Dynamic Mode Decomposition with Control Liouville Operators ⋆

Joel A. Rosenfeld * **Rushikesh Kamalapurkar** **

* *Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620 USA (e-mail: rosenfeldj@usf.edu).*

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

Abstract: This manuscript provides a theoretical foundation for the Dynamic Mode Decomposition (DMD) of control affine dynamical systems through vector valued reproducing kernel Hilbert spaces (RKHSs). Specifically, control Liouville operators and control occupation kernels are introduced to separate the drift dynamics from the control effectiveness components. Given a known feedback controller that is represented through a multiplication operator, a DMD analysis may be performed on the composition of these operators to make predictions concerning the system controlled by the feedback controller.

Keywords: system identification, spectral analysis, operators, model approximation, control system analysis

1. INTRODUCTION

Dynamic Mode Decomposition (DMD) is a method of system identification that casts unknown discrete or continuous time dynamics over finite dimensions into a linear operator over an infinite dimensional space (cf. [1]). DMD collects trajectory data from observations, or *snapshots*, of a dynamical system, and the method constructs a finite rank representation of an operator that describes the evolution of the state. In the discrete time case, this linear operator is a composition operator called the Koopman operator [2], popularized by [3]. The finite rank representation of the dynamics is then diagonalized, and the resultant eigenfunction and eigenvalues are used to provide a representation of the identity function, which in turn provides the dynamic modes via vector valued coefficients attached to the eigenfunctions. Thereafter, a state trajectory can be represented as a sum of exponential functions multiplied by the dynamic modes (cf. [4,1,5]).

The primary application area of Koopman spectral analysis of dynamical systems has been fluid dynamics, where DMD is compared with proper orthogonal decompositions (POD) for nonlinear fluid equations (cf. [6]). DMD has also been employed in the study of stability properties of dynamical systems [7,8], neuroscience [9], financial trading [10], feedback stabilization [11], optimal control [12], modeling of dynamical systems [13–15], and model-predictive control [16]. For a generalized treatment of DMD as a Markov model, see [17].

In each of the above examples, continuous time dynamics are studied using DMD via the discrete time dynamics determined through a fixed time-step. Recently, [18] imported the notion of occupation measures from the Banach spaces of continuous function to that of reproducing kernel Hilbert spaces (RKHSs). It was subsequently demonstrated in [5] that when combined with the Liouville operator (also known as the Koopman generator), the so-called occupation kernels provide a method for performing a DMD analysis on the continuous time dynamic directly.

The paradigm shift afforded by occupation kernels arises through the consideration of the state trajectory as the fundamental unit of data, as was done in [18]. Computationally, the DMD analysis of the Liouville operator is remarkably unchanged to that of the Koopman operator. The only adjustment is in the computation of the Gram matrix of occupation kernels, which requires the computation of double integrals for each entry as the inner product of the occupation kernels with each other.

This manuscript builds on the concept of Liouville operators and occupation kernels through the incorporation of control-affine dynamics, where the trajectory data given by the system is accompanied by a collection of control signals. In this context, occupation kernels are augmented by the control signals resulting in what are called control occupation kernels, and the Liouville operator now includes the control effectiveness matrix valued function. The augmentations are possible through vector valued RKHSs, which have been extensively studied in [19] and [20].

⋆ This research was supported in part by the Air Force Research Laboratory (FA8651-19-2-0009). Any opinions, findings, or conclusions in this manuscript are those of the author(s) and do not necessarily reflect the views of the sponsoring agencies.

2. VECTOR VALUED REPRODUCING KERNEL HILBERT SPACES

This section reviews important properties of vector valued RKHSs, and relies heavily on the discussion given in [20].

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 (vector valued) reproducing kernel Hilbert space (RKHS) if for every $v \in \mathcal{Y}$ and $x_0 \in X$, the functional $f \mapsto \langle f(x_0), v\rangle_{\mathcal{Y}}$ is bounded.

A vector valued RKHS is a direct generalization of a "scalar" valued RKHS, where for a fixed $v \in \mathcal{Y}$, the collection of functions $\{g(x) = \langle f(x), v\rangle_{\mathcal{Y}} : f \in H\}$ forms a RKHS of scalar valued functions. Moreover, for a basis of $\mathcal{Y}$, $\{v_1, v_2, \ldots\} \subset \mathcal{Y}$, the values of functions in H may be recovered via their individual scalar valued projections.

