INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Wu Z, Brunton SL, Revzen S. 2021 Challenges in dynamic mode decomposition. *J. R. Soc. Interface* **18**: 20210686.
https://doi.org/10.1098/rsif.2021.0686

Received: 1 September 2021
Accepted: 30 November 2021

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
biomathematics, computational biology

Keywords:
dynamic mode decomposition dynamical systems, locomotion

Author for correspondence:
Ziyou Wu
e-mail: wuziyou@umich.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5754119.

THE ROYAL SOCIETY
PUBLISHING

Challenges in dynamic mode decomposition

Ziyou Wu[1], Steven L. Brunton[2] and Shai Revzen[1]

[1]University of Michigan, Ann Arbor, USA
[2]University of Washington, Mechanical Engineering, Seattle, WA, USA

ZW, 0000-0002-4543-6815; SLB, 0000-0002-6565-5118; SR, 0000-0002-2989-0356

Dynamic mode decomposition (DMD) is a powerful tool for extracting spatial and temporal patterns from multi-dimensional time series, and it has been used successfully in a wide range of fields, including fluid mechanics, robotics and neuroscience. Two of the main challenges remaining in DMD research are noise sensitivity and issues related to Krylov space closure when modelling nonlinear systems. Here, we investigate the combination of noise and nonlinearity in a controlled setting, by studying a class of systems with linear latent dynamics which are observed via multinomial observables. Our numerical models include system and measurement noise. We explore the influences of dataset metrics, the spectrum of the latent dynamics, the normality of the system matrix and the geometry of the dynamics. Our results show that even for these very mildly nonlinear conditions, DMD methods often fail to recover the spectrum and can have poor predictive ability. Our work is motivated by our experience modelling multilegged robot data, where we have encountered great difficulty in reconstructing time series for oscillatory systems with intermediate transients, which decay only slightly faster than a period.

1. Introduction

Dynamic mode decomposition (DMD) is a dimensionality reduction and modelling approach that was initially developed in the fluids community [1,2]. DMD has also been used extensively in modelling complex systems in the life sciences, such as human locomotion [3], brain neural activity [4,5], blood flow [6], epidemiology [7] and the kinematics of Crevalle jack fish [8]. DMD is a modal decomposition technique [9,10] that decomposes a matrix of high-dimensional time series data into a set of spatial coherent structures that exhibit the same linear dynamics in time (e.g. oscillations, exponential growth/decay). In this way, the DMD may be thought of as an algorithmic descendent of principal components analysis (PCA) in space with the Fourier transform in time [11–14]. Since its introduction, DMD has been rigorously connected to nonlinear dynamical systems [15], as an approximation of the infinite-dimensional Koopman operator [16,17]. The Koopman operator, introduced in 1931, provides an alternative operator-theoretic perspective for dynamical systems, and it has experienced a resurgence because of its utility for data-driven analysis [18–23].

Since its introduction, DMD has been applied to a wide range of fields beyond fluid mechanics [2,24], including robotics [25], disease modelling [7], neuroscience [4], finance [26] and plasma physics [27,28], to name only a few [13]. One of the key reasons for the broad success of DMD is its formulation in terms of linear algebra, statistical regression and optimization, making it relatable to the recent work in machine learning [29] and highly extensible. Within a short time, DMD has seen dozens of algorithmic extensions, including for control [30,31], for large and streaming datasets [32–35] incorporating ideas of randomized linear algebra [36,37], for model reduction [38,39], for limited measurements [40–43] and time delays [44,45], for multiresolution analysis [46], for selecting a sparse subset of dominant modes [47], and for Bayesian formulations [48], among others [49,50]. The PyDMD [51] software package has emerged as an open source tool for several of these extensions.

Two of the main avenues of DMD research have centred around noise robustness and closure issues related to modelling nonlinear systems. It was recognized early that DMD is quite sensitive to noise [52], and several approaches have been proposed to address this sensitivity, including a forward–backward (FB) averaging [53], total least-squares regression [54], variable projection [55], variational approaches [49], subspace DMD [56], and methods based on the Kalman filter [57,58]. A robust approach, based on robust principal component analysis [59] has also been developed to handle non-Gaussian noise, such as sparse outliers and corruption. However, noise sensitivity is still a leading challenge, and is one that we directly investigate in this paper. Similarly, soon after the initial connection between DMD and Koopman theory [15], there were several follow-up studies addressing how and when a linear regression model can capture essential features of a nonlinear system. The extended DMD algorithm was developed to augment the DMD state with nonlinear observables to enrich the regression [60–62], and theoretical convergence results prove that in the infinite data limit, these models converge to the projection of the Koopman operator on this observable subspace [63]. However, often the observable subspace is not closed under action of the Koopman operator, so that it does not form an invariant subspace, resulting in poorly predicted dynamics [64].

In our own work, we have sometimes encountered great difficulty in reconstructing bio-locomotion time series with DMD. A particularly difficult class seems to be the reconstruction of *intermediate transients*—transient phenomena that decay faster than the slowest mode, but not much faster. This phenomenon is common in both animal and robot locomotion, when springy legs dissipate energy during ground contact. The transient effect of a perturbation can last for several gait cycles before decaying below the noise floor [65,66], which should be characterized by modes inside the unit circle. Limited dataset size in animal experiments also prevents us from inferring the true spectrum of the dynamics. Different from the commonly studied long single trajectory, the animal data are usually composed of multiple short trajectories.

Here, we investigate the possible sources of these difficulties by studying special class of systems for which both the latent linearization and the linearizing observables are completely known—the class systems with linear latent dynamics which are observed via multinomial observables:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k; \quad \mathbf{y} = P(\mathbf{x}), \tag{1.1}$$

where $P(\mathbf{x})$ is a (typically low order) multinomial. However, since we are considering models of physical systems, we will assume both system noise, and measurement noise:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + (\text{system noise}) \quad \text{and} \\ \mathbf{y} &= P(\mathbf{x}) + (\text{measurement noise}). \end{aligned} \tag{1.2}$$

In this paper, we explore the combination of noise and nonlinearity in this controlled setting, where it is possible to isolate and analyse each effect with incrementally increasing severity. Specifically, we explore a particularly benign class of such systems, having three-dimensional latent dynamics and observables that are monotone polynomials in individual coordinates. We demonstrate that even for these very mildly nonlinear systems, observed over a subspace closed under the Koopman operator, DMD methods often fail to recover the spectrum and can have poor predictive value. We also explore the influences of dataset metrics, the spectrum of the latent dynamics, the normality of the system matrix and the geometry of the dynamics. Based on results from this self-contained framework, we give generalizable recommendations regarding dataset properties for DMD analysis. All code used to generate these results are available as open source at: 10.7302/nzq9-4715.

2. Background

Here, we will introduce the DMD, which is a recent technique for linear system identification. DMD is closely related to Koopman operator theory, which provides a representation of a dynamical system in terms of the evolution of its 'observables'. We will also introduce Koopman theory, as it relates to performing DMD on data that is augmented with nonlinear observables.

2.1. Modal decomposition of dynamical systems

Consider equation (1.1), representing a linear dynamical system. Assuming we can observe the state $\mathbf{x} \in \mathbb{R}^n$ directly, the most natural modelling question would be to ask 'can we recover the system matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ that updates the state $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$?' Modal decompositions are a class of solutions to this estimation problem that are particularly suited to high-dimensional data where $n \gg 1$. The decomposition they use represents the dynamics in terms of '*modes*' or '*coherent structures*', i.e. components of $\mathbf{x}$, each of which evolves linearly, and which evolve independently of each other. The eigenvectors ϕ of the system matrix $\mathbf{A}$ represent these modes. For non-degenerate systems, they have the dimension of the original data $\mathbf{x}$, and describe spatial structures that behave coherently in time according to the corresponding eigenvalue λ, i.e. we can write

$$\mathbf{x}_k \approx \Phi\,\mathrm{diag}(\boldsymbol{\lambda})^{k-1}\mathbf{b} = \sum_i \phi_i \lambda_i^{k-1} b_i, \tag{2.1}$$

where each mode ϕ_i evolves exponentially with its own eigenvalue λ_i, and the initial condition $\mathbf{x}_0 = \Phi\boldsymbol{b}$ merely sets $\mathbf{b}$, the weights of the various modes.

This modelling approach may seem to be highly limited—after all, it requires linear systems with an observable state. However, the linearity of the underlying system turns out to be merely a matter of dimension. Koopman theory teaches us that nonlinear dynamics can be represented by an infinite-dimensional linear operator, raising the hope of modelling such systems with (possibly large) finite-dimensional approximate representations of the infinite-dimensional operator.

2.1.1. Koopman theory

Consider the flow $\mathbf{F}: \mathbb{X} \times \mathbb{R} \to \mathbb{X}$ of a continuous- or discrete-time dynamical system, denoted $\mathbf{x}(t) = \mathbf{F}^t(\mathbf{x}(0))$ where $\mathbf{x}(0), \mathbf{x}(t) \in \mathbb{X}$. By 'flow' we mean that $\mathbf{F}$ under composition has the same group structure as the group $(\mathbb{R}, +)$, i.e. $\forall\, t, s \in \mathbb{R}: \mathbf{F}^t \circ \mathbf{F}^s = \mathbf{F}^{t+s}$ and $\mathbf{F}^0$ is the identity map on $\mathbb{X}$.

