

SINGULAR DYNAMIC MODE DECOMPOSITION*

JOEL A. ROSENFELD[†] AND RUSHIKESH KAMALAPURKAR[‡]

Abstract. This manuscript is aimed at addressing several long standing challenges in the application of Koopman analysis to dynamic mode decomposition (DMD). Principle among these challenges are the convergence of the algorithms and existence of the associated Koopman modes. This paper introduces two major innovations to address these challenges, where Koopman operators are removed from the analysis in favor of Liouville operators (known as Koopman generators in special cases), and these operators are shown to be compact for appropriately selected pairs of Hilbert spaces as the domain and range of the operator. While these operators no longer admit eigenfunctions in the general analysis due to the domain and the range being different, reconstruction of system trajectories using a suitably modified DMD algorithm is still possible. Furthermore, the sacrifice of eigenfunctions realizes theoretical goals of DMD analysis, such as convergence and existence of Koopman modes, that have yet to be achieved in existing literature. In the case where the domain is embedded in the range, an eigenfunction approach is recovered, along with a more typical DMD routine that leverages a norm-convergent finite rank representation. The manuscript concludes with a description of two DMD algorithms that converge when a dense collection of occupation kernels, arising from the data, are leveraged in the analysis.

1. Introduction. This manuscript is aimed at addressing several long standing challenges in the application of Koopman analysis in dynamic mode decomposition (DMD). Principle among these challenges are the convergence of associated DMD algorithms and the existence of Koopman modes, where the former has only been established with respect to the strong operator topology (SOT) (which does not guarantee the convergence of the spectrum), and the latter requires the Koopman operator to be compact as well as self-adjoint or normal, which is a rare occurrence in typical applications of DMD.

DMD methods are data analysis methods that aim to decompose a time series corresponding to a nonlinear dynamical system into a collection of dynamic modes [14, 4, 16, 13]. When they are effective, a given time series can be expressed as a linear combination of dynamic modes and exponential functions whose growth rates are derived from the spectrum of a finite rank representation of a particular operator, usually the Koopman operator.

In [26], DMD methods were extended to infinite feature sets, and this included the first instance of kernel methods appearing in connection with DMD. Kernel functions were leveraged there to evaluate inner products between two infinite feature representations of finite dimensional data, a machine learning methodology frequently referred to as the “kernel trick.” However, [26] does not analyze the impact a space has on the collection of available Koopman operators, which are frequently unbounded operators over the kernel spaces. Unboundedness has been well explored in the literature under the guise of composition operators, of which Koopman operators are a special case (cf. [5, 6]).

*A YouTube playlist to accompany this work may be found at: https://youtube.com/playlist?list=PLldiDnQu2phsZdFP3nHoGnk_Aq-kp_4nE

Funding: 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 award numbers 2027976 and 1900364. 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.

[†]Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620 USA (rosenfeldj@usf.edu, <http://thelearningdock.org>)

[‡]Department of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK 74078 USA (rushikesh.kamalapurkar@okstate.edu, <https://scc-lab.github.io>)

The use of Koopman operators places certain constraints on the continuous time dynamics that can be studied with DMD methods. In particular, Koopman operators analyze continuous time dynamics through a discrete time proxy obtained by fixing a time-step for a continuous time system [15]. However, only a small subset of continuous time dynamics satisfy the forward complete property necessary to obtain a discretization [19]. Moreover, to establish convergence guarantees for DMD routines, additional structure is required of Koopman operators, where a sequence of finite rank operators converges to the Koopman operator in norm only if the Koopman operator is compact (see, e.g., [17, Theorem 3.3.3]). Compactness is rarely satisfied for Koopman operators, where even the Koopman operator obtained through discretization of the simplest dynamical system $\dot{x} = 0$ is the identity operator and is not compact. More generally, it was established in [23] that no composition operator (including Koopman operators) is compact over an L^2 space having non-atomic measure. A partial result has been demonstrated for when Koopman operators are bounded in [13], where a sequence of finite rank operators converge to a Koopman operator in the Strong Operator Topology (SOT). However, SOT convergence does not guarantee convergence of the spectra [13], which is necessary for a DMD routine.

There are stronger theoretical difficulties associated with Koopman operators. It has been demonstrated that among the typical Hilbert spaces leveraged in sampling theory, such as the exponential dot product's [5], the Gaussian radial basis function (RBF)'s [12], and the polynomial kernel's native spaces as well as the classical Paley Wiener space [6], the only discrete time dynamics that yield a bounded Koopman operator are those dynamics that are affine. Hence, depending on the kernel function selected for the approximation of a Koopman operator, a given Koopman operator can at best be expected to be a densely defined operator, which obviates the aforementioned convergence properties.

Another motivation for the use of Koopman operators in the study of continuous time dynamical systems is a heuristic that for small time steps the spectra and eigenfunctions of the resultant Koopman operator should be close to that of the Liouville operator representing the continuous time systems [3]. However, for two fixed time steps, the corresponding Koopman operators can have different collections of eigenfunctions and eigenvalues, and these are artifacts of the discretization itself [12]. Since in most cases the Koopman operators are used for this analysis, it is not clear if there is a method for distinguishing which of these eigenfunctions and eigenvalues are a product of the discretization and which are fundamental to the dynamics themselves.

Finally, and perhaps most alarming, is that Koopman modes themselves exist for only a small subset of Koopman operators [12]. Specifically, if a Koopman operator is self-adjoint, then it admits an orthonormal basis of eigenfunctions [3], and the projection of the full state observable onto this basis yields a collection of (vector valued) coefficients attached to these basis functions. These coefficients are known as Koopman Modes or Dynamic Modes. Koopman operators are not necessarily diagonalizable over a given Hilbert space, and when they are diagonalizable, their complete eigenbasis is not always an orthogonal basis. Hence, as the full state observable is projected on larger and larger finite collections of eigenfunctions, the weights attached to each eigenfunction will change as more are added. This adjustment to the weights with the addition of more eigenfunctions is why a series expansion is only ever given in Hilbert space theory when there is an orthonormal basis of eigenfunctions, otherwise an expansion is written as limit of finite linear combinations of eigenfunctions.¹

¹There are notable exceptions, such as in atomic decomposition [27, Section 2.5]. However, this

To address these limitations, two major modifications are made, where Koopman operators are removed from the analysis in light of Liouville operators (known as Koopman generators in special cases), and these operators are shown to be compact for certain pairs of Hilbert spaces selected separately as the domain and range of the operator. While eigenfunctions are discarded in the general analysis, a viable reconstruction algorithm is still achievable, and the sacrifice of eigenfunctions realizes the theoretical goals of DMD analysis that have yet to be achieved in other contexts. It should be noted that Liouville and Koopman operators rarely admit a diagonalization, and as such, this approach discards that additional assumption on the operators.

However, at the cost of well defined Dynamic Modes, an eigenfunction approach is still achievable when the domain is embedded in the range of the operator. This allows for the search of eigenfunctions through finite rank approximations that converge to the Liouville operator. The result is a norm convergence DMD routine (using eigenfunctions), which is an achievement over the SOT convergent results previously established in the field [13]. This gives a balance between the two convergence methods presented in this manuscript, where well defined modes come at the price of ease of reconstruction, and a straightforward reconstruction algorithm may not have well defined limiting dynamic modes (a problem shared with all other DMD routines).

To be explicit, the singular DMD approach (Section 4 and Section 6) yields the following benefits:

1. Eliminates the requirement of forward completeness (similar to the method given in [20]).
2. Provides well defined dynamic modes.
3. Approximates a compact operator, thereby achieving convergence.
4. Yields an orthonormal basis through which the full state observable may be decomposed.

However, this achievement comes at the expense of eigenfunctions of the given operator. As it turns out, the abandonment of eigenfunctions for the analysis does not actually limit the applicability, where even for very simple dynamics, such as $f(x) = x^2$ in the one dimensional setting, the corresponding Liouville operators will have no eigenfunctions over any space of continuous functions. For the present example, the solution to the eigenfunction equation, $g'(x)x^2 = \lambda g(x)$, gives $g(x) = e^{\lambda/x}$ for $\lambda \neq 0$, a discontinuous function on the real line. Additionally, reconstruction of the original time series may still be achieved using Runge-Kutta like methods.

The DMD routine leveraging the case where the domain is embedded in the range provides the following advantages (Section 5 and Section 7):

1. Eliminates the requirement of forward completeness (similar to the method given in [20]).
2. Approximates a compact operator, thereby achieving convergence.
3. Yields an approximate eigenbasis through which the full state observable may be decomposed.

It should be noted that there have been several attempts at providing compact operators for the study of DMD. The approaches [9] and [19] find compact operators through the multiplication of auxiliary operator against Koopman and Liouville operators respectively. However, the resultant operators are not the operators that truly correspond to the dynamics in question, and as such, the decomposition of those operators can only achieve heuristic results. The approach taken presently gives compact Liouville operators directly connected with the continuous time dynamics. Specific

is another rare property of a basis.

examples of compact Liouville operators are given in Section 3. However, these examples are not meant to be exhaustive, and other pairs of Hilbert spaces and dynamics may still determine more compact operators.

2. Reproducing Kernel Hilbert Spaces. A reproducing kernel Hilbert space (RKHS), H , over a set X is a space of functions from X to $\mathbb{R}$ such that the functional of evaluation, $E_x g := g(x)$ is bounded for every $x \in X$. By the Riesz theorem, this means for each $x \in X$ there exists a function $K_x \in H$ such that $\langle f, K_x \rangle_H = f(x)$ for all f . The function K_x is called the kernel function centered at X , and the function $K(x, y) := \langle K_y, K_x \rangle_H$ is called the kernel function corresponding to H . Note that $K_y(x) = K(x, y)$. Classical examples of kernel functions in data science are the Gaussian radial basis function for $\mu > 0$, $K(x, y) = \exp(-\frac{1}{\mu}\|x - y\|^2)$, and the exponential dot product kernel, $\exp(\frac{1}{\mu}x^\top y)$ [24, Section 4.1].

The function $K(x, y)$ is a positive definite kernel function, which means that for every finite collection of points, $\{x_1, \dots, x_M\} \subset X$, the Gram matrix $(K(x_i, x_j))_{i,j=1}^M$ is positive definite. For each positive definite kernel function, there exists a unique RKHS for which K is the kernel function for that space by the Aronszajn-Moore theorem in [1].

Given a RKHS, H , over a compact set $X \in \mathbb{R}^n$ consisting of continuous functions and given a continuous signal, $\theta : [0, T] \rightarrow X$, the linear functional $g \mapsto \int_0^T g(\theta(t))dt$ is bounded. Hence, there exist a function, $\Gamma_\theta \in H$, such that $\langle g, \Gamma_\theta \rangle_H = \int_0^T g(\theta(t))dt$ for all $g \in H$. The function Γ_θ is called the occupation kernel in H corresponding to θ . These occupation kernels were first introduced in [20, 21].

2.1. Dynamic Mode Decomposition from the Perspective of RKHSs.

The motivation for DMD methods arise from the behavior of eigenfunctions of certain dynamic operators, and when the full state observable is projected onto those eigenfunctions, a model is obtained in the form of a linear combination of exponential functions. While dynamic operators over RKHSs may not admit a complete eigen-decomposition, DMD methods aim to construct a finite rank approximation of the dynamic operator and to leverage the eigenfunctions of the approximating operator for modeling the state of the unknown dynamic operator. That is, determining eigenfunctions for the generator itself is not the purpose of DMD methods. DMD methods first approximate an operator, preferably in the operator norm, and then the eigenfunctions that are leveraged for the model are the eigenfunctions of the finite rank approximation.

