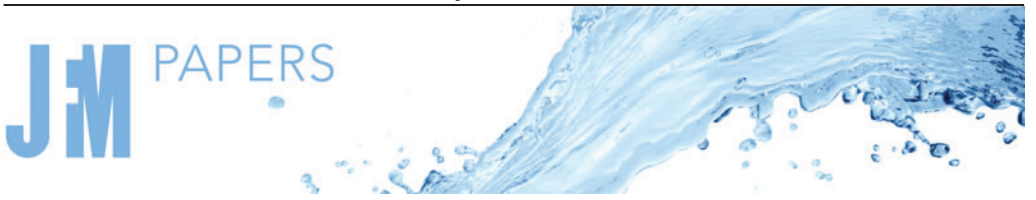
J. Fluid Mech. (2023), *vol*. 955, A21, doi:10.1017/jfm.2022.1052



Residual dynamic mode decomposition: robust and verified Koopmanism

Matthew J. Colbrook[1],†, **Lorna J. Ayton**[1] **and Máté Szőke**[2]

[1]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK

[2]Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24060, USA

(Received 30 April 2022; revised 27 September 2022; accepted 7 December 2022)

Dynamic mode decomposition (DMD) describes complex dynamic processes through a hierarchy of simpler coherent features. DMD is regularly used to understand the fundamental characteristics of turbulence and is closely related to Koopman operators. However, verifying the decomposition, equivalently the computed spectral features of Koopman operators, remains a significant challenge due to the infinite-dimensional nature of Koopman operators. Challenges include spurious (unphysical) modes and dealing with continuous spectra, which both occur regularly in turbulent flows. Residual dynamic mode decomposition (ResDMD), introduced by Colbrook & Townsend (Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems. 2021. arXiv:2111.14889), overcomes such challenges through the data-driven computation of residuals associated with the full infinite-dimensional Koopman operator. ResDMD computes spectra and pseudospectra of general Koopman operators with error control and computes smoothed approximations of spectral measures (including continuous spectra) with explicit high-order convergence theorems. ResDMD thus provides robust and verified Koopmanism. We implement ResDMD and demonstrate its application in various fluid dynamic situations at varying Reynolds numbers from both numerical and experimental data. Examples include vortex shedding behind a cylinder, hot-wire data acquired in a turbulent boundary layer, particle image velocimetry data focusing on a wall-jet flow and laser-induced plasma acoustic pressure signals. We present some advantages of ResDMD: the ability to resolve nonlinear and transient modes verifiably; the verification of learnt dictionaries; the verification of Koopman mode decompositions; and spectral calculations with reduced broadening effects. We also discuss how a new ordering of modes via residuals enables greater accuracy than the traditional modulus ordering

† Email address for correspondence: m.colbrook@damtp.cam.ac.uk

(e.g. when forecasting) with a smaller dictionary. This result paves the way for more significant dynamic compression of large datasets without sacrificing accuracy.

Key words: computational methods, turbulent boundary layers, big data

1. Introduction

We consider dynamical systems whose state $\boldsymbol{x}$ evolves over a state space $\Omega \subseteq \mathbb{R}^d$ in discrete time steps according to a function $F : \Omega \to \Omega$. That is,

$$\boldsymbol{x}_{n+1} = F(\boldsymbol{x}_n), \quad n \geq 0, \tag{1.1}$$

for an initial state $\boldsymbol{x}_0$. At the nth time step the system is in state $\boldsymbol{x}_n$, and such a dynamical system forms a trajectory of iterates $\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ in Ω. We consider the discrete-time setting since we are interested in analysing data collected from experiments and numerical simulations, where we cannot practically obtain a continuum of data. The dynamical system in (1.1) could, for example, be experimental fluid flow measured in a particle image velocimetry (PIV) window over a grid of discrete spatial points (see § 7 for an example). In that case, the function F corresponds to evolving the flow one time step forward and can, of course, be nonlinear.

Koopman operators (Koopman 1931; Koopman & von Neumann 1932) provide a 'linearisation' of nonlinear dynamical systems (1.1) (Mezić 2021). Sparked by Mezić (2005) and Mezić & Banaszuk (2004), recent attention has shifted to using Koopman operators in data-driven methods for studying dynamical systems from trajectory data (Brunton *et al.* 2022). The ensuing explosion in popularity, dubbed 'Koopmanism' (Budišić, Mohr & Mezić 2012), includes thousands of articles over the last decade. The Koopman operator for the system (1.1), $\mathcal{K}$, is defined by its action on 'observables' or functions $g : \Omega \to \mathbb{C}$, via the composition formula

$$\mathcal{K}g = g \circ F, \quad g \in \mathcal{D}(\mathcal{K}), \tag{1.2}$$

where $\mathcal{D}(\mathcal{K})$ is a suitable domain of observables. Note that observables typically differ from the full state of the system in that they are only a given measure of the system. For example, an observable $g(\boldsymbol{x})$ could be the PIV-measured horizontal fluid velocity at a single point, whereas $\boldsymbol{x}$ could correspond to the velocity values at all spatial grid points at a given time. When applied to a state of the system

$$[\mathcal{K}g](\boldsymbol{x}_n) = g(F(\boldsymbol{x}_n)) = g(\boldsymbol{x}_{n+1}). \tag{1.3}$$

Thus, $\mathcal{K}$ acts on observables by advancing them one step forward in time. By considering $g_j(\boldsymbol{x}) = [\boldsymbol{x}]_j$ for $j = 1, \ldots, d$, we see that $\mathcal{K}$ encodes the action of F on $\boldsymbol{x}_n$. Nevertheless, from (1.2), one can show that $\mathcal{K}$ is a linear operator, regardless of whether the dynamics are linear or nonlinear.

The Koopman operator transforms the nonlinear dynamics in the state variable $\boldsymbol{x}$ into an equivalent linear dynamics in the observables g. Hence, the spectral information of $\mathcal{K}$ determines the behaviour of the dynamical system (1.1). For example, Koopman modes are projections of the physical field onto eigenfunctions of the Koopman operator. For fluid flows, Koopman modes provide collective motions of the fluid that are occurring at the same spatial frequency, growth or decay rate, according to an (approximate) eigenvalue of the Koopman operator. Since the vast majority of turbulent fluid interactions have nonlinear properties, this ability to transition to a linear operator is exceedingly useful.

For example, obtaining linear representations for nonlinear systems has the potential to revolutionise our ability to predict and control such systems. However, there is price to pay for this linearisation – $\mathcal{K}$ acts on an infinite-dimensional space, and this is where computational difficulties arise.

In the data-driven setting of this paper, we do not assume knowledge of the function F in (1.1). Instead, we aim to recover the spectral properties of $\mathcal{K}$ from measured trajectory data. The leading numerical algorithm used to approximate $\mathcal{K}$ from trajectory data is dynamic mode decomposition (DMD) (Tu *et al.* 2014; Kutz *et al.* 2016*a*). First introduced by Schmid (2009, 2010) in the fluids community and connected to the Koopman operator in Rowley *et al.* (2009), there are now numerous extensions and variants of DMD (Chen, Tu & Rowley 2012; Wynn *et al.* 2013; Jovanović, Schmid & Nichols 2014; Kutz, Fu & Brunton 2016*b*; Noack *et al.* 2016; Proctor, Brunton & Kutz 2016; Korda & Mezić 2018; Loiseau & Brunton 2018; Deem *et al.* 2020; Klus *et al.* 2020; Baddoo *et al.* 2021, 2022; Herrmann *et al.* 2021; Wang & Shoele 2021; Colbrook 2022). Of particular interest is extended DMD (EDMD) (Williams, Kevrekidis & Rowley 2015*a*), which generalised DMD to a broader class of nonlinear observables. The reader is encouraged to consult Schmid (2022) for a recent overview of DMD-type methods, and Taira *et al.* (2017, 2020) and Towne, Schmidt & Colonius (2018) for an overview of modal-decomposition techniques in fluid flows. DMD breaks apart a high-dimensional spatio-temporal signal into a triplet of Koopman modes, scalar amplitudes and purely temporal signals. This decomposition allows the user to describe complex flow patterns by a hierarchy of simpler processes. When linearly superimposed, these simpler processes approximately recover the full flow.

DMD is undoubtedly a powerful data-driven algorithm that has led to rapid progress over the past decade (Grilli *et al.* 2012; Mezić 2013; Motheau, Nicoud & Poinsot 2014; Sarmast *et al.* 2014; Sayadi *et al.* 2014; Subbareddy, Bartkowicz & Candler 2014; Chai, Iyer & Mahesh 2015; Brunton *et al.* 2016; Hua *et al.* 2016; Priebe *et al.* 2016; Pasquariello, Hickel & Adams 2017; Page & Kerswell 2018, 2019, 2020; Brunton & Kutz 2019; Korda, Putinar & Mezić 2020). However, the infinite-dimensional nature of $\mathcal{K}$ leads to several fundamental challenges. The spectral information of infinite-dimensional operators can be far more complicated and challenging to compute than that of a finite matrix (Ben-Artzi *et al.* 2020; Colbrook 2020). For example, $\mathcal{K}$ can have both discrete and continuous spectra (Mezić 2005). Recently, Colbrook & Townsend (2021) introduced Residual DMD (ResDMD) aimed at tackling some of the challenges associated with infinite dimensions. The idea of ResDMD is summarised in figure 1 (see § 3). ResDMD is easily used in conjunction with existing DMD methods – it does not require additional data and is no more computationally expensive than EDMD. ResDMD addresses the challenges of:

- *Spectral pollution (spurious modes).* A well-known difficulty of computing spectra of infinite-dimensional operators is spectral pollution, where discretisations to a finite matrix cause the appearance of spurious eigenvalues that have nothing to do with the operator and typically have no physical meaning (Lewin & Séré 2010; Colbrook, Roman & Hansen 2019). DMD-type methods typically suffer from spectral pollution and can produce modes with no physical relevance to the system being investigated (Budišić *et al.* 2012; Williams *et al.* 2015*a*). It is highly desirable to have a principled way of detecting spectral pollution with as few assumptions as possible. By computing residuals associated with the spectrum with error control, ResDMD allows the computation of spectra of general Koopman operators with rigorous convergence guarantees and without spectral pollution. ResDMD verifies

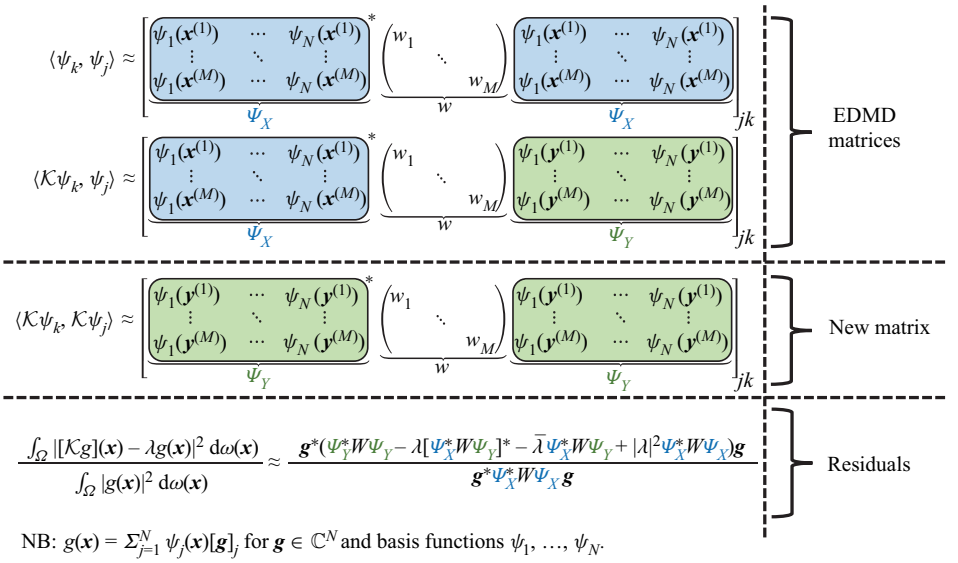


Figure 1. The basic idea of ResDMD – by introducing an additional matrix $\Psi_Y^* W \Psi_Y$ (compared with EDMD), we compute a residual in infinite dimensions. The matrices Ψ_X and Ψ_Y are defined in (2.8*a*,*b*) and correspond to the dictionary evaluated at the snapshot data. The matrix $W = \mathrm{diag}(w_1, \ldots, w_M)$ is a diagonal weight matrix. The approximation of the residual becomes exact in the large data limit $M \to \infty$.

computations of spectral properties and provides a data-driven study of dynamical systems with error control (e.g. see Koopman modes in figure 16).

- *Invariant subspaces.* A finite-dimensional invariant subspace of $\mathcal{K}$ is a space of observables $\mathcal{G} = \mathrm{span}\{g_1, \ldots, g_k\}$ such that $\mathcal{K}g \in \mathcal{G}$ for all $g \in \mathcal{G}$. Non-trivial finite-dimensional invariant spaces need not exist (e.g. when the system is mixing), can be challenging to compute, or may not capture all of the dynamics of interest. Typically, one must settle for approximate invariant subspaces, and DMD-type methods can result in closure issues (Brunton *et al.* 2016). By computing upper bounds on residuals, ResDMD provides a way of measuring how invariant a candidate subspace is. It is important to stress that ResDMD computes residuals associated with the underlying infinite-dimensional operator $\mathcal{K}$ (the additional matrix in figure 1 allows us to do this using finite matrices), in contrast to earlier work that computes residuals of observables with respect to a finite DMD discretisation matrix (Drmac, Mezic & Mohr 2018). In contrast to ResDMD, residuals with respect to a finite DMD discretisation of $\mathcal{K}$ can never be used as error bounds for spectral information of $\mathcal{K}$ and suffer from issues such as spectral pollution. Residuals computed by ResDMD also allow a natural ranking or ordering of approximated eigenvalues and Koopman modes, that can be exploited for efficient system compression (see § 8).
- *Continuous spectra.* The operator $\mathcal{K}$ can have a continuous spectrum, which is a generic feature of chaotic or turbulent flow (Basley *et al.* 2011; Arbabi & Mezić 2017*b*). A formidable challenge is dealing with the continuous spectrum (Mezić 2013; Colbrook 2021; Colbrook, Horning & Townsend 2021), since discretising to a finite-dimensional operator destroys its presence. Most existing non-parametric approaches for computing continuous spectra of $\mathcal{K}$ are restricted to ergodic systems (Arbabi & Mezić 2017*b*; Giannakis 2019; Korda *et al.* 2020; Das, Giannakis &

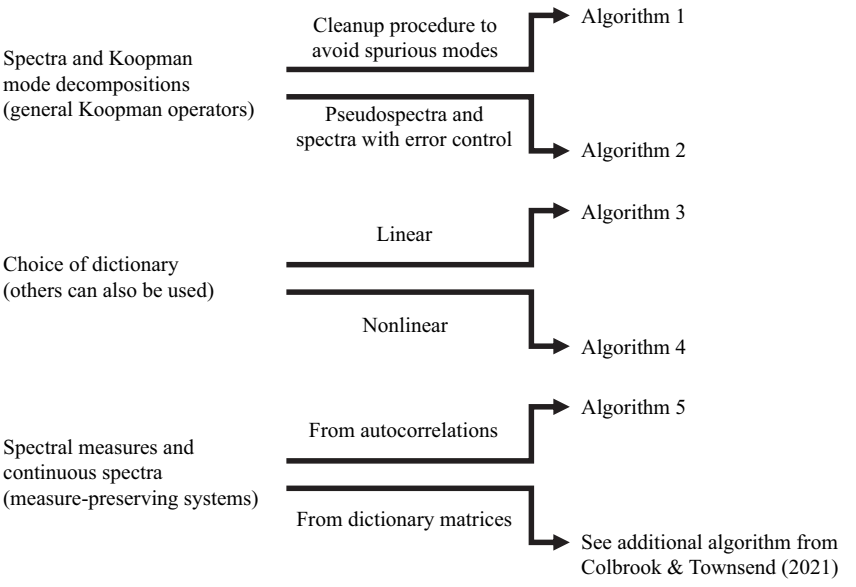


Figure 2. A diagrammatic chart for the algorithms used in this paper. The computational problem is shown on the left, and the relevant algorithms on the right.

Slawinska 2021), as this allows relevant integrals to be computed using long-time averages. ResDMD (Colbrook & Townsend 2021) provides a general computational framework to deal with continuous spectra through smoothed approximations of spectral measures, leading to explicit and rigorous high-order convergence (see the examples in § 6). The methods deal jointly with discrete and continuous spectra, do not assume ergodicity, and can be applied to data collected from either short or long trajectories.

- *Nonlinearity and high-dimensional state space.* For many fluid flows, e.g. turbulent phenomena, the corresponding dynamical system is strongly nonlinear and has a very large state-space dimension. ResDMD can be combined with learned dictionaries to tackle this issue. A key advantage of ResDMD compared with traditional learning methods is that one still has convergence theory and can perform *a posteriori* verification that the learned dictionary is suitable. In this paper, we demonstrate this for two choices of dictionary: kernelised extended dynamic mode decomposition (kEDMD) (Williams, Rowley & Kevrekidis 2015*b*) and (rank-reduced) DMD.

This paper demonstrates ResDMD in a wide range of fluid dynamic situations, at varying Reynolds numbers, from both numerical and experimental data. We discuss how the new ResDMD methods can be reliably used for turbulent flows such as aerodynamic boundary layers, and we include, for the first time, a link between the spectral measures of Koopman operators and turbulent power spectra. By windowing in the frequency domain, ResDMD avoids the problem of broadening and also allows convergence theory. We illustrate this link explicitly for experimental measurements of boundary layer turbulence. We also discuss how a new modal ordering based on the residual permits greater accuracy with a smaller DMD dictionary than traditional modal modulus ordering. This result paves the way for greater dynamic compression of large data sets without sacrificing accuracy.

The paper is outlined as follows. In § 2, we recap EDMD, for which DMD is a special case, and interpret the algorithm as a Galerkin method. ResDMD is then introduced in § 3, where we recall and expand upon the results of Colbrook & Townsend (2021). We stress that ResDMD does not make any assumptions about the Koopman operator or dynamical system. In § 4, we recall one of the algorithms of Colbrook & Townsend (2021) for computing spectral measures (dealing with continuous spectra) under the assumption that the dynamical system is measure preserving, and make a new link between an algorithm for spectral measures of Koopman operators and the power spectra of signals. We then validate and apply our methods to four different flow cases. We treat flow past a cylinder (numerical data) in § 5, turbulent boundary layer flow (hot-wire experimental data) in § 6, wall-jet boundary layer flow (PIV experimental data) in § 7 and laser-induced plasma (experimental data collected with a microphone) in § 8. In each case, only the flow fields or pressure fields are used to extract relevant dynamical information. We end with a discussion and conclusions in § 9. Code for our methods can be found at https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition, and we provide a diagrammatic chart for implementation in figure 2.

2. Recalling DMD

Here, we recall the definition of EDMD from Williams *et al.* (2015*a*) and its interpretation as a Galerkin method. As a particular case of EDMD, DMD can be interpreted as producing a Galerkin approximation of the Koopman operator using linear basis functions.

2.1. *Trajectory data*

We assume that (1.1) is observed for M_2 time steps, starting at M_1 initial conditions. It is helpful to view the trajectory data as an $M_1 \times M_2$ matrix

$$B_{data} = \begin{bmatrix} \boldsymbol{x}_0^{(1)} & \cdots & \boldsymbol{x}_{M_2-1}^{(1)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{x}_0^{(M_1)} & \cdots & \boldsymbol{x}_{M_2-1}^{(M_1)} \end{bmatrix}. \tag{2.1}$$

Each row of B_{data} corresponds to an observation of a single trajectory of the dynamical system that is witnessed for M_2 time steps. In particular, $\boldsymbol{x}_{i+1}^{(j)} = F(\boldsymbol{x}_i^{(j)})$ for $0 \leq i \leq M_2 - 2$ and $1 \leq j \leq M_1$. For example, these snapshots could be measurements of unsteady velocity field across M_1 initial state configurations. One could therefore use just consider $M_2 = 2$. One could also consider a single trajectory, $M_1 = 1$, and large M_2. The exact structure depends on the data acquisition method. In general, (2.1) could represent measurements of the velocity along M_1 trajectories for M_2 time steps, a total of $M_1 M_2$ measurements. One can also use our algorithms with data consisting of trajectories of different lengths (i.e. general snapshots).