Koopman theory governs the time evolution of observables of the state $\mathbf{x}(t)$. An 'observable' is a map $\mathbf{g}: \mathbb{X} \to \mathbb{V}$ into some space of observation values $\mathbb{V}$, which we will assume is of the form $\mathbb{C}^d$: a vector space over the field of complex numbers. The 'Koopman operators' (a.k.a. 'composition operators') $\mathcal{K}^t$ act on observables (e.g. $\mathbf{g}$) by

$$\forall\, t \in \mathbb{R}, \quad \mathbf{x} \in \mathbb{X}: \ (\mathcal{K}^t\mathbf{g})(\mathbf{x}) = \mathbf{g}(\mathbf{F}^t(\mathbf{x})), \tag{2.2}$$

i.e. $\mathcal{K}^t$ takes the observable $\mathbf{g}$ to the new observable $\mathbf{g} \circ \mathbf{F}^t$ which describes the observation each state would produce after first evolving under the dynamics for t units of time.

Our interest lies in observables $\mathbf{g}$ such that $\mathbf{g} \circ \mathbf{F}^t$ is a linear system in $\mathbb{V}$, i.e. there exists a matrix $\mathbf{M} \in \mathrm{Aut}(\mathbb{V})$ such that $\exp(t\mathbf{M}) \cdot \mathbf{g} = \mathbf{g} \circ \mathbf{F}^t = \mathcal{K}^t\mathbf{g}$. Such $\mathbf{g}$ spans an 'eigenspace' of $\mathcal{K}$ over which the finite-dimensional linear evolution operator $\mathbf{M}$ is a faithful representation. In the case where $\mathbf{M}$ is a scalar, this coincides with the familiar definition of eigenvectors of the linear operator $\mathcal{K}^t$. Furthermore, if $\mathbf{g}$ is injective it has a left inverse and the nonlinear dynamics $\mathbf{F}^t$ can be rewritten as

$$\mathbf{F}^t = \mathbf{g}^{-1} \circ \exp(t\mathbf{M}) \cdot \mathbf{g}, \tag{2.3}$$

i.e. as latent linear dynamics observed through a change of coordinates. This also implies that by diagonalizing $\mathbf{M} = \mathbf{V}^{-1}\boldsymbol{\Lambda}\mathbf{V}$ we obtain that $\mathbf{Vg}$ is a modal decomposition of the nonlinear dynamics of $\mathbf{x}$ under $\mathbf{F}^t$.

Note that for any two eigenfunctions $g_{\mathbf{A}}$, $g_{\mathbf{B}}$ with commuting (matrix-valued) eigenvalues $\mathbf{A}$ and $\mathbf{B}$, the function $g(\mathbf{x}) := g_{\mathbf{A}}(\mathbf{x}) \cdot g_{\mathbf{B}}(\mathbf{x})$ satisfies

$$\begin{aligned} g \circ \mathbf{F}^t &= (g_{\mathbf{A}} \circ \mathbf{F}^t) \cdot (g_{\mathbf{B}} \circ \mathbf{F}^t) = \exp(t\mathbf{A}) \cdot g_{\mathbf{A}} \cdot \exp(t\mathbf{B}) \cdot g_{\mathbf{B}} \\ &= \exp(t(\mathbf{A}+\mathbf{B})) \cdot (g_{\mathbf{A}} \cdot g_{\mathbf{B}}) = \exp(t(\mathbf{A}+\mathbf{B})) \cdot g, \end{aligned} \tag{2.4}$$

and thus the pointwise product of eigenfunctions whose eigenvalues commute is itself an eigenfunction whose eigenvalues are the sum of the constituent eigenvalues. This also implies (trivially) that non-negative integer powers of eigenfunctions are eigenfunctions. Together these observations imply that any monomial of eigenfunctions $g_1^{\alpha_1}, \dots, g_m^{\alpha_m}$ whose respective matrix eigenvalues were $\mathbf{M}_k$ for g_k, and for whom all $\mathbf{M}_k$ commute, is itself an eigenfunction which has the matrix eigenvalue $\alpha_1\mathbf{M}_1 + \cdots + \alpha_m\mathbf{M}_m$. Clearly, from the Koopman perspective the modal decomposition is not unique without additional conditions (see [67–69]).

However, as we show in electronic supplementary material, A.1, it turns out that the existence of Koopman linearizations for all smooth nonlinear systems with a stable fixed point implies the existence of finite-dimensional linear representations of arbitrary precision (theorem 1). Thus one could reasonably hope that modal decomposition algorithms will work directly on the nonlinear data, provided a high enough operating dimension. This hope provides a formal rationale for trying to use DMD on nonlinear systems.

2.1.2. Dynamic mode decomposition

DMD algorithms seek to obtain the modal decomposition of the dynamics observed in time-series data $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$. For purposes of our discussion will arrange the time-series data into two data matrices, with $\mathbf{x}_k$ arrayed in columns

$$\mathbf{X} = [\mathbf{x}_1, \quad \mathbf{x}_2, \quad \cdots, \quad \mathbf{x}_{m-1}] \quad \mathbf{X}' = [\mathbf{x}_2, \quad \mathbf{x}_3, \quad \cdots, \quad \mathbf{x}_m]. \tag{2.5}$$

These two matrices contain the data one time-step apart. DMD now aims to find the best-fit linear operator $\mathbf{A}$ for the data equation

$$\mathbf{X}' \approx \mathbf{AX}, \tag{2.6}$$

where different DMD algorithms differ primarily in what they define as 'best-fit' in the attempt of making the estimation of $\mathbf{A}$ more accurate and more robust.

Exact DMD. We will begin our brief exposition of DMD methods with the *exact DMD* formulation of Tu *et al.* [12], and introduce several other leading variants of DMD which have shown promise in handling noisy data [53–55]. It should be noted that exact DMD is modified from the original DMD algorithm which was described by Schmid [1].

In exact DMD, this solution comes from least-squares minimization of the Frobenius norm of the error

$$\mathbf{A} = \underset{\mathbf{A}}{\mathrm{argmin}}\, \|\mathbf{X}' - \mathbf{AX}\|_F = \mathbf{X}'\mathbf{X}^{\dagger}, \tag{2.7}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\mathbf{X}^{\dagger}$ is a pseudoinverse of $\mathbf{X}$, which in exact DMD is computed via the singular value decomposition (SVD)

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T \Longrightarrow \mathbf{X}^{\dagger} = \mathbf{V}\Sigma^{-1}\mathbf{U}^T. \tag{2.8}$$

In practice, the SVD of $\mathbf{X}$ is typically truncated so that only $r \ll \min(m, n)$ modes are retained: $\mathbf{X} \approx \mathbf{U}_r\Sigma_r\mathbf{V}_r^T$.

DMD was introduced for high-dimensional systems, such as fluid dynamics, where the state dimension n may be in the millions. In this case, the matrix $\mathbf{A}$ may be intractably large, and instead of constructing $\mathbf{A}$ directly, its low-dimensional representation $\tilde{\mathbf{A}}$ is analysed in the r-dimensional subspace given by the columns of $\mathbf{U}_r$

$$\tilde{\mathbf{A}} = \mathbf{U}_r^T\mathbf{A}\mathbf{U}_r = \mathbf{U}_r^T\mathbf{X}'\mathbf{V}_r\Sigma_r^{-1}. \tag{2.9}$$

The eigenvalues of $\tilde{\mathbf{A}}$ are the same as the eigenvalues of $\mathbf{A}$, and are given by the elements of the diagonal matrix $\boldsymbol{\Lambda}$

$$\tilde{\mathbf{A}}\mathbf{W} = \mathbf{W}\boldsymbol{\Lambda}. \tag{2.10}$$

The corresponding DMD modes, or eigenvectors of the high-dimensional system, are computed from the projected eigenvectors $\mathbf{W}$ as

$$\boldsymbol{\Phi} = \mathbf{X}'\mathbf{V}_r^T\Sigma_r^{-1}\mathbf{W}. \tag{2.11}$$

Tu *et al.* [12] showed that the columns of $\boldsymbol{\Phi}$ are exact eigenvectors of the high-dimensional matrix $\mathbf{A}$.

Once the DMD eigenvalues and eigenvectors are computed, it is possible to approximate the time-series data through the DMD expansion, which is closely related to the Koopman mode decomposition

$$\mathbf{x}_k \approx \sum_{j=1}^{r} \boldsymbol{\phi}_j \lambda^{k-1} b_j = \boldsymbol{\Phi}\boldsymbol{\Lambda}^{k-1}\mathbf{b}. \tag{2.12}$$

The vector $\mathbf{b}$ contains the amplitudes of each DMD mode, and is often computed as

$$\mathbf{b} = \boldsymbol{\Phi}^{\dagger}\mathbf{x}_1. \tag{2.13}$$

Written in matrix form, the DMD expansion becomes

$$\mathbf{X} \approx \underbrace{\begin{bmatrix} | & & | \\ \boldsymbol{\phi}_1 & \cdots & \boldsymbol{\phi}_r \\ | & & | \end{bmatrix}}_{\boldsymbol{\Phi}} \underbrace{\begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_r \end{bmatrix}}_{\mathrm{diag}(\mathbf{b})} \underbrace{\begin{bmatrix} 1 & \cdots & \lambda_1^{m-2} \\ \vdots & \ddots & \vdots \\ 1 & \cdots & \lambda_r^{m-2} \end{bmatrix}}_{\mathbf{T}(\boldsymbol{\Lambda})}, \tag{2.14}$$

where $\mathbf{T}(\boldsymbol{\Lambda})$ is the Vandermonde matrix generated from $\boldsymbol{\Lambda}$ and representing the mode evolution over time.

