

On Dynamic Mode Decomposition of Control-affine Systems

Moad Abudia, Joel A. Rosenfeld, and Rushikesh Kamalapurkar

Abstract—This paper builds on 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 used to separate the drift dynamics from the input dynamics. A provably convergent finite-rank estimation of a compact control Liouville operator is obtained, provided sufficiently rich data are available. A matrix representation of the finite-rank operator is used to construct a data-driven representation of its singular values, left singular functions, and right singular functions. The singular value decomposition is used to generate a data-driven model of the control-affine nonlinear system. The developed method generates a model that can be used to predict the trajectories of the system in response to any admissible control input.

I. INTRODUCTION

Dynamic mode decomposition (DMD) is a data analysis method that aims to generate a finite-rank representation of a transfer operator corresponding to a nonlinear dynamical system using time series measurements [1]–[3]. The time series is expressed as a linear combination of the dynamic modes. The coefficients in the linear combination are given by exponential functions of time. The dynamic modes and the growth rates of the exponential functions are derived from the spectrum of a finite-rank representation of the Koopman operator (or, in the continuous-time case, Koopman generator). In [3], it was established that the finite rank representation of the Koopman operator converges, in the strong operator topology (SOT), to the true Koopman operator as the amount of data increases. However, convergence of the spectrum doesn't necessarily follow from convergence in the SOT [4]. Since the DMD procedure relies on the spectrum to construct a model, the constructed model is not guaranteed to converge to the true dynamical system model that is being identified.

Kernel methods developed by the machine learning community have been adopted for system identification purposes

This research was supported by the Air Force Office of Scientific Research (AFOSR) under contract numbers FA9550-20-1-0127 and FA9550-21-1-0134, and the National Science Foundation (NSF) under awards 2027976 and 2027999. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsoring agencies.

M. Abudia is with the School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK, USA, email: moad.abudia@okstate.edu

J. A. Rosenfeld is with the Department of Mathematics and Statistics, University of South Florida, Tampa, FL, USA, email: rosenfeldj@usf.edu

R. Kamalapurkar is with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA, email: rkamalapurkar@ufl.edu

by the control community [5]–[7]. The complexity of kernel methods is further investigated in the system identification context in [8], where it shows that tuning of the hyperparameters is a problem that still persists today.

In [9], Liouville operators (or Koopman generators) were used instead of the Koopman operator, where examples of RKHSs and dynamical systems for which the Liouville operators are compact are provided. Furthermore, a finite-rank estimation of the compact Liouville operator is shown to converge in norm to the true Liouville operator. However, [9] does not consider controlled systems, which is the focus of this paper.

In [10] a DMD routine to represent a general nonlinear system with control as a control-affine linear system is developed. This idea is generalized in [11] with extended DMD (eDMD), which provides good predictions but with no spectral convergence guarantees. Additionally, for a general discrete-time, nonlinear dynamical system with control, [11] utilizes the shift operator to describe the time evolution of the control signal. Also, in discrete-time, separation of the control input from the state can be achieved via first order approximations [12]. For continuous-time dynamical systems, the Koopman canonical transform (see [13]) is used in [14] to lift the nonlinear dynamical system and approximate it as a control-affine, bilinear system, called the Koopman bilinear form (KBF), which depends on the usage of a dictionary of functions, and the convergence analysis relies on a non-recurrence assumption.

More recently, in [15] a lifted bilinear representation of the nonlinear system is identified and is used to develop a feedback controller that is guaranteed to stabilize the system given sufficiently rich data and an appropriate region of attraction. In [16] probabilistic bounds for the approximation error and the prediction error as a function of the number of training data points is obtained, using either ergodic trajectories or i.i.d. samples. Both [15] and [16] rely on a user defined library of functions to be used as observables, which have to be rich enough to approximate the system, and [15] assumes that the library is invariant under the infinitesimal generator.

In this paper, we present a method to identify a control-affine nonlinear system via a finite-rank operator, which converges (in norm) to the true compact control Liouville operator provided sufficiently rich data are available. This work uses control occupation kernels, which were introduced in [17] as augmentations of occupation kernels, that incorporate control signal information by leveraging vector valued reproducing kernel Hilbert spaces (vvRKHS). Occupation kernels themselves were introduced in [18], where system

identification problems are addressed not through numerical differentiation, but rather through integration. Using numerical integration is less sensitive to noise [19] as opposed to numerical differentiation used in popular system identification methods such as SINDYc [20].

The main contribution of this paper is the development of a DMD routine for nonlinear control-affine dynamical systems using a compact control Liouville operator. The finite-rank approximation is provably convergent, in norm, to the true control Liouville operator on Bargmann-Fock spaces restricted to the set of real numbers.

