Dynamic Mode Decomposition with Control Liouville Operators

Joel A. Rosenfeld and Rushikesh Kamalapurkar

Abstract—**This paper builds the theoretical foundations for dynamic mode decomposition (DMD) of control-affine dynamical systems by leveraging the theory of vector-valued reproducing kernel Hilbert spaces (RKHSs). Specifically, control Liouville operators and control occupation kernels are introduced to separate the drift dynamics from the input dynamics. A given feedback controller is represented through a multiplication operator and a composition of the control Liouville operator and the multiplication operator is used to express the nonlinear closed-loop system as a linear total derivative operator on RKHSs. A spectral decomposition of a finite-rank representation of the total derivative operator yields a DMD of the closed-loop system. The DMD generates a model that can be used to predict the trajectories of the closed-loop system. For a large class of systems, the total derivative operator is shown to be compact provided the domain and the range RKHSs are selected appropriately. The sequence of models, resulting from increasing-rank finite-rank representations of the compact total derivative operator, are shown to converge to the true system dynamics, provided sufficiently rich data are available. Numerical experiments are included to demonstrate the efficacy of the developed technique.**

Index Terms—**dynamic mode decomposition, NL system identification, Computational methods, Reduced order modeling, Nonlinear systems**

I. INTRODUCTION

SPECTRAL methods for identification of nonlinear systems utilize representations of unknown, finite-dimensional nonlinear dynamics, in discrete or continuous time, as linear operators over infinite dimensional spaces (cf. [1]). In the discrete-time case, this linear operator is a composition operator called the Koopman operator [2]. In the continuous time case, it is a total derivative operator called the Liouville operator [3] (or the Koopman generator, in special cases where it can be obtained as the limit of a sequence of Koopman operators with decreasing sample times [4, Section 7.5]). In dynamic mode decomposition (DMD), trajectories of a dynamical system are used to construct a finite-rank

Joel A. Rosenfeld is with the Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620 USA (e-mail: rosenfeldj@usf.edu).

Rushikesh Kamalapurkar is with the School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: rushikesh.kamalapurkar@okstate.edu)

representation of the aforementioned linear operator [5]. The finite-rank representation is then diagonalized and the resultant eigenfunction and eigenvalues are used to provide a representation of the identity function. This representation provides the dynamic modes of the system as vector-valued coefficients attached to the eigenfunctions. Thereafter, a state trajectory can be predicted as a sum of exponential functions multiplied by the dynamic modes (cf. [3], [5], [6]).

The primary application area for Koopman spectral analysis of dynamical systems has been fluid dynamics, where DMD is compared with proper orthogonal decomposition (POD) for nonlinear fluid equations (cf. [7]). DMD has also been employed in the study of stability properties of dynamical systems [8], [9], neuroscience [10], financial trading [11], feedback stabilization [12], optimal control [13], modeling of dynamical systems [14]–[16], and model-predictive control [17]. For a generalized treatment of DMD as a Markov model, see [18].

Extensions of the idea of Koopman operator-based DMD to systems with control can be loosely categorized in three categories: spectral analysis of the drift (zero-input) dynamics [19], input-parameterized Koopman operators [20], and reformulation as an autonomous state-control dynamical system [21]. These methods rely on discretization of continuous-time systems, either for computation (when Koopman operators are used), or for analysis (when Koopman generators are used), and as such, are only applicable to systems that admit a globally well-defined discretization (i.e., systems that cannot escape to infinity in finite time starting from any initial condition). When dealing with Koopman generators, the data required for a spectral decomposition typically include the time derivative of the state, which is not generally available. Recently, inspired by the notion of occupation measures [22] defined on Banach spaces of continuous functions, the authors in [23] defined analogous objects on reproducing kernel Hilbert spaces (RKHSs). The so-called occupation kernels, when combined with operators such as the Liouville operator, provide a method for spectral analysis of continuous-time systems directly, without the need for discretization.

The paradigm shift afforded by occupation kernels arises through the consideration of the state trajectory as the fundamental unit of data [23]. This paper, along with the preliminary results reported in [24], build on the foundations developed in [23] to address DMD of control-affine dynamical systems. To address systems with control, the occupation kernels are augmented by the control signals, resulting in the so-called *control occupation kernels*, and the Liouville operator is

extended to include the input dynamics, to yield the so-called *control Liouville operator* [24]. The extension in [24] utilizes the theory of vector-valued RKHSs (vvRKHSs), introduced in [25] and [26], and extensively studied in a machine learning context in [27], [28], and [29]. Multiplication operators that map between scalar-valued and vector-valued RKHSs are also utilized to define a total derivative operator that represents the dynamics of the closed-loop system controlled using a feedback controller. Using the control occupation kernels, the control Liouville operators, and the multiplication operators, a technique for discretization-free DMD of control-affine, continuous-time, nonlinear systems is developed. The developed control-Liouville DMD (CLDMD) and singular control-Liouville DMD (SCLDMD) methods yield a predictor that can predict the closed-loop behavior of a system under any given locally Lipschitz continuous feedback controller by measuring its response to different open-loop control signals.

The definitions of control occupation kernels and control Liouville operators used in this paper were first reported in the conference paper [24]. In that paper, a finite-rank representation of the closed-loop total derivative operator is indirectly derived through its adjoint. In this paper, a finite-rank representation of the closed-loop total differential operator is obtained directly, resulting in a simpler DMD algorithm. Furthermore, this paper includes a novel singular value decomposition (SVD)-based finite-rank representation of a new total derivative operator which converges in norm to the true operator with increasing rank under a compactness assumption and given sufficiently rich data. Examples of general classes of nonlinear systems where the new total derivative operators are compact are also provided to justify the compactness assumptions.

II. Problem Statement

Given Carathéodory solutions $\{\gamma_{u_i} : [0, T_i] \to \mathbb{R}^n\}_{i=1}^M$ of a nonlinear control-affine system of the form

$$\dot{x} = f(x) + g(x)u_i(t), \quad x(0) = \gamma_{u_i}(0) \qquad (1)$$

under Lebesgue measurable, bounded control inputs $\{u_i : [0, T_i] \to \mathbb{R}^m\}_{i=1}^M$, the objective of this paper is to provide an operator theoretic approach for the analysis of the *closed loop* system

$$\dot{x} = f(x) + g(x)\mu(x) =: F_\mu(x), \qquad (2)$$

where $x \in \mathbb{R}^n$ is the state, $\mu : \mathbb{R}^n \to \mathbb{R}^m$ is a locally Lipschitz continuous feedback controller, $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are locally Lipschitz continuous functions corresponding to the drift dynamics and the control effectiveness matrix, respectively, and $\dot{x}$ denotes the time derivative of x. The observed control trajectories and control inputs will allow for the construction of a finite-rank representation of the so-called *control Liouville operator*, which is a generalization of the Liouville operator introduced in [3].

Euler-Lagrange systems, which encompass a large class of physical systems (see, e.g., [30]–[33]), can be expressed in the control-affine form if their inertia matrices are invertible [34]. Since most physical systems of practical importance

such as robot manipulators and ground, air, and maritime vehicles and vessels have inertia matrices that are invertible over large operating regions [34], [35], control-affine models also encompass a large class of physical systems.

III. Operators and Dynamic Mode Decomposition

In this section, the general idea behind the developed operator-theoretic DMD approach is introduced. The approach relies on representation of a closed loop dynamical system as an operator that maps between suitable function spaces. For the motivational discussion in this section, assume that given functions f, g, and μ, and RKHSs $\tilde{H}_d$ and $\tilde{H}_r$ defined on a compact set $X \subset \mathbb{R}^n$, there exist a set, $\mathcal{D}\left(A_{F_\mu}\right) \subset \tilde{H}_d$ and a *total derivative operator* $A_{F_\mu} : \mathcal{D}\left(A_{F_\mu}\right) \to \tilde{H}_r$ such that

(R1) for all $h \in \mathcal{D}\left(A_{F_\mu}\right)$, $A_{F_\mu} h := \frac{\partial h}{\partial x} F_\mu \in \tilde{H}_r$, where $\frac{\partial h}{\partial x}$ is a row vector, and

(R2) $h_{\mathrm{id},j} \in \mathcal{D}\left(A_{F_\mu}\right)$ for all $j = 1, \cdots, n$, where $h_{\mathrm{id}} = \begin{bmatrix} h_{\mathrm{id},1} & \cdots & h_{\mathrm{id},n} \end{bmatrix}^\top$ is the identity function, with components defined as $h_{\mathrm{id},j}(x) = x_j$ for all $x \in X$.

Note that the total derivative operator A_{F_μ} above is the Liouville operator (or the Koopman generator) with symbol $F_\mu = f + g\mu$ as defined in [3]. As such, a DMD of the closed loop system could be obtained using the methods presented in [3] *provided data generated by the closed loop system $\dot{x} = F_\mu(x)$ is available*. The objective in this paper is to develop a model of the system using a feedback-agnostic data set. That is, given any feedback controller $\mu : \mathbb{R}^n \to \mathbb{R}^m$ and a data set recorded by exciting the open-loop system $\dot{x} = f(x) + g(x)u$ using control signals $u = u_i : [0, T_i] \to \mathbb{R}^n$, $i = 1, \ldots, M$, we aim to build a predictive model of the closed loop system $\dot{x} = F_\mu(x)$.

A. The Eigendecomposition Approach

If $\tilde{H}_d = \tilde{H}_r$, ϕ is an eigenfunction of A_{F_μ} with eigenvalue λ, and γ_μ is a controlled trajectory arising from (2), then it follows that

$$\frac{\mathrm{d}(\phi(\gamma_\mu(t)))}{\mathrm{d}t} = \frac{\partial \phi}{\partial x}(\gamma_\mu(t))\Big(f(\gamma_\mu(t)) + g(\gamma_\mu(t))\mu(\gamma_\mu(t))\Big)$$
$$= \big[A_{F_\mu}\phi\big](\gamma_\mu(t)) = \lambda\phi(\gamma_\mu(t)).$$

Hence, $\phi(\gamma_\mu(t)) = \mathrm{e}^{\lambda t}\phi(\gamma_\mu(0))$.

If the span of the eigenfunctions $\{\phi_i\}_{i=1}^\infty$ of A_{F_μ} is dense in $\tilde{H}_d$, then the identity function can be decomposed using the eigenfunctions as $h_{\mathrm{id}}(x) = \lim_{M \to \infty} \sum_{i=1}^M \xi_{i,M}\phi_i(x)$, where $\xi_{i,M} \in \mathbb{C}^n$ are the *dynamic modes* of the closed loop system. Moreover, it follows that

$$\gamma_\mu(t) = h_{\mathrm{id}}(\gamma_\mu(t)) = \lim_{M \to \infty} \sum_{i=1}^M \xi_{i,M}\phi_i(\gamma_\mu(0))\mathrm{e}^{\lambda_i t}, \qquad (3)$$

where λ_i denotes the eigenvalue corresponding to the eigenfunction ϕ_i, and the coefficients $\xi_{i,M}$ depend on M because the eigenfunctions are not generally orthogonal.

If the eigenfunctions, the eigenvalues, and the modes could be computed from data, then a finite truncation of (3) could be used as a predictive model. However, since the operator

A_{F_μ} cannot generally be expected to be bounded, even the existence of eigenfunctions cannot be guaranteed.

The idea in DMD is to construct a rank M approximation $\hat{A}_{F_\mu,M}$ of A_{F_μ}. The eigenfunctions $\{\hat{\phi}_{i,M}\}_{i=1}^M$, the eigenvalues $\{\hat{\lambda}_{i,M}\}_{i=1}^M$, and the modes $\{\hat{\xi}_{i,M}\}_{i=1}^M$ of $\hat{A}_{F_\mu,M}$ are then used as proxies in a finite truncation of (3) to generate a predictive model.

If the operators $\hat{A}_{F_\mu,M}$ can be shown to converge to A_{F_μ} *in the norm topology*, then given any $\epsilon > 0$, there exists M such that for all $i = 1, \ldots, M$, the pairs $(\hat{\phi}_{i,M}, \hat{\lambda}_{i,M})$ are approximate eigenpairs for the true operator A_{F_μ}. That is, for all $i = 1, \ldots, M$ and for all $x \in X$, $\left| \left[A_{F_\mu} \hat{\phi}_{i,M} \right](x) - \hat{\lambda}_{i,M} \hat{\phi}_{i,M}(x) \right| < \epsilon$. The approximate eigenpairs can then be used to obtain a model that, given rich enough data and a large enough M, can accurately predict the system trajectories in X over a finite horizon.

While requirements (R1) and (R2) above, compactness of the Liouville operator, and density of the eigenfunctions in $\tilde{H}_d$ are difficult to guarantee in general, empirical evidence suggests that the eigenfunctions $\hat{\phi}_{i,M}$ are expressive enough to approximate $h_{\mathrm{id},j}$ in a variety of applications [3]. Since $\hat{\phi}_{i,M}$ are computed as linear combinations of reproducing kernels or occupation kernels, the empirical evidence could be explained by the postulate that the approximate eigenfunctions inherit universality properties of the reproducing kernels and the occupation kernels [23]. A theoretical examination of the expressiveness of the approximate eigenfunctions for a specific operators, Hilbert spaces, and data set is out of the scope of this article.

Convergence of the finite-rank representation to the true operator in the norm topology is also typically impossible to guarantee in the eigendecomposition-based DMD framework [3], [21]. As such, similar to most DMD techniques, the eigendecomposition approach, while well-motivated by the theory presented in this paper, is a heuristic technique. On the other hand, as shown in [36], obtaining norm convergence of finite rank representations to the true Liouville operator is possible in an SVD-based framework.

B. The Singular Value Decomposition Approach

In the SVD-based framework, two different RKHSs $\tilde{H}_d$ and $\tilde{H}_r$ are selected as the domain and the co-domain of A_{F_μ}, respectively. If the domain and the range RKHSs are selected carefully, then for a large class of nonlinear systems, the operator A_{F_μ} can be shown to be compact. Compactness trivially ensures satisfaction of Requirement (R1) above. Requirement (R2) can be met by proper selection of $\tilde{H}_d$ (see Section VI). Compactness also allows for the construction of the needed sequence $\hat{A}_{F_\mu,M}$ that converges to A_{F_μ} in the norm topology. The left and right singular functions of $\hat{A}_{F_\mu,M}$ can then be used to generate a sequence of system models that converges to the true system model.

In particular, the closed-loop model $\dot{x} = f(x) + g(x)\mu(x)$ can be expressed in terms of the total derivative operator as

$$\dot{x} = \frac{\partial h_{\mathrm{id}}}{\partial x}(x) \begin{bmatrix} f(x) & g(x) \end{bmatrix} \begin{bmatrix} 1 \\ \mu(x) \end{bmatrix} = [A_{F_\mu} h_{\mathrm{id}}](x), \quad (4)$$

where the notation $A_{F_\mu} h_{\mathrm{id}}$ is used to denote the operator A_{F_μ} acting on every row of the vector-valued function h_{id}. If $A_{F_\mu} : \tilde{H}_d \to \tilde{H}_r$ is a compact operator, then there exist singular values $\{\sigma_i\}_{i=1}^\infty \subset \mathbb{R}$, left singular functions $\{\phi_i\}_{i=1}^\infty \subset \tilde{H}_d$, and right singular functions $\{\psi_i\}_{i=1}^\infty \subset \tilde{H}_r$ such that

$$\dot{x} = \sum_{i=1}^\infty \sigma_i \langle h_{\mathrm{id}}, \phi_i \rangle_{\tilde{H}_d} \psi_i(x), \quad (5)$$

where the notation $\langle h_{\mathrm{id}}, \phi_i \rangle_{\tilde{H}_d}$ is used to denote the $n-$vector $\left[\langle h_{\mathrm{id},1}, \phi_i \rangle_{\tilde{H}_d}, \quad \ldots, \quad \langle h_{\mathrm{id},n}, \phi_i \rangle_{\tilde{H}_d} \right]^\top$. In singular DMD, the idea is to use the SVD of $\hat{A}_{F_\mu,M}$ as a proxy in a finite truncation of (5) to construct a predictive model.

C. Related Work

Operator-based DMD methods for systems with control can be loosely categorized in three categories: spectral analysis of the drift (zero-input) dynamics [19], input-parameterized Koopman operators [20], and reformulation as an autonomous state-control dynamical system [21].