The objective is to find functions for which

$$(2.1) \quad |A_f \phi(x) - \lambda \phi(x)| < \epsilon$$

for some λ and some small positive ϵ and all x within some workspace, which contrasts with other methods that achieve almost everywhere or almost sure convergence [11, 7, 8, 10]. Once accomplished, given a trajectory, $\dot{x} = f(x)$, the eigenfunction behaves approximately as $\phi(x(t)) = x(0)e^{\lambda t}$. The significance of RKHSs is that norm convergence implies pointwise convergence, which has been leveraged in approximation and machine learning frameworks since the 1990s [24, 25]. Since norm convergence in a RKHS of continuous functions yields uniform convergence over compact sets, kernel methods allow for a relaxation of (2.1), where it is sufficient to satisfy $\|A_f \phi - \lambda \phi\|_H < \epsilon$. In turn, if a finite rank approximation of A_f , call it $\hat{A}_f$, is close

enough, it is sufficient to satisfy $\|A_f - \tilde{A}_f\| < \epsilon$, and the rest follows as

$$\begin{aligned} |A_f\phi(x) - \lambda\phi(x)| &< C\|A_f\phi - \lambda\phi\|_H \\ C\|A_f\phi - \tilde{A}_f\phi\|_H &< C\|A_f - \tilde{A}_f\|_H < C\epsilon, \end{aligned}$$

where C is a positive constant that depends on the workspace and the kernel function, and the function ϕ is assumed to be normalized. In the case where the kernel is the Gaussian RBF, C may be taken to be 1.

It is important to note that finite rank operators themselves are almost always diagonalizable, and determining a collection of approximate eigenfunctions from the eigenfunctions of the finite rank operators is certainly well defined. However, to obtain a close approximation of a dynamic operator using a finite rank operator requires compactness (see, e.g., [17, Theorem 3.3.3]), which motivates the investigation of the present manuscript.

3. Compact Liouville Operators. This section demonstrates the existence of compact Liouville operators, given formally as $A_f g(x) = \nabla g(x)f(x)$, where compactness is achieved through the consideration of differing spaces for the domain and range of the operator. Section 3.1 builds on a classical result where differentiation between differing weighted Hardy spaces can be readily shown to be compact. Following a similar argument, Section 3.2 presents several examples of compact Liouville operators over spaces of functions of several variables. We would like to emphasize that the collections of compact Liouville operators are not restricted to these particular pairs of functions spaces, but rather this section provides several examples demonstrating the existence of such operators, thereby validating the approach in the sequel.

3.1. Inspirations from Classical Function Theory. Consider the weighted Hardy spaces (cf. [2]), H_ω^2 , where $\omega = \{\omega_m\}_{m=0}^\infty$ is a sequence of positive real numbers such that $|\omega_{m+1}/\omega_m| \rightarrow 1$, and $g(z) = \sum_{m=0}^\infty a_m z^m$ is a function in H_ω^2 if the coefficients of g satisfy $\|g\|_{H_\omega^2}^2 := \sum_{m=0}^\infty \omega_m |a_m|^2 < \infty$. Each weighted Hardy space is a RKHS over the complex unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| = 1\}$ with kernel function given as $K_\omega(z, w) = \sum_{m=0}^\infty \omega_m z^m \bar{w}^m$, and the monomials $\left\{ \frac{z^m}{\sqrt{\omega_m}} \right\}_{m=0}^\infty$ form an orthonormal basis for each space.

The weighted Hardy space corresponding to the sequence $\omega_{(0)} := \{1, 1, \dots\}$ is the classical Hardy space, H^2 , that was introduced by Riesz in 1923 [18]. The Dirichlet space corresponds to the weight sequence $\omega_{(1)} = \{(m+1)\}_{m=0}^\infty$, and the Bergman space corresponds to $\omega_{(-1)} = \{(m+1)^{-1}\}_{m=0}^\infty$. Of interest here is the weighted Hardy space corresponding to $\omega_{(3)} := \{m^3\}_{m=0}^\infty$, which will be denoted as H_3^2 for convenience.

It is immediately evident that the operation of differentiation on elements of H_3^2 is bounded as an operator from H_3^2 to H^2 . The reason for this inclusion can be seen directly through the power series for these function spaces. In particular, a function $h(z) = \sum_{m=0}^\infty a_m z^m$ is in H_3^2 if $\|h\|_{H_3^2}^2 = \sum_{m=0}^\infty (m+1)^3 |a_m|^2 < \infty$, and in the Hardy space if $\|h\|_{H^2}^2 = \sum_{m=0}^\infty |a_m|^2 < \infty$.

A function g in H_3^2 has derivative $g'(z) = \sum_{m=1}^\infty m a_m z^{m-1} = \sum_{m=0}^\infty (m+1) a_{m+1} z^m$, and by considering the Hardy space norm,

$$\left\| \frac{d}{dz} g \right\|_{H^2} = \sum_{m=0}^\infty (m+1)^2 |a_{m+1}|^2 \leq \sum_{m=0}^\infty (m+1)^3 |a_{m+1}|^2,$$

but this is exactly the H_3^2 norm on g less the constant term. Hence differentiation

is a bounded operator from the space H_3^2 to the Hardy space with operator norm at most 1.

PROPOSITION 3.1. *The operator $\frac{d}{dz} : H_3^2 \rightarrow H^2$ is compact. Moreover, if $f : \mathbb{D} \rightarrow \mathbb{D}$ is a bounded analytic function corresponding to a bounded multiplication operator, $M_f g := g(x)f(x)$, over the Hardy space, then the Liouville operator, $A_f := M_f \frac{d}{dz}$, is compact from H_3^2 to H^2 .*

Proof. To see that differentiation is a compact operator from the H_3^2 to the Hardy space, we may select a sequence of finite rank operators that converge in norm to differentiation. In particular, note that the monomials form an orthonormal basis of the Hardy space as is evident from the given norm. Let $\alpha_M := \{1, z, \dots, z^M\}$ be the first M monomials in z , and let P_{α_M} be the projection onto the span of these monomials. The operator $P_{\alpha_M} \frac{d}{dz}$ is a finite rank operator, where the image of this operator is a polynomial of degree up to M .

To demonstrate that this sequence of finite rank operators converges to differentiation in the operator norm it must be shown that the difference under the operator norm,

$$\left\| P_{\alpha_M} \frac{d}{dz} - \frac{d}{dz} \right\|_{H_3^2}^{H^2} := \sup_{g \in H_3^2} \frac{\|P_{\alpha_M} \frac{d}{dz} g - \frac{d}{dz} g\|_{H^2}}{\|g\|_{H_3^2}},$$

goes to zero. Note that

$$\begin{aligned} \|P_{\alpha_M} \frac{d}{dz} g - \frac{d}{dz} g\|_{H^2}^2 &= \sum_{m=M+1}^{\infty} (m+1)^2 |a_{m+1}|^2 \\ &= \sum_{m=M+1}^{\infty} \frac{1}{m+1} (m+1)^3 |a_{m+1}|^2 \leq \frac{1}{M+1} \sum_{m=M+1}^{\infty} (m+1)^3 |a_{m+1}|^2 \leq \frac{1}{M+1} \|g\|_{H_3^2}^2. \end{aligned}$$

Hence $\|P_{\alpha_M} \frac{d}{dz} - \frac{d}{dz}\|_{H_3^2}^{H^2} \leq \frac{1}{M+1} \rightarrow 0$. This proves that differentiation is a compact operator from H_3^2 to H^2 .

If a function, f , is a bounded analytic function on the closed unit disc, then it is the symbol for a bounded multiplier over H^2 . Hence, the $M_f \frac{d}{dz}$ is a compact operator from H_3^2 to H^2 . To be explicit, since $P_{\alpha_M} \frac{d}{dz}$ has finite rank, $M_f (P_{\alpha_M} \frac{d}{dz})$ also has finite rank. Moreover, $\|M_f P_{\alpha_M} \frac{d}{dz} - M_f \frac{d}{dz}\|_{H_3^2}^{H^2} = \|M_f (P_{\alpha_M} \frac{d}{dz} - \frac{d}{dz})\|_{H_3^2}^{H^2} \leq \|M_f\|_{H^2}^{H^2} \|P_{\alpha_M} \frac{d}{dz} - \frac{d}{dz}\|_{H_3^2}^{H^2} \rightarrow 0$. Hence, $M_f \frac{d}{dz}$ is an operator norm limit of finite rank operators, and is compact. Finally, it can be seen that $M_f \frac{d}{dz} g(z) = g'(z)f(z) = A_f g(z)$, and A_f is a compact Liouville operator from H_3^2 to H^2 . $\square$

3.2. Compact Liouville Operators of Several Variables. The example of the previous section demonstrated that compact Liouville operators may be obtained in one dimension. However, this is readily extended to higher dimensions through similar arguments, and in particular can be demonstrated for dot product kernels of the form $K(x, y) = (1 + \mu x^\top y)^{-1}$. In some cases, such as with the exponential dot product kernel and the Gaussian RBF, where the kernel functions over $\mathbb{R}^n$ decompose as a product of kernel functions over $\mathbb{R}$ for the individual variables, the establishment of compact Liouville operators from the single variable spaces to an auxiliary range RKHSs yields compact Liouville operators through tensor products of the respective spaces.

The exponential dot product kernel, with parameter $\mu > 0$, is given as $K(x, y) = \exp(\mu x^\top y)$. In the single variable case, the native space for this kernel may be expressed as $F_\mu^2(\mathbb{R}) = \left\{ f(x) = \sum_{m=0}^{\infty} a_m x^m : \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\mu^m} < \infty \right\}$. This definition can be readily extended to higher dimensions, where collection of monomials, $x^\alpha \frac{\mu^{|\alpha|}}{\sqrt{\alpha!}}$, with multi-indices $\alpha \in \mathbb{N}^n$ form an orthonormal basis. The norm of functions in $F_\mu^2(\mathbb{R}^n)$ will be denoted by $\|g\|_\mu$.

In this setting, if $\mu_2 > \mu_1$ (i.e. $1/\mu_1 > 1/\mu_2$), then by arguments similar to those given in the previous section, it follows that partial differentiation with respect to each variable is a compact operator from $F_{\mu_1}^2$ to $F_{\mu_2}^2$. However, since multiplication operators are unbounded from F_μ^2 to itself for every $\mu > 0$, another step is necessary to ensure compactness.

LEMMA 3.2. *Suppose that $\eta < \mu$, then given any polynomial of several variables, f , the multiplication operator $M_f : F_\eta^2(\mathbb{R}^n) \rightarrow F_\mu^2(\mathbb{R}^n)$ is bounded.*

Proof. To facilitate a clarity of exposition, this will be proven with respect to functions of a single variable. The same arguments extend to the spaces of several variables, albeit with more bookkeeping.

Let $g \in F_\eta^2$. Then $g(x) = \sum_{m=0}^{\infty} a_m x^m$, and $\|g\|_\eta^2 = \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\eta^m}$.