II. PRELIMINARIES

Given a Hilbert space \mathcal{Y} and a set X , a vector valued RKHS (vvRKHS), H , is a Hilbert space of functions from X to \mathcal{Y} , such that for each $v \in \mathcal{Y}$ and $x \in X$, the functional $f \mapsto \langle f(x), v \rangle_{\mathcal{Y}}$ is bounded. Hence for each $x \in X$ and $v \in \mathcal{Y}$, there is a function $K_{x,v} \in H$ such that $\langle f(x), v \rangle_{\mathcal{Y}} = \langle f, K_{x,v} \rangle_H$. The mapping $v \mapsto K_{x,v}$ is linear over \mathcal{Y} ; hence, $K_{x,v}$ may be expressed as an operator over \mathcal{Y} as $K_x v := K_{x,v}$. The operator $K(x, y) := K_y^* K_x$ is called the reproducing kernel operator corresponding to H .

In the present context, we define $\mathcal{Y} = \mathbb{R}^{m+1}$ (viewed as row vectors), $X = \mathbb{R}^n$, and H consists of continuously differentiable functions. Since \mathcal{Y} is the space of row-vectors, the operation on $v \in \mathbb{R}^{m+1}$ by the kernel operator K_x will be expressed as $v K_x$. Given two continuous signals, $\theta : [0, T] \rightarrow \mathbb{R}^n$ and $u : [0, T] \rightarrow \mathbb{R}^m$, the control occupation kernel corresponding to this pair of signals is the unique function, $\Gamma_{\theta,u} \in H$, that represents the bounded functional $h \mapsto \int_0^T h(\theta(t)) \begin{pmatrix} 1 \\ u(t) \end{pmatrix} dt$ as $\langle h, \Gamma_{\theta,u} \rangle_H = \int_0^T h(\theta(t)) \begin{pmatrix} 1 \\ u(t) \end{pmatrix} dt$ via the following proposition.

Proposition 1: [19] If H is a vvRKHS consisting of \mathbb{R}^{m+1} (row) valued continuous functions over \mathbb{R}^n , $\gamma_u : [0, T] \rightarrow \mathbb{R}^n$ is a controlled trajectory with continuous controller $u : [0, T] \rightarrow \mathbb{R}^m$ satisfying, for all $t \in [0, T]$, $\frac{d^s \gamma_u}{dt^s}(t) = f(\gamma_u(t)) + g(\gamma_u(t))u(t)$, and $((f)_j \ (g)_j) \in H$ for each $j = 1, \dots, n$, then $\langle ((f)_j \ (g)_j), \Gamma_{\theta,u}^{(s)} \rangle_H = \left(\gamma_u(T) - \sum_{\ell=0}^{s-1} \frac{T^\ell}{\ell!} \frac{d^\ell \gamma_u}{dt^\ell}(0) \right)_j$.

Proof: See the proof of Proposition 1 in [19]. ■

While Proposition 1 follows from a simple application of the fundamental theorem of calculus, it sets the stage for a powerful approximation routine leveraging higher order control occupation kernels. These kernels can be implemented directly using a direct interpolation approach, or they can arise naturally in a regularized regression problem.

Proposition 2: [19] Fix $s \in \mathbb{N}$, let H be a vvRKHS with the kernel operator K_x . If θ and u are continuous signals from $[0, T]$ to \mathbb{R}^n and \mathbb{R}^m respectively. Then the control occupation kernel $\Gamma_{\theta,u}^{(s)}(x)$ can be obtained by

$$\Gamma_{\theta,u}^{(s)}(x) = \frac{1}{(s-1)!} \int_0^T (T-t)^{s-1} \begin{pmatrix} 1 & u^\top(t) \end{pmatrix} K_{\theta(t)}(x) dt. \quad (1)$$

Proof: See the proof of Proposition 2 in [19]. ■

If variable length trajectories are admitted, then it can be shown that the span of the occupation kernels corresponding to trajectories θ_u that result from the application of the control u to the system in (3) is dense in H .

Proposition 3: For any order s , the span of the set

$$A_s := \left\{ \Gamma_{\gamma_u, \gamma_0, u}^{(s)} \mid \begin{array}{l} \gamma_{u, \gamma_0} \in C([0, T_u]; \mathbb{R}^n), \gamma_0 \in \mathbb{R}^n, \\ u \in C([0, T_u]; \mathbb{R}^m), T_u \in [0, T] \end{array} \right\}$$

is dense in H , where γ_{u, γ_0} is a solution of (3), under the control u and starting from γ_0 .

Proof: Select γ_0 such that $h(\gamma_0) \neq 0$. Continuity of h and γ_{u, γ_0} can then be invoked to conclude that for any constant control signal $u(t) = b$ such that $h(\gamma_0) \begin{pmatrix} 1 \\ b \end{pmatrix} \neq 0$, there exists a $T_u > 0$ for which $\int_0^{T_u} h(\gamma_{u, \gamma_0}(t)) \begin{pmatrix} 1 \\ b \end{pmatrix} dt \neq 0$. For example, one can select $T_u = \min \left\{ T, \inf_t \left\{ h(\gamma_{u, \gamma_0}(t)) \begin{pmatrix} 1 \\ b \end{pmatrix} = 0 \right\} \right\}$. A straightforward extension of the above argument to the higher order case allows us to conclude that for any order s and any nonzero h , there exist γ_0 , u , and T_u such that $\langle h, \Gamma_{\gamma_u, \gamma_0, u}^{(s)} \rangle_H \neq 0$. That is, $H \cap A_s^\perp = \{0\}$, and as a result, $(A_s^\perp)^\perp = H$. Since $(A_s^\perp)^\perp = \overline{\text{span } A_s}$, we conclude that $H = \overline{\text{span } A_s}$. ■

Proposition 3 motivates the use of control occupation kernels for system identification. If H is universal, then any continuous function can be approximated, uniformly over any compact set, by a function in H , and any function in H can be approximated using a linear combination of sufficiently many control occupation kernels.