Through the Riesz representation theorem, for each $x_0 \in X$ and $v \in \mathcal{Y}$, there is a function $K_{x_0,v} \in H$ such that $\langle f, K_{x_0,v}\rangle_H = \langle f(x_0), v\rangle_{\mathcal{Y}}$ for all $f \in H$. It can be seen that the mapping $v \mapsto K_{x_0,v}$ is linear over $\mathcal{Y}$. Hence, $K_{x_0} v := K_{x_0,v}$ where $K_{x_0} : \mathcal{Y} \to H$ is a linear operator. The function K_{x_0} is called the kernel centered at x_0 and is an operator from $\mathcal{Y}$ to H.

The kernel operator has a particular representation as a sum of rank one tensors of evaluations of an orthonormal basis for H, $\{e_m\}_{m=1}^{\infty}$, where

$$K_{x_0,v}(x) = \sum_{m=1}^{\infty} e_m(x)\langle K_{x_0} v, e_m\rangle_H$$
$$= \sum_{m=1}^{\infty} e_m(x)\langle e_m(x_0), v\rangle_{\mathcal{Y}} = \left(\sum_{m=1}^{\infty} e_m(x) \otimes e_m(x_0)\right) v.$$

Hence, the *kernel operator* associated with H, $K : X \times X \to \mathcal{L}(\mathcal{Y}, \mathcal{Y})$, is defined as $K(x, x_0) := \sum_{m=1}^{\infty} e_m(x) \otimes e_m(x_0)$, where e_m is an orthonormal basis for H. In the particular case that $\mathcal{Y} = \mathbb{R}^n$, $K(x, x_0)$ is a real valued $n \times n$ matrix for fixed $x, x_0 \in X$. This suggests several examples of vector valued kernels, and indeed, for any positive definite matrix, A, and scalar valued kernel $\tilde{K}$, the kernel given as $K(x, x_0) = A\tilde{K}(x, x_0)$ is a kernel operator.

Just as will scalar valued kernels, the span of the set $E := \{K_{x,v} : v \in \mathcal{Y} \text{ and } x \in X\}$ is dense in H.

Proposition 1. Let E be given as above, then $(E^{\perp})^{\perp} = \overline{\operatorname{span} E} = H$.

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 $(E^{\perp})^{\perp} = H$. □

Hence, given $\epsilon > 0$ and $h \in H$, there is a linear combination of vector valued kernels that approximate h to within ϵ.

Remark 1. The implementation of vector valued RKHSs leveraged in this manuscript uses the vector space $\mathbb{R}^{1\times(1+m)}$ of row vectors, which arises from the gradient of observables. Thus, the linear operation of K_{x_0} on $v \in \mathbb{R}^{1\times(1+m)}$ will be expressed as $K_{x_0,v} = vK_{x_0}$ henceforth.

3. PROBLEM STATEMENT

The objective of this paper is to provide an operator theoretic approach for the analysis of a closed loop nonlinear control affine systems,

$$\dot{x} = f(x) + g(x)\mu(x), \tag{1}$$

with the state feedback controller $\mu : \mathbb{R}^n \to \mathbb{R}^m$ by using data provided by observed absolutely continuous controlled trajectories, $\{\gamma_{u_i} : [0, T] \to \mathbb{R}^n\}_{i=1}^{M}$, and corresponding bounded measurable (with respect to Lebesgue measure) control inputs, $\{u_i : [0, T] \to \mathbb{R}^m\}_{i=1}^{M}$, satisfying

$$\dot{\gamma}_{u_i}(t) = f(\gamma_{u_i}(t)) + g(\gamma_{u_i}(t))u_i(t)$$

in the Carathéodory sense. The observed control trajectories and control inputs will allow for the construction of a finite rank representation for a generalized Liouville operator, called the control Liouville operator, that is similar to that presented in [5]. Note that the terminal time, T, may vary between trajectories with little change to the implemenation, but it is kept constant for the sake of clarity of presentation.

The resolution of the this problem provided by this manuscript will yield a DMD analysis of the closed loop system via a composition of operators that separate the dynamics learning through observations and the known feedback controller.

Any Euler-Lagrange system with an invertible inertia matrix can be expressed in the control-affine form. The Euler-Lagrange equations are used to describe a large class of physical systems (cf. [21]), and as such, various methods for control and identification of nonlinear systems in the Euler-Lagrange form have been studied in detail over the years (see, e.g., [22–24]). Since most physical systems of practical importance; such as robot manipulators [25] and ground, air, and maritime vehicles and vessels [26]; admit invertible inertia matrices over large operating regions; control-affine models encompass a large class of physical systems.