For $f \equiv 1$, M_1 is the identity operator. Thus, the boundedness of M_1 is equivalent to demonstrating that F_η^2 is boundedly included in F_μ^2 . In particular, note that

$$\begin{aligned} \|M_1 g\|_\mu^2 &= \|g\|_\mu^2 = \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\mu^m} = \sum_{m=0}^{\infty} |a_m|^2 \left(\frac{\eta}{\mu}\right)^m \frac{m!}{\eta^m} \\ &< \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\eta^m} = \|g\|_\eta^2 \end{aligned}$$

Fix $k \in \mathbb{N}$ and consider the multiplication operator $M_{x^k} : F_\eta^2 \rightarrow F_\mu^2$ defined as $M_{x^k} g := xg$ for all $g \in F_\eta^2$. Note that the power series of $M_{x^k} g$ is given as $xg(x) = \sum_{m=0}^{\infty} a_m x^{m+k} = \sum_{m=k}^{\infty} a_{m-k} x^m$. Hence,

$$\begin{aligned} \|x^k g(x)\|_\mu^2 &= \sum_{m=k}^{\infty} |a_{m-k}|^2 \frac{m!}{\mu^m} = \sum_{m=0}^{\infty} |a_m|^2 \frac{(m+k)!}{\mu^{m+k}} \\ &= \sum_{m=0}^{\infty} |a_m|^2 \frac{(m+k)!}{m! \mu^k} \frac{m!}{\mu^m} = \sum_{m=0}^{\infty} |a_m|^2 \left(\frac{m+k}{m! \mu^k}\right) \left(\frac{\eta}{\mu}\right)^m \frac{m!}{\eta^m}, \end{aligned}$$

and as $\left(\frac{m+k}{m! \mu^k}\right) \left(\frac{\eta}{\mu}\right)^m$ is bounded as a function of m by some constant $C > 0$ (owing to the exponential decay of $(\eta/\mu)^m$), it follows that $\|M_{x^k}\|_{F_\eta^2 \rightarrow F_\mu^2} < C$.

Hence, by linear combinations of monomials it has been demonstrated that a multiplication operator with polynomial symbol is a bounded operator. $\square$

Remark 3.3. The authors emphasize that the collection of bounded multiplication operators between these spaces is strictly larger than the those with polynomial symbols. The purpose of this lemma is to simply support the existence of compact Liouville operators, rather than to provide a complete classification.

THEOREM 3.4. *Let $\mu_3 > \mu_1$, and suppose that f is a vector valued function over several variables, where each entry is a polynomial. Then the Liouville operator $A_f : F_{\mu_1}^2(\mathbb{R}^n) \rightarrow F_{\mu_3}^2(\mathbb{R}^n)$ defined as $A_f g = \nabla g \cdot f$ is a compact operator.*

Proof. Let $f = (f_1, f_2, \dots, f_n)^\top$, and select μ_2 such that $\mu_1 < \mu_2 < \mu_3$. For each $i = 1, \dots, n$, the operator of partial differentiation $\frac{\partial}{\partial x_i} : F_{\mu_1}^2(\mathbb{R}^n) \rightarrow F_{\mu_2}^2(\mathbb{R}^n)$ is a compact operator, and the multiplication operator $M_{f_i} : F_{\mu_2}^2(\mathbb{R}^n) \rightarrow F_{\mu_3}^2(\mathbb{R}^n)$ is bounded. Hence, the operator $M_{f_i} \frac{\partial}{\partial x_i}$ is compact. As $A_f = M_{f_1} \frac{\partial}{\partial x_1} + \dots + M_{f_n} \frac{\partial}{\partial x_n}$, it follows that A_f is a compact operator from $F_{\mu_1}^2(\mathbb{R}^n)$ to $F_{\mu_3}^2(\mathbb{R}^n)$. $\square$

This section has thus established the existence of compact Liouville operators between various pairs of spaces. It is emphasized that these are not the only pairs for which a compact Liouville operator may be determined.

4. Singular Dynamic Mode Decomposition for Compact Liouville Operators. The objective of this section is to determine a decomposition of the full state observable, $g_{id}(x) := x$, with respect to an orthonormal basis obtained from a Liouville operator corresponding to a continuous time dynamical system $\dot{x} = f(x)$. We will let H and $\tilde{H}$ be two RKHSs over $\mathbb{R}^n$ such that the Liouville operator, $A_f g(x) = \nabla g(x) f(x)$ is compact as an operator from H to $\tilde{H}$. To obtain an orthonormal basis, a singular value decomposition for the compact operator A_f is obtained. Specifically, note that as A_f is compact, so is A_f^* . Hence, $A_f^* A_f$ is diagonalizable as a self adjoint compact operator. Thus, there is a countable collection of non-negative eigenvalues $\sigma_m^2 \geq 0$ and eigenfunctions φ_m corresponding to $A_f^* A_f$, such that $A_f^* A_f \varphi_m = \sigma_m^2 \varphi_m$. Since $A_f^* A_f$ is self adjoint, $\{\varphi_m\}_{m=0}^\infty$ may be selected in such a way that they form an orthonormal basis of H . The functions φ_m are the right singular vectors of A_f .

For $\sigma_m \neq 0$, the left singular vectors may be determined as $\psi_m := \frac{A_f \varphi_m}{\sigma_m}$, and the collection of nonzero ψ_m form an orthonormal set in $\tilde{H}$. This may be seen via

$$\begin{aligned} \langle \psi_m, \psi_{m'} \rangle_{\tilde{H}} &= \frac{1}{\sigma_m \sigma_{m'}} \langle A_f \varphi_m, A_f \varphi_{m'} \rangle_{\tilde{H}} \\ &= \frac{1}{\sigma_m \sigma_{m'}} \langle \varphi_m, A_f^* A_f \varphi_{m'} \rangle_H = \frac{\sigma_{m'}^2}{\sigma_m \sigma_{m'}} \langle \varphi_m, \varphi_{m'} \rangle = \frac{\sigma_{m'}^2}{\sigma_m \sigma_{m'}} \delta_{m,m'}, \end{aligned}$$

where $\delta_{\cdot,\cdot}$ is the Kronecker delta function.

Finally,

$$A_f g = \sum_{\sigma_m \neq 0} \langle g, \varphi_m \rangle_H \sigma_m \psi_m$$

for all $g \in H$, and

$$A_f^* h = \sum_{\sigma_m \neq 0} \langle h, \psi_m \rangle_{\tilde{H}} \sigma_m \varphi_m.$$

To find a decomposition for the full state observable, g_{id} , first note that the full state observable is vector valued, whereas the Hilbert spaces consist of scalar valued functions. To ameliorate this discrepancy, we will work with the individual entries of the full state observable, namely the maps $x \mapsto (x)_i$, for $i = 1, \dots, n$, which are the mappings of x to its individual components. When $(x)_i$ resides in the Hilbert space, such as with the space $F_\mu^2(\mathbb{R}^n)$, and $(x)_i$ may be directly expanded with respect to the right singular vectors of A_f . If $(x)_i$ is not in the space, as in the case with the Gaussian RBF, if the space is universal, then a suitable approximation may be determined over a fixed compact subset, and the approximation will be expanded instead. Performing the entry wise decomposition of the full state observable is equivalent to performing the decomposition over vector valued RKHSs with diagonal kernel operators, and replacing the gradient of g with the matrix valued derivative.

Hence, for each $i = 1, \dots, n$, we have $(x)_i = \sum_{m=0}^{\infty} (\xi_m)_i \varphi_m(x)$, where $(\xi_m)_i = \langle (x)_i, \varphi_m \rangle_H$. The vectors ξ_m are called the *singular Liouville modes* of the dynamical system with respect to the pair of Hilbert space H and $\tilde{H}$.

Note that for a trajectory of the system, given as $x(t)$, it can be seen that

$$\begin{aligned} \dot{x}(t) &= f(x(t)) = \nabla g_{id}(x(t))f(x(t)) = A_f g_{id}(x(t)) \\ &= \sum_{m=0}^{\infty} \langle g_{id}, \varphi_m \rangle_H \sigma_m \psi_m(x(t)) = \sum_{m=0}^{\infty} \xi_m \sigma_m \psi_m(x(t)). \end{aligned}$$

Hence, $x(t)$ satisfies a differential equation with respect to the left singular vectors of the Liouville operator and the singular Liouville modes. Given these quantities, reconstruction of $x(t)$ is possible using tools from the solution of initial value problems. In particular, the following form of the equation may be exploited:

$$x(t) = x(0) + \sum_{m=0}^{\infty} \xi_m \sigma_m \int_0^t \psi_m(x(\tau)) d\tau.$$

5. Recovering an Eigenfunction Approach in Special Cases. While the majority of this manuscript is aimed at the singular DMD, where the domain and range are different for the compact Liouville operator, there is still a possibility of obtaining an eigendecomposition in special cases. In particular, for many of the examples shown above, the domain and range spaces have similar structure and the range space has less stringent requirement for the functions it contains. This means that the domain itself may be embedded in the range space, and if there is a complete set of eigenfunctions in this embedded space, then the operator may still be diagonalized.

Note that the operator is still mapping between two different Hilbert spaces, which means that the inner product on the embedding is different than the inner product on the domain. This difference will appear in the numerical methods given in subsequent sections.

The following is a well known result (cf. [27, Theorem 2.10]), and is included here for illustration purposes.

PROPOSITION 5.1. *If $\mu_1 < \mu_2$, then $F_{\mu_1}^2(\mathbb{R}^n) \subset F_{\mu_2}^2(\mathbb{R}^n)$.*

Proof. Again this is shown for the single variable case, where the multivariate case follows by an identical argument, but with more bookkeeping.

Suppose that $g \in F_{\mu_1}^2(\mathbb{R})$ with $g(z) = \sum_{m=0}^{\infty} a_m z^m$. Then

$$\|g\|_{F_{\mu_2}^2(\mathbb{R})}^2 = \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\mu_2^m} = \sum_{m=0}^{\infty} |a_m|^2 \left(\frac{\mu_1}{\mu_2}\right)^m \frac{m!}{\mu_1^m} \leq \sum_{m=0}^{\infty} |a_m|^2 \frac{m!}{\mu_1^m} = \|g\|_{F_{\mu_1}^2(\mathbb{R})}^2.$$

Since the quantity on the right is bounded, so is the quantity on the left. Hence $g \in F_{\mu_2}^2(\mathbb{R})$. $\square$

EXAMPLE 1. *A simple example demonstrating that an eigenbasis may be found between the two spaces arises in the study of $A_x : F_{\mu_1}^2(\mathbb{R}) \rightarrow F_{\mu_2}^2(\mathbb{R})$ for $\mu_1 < \mu_2$. Note that an eigenfunction, φ , for A_x must reside in $F_{\mu_1}^2(\mathbb{R}) \cap F_{\mu_2}^2(\mathbb{R}) = F_{\mu_1}^2(\mathbb{R})$, and satisfy $\varphi'(x)x = \lambda \varphi(x)$. Consequently, takes the form $\varphi(x) = x^\lambda$, and is in $F_{\mu_1}^2(\mathbb{R})$ only for $\lambda = 0, 1, 2, \dots$. Hence, the eigenfunctions of A_x are the monomials. Monomials are contained in $F_{\mu_1}^2(\mathbb{R})$ and form a complete eigenbasis for both spaces. Note that the norm of x^m is $\sqrt{\frac{m!}{\mu_1^m}}$ in $F_{\mu_1}^2(\mathbb{R})$ and $\sqrt{\frac{m!}{\mu_2^m}}$ in $F_{\mu_2}^2(\mathbb{R})$.*

The following proposition is obtained in the same manner as in the classical case.