Letting $\{\boldsymbol{x}^{(m)}\}_{m=1}^{M}$ and $\{\boldsymbol{y}^{(m)}\}_{m=1}^{M}$ be enumerations of the first $M_2 - 1$ columns and the second to final columns of B_{data}, respectively, with $M = M_1(M_2 - 1)$, we can conveniently represent the data as a finite set of M pairs of measurements of the state

$$\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^{M} \text{ such that } \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)}), \quad m = 1, \ldots, M. \tag{2.2}$$

One could also consider measurements of certain observables g, $\{g(\boldsymbol{x}^{(m)}), g(\boldsymbol{y}^{(m)})\}_{m=1}^{M}$, and interpret the following algorithms as such. EDMD provides a way of using these measurements to approximate the operator $\mathcal{K}$ via a matrix $\mathbb{K} \in \mathbb{C}^{N \times N}$.

2.2. *Extended DMD*

We work in the Hilbert space $L^2(\Omega, \omega)$ of square-integrable observables with respect to a positive measure ω on Ω. We do not assume that this measure is invariant. A common choice of ω is the standard Lebesgue measure. This choice is natural for Hamiltonian systems, for which the Koopman operator is unitary on $L^2(\Omega, \omega)$ (Arnold 1989). Another common choice is for ω to correspond to an ergodic measure on an attractor (Mezić 2005).

Given a dictionary (a set of basis functions) $\{\psi_1, \dots, \psi_N\} \subset \mathcal{D}(\mathcal{K}) \subseteq L^2(\Omega, \omega)$ of observables, EDMD constructs a matrix $\mathbb{K} \in \mathbb{C}^{N\times N}$ from the snapshot data (2.2) that approximates $\mathcal{K}$ on the finite-dimensional subspace $V_N = \operatorname{span}\{\psi_1, \dots, \psi_N\}$. The choice of the dictionary is up to the user, with some common hand-crafted choices given in Williams *et al.* (2015*a*, table 1). When the state-space dimension d is large, as in this paper, it is beneficial to use a data-driven dictionary. We discuss this in § 3.3, where we present DMD (Kutz *et al.* 2016*a*) and kEDMD (Williams *et al.* 2015*b*) in a unified framework.

Given a dictionary of observables, we define the vector-valued observable or 'quasimatrix'

$$\Psi(\boldsymbol{x}) = [\psi_1(\boldsymbol{x}) \quad \cdots \quad \psi_N(\boldsymbol{x})] \in \mathbb{C}^{1\times N}. \tag{2.3}$$

Any new observable $g \in V_N$ can then be written as $g(\boldsymbol{x}) = \sum_{j=1}^{N} \psi_j(\boldsymbol{x}) g_j = \Psi(\boldsymbol{x})\,\boldsymbol{g}$ for some vector of constant coefficients $\boldsymbol{g} = [g_1 \cdots g_N]^\top \in \mathbb{C}^N$. It follows from (1.2) that

$$[\mathcal{K}g](\boldsymbol{x}) = \Psi(F(\boldsymbol{x}))\boldsymbol{g} = \Psi(\boldsymbol{x})(\mathbb{K}\boldsymbol{g}) + R(\boldsymbol{g}, \boldsymbol{x}), \tag{2.4}$$

where

$$R(\boldsymbol{g}, \boldsymbol{x}) = \sum_{j=1}^{N} \psi_j(F(\boldsymbol{x}))g_j - \Psi(\boldsymbol{x})(\mathbb{K}\boldsymbol{g}). \tag{2.5}$$

Typically, the subspace V_N generated by the dictionary is not an invariant subspace of $\mathcal{K}$, and hence there is no choice of $\mathbb{K}$ that makes the error $R(\boldsymbol{g}, \boldsymbol{x})$ zero for all choices of $g \in V_N$ and $\boldsymbol{x} \in \Omega$. Instead, it is natural to select $\mathbb{K}$ as to minimise

$$\int_\Omega \max_{\|\boldsymbol{g}\|=1} |R(\boldsymbol{g}, \boldsymbol{x})|^2 \,\mathrm{d}\omega(\boldsymbol{x}) = \int_\Omega \|\Psi(F(\boldsymbol{x})) - \Psi(\boldsymbol{x})\mathbb{K}\|^2 \,\mathrm{d}\omega(\boldsymbol{x}), \tag{2.6}$$

where $\|\boldsymbol{g}\|$ denotes the standard Euclidean norm of a vector $\boldsymbol{g}$. Given a finite amount of snapshot data, we cannot directly evaluate the integral in (2.6). Instead, we approximate it via a quadrature rule by treating the data points $\{\boldsymbol{x}^{(m)}\}_{m=1}^{M}$ as quadrature nodes with weights $\{w_m\}_{m=1}^{M}$. The discretised version of (2.6) is the following weighted least-squares problem:

$$\mathbb{K} \in \operatorname*{argmin}_{B\in\mathbb{C}^{N\times N}} \sum_{m=1}^{M} w_m \|\Psi(\boldsymbol{y}^{(m)}) - \Psi(\boldsymbol{x}^{(m)})B\|^2. \tag{2.7}$$

We define the following two matrices:

$$\Psi_X = \begin{pmatrix} \Psi(\boldsymbol{x}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{x}^{(M)}) \end{pmatrix} \in \mathbb{C}^{M\times N}, \quad \Psi_Y = \begin{pmatrix} \Psi(\boldsymbol{y}^{(1)}) \\ \vdots \\ \Psi(\boldsymbol{y}^{(M)}) \end{pmatrix} \in \mathbb{C}^{M\times N}, \tag{2.8a,b}$$

and let $W = \operatorname{diag}(w_1, \dots, w_M)$ be the diagonal weight matrix of the quadrature rule. A solution to (2.7) is

$$\mathbb{K} = (\Psi_X^* W \Psi_X)^\dagger (\Psi_X^* W \Psi_Y) = (\sqrt{W}\Psi_X)^\dagger \sqrt{W}\Psi_Y, \tag{2.9}$$

where '†' denotes the pseudoinverse. In some applications, the matrix $\Psi_X^* W \Psi_X$ may be il -conditioned, so it is common to consider a regularisation such as a truncated singular value decomposition (SVD).

2.3. *Quadrature and Galerkin methods*

We can view EDMD as a Galerkin method. Note that

$$\Psi_X^* W \Psi_X = \sum_{m=1}^{M} w_m \Psi(\boldsymbol{x}^{(m)})^* \Psi(\boldsymbol{x}^{(m)}), \quad \Psi_X^* W \Psi_Y = \sum_{m=1}^{M} w_m \Psi(\boldsymbol{x}^{(m)})^* \Psi(\boldsymbol{y}^{(m)}). \quad (2.10a,b)$$

If the quadrature approximation converges, it follows that

$$\lim_{M\to\infty} [\Psi_X^* W \Psi_X]_{jk} = \langle \psi_k, \psi_j \rangle \quad \text{and} \quad \lim_{M\to\infty} [\Psi_X^* W \Psi_Y]_{jk} = \langle \mathcal{K}\psi_k, \psi_j \rangle, \quad (2.11a,b)$$

where $\langle \cdot, \cdot \rangle$ is the inner product associated with the Hilbert space $L^2(\Omega, \omega)$. Let $\mathcal{P}_{V_N}$ denote the orthogonal projection onto V_N. As $M \to \infty$, the above convergence means that $\mathbb{K}$ approaches a matrix representation of $\mathcal{P}_{V_N} \mathcal{K} \mathcal{P}_{V_N}^*$. Thus, EDMD is a Galerkin method in the large data limit $M \to \infty$. There are typically three scenarios where (2.11*a*,*b*) holds:

(i) Random sampling: in the initial definition of EDMD, ω is a probability measure and $\{\boldsymbol{x}^{(m)}\}_{m=1}^M$ are drawn independently according to ω with the quadrature weights $w_m = 1/M$. The strong law of large numbers shows that (2.11*a*,*b*) holds with probability one (Klus, Koltai & Schütte 2016, § 3.4), provided that ω is not supported on a zero level set that is a linear combination of the dictionary (Korda & Mezić 2018, § 4). Convergence is typically at a Monte Carlo rate of $O(M^{-1/2})$ (Caflisch 1998). From an experimental point of view, an example of random sampling could be $\{\boldsymbol{x}^{(m)}\}$ observed with a sampling rate that is lower than the characteristic time period of the system of interest.

(ii) Highorder quadrature: if the dictionary and F are sufficiently regular and we are free to choose the $\{\boldsymbol{x}^{(m)}\}_{m=1}^M$, then it is beneficial to select $\{\boldsymbol{x}^{(m)}\}_{m=1}^M$ as an M-point quadrature rule with weights $\{w_m\}_{m=1}^M$. This can lead to much faster convergence rates in (2.11*a*,*b*) (Colbrook & Townsend 2021), but can be difficult if d is large.

(iii) Ergodic sampling: for a single fixed initial condition $\boldsymbol{x}_0$ and $\boldsymbol{x}^{(m)} = F^{m-1}(\boldsymbol{x}_0)$ (i.e. data collected along one trajectory with $M_1 = 1$ in (2.1)), if the dynamical system is ergodic, then one can use Birkhoff's ergodic theorem to show (2.11*a*,*b*) (Korda & Mezić 2018). One chooses $w_m = 1/M$ but the convergence rate is problem dependent (Kachurovskii 1996). An example of ergodic sampling could be a time-resolved PIV dataset of a post-transient flow field over a long time period.

In this paper, we use (i) random sampling, and (iii) ergodic sampling, which are typical for experimental data collection since they arise from long-time trajectory measurements.

The convergence in (2.11*a*,*b*) implies that the EDMD eigenvalues approach the spectrum of $\mathcal{P}_{V_N} \mathcal{K} \mathcal{P}_{V_N}^*$ as $M \to \infty$. Thus, approximating the spectrum of $\mathcal{K}$, $\sigma(\mathcal{K})$, by the eigenvalues of $\mathbb{K}$ is closely related to the so-called finite section method (Böttcher & Silbermann 1983). Since the finite section method can suffer from spectral pollution (spurious modes), spectral pollution is also a concern for EDMD (Williams *et al.* 2015*a*). It is important to have an independent way to measure the accuracy of the candidate eigenvalue–eigenvector pairs, which is what we propose in our ResDMD method presented in § 3.

2.3.1. *Relationship with DMD*

Williams *et al.* (2015*a*) observed that if we choose a dictionary of observables of the form $\psi_j(\boldsymbol{x}) = e_j^*\boldsymbol{x}$ for $j = 1, \ldots, d$, where $\{e_j\}$ denote the canonical basis, then the matrix $\mathbb{K} = (\sqrt{W}\Psi_X)^\dagger\sqrt{W}\Psi_Y$ with $w_m = 1/M$ is the transpose of the usual DMD matrix

$$\mathbb{K}_{DMD} = \Psi_Y^\top \Psi_X^{\top\dagger} = \Psi_Y^\top\sqrt{W}(\Psi_X^\top\sqrt{W})^\dagger = ((\sqrt{W}\Psi_X)^\dagger\sqrt{W}\Psi_Y)^\top = \mathbb{K}^\top. \tag{2.12}$$

Thus, DMD can be interpreted as producing a Galerkin approximation of the Koopman operator using the set of linear monomials as basis functions. EDMD can be considered an extension allowing nonlinear functions in the dictionary.

2.4. *Koopman mode decomposition*

Given an observable $h : \Omega \to \mathbb{C}$, we approximate it by an element of V_N via

$$h(\boldsymbol{x}) \approx \Psi(\boldsymbol{x})(\Psi_X^* W \Psi_X)^\dagger \Psi_X^* W (h(\boldsymbol{x}^{(1)}) \quad \cdots \quad h(\boldsymbol{x}^{(M)}))^\top. \tag{2.13}$$

Assuming that the quadrature rule converges, the right-hand side converges to the projection $\mathcal{P}_{V_N}h$ in $L^2(\Omega, \omega)$ as $M \to \infty$. Assuming that $\mathbb{K}V = V\Lambda$ for a diagonal matrix Λ of eigenvalues and a matrix V of eigenvectors, we obtain

$$h(\boldsymbol{x}) \approx \Psi(\boldsymbol{x})V[V^{-1}(\sqrt{W}\Psi_X)^\dagger\sqrt{W}(h(\boldsymbol{x}^{(1)}) \quad \cdots \quad h(\boldsymbol{x}^{(M)}))^\top]. \tag{2.14}$$

As a special case, and vectorising, we have the Koopman mode decomposition

$$\boldsymbol{x} \approx \underbrace{\Psi(\boldsymbol{x})V}_{\text{Koopman e-functions}} \underbrace{[V^{-1}(\sqrt{W}\Psi_X)^\dagger\sqrt{W}(\boldsymbol{x}^{(1)} \quad \cdots \quad \boldsymbol{x}^{(M)})^\top]}_{N\times d \text{ matrix of Koopman modes}}. \tag{2.15}$$

The jth Koopman mode corresponds to the jth row of the matrix in square brackets, and ΨV is a quasimatrix of approximate Koopman eigenfunctions.

3. Residual DMD

We now present ResDMD for computing spectral properties of Koopman operators, that in turn allows us to analyse fluid flow structures such as turbulence. ResDMD, first introduced in Colbrook & Townsend (2021), uses an additional matrix constructed from the measured data, $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^M$. The key difference between ResDMD and other DMD-type algorithms is that we construct Galerkin approximations for not only $\mathcal{K}$, but also $\mathcal{K}^*\mathcal{K}$. This difference allows us to have rigorous convergence guarantees for ResDMD and obtain error guarantees on the approximation. In other words, we can tell *a posteriori* which parts of the computed spectra and Koopman modes are reliable, thus rectifying issues such as spectral pollution that arise in previous DMD-type methods.

3.1. *ResDMD and a new data matrix*

Whilst EDMD obtains eigenvalue–eigenvector pairs for an approximation of the Koopman operator, it cannot verify the accuracy of the computed pairs. However, we show below how one can confidently identify true physical turbulent flow structures by rigorously rejecting inaccurate predictions (spurious modes). This rigorous measure of accuracy is the linchpin of our new ResDMD method and is shown pictorially in figure 1.

We approximate residuals to measure the accuracy of a candidate eigenvalue–eigenvector pair (λ, g) (which could, for example, be computed from $\mathbb{K}$ or some other method). For $\lambda \in \mathbb{C}$ and $g = \Psi \boldsymbol{g} \in V_N$, the squared relative residual is

$$\frac{\displaystyle\int_\Omega |[\mathcal{K}g](\boldsymbol{x}) - \lambda g(\boldsymbol{x})|^2 \,\mathrm{d}\omega(\boldsymbol{x})}{\displaystyle\int_\Omega |g(\boldsymbol{x})|^2 \,\mathrm{d}\omega(\boldsymbol{x})} = \frac{\displaystyle\sum_{j,k=1}^{N} \overline{g_j} g_k [\langle \mathcal{K}\psi_k, \mathcal{K}\psi_j\rangle - \lambda\langle \psi_k, \mathcal{K}\psi_j\rangle - \bar{\lambda}\langle \mathcal{K}\psi_k, \psi_j\rangle + |\lambda|^2\langle \psi_k, \psi_j\rangle]}{\displaystyle\sum_{j,k=1}^{N} \overline{g_j} g_k \langle \psi_k, \psi_j\rangle}. \tag{3.1}$$

We approximate this residual using the same quadrature rule in § 2.2

$$res(\lambda, g)^2 = \frac{\boldsymbol{g}^*[\Psi_Y^* W \Psi_Y - \lambda[\Psi_X^* W \Psi_Y]^* - \bar{\lambda}\Psi_X^* W \Psi_Y + |\lambda|^2 \Psi_X^* W \Psi_X]\boldsymbol{g}}{\boldsymbol{g}^*[\Psi_X^* W \Psi_X]\boldsymbol{g}}. \tag{3.2}$$

As well as the matrices $\Psi_X^* W \Psi_X$ and $\Psi_X^* W \Psi_Y$ found in EDMD, (3.2) has the additional matrix $\Psi_Y^* W \Psi_Y$. Note that this additional matrix is no more expensive to compute than either of the EDMD matrices, and uses the same data for its construction as EDMD. Since $\Psi_Y^* W \Psi_Y = \sum_{m=1}^{M} w_m \Psi(\boldsymbol{y}^{(m)})^* \Psi(\boldsymbol{y}^{(m)})$, if the quadrature rule converges then

$$\lim_{M\to\infty} [\Psi_Y^* W \Psi_Y]_{jk} = \langle \mathcal{K}\psi_k, \mathcal{K}\psi_j\rangle. \tag{3.3}$$

Hence, $\Psi_Y^* W \Psi_Y$ formally corresponds to an approximation of $\mathcal{K}^*\mathcal{K}$. We also have

$$\lim_{M\to\infty} res(\lambda, g)^2 = \frac{\displaystyle\int_\Omega |[\mathcal{K}g](\boldsymbol{x}) - \lambda g(\boldsymbol{x})|^2 \,\mathrm{d}\omega(\boldsymbol{x})}{\displaystyle\int_\Omega |g(\boldsymbol{x})|^2 \,\mathrm{d}\omega(\boldsymbol{x})}. \tag{3.4}$$

In Colbrook & Townsend (2021), it was shown that the quantity $res(\lambda, g)$ can be used to rigorously compute spectra and pseudospectra of $\mathcal{K}$. Next, we summarise some of the algorithms and results of Colbrook & Townsend (2021).

3.2. *ResDMD for computing spectra and pseudospectra*

3.2.1. *Avoiding spurious eigenvalues*

Algorithm 1 uses the residual defined in (3.2) to avoid spectral pollution (spurious modes). As is usually done, we assume that $\mathbb{K}$ is diagonalisable. We first find the eigenvalues and eigenvectors of $\mathbb{K}$, i.e. we solve $(\Psi_X^* W \Psi_X)^\dagger (\Psi_X^* W \Psi_Y)\boldsymbol{g} = \lambda \boldsymbol{g}$. One can solve this eigenproblem directly, but it is often numerically more stable to solve the generalised eigenproblem $(\Psi_X^* W \Psi_Y)\boldsymbol{g} = \lambda(\Psi_X^* W \Psi_X)\boldsymbol{g}$. Afterwards, to avoid spectral pollution, we discard eigenpairs with a relative residual larger than a specified accuracy goal $\epsilon > 0$.

The procedure is a simple modification of EDMD, as the only difference is a clean-up step where spurious eigenpairs are discarded based on their residual. This clean-up step is typically faster to execute than the eigendecomposition in step 2 of algorithm 1. The total

Algorithm 1 : ResDMD for computing eigenpairs without spectral pollution. The corresponding Koopman mode decomposition is given in (3.6).

Input: snapshot data $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}$, quadrature weights $\{w_m\}_{m=1}^{M}$, a dictionary of observables $\{\psi_j\}_{j=1}^{N}$ and an accuracy goal $\epsilon > 0$.

1: Compute $\Psi_X^* W \Psi_X$, $\Psi_X^* W \Psi_Y$, and $\Psi_Y^* W \Psi_Y$, where Ψ_X, Ψ_Y are given in (2.8*a*,*b*).
2: Solve $(\Psi_X^* W \Psi_Y)\boldsymbol{g} = \lambda(\Psi_X^* W \Psi_X)\boldsymbol{g}$ for eigenpairs $\{(\lambda_j, g_{(j)} = \Psi \boldsymbol{g}_j)\}$.
3: Compute $res(\lambda_j, g_{(j)})$ for all j (see (3.2)) and discard $(\lambda_j, g_{(j)})$ if $res(\lambda_j, g_{(j)}) > \epsilon$.

