

Research Article

Applied Koopman Theory for Partial Differential Equations and Data-Driven Modeling of Spatio-Temporal Systems

J. Nathan Kutz¹, J. L. Proctor², and S. L. Brunton³

¹Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

²Institute for Disease Modeling, 3150 139th Ave SE, Bellevue, WA 98005, USA

³Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, USA

Correspondence should be addressed to J. Nathan Kutz; kutz@uw.edu

Received 1 June 2018; Accepted 27 August 2018; Published 2 December 2018

Academic Editor: Qingdu Li

Copyright © 2018 J. Nathan Kutz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the application of Koopman theory to nonlinear partial differential equations and data-driven spatio-temporal systems. We demonstrate that the observables chosen for constructing the Koopman operator are critical for enabling an accurate approximation to the nonlinear dynamics. If such observables can be found, then the dynamic mode decomposition (DMD) algorithm can be enacted to compute a finite-dimensional approximation of the Koopman operator, including its eigenfunctions, eigenvalues, and Koopman modes. We demonstrate simple rules of thumb for selecting a parsimonious set of observables that can greatly improve the approximation of the Koopman operator. Further, we show that the clear goal in selecting observables is to place the DMD eigenvalues on the imaginary axis, thus giving an objective function for observable selection. Judiciously chosen observables lead to physically interpretable spatio-temporal features of the complex system under consideration and provide a connection to manifold learning methods. Our method provides a valuable intermediate, yet interpretable, approximation to the Koopman operator that lies between the DMD method and the computationally intensive extended DMD (EDMD). We demonstrate the impact of observable selection, including kernel methods, and construction of the Koopman operator on several canonical nonlinear PDEs: Burgers' equation, the nonlinear Schrödinger equation, the cubic-quintic Ginzburg-Landau equation, and a reaction-diffusion system. These examples serve to highlight the most pressing and critical challenge of Koopman theory: a principled way to select appropriate observables.

1. Introduction

Data-driven mathematical methods are increasingly important for characterizing complex systems across the physical, engineering, social, and biological sciences. These methods are aimed at discovering and exploiting a relatively small subset of the full space where low-dimensional models can be used to describe the evolution of the system. Thus, solutions can often be approximated through dimensionality reduction methods where if n is the dimension of the original high-dimensional system and r is the dimension of the subspace (or slow-manifold) where the dynamics is embedded, then $r \ll n$. The reduced order modeling (ROM) community has used this to great effect in applications such as large-scale patterns of atmospheric variability [1], turbulent flow

control architectures [2], and/or spatio-temporal encodings in neurosensory systems [3]. Traditionally, the large-scale dynamics may be embedded in the low-dimensional space using, for instance, the proper orthogonal decomposition (POD) in conjunction with Galerkin projection. More recently, the *Dynamic Mode Decomposition* (DMD) and its Koopman generalization have garnered attention due to the fact that they can (i) discover low-rank spatio-temporal patterns of activity and (ii) they can embed the dynamics in the subspace in an equation-free manner, unlike the Galerkin-POD method of ROMs. In this manuscript, we demonstrate that the judicious, and parsimonious, selection of observables for the Koopman architecture can yield accurate low-dimensional embedding for nonlinear partial differential equations (PDEs) while keeping computational

costs down and avoiding costly cross-validation. Critical to its success is an appropriate choice of observables, which is demonstrated to act as a nonlinear manifold learning method. We demonstrate the success of the method, and compare it to traditional DMD, on several canonical PDE models: Burgers' equation, the nonlinear Schrödinger equation, the cubic-quintic Ginzburg-Landau equation, and a $\lambda - \omega$ reaction-diffusion system.

Historically, the DMD method originated in the fluid dynamics community as a principled technique to decompose complex flows into a simple representation based on low-rank spatio-temporal coherent structures. Schmid [4] first defined the DMD algorithm and demonstrated its ability to provide physically interpretable insights from high-dimensional fluid data. The growing success of DMD stems from the fact that it is an *equation-free* data-driven method capable of providing an accurate decomposition of a complex system into spatio-temporal coherent structures that may be used for diagnostic analysis, short-time future state prediction, and control. Importantly, Rowley et al. [5] showed that DMD is connected to the underlying nonlinear dynamics through Koopman operator theory [6] and is readily interpretable using standard dynamical system techniques [7–10]. Specifically, the DMD algorithm is a manifestation of Koopman theory when the observable functions are the identity or a linear transformation of the underlying state space. Thus, DMD is a principled algorithmic architecture allowing for an explicit approximation of the Koopman operator. For more details, there are numerous detailed references [5, 11, 12].

The approximation of the Koopman operator via DMD is critically important for enabling evaluation of the operator from data. Indeed, it transforms Koopman theory from an abstract mathematical conception to a readily tractable computation. It also highlights the important role played by observables and their associated evolution manifolds. In particular, nonlinear PDEs can be thought to evolve on manifolds which are often difficult to characterize and are rarely known analytically. A correct choice of observables can, in some cases, *linearize* the nonlinear manifold. For instance, the nonlinear evolution governed by Burgers' PDE equation can be linearized by the Cole-Hopf transformation, thus providing a linear manifold which can trivially describe the evolution dynamics. Such exact solutions to nonlinear PDEs are extremely rare and do not often exist in practice, with the inverse scattering transform (IST) for Korteweg-deVries, nonlinear Schrödinger, and other integrable PDEs being the notable exceptions [13]. Regardless, judiciously chosen observables can help transform a PDE evolving on a strongly nonlinear manifold to a weakly nonlinear manifold, enabling a more accurate and broader range of applicability of the Koopman approximation.

The selection of appropriate observables remains one of the most important and open challenges for Koopman theory. The widely used DMD algorithm simply takes the state space variable as the observable. This provides the simplest approximation to the Koopman operator. An alternative approach has been advocated by Williams et al. and Kevrekidis et al. [14, 15] whereby machine learning concepts

(e.g., support vector machines (SVM)) are used to project the data to a large number of variables using the so-called extended DMD and kernel DMD (EDMD) methods. Thus, the DMD approximation is computed in a large set of nonlinear observables of the original data. Recently, it has been shown that the EDMD method is equivalent to the *variational approach of conformation dynamics* (VAC) [16, 17], which was first derived by Noé and Nüske in 2013 for simulating slow processes in stochastic dynamical systems applied to molecular kinetics [18, 19]. The authors further show that *time-lagged independent component analysis* (TICA) [20, 21], which was originally developed in 1994, is closely related to the DMD algorithm [17]. Regardless of the EDMD/VAC strategy, it is well known in the machine learning literature, such projections into higher dimensional space, through SVM or deep neural nets, can lead to improved predictions at the cost of loss of interpretability. It also projects to variables that may have no natural association with the underlying physics or nonlinear manifold of the dynamics being considered. Importantly, Klus et al. [17] show that the EDMD/VAC method requires a principled cross-validation strategy in order to make the technique useful.

Our approach is aimed at improving the straightforward DMD approximation by adding a parsimonious set of judiciously chosen variables which are motivated by the governing equations, i.e., it is a version of EDMD with only a few extra variables. Thus, simple choices of nonlinear observables can greatly improve the Koopman approximation. Moreover, we show that a clear goal in selecting observables is to move DMD eigenvalues onto the imaginary axis. We show that selecting a parsimonious set of observables allows us to capitalize on the EDMD architecture while only incurring a marginal increase in computational costs and avoiding costly cross-validation for computing an improved approximation to the Koopman operator. Further, the judicious choice of observables can also be used to help understand the nonlinear manifold on which the dynamics evolve. Ultimately, this provides a valuable third option for variable selection that sits between the standard application of the DMD and EDMD methods.

2. Koopman Theory, Observables, and Dynamic Mode Decomposition

The original work of Koopman in 1931 [6] considered Hamiltonian systems and formulated the Koopman operator as a discrete-time mapping. In the following year, Koopman and von Neumann extended these results to dynamical systems with continuous spectra [22]. Critical to implementing this definition numerically is understanding how to choose a finite set of observables $g(x)$. This remains an open challenge today and will be addressed in our PDE examples.

By construction, the Koopman operator is a *linear* infinite-dimensional operator that acts on the Hilbert space \mathcal{H} of *all* scalar measurement functions g . The Koopman operator acts on functions of the state space of the dynamical system, trading nonlinear finite-dimensional dynamics for

linear infinite-dimensional dynamics. It can be further generalized to map infinite-dimensional nonlinear dynamics to infinite-dimensional linear dynamics by appropriate choice of observables. In practice, the computation of the Koopman operator will require a finite-dimensional representation. The advantage of the Koopman representation is compelling: linear problems can be solved using standard linear operator theory and spectral decompositions. With such methods, the infinite dimensional representation is handled by considering a sufficiently large, but finite, sum of modes to approximate the Koopman spectral solution.