Definition 1: [21] The control Liouville operator with symbol f, g denoted by $A_{f,g} : \mathcal{D}(A_{f,g}) \rightarrow H$ is defined $A_{f,g}h(x) := \nabla_x h(x) (f(x) \ g(x))$ where the domain $\mathcal{D}(A_{f,g})$ is defined canonically as

$$\mathcal{D}(A_{f,g}) = \{h \in H : A_{f,g}h \in H\}. \quad (2)$$

There is a relation between the adjoint of the control Liouville operator and the control occupation kernel, which is presented in the following proposition

Proposition 4: ([21, Proposition 3]) Suppose that f and g correspond to a control Liouville operator, $A_{f,g} : \mathcal{D}(A_{f,g}) \rightarrow H$, and let u be an admissible control signal for the control-affine dynamical system in (3) with a corresponding controlled trajectory, γ_u . Then, $\Gamma_{\gamma_u, u} \in \mathcal{D}(A_{f,g}^*)$ and $A_{f,g}^* \Gamma_{\gamma_u, u} = K(\cdot, \gamma_u(T)) - K(\cdot, \gamma_u(0))$.

Proof: See the proof of Proposition 3 in [21] ■

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 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 multi-variable 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. 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 [17, see Proposition 8].

III. PROBLEM STATEMENT

The objective in this manuscript is to learn an unknown control affine system from observed control signals and controlled trajectories of different lengths, $\{u_j : [0, T_j] \rightarrow \mathbb{R}^m\}_{j=1}^M$ and $\{\gamma_{u_j} : [0, T_j] \rightarrow \mathbb{R}^n\}_{j=1}^M$, respectively, satisfying

$$\dot{x} = F(x, u) = f(x) + g(x)u, \quad (3)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the drift function and $g : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ is the control effectiveness matrix.

Systems of the form (3) encompass linear systems and Euler-Lagrange models with invertible inertia matrices, and hence, represent a wide class of physical plants, including but not limited to robotic manipulators and autonomous ground, aerial, and underwater vehicles.

In order to facilitate the description of the controlled dynamical system in terms of operators, the vvRKHS framework from section II is utilized in this paper.

Definition 2: Given a compact control Liouville operator $A_{f,g} : \tilde{H}_d \rightarrow H$, where \tilde{H}_d is a scalar valued RKHS. The tuples $\{(\sigma_i, \phi_i, \psi_i)\}_{i=1}^\infty$, with $\sigma_i \in \mathbb{R}^n$, $\phi_i \in \tilde{H}_d$, and $\psi_i \in H$, are singular values, left singular vectors, and right singular vectors of $A_{f,g}$, respectively, if $\forall h \in \text{span } d$, $A_{f,g}h = \sum_{i=1}^\infty \sigma_i \psi_i \langle h, \phi_i \rangle_{\tilde{H}_d}$. Let $h_{\text{id}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the identity function. Given singular values, left singular vectors, and right singular vectors of $A_{f,g}$ and a control signal u , the dynamics of the system can be expressed as

$$\begin{aligned} \dot{x} &= \begin{pmatrix} A_{f,g}(h_{\text{id}})_1 \\ \vdots \\ A_{f,g}(h_{\text{id}})_n \end{pmatrix} \begin{pmatrix} 1 \\ u \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^\infty \sigma_i \psi_i \langle (h_{\text{id}})_1, \phi_i \rangle_{\tilde{H}_d} \\ \vdots \\ \sum_{i=1}^\infty \sigma_i \psi_i \langle (h_{\text{id}})_M, \phi_i \rangle_{\tilde{H}_d} \end{pmatrix} \begin{pmatrix} 1 \\ u \end{pmatrix} \quad (4) \end{aligned}$$

The objective of this work is to generate a provably convergent finite truncation of the above model. To that end, we derive a finite rank representation of the control Liouville operator.

IV. FINITE-RANK REPRESENTATION OF THE CONTROL LIOUVILLE OPERATOR

To facilitate computation, an explicit finite-rank representation of $A_{f,g}$ is needed to determine the dynamic modes of the resultant system. In the following, finite collections of linearly independent vectors, d^M and β^M are selected to establish the needed finite-rank representation. Since the adjoint of $A_{f,g}$ maps control occupation kernels to kernel

differences ([21, Proposition 3]), the span of the collection of kernel differences

$$d^M = \{K_d(\cdot, \gamma_{u_i}(T_i)) - K_d(\cdot, \gamma_{u_i}(0))\}_{i=1}^M \subset \tilde{H}_d \quad (5)$$

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

$$\beta^M = \{\Gamma_{\gamma_{u_i}, u_i}\}_{i=1}^M \subset H. \quad (6)$$

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

A rank- M (or less) representation of the operator $A_{f,g}$ is then given by $P_{\beta^M} A_{f,g} P_{d^M} : \tilde{H}_d \rightarrow \text{span } \beta^M$, where P_{d^M} and P_{β^M} denote projection operators onto $\text{span } d^M$ and $\text{span } \beta^M$, respectively.