If data can be collected for the system with zero inputs, or if the system is affine in control, then techniques such as dynamic mode decomposition with control (DMDc) [14], sparse nonlinear system identification with control (SINDYc) [37], extended dynamic mode decomposition with control (EDMDc) [21], bilinearization [38], etc., can be utilized to estimate eigenvalues and eigenfunctions of the Koopman operator, or the Koopman generator, of the drift (zero-input) dynamics. In the case of control-affine systems, the eigenvalues and eigenfunctions can also be utilized to solve a wide variety of problems including, but not limited to, reachability [38], optimal control [39], model-based predictive control, [40], and observer synthesis [19].

A different approach to operator theoretic analysis of systems with control is via input-parameterized Koopman operators [20]. The central idea in this family of methods is that if the input is constant, then the dynamical system is autonomous, and as such, admits a Koopman operator. Given a set of possible input levels, an input-parameterized family of Koopman operators (or generators) can thus be constructed [39]. This observation is particularly useful when utilized for spectral analysis of control-affine systems, where Koopman generators are themselves affine in control. The state of the system can thus be predicted using a linear combination of a finite number of input-parameterized Koopman generators [41]. In addition to motivating DMDc and EDMDc, input-parameterized Koopman generators can also be used for various control and estimation tasks [42].

Systems with control can also be analyzed by studying operators that operate on a more general set of observables. Instead of observables that are functions of the state in the typical Koopman framework, the observables here are functions of the state and the control [20], [21]. The methods in this category include Koopman with inputs and control (KIC) [20] and linear and bilinear predictors [21]. The KIC approach is cogent if the control signal itself is produced by a dynamical system, and leads to useful heuristics when it is not. In [21],

the shift operator is used as the dynamics of the control signal to develop a Koopman operator that operates on observables defined on an infinite-dimensional state space that includes the space of all possible control *sequences*. Spectral analysis of this operator with carefully selected observables yields linear and bilinear predictors for the underlying nonlinear system. Applications of this approach include model-based predictive control [21], robust model-based predictive control [43], and system identification [44].

In this paper, the operator A_{F_μ} is constructed as a composition of two operators, a differential operator that maps from $\tilde{H}_d$ into a vvRKHS and a multiplication operator that maps from the vvRKHS either back into $\tilde{H}_d$ (the eigendecomposition approach) or into $\tilde{H}_r$ (the SVD approach) (see Fig. 1).

IV. Vector-valued Reproducing Kernel Hilbert Spaces

In this section, properties of vvRKHSs relevant to topic under consideration are reviewed. The review relies heavily on the discussion given in [29].

Definition 1: 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 $v \in \mathcal{Y}$ and $x \in X$, the functional $f \mapsto \langle f(x), v \rangle_{\mathcal{Y}}$ is bounded.
A vvRKHS is a direct generalization of a "scalar-valued" RKHS, since for a fixed $v \in \mathcal{Y}$, the collection of functions $\{g(x) = \langle f(x), v \rangle_{\mathcal{Y}} : f \in H\}$ forms an RKHS of scalar-valued functions.

By the Riesz representation theorem, for each $x \in X$ and $v \in \mathcal{Y}$, there exists a unique function $K_{x,v} \in H$ such that $\langle f, K_{x,v} \rangle_H = \langle f(x), v \rangle_{\mathcal{Y}}$ for all $f \in H$. The fact that the mapping $v \mapsto K_{x,v}$ is linear over $\mathcal{Y}$ yields a linear operator $K_x : \mathcal{Y} \to H$, defined as $K_x := v \mapsto K_{x,v}$, called *the kernel centered at x, associated with H*. The operator $K : X \times X \to \mathcal{L}(\mathcal{Y}, \mathcal{Y})$, defined as $K(x,y) := K_x^* K_y$, where $K_x^* : H \to \mathcal{Y}$ is the adjoint of K_x and $\mathcal{L}(\mathcal{Y}, \mathcal{Y})$ is the space of linear operators from $\mathcal{Y}$ to $\mathcal{Y}$, is called the *reproducing kernel of H*. For any $f \in H$, $x \in X$, and $v \in \mathcal{Y}$, we have $\langle K_x^* f, v \rangle_{\mathcal{Y}} = \langle f, K_x v \rangle_H = \langle f(x), v \rangle_{\mathcal{Y}}$, and as a result, $K_x^* f = f(x)$. With $f = K_y v$, we see that for all $v \in \mathcal{Y}$, $[K_y v](x) = K_x^* K_y v = K(x,y)v$.

In the particular case that $\mathcal{Y} = \mathbb{R}^n$, $K(x,y)$ is a real-valued $n \times n$ matrix for fixed $x, y \in X$. As a result one can construct several examples of vector-valued kernels. Indeed, given a scalar-valued RKHS $\tilde{H}$ over X, with the corresponding reproducing kernel $\tilde{K} : X \times X \to \mathbb{R}$, and a positive definite matrix, $A \in \mathbb{R}^{n \times n}$, the operator $(x,y) \mapsto A\tilde{K}(x,y)$ that maps from $X \times X$ to $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a reproducing kernel of a vvRKHS.

Similar to scalar-valued kernels, it can be shown that the span of vector-valued kernels is dense in H.

Proposition 1: The span of the set $E := \{K_{x,v} : v \in \mathcal{Y} \text{ and } x \in X\}$, is dense in H.

Proof: Suppose that $h \in E^\perp$, then given a fixed $x \in X$, $\langle h, K_{x,v} \rangle_H = \langle h(x), v \rangle_{\mathcal{Y}} = 0$ for all $v \in \mathcal{Y}$. Hence, $h(x) = 0 \in \mathcal{Y}$. Since x was arbitrarily selected, $h \equiv 0 \in H$. Thus, $E^\perp = \{0\}$ and $\overline{\text{span}(E)} = (E^\perp)^\perp = H$. ∎

As a consequence of Proposition 1, given $\epsilon > 0$ and $h \in H$, there is a finite linear combination of vector-valued kernels that approximates h with an error smaller than ϵ in the Hilbert space norm.

In the following development, unless otherwise specified, it is assumed that $X \subset \mathbb{R}^n$ is compact, the Hilbert space $\mathcal{Y}$ is selected to be $\mathbb{R}^{1 \times (m+1)}$ with the usual definitions of vector norms and inner products, $\tilde{H}_d$ and $\tilde{H}_r$ are RKHSs of continuously differentiable functions from X to $\mathbb{R}$, and H is a vvRKHS of continuous functions from X to $\mathbb{R}^{1 \times (m+1)}$. The reproducing kernel of H is denoted by $K : X \times X \to \mathcal{L}(\mathbb{R}^{1 \times (m+1)}, \mathbb{R}^{1 \times (m+1)})$ and the reproducing kernels of $\tilde{H}_d$ and $\tilde{H}_r$ are denoted by $\tilde{K}_d : X \times X \to \mathbb{R}$ and $\tilde{K}_r : X \times X \to \mathbb{R}$, respectively. When the domain and the range RKHSs are identical, the subscripts d and r are omitted. The Hilbert space $\mathcal{Y}$ is selected to be a space of row vectors to accommodate the row vector convention for partial derivatives. As such, the linear operation of K_x on $v \in \mathcal{Y}$ is expressed as $K_{x,v} = vK_x$.

V. Closed Loop Nonlinear Systems as Operators over RKHSS

To solve the problem as stated in Section II in a vvRKHS framework, the closed-loop nonlinear system is expressed in terms of operators over two RKHSs and a vvRKHS. A majority of the definitions and propositions in this section were first introduced in [24]. The definitions are included here for completeness and the proofs of most of the propositions are more detailed than the corresponding proofs in [24].

A. Control Liouville Operators and Multiplication Operators

Representation of a controlled system in terms of operators can be realized using the so-called control Liouville Operator.

Definition 2: Let $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ be locally Lipschitz continuous functions and the set

$$\mathcal{D}(A_{f,g}) := \left\{ h \in \tilde{H}_d : x \mapsto \frac{\partial h}{\partial x}(x) \begin{bmatrix} f(x) & g(x) \end{bmatrix} \in H \right\}$$

be the domain of the operator, $A_{f,g} : \mathcal{D}(A_{f,g}) \to H$, given as

$$[A_{f,g}h](x) := \frac{\partial h}{\partial x}(x) \begin{bmatrix} f(x) & g(x) \end{bmatrix}.$$

The operator $A_{f,g}$ is called *the control Liouville operator corresponding to f and g over H*.
Control Liouville operators generalize Liouville operators, where the drift dynamics and control effectiveness components of the dynamics are separated on the operator theoretic level. Vector-valued RKHSs arise naturally in this context, where the partial derivative of $h \in \mathcal{D}(A_{f,g})$ with respect to x is a row vector of dimension n, and through a dot product with f and multiplication by the matrix g, the result of the operation of $A_{f,g}$ on h is a row vector with dimension $m + 1$.

The control Liouville operator does not depend on the control input, and as such, is not sufficient by itself for prediction of system behavior. An additional operator is thus required to complete the construction of the operator A_{F_μ} alluded to in Section III. The controller is incorporated in the developed framework via a multiplication operator. The
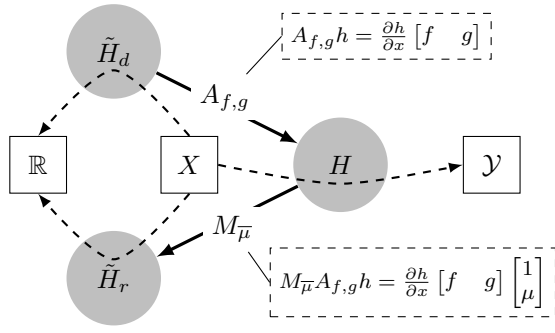
Fig. 1. A schematic diagram of the construction presented in Section V. The RKHSs are represented by filled circles. The squares at the endpoints of the dashed arrows passing through the circles indicate the domains and co-domains of the functions contained in the RKHSs. The thick arrows between the RKHSs indicate operators.

inclusion of this multiplication operator, in addition to the newly defined control Liouville operator, sets the theoretical foundations of DMD of controlled systems apart from the uncontrolled case studied in [3] and [36].

Definition 3: For a continuous function $\nu : X \to \mathcal{Y}$, the *multiplication operator with symbol ν*, denoted by $M_\nu : \mathcal{D}(M_\nu) \to \tilde{H}_r$, is defined as

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

where $\mathcal{D}(M_\nu) := \{h \in H : x \mapsto \langle h(x), \nu(x) \rangle_\mathcal{Y} \in \tilde{H}_r\}$.

Given the continuous function $\overline{\mu} : \mathbb{R}^n \to \mathbb{R}^{1 \times m+1}$ derived from a feedback controller $\mu : \mathbb{R}^n \to \mathbb{R}^m$ as $\overline{\mu}(x) := \begin{bmatrix} 1 & \mu(x)^\top \end{bmatrix}$, the corresponding multiplication operator $M_{\overline{\mu}} : \mathcal{D}(M_{\overline{\mu}}) \to \tilde{H}_r$ is given as $M_{\overline{\mu}} h = x \mapsto h(x) \begin{bmatrix} 1 & \mu(x)^\top \end{bmatrix}^\top$ and $\mathcal{D}(M_{\overline{\mu}}) = \{h \in H : x \mapsto h(x) \begin{bmatrix} 1 & \mu(x)^\top \end{bmatrix}^\top \in \tilde{H}_r\}$. The operator $M_{\overline{\mu}} A_{f,g}$ maps from $\mathcal{D}(A_{f,g}) \subset \tilde{H}_d$ to $\tilde{H}_r$, and plays the role of the operator A_{F_μ} described in Section III. In the above construction, it is assumed that the image of $A_{f,g}$ falls within the domain of $M_{\overline{\mu}}$. This assumption is not easy to verify in general, but it is trivially met in the example presented in Section VI, where the RKHSs are Bargmann-Fock spaces restricted to the set of real numbers. The control Liouville operator will be assumed to be compact in Section VI and densely defined in Section VII. Further comments on the density and the compactness assumptions, including examples of systems and RKHSs for which these assumptions are met, are provided in the respective sections.

B. Control Occupation Kernels

To facilitate the computation of a finite-rank representation of $A_{F_\mu} = M_{\overline{\mu}} A_{f,g}$, and subsequently, the approximate eigenfunctions required for DMD, trajectories of controlled dynamical systems are embedded within vvRKHSs using the so-called *control occupation kernels*. Control occupation kernels arise from a generalization of the idea of *occupation kernels* introduced in [23] as follows.

Definition 4 ([23]): Given a continuous function $\gamma : [0, T] \to X$, an RKHS of continuous functions $\tilde{H}$, and the bounded functional $\mathcal{T} : \tilde{H} \to \mathbb{C}$ defined as $\mathcal{T} h =$

$\int_0^T h(\gamma(\tau)) \mathrm{d}\tau$ for all $h \in \tilde{H}$, the unique function[1] $\Gamma_\gamma \in \tilde{H}$ that satisfies $\mathcal{T} h = \langle h, \Gamma_\gamma \rangle_{\tilde{H}}$ for all $h \in \tilde{H}$ is called the *occupation kernel corresponding to γ in $\tilde{H}$*.

Boundedness of $\mathcal{T}$ in Definition 4 follows from arguments similar to the vector-valued case below.

Proposition 2: If H is a vvRKHS of continuous functions, defined on $X \subseteq \mathbb{R}^n$, with reproducing kernel K, then for all continuous functions $\gamma : [0, T] \to \mathbb{R}^n$ and bounded measurable functions $u : [0, T] \to \mathbb{R}^m$, the functional $\mathcal{T} : H \to \mathbb{C}$, defined as

$$\mathcal{T} h := \int_0^T h(\gamma(t)) \begin{bmatrix} 1 \\ u(t) \end{bmatrix} \mathrm{d}t, \quad \forall h \in H, \qquad (6)$$

is bounded.

Proof: Let $\bar{u}(t) = \begin{bmatrix} 1 & u(t)^\top \end{bmatrix}$. Using the defining properties of H,

$$\mathcal{T} h = \int_0^T \langle h(\gamma(t)), \bar{u}(t) \rangle_\mathcal{Y} \, \mathrm{d}t = \int_0^T \langle h, \bar{u}(t) K_{\gamma(t)} \rangle_H \, \mathrm{d}t.$$

Using the Cauchy-Schwarz inequality and the fact that $K_{\gamma(t)}^* K_{\gamma(t)} = K(\gamma(t), \gamma(t))$,

$$|\mathcal{T} h| \le \|h\|_H \int_0^T \sqrt{\|K(\gamma(t), \gamma(t)) \bar{u}(t)\|_\mathcal{Y} \|\bar{u}(t)\|_\mathcal{Y}} \mathrm{d}t.$$

Since γ is continuous and $x \mapsto K(x, x)$ is locally bounded [29, Proposition 2], there exists $\overline{K} < \infty$ such that for all $x \in \gamma([0, T])$, the operator norm of $K(x, x)$ satisfies $\|K(x, x)\| \le \overline{K}$. Since $u(\cdot)$ is bounded, we conclude that $\mathcal{T}$ is bounded. ∎

An extension of Definition 4 to systems with control results in control occupation kernels.

Definition 5 ([24]): Given a bounded measurable function $u : [0, T] \to \mathbb{R}^m$, a continuous function $\gamma : [0, T] \to X$, and the bounded functional $\mathcal{T} : H \to \mathbb{C}$ defined in (6), the unique function $\Gamma_{\gamma, u} \in H$ that satisfies $\mathcal{T} h = \langle h, \Gamma_{\gamma, u} \rangle_H$ for all $h \in H$ is called *the control occupation kernel corresponding to u and γ in H*.

Control occupation kernels can be expressed in terms of the reproducing kernels of H to facilitate computation.

