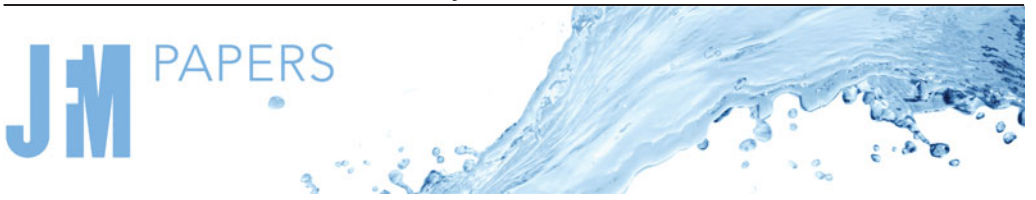
J. Fluid Mech. (2021), *vol*. 918, A10, doi:10.1017/jfm.2021.337



Data-driven resolvent analysis

Benjamin Herrmann[1,2,†], Peter J. Baddoo[3], Richard Semaan[2], Steven L. Brunton[1] and Beverley J. McKeon[4]

[1]Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, USA

[2]Institute of Fluid Mechanics, Technische Universität Braunschweig, 38108 Braunschweig, Germany

[3]Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom

[4]Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, USA

(Received 15 October 2020; revised 2 March 2021; accepted 11 April 2021)

Resolvent analysis identifies the most responsive forcings and most receptive states of a dynamical system, in an input–output sense, based on its governing equations. Interest in the method has continued to grow during the past decade due to its potential to reveal structures in turbulent flows, to guide sensor/actuator placement and for flow control applications. However, resolvent analysis requires access to high-fidelity numerical solvers to produce the linearized dynamics operator. In this work, we develop a purely data-driven algorithm to perform resolvent analysis to obtain the leading forcing and response modes, without recourse to the governing equations, but instead based on snapshots of the transient evolution of linearly stable flows. The formulation of our method follows from two established facts: (i) dynamic mode decomposition can approximate eigenvalues and eigenvectors of the underlying operator governing the evolution of a system from measurement data, and (ii) a projection of the resolvent operator onto an invariant subspace can be built from this learned eigendecomposition. We demonstrate the method on numerical data of the linearized complex Ginzburg–Landau equation and of three-dimensional transitional channel flow, and discuss data requirements. Presently, the method is suitable for the analysis of laminar equilibria, and its application to turbulent flows would require disambiguation between the linear and nonlinear dynamics driving the flow. The ability to perform resolvent analysis in a completely equation-free and adjoint-free manner will play a significant role in lowering the barrier of entry to resolvent research and applications.

Key words: low-dimensional models, machine learning

† Email address for correspondence: benherrm@uw.edu

1. Introduction

The resolvent is a linear operator that governs how harmonic forcing inputs are amplified by the linear dynamics of a system and mapped onto harmonic response outputs. Resolvent analysis refers to the inspection of this operator to find the most responsive inputs, their gains and the most receptive outputs. The resulting low-rank approximation of the forcing-response dynamics of the full system is extremely valuable for modelling, controlling and understanding the physics of fluid flows. Interest in the approach has continued to grow since McKeon & Sharma (2010) showed that, by interpreting the nonlinear term in the Fourier-transformed Navier–Stokes equations as an exogenous harmonic forcing, resolvent analysis can uncover elements of the structure in wall turbulence.

Two decades before it was coined as such, resolvent analysis was first used in the seminal work of Trefethen *et al.* (1993) to study the response of linearly stable flows to deterministic external disturbances, such as those coming from wall roughness, acoustic perturbations, body forces or free-stream turbulence. A key result was to identify the non-normality of the linearized operator as the cause of transient energy amplification of disturbances, even for cases deemed as stable by eigenvalue analysis. During the 1990s, non-modal stability theory emerged to provide a more complete picture of the linear perturbation dynamics for fluid flows using an initial-value problem formulation, as a complement to the eigenproblem from classic hydrodynamic stability theory (Schmid & Henningson 2001; Schmid 2007; Schmid & Brandt 2014). This formulation allowed the study of the response of fluid flows to initial conditions (Gustavsson 1991; Butler & Farrell 1992; Farrell & Ioannou 1993*a*; Reddy & Henningson 1993; Hwang & Cossu 2010*b*; Herrmann, Calderón-Muñoz & Soto 2018*b*), stochastic inputs (Farrell & Ioannou 1993*b*; Del Alamo & Jimenez 2006; Hwang & Cossu 2010*a*,*b*) and harmonic forcing (Jovanović & Bamieh 2005; Hwang & Cossu 2010*a*,*b*; Herrmann *et al.* 2018*b*). Another landmark is the framework adopted by Jovanović & Bamieh (2005) that focuses on the response of certain outputs of interest to forcing of specific input components. This input–output viewpoint (Jovanović 2021) allows the examination of localized disturbances and provides mechanistic insight into multi-physics systems (Jeun, Nichols & Jovanović 2016; Herrmann *et al.* 2018*a*), making it particularly relevant for control applications. Reduced-order models based on resolvent modes have been studied for turbulent channel flows (Moarref *et al.* 2013, 2014; McKeon 2017; Abreu *et al.* 2020), laminar and turbulent cavity flows (Gómez *et al.* 2016; Sun *et al.* 2020) and turbulent jets (Schmidt *et al.* 2018; Lesshafft *et al.* 2019). Other exciting recent developments include the design of an airfoil separation control strategy based on resolvent analysis by Yeh & Taira (2019), the harmonic resolvent formalism to capture cross-frequency interactions in periodic flows by Padovan, Otto & Rowley (2020), and the application of a nonlinear input–output analysis to boundary layer transition by Rigas, Sipp & Colonius (2021).

Even though there is a growing interest in the community to study flows with two and three inhomogeneous directions, global resolvent analysis is still far from commonplace. The main reasons for this are the requirement of a high-fidelity solver for the linearized governing equations, and the computational cost and memory allocation associated with handling a very large operator. The latter challenge has been addressed using matrix-free iterative techniques (Bagheri *et al.* 2009*a*; Monokrousos *et al.* 2010; Loiseau *et al.* 2019; Martini *et al.* 2020). However, this approach requires having access to a high-fidelity solver for the adjoint equations, which adds to the first challenge. A promising alternative, first used by Moarref *et al.* (2013) and further investigated by Ribeiro, Yeh & Taira (2020), is the use of randomized numerical linear algebra techniques (Liberty *et al.* 2007; Halko,

Martinsson & Tropp 2011). To simultaneously address both challenges, we propose a purely data-driven approach to obtain the resolvent operator that does not rely on access to the governing equations and can be combined with randomized methods to alleviate the computational expense if needed.

The unprecedented availability of high-fidelity numerical simulations and experimental measurements has led to incredible growth of research in data-driven modelling of dynamical systems during the past decade (Schmid 2010; Williams, Kevrekidis & Rowley 2015; Brunton, Proctor & Kutz 2016; Rudy *et al.* 2017; Loiseau & Brunton 2018; Blanchard & Sapsis 2019; Brunton & Kutz 2019; Duraisamy, Iaccarino & Xiao 2019; Raissi, Perdikaris & Karniadakis 2019; Qian *et al.* 2020; Raissi, Yazdani & Karniadakis 2020; Li *et al.* 2021). In fluid dynamics, this has led to the development and application of machine learning algorithms to extract dominant coherent structures from flow data (Taira *et al.* 2017; Brenner, Eldredge & Freund 2019; Taira *et al.* 2019; Brunton, Noack & Koumoutsakos 2020). Dynamic mode decomposition (DMD) is a particularly relevant technique introduced by Schmid (2010) to learn spatio-temporal patterns from time-resolved data that are each associated with a single frequency and growth/decay rate (Kutz *et al.* 2016). During the last decade, considerable effort has been devoted to interpret its application to nonlinear dynamical systems, based on a deep connection to Koopman theory (Rowley *et al.* 2009; Mezić 2013), and to develop numerous extensions to allow the use of non-sequential measurements (Tu *et al.* 2014*b*), promote sparsity in the solution (Jovanović, Schmid & Nichols 2014), work with streaming datasets (Hemati, Williams & Rowley 2014), improve its accuracy with nonlinear observables (Williams *et al.* 2015), incorporate the effect of control inputs (Proctor, Brunton & Kutz 2016), add robustness using nonlinear optimization (Askham & Kutz 2018) and enable its application to massive datasets with randomized linear algebra techniques (Erichson *et al.* 2019*b*). Recently, Towne, Schmidt & Colonius (2018) showed that, for statistically stationary flow data, modes obtained from spectral proper orthogonal decomposition (Lumley 1970) are equivalent to the output resolvent modes if the nonlinear forcing exhibits no preferential direction. Nonetheless, it is the input resolvent modes that provide insight into the most amplified flow structures, the most sensitive actuator locations and the most responsive control inputs. Also recently, Gómez & Blackburn (2017) developed a promising data-driven method to identify sensitive spatial locations in unsteady flows and demonstrated its use for the design of passive flow control strategies. However, this approach requires using time-resolved snapshots of both, the state (velocity field) and of the nonlinear forcing (Reynolds stress divergence).