Output: a collection of accurate eigenpairs $\{(\lambda_j, \Psi \boldsymbol{g}_j) : res(\lambda_j, g_{(j)}) \leq \epsilon\}$.

computational cost of algorithm 1 is $O(N^2M + N^3)$, which is the same as EDMD. The clean up avoids spectral pollution and also removes eigenpairs that are inaccurate due to numerical errors associated with non-normal operators, up to the relative tolerance ϵ. The following result (Colbrook & Townsend 2021, theorem 4.1) makes this precise.

THEOREM 3.1. *Let $\mathcal{K}$ be the associated Koopman operator of the dynamical system (1.1) from which snapshot data are collected. Let Λ_M denote the eigenvalues in the output of algorithm 1. Then, assuming convergence of the quadrature rule in § 2.2*

$$\limsup_{M\to\infty} \max_{\lambda\in\Lambda_M} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \epsilon. \tag{3.5}$$

We can also use algorithm 1 to clean up the Koopman mode decomposition in (2.15). To do this, we simply let $V^{(\epsilon)}$ denote the matrix whose columns are the eigenvectors $\boldsymbol{g}_j$ with $res(\lambda_j, g_{(j)}) \leq \epsilon$ and compute the Koopman mode decomposition with respect to $\Psi_X^{(\epsilon)} = \Psi_X V^{(\epsilon)}$ and $\Psi_Y^{(\epsilon)} = \Psi_Y V^{(\epsilon)}$. Since $(\sqrt{W}\Psi_X V^{(\epsilon)})^{\dagger}\sqrt{W}\Psi_Y V^{(\epsilon)}$ is diagonal, the Koopman mode decomposition now becomes

$$\boldsymbol{x} \approx \underbrace{\Psi(\boldsymbol{x})V^{(\epsilon)}}_{\text{Koopman e-functions}} \underbrace{[(\sqrt{W}\Psi_X V^{(\epsilon)})^{\dagger}\sqrt{W}(\boldsymbol{x}^{(1)} \quad \cdots \quad \boldsymbol{x}^{(M)})^{\top}]}_{N\times d \text{ matrix of Koopman modes}}. \tag{3.6}$$

3.2.2. *Computing the full spectrum and pseudospectra*

Theorem 3.1 tells us that, in the large data limit, algorithm 1 computes eigenvalues inside the ϵ-pseudospectrum of $\mathcal{K}$. Hence, algorithm 1 avoids spectral pollution and returns reasonable eigenvalues. The ϵ-pseudospectrum of $\mathcal{K}$ is (Trefethen & Embree 2005)

$$\sigma_\epsilon(\mathcal{K}) := \mathrm{cl}(\{\lambda \in \mathbb{C} : \|(\mathcal{K} - \lambda)^{-1}\| > 1/\epsilon\}) = \mathrm{cl}(\cup_{\|\mathcal{B}\|<\epsilon}\sigma(\mathcal{K} + \mathcal{B})), \tag{3.7}$$

where cl denotes the closure of a set, and $\lim_{\epsilon\downarrow 0}\sigma_\epsilon(\mathcal{K}) = \sigma(\mathcal{K})$. Despite theorem 3.1, algorithm 1 may not approximate the whole of $\sigma_\epsilon(\mathcal{K})$, even as $M \to \infty$ and $N \to \infty$. This is because the eigenvalues of $\mathcal{P}_{V_N}\mathcal{K}\mathcal{P}_{V_N}^*$ may not approximate the whole of $\sigma(\mathcal{K})$ as $N \to \infty$ (Colbrook & Townsend 2021; Mezic 2022), even if $\cup_N V_N$ is dense in $L^2(\Omega, \omega)$.

To overcome this issue, algorithm 2 computes practical approximations of ϵ-pseudospectra with rigorous convergence guarantees. Assuming convergence of the quadrature rule in § 2.2, in the limit $M \to \infty$, the key quantity

$$\tau_N(\lambda) = \min_{\boldsymbol{g}\in\mathbb{C}^N} res(\lambda, \Psi \boldsymbol{g}), \tag{3.8}$$

is an upper bound for $\|(\mathcal{K} - \lambda)^{-1}\|^{-1}$. The output of algorithm 2 is guaranteed to be inside the ϵ-pseudospectrum of $\mathcal{K}$. As $N \to \infty$ and the grid of points is refined, algorithm 2

Algorithm 2 : ResDMD for estimating ϵ-pseudospectra.

Input: snapshot data $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}$, quadrature weights $\{w_m\}_{m=1}^{M}$, a dictionary of observables $\{\psi_j\}_{j=1}^{N}$, an accuracy goal $\epsilon > 0$ and a grid $z_1, \ldots, z_k \in \mathbb{C}$.

1: Compute $\Psi_X^* W \Psi_X$, $\Psi_X^* W \Psi_Y$, and $\Psi_Y^* W \Psi_Y$, where Ψ_X, Ψ_Y are given in (2.8*a*,*b*).
2: For each z_j, compute $\tau_N(z_j) = \min_{\boldsymbol{g} \in \mathbb{C}^N} res(z_j, \Psi \boldsymbol{g})$ (see (3.2)), which is a generalised SVD problem, and the corresponding singular vectors $\boldsymbol{g}_j$.

Output: estimate of the ϵ-pseudospectrum $\{z_j : \tau_N(z_j) < \epsilon\}$ and ϵ-approximate eigenfunctions $\{\boldsymbol{g}_j : res(z_j, \Psi \boldsymbol{g}) < \epsilon\}$.

converges to the pseudospectrum uniformly on compact subsets of $\mathbb{C}$ (Colbrook & Townsend 2021, theorems B.1 and B.2). In practice the grid $\{z_j\}_{j=1}^{k}$ is chosen for plotting purposes (e.g. in figure 3). Strictly speaking, we converge to the approximate point pseudospectrum, a more complicated algorithm leads to computation of the full pseudospectrum – see Colbrook & Townsend (2021, appendix B). For brevity, we have not included a statement of the results. We can then compute the spectrum by taking $\epsilon \downarrow 0$. Algorithm 2 also computes observables g with $res(\lambda, g) < \epsilon$, known as ϵ-approximate eigenfunctions. The computational cost of algorithm 2 is $O(N^2 M + kN^3)$. However, the computation of τ_N can be done in parallel over the k grid points, and we can refine the grid adaptively near regions of interest.

3.3. *Choice of dictionary*

When d is large, it can be impractical to store or form the matrix $\mathbb{K}$ since the initial value of N may be huge. In other words, we run into the curse of dimensionality. We consider two common methods to overcome this issue:

(i) DMD: in this case, the dictionary consists of linear functions (see § 2.3.1). It is standard to form a low-rank approximation of $\sqrt{W}\Psi_X$ via a truncated SVD as

$$\sqrt{W}\Psi_X \approx U_r \Sigma_r V_r^*. \tag{3.9}$$

Here, $\Sigma_r \in \mathbb{C}^{r \times r}$ is diagonal with strictly positive diagonal entries, and $V_r \in \mathbb{C}^{N \times r}$ and $U_r \in \mathbb{C}^{M \times r}$ have $V_r^* V_r = U_r^* U_r = I_r$. We then form the matrix

$$\tilde{\mathbb{K}} = (\sqrt{W}\Psi_X V_r)^{\dagger} \sqrt{W}\Psi_Y V_r = \Sigma_r^{-1} U_r^* \sqrt{W}\Psi_Y V_r = V_r^* \mathbb{K} V_r \in \mathbb{C}^{r \times r}. \tag{3.10}$$

Note that to fit into our Galerkin framework, this matrix is the transpose of the DMD matrix that is commonly computed in the literature.

(ii) Kernelised EDMD (kEDMD): kEDMD (Williams *et al.* 2015*b*) aims to make EDMD practical for large d. Supposing that Ψ_X is of full rank, kEDMD constructs a matrix with an identical formula to (3.10) when $r = M$, for which we have the equivalent form

$$\tilde{\mathbb{K}} = (\Sigma_M^{\dagger} U_M^*)(\sqrt{W}\Psi_Y \Psi_X^* \sqrt{W})(U_M \Sigma_M^{\dagger}). \tag{3.11}$$

Suitable matrices U_M and Σ_M can be recovered from the eigenvalue decomposition $\sqrt{W}\Psi_X \Psi_X^* \sqrt{W} = U_M \Sigma_M^2 U_M^*$. Moreover, both matrices $\sqrt{W}\Psi_X \Psi_X^* \sqrt{W}$ and $\sqrt{W}\Psi_Y \Psi_X^* \sqrt{W}$ can be computed using inner products. The kEDMD applies the kernel trick to compute the inner products in an implicitly defined reproducing

Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ (Scholkopf 2001). A positive–definite kernel function $\mathcal{S} : \Omega \times \Omega \to \mathbb{R}$ induces a feature map $\varphi : \mathbb{R}^d \to \mathcal{H}$ so that $\langle \varphi(\boldsymbol{x}), \varphi(\boldsymbol{y}) \rangle_{\mathcal{H}} = \mathcal{S}(\boldsymbol{x}, \boldsymbol{y})$. This leads to an implicit choice of (typically nonlinear) dictionary $\Psi(\boldsymbol{x})$, that is dependent on the kernel function, so that $\Psi(\boldsymbol{x})\Psi(\boldsymbol{y})^* = \langle \varphi(\boldsymbol{x}), \varphi(\boldsymbol{y}) \rangle_{\mathcal{H}} = \mathcal{S}(\boldsymbol{x}, \boldsymbol{y})$. Often $\mathcal{S}$ can be evaluated in $O(d)$ operations, meaning that $\tilde{\mathbb{K}}$ is constructed in $O(dM^2)$ operations.

In either of these two cases, the approximation of $\mathcal{K}$ is equivalent to using a new dictionary with feature map $\Psi(\boldsymbol{x})V_r \in \mathbb{C}^{1\times r}$. In the case of DMD, we found it beneficial to use the mathematically equivalent choice $\Psi(\boldsymbol{x})V_r\Sigma_r^{-1}$, which is numerically better conditioned. To see why, note that $\sqrt{W}\Psi_X V_r \Sigma_r^{-1} \approx U_r$ and U_r has orthonormal columns.

3.3.1. *The problem of vanishing residual estimates*

Suppose that $\sqrt{W}\Psi_X V_r$ has full row rank, so that $r = M$, and that $\boldsymbol{v} \in \mathbb{C}^M$ is an eigenvector of $\tilde{\mathbb{K}}$ with eigenvalue λ. This means that $(\sqrt{W}\Psi_X V_M)^\dagger \sqrt{W}\Psi_Y V_M \boldsymbol{v} = \lambda \boldsymbol{v}$. Since $\sqrt{W}\Psi_X V_M$ has independent rows, $\sqrt{W}\Psi_X V_M (\sqrt{W}\Psi_X V_M)^\dagger = I_M$ and hence $\sqrt{W}\Psi_Y V_M \boldsymbol{v} = \lambda \sqrt{W}\Psi_X V_M \boldsymbol{v}$. The corresponding observable is $g(\boldsymbol{x}) = \Psi(\boldsymbol{x})V_M\boldsymbol{v}$ and the numerator of $res(\lambda, g)^2$ in (3.2) is equal to $\|\sqrt{W}\Psi_Y V_M \boldsymbol{v} - \lambda\sqrt{W}\Psi_X V_M \boldsymbol{v}\|^2$. It follows that $res(\lambda, g) = 0$. Similarly, if r is too large, $res(\lambda, g)$ will be a bad approximation of the true residual.

In other words, the regime $r \sim M$ prevents the large data convergence ($M \to \infty$) of the quadrature rule and (3.4), which holds for a fixed basis and hence a fixed basis size. In turn, this prevents us from being able to apply the results of § 3.2. We next discuss overcoming this issue by using two sets of snapshot data; these could arise from two independent tests of the same system, or by partitioning the measured data into two groups.

3.3.2. *Using two subsets of the snapshot data*

A simple remedy to avoid the problem in § 3.3.1 is to consider two sets of snapshot data. We consider an initial set $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$, that we use to form our dictionary. We then apply ResDMD to the computed dictionary with a second set of snapshot data $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, allowing us to prove convergence as $M'' \to \infty$.

How to acquire a second set of snapshot data depends on the problem and method of data collection. Given snapshot data with random and independent $\{\boldsymbol{x}^{(m)}\}$, one can split up the snapshot data into two parts. For initial conditions that are distributed according to a high-order quadrature rule, if one already has access to M' snapshots, then one must typically go back to the original dynamical system and request M'' further snapshots. For ergodic sampling along a trajectory, we can let $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ correspond to the initial $M' + 1$ points of the trajectory ($\tilde{\boldsymbol{x}}^{(m)} = F^{m-1}(\boldsymbol{x}_0)$ for $m = 1, \ldots, M'$) and let $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$ correspond to the initial $M'' + 1$ points of the trajectory ($\hat{\boldsymbol{x}}^{(m)} = F^{m-1}(\boldsymbol{x}_0)$ for $m = 1, \ldots, M''$).

In the case of DMD, algorithm 3 summarises the two-stage process. Often a suitable choice of N can be obtained by studying the decay of the singular values of the data matrix or the condition number of the matrix $\Psi_X^* W \Psi_X$. When $M' \leq d$, the computational cost of steps 2 and 3 of algorithm 3 are $O(dM'^2)$ and $O(NdM'')$, respectively.

In the case of kEDMD, we follow Colbrook & Townsend (2021), and algorithm 4 summarises the two-stage process. The choice of kernel $\mathcal{S}$ determines the dictionary and the best choice depends on the application. In the following experiments, we use

Algorithm 3 : **ResDMD with DMD selected observables.**

Input: snapshot data $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ and $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, positive integer $N \leq M'$.

1: Set $\Psi_{\mathrm{DMD}}(\boldsymbol{x}) = \begin{bmatrix} e_1^* \boldsymbol{x} & \cdots & e_d^* \boldsymbol{x} \end{bmatrix}$.

2: Compute a truncated SVD

$$\frac{1}{\sqrt{M'}} \left(\Psi_{\mathrm{DMD}}(\tilde{\boldsymbol{x}}^{(1)})^\top \quad \cdots \quad \Psi_{\mathrm{DMD}}(\tilde{\boldsymbol{x}}^{(M')})^\top \right)^\top \approx U_N \Sigma_N V_N^*.$$

3: Apply algorithms 1 and 2 with the matrices

$$\Psi_X = \begin{pmatrix} \Psi_{\mathrm{DMD}}(\hat{\boldsymbol{x}}^{(1)}) \\ \vdots \\ \Psi_{\mathrm{DMD}}(\hat{\boldsymbol{x}}^{(M'')}) \end{pmatrix} V_N \Sigma_N^\dagger, \quad \Psi_Y = \begin{pmatrix} \Psi_{\mathrm{DMD}}(\hat{\boldsymbol{y}}^{(1)}) \\ \vdots \\ \Psi_{\mathrm{DMD}}(\hat{\boldsymbol{y}}^{(M'')}) \end{pmatrix} V_N \Sigma_N^\dagger. \tag{1}$$

Output: spectral properties of Koopman operator according to algorithms 1 and 2.

Algorithm 4 : **ResDMD with kEDMD selected observables.**

Input: snapshot data $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ and $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, positive–definite kernel function $\mathcal{S} : \Omega \times \Omega \to \mathbb{R}$, and positive integer $N \leq M'$.

1: Apply kEDMD to $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ with kernel $\mathcal{S}$ to compute the matrices $\tilde{\mathbb{K}}$, $U_{M'}$ and $\Sigma_{M'}$ using the kernel trick.

2: Compute the dominant N eigenvalues of $\tilde{\mathbb{K}}$ and stack the corresponding eigenvectors column by column into $Z \in \mathbb{C}^{M' \times N}$.

3: Apply a QR decomposition to orthogonalise Z to $Q = \begin{bmatrix} Q_1 & \cdots & Q_N \end{bmatrix} \in \mathbb{C}^{M' \times N}$.

4: Apply algorithms 1 and 2 with $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$ and the dictionary $\{\psi_j\}_{j=1}^{N}$, where

$$\psi_j(\boldsymbol{x}) = \begin{bmatrix} \mathcal{S}(\boldsymbol{x}, \tilde{\boldsymbol{x}}^{(1)}) & \mathcal{S}(\boldsymbol{x}, \tilde{\boldsymbol{x}}^{(2)}) & \cdots & \mathcal{S}(\boldsymbol{x}, \tilde{\boldsymbol{x}}^{(M')}) \end{bmatrix} (U_{M'} \Sigma_{M'}^\dagger) Q_j, \qquad 1 \leq j \leq N.$$

Output: spectral properties of Koopman operator according to algorithms 1 and 2.

the Laplacian kernel $\mathcal{S}(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\gamma \|\boldsymbol{x} - \boldsymbol{y}\|)$, where γ is the reciprocal of the average ℓ^2-norm of the snapshot data after it is shifted to have mean zero.

We can now apply the theory of § 3.2 in the limit $M'' \to \infty$. It is well known that the eigenvalues computed by DMD and kEDMD may suffer from spectral pollution. However, crucially in our setting, we do not directly use these methods to compute spectral properties of $\mathcal{K}$. Instead, we only use them to select a good dictionary of size N, after which our rigorous ResDMD algorithms can be used. Moreover, we use $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$ to check the quality of the constructed dictionary. By studying the residuals and using the error control in ResDMD, we can tell *a posteriori* whether the dictionary is satisfactory and whether N is sufficiently large.

Finally, it is worth pointing out that the above choices of dictionaries are certainly not the only choices. ResDMD can be applied to any suitable choice. For example, one could use other data-driven methods such as diffusion kernels (Giannakis *et al.* 2018) or trained neural networks (Li *et al.* 2017; Murata, Fukami & Fukagata 2020).

4. Spectral measures

Many physical systems described by (1.1) are measure preserving (preserve volume). Examples include Hamiltonian flows (Arnold 1989), geodesic flows on Riemannian manifolds (Dubrovin, Fomenko & Novikov 1991, Chapter 5), Bernoulli schemes in probability theory (Shields 1973) and ergodic systems (Walters 2000). The Koopman operator $\mathcal{K}$ associated with a measure-preserving dynamical system is an isometry, i.e. $\|\mathcal{K}g\| = \|g\|$ for all observables $g \in \mathcal{D}(\mathcal{K}) = L^2(\Omega, \omega)$. In this case, spectral measures provide a way of including continuous spectra in the Koopman mode decomposition. This inclusion is beneficial in the case of turbulent flows where *a priori* knowledge of the spectra (e.g. whether it contains a continuous part) may be unknown. The methods described in this section allow us to compute continuous spectra.

4.1. *Spectral measures and Koopman mode decompositions*

Given an observable $g \in L^2(\Omega, \omega)$ of interest, the spectral measure of $\mathcal{K}$ with respect to g is a measure ν_g defined on the periodic interval $[-\pi, \pi]_{per}$. If g is normalised so that $\|g\| = 1$, then ν_g is a probability measure, otherwise ν_g is a positive measure of total mass $\|g\|^2$. We provide a mathematical description of spectral measures in § A for completeness. The reader should think of these measures as supplying a diagonalisation of $\mathcal{K}$.