Classically, the estimation of $\mathbf{b}$ was only based on the initial time-step data $\mathbf{x}_1$. It is possible to improve this estimate with the following optimization over all snapshots:

$$\underset{\mathbf{b}}{\mathrm{argmin}}\, \|\mathbf{X} - \boldsymbol{\Phi}\mathrm{diag}(\mathbf{b})\mathbf{T}(\boldsymbol{\Lambda})\|_F. \tag{2.15}$$

Jovanovic *et al.* [47] further add a sparsity promoting penalty to make the vector $\mathbf{b}$ as sparse as possible.

Optimized DMD. Askham & Kutz [55] introduced the *optimized* DMD algorithm that simultaneously optimizes over the modes and eigenvalues using variable projection. In this approach, they combine the modes and amplitudes $\Phi_\mathbf{b} = \Phi \operatorname{diag}(\mathbf{b})$ and solve the following optimization:

$$\min_{\Lambda,\Phi_\mathbf{b}} \|\mathbf{X} - \Phi_\mathbf{b}\mathbf{T}(\Lambda)\|_F. \tag{2.16}$$

The optimized DMD approach has been proven to be quite robust to noisy data and data that comes from a system which is not perfectly linear.

Forward–backward DMD. The sensitivity of DMD to noise was recognized early by Bagheri [52], and several DMD variants have been proposed to improve the robustness of the algorithm.

FB DMD [53] averages the standard DMD operator $\mathbf{A}$ and the inverse of a reverse-time DMD operator $\mathbf{A}^b$ computed by switching the order of $\mathbf{X}$ and $\mathbf{X}'$ in the algorithm. The averaged operator

$$\mathbf{A}_{\text{FB}} = (\mathbf{A}(\mathbf{A}^b)^{-1})^{1/2} \tag{2.17}$$

cancels out the bias introduced from noise.

Total least-squares DMD. Similarly, total least-squares (TLS) DMD [54] makes note of the fact that if DMD is viewed as a regression from $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$, then noise will affect both the dependent and independent variables. Thus, replacing the standard SVD-based least-squares regression with a total least-squares algorithm will improve the robustness to noise and remove bias. In the TLS algorithm, the data $\mathbf{X}$ and $\mathbf{X}'$ are stacked into a larger matrix that is used to compute the low-dimensional subspace

$$\mathbf{Z} = \begin{bmatrix} \mathbf{X} \\ \mathbf{X}' \end{bmatrix} = \mathbf{U_Z}\Sigma_\mathbf{Z}\mathbf{V_Z}^T. \tag{2.18}$$

The $\mathbf{U_Z}$ matrix is partitioned into four sub-matrices

$$\mathbf{U_Z} = \begin{bmatrix} \mathbf{U}_{\mathbf{Z},a} & \mathbf{U}_{\mathbf{Z},b} \\ \mathbf{U}_{\mathbf{Z},c} & \mathbf{U}_{\mathbf{Z},d} \end{bmatrix}, \tag{2.19}$$

and the debiased estimate of the DMD operator $\mathbf{A}$ is given by

$$\mathbf{A}_{\text{TLS}} = \mathbf{U}_{\mathbf{Z},c}\mathbf{U}^\dagger_{\mathbf{Z},a}. \tag{2.20}$$

In this work, we will compare the exact DMD algorithm, along with the FB, TLS and optimized DMD variants.

Extended DMD. It is also possible to compute DMD on an augmented state that includes nonlinear functions of the measured state $\mathbf{x}$, in a procedure called *extended DMD* (eDMD) [60–62]. By including nonlinear functions, it is possible to approximate the projection of the Koopman operator onto a larger space of functions, where it may be possible to better approximate the relevant eigenfunctions. It was recently shown that extended DMD is equivalent to the earlier *variational approach of conformation dynamics* (VAC) [70–72], introduced by Noé and Nüske to simulate molecular dynamics with a broad separation of timescales. Further connections between eDMD and VAC and between DMD and the time-lagged independent component analysis (TICA) are explored in a recent review [73].

Here, we also explored an extended DMD where the state was augmented with monomial functions that were known by construction to be sufficient for producing a correct model.

3. Methods

Generic stable linear systems can exhibit a great many behaviours; in fact, as we have shown in the appendix, to any finite precision they can exhibit any behaviour a stable smooth nonlinear system can exhibit. To explore several of these, including some of the interactions between different eigenvalues and eigenspace geometries, we constructed a general three-dimensional linear system.

3.1. A 'simple' linear system

We created a stable real linear system, with states $\mathbf{x} \in \mathbb{R}^3$, and considered specifically the case of a spectrum that consists of a (stable) complex-conjugate pair and a single (stable) real eigenvalue. Such a system has two invariant subspaces—a two-dimensional subspace in which the complex-conjugate pair generate a spiral sink, and a one-dimensional subspace in which the real eigenvalue induces exponential convergence. We chose, without loss of generality, to make the two-dimensional invariant space the XY plane with the major and minor axes of the ellipsoidal spiral to be aligned with the X and Y axes. Thus, the most general 3D system matrix $\mathbf{A}$, with a spectrum consisting of one complex-conjugate pair $\alpha \pm i\beta$ and one real eigenvalue λ, can be constructed as follows. We took $\mathbf{A}$ to be constructed from an eigenvector matrix $\mathbf{Q}$, a diagonal matrix $\mathbf{S}$ to rescale the major versus the minor axis in the XY plane, and a block diagonal eigenvalue matrix Λ,

$$\mathbf{A} := \mathbf{QS}\Lambda\mathbf{S}^{-1}\mathbf{Q}^{-1}. \tag{3.1}$$

We parameterized $\mathbf{Q}$, $\mathbf{S}$, Λ by

$$\mathbf{Q} := \begin{bmatrix} 1 & 0 & \sin(\theta)\cos(\phi) \\ 0 & 1 & \sin(\theta)\sin(\phi) \\ 0 & 0 & \cos(\theta) \end{bmatrix},$$
$$\mathbf{S} := \begin{bmatrix} s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Lambda := \begin{bmatrix} \alpha & \beta & 0 \\ -\beta & \alpha & 0 \\ 0 & 0 & \lambda \end{bmatrix}. \tag{3.2}$$

The parameters ϕ, θ and s exhaust all possible 3D linear dynamics with the spectrum of Λ up to an orthogonal similarity transform, as proved in electronic supplementary material, A.3.

3.2. Nonlinear observations of the system

We took as our nonlinear observations of the system the output

$$\begin{aligned} \mathbf{y} := [&x_1 + 0.1(x_1^2 + x_2x_3), \quad x_2 + 0.1(x_2^2 + x_1x_3), \\ &x_3 + 0.1(x_3^2 + x_1x_2)]^T. \end{aligned} \tag{3.3}$$

For the range of $\mathbf{x}$ values, we considered (initial conditions on the unit sphere), $\mathbf{y}$ coordinates were strictly monotone in the corresponding $\mathbf{x}$ coordinates, and observations corresponded to unique states. As figure 1 illustrates, the trajectories of $\mathbf{y}$ and of $\mathbf{x}$ are so similar that for a casual observer they would seem nearly indistinguishable. However, as we will show in 4, these nonlinearities can have a profound impact on our ability to model the system using DMD.

Because $\mathbf{A}$ had distinct eigenvalues, it was diagonalizable over $\mathbb{C}$. Consider the diagonal matrix $\tilde{\Lambda}$ with nine diagonal elements comprising the spectrum of $\mathbf{A}$ and all six pairwise sums of eigenvalues of $\mathbf{A}$. Every monomial of the form $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ with $1 \le k_1 + k_2 + k_3 \le 2$ evolves exponentially over time as one of the diagonal elements of $\tilde{\Lambda}$. We conclude that the dynamics of $\mathbf{y}$ could just as well have been written as $\mathbf{z}_{n+1} := \mathbf{P}\tilde{\Lambda}\mathbf{P}^{-1}\mathbf{z}_n$, $\mathbf{y} := \mathbf{Wz}$ for matrices $\mathbf{P}$ and $\mathbf{W}$.

This is of utmost importance, since it implies that *this nonlinearly observed 3D linear system is exactly reproduced in a 9D linear system observed linearly*. More generally, all monomial observables of order m or less can be represented using eigenspaces whose eigenvalues can be obtained as a sum of no more than

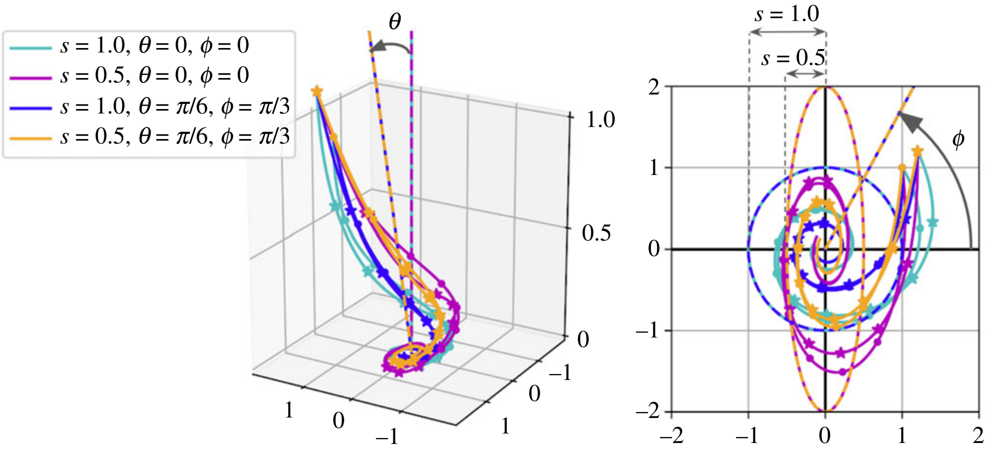

Figure 1. Example trajectories of our 3D system equation (1.2) for various parameters of equation (3.2) (colours). Eigenvalues were $0.9\left(\sqrt{3}/2 \pm 0.5i\right)$ and 0.6 throughout, and we used the same initial condition. We plotted trajectories without noise added for both the latent $\mathbf{x}$ (dots), and the mildly nonlinear observable $\mathbf{y}$ (equation (3.3); stars), in 3D perspective (left) and in an *XY* plane projection (right). In the 3D plot, we indicated the direction of the 0.6 eigenspace controlled by θ and ϕ and given by the last column of Q in equation (3.2) (dashed line). We showed the effects of s in the *XY* plane using ellipsoids (dashed line) whose major and minor axes have ratio s^2. The trajectories demonstrate that our nonlinear observations are quite similar to the latent linear dynamics, and that trajectories vary their geometry with the parameters of equation (3.2) even while the spectrum remains the same.

m of the original eigenvalues; see electronic supplementary material, A.2 for details. We may therefore rest assured that a DMD method that allowed for nine modes could in fact predict the values of $\mathbf{y}$ to without any truncation errors due to the nonlinearity of the observations.