PROPOSITION 5.2. *Suppose that H and $\tilde{H}$ are two RKHSs over $\mathbb{R}^n$, and that $H \subset \tilde{H}$. If $\varphi \in H$ is an eigenfunction for A_f as $A_f\varphi = \lambda\varphi$, then given a trajectory $x : [0, T] \rightarrow \mathbb{R}^n$ satisfying $\dot{x} = f(x)$ the following holds $\varphi(x(t)) = e^{\lambda t}\varphi(x(0))$.*

Proof. Since $A_f\varphi = \nabla\varphi f$, it follows that

$$\frac{d}{dt}\varphi(x(t)) = \nabla\varphi(x(t))\dot{x}(t) = \nabla\varphi(x(t))f(x(t)) = A_f\varphi(x(t)) = \lambda\varphi(x(t)).$$

That is, $\frac{d}{dt}\varphi(x(t)) = \lambda\varphi(x(t))$. Thus, the conclusion follows. $\square$

Suppose that $A_f : H \rightarrow \tilde{H}$ has a complete eigenbasis in the sense that the span of the eigenfunctions, $\{\varphi_m\}_{m=1}^\infty$, are dense in H . If g_{id} , is the full state observable, then each entry of g_{id} , $(x)_i$ for $i = 1, \dots, n$, may be expressed as

$$(x)_i = \lim_{M \rightarrow \infty} \sum_{m=1}^M (\xi_{m,M})_i \varphi_m(x),$$

where $(\xi_{m,M})_i$ is the m -th coefficient obtained from projecting $(x)_i$ onto the span of the first M eigenfunctions. If the eigenfunctions are orthogonal, then the dependence on M may be removed from $\xi_{m,M}$. Hence, the full state observable is obtained from

$$(5.1) \quad g_{id}(x) = \lim_{M \rightarrow \infty} \sum_{m=1}^M \xi_{m,M} \varphi_m(x),$$

with $\xi_{m,M}$ being the vector obtained by stacking $(\xi_{m,M})_i$. Finally, by substituting $x(t)$ into this representation (where $\dot{x} = f(x)$), the following holds

$$(5.2) \quad x(t) = g_{id}(x(t)) = \lim_{M \rightarrow \infty} \sum_{m=1}^M \xi_{m,M} e^{\lambda t} \varphi_m(x(0)).$$

Hence, this methodology yields a DMD routine, where the finite rank representations will converge to the compact Liouville operators, following the proof given in the Appendix of [19].

6. Singular Dynamic Mode Decomposition Algorithm. This section is aimed at determining a convergent algorithm that can determine approximations of the singular Liouville modes and the singular vectors of A_f . While an eigenfunction expansion is still possible in the case of nested spaces, the Singular DMD algorithm is technically more general. Moreover, the SVD ensures the existence of dynamic modes, which may not be well defined fixed concepts for the eigenfunction case.

From the data perspective, a collection of trajectories, $\{\gamma_j : [0, T_j] \rightarrow \mathbb{R}^n\}_{j=1}^M$, corresponding to an unknown dynamical system, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, as $\dot{\gamma}_j = f(\gamma_j)$ have been observed. The objective of DMD is to get an approximation of the dynamic modes of the system, and to obtain an approximate reconstruction of a given trajectory. Once a reconstruction is determined, then data driven predictions concerning future states of the trajectory may be determined. A DMD routine is somewhat like a Fourier series representation, which can reproduce a continuous trajectory exactly, however DMD methods exploit a trajectory's underlying dynamic structure.

This routine effectively interpolates the action of the Liouville operator on a collection of basis functions. When these basis functions form a complete set within

the Hilbert space, which can be achieved by selecting a dense collection of short trajectories throughout the workspace, then a sequence of finite rank approximations determined by this routine converges to the compact Liouville operator in norm. Which means that the left and right singular functions of the finite rank operators in the sequence converge to those of the Liouville operator, and that the singular values converge as well.

DMD routines involving the Koopman operator add the additional requirement of forward completeness for the sake of discretization. This method as well as that of [19] sidestep that requirement by accessing the Liouville operators directly through their connection with the occupation kernels of the RKHSs. To wit, given two RKHSs of continuously differentiable functions, H and $\tilde{H}$, with kernels K and $\tilde{K}$ respectively, and a compact Liouville operator, $A_f : H \rightarrow \tilde{H}$, the occupation kernel, $\Gamma_{\gamma_j} \in \tilde{H}$ corresponding to the trajectory γ_j satisfies

$$(6.1) \quad A_f^* \Gamma_{\gamma_j} = K(\cdot, \gamma_j(T_j)) - K(\cdot, \gamma_j(0)),$$

where K is the kernel function for the space H . In particular, given $g \in H$,

$$\begin{aligned} \langle A_f g, \Gamma_{\gamma_j} \rangle_{\tilde{H}} &= \int_0^{T_j} \nabla g(\gamma_j(t)) f(\gamma_j(t)) dt \\ &= \int_0^{T_j} \dot{g}(\gamma_j(t)) dt = g(\gamma_j(T_j)) - g(\gamma_j(0)) = \langle g, K_{\gamma_j(T)} - K_{\gamma_j(0)} \rangle_H. \end{aligned}$$

The objective is to construct a finite rank approximation of A_f through which an SVD may be performed to find approximate singular values and singular vectors, and to ultimately approximate the singular Liouville modes. Note that since the dynamics are unknown, the adjoint must be approximated instead, where the action of the adjoint on the occupation kernels provides a sample of the operator. Thus, the finite rank representation will be determined through the restriction of $\tilde{H}$ to the span of the ordered basis $\alpha = \{\Gamma_{\gamma_j}\}_{j=1}^M$. A corresponding basis for H must also be selected, and given the available information, $\beta = \{K(\cdot, \gamma_j(T_j)) - K(\cdot, \gamma_j(0))\}$ is most reasonable. Of course, this leads to a rather benign matrix representation of

$$[A_f^*]_{\alpha}^{\beta} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Moreover, if this matrix is input into an SVD routine, typical algorithms would not be aware of the non-orthogonal inner products between the basis elements. To rectify this, two orthonormal bases α' and β' may be obtained from an eigendecomposition of the Gram matrices (which are assumed to be strictly positive definite)

$$\begin{aligned} (6.2) \quad \tilde{G} &:= \begin{pmatrix} \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \\ &= V \Lambda V^* := \begin{pmatrix} | & & | \\ \tilde{v}_1 & \cdots & \tilde{v}_M \\ | & & | \end{pmatrix} \begin{pmatrix} \tilde{\lambda}_1 & & \\ & \cdots & \\ & & \tilde{\lambda}_M \end{pmatrix} \begin{pmatrix} - & \tilde{v}_1^* & - \\ & \vdots & \\ - & \tilde{v}_M^* & - \end{pmatrix} \end{aligned}$$

for α and

$$(6.3) \quad G = \begin{pmatrix} \langle A_f^* \Gamma_{\gamma_1}, A_f^* \Gamma_{\gamma_1} \rangle_H & \cdots & \langle A_f^* \Gamma_{\gamma_1}, A_f^* \Gamma_{\gamma_M} \rangle_H \\ \vdots & \ddots & \vdots \\ \langle A_f^* \Gamma_{\gamma_M}, A_f^* \Gamma_{\gamma_1} \rangle_H & \cdots & \langle A_f^* \Gamma_{\gamma_M}, A_f^* \Gamma_{\gamma_M} \rangle_H \end{pmatrix}$$

$$= \tilde{V} \tilde{\Lambda} \tilde{V}^* := \begin{pmatrix} | & & | \\ v_1 & \cdots & v_M \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \cdots & \\ & & \lambda_M \end{pmatrix} \begin{pmatrix} - & v_1^* & - \\ & \vdots & \\ - & v_M^* & - \end{pmatrix}$$

for β . A more meaningful representation of A_f^* may be found by re-expressing $[A_f^*]_{\alpha}^{\beta}$ in terms of the orthonormal sets $\alpha' = \{q_j\}_{j=1}^M$ and $\beta' = \{p_j\}_{j=1}^M$ where

$$p_j = \frac{1}{\sqrt{v_j^* G v_j}} \sum_{\ell=1}^M (v_j)_{\ell} (K(\cdot, \gamma_{\ell}(T_{\ell})) - K(\cdot, \gamma_{\ell}(0))), \text{ and}$$

$$q_j = \frac{1}{\sqrt{\tilde{v}_j^* \tilde{G} \tilde{v}_j}} \sum_{\ell=1}^M (\tilde{v}_j)_{\ell} \Gamma_{\gamma_{\ell}}.$$

In other words the orthonormal set α' is given by,

$$(6.4) \quad \begin{pmatrix} q_1(x) \\ \vdots \\ q_M(x) \end{pmatrix} = \begin{pmatrix} \left(\sqrt{\tilde{v}_1^* \tilde{G} \tilde{v}_1} \right)^{-1} & & \\ & \ddots & \\ & & \left(\sqrt{\tilde{v}_M^* \tilde{G} \tilde{v}_M} \right)^{-1} \end{pmatrix} \tilde{V}^{\top} \begin{pmatrix} \Gamma_{\gamma_1}(x) \\ \vdots \\ \Gamma_{\gamma_M}(x) \end{pmatrix},$$

and a similar expression may be written for the orthonormal set β' . Write

$$(6.5) \quad \tilde{V}_0 = \tilde{V} \text{diag} \left(\left(\sqrt{\tilde{v}_1^* \tilde{G} \tilde{v}_1} \right)^{-1}, \dots, \left(\sqrt{\tilde{v}_M^* \tilde{G} \tilde{v}_M} \right)^{-1} \right),$$

and

$$(6.6) \quad V_0 = V \text{diag} \left(\left(\sqrt{v_1^* G v_1} \right)^{-1}, \dots, \left(\sqrt{v_M^* G v_M} \right)^{-1} \right),$$

where the coefficients of each column of V_0 and $\tilde{V}_0$ correspond to functions of norm 1 for their respective spaces. It follows that

$$(6.7) \quad [A_f^*]_{\alpha'}^{\beta'} = V_0^{-1} [A_f^*]_{\alpha}^{\beta} \tilde{V}_0 = V_0^{-1} \tilde{V}_0.$$

That is, the matrix representation with respect to the bases β' and α' are obtained by sending elements of α' to α , computing the action of $[A_f^*]_{\alpha}^{\beta}$ on this transformation, and then sending the result expressed in terms of the β basis to β' .

Now the approximate singular vectors may be obtained for A_f by taking the SVD of $[A_f^*]_{\alpha'}^{\beta'}$. In particular, the right singular vectors of $[A_f^*]_{\alpha'}^{\beta'}$ will correspond to the approximate left singular functions of A_f and vice versa. In particular, writing the

SVD of $[A_f^*]_{\alpha'}^{\beta'}$ as

$$[A_f^*]_{\alpha'}^{\beta'} = \hat{U} \hat{\Sigma} \hat{V}^* = \begin{pmatrix} | & & | \\ \hat{u}_1 & \cdots & \hat{u}_M \\ | & & | \end{pmatrix} \begin{pmatrix} \hat{\sigma}_1^2 & & \\ & \ddots & \\ & & \hat{\sigma}_M^2 \end{pmatrix} \begin{pmatrix} - & \hat{v}_1^* & - \\ & \vdots & \\ - & \hat{v}_M^* & - \end{pmatrix},$$

and the approximate right singular vector for A_f is $\hat{\varphi}_j = \frac{1}{\sqrt{\hat{u}_j^* G_p \hat{u}_j}} \sum_{\ell} (\hat{u}_j)_{\ell} p_{\ell}$, and the approximate left singular vector for A_f is $\hat{\psi}_j = \frac{1}{\sqrt{\hat{v}_j^* G_q \hat{v}_j}} \sum_{\ell} (\hat{v}_j)_{\ell} q_{\ell}$, where G_p and G_q are the Gram matrices for the ordered bases β' and α' respectively.