Under the compactness assumptions and given rich enough data so that the spans of $\{d_i\}_{i=1}^\infty$ and $\{\beta_i\}_{i=1}^\infty$ are dense in \tilde{H}_d and H , respectively, the sequence of finite-rank operators $\{P_{\beta^M} A_{f,g} P_{d^M}\}_{M=1}^\infty$ can be shown to converge, in norm, to $A_{f,g}$. To facilitate the proof of convergence, we recall the following result from [22].

Lemma 1: Let H and G be RKHSs defined on $X \subset \mathbb{R}^n$ and let $A_N : H \rightarrow G$ be a finite-rank operator with rank N . If the spans of $\{d_i\}_{i=1}^\infty$ and $\{\beta_i\}_{i=1}^\infty$ 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 $\|A_N h - P_{\beta^i} A_N h\|_G \leq \epsilon \|h\|_H$.

Proof: See the proof of [22, Theorem 2]. ■

The convergence result for Liouville operators on Bargmann-Fock spaces restricted to the set of real numbers, which will be used for H and \tilde{H}_d , follows from the following more general result.

Proposition 5: If $A : \tilde{H}_d \rightarrow H$ is a compact operator and the spans of $\{d_i\}_{i=1}^\infty$ and $\{\beta_i\}_{i=1}^\infty$ are dense in \tilde{H}_d and H , respectively, then $\lim_{M \rightarrow \infty} \|A - P_{\beta^M} A P_{d^M}\|_{\tilde{H}_d}^H = 0$, where $\|\cdot\|_{\tilde{H}_d}^H$ denotes the operator norm of operators from \tilde{H}_d to H .

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

$$\begin{aligned} \|Ah - P_{\beta^M} A P_{d^M} h\|_H &\leq \|Ah - A_N h\|_H + \|A_N h - A_N P_{d^M} h\|_H \\ &\quad + \|A_N P_{d^M} h - P_{\beta^M} A_N P_{d^M} h\|_H \\ &\quad + \|P_{\beta^M} A_N P_{d^M} h - P_{\beta^M} A P_{d^M} h\|_H. \end{aligned}$$

Using the fact that A_N and $P_{\beta^M} A_N P_{d^M}$ 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 $i \geq M(N)$

$$\begin{aligned} \|Ah - P_{\beta^i} A P_{d^i} h\|_H &\leq \|Ah - A_N h\|_H \\ &\quad + 2\epsilon \|h\|_{\tilde{H}_d} + \|A_N P_{d^i} h - A P_{d^i} h\|_H. \end{aligned}$$

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 $i \geq M(j)$, $\|Ah - P_{\beta^i} A P_{d^i} h\|_H \leq 4\epsilon \|h\|_{\tilde{H}_d}$. ■

The convergence result can then be stated as follows.

Theorem 1: Let $\rho_d \in \mathbb{R}$, $\varrho_d \in \mathbb{R}$, and $\rho = [\rho_1 \ \dots \ \rho_{m+1}]^\top \in \mathbb{R}^{m+1}$ be parameters such that $\rho_i < \varrho_d$, and $\varrho_d < \rho_d$ for $i = 1, \dots, m+1$. Let $\tilde{H}_d = F_{\rho_d}^2(\mathbb{R}^n)$, $\tilde{G}_d = F_{\varrho_d}^2(\mathbb{R}^n)$, and $H = F_{\rho}^2(\mathbb{R}^n)$. If f and g are component-wise polynomial, and if the spans of the collections $\{d_i\}_{i=1}^\infty$ and $\{\beta_i\}_{i=1}^\infty$ are dense in \tilde{H}_d and H , respectively, then $\lim_{M \rightarrow \infty} \|A_{f,g} - P_{\beta^M} A_{f,g} P_{d^M}\|_{\tilde{H}_d}^H = 0$.

Proof: [21, Propositions 7, 8, and 9] imply that $A_{f,g}$ is compact and hence, the theorem follows from Proposition 5. ■

V. MATRIX REPRESENTATION OF THE FINITE-RANK OPERATOR

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

Proposition 6: If $h = \delta^\top d \in \text{span } d$ is a function with coefficients $\delta \in \mathbb{R}^M$ and if $g = P_{\beta} A_{f,g} h$, then $g = a^\top \beta$, where $a = G_{\beta}^+ G_d \delta$ and $(\cdot)^+$ denotes the Moore-Penrose pseudoinverse.