4. CONTROL LIOUVILLE OPERATORS AND CONTROL OCCUPATION KERNEL OPERATORS

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be the drift dynamics and the control effectiveness matrix, respectively, for a control affine dynamical system. Let $\tilde{H}$ be a scalar valued RKHS of continuously differentiable functions over $\mathbb{R}^n$ and let H be a $\mathbb{R}^{m+1}$ (row) vector valued RKHS of continuous functions over $\mathbb{R}^n$. Let $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^{1\times(m+1)}, \mathbb{R}^{1\times(m+1)})$ and $\tilde{K} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the kernel functions of H and $\tilde{H}$, respectively.

Definition 2. Let the set

$$\mathcal{D}(A_{f,g}) := \{h \in \tilde{H} : \nabla h(x) \left(f(x)\ \ g(x)\right) \in H\}$$

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

$$A_{f,g}h(x) = \nabla h(x) \left(f(x)\ \ g(x)\right).$$

The operator $A_{f,g}$ is called the control Liouville operator corresponding to f and g over H.

Control Liouville operators are a direct generalization of more traditional Liouville operators, where the drift dynamics and control effectiveness components of the dynamics are separated on the operator theoretic level. It can be seen here that vector valued RKHSs arise naturally in the context, where the gradient of $h \in \mathcal{D}(A_{f,g})$ is a row vector, 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 an $m + 1$ dimension row vector.

For the purposes of this manuscript, control Liouville operators are assumed to be densely defined over their respective Hilbert spaces, and it follows from the methods of [18] that control Liouville operators are closed. Hence, the adjoint of control Liouville operators are also densely defined [27].

Given a bounded measurable control input, $u : [0,T] \to \mathbb{R}^m$, and an absolutely continuous trajectory, $\gamma : [0,T] \to \mathbb{R}^n$, the functional, $T : H \to \mathbb{R}$, given as

$$Tp = \int_0^T p(\gamma(t)) \begin{pmatrix} 1 \\ u(t) \end{pmatrix} dt$$

is bounded. Hence, there is a function $\Gamma_{\gamma,u} \in H$ such that $Tp = \langle p, \Gamma_{\gamma,u} \rangle_H$ for all $p \in H$.

Definition 3. For a bounded measurable control input u and absolutely continuous trajectory γ, the function $\Gamma_{\gamma,u} \in H$ is called the control occupation kernel corresponding to u and γ in H.

Proposition 2. The occupation kernel corresponding to u and γ (given above) may be expressed as

$$\Gamma_{\gamma,u}(x) = \int_0^T \begin{pmatrix} 1 & u(t)^T \end{pmatrix} K(x, \gamma(t)) dt, \qquad (2)$$

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

$$\|\Gamma_{\gamma,u}\|_2^2 = \int_0^T \int_0^T \begin{pmatrix} 1 & u(t)^T \end{pmatrix} K(\gamma(\tau), \gamma(t)) \begin{pmatrix} 1 \\ u(\tau) \end{pmatrix} dt d\tau.$$

Proof. Consider for $x \in X$ and $v \in \mathbb{R}^{1\times(m+1)}$,

$$\begin{aligned}\langle \Gamma_{\gamma,u}(x), v \rangle_{\mathbb{R}^{1\times(m+1)}} &= \langle \Gamma_{\gamma,u}, vK_x \rangle_H = \langle vK_x, \Gamma_{\gamma,u} \rangle_H \\ &= \int_0^T vK(x, \gamma(t)) \begin{pmatrix} 1 \\ u(t) \end{pmatrix} dt. \qquad (3)\end{aligned}$$

As (4) holds for all $v \in \mathbb{R}^{1\times(m+1)}$, (2) follows. The norm of $\Gamma_{\gamma,u}$ follows from $\|\Gamma_{\gamma,u}\|^2 = \langle \Gamma_{\gamma,u}, \Gamma_{\gamma,u} \rangle_H$, and the defining properties of $\Gamma_{\gamma,u}$. □

As with Liouville operators over RKHSs corresponding to nonlinear dynamical systems, there is a direction connection between the adjoints of control Liouville operators and control occupation kernels that correspond to admissible control signals, u, and their corresponding controlled trajectories, γ_u, that satisfy (1).