Proposition 3 ([24]): The control occupation kernel $\Gamma_{\gamma, u} \in H$, corresponding to u and γ, can be expressed as

$$\Gamma_{\gamma, u}(x) = \int_0^T \left[\begin{bmatrix} 1 & u(t)^\top \end{bmatrix} K_{\gamma(t)} \right](x) \, \mathrm{d}t, \qquad (7)$$

and the norm of $\Gamma_{\gamma, u}$ is given as

$$\|\Gamma_{\gamma, u}\|_H^2 = \int_0^T \int_0^T \begin{bmatrix} 1 & u(t)^\top \end{bmatrix} K(\gamma(\tau), \gamma(t)) \begin{bmatrix} 1 \\ u(\tau) \end{bmatrix} \mathrm{d}t \mathrm{d}\tau.$$

Proof: For $x \in X$ and $v \in \mathbb{C}^{1 \times (m+1)}$,

$$\langle \Gamma_{\gamma, u}(x), v \rangle_{\mathbb{C}^{1 \times (m+1)}} = \langle \Gamma_{\gamma, u}, v K_x \rangle_H$$

$$= \int_0^T \langle [v K_x](\gamma(t)), \begin{bmatrix} 1 & u(t)^\top \end{bmatrix} \rangle_{\mathbb{C}^{1 \times (m+1)}} \mathrm{d}t$$

$$= \int_0^T \langle v K_x, \begin{bmatrix} 1 & u(t)^\top \end{bmatrix} K_{\gamma(t)} \rangle_H \mathrm{d}t$$

$$= \int_0^T \langle \left[\begin{bmatrix} 1 & u(t)^\top \end{bmatrix} K_{\gamma(t)} \right](x), v \rangle_{\mathbb{C}^{1 \times (m+1)}} \mathrm{d}t$$

---

[1] Such a function exists by the Riesz representation theorem.

$$= \left\langle \int_0^T \left[\begin{matrix}1 & u(t)^\top\end{matrix}\right] K_{\gamma(t)} \, (x)\mathrm{d}t, v \right\rangle_{\mathbb{C}^{1\times(m+1)}}. \quad (8)$$

As (8) holds for all $v \in \mathbb{C}^{1\times(m+1)}$, (7) follows. The expression for the norm of $\Gamma_{\gamma,u}$ follows from $\|\Gamma_{\gamma,u}\|_H^2 = \langle \Gamma_{\gamma,u}, \Gamma_{\gamma,u}\rangle_H$ and the defining properties of $\Gamma_{\gamma,u}$. $\blacksquare$

C. Control Liouville Operators and Control Occupation Kernels

There is a direct connection between the adjoints of *densely defined* control Liouville operators and control occupation kernels that correspond to *admissible* control signals, u, and their corresponding controlled trajectories, γ_u, that satisfy (2).

Definition 6: A bounded, measurable control signal $u : [0, T] \to \mathbb{R}^m$ is called *admissible for the initial value problem* (1) *over the time interval* $[0, T]$ *and the domain* X, if the corresponding Carathéodory solution $\gamma_u : [0, T] \to \mathbb{R}^n$ is contained within X.

To illustrate the connection, the construction of adjoints of densely defined operators is revisited in the following.

Definition 7: The *domain of the adjoint of* $A : \mathcal{D}(A) \to H$, with $\mathcal{D}(A) \subseteq \tilde{H}_d$, is defined as

$$\mathcal{D}\left(A^*\right) \coloneqq \{h \in H \mid \phi \mapsto \langle A\phi, h\rangle_H \text{ is bounded on } \mathcal{D}\left(A\right)\}.$$

If $\mathcal{D}(A)$ is dense in $\tilde{H}_d$, then the functionals $\phi \mapsto \langle A\phi, h\rangle_H$ may be extended uniquely to functionals that are bounded over all of $\tilde{H}_d$. As a result, for each $h \in \mathcal{D}\left(A^*\right)$, the Riesz representation theorem guarantees the existence of a unique function $A^*h \in \tilde{H}_d$ such that $\langle A^*h, \phi\rangle_{\tilde{H}} = \langle A\phi, h\rangle_H$ for all $\phi \in \mathcal{D}(A)$. The operator $h \mapsto A^*h$ is defined as *the adjoint of* A.

The following proposition formalizes the relationship between control occupation kernels and control Liouville operators for trajectories of the system under admissible control signals.

Proposition 4 ([24]): If f and g correspond to a densely defined control Liouville operator, $A_{f,g} : \mathcal{D}(A_{f,g}) \to H$, with $\mathcal{D}(A_{f,g}) \subset \tilde{H}_d$ and if u is an admissible control signal for the initial value problem (1) over the time interval $[0, T]$, with a corresponding controlled trajectory γ_u, then $\Gamma_{\gamma_u,u} \in \mathcal{D}(A_{f,g}^*)$ and

$$A_{f,g}^* \Gamma_{\gamma_u,u} = \tilde{K}_d(\cdot, \gamma_u(T)) - \tilde{K}_d(\cdot, \gamma_u(0)). \quad (9)$$

Proof: To demonstrate that $\Gamma_{\gamma_u,u}$ is in $\mathcal{D}(A_{f,g}^*)$ it must be shown that the functional $h \mapsto \langle A_{f,g}h, \Gamma_{\gamma_u,u}\rangle_H$ is a bounded. Note that

$$\langle A_{f,g}h, \Gamma_{\gamma_u,u}\rangle_H$$
$$= \int_0^T \frac{\partial}{\partial x} h(\gamma_u(t)) \left[\begin{matrix}f(\gamma_u(t)) & g(\gamma_u(t))\end{matrix}\right] \left[\begin{matrix}1 \\ u(t)\end{matrix}\right] \mathrm{d}t$$
$$= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} h(\gamma_u(t))\mathrm{d}t = h(\gamma_u(T)) - h(\gamma_u(0))$$
$$= \left\langle h, \tilde{K}_d(\cdot, \gamma_u(T)) - \tilde{K}_d(\cdot, \gamma_u(0))\right\rangle_{\tilde{H}_d}. \quad (10)$$

The functional $h \mapsto \langle A_{f,g}h, \Gamma_{\gamma_u,u}\rangle_H$ is thus bounded with norm not exceeding $\|\tilde{K}_d(\cdot, \gamma_u(T)) - \tilde{K}_d(\cdot, \gamma_u(0))\|_{\tilde{H}_d}$. By the definition of the adjoint, (10) implies (9). $\blacksquare$

D. Properties of Multiplication Operators

In this section, multiplication operators that map from vvRKHSs to scalar-valued RKHSs are studied. Many of the propositions in this section have been established for scalar-valued RKHSs (cf. [45]–[47]), and are proved using similar methods.

The following proposition investigates the interaction between adjoints of multiplication operators and reproducing kernels of scalar-valued RKHSs.

Proposition 5 ([24]): If $\nu : X \to \mathcal{Y}$ corresponds to a densely defined multiplication operator $M_\nu : \mathcal{D}\left(M_\nu\right) \to \tilde{H}_r$ with $\mathcal{D}\left(M_\nu\right) \subset H$, then for each $x \in X$, $\tilde{K}_r(\cdot, x)$ is in the domain of M_ν^* and

$$M_\nu^* \tilde{K}_r(\cdot, x) = K_{x,\nu(x)}. \quad (11)$$

Proof: If $h \in \mathcal{D}(M_\nu)$, then

$$\langle M_\nu h, \tilde{K}_r(\cdot, x)\rangle_{\tilde{H}_r} = \langle h(x), \nu(x)\rangle_{\mathcal{Y}} = \langle h, K_{x,\nu(x)}\rangle_H. \quad (12)$$

Hence, the mapping $h \mapsto \langle M_\nu h, \tilde{K}_r(\cdot, x)\rangle_{\tilde{H}_r}$ is a bounded functional with norm bounded by $\|K_{x,\nu(x)}\|_H$, and as such, $\tilde{K}_r(\cdot, x)$ is in the domain of M_ν^*. The proposition then follows from (10). $\blacksquare$

Proposition 5, along with the density of the span of the kernels $\tilde{K}(\cdot, x)$ in $\tilde{H}$ implies that the adjoint of the multiplication operator is also densely defined.

Proposition 6 ([24]): Multiplication operators are closed operators.

Proof: Suppose that $\{h_n\}_{n=1}^\infty \subset \mathcal{D}(M_\nu)$, $h_n \to h \in H$, and $M_\nu h_n \to W \in \tilde{H}_r$. To show that M_ν is a closed operator, it must be shown that $W(x) = \langle h(x), \nu(x)\rangle_{\mathcal{Y}}$ for all $x \in X$, and thus $h \in \mathcal{D}(M_\nu)$ by definition and $M_\nu h = W$. Let $v \in \mathcal{Y}$ and $x \in X$, then

$$W(x) = \lim_{n\to\infty} \langle M_\nu h_n, \tilde{K}_r(\cdot, x)\rangle_{\tilde{H}_r}$$
$$= \lim_{n\to\infty} \langle h_n, K_{x,\nu(x)}\rangle_H = \langle h, K_{x,\nu(x)}\rangle_H = \langle h(x), \nu(x)\rangle_{\mathcal{Y}},$$

where the first equality follows since norm convergence in $\tilde{H}_r$ implies pointwise convergence and the third inequality follows from continuity of the inner product on H. $\blacksquare$

The following proposition demonstrates how multiplication operators $M_{\overline{\mu}}$, with symbols $\overline{\mu}$ given as $\overline{\mu}(x) = \left[\begin{matrix}1 & \mu(x)^\top\end{matrix}\right], \forall x \in \mathbb{R}^n$, connect occupation kernels Γ_γ with feedback control occupation kernels $\Gamma_{\gamma,\mu\circ\gamma}$.

Proposition 7: If $\mu : X \to \mathbb{R}^m$ is a continuous function, $\overline{\mu} : X \to \mathcal{Y}$ is defined as $\overline{\mu}(x) \coloneqq \left[\begin{matrix}1 & \mu(x)^\top\end{matrix}\right], \forall x \in X$, the corresponding multiplication operator $M_{\overline{\mu}} : \mathcal{D}\left(M_{\overline{\mu}}\right) \to \tilde{H}_r$, with $\mathcal{D}\left(M_{\overline{\mu}}\right) \subset H$, is densely defined, $\Gamma_\gamma \in \tilde{H}_r$ is the occupation kernel corresponding to a continuous function $\gamma : [0, T] \to X$ in $\tilde{H}_r$, then Γ_γ is in the domain of $M_{\overline{\mu}}^*$ and

$$M_{\overline{\mu}}^* \Gamma_\gamma = \Gamma_{\gamma,\mu\circ\gamma}. \quad (13)$$

Proof: Let $h \in \mathcal{D}(M_{\overline{\mu}})$. Using definitions 4 and 5, it can be concluded that

$$\langle M_{\overline{\mu}}h, \Gamma_\gamma\rangle_{\tilde{H}} = \int_0^T \langle h(\gamma(t)), \overline{\mu}(\gamma(t))\rangle_{\mathcal{Y}} \, \mathrm{d}t$$
$$= \int_0^T h\left(\gamma\left(t\right)\right) \left[\begin{matrix}1 \\ \mu(\gamma(t))\end{matrix}\right] \mathrm{d}t = \langle h, \Gamma_{\gamma,\mu\circ\gamma}\rangle_H. \quad (14)$$

Hence, the norm of the functional $h \mapsto \langle M_{\overline{\mu}} h, \Gamma_\gamma \rangle_{\tilde{H}_r}$ is bounded by $\|\Gamma_{\gamma, \mu \circ \gamma}\|_H$. It can thus be concluded that $\Gamma_\gamma \in \mathcal{D}(M_{\overline{\mu}}^*)$, and as a result, (13) follows. ∎

E. Compact and Densely Defined Operators for DMD

As noted in Section III, DMD relies on computation of eigenfunctions of a finite-rank representation of an operator that represents the dynamical system. The eigenfunctions of the finite-rank representations can be shown to converge to the eigenfunctions of the true operator if the finite-rank representations themselves converge to the true operator in the norm topology and the true operator is compact [21]. Koopman operators, Koopman generators, and Liouville operators are typically not compact if their domains and co-domains are viewed as subsets of the same RKHS [21], [36], [48] (see Remark 1).

As noted in [36], Liouville operators corresponding to a large class of dynamical systems are compact provided the domain and the range RKHSs are selected appropriately. However, since the domain and the range RKHSs need to be different, the resulting operators do not admit eigenfunctions. In Section VI, it is shown that when the domain and the range RKHSs are different, an SVD-based approach can be used to estimate the system dynamics. The SVD-based approach relies on compactness of the total derivative operator and generates sequences of singular values and singular functions that converge to the true singular values and singular functions. As shown in Section VI, compact total derivative operators result from bounded multiplication operators and compact control Liouville operators, both of which exist for a large class of dynamical systems and feedback laws.

In Section VII, an eigendecomposition-based DMD approach is developed that lacks convergence guarantees but generates useful heuristic approximations of the eigenfunctions under the weaker assumption that the total derivative operator is densely defined. As shown in Section VII, a densely defined total derivative operator results from a densely defined multiplication operator whose range is a subset of the domain of a densely defined control Liouville operator. As discussed in Section VII, such multiplication operators and control Liouville operators also exist for a large class of dynamical systems and feedback laws.

In the following, for an operator A and finite collections of functions d and r, in the domain and the range of the operator, respectively, the notation $A|_d$ is used to denote the operator A restricted to the set $\operatorname{span} d$, and the notation $[A]_d^r$ is used to denote a matrix representation of the finite-rank operator $P_r A|_d$, where P_r denotes the projection operator onto $\operatorname{span} r$.

VI. A SINGULAR VALUE DECOMPOSITION APPROACH TO DMD

With careful selection of the domain and range RKHSs, the total derivative operator $M_{\overline{\mu}} A_{f,g}$ can be made to be compact. While the results in this section apply to large class of dynamical systems, a complete characterization of RKHSs and symbols that yield compact differential operators and bounded multiplication operators is out of the scope of this paper.

A. Existence of Bounded Multiplication Operators and Compact Differential Operators

The discussion in this section closely follows [36], where a similar result is obtained for systems without control. Consider the exponential dot product kernel with parameter $\tilde{\rho}$, defined as $\tilde{K}_{\tilde{\rho}}(x, y) = \exp\left(\frac{x^\top y}{\tilde{\rho}}\right)$. In the single variable case, the RKHS[2] of this kernel is the restriction of the Bargmann-Fock space to real numbers, denoted by $F_{\tilde{\rho}}^2(\mathbb{R})$. This space consists of the set of functions of the form $h(x) = \sum_{k=0}^\infty a_k x^k$, where the coefficients satisfy $\sum_{k=0}^\infty |a_k|^2 \tilde{\rho}^k k! < \infty$, and the norm is given by $\|h\|_{\tilde{\rho}}^2 = \sum_{k=0}^\infty |a_k|^2 \tilde{\rho}^k k!$. Note that the set of polynomials in x is a subset of $F_{\tilde{\rho}}^2(\mathbb{R})$. Extension of this definition to the multivariable case yields the space $F_{\tilde{\rho}}^2(\mathbb{R}^n)$ where the collection of monomials, $x^\alpha \frac{\tilde{\rho}^{|\alpha|}}{\sqrt{\alpha!}}$, with multi-indices $\alpha \in \mathbb{N}^n$ forms an orthonormal basis[3]. In this setting, provided $\tilde{\rho}_2 < \tilde{\rho}_1$, differential operators from $F_{\tilde{\rho}_1}^2(\mathbb{R}^n)$ to $F_{\tilde{\rho}_2}^2(\mathbb{R}^n)$ can be shown to be compact.

Proposition 8: If $\tilde{\rho}_2 < \tilde{\rho}_1$, then the differential operators $\frac{\partial}{\partial x_i} : F_{\tilde{\rho}_1}^2(\mathbb{R}^n) \to F_{\tilde{\rho}_2}^2(\mathbb{R}^n)$, are compact for $i = 1, \ldots, n$.