Proof: [21, Propositions 3 and 6] imply that that for all $j = 1, \dots, M$, $A_{f,g}^* \beta_j = d_j$. Note that since g is a projection of $A_{f,g} h$ onto $\text{span } \beta$, $g = a^\top \beta$ for any a that solves

$$G_{\beta} a = \begin{bmatrix} \langle A_{f,g} h, \beta_1 \rangle_H \\ \vdots \\ \langle A_{f,g} h, \beta_M \rangle_H \end{bmatrix} = \begin{bmatrix} \langle h, A_{f,g}^* \beta_1 \rangle_{\tilde{H}_d} \\ \vdots \\ \langle h, A_{f,g}^* \beta_M \rangle_{\tilde{H}_d} \end{bmatrix} \quad (7)$$

Using the adjoint relationship,

$$G_{\beta} a = \begin{bmatrix} \langle h, d_1 \rangle_{\tilde{H}_d} \\ \vdots \\ \langle h, d_M \rangle_{\tilde{H}_d} \end{bmatrix} = \begin{bmatrix} \langle \delta^\top d, d_1 \rangle_{\tilde{H}_d} \\ \vdots \\ \langle \delta^\top d, d_M \rangle_{\tilde{H}_d} \end{bmatrix} = G_d \delta \quad (8)$$

Selecting the solution

$$a = G_{\beta}^+ G_d \delta \quad (9)$$

of (8), a matrix representation $[P_{\beta} A_{f,g}]_d^\beta$ of the operator $P_{\beta} A_{f,g}|_d$ is obtained as $G_{\beta}^+ G_d$. ■

Note that matrix representations are generally not unique. Different representations may be obtained by selecting different solutions of (7) and (9). In the case where the Gram matrix G_{β} is nonsingular, equation (7) has a unique solution, resulting in the unique matrix representation $G_{\beta}^{-1} G_d$.

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

VI. PSEUDO-SINGULAR FUNCTIONS OF THE FINITE-RANK OPERATOR

The following proposition states that an SVD-like decomposition of $P_{\beta} A_{f,g}|_d$ can be computed using the matrix representation $[P_{\beta} A_{f,g}]_d^\beta$ derived in the previous section.

Proposition 7: If (W, Σ, V) is the SVD of G_{β}^+ with $W = [w_1, \dots, w_M]$, $V = [v_1, \dots, v_M]$, and $\Sigma = \text{diag}([\sigma_1, \dots, \sigma_M])$, then for all $i = 1, \dots, M$, σ_i are pseudo-singular values of $P_{\beta} A_{f,g}|_d$ with left pseudo-singular functions $\phi_i := v_i^\top d$ and right pseudo-singular functions $\psi_i := w_i^\top \beta$.

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

$$\begin{aligned} P_{\beta} A_{f,g} h &= \sum_{i=1}^M \sigma_i \psi_i \langle h, \phi_i \rangle_{\tilde{H}_d} \\ \iff P_{\beta} A_{f,g} \delta^\top d &= \sum_{i=1}^M \sigma_i w_i^\top \beta \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d} \end{aligned}$$

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

$$(G_{\beta}^+ G_d \delta)^\top \beta = \left(\sum_{i=1}^M \sigma_i \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d} w_i^\top \right) \beta. \quad (10)$$

Simple matrix manipulations yield the chain of implications

$$\begin{aligned} (10) &\iff \forall \delta \in \mathbb{R}^M, G_{\beta}^+ G_d \delta = \sum_{i=1}^M \sigma_i \langle \delta^\top d, v_i^\top d \rangle_{\tilde{H}_d} w_i \\ &\iff \forall \delta \in \mathbb{R}^M, G_{\beta}^+ G_d \delta = \sum_{i=1}^M \sigma_i (w_i v_i^\top G_d) \delta \\ &\iff G_{\beta}^+ G_d = \sum_{i=1}^M \sigma_i (w_i v_i^\top) G_d \\ &\iff G_{\beta}^+ = \sum_{i=1}^M \sigma_i w_i v_i^\top = W \Sigma V^\top, \end{aligned}$$

which proves the proposition. ■

Remark 1: The standard usage of the term SVD refers to a decomposition using orthonormal left and right singular vectors. The left singular functions $\{\phi_i\}_{i=1}^M$ and right singular functions $\{\psi_i\}_{i=1}^M$ are not necessarily orthonormal. Therefore, the decomposition in the previous proposition is not an SVD of the finite-rank operator $P_{\beta} A_{f,g}|_d$.

In the following section, an orthonormalization process is applied to the domain and range of the finite-rank representation. The orthonormalization process will allow us to compute the SVD of $P_{\beta} A_{f,g}|_d$.