Proposition 3. Suppose that f and g correspond to a control Liouville operator, $A_{f,g} : \mathcal{D}(A_{f,g}) \to H$, and let u be an admissible control signal for the control affine dynamical system in (1) with a corresponding controlled trajectory, γ_u. Then, $\Gamma_{\gamma_u,u} \in \mathcal{D}(A^*_{f,g})$ and $A^*_{f,g}\Gamma_{\gamma_u,u} = \tilde{K}(\cdot, \gamma_u(T)) - \tilde{K}(\cdot, \gamma_u(0))$.

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

$$\begin{aligned}&\langle A_{f,g}h, \Gamma_{\gamma_u,u} \rangle_H \\ &= \int_0^T \nabla h(\gamma_u(t)) \left(f(\gamma_u(t)) \; g(\gamma_u(t)) \right) \begin{pmatrix} 1 \\ u(t) \end{pmatrix} dt \\ &= \int_0^T \frac{d}{dt} h(\gamma_u(t)) dt = h(\gamma_u(T)) - h(\gamma_u(0)) \\ &= \langle h, \tilde{K}(\cdot, \gamma_u(T)) - \tilde{K}(\cdot, \gamma_u(0)) \rangle_{\tilde{H}}.\end{aligned}$$

Hence, the functional $h \mapsto \langle A_{f,g}h, \Gamma_{\gamma_u,u} \rangle_H$ is bounded with norm not exceeding $\|\tilde{K}(\cdot, \gamma_u(T)) - \tilde{K}(\cdot, \gamma_u(0))\|_{\tilde{H}}$. Moreover, the above equations yield

$$A^*_{f,g}\Gamma_{\gamma_u,u} = \tilde{K}(\cdot, \gamma_u(T)) - \tilde{K}(\cdot, \gamma_u(0)).$$

□

5. MULTIPLICATION OPERATORS FROM VECTOR VALUED TO SCALAR VALUED REPRODUCING KERNEL HILBERT SPACES

The inclusion of a second operator in addition to the control Liouville operator sets the theoretical foundations of DMD for control-affine systems apart from the uncontrolled case in [5]. As will be seen in subsequent sections, after the determination of a representation of $A_{f,g}$, the inclusion of a state feedback controller is necessary for a DMD procedure. The state feedback controller is implemented via a multiplication operator. This section established some results on multiplication operators from vector valued RKHSs to scalar valued RKHSs. Many of these theorems have been established for scalar valued RKHSs (cf. [28–30]), and the results follows from similar methods.

Let $\mu : X \to \mathcal{Y}$ be a vector valued function, and define the multiplication operator, $M_\mu : \mathcal{D}(M_\mu) \to \tilde{H}$, with symbol μ as $M_\mu h = \langle h, \mu \rangle_{\mathcal{Y}}$, where $\mathcal{D}(M_\mu) := \{h \in H : \langle h, \mu \rangle_{\mathcal{Y}} \in \tilde{H}\}$ is a collection of (row) vector valued functions in H and $\tilde{H}$ is a scalar valued RKHS.

Proposition 4. The function $\tilde{K}(\cdot, x)$ is in the domain of M^*_μ, $\mathcal{D}(M^*_\mu)$. Moreover, $M^*_\mu \tilde{K}(\cdot, x) = K_{x,\mu(x)}$

Proof. Let $h \in \mathcal{D}(M_\mu)$, then

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

Hence, the mapping $h \mapsto \langle M_\mu h, \tilde{K}(\cdot, x) \rangle_{\tilde{H}}$ is a bounded functional with norm bounded by $\|K_{x,\mu(x)}\|_H$, and $\tilde{K}(\cdot, x)$ is in the domain. Moreover, this demonstrates that

$$\langle M_\mu h, \tilde{K}(\cdot, x) \rangle_{\tilde{H}} = \langle h, K_{x,\mu(x)} \rangle_H,$$

which establishes the adjoint formula. □

Thus, the adjoint of a multiplication operator connects occupation kernels with control occupation kernels.

Proposition 5. Multiplication operators are closed operators.