Proof: To facilitate the clarity of exposition, the proof is written for functions of a single variable. Extension to functions of several variables using multi-indices is conceptually straightforward. Let $h \in F_{\tilde{\rho}_1}^2(\mathbb{R})$ be given by $h(x) = \sum_{k=0}^\infty a_k x^k$, with $\frac{\partial h}{\partial x} = \sum_{k=0}^\infty b_k x^k$, where $b_k = (k+1)a_{k+1}$. The norm of the derivative in $F_{\tilde{\rho}_2}^2(\mathbb{R})$ is given by

$$\left\| \frac{\partial h}{\partial x} \right\|_{\tilde{\rho}_2}^2 = \sum_{k=0}^\infty |b_k|^2 \tilde{\rho}_2^k k!$$

$$= \sum_{k=0}^\infty |a_{k+1}|^2 \frac{k+1}{\tilde{\rho}_1} \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^k \tilde{\rho}_1^{(k+1)} (k+1)!$$

$$= \sum_{k=0}^\infty |a_k|^2 \frac{k}{\tilde{\rho}_1} \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^{(k-1)} \tilde{\rho}_1^k k!$$

If $\tilde{\rho}_2 < \tilde{\rho}_1$ then there exists a constant $C < \infty$ such that $\frac{k}{\tilde{\rho}_1} \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^{(k-1)} \leq C$ for all $k \in \mathbb{N}$. As a result, $\left\| \frac{\partial h}{\partial x} \right\|_{\tilde{\rho}_2}^2 \leq C \|h\|_{\tilde{\rho}_1}^2$, which establishes boundedness of the differential operator $\frac{\partial}{\partial x} : F_{\tilde{\rho}_1}^2(\mathbb{R}) \to F_{\tilde{\rho}_2}^2(\mathbb{R})$.

To prove compactness, we construct a sequence of finite-rank operators that converge, in norm, to $\frac{\partial}{\partial x}$. Let $\alpha_M := \{1, x, \ldots, x^M\}$ be the first M monomials in x, and let P_{α_M} be the projection onto the span of these monomials. Consider the sequence $\{P_{\alpha_M} \frac{\partial}{\partial x}\}$ of finite-rank operators. Let $h \in F_{\tilde{\rho}_1}^2(\mathbb{R})$ be given by $h(x) = \sum_{k=0}^\infty a_k x^k$. Then,

$$\left\| P_{\alpha_M} \frac{\partial h}{\partial x} - \frac{\partial h}{\partial x} \right\|_{\tilde{\rho}_2}^2 = \sum_{k=M+1}^\infty (k+1)^2 |a_{k+1}|^2 \tilde{\rho}_2^{(k+1)} (k+1)!$$

$$\leq \sum_{k=M+1}^\infty (k+1)^2 \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^{(k+1)} \sum_{k=M+1}^\infty |a_{k+1}|^2 \tilde{\rho}_1^{(k+1)} (k+1)!$$

---

[2] The RKHS of a symmetric positive semidefinite kernel K is the unique RKHS of which K is the reproducing kernel. Such an RKHS is guaranteed to exist by the Moore–Aronszajn theorem [49].

[3] For $\alpha \in \mathbb{N}^n$, $\alpha! = \prod_{i=1}^n \alpha_i!$, $|\alpha| = \sum_{i=1}^n \alpha_i$, and $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$.

$$\leq \sum_{k=M+1}^{\infty} (k+1)^2 \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^{(k+1)} \|h\|_{\tilde{\rho}_1}^2 .$$

If $\tilde{\rho}_2 < \tilde{\rho}_1$ then $\lim_{M\to\infty} \sum_{k=M+1}^{\infty} (k+1)^2 \left(\frac{\tilde{\rho}_2}{\tilde{\rho}_1}\right)^{(k+1)} = 0$, and as a result, $\lim_{M\to\infty} \left\| P_{\alpha_M} \frac{\partial h}{\partial x} - \frac{\partial h}{\partial x} \right\|_{\tilde{\rho}_2}^2 = 0$. Therefore, the operator norm

$$\left\| P_{\alpha_M} \frac{\partial}{\partial x} - \frac{\partial}{\partial x} \right\|_{F_{\tilde{\rho}_1}^2(\mathbb{R})}^{F_{\tilde{\rho}_2}^2(\mathbb{R})} := \sup_{h \in F_{\tilde{\rho}_1}^2(\mathbb{R})} \frac{\| P_{\alpha_M} \frac{\partial h}{\partial x} - \frac{\partial h}{\partial x} \|_{\tilde{\rho}_2}}{\|h\|_{\tilde{\rho}_1}}$$

converges to zero as $M \to \infty$, which establishes compactness of $\frac{\partial}{\partial x} : F_{\tilde{\rho}_1}^2(\mathbb{R}) \to F_{\tilde{\rho}_2}^2(\mathbb{R})$. ∎

Remark 1: Note that if $\tilde{\rho}_2 < \tilde{\rho}_1$ then $F_{\tilde{\rho}_1}^2(\mathbb{R}^n) \subset F_{\tilde{\rho}_2}^2(\mathbb{R}^n)$ [36, Proposition 5.1]. In this case, one can view the differential operators as maps from $F_{\tilde{\rho}_2}^2(\mathbb{R}^n)$ to itself. However, when viewed as such, the differential operators may not be compact.

As shown in [36], multiplication operators can be shown to be bounded provided their symbols are polynomial.

Proposition 9: If $\tilde{\rho}_2 < \tilde{\rho}_1$, then for any polynomial function $p : \mathbb{R}^n \to \mathbb{R}$, the multiplication operator $M_p : F_{\tilde{\rho}_1}^2(\mathbb{R}^n) \to F_{\tilde{\rho}_2}^2(\mathbb{R}^n)$, defined as $[M_p h](x) = p(x)h(x)$, is bounded.

Proof: See [36, Lemma 3.2]. ∎

Proposition 9 trivially extends to vvRKHSs defined using diagonal reproducing kernels.

Proposition 10: Let $F_{\tilde{\rho}}^2(\mathbb{R}^n)$ denote the native (row) vvRKHS of a diagonal reproducing kernel defined as $K(x,y) := \text{diag}\left(\left[\tilde{K}_{\rho_1}(x,y), \ldots, \tilde{K}_{\rho_{m+1}}(x,y)\right]\right)$. If $\rho_i < \tilde{\rho}$ for $i = 1, \ldots, m+1$, then given any set of polynomials p_i, $i = 1, \ldots, m+1$, the multiplication operator $M_{p_1,\ldots,p_{m+1}} : F_{\tilde{\rho}}^2(\mathbb{R}^n) \to F_{\rho}^2(\mathbb{R}^n)$, defined as $\left[M_{p_1,\ldots,p_{m+1}} h\right](x) = h(x)\left[p_1(x), \ldots, p_{m+1}(x)\right]$, is bounded. On the other hand, if $\tilde{\rho} < \rho_i$ for $i = 1, \ldots, m+1$, then for any component-wise polynomial function $\mu = \left[\mu_1, \ldots, \mu_m\right]^\top : \mathbb{R}^n \to \mathbb{R}^m$, the multiplication operator $M_{\overline{\mu}} : F_{\rho}^2(\mathbb{R}^n) \to F_{\tilde{\rho}}^2(\mathbb{R}^n)$, defined as $[M_{\overline{\mu}} h](x) = h(x)\left[1, \mu_1(x), \ldots, \mu_m(x)\right]^\top$, is bounded.

Proof: Follows from arguments similar to Lemma 3.2 from [36]. ∎

Since Koopman operators are generally unbounded for nonlinear systems [48], the above propositions make a strong case for spectral analysis of continuous-time systems in the Liouville operator (or Koopman generator) framework as opposed to the Koopman operator framework via discretization.

B. Finite-rank Representation of the Closed Loop Total Derivative Operator

Since the dynamic modes may only be extracted from the composition of $M_{\overline{\mu}}$ with $A_{f,g}$, an explicit finite-rank representation of $A_{f,g}$ and $M_{\overline{\mu}}$ is needed to determine the dynamic modes of the resultant system. In the following, finite collections of linearly independent vectors, d^M, ϖ^M, β^M, and r^M are selected to establish the needed finite-rank representation. Since the adjoint of $A_{f,g}$ maps control occupation kernels to kernel differences (Proposition 4), the span of the collection of kernel differences

$$d^M = \left\{ K_d(\cdot, \gamma_{u_i}(T_i)) - K_d(\cdot, \gamma_{u_i}(0)) \right\}_{i=1}^M \subset \tilde{H}_d \quad (15)$$

is selected to be the domain of $A_{f,g}$. The corresponding Gram matrix is denoted by $G_{d^M} = \left(\langle d_i, d_j \rangle_{\tilde{H}_d}\right)_{i,j=1}^M$. The output of $A_{f,g}$ is projected onto the span of the control occupation kernels

$$\beta^M = \left\{ \Gamma_{\gamma_{u_i}, u_i} \right\}_{i=1}^M \subset H \quad (16)$$

before application of $M_{\overline{\mu}}$. The corresponding Gram matrix is denoted by $G_{\beta^M} = \left(\langle \beta_i, \beta_j \rangle_H\right)_{i,j=1}^M$.
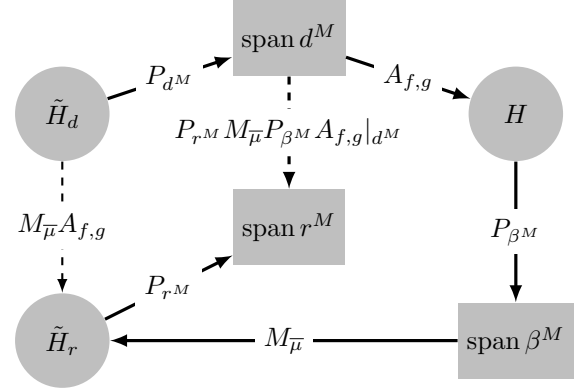

Fig. 2. A schematic diagram of the finite-rank representation of the total derivative operator.

Since the adjoint of $M_{\overline{\mu}}$ maps occupation kernels to control occupation kernels of the form $\Gamma_{\gamma_{u_i}, \mu \circ \gamma_{u_i}}$ (Proposition 7), the derivation also requires the collection

$$\varpi^M = \left\{ \Gamma_{\gamma_{u_i}, \mu \circ \gamma_{u_i}} \right\}_{i=1}^M \subset H \quad (17)$$

of feedback control occupation kernels in H corresponding to the trajectories γ_{u_i} and control signals $\mu \circ \gamma_{u_i}$. Finally, the result of $M_{\overline{\mu}}$ is projected onto the span of the occupation kernels

$$r^M = \left\{ \Gamma_{\gamma_{u_i}} \right\}_{i=1}^M \subset \tilde{H}_r. \quad (18)$$

The corresponding Gram matrix is denoted by $G_{r^M} = \left(\langle r_i, r_j \rangle_{\tilde{H}_r}\right)_{i,j=1}^M$.

A rank-M representation of the operator $M_{\overline{\mu}} A_{f,g}$ is then given by $P_{r^M} M_{\overline{\mu}} P_{\beta^M} A_{f,g} P_{d^M} : \tilde{H}_d \to \text{span}\, r^M$, where P_{r^M}, P_{d^M}, and P_{β^M} denote projection operators onto $\text{span}\, r^M$, $\text{span}\, d^M$, and $\text{span}\, \beta^M$, respectively. The construction is illustrated in Fig. 2.

Under the compactness assumptions and given rich enough data so that the spans of $\{d_i\}_{i=1}^\infty$, $\{r_i\}_{i=1}^\infty$, and $\{\beta_i\}_{i=1}^\infty$ are dense in $\tilde{H}_d$, $\tilde{H}_r$, and H, respectively, the sequence of finite-rank operators $\{P_{r^M} M_{\overline{\mu}} P_{\beta^M} A_{f,g} P_{d^M}\}_{M=1}^\infty$ can be shown to converge, in norm, to $M_{\overline{\mu}} A_{f,g}$. To facilitate the proof of convergence, we recall the following result from [3].

Lemma 1: Let H and G be RKHSs defined on $X \subset \mathbb{R}^n$ and let $A_N : H \to G$ be a finite-rank operator with rank N. If the spans of $\{d_i\}_{i=1}^\infty$ and $\{r_i\}_{i=1}^\infty$ are dense in H and G, respectively, then for all $\epsilon > 0$, there exists $M(N) \in \mathbb{N}$ such that for all $i \geq M(N)$ and $h \in H$, $\|A_N h - A_N P_{d^i} h\|_G \leq \epsilon \|h\|_H$ and $\|A_N h - P_{r^i} A_N h\|_G \leq \epsilon \|h\|_H$.

Proof: See the proof of [3, Theorem 2]. ∎

The convergence result for Liouville operators on Bargmann-Fock spaces restricted to the set of real numbers follows from

the following more general result.

Proposition 11: If $B : H \to \tilde{H}_r$ is a bounded linear operator, $A : \tilde{H}_d \to H$ is a compact operator, and the spans of $\{d_i\}_{i=1}^{\infty}$, $\{r_i\}_{i=1}^{\infty}$, and $\{\beta_i\}_{i=1}^{\infty}$ are dense in $\tilde{H}_d$, $\tilde{H}_r$, and H, respectively, then $\lim_{M \to \infty} \left\| BA - P_{r^M} BP_{\beta^M} AP_{d^M} \right\|_{\tilde{H}_d}^{\tilde{H}_r} = 0$, where $\|\cdot\|_{\tilde{H}_d}^{\tilde{H}_r}$ denotes the operator norm of operators from $\tilde{H}_d$ to $\tilde{H}_r$.

Proof: Let $\{A_N\}_{N=1}^{\infty}$ be a sequence of rank-N operators converging, in norm, to A. For an arbitrary $h \in \tilde{H}_d$,

$$\left\| BAh - P_{r^M} BP_{\beta^M} AP_{d^M} h \right\|_{\tilde{H}_r}$$
$$\leq \|BAh - BA_N h\|_{\tilde{H}_r} + \|BA_N h - BA_N P_{d^M} h\|_{\tilde{H}_r}$$
$$+ \left\| BA_N P_{d^M} h - BP_{\beta^M} A_N P_{d^M} h \right\|_{\tilde{H}_r}$$
$$+ \left\| BP_{\beta^M} A_N P_{d^M} h - P_{r^M} BP_{\beta^M} A_N P_{d^M} h \right\|_{\tilde{H}_r}$$
$$+ \left\| P_{r^M} BP_{\beta^M} A_N P_{d^M} h - P_{r^M} BP_{\beta^M} AP_{d^M} h \right\|_{\tilde{H}_r}.$$

Assuming that the operator norm of B is $\overline{B}$,

$$\left\| BAh - P_{r^M} BP_{\beta^M} AP_{d^M} h \right\|_{\tilde{H}_r} \leq \overline{B} \|Ah - A_N h\|_H +$$
$$\overline{B} \|A_N h - A_N P_{d^M} h\|_H + \overline{B} \|A_N P_{d^M} h - P_{\beta^M} A_N P_{d^M} h\|_H$$
$$+ \left\| BP_{\beta^M} A_N P_{d^M} h - P_{r^M} BP_{\beta^M} A_N P_{d^M} h \right\|_{\tilde{H}_r}$$
$$+ \overline{B} \|A_N P_{d^M} h - AP_{d^M} h\|_H.$$

Using the fact that A_N and $BP_{\beta^M} A_N P_{d^M}$ are finite-rank operators, Lemma 1, can be used to conclude that for all $\epsilon > 0$, there exists $M(N) \in \mathbb{N}$ such that for all $i \geq M(N)$

$$\left\| BAh - P_{r^i} BP_{\beta^i} AP_{d^i} h \right\|_{\tilde{H}_r} \leq \overline{B} \|Ah - A_N h\|_H$$
$$+ 3\overline{B}\epsilon \|h\|_{\tilde{H}_d} + \overline{B} \|A_N P_{d^i} h - AP_{d^i} h\|_H.$$

Since A_N converges to A in norm, given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $j \geq N$, and $g \in \tilde{H}_d$ $\|Ag - A_j g\|_H \leq \epsilon \|g\|_{\tilde{H}_d}$. Thus, for all $j \geq N$ and $i \geq M(j)$, $\left\| BAh - P_{r^i} BP_{\beta^i} AP_{d^i} h \right\|_{\tilde{H}_r} \leq 5\overline{B}\epsilon \|h\|_{\tilde{H}_d}$. ∎
The convergence result can then be stated as follows.