VII. ORTHONORMALIZATION AND SVD

Let $\{\lambda_{\beta,i}, v_{\beta,i}\}_{i=1}^{N_{\beta}}$ and $\{\lambda_{d,i}, v_{d,i}\}_{i=1}^{N_d}$ be nonzero eigenvalues and eigenvectors of G_{β} and G_d , respectively. These eigenvectors can be used to construct orthonormal basis

$\beta' = O_\beta \beta$ and $d' = O_d d$. Where the range orthogonalization matrix is given by

$$O_\beta = \text{diag} \left(\frac{1}{\sqrt{v_{\beta,1}^* G_\beta v_{\beta,1}}} \quad \cdots \quad \frac{1}{\sqrt{v_{\beta,N_\beta}^* G_\beta v_{\beta,N_\beta}}} \right) \begin{pmatrix} v_{\beta,1} \\ \vdots \\ v_{\beta,N_\beta} \end{pmatrix},$$

and the domain orthogonalization matrix is given by

$$O_d = \text{diag} \left(\frac{1}{\sqrt{v_{d,1}^* G_d v_{d,1}}} \quad \cdots \quad \frac{1}{\sqrt{v_{d,N_d}^* G_d v_{d,N_d}}} \right) \begin{pmatrix} v_{d,1} \\ \vdots \\ v_{d,N_d} \end{pmatrix}.$$

The following proposition is used to formulate the matrix representation of the finite-rank representation using the orthonormal basis β' and d' .

Proposition 8: If $h = \delta^\top d' \in \text{span } d$ is a function with coefficients $\delta \in \mathbb{R}^{N_d}$ and if $g = P_\beta A_{f,g} h$, then $g = a^\top \beta'$, where $a = G_{\beta'}^{-1} O_\beta G_d O_d^\top \delta$.

Proof: Note that since g is a projection of $A_{f,g} h$ onto $\text{span } \beta$, $g = a^\top \beta'$ for any a that solves

$$G_{\beta'} a = \begin{bmatrix} \langle A_{f,g} h, \beta'_1 \rangle_H \\ \vdots \\ \langle A_{f,g} h, \beta'_M \rangle_H \end{bmatrix} = \begin{bmatrix} \langle h, A_{f,g}^* \beta'_1 \rangle_{\tilde{H}_d} \\ \vdots \\ \langle h, A_{f,g}^* \beta'_M \rangle_{\tilde{H}_d} \end{bmatrix} \quad (11)$$

Using $\beta'_k = \sum_{i=1}^{N_\beta} \frac{1}{\sqrt{v_{\beta,i}^* G_\beta v_{\beta,i}}} (v_{\beta,i})_i \beta_i$, the adjoint relationship, and $h = \delta^\top d' = \delta^\top O_d d$,

$$G_{\beta'} a = \begin{bmatrix} \delta^\top O_d G_d \frac{1}{\sqrt{v_{\beta,1}^* G_\beta v_{\beta,1}}} v_{\beta,1} \\ \vdots \\ \delta^\top O_d G_d \frac{1}{\sqrt{v_{\beta,N_\beta}^* G_\beta v_{\beta,N_\beta}}} v_{\beta,N_\beta} \end{bmatrix} = O_\beta G_d O_d^\top \delta. \quad (12)$$

Solving for a

$$a = G_{\beta'}^{-1} O_\beta G_d O_d^\top \delta. \quad (13)$$

The matrix representation $[P_\beta A_{f,g}]_{d'}^{\beta'}$ of the operator $P_\beta A_{f,g}|_d$ is thus obtained as $G_{\beta'}^{-1} O_\beta G_d O_d^\top$. ■

The following proposition states that the SVD of $P_\beta A_{f,g}|_d$ can be computed using the matrix representation $[P_\beta A_{f,g}]_{d'}^{\beta'}$.

Proposition 9: If (W, Σ, V) is the SVD of $G_{\beta'}^{-1} O_\beta G_d O_d^\top G_{d'}^{-1}$ with $W = [w_1, \dots, w_{N_d}]$, $V = [v_1, \dots, v_{N_\beta}]$, and $\Sigma = \text{diag}([\sigma_1, \dots, \sigma_N])$, where $N = \min\{N_\beta, N_d\}$, then for all $i = 1, \dots, N$, σ_i are singular values of $P_\beta A_{f,g}|_d$ with left singular functions $\phi_i := v_i^\top d'$ and right singular functions $\psi_i := w_i^\top \beta'$.

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

$$\begin{aligned} P_\beta A_{f,g} h &= (G_{\beta'}^{-1} O_\beta G_d O_d^\top \delta)^\top \beta' \\ &= \sum_{i=1}^N \sigma_i w_i^\top \beta' \langle \delta^\top d', v_i^\top d' \rangle_{\tilde{H}_d} \\ &\Leftarrow G_{\beta'}^{-1} O_\beta G_d O_d^\top \delta = \sum_{i=1}^N \sigma_i w_i w_i^\top G_{d'} \delta \end{aligned}$$

$$\Leftarrow G_{\beta'}^{-1} O_\beta G_d O_d^\top G_{d'}^{-1} = \sum_{i=1}^N \sigma_i w_i v_i^\top,$$

which proves the proposition. ■

For ease of exposition, in the following section, the pseudo-singular values, the pseudo-left singular vectors, and pseudo-right singular vectors are used, along with a finite truncation of (2) to generate a data-driven model. A similar procedure can be carried out using the SVD, resulting in an equivalent model.