In this work, we present an algorithm to obtain the leading input and output resolvent modes, and the associated gains, of linearly stable flows directly from data by following the procedure shown in figure 1. The method relies on DMD to approximate the eigenvalues and eigenfunctions of the system based on snapshots from one or more transient trajectories of the flow. We show that, using an appropriate inner product, the resolvent of the matrix of DMD eigenvalues is the resolvent of the system projected onto the span of the DMD eigenvectors, which are subsequently used to synthesize the resolvent modes in physical coordinates. Our method is able to find the optimal forcing, response and gain of a linear dynamical system in an equation-free manner, and without access to data from adjoint simulations, therefore opening the possibility of resolvent analysis purely based on experimental measurements. Further work is required to allow data-driven resolvent analysis of turbulent flows, which calls for another new method that is capable of distinguishing between linear and nonlinear dynamics driving the evolution of spatio-temporal measurements of a system. To promote ease of reproducibility of our

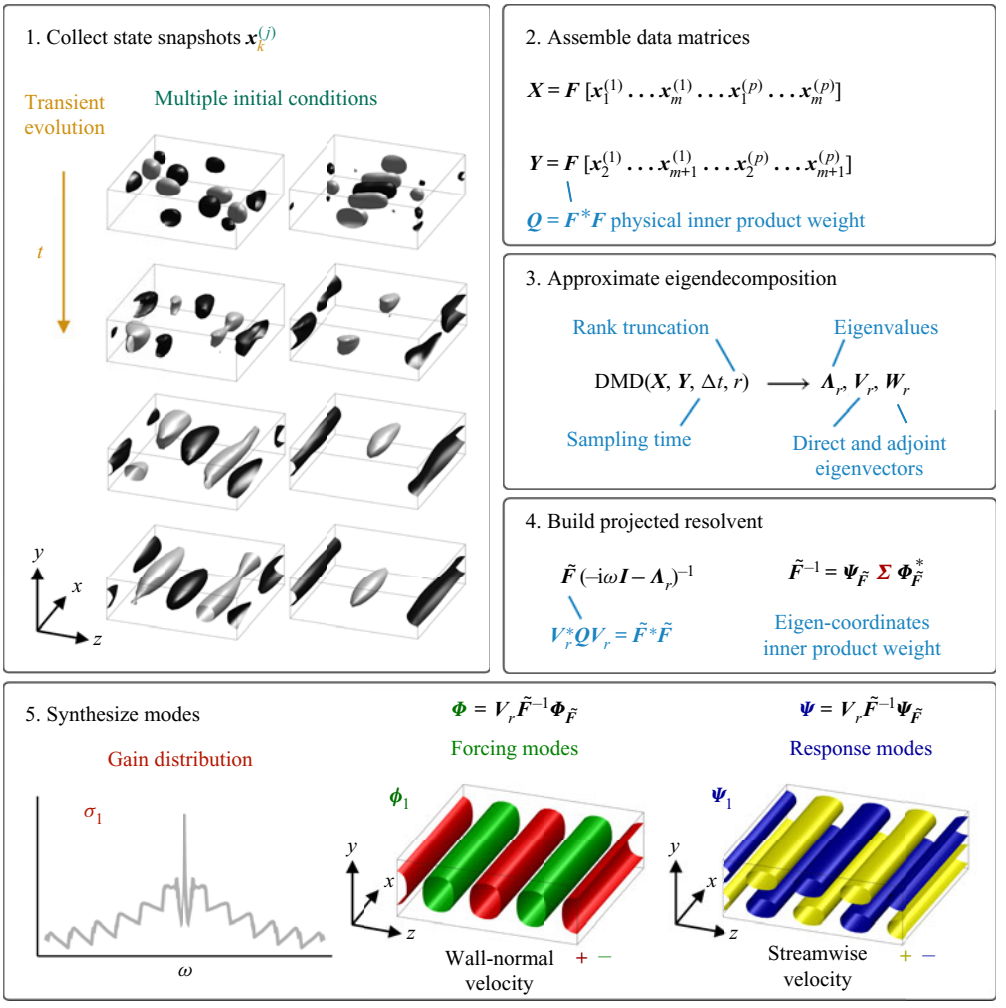


Figure 1. Schematic of the data-driven resolvent analysis algorithm demonstrated on the transitional channel flow example detailed in § 4.2. Data are collected from time recordings of the system of interest, where one or more initial conditions are used to generate the transient dynamics. Measurements are stacked into data matrices that are used to approximate an eigendecomposition of the underlying system via DMD. A projection of the resolvent operator onto the span of the learned eigenvectors is analysed, and, finally, the produced modes are lifted to physical coordinates.

results, all of the code developed for this work is available on github.com/ben-herrmann, and all the data are available for sharing upon request to the corresponding author.

The remainder of the paper is organized as follows. A formulation of resolvent analysis is presented in § 2, followed by a description of our proposed method in § 3. We demonstrate the approach on two examples in § 4 and discuss its main limitations and possible extensions in § 5. Our conclusions are offered in § 6.

2. Resolvent analysis

In this section, we present a brief and practical formulation of the conventional resolvent and input–output analyses.

2.1. *General description*

Let us consider a forced linear dynamical system

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{f}, \tag{2.1}$$

where the dot denotes time differentiation, $\boldsymbol{x} \in \mathbb{C}^n$ is the state vector, $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ is the linear dynamics matrix and $\boldsymbol{f} \in \mathbb{C}^n$ is an external driving force. Such a system may arise from a semi-discretized partial differential equation, and in the case of fluid flows, the incompressible Navier–Stokes equations can be written in this form by projecting the velocity field onto a divergence-free basis to eliminate the pressure variable. The state $\boldsymbol{x}$ may either represent the deviation from a steady state of a laminar flow, or fluctuations about the temporal mean of a statistically stationary unsteady flow. In both cases the matrix $\boldsymbol{A}$ is the linearization of the underlying nonlinear system about the corresponding base flow, either the equilibrium or the mean. However, if we are dealing with an unsteady flow, it is important to note that a Reynolds decomposition yields a nonlinear term that typically cannot be neglected. In this scenario, the nonlinearity is lumped into $\boldsymbol{f}$ and considered as an external forcing, with the caveat that the system may exhibit preferred input directions due to internal feedback between linear amplification and nonlinear interactions (McKeon & Sharma 2010; Beneddine *et al.* 2016).

In this work, we focus on the case where $\boldsymbol{x}$ is the deviation from a stable steady state and $\boldsymbol{f}$ is an exogenous input with no preferential direction, representing disturbances from the environment, model discrepancy, or an open-loop control actuation. For a harmonic forcing $\boldsymbol{f}(t) = \hat{\boldsymbol{f}}\,\mathrm{e}^{-\mathrm{i}\omega t} + \text{c.c.}$, where c.c. means complex conjugate, $\omega \in \mathbb{R}$ is the angular driving frequency and $t \in \mathbb{R}$ the time variable, the long-term response is also harmonic, $\boldsymbol{x}(t) = \hat{\boldsymbol{x}}\,\mathrm{e}^{-\mathrm{i}\omega t} + \text{c.c.}$, and is governed by the particular solution to (2.1) given by

$$\hat{\boldsymbol{x}} = (-\mathrm{i}\omega\boldsymbol{I} - \boldsymbol{A})^{-1}\hat{\boldsymbol{f}}, \tag{2.2}$$

where $\boldsymbol{I}$ is the $n \times n$ identity matrix. Let $\boldsymbol{H}(\omega) = (-\mathrm{i}\omega\boldsymbol{I} - \boldsymbol{A})^{-1}$ be the matrix approximation of the resolvent operator, where the negative sign accompanying the ω-term follows the convention used for travelling waves. We seek the largest input–output gain, $\sigma_1(\omega)$, optimized over all possible forcing vectors $\hat{\boldsymbol{f}}$, or more formally

$$\sigma_1(\omega) = \max_{\hat{\boldsymbol{f}}\,\neq\,\boldsymbol{0}} \frac{\|\hat{\boldsymbol{x}}\|_{\boldsymbol{Q}}^2}{\|\hat{\boldsymbol{f}}\|_{\boldsymbol{Q}}^2} = \max_{\hat{\boldsymbol{f}}\,\neq\,\boldsymbol{0}} \frac{\|\boldsymbol{H}(\omega)\hat{\boldsymbol{f}}\|_{\boldsymbol{Q}}^2}{\|\hat{\boldsymbol{f}}\|_{\boldsymbol{Q}}^2}, \tag{2.3}$$

where $\|\hat{\boldsymbol{x}}\|_{\boldsymbol{Q}}^2 = \hat{\boldsymbol{x}}^*\boldsymbol{Q}\hat{\boldsymbol{x}}$, with $(\,)^*$ denoting the Hermitian transpose, measures the size of the state based on a physically meaningful metric given by the positive–definite weighting matrix $\boldsymbol{Q}$. This weighting accounts for integration quadratures, non-uniform spatial discretizations and appropriate scaling of heterogeneous variables in multi-physics systems (Herrmann *et al.* 2018*a*). The Cholesky decomposition is used to factorize $\boldsymbol{Q} = \boldsymbol{F}^*\boldsymbol{F}$ and relate the physically relevant norm to the standard Euclidean 2-norm via $\|\hat{\boldsymbol{x}}\|_{\boldsymbol{Q}}^2 = \|\boldsymbol{F}\hat{\boldsymbol{x}}\|_2^2$. With this scaling of the states, and using the definition of the vector-induced matrix norm, the optimal gain optimization problem (2.3) is equivalent to

$$\sigma_1(\omega) = \|\boldsymbol{F}\boldsymbol{H}(\omega)\boldsymbol{F}^{-1}\|_2^2. \tag{2.4}$$

The solution to (2.4), along with a hierarchy of optimal and sub-optimal forcing and response vectors, is given by the singular value decomposition (SVD) of the weighted

resolvent

$$\boldsymbol{F}\boldsymbol{H}(\omega)\boldsymbol{F}^{-1} = \boldsymbol{\Psi}_{\boldsymbol{F}}(\omega)\boldsymbol{\Sigma}(\omega)\boldsymbol{\Phi}_{\boldsymbol{F}}^{*}(\omega), \tag{2.5}$$

where $\boldsymbol{\Sigma} \in \mathbb{R}^{n}$ is a diagonal matrix containing the gains $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$, also known as singular values, and $\boldsymbol{\Phi}_{\boldsymbol{F}} = \boldsymbol{F}\ [\boldsymbol{\phi}_1\ \boldsymbol{\phi}_2 \cdots \boldsymbol{\phi}_n] \in \mathbb{C}^{n\times n}$ and $\boldsymbol{\Psi}_{\boldsymbol{F}} = \boldsymbol{F}[\boldsymbol{\psi}_1\ \boldsymbol{\psi}_2 \cdots \boldsymbol{\psi}_n] \in \mathbb{C}^{n\times n}$ are unitary matrices whose columns, when left multiplied by $\boldsymbol{F}^{-1}$, yield the input and output resolvent modes, $\boldsymbol{\phi}_j$ and $\boldsymbol{\psi}_j$, respectively.