Translating this to the original bases α and β , we find the following:

$$(6.8) \quad \hat{\varphi}_j = \frac{1}{\sqrt{\hat{u}_j^* G_p \hat{u}_j}} \hat{u}_j^{\top} V_0^{\top} \begin{pmatrix} K(\cdot, \gamma_1(T_1)) - K(\cdot, \gamma_1(0)) \\ \vdots \\ K(\cdot, \gamma_M(T_M)) - K(\cdot, \gamma_M(0)) \end{pmatrix},$$

and

$$(6.9) \quad \hat{\psi}_j = \frac{1}{\sqrt{\hat{v}_j^* G_q \hat{v}_j}} \hat{v}_j^{\top} \tilde{V}_0^{\top} \begin{pmatrix} \Gamma_{\gamma_1} \\ \vdots \\ \Gamma_{\gamma_M} \end{pmatrix}.$$

Thus, if $x : [0, T] \rightarrow \mathbb{R}^n$ satisfies $\dot{x} = f(x)$, then it may be approximately expressed through the integral equation

$$(6.10) \quad x(t) \approx x(0) + \int_0^t \sum_{j=1}^M \sigma_j \hat{\xi}_j \hat{\psi}_j(x(\tau)) d\tau,$$

where

$$(6.11) \quad f(x) \approx \sum_{j=1}^M \sigma_j \hat{\xi}_j \hat{\psi}_j(x)$$

serves as an approximation of the vector field f and

$$(6.12) \quad \hat{\xi}_j = \begin{pmatrix} \langle (x)_1, \hat{\varphi}_j \rangle_H \\ \vdots \\ \langle (x)_n, \hat{\varphi}_j \rangle_H \end{pmatrix} = \text{diag} \left(\frac{1}{\sqrt{\hat{u}_1^* G_p \hat{u}_1}}, \dots, \frac{1}{\sqrt{\hat{u}_M^* G_p \hat{u}_M}} \right) \\ \times \begin{pmatrix} - & \hat{u}_1^{\top} & - \\ & \vdots & \\ - & \hat{u}_M^{\top} & - \end{pmatrix} V_0^{\top} \begin{pmatrix} (\gamma_1(T_1))_j - (\gamma_1(0))_j \\ \vdots \\ (\gamma_M(T_M))_j - (\gamma_M(0))_j \end{pmatrix}$$

are the DMD modes. The singular DMD technique is summarized in Algorithm 6.1.

7. The Eigenfunction based DMD Algorithm. In this section it will be assumed that $A_f : H \rightarrow \tilde{H}$ is a compact operator, and that $H \subset \tilde{H}$. Since A_f is compact, it is bounded, which means that unlike [12] and [19], no additional assumptions are needed concerning the domain of this operator.

Algorithm 6.1 Pseudocode for the singular DMD routine of Section 6. Once the singular DMD modes, the right and left singular vectors, and the singular values are returned, (6.10) can be used, along with a numerical integration routine, to reconstruct trajectories of the system starting from any given initial condition $x(0)$. Evaluation of the left singular vector $\hat{\psi}$ at $x(\tau)$ itself requires evaluation of the occupation kernels $\Gamma_{\gamma_j}(x(\tau))$, which can be achieved using the integral representation of the occupation kernel as $\Gamma_{\gamma_j}(x(\tau)) = \int_0^{T_j} \tilde{K}(x(\tau), \gamma_j(t)) dt$ (cf. [19]). The choice of numerical integration routine can have a significant impact on the overall results, and it is advised that a high accuracy method is leveraged in practice.

Require: Sampled trajectories $\{\gamma_j : [0, T] \rightarrow \mathbb{R}^n\}_{j=1}^M$

Require: Kernel function $K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ of the domain RKHS H

Require: Kernel function $\tilde{K} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ of the range RKHS $\tilde{H}$

Require: A numerical integration routine

- 1: Construct the matrix $\tilde{G}$ in (6.2) using $\langle \Gamma_{\gamma_j}, \Gamma_{\gamma_i} \rangle_{\tilde{H}} = \int_0^{T_i} \int_0^{T_j} \tilde{K}(\gamma_i(\tau), \gamma_j(t)) dt d\tau$ and a numerical integration routine (cf. [19])
 - 2: Construct the matrix G in (6.3) using (6.1)
 - 3: Compute eigenvalues, $\tilde{\lambda}_i$, and eigenvectors, $\tilde{v}_i$, of $\tilde{G}$
 - 4: Compute eigenvalues, λ_i , and eigenvectors, v_i , of G
 - 5: Construct the matrices $\tilde{V}_0$ and V_0 using (6.5) and (6.6), respectively
 - 6: Construct the matrices $G_p = [\langle p_i, p_j \rangle_H]_{i,j=1}^M$ and $G_q = [\langle q_i, q_j \rangle_{\tilde{H}}]_{i,j=1}^M$ using (6.4) and a similar expression for β'
 - 7: Construct the finite rank representation $[A_f^*]_{\alpha'}^{\beta'}$ using (6.7)
 - 8: Construct the approximate right and left singular vectors of A_f using the left and right singular vectors of $[A_f^*]_{\alpha'}^{\beta'}$ and (6.8) and (6.9), respectively
 - 9: Construct the singular DMD modes $\hat{\xi}_j$ using (6.12)
 - 10: **return** Singular DMD modes, $\hat{\xi}_j$, approximate normalized right and left singular vectors $\hat{\varphi}_j, \hat{\psi}_j$, respectively, and approximate singular values σ_j for $j = 1, \dots, M$
-

7.1. Derivation of the Eigenfunction Method. For a collection of observed trajectories, $\{\gamma_1, \dots, \gamma_M\}$ consider the collection of occupation kernels for the space H , denoted by $\alpha = \{\Gamma_{\gamma_1}, \dots, \Gamma_{\gamma_M}\}_{m=1}^M$, and the occupation kernels for the space $\tilde{H}$, denoted by $\beta = \{\tilde{\Gamma}_{\gamma_1}, \dots, \tilde{\Gamma}_{\gamma_M}\}$. Let P_α be the projection from H to H onto the span of α , and let $\tilde{P}_\alpha$ and $\tilde{P}_\beta$ be the corresponding projections onto the spans of α and β respectively (viewed as subspaces of $\tilde{H}$). The numerical method presented in this section will construct a matrix representation for the operator $\tilde{P}_\alpha \tilde{P}_\beta A_f P_\alpha$, where the matrix, $[\tilde{P}_\alpha \tilde{P}_\beta A_f P_\alpha]_\alpha^\alpha$, represents this operator on the span of α in the domain and range respectively. Note that since the matrix representation is defined over α , $[\tilde{P}_\alpha \tilde{P}_\beta A_f P_\alpha]_\alpha^\alpha = [\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha$.

Recall that for a function $g \in \tilde{H}$, $\tilde{P}_\beta g$ is a linear combination of the functions of β as $\sum_{m=1}^M w_m \tilde{\Gamma}_{\gamma_m}$, where the weights are obtained via

$$\begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_M \end{pmatrix} = \begin{pmatrix} \langle g, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots \\ \langle g, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix},$$

and the matrix $G_\beta := [\langle \tilde{\Gamma}_{\gamma_i}, \tilde{\Gamma}_{\gamma_j} \rangle_{\tilde{H}}]_{i,j=1}^M$ is the Gram matrix for the basis β in the

space $\tilde{H}$.

Hence, for each Γ_{γ_j} , the weights for the projection, $\sum_{m=1}^M w_m \tilde{\Gamma}_{\gamma_m}$, of $A_f \Gamma_{\gamma_j}$ onto the span of β may be obtained as

$$\begin{aligned} & \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_M \end{pmatrix} \\ &= \begin{pmatrix} \langle A_f \Gamma_{\gamma_j}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots \\ \langle A_f \Gamma_{\gamma_j}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} = \begin{pmatrix} \langle \Gamma_{\gamma_j}, A_f^* \tilde{\Gamma}_{\gamma_1} \rangle_H \\ \vdots \\ \langle \Gamma_{\gamma_j}, A_f^* \tilde{\Gamma}_{\gamma_M} \rangle_H \end{pmatrix} \\ &= \begin{pmatrix} \langle \Gamma_{\gamma_j}, K_{\gamma_1(T_1)} - K_{\gamma_1(0)} \rangle_H \\ \vdots \\ \langle \Gamma_{\gamma_j}, K_{\gamma_M(T_M)} - K_{\gamma_M(0)} \rangle_H \end{pmatrix} = \begin{pmatrix} \Gamma_{\gamma_j}(\gamma_1(T_1)) - \Gamma_{\gamma_j}(\gamma_1(0)) \\ \vdots \\ \Gamma_{\gamma_j}(\gamma_M(T_M)) - \Gamma_{\gamma_j}(\gamma_M(0)) \end{pmatrix}. \end{aligned}$$

Next, a projection of $\sum_{m=1}^M w_m \tilde{\Gamma}_{\gamma_m}$ onto the span of α within $\tilde{H}$ must be performed. For each $\tilde{\Gamma}_{\gamma_j}$, the weights corresponding to its projection, $\sum_{l=1}^M v_{l,j} \Gamma_{\gamma_l}$, onto the span of α are given via

$$\begin{pmatrix} \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \begin{pmatrix} v_{1,j} \\ \vdots \\ v_{M,j} \end{pmatrix} = \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_j}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots \\ \langle \tilde{\Gamma}_{\gamma_j}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}.$$

Computation the inner products $\langle \Gamma_{\gamma_i}, \Gamma_{\gamma_j} \rangle_{\tilde{H}}$ and $\langle \tilde{\Gamma}_{\gamma_j}, \Gamma_{\gamma_i} \rangle_{\tilde{H}}$ in the equation above is detailed in Section 7.2.

Hence, the projection of $A_f \Gamma_{\gamma_j}$ is given as

$$\begin{aligned} & \tilde{P}_\alpha \tilde{P}_\beta A_f \Gamma_{\gamma_j} = \sum_{m=1}^M w_m \sum_{\ell=1}^M v_{\ell,m} \Gamma_{\gamma_\ell} = \\ & \sum_{m=1}^M w_m \sum_{\ell=1}^M \left(\begin{pmatrix} \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}^{-1} \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_m}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots \\ \langle \tilde{\Gamma}_{\gamma_m}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \right)^\top \begin{pmatrix} \Gamma_{\gamma_1} \\ \vdots \\ \Gamma_{\gamma_M} \end{pmatrix} \\ &= \left(\begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}^{-1} \begin{pmatrix} \Gamma_{\gamma_j}(\gamma_1(T_1)) - \Gamma_{\gamma_j}(\gamma_1(0)) \\ \vdots \\ \Gamma_{\gamma_j}(\gamma_M(T_M)) - \Gamma_{\gamma_j}(\gamma_M(0)) \end{pmatrix} \right)^\top \\ & \times \left(\begin{pmatrix} \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}^{-1} \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \tilde{\Gamma}_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \right)^\top \begin{pmatrix} \Gamma_{\gamma_1} \\ \vdots \\ \Gamma_{\gamma_M} \end{pmatrix}, \end{aligned}$$

and the final representation, $[\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha$ is given as

$$(7.1) \quad [\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha = \begin{pmatrix} \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \Gamma_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}^{-1} \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \Gamma_{\gamma_1} \rangle_{\tilde{H}} \\ \vdots & & \vdots \\ \langle \tilde{\Gamma}_{\gamma_1}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \Gamma_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix} \\ \times \begin{pmatrix} \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_1}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \\ \vdots & \ddots & \vdots \\ \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_1} \rangle_{\tilde{H}} & \cdots & \langle \tilde{\Gamma}_{\gamma_M}, \tilde{\Gamma}_{\gamma_M} \rangle_{\tilde{H}} \end{pmatrix}^{-1} \\ \times \begin{pmatrix} \Gamma_{\gamma_1}(\gamma_1(T_1)) - \Gamma_{\gamma_1}(\gamma_1(0)) & \cdots & \Gamma_{\gamma_M}(\gamma_1(T_1)) - \Gamma_{\gamma_M}(\gamma_1(0)) \\ \vdots & & \vdots \\ \Gamma_{\gamma_1}(\gamma_M(T_M)) - \Gamma_{\gamma_1}(\gamma_M(0)) & \cdots & \Gamma_{\gamma_M}(\gamma_M(T_M)) - \Gamma_{\gamma_M}(\gamma_M(0)) \end{pmatrix}.$$

Note that when $H = \tilde{H}$ and the occupation kernels are assumed to be in the domain of the Liouville operator, the first two matrices cancel, and the representation reduces to that of [19].