VIII. THE SCLDMD ALGORITHM

Motivated by (2), assuming that $h_{\text{id},j} \in \tilde{H}_d$ for $j = 1, \dots, n$, the system dynamics are approximated using the finite-rank representation, with rank at most M , as

$$\dot{x} \approx \hat{F}_M(x, u) := [P_\beta A_{f,g} P_d h_{\text{id}}](x) \begin{pmatrix} 1 \\ u \end{pmatrix},$$

where $P_\beta A_{f,g} P_d h_{\text{id}}$ denotes row-wise operation of the operator $P_\beta A_{f,g}$ on the function $P_d h_{\text{id}}$. Using the definition of singular values and singular functions of $P_\beta A_{f,g}|_d$,

$$\dot{x} \approx \sum_{i=1}^M \sigma_i \xi_i w_i^\top \beta(x) \begin{pmatrix} 1 \\ u \end{pmatrix} = \xi \Sigma W^\top \beta(x) \begin{pmatrix} 1 \\ u \end{pmatrix}, \quad (14)$$

where $\xi_i := \langle P_d h_{\text{id}}, \phi_i \rangle_{\tilde{H}_d}$ and $\xi := [\xi_1, \dots, \xi_M]$.

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

$$\begin{aligned} \xi &= \begin{bmatrix} \langle P_d h_{\text{id},1}, v_1^\top d \rangle_{\tilde{H}_d} & \dots & \langle P_d h_{\text{id},1}, v_M^\top d \rangle_{\tilde{H}_d} \\ \vdots & \ddots & \vdots \\ \langle P_d h_{\text{id},n}, v_1^\top d \rangle_{\tilde{H}_d} & \dots & \langle P_d h_{\text{id},n}, v_M^\top d \rangle_{\tilde{H}_d} \end{bmatrix} \\ &= \begin{bmatrix} \langle \delta_1^\top d, d_1 \rangle_{\tilde{H}_d} & \dots & \langle \delta_1^\top d, d_M \rangle_{\tilde{H}_d} \\ \vdots & \ddots & \vdots \\ \langle \delta_n^\top d, d_1 \rangle_{\tilde{H}_d} & \dots & \langle \delta_n^\top d, d_M \rangle_{\tilde{H}_d} \end{bmatrix} V = \delta^\top G_d V, \end{aligned}$$

where $\delta := [\delta_1, \dots, \delta_n]$. 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 $D := ((\gamma_{u_j}(T_j)) - (\gamma_{u_j}(0))_i)_{i,j=1}^{n,M}$ it can be concluded that $\delta^\top G_d = D$. Finally, the modes ξ are given by $\xi = DV$ and the estimated open-loop model is given by

$$\dot{x} \approx \hat{F}_M(x, u) = DV \Sigma W^\top \beta(x) \begin{pmatrix} 1 \\ u \end{pmatrix} = DG_\beta^+ \beta(x) \begin{pmatrix} 1 \\ u \end{pmatrix}. \quad (15)$$

The approximation $\hat{f}_M(x)$, an approximation of the drift dynamics, $f(x)$, is given by the first column of $DG_\beta^+ \beta(x)$ and $\hat{g}_M(x)$, an approximation of the control-effectiveness matrix, $g(x)$, is given by the last m columns of $DG_\beta^+ \beta(x)$.

Since $P_\beta A_{f,g} P_d$ converges to $A_{f,g}$ in norm as $M \rightarrow \infty$, and since the space $F_{\rho_d}^2(\mathbb{R}^n)$ contains $h_{\text{id},j}$ for $j = 1, \dots, n$, the following result is immediate.

Corollary 1: Under the hypothesis of Theorem 1, for all $u \in \mathcal{U}$, where for all $t \in [0, T]$, $\|(1 \ u^\top(t))\| < \bar{U}$, $\lim_{M \rightarrow \infty} \left(\sup_{x \in X} \left\| \hat{F}_M(x, u) - F(x, u) \right\|_2 \right) = 0$.

Proof: The proof is similar to the proof of [21, Corollary 1]. It is omitted from this paper for the sake of brevity. ■

The SCLDMD technique is summarized in Algorithm 1. The characterization $\Gamma_{\gamma_{u_j}} = \int_0^{T_j} \tilde{K}(\cdot, \gamma_{u_j}(t)) dt$ of occupation kernels is used on line 6. The inner product of two control occupation kernels is given in [21] as

$$\begin{aligned} & \left\langle \Gamma_{\gamma_{u_i}, u_i}, \Gamma_{\gamma_{u_j}, u_j} \right\rangle_H \\ &= \int_0^{T_j} \int_0^{T_i} [1 \ u_i^\top(\tau)] K(\gamma_{u_j}(t), \gamma_{u_i}(\tau)) \begin{bmatrix} 1 \\ u_j(t) \end{bmatrix} d\tau dt. \end{aligned} \quad (16)$$

Algorithm 1 The SCLDMD algorithm

Input: Trajectories $\{\gamma_{u_i}\}_{i=1}^M$, a control signal u , a numerical integration procedure, and reproducing kernels \tilde{K}_d and K of \tilde{H}_d and H , respectively.

