Sparse reduced-order modeling: Sensor-based dynamics to full-state estimation

Jean-Christophe Loiseau\textsuperscript{1,}†, Bernd R. Noack\textsuperscript{2,3,4} and Steven L. Brunton\textsuperscript{5}

\textsuperscript{1}Laboratoire DynFluid, Arts et Métiers ParisTech, 75013 Paris, France
\textsuperscript{2}Laboratoire d’Informatique pour la Mécanique et les Sciences de l’Ingénieur, LIMSI-CNRS, Rue John von Neumann, Campus Universitaire d’Orsay, Bât 508, F-91403 Orsay, France
\textsuperscript{3}Institute for Turbulence-Noise-Vibration Interaction and Control, Harbin Institute of Technology, Shenzhen Campus, Shenzhen, Peoples Republic of China
\textsuperscript{4}Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin, Müller-Breslau-Straße 8, D-10623 Berlin, Germany
\textsuperscript{5}Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, USA

(Received xx; revised xx; accepted xx)

We propose a general dynamic reduced-order modeling framework for typical experimental data: time-resolved sensor data and optional non-time-resolved particle image velocimetry (PIV) snapshots. This framework can be decomposed into four building blocks. First, the sensor signals are lifted to a dynamic feature space without false neighbours. Second, we identify a sparse human-interpretable nonlinear dynamical system for the feature state based on the sparse identification of nonlinear dynamics (SINDy). Third, if PIV snapshots are available, a local linear mapping from the feature state to the velocity field is performed to reconstruct the full-state of the system. Fourth, a generalized feature-based modal decomposition identifies coherent structures that are most dynamically correlated with the linear and nonlinear interaction terms in the sparse model, adding interpretability. Steps 1 and 2 define a black-box model. Optional steps 3 and 4 lift the black-box dynamics to a gray-box model in terms of the identified coherent structures, if non-time-resolved full-state data is available. This gray-box modeling strategy is successfully applied to the transient and post-transient laminar cylinder wake, and compares favorably with a POD model. We foresee numerous applications of this highly flexible modeling strategy, including estimation, prediction, and control. Moreover, the feature space may be based on intrinsic coordinates, which are unaffected by a key challenge of modal expansion: the slow change of low-dimensional coherent structures with changing geometry and varying parameters.

1. Introduction

Understanding, modeling and controlling complex fluid flows is a central focus in many scientific, technological, and industrial applications, including energy (e.g. wind, tidal, and combustion), transportation (e.g. planes, trains, and automobiles), security (e.g. airborne contamination), and medicine (e.g. artificial hearts and artificial respiration). Improved models of engineering flows have the potential to dramatically improve performance in these systems through optimization and control, resulting in practical gains such as drag reduction, lift increase, and mixing enhancement (Fabbiane \textit{et al.} 2014; Brunton \& Noack 2015; Sipp \& Schmid 2016; Rowley \& Dawson 2016).

† Email address for correspondence: loiseau.jc@gmail.com
Although the Navier-Stokes equations provide a detailed mathematical model, this representation may be difficult to use for engineering design, optimization, and control. Instead, they are commonly discretized into a high-dimensional, nonlinear dynamical system with many degrees of freedom and multi-scale interactions. These equations are nonetheless expensive to simulate, making them unwieldy for iterative optimization or in-time control. They may also obscure the underlying physics, which often evolves on a low-dimensional attractor (Holmes et al. 2012; Noack et al. 2003). The various fidelities of model description were described by Wiener (1948): ‘white-box’ describes an accurate evolution equation based on first principles (e.g. Navier-Stokes discretization), ‘gray-box’ describes a low-dimensional model approximating the full state (e.g. POD-Galerkin models), and ‘black-box’ describes input-output models that lack a connection to the full state space (e.g. neural networks).

In the following, we outline related reduced-order models as our point of departure in §1.1 and foreshadow proposed innovations of this study in §1.2.

1.1. Related reduced-order models as point of departure

Reduced-order models provide low-dimensional descriptions of the underlying fluid behavior in a compact and computationally efficient representation. There are many techniques for reduced-order modeling, ranging from physical reductions to purely data-driven methods, and nearly everything in between. Proper orthogonal decomposition (POD) (Sirovich 1987; Berkooz et al. 1993; Holmes et al. 2012) provides a low-rank modal decomposition of fluid flow field data, extracting the most energetic modes. It is then possible to Galerkin project the Navier-Stokes equations onto these modes, resulting in an approximate, low-dimensional model in terms of mode coefficients (Noack et al. 2011; Carlberg et al. 2015). POD-Galerkin models are widely used, as they are interpretable, gray-box models, and it is straightforward to reconstruct the high-dimensional flow field from the low-dimensional model via POD modes. The first pioneering example of Aubry et al. (1988) featured wall turbulence, almost three decades ago. Subsequent POD models have been developed for the transitional boundary layer (Rempfer & Fasel 1994), the mixing layer (Ukeiley et al. 2001; Wei & Rowley 2009), the cylinder wake (Deane et al. 1991; Galletti et al. 2004), and the Ahmed body wake (Östh et al. 2014), to name only a few.