2.2. *Input–output analysis*

The more general input–output analysis (Jovanović & Bamieh 2005; Jovanović 2021) follows from considering that only measurements of the state $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$ are observed, and that the forcing is constrained to be of the form $\boldsymbol{f} = \boldsymbol{B}\boldsymbol{u}$, where $\boldsymbol{y} \in \mathbb{C}^{n_y}$ and $\boldsymbol{u} \in \mathbb{C}^{n_u}$. In this framework, the matrices $\boldsymbol{B} \in \mathbb{C}^{n\times n_u}$ and $\boldsymbol{C} \in \mathbb{C}^{n_y\times n}$ restrict the input and output subspaces, allowing the analysis of scenarios where only certain measurements are available (e.g. on a specific region of the spatial domain) and where the possible types of actuation are limited. In this case, the long-term response to harmonic forcing of the measured variables is governed by

$$\hat{\boldsymbol{y}} = \boldsymbol{C}\left(-\mathrm{i}\omega\boldsymbol{I} - \boldsymbol{A}\right)^{-1}\boldsymbol{B}\hat{\boldsymbol{u}}, \tag{2.6}$$

where, instead of the resolvent, the transfer function $\boldsymbol{C}\boldsymbol{H}(\omega)\boldsymbol{B}$ governs the amplification of the allowed inputs to produce the observed outputs. The analysis of the latter operator requires introducing norms based on inner products for both the input and output spaces, $\|\hat{\boldsymbol{u}}\|_{\boldsymbol{Q}_u}^2 = \hat{\boldsymbol{u}}^*\boldsymbol{Q}_u\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}^*\boldsymbol{F}_u^*\boldsymbol{F}_u\hat{\boldsymbol{u}} = \|\boldsymbol{F}_u\hat{\boldsymbol{u}}\|_2^2$ and $\|\hat{\boldsymbol{y}}\|_{\boldsymbol{Q}_y}^2 = \hat{\boldsymbol{y}}^*\boldsymbol{Q}_y\hat{\boldsymbol{y}} = \hat{\boldsymbol{y}}^*\boldsymbol{F}_y^*\boldsymbol{F}_y\hat{\boldsymbol{y}} = \|\boldsymbol{F}_y\hat{\boldsymbol{y}}\|_2^2$, respectively. Input–output analysis then amounts to taking the SVD of the weighted transfer function

$$\boldsymbol{F}_y\boldsymbol{C}\boldsymbol{H}(\omega)\boldsymbol{B}\boldsymbol{F}_u^{-1} = \boldsymbol{\Psi}_{\boldsymbol{F}_y}(\omega)\boldsymbol{\Sigma}(\omega)\boldsymbol{\Phi}_{\boldsymbol{F}_u}^{*}(\omega), \tag{2.7}$$

where $\boldsymbol{\Sigma}$ contains the input–output gains in its diagonal, and the input and output modes are given by $\boldsymbol{\Phi}_u = \boldsymbol{F}_u^{-1}\boldsymbol{\Phi}_{\boldsymbol{F}_u}$ and $\boldsymbol{\Psi}_y = \boldsymbol{F}_y^{-1}\boldsymbol{\Psi}_{\boldsymbol{F}_y}$, respectively.

3. Data-driven resolvent analysis

In this section, we present a general description of the proposed method to perform resolvent analysis based on data followed by comments on its extension to the input–output framework.

3.1. *General description*

Data-driven resolvent analysis relies on DMD to approximate the eigenvalues and eigenvectors of the underlying dynamical system. Among the many choices of DMD variants, we use the exact DMD approach of Tu *et al.* (2014*b*). In the absence of forcing, the evolution of measurements of the dynamical system of interest (2.1) is governed by

$$\boldsymbol{x}_{k+1} = \exp(\boldsymbol{A}\Delta t)\boldsymbol{x}_k, \tag{3.1}$$

where $\boldsymbol{x}_k$ is the measurement at time $t_k = k\Delta t$, and Δt is the sampling time. As in the previous section, let the the weight matrix $\boldsymbol{Q} = \boldsymbol{F}^*\boldsymbol{F}$ define a physically meaningful norm

to quantify the size of the state vector. The following transformation allows us to work in the 2-norm framework,

$$\boldsymbol{F}\boldsymbol{x}_{k+1} = \boldsymbol{F}\exp(\boldsymbol{A}\Delta t)\boldsymbol{F}^{-1}\boldsymbol{F}\boldsymbol{x}_k = \boldsymbol{\Theta}\boldsymbol{F}\boldsymbol{x}_k, \tag{3.2}$$

where $\boldsymbol{\Theta} = \boldsymbol{F}\exp(\boldsymbol{A}\Delta t)\boldsymbol{F}^{-1}$ evolves the weighted measurements $\boldsymbol{F}\boldsymbol{x}_k$ one time step into the future. Under this transformation, the adjoint of $\boldsymbol{\Theta}$ based on the $\boldsymbol{Q}$-norm is equivalent to the Hermitian adjoint. Thus, using the weighted measurements, we can proceed using readily available DMD codes. To begin, we collect snapshots of the state denoted by $\boldsymbol{x}_k^{(j)}$, where the subscript $k \in \{1, \ldots, m+1\}$ denotes the sample number, and the superscript $j \in \{1, \ldots, p\}$ denotes different trajectories started from $p \geq 1$ initial conditions. The next step is to assemble the weighted data matrices

$$\boldsymbol{X} = \boldsymbol{F}\left[\boldsymbol{x}_1^{(1)}\ \boldsymbol{x}_2^{(1)}\cdots\boldsymbol{x}_m^{(1)}\ \middle|\ \boldsymbol{x}_1^{(2)}\ \boldsymbol{x}_2^{(2)}\cdots\boldsymbol{x}_m^{(2)}\ \middle|\cdots\middle|\ \boldsymbol{x}_1^{(p)}\ \boldsymbol{x}_2^{(p)}\cdots\boldsymbol{x}_m^{(p)}\right], \tag{3.3a}$$

$$\boldsymbol{Y} = \boldsymbol{F}\left[\boldsymbol{x}_2^{(1)}\ \boldsymbol{x}_3^{(1)}\cdots\boldsymbol{x}_{m+1}^{(1)}\ \middle|\ \boldsymbol{x}_2^{(2)}\ \boldsymbol{x}_3^{(2)}\cdots\boldsymbol{x}_{m+1}^{(2)}\ \middle|\cdots\middle|\ \boldsymbol{x}_2^{(p)}\ \boldsymbol{x}_3^{(p)}\cdots\boldsymbol{x}_{m+1}^{(p)}\right], \tag{3.3b}$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are of size $n \times pm$. Based on these data matrices, the DMD framework with a rank-r truncated SVD yields the matrices $\boldsymbol{D}_r = \mathrm{diag}(\rho_1, \rho_2, \ldots, \rho_r) \in \mathbb{C}^{r\times r}$ containing the approximated eigenvalues, and $\boldsymbol{V}_{r,\boldsymbol{F}} \in \mathbb{C}^{n\times r}$ and $\boldsymbol{W}_{r,\boldsymbol{F}} \in \mathbb{C}^{n\times r}$, whose columns are the approximated direct and adjoint eigenvectors of the underlying operator $\boldsymbol{\Theta}$. The reader is referred to the work of Tu *et al.* (2014*b*) for the derivation of the adjoint DMD modes. These eigenvalues and vectors are related to those corresponding to $\boldsymbol{A}$ via $\lambda_j = \log(\rho_j)/\Delta t$, $\boldsymbol{V}_r = \boldsymbol{F}^{-1}\boldsymbol{V}_{r,\boldsymbol{F}}$ and $\boldsymbol{W}_r = \boldsymbol{F}^{-1}\boldsymbol{W}_{r,\boldsymbol{F}}$. Hence, we obtain $\boldsymbol{\Lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r) \in \mathbb{C}^{r\times r}$, $\boldsymbol{V}_r = [\boldsymbol{v}_1\ \boldsymbol{v}_2\cdots\boldsymbol{v}_r] \in \mathbb{C}^{n\times r}$, and $\boldsymbol{W}_r = [\boldsymbol{w}_1\ \boldsymbol{w}_2\cdots\boldsymbol{w}_r] \in \mathbb{C}^{n\times r}$, that satisfy

$$\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}, \quad \text{and} \quad \boldsymbol{A}^{+}\boldsymbol{W} = \boldsymbol{W}\boldsymbol{\Lambda}^{*}, \tag{3.4a,b}$$

for an unknown underlying operator $\boldsymbol{A}$, where $\boldsymbol{A}^{+} = \boldsymbol{Q}^{-1}\boldsymbol{A}^{*}\boldsymbol{Q}$ is its $\boldsymbol{Q}$-norm adjoint. In other words, we use DMD as a data-driven eigendecomposition, which is not surprising considering that its connection to Arnoldi methods and Krylov subspaces has been clear since the origins of the algorithm (Schmid 2010).