Under the assumption of diagonalizability for (7.1), which holds for almost all matrices, an eigendecomposition for (7.1) may be determined as

$$[\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha = \begin{pmatrix} | & & | \\ V_1 & \cdots & V_M \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_M \end{pmatrix} \begin{pmatrix} | & & | \\ V_1 & \cdots & V_M \\ | & & | \end{pmatrix}^{-1},$$

where each column, V_j , is an eigenvector of $[\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha$ with eigenvalue λ_j . The corresponding normalized eigenfunction is given as

$$(7.2) \quad \hat{\varphi}_j(x) = \frac{1}{\sqrt{V_j^\top G_\alpha V_j}} V_j^\top \begin{pmatrix} \Gamma_{\gamma_1} \\ \vdots \\ \Gamma_{\gamma_M} \end{pmatrix},$$

where the normalization is performed in the Hilbert space H through the Gram matrix for α , G_α , according to H 's inner product. Set $\bar{V}_j := \frac{1}{\sqrt{V_j^\top G_\alpha V_j}} V_j$, and let $\bar{V} := (\bar{V}_1 \cdots \bar{V}_M)$.

The Gram matrix for the normalized eigenbasis may be quickly computed as $\bar{V}^\top G_\alpha \bar{V}$, and the weights for the projection of the full state observable onto this

Algorithm 7.1 Pseudocode for the eigenfunction based DMD routine of Section 7. Once the singular DMD modes, the normalized eigenfunctions, and the eigenvalues are returned, (7.5) can be used along with a numerical integration routine to reconstruct trajectories of the system starting from any given initial condition $x(0)$. Similar to Algorithm 6.1, evaluation of the eigenfunctions at $x(0)$ requires the integral representation $\Gamma_{\gamma_j}(\cdot) = \int_0^{T_j} K(\cdot, \gamma_j(t)) dt$. The choice of numerical integration routine can have a significant impact on the overall results, and it is advised that a high accuracy method is leveraged in practice. If the matrices in steps 1, 2, and 3 are close to singular, they can be regularized by adding $\epsilon I_{M \times M}$ where $\epsilon > 0$ is a small constant.

Require: Sampled trajectories $\{\gamma_j : [0, T] \rightarrow \mathbb{R}^n\}_{j=1}^M$

Require: Kernel function $K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ of the domain RKHS H

Require: Kernel function $\tilde{K} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ of the range RKHS $\tilde{H}$

Require: A numerical integration routine

- 1: Construct the matrix $G_\beta := \left[\langle \tilde{\Gamma}_{\gamma_i}, \tilde{\Gamma}_{\gamma_j} \rangle_{\tilde{H}} \right]_{i,j=1}^M$ using $\langle \tilde{\Gamma}_{\gamma_j}, \tilde{\Gamma}_{\gamma_i} \rangle_{\tilde{H}} = \int_0^{T_i} \int_0^{T_j} \tilde{K}(\gamma_i(\tau), \gamma_j(t)) dt d\tau$ and a numerical integration routine (cf. [19])
 - 2: Construct the matrix $[\langle \Gamma_{\gamma_i}, \Gamma_{\gamma_j} \rangle_{\tilde{H}}]_{i,j=1}^M$ using (7.8) and a numerical integration routine
 - 3: Construct the matrix $[\langle \tilde{\Gamma}_{\gamma_i}, \Gamma_{\gamma_j} \rangle_{\tilde{H}}]_{i,j=1}^M$ using (7.7) and a numerical integration routine
 - 4: Construct the matrix $[\Gamma_{\gamma_j}(\gamma_i(T_i)) - \Gamma_{\gamma_j}(\gamma_i(0))]_{i,j=1}^M$ using the integral representation $\Gamma_{\gamma_j}(\cdot) = \int_0^{T_j} K(\cdot, \gamma_j(t)) dt$ and a numerical integration routine
 - 5: Construct the matrix $[\tilde{P}_\alpha \tilde{P}_\beta A_f]_\alpha^\alpha$ using (7.1) and compute its eigenvalues, λ_i , and eigenvectors, V_i
 - 6: Use (7.2) and a numerical integration routine to compute the eigenfunctions $\hat{\varphi}_j$
 - 7: Use (7.3) and a numerical integration routine to compute the singular DMD modes $\hat{\xi}_j$
 - 8: **return** Singular DMD modes, ξ_j , eigenfunctions, $\hat{\varphi}_j$, and eigenvalues λ_j for $j = 1, \dots, M$
-

eigenbasis may be written as

$$\begin{aligned}
 (7.3) \quad \begin{pmatrix} - & \hat{\xi}_1^\top & - \\ & \vdots & \\ - & \hat{\xi}_M^\top & - \end{pmatrix} &= (\bar{V}^\top G_\alpha \bar{V})^{-1} \begin{pmatrix} \langle (x)_1, \hat{\varphi}_1 \rangle_H & \cdots & \langle (x)_n, \hat{\varphi}_1 \rangle_H \\ \vdots & \ddots & \vdots \\ \langle (x)_1, \hat{\varphi}_M \rangle_H & \cdots & \langle (x)_n, \hat{\varphi}_M \rangle_H \end{pmatrix} \\
 &= (\bar{V}^\top G_\alpha \bar{V})^{-1} \bar{V}^\top \begin{pmatrix} \langle (x)_1, \Gamma_{\gamma_1} \rangle_H & \cdots & \langle (x)_n, \Gamma_{\gamma_1} \rangle_H \\ \vdots & \ddots & \vdots \\ \langle (x)_1, \Gamma_{\gamma_M} \rangle_H & \cdots & \langle (x)_n, \Gamma_{\gamma_M} \rangle_H \end{pmatrix} \\
 &= (\bar{V}^\top G_\alpha \bar{V})^{-1} \bar{V}^\top \begin{pmatrix} \int_0^{T_1} \gamma_1(t)^\top dt \\ \vdots \\ \int_0^{T_1} \gamma_M(t)^\top dt \end{pmatrix}
 \end{aligned}$$

and thus,

$$(7.4) \quad g_{id}(x) \approx \sum_{m=1}^M \hat{\xi}_m \hat{\varphi}_m(x).$$

The approximation error (with respect to the norm of the RKHS) approaches zero if the number of trajectories increases and the corresponding collection of occupation kernels forms a dense set. Convergence in the norm of the RKHS implies uniform convergence on compact subsets of the domain.

Consequently, a trajectory $x : [0, T] \rightarrow \mathbb{R}^n$ satisfying $\dot{x} = f(x)$ may be approximately expressed as

$$(7.5) \quad x(t) = g_{id}(x(t)) \approx \sum_{m=1}^M \hat{\xi}_m e^{\lambda_m t} \hat{\varphi}_m(x(0)),$$

where the eigenfunctions for the finite rank approximation of A_f play the role of eigenfunctions for the original operator, A_f . Furthermore, the vector field f may be approximated as

$$(7.6) \quad f(x) \approx \sum_{m=1}^M \lambda_m \hat{\xi}_m \hat{\varphi}_m(x).$$

Note that for a given $\epsilon > 0$ there is a sufficiently large collection of trajectories and occupation kernels such that $\|\tilde{P}_\alpha \tilde{P}_\beta A_f P_\alpha - A_f\|_{\tilde{H}}^H < \epsilon$. Hence, if $\hat{\varphi}$ is a normalized eigenfunction for the finite rank representation with eigenvalue λ , then

$$\|\lambda \hat{\varphi} - A_f \hat{\varphi}\|_{\tilde{H}} = \|\tilde{P}_\alpha \tilde{P}_\beta A_f P_\alpha \hat{\varphi} - A_f \hat{\varphi}\|_{\tilde{H}} \leq \epsilon \|\hat{\varphi}\|_H = \epsilon.$$

Consequently, given a compact subset of $\mathbb{R}^n$ and a given tolerance, ϵ_0 , a finite rank approximation may be selected such that for each normalized eigenfunction the relation $|\frac{d}{dt} \hat{\varphi}(x(t)) - \lambda \hat{\varphi}(x(t))| < \epsilon_0$ for all $x(t)$ in the compact set. Hence, for sufficiently rich information, $\hat{\varphi}(x(t)) \approx e^{\lambda t} \hat{\varphi}(x(0))$.

7.2. Computational Remarks for the Eigenfunction Method. Some entries for the matrices in the above computations require a bit more analysis. Namely, this includes the inner products, $\langle \Gamma_{\gamma_i}, \Gamma_{\gamma_j} \rangle_{\tilde{H}}$ and $\langle \Gamma_{\gamma_i}, \tilde{\Gamma}_{\gamma_j} \rangle_{\tilde{H}}$. All the other quantities have been discussed at length in [20, 21, 19].

The second quantity simply utilizes the functional definition of the function $\tilde{\Gamma}_{\gamma_j}$ as a function in $\tilde{H}$, $\langle \Gamma_{\gamma_i}, \tilde{\Gamma}_{\gamma_j} \rangle_{\tilde{H}} = \int_0^{T_j} \Gamma_{\gamma_i}(\gamma_j(t)) dt = \int_0^{T_j} \int_0^{T_i} K(\gamma_j(t), \gamma_i(\tau)) d\tau dt$, where K is the kernel function for H . Note that this means

$$(7.7) \quad \langle \Gamma_{\gamma_i}, \tilde{\Gamma}_{\gamma_j} \rangle_{\tilde{H}} = \langle \Gamma_{\gamma_i}, \Gamma_{\gamma_j} \rangle_H.$$

However, the first quantity is more complicated and is context dependent. In particular, Γ_{γ_i} is not the occupation kernel corresponding to $\tilde{H}$, so it's functional relationship cannot be exploited in the same manner. On the other hand, $\Gamma_{\gamma_i}(x) = \int_0^{T_i} K(x, \gamma_i(t))$. To compute the inner product in $\tilde{H}$, a specific selection of spaces must be considered.