Theorem 1: Let $\rho_d \in \mathbb{R}$, $\varrho_d \in \mathbb{R}$, $\rho_r \in \mathbb{R}$, and $\rho = \begin{bmatrix} \rho_1 & \cdots & \rho_{m+1} \end{bmatrix}^{\top} \in \mathbb{R}^{m+1}$ be parameters such that $\rho_r < \rho_i$, $\rho_i < \varrho_d$, and $\varrho_d < \rho_d$ for $i = 1, \ldots, m+1$. Let $\tilde{H}_d = F_{\tilde{\rho}_d}^2(\mathbb{R}^n)$, $\tilde{G}_d = F_{\tilde{\varrho}_d}^2(\mathbb{R}^n)$, $\tilde{H}_r = F_{\tilde{\rho}_r}^2(\mathbb{R}^n)$, and $H = F_\rho^2(\mathbb{R}^n)$. If f, g, and μ are component-wise polynomial, and if the spans of the collections $\{d_i\}_{i=1}^{\infty}$, $\{r_i\}_{i=1}^{\infty}$, and $\{\beta_i\}_{i=1}^{\infty}$ are dense in $\tilde{H}_d$, $\tilde{H}_r$, and H, respectively, then $\lim_{M \to \infty} \left\| M_{\overline{\mu}} A_{f,g} - P_{r^M} M_{\overline{\mu}} P_{\beta^M} A_{f,g} P_{d^M} \right\|_{\tilde{H}_d}^{\tilde{H}_r} = 0$.

Proof: Propositions 8, 9, and 10 imply that $M_{\overline{\mu}}$ is bounded and $A_{f,g}$ is compact. Since multiplication operators are linear by definition, the theorem follows from Proposition 11. ∎

C. Matrix Representation of the Finite-rank Operator

To formulate a matrix representation of the finite-rank operator $P_{r^M} M_{\overline{\mu}} P_{\beta^M} A_{f,g} P_{d^M}$, the operator is restricted to $\text{span } d^M$ to yield the operator $P_{r^M} M_{\overline{\mu}} P_{\beta^M} A_{f,g}|_{d^M} : \text{span } d^M \to \text{span } r^M$. For brevity of exposition, the superscript M is suppressed hereafter and d, β, ϖ, and r are interpreted as $M-$dimensional vectors.

Proposition 12: If $h = \delta^{\top} d \in \text{span } d$ is a function with coefficients $\delta \in \mathbb{R}^M$ and if $g = P_r M_{\overline{\mu}} P_\beta A_{f,g} h$, then $g = a^{\top} r$, where $a = G_r^+ I G_\beta^+ G_d \delta$, $I := \left(\langle \varpi_i, \beta_j \rangle_H \right)_{i,j=1}^M$, and $(\cdot)^+$ denotes the Moore-Penrose pseudoinverse.

Proof: Propositions 4 and 7 imply that that for all $j = 1, \cdots, M$, $A_{f,g}^* \beta_j = d_j$, and $M_{\overline{\mu}}^* r_j = \varpi_j$, respectively. Note that since g is a projection of $M_{\overline{\mu}} P_\beta A_{f,g} h$ onto $\text{span } r$, $g = a^{\top} r$ for any a that solves

$$G_r a = \begin{bmatrix} \langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_1 \rangle_{\tilde{H}_r} \\ \vdots \\ \langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_M \rangle_{\tilde{H}_r} \end{bmatrix} = \begin{bmatrix} \langle A_{f,g} h, P_\beta M_{\overline{\mu}}^* r_1 \rangle_H \\ \vdots \\ \langle A_{f,g} h, P_\beta M_{\overline{\mu}}^* r_M \rangle_H \end{bmatrix}. \tag{19}$$

Furthermore, for all $j = 1, \cdots, M$, $P_\beta M_{\overline{\mu}}^* r_1 = b_j^{\top} \beta$, for any b_j that solves

$$G_\beta b_j = \begin{bmatrix} \langle M_{\overline{\mu}}^* r_j, \beta_1 \rangle_H \\ \vdots \\ \langle M_{\overline{\mu}}^* r_j, \beta_M \rangle_H \end{bmatrix} = \begin{bmatrix} \langle \varpi_j, \beta_1 \rangle_H \\ \vdots \\ \langle \varpi_j, \beta_M \rangle_H \end{bmatrix}. \tag{20}$$

As a result, $b := \begin{bmatrix} b_1, & \ldots, & b_M \end{bmatrix}$ is a solution of

$$G_\beta b = \left(\langle \varpi_j, \beta_i \rangle_H \right)_{i,j=1}^M = I^{\top}. \tag{21}$$

Substituting (20) into (19),

$$G_r a = \begin{bmatrix} \langle A_{f,g} h, b_1^{\top} \beta \rangle_H \\ \vdots \\ \langle A_{f,g} h, b_M^{\top} \beta \rangle_H \end{bmatrix} = \begin{bmatrix} \langle h, b_1^{\top} A_{f,g}^* \beta \rangle_{\tilde{H}_d} \\ \vdots \\ \langle h, b_M^{\top} A_{f,g}^* \beta \rangle_{\tilde{H}_d} \end{bmatrix},$$

where $A_{f,g}^* \beta$ is interpreted as $A_{f,g}^* \beta = \begin{bmatrix} A_{f,g}^* \beta_1, & \ldots, & A_{f,g}^* \beta_M \end{bmatrix}^{\top}$. Using $A_{f,g}^* \beta_j = d_j$ and $h = \delta^{\top} d$,

$$G_r a = \begin{bmatrix} \langle \delta^{\top} d, b_1^{\top} d \rangle_{\tilde{H}_d} \\ \vdots \\ \langle \delta^{\top} d, b_M^{\top} d \rangle_{\tilde{H}_d} \end{bmatrix} = \begin{bmatrix} b_1^{\top} G_d d \\ \vdots \\ b_M^{\top} G_d d \end{bmatrix}$$

Selecting solutions of (19) and (21) that minimize the 2-norm of a and b_j, respectively,

$$a = G_r^+ b^{\top} G_d \delta = G_r^+ I G_\beta^+ G_d \delta = G_r^+ I G_\beta^+ G_d \delta. \tag{22}$$

That is, a matrix representation $[M_{\overline{\mu}} P_\beta A_{f,g}]_d^r$ of the operator $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$ is given by $G_r^+ I G_\beta^+ G_d$. ∎
Note that matrix representations are generally not unique. Different representations may be obtained by selecting different solutions of (19) and (21). In the case where the Gram matrices G_r and G_β are nonsingular, equations (19) and (21) have unique solutions, resulting in the unique matrix representation $G_r^{-1} I G_\beta^{-1} G_d$.

In the following section, the matrix representation $[M_{\overline{\mu}} P_\beta A_{f,g}]_d^r$ is used to construct a data-driven representation of the singular values and the left and right singular functions of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$.

D. Singular Functions of the Finite-rank Operator

Recall that the tuples $\{(\sigma_i, \phi_i, \psi_i)\}_{i=1}^M$, with $\sigma_i \in \mathbb{R}^n$, $\phi_i \in \tilde{H}_d$, and $\psi_i \in \tilde{H}_r$, are singular values, left singular vectors,

and right singular vectors of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$, respectively, if $\forall h \in \text{span}\, d$, $P_r M_{\overline{\mu}} P_\beta A_{f,g} h = \sum_{i=1}^{M} \sigma_i \psi_i \langle h, \phi_i \rangle_{\tilde{H}_d}$. The following proposition states that the SVD of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$ can be computed using matrices in the matrix representation $[M_{\overline{\mu}} P_\beta A_{f,g}]_d^r$ derived in the previous section.

Proposition 13: If (W, Σ, V) is the SVD of $G_r^+ I G_\beta^+$ with $W = \begin{bmatrix} w_1, & \ldots, & w_M \end{bmatrix}$, $V = \begin{bmatrix} v_1, & \ldots, & v_M \end{bmatrix}$, and $\Sigma = \text{diag}\left(\begin{bmatrix} \sigma_1, & \ldots, & \sigma_M \end{bmatrix}\right)$, then for all $i = 1, \ldots, M$, σ_i are singular values of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$ with left singular functions $\phi_i := v_i^\top d$ and right singular functions $\psi_i := w_i^\top r$.

Proof: Let $\phi_i = v_i^\top d$ and $\psi_i = w_i^\top r$ and $h = \delta^\top d$. Then,

$$P_r M_{\overline{\mu}} P_\beta A_{f,g} h = \sum_{i=1}^{M} \sigma_i \psi_i \langle h, \phi_i \rangle_{\tilde{H}_d}$$

$$\iff P_r M_{\overline{\mu}} P_\beta A_{f,g} \delta^\top d = \sum_{i=1}^{M} \sigma_i w_i^\top r \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d}$$

Using the finite-rank representation, the collection $\{(\sigma_i, \phi_i, \psi_i)\}_{i=1}^{M}$, is an SVD of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_d$, if for all $\delta \in \mathbb{R}^M$,

$$\left(G_r^+ I G_\beta^+ G_d \delta \right)^\top r = \left(\sum_{i=1}^{M} \sigma_i \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d} w_i^\top \right) r. \quad (23)$$

Simple matrix manipulations yield the chain of implications

$$(23) \Longleftarrow \forall \delta \in \mathbb{R}^M, G_r^+ I G_\beta^+ G_d \delta = \sum_{i=1}^{M} \sigma_i \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d} w_i$$

$$\iff \forall \delta \in \mathbb{R}^M, G_r^+ I G_\beta^+ G_d \delta = \sum_{i=1}^{M} \sigma_i \left(w_i v_i^\top G_d \right) \delta$$

$$\Longleftarrow G_r^+ I G_\beta^+ G_d = \sum_{i=1}^{M} \sigma_i \left(w_i v_i^\top \right) G_d$$

$$\Longleftarrow G_r^+ I G_\beta^+ = \sum_{i=1}^{M} \sigma_i w_i v_i^\top = W \Sigma V^\top,$$

which proves the proposition. ∎

In the following section, the singular values and the left and right singular vectors are used, along with a finite truncation of (5) to generate a data-driven model.

E. The SCLDMD Algorithm

Motivated by (4), assuming that $h_{\text{id},j} \in \tilde{H}_d$ for $j = 1, \cdots, n$, the system dynamics are approximated using the rank-M representation as $\dot{x} \approx \hat{F}_{\mu,M}(x) := [P_r M_{\overline{\mu}} P_\beta A_{f,g} P_d h_{\text{id}}](x)$, where $P_r M_{\overline{\mu}} P_\beta A_{f,g} P_d h_{\text{id}}$ denotes row-wise operation of the operator $P_r M_{\overline{\mu}} P_\beta A_{f,g} P_d$ on the function h_{id}. Since $P_r M_{\overline{\mu}} P_\beta A_{f,g} P_d$ converges to $M_{\overline{\mu}} A_{f,g}$ in norm as $M \to \infty$, and since the space $F_{\tilde{\rho}_d}^2(\mathbb{R}^n)$ contains $h_{\text{id},j}$ for $j = 1, \cdots, n$, the following result is immediate.

Corollary 1: Under the hypothesis of Theorem 1, $\lim_{M \to \infty} \left(\sup_{x \in X} \left\| \hat{F}_{\mu,M}(x) - F_\mu(x) \right\|_2 \right) = 0$.

Proof: Since the space $F_{\tilde{\rho}_d}^2(\mathbb{R}^n)$ contains $h_{\text{id},j}$ for $j = 1, \cdots, n$, the functions $\hat{F}_{\mu,M,j} := P_{rj} M_{\overline{\mu}} P_{\beta j} A_{f,g} P_{dj} h_{\text{id},j}$ and $F_{\mu,j} := M_{\overline{\mu}} A_{f,g} h_{\text{id},j}$ that denote the $j-$th row of $\hat{F}_{\mu,M}$

and F_μ, respectively, exist as members of $\tilde{H}_r$. Since $x \mapsto \tilde{K}_r(x,x) = \exp\left(\frac{x^\top x}{\tilde{\rho}_r}\right)$ is continuous and X is compact, there exists a real number $\overline{K}$ such that $\sup_{x \in X} \tilde{K}_r(x,x) = \overline{K}$. Theorem 1 can then be used to conclude that for all $\epsilon > 0$ and $j = 1, \ldots, n$, there exists $M(j) \in \mathbb{N}$ such that for all $i \geq M(j)$, $\left\| \hat{F}_{\mu,i,j} - F_{\mu,j} \right\|_{\tilde{H}_r}^2 \leq \frac{\epsilon^2}{n\overline{K}^2}$. Using the reproducing property, for $i \geq \overline{M} := \max_j M(j)$,

$$\left\| \hat{F}_{\mu,i}(x) - F_\mu(x) \right\|_2^2 = \sum_{j=1}^{n} \left\langle \left(\hat{F}_{\mu,i,j} - F_{\mu,j} \right), \tilde{K}_r(\cdot,x) \right\rangle_{\tilde{H}_r}^2$$

$$\leq \sum_{j=1}^{n} \left\| \hat{F}_{\mu,i,j} - F_{\mu,j} \right\|_{\tilde{H}_r}^2 \left\| \tilde{K}_r(\cdot,x) \right\|_{\tilde{H}_r}^2$$

$$\leq \sum_{j=1}^{n} \frac{\epsilon^2}{n\overline{K}^2} \left\langle \tilde{K}_r(\cdot,x), \tilde{K}_r(\cdot,x) \right\rangle_{\tilde{H}_r}^2 = \frac{\epsilon^2}{\overline{K}^2} \tilde{K}_r(x,x)^2.$$

As a result, for all $\epsilon \geq 0$ there exists $\overline{M}$ such that for all $i \geq \overline{M}$,

$$\sup_{x \in X} \left\| \hat{F}_{\mu,i}(x) - F_\mu(x) \right\|_2 \leq \sqrt{\frac{\epsilon^2}{\overline{K}^2} \sup_{x \in X} \tilde{K}_r(x,x)^2} = \epsilon,$$

which completes the proof. ∎

Using the definition of singular values and singular functions,

$$\dot{x} \approx \sum_{i=1}^{M} \sigma_i \xi_i w_i^\top r(x) = \xi \Sigma W^\top r(x), \quad (24)$$

where $\xi_i := \langle P_d h_{\text{id}}, \phi_i \rangle_{\tilde{H}_d}$ and $\xi := \begin{bmatrix} \xi_1, & \ldots, & \xi_M \end{bmatrix}$.

The modes ξ can be computed using $\phi_i = v_i^\top d$ as

$$\xi = \begin{bmatrix} \langle P_d h_{\text{id},1}, v_1^\top d \rangle_{\tilde{H}_d}, & \ldots, & \langle P_d h_{\text{id},1}, v_M^\top d \rangle_{\tilde{H}_d} \\ \vdots & \ddots & \vdots \\ \langle P_d h_{\text{id},n}, v_1^\top d \rangle_{\tilde{H}_d}, & \ldots, & \langle P_d h_{\text{id},n}, v_M^\top d \rangle_{\tilde{H}_d} \end{bmatrix}$$

$$= \begin{bmatrix} \langle \delta_1^\top d, d_1 \rangle_{\tilde{H}_d}, & \ldots, & \langle \delta_1^\top d, d_M \rangle_{\tilde{H}_d} \\ \vdots & \ddots & \vdots \\ \langle \delta_n^\top d, d_1 \rangle_{\tilde{H}_d}, & \ldots, & \langle \delta_n^\top d, d_M \rangle_{\tilde{H}_d} \end{bmatrix} V = \delta^\top G_d V,$$

where $\delta := \begin{bmatrix} \delta_1, & \ldots, & \delta_n \end{bmatrix}$. Using the reproducing property of the reproducing kernel of $\tilde{H}_d$, the coefficients δ_i in the projection of $h_{\text{id},i}$ onto d satisfy

$$G_d \delta_i = \begin{bmatrix} \langle (h_{\text{id}})_i, d_1 \rangle_{\tilde{H}_d} \\ \vdots \\ \langle (h_{\text{id}})_i, d_M \rangle_{\tilde{H}_d} \end{bmatrix} = \begin{bmatrix} (\gamma_{u_1}(T_1))_i - (\gamma_{u_1}(0))_i \\ \vdots \\ (\gamma_{u_M}(T_M))_i - (\gamma_{u_M}(0))_i \end{bmatrix}.$$

Letting $D := \left((\gamma_{u_j}(T_j))_i - (\gamma_{u_j}(0))_i \right)_{i,j=1}^{n,M}$ it can be concluded that $\delta^\top G_d = D$. Finally, the modes ξ are given by $\xi = DV$ and the estimated closed-loop model is given by

$$\dot{x} \approx \hat{F}_{\mu,M}(x) = DV \Sigma W^\top r(x) = D G_\beta^+ I^\top G_r^+ r(x) \quad (25)$$

The SCLDMD technique is summarized in Algorithm 1. The characterization $\Gamma_{\gamma_{u_j}} = \int_0^{T_j} \tilde{K}\left(\cdot, \gamma_{u_j}(t)\right) dt$ of occupation kernels, introduced in [23], is used on line 8.