For example, the decomposition in (A3) provides important information on the evolution of dynamical systems. Suppose that there is no singular continuous spectrum, then any $g \in L^2(\Omega, \omega)$ can be written as

$$g = \sum_{\lambda \in \sigma_p(\mathcal{K})} c_\lambda \varphi_\lambda + \int_{[-\pi,\pi]_{per}} \phi_{\theta,g}\, \mathrm{d}\theta, \tag{4.1}$$

where $\sigma_p(\mathcal{K})$ is the set of eigenvalues of $\mathcal{K}$, the φ_λ are the eigenfunctions of $\mathcal{K}$, c_λ are expansion coefficients and the density of the absolutely continuous part of ν_g is $\rho_g(\theta) = \langle \phi_{\theta,g}, g\rangle$. One should think of $\phi_{\theta,g}$ as a 'continuously parametrised' collection of eigenfunctions. Using (4.1), one obtains the Koopman mode decomposition (Mezić 2005)

$$g(\boldsymbol{x}_n) = [\mathcal{K}^n g](\boldsymbol{x}_0) = \sum_{\lambda \in \sigma_p(\mathcal{K})} c_\lambda \lambda^n \varphi_\lambda(\boldsymbol{x}_0) + \int_{[-\pi,\pi]_{per}} \mathrm{e}^{\mathrm{i}n\theta} \phi_{\theta,g}(\boldsymbol{x}_0)\, \mathrm{d}\theta. \tag{4.2}$$

For characterisations of the dynamical system in terms of these decompositions, see Halmos (2017), Mezić (2013) and Zaslavsky (2002). Typically, the eigenvalues correspond to isolated oscillation frequencies in the fluid flow and the growth rates of stable and unstable modes, whilst the continuous spectrum corresponds to chaotic motion. Computing the measures ν_g provides us with a diagonalisation of the nonlinear dynamical system in (1.1).

4.2. *Spectral measures and autocorrelations*

The Fourier coefficients of ν_g are given by

$$\widehat{\nu_g}(n) := \frac{1}{2\pi} \int_{[-\pi,\pi]_{per}} \mathrm{e}^{-\mathrm{i}n\theta}\, \mathrm{d}\nu_g(\theta), \quad n \in \mathbb{Z}. \tag{4.3}$$

These Fourier coefficients can be expressed in terms of correlations $\langle \mathcal{K}^n g, g\rangle$ and $\langle g, \mathcal{K}^n g\rangle$ (Colbrook & Townsend 2021). That is, for $g \in L^2(\Omega, \omega)$

$$\widehat{\nu_g}(n) = \frac{1}{2\pi}\langle \mathcal{K}^{-n} g, g\rangle, \quad n < 0, \quad \widehat{\nu_g}(n) = \frac{1}{2\pi}\langle g, \mathcal{K}^n g\rangle, \quad n \geq 0. \tag{4.4a,b}$$

From (4.4a,b), we see that $\widehat{\nu_g}(-n) = \overline{\widehat{\nu_g}(n)}$ for $n \in \mathbb{Z}$. In particular, ν_g is completely determined by the forward-time dynamical autocorrelations $\langle g, \mathcal{K}^n g\rangle$ with $n \geq 0$. Equivalently, the spectral measure of $\mathcal{K}$ with respect to $g \in L^2(\Omega, \omega)$ is a signature for the forward-time dynamics of (1.1). This connection allows us to interpret spectral measures as an infinite-dimensional version of power spectra in § 4.5.

4.3. *Computing autocorrelations*

To approximate spectral measures, we make use of the relation (4.4a,b) between the Fourier coefficients of ν_g and the autocorrelations $\langle g, \mathcal{K}^n g\rangle$. There are typically three ways to compute the autocorrelations, corresponding to the three scenarios discussed in § 2.3:

1. Random sampling: the autocorrelations can be approximated as

$$\langle g, \mathcal{K}^n g\rangle \approx \frac{1}{M}\sum_{m=1}^{M} g(\boldsymbol{x}^{(m)})\overline{[\mathcal{K}^n g](\boldsymbol{x}^{(m)})}. \tag{4.5}$$

2. High-order quadrature: the autocorrelations can be approximated as

$$\langle g, \mathcal{K}^n g\rangle = \int_{\Omega} g(\boldsymbol{x})\overline{[\mathcal{K}^n g](\boldsymbol{x})}\,\mathrm{d}\omega(\boldsymbol{x}) \approx \sum_{m=1}^{M} w_m g(\boldsymbol{x}^{(m)})\overline{[\mathcal{K}^n g](\boldsymbol{x}^{(m)})}. \tag{4.6}$$

3. Ergodic sampling: the autocorrelations can be approximated as

$$\langle g, \mathcal{K}^n g\rangle \approx \frac{1}{M-n}\sum_{m=0}^{M-n-1} g(\boldsymbol{x}_m)\overline{[\mathcal{K}^n g](\boldsymbol{x}_m)} = \frac{1}{M-n}\sum_{m=0}^{M-n-1} g(\boldsymbol{x}_m)\overline{g(\boldsymbol{x}_{m+n})}. \tag{4.7}$$

The first two methods require multiple snapshots of the form $\{\boldsymbol{x}_0^{(m)}, \ldots, \boldsymbol{x}_n^{(m)}\}$, with $[\mathcal{K}^n g](\boldsymbol{x}_0^{(m)}) = g(\boldsymbol{x}_n^{(m)})$. Ergodic sampling only requires a single trajectory, and the ergodic averages in (4.7) can be rewritten as discrete (non-periodic) convolutions. Thus, by zero padding, the fast Fourier transform computes all averages simultaneously and rapidly.

4.4. *Computing spectral measures*

We suppose now that one has already computed the autocorrelations $\langle g, \mathcal{K}^n g\rangle$ for $0 \leq n \leq N_{ac}$. In practice, given a fixed data set of M snapshots, we choose a suitable N_{ac} by checking for convergence of the autocorrelations by comparing with smaller values of M. Following Colbrook & Townsend (2021), we define an approximation to ν_g as

$$\nu_{g,N_{ac}}(\theta) = \sum_{n=-N_{ac}}^{N_{ac}} \varphi\left(\frac{n}{N_{ac}}\right)\widehat{\nu_g}(n)\,\mathrm{e}^{\mathrm{i}n\theta}. \tag{4.8}$$

The function $\varphi : [-1, 1] \to \mathbb{R}$ is often called a filter function (Tadmor 2007; Hesthaven 2017) and $\varphi(x)$ is close to 1 when x is close to 0, but tapers to 0 near $x = \pm 1$. Algorithm 5 summarises the approach, and the choice of φ affects the convergence.

Algorithm 5 : Approximating spectral measures from autocorrelations.

Input: trajectory data, a filter φ, and an observable $g \in L^2(\Omega, \omega)$.

1: Approximate the autocorrelations $\widehat{v_g}(n) = \frac{1}{2\pi}\langle g, \mathcal{K}^n g\rangle$ for $0 \le n \le N_{ac}$. (The precise value of N_{ac} and the approach depends on the trajectory data.)
2: Set $\widehat{v_g}(-n) = \overline{\widehat{v_g}(n)}$ for $1 \le n \le N_{ac}$.

Output: the function $\nu_{g,N_{ac}}(\theta) = \sum_{n=-N_{ac}}^{N_{ac}} \varphi\left(\frac{n}{N_{ac}}\right) \widehat{v_g}(n)\, \mathrm{e}^{\mathrm{i}n\theta}$.

For $m \in \mathbb{N}$, suppose that φ is even, continuous and compactly supported on $[-1, 1]$ with

(a) $\varphi \in \mathcal{C}^{m-1}([-1, 1])$;
(b) $\varphi(0) = 1$ and $\varphi^{(n)}(0) = 0$ for any integer $1 \le n \le m - 1$;
(c) $\varphi^{(n)}(1) = 0$ for any integer $0 \le n \le m - 1$,
(d) $\varphi|_{[0,1]} \in \mathcal{C}^{m+1}([0, 1])$.

Then, we have the following forms of convergence (Colbrook & Townsend 2021, § 3):

(i) Under suitable (local) smoothness assumptions, $|\nu_{g,N_{ac}}(\theta) - \rho_g(\theta)| = O(N_{ac}^{-m} \log(N_{ac}))$,
(ii) A suitable rescaling of $\nu_{g,N_{ac}}(\theta_0)$ approximates the point spectrum at θ_0,
(iii) For any $\phi \in \mathcal{C}^{n,\alpha}([-\pi, \pi]_{per})$,

$$\left| \int_{[-\pi,\pi]_{per}} \phi(\theta) \nu_{g,N_{ac}}(\theta)\, \mathrm{d}\theta - \int_{[-\pi,\pi]_{per}} \phi(\theta)\, \mathrm{d}v_g(\theta) \right| \lesssim \|\phi\|_{\mathcal{C}^{n,\alpha}} (N_{ac}^{-(n+\alpha)} + N_{ac}^{-m} \log(N_{ac})). \tag{4.9}$$

This last property is known as weak convergence. One should think of $\nu_{g,N_{ac}}$ as a smooth function that approximates the spectral measure v_g to order $O(N_{ac}^{-m} \log(N_{ac}))$, with a frequency smoothing scale of $O(N_{ac}^{-m})$.

One can also compute spectral measures without autocorrelations. Colbrook & Townsend (2021) also develop high-order methods based on rational kernels that approximate spectral measures using the ResDMD matrices. Computing suitable residuals allows an adaptive and rigorous selection of the smoothing parameter used in the convolution. In particular, this set-up allows us to deal with general snapshot data $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}$ without the need for long trajectories. For brevity, we omit the details.

4.5. *Interpretation of Koopman spectral measures as power spectra*

We can interpret algorithm 5 as an approximation of the power spectrum of the signal $g(\boldsymbol{x}(t))$, given by Glegg & Devenport (2017, chapter 8)

$$S_{gg}(f) = \int_{-T}^{T} R_{gg}(t)\, \mathrm{e}^{2\pi \mathrm{i} f t}\, \mathrm{d}t, \tag{4.10}$$

over a time window $[-T, T]$ and for frequency f (measured in Hertz). Here, $R_{gg}(t)$ is the delay autocorrelation function, defined for $t \ge 0$ as

$$R_{gg}(t) = \langle g, g \circ F_t \rangle, \tag{4.11}$$

where F_t is the forward-time propagator for a time step t. In particular, we have

$$R_{gg}(n\Delta t) = \begin{cases} \langle g, \mathcal{K}^n g\rangle & \text{if } n \geq 0 \\ \overline{\langle g, \mathcal{K}^{-n} g\rangle} = \langle \mathcal{K}^{-n} g, g\rangle & \text{otherwise} \end{cases}. \tag{4.12}$$

We apply windowing and multiply the integrand in (4.10) by the filter function $\varphi(t/T)$. We then discretise the resulting integral using the trapezoidal rule (noting that the endpoint contributions vanish) with step size $\Delta t = T/N_{ac}$ to obtain the approximation

$$\frac{S_{gg}(f)}{2\pi\Delta t} \approx \sum_{n=-N_{ac}}^{N_{ac}} \varphi\left(\frac{n}{N_{ac}}\right) \frac{R_{gg}(n\Delta t)}{2\pi} \mathrm{e}^{\mathrm{i}n(2\pi f\Delta t)} = \sum_{n=-N_{ac}}^{N_{ac}} \varphi\left(\frac{n}{N_{ac}}\right) \widehat{\nu_g}(n)\, \mathrm{e}^{\mathrm{i}n(2\pi f\Delta t)}. \tag{4.13}$$

It follows that $\nu_{g,N_{ac}}(2\pi f\Delta t)$ can be understood as a discretised version of $S_{gg}(f)/(2\pi\Delta t)$ over the time window $[-T, T]$. Taking the limit $N_{ac} \to \infty$ with $(2\pi\Delta t)f = \theta$, we see that ν_g is an appropriate limit of the power spectrum with time resolution Δt. There are two key benefits of using ν_g. First, we do not explicitly periodically extend the signal to compute autocorrelations and thus avoid the problem of broadening. Instead, we window in the frequency domain. Second, we have rigorous convergence theory as $N_{ac} \to \infty$. We compare spectral measures with power spectra in § 6.

5. Example I: flow past a cylinder wake

We first verify our method by considering the classic example of low Reynolds number flow past a circular cylinder, as focused on by Bagheri (2013). Due to its simplicity and its relevance in engineering, this is one of the most studied examples in modal-analysis techniques (Rowley & Dawson 2017, table 3), (Chen *et al.* 2012; Taira *et al.* 2020). We consider the post-transient regime with $Re = 100$, corresponding to periodic oscillations on a stable limit cycle. The Koopman operator of the flow has a pure point spectrum with a lattice structure on the unit circle (Bagheri 2013) (this is not to be confused with the two-dimensional lattice for the regime on and near the limit cycle that includes transient modes inside the unit disc).

5.1. *Computational set-up*

The flow around a circular cylinder of diameter D is obtained using an incompressible, two-dimensional lattice-Boltzmann computational fluid dynamics (CFD) flow solver. The solver uses the two-dimensional 9-velocity lattice model (commonly referred to as D2Q9) lattice and BGKW (Bhatnagar, Gross & Krook 1954) collision models to calculate the velocity field. The temporal resolution (time step) of the flow simulations is such that the Courant–Friedrichs–Lewy number is unity on the uniform numerical grid. However, this results in a very finely resolved flow field and hence a large volume of data. An initial down-sampling study revealed that storing 12 snapshots of flow field data within a period of vortex shedding still enables us to use our analysis tools without affecting the results. We also verified our results against higher grid resolution for the CFD solver. The computational domain size is $18D$ in length and $5D$ in height. The cylinder is positioned $2D$ downstream of the inlet at the mid-height of the domain. The cylinder walls and the side walls are defined as bounce-back no-slip walls, and a parabolic velocity inlet profile is defined at the inlet of the domain such that $Re = 100$. The outlet is defined as a non-reflecting outflow. For a detailed description of the solver, we refer the reader to Józsa *et al.* (2016) and Szőke *et al.* (2017).

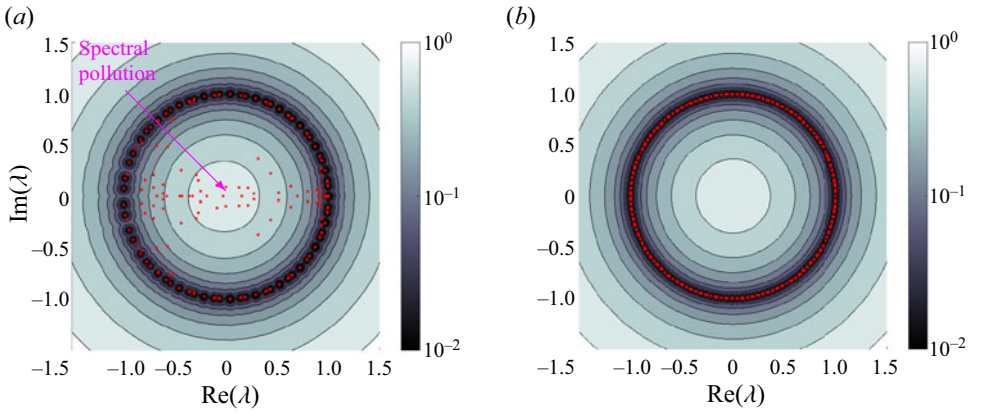


Figure 3. Pseudospectral contours computed using algorithm 2 for the cylinder wake example, using a linear dictionary (*a*) and a nonlinear dictionary (*b*). The eigenvalues of the finite Galerkin matrices $\mathbb{K}$ are shown as red dots in both cases. The computed residuals allow ResDMD to detect spectral pollution (spurious modes); (*a*) $\tau_{200}(\lambda)$, linear dictionary and (*b*) $\tau_{200}(\lambda)$, nonlinear dictionary.

5.2. *Results*

We collect snapshot data of the velocity field along a single trajectory in the post-transient regime of the flow and split the data into two data sets according to § 3.3.2. The first set $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ corresponds to $M' = 500$ snapshots. We then collect $M'' = 1000$ further snapshots $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, where $\hat{\boldsymbol{x}}^{(1)}$ corresponds to approximately 40 time periods of vortex shedding later than $\tilde{\boldsymbol{x}}^{(M')}$. We use algorithms 3 and 4, which we refer to as a linear dictionary (obtained using DMD) and a nonlinear dictionary (obtained using kEDMD), respectively. For both dictionaries, we use $N = 200$ functions. Similar results and conclusions are obtained when varying N, M' and M''.

Figure 3 shows values of τ_N (pseudospectral contours) computed using algorithm 2, where τ_N is the minimal residual in (3.8). The circular contours show excellent agreement with the distance to the spectrum, which is the unit circle in this example. The spectrum corresponds to the closure of the set of eigenvalues which wrap around the circle. The eigenvalues of the finite $N \times N$ Galerkin matrix $\mathbb{K}$ in each case are shown as red dots. The linear dictionary demonstrates spectral pollution, i.e. 'spurious modes', all of which are easily detected by ResDMD (e.g. algorithm 1). Generically, spectral pollution can occur anywhere inside the essential numerical range of $\mathcal{K}$ (Pokrzywa 1979), which in this case is the complex unit disc. Whether spectral pollution occurs is a subtle issue that is highly dependent on the dictionary. In our case, we observed a reduction in spectral pollution for smaller N, with pollution occurring once the linear dictionary cannot approximate the higher-order modes. Detecting when this occurs is a considerable challenge; hence, methods such as algorithm 2 are needed, particularly when we do not know the spectrum *a priori*.

Figure 4 shows the convergence of $\tau_N(-0.5)$ (see (3.8)) as we vary N and M''. As expected, we see convergence as $M'' \to \infty$. Moreover, $\tau_N(-0.5)$ decreases and converges monotonically down to $\|(\mathcal{K} + 0.5)^{-1}\|^{-1} = 0.5$ as N increases. Although both choices of dictionary behave similarly in figure 4, this is not the case for other choices of λ, as is evident near $\lambda = -1$ in figure 3.

Figure 5 shows the eigenvalues of the finite $N \times N$ Galerkin $\mathbb{K}$ as a phase-residual plot. Some eigenvalues computed using the linear dictionary have minimal relative residuals

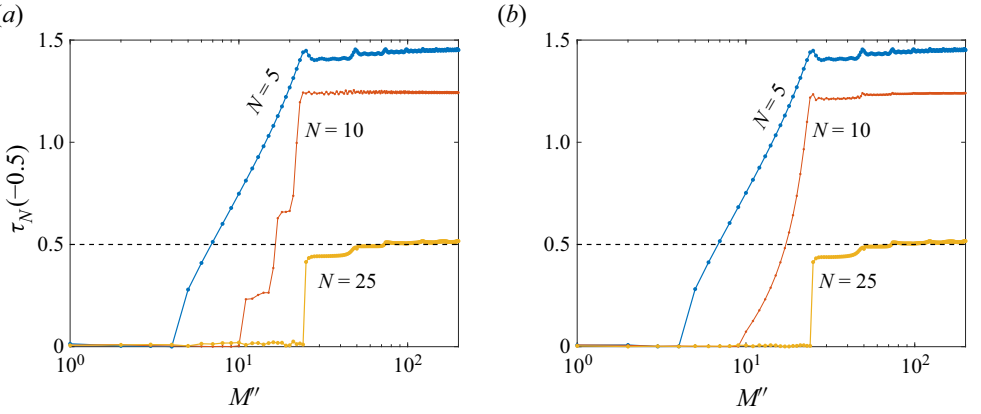


Figure 4. Convergence of $\tau_N(-0.5)$ as M' (the number of snapshots used to build the ResDMD matrices) increases. The plots show the clear monotonicity of $\tau_N(-0.5)$, which decreases to $\|(\mathcal{K}+0.5)^{-1}\|^{-1} = 0.5$ as N increases. (*a*) linear dictionary; (*b*) nonlinear dictionary.