In the particular setting where $H = F_{\mu_1}^2(\mathbb{R}^n)$ and $\tilde{H} = F_{\mu_2}^2(\mathbb{R}^n)$, with $\mu_1 < \mu_2$, it follows that $\Gamma_{\gamma_i}(x) = \int_0^T e^{\mu_1 x^\top \gamma_i(t)} dt$. Moreover, $K(x, \gamma_i(t)) = e^{\mu_1 x^\top \gamma_i(t)} =$

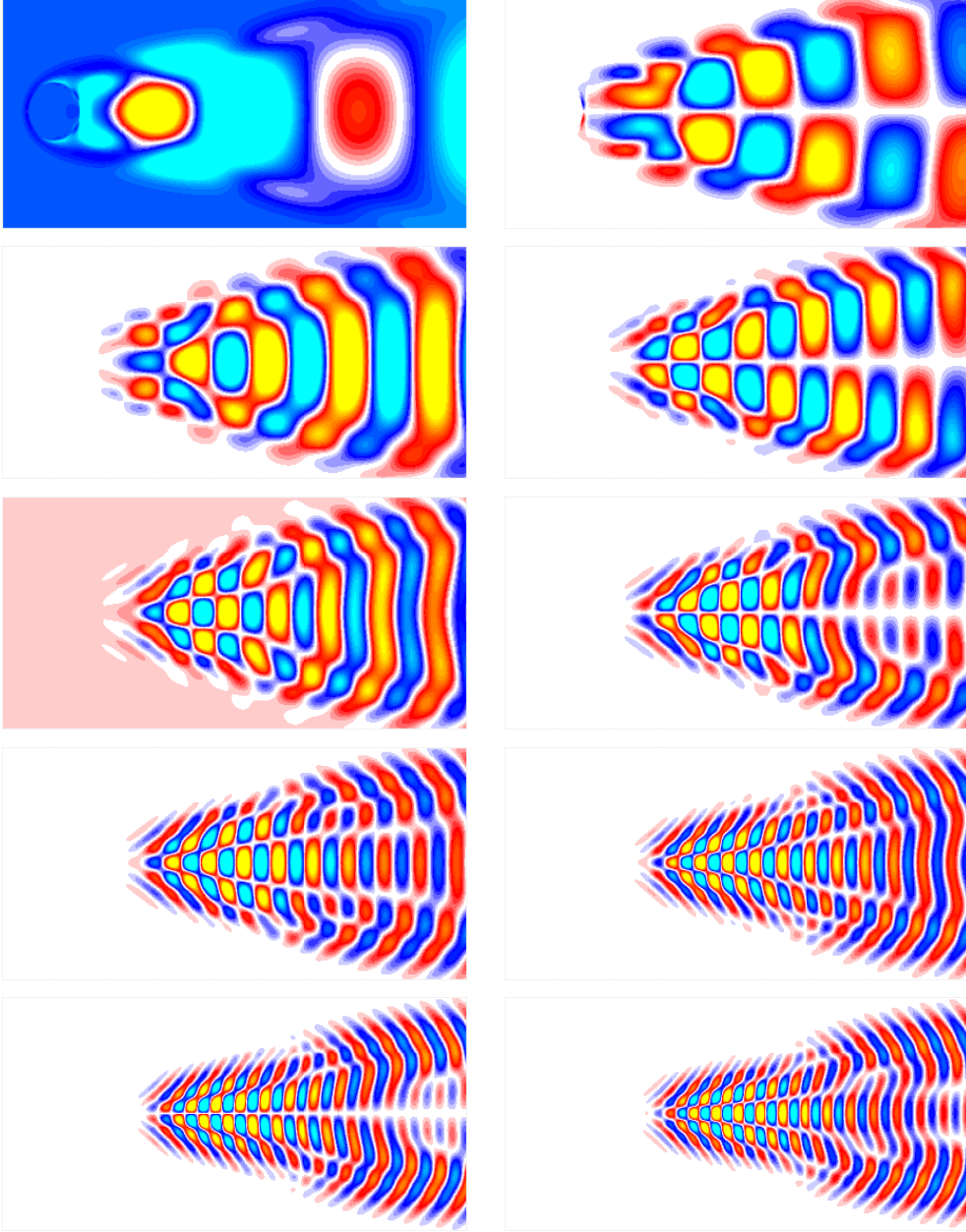


FIG. 1. A selection of the real parts of approximate Liouville modes obtained using the exponential dot product kernel, where the domain corresponds to $\mu_1 = 1/10000$ and the range corresponds to $\mu_2 = 1/9999$.

$e^{\mu_2 x^\top \left(\frac{\mu_1}{\mu_2} \gamma(t) \right)} = \tilde{K}(x, (\mu_1/\mu_2) \gamma_i(t))$. Hence, $\Gamma_{\gamma_i}(x) = \tilde{\Gamma}_{(\mu_1/\mu_2) \gamma_i}(x)$, and

$$\begin{aligned}
 (7.8) \quad \langle \Gamma_{\gamma_i}, \Gamma_{\gamma_j} \rangle_{\tilde{H}} &= \langle \tilde{\Gamma}_{(\mu_1/\mu_2) \gamma_i}, \tilde{\Gamma}_{(\mu_1/\mu_2) \gamma_j} \rangle_{\tilde{H}} \\
 &= \int_0^{T_i} \int_0^{T_j} \tilde{K}((\mu_1/\mu_2) \gamma_i(t), (\mu_1/\mu_2) \gamma_j(\tau)) d\tau dt.
 \end{aligned}$$

The eigenfunction approach to singular DMD is summarized in Algorithm 7.1.

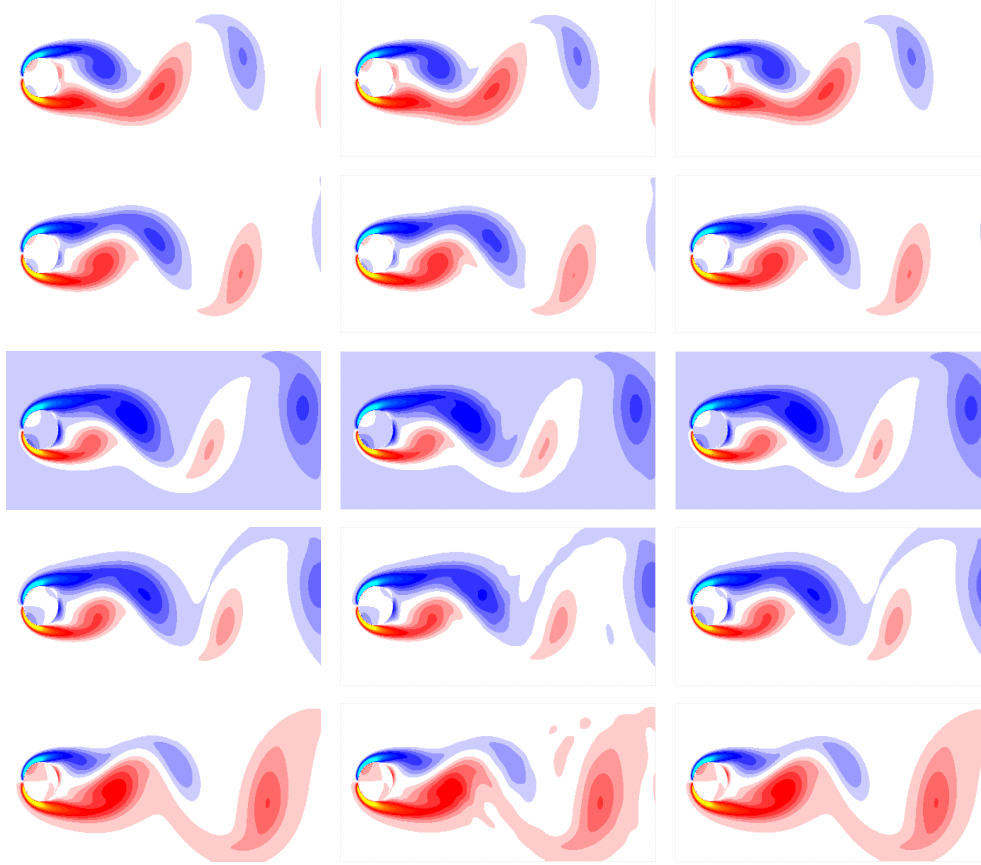


FIG. 2. A selection of true and reconstructed snapshots for the cylinder flow example. From top to bottom, the first column presents the true snapshots at $t = 0.18s$, $t = 1.08s$, $t = 1.58s$, $t = 2.24s$, and $t = 2.98s$, respectively, the second column presents the corresponding reconstructed snapshots, and the third column presents snapshots reconstructed using kernel DMD (see [26]). Parameters for the two methods are tuned manually to yield the smallest reconstruction error.

8. Numerical Results. This section presents the results obtained through implementation of the eigenfunction method in Section 7 with the domain viewed as embedded in the range of the operator.

8.1. Experiment 1: Cylinder flow. This experiment uses the benchmark cylinder flow data set found in [14]. The dataset is a time series of fluid vorticity field snapshots for the wake behind a circular cylinder at Reynolds number $Re = 100$. The dataset comprises 151 snapshots sampled at $h = 0.02s$. To implement the developed method, the dataset was segmented into strings of adjacent snapshots of length 5 yielding 147 trajectories. The exponential dot product kernel is used with parameters selected by trial and error as $\mu_1 = 1/10000$ and $\mu_2 = 1/9999$. Numerical integration is performed using Simpson's Rule. To improve numerical stability, each Gram matrix is regularized by adding $\epsilon I_{M \times M}$ with $\epsilon = 1 \times 10^{-5}$.

Presented in Figure 1 are a selection of approximate Liouville modes obtained for this operator through the finite rank approximation determined by Section 7.

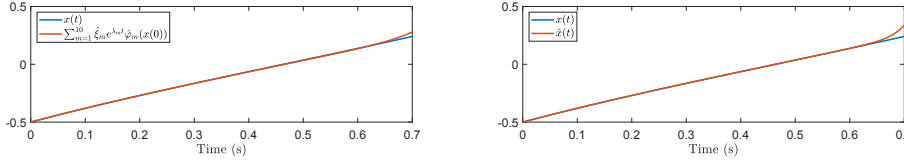


FIG. 3. The plot on the left shows true sample trajectory of the scalar system $\dot{x} = 1 + x^2$ starting from $x(0) = -0.5$ and the corresponding approximation generated using generated using (7.5). The plot on the right shows the same true sample trajectory along with the approximation generated using kernel DMD (see [26]). Parameters for the two methods are tuned manually to yield the smallest reconstruction error.

Examples of the reconstructed and original data are shown in Figure 2.

The results are compared with an implementation of the kernel DMD method developed in [26], implemented with snapshots 1 to 150 as the input matrix and snapshots 2 to 151 as the output matrix, and using the exponential dot product kernel, with parameter $\mu = 500$.

8.2. Experiment 2: Systems with finite escape time. This experiment demonstrates applicability of the developed method to systems with finite escape time. Ten trajectories of the scalar system $\dot{x} = 1 + x^2$ are generated, starting from initial conditions equally spaced between -1 and 1 over a time horizon of 0.5 s. The trajectories are sampled every $h = 0.02$ s to generate a dataset. The exponential dot product kernel is used with parameters selected by trial and error as $\mu_1 = 1/4.71$ and $\mu_2 = 1/4.7$. Numerical integration is performed using Simpson's Rule. To improve numerical stability, each Gram matrix is regularized by adding $\epsilon I_{M \times M}$ with $\epsilon = 1 \times 10^{-7}$.

The resulting model is used to generate a trajectory of the system, starting from $x(0) = -0.5$ over a longer time horizon of 0.7 s. The developed method is compared against the kernel DMD method from [26]. Kernel DMD is implemented using a dataset containing 45 snapshots of the system, corresponding to the values of x at 0 s, 0.25 s, and 0.5 s in 15 trajectories of the system starting from initial conditions equally spaced between -1 and 1 . The exponential dot product kernel is used with parameter selected by trial and error as $\mu = 1/1.7$.