The Koopman operator is defined for discrete-time dynamical systems. A continuous dynamical system will induce a discrete-time dynamical system given by the flow map $\mathbf{F}_t : \mathcal{M} \rightarrow \mathcal{M}$, which maps the state $\mathbf{x}(t_0)$ to a future time $\mathbf{x}(t_0 + t)$:

$$\mathbf{F}_t(\mathbf{x}(t_0)) = \mathbf{x}(t_0 + t) = \mathbf{x}(t_0) + \int_{t_0}^{t_0+t} \mathbf{f}(\mathbf{x}(\tau)) d\tau. \quad (1)$$

This induces the discrete-time dynamical system

$$\mathbf{x}_{k+1} = \mathbf{F}_t(\mathbf{x}_k), \quad (2)$$

where $\mathbf{x}_k = \mathbf{x}(kt)$. The analogous discrete-time Koopman operator is given by \mathcal{K}_t such that $\mathcal{K}_t g = g \circ \mathbf{F}_t$. Thus, the Koopman operator sets up a discrete-time dynamical system on the observable function g :

$$\mathcal{K}_t g(\mathbf{x}_k) = g(\mathbf{F}_t(\mathbf{x}_k)) = g(\mathbf{x}_{k+1}). \quad (3)$$

If an appropriate Koopman operator can be constructed, then linear operator theory provides the spectral decomposition required to represent the dynamical solutions of interest. Specifically, the eigenfunctions and eigenvalues of the Koopman operator \mathcal{K} give a complete characterization of the dynamics. We consider the eigenvalue problem

$$\mathcal{K} \varphi_k = \lambda_k \varphi_k. \quad (4)$$

The functions $\varphi_k(\mathbf{x})$ are Koopman eigenfunctions, and they define a set of intrinsic measurement coordinates, on which it is possible to advance these measurements with a *linear* dynamical system. The low-dimensional embedding of the dynamics is ultimately extracted from the Koopman eigenfunctions. More precisely, a reduced-order linear model can be constructed by a rank- r truncation of the dominant eigenfunctions φ_k .

A vector of observables \mathbf{g} , which is in our new *measurement space*, may be expressed in terms of Koopman eigenfunctions φ as

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_p(\mathbf{x}) \end{bmatrix} = \sum_{k=1}^{\infty} \varphi_k(\mathbf{x}) \mathbf{v}_k, \quad (5)$$

where \mathbf{v}_k is the k th Koopman mode associated with the k th Koopman eigenfunction φ_k , i.e., it is the weighting of each observable on the eigenfunction. In the original theory [6], Koopman considered Hamiltonian flows that are measure preserving, so that the Koopman operator is unitary. In this case, the eigenfunctions are all orthonormal, and (5) may be written explicitly as

$$\mathbf{g}(\mathbf{x}) = \sum_{k=1}^{\infty} \varphi_k(\mathbf{x}) \begin{bmatrix} \langle \varphi_k, g_1 \rangle \\ \langle \varphi_k, g_2 \rangle \\ \vdots \\ \langle \varphi_k, g_p \rangle \end{bmatrix} = \sum_{k=1}^{\infty} \varphi_k(\mathbf{x}) \mathbf{v}_k. \quad (6)$$

The dynamic mode decomposition algorithm is used to compute an approximation to the Koopman eigenvalues λ_k and modes \mathbf{v}_k .

The nonlinear dynamical system defined by f and the infinite-dimensional linear dynamics defined by \mathcal{K} are equivalent representations of a dynamical system. One can either evolve the system in the original state space, requiring computational effort since it is nonlinear, or one can instead evolve using (5) so that the time dynamics are trivially computed

$$\mathcal{K} g(\mathbf{x}) = \mathcal{K} \sum_{k=1}^{\infty} \varphi_k(\mathbf{x}) \mathbf{v}_k = \sum_{k=1}^{\infty} \mathcal{K} \varphi_k(\mathbf{x}) \mathbf{v}_k = \sum_{k=1}^{\infty} \lambda_k \varphi_k(\mathbf{x}) \mathbf{v}_k. \quad (7)$$

Thus, future solutions can be computed by simple multiplication with the Koopman eigenvalue. Such a mathematical strategy for evolving nonlinear dynamical systems would always seem to be advantageous. However, it remains an open challenge how to systematically link the observables g and the associated Koopman mode expansion to the original evolution defined by f . For a limited class of nonlinear dynamics, this can be done explicitly [23].

In theory, the modification of the Koopman operator to PDEs would generate eigenfunctionals of the Koopman operator. In practice, the discretization of space and time, either in experiment or simulation, yields a high-dimensional system of ODEs. The PDE itself imposes clear relations between the high-dimensional data which correspond to spatial locations. Specifically, the data generated from the PDE most certainly inherit the underlying dynamics enforced by spatial relations and their spatial derivatives, leading to dimensionality-reduction and low-rank truncation possibilities. In our examples, the spatio-temporal patterns can be represented in POD/DMD modes with a truncation using the dominant r modes. This truncation to a low-dimensional space is a direct consequence of the PDE nature of the solutions. For a thorough discussion of the difference simply between high-dimensional ODEs and PDEs, please see [24].

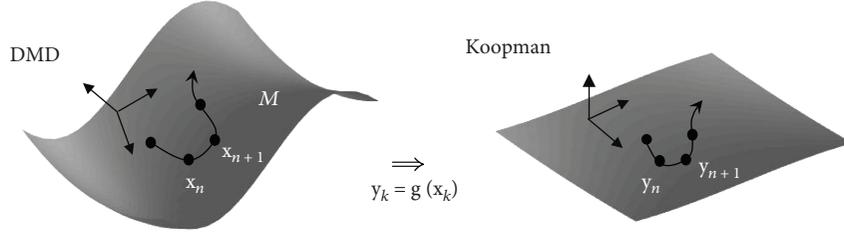


FIGURE 1: The left panel illustrates the nonlinear manifold on which the dynamical system defined by F_t in (2) generates a solution. DMD approximates the evolution on this manifold by a least-square fit linear dynamical system. In contrast, the selection of appropriate observables $g(x)$ define a Koopman operator that helps *linearize* the manifold so that a least-square fit linear dynamical system provides a much better approximation to the system (right panel).

Figure 1 illustrates the underlying concept in the Koopman approach. A dynamical system consisting of snapshots x_k evolves according to the nonlinear dynamical system defined by F_t in (2). In the state space, the nonlinearity generates a nonlinear manifold in which the data are embedded. The DMD approximation produces a least-square fit *linear dynamical system* approximating the flow map and the low-dimensional embedding (left panel of Figure 1). Koopman theory ideally defines an operator that attempts to *linearize* the space in which the data are embedded. The Koopman operator then produces a linear flow map and low-dimensional embedding that approximates the full nonlinear dynamics (right panel of Figure 1).

The nonlinear manifold on which the dynamics evolve can change due to parameter changes in the PDE. For instance, the dynamics can undergo bifurcation and generate a new nonlinear manifold. This requires building a new Koopman operator to characterize the dynamics. The Koopman embedding can be used to build libraries of low-rank representations for the dynamics. The concept of library building of low-rank “features” from data is well established in the computer science community. In the reduced-order modeling community, it has recently become an issue of intense investigation. Indeed, a variety of recent works have produced libraries of ROM models that can be selected and/or interpolated through measurement and classification [25–31]. Alternatively, cluster-based reduced order models use a k-means clustering to build a Markov transition model between dynamical states [32]. More recently, such techniques have been applied using the DMD approximation for the Koopman operator [33]. The modeling of parametric systems remains an open challenge for model reduction frameworks.

3. The DMD and Koopman Algorithms

The DMD algorithm underlies the computation of the Koopman eigenvalues and modes directly from data. Its effectiveness depends sensitively on the choice of observables. Rowley et al. [5] showed that DMD approximates the Koopman operator for the set of observables $g(x) = x$. We will use this fact in constructing a DMD algorithm for observables of x instead of the state variable itself. To start, we use the following definition of the DMD decomposition [11]:

Definition 1. Dynamic mode decomposition [11]: suppose we have a dynamical system (2) and two sets of data

$$X = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & \cdots & | \end{bmatrix}, \quad (8a)$$

$$X' = \begin{bmatrix} | & | & \cdots & | \\ x'_1 & x'_2 & \cdots & x'_m \\ | & | & \cdots & | \end{bmatrix}, \quad (8b)$$

with x_k an initial condition to (2) and x'_k its corresponding output after some prescribed evolution time Δt with there being m initial conditions considered. The DMD modes are eigenvectors of

$$A_X = X'X^\dagger, \quad (9)$$

where \dagger denotes the Moore-Penrose pseudoinverse.

The above definition provides a computational method for evaluating the Koopman operator for a linear observable. In practice, three practical constraints must be considered: (i) We have data X and X' , but we do not necessarily know $F_t(\cdot)$, (ii) We will have to make a finite-dimensional approximation to the infinite-dimensional Koopman operator \mathcal{K} , and (iii) We will have to judiciously select the observables $g(x)$ in order to have confidence that that Koopman operator will approximate the nonlinear dynamics of $F_t(\cdot)$. Points (i) and (ii) go naturally together. Specifically, the number of measurements in each column of X and X' is n , while the number of total columns (time measurements) is m . Thus, finite-dimensionality is imposed simply from the data collection limitations. The dimension can be increased with a large set of observables, or it can be decreased via a low-rank truncation during the DMD process. The observables are more difficult to deal with in a principled way. Indeed, a good choice of observables can make the method extremely effective, but it would also require expert knowledge of the system at hand [23]. This will be discussed further in the examples.

The following gives a practical demonstration of how to use the data, the DMD algorithm, and the observables to produce a Koopman operator and a future state prediction of the nonlinear evolution. The Koopman algorithm simply applies DMD on the space of observables.