Proof. Suppose that $\{h_n\} \in \mathcal{D}(M_\mu)$, $h_n \to h \in H$, and $M_\mu h_n \to W \in H$. To show that M_μ is a closed operator, it must be shown that $W(x) = \langle h(x), \mu(x) \rangle_{\mathcal{Y}}$ for all $x \in X$, and thus $h \in \mathcal{D}(M_\mu)$ by definition and $M_\mu h = W$. Let $v \in \mathcal{Y}$ and $x \in X$, then

$$\begin{aligned}W(x) &= \lim_{n\to\infty} \langle M_\mu h_n, \tilde{K}(\cdot, x) \rangle_H \\ &= \lim_{n\to\infty} \langle h_n, K_{x,\mu(x)} \rangle_H = \langle h, K_{x,\mu(x)} \rangle_H \\ &= \langle h(x), \mu(x) \rangle_{\mathcal{Y}},\end{aligned}$$

where the convergence follows as norm convergence implies weak convergence. □

Hence, the adjoint of a multiplication operator is densely defined by [27, Proposition 5.1.7].

Let for a trajectory $\gamma : [0,T] \to X$, let Γ_γ be the occupation kernel corresponding to γ within $\tilde{H}$ (cf. [18]).

Proposition 6. The function Γ_γ is in the domain of M_μ, $\mathcal{D}(M_\mu)$. Moreover, $M_\mu^* \Gamma_\gamma = \Gamma_{\gamma,\mu(\gamma(\cdot))}$.

Proof. This follows from a similar method as Proposition 4. □

6. DYNAMIC MODES FOR CLOSED LOOP CONTROL SYSTEMS

Dynamic mode decomposition aims to determine eigenfunctions and eigenvalues for the Liouville or Koopman operators that correspond to a particular dynamical system. For control Liouville operators, the dynamics given in terms of f and g correspond to an operator whose range is a vector valued RKHS and domain a scalar valued RKHS. As the domain and range are different vector spaces, an eigendecomposition of the control Liouville operator does not make sense. Indeed, for the dynamical system itself, a control input is required in addition to f and g to simulate the system.

After a finite rank representation of a control Liouville operator is determined, an additional operator is required before a dynamic mode decomposition may be determined. In particular, given a feedback controller, $\mu : \mathbb{R}^n \to \mathbb{R}^m$, the corresponding multiplication operator $M_\mu : \mathcal{D}(M_\mu) \to \tilde{H}$ is given as $M_\mu h := h \begin{pmatrix} 1 \\ \mu \end{pmatrix}$ and $\mathcal{D}(M_\mu) := \{h \in H : h \begin{pmatrix} 1 & \mu^T \end{pmatrix}^T \in \tilde{H}\}$. What differs from the previous controllers that were implemented via the control occupation kernels is that μ is a function of the state variable, $x \in \mathbb{R}^n$, rather than a function of time. It will be assumed that the image of $A_{f,g}$ falls within the domain of M_μ.

Note that $M_\mu A_{f,g} : \mathcal{D}(A_{f,g}) \to \tilde{H}$, that is the composition of M_μ with $A_{f,g}$ is a mapping with domain and range within $\tilde{H}$. If ϕ is an eigenfunction of $M_\mu A_{f,g}$ with eigenvalue λ, then if γ_μ is a controlled trajectory arising from (1) it follows that

$$\begin{aligned}\dot{\phi}(\gamma_\mu(t)) &= \nabla\phi(\gamma_\mu(t))(f(\gamma_\mu(t)) + g(\gamma_\mu(t))\mu(\gamma_\mu(t))) \\ &= M_\mu A_{f,g}\phi(\gamma_\mu(t)) = \lambda\phi(\gamma_\mu(t)).\end{aligned}$$

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

Thus, the identity function, $g_{id} : \mathbb{R}^n \to \mathbb{R}^n$ given as $g_{id}(x) := x$, may be decomposed using the eigenbasis for $M_\mu A_{f,g}$, denoted as ϕ_i with eigenvalue λ_i, as $g_{id}(x) = \sum_{i=1}^\infty \xi_i \phi_i(x)$, where $\xi_i \in \mathbb{R}^n$ are the dynamic modes of the closed loop system. Moreover, it follows that

$$\gamma_\mu(t) = g_{id}(\gamma_\mu(t)) = \sum_{i=1}^\infty \xi_i \phi_i(\gamma_\mu(0)) e^{\lambda_i t}.$$

The above construction makes several assumptions about the composition of M_μ with $A_{f,g}$ that are common in the study of DMD (cf. [5,4]). It assumes that there is a complete eigenbasis of $M_\mu A_{f,g}$ that may be accessed through finite rank representations of $M_\mu A_{f,g}$. Moreover, it assumes that for each $i = 1, \ldots, n$ that the mapping $x \mapsto x_i$ is within $\tilde{H}$ or at least it may be approximated by functions in $\tilde{H}$. The validity of each of these assumption depends on both the dynamics and the respective RKHSs, and may not always hold. Thus, DMD gives a data-driven heuristic and operator motivated methods for the analysis of a dynamical system.