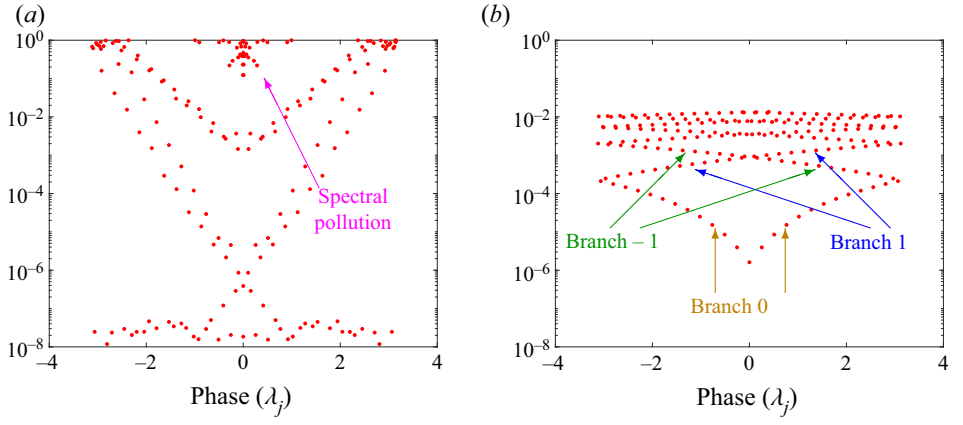


Figure 5. Phase-residual plot of the eigenvalues of $\mathbb{K}$. The linear dictionary has very small residuals for the lower modes, yet also suffers from severe spectral pollution. The nonlinear dictionary demonstrates the lattice structure of the Koopman eigenvalues. Branches of the phase as the eigenvalues wrap around the unit circle are labelled. (*a*) $res(\lambda_j, g_{(j)})$, linear dictionary and (*b*) $res(\lambda_j, g_{(j)})$, nonlinear dictionary.

of approximately 10^{-8}. Since we compute the squared relative residual in (3.2) and then take a square root, and due to the floating point arithmetic used in the software, these small residuals are effectively the level of $\sqrt{\epsilon_{mach}}$. The linear dictionary also has severe spectral pollution. In contrast, the nonlinear dictionary captures the lattice structure of the eigenvalues much better. We have labelled the different branches of the phase as the eigenvalues (powers of the fundamental eigenvalues) wrap around the unit circle.

To investigate these points further, figure 6 plots the errors of the eigenvalues, the relative residuals, and the errors of the associated eigenspaces. For each case of dictionary, we first compute a reference eigenvalue $\lambda_1 \approx 0.9676 + 0.2525i$ corresponding to the first mode, and the corresponding eigenfunction $g_{(1)}$. For each j, we compare the computed eigenvalue λ_j with λ_1^j. The computed residual satisfies $res(\lambda_j, g_{(j)}) \geq \tau_{200}(\lambda_j) \geq |\lambda_j - \lambda_1^j|$,

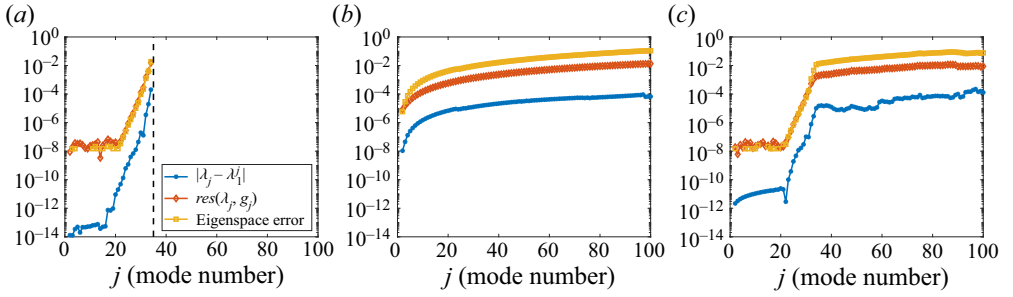


Figure 6. Errors of the computed spectral information of each mode. (*a*) linear dictionary; (*b*) nonlinear dictionary; (*c*) union of dictionaries.

confirming that algorithm 1 provides error bounds on the computed spectra. We evaluate $g^j_{(1)}$ and $g_{(j)}$ at the points $\{\hat{\boldsymbol{x}}^{(m)}\}_{m=1}^{M''}$, and the 'eigenspace error' corresponds to the subspace angle between the linear spans of the resulting vectors (Colbrook & Hansen 2019, (1.4)). The linear dictionary does an excellent job at capturing the lower spectral content, up to about $j = 35$, but is unable to capture the high-order eigenfunctions that are strongly nonlinear (as functions on the state space). In contrast, the nonlinear dictionary does a much better job capturing the higher-order spectral information. For this problem, only a few Koopman modes are needed to reconstruct the flow. However, for some problems, having nonlinear functions in the dictionary is essential to capture the dynamics (e.g. see Brunton *et al.* (2016), and our examples in §§ 7 and 8). One is completely free to choose the dictionary used in ResDMD. For example, one could also use a mixture of the DMD and kEDMD computed dictionaries, or other methods. For example, figure 6(*c*) shows that taking a union of the two dictionaries can capture the low order modes with high accuracy and also cope with the nonlinearities of the high order modes.

Figures 7 and 8 show examples of Koopman modes (see (2.15)) for the x component of the velocity, computed using the linear and nonlinear dictionaries, respectively. Each Koopman eigenfunction is normalised so that the vector $\sqrt{W}\Psi_X V(:, j)$ has norm 1. Modes 1 and 2 show excellent agreement between the two dictionary choices and can also be compared with Taira *et al.* (2020, figure 3). However, for the higher mode (mode 20), the two results bear little similarity. Mode 20 may be seen as a turbulence–turbulence interaction mode that potentially stems from the nonlinear convection term of the Navier–Stokes equation. From a mathematical aspect, the difference in the 20th Koopman mode is because one cannot accurately capture highly nonlinear behaviour with a linear approximation; eventually, there must be a breakdown in the approximation. Whilst the low-order linear approximation does capture the (weakly) nonlinear low-order behaviour, at higher order, the nonlinearity is strengthened. Thus the attempt to capture it linearly is poor. Recall that the Koopman modes correspond to the vector-valued expansion coefficients of the state in the approximate eigenfunctions, as opposed to the eigenfunctions themselves. Thus, the difference indicates that for high-order modes, nonlinearity becomes important in this expansion. We can be assured by the residual measure in the ResDMD algorithm that the modes arising from using the nonlinear dictionary are physical features of the flow and not spurious. ResDMD, therefore, brings certainty to the high-order modal analysis of this classic example, which can be lacking in prior DMD or EDMD approaches.

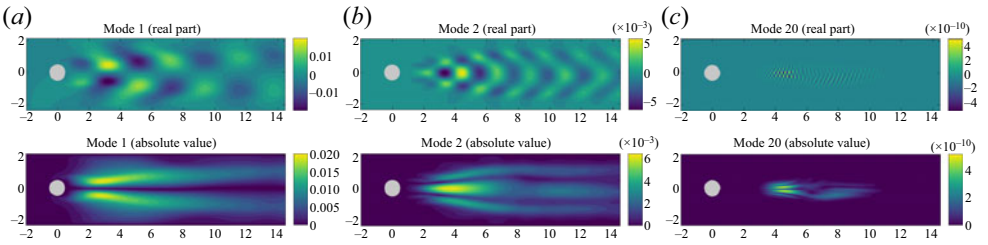


Figure 7. Koopman modes for the cylinder wake, computed using the linear dictionary.

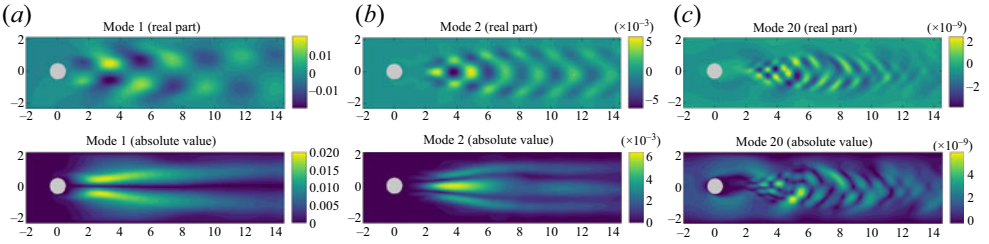


Figure 8. Koopman modes for the cylinder wake, computed using the nonlinear dictionary (see graphical abstract).

6. Example II: embedded shear layer in a turbulent boundary layer flow

We now use our method to analyse a high Reynolds number turbulent flow with a fine temporal resolution. We consider the turbulence structure within a turbulent boundary layer over a flat plate both with and without an embedded shear profile. A shear profile is achieved by permitting a steady flow to be injected through a section of the plate. Turbulent boundary layer flow can be considered a challenging test case for DMD algorithms, particularly when assessing potential nonlinear flow behaviours, and the wide range of length scales present in a high Reynolds number flow problem. On the other hand, the embedded shear layer is anticipated to have a set of characteristic length scales with a broadband energy distribution.

6.1. *Experimental set-up*

We consider two cases of turbulent boundary layer flow, a canonical (baseline) and one with an embedded shear layer. Both cases are generated experimentally and detailed in Szőke, Fiscaletti & Azarpeyvand (2020). For clarity, we briefly recall the key features and illustrate the experimental set-up in figure 9. The baseline case consists of a canonical, zero-pressure-gradient turbulent boundary layer (TBL) passing over a flat plate (no flow control turned on). The friction Reynolds number is $Re_\tau = 1400$. In the experiments, the development of a shear layer is triggered using perpendicular flow injection (active flow control) through a finely perforated sheet, denoted by u_{AFC}. The velocity field is sensed at several positions downstream of the flow control area by traversing a hot-wire probe across the entire boundary layer. The data gathered using the hot-wire sensor provides a fine temporal description of the flow. For a more detailed description of the experimental set-up and the flow field, see Szőke *et al.* (2020). A wide range of downstream positions and flow control severity values were considered in the original study to assess flow control's effects on trailing edge noise. Here, the same data are used, but we focus our attention on one streamwise position (labelled as BL3 in figure 9) and consider a flow injection case

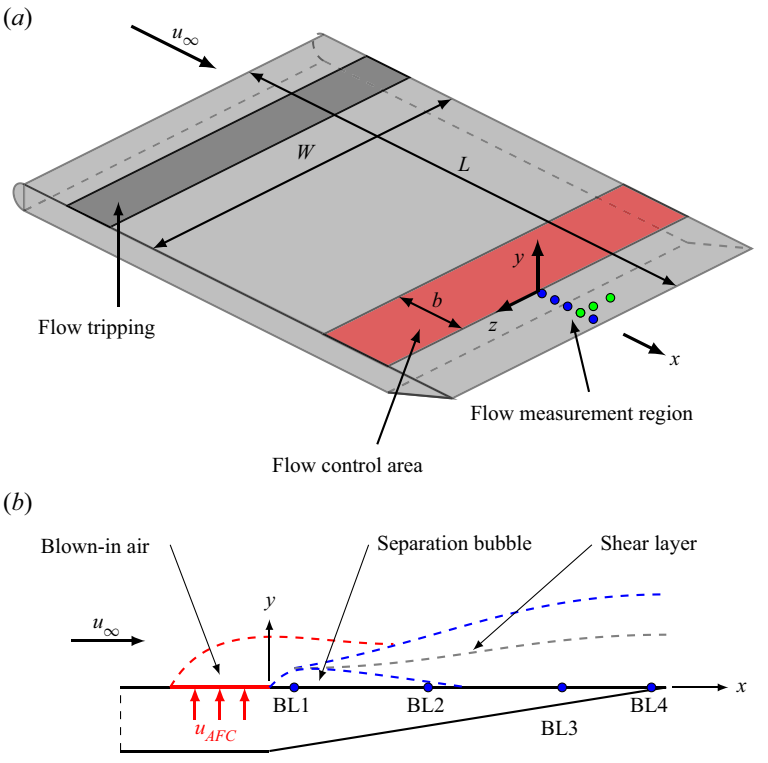


Figure 9. Schematic of the experimental set-up for a boundary layer with embedded shear profile, and side view of the boundary layer near the flow control region.

where a shear layer develops as a result of perpendicular flow injection. Similar qualitative comparisons (against power spectral densities) were found for different sensor locations and different injection angles.

6.2. *Results*

The fine temporal resolution of the flow field enables us to assess the spectral properties of the flow. We use algorithm 5 to calculate the spectral measures of the flow and compare our results with the power spectral density (PSD) obtained using Welch's PSD estimate as described in Szőke *et al.* (2020). For this example, DMD-type methods cannot robustly approximate the Koopman operator of the underlying dynamical system since data are only collected along a single line. However, algorithm 5 can still be used to compute the spectral measures of the Koopman operator.

As a first numerical test, we compute spectral measures of the data collected at 10 % of the undisturbed boundary layer height ($y/\delta_0 = 0.1$). We then integrate this spectral measure against the test function $\phi(\theta) = \exp(\sin(\theta))$. Rather than being a physical example, this is chosen to demonstrate the convergence of our method. We consider the integral computed using the PSD (4.10) with a window size of N_{ac} for direct comparison, and algorithm 5 for choices of filter $\varphi_{hat}(x) = 1 - |x|$, $\varphi_{\cos}(x) = \frac{1}{2}(1 - \cos(\pi x))$, $\varphi_{four}(x) = 1 - x^4(-20|x|^3 + 70x^2 - 84|x| + 35)$ and $\varphi_{bump}(x) = \exp(-(2/(1 - |x|))\exp(-c/x^4))$ ($c \approx 0.109550455106347$). These filters are first, second, fourth and infinite order, respectively. Figure 10 shows the results, where we

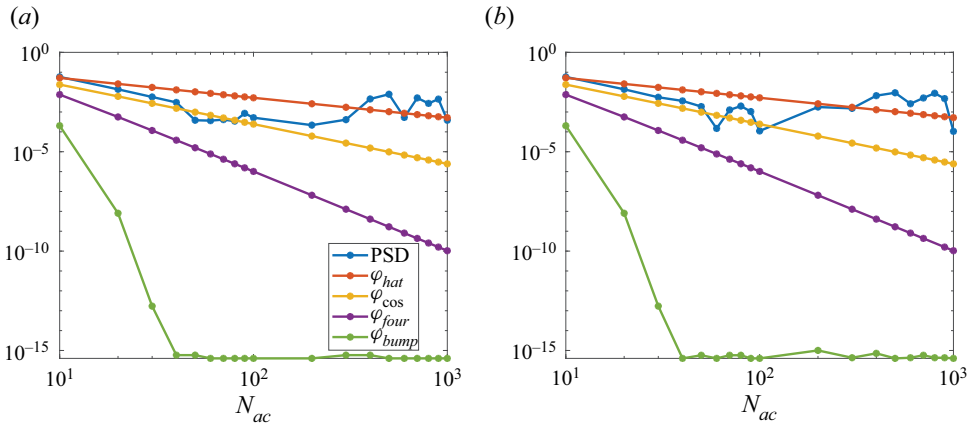


Figure 10. Relative error in integration against the test function $\phi(\theta) = \exp(\sin(\theta))$ for data collected at $y/\delta_0 = 0.1$. Algorithm 5 converges at the expected rates for various filter functions. In contrast, the PSD approximation appears to converge at a second-order rate initially but then stagnates. (*a*) relative error (injection); (*b*) relative error (baseline).

see the expected rates of convergence of algorithm 5. In contrast, the PSD approximation initially appears to converge at a second-order rate and then stagnates.

Figure 11 compares the spectral measure, as found using algorithm 5, with the PSD (4.10) obtained from direct processing of the experimental data, where we have used a window size of N_{ac} for direct comparison. To directly compare with PSD, we use algorithm 5 with the second-order filter $\varphi_{\cos}$. A range of different vertical locations y/δ_0 are considered, where δ_0 is the boundary layer thickness of the baseline case (i.e. $u_{AFC} = 0$ m s^{-1}). Whilst the high-frequency behaviour is almost identical between the two methods, at low frequencies ($<$10 Hz) the spectral measure returns values approximately $\sim$1 dB greater than the PSD processing because the conventional PSD calculation results observe broadening at low frequencies. Figure 12 confirms this and shows the low-frequency values for $N_{ac} = 4000$ and various choices of filter function. In general, higher-order filters lead to a sharper peak at low frequencies. This comparison also holds for different sensor locations (in figure 9) and different injection angles. As we are assured that the spectral measure rigorously converges, this new method provides the more accurate measure of the power at low frequencies as $N_{\text{ac}} \to \infty$.

7. Example III: wall-jet boundary layer flow

As a further example of TBL flow, we now consider a wall-jet boundary layer (George *et al.* 2000; Gersten 2015; Kleinfelter *et al.* 2019). Whilst hot-wire probes enable a very fine temporal resolution of the flow field, they are usually restricted to a single point or line in space, and thus preclude the use of many DMD-type methods. On the other hand, time-resolved (TR) PIV offers both spatial and temporal descriptions of the flow. For this example, we assess the performance of the ResDMD algorithm on a set of TR-PIV data. We consider the boundary layer generated by a thin jet ($h_{jet} = 12.7$ mm) injecting air onto a smooth flat wall. As in example II, this case is challenging for regular DMD approaches due to multiple turbulent scales expected within the boundary layer. This section demonstrates the use of ResDMD for a high Reynolds number, turbulent, complicated flow field.

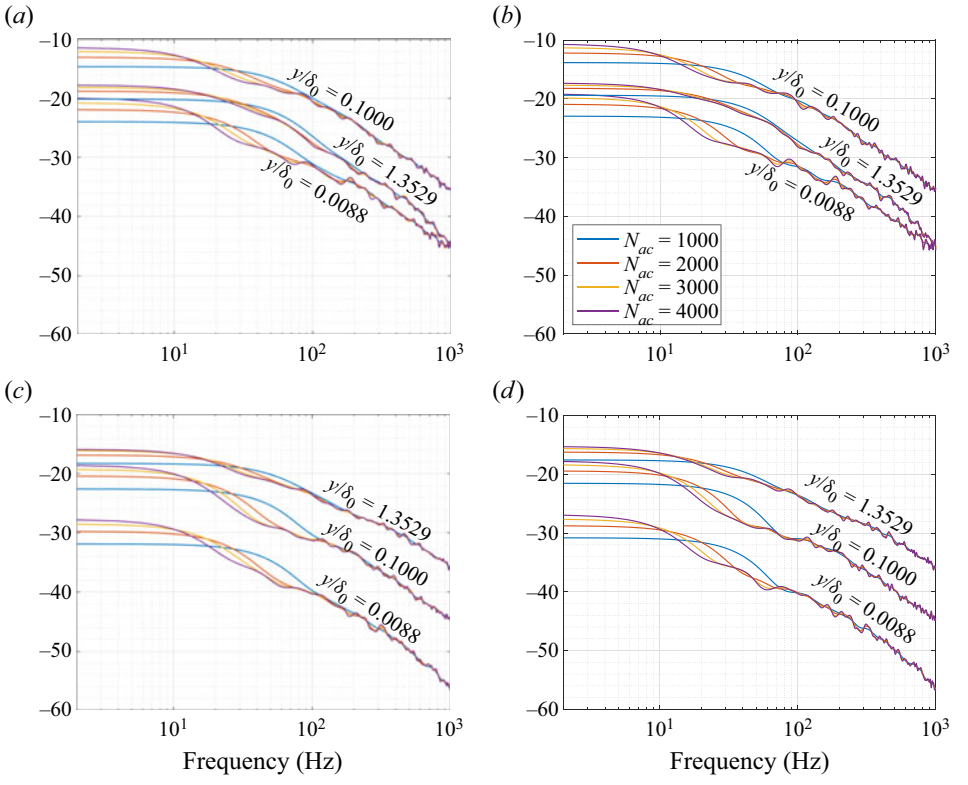


Figure 11. Comparison of the traditional PSD definition (4.10) with spectral measure (4.13) (referenced to 1 m^2 s^{-2}), for baseline and injection flows, at a range of vertical heights within the boundary layer. (*a*) PSD (injection); (*b*) spectral measure (injection); (*c*) PSD (baseline); (*d*) spectral measure (baseline).