Next, we seek an approximation of the resolvent operator built on $\boldsymbol{\Lambda}_r$, $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$. Our approach leverages an operator-based dimensionality reduction technique first used in the context of non-modal stability analysis by Reddy & Henningson (1993). We now return our attention to the forced system (2.1), and consider an eigenvector expansion of $\boldsymbol{x}$ and $\boldsymbol{f}$, as follows

$$\boldsymbol{x}(t) = \boldsymbol{V}_r\boldsymbol{a}(t), \quad \text{and} \quad \boldsymbol{f}(t) = \boldsymbol{V}_r\boldsymbol{b}(t), \tag{3.5a,b}$$

where $\boldsymbol{a} = [a_1\ a_2\cdots a_r]^{\mathrm{T}} \in \mathbb{C}^r$, and $\boldsymbol{b} = [b_1\ b_2\cdots b_r]^{\mathrm{T}} \in \mathbb{C}^r$ are the vector of expansion coefficients in eigenvector coordinates. In the work of Reddy & Henningson (1993), $\boldsymbol{V}_r$ are the eigenvectors associated with the first r eigenvalues with largest real part of a known operator $\boldsymbol{A}$, whereas here they are the DMD modes. Substitution of (3.5*a*,*b*) in (2.1), and taking the inner product with $\boldsymbol{W}_r$ at both sides yields

$$\dot{\boldsymbol{a}} = \boldsymbol{\Lambda}_r\boldsymbol{a} + \boldsymbol{b}, \tag{3.6}$$

where we have used the bi-orthogonality property between the sets of direct and adjoint eigenvectors, and assumed that they have been normalized such that $\boldsymbol{W}_r^{*}\boldsymbol{Q}\boldsymbol{V}_r = \boldsymbol{I} \in \mathbb{R}^{r\times r}$. Because we are now working in different coordinates, if we want to retain the physical meaning of the norm, we need to adjust our inner product accordingly. The new weighting matrix is derived as $\|\boldsymbol{x}\|_{\boldsymbol{Q}}^2 = \boldsymbol{x}^{*}\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{a}^{*}\boldsymbol{V}_r^{*}\boldsymbol{Q}\boldsymbol{V}_r\boldsymbol{a} = \|\widetilde{\boldsymbol{F}}\boldsymbol{a}\|_2^2$, where we have defined the

new matrix $\widetilde{\boldsymbol{F}} \in \mathbb{C}^{r\times r}$ from the Cholesky factorization of $\boldsymbol{V}_r^*\boldsymbol{Q}\boldsymbol{V}_r = \widetilde{\boldsymbol{F}}^*\widetilde{\boldsymbol{F}}$. We are now ready to proceed with the resolvent analysis for the system (3.6). As presented in the previous section, the weighted resolvent modes and gains are obtained from the SVD of

$$\widetilde{\boldsymbol{F}}(-\mathrm{i}\omega\boldsymbol{I} - \boldsymbol{\Lambda}_r)^{-1}\widetilde{\boldsymbol{F}}^{-1} = \boldsymbol{\Psi}_{\widetilde{\boldsymbol{F}}}(\omega)\boldsymbol{\Sigma}(\omega)\boldsymbol{\Phi}^*_{\widetilde{\boldsymbol{F}}}(\omega). \tag{3.7}$$

The final step is to synthesize the resolvent modes in physical coordinates, as follows

$$\boldsymbol{\Phi} = \boldsymbol{V}_r\widetilde{\boldsymbol{F}}^{-1}\boldsymbol{\Phi}_{\widetilde{\boldsymbol{F}}}, \quad \text{and} \quad \boldsymbol{\Psi} = \boldsymbol{V}_r\widetilde{\boldsymbol{F}}^{-1}\boldsymbol{\Psi}_{\widetilde{\boldsymbol{F}}}. \tag{3.8a,b}$$

A schematic that summarizes the entire procedure is shown in figure 1. It is worth pointing out that an analogous procedure can be carried out to obtain data-driven transient growth modes, simply by replacing the resolvent operator with the matrix exponential propagator evaluated at a finite time horizon. In addition, notice that the reduced-order resolvent matrix in (3.7) is of size $r \times r$, meaning that its full SVD requires $O(r^3)$ operations instead of $O(n^3)$, and therefore is considerably cheaper to compute. This projection onto the span of the eigenvectors has been successfully used in operator-based non-modal stability analyses to achieve computational speedups of orders of magnitude (Reddy & Henningson 1993; Herrmann *et al.* 2018*b*).

3.2. *Data-driven input–output analysis*

The extension of the above presented method to the input–output framework is quite straightforward if the dataset is composed from full-state measurements. Just as in the operator-based approach, this allows exploration of what the forcing-response dynamics would be like if the inputs and outputs were restricted by the linear mappings $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$ and $\boldsymbol{f} = \boldsymbol{B}\boldsymbol{u}$. In a similar fashion to (2.7), data-driven input–output analysis follows from the SVD of

$$\boldsymbol{F}_y\boldsymbol{C}\boldsymbol{V}_r(-\mathrm{i}\omega\boldsymbol{I} - \boldsymbol{\Lambda}_r)^{-1}\boldsymbol{W}^*\boldsymbol{Q}\boldsymbol{B}\boldsymbol{F}_u^{-1} = \boldsymbol{\Psi}_{\boldsymbol{F}_y}(\omega)\boldsymbol{\Sigma}(\omega)\boldsymbol{\Phi}^*_{\boldsymbol{F}_u}(\omega), \tag{3.9}$$

where, again, $\boldsymbol{\Sigma}$ contains the input–output gains in its diagonal, and the input and output modes are given by $\boldsymbol{\Phi}_u = \boldsymbol{F}_u^{-1}\boldsymbol{\Phi}_{\boldsymbol{F}_u}$ and $\boldsymbol{\Psi}_y = \boldsymbol{F}_y^{-1}\boldsymbol{\Psi}_{\boldsymbol{F}_y}$, respectively. In (3.9) we have once more leveraged the projection onto the eigen-basis learned from DMD.

We stress the fact that the presented framework considers full-state measurement data, which are generally required to approximate $\boldsymbol{\Lambda}_r$, $\boldsymbol{V}_r$ and $\boldsymbol{W}_r$. Learning the spectrum and eigenvectors of a linear system from a dataset composed of partial state recordings of the dynamics generated using limited actuation remains an open challenge. This is further discussed in § 5 with our results at hand.

4. Examples and discussion

In this section, we demonstrate the application of data-driven resolvent analysis on two example problems.

4.1. *Complex Ginzburg–Landau equation*

Our first example is the linearized complex Ginzburg–Landau equation, which is a typical model for instabilities in spatially evolving flows. The system is governed by the linear operator

$$\boldsymbol{A} = -\nu\boldsymbol{D}_x + \gamma\boldsymbol{D}_x^2 + \mu(x), \tag{4.1}$$

where x is the spatial coordinate, and $\boldsymbol{D}_x$ and $\boldsymbol{D}_x^2$ are the first- and second-order spatial differentiation matrices with homogeneous boundary conditions at $x \to \pm\infty$. We choose

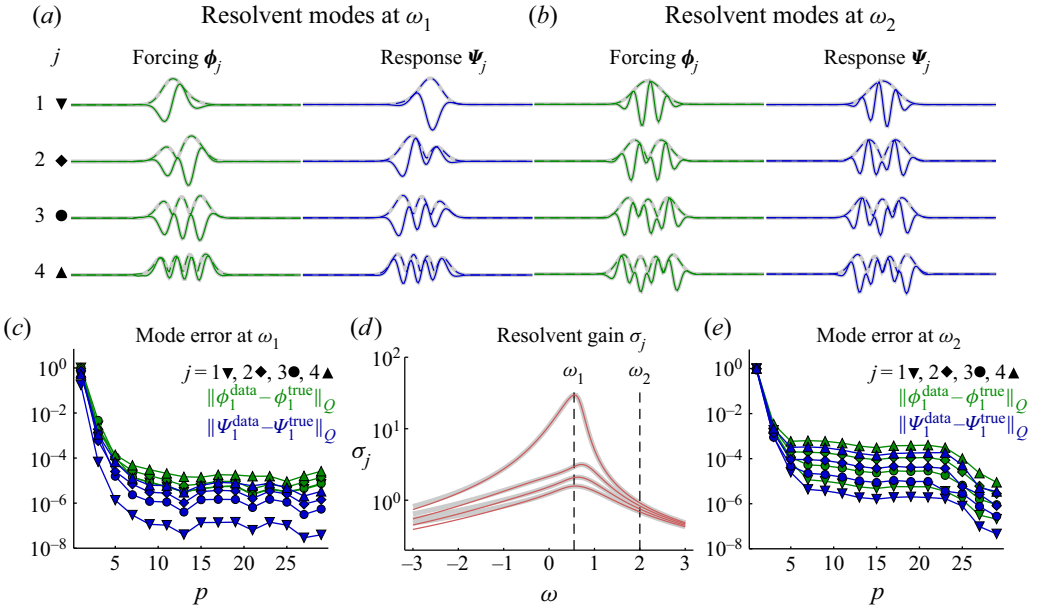


Figure 2. Data-driven resolvent analysis of the linearized complex Ginzburg–Landau equation. (*a*) The first four forcing and response modes at $\omega_1 = 0.55$, where solid and dashed lines show the real part and magnitude of the modes. (*b*) The same as (*a*), but for a frequency $\omega_2 = 2$ where there is much less gain separation. (*c*) The $\boldsymbol{Q}$-norm error between the operator-based and the data-driven resolvent modes at ω_1 as a function of the number of trajectories p considered in the dataset. (*d*) Resolvent gain distribution for the first four modes as a function of frequency. (*e*) The same as (*c*), but for ω_2. In (*a*), (*b*,*d*), the thick grey lines show operator-based quantities for a ground-truth comparison.

a quadratic spatial dependence for the parameter $\mu(x) = (\mu_0 - c_\mu^2) + (\mu_2/2)x^2$, that has been used previously by several authors, e.g. Bagheri *et al.* (2009*b*). The other parameters are set to $\mu_0 = 0.23$, $\mu_2 = -0.01$, $\nu = 2 + 0.4\text{i}$ and $\gamma = 1 - \text{i}$, giving rise to a linearly stable dynamics. As in Bagheri *et al.* (2009*b*), we use spectral collocation based on Gauss-weighted Hermite polynomials to build the differentiation matrices $\boldsymbol{D}_x$ and $\boldsymbol{D}_x^2$ and the integration quadrature $\boldsymbol{Q}$. The spatial coordinate is discretized into $n = 220$ collocation points, and the domain is truncated to $x \in [-85, 85]$, which is sufficient to enforce the far-field boundary conditions.