7. FINITE RANK REPRESENTATIONS

As M_μ and $A_{f,g}$ are modally unbounded operators, the use of data to give empirical representations of these operators are not expected to give approximations M_μ and $A_{f,g}$ under the operator norm. However, the only data available concerning these operators is given through the trajectories themselves, thus DMD produces a finite rank representation under a data driven heuristic. The kernel based extended DMD procedure was developed in [4], where kernel functions were seen as a means of computing inner products of feature maps in $\ell^2(\mathbb{N})$. The infinite dimensional representation selected in [4] was with respect to the feature basis itself. As the feature space is infinite in size the final matrix for the Koopman operator's representation is never fully computed, and the work perfomed is executed on a finite sized proxy. The use of a proxy representation is also leveraged in [5] for continuous time DMD.

However, as dynamic modes may only be extracted from the product of M_μ with $A_{f,g}$, it is necessary to give an explicit finite rank representation of $A_{f,g}$ and M_μ to determine the dynamic modes of the resultant system. It should be noted that in the above development the feature space has not been invoked. Indeed, all of the development thus far occurs at the level of RKHSs.

The following subsections give two approaches to generate finite rank representations using basis functions within the RKHS, $\tilde{H}$. The results will yield finite rank representations of $M_\mu A_{f,g}$ as a product of four matrices. The SVD for a product of matrices may be computed using iterative methods such as [31], to complete the DMD computations.

In what follows, finite collections of linearly independent vectors, α and β, are selected for establishing finite rank matrices over a domain, $\operatorname{span}\alpha$, and range, $\operatorname{span}\beta$. For an operator T, the notation $[T]_\alpha^\beta$ is to be read as the matrix representation of $P_\beta T$ with respect to the domain, $\operatorname{span}\alpha$, and range $\operatorname{span}\beta$. Here P_β is the projection onto $\operatorname{span}\beta$.

7.1 A Kernel Basis

As kernels are dense in both the vector valued and scalar valued cases, finite rank representation of each operator is possible with respect to the kernels themselves. As only a finite number of kernels may be leveraged in any computational procedure, there is an element of imprecision in the representation of the functions in the domain and range of these operators. However, as kernels are dense in their respective RKHSs, the use of an increasing number of kernels reduces the overall error of the estimations.

The objective of this section is to select a collection of centers, $\{c_i\}_{i=1}^{\tilde{N}} \subset \Omega$, from a compact set, $\Omega \subset \mathbb{R}^n$, containing the sampled trajectories of the system, and to use the collection of corresponding kernel functions, $\alpha = \{\tilde{K}(\cdot, c_i)\}_{i=1}^{\tilde{N}} \subset \tilde{H}$, as a basis for the matrix representation of $M_\mu A_{f,g}$, or indirectly of $A_{f,g}^* M_\mu^*$. The control occupation kernels corresponding to the sampled trajectories will be denoted as $\beta = \{\Gamma_{\gamma_{u_i}, u_i}\}_{i=1}^M$.

To form the various required matrices, projections will be computed with respect to the individual bases. If $\tilde{h} \in \tilde{H}$, then the coefficients mapping $\tilde{h}$ to its projection onto $\text{span}\{\tilde{K}(\cdot, c_i)\}_{i=1}^{\tilde{N}}$ are given as the solution to the following system:

$$\begin{pmatrix} \tilde{K}(c_1, c_1) & \cdots & \tilde{K}(c_1, c_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ \tilde{K}(c_{\tilde{N}}, c_1) & \cdots & \tilde{K}(c_{\tilde{N}}, c_{\tilde{N}}) \end{pmatrix} \begin{pmatrix} \tilde{w}_1 \\ \vdots \\ \tilde{w}_{\tilde{N}} \end{pmatrix} = \begin{pmatrix} \tilde{h}(c_1) \\ \vdots \\ \tilde{h}(c_{\tilde{N}}) \end{pmatrix},$$