Algorithm 1 The SCLDMD algorithm

Input: Trajectories $\{\gamma_{u_i}\}_{i=1}^M$, a feedback law μ, a numerical integration procedure, reproducing kernels $\tilde{K}_d$, $\tilde{K}_r$, and K of $\tilde{H}_d$, $\tilde{H}_r$, and H, respectively, and regularization parameters ϵ_r and $\tilde{\epsilon}$.

Output: $\{\xi_j, \sigma_j, \varphi_j, \phi_j\}_{j=1}^M$

1: $G_\beta \leftarrow \left(\left\langle \Gamma_{\gamma_{u_i}, u_i}, \Gamma_{\gamma_{u_j}, u_j} \right\rangle_H \right)_{i,j=1}^M$ (see (34))

2: $G_r \leftarrow \left(\left\langle \Gamma_{\gamma_{u_i}}, \Gamma_{\gamma_{u_j}} \right\rangle_{\tilde{H}_r} \right)_{i,j=1}^M$ (see (35))

3: $I \leftarrow \left(\left\langle \Gamma_{\gamma_{u_i}, \mu \circ \gamma_{u_i}}, \Gamma_{\gamma_{u_j}, u_j} \right\rangle_H \right)_{i,j=1}^M$ (see (37))

4: $D \leftarrow \left((\gamma_{u_j}(T_j))_i - (\gamma_{u_j}(0))_i \right)_{i,j=1}^{n,M}$

5: $(W, \Sigma, V) \leftarrow$ SVD of $G_r^+ I G_\beta^+$ (See Remark 2)

6: $\xi \leftarrow DV$

7: $\phi_j \leftarrow \sum_{i=1}^M \leftarrow (V)_{i,j} \left(K_d(\cdot, \gamma_{u_i}(T_i)) - K_d(\cdot, \gamma_{u_i}(0)) \right)$

8: $\psi_j \leftarrow \sum_{i=1}^M \int_0^{T_i} (W)_{i,j} \tilde{K}\left(\cdot, \gamma_{u_i}(t) \right) \mathrm{d}t$

9: **return** $\{\xi_j, \sigma_j, \varphi_j, \phi_j\}_{j=1}^M$

VII. EIGENDECOMPOSITION APPROACH TO DMD

In this section, an alternative finite-rank representation of the operator $M_{\overline{\mu}} A_{f,g}$ is presented, where its domain and range are assumed to be subsets of the same RKHS $\tilde{H}$ of complex-valued continuously differentiable functions, with a real-valued reproducing kernel $\tilde{K}$. In particular, the finite-rank representation of $M_{\overline{\mu}} A_{f,g}$ is selected to be $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r$, where the domain and the range are both $\operatorname{span} r$. A consequence of this choice is that the finite-rank representation admits eigenfunctions which could potentially generate an approximate invariant subspace of the closed-loop system.

While eigenfunctions of the finite-rank representation exist, they generally cannot be shown to converge to eigenfunctions of the original operator, since the operators $M_{\overline{\mu}}$ and $A_{f,g}$ can no longer be assumed to be bounded and compact, respectively. Instead, they are assumed to be densely defined. Since the operators are not defined everywhere, we need the additional assumptions that 1) the image of $A_{f,g}$ is contained within the domain of $M_{\overline{\mu}}$, 2) the span of r is a subset of the domain of $A_{f,g}$, and 3) the functions $h_{\mathrm{id},j}$ can be well-approximated by linear combinations of the eigenfunctions of the finite-rank representation for $j = 1, \ldots, n$. Due to the lack of convergence guarantees and since the assumptions on $\operatorname{span} r$ and $h_{\mathrm{id},j}$ are difficult to verify, the resulting algorithm, while useful, is heuristic in nature. Since unbounded operators over Hilbert spaces of real-valued functions can have empty spectra, in this section, the RKHS $\tilde{H}$ is assumed to be composed of complex-valued functions of real variables of the form $h : X \to \mathbb{C}$.

The operators $M_{\overline{\mu}}$ and $A_{f,g}$ are densely defined in a large class of problems. For example, if the domain and range spaces in Section VI-A are selected to have identical kernel parameters, then the resulting operators are densely defined [23], and the image of $A_{f,g}$ is also contained within the domain of $M_{\overline{\mu}}$. The assumption that $\operatorname{span} r \subset \mathcal{D}(A_{f,g})$ can be removed in favor of the assumption that the matrix that encodes the

finite rank representation of $M_{\overline{\mu}} A_{f,g}$ is approximately equal to the transpose of the matrix that encodes the finite rank representation of the adjoint $A_{f,g}^* M_{\overline{\mu}}^*$ (see [24]).

A. Matrix Representation of the Finite-rank Operator

In this section, a matrix representation of the finite-rank representation is developed.

Proposition 14: If $h = \delta^\top r \in \operatorname{span} r$ is a function with coefficients $\delta \in \mathbb{C}^M$, $A_{f,g}$ and $M_{\overline{\mu}}$ are densely defined, $\operatorname{span} r \subset \mathcal{D}(A_{f,g})$, $\operatorname{span} \beta \subset \mathcal{D}(M_{\overline{\mu}})$, and $g = P_r M_{\overline{\mu}} P_\beta A_{f,g} h$, then $g = a^\top r$ with $a = G_r^+ I G_\beta^+ \tilde{I}^\top \delta$, where $I := \left(\left\langle M_{\overline{\mu}}^* r_j, \Gamma_{\gamma_{u_k}, u_k} \right\rangle_H \right)_{j,k=1}^M$ and $\tilde{I} := \left(\left\langle r_i, A_{f,g}^* \Gamma_{\gamma_{u_k}, u_k} \right\rangle_{\tilde{H}} \right)_{i,k=1}^M$.

Proof: The coefficients $a = \begin{bmatrix} a_1 & \cdots & a_M \end{bmatrix}^\top \in \mathbb{C}^M$ in the projection of $M_{\overline{\mu}} P_\beta A_{f,g} h \in \tilde{H}$ onto $\operatorname{span} r$ are given by the solution of the linear system

$$G_r a = \begin{bmatrix} \langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_1 \rangle_{\tilde{H}} \\ \vdots \\ \langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_M \rangle_{\tilde{H}} \end{bmatrix}. \quad (26)$$

A matrix representation of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r$ relates the coefficients $\delta = \begin{bmatrix} \delta_1 & \cdots & \delta_M \end{bmatrix}^\top \in \mathbb{C}^M$ of a function $h = \delta^\top r \in \operatorname{span} r$, with the coefficients a above. Using the properties of the multiplication operator and the control Liouville operator established in the previous sections, the inner products $\langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_j \rangle_{\tilde{H}}$ on the right hand side can be evaluated as

$$\langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_j \rangle_{\tilde{H}} = \sum_{i=1}^M \delta_i \left\langle A_{f,g} r_i, P_\beta M_{\overline{\mu}}^* r_j \right\rangle_H$$

$$= \sum_{i=1}^M \delta_i \left\langle A_{f,g} r_i, \sum_{k=1}^M b_{k,j} \Gamma_{\gamma_{u_k}, u_k} \right\rangle_H,$$

where $\{b_{k,j}\}_{k=1}^M \subset \mathbb{R}$ are the coefficients in the projection of $M_{\overline{\mu}}^* r_j \in H$ onto $\operatorname{span} \beta$, which can be computed by solving

$$G_\beta \begin{bmatrix} b_{1,j} \\ \vdots \\ b_{M,j} \end{bmatrix} = \begin{bmatrix} \left\langle M_{\overline{\mu}}^* r_j, \Gamma_{\gamma_{u_1}, u_1} \right\rangle_H \\ \vdots \\ \left\langle M_{\overline{\mu}}^* r_j, \Gamma_{\gamma_{u_M}, u_M} \right\rangle_H \end{bmatrix}. \quad (27)$$

Note that since the control occupation kernels $\beta_i = \Gamma_{\gamma_{u_i}, u_i}$ the occupation kernels $r_i = \Gamma_{\gamma_{u_i}}$, and the symbol $\overline{\mu}$ are all real-valued functions, the coefficients $b_{i,j}$ are real numbers. The inner product can thus be further simplified as $\langle M_{\overline{\mu}} P_\beta A_{f,g} h, r_j \rangle_{\tilde{H}} = \sum_{i=1}^M \delta_i \sum_{k=1}^M b_{k,j} \left\langle r_i, A_{f,g}^* \Gamma_{\gamma_{u_k}, u_k} \right\rangle_{\tilde{H}} = \delta^\top \tilde{I} b_j$, where $b_j := \begin{bmatrix} b_{1,j} & \cdots & b_{M,j} \end{bmatrix}^\top$. Stacking the inner products on the left hand side in a column and selecting solutions of (26) and (27) that minimize the 2-norm of a and b_j, respectively, it can be concluded that $a = G_r^+ I G_\beta^+ \tilde{I}^\top \delta$. A matrix representation $[M_{\overline{\mu}} P_\beta A_{f,g}]_r^r \in \mathbb{R}^M$ of the finite-rank representation $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r$ of the operator $M_{\overline{\mu}} A_{f,g}$ is thus given by $[M_{\overline{\mu}} P_\beta A_{f,g}]_r^r = G_r^+ I G_\beta^+ \tilde{I}^\top$. ∎

In the following section, the matrix representation $[M_{\overline{\mu}}P_\beta A_{f,g}]^r_r$ is used to construct a data-driven representation of the eigenvalues and the eigenfunctions of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r$.

B. Eigenfunctions of the finite-rank representation

Given an eigenvalue $\tilde{\lambda}_j \in \mathbb{C}$ and the corresponding eigenvector $\tilde{v}_j := \begin{bmatrix} \tilde{v}_{1,j} & \cdots & \tilde{v}_{M,j} \end{bmatrix}^\top \in \mathbb{C}^M$ of $[M_{\overline{\mu}}P_\beta A_{f,g}]^r_r$ and the vector $r := \begin{bmatrix} r_1 & \cdots & r_M \end{bmatrix}^\top$ of occupation kernels in $\tilde{H}$, it is straightforward to show that $\varphi_j = \left(1/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right) \tilde{v}_j^\top r$ is an eigenfunction of $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r$, where $(\cdot)^\dagger$ denotes the conjugate transpose. Indeed, by the definition of the matrix $[M_{\overline{\mu}}P_\beta A_{f,g}]^r_r$, it can be seen that $P_r M_{\overline{\mu}} P_\beta A_{f,g}|_r \varphi_j = \left(1/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right) \left([M_{\overline{\mu}}P_\beta A_{f,g}]^r_r \tilde{v}_j \right)^\top r = \tilde{\lambda}_j \left(1/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right) \tilde{v}_j^\top r$.

Using the fact that $r_i(x) = \Gamma_{\gamma_{u_i}}(x) = \int_0^{T_i} \tilde{K}(x, \gamma_{u_i}(t)) \mathrm{d}t$, the eigenfunctions, evaluated at a point $x \in X$, can be computed as

$$\varphi_j(x) = \frac{1}{\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j}} \sum_{i=1}^M \tilde{v}_{i,j} \int_0^{T_i} \tilde{K}(x, \gamma_{u_i}(t)) \, \mathrm{d}t. \quad (28)$$

In the following section, the eigenvalues and the eigenfunctions are used to generate a data-driven model.

C. The CLDMD Algorithm

Let $\tilde{W} = \left(\tilde{v}_{i,j}/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right)_{i,j=1}^M \in \mathbb{C}^{M \times M}$ be the matrix of coefficients of the normalized eigenfunctions, arranged so that each column corresponds to an eigenfunction. Assuming that $h_{\mathrm{id},j}$ is in the span of the above eigenfunctions for each $j = 1, \cdots, n$, a representation of the identity function as a linear combination of a fixed number of eigenfunctions is given as $h_{\mathrm{id}}(x) \approx \sum_{i=1}^M \xi_i \varphi_i(x)$, where $\{\xi_i\}_{i=1}^M \subset \mathbb{C}^n$ are the so-called *control-Liouville modes*. Similar to [3, Section 4.2], by examining the inner products $\langle h_{\mathrm{id},j}, r_i \rangle_{\tilde{H}}$, the matrix $\xi := \begin{bmatrix} \xi_1, & \cdots, & \xi_M \end{bmatrix}$ can be shown to be a solution of the linear system of equations

$$\xi \left(\tilde{W}^\top G_r \overline{\tilde{W}} \right) = R \overline{\tilde{W}}, \quad (29)$$

where $R := \left(\langle h_{\mathrm{id},j}, r_i \rangle_{\tilde{H}} \right)_{j,i=1}^{n,M}$ and $\overline{\tilde{W}}$ denotes the complex conjugate of $\tilde{W}$. Indeed, letting $\xi_{i,j}$ denote the $j-$th element of the vector ξ_i, the row of coefficients $\xi^j := \begin{bmatrix} \xi_{1,j}, & \ldots, \xi_{M,j} \end{bmatrix} \in \mathbb{C}^{1 \times M}$ in the projection of $h_{\mathrm{id},j}$ onto the span of the eigenfunctions $\{\varphi_i\}_{i=1}^M$ is a solution of

$$\xi^j G_\varphi^\top = \begin{bmatrix} \langle h_{\mathrm{id},j}, \varphi_1 \rangle_{\tilde{H}}, & \ldots, & \langle h_{\mathrm{id},j}, \varphi_M \rangle_{\tilde{H}} \end{bmatrix}, \quad (30)$$

where $G_\varphi = \left(\langle \varphi_i, \varphi_j \rangle_{\tilde{H}} \right)_{i,j=1}^M \in \mathbb{C}^{M \times M}$. Using the fact that $\langle \varphi_i, \varphi_j \rangle_{\tilde{H}} = \langle \tilde{w}_i^\top r, \tilde{w}_j^\top r \rangle_{\tilde{H}} = \tilde{w}_j^\dagger G_r \tilde{w}_i$, where $\tilde{w}_j := \left(1/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right) \tilde{v}_j$, denotes the $j-$th column of $\tilde{W}$, the Gram matrix G_φ can be expressed as $G_\varphi = \tilde{W}^\dagger G_r \tilde{W}$. Furthermore, using the fact that $\langle h_{\mathrm{id},j}, \varphi_i \rangle_{\tilde{H}} = \langle h_{\mathrm{id},j}, \tilde{w}_i^\top r \rangle_{\tilde{H}} = \begin{bmatrix} \langle h_{\mathrm{id},j}, r_1 \rangle_{\tilde{H}}, & \ldots, & \langle h_{\mathrm{id},j}, r_M \rangle_{\tilde{H}} \end{bmatrix} \overline{\tilde{w}_i}$, where $\overline{\tilde{w}_i}$ denotes the complex conjugate of $\tilde{w}_i$, the right hand side of (30) can be expressed as $\begin{bmatrix} \langle h_{\mathrm{id},j}, \varphi_1 \rangle_{\tilde{H}}, & \ldots, & \langle h_{\mathrm{id},j}, \varphi_M \rangle_{\tilde{H}} \end{bmatrix} = \begin{bmatrix} \langle h_{\mathrm{id},j}, r_1 \rangle_{\tilde{H}}, & \ldots, & \langle h_{\mathrm{id},j}, r_M \rangle_{\tilde{H}} \end{bmatrix} \overline{\tilde{W}}$. Concatenating (30) for $j = 1, \ldots, n$ into a column vector, the matrix ξ is seen

to be a solution of (29). Using the fact that any solution of $\xi \tilde{W}^\top G_r = R$ is also a solution of (29), selecting the solution of $\xi \tilde{W}^\top G_r = R$ that minimizes the 2-norm of ξ_i for $i = 1, \cdots, M$, and using the relationship $\langle h_{\mathrm{id},j}, r_i \rangle_{\tilde{H}} = \langle h_{\mathrm{id},j}, \Gamma_{\gamma_{u_i}} \rangle_{\tilde{H}} = \int_0^{T_i} \gamma_{u_i,j}(t) \mathrm{d}t$, where $\gamma_{u_i,j}(t)$ denotes the $j-$th component of $\gamma_{u_i}(t)$, a set of control Liouville modes can be obtained as

$$\xi = \begin{bmatrix} \int_0^{T_1} \gamma_{u_1}(t) \mathrm{d}t & \cdots & \int_0^{T_M} \gamma_{u_M}(t) \mathrm{d}t \end{bmatrix} \left(\tilde{W}^\top G_r \right)^+. \quad (31)$$

The response $t \mapsto \gamma_\mu(t)$ of the system, starting from the initial condition $\gamma_\mu(0) = \gamma_0$, under the feedback control law μ, can then be predicted as

$$\gamma_\mu(t) \approx \sum_{j=1}^M \xi_j \varphi_j(\gamma_0) \mathrm{e}^{\tilde{\lambda}_j t}. \quad (32)$$

Furthermore, a pointwise approximation of the closed-loop model can also be obtained as

$$\dot{x} \approx \hat{F}_{\mu,M}(x) := \sum_{j=1}^M \tilde{\lambda}_j \xi_j \varphi_j(x). \quad (33)$$

The CLDMD method is summarized in Algorithm 2.