Selection of parameters for the exponential dot product kernel lacks the theoretical backing that the Gaussian RBF kernels enjoy. As shown in [22], if Gaussian RBF kernels are used, then the fill distance of the data provides guidance for selection of the parameter. Both the exponential dot product kernel and the Gaussian RBF are universal kernels, however, and are can be used effectively for approximation purposes. Heuristically, the selection of the parameter for the exponential dot product kernel can be made to prevent large data points from appearing directly inside the exponential function. Such selection can be achieved, for example, by setting the parameter so that the exponent in the kernel is within $[-1, 1]$.

Figure 3 shows a representative trajectory of the system starting from $x(0) = -0.5$, reconstructed using the developed method. Figure 4 shows the error between approximation of the vector field $1 + x^2$, generated using (7.6), and the true vector field.

9. Discussion. The methods presented in this manuscript give two algorithms for performing a DMD. Together with the compactness of the Liouville operators, the

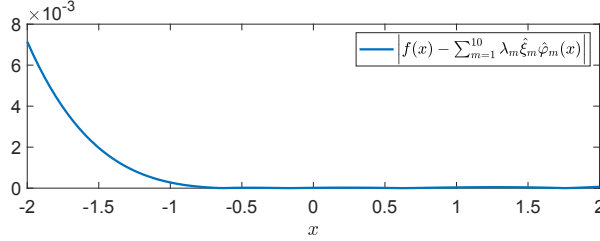


FIG. 4. The error between the vector field of the scalar system $\dot{x} = 1 + x^2$ and the corresponding approximation generated using generated using (7.6).

singular DMD approach guarantees the existence of dynamic modes and convergence through singular value decomposition of compact operators. Singular DMD is a general purpose approach to performing a DMD for when the domain and range of the operators disagree. The major drawback of this approach is that even though it can guarantee the existence of dynamic modes, which cannot be done for eigenfunction methods, the reconstruction involves the solution of an initial value problem, which is technically more involved than the eigenfunction approach.

The second method adds an additional assumption to the problem, where the domain is assumed to be embedded in the range of the operator. These embeddings frequently occur in the study of RKHSs, where the adjustment of a parameter loosens the requirement on functions within that space. It was demonstrated that this embedding may be established for the exponential dot product kernel, and it also holds for the native spaces of Gaussian RBFs with differing parameters.

Convergence of these routines follow the proof found in [19], which is a general purpose approach for showing convergence of operator level interpolants to the compact operators they are approximating. In particular, given an infinite collection of trajectories for a dynamical system, if the span of the occupation kernels form a dense subset of their respective Hilbert spaces, convergence of the overall algorithm is achieved.

The density of the occupation kernels corresponding to trajectories are easily established for Lipschitz continuous dynamics. This follows since, given any initial point, x_0 in $\mathbb{R}^n$, there is a T_0 such that the trajectory starting at x_0 , γ_{x_0} , exists over the interval $[0, T_0]$. Consider the sequence of occupation kernels indexed by $\delta \in [0, T_0]$, $\Gamma_{\gamma_{x_0}, \delta}(x) := \int_0^\delta K(x, \gamma_{x_0}(t)) dt$. Then $\frac{1}{\delta} \Gamma_{\gamma_{x_0}, \delta} \rightarrow K(x, x_0)$ in the Hilbert space norm. Hence, as x_0 was arbitrary, every kernel may be approximated by an occupation kernel corresponding to a trajectory, and since kernels are dense in H , so are these occupation kernels. Finally, if H and $\hat{H}$ are spaces of real analytic functions, the dynamics must also be real analytic by the same proof found in [20]. Spaces of real analytic functions include the Gaussian RBF and the exponential dot product kernel space.

Simulation results demonstrate the efficacy of the developed method. As shown in Figure 2, the developed model reproduces the vorticity field of the flow past a cylinder. A 3 seconds long trajectory is reproduced from a dataset that uses 147 trajectories, each 0.1 seconds long. The developed method can also be used to model systems with non-Lipschitz dynamics as indicated by Figure 4, where the vector field $1 + x^2$ is approximated using 10 trajectories, each 0.5 seconds long. While the approximation results in Figures 3 and 4 indicate divergence of the reproduced trajectory and vector

field from the true trajectory and vector field for large t and large x , respectively, such divergence is to be expected when a numerical approximation technique is used to approximate non-Lipschitz models.

Figure 2 indicates that the performance of singular DMD, when applied to the cylinder flow problem, is slightly worse than the kernel DMD method in [26]. On the other hand, as indicated by Figure 3, when applied to a system that exhibits finite escape time, the developed method marginally outperforms kernel DMD. It should be noted that the performance of both methods is sensitive to the selection of kernels, kernel parameters, and hyperparameters such as sample times, number of trajectories, and lengths of trajectories. In the numerical experiments above, the parameters and the hyperparameters were manually selected to minimize reconstruction errors. As such, the comparison is not quantitatively meaningful, it only serves an illustrative purpose. While the numerical results are interesting, the authors do not expect that the developed algorithm will perform better in practice. The operator norm convergence guarantees, absent from results such as [13, 19, 26] and related literature on Koopman-based DMD, and well-posedness of the Liouville operator for systems with finite escape time, where Koopman operators are not well-defined, are the key contributions of this work.

One interesting result of the structure of the finite rank approximation given in Section 7 is that as $\mu_1 \rightarrow \mu_2$, the first two matrices cancel. The matrix computations then approach the computations in [19]. Hence, for close enough μ_1 and μ_2 the computations are computationally indistinguishable from [19] over a fixed compact set containing the trajectories.

Finally, it should be noted that this methodology is not restricted to spaces of analytic functions, but rather it can work for a large collection of pairs of spaces. As a rule, the range space should be less restrictive as to the collection of functions in that space than the domain space. With this in mind, for many of the cases where compact Liouville operators may be established, the domain will embed into the range. The complications arise in computing the first matrix in (7.1), where the inner product of the occupation kernels for the domain are computed in the range space. Hence, the explicit description for spaces of real analytic functions help resolve that computation.

10. Conclusion. This paper presented a theoretical and algorithmic framework that achieves many long standing goals of DMD. To wit, by selecting differing domains and ranges for the Liouville operators (sometimes Koopman generators), the resulting operators are compact. This comes at the sacrifice of eigenfunctions when the domain is not embedded in the range of the operator, but achieves well defined dynamic modes and convergence. Reconstruction can then be determined using typical numerical methods for initial value problems. However, in the case of an embedding between the spaces, an algorithm may be established to determine approximate eigenfunctions for the operators, resulting in a more typical DMD routine that also converges.

REFERENCES

- [1] N. ARONSZAJN, *Theory of reproducing kernels*, Transactions of the American mathematical society, 68 (1950), pp. 337–404.
- [2] C. BÉNÉTEAU, M. C. FLEEMAN, D. S. KHAVINSON, D. SECO, AND A. A. SOLA, *Remarks on inner functions and optimal approximants*, Canadian Mathematical Bulletin, 61 (2018), pp. 704–716.
- [3] S. L. BRUNTON AND J. N. KUTZ, *Data-driven science and engineering: Machine learning, dynamical systems, and control*, Cambridge University Press, 2019.
- [4] M. BUDIŠIĆ, R. MOHR, AND I. MEZIĆ, *Applied koopmanism*, Chaos: An Interdisciplinary Jour-

- nal of Nonlinear Science, 22 (2012), p. 047510.
- [5] B. CARSWELL, B. D. MACCLUER, AND A. SCHUSTER, *Composition operators on the Fock space*, Acta Scientiarum Mathematicarum, 69 (2003), pp. 871–887.
 - [6] G. CHACÓN AND J. GIMÉNEZ, *Composition operators on spaces of entire functions*, Proceedings of the American Mathematical Society, 135 (2007), pp. 2205–2218.
 - [7] D. CROMMELIN AND E. VANDEN-EIJNDEN, *Reconstruction of diffusions using spectral data from timeseries*, (2006).
 - [8] D. CROMMELIN AND E. VANDEN-EIJNDEN, *Diffusion estimation from multiscale data by operator eigenpairs*, Multiscale Modeling & Simulation, 9 (2011), pp. 1588–1623.
 - [9] S. DAS, D. GIANNAKIS, AND J. SLAWINSKA, *Reproducing kernel Hilbert space compactification of unitary evolution groups*, Applied and Computational Harmonic Analysis, 54 (2021), pp. 75–136.
 - [10] G. FROYLAND, G. A. GOTTWALD, AND A. HAMMERLINDL, *A trajectory-free framework for analysing multiscale systems*, Physica D: Nonlinear Phenomena, 328 (2016), pp. 34–43.
 - [11] E. GOBET, M. HOFFMANN, AND M. REISS, *Nonparametric estimation of scalar diffusions based on low frequency data*, (2004).
 - [12] E. GONZALEZ, M. ABUDIA, M. JURY, R. KAMALAPURKAR, AND J. A. ROSENFELD, *The kernel perspective on dynamic mode decomposition*. arXiv:2106.00106, 2021.
 - [13] M. KORDA AND I. MEZIĆ, *On convergence of extended dynamic mode decomposition to the Koopman operator*, Journal of Nonlinear Science, 28 (2018), pp. 687–710.
 - [14] J. N. KUTZ, S. L. BRUNTON, B. W. BRUNTON, AND J. L. PROCTOR, *Dynamic mode decomposition: data-driven modeling of complex systems*, SIAM, 2016.
 - [15] A. MAUROY AND I. MEZIĆ, *Global stability analysis using the eigenfunctions of the Koopman operator*, IEEE Transactions on Automatic Control, 61 (2016), pp. 3356–3369.
 - [16] I. MEZIĆ, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dynamics, 41 (2005), pp. 309–325.
 - [17] G. K. PEDERSEN, *Analysis now*, vol. 118, Springer Science & Business Media, 2012.
 - [18] F. RIESZ, *Über die Randwerte einer analytischen Funktion*, Mathematische Zeitschrift, 18 (1923), pp. 87–95.
 - [19] J. A. ROSENFELD, R. KAMALAPURKAR, L. F. GRUSS, AND T. T. JOHNSON, *Dynamic mode decomposition for continuous time systems with the Liouville operator*, Journal of Nonlinear Science, 32 (2022), pp. 1–30.
 - [20] J. A. ROSENFELD, R. KAMALAPURKAR, B. RUSSO, AND T. T. JOHNSON, *Occupation kernels and densely defined Liouville operators for system identification*, in Proceedings of the IEEE Conference on Decision and Control, 2019, pp. 6455–6460.
 - [21] J. A. ROSENFELD, B. RUSSO, R. KAMALAPURKAR, AND T. T. JOHNSON, *The occupation kernel method for nonlinear system identification*. arXiv:1909.11792, 2019.
 - [22] B. P. RUSSO, R. KAMALAPURKAR, D. CHANG, AND J. A. ROSENFELD, *Motion tomography via occupation kernels*, Journal of Computational Dynamics, 9 (2022), pp. 27–45.
 - [23] R. K. SINGH AND A. KUMAR, *Compact composition operators*, Journal of the Australian Mathematical Society, 28 (1979), pp. 309–314.
 - [24] I. STEINWART AND A. CHRISTMANN, *Support vector machines*, Springer Science & Business Media, 2008.
 - [25] H. WENDLAND, *Scattered data approximation*, vol. 17, Cambridge university press, 2004.
 - [26] M. O. WILLIAMS, C. W. ROWLEY, AND I. G. KEVREKIDIS, *A kernel-based method for data-driven koopman spectral analysis*, Journal of Computational Dynamics, 2 (2015), p. 247.
 - [27] K. ZHU, *Analysis on Fock spaces*, vol. 263, Springer Science & Business Media, 2012.