Data are generated from 30 simulations that are each started from different initial conditions which we choose to be the first 30 Gauss-weighted Hermite polynomials. We record $m = 100$ snapshots that are sampled every $\Delta t = 0.5$ time units. Data-driven resolvent analysis is performed using snapshot matrices assembled considering measurements from the first trajectory only. Subsequently, the method is applied on snapshot matrices where measurements from the other simulations are sequentially concatenated one by one, to investigate the convergence behaviour in regards to the amount of data required. The $\boldsymbol{Q}$-norm error between the first four operator-based and data-driven resolvent modes as a function of the number of trajectories considered is shown in figures 2(*c*) and 2(*e*) for two different input frequencies $\omega_1 = 0.55$ and $\omega_2 = 2$, respectively. As more data are included, the abrupt drop in the error is expected, since DMD is able to accurately approximate a larger number of eigenvectors, therefore enriching the basis we use to represent the resolvent modes.

The first four resolvent modes at ω_1 and at ω_2, as well as their gains as a function the input frequency are shown in figures 2(*a*), 2(*b*) and 2(*d*). In this case, the data-driven analysis provides an accurate approximation of the true leading and sub-optimal resolvent modes, even at frequencies where there is a relatively small gain separation, albeit with a larger mode error. We also note that the error in the approximation of the forcing modes is consistently larger than that of the response modes by approximately an order of magnitude. A reason for this may be that, in the case of non-normal systems, the spatial support of the direct eigenvectors is similar to that of the response modes, but it is typically separated from that of the forcing modes. For all results presented in this section, we used $r = 24$ for the rank truncation in both DMD and the data-driven eigenbasis $\boldsymbol{V}_r$. Note that caution is needed when retaining more vectors in $\boldsymbol{V}_r$ to not include spurious eigenvectors, which instead of enriching the basis can be detrimental to the performance of the method.

4.2. *Transitional channel flow*

Our second example is the three-dimensional flow in a plane channel of finite length and depth, and with periodic streamwise and spanwise boundary conditions. The system is governed by the incompressible Navier–Stokes equations, and we consider a Reynolds number $Re = 2000$ based on the channel half-height and the centreline velocity, and a domain size of $2\pi \times 2 \times 2\pi$ dimensionless length units along the x, y and z coordinates that indicate the streamwise, wall-normal and spanwise directions, respectively. The state vector in this case is composed of the three-dimensional flow field of disturbances about the base parabolic velocity profile.

In the case of single-wavenumber perturbations, the dynamics is described by the traditional Orr–Sommerfeld and Squire equations (Schmid & Henningson 2001). The corresponding linear operator is built using Chebyshev spectral collocation to discretize the wall-normal direction. The Orr–Sommerfeld/Squire operator is then used to compute the operator-based spectrum and the leading resolvent gain of the three-dimensional flow, as shown in figure 1. This is achieved looping over wavenumber combinations that are compatible with the finite channel dimensions, i.e. integers for the current set-up. We consider $N_y = 101$ collocation points, and wavenumbers in the range $|\alpha| \leq 7$ for the streamwise component, and $|\beta| \leq 7$ for the spanwise component. The operator-based leading resolvent modes, shown in figure 3, are computed using $\omega = 0,\ \alpha = 0,\ \beta = 2$, for which the maximum gain is observed to occur. The optimal forcing and response correspond to the familiar streamwise vortices that excite streamwise streaks (Trefethen *et al.* 1993).

To demonstrate our data-driven resolvent analysis, we need snapshots from the transient evolution of the full three-dimensional flow field. We use the spectral code *Channelflow* (Gibson, Halcrow & Cvitanović 2008; Gibson 2014) to perform direct numerical simulations of the incompressible Navier–Stokes equations. The code uses Chebyshev and Fourier expansions of the flow field in the wall-normal and horizontal directions, and a third-order Adams–Bashforth backward differentiation scheme for the time integration. We find that a grid with $N_y = 65$ and $N_x = N_z = 32$ points is sufficient to discretize the domain for the cases studied, and a time step of 0.01 time units is selected, keeping the Courant–Friedrichs–Lewy (CFL) number at 0.32. All cases described below are simulated for 400 time units, and snapshots are saved every $\Delta t = 0.5$ time units. The perturbation kinetic energy of all initial conditions simulated is set to 10^{-5} to ensure that the effect of nonlinearity is negligible.

In order to demonstrate our data-driven method on a higher-dimensional system than in the previous example, rather than Fourier transforming in the streamwise and spanwise

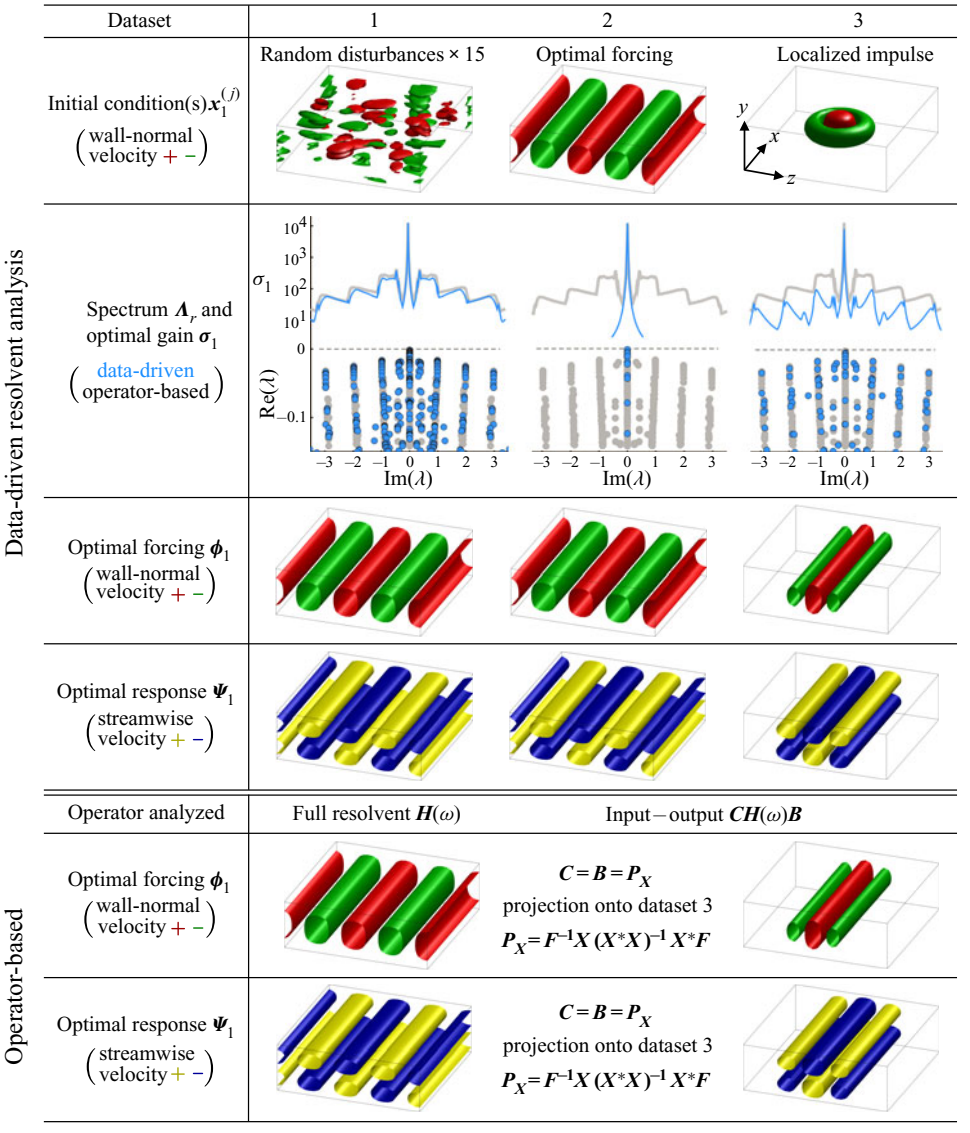


Figure 3. Data-driven resolvent analysis of three-dimensional plane channel flow at $Re = 2000$ based on the channel half-height and the centreline velocity. The method is demonstrated using three datasets obtained from DNS initialized with: (i) small-wavenumber random disturbances, where 15 trajectories are considered, (ii) the optimal forcing and (iii) localized actuation. Operator-based results are also shown for comparison, including the resolvent modes obtained when the input and output are restricted to lie in the span of the snapshots from dataset 3.

directions to perform a separate local analysis for each wavenumber tuple, we directly use the three-dimensional flow field data to perform a global analysis. This provides a more challenging test bed, requiring that we learn a much larger eigen-basis.