- (1) From the data matrices \mathbf{X} and \mathbf{X}' , create the data matrices of observables \mathbf{Y} and \mathbf{Y}' :

$$\mathbf{Y} = \begin{bmatrix} | & | & & | \\ \mathbf{g}(x_1) & \mathbf{g}(x_2) & \cdots & \mathbf{g}(x_{m-1}) \\ | & | & & | \end{bmatrix}, \quad (10a)$$

$$\mathbf{Y}' = \begin{bmatrix} | & | & & | \\ \mathbf{g}(x'_1) & \mathbf{g}(x'_2) & \cdots & \mathbf{g}(x'_{m-1}) \\ | & | & & | \end{bmatrix}, \quad (10b)$$

where each column is given by $\mathbf{y}_k = \mathbf{g}(\mathbf{x}_k)$ or $\mathbf{y}'_k = \mathbf{g}(\mathbf{x}'_k)$

- (2) Perform the DMD algorithm on the pair \mathbf{Y} and \mathbf{Y}' to compute

$$\mathbf{A}_Y = \mathbf{Y}'\mathbf{Y}^\dagger \quad (11)$$

along with the low-rank counterpart $\tilde{\mathbf{A}}_Y$ obtained by projection onto a truncated POD subspace. The eigenvalues and eigenvectors of \mathbf{A}_Y may approximate Koopman eigenvalues and modes if the observables are well chosen

- (3) DMD can be used to compute the augmented modes Φ_Y , which may approximate the Koopman modes, by

$$\Phi_Y = \mathbf{Y}'\mathbf{V}\Sigma^{-1}\mathbf{W}, \quad (12)$$

where \mathbf{W} comes from the eigenvalue problem $\tilde{\mathbf{A}}_Y\mathbf{W} = \mathbf{W}\Lambda$ and $\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}^*$. Note that an r -rank truncation of the SVD is performed at this stage

- (4) The future state in the space of observables is given by the linear evolution

$$\mathbf{y}(t) = \Phi_Y \text{diag}(\exp(\omega t)) \mathbf{b}, \quad (13)$$

where $\mathbf{b} = \Phi_Y^\dagger \mathbf{y}_1$ is determined by projecting back to the initial data observable. The continuous-time eigenvalues ω are obtained from the discrete-time eigenvalues λ_k (i.e., diagonal elements of the matrix Λ) where $\omega_k = \ln(\lambda_k)/\Delta t$

- (5) Transform from observables to state space

$$\mathbf{y}_k = \mathbf{g}(\mathbf{x}_k) \longrightarrow x_k = \mathbf{g}^{-1}(\mathbf{y}_k). \quad (14)$$

This last step is trivial if one of the observables selected to comprise $\mathbf{g}(\mathbf{x}_k)$ is the state variable x_k itself. If only nonlinear observables of x_k are chosen, then the inversion process can be difficult.

This process shows that the DMD algorithm is closely related to the Koopman operator. Indeed, it is the foundational piece for practical evaluation of the finite-dimensional Koopman operator. It is stressed once again here: selection of appropriate observables is critical for the algorithm to generate good reconstructions and approximations to the future state. We can also now introduce the following theorem [5, 11, 14, 34].

Theorem 1. *Koopman and dynamic mode decomposition: let φ_k be an eigenfunction of \mathcal{X} with eigenvalue λ_k , and suppose $\varphi_k \in \text{span}\{g_j\}$, so that*

$$\varphi_k(\mathbf{x}) = w_1 g_1(\mathbf{x}) + w_2 g_2(\mathbf{x}) + \cdots + w_p g_p(\mathbf{x}) = \mathbf{w} \cdot \mathbf{g} \quad (15)$$

for some $\mathbf{w} = [w_1 w_2 \cdots w_p]^T \in \mathbb{C}^p$. If $\mathbf{w} \in R(\mathbf{Y})$, where R is the range, then \mathbf{w} is a left eigenvector of \mathbf{A}_Y with eigenvalue λ_k so that $\tilde{\mathbf{w}}^* \mathbf{A}_Y = \lambda_k \tilde{\mathbf{w}}^*$.

Note here that the observables $g_j(x)$ as introduced in the theorem [5, 11, 14, 34] are not denoted as vectors since the theorem applies to functions. However, in practice, when a system is discretized, then $g_j(x) \longrightarrow g_j(x)$ as is explicitly constructed in the algorithm above. Thus, the Koopman eigenvalues are the DMD eigenvalues provided (i) the set of observables is sufficiently large so that $\varphi_k(x) \in \text{span}\{g_j\}$ and (ii) the data is sufficiently *rich* so that $w \in R(X)$. This directly shows that the choice of observables is critical in allowing one to connect DMD theory to Koopman spectral analysis. If this can be done, then one can simply take data snapshots of a finite-dimensional nonlinear dynamical system in time and reparameterize it as a linear system in the observable coordinates, which is amenable to a simple eigenfunction (spectral) decomposition. This representation diagonalizes the dynamics and shows that the time evolution of each eigenfunction corresponds to multiplication by its corresponding eigenvalue.

4. Koopman Observables and Kernel Methods

The effectiveness of Koopman theory hinges on one thing: selecting appropriate observables. Once observables are selected, the previous section defines a DMD-based algorithm for computing the Koopman operator whose spectral decomposition completely characterizes the approximation. In the machine learning literature, observables are often the basis of generating *features*, and we will build upon this concept to generate appropriate observables. An important practical consideration becomes the computational cost in generating the DMD approximation as the number of rows in the matrices \mathbf{Y} and \mathbf{Y}' gets progressively larger with each additional observable.

To be more precise about the distinction between the EDMD/VAC and the present work, we emphasize that the state-of-the-art machine learning approach to EDMD is given by Williams et al. [14] and Klus et al. [17]. These works provide a framework for projection to observables through kernel SVM-like methods, as well as showing how to cross-validate the results. As is common in such schemes, interpretability and generalizability is typically lost. In the current work, our goal is to select a parsimonious set of observables based upon knowledge of the physics or its constraints. It provides a viable strategy whereby expert knowledge can be leveraged to guide the selection of observables. In the previous works [14, 17], parsimony was not used in the regression framework. Here, parsimony to produce interpretable models, and potentially generalizable models, is the critical innovation advocated. It should be noted that we illustrate some of the kernel methods here in order to simply show the lack of robustness of EDMD in the absence of parameter tuning.

In the absence of expert-in-the-loop knowledge of the dynamical system, one might consider, for instance, the support vector machine (SVM) literature and associated kernel methods [35–38] for feature selection (observables). The SVM architecture suggests a number of techniques for constructing the feature space $g(x)$, with a common choice being the set of polynomials such that

$$g_j(x) = \{x, x^2, x^3, x^4, \dots, x^p\}. \quad (16)$$

Using a large number of polynomials can generate an extremely large vector of observables for each snapshot in time. This is closely related to the Carleman linearization technique in dynamical systems [39–41]. Alternatively, kernel methods have found a high degree of success using (i) radial basis functions, typically for problems defined on irregular domains, (ii) Hermite polynomials for problems defined on \mathbb{R}^n , and (iii) discontinuous spectral elements for large problems with block diagonal structures. Regardless of the specific choice of feature space, the goal is to choose a sufficiently rich and diverse set of observables that allow an accurate approximation of the Koopman operator \mathcal{K} . Instead of choosing the correct observables, one then simply chooses a large set of candidate observables with the expectation that a sufficiently diverse set will include enough features for an accurate reconstruction of the Koopman modes, eigenfunctions, and eigenvalues, which intrinsically characterize the nonlinear dynamical system.

Williams et al. and Kevrekidis et al. [14, 15] have recently capitalized on the ideas of machine learning by implementing the so-called extended DMD and kernel DMD method on extended observables (16) within the DMD architecture. Moreover, they have developed an efficient way to compute $\tilde{\mathbf{A}}_Y$ even for a large observable space. The kernel DMD method is the most relevant in practice as the number of observables (features) can rapidly grow so as to make n extremely high-dimensional. In the context of the Koopman operator, the kernel trick [35–38] will define a function $h(x, x')$ that can be related to the observables $g_j(x)$ used

for constructing \mathbf{Y} and \mathbf{Y}' . Consider the simple example of a quadratic polynomial kernel

$$h(x, x') = (1 + x^T x')^2, \quad (17)$$

where x and x' are data points in \mathbb{R}^2 . When expanded, the kernel function takes the form

$$\begin{aligned} h(\mathbf{x}, \mathbf{x}') &= (1 + x_1 x_1' + x_2 x_2')^2 \\ &= (1 + 2x_1 x_1' + 2x_2 x_2' + 2x_1 x_2 x_1' x_2' + x_1^2 x_1'^2 + x_2^2 x_2'^2) \\ &= 1 + \mathbf{Y}^T(\mathbf{x}') \mathbf{Y}(\mathbf{x}), \end{aligned} \quad (18)$$

where $\mathbf{Y}(\mathbf{x}) = [\sqrt{2}x_1 \quad \sqrt{2}x_2 \quad \sqrt{2}x_1 x_2 \quad x_1^2 \quad x_2^2]^T$. Note that for this case, both the Koopman observables and the kernel function (17) are equivalent representations that are paired together through the expansion (18). The so-called kernel trick posits that (17) is a significantly more efficient representation of the polynomial variables that emerge from the expansion (18). Instead of defining the Koopman observables $g_i(x)$, we instead define the kernel function (17) as it provides a compact representation of the infinite-dimensional feature space and an implicit computation of the inner products required for the Koopman operator.