3.3. Dataset preparation

Our numerical experiments used the 3D systems, with Gaussian system noise and Gaussian measurement noise both set to zero mean and a standard deviation of 0.05. The initial conditions were chosen uniformly distributed on the unit sphere.

We explored all combinations of $s \in \{0.1, 0.5, 1.0\}$, $\phi \in \{0, 0.79, \pi/2\}$, $\theta \in \{0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.52, 1.53, 1.560\}$ with a fixed eigenvalue spectrum $0.5\left(\sqrt{3}/2 \pm 0.5i\right)$ and 0.8. We also explored different spectra with complex conjugate eigenvalue pair $0.5\left(\sqrt{3}/2 \pm 0.5i\right)$ and a real eigenvalue varying from 0.01 to 1, with orthogonal eigenvector space $s = 1$, $\theta = 0$. Furthermore, we considered datasets consisting of 100 points each with: (i) $N = 50$ initial conditions for trajectories of length $L = 2$; (ii) $N = 10$ initial conditions for trajectories of length $L = 10$; (iii) $N = 2$ initial conditions for trajectories of length $L = 50$. Because of the presence of system noise, all of these datasets meet the condition of *persistence of excitation* [74].

In the following, we present contour plots showing the distribution of eigenvalue estimates. Wherever we show such plots, they were smoothed with a Gaussian kernel of $\sigma = 0.01$. We first sorted all eigenvalues by their imaginary parts, and then registered each to its nearest true eigenvalue bin. To ensure the spectrum symmetry, we discarded estimates with more real eigenvalues than the true spectrum. If there were any extra complex eigenvalues pairs, we registered two copies of the average value. Furthermore, we added data to these contour plots in batches of 300 numerical experiments. This number was sufficient so that the KL-divergence in P_1 and P_2, the probability distributions before and after a new estimated eigenvalue $\hat{\lambda}$ was added, was smaller than 0.001 for all experiments, i.e. $\sum_{\hat{\lambda}} P_1(\hat{\lambda}) \log(P_1(\hat{\lambda})/P_2(\hat{\lambda})) < 0.001$.

4. Results

In this section, we present the eigenvalue estimation by exact DMD on a 3D linear system with mild nonlinear measurements and observation noise. We examine the effects of system normality, spectrum, mild nonlinearity in the observables, and varying number of initial conditions through their effect on the recovered spectrum and the prediction of the dynamics. We found that mildly nonlinear observations and slight non-normality caused large eigenvalue estimation errors. For a system with a well-posed system matrix, having a dataset with more initial conditions and shorter trajectories can significantly improve the prediction. With a slightly ill-conditioned system matrix, a moderate trajectory length improves the spectrum recovery.

4.1. Fixed spectrum with different system matrix normality

We fixed the spectrum of the system matrix to be $0.5\left(\sqrt{3}/2 \pm 0.5i\right)$ and 0.8, and varied its ϕ, s and θ as described in equation (3.1). The linear states evolved with Gaussian system noise $\sigma = 0.05$, and they were observed via monomials with Gaussian observation noise $\sigma = 0.05$. In figures 2 and 4, we show the DMD eigenvalue density plot under five different normality settings, using first- and second-order monomial observables. In the first column, the system matrix $\mathbf{A}$ is well-conditioned, with orthogonal eigenvectors and a centrosymmetric spiral sink. In the second column, the eigenvectors remain orthogonal, and we stretch the spiral sink by varying the parameter s, so that the resulting trajectory converges following an ellipsoidal spiral with a minor-to-major axes ratio of s^2. The eigenvalue density contour is stretched vertically when s decreases, which yields a larger variance in the complex frequency estimation. In the third and fifth column, we varied the angle between the one-dimensional stable subspace corresponding to the real eigenvalue and the two-dimensional subspace corresponding to the spiral sink, with the spiral sink remaining centrosymmetric. The fourth column exhibits both non-orthogonal eigenvectors and a non-centrosymmetric spiral sink. When eigenvectors are non-orthogonal, a complicated multi-modal error structure shows up in the $L = 10, 50$, $\theta = 1.40, 1.56$ cases.

To thoroughly explore the effect of system matrix normality on spectrum recovery and dynamics prediction, in figures 3 and 5, we plotted the standard deviation of each estimated DMD eigenvalue versus the condition number. The complex eigenvalues always come out as conjugate pairs from exact DMD, so we give one plot per pair. The discarded trials plot

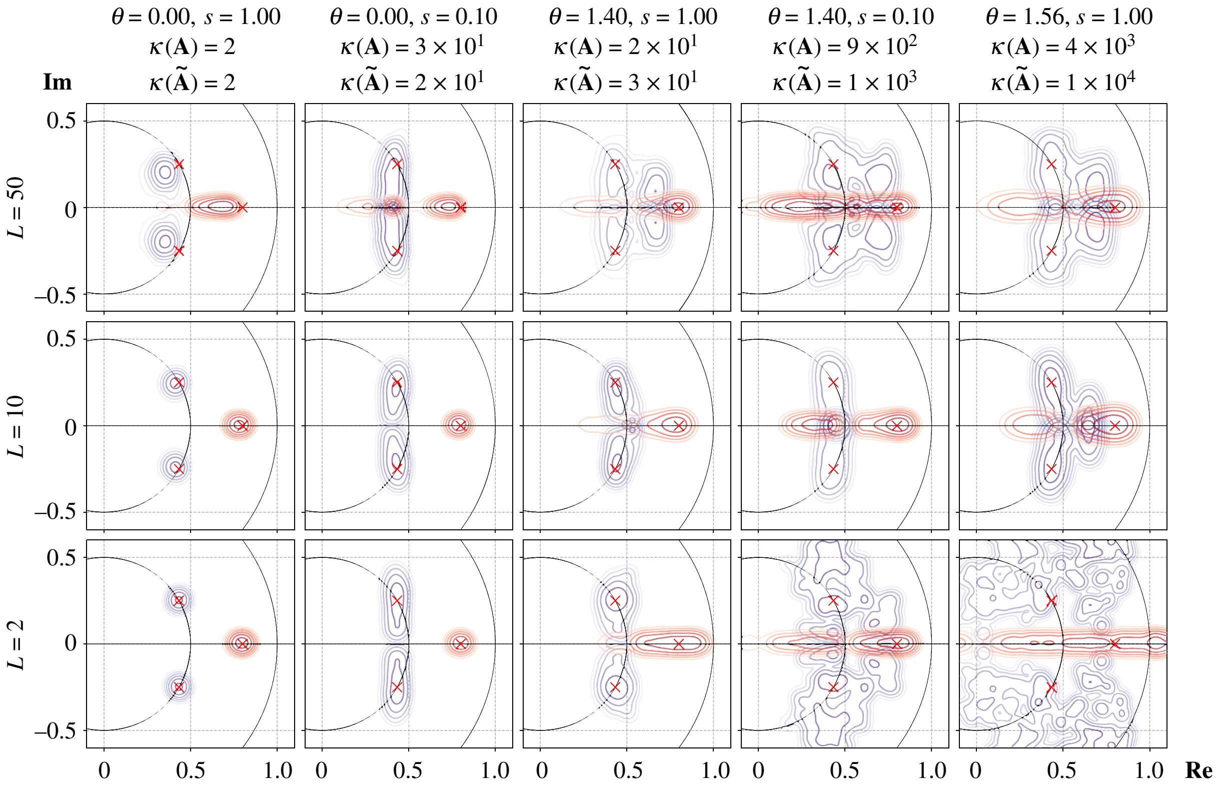

Figure 2. Eigenvalue density plot with first-order linear observables. Different rows have different trajectory lengths and number of trajectories, while holding the same total data points. The system **A** matrix is constructed by $\phi = 0$ and θ, s specified on top of each column. $\kappa(\mathbf{A})$ is the condition number of ground truth system matrix, and $\kappa(\tilde{\mathbf{A}})$ is the mean of estimated condition number found by exact DMD. Red crosses are ground truth eigenvalues.