First, we consider initial disturbances of the form

$$\boldsymbol{u}(x, y, z, 0) = \sum_{i,j,k,l} \left(c_{ijkl} T_k(y) \exp(\mathrm{i}(ix + jz)) + \text{c.c.} \right) \mathbf{e}_l, \tag{4.2}$$

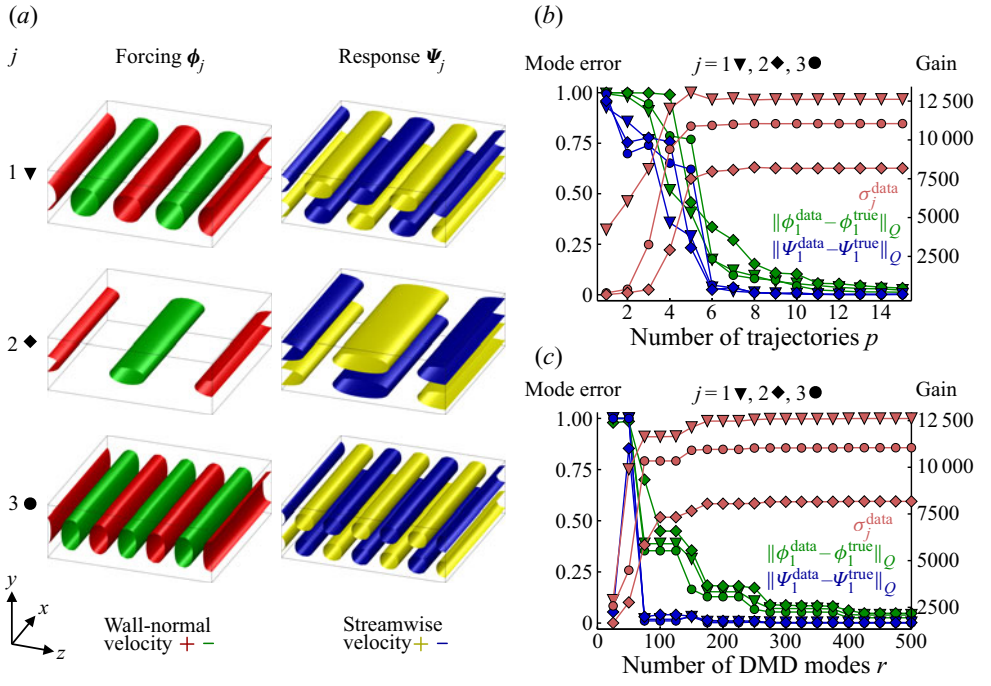


Figure 4. Convergence of data-driven resolvent modes of three-dimensional plane channel flow at $Re = 2000$ based on the channel half-height and the centreline velocity. (*a*) The first three forcing and response modes at the dominant frequency $\omega = 0$ computed with $r = 400$ DMD eigenvectors learned from a dataset composed of $p = 15$ simulations initialized with random disturbances of the form shown in (4.2). (*b*) Data-driven resolvent gains and $\boldsymbol{Q}$-norm error between the operator-based and the data-driven resolvent modes as a function of p with $r = 400$. (*c*) The same, but with $p = 15$ fixed, and as a function of r instead. In (*b*,*c*) the superscript true denotes the operator-based modes obtained using direct computation of the resolvent.

where $i, j \in \{-3, \ldots, 3\}$, $T_k(y)$ are Chebyshev polynomials with $k \in \{1, \ldots, 65\}$, $l \in \{1, \ldots, 3\}$, and c_{ijkl} are randomly sampled complex numbers from inside the unit disk. Once generated, the disturbances are corrected to satisfy the boundary and divergence-free conditions, and scaled to have the desired kinetic energy. To investigate the convergence of the method with the amount of data, we consider growing datasets aggregating the snapshots from 1 up to 15 simulations initialized with these random small-wavenumber disturbances. The first three forcing and response modes at $\omega = 0$ are computed retaining $r = 400$ DMD modes and $p = 15$ initial conditions, and are shown in figure 4(*a*). With $r = 400$ fixed, as the number of trajectories in the dataset p increases, the $\boldsymbol{Q}$-norm error between data-driven and operator-based resolvent modes slowly decreases, as shown in figure 4(*b*). To put in perspective the large amount of data used, recall that the state of the system is comprised of the three components of the velocity field at every grid point, yielding a vector of dimension $\sim 2 \times 10^5$, and for every trajectory simulated we record 800 snapshots, resulting in a total of 12 000 snapshots for the case with $p = 15$. Moreover, considering $p = 15$ trajectories, the convergence of the leading and sub-optimal modes with the number of DMD eigenvectors retained r is shown in figure 4(*c*). It is important to notice that the respective resolvent gains also converge around the same values of p and r, as shown in figures 4(*b*) and 4(*c*), therefore they could be used as an alternative to assess convergence when the reference modes are not available.

In addition to how much data are required for the data-driven resolvent discussed above, we now investigate the more interesting question of which data. To do so, we learn the leading resolvent modes, spectra and resolvent gain distributions from three datasets obtained from qualitatively different initial perturbations. The first dataset is the one described above for the convergence study, composed of $p = 15$ simulations initialized with random disturbances of the form of (4.2). Data-driven resolvent analysis is applied retaining $r = 400$ DMD eigenvectors, and the learned spectrum, gains and modes are shown in figure 3. The data-driven approximation of the optimal gain is very accurate in a narrow band around the optimal frequency $\omega = 0$, and closely follows the trend of the operator-based gain distribution over a broad range of frequencies. The quality of the approximation degrades for larger frequencies where the magnitude of the optimal gain decreases.

The second dataset considers a single simulation started using the optimal forcing mode as the initial condition. In this scenario, the optimal gain and modes can be learned with only $r = 20$ DMD eigenvectors, as shown in figure 3. It is interesting to look at the DMD eigenvalues, which for this case form a small subset of those learned from the first dataset, as shown in figure 3. In the previous case we learned the eigenvalues with largest real part, now we discover ones from a very specific subset of the complex plane. This highlights that, although the resolvent modes can be accurately represented by very few eigenvectors, the amount of data required to learn those eigenvectors that form an efficient basis is highly dependent on the dynamic trajectories sampled. Moreover, this opens up exciting research directions, emphasizing the importance of finding principled disturbance designs to effectively probe dynamical systems.

Lastly, the third dataset considers a localized disturbance as initial condition, which was previously studied by Ilak & Rowley (2008). The exact form of the wall-normal velocity component is

$$v(x, y, z, 0) = \left(1 - \frac{r^2}{c_r^2}\right)(\cos(\pi y) + 1)\exp\left(-r^2/c_r^2 - y^2/c_y^2\right), \tag{4.3}$$

where $r^2 = (x - \pi)^2 + (z - \pi)^2$, and the parameters are set to $c_r = 0.7$ and $c_y = 0.6$. This type of disturbance is close to what could be generated in experiments using a spanwise and streamwise periodic array of axisymmetric jets injecting fluid perpendicular to the wall. Data-driven resolvent analysis is performed using $r = 200$ DMD eigenvectors, and the resulting optimal forcing and response modes do not resemble the operator-based ones, as shown in figure 3. In this scenario, the true resolvent modes of the system do not lie on the span of the learned DMD eigenvectors. More importantly, this occurs because some of the eigenvectors required to represent the true resolvent modes are not in the span of the data snapshots, and thus cannot be learned by DMD. In fact, we show in figure 3 that the learned resolvent modes for this dataset coincide with the operator-based ones when the input and output subspaces are constrained to lie on the span of the data snapshots.

5. Limitations and outlook

In this section we discuss the main limitations of the presented method and offer our perspective on future developments and applications. Specifically, we comment on computational cost and memory allocation, data requirements in terms of spatio-temporal resolution, inherent difficulties due to non-normality, the extension to nonlinear flows and the application of the method using partial measurements and limited actuation.

5.1. *Computational cost and memory allocation*

As already mentioned in § 3, because of the projection onto the DMD learned eigen-basis, our method builds a reduced-order resolvent operator of size $r \times r$. The subsequent analysis requires taking an SVD with an operation count of up to $O(r^3)$, instead of the $O(n^3)$ operations taxed by the standard matrix-forming operator-based approach, where n is the number of state variables. Although the potential time savings are promising, the computational bottleneck of the presented data-driven method is the truncated SVD in the DMD step, with an operation count of $O(np^2m^2)$, where p is the number of trajectories evolving from different initial conditions and m is the number of snapshots recorded along each of them. This is typically greater than $O(r^3)$, but still lower than $O(n^3)$. For a very high-dimensional system in which many transient trajectories need to be recorded, the memory allocation of the data matrices can also become challenging. For instance, in our channel flow example, dataset 1 considers 12 000 snapshots of state vectors with $\sim 2 \times 10^5$ entries. Fortunately, our approach benefits from all past and future innovations to improve the accuracy, robustness, flexibility, and speed of DMD (Kutz *et al.* 2016). In particular, the randomized DMD technique developed by Erichson *et al.* (2019*b*) may be useful to reduce the computational cost, and the streaming DMD algorithm by Hemati *et al.* (2014) can be used to avoid allocating the full data matrices.

5.2. *Spatio-temporal resolution*

A challenging aspect for the application of our method in an experimental scenario is the requirement of measurements that are spatially and temporally resolved. As in the operator-based approach, an accurate spatial discretization is required, since we need to be able to resolve a large number of DMD modes that conform the basis in which the computed resolvent modes are synthesized. Regarding temporal resolution, the sampling frequency for DMD is lower bounded by the Nyquist criterion, meaning that the signal of interest should be sampled with at least twice its inherent frequency.

It is worth mentioning that there are several extensions and modifications to DMD that may help alleviate these requirements by exploiting low-rank structure and sparsity in the data. If there are only a few dominant modes in the dataset, compressed sensing techniques enable the use of DMD with data that are heavily sub-sampled in space (Brunton *et al.* 2015; Erichson, Brunton & Kutz 2019*a*). Moreover, non-uniform sampling in time can be performed at an average rate that is below the Nyquist cutoff, provided that the time dynamics of the dominant modes admits a sparse Fourier representation (Tu *et al.* 2014*a*; Guéniat, Mathelin & Pastur 2015).