The computational advantages of the kernel trick are considerable. For example, a polynomial kernel of degree p acting on data vectors \mathbf{x} and \mathbf{x}' in \mathbb{R}^n is given by

$$h(\mathbf{x}, \mathbf{x}') = (1 + \mathbf{x}^T \mathbf{x}')^p, \quad (19)$$

which requires a single computation of the inner product $\alpha = \mathbf{x}^T \mathbf{x}'$. This requires $\mathcal{O}(n)$ and produces $f(x, x') = (1 + \alpha)^p$ where α is a constant. The resulting computational cost for this p th degree polynomial kernel remains $\mathcal{O}(n)$. In contrast, the alternative observable space using Y requires construction of a vector of length

$$O\binom{n+p}{p} \quad (20)$$

taking the form

$$\mathbf{Y}(\mathbf{x}) = [1 \quad x_1 \quad \dots \quad x_n \quad x_1^2 \quad x_1 x_2 \quad \dots \quad x_n^2 \quad \dots \quad x_1^p \quad \dots \quad x_n^p]^T. \quad (21)$$

Computing the inner product $\mathbf{Y}^T(\mathbf{x}') \mathbf{Y}(\mathbf{x})$ is a significantly larger computation than the kernel form (19). Thus, the kernel trick enables an advantageous representation of the various polynomial terms and circumvents the formation and computation associated with (21). It should be noted that these two representations are not equivalent. Rather, they give two different representations of nonlinear observables from which a feature space can be extracted. In the former

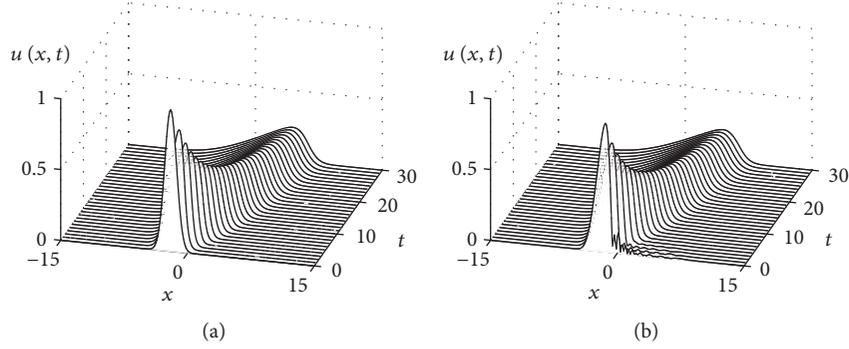


FIGURE 2: (a) Evolution dynamics of Burgers' equation with initial condition $u(x, 0) = \exp(-(x+2)^2)$. (b) Fifteen mode DMD approximation of the Burgers' evolution. The simulation of (24) was performed over $t \in [0, 30]$ where the sampling was taken at every $\Delta t = 1$. The domain was discretized with $n = 256$ points on a domain $x \in [-15, 15]$.

(19), the computations are tractable, while in the latter (21), the computation quickly becomes intractable.

The choice of kernel is important and in practice is not robust for Koopman methods. Some standard choices are often used, including the three most common kernels of SVM-based data methods:

$$\text{polynomial kernel (degree } p) h(\mathbf{x}, \mathbf{x}') = (a + \mathbf{x}^T \mathbf{x}')^p, \quad (22a)$$

$$\text{radial basis functions } h(\mathbf{x}, \mathbf{x}') = \exp(-a|\mathbf{x} - \mathbf{x}'|^2), \quad (22b)$$

$$\text{sigmoid kernel } h(\mathbf{x}, \mathbf{x}') = \tanh(\mathbf{x}^T \mathbf{x}' + a). \quad (22c)$$

The advantage of the kernel trick is quite clear, providing a compact representation of a very large feature space. For the polynomial kernel, for instance, a 20th-degree polynomial ($p=20$) using (22a) is trivial and does not compute all the inner products directly. In contrast, using our standard Koopman observables $g(x_j)$ would require one to explicitly write out all the terms generated from a 20th-degree polynomial on an n -dimensional data set, which is computationally intractable for even moderately large n . The tuning parameter a must be carefully chosen in practice for reasonable results.

In practice, the observables for \mathbf{Y} are implicitly embedded in the kernel $h(x, x')$. Specifically, we consider the observable matrix elements defined by

$$\mathbf{Y}^T \mathbf{Y}'(j, k) = h(\mathbf{x}_j, \mathbf{x}'_k), \quad (23)$$

where (j, k) denotes the j th row and k th column of the correlation matrix, and x_j and x'_k are the j th and k th columns of data. The kernel DMD formulation still requires the computation of the matrices \mathbf{V} and $\mathbf{\Sigma}$ which can be produced from $\mathbf{Y}^* \mathbf{Y} \mathbf{V} = \mathbf{\Sigma}^2 \mathbf{V}$. As before, the matrix elements of $\mathbf{Y}^* \mathbf{Y}$ are computed from $\mathbf{Y}^* \mathbf{Y}(j, k) = h(x_j, x_k)$. Thus, all the required inner products are computed by projecting directly to the new feature space defined by the specific kernel used. Note

that if the linear kernel function $h(x, y) = x^T y$ is chosen, the kernel DMD reduces to the standard DMD algorithm.

5. Application to PDEs

To demonstrate the Koopman operator concepts, we apply the methodology to various illustrative and canonical PDEs: Burgers' equation, nonlinear Schrödinger equation, the cubic-quintic Ginzburg-Landau equation, and a reaction-diffusion model. With these examples, we can (i) illustrate a scenario where the Koopman operator can exactly (analytically) linearize a dynamical system, (ii) demonstrate how to judiciously select observables, and (iii) show that kernel methods are highly sensitive as an observable selection technique.

The simulations are based upon a pseudospectral technique whereby the spatial domain and its derivatives are computed in the Fourier domain using the Fast Fourier Transform (FFT). The time-stepping algorithm is based upon an adaptive 4th-order Runge-Kutta scheme, i.e., *ode45* in MATLAB. By default, the FFT-based strategy imposes periodic boundary conditions.

5.1. Burgers' Equation. To demonstrate the construction of a specific and exact Koopman operator, we consider the canonical nonlinear PDE: Burgers' equation with diffusive regularization. The evolution, as illustrated in Figure 2(a), is governed by diffusion with a nonlinear advection [42]:

$$u_t + uu_x - \epsilon u_{xx} = 0 \quad \epsilon > 0, \quad x \in [-\infty, \infty]. \quad (24)$$

When $\epsilon = 0$, the evolution can lead to shock formation in finite time. The presence of the diffusion term regularizes the PDE, ensuring continuous solutions for all time.

Burgers' equation is one of the few nonlinear PDEs whose analytic solution form can be derived. In independent seminal contributions, Hopf [43] and Cole [44] derived a transformation that linearizes the PDE. The Cole-Hopf transformation is defined as follows

$$u = -2\epsilon v_x / v. \quad (25)$$

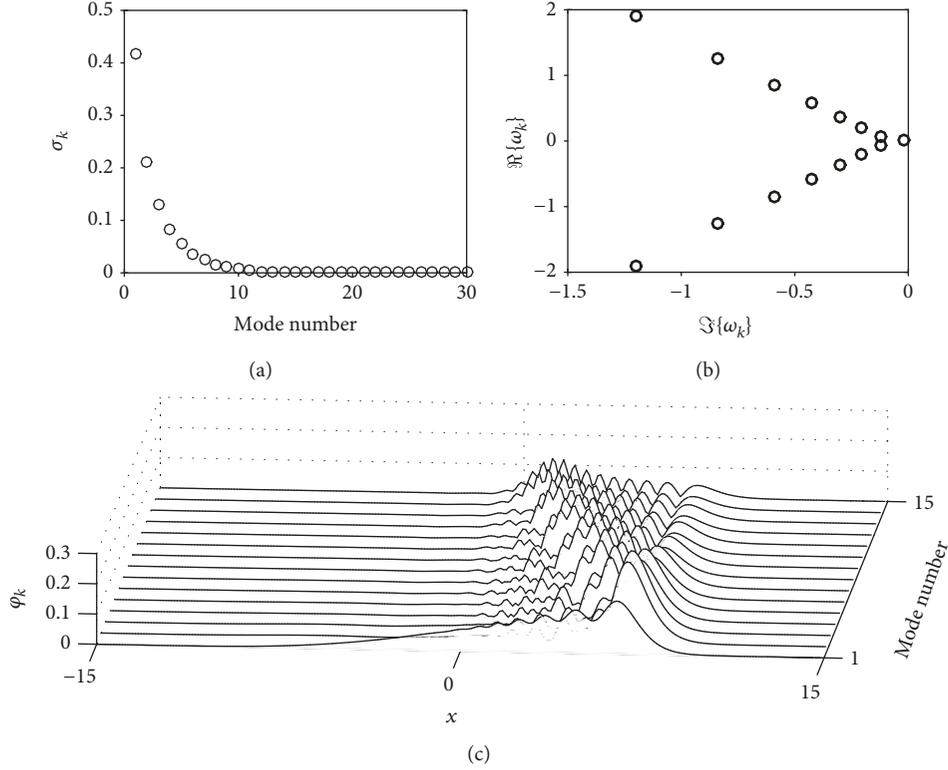


FIGURE 3: DMD of the Burgers' equation. (a) The singular value spectrum demonstrates that a rank $r = 15$ truncation should be adequate to capture the dynamics of the front propagation in Figure 2. (b) The eigendecomposition in the DMD algorithm produces a DMD spectra whose eigenvalues are decaying. (c) The DMD modes used for reconstructing the solution in Figure 2 ordered according to the smallest (in magnitude) eigenvalues. The first mode is like a *background* mode since the eigenvalue is almost zero.

The transformation to the new variable $v(x, t)$ replaces the nonlinear PDE (24) with the linear diffusion equation

$$v_t = \epsilon v_{xx}, \quad (26)$$

where it is noted that $\epsilon > 0$ in (24) in order to produce a well-posed PDE.

The diffusion equation can be easily solved using Fourier transforms. Fourier transforming in x gives the ODE system

$$\hat{v}_t = -\epsilon k^2 \hat{v}, \quad (27)$$

where $\hat{v} = \hat{v}(k, t)$ denotes the Fourier transform of $v(x, t)$ and k is the wavenumber. The solution in the Fourier domain is easily found to be

$$\hat{v} = \hat{v}_0 \exp(-\epsilon k^2 t), \quad (28)$$

where $\hat{v}_0 = \hat{v}(k, 0)$ is the Fourier transform of the initial condition $v(x, 0)$.

To construct the Koopman operator, we can then combine the transform to the variable $v(x, t)$ from (25)

$$v(x, t) = \exp \left[-\frac{\int_{-\infty}^x u(\xi, t) d\xi}{2\epsilon} \right], \quad (29)$$

with the Fourier transform to define the observables

$$g(u) = \hat{v}. \quad (30)$$

The Koopman operator is then constructed from (28) so that

$$\mathcal{K} = \exp(-\epsilon k^2 t). \quad (31)$$

This is one of the rare instances where an explicit expression for the Koopman operator and the observables can be constructed analytically. The inverse scattering transform [13] for other canonical PDEs, Korteweg-deVries (KdV) and nonlinear Schrödinger (NLS) equations, also can lead to an explicit expression for the Koopman operator, but the scattering transform and its inversion are much more difficult to construct in practice.

