. Dyn*, vol. 1, no. 2, pp. 391–421, 2014.

[7] E. Gonzalez, M. Abudia, M. Jury, R. Kamalapurkar, and J. A. Rosenfeld, "The kernel perspective on dynamic mode decompositions," arXiv:2106.00106, 2021.

[8] P. Sharma, B. Huang, V. Ajjarapu, and U. Vaidya, "Data-driven identification and prediction of power system dynamics using linear operators," in *Proc. IEEE Power Energy Soc. Gen. Meet.*, 2019, pp. 1–5.

[9] D. Goswami and D. A. Paley, "Bilinearization, reachability, and optimal control of control-affine nonlinear systems: A Koopman spectral approach," *IEEE Trans. Autom. Control*, vol. 67, no. 6, pp. 2715–2728, 2022.

[10] M. Korda and I. Mezić, "Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control," *Automatica*, vol. 93, pp. 149–160, 2018.

[11] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Dynamic mode decomposition with control," *SIAM J. Appl. Dyn. Syst.*, vol. 15, no. 1, pp. 142–161, 2016.

[12] M. O. Williams, M. S. Hemati, S. T. Dawson, I. G. Kevrekidis, and C. W. Rowley, "Extending data-driven koopman analysis to actuated systems," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 704–709, 2016.

[13] J. L. Proctor, S. L. Brunton, and J. N. Kutz, "Generalizing Koopman theory to allow for inputs and control," *SIAM J. Appl. Dyn. Syst.*, vol. 17, no. 1, pp. 909–930, 2018.

[14] B. Huang, X. Ma, and U. Vaidya, "Feedback stabilization using Koopman operator," in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 6434–6439.

[15] R. Strässer, J. Berberich, and F. Allgöwer, "Robust data-driven control for nonlinear systems using the Koopman operator," arXiv.2304.03519, 2023.

[16] A. Surana, "Koopman operator based observer synthesis for control-affine nonlinear systems," in *Proc. IEEE Conf. Decis. Control*. IEEE, 2016, pp. 6492–6499.

[17] C. Carmeli, E. De Vito, A. Toigo, and V. Umanitá., "Vector valued reproducing kernel Hilbert spaces and universality," *Anal. Appl.*, vol. 08, no. 01, pp. 19–61, 2010.

[18] G. K. Pedersen, *Analysis now*, ser. Graduate Texts in Mathematics. Springer-Verlag, New York, 1989, vol. 118.

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