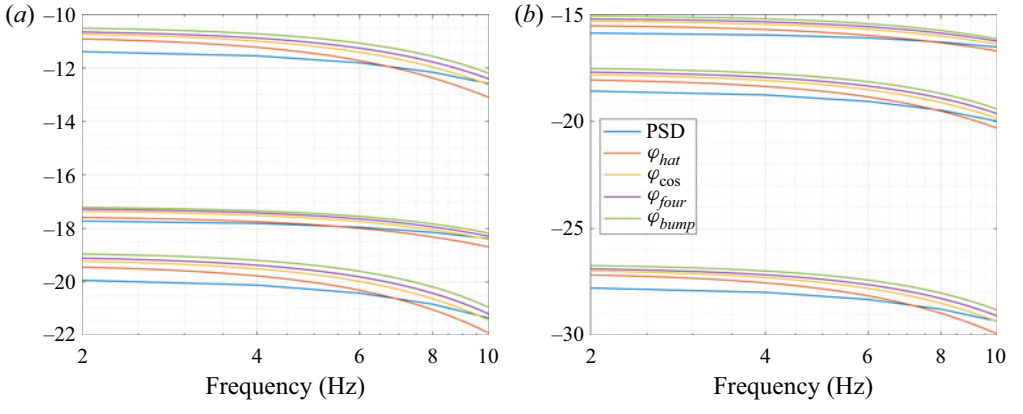


Figure 12. Zoomed in values for low frequencies computed using $N_{ac} = 4000$ and various choices of filter function. In general, higher order filters lead to a sharper peak at low frequencies. (*a*) injection; (*b*) baseline.

7.1. *Experimental set-up*

Experiments using TR-PIV are performed at the Wall Jet Wind Tunnel of Virginia Tech as schematically shown in figure 13. For a detailed description of the facility, we refer to Kleinfelter *et al.* (2019). A two-dimensional two-component TR-PIV system is used

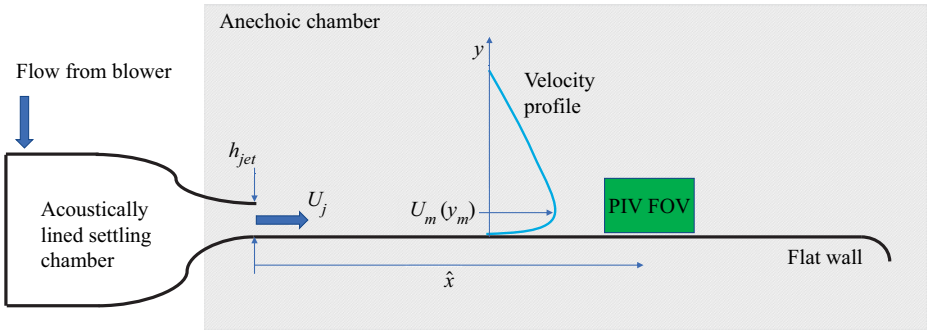


Figure 13. The schematic of the TR-PIV experiments conducted in the Wall Jet Wind Tunnel of Virginia Tech.

to capture the wall-jet flow, and the streamwise origin of the field-of-view (FOV) is $\hat{x} = 1282.7$ mm downstream of the wall-jet nozzle. We use a jet velocity of $U_j = 50\ \mathrm{m\,s^{-1}}$, corresponding to a jet Reynolds number of $Re_{jet} = h_{jet}U_j/\nu = 63.5 \times 10^3$. The length and height of the FOV are approximately 75mm $\times$ 40 mm, and the spatial resolution of the velocity vector field is 0.24 mm. This resolution corresponds to dimension $d = 102, 300$ in (1.1). The high-speed cameras are operated in a double frame mode, with a rate of 12 000 frame pairs per second, resulting in a fine temporal resolution of 0.083 ms. For a more detailed description of the experimental set-up and flow field, see Kleinfelter *et al.* (2019) and Szőke *et al.* (2021).

The associated flow has some special properties. It is self-similar, and its main characteristics (boundary layer thickness, edge velocity, skin friction coefficient, etc.) can be accurately calculated *a priori* through power-law curve fits (Kleinfelter *et al.* 2019). The flow consists of two main regions. Within the region bounded by the wall and the peak in the velocity profile, the flow exhibits the properties of a zero-pressure-gradient TBL. Above this fluid portion, the flow is dominated by a two-dimensional shear layer consisting of rather large, energetic flow structures. While the peak in the velocity profile is $y_m \approx 18$ mm from the wall in our case, the overall thickness of the wall-jet flow is of the order of 200 mm. The PIV experiments must compromise between a good spatial resolution and capturing the entire flow field. In our case, the FOV was not tall enough to capture the entire wall-jet flow field. For this reason, the standard DMD algorithm under-predicts the energies corresponding to the shear-layer portion of the wall-jet flow as the corresponding length scales fall outside of the limits of the FOV. We also verified our results, in particular the presence of the modes in figure 16, by comparing with different spatial and temporal resolutions.

7.2. *Results*

We collect snapshot data of the velocity field from two separate realisations of the experiment. We use the first experiment to generate data $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ with $M' = 2000$, corresponding to 121 boundary layer turnover times. This data is used to select our dictionary of functions. We then use the second experiment to generate data $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$ with $M'' = 12\,000$ (a single trajectory of one second of physical flow time and 728 boundary layer turnover times), which we use to generate the ResDMD matrices, as outlined in § 3.3.2.

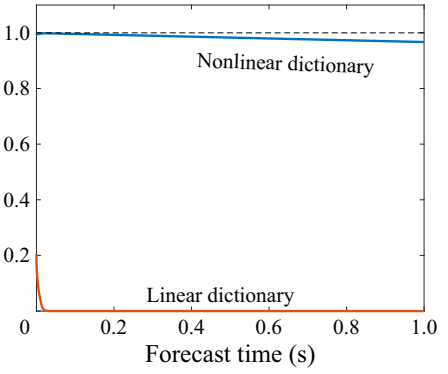


Figure 14. Forecast of total kinetic energy (normalised by the time average of the kinetic energy), averaged over the 12 000 initial conditions. Values closer to 1 correspond to better predictions.

To demonstrate the need for nonlinear functions in our dictionary, we compute the Koopman mode decomposition of the total kinetic energy of the domain using (2.14). Using this decomposition, we compute forecasts of the total energy from a given initial condition of the system. Figure 14 shows the results, where we average over the 12 000 initial conditions in the data set and normalise by the true time-averaged kinetic energy. We use algorithm 3 and 4 with $N = 2000$, which we refer to as a linear dictionary and nonlinear dictionary, respectively. The importance of including nonlinear functions in the dictionary is clear, and corresponds to a much better approximation of the spectral content of $\mathcal{K}$ near 0. This is confirmed in figure 15, where we show the eigenpairs and residuals for each choice of dictionary. The eigenvalues of non-normal matrices can be severely unstable to perturbations, particularly for large N, so we checked the computation of the eigenvalues of $\mathbb{K}$ by comparing with extended precision and predict a bound of approximately 10^{-10} on the numerical error in figure 15. The nonlinear dictionary has smaller residuals and a more resolved spectrum near 1 with less spectral pollution. Unlike the example in § 5, the boundary of the spectra are not circular. Instead, they appear to be centred around a curve of the form $r = \exp(-c|\theta|)$ (shown as a black curve in the plot), corresponding to successive powers of transient modes. Therefore, we only use the nonlinear dictionary for the rest of this section.

To investigate the Koopman modes, we compute the ResDMD Koopman mode decomposition corresponding to (3.6) with an error tolerance of $\epsilon = 0.5$ to get rid of the most severe spectral pollution. The total number of modes used is 656. Figure 16 illustrates a range of Koopman modes that are long-lasting (left-hand column) and transient (right-hand column). Within each figure, the arrows dictate the unsteady fluid structure (computed from the Koopman modes of the velocity fields), with the magnitude of the arrow indicating the local flow speed, and the colour bar denotes the Koopman mode of the velocity magnitude. The corresponding approximate eigenvalues, λ, and residual bound, τ_N, are provided for each mode. Due to residuals, we can accurately select physical transient modes. We also remark that many of these modes were not detected when using a linear dictionary.

The modes in the left column of figure 16 illustrate the range of rolling eddies within the boundary layer, with the smaller structures containing less energy than the largest structures. Interestingly, the third mode in the left column resembles the shape of ejection-like motions within the boundary layer flow ($y/y_m < 1$). At the same time, larger-scale structures above the boundary layer ($y/y_m > 1$) are also visible. This may be

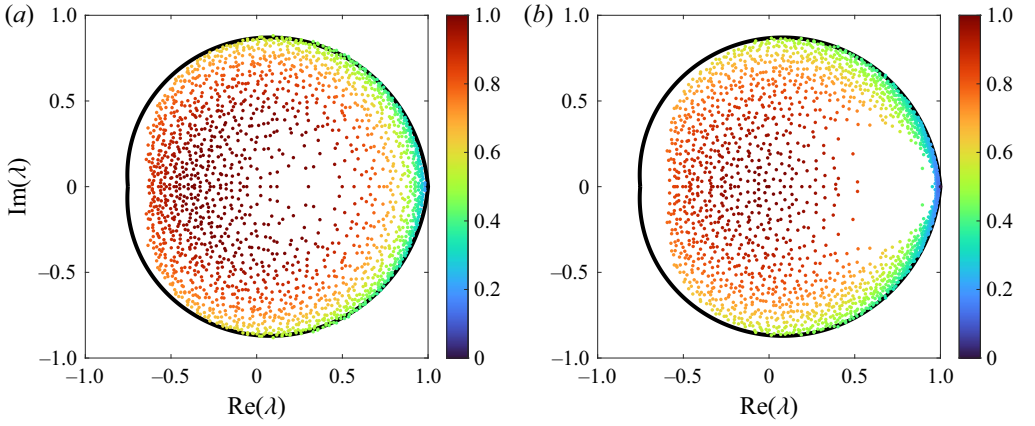


Figure 15. The eigenvalues of finite Galerkin matrix $\mathbb{K}$ for the wall-jet example. For each eigenvalue–eigenvector pair $(\lambda_j, g_{(j)})$, the residual $res(\lambda_j, g_{(j)})$ is shown on a colour scale. When truncating the Koopman mode decomposition for figure 16, we kept those modes with residual at most 0.5. The black curve corresponds to a fit $r = \exp(-c|\theta|)$ of the boundary of the eigenvalues and represents successive powers of modes. (*a*) $res(\lambda_j, g_{(j)})$, linear dictionary and (*b*) $res(\lambda_j, g_{(j)})$, nonlinear dictionary.

interpreted as a nonlinear interaction in the turbulent flow field that is efficiently captured using the ResDMD algorithm. The transient modes in the right column of figure 16 show a richer structure. We interpret these modes as transient, short-lived turbulence behaviour based on our analysis. Again, the use of the conventional (linear) DMD method does not enable the capture of these modes. The uppermost panel may be seen as the shear layer travelling over the boundary layer ($y/y_m > 1$), with the following panel potentially seen as the breakdown of this transient structure into smaller structures. The third panel may be seen as an interaction between an ejection-type vortex and the shear layer; note the ejection-like shape of negative contours below $y/y_m = 1.5$ with a height-invariant positive island of contour at $y/y_m \approx 1.75$. Finally, the bottom-most panel could be seen as a flow uplift out of the boundary layer and further turbulent streaks with counter-rotating properties.

Using the autocorrelation functions of both streamwise (u) and wall-normal (v) velocity components and algorithm 5, we obtain the streamwise wavenumber spectra of the velocity fluctuations, shown in the top row of figure 17. The streamwise-averaged spectral measure of the ergodic component of the flow for both streamwise and vertical unsteady velocities are shown in the bottom row of figure 17. When observing either the wavenumber spectra or the spectral measures of the streamwise velocity component, we see consistent spectra across the different boundary layer heights due to the dominance of the strong shear layer in the wall-jet flow. For the vertical velocity spectra, as one might expect, more energy is observed for both wavenumber spectra and spectral measures with increasing y/y_m.

8. Example IV: acoustic signature of laser-induced plasma

So far, our examples of ResDMD have focused on fluid flows. However, the capability of ResDMD to capture nonlinear mechanics can be applied more broadly. Moreover, the computation of residuals allows an efficient compression of the Koopman mode decomposition through (3.6). As our final example, we demonstrate the use of ResDMD on an acoustic example where the sound source of interest exhibits highly nonlinear properties.

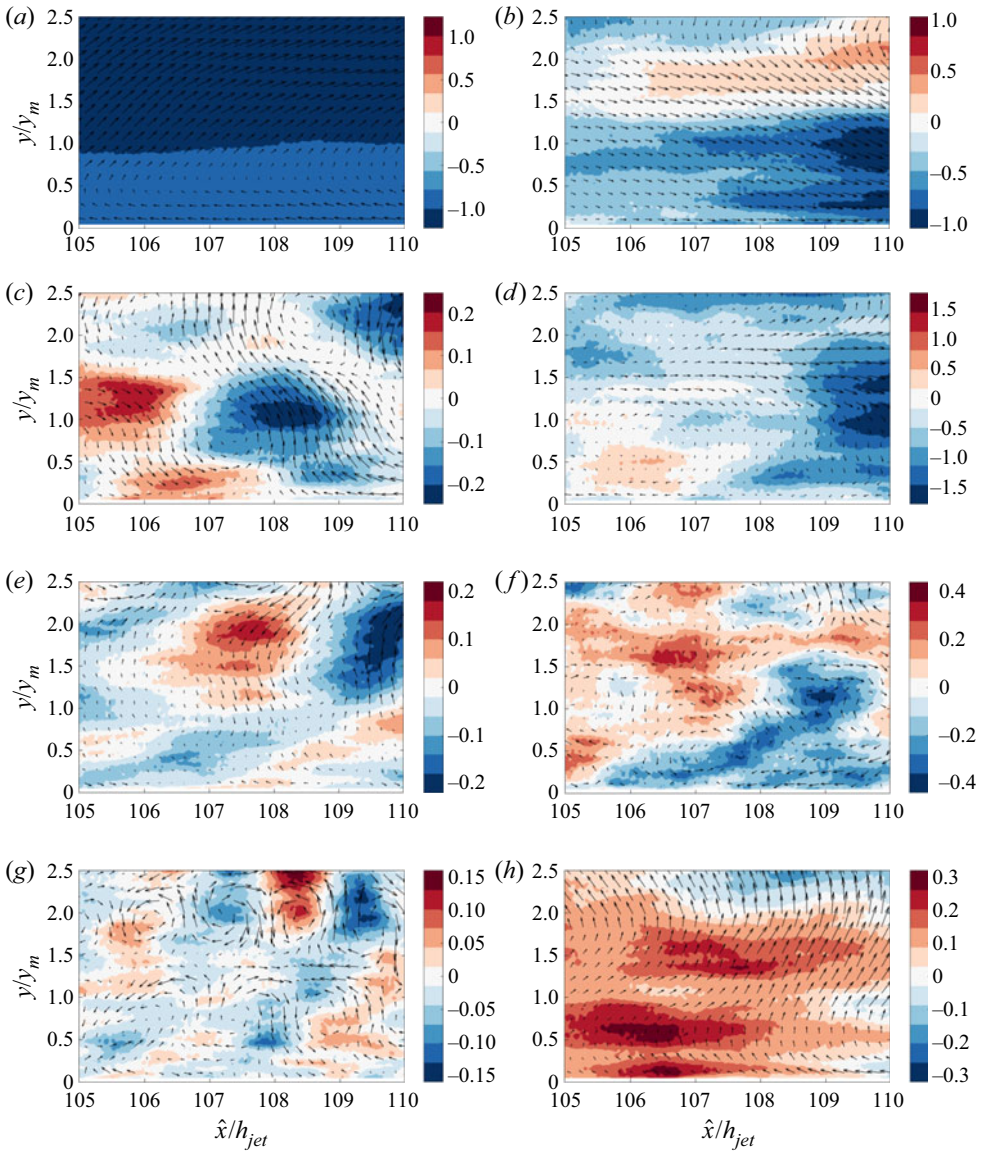


Figure 16. (*a*,*c*,*e*,*g*) A range of long-lasting modes from the ResDMD Koopman mode decomposition. (*b*,*d*,*f*,*h*) A range of transient modes from the ResDMD Koopman mode decomposition. The arrows dictate the unsteady fluid structure (computed from the Koopman modes of the velocity fields), with the magnitude of the arrow indicating the local flow speed, and the colour bar denotes the Koopman mode of the velocity magnitude; (*a*) $\lambda = 1.0000 + 0.0000i$, $\tau_{2000}(\lambda) = 0.0024$, (*b*) $\lambda = 0.9837 + 0.0057i$, $\tau_{2000}(\lambda) = 0.0175$, (*c*) $\lambda = 0.9760 + 0.1132i$, $\tau_{2000}(\lambda) = 0.0539$, (*d*) $\lambda = 0.9528 + 0.0000i$, $\tau_{2000}(\lambda) = 0.0472$, (*e*) $\lambda = 0.9700 + 0.1432i$, $\tau_{2000}(\lambda) = 0.0602$, (*f*) $\lambda = 0.8948 + 0.1065i$, $\tau_{2000}(\lambda) = 0.1105$, (*g*) $\lambda = 0.9439 + 0.2458i$, $\tau_{2000}(\lambda) = 0.0765$ and (*h*) $\lambda = 0.9888 + 0.0091i$, $\tau_{2000}(\lambda) = 0.0146$.

8.1. *Experimental set-up*

We investigate a near-ideal acoustic monopole source that is generated using the laser-optical set-up illustrated in figure 18. When a high-energy laser beam is focused on a point, the air ionises and plasma is generated due to the extremely high electromagnetic

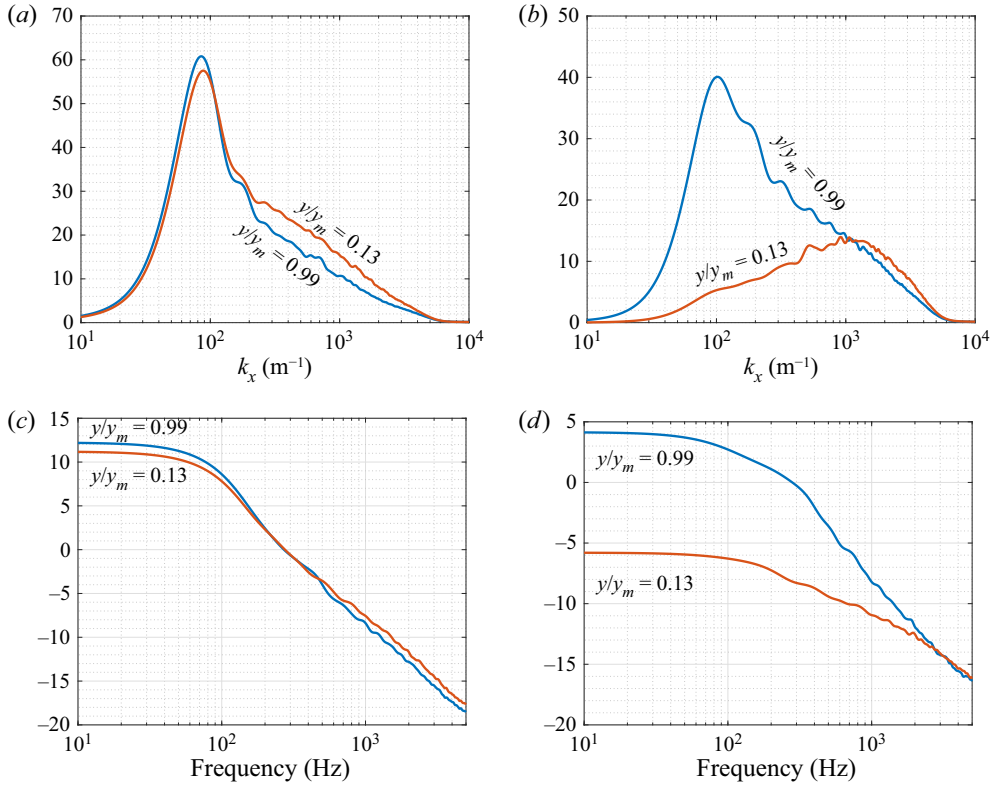


Figure 17. Pre-multiplied streamwise wavenumber spectra (*a*,*b*) and spectral measures (*c*,*d*) for streamwise (u) and wall-normal (v) velocity components; (*a*) $k_x|\phi uu|$, (*b*) $k_x|\phi vv|$, (*c*) spectral measure (u field) and (*d*) spectral measure (v field).