To make comparison between Koopman theory and DMD, we consider the DMD method applied to governing (24). Applying the algorithm of Section 3 to the observables $g(x) = x$ gives the DMD approximation to the Burgers' dynamics as shown in Figure 2(b). For this simulation, data snapshots were collected at intervals of $\Delta t = 1$ for the time range $t \in [0, 30]$. The singular value decay for the dynamics is shown in Figure 3(a), suggesting that a rank $r = 15$ truncation is appropriate. The DMD spectra and DMD modes are illustrated in Figures 3(b) and 3(c), respectively. Thus, using

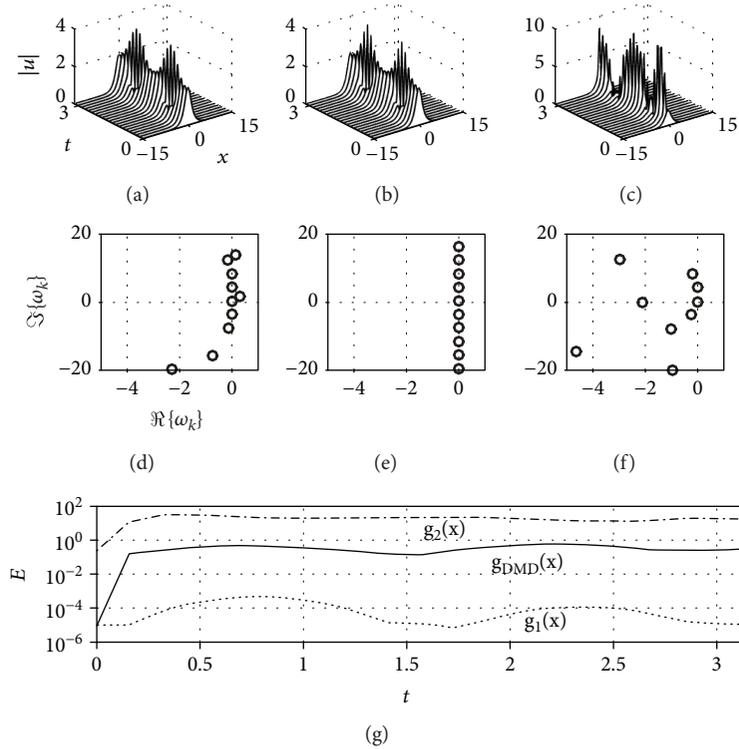


FIGURE 4: Reconstruction of the NLS dynamics using (a) a standard DMD approximation $g_{\text{DMD}}(x)$, (b) the NLS motivated $g_1(x)$, and (c) a quadratic observable $g_2(x)$. The Koopman spectra for each observable are demonstrated in (d), (e), and (f) which accompany the observables of (a), (b), and (c), respectively. Note that the observable $g_1(x)$ produces spectra which are approximately purely imaginary which is expected of the 2-soliton evolution. The error between the three observables and the full simulation is shown (g). Note that the observable $g_1(x)$ gives an error reduction of four-orders of magnitude over DMD, while $g_2(x)$ is an order of magnitude worse. This highlights the importance of selecting good observables. The simulation was performed over $t \in [0, \pi]$ with 41 equally spaced snapshots taken. The domain was discretized with $n = 512$ points on a domain $x \in [-15, 15]$.

$u(x, t)$ directly as an observable produces a low-rank model with fifteen modes. In contrast, by working with the observable (30), the Koopman operator can be trivially computed (31) and the dynamics analytically produced without need of approximation. In this case, the Koopman operator exactly linearizes the dynamics. This is the ideal which is hoped for but rarely achieved with nonlinear PDEs (or nonlinear dynamical systems in general).

5.2. Nonlinear Schrödinger Equation. The example of Burgers' equation was easy to quantify and understand since the Cole-Hopf transformation was discovered nearly seven decades ago. Thus, the observables chosen were easily motivated from knowledge of the analytic solution. Unfortunately, it is rarely the case that such linearizing transformations are known. In our second example, we consider the Koopman operator applied to a second canonical nonlinear PDE: the nonlinear Schrödinger equation

$$iu_t + \frac{1}{2}u_{xx} + |u|^2u = 0, \quad (32)$$

where $u(x, t)$ is a function of space and time modeling slowly varying optical fields or deep water waves, for instance. Discretizing in the spatial variable x , we can Fourier transform the solution in space and use a standard time-stepping

algorithm, such as a fourth-order Runge-Kutta, to integrate the solution forward in time.

As with Burgers' equation, we can compute the DMD by collecting snapshots of the dynamics over a specified time window. Specifically, we consider simulations of the equation with initial data

$$u(x, 0) = 2 \operatorname{sech}(x) \quad (33)$$

over the time interval $t \in [0, \pi]$. Twenty-one snapshots of the dynamics are collected during the evolution, allowing us to create the snapshot matrix \mathbf{X} and \mathbf{X}' . The DMD reconstruction of the dynamics is demonstrated in Figure 4(a). The low-rank DMD reconstruction provides a good approximation to the dynamics of the PDE.

To be more precise, it is explicitly assumed in the DMD reduction that the observables are simply the state variables x where $x = u(x, t)$ at discrete space and time points. The DMD observables are then given by

$$\mathbf{g}_{\text{DMD}}(\mathbf{x}) = \mathbf{x}. \quad (34)$$

Thus, as previously noted, the DMD approximation is a special case of Koopman. The DMD spectrum for a rank $r = 10$ approximation is shown in Figure 4(d). An ideal

approximation would have the eigenvalues aligned along the imaginary axis since the evolution with the initial condition given by (33) is known as the 2-soliton solution which is purely oscillatory.

Koopman theory allows us a much broader set of observables. In what follows, we consider two additional observables

$$\mathbf{g}_1(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ |\mathbf{x}|^2 \mathbf{x} \end{bmatrix}, \quad (35a)$$

$$\mathbf{g}_2(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ |\mathbf{x}|^2 \end{bmatrix} \quad (35b)$$

The first observable $g_1(x)$ is motivated by the form of the nonlinearity in the NLS equation. The second, $g_2(x)$, is chosen to have a simple quadratic nonlinearity. It has no special relationship to the governing equations. Note that the choice of the observable $|\mathbf{x}|^2$ in $g_2(x)$ is relatively arbitrary. For instance, one could consider instead $|\mathbf{x}|^5 \mathbf{x}$, x^2 , x^3 , or x^5 . These all produce similar results to $g_2(x)$ selected in (35b). Specifically, the observable $g_2(x)$ is inferior to either the DMD or judiciously selected $g_1(x)$ for the Koopman reconstruction.

As has been repeatedly stated, the success of the Koopman decomposition relies almost exclusively on the choice of observables. To demonstrate this in practice, we compute the Koopman decomposition of the NLS (32) using the two observables (35a) and (35b). The required data matrices have $2n$ rows of data, and only the state variables need to be recovered at the end of the procedure. Note that the algorithm produces both a state approximation since the first n components are actually the state vector x and approximations to the nonlinearity. The Koopman eigenfunctions and eigenvalues provide information about the evolution on the observable space.

Figures 4(b) and 4(d) show the Koopman reconstruction of the simulated data for the observables (35a) and (35b). The observable $g_1(x)$ provides an exceptional approximation to the evolution while $g_2(x)$ is quite poor. Indeed, the error of the DMD approximation and two nonlinear observables (35a) and (35b) are shown in Figure 4(g) where the following error metric is used:

$$E(t_k) = \|\mathbf{x}(t_k) - \tilde{\mathbf{x}}(t_k)\| \quad k = 1, 2, \dots, m, \quad (36)$$

where x is the full simulation and \tilde{x} is the DMD or Koopman approximation. With the choice of observable $g_1(x)$, which was judiciously chosen to match the nonlinearity of the NLS, the Koopman approximation of the dynamics is four-orders of magnitude better than a DMD approximation. A poor choice of observables, given by $g_2(x)$, gives the worse performance of all, an order of magnitude worse than DMD. Note also the difference in the Koopman spectra as shown in Figures 4(d)–4(f). In particular, note that the judicious observable $g_1(x)$ aligns the eigenvalues along the imaginary axis as is expected from the dynamics. It further

suggests that much better long-time predictions can be achieved with the Koopman decomposition using $g_1(x)$.

Observable selection in this case was facilitated by knowledge of the governing equations. However, in many cases, no such expert knowledge is available, and we must rely on data. The kernel DMD method allows one to use the kernel trick to consider a vast range of potential observables. As already highlighted, the kernel method allows for an efficient method to consider a large class of potential observables without making the observation vector $g(x)$ computationally intractable. For instance, one can consider a radial basis function kernel

$$f(\mathbf{x}, \mathbf{x}') = \exp(-|\mathbf{x} - \mathbf{x}'|^2). \quad (37)$$

The absolute value is conjectured to be important for the case of the NLS equation considered due to the nonlinear evolution of the phase. This radial basis-type function is one of the more commonly considered kernels. Other kernels that we might consider include the three following observables

$$f(\mathbf{x}, \mathbf{x}') = (1 + \mathbf{x}^T \mathbf{x}')^{20}, \quad (38a)$$

$$f(\mathbf{x}, \mathbf{x}') = (1 + |\mathbf{x}^T \mathbf{x}'|)^{20}, \quad (38b)$$

$$f(\mathbf{x}, \mathbf{x}') = \exp(-\mathbf{x}^T \mathbf{x}'). \quad (38c)$$

The first function is the standard polynomial kernel of 20th degree. The second instead takes the absolute value of the variable in a polynomial in order to remove the phase, and the third is a Gaussian-type kernel that uses the same inner product as the polynomial kernel. Note that it uses the same inner product in state space but a completely different inner product in feature space. The selection of these kernels is motivated by well-known and often used kernels for real valued data. Other kernels can easily be considered. Our objective is not so much to evaluate a specific kernel but to demonstrate that kernel selection produces highly variable results so that kernel tuning via cross validation is of critical importance. It is well known that SVM and deep neural nets require significant cross-validation in order to work well. Moreover, note that complex data are rarely considered for kernel selection, so there is ambiguity about what impact this may have. Indeed, this remains an open research question in its own right. However, the NLS has a specific form of nonlinearity which is phase-independent, thus motivating some of our choices of potential kernels.