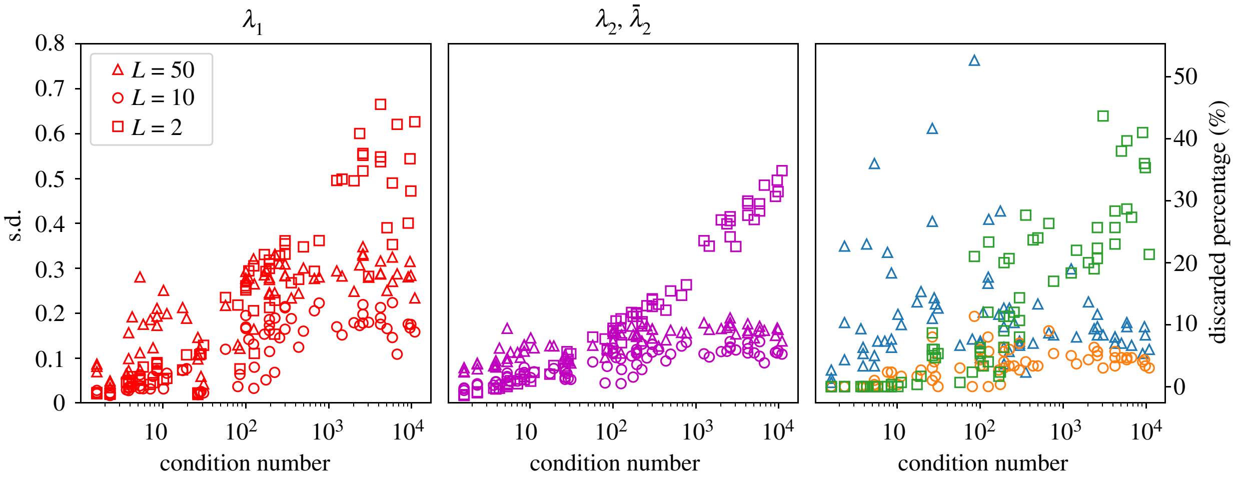

Figure 3. The standard deviation of each estimated eigenvalue versus system **A** matrix condition number plot. The colours of different eigenvalues match the colour of contours in figure 2.

(right lower corner in figures 3 and 5) gives the percentage of DMD estimates resulting in a structurally incorrect estimation, i.e. where the number of real eigenvalues is larger than the ground truth system, because it breaks the symmetry pattern of the complex conjugate pairs. When the condition number is below 100, the eigenvalue estimation from a dataset with more initial conditions and short trajectories (squares) has lower standard deviation, and less discarded trials. When the condition number is larger than 100, the cases with moderate trajectory length (circles) performs better.

The estimation error caused by system matrix non-normality can be amplified through higher order observables. In the first two columns of figure 2, we find that the eigenvalue contours are close to the true eigenvalues. The second-order observables in figure 4 amplified the effect of slight changes in the condition number of the system matrix. Only the slowest modes (red and purple contours) were identifiable. Furthermore, in figure 5, the standard deviation of the eigenvalue estimates increases to 0.25 for most eigenvalues with a modest condition number of 4.

4.2. Different spectrum with orthogonal eigenspace

In this section, we kept an orthogonal eigenvector space with $s = 1$, $\theta = 0$ and a complex conjugate eigenvalue pair $0.5(\sqrt{3}/2 \pm 0.5i)$. We varied the real eigenvalue from 0 to 1. The linear states evolved with Gaussian system noise $\sigma = 0.05$, and were observed via monomials having Gaussian observation noise $\sigma = 0.05$. In figure 6, we show the DMD eigenvalue density plot under three different spectra using

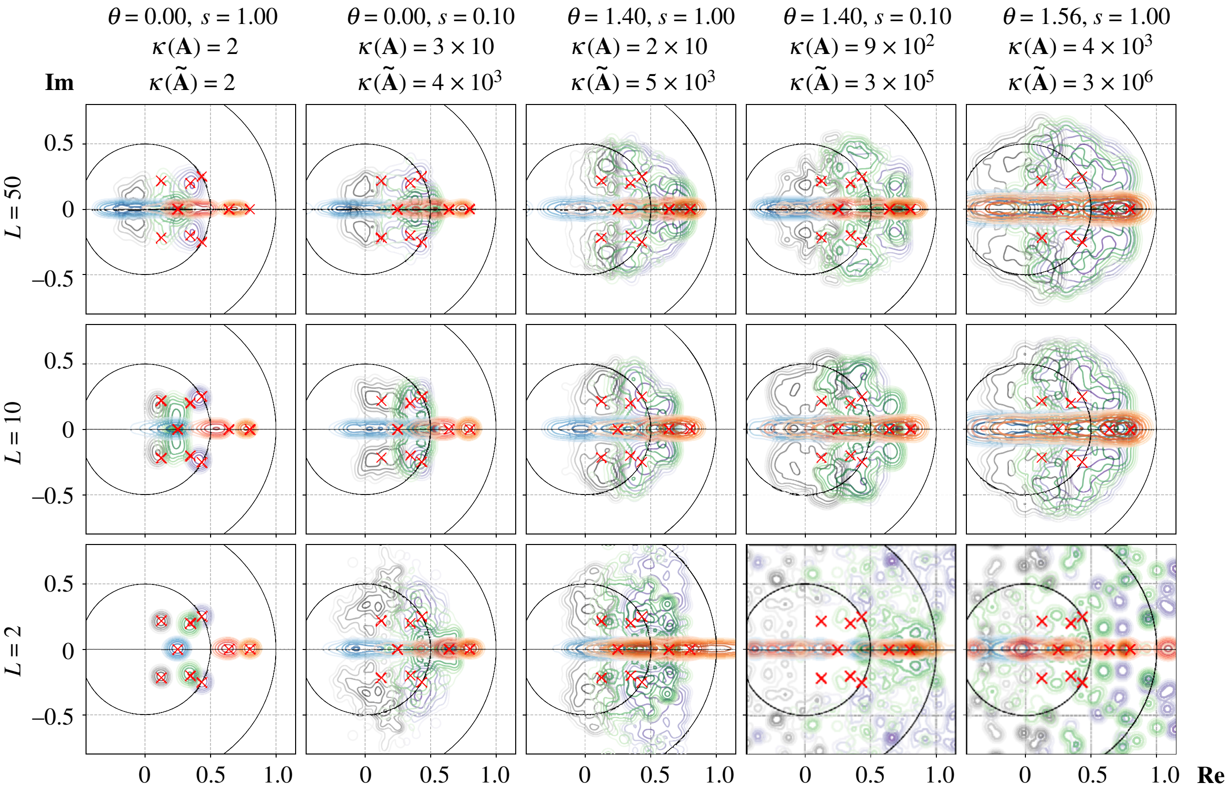

Figure 4. Eigenvalue density plot with monomial observables up to second order. Different rows have different trajectory lengths and number of trajectories, while holding the same total data points. The system **A** matrix is constructed by $\phi = 0$ and θ, s specified on top of each column. $\kappa(\mathbf{A})$ is the condition number of ground truth system matrix, and $\kappa(\tilde{\mathbf{A}})$ is the mean of estimated condition number found by exact DMD. Red crosses are ground truth eigenvalues.

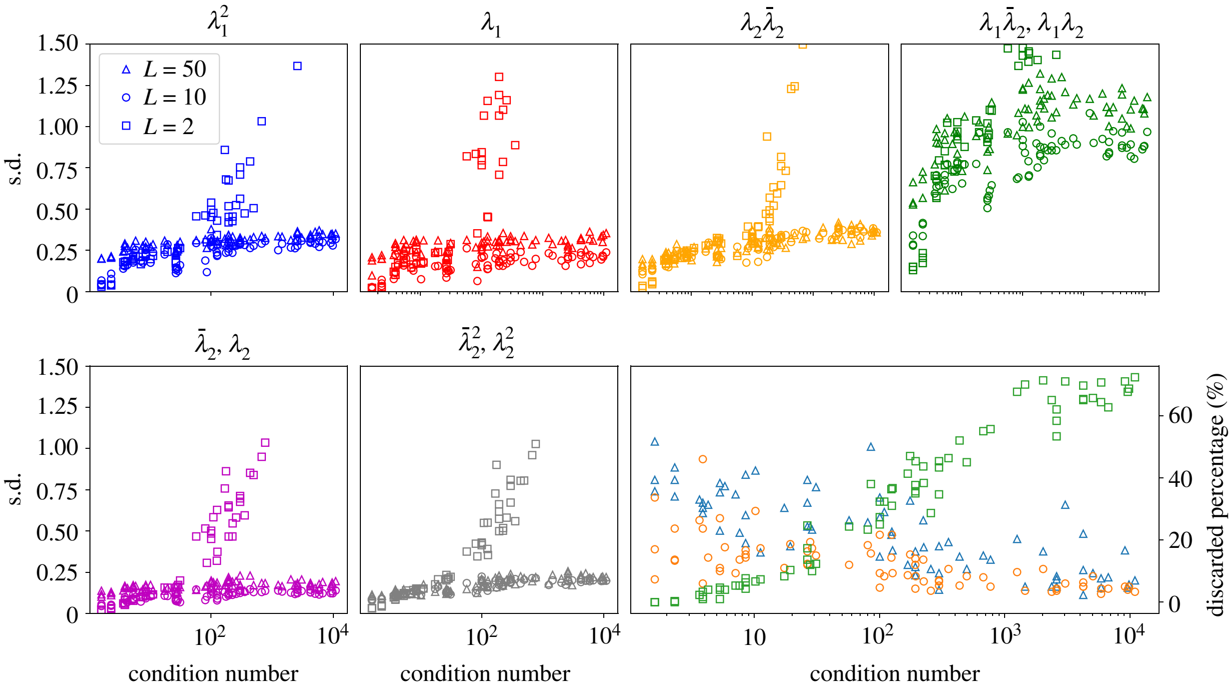

Figure 5. The standard deviation of each estimated eigenvalue versus system **A** matrix condition number plot. The colours of different eigenvalues match the colour of contours in figure 4.

data prepared as described in §3.3. In figures 7 and 8, we show the standard deviation of the estimated DMD eigenvalues versus varying the eigenvalue λ.