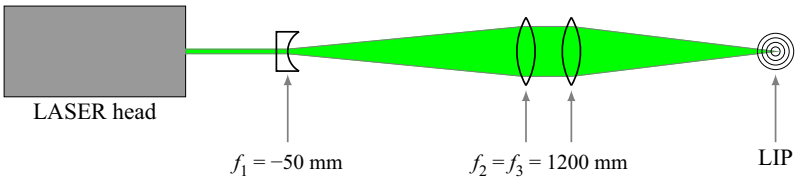


Figure 18. Schematic diagram of the laser beam set-up used to generate the laser-induced plasma.

energy density (of the order of 10^{12} W m^{-2}). As a result of the sudden deposit of energy, the air volume undergoes a sudden expansion that generates a shockwave. The initial propagation characteristics can be modelled using von Neumann's point strong explosion theory (Von Neumann 1963), which was originally developed for nuclear explosion modelling. For our ResDMD analysis, we use laser-induced plasma (LIP) sound signature data measured using a 1/8 inch, Bruel & Kjaer (B&K) type 4138 microphone operated using a B&K Nexus module Szőke *et al.* (2022). The data from the microphone are acquired using an NI-6358 module at a sampling rate of $f_s = 1.25$ MS s^{-1} (million samples per second). With this apparatus, we can resolve the high-frequency nature of the LIP up to 100 kHz. For a detailed description of the laser-optical set-up, experimental apparatus, and data processing procedures, see Szőke *et al.* (2022) and Szőke & Devenport (2021).

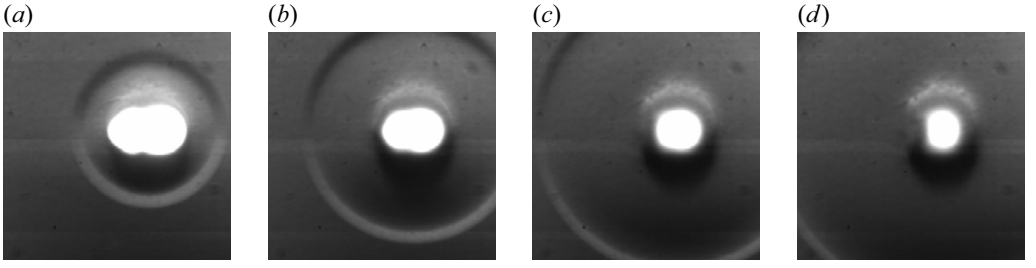


Figure 19. Schlieren images of the initial LIP illustrating the shock wave formation and propagation; (*a*) $t = 5$ μs, (*b*) $t = 10$ μs, (*c*) $t = 15$ μs and (*d*) $t = 20$ μs.

The important acoustic characteristic of the LIP is a short time period of initial supersonic propagation speed, shown as Schlieren images taken over a 15 μs window in figure 19. When observed from the far field, this initial supersonic propagation is observed as a nonlinear characteristic even though the wave speed is supersonic only until approximately 3–4 mm from the optical focal point. During the experiments, 65 realisations of LIP are captured using microphones. Each realisation of LIP is then gated in time such that only the direct propagation path of the LIP remains in the signal (Szőke & Devenport 2021). We use this gated data for our ResDMD analysis.

8.2. *Results*

For a positive integer d, we take the state at time step n to be

$$\boldsymbol{x}_n = (p(n) \quad p(n-1) \quad \cdots \quad p(n-d+1))^{\top} \in \mathbb{R}^d, \tag{8.1}$$

where p is the acoustic pressure. This choice of state corresponds to time-delay embedding, a popular method for DMD-type algorithms (Arbabi & Mezic 2017*a*; Kamb *et al.* 2020; Pan & Duraisamy 2020). When we make future state predictions using the Koopman mode decomposition, there is an additional interpretation of d. The value of d corresponds to the initial time interval we use to make future state predictions. This interval is shown as vertical dashed lines in the plots below.

We split the data into three parts. The first 10 realisations of LIP correspond to $\{\tilde{\boldsymbol{x}}^{(m)}, \tilde{\boldsymbol{y}}^{(m)}\}_{m=1}^{M'}$ and are used to train the dictionary. The next 50 realisations correspond to $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, and are used to construct the ResDMD matrices. The final five realisations are used to test the resulting Koopman mode decomposition. We verified our results by comparing different splittings of the data into sets for dictionary learning, matrix construction and testing. We consider two choices of the dictionary. The first is a linear dictionary computed using algorithm 3. The second is the union of the linear dictionary and the dictionary computed using algorithm 4 with $N = 200$. We refer to this combined dictionary as the nonlinear dictionary.

Figure 20(*a*) shows the results of the Koopman mode decomposition (2.15), applied to the first realisation of the experiment in the test set, with $d = 10$. Namely, we approximate the state as

$$\begin{aligned} \boldsymbol{x}_n &\approx \Psi(\boldsymbol{x}_0)\mathbb{K}^n V[V^{-1}(\sqrt{W}\Psi_X)^{\dagger}\sqrt{W}[\hat{\boldsymbol{x}}^{(1)} \quad \cdots \quad \hat{\boldsymbol{x}}^{(M'')}]^{\top}] \\ &= \Psi(\boldsymbol{x}_0)V\Lambda^n[V^{-1}(\sqrt{W}\Psi_X)^{\dagger}\sqrt{W}[\hat{\boldsymbol{x}}^{(1)} \quad \cdots \quad \hat{\boldsymbol{x}}^{(M'')}]^{\top}]. \end{aligned} \tag{8.2}$$

In particular, we test the Koopman mode decomposition on unseen data corresponding to the test set. The values of p to the left of the vertical dashed line correspond

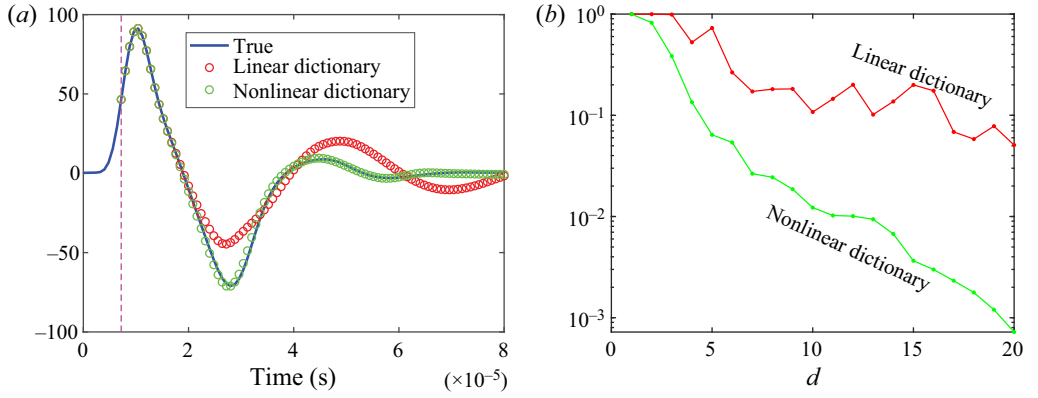


Figure 20. (*a*) Prediction using (8.2) on the first experiment in the test set. The values of p to the left of the vertical dashed line correspond to $\boldsymbol{x}_0$. (*b*) The RMSE averaged over the test set for different values of d.

to $\boldsymbol{x}_0$. The nonlinear dictionary does a much better job of representing the system's nonlinear behaviour. While the linear dictionary can predict the positive pressure peak, it fails to capture the smaller, high-frequency oscillations following the first two large oscillations. These discrepancies between the linear and nonlinear dictionary-based results also pinpoint where nonlinearity in the signal resides. In other words, the nonlinear signature of the pressure wave resides in the negative portion of the wave. Figure 20(*b*) plots the relative mean squared error (RMSE) averaged over the test set for different values of d. The nonlinear dictionary allows an average relative L^2 error of around 6 % for $d = 15$.

Figure 21(*a*) shows the corresponding pseudospectral contours, computed using algorithm 2 with $d = 10$. We can use ResDMD to compress the representation of the dynamics, by ordering the Koopman eigenvalues λ_j, eigenfunctions $g_{(j)}$ and modes according to their (relative) residuals $res(\lambda_j, g_{(j)})$ (defined in (3.2)), which provide a measure of error in the approximate eigenfunctions. For a prescribed value of N', we can alter ϵ in (3.6) to produce a Koopman mode decomposition of the N' eigenfunctions with the smallest residual. In figure 21(*b*), we compare this with a compression based on the modulus of the eigenvalues using 40 modes in each expansion. Ordering the eigenvalues by their residuals gives a much better compression of the dynamics. To investigate this further, figure 22 shows the error curves of the two different compressions for various dictionary sizes and choices of d. This figure suggests that ResDMD may be effective in the construction of reduced-order models.

9. Discussion and conclusions

This paper has implemented a new form of DMD, ResDMD, that permits accurate calculation of residuals during the modal decomposition of general Koopman operators. ResDMD computes spectra and pseudospectra of general Koopman operators with error control, and computes smoothed approximations of spectral measures (including continuous spectra) with explicit high-order convergence theorems. We have illustrated the strengths of this new method through four examples of turbulent flow. For the canonical example of flow past a circular cylinder, the residual allows accurate identification of spurious modes and verification of nonlinear modes. For TBL flow, the residual ensures we can accurately identify physical transient modes (and filter out spurious modes), which leads to a greater understanding of the turbulent structures within the boundary layer.

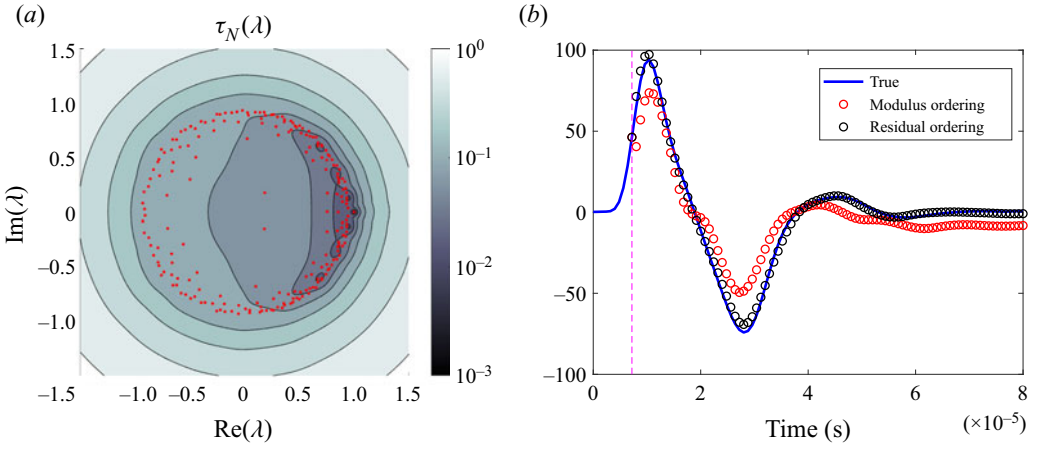


Figure 21. (*a*) Pseudospectral contours, computed using algorithm 2 with the nonlinear dictionary and $d = 10$. The eigenvalues of $\mathbb{K}$ are shown as red dots. (*b*) Prediction using (3.6) on the first experiment in the test set. The values of p to the left of the vertical dashed line correspond to $\boldsymbol{x}_0$. For each type of ordering, we use 40 modes.

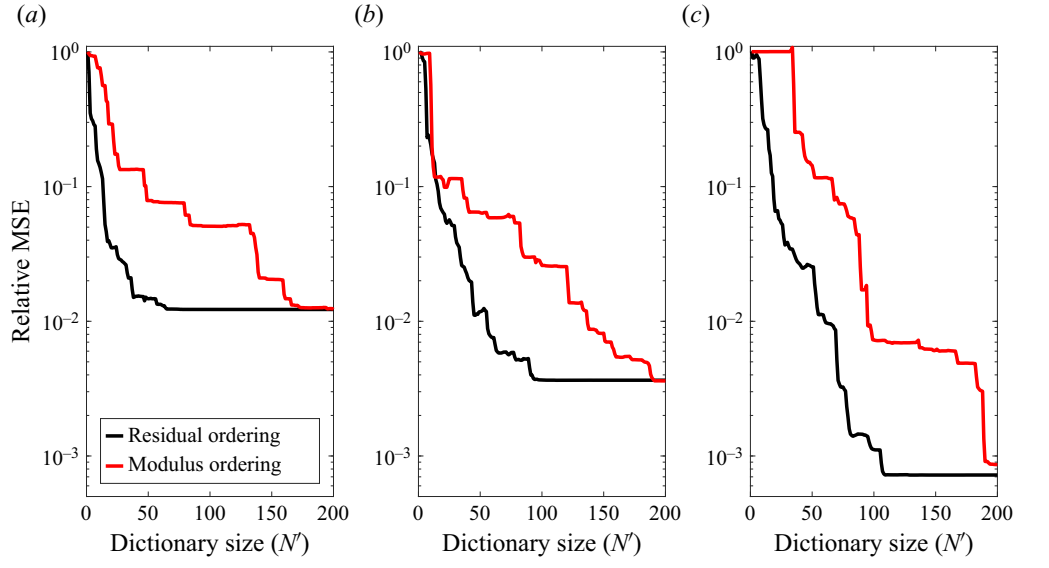


Figure 22. The RMSE averaged over the test set for two types of compression. The 'residual ordering' (black curves) corresponds to ordering approximate eigenvalues based on their residual. The 'modulus ordering' (red curves) corresponds to ordering approximate eigenvalues based on their modulus; (*a*) $d = 10$, (*b*) $d = 15$ and (*c*) $d = 20$.

Further, by relating the Koopman spectral measure to the PSD, we show that processing experimental data via the new ResDMD algorithm reduces low-frequency broadening that is commonplace in conventional PSD calculations. ResDMD also provides convergence guarantees. Finally, by analysing a LIP's acoustic signature, we show that ordering modes by their residual enables greater data compression than ordering modes by their modulus. There are many potential extensions of these ideas. For example, ResDMD may open the door to computing more exotic features of spectra of Koopman operators

(Colbrook; Colbrook & Hansen), and physical questions such as studying and controlling the transition to turbulence (Herrmann *et al.* 2021) or obtaining closure for turbulent transport (Souza, Lutz & Flierl 2022). Other generalisations could include continuous-time dynamical systems, control and stochastic systems. Since calculating the residual is as easy and computationally efficient as traditional DMD, these new methods provide a valuable toolbox for robust and verified Koopmanism.

Acknowledgements. We would like to thank W. Devenport and M. Azarpeyvand for valuable discussions.

Funding. This work was supported by a FSMP Fellowship at École Normale Supérieure (M.J.C.); EPSRC Early-Career Fellowship EP/P015980/1 (L.J.A.); and NSF Grant CBET-1802915 (M.S.).

M.J.C. would like to thank both the Cecil King Foundation and the London Mathematical Society for a Cecil King Travel Scholarship, part of which funded a visit to Virginia Tech. M.S. would like to thank the National Science Foundation, in particular Dr R. Joslin, for their support (grant CBET-1802915) that enabled the capture of the PIV dataset.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.
Matthew J. Colbrook https://orcid.org/0000-0003-4964-9575;
Lorna J. Ayton https://orcid.org/0000-0001-6280-9460;
Máté Szőke https://orcid.org/0000-0002-3768-7956.

Appendix A. Spectral measures of Koopman operators

Any normal matrix $A \in \mathbb{C}^{n\times n}$, i.e. such that $A^*A = AA^*$, has an orthonormal basis of eigenvectors $v_1, \dots, v_n$ for $\mathbb{C}^n$ such that

$$v = \left(\sum_{k=1}^{n} v_k v_k^*\right) v, \quad v \in \mathbb{C}^n \quad \text{and} \quad Av = \left(\sum_{k=1}^{n} \lambda_k v_k v_k^*\right) v, \quad v \in \mathbb{C}^n, \tag{A1a,b}$$

where $\lambda_1, \dots, \lambda_n$ are eigenvalues of A, i.e. $Av_k = \lambda_k v_k$ for $1 \le k \le n$. The projections $v_k v_k^*$ simultaneously decompose the space $\mathbb{C}^n$ and diagonalise the operator A. This idea carries over to our infinite-dimensional setting with two changes.

First, a Koopman operator that is an isometry does not necessarily commute with its adjoint (Colbrook & Townsend 2021) and hence is not normal. Therefore, we must consider a unitary extension of $\mathcal{K}$. Namely, a unitary operator $\mathcal{K}'$ defined on an extended Hilbert space $\mathcal{H}'$ with $L^2(\Omega, \omega) \subset \mathcal{H}'$ Nagy *et al.* (2010, proposition I.2.3). Even though such an extension is not unique, it allows us to understand the spectral information of $\mathcal{K}$. The spectral measures discussed in this paper are canonical and do not depend on the choice of extension Colbrook (2022, proposition 2.1).

Second, if $\mathcal{K}'$ has non-empty continuous spectrum, then the eigenvectors of $\mathcal{K}'$ do not form a basis for $\mathcal{H}'$ or diagonalise $\mathcal{K}'$. Instead, the projections $v_k v_k^*$ in (A1*a*,*b*) can be replaced by a projection-valued measure $\mathcal{E}$ supported on the spectrum of $\mathcal{K}'$ (Reed & Simon 1980, theorem VIII.6). $\mathcal{K}'$ is unitary and hence $\sigma(\mathcal{K}')$ lies inside the unit circle $\mathbb{T}$. The measure $\mathcal{E}$ assigns an orthogonal projector to each Borel measurable subset of $\mathbb{T}$ such that

$$f = \left(\int_{\mathbb{T}} \mathrm{d}\mathcal{E}(y)\right) f \quad \text{and} \quad \mathcal{K}'f = \left(\int_{\mathbb{T}} y\,\mathrm{d}\mathcal{E}(y)\right) f, \quad f \in \mathcal{H}'. \tag{A2a,b}$$

Analogous to (A1*a*,*b*), $\mathcal{E}$ decomposes $\mathcal{H}'$ and diagonalises the operator $\mathcal{K}'$. For example, if $U \subset \mathbb{T}$ contains only eigenvalues of $\mathcal{K}'$ and no other types of spectra, then $\mathcal{E}(U)$ is simply the spectral projector onto the invariant subspace spanned by the corresponding

eigenfunctions. More generally, $\mathcal{E}$ decomposes elements of $\mathcal{H}'$ along the discrete and continuous spectrum of $\mathcal{K}'$ (see (A3)).