These three new kernels are compared to each other and the radial basis function. Figure 5 shows the spectra generated by these four kernels along with a comparison to the Koopman spectra generated by $g_1(x)$. Note the tremendous variability of the results based upon the choice of kernel. More precisely, it simply highlights the critical importance of calibrating the kernel through cross-validation.

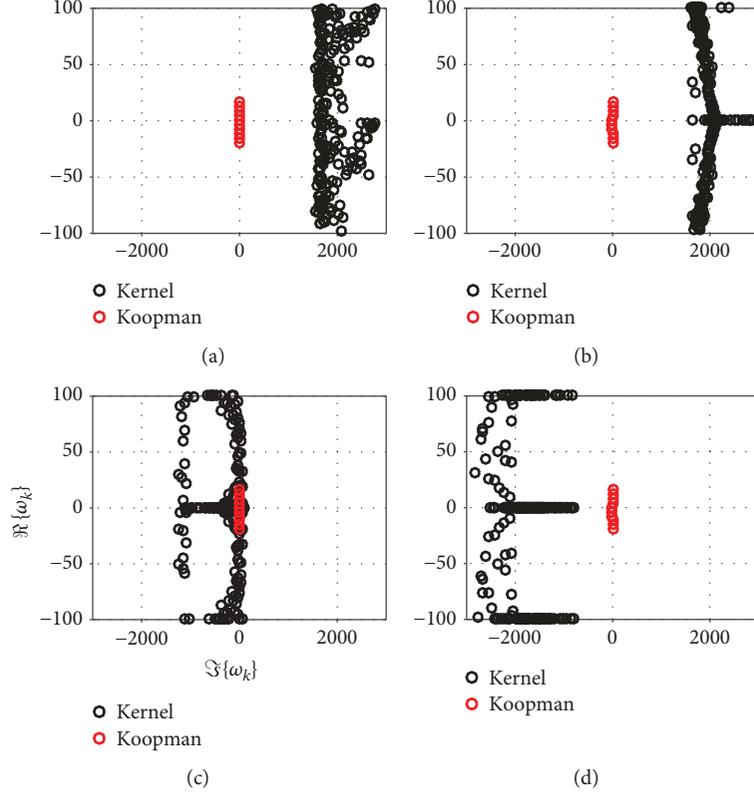


FIGURE 5: (a)–(d) Koopman spectra of the four kernels considered in (37) and (38a)–(38c), respectively. The red spectra are the Koopman spectra generated from the rank $r = 10$ observable $g_1(x)$ which provides an exceptionally accurate reconstruction of the NLS dynamics.

Specifically, the choice of kernel must be carefully selected for either the extended or kernel DMD to give anything reasonable. Cross-validation techniques could potentially be used to select a suitable kernel for applications of interest. It could also ensure that overfitting of the data does not occur. In either case, this simple example should serve as a strong cautionary tale for using kernel techniques in Koopman theory unless results are carefully cross validated. In contrast, our judiciously selected variables produce reasonable and parsimonious results which can be easily cross validated, saving a great deal of computational effort in producing improved performance over DMD.

5.3. Cubic Quintic Ginzburg-Landau Equation. A second and more difficult example for the DMD and Koopman theory to characterize is the cubic-quintic Ginzburg-Landau equation (CQGLE). The CQGLE is a canonical model from mathematical physics that exhibits a wide range of nonlinear spatio-temporal dynamics, including spatio-temporal periodicity and chaos. The evolution equation for CQGLE is given by [45]

$$iu_t + \left(\frac{1}{2} - i\tau\right)u_{xx} - iku_{xxx} + (1 - i\beta)|u|^2u + (\nu - i\sigma)|u|^4u - i\gamma u = 0, \quad (39)$$

where the state variable $u(x, t)$ is a function of space and time. Unlike the NLS equation, the CQGLE is not Hamiltonian and integrable, rather there are significant effects from

dissipation and gain effects, both linear and nonlinear. An example solution generated from the CQGLE is illustrated in Figure 6. This breather-type solution, although simple looking, does not have a simple low-rank representation like the NLS two-soliton breather. Indeed, the singular value decay suggests that a large number of modes are required for an accurate reconstruction. Importantly, the temporal evolution of the POD modes, which can be extracted from the columns of the \mathbf{V} matrix of the SVD, shows that the temporal evolution of the dynamics is quite complicated. This makes it difficult for the DMD approximation since it relies on approximating the temporal evolution by simple Fourier modes in time.

We again consider two additional observables

$$\mathbf{g}_3(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ |\mathbf{x}|^2\mathbf{x} \\ |\mathbf{x}|^4\mathbf{x} \end{bmatrix}, \quad (40a)$$

$$\mathbf{g}_4(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ |\mathbf{x}|^4 \end{bmatrix}. \quad (40b)$$

The first observable $g_3(x)$ is motivated by the form of the nonlinearity in the CQGLE equation. The second, $g_4(x)$, is chosen to have a quartic nonlinearity. The latter of the observables has no special relationship to the governing equations. And as before, a wide range of other

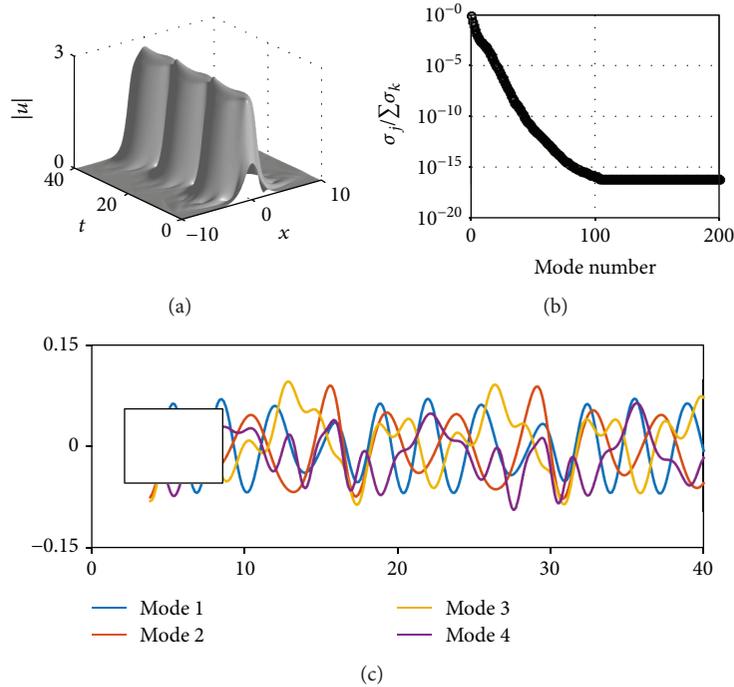


FIGURE 6: Spatio-temporal breather solution of the CQGLE equation for the parameter regime $\tau = 0.08$, $\kappa = 0$, $\beta = 0.66$, $\nu = -0.1$, $\sigma = -0.1$, and $\gamma = -0.1$. Although the dynamics illustrated (a) look relatively simple, the singular value decay (b) shows that a large number of modes are required to reconstruct the fine spatio-temporal features of the nonlinear evolution. Moreover, the temporal dynamics of the first four POD modes, which are extracted from the columns of the \mathbf{V} matrix of the SVD (c), characterize a complicated temporal behavior for the individual modes. Unlike the NLS example, a low-rank approximation does not work well for reconstructing the dynamics. The simulation was performed over $t \in [0, 40]$ with 301 equally spaced snapshots taken. The domain was discretized with $n = 512$ points on a domain $x \in [-10, 10]$.

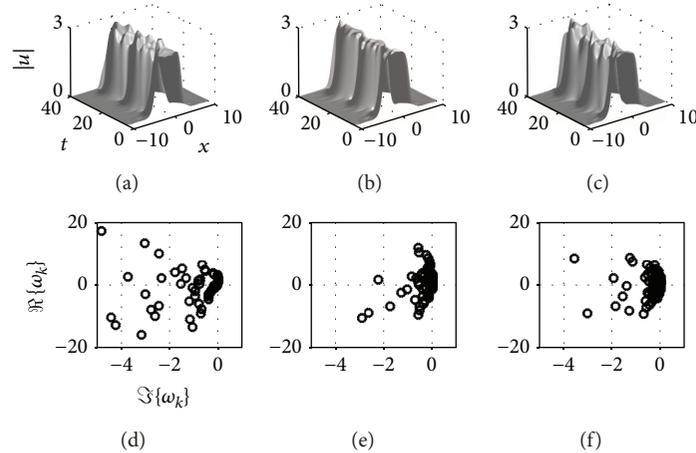


FIGURE 7: Reconstruction of the CQGLE dynamics of Figure 6(a) using a rank $r = 60$ expansion. Shown are (a) a standard DMD approximation $g_{\text{DMD}}(x)$, (b) the CQGLE motivated $g_3(x)$, and (c) a quartic observable $g_4(x)$. The Koopman spectra for each observable are demonstrated in (d), (e), and (f) which accompany the observables of (a), (b), and (c), respectively. Note that the observable $g_3(x)$ produces spectra which are most approximately purely imaginary which is expected of the periodic evolution. A visual inspection shows that the observable $g_3(x)$ produces the best reconstruction.

randomly selected nonlinear observables produce similar results to $g_4(x)$.