More initial conditions with shorter trajectories uniformly outperforms less initial conditions with longer trajectories, resulting in more accurate and precise estimations. In the case of second-order monomial observables with a trajectory length of $L = 2$, the standard deviation of the estimated eigenvalues increases when λ moves closer to the conjugate pair. This suggests a dataset with short trajectories might be more sensitive to a small spectral distance. Nonetheless, this case still has lower standard deviation comparing to less initial conditions with longer trajectories. When λ moves from 0.2 to 1.0, the percentage of discarded trials for $L = 2$ decreases and stays at 0, for $L = 10$ this percentage remains about the same, and for $L = 50$ it increases to 50%.

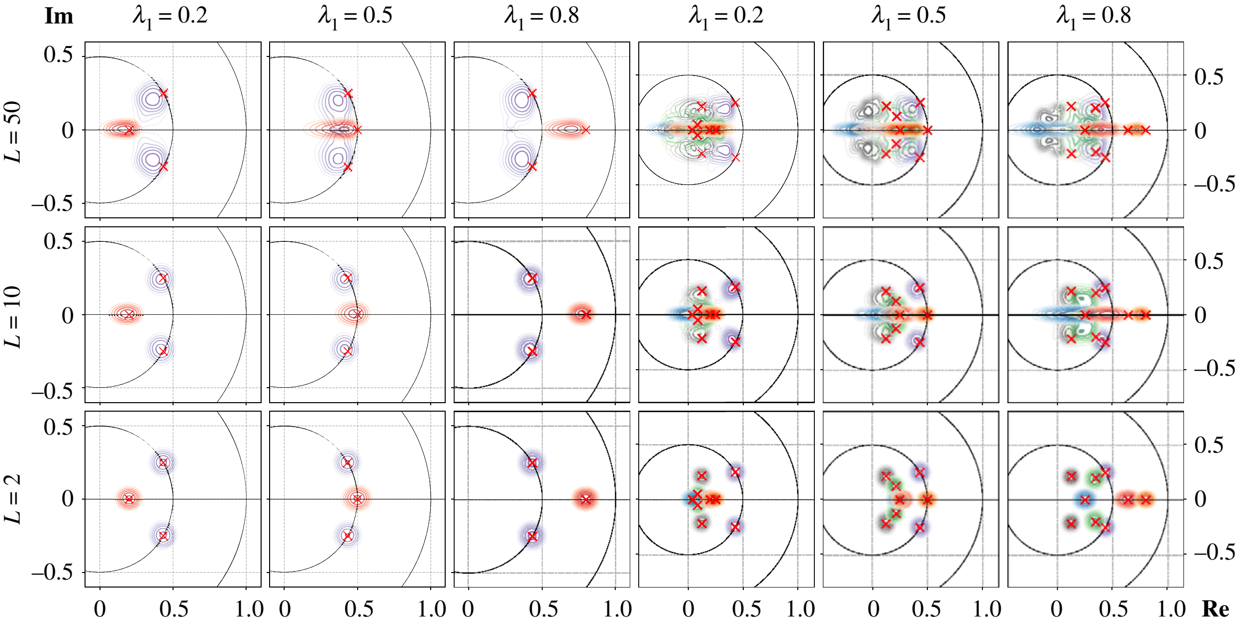

Figure 6. Eigenvalue density plot with linear observables (left nine plots) and observables up to second-order monomials (right nine plots). Different rows have different trajectory length, and different columns have different real eigenvalue. Red crosses are ground truth eigenvalues.

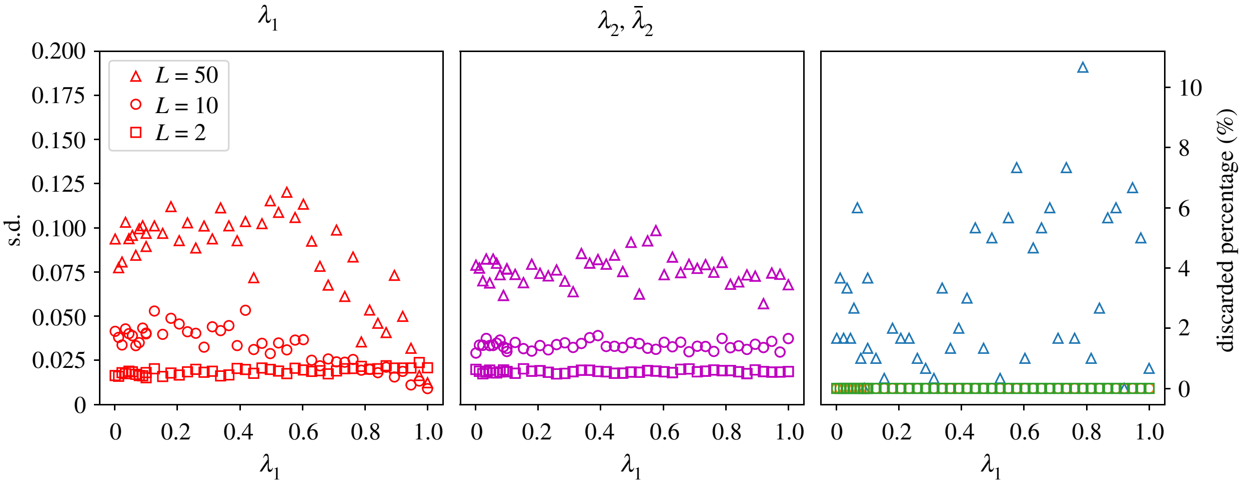

Figure 7. The standard deviation versus real eigenvalue. The colours of different eigenvalues match the colour of contours in the left nine plots of figure 6.

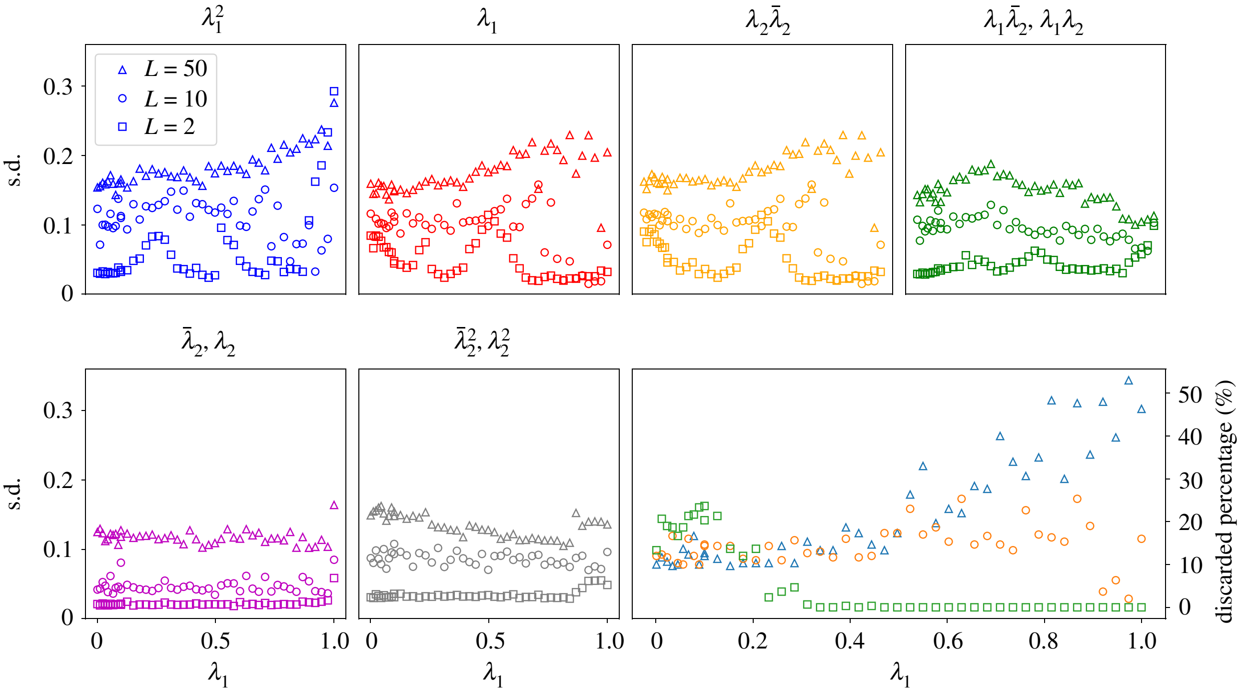

Figure 8. The standard deviation versus real eigenvalue. The colours of different eigenvalues match the colour of contours in the right nine plots of figure 6.

5. Discussion and conclusion

We have shown that under conditions of mild system and observation noise, and mild nonlinearity, DMD can encounter significant difficulties in recovering a linear model with the correct spectrum and an accurate prediction of the dynamics, even when an exact linear model is known to exist. To study these effects, we constructed a self-contained 3D linear system where it is possible to independently control the eigenvalue spectrum, the normality of the system matrix, the nonlinearity of the observations, the system and observation noise, and the number and length of the trajectories used as training data.

In this study, we find that DMD is quite sensitive to non-normality (see figures 2–5; electronic supplementary material, figures 12, 13), and to the existence of higher order nonlinear terms (see figures 4 and 6; electronic supplementary material, figures 11, 13). As to the question of which dataset metrics govern this sensitivity, in some cases shorter data with more initial conditions performs better (as in figures 6–8), and in some data fewer, longer time series produce better results (see 2, 4). The challenge of non-normality has implications for the identification of fluid systems, as highly sheared flows are generally characterized by non-normal linearized dynamics, which was an original motivation for DMD [1].