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 (16))
 - 2: $D \leftarrow ((\gamma_{u_j}(T_j))_i - (\gamma_{u_j}(0))_i)_{i,j=1}^{n,M}$
 - 3: $(W, \Sigma, V) \leftarrow \text{SVD of } G_\beta^+$
 - 4: $\xi \leftarrow DV$
 - 5: $\phi_j \leftarrow \sum_{i=1}^M (V)_{i,j} (K_d(\cdot, \gamma_{u_i}(T_i)) - K_d(\cdot, \gamma_{u_i}(0)))$
 - 6: $\psi_j \leftarrow \sum_{i=1}^M \int_0^{T_i} (W)_{i,j} \begin{bmatrix} 1 \\ u_i(t)^\top \end{bmatrix} K_{\gamma_{u_i}(t)}(\cdot) dt$
 - 7: **return** $\{\xi_j, \sigma_j, \varphi_j, \phi_j\}_{j=1}^M$
-

IX. CONCLUSION

This paper introduces a novel approach towards the construction of a finite-rank representation of the control Liouville operator. New results on the construction of singular values and functions of the finite rank operator using singular values and vectors of a matrix representation are also obtained. Once the singular values and functions are at hand, the drift dynamics and the control effectiveness can be approximated, which facilitates the prediction of the states of the dynamical system in response to an open-loop control signal. Moreover, the finite-rank representation of the control Liouville operator is shown to be convergent (in norm) to the true control Liouville operator, provided sufficiently rich data are available.

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: Society for Industrial and Applied Mathematics, 2016.
- [2] I. Mezić, “Spectral properties of dynamical systems, model reduction and decompositions,” *Nonlinear Dyn.*, vol. 41, no. 1, pp. 309–325, 2005.
- [3] M. Korda and I. Mezić, “On convergence of extended dynamic mode decomposition to the Koopman operator,” *J. Nonlinear Sci.*, vol. 28, no. 2, pp. 687–710, 2018.
- [4] G. K. Pedersen, *Analysis now*. Springer Science & Business Media, 2012, vol. 118.
- [5] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, “Kernel methods in system identification, machine learning and function estimation: A survey,” *Automatica*, vol. 50, no. 3, pp. 657–682, 2014.
- [6] G. Pillonetto and G. De Nicolao, “A new kernel-based approach for linear system identification,” *Automatica*, vol. 46, no. 1, pp. 81–93, 2010.
- [7] G. Pillonetto, M. H. Quang, and A. Chiuso, “A new kernel-based approach for nonlinear system identification,” *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2825–2840, 2011.
- [8] G. Pillonetto and A. Chiuso, “Tuning complexity in regularized kernel-based regression and linear system identification: The robustness of the marginal likelihood estimator,” *Automatica*, vol. 58, pp. 106–117, 2015.
- [9] M. Abudia, J. A. Rosenfeld, and R. Kamalapurkar, “On convergent dynamic mode decomposition and its equivalence with occupation kernel regression,” in *IFAC-PapersOnLine*, 2024, to appear.
- [10] 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.
- [11] M. Korda and I. Mezić, “Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control,” *Automatica*, vol. 93, pp. 149–160, 2018.
- [12] R. Strässer, J. Berberich, and F. Allgöwer, “Robust data-driven control for nonlinear systems using the koopman operator*,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 2257–2262, 2023, 22nd IFAC World Congress.
- [13] A. Surana, “Koopman operator based observer synthesis for control-affine nonlinear systems,” in *Proc. IEEE Conf. Decis. Control*. IEEE, 2016, pp. 6492–6499.
- [14] 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.
- [15] R. Strässer, M. Schaller, K. Worthmann, J. Berberich, and F. Allgöwer, “Koopman-based feedback design with stability guarantees,” *IEEE Trans. Autom. Control*, 2024.
- [16] F. Nüske, S. Peitz, F. Philipp, M. Schaller, and K. Worthmann, “Finite-data error bounds for koopman-based prediction and control,” *J. Nonlinear Sci.*, vol. 33, no. 1, p. 14, 2023.
- [17] J. A. Rosenfeld and R. Kamalapurkar, “Dynamic mode decomposition with control Liouville operators,” in *IFAC-PapersOnLine*, vol. 54, no. 9, 2021, pp. 707–712.
- [18] J. A. Rosenfeld, B. Russo, R. Kamalapurkar, and T. Johnson, “The occupation kernel method for nonlinear system identification,” *SIAM J. Control Optim.*, vol. 62, no. 3, pp. 1643–1668, 2024.
- [19] M. Abudia, T. Channagiri, J. A. Rosenfeld, and R. Kamalapurkar, “Control occupation kernel regression for nonlinear control-affine systems,” arXiv:2106.00103, in preparation.
- [20] 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.
- [21] J. A. Rosenfeld and R. Kamalapurkar, “Dynamic mode decomposition with control liouville operators,” *IEEE Trans. Autom. Control*, vol. 69, no. 12, pp. 8571–8586, 2024.
- [22] 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.