The DMD and Koopman reconstructions of the dynamics of the CQGLE are illustrated in Figure 7. As with the NLS example, the CQGLE motivated $g_3(x)$ gives the best

reconstruction. Importantly, the spectra generated are closest to the imaginary axis, which is expected for the periodic spatio-temporal dynamics observed. Indeed, the DMD algorithm, or its Koopman variant applied to observables, ideally generates a purely imaginary spectrum.

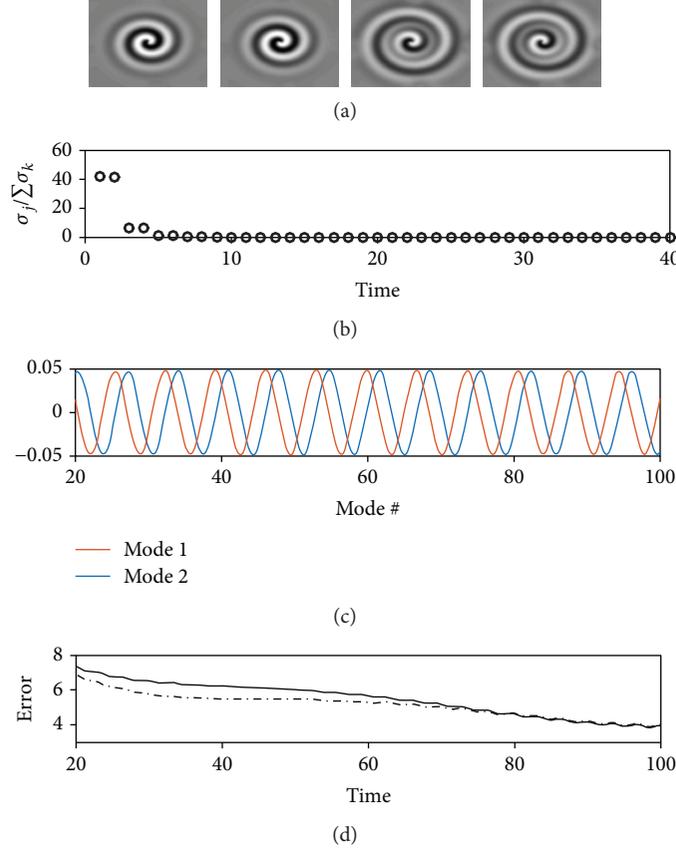


FIGURE 8: Diagnostic features of the reaction diffusion system (41a) and (41b) with $\beta = 1$ and $D = 0.1$. (a) The first four POD modes from sampling snapshots of the spiral wave dynamics. For this data, the snapshots were collected once the spiral wave was formed. Additionally, a spatial filter of the form $f(x, y) = \exp[-0.01(x^2 + y^2)]$ was applied in order to remove the effects of the periodic boundary conditions on the dynamics. (b) There is a two-mode dominance in the dynamics. (c) The temporal dynamics of the first two POD modes, which are extracted from the columns of the \mathbf{V} matrix of the SVD. For the spiral waves, these temporal modes are nearly perfect sinusoids. (d) A comparison of the error between the DMD reconstruction (solid line) and reaction-diffusion-inspired observable $g_5(u, v)$ (dash-dot line). Note that there is very little difference in the reconstruction error in contrast to the NLS and CQGLE examples. The simulation was performed over $t \in [0, 100]$ with 1001 equally spaced snapshots taken. The domain was discretized with $n = 128$ points on the domains $x \in [-20, 20]$ and $y \in [-20, 20]$.

5.4. *Reaction Diffusion System.* As a final example, we consider the $\lambda - \omega$ reaction-diffusion system [46]

$$u_t = D\nabla^2 u + \lambda(A)u - \omega(A)v, \quad (41a)$$

$$v_t = D\nabla^2 v + \omega(A)u + \lambda(A)v, \quad (41b)$$

where $A^2 = u^2 + v^2$, $\omega(A) = -\beta A^2$, $\lambda(A) = 1 - A^2$, $D = 0.1$, $\beta = 1$, and periodic boundaries are applied. This model generates spiral wave solutions which are sustained in the reaction-diffusion process. For processing the data, a spatial filter of the form $f(x, y) = \exp[-0.01(x^2 + y^2)]$ is applied to the snapshots. This removes the boundary effects from computing on a square domain.

Figure 8 shows key characteristics of the evolution dynamics and the decomposition architecture. In Figure 8(a), the first four POD modes are shown. These are the dominant modes of the dynamics associated with the spiral wave, containing approximately 99% of the total variance.

In the low-rank approximation applied, only the first two modes $r = 2$ are used. Importantly, the temporal evolution of the POD modes, which can be extracted from the columns of the \mathbf{V} matrix of the SVD, shows that the temporal evolution of the dynamics is almost purely sinusoidal. This makes it exceptionally easy for the DMD approximation since it relies on approximating the temporal evolution by simple Fourier modes in time. Indeed, given the simple sinusoidal evolution in time, the direct application of DMD is not improved upon by using additional observables. Specifically, we can consider the reaction-diffusion motivated observable

$$\mathbf{g}_5(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ (\mathbf{u}^2 + \mathbf{v}^2)\mathbf{u} \\ (\mathbf{u}^2 + \mathbf{v}^2)\mathbf{v} \end{bmatrix}, \quad (42)$$

where \mathbf{u} and \mathbf{v} are the discretized and reshaped vectors formed from numerically solving the reaction diffusion system. Figure 8(d) shows the reconstruction error for the DMD algorithm compared with the Koopman reconstruction using the observable (42). The error for both is comparable, which is in contrast to the previous examples of NLS and CQGLE where a significant increase in accuracy was achieved using well-selected variables. The fact is that given the almost perfectly sinusoidal low-rank nature of the temporal dynamics, the DMD algorithm simply does not require additional observables to produce an exceptional approximation.

6. Outlook on Koopman Theory for PDEs

Koopman analysis is a remarkable theoretical architecture with applicability to a wide range of nonlinear dynamical systems and PDEs. It combines a number of innovations across disciplines, including dimensionality-reduction techniques, manifold learning, linear operator theory, and dynamical systems. Although the abstract architecture provides a tremendously compelling viewpoint on how to transform nonlinear dynamical systems to infinite-dimensional linear dynamics, significant challenges remain in positing an appropriate set of observables for construction of the Koopman operator. If good candidate observables can be found, then the DMD algorithm can be enacted to compute a finite-dimensional approximation of the Koopman operator, including its eigenfunctions, eigenvalues, and Koopman modes. With a judicious choice of observables, these computed quantities can often lead to physically interpretable spatio-temporal features of the complex system under consideration.

We have demonstrated the application of Koopman theory on several canonical nonlinear PDEs: Burgers' equation, the nonlinear Schrödinger equation, the cubic-quintic Ginzburg-Landau equation, and a reaction-diffusion system. For Burgers' equation, the well-known Cole-Hopf transformation provides a critical link to an explicit calculation of the Koopman operator for a nonlinear PDE. Indeed, we show that the Koopman operator and associated observables can be trivially constructed from knowledge of the Cole-Hopf transformation. In contrast, choosing linear state observables for Burgers' yields a DMD approximation which is accurate but lacks the clear physical interpretation of the exact Koopman reduction. Although the NLS equation can similarly be linearized via the inverse scattering transform, the transform and its inverse are technically difficult to compute for arbitrary initial conditions. Instead, we demonstrate that the selection of an observable that is motivated by the nonlinearity of the governing PDE gives a remarkably accurate Koopman reduction. Indeed, the Koopman eigenfunctions and eigenvalues provide an approximation that is nearly equivalent to the accuracy of the numerical simulation itself. Importantly, for the NLS example, we also demonstrate that poor choices of observables are significantly worse than the DMD approximation. And for the case of observables chosen with a kernel method, the resulting spectra and eigenfunctions are highly inaccurate and nonrobust, suggesting that

such generic techniques as kernel methods may face challenges for use in observable selection. Importantly, the cross validation of the EDMD methods is critically important as the large number of variables used to describe the dynamics, most of which do not have any physical interpretability, must be carefully tuned. Like NLS, the CQGLE model can be similarly improved by variables motivated by the nonlinearity of the PDE. In contrast, the reaction-diffusion system shows that the standard DMD approximation is difficult to improve upon given that the temporal dynamics are almost purely sinusoidal. Such sinusoidal temporal evolution is ideal for the DMD approximation. Even if the governing PDE is not known, symmetries and/or conservation laws can help inform appropriate choices of a parsimonious set of observables. If nothing is known of the physics, then the standard EDMD remains a viable strategy for producing a model, even if interpretability and generalizability is typically lost. This remains an open direction of research which is beyond the current manuscript.

The results presented here provide a prescriptive algorithm for variable selection. Specifically, we recommend the following heuristic measures for a variable selection algorithm: (i) Upon performing the SVD of the data matrix $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ of snapshots of the state space x_j , evaluate the temporal nature of the dominant modes via the columns of the matrix \mathbf{V} . If the dominant columns of \mathbf{V} are approximately sinusoidal, then the standard DMD algorithm should be used (see the reaction-diffusion example). (ii) If the dominant temporal behavior is not sinusoidal, then select observables motivated by the nonlinearity of the governing PDE (see NLS and CQGLE examples). (iii) If the governing PDE is unknown or the method in (ii) is performing poorly, then enact the cross-validated EDMD architecture. All three methods, DMD, judiciously chosen observables with DMD, and EDMD, are all important components of a robust strategy for evaluating nonlinear PDE dynamics.