Given an observable $g \in L^2(\Omega, \omega) \subset \mathcal{H}'$ of interest, the spectral measure of $\mathcal{K}'$ with respect to g is a scalar measure defined as $\mu_g(U) := \langle \mathcal{E}(U)g, g\rangle$, where $U \subset \mathbb{T}$ is a Borel measurable set. It is convenient to equivalently consider the corresponding measures ν_g defined on the periodic interval $[-\pi, \pi]_{per}$ after a change of variables $\lambda = \exp(\mathrm{i}\,\theta)$. Lebesgue's decomposition (Stein & Shakarchi 2009) splits ν_g into discrete and continuous parts

$$d\nu_g(y) = \underbrace{\sum_{\lambda=\exp(\mathrm{i}\,\theta)\in\sigma_p(\mathcal{K})} \langle \mathcal{P}_\lambda g, g\rangle \delta(y-\theta)\mathrm{d}y}_{\text{discrete part}} + \underbrace{\rho_g(y)\,\mathrm{d}y + d\nu_g^{(sc)}(y)}_{\text{continuous part}}. \tag{A3}$$

Throughout this paper, we use the term 'discrete spectra' to mean the eigenvalues of $\mathcal{K}$, also known as the point spectrum. This set also includes eigenvalues embedded in the essential or continuous spectrum, in contrast to the usual definition of the discrete spectrum. The discrete (or atomic) part of ν_g is a sum of Dirac delta distributions, supported on $\sigma_p(\mathcal{K})$, the set of eigenvalues of $\mathcal{K}$. The coefficient of each δ in the sum is $\langle \mathcal{P}_\lambda g, g\rangle = \|\mathcal{P}_\lambda g\|^2$, where $\mathcal{P}_\lambda$ is the orthogonal spectral projection associated with the eigenvalue λ. The continuous part of ν_g consists of a part that is absolutely continuous with density $\rho_g \in L^1([-\pi, \pi]_{per})$, and a singular continuous component $\nu_g^{(sc)}$.

References

Arbabi, H. & Mezic, I. 2017*a* Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator. *SIAM J. Appl. Dyn. Syst.* **16** (4), 2096–2126.

Arbabi, H. & Mezić, I. 2017*b* Study of dynamics in post-transient flows using Koopman mode decomposition. *Phys. Rev. Fluids* **2** (12), 124402.

Arnold, V.I. 1989 *Mathematical Methods of Classical Mechanics*. Springer-Verlag.

Baddoo, P.J., Herrmann, B., McKeon, B.J. & Brunton, S.L. 2022 Kernel learning for robust dynamic mode decomposition: linear and nonlinear disambiguation optimization. *Proc. R. Soc.* A **478** (2260), 20210830.

Baddoo, P.J., Herrmann, B., McKeon, B.J., Kutz, J.N. & Brunton, S.L. 2021 Physics-informed dynamic mode decomposition (piDMD). Preprint, arXiv:2112.04307.

Bagheri, S. 2013 Koopman-mode decomposition of the cylinder wake. *J. Fluid Mech.* **726**, 596–623.

Basley, J., Pastur, L.R., Lusseyran, F., Faure, T.M. & Delprat, N. 2011 Experimental investigation of global structures in an incompressible cavity flow using time-resolved PIV. *Exp. Fluids* **50** (4), 905–918.

Ben-Artzi, J., Colbrook, M.J., Hansen, A.C., Nevanlinna, O. & Seidel, M. 2020 Computing spectra - On the solvability complexity index hierarchy and towers of algorithms. arXiv:1508.03280.

Bhatnagar, P.L., Gross, E.P. & Krook, M. 1954 A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. *Phys. Rev.* **94** (3), 511–525.

Böttcher, A. & Silbermann, B. 1983 The finite section method for Toeplitz operators on the quarter-plane with piecewise continuous symbols. *Math. Nachr.* **110** (1), 279–291.

Brunton, S.L., Brunton, B.W., Proctor, J.L. & Kutz, J.N. 2016 Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS One* **11** (2), e0150171.

Brunton, S.L., Budišić, M., Kaiser, E. & Kutz, J.N. 2022 Modern Koopman theory for dynamical systems. *SIAM Rev.* **64** (2), 229–340.

Brunton, S.L. & Kutz, J.N. 2019 *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press.

Budišić, M., Mohr, R. & Mezić, I. 2012 Applied Koopmanism. *Chaos* **22** (4), 047510.

Caflisch, R.E. 1998 Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **7**, 1–49.

Chai, X., Iyer, P.S. & Mahesh, K. 2015 Numerical study of high speed jets in crossflow. *J. Fluid Mech.* **785**, 152–188.

Chen, K.K., Tu, J.H. & Rowley, C.W. 2012 Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses. *J. Nonlinear Sci.* **22** (6), 887–915.

Colbrook, M.J. 2020 The foundations of infinite-dimensional spectral computations. PhD thesis, University of Cambridge.

Colbrook, M.J. 2021 Computing spectral measures and spectral types. *Commun. Math. Phys.* **384** (1), 433–501.

Colbrook, M.J. 2022 The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems. Preprint, arXiv:2209.02244.

Colbrook, M.J. 2022 On the computation of geometric features of spectra of linear operators on Hilbert spaces. *Found. Comput. Math.* https://doi.org/10.1007/s10208-022-09598-0.

Colbrook, M.J. & Hansen, A.C. 2019 On the infinite-dimensional QR algorithm. *Numer. Math.* **143** (1), 17–83.

Colbrook, M.J. & Hansen, A.C. The foundations of spectral computations via the solvability complexity index hierarchy. *J. Eur. Math. Soc.* doi:10.4171/JEMS/1289.

Colbrook, M.J., Horning, A. & Townsend, A. 2021 Computing spectral measures of self-adjoint operators. *SIAM Rev.* **63** (3), 489–524.

Colbrook, M.J., Roman, B. & Hansen, A.C. 2019 How to compute spectra with error control. *Phys. Rev. Lett.* **122**, 250201.

Colbrook, M.J. & Townsend, A. 2021 Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems. arXiv:2111.14889.

Das, S., Giannakis, D. & Slawinska, J. 2021 Reproducing kernel Hilbert space compactification of unitary evolution groups. *Appl. Comput. Harmon. Anal.* **54**, 75–136.

Deem, E.A., Cattafesta, L.N., Hemati, M.S., Zhang, H., Rowley, C. & Mittal, R. 2020 Adaptive separation control of a laminar boundary layer using online dynamic mode decomposition. *J. Fluid Mech.* **903**, A21.

Drmac, Z., Mezic, I. & Mohr, R. 2018 Data driven modal decompositions: analysis and enhancements. *SIAM J. Sci. Comput.* **40** (4), A2253–A2285.

Dubrovin, B.A., Fomenko, A.T. & Novikov, S.P. 1991 *Modern Geometry – Methods and Applications Part I. The Geometry of Surfaces, Transformation Groups, and Fields*, vol. 104. Springer Science & Business Media.

George, W.K., Abrahamsson, H., Eriksson, J., Karlsson, R.I., Löfdahl, L. & Wosnik, M. 2000 A similarity theory for the turbulent plane wall jet without external stream. *J. Fluid Mech.* **425**, 367–411.

Gersten, K. 2015 The asymptotic downstream flow of plane turbulent wall jets without external stream. *J. Fluid Mech.* **779**, 351–370.

Giannakis, D. 2019 Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Appl. Comput. Harmon. Anal.* **47** (2), 338–396.

Giannakis, D., Kolchinskaya, A., Krasnov, D. & Schumacher, J. 2018 Koopman analysis of the long-term evolution in a turbulent convection cell. *J. Fluid Mech.* **847**, 735–767.

Glegg, S. & Devenport, W. 2017 *Aeroacoustics of Low Mach Number Flows: Fundamentals, Analysis, and Measurement*. Academic Press.

Grilli, M., Schmid, P.J., Hickel, S. & Adams, N.A. 2012 Analysis of unsteady behaviour in shockwave turbulent boundary layer interaction. *J. Fluid Mech.* **700**, 16–28.

Halmos, P.R. 2017 *Lectures on Ergodic Theory*. Courier Dover Publications.

Herrmann, B., Baddoo, P.J., Semaan, R., Brunton, S.L. & McKeon, B.J. 2021 Data-driven resolvent analysis. *J. Fluid Mech.* **918**, A10.

Hesthaven, J.S. 2017 *Numerical Methods for Conservation Laws: From Analysis to Algorithms*. SIAM.

Hua, J.-C., Gunaratne, G.H., Talley, D.G., Gord, J.R. & Roy, S. 2016 Dynamic-mode decomposition based analysis of shear coaxial jets with and without transverse acoustic driving. *J. Fluid Mech.* **790**, 5–32.

Jovanović, M.R., Schmid, P.J. & Nichols, J.W. 2014 Sparsity-promoting dynamic mode decomposition. *Phys. Fluids* **26** (2), 024103.

Józsa, T., Szőke, M., Teschner, T.-R., Könözsy, L.Z. & Moulitsas, I. 2016 Validation and verification of a 2D lattice Boltzmann solver for incompressible fluid flow. In *European Congress on Computational Methods in Applied Sciences and Engineering*. National Technical University of Athens.

Kachurovskii, A.G. 1996 The rate of convergence in ergodic theorems. *Russian Math. Sur.* **51** (4), 653–703.

Kamb, M., Kaiser, E., Brunton, S.L. & Kutz, J.N. 2020 Time-delay observables for Koopman: theory and applications. *SIAM J. Appl. Dyn. Syst.* **19** (2), 886–917.

Kleinfelter, A.W., Repasky, R., Hari, N., Letica, S., Vishwanathan, V., Organski, L., Schwaner, J., Alexander, W.N. & Devenport, W.J. 2019 Development and calibration of a new anechoic wall jet wind tunnel. In *AIAA Scitech 2019 Forum, AIAA Paper* 2019-1936.

KLUS, S., KOLTAI, P. & SCHÜTTE, C. 2016 On the numerical approximation of the Perron-Frobenius and Koopman operator. *J. Comput. Dyn.* **3** (1), 51–79.

KLUS, S., NÜSKE, F., PEITZ, S., NIEMANN, J.-H., CLEMENTI, C. & SCHÜTTE, C. 2020 Data-driven approximation of the Koopman generator: model reduction, system identification, and control. *Phys.* D*: Nonlinear Phenom.* **406**, 132416.

KOOPMAN, B.O. 1931 Hamiltonian systems and transformation in Hilbert space. *Proc. Natl Acad. Sci. USA* **17** (5), 315.

KOOPMAN, B.O. & VON NEUMANN, J. 1932 Dynamical systems of continuous spectra. *Proc. Natl Acad. Sci. USA* **18** (3), 255.

KORDA, M. & MEZIĆ, I. 2018 On convergence of extended dynamic mode decomposition to the Koopman operator. *J. Nonlinear Sci.* **28** (2), 687–710.

KORDA, M., PUTINAR, M. & MEZIĆ, I. 2020 Data-driven spectral analysis of the Koopman operator. *Appl. Comput. Harmon. Anal.* **48** (2), 599–629.

KUTZ, J.N., BRUNTON, S.L., BRUNTON, B.W. & PROCTOR, J.L. 2016*a Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*. SIAM.

KUTZ, J.N., FU, X. & BRUNTON, S.L. 2016*b* Multiresolution dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **15** (2), 713–735.

LEWIN, M. & SÉRÉ, É. 2010 Spectral pollution and how to avoid it. *Proc. Lond. Math. Soc.* **100** (3), 864–900.

LI, Q., DIETRICH, F., BOLLT, E.M. & KEVREKIDIS, I.G. 2017 Extended dynamic mode decomposition with dictionary learning: a data-driven adaptive spectral decomposition of the Koopman operator. *Chaos* **27** (10), 103111.

LOISEAU, J.-C. & BRUNTON, S.L. 2018 Constrained sparse Galerkin regression. *J. Fluid Mech.* **838**, 42–67.

MEZIĆ, I. 2005 Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **41** (1), 309–325.

MEZIĆ, I. 2013 Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45**, 357–378.

MEZIĆ, I. 2021 Koopman operator, geometry, and learning of dynamical systems. *Not. Am. Math. Soc*. **68** (7), 1087–1105.

MEZIC, I. 2022 On numerical approximations of the Koopman operator. *Mathematics* **10** (7), 1180.

MEZIĆ, I. & BANASZUK, A. 2004 Comparison of systems with complex behavior. *Physica* D*: Nonlinear Phenom.* **197** (1–2), 101–133.

MOTHEAU, E., NICOUD, F. & POINSOT, T. 2014 Mixed acoustic–entropy combustion instabilities in gas turbines. *J. Fluid Mech.* **749**, 542–576.

MURATA, T., FUKAMI, K. & FUKAGATA, K. 2020 Nonlinear mode decomposition with convolutional neural networks for fluid dynamics. *J. Fluid Mech.* **882**, A13.

NAGY, B.S., FOIAS, C., BERCOVICI, H. & KÉRCHY, L. 2010 *Harmonic Analysis of Operators on Hilbert Space*. Springer Science & Business Media.

NOACK, B.R., STANKIEWICZ, W., MORZYŃSKI, M. & SCHMID, P.J. 2016 Recursive dynamic mode decomposition of transient and post-transient wake flows. *J. Fluid Mech.* **809**, 843–872.

PAGE, J. & KERSWELL, R.R. 2018 Koopman analysis of Burgers equation. *Phys. Rev. Fluids* **3** (7), 071901.

PAGE, J. & KERSWELL, R.R. 2019 Koopman mode expansions between simple invariant solutions. *J. Fluid Mech.* **879**, 1–27.

PAGE, J. & KERSWELL, R.R. 2020 Searching turbulence for periodic orbits with dynamic mode decomposition. *J. Fluid Mech.* **886**, A28.

PAN, S. & DURAISAMY, K. 2020 On the structure of time-delay embedding in linear models of non-linear dynamical systems. *Chaos* **30** (7), 073135.

PASQUARIELLO, V., HICKEL, S. & ADAMS, N.A. 2017 Unsteady effects of strong shock-wave/boundary-layer interaction at high Reynolds number. *J. Fluid Mech.* **823**, 617–657.

POKRZYWA, A. 1979 Method of orthogonal projections and approximation of the spectrum of a bounded operator. *Studia Math.* **65**, 21–29.

PRIEBE, S., TU, J.H., ROWLEY, C.W. & MARTÍN, M.P. 2016 Low-frequency dynamics in a shock-induced separated flow. *J. Fluid Mech.* **807**, 441–477.

PROCTOR, J.L., BRUNTON, S.L. & KUTZ, J.N. 2016 Dynamic mode decomposition with control. *SIAM J. Appl. Dyn. Syst.* **15** (1), 142–161.

REED, M. & SIMON, B. 1980 *Methods of Modern Mathematical Physics. I*, 2nd edn. Academic Press.

ROWLEY, C.W. & DAWSON, S.T.M. 2017 Model reduction for flow analysis and control. *Annu. Rev. Fluid Mech.* **49**, 387–417.

ROWLEY, C.W., MEZIĆ, I., BAGHERI, S., SCHLATTER, P. & HENNINGSON, D.S. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127.

SARMAST, S., DADFAR, R., MIKKELSEN, R.F., SCHLATTER, P., IVANELL, S., SØRENSEN, J.N. & HENNINGSON, D.S. 2014 Mutual inductance instability of the tip vortices behind a wind turbine. *J. Fluid Mech.* **755**, 705–731.

SAYADI, T., SCHMID, P.J., NICHOLS, J.W. & MOIN, P. 2014 Reduced-order representation of near-wall structures in the late transitional boundary layer. *J. Fluid Mech.* **748**, 278–301.

SCHMID, P.J. 2009 Dynamic mode decomposition of experimental data. In *8th International Symposium on Particle Image Velocimetry (PIV09)*.

SCHMID, P.J. 2010 Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.* **656**, 5–28.

SCHMID, P.J. 2022 Dynamic mode decomposition and its variants. *Annu. Rev. Fluid Mech.* **54**, 225–254.

SCHOLKOPF, B. 2001 The kernel trick for distances. In *Advances in Neural Information Processing Systems*, pp. 301–307.

SHIELDS, P.C. 1973 *The Theory of Bernoulli Shifts*. University of Chicago Press.

SOUZA, A.N., LUTZ, T. & FLIERL, G.R. 2022 Statistical nonlocality of dynamically coherent structures. Preprint, arXiv:2211.16608.

STEIN, E.M. & SHAKARCHI, R. 2009 *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press.

SUBBAREDDY, P.K., BARTKOWICZ, M.D. & CANDLER, G.V. 2014 Direct numerical simulation of high-speed transition due to an isolated roughness element. *J. Fluid Mech.* **748**, 848–878.

SZŐKE, M. & DEVENPORT, W.J. 2021 Propagation characteristics of laser-induced acoustic sources in hybrid anechoic wind tunnels. *J. Sound Vib.* **510**, 116294.

SZŐKE, M., DEVENPORT, W.J., BORGOLTZ, A., ALEXANDER, W.N., HARI, N., GLEGG, S.A.L., LI, A., VALLABH, R. & SEYAM, A.-F.M. 2022 Investigating the aeroacoustic properties of porous fabrics. *AIAA J.* **60** (6) 1–10.

SZŐKE, M., FISCALETTI, D. & AZARPEYVAND, M. 2020 Uniform flow injection into a turbulent boundary layer for trailing edge noise reduction. *Phys. Fluids* **32** (8), 085104.

SZŐKE, M., JOZSA, T.I., KOLESZÁR, Á., MOULITSAS, I. & KÖNÖZSY, L. 2017 Performance evaluation of a two-dimensional lattice Boltzmann solver using CUDA and PGAS UPC based parallelisation. *ACM Trans. Math. Softw.* **44** (1), 1–22.

SZŐKE, M., NURANI HARI, N., DEVENPORT, W.J., GLEGG, S.A. & TESCHNER, T.-R. 2021 Flow field analysis around pressure shielding structures. In *AIAA AVIATION 2021 FORUM, AIAA Paper* 2021-2293.

TADMOR, E. 2007 Filters, mollifiers and the computation of the Gibbs phenomenon. *Acta Numer.* **16**, 305–378.

TAIRA, K., BRUNTON, S.L., DAWSON, S.T.M., ROWLEY, C.W., COLONIUS, T., MCKEON, B.J., SCHMIDT, O.T., GORDEYEV, S., THEOFILIS, V. & UKEILEY, L.S. 2017 Modal analysis of fluid flows: an overview. *AIAA J.* **55** (12), 4013–4041.

TAIRA, K., HEMATI, M.S., BRUNTON, S.L., SUN, Y., DURAISAMY, K., BAGHERI, S., DAWSON, S.T.M. & YEH, C.-A. 2020 Modal analysis of fluid flows: applications and outlook. *AIAA J.* **58** (3), 998–1022.

TOWNE, A., SCHMIDT, O.T. & COLONIUS, T. 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.* **847**, 821–867.

TREFETHEN, L.N. & EMBREE, M. 2005 *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press.

TU, J.H., ROWLEY, C.W., LUCHTENBURG, D.M., BRUNTON, S.L. & KUTZ, J.N. 2014 On dynamic mode decomposition: theory and applications. *J. Comput. Dyn.* **1** (2), 391–421.

VON NEUMANN, J. 1963 The point source solution. In *John Von Neumann Collected Works*, vol. 6, pp. 219–237.

WALTERS, P. 2000 *An Introduction to Ergodic Theory*, vol. 79. Springer Science & Business Media.

WANG, T.-K. & SHOELE, K. 2021 Geometrically weighted modal decomposition techniques. *J. Fluid Mech.* **911**, A41.

WILLIAMS, M.O., KEVREKIDIS, I.G. & ROWLEY, C.W. 2015*a* A data–driven approximation of the Koopman operator: extending dynamic mode decomposition. *J. Nonlinear Sci.* **25** (6), 1307–1346.

WILLIAMS, M.O., ROWLEY, C.W. & KEVREKIDIS, I.G. 2015*b* A kernel-based method for data-driven Koopman spectral analysis. *J. Comput. Dyn.* **2** (2), 247.

WYNN, A., PEARSON, D.S., GANAPATHISUBRAMANI, B. & GOULART, P.J. 2013 Optimal mode decomposition for unsteady flows. *J. Fluid Mech.* **733**, 473–503.

ZASLAVSKY, G.M. 2002 Chaos, fractional kinetics, and anomalous transport. *Phys. Rep.* **371** (6), 461–580.