which is solvable since $G_1 = (\tilde{K}(c_i, c_j))_{i,j=1}^{\tilde{N}}$, (i.e., the Gram matrix) is positive definite for many choices of kernel functions, and is always positive semidefinite. The projection onto the control occupation kernels may be computed similarly when $h \in H$ by solving,

$$\begin{pmatrix} \langle \Gamma_{\gamma_{u_1}, u_1}, \Gamma_{\gamma_{u_1}, u_1} \rangle_H & \cdots & \langle \Gamma_{\gamma_{u_1}, u_1}, \Gamma_{\gamma_{u_M}, u_M} \rangle_H \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_{u_M}, u_M}, \Gamma_{\gamma_{u_1}, u_1} \rangle_H & \cdots & \langle \Gamma_{\gamma_{u_M}, u_M}, \Gamma_{\gamma_{u_M}, u_M} \rangle_H \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_{\tilde{N}} \end{pmatrix} = \begin{pmatrix} \int_0^T h(\gamma_{u_1}(t)) \begin{pmatrix} 1 \\ u_1(t) \end{pmatrix} dt \\ \vdots \\ \int_0^T h(\gamma_{u_M}(t)) \begin{pmatrix} 1 \\ u_M(t) \end{pmatrix} dt \end{pmatrix},$$

where each element of $G_2 = (\langle \Gamma_{\gamma_{u_i}, u_i}, \Gamma_{\gamma_{u_j}, u_j} \rangle_H)_{i,j=1}^M$ may be computed through a double integral.

Each kernel function, $\tilde{K}(\cdot, c_i)$, in $\tilde{H}$ satisfies $M_\mu^* \tilde{K}(\cdot, c_i) = K_{c_i, \mu(c_i)}$. As only the action of the control Liouville operator on the occupation kernels is known, computation of $A_{f,g}^* M_\mu^* \tilde{K}(\cdot, c_i)$ requires the result to be projected onto the control occupation kernel basis. Hence, the finite rank representation of M_μ with domain basis α and range basis β is given as,

$$[M_\mu^*]_\alpha^\beta = G_2^{-1} \left(\int_0^T K_{c_i, \mu(c_i)}(\gamma_{u_j}(t)) \begin{pmatrix} 1 \\ u_j(t) \end{pmatrix} dt \right)_{i=1,j=1}^{i=\tilde{N}, j=M}.$$

Similarly, the operation of $A_{f,g}^*$ on the basis β will be projected onto α. The resulting finite rank representation of $A_{f,g}^*$ is then given as

$$[A_{f,g}^*]_\beta^\alpha = G_1^{-1} \left(\tilde{K}(c_j, \gamma_{u_i}(T)) - \tilde{K}(c_j, \gamma_{u_i}(0)) \right)_{i=1,j=1}^{i=M, j=\tilde{N}}.$$

Hence, the matrix representation of the product of M_μ and $A_{f,g}$ is given as

$$[M_\mu A_{f,g}]_\alpha^\alpha \approx ([A_{f,g}^*]_\beta^\alpha [M_\mu^*]_\alpha^\beta)^T$$

and is a product of four explicit matrices.

7.2 A data-centric basis

The disadvantage of the kernel basis is that the centers are selected independently of the data. A large number of centers may be selected, however, the rank of the resultant representation is bottlenecked by the control occupation kernels representing the trajectories and controllers. A data-centric selection of basis functions can be made by collecting the (scalar valued) occupation kernels representing the trajectories in $\tilde{H}$, $\delta := \{\Gamma_{\gamma_{u_i}}\}_{i=1}^M \subset \tilde{H}$.

To compute the projection of a function $\tilde{h} \in \tilde{H}$ with respect to the δ basis, the following system may be solved

$$\begin{pmatrix} \langle \Gamma_{\gamma_{u_1}}, \Gamma_{\gamma_{u_1}} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_{u_1}}, \Gamma_{\gamma_{u_M}} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_{u_M}}, \Gamma_{\gamma_{u_1}} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_{u_M}}, \Gamma_{\gamma_{u_M}} \rangle_{\tilde{H}} \end{pmatrix} \begin{pmatrix} \tilde{w}_1 \\ \vdots \\ \tilde{w}_{\tilde{N}} \end{pmatrix} = \begin{pmatrix} \int_0^T \tilde{h}(\gamma_{u_1}(t)) dt \\ \vdots \\ \int_0^T \tilde{h}(\gamma_{u_M}(t)) dt \end{pmatrix},$$

where the Gram matrix for the δ basis is given as $\tilde{G}_1 = (\langle \Gamma_{\gamma_{u_i}}, \Gamma_{\gamma_{u_j}} \rangle_{\tilde{H}})_{i,j=1}^M$.