Algorithm 2 The CLDMD algorithm

Input: Trajectories $\{\gamma_{u_i}\}_{i=1}^M$, a feedback law μ, a numerical integration procedure, Reproducing kernel $\tilde{K}$ of $\tilde{H}$, Reproducing kernel K of H, and if needed, regularization parameters ϵ and $\tilde{\epsilon}$.
Output: $\{\xi_j, \lambda_j, \varphi_j\}_{j=1}^M$

1: $G_\beta \leftarrow \left(\left\langle \Gamma_{\gamma_{u_i}, u_i}, \Gamma_{\gamma_{u_j}, u_j} \right\rangle_H \right)_{i,j=1}^M$ (see (34))
2: $G_r \leftarrow \left(\langle r_i, r_j \rangle_{\tilde{H}} \right)_{i,j=1}^M$ (see (35))
3: $I \leftarrow \left(\left\langle M_{\overline{H}}^* r_j, \Gamma_{\gamma_{u_k}, u_k} \right\rangle_H \right)_{j,k=1}^M$ (see (37))
4: $\tilde{I} \leftarrow \left(\left\langle r_i, A_{f,g}^* \Gamma_{\gamma_{u_k}, u_k} \right\rangle_{\tilde{H}} \right)_{i,k=1}^M$ (see (36))
5: $[M_{\overline{\mu}}P_\beta A_{f,g}]^r_r \leftarrow G_r^+ I G_\beta^+ \tilde{I}^\top$ (See Remark 2)
6: $\{\lambda_j, \tilde{v}_j\}_{j=1}^M \leftarrow$ eigendecomposition of $[M_{\overline{\mu}}P_\beta A_{f,g}]^r_r$
7: $\tilde{W} \leftarrow \left(\tilde{v}_{i,j}/\sqrt{\tilde{v}_j^\dagger G_r \tilde{v}_j} \right)_{i,j=1}^M$
8: Compute $\{\xi_j, \varphi_j\}_{j=1}^M$ using (31) and (28) (See Remark 2)
9: **return** $\{\xi_j, \lambda_j, \varphi_j\}_{j=1}^M$

VIII. COMPUTATION OF INNER PRODUCTS

The elements of the Gram matrix G_β, corresponding to β, can be computed using Proposition 3 as

$$\left\langle \Gamma_{\gamma_{u_i}, u_i}, \Gamma_{\gamma_{u_j}, u_j} \right\rangle_H$$
$$= \int_0^{T_j} \int_0^{T_i} \begin{bmatrix} 1 & u_i^\top(\tau) \end{bmatrix} K\left(\gamma_{u_j}(t), \gamma_{u_i}(\tau) \right) \begin{bmatrix} 1 \\ u_j(t) \end{bmatrix} \mathrm{d}\tau \mathrm{d}t \quad (34)$$

The elements of the Gram matrix G_r can be computed using

the double integral (cf. [3])

$$\left\langle \Gamma_{\gamma_{u_i}}, \Gamma_{\gamma_{u_j}} \right\rangle_{\tilde{H}} = \int_0^{T_j} \int_0^{T_i} \tilde{K}\left(\gamma_{u_j(t)}, \gamma_{u_i(\tau)}\right) \mathrm{d}\tau \mathrm{d}t. \quad (35)$$

Using Proposition 4, the elements of the interaction matrix $\tilde{I}$ can be evaluated as

$$\left\langle \Gamma_{\gamma_{u_i}}, A^*_{f,g} \Gamma_{\gamma_{u_k}, u_k} \right\rangle_{\tilde{H}} = \left\langle \Gamma_{\gamma_{u_i}}, \tilde{K}(\cdot, \gamma_{u_k}(T_k)) - \tilde{K}(\cdot, \gamma_{u_k}(0)) \right\rangle_{\tilde{H}}$$

$$= \int_0^{T_i} \left(\tilde{K}(\gamma_{u_i}(t), \gamma_{u_k}(T_k)) - \tilde{K}(\gamma_{u_i}(t), \gamma_{u_k}(0)) \right) \mathrm{d}t. \quad (36)$$

Using Proposition 7, the elements of the interaction matrix I can be evaluated as

$$\left\langle M^*_{\tilde{\mu}} \Gamma_{\gamma_{u_i}}, \Gamma_{\gamma_{u_k}, u_k} \right\rangle_H = \left\langle \Gamma_{\gamma_{u_i}, \mu \circ \gamma_{u_i}}, \Gamma_{\gamma_{u_k}, u_k} \right\rangle_H$$

$$= \int_0^{T_j} \int_0^{T_i} \left[1 \quad \mu^\top(\gamma_{u_i}(\tau)) \right] K\left(\gamma_{u_j}(t), \gamma_{u_i}(\tau)\right) \begin{bmatrix} 1 \\ u_j(t) \end{bmatrix} \mathrm{d}\tau \mathrm{d}t. \quad (37)$$

Assuming that each trajectory is sampled at N points in time, the computation of G_β and I is $O(nN^2M^2(m+1)^2)$, the computation of G_r is $O(nN^2M^2)$, and the computation of $\tilde{I}$ is $O(nNM^2)$. Computation of the finite-rank representation and its decomposition are $O(M^3)$. Evaluation of the occupation kernel is $O(nN)$.

Remark 2: The SCLDMD and CLDMD algorithms can also be implemented using regularization, where the Gram matrices G_β and G_r, if singular, are replaced by $G_\beta + \epsilon I_M$ and $G_r + \tilde{\epsilon} I_M$, respectively, where I_M denotes the $M \times M$ identity matrix, and $\epsilon > 0$ and $\tilde{\epsilon} > 0$ are regularization coefficients.

IX. NUMERICAL EXPERIMENTS

Two numerical experiments are performed to evaluate the developed SCLDMD and CLDMD methods, one using a simulated controlled Duffing oscillator and another using a simulated two-link robot manipulator.

A. Controlled Duffing oscillator

This experiment concerns the controlled Duffing oscillator

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = x_1 - x_1^3 + (2 + \sin(x_1))u,$$

where $x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$ is the state and $u \in \mathbb{R}$ is the control. A total of 225 open-loop trajectories of the controlled Duffing oscillator are generated using the MATLAB® ode45 solver, starting from initial conditions on a 15×15 regular grid on a 6×6 square centered at the origin of the state space, $\mathbb{R}^2$. The control signal used for trajectory generation is of the form $u(t) = \sum_{i=1}^{15} b_i \sin(\omega_i t + \varphi_i)$, where the magnitudes b_i, the frequencies ω_i, and the phase differences φ_i are generated randomly from a uniform distribution on the interval $[-1, 1]$. All trajectories are recorded over a duration of 1s, and are sampled at a frequency of 20 Hz.

The trajectories are then utilized to predict the behavior of the oscillator under the state feedback controller $\mu(x) =$

$\begin{bmatrix} -2 & -2 \end{bmatrix} x$. CLDMD is implemented using the exponential dot product reproducing kernel $\tilde{K}_{\tilde{\rho}} = \exp\left(\frac{x^\top y}{\tilde{\rho}}\right)$ with parameter $\tilde{\rho} = 5$, and a diagonal kernel given by $K = \mathrm{diag}\begin{bmatrix} \tilde{K}_{\tilde{\rho}} & \tilde{K}_{\tilde{\rho}} \end{bmatrix}$. SCLDMD is implemented using $\tilde{K}_r = \tilde{K}_5$, $K = \mathrm{diag}\begin{bmatrix} \tilde{K}_6 & \tilde{K}_6 \end{bmatrix}$, and $\tilde{K}_d = \tilde{K}_7$. Simpson's 1/3 rule is used to compute the integrals involved in algorithms 1 and 2.

1) Vector Field Reconstruction: Fig. 3 shows a side by side comparison of the pointwise 2-norm of the relative error between the approximated vector field $\hat{F}_{\mu,M}$ (generated using (25) for SCLDMD and (33) for CLDMD), and the true vector field, F_μ. The results in Fig. 3 indicate that both the CLDMD and the SCLDMD methods are able to obtain accurate estimates of the closed-loop vector field on a domain contained within the grid of initial conditions of the data.
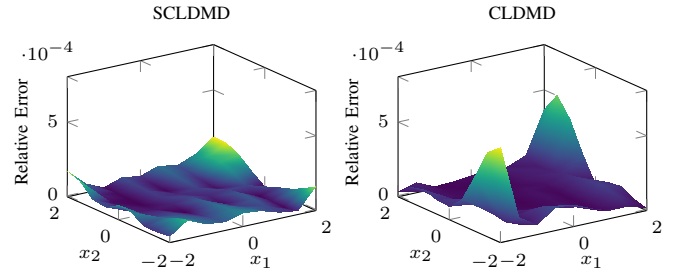

Fig. 3. Relative error $\frac{\|F_\mu(x) - \hat{F}_\mu(x)\|_2}{\max_{x \in [-2,2] \times [-2,2]} \|F_\mu(x)\|_2}$ in the estimation of the vector field F_μ of the controlled Duffing oscillator as a function of x, obtained using SCLDMD (left) and CLDMD (right).

2) Indirect Closed-loop Response Prediction: The closed loop response can be predicted using either SCLDMD or CLDMD by numerically solving the initial value problems in (25) and (33), respectively, starting from the desired initial condition. Fig. 4 shows the prediction error resulting from this indirect approach, starting from $x_0 = \begin{bmatrix} 2 & -2 \end{bmatrix}^\top$. The results in Fig. 4 indicate that both the CLDMD and the SCLDMD methods, when coupled with indirect prediction, accurately predict the desired closed-loop trajectory.
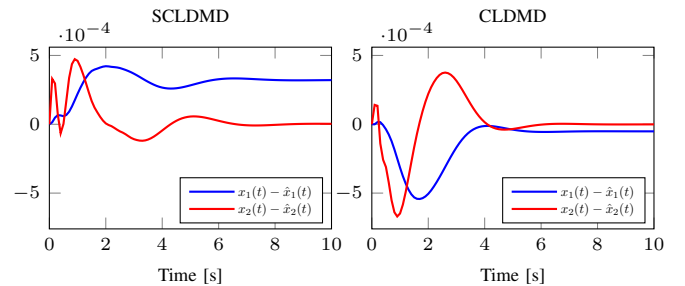

Fig. 4. Error between predicted and true trajectories of the controlled duffing oscillator for the experiment in Section IX-A. The figure on the left is obtained using SCLDMD indirect prediction by solving (25) and the figure on the right is obtained using CLDMD indirect prediction by solving (33), both using the MATLAB® ode45 solver.

3) Direct Closed-loop Response Prediction: The CLDMD method can also be used to predict the behavior of the closed-loop system starting from a given initial condition, and under the given feedback controller. Direct reconstruction is implemented using (32). Fig. 5 shows the true and the predicted

trajectories starting from the initial condition $x_0 = \begin{bmatrix} 2 & -2 \end{bmatrix}^\top$. The predicted trajectory is denoted by $\hat{x}$. The results in Fig. 5 indicate that the CLDMD method, when coupled with direct prediction, fails to obtain accurate prediction of the closed-loop trajectories.
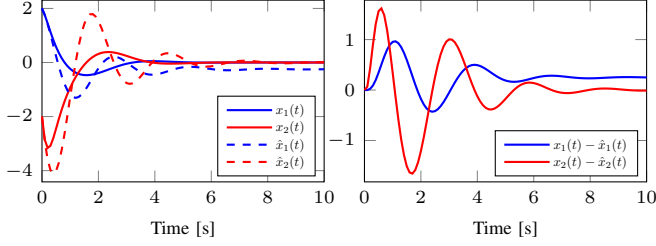

Fig. 5. Predicted and true trajectories (left) and the corresponding prediction errors (right) of the controlled duffing oscillator for the experiment in Section IX-A. This result is obtained using CLDMD direct prediction (32) with kernel parameter $\tilde{\rho} = 1e8$.

B. Two-link Robot Manipulator

This experiment concerns a planar two-link robot manipulator described by Euler-Lagrange dynamics

$$M(q)\ddot{q} + V_m(q,\dot{q})\dot{q} + F(\dot{q}) = \tau,$$

where $q = (q_1 \ q_2)^\top \in \mathbb{R}^2$ and $\dot{q} = (\dot{q}_1 \ \dot{q}_2)^\top$ are the angular positions (rad) and angular velocities ($\mathrm{rad\,s^{-1}}$) of the two links, respectively, $\tau = (\tau_1 \ \tau_2)^\top$ is the torque ($\mathrm{N\,m}$) produced by the motors that drive the joints, $M(q)$ is the inertia matrix, and $V_m(q,\dot{q})$ is the centripetal-Coriolis matrix, defined as

$$M(q) := \begin{bmatrix} p_1 + 2p_3 c_2(q) & p_2 + p_3 c_2(q) \\ p_2 + p_3 c_2(q) & p_2 \end{bmatrix},$$

and

$$V_m(q,\dot{q}) = \begin{bmatrix} p_3 s_2(q)\dot{q}_2 & -p_3 s_2(q)(\dot{q}_1 + \dot{q}_2) \\ p_3 s_2(q)\dot{q}_1 & 0 \end{bmatrix},$$

where $p_1 = 3.473\,\mathrm{kg\,m^2}$, $p_2 = 0.196\,\mathrm{kg\,m^2}$, $p_3 = 0.242\,\mathrm{kg\,m^2}$, $c_2(q) = \cos(q_2)$, $s_2(q) = \sin(q_2)$, and $F(\dot{q}) = \begin{bmatrix} f_{d1}\dot{q}_1 + f_{s1}\tanh(\dot{q}_1) & f_{d2}\dot{q}_2 + f_{s2}\tanh(\dot{q}_2) \end{bmatrix}^\top$ is the model for friction, where $f_{d1} = 5.3\,\mathrm{kg\,m^2\,s^{-1}}$, $f_{d2} = 1.1\,\mathrm{kg\,m^2\,s^{-1}}$, $f_{s1} = 8.45\,\mathrm{kg\,m^2\,s^{-1}}$, and $f_{s2} = 2.35\,\mathrm{kg\,m^2\,s^{-1}}$. The model can be expressed in the form $\dot{x} = f(x) + g(x)u$ with $x = \begin{bmatrix} q^\top & \dot{q}^\top \end{bmatrix}^\top$, $u = \tau$, $f(x) = \begin{bmatrix} \dot{q}^\top & (M^{-1}(q)(-V_m(q,\dot{q})\dot{q} + F(\dot{q})))^\top \end{bmatrix}$, and $g(x) = \begin{bmatrix} 0_{2\times2} & (M^{-1}(q))^\top \end{bmatrix}^\top$, where $0_{2\times2}$ denotes a 2×2 matrix of zeros.