However, the sparsity and low-rank assumptions do not typically hold for transient dynamics of non-normal systems, as discussed in the next section. Because of this, we advise caution in the implementation of these DMD extensions to perform data-driven resolvent analysis. In the setting of conventional DMD, a good practice to select the time resolution is sampling the system at approximately three times the Nyquist cutoff (Schmid 2010). A final recommendation is to use a high-order integration quadrature to define the inner-product weighting, which may help alleviate the spatial resolution requirement.

5.3. *Non-normality*

As highlighted in our results for both examples, an accurate approximation of the resolvent modes requires a dataset composed of several transient trajectories of the system initialized from a large enough set of different initial disturbances to ensure coverage of the dynamical

features of the system. However, for highly non-normal systems, enlarging or enriching the dataset may not be enough.

Perhaps the most critical potential pitfall of the application of our method is the intrinsic difficulty to get a numerically robust spectrum in highly non-normal flows (Trefethen & Embree 2005). These systems are characterized by eigenvectors that are close to being parallel and are therefore extremely sensitive to perturbations of the dynamics operator (Trefethen & Embree 2005; Sipp *et al.* 2010). At the same time, it is well established that, for non-normal flows, individual global modes carry no physical meaning, and only a linear combination of a great many of them yields physically relevant features, such as the resolvent modes (Sipp *et al.* 2010). As a consequence, machine precision may not be enough to accurately compute an eigen-basis that is sufficiently large to reconstruct the dynamics, as shown by Garnaud *et al.* (2013) for the case of a two-dimensional jet. Moreover, in unbounded flows that are advection dominated, the introduction of artificial upstream or downstream boundaries to truncate the computational domain may give rise to spurious global modes (Lesshafft 2018).

These difficulties translate to an even more challenging approximation of the eigenvalues and eigenvectors of the system via DMD. Further, using an eigen-basis containing spurious vectors to perform data-driven resolvent analysis poses the risk of producing unreliable results. Due to the ubiquity of unbounded and highly non-normal flows, we expect future research to focus on these aspects.

5.4. *Nonlinearity*

An inherent limitation of the present method is that it is only suitable for systems in a linear regime, i.e. nonlinear systems close to an equilibrium point or actual linear systems. In the case of disturbances close to an equilibrium point, a simple test to ensure that the system is evolving linearly and that use of our method is appropriate, would be checking if DMD is able to reconstruct the recorded dynamics. Further, the extension to nonlinear systems would enable the analysis of statistically stationary turbulent flows provided that the mean-flow linearized operator was inferred from data. However, this requires yet another new method that is capable of distinguishing between the linear and nonlinear dynamics governing the evolution of spatio-temporal measurements, and is therefore outside of the scope of the present work.

5.5. *Partial measurements and limited actuation*

In an experimental setting, full-state measurements will rarely be available, and most likely we will only have access to measurements in a bounded spatial region and perhaps of only certain state variables. In such a scenario, we cannot expect our data-driven method to infer input–output relations dominated by unobserved variables, but instead it would produce forcing and response modes that are both restricted to the measurement subspace. More formally, if the state is represented in terms of an eigen-basis, the partial measurements can be expressed as $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} = \boldsymbol{C}\boldsymbol{V}_r\boldsymbol{a}$. If this basis contains the minimal set of eigenvectors needed to represent the resolvent modes with a prescribed level of accuracy, then we require that all columns of $\boldsymbol{V}_r$ are in the row space of $\boldsymbol{C}$, otherwise those eigenvectors cannot be learned from the measurements. Therefore, success of data-driven resolvent analysis using partial measurements is possible, but highly case dependent.

Moreover, in an experiment the types of actuation to generate initial disturbances will be limited by hardware. Nevertheless, as shown in our channel flow example with localized forcing (dataset 3), even though the flow is forced by a specific input, we

learn about the response to inputs living in a much larger subspace. Specifically, the learned forcing modes are constrained to the subspace spanned by the whole sequence of data snapshots, whereas the actual impulse corresponds to the first snapshot only. More formally, if initial disturbances are generated by an actuator configuration limited to producing inputs $\boldsymbol{f} = \boldsymbol{Bu}$, then the learned forcing and response modes are constrained to live in the controllable subspace of the system, rather than just the column space of $\boldsymbol{B}$. This result is very encouraging, demonstrating that we are able to learn a highly amplified, albeit sub-optimal, forcing mode and its corresponding response mode from an experimentally plausible disturbance. Interestingly, this result implies that, with enough measurements, system controllability is a sufficient, although not necessary, condition to guarantee convergence of data-driven results to the operator-based ones from full resolvent analysis.

The limitations described above make it challenging for this method to approximate the forcing and response modes of the full resolvent operator in an experimental setting. However, even with partial measurements and a limited set of initial disturbances, we envisage experimental applications of data-driven resolvent analysis to learn high-gain input–output pairs that, although will not correspond to the optimal forcing and response, may provide valuable 'and otherwise unavailable' information to guide controller design in linear non-normal systems.

6. Summary and conclusions

In this work, we have developed an algorithm to perform resolvent analysis based on time-resolved data of dynamical systems. Unlike other modal decompositions, resolvent modes provide insight into the most amplified states, the most sensitive actuator locations and the most responsive control inputs. Our method relies on DMD to learn approximate eigenvalues and eigenvectors of the underlying linear operator from snapshots of the transient dynamics of the system. Subsequently, we are able to compute the resolvent operator projected onto the learned eigenbasis. We perform data-driven resolvent analysis on numerical data of the linearized complex Ginzburg–Landau equation and of disturbances in a three-dimensional plane channel flow, demonstrating agreement between the leading data-driven and operator-based resolvent modes and gain distribution. A critical requirement is the design of initial disturbances to generate transients that are dynamically rich. We show that, using disturbances from a localized actuator, our method recovers the optimal forcing and response of the underlying system projected onto the span of the measured snapshots. This stresses the need for strategies to effectively explore the state space of a dynamical system.

The proposed algorithm performs resolvent analysis in an equation-free and adjoint-free manner, therefore opening the possibility of only using experimental measurements. Data-driven resolvent analysis will play a significant role in lowering the barrier of entry to resolvent research and applications. Our results are encouraging for linearly stable flows; however, more work is required before applying this technique to turbulent flows, where the linear and nonlinear contributions to the transient dynamics of the system must be disambiguated. All of the code developed for this work is available on github.com/ben-herrmann, and all the data are available for sharing upon request to the corresponding author.

Acknowledgements. We gratefully acknowledge L. Cordier, as well as the anonymous referees, for helpful comments and suggestions.

Funding. This work was supported by the PRIME programme of the German Academic Exchange Service (DAAD) with funds from the German Federal Ministry of Education and Research (BMBF) and by the U.S. Army Research Office (ARO W911NF-17-1-0306).

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

Author ORCIDs.
Benjamin Herrmann https://orcid.org/0000-0001-9265-7437;
Peter J. Baddoo https://orcid.org/0000-0002-8671-6952;
Richard Semaan https://orcid.org/0000-0002-3219-0545;
Steven L. Brunton https://orcid.org/0000-0002-6565-5118;
Beverley J. McKeon https://orcid.org/0000-0003-4220-1583.

References

Abreu, L.I., Cavalieri, A.V., Schlatter, P., Vinuesa, R. & Henningson, D.S. 2020 Resolvent modelling of near-wall coherent structures in turbulent channel flow. *Intl J. Heat Fluid Flow* **85**, 108662.

Askham, T. & Kutz, J.N. 2018 Variable projection methods for an optimized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **17** (1), 380–416.

Bagheri, S., Åkervik, E., Brandt, L. & Henningson, D.S. 2009*a* Matrix-free methods for the stability and control of boundary layers. *AIAA J.* **47** (5), 1057–1068.

Bagheri, S., Henningson, D., Hoepffner, J. & Schmid, P. 2009*b* Input-output analysis and control design applied to a linear model of spatially developing flows. *Appl. Mech. Rev.* **62** (2), 020803.

Beneddine, S., Sipp, D., Arnault, A., Dandois, J. & Lesshafft, L. 2016 Conditions for validity of mean flow stability analysis. *J. Fluid Mech.* **798**, 485–504.

Blanchard, A. & Sapsis, T.P. 2019 Learning the tangent space of dynamical instabilities from data. *Chaos* **29** (11), 113120.

Brenner, M., Eldredge, J. & Freund, J. 2019 Perspective on machine learning for advancing fluid mechanics. *Phys. Rev. Fluids* **4** (10), 100501.

Brunton, S.L. & Kutz, J.N. 2019 *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press.

Brunton, S.L., Noack, B.R. & Koumoutsakos, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508.

Brunton, S.L., Proctor, J.L. & Kutz, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **113** (15), 3932–3937.

Brunton, S.L., Proctor, J.L., Tu, J.H. & Kutz, J.N. 2015 Compressed sensing and dynamic mode decomposition. *J. Comput. Dyn.* **2** (2), 165–191.

Butler, K.M. & Farrell, B.F. 1992 Three-dimensional optimal perturbations in viscous shear flow. *Phys. Fluids* A **4** (8), 1637–1650.

Del Alamo, J.C. & Jimenez, J. 2006 Linear energy amplification in turbulent channels. *J. Fluid Mech.* **559**, 205–213.

Duraisamy, K., Iaccarino, G. & Xiao, H. 2019 Turbulence modeling in the age of data. *Annu. Rev. Fluid Mech.* **51**, 357–377.

Erichson, N.B., Brunton, S.L. & Kutz, J.N. 2019*a* Compressed dynamic mode decomposition for background modeling. *J. Real-Time Image Process.* **16** (5), 1479–1492.