Further leveraging the established mappings for the multiplication operator, the action of M_μ^* on the δ basis can be expressed in terms of the β basis as

$$[M_\mu^*]_\delta^\beta = G_2^{-1} \left(\int_0^T \Gamma_{\gamma_{u_i}, \mu(\gamma_{u_i}(\cdot))}(\gamma_{u_j}(t)) \begin{pmatrix} 1 \\ u_j(t) \end{pmatrix} dt \right)_{i,j=1}^M.$$

The mapping of $A_{f,g}$ from the β basis to the projection onto the δ basis is given as

$$[A_{f,g}^*]_\beta^\delta = \tilde{G}_1^{-1} \left(\int_0^T K(\gamma_{u_j}(t), \gamma_{u_i}(T)) - K(\gamma_{u_j}(t), \gamma_{u_i}(0)) dt \right)_{i,j=1}^M.$$

In this case, the finite rank representation of $M_\mu A_{f,g}$ is given as

$$[M_\mu A_{f,g}]_\delta^\delta \approx ([A_{f,g}^*]_\beta^\delta [M_\mu^*]_\delta^\beta)^T$$

and has the advantage of avoiding the user selection of centers for kernel functions as in Section 7.1.

8. DISCUSSION

The accuracy of the estimation of the action of $M_\mu A_{f,g}$ on the selected basis functions, whether the kernel basis, the occupation kernel basis, or another collection of basis functions, depends directly on the expressiveness of the collection control occupation kernels corresponding to the controlled trajectories. It can be seen through the invocation of G_2^{-1} in the finite rank representations of the previous section that the rank of the representation is bottlenecked by the control occupation kernels. The rank of G_2 and expressiveness of the collection of kernels can be increased through the segmentation of the trajectories.

The condition number of G_2 depends not only on the trajectories but also of the selected kernel functions. For example, the Gaussian Radial Basis Functions (RBFs), given as $\tilde{K}(x, y) = \exp(-\frac{1}{\mu}\|x - y\|_2^2)$ for $\mu > 0$, are more poorly conditioned for large μ than for μ small. However, large μ values correspond to faster convergence of interpolation problems within the native space of the kernel (cf. [32]). 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_2, however, such formulation is out of the scope of this paper.

The finite rank representation utilizes trajectories of a system, observed under the controllers u_j, to predict its behavior in response to any arbitrary feedback controller μ. Since the control Liouville operator is described only through its action on the control occupation kernels, implicit in the computation is a projection of the feedback $\mu(x)$ onto the span of the control occupation kernels. As a result, knowledge of the behavior of the system in response to different control inputs near a state is required to resolve the action of the feedback controller at that state. Formulation of explicit excitation conditions that relate μ and the control occupation kernels is out of the scope of this paper.

REFERENCES

[1] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems.* Philadelphia, PA, USA: SIAM, 2016.

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

[4] 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.

[5] J. A. Rosenfeld, R. Kamalapurkar, L. Gruss, and T. T. Johnson, "Dynamic mode decomposition for continuous time systems with the liouville operator," *arXiv preprint arXiv:1910.03977*, 2019.

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

[7] 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, Jun. 2010.

[8] 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.

[9] 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.

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

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

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

[13] 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.

[14] 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, 2018.

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

[16] 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*, Dec 2018, pp. 6409–6414.

[17] 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.

[18] J. A. Rosenfeld, B. Russo, R. Kamalapurkar, and T. T. Johnson, "The occupation kernel method for nonlinear system identification," *arXiv preprint arXiv:1909.11792*, 2019.

[19] 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.

[20] 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.

[21] H. Goldstein, C. Poole, and J. Safko, "Classical mechanics," 2002.

[22] 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.

[23] 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.

[24] 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, Jun. 2018.

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

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

[27] G. K. Pedersen, *Analysis now.* Springer Science & Business Media, 2012, vol. 118.

[28] 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.

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

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

[31] G. Golub, K. Solna, and P. Van Dooren, "Computing the svd of a general matrix product/quotient," *SIAM J. Matrix Anal. Appl.*, vol. 22, no. 1, pp. 1–19, 2000.

[32] G. E. Fasshauer, *Meshfree approximation methods with MATLAB.* World Scientific, 2007, vol. 6.