Ultimately, the selected observables do not need to exactly linearize the system, but they should provide a method for transforming a strongly nonlinear dynamical system to a weakly nonlinear dynamical system. In practice, this is all that is necessary to make the method viable and informative. The results presented here are simultaneously compelling and concerning, highlighting the broader outlook of the Koopman method in general. Specifically, the success of the method will hinge on one issue: selection of observables. If principled techniques, from expert-in-the-loop knowledge, the form of the governing equation, or information about the manifold on which the data exists, can be leveraged to construct suitable observables, then Koopman theory should provide a transformative method for nonlinear dynamical systems and PDEs. We posit that sparse statistical regression techniques from machine learning may provide a path forward towards achieving this goal of selecting quality observables [23, 47]. Failing this, the Koopman architecture may have a limited impact in the mathematical sciences. Because of the importance of identifying meaningful observables, this is an exciting and growing area of research, especially given new developments in machine learning that may provide a robust and principled approach to observable

selection. For those interested in pursuing the EDMD architecture further, we recommend the recent text [12] which highlights many aspects of the current work and some of the structure of the EDMD algorithm that makes in computationally tractable. Included in the book is a link to all codes used in this manuscript.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

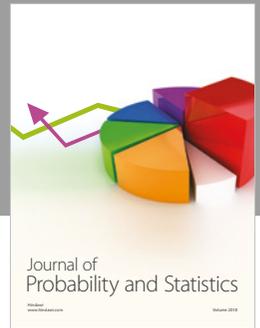
Acknowledgments

J. N. Kutz would like to acknowledge support from the Air Force Office of Scientific Research (FA9550-17-1-0329). S. L. Brunton and J. N. Kutz acknowledge support from the Defense Advanced Research Projects Agency (DARPA HR0011-16-C-0016). J.L. Proctor would like to thank Bill and Melinda Gates for their active support of the Institute for Disease Modeling and their sponsorship through the Global Good Fund.

References

- [1] A. J. Majda, "Challenges in climate science and contemporary applied mathematics," *Communications on Pure and Applied Mathematics*, vol. 65, no. 7, pp. 920–948, 2012.
- [2] S. L. Brunton and B. R. Noack, "Closed-loop turbulence control: progress and challenges," *Applied Mechanics Reviews*, vol. 67, article 50801, no. 5, 2015.
- [3] S. Ganguli and H. Sompolinsky, "Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis," *Annual Review of Neuroscience*, vol. 35, no. 1, pp. 485–508, 2012.
- [4] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010.
- [5] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," *Journal of Fluid Mechanics*, vol. 641, p. 115, 2009.
- [6] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315–318, 1931.
- [7] I. Mezić and A. Banaszuk, "Comparison of systems with complex behavior," *Physica D: Nonlinear Phenomena*, vol. 197, no. 1–2, pp. 101–133, 2004.
- [8] I. Mezić, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, no. 1–3, pp. 309–325, 2005.
- [9] M. Budišić, R. Mohr, and I. Mezić, "Applied Koopmanism," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 4, article 047510, 2012.
- [10] I. Mezić, "Analysis of fluid flows via spectral properties of the Koopman operator," *Annual Review of Fluid Mechanics*, vol. 45, no. 1, pp. 357–378, 2013.
- [11] J. Nathan Kutz, S. L. Brunton, D. M. Luchtenburg, C. W. Rowley, and J. H. Tu, "On dynamic mode decomposition: theory and applications," *Journal of Computational Dynamics*, vol. 1, no. 2, pp. 391–421, 2014.
- [12] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, Society for Industrial and Applied Mathematics, 2016.
- [13] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform*, Society for Industrial and Applied Mathematics, 1981.
- [14] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the Koopman operator: extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.
- [15] I. G. Kevrekidis, C. W. Rowley, and M. O. Williams, "A kernel-based method for data-driven koopman spectral analysis," *Journal of Computational Dynamics*, vol. 2, no. 2, pp. 247–265, 2015.
- [16] C. Schütte, P. Koltai, and S. Klus, "On the numerical approximation of the Perron-Frobenius and Koopman operator," *Journal of Computational Dynamics*, vol. 3, no. 1, pp. 1–12, 2016.
- [17] S. Klus, F. Nüske, P. Koltai et al., "Data-driven model reduction and transfer operator approximation," *Journal of Nonlinear Science*, vol. 28, no. 3, pp. 985–1010, 2018.
- [18] F. Noé and F. Nüske, "A variational approach to modeling slow processes in stochastic dynamical systems," *Multiscale Modeling & Simulation*, vol. 11, no. 2, pp. 635–655, 2013.
- [19] F. Nüske, B. G. Keller, G. Pérez-Hernández, A. S. J. S. Mey, and F. Noé, "Variational approach to molecular kinetics," *Journal of Chemical Theory and Computation*, vol. 10, no. 4, pp. 1739–1752, 2014.
- [20] L. Molgedey and H. G. Schuster, "Separation of a mixture of independent signals using time delayed correlations," *Physical Review Letters*, vol. 72, no. 23, pp. 3634–3637, 1994.
- [21] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noé, "Identification of slow molecular order parameters for Markov model construction," *The Journal of Chemical Physics*, vol. 139, no. 1, article 15102, 2013.
- [22] B. O. Koopman and J. v. Neumann, "Dynamical systems of continuous spectra," *Proceedings of the National Academy of Sciences*, vol. 18, no. 3, pp. 255–263, 1932.
- [23] S. L. Brunton, B. W. Brunton, J. L. Proctor, and J. N. Kutz, "Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control," *PLoS One*, vol. 11, no. 2, article e0150171, 2016.
- [24] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Society for Industrial and Applied Mathematics, 2005.
- [25] S. L. Brunton, J. H. Tu, I. Bright, and J. N. Kutz, "Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems," *SIAM Journal on Applied Dynamical Systems*, vol. 13, no. 4, pp. 1716–1732, 2014.
- [26] S. Sargsyan, S. L. Brunton, and J. N. Kutz, "Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries," *Physical Review E*, vol. 92, no. 3, article 33304, 2015.

- [27] D. Amsallem, R. Tezaur, and C. Farhat, "Real-time solution of linear computational problems using databases of parametric reduced-order models with arbitrary underlying meshes," *Journal of Computational Physics*, vol. 326, pp. 373–397, 2016.
- [28] Y. Choi, D. Amsallem, and C. Farhat, "Gradient-based constrained optimization using a database of linear reduced-order models," 2015, <http://arxiv.org/abs/1506.07849>.
- [29] B. Peherstorfer and K. Willcox, "Online adaptive model reduction for nonlinear systems via low-rank updates," *SIAM Journal on Scientific Computing*, vol. 37, no. 4, pp. A2123–A2150, 2015.
- [30] B. Peherstorfer and K. Willcox, "Dynamic data-driven reduced-order models," *Computer Methods in Applied Mechanics and Engineering*, vol. 291, pp. 21–41, 2015.
- [31] B. Peherstorfer and K. Willcox, "Detecting and adapting to parameter changes for reduced models of dynamic data-driven application systems," *Procedia Computer Science*, vol. 51, pp. 2553–2562, 2015.
- [32] E. Kaiser, B. R. Noack, L. Cordier et al., "Cluster-based reduced-order modelling of a mixing layer," *Journal of Fluid Mechanics*, vol. 754, pp. 365–414, 2014.
- [33] B. Kramer, P. Grover, P. Boufounos, S. Nabi, and M. Benosman, "Sparse sensing and DMD-based identification of flow regimes and bifurcations in complex flows," *SIAM Journal on Applied Dynamical Systems*, vol. 16, no. 2, pp. 1164–1196, 2017.
- [34] C. W. Rowley, M. O. Williams, and I. G. Kevrekidis, "Dynamic mode decomposition and the Koopman operator: algorithms and applications," *IPAM, UCLA*, 2014.
- [35] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [36] S. M. J. Ham, D. D. Lee, and B. Schölkopf, "A kernel view of the dimensionality reduction of manifolds," in *Twenty-first international conference on Machine learning - ICML '04*, Banff, Alberta, Canada, July 2004.
- [37] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.
- [38] G. Baudat and F. Anouar, "Kernel-based methods and function approximation," in *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, pp. 1244–1249, Washington, D.C., 2001.
- [39] W.-H. Steeb and F. Wilhelm, "Non-linear autonomous systems of differential equations and Carleman linearization procedure," *Journal of Mathematical Analysis and Applications*, vol. 77, no. 2, pp. 601–611, 1980.
- [40] K. Kowalski and W.-H. Steeb, *Nonlinear Dynamical Systems and Carleman Linearization*, World Scientific, 1991.
- [41] S. P. Banks, "Infinite-dimensional Carleman linearization, the Lie series and optimal control of non-linear partial differential equations," *International Journal of Systems Science*, vol. 23, no. 5, pp. 663–675, 1992.
- [42] J. M. Burgers, "A mathematical model illustrating the theory of turbulence," *Advances in Applied Mechanics*, vol. 1, pp. 171–199, 1948.
- [43] E. Hopf, "The partial differential equation $u_t + uu_x = \mu_{xx}$," *Communications on Pure and Applied Mathematics*, vol. 3, no. 3, pp. 201–230, 1950.
- [44] J. D. Cole, "On a quasi-linear parabolic equation occurring in aerodynamics," *Quarterly of Applied Mathematics*, vol. 9, no. 3, pp. 225–236, 1951.
- [45] J. N. Kutz, "Mode-locked soliton lasers," *SIAM Review*, vol. 48, no. 4, pp. 629–678, 2006.
- [46] J. D. Murray, *Mathematical Biology. II Spatial Models and Biomedical Applications {Interdisciplinary Applied Mathematics V. 18}*, Springer-Verlag New York Incorporated, 2001.
- [47] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.



Hindawi

Submit your manuscripts at
www.hindawi.com