A total of 200 open-loop trajectories of the manipulator are generated using the MATLAB® ode45 solver, starting from initial conditions selected to fill a hypercube of side 1, centered at the origin of the state space, $\mathbb{R}^4$, using a Halton sequence. The control signal used for trajectory generation is of the form $u = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^\top$ with $u_j(t) = \sum_{i=1}^{15} b_{j,i}\sin(\omega_{j,i}t + \varphi_{j,i})$, for $j = 1,2$, where the magnitudes $b_{j,i}$, the frequencies $\omega_{j,i}$, and the phase differences $\varphi_{j,i}$ are generated randomly from a uniform distribution on the interval $[-1,1]$. All trajectories are
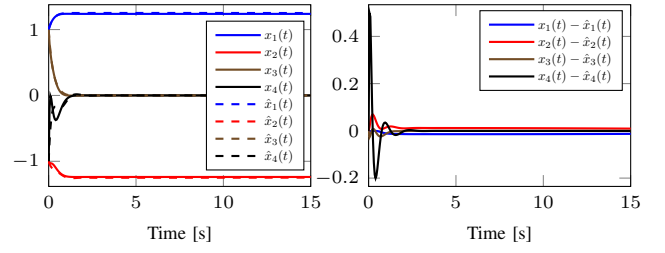

Fig. 6. Predicted and true trajectories (left) and the corresponding prediction errors (right) of the 2-link robot manipulator for the experiment in Section IX-B. This result is obtained using CLDMD direct prediction (32) with kernel parameter $\tilde{\rho} = 1e5$ and regularization parameter $\tilde{\epsilon} = \epsilon = 1e-7$.
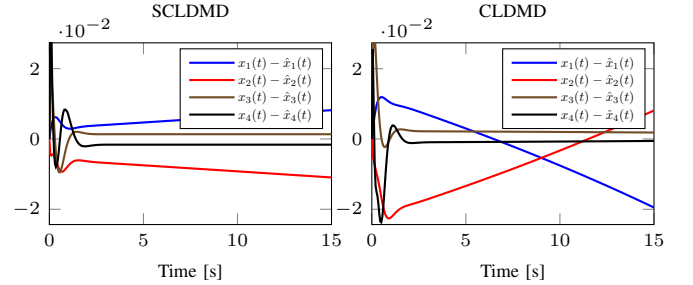

Fig. 7. Error between the predicted and the true trajectories of the 2-link manipulator for the experiment in Section IX-B. The figure on the left is obtained using SCLDMD indirect prediction by solving (25) and the figure on the right is obtained using CLDMD indirect prediction by solving (33), both using the MATLAB® ode45 solver.

recorded over a duration of 1s, and are sampled at a frequency of 10 Hz.

The trajectories are then utilized to predict the behavior of the oscillator under the state feedback controller $\mu(x) = \begin{bmatrix} -5 & -5 \\ -15 & -15 \end{bmatrix} x$, starting from $x_0 = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}^\top$. CLDMD is implemented using the exponential dot product reproducing kernel with parameter 10 and a diagonal kernel given by $K = \mathrm{diag}\begin{bmatrix} \tilde{K}_{10} & \tilde{K}_{10} & \tilde{K}_{10} \end{bmatrix}$. SCLDMD is implemented using $\tilde{K}_r = \tilde{K}_5$, $K = \mathrm{diag}\begin{bmatrix} \tilde{K}_{10} & \tilde{K}_{10} & \tilde{K}_{10} \end{bmatrix}$, and $\tilde{K}_d = \tilde{K}_{15}$. Gram matrices are regularized as described in Remark 2 using regularization coefficients $\epsilon = \tilde{\epsilon} = 1e-3$. Simpson's 1/3 rule is used to compute the integrals involved in algorithms 1 and 2. Since the vector field is now a function of 4 variables in each dimension, direct visualization of the true and approximate vector fields is not possible. However, the reconstruction accuracy may be indirectly gauged through indirect prediction of trajectories of the system.

Fig. 6 shows the true and the predicted trajectories using the direct reconstruction method, implemented using (32). The results in Fig. 6 indicate that the CLDMD method, when coupled with indirect prediction, is able to predict the desired closed-loop trajectory much better in this experiment than the Duffing oscillator experiment in Fig. 5. Fig. 7 shows the predicted trajectories and the prediction error resulting from the indirect approach. The results in Fig. 7 indicate that both the CLDMD and the SCLDMD methods, when coupled with indirect prediction, accurately predict the desired closed-loop trajectory.

X. Discussion

As is evident from Fig. 3, the methods developed in algorithms 1 and 2 can effectively utilize data collected under open-loop control signals to construct the closed-loop vector field under a given feedback policy.

Figures 5 and 4 indicate that the prediction error for the highly nonlinear controlled Duffing oscillator is significantly higher in direct prediction as compared to indirect prediction. We postulate that this is due to the linear nature of the model in (32). Since (32) is a solution of a system of linear ordinary differential equations, the resulting reconstruction diverges quickly from the trajectories of the nonlinear model. On the other hand, we postulate that due to the presence of the eigenfunctions in (33), the model used in the indirect approach includes nonlinear effects and as a result, generates a better prediction. When the nonlinearities in the original system are mild, like the trigonometric nonlinearities in the two link robot model, the predictions from the direct and the indirect method are close, as seen in figures 6 and 7.

The theory and the computations that support the developed algorithms require data-richness. In the convergence proofs, data-richness manifests as the density of the kernel differences, the occupation kernels, and the control occupation kernels in their respective RKHSs. In the computations, data-richness is required for the Gram matrices G_r and G_β of the occupation kernels and the control occupation kernels, respectively to be invertible. It is shown in [50] that for Gaussian radial basis function reproducing kernels, the rank of the occupation kernel Gram matrix G_r can be characterized using the so-called trajectory separation distance. Roughly, the trajectory separation distance is the largest radius q such that when all trajectories are inflated to tubes of radius q, the resulting tubes are disjoint. Since multiple, shorter trajectories, starting from initial conditions that are well-separated, would generally result in better separation distances, such a dataset would be preferred. However, if the trajectories are too short, then the matrix D of trajectory endpoint differences can reduce to a zero matrix, resulting in poor performance. Obtaining similar results for other reproducing kernels and for characterization of the rank of the control occupation kernel Gram matrix G_β is a topic for future research.

The condition number of G_r and G_β depends not only on the trajectories but also on the selected reproducing kernels. For example, the condition number of Gram matrices corresponding to Gaussian radial basis functions, given as $\tilde{K}(x, y) = \exp(-\frac{1}{\tilde{\rho}}\|x - y\|_2^2)$ for $\tilde{\rho} > 0$ is larger for larger $\tilde{\rho}$. However, large $\tilde{\rho}$ values correspond to faster convergence of interpolation problems within the RKHS of the kernel (cf. [51]). Data-richness conditions similar to the persistence of excitation (PE) condition in adaptive control that relate the trajectories and the kernels can potentially be formulated to ensure a well-conditioned G_r and G_β, however, such formulation is a topic for future research.

Numerical experiments indicate that while direct trajectory reconstruction can be poor for systems with severe nonlinearities, the developed techniques generate accurate estimates of the closed-loop vector field from data. Unlike traditional sys-

tem identification techniques, the algorithms developed in this paper do not require careful selection of basis functions. While the implementation can be done using any universal kernels, careful tuning of the kernel parameter is often necessary.

XI. Conclusion

In this paper, a novel operator-theoretic framework is developed for the study of controlled nonlinear systems. The framework utilizes RKHSs, where feedback-controlled nonlinear systems are expressed using a composition of infinite dimensional multiplication operator and an infinite dimensional control Liouville operator. A provably convergent finite-rank representation of the composition, that utilizes trajectories of a system, observed under open-loop control inputs u_j, is developed. Eigendecomposition and SVD of the finite-rank representation is utilized to predict the behavior of the system response to a query feedback controller, μ. The same dataset can be used to predict the system behavior in response to a multitude of query feedback controllers.

To the best of our knowledge, this paper, along with the conference paper [24], are the first to study spectral decomposition of continuous-time feedback-controlled nonlinear systems in a provably convergent manner. While this paper focuses solely on system identification, the uniform convergence guarantees established by Corollary 1, makes the developed modeling technique an attractive candidate for use in a variety of applications, including, but not limited to, data-driven control synthesis and data-driven analysis and validation of feedback controllers.

References

[1] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dyn.*, vol. 41, no. 1, pp. 309–325, 2005.

[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. A. Rosenfeld, R. Kamalapurkar, L. F. Gruss, and T. T. Johnson, "Dynamic mode decomposition for continuous time systems with the Liouville operator," *J. Nonlinear Sci.*, vol. 32, no. 1, pp. 1–30, 2022.

[4] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, 2nd ed. New York: Springer, 1994.

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

[6] M. O. Williams, C. W. Rowley, and I. G. Kevrekidis, "A kernel-based method for data-driven Koopman spectral analysis," *J. Comput. Dyn.*, vol. 2, no. 2, pp. 247–265, 2015.

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

[8] U. Vaidya, P. G. Mehta, and U. V. Shanbhag, "Nonlinear stabilization via control Lyapunov measure," *IEEE Trans. Autom. Control*, vol. 55, no. 6, pp. 1314–1328, 2010.

[9] 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, 2016.

[10] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz, "Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition," *J. Neurosci. Methods*, vol. 258, pp. 1–15, 2016.

[11] J. Mann and J. N. Kutz, "Dynamic mode decomposition for financial trading strategies," *Quant. Finance*, vol. 16, no. 11, pp. 1643–1655, 2016.

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

[13] A. Sootla, A. Mauroy, and D. Ernst, "Optimal control formulation of pulse-based control using Koopman operator," *Automatica*, vol. 91, pp. 217–224, 2018.

[14] 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.

[15] M. Quade, M. Abel, J. Nathan Kutz, and S. L. Brunton, "Sparse identification of nonlinear dynamics for rapid model recovery," *Chaos*, vol. 28, no. 6, p. 063116, 2018.

[16] S. Sinha, B. Huang, and U. Vaidya, "On robust computation of Koopman operator and prediction in random dynamical systems," *J. Nonlinear Sci.*, 2019.

[17] H. Arbabi, M. Korda, and I. Mezić, "A data-driven Koopman model predictive control framework for nonlinear partial differential equations," in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 6409–6414.

[18] B. Jayaraman, C. Lu, J. Whitman, and G. Chowdhary, "Sparse feature map-based Markov models for nonlinear fluid flows," *Comput. Fluids*, vol. 191, p. 104252, 2019.

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

[20] 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.

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

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

[23] J. A. Rosenfeld, B. Russo, R. Kamalapurkar, and T. Johnson, "The occupation kernel method for nonlinear system identification," *SIAM J. Control Optim.*, to appear, see arXiv:1909.11792.

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

[25] G. Pedrick, "Theory of reproducing kernels for Hilbert spaces of vector valued functions," Ph.D. dissertation, University of Kansas, 1957.

[26] L. Schwartz, "Sous-espaces Hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants)," *J. Anal. Math.*, vol. 13, pp. 115–256, 1964.

[27] C. A. Micchelli and M. Pontil, "On learning vector-valued functions," *Neural Comput.*, vol. 17, no. 1, pp. 177–204, 2005.

[28] C. Carmeli, E. De Vito, and A. Toigo, "Vector valued reproducing kernel Hilbert spaces of integrable functions and Mercer theorem," *Anal. Appl.*, vol. 4, no. 04, pp. 377–408, 2006.

[29] 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.

[30] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*. Addison-Wesley, 2002.

[31] R. Ortega, A. Loría, P. J. Nicklasson, and H. J. Sira-Ramirez, *Passivity-based control of Euler-Lagrange systems: mechanical, electrical and electromechanical applications*. Springer, 1998.

[32] F. Morabito, A. R. Teel, and L. Zaccarian, "Nonlinear antiwindup applied to Euler-Lagrange systems," *IEEE Trans. Robot. Autom.*, vol. 20, no. 3, pp. 526–537, 2004.

[33] Z. Feng, G. Hu, W. Ren, W. E. Dixon, and J. Mei, "Distributed coordination of multiple unknown Euler-Lagrange systems," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 1, pp. 55–66, 2018.

[34] R. Kamalapurkar, P. Walters, J. A. Rosenfeld, and W. E. Dixon, *Reinforcement learning for optimal feedback control: A Lyapunov-based approach*, ser. Communications and Control Engineering. Springer International Publishing, 2018.

[35] A. Behal, W. E. Dixon, B. Xian, and D. M. Dawson, *Lyapunov-based control of robotic systems*. Taylor and Francis, 2009.

[36] J. A. Rosenfeld and R. Kamalapurkar, "Singular dynamic mode decomposition," *SIAM J. Appl. Dyn. Syst.*, vol. 22, no. 3, pp. 2357–2381, 2023.

[37] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Sparse identification of nonlinear dynamics with control (SINDYc)," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 710–715, 2016.

[38] 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.

[39] E. Kaiser, J. N. Kutz, and S. L. Brunton, "Data-driven discovery of Koopman eigenfunctions for control," *Machine Learning: Science and Technology*, vol. 2, no. 3, p. 035023, 2021.

[40] C. Folkestad, D. Pastor, I. Mezic, R. Mohr, M. Fonoberova, and J. Burdick, "Extended dynamic mode decomposition with learned Koopman eigenfunctions for prediction and control," in *Proc. Am. Control Conf.* IEEE, 2020, pp. 3906–3913.

[41] S. Peitz, S. E. Otto, and C. W. Rowley, "Data-driven model predictive control using interpolated Koopman generators," *SIAM J. Appl. Dyn. Sys.*, vol. 19, no. 3, pp. 2162–2193, 2020.

[42] S. E. Otto and C. W. Rowley, "Koopman operators for estimation and control of dynamical systems," *Annu. Rev. Control Robot. Auton. Sys.*, vol. 4, pp. 59–87, 2021.

[43] X. Zhang, W. Pan, R. Scattolini, S. Yu, and X. Xu, "Robust tube-based model predictive control with Koopman operators," *Automatica*, vol. 137, p. 110114, 2022.

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

[45] J. A. Rosenfeld, "Densely defined multiplication on several Sobolev spaces of a single variable," *Complex Anal. Oper. Theory*, vol. 9, no. 6, pp. 1303–1309, 2015.

[46] ——, "Introducing the polylogarithmic Hardy space," *Integral Equ. Oper. Theory*, vol. 83, no. 4, pp. 589–600, 2015.

[47] F. H. Szafraniec, "The reproducing kernel Hilbert space and its multiplication operators," in *Complex Analysis and Related Topics*. Springer, 2000, pp. 253–263.

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

[49] N. Aronszajn, "Theory of reproducing kernels," *Trans. Am. Math. Soc.*, vol. 68, pp. 337–404, 1950.

[50] B. P. Russo, R. Kamalapurkar, D. Chang, and J. A. Rosenfeld, "Motion tomography via occupation kernels," *J. Comput. Dyn.*, vol. 9, no. 1, pp. 27–45, 2022.

[51] G. E. Fasshauer, *Meshfree approximation methods with MATLAB*, ser. Interdisciplinary Mathematical Sciences. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007, vol. 6.

Joel A. Rosenfeld is an Associate Professor in the Department of Mathematics and Statistics at the University of South Florida and the Principal Investigator of the Learning DOCK Group. He received his Ph.D. from the mathematics department at the University of Florida in 2013 under the advisement of Dr. Michael T. Jury studying Operator Theory and Functional Analysis. Following graduate school, he was a postdoc in Mechanical Engineering at the University of Florida under Dr. Warren E. Dixon studying numerical methods in optimal control theory and fractional calculus. Subsequently, he joined the Department of Electrical Engineering and Computer Science at Vanderbilt University under Dr. Taylor T. Johnson studying numerical methods in formal methods for computing, which ultimately led to a position as a Senior Research Scientist Engineer within the Institute for Software Integrated Systems. He is a 2020 Young Investigator Research Program awardee through the Air Force Office of Scientific Research.

Rushikesh Kamalapurkar received his M.S. and his Ph.D. degree in 2011 and 2014, respectively, from the Department of Mechanical and Aerospace Engineering at the University of Florida. After working for a year as a postdoctoral researcher with Dr. Warren E. Dixon, he was appointed as the 2015-16 MAE postdoctoral teaching fellow. From 2016 to 2024 he served on the faculty of of Mechanical and Aerospace Engineering at the Oklahoma State University. In 2024, he joined the faculty of the Department of Mechanical and Aerospace Engineering at the University of Florida where he he directs the Systems, Cognition, and Control laboratory. His primary research interest is intelligent, learning-based optimal control of uncertain nonlinear dynamical systems.