Erichson, N.B., Mathelin, L., Kutz, J.N. & Brunton, S.L. 2019*b* Randomized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **18** (4), 1867–1891.

Farrell, B.F. & Ioannou, P.J. 1993*a* Optimal excitation of three-dimensional perturbations in viscous constant shear flow. *Phys. Fluids* A **5** (6), 1390–1400.

Farrell, B.F. & Ioannou, P.J. 1993*b* Stochastic forcing of the linearized Navier–Stokes equations. *Phys. Fluids* A **5** (11), 2600–2609.

Garnaud, X., Lesshafft, L., Schmid, P. & Huerre, P. 2013 Modal and transient dynamics of jet flows. *Phys. Fluids* **25** (4), 044103.

Gibson, J.F. 2014 Channelflow: a spectral Navier–Stokes simulator in C++. *Tech. Rep.* University of New Hampshire.

Gibson, J.F., Halcrow, J. & Cvitanović, P. 2008 Visualizing the geometry of state space in plane Couette flow. *J. Fluid Mech.* **611**, 107–130.

GÓMEZ, F. & BLACKBURN, H. 2017 Data-driven approach to design of passive flow control strategies. *Phys. Rev. Fluids* **2** (2), 021901.

GÓMEZ, F., BLACKBURN, H.M., RUDMAN, M., SHARMA, A.S. & MCKEON, B.J. 2016 A reduced-order model of three-dimensional unsteady flow in a cavity based on the resolvent operator. *J. Fluid Mech.* **798**, R2.

GUÉNIAT, F., MATHELIN, L. & PASTUR, L.R. 2015 A dynamic mode decomposition approach for large and arbitrarily sampled systems. *Phys. Fluids* **27** (2), 025113.

GUSTAVSSON, L.H. 1991 Energy growth of three-dimensional disturbances in plane Poiseuille flow. *J. Fluid Mech.* **224**, 241–260.

HALKO, N., MARTINSSON, P.-G. & TROPP, J.A. 2011 Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **53** (2), 217–288.

HEMATI, M.S., WILLIAMS, M.O. & ROWLEY, C.W. 2014 Dynamic mode decomposition for large and streaming datasets. *Phys. Fluids* **26** (11), 111701.

HERRMANN, B., CALDERÓN-MUÑOZ, W.R., DIAZ, G. & SOTO, R. 2018*a* Heat transfer enhancement strategies in a swirl flow minichannel heat sink based on hydrodynamic receptivity. *Intl J. Heat Mass Transfer* **127**, 245–256.

HERRMANN, B., CALDERÓN-MUÑOZ, W.R. & SOTO, R. 2018*b* Stability and receptivity of boundary layers in a swirl flow channel. *Acta Mechanica* **229** (10), 4005–4015.

HWANG, Y. & COSSU, C. 2010*a* Amplification of coherent streaks in the turbulent Couette flow: an input-output analysis at low Reynolds number. *J. Fluid Mech.* **643**, 333–348.

HWANG, Y. & COSSU, C. 2010*b* Linear non-normal energy amplification of harmonic and stochastic forcing in turbulent channel flow. *J. Fluid Mech.* **664**, 51–73.

ILAK, M. & ROWLEY, C.W. 2008 Modeling of transitional channel flow using balanced proper orthogonal decomposition. *Phys. Fluids* **20** (3), 034103.

JEUN, J., NICHOLS, J.W. & JOVANOVIĆ, M.R. 2016 Input-output analysis of high-speed axisymmetric isothermal jet noise. *Phys. Fluids* **28** (4), 047101.

JOVANOVIĆ, M.R. 2021 From bypass transition to flow control and data-driven turbulence modeling: an input–output viewpoint. *Annu. Rev. Fluid Mech.* **53**, 311–345.

JOVANOVIĆ, M.R. & BAMIEH, B. 2005 Componentwise energy amplification in channel flows. *J. Fluid Mech.* **534**, 145–183.

JOVANOVIĆ, M.R., SCHMID, P.J. & NICHOLS, J.W. 2014 Sparsity-promoting dynamic mode decomposition. *Phys. Fluids* **26** (2), 024103.

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

LESSHAFFT, L. 2018 Artificial eigenmodes in truncated flow domains. *Theor. Comput. Fluid Dyn.* **32** (3), 245–262.

LESSHAFFT, L., SEMERARO, O., JAUNET, V., CAVALIERI, A.V. & JORDAN, P. 2019 Resolvent-based modeling of coherent wave packets in a turbulent jet. *Phys. Rev. Fluids* **4** (6), 063901.

LI, H., FERNEX, D., SEMAAN, R., TAN, J., MORZYŃSKI, M. & NOACK, B.R. 2021 Cluster-based network model. *J. Fluid Mech.* **906**, A21.

LIBERTY, E., WOOLFE, F., MARTINSSON, P.-G., ROKHLIN, V. & TYGERT, M. 2007 Randomized algorithms for the low-rank approximation of matrices. *Proc. Natl Acad. Sci. USA* **104** (51), 20167–20172.

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

LOISEAU, J.-C., BUCCI, M.A., CHERUBINI, S. & ROBINET, J.-C. 2019 Time-stepping and Krylov methods for large-scale instability problems. In *Computational Modelling of Bifurcations and Instabilities in Fluid Dynamics* (ed. A. Gelfgat), pp. 33–73. Springer.

LUMLEY, J.L. 1970 *Stochastic tools in Turbulence*. Academic Press.

MARTINI, E., RODRÍGUEZ, D., TOWNE, A. & CAVALIERI, A.V.G. 2020 Efficient computation of global resolvent modes. arXiv:2008.10904.

MCKEON, B.J. 2017 The engine behind (wall) turbulence: perspectives on scale interactions. *J. Fluid Mech.* **817**, P1.

MCKEON, B.J. & SHARMA, A.S. 2010 A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.* **658**, 336–382.

MEZIĆ, I. 2013 Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45** (1), 357–378.

MOARREF, R., JOVANOVIĆ, M., TROPP, J., SHARMA, A. & MCKEON, B. 2014 A low-order decomposition of turbulent channel flow via resolvent analysis and convex optimization. *Phys. Fluids* **26** (5), 051701.

MOARREF, R., SHARMA, A.S., TROPP, J.A. & MCKEON, B.J. 2013 Model-based scaling of the streamwise energy density in high-Reynolds-number turbulent channels. *J. Fluid Mech.* **734**, 275–316.

MONOKROUSOS, A., ÅKERVIK, E., BRANDT, L. & HENNINGSON, D.S. 2010 Global three-dimensional optimal disturbances in the Blasius boundary-layer flow using time-steppers. *J. Fluid Mech.* **650**, 181–214.

PADOVAN, A., OTTO, S.E. & ROWLEY, C.W. 2020 Analysis of amplification mechanisms and cross-frequency interactions in nonlinear flows via the harmonic resolvent. *J. Fluid Mech.* **900**, A14.

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

QIAN, E., KRAMER, B., PEHERSTORFER, B. & WILLCOX, K. 2020 Lift and learn: physics-informed machine learning for large-scale nonlinear dynamical systems. *Physica* D **406**, 132401.

RAISSI, M., PERDIKARIS, P. & KARNIADAKIS, G.E. 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707.

RAISSI, M., YAZDANI, A. & KARNIADAKIS, G.E. 2020 Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* **367** (6481), 1026–1030.

REDDY, S.C. & HENNINGSON, D.S. 1993 Energy growth in viscous channel flows. *J. Fluid Mech.* **252**, 209–238.

RIBEIRO, J.H.M., YEH, C.-A. & TAIRA, K. 2020 Randomized resolvent analysis. *Phys. Rev. Fluids* **5** (3), 033902.

RIGAS, G., SIPP, D. & COLONIUS, T. 2021 Nonlinear input/output analysis: application to boundary layer transition. *J. Fluid Mech.* **911**, A15.

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

RUDY, S.H., BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2017 Data-driven discovery of partial differential equations. *Sci. Adv.* **3** (4), e1602614.

SCHMID, P.J. 2007 Nonmodal stability theory. *Annu. Rev. Fluid Mech.* **39**, 129–162.

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

SCHMID, P.J. & BRANDT, L. 2014 Analysis of fluid systems: stability, receptivity, sensitivity. *Appl. Mech. Rev.* **66** (2), 024803.

SCHMID, P.J. & HENNINGSON, D.S. 2001 *Stability and Transition in Shear Flows*, vol. 142. Springer.

SCHMIDT, O.T., TOWNE, A., RIGAS, G., COLONIUS, T. & BRÈS, G.A. 2018 Spectral analysis of jet turbulence. *J. Fluid Mech.* **855**, 953–982.

SIPP, D., MARQUET, O., MELIGA, P. & BARBAGALLO, A. 2010 Dynamics and control of global instabilities in open-flows: a linearized approach. *Appl. Mech. Rev.* **63** (3), 030801.

SUN, Y., LIU, Q., CATTAFESTA III, L.N., UKEILEY, L.S. & TAIRA, K. 2020 Resolvent analysis of compressible laminar and turbulent cavity flows. *AIAA J.* **58** (3), 1046–1055.

TAIRA, K., BRUNTON, S.L., DAWSON, S., 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. 2019 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.

TREFETHEN, L.N., TREFETHEN, A.E., REDDY, S.C. & DRISCOLL, T.A. 1993 Hydrodynamic stability without eigenvalues. *Science* **261** (5121), 578–584.

TU, J.H., ROWLEY, C.W., KUTZ, J.N. & SHANG, J.K. 2014*a* Spectral analysis of fluid flows using sub-Nyquist-rate PIV data. *Exp. Fluids* **55** (9), 1–13.

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

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

YEH, C.-A. & TAIRA, K. 2019 Resolvent-analysis-based design of airfoil separation control. *J. Fluid Mech.* **867**, 572–610.