We first fixed the spectrum of our linear system, varying the normality of system matrix and the geometry of trajectories. The results suggest that an ill-condition system matrix and mildly nonlinear observables cause DMD to fail to recover the full spectrum; this phenomenon is especially evident in figure 4. Furthermore, DMD estimation can be very sensitive to noise when the angle between eigenspaces is small. For the case of linear observables, when we decrease the angle between invariant subspaces, a multi-modal error structure appears in figure 2. The complex eigenvalue pair and the real eigenvalue appear to have switched real parts. The cause of this error structure is not understood and remains an interesting avenue of future work.

We then explored systems with different eigenvalue spectra, where we changed the distance between the eigenvalues. For short trajectories ($L = 2$), closer eigenvalues result in higher standard deviation in the predicted spectrum, using second-order observables. Generally, with enough initial conditions, DMD yields descent prediction of the spectra for all well-conditioned systems explored here. The data requirements of DMD have been widely studied [13,54], and there are some key heuristics and guidelines that have emerged. First, even in the infinite data limit, it is known that DMD is a biased estimator of the dynamics [53,54], and so noisy data will result in bias in the predicted eigenvalues even for large volumes of data. The balance between L and N has also been studied in DMD, and in system identification more generally [75]. Typically, a larger diversity of initial conditions (larger N) for smaller trajectory bursts (smaller L) will be more useful for building accurate and efficient models. However, in many instances, it is impractical or infeasible to reinitialize multiple diverse initial conditions, and instead a single long trajectory is available. Similarly, for clean data with low noise levels, generally it is helpful to reduce the sampling time until it is smaller than the fastest relevant time-scales. However, decreasing the sampling time further typically does not help and may in fact increase the condition number and therefore the sensitivity to noise.

DMD under second-order observations generally does not provide good estimations for transient oscillatory modes. To ensure this is not the effect of spectrum resonance caused by first- and second-order monomial observations, we compute the same contour plot on a 9D non-resonant linear system with similar spectrum in electronic supplementary material, figure 9. The result is similar to the right nine plots of figure 6. Furthermore, we show that using the latest denoizing DMD algorithms in electronic supplementary material, A.5, the difficulty in spectrum recovery often persists, although each method performs well for certain cases.

The 3D linear system analysed here is a special case of a larger class of systems designed specifically to test the sensitivity of DMD to the challenges listed above. The class of diagonal linear systems with multinomial observables provides an algebraically convenient yet sufficiently general class of systems to use to study these problems. As the theorems in electronic supplementary material, A.1 and A.2 show, diagonal linear systems with multinomial observables are universal in the sense of being able to approximate any smooth, stable dynamical system (details in electronic supplementary material, theorem 1). Furthermore, these systems have a finite-dimensional linear representation which is exact—a linear observation of a linear dynamical system (see electronic supplementary material, theorem 2)—and computable in close form. Thus, the class of diagonal linear systems with multinomial observables is a useful space to test any technique that identifies linear models, including the many variants of DMD.

There are several future directions that are motivated by this work. It will be important to explore how these challenges scale to higher dimensional systems. There is also the potential to design sampling strategies to improve the conditioning of the DMD procedure. In addition to understanding and characterizing these challenges on numerical examples, it will be interesting to apply these careful studies to low-dimensional mechanical systems.

Data accessibility. Code and figures repository: https://doi.org/10.7302/nzq9-4715. Data are available as electronic supplementary material [76].

Competing interests. We declare we have no competing interests.

Funding. The authors acknowledge funding support from the Army Research Office (W911NF-17-1-0306).

References

1. Schmid PJ. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28. (doi:10.1017/S0022112010001217)
2. Schmid P, Li L, Juniper M, Pust O. 2011 Applications of the dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **25**, 249–259. (doi:10.1007/s00162-010-0203-9)
3. Fujii K, Takeishi N, Kibushi B, Kouzaki M, Kawahara Y. 2019 Data-driven spectral analysis for coordinative structures in periodic human locomotion. *Sci. Rep.* **9**, 16755. (doi:10.1038/s41598-019-53187-1)
4. Brunton BW, Johnson LA, Ojemann JG, Kutz JN. 2016 Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *J. Neurosci. Methods* **258**, 1–15. (doi:10.1016/j.jneumeth.2015.10.010)
5. Linden NJ, Tabuena DR, Steinmetz NA, Moody WJ, Brunton SL, Brunton BW. 2021 Go with the flow: visualizing spatiotemporal dynamics in optical

widefield calcium imaging. *J. R. Soc. Interface* **18**, 20210523. (doi:10.1098/rsif.2021.0523)

6. Arzani A, Dawson STM. 2021 Data-driven cardiovascular flow modelling: examples and opportunities. *J. R. Soc. Interface* **18**, 20200802. (doi:10.1098/rsif.2020.0802)
7. Proctor JL, Eckhoff PA. 2015 Discovering dynamic patterns from infectious disease data using dynamic mode decomposition. *Int. Health* **7**, 139–145. (doi:10.1093/inthealth/ihv009)
8. Khalid MSU, Wang J, Akhtar I, Dong H, Liu M. 2020 Modal decompositions of the kinematics of Crevalle Jack and the fluid–caudal fin interaction. *Bioinspir. Biomim.* **16**, 016018. (doi:10.1088/1748-3190/abc294)
9. Taira K *et al.* 2017 Modal analysis of fluid flows: an overview. *AIAA J.* **55**, 4013–4041. (doi:10.2514/1.J056060)
10. Taira K, Hemati MS, Brunton SL, Sun Y, Duraisamy K, Bagheri S, Dawson S, Yeh C-A. 2020 Modal analysis of fluid flows: applications and outlook. *AIAA J.* **58**, 998–1022. (doi:10.2514/1.J058462)
11. Chen KK, Tu JH, Rowley CW. 2012 Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses. *J. Nonlinear Sci.* **22**, 887–915. (doi:10.1007/s00332-012-9130-9)
12. Tu JH, Rowley CW, Luchtenburg DM, Brunton SL, Kutz JN. 2014 On dynamic mode decomposition: theory and applications. *J. Comput. Dyn.* **1**, 391–421. (doi:10.3934/jcd.2014.1.391)
13. Kutz JN, Brunton SL, Brunton BW, Proctor JL. 2016 *Dynamic mode decomposition: data-driven modeling of complex systems*. Philadelphia, PA: SIAM.
14. Brunton SL, Kutz JN. 2019 *Data-driven science and engineering: machine learning, dynamical systems, and control*. Cambridge, UK: Cambridge University Press.
15. Rowley CW, Mezić I, Bagheri S, Schlatter P, Henningson D. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **645**, 115–127. (doi:10.1017/S0022112009992059)
16. Koopman BO. 1931 Hamiltonian systems and transformation in Hilbert space. *Proc. Natl Acad. Sci. USA* **17**, 315–318. (doi:10.1073/pnas.17.5.315)
17. Koopman B, Neumann J v.. 1932 Dynamical systems of continuous spectra. *Proc. Natl Acad. Sci. USA* **18**, 255. (doi:10.1073/pnas.18.3.255)
18. Mezić I, Banaszuk A. 2004 Comparison of systems with complex behavior. *Physica D* **197**, 101–133. (doi:10.1016/j.physd.2004.06.015)
19. Mezić I. 2005 Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **41**, 309–325. (doi:10.1007/s11071-005-2824-x)
20. Budišić M, Mohr R, Mezić I. 2012 Applied Koopmanism (a). *Chaos* **22**, 047510. (doi:10.1063/1.4772195)
21. Mezic I. 2013 Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45**, 357–378. (doi:10.1146/annurev-fluid-011212-140652)
22. Bagheri S. 2013 Koopman-mode decomposition of the cylinder wake. *J. Fluid Mech.* **726**, 596–623. (doi:10.1017/jfm.2013.249)
23. Brunton SL, Budišić M, Kaiser E, Kutz JN. 2021 Modern Koopman theory for dynamical systems. (http://arxiv.org/abs/2102.12086)
24. Towne A, Schmidt OT, Colonius T. 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847**, 821–867. (doi:10.1017/jfm.2018.283)
25. Berger E, Sastuba M, Vogt D, Jung B, Amor HB. 2015 Estimation of perturbations in robotic behavior using dynamic mode decomposition. *J. Adv. Rob.* **29**, 331–343. (doi:10.1080/01691864.2014.981292)
26. Mann J, Kutz JN. 2016 Dynamic mode decomposition for financial trading strategies. *Quant. Finance* **16**, 1–13. (doi:10.1080/14697688.2016.1170194)
27. Taylor R, Kutz JN, Morgan K, Nelson B. 2018 Dynamic mode decomposition for plasma diagnostics and validation. *Rev. Sci. Instrum.* **89**, 053501. (doi:10.1063/1.5027419)
28. Kaptanoglu AA, Morgan KD, Hansen CJ, Brunton SL. 2020 Characterizing magnetized plasmas with dynamic mode decomposition. *Phys. Plasmas* **27**, 032108. (doi:10.1063/1.5138932)
29. Brunton SL, Noack BR, Koumoutsakos P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508. (doi:10.1146/annurev-fluid-010719-060214)
30. Proctor JL, Brunton SL, Kutz JN. 2016 Dynamic mode decomposition with control. *SIAM J. Appl. Dyn. Syst.* **15**, 142–161. (doi:10.1137/15M1013857)
31. Kaiser E, Kutz JN, Brunton SL. 2018 Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. *Proc. R. Soc. A* **474**, 20180335. (doi:10.1098/rspa.2018.0335)
32. Hemati MS, Williams MO, Rowley CW. 2014 Dynamic mode decomposition for large and streaming datasets. *Phys. Fluids* **26**, 111701. (doi:10.1063/1.4901016)
33. Sayadi T, Schmid PJ. 2016 Parallel data-driven decomposition algorithm for large-scale datasets: with application to transitional boundary layers. *Theor. Comput. Fluid Dyn.* **30**, 1–14. (doi:10.1007/s00162-016-0385-x)
34. Pendergrass SD, Kutz JN, Brunton SL. 2016 Streaming GPU singular value and dynamic mode decompositions. arXiv, 1612.07875. (http://arxiv.org/abs/1612.07875)
35. Zhang H, Rowley CW, Deem EA, Cattafesta LN. 2017 Online dynamic mode decomposition for time-varying systems. arXiv, 1707.02876. (http://arxiv.org/abs/1707.02876)
36. Bistrian D, Navon I. 2016 Randomized dynamic mode decomposition for non-intrusive reduced order modelling. *Int. J. Numer. Methods Eng.* **112**, 3–25. (doi:10.1002/nme.5499)
37. Erichson NB, Mathelin L, Kutz JN, Brunton SL. 2019 Randomized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **18**, 1867–1891. (doi:10.1137/18M1215013)
38. Tissot G, Cordier L, Benard N, Noack BR. 2014 Model reduction using dynamic mode decomposition. *Comptes Rendus Mécanique* **342**, 410–416. (doi:10.1016/j.crme.2013.12.011)
39. Noack BR, Stankiewicz W, Morzynski M, Schmid PJ. 2016 Recursive dynamic mode decomposition of a transient cylinder wake. *J. Fluid Mech.* **809**, 843–872. (doi:10.1017/jfm.2016.678)
40. Brunton SL, Proctor JL, Tu JH, Kutz JN. 2015 Compressed sensing and dynamic mode decomposition. *J. Comput. Dyn.* **2**, 165–191. (doi:10.3934/jcd.2015002)
41. Gueniat F, Mathelin L, Pastur L. 2015 A dynamic mode decomposition approach for large and arbitrarily sampled systems. *Phys. Fluids* **27**, 025113. (doi:10.1063/1.4908073)
42. Bai Z, Kaiser E, Proctor JL, Kutz JN, Brunton SL. 2017 Dynamic mode decomposition for compressive system identification. *AIAA J.* **58**, 561–574. (doi:10.2514/1.J057870)
43. Kramer B, Grover P, Boufounos P, Benosman M, Nabi S. 2017 Sparse sensing and DMD based identification of flow regimes and bifurcations in complex flows. *SIAM J. Appl. Dyn. Syst.* **16**, 1164–1196. (doi:10.1137/15M104565X)
44. Brunton SL, Brunton BW, Proctor JL, Kaiser E, Kutz JN. 2017 Chaos as an intermittently forced linear system. *Nat. Commun.* **8**, 1–9. (doi:10.1038/s41467-017-00030-8)
45. Arbabi H, Mezic I. 2017 Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator. *SIAM J. Appl. Dyn. Syst.* **16**, 2096–2126. (doi:10.1137/17M1125236)
46. Kutz JN, Fu X, Brunton SL. 2016 Multi-resolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **15**, 713–735. (doi:10.1137/15M1023543)
47. Jovanović MR, Schmid PJ, Nichols JW. 2014 Sparsity-promoting dynamic mode decomposition. *Phys. Fluids* **26**, 024103. (doi:10.1063/1.4863670)
48. Takeishi N, Kawahara Y, Tabei Y, Yairi T. 2017 Bayesian dynamic mode decomposition. In *26th Int. Joint Conf. on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, 19–25 August 2017* (ed. C Sierra), pp. 2814–2821. (doi:10.24963/ijcai.2017/392)
49. Azencot O, Yin W, Bertozzi A. 2019 Consistent dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **18**, 1565–1585. (doi:10.1137/18M1233960)
50. Nair AG, Strom B, Brunton BW, Brunton SL. 2020 Phase-consistent dynamic mode decomposition from multiple overlapping spatial domains. *Phys. Rev. Fluids* **5**, 074702. (doi:10.1103/PhysRevFluids.5.074702)
51. Demo N, Tezzele M, Rozza G. 2018 PyDMD: Python dynamic mode decomposition. *J. Open Source Softw.* **3**, 530. (doi:10.21105/joss.00530)
52. Bagheri S. 2014 Effects of weak noise on oscillating flows: linking quality factor, Floquet modes, and Koopman spectrum. *Phys. Fluids* **26**, 094104. (doi:10.1063/1.4895898)
53. Dawson ST, Hemati MS, Williams MO, Rowley CW. 2016 Characterizing and correcting for the effect of sensor noise in the dynamic mode decomposition. *Exp. Fluids* **57**, 1–19. (doi:10.1007/s00348-016-2127-7)

54. Hemati MS, Rowley CW, Deem EA, Cattafesta LN. 2017 De-biasing the dynamic mode decomposition for applied Koopman spectral analysis. *Theor. Comput. Fluid Dyn.* **31**, 349–368. (doi:10.1007/s00162-017-0432-2)
55. Askham T, Kutz JN. 2018 Variable projection methods for an optimized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **17**, 380–416. (doi:10.1137/M1124176)
56. Takeishi N, Kawahara Y, Yairi T. 2017 Subspace dynamic mode decomposition for stochastic Koopman analysis. *Phys. Rev. E* **96**, 033310. (doi:10.1103/PhysRevE.96.033310)
57. Nonomura T, Shibata H, Takaki R. 2018 Dynamic mode decomposition using a Kalman filter for parameter estimation. *AIP Adv.* **8**, 105106. (doi:10.1063/1.5031816)
58. Nonomura T, Shibata H, Takaki R. 2019 Extended-Kalman-filter-based dynamic mode decomposition for simultaneous system identification and denoising. *PLoS ONE* **14**, e0209836. (doi:10.1371/journal.pone.0209836)
59. Scherl I, Strom B, Shang JK, Williams O, Polagye BL, Brunton SL. 2020 Robust principal component analysis for modal decompositionn of corrupt fluid flows. *Phys. Rev. Fluids* **5**, 054401. (doi:10.1103/PhysRevFluids.5.054401)
60. Williams MO, Rowley CW, Mezić I, Kevrekidis IG. 2015 Data fusion via intrinsic dynamic variables: an application of data-driven Koopman spectral analysis. *EPL* (*Europhys. Lett.*) **109**, 40007. (doi:10.1209/0295-5075/109/40007)
61. Williams MO, Rowley CW, Kevrekidis IG. 2015 A kernel approach to data-driven Koopman spectral analysis. *J. Comput. Dyn.* **2**, 247–265. (doi:10.3934/jcd.2015005)
62. Li Q, Dietrich F, Bollt EM, Kevrekidis IG. 2017 Extended dynamic mode decomposition with dictionary learning: a data-driven adaptive spectral decomposition of the Koopman operator. *Chaos* **27**, 103111. (doi:10.1063/1.4993854)
63. Korda M, Mezić I. 2017 On convergence of extended dynamic mode decomposition to the Koopman operator. arXiv, 1703.04680. (http://arxiv.org/abs/1703.04680)
64. Brunton SL, Brunton BW, Proctor JL, Kutz JN. 2016 Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE* **11**, e0150171. (doi:10.1371/journal.pone.0150171)
65. Revzen S, Burden SA, Moore TY, Mongeau J-M, Full RJ. 2012 Instantaneous kinematic phase reflects neuromechanical response to lateral perturbations of running cockroaches. *Biol. Cybern* **107**, 179–200. (doi:10.1007/s00422-012-0545-z)
66. Maus H-M, Revzen S, Guckenheimer J, Ludwig C, Reger J, Seyfarth A. 2015 Constructing predictive models of human running. *J. R. Soc. Interface* **12**, 20140899. (doi:10.1098/rsif.2014.0899)
67. Kvalheim MD, Revzen S. 2019 Existence and uniqueness of global Koopman eigenfunctions for stable fixed points and periodic orbits. (http://arxiv.org/abs/1911.11996)
68. Lan Y, Mezic I. 2013 Linearization in the large of nonlinear systems and Koopman operator spectrum. *Physica D* **242**, 42–53. (doi:10.1016/j.physd.2012.08.017)
69. Kvalheim MD, Hong D, Revzen S. 2020 Generic properties of Koopman eigenfunctions for stable fixed points and periodic orbits. arXiv, 2010.04008. (http://arxiv.org/abs/2010.04008)
70. Noé F, Nuske F. 2013 A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Model. Simul.* **11**, 635–655. (doi:10.1137/110858616)
71. Nüske F, Keller BG, Pérez-Hernández G, Mey AS, Noé F. 2014 Variational approach to molecular kinetics. *J. Chem. Theory Comput.* **10**, 1739–1752. (doi:10.1021/ct4009156)
72. Nüske F, Schneider R, Vitalini F, Noé F. 2016 Variational tensor approach for approximating the rare-event kinetics of macromolecular systems. *J. Chem. Phys.* **144**, 054105. (doi:10.1063/1.4940774)
73. Klus S, Nüske F, Koltai P, Wu H, Kevrekidis I, Schütte C, Noé F. 2018 Data-driven model reduction and transfer operator approximation. *J. Nonlinear Sci.* **28**, 985–1010. (doi:10.1007/s00332-017-9437-7)
74. Green M, Moore JB. 1985 Persistence of excitation in linear systems. *Proceedings of the American Control Conference, Boston, 1984*, pp. 412–417. *Systems Control Lett.* **7**, 351–360.
75. Wu K, Xiu D. 2018 Numerical aspects for approximating governing equations using data. arXiv, 1809.09170. (http://arxiv.org/abs/1809.09170)
76. Wu Z, Brunton SL, Revzen S. 2021 Data from: Challenges in dynamic mode decomposition. Figshare.