POD-Galerkin modeling is challenging for changing domains (Bourguet et al. 2011), changing boundary conditions (Graham et al. 1999) and slow deformation of the modal basis (Babaee & Sapsis 2016). Standard Galerkin projection can also be expected to suffer from stability issues (Rempfer 2000; Schlegel & Noack 2015; Carlberg et al. 2017), although including energy-preserving constraints may improve the long-time stability and performance of nonlinear models (Balajewicz et al. 2013; Cordier et al. 2013). POD-Galerkin models tend to be valid for a narrow range of operating conditions, near those of the data set used to generate the POD modes. Transients also pose a challenge to POD modeling. Noack et al. (2003) and Tadmor et al. (2010) demonstrate the ability of a low-dimensional model to reproduce nonlinear transients of the von Kármán vortex shedding past a two-dimensional cylinder, provided the projection basis includes a **shift mode** quantifying the distortion between the linearly unstable base flow and marginally stable mean flow. These techniques have been extended to include the effect of wall actuation (Graham et al. 1999; Rediniotis et al. 2002).

In addition to the physics-informed Galerkin projection, data-driven modeling approaches are prevalent in fluid dynamics (Brunton & Noack 2015; Rowley & Dawson 2016). For example, dynamic mode decomposition (DMD) (Schmid 2010; Rowley et al. 2009; Kutz et al. 2016), the eigensystem realization algorithm (ERA) (Juang & Pappa...
Sparse reduced-order modeling

1985), Koopman analysis (Mezić 2005, 2013; Tu et al. 2014; Williams et al. 2015), cluster-based reduced order models (CROM) (Kaiser et al. 2014), NARMAX models (Billings 2013; Semeraro et al. 2016; Zhang et al. 2012; Glaz et al. 2010), and network analysis (Nair & Taira 2015) have all been used to identify dynamical systems models from fluids data, without relying on prior knowledge of the underlying Navier-Stokes equations. DMD models are readily obtained directly from data, and they provide interpretability in terms of flow structures, but the resulting models are linear, and the connection to nonlinear systems is tenuous unless DMD is enriched with nonlinear functions of the data (Williams et al. 2015; Kutz et al. 2016). Neural networks have long been used for flow modeling and control (Milano & Koumoutsakos 2002; Zhang & Duraisamy 2015; Lee et al. 1997; Krizhevsky et al. 2012), and recently deep neural networks have been used for Reynolds averaged turbulence modeling (Ling et al. 2016; Kutz 2017). However, many machine learning methods may be prone to overfitting, have limited interpretability, and make it difficult to incorporate known physical constraints. Parsimony has thus become an overarching goal when using machine learning to model nonlinear dynamics.

In the seminal work of Bongard & Lipson (2007) and Schmidt & Lipson (2009) governing dynamics and conservation laws are discovered using genetic programming along with a Pareto analysis to balance model accuracy and complexity, preventing overfitting. Recently, Brunton et al. (2016b) introduced the sparse identification of nonlinear dynamics (SINDy), which identifies parsimonious nonlinear models from data. SINDy follows the principle of Ockham’s razor, resting on the assumption that there are only a few important terms that govern the dynamics of a system, so that the equations are sparse in the space of possible functions. Sparse regression is then used to efficiently determine the fewest terms in the dynamics required to accurately represent the data, preventing overfitting. Because SINDy is based on linear algebra (i.e. the nonlinear dynamics are represented as a linear combination of candidate nonlinear functions), the method is readily extended to incorporate known physical constraints (Loiseau & Brunton 2016). In general, it is possible to obtain nonlinear models using genetic programming or SINDy on POD or DMD mode coefficients, which make these methods gray box, having a transformation from the model back to the high-dimensional, interpretable state-space. However, models developed on POD/DMD mode coefficients may still suffer from fundamental challenges of traditional POD-Galerkin models, such as capturing changing boundary conditions, moving geometry, and varying operating condition.

1.2. Contribution of this work

In this work, we introduce a new gray-box modeling procedure that yields interpretable nonlinear models from measurement data. The method is applied to the well-investigated two-dimensional transient flow past a circular cylinder with slow change of the base flow and varying coherent structures (Tadmor et al. 2011). In particular, we develop nonlinear models only from lift measurements that accurately capture steady-state and transient flow behavior. First, a feature vector is constructed from the lift signal, including a time-delayed value. This lifting shall avoid false neighbours, i.e. neighboring states with significantly different temporal evolution. Second, a sparse dynamical model is identified in this feature space. For the following steps, full-state measurement data is assumed to be available. Third, a local linear mapping from feature vector to velocity field is constructed using a K-nearest neighbors (KNN) approach. This mapping provides significantly more accurate flow reconstruction, as compared to a POD-Galerkin model of the same order. Technically, we mitigate the significant challenges of using an ‘elliptic’ Galerkin modeling approach to an ‘hyperbolic’ dynamics: the global modes connect all locations in all
Figure 1. Schematic overview of the proposed sparse modeling procedure. A sparse dynamical system (Brunton et al. 2016b) is identified based on features obtained from sensor signals $s$, and the full state $u$ may also be estimated with the availability of PIV snapshots (optional).

To summarize, the resulting gray-box modeling procedure has the following beneficial features: (i) it captures nonlinear physics, (ii) it is based on a simple, non-invasive computational algorithm, (iii) the resulting model is interpretable in terms of non-linear interaction physics and generalized modes (optional with full-state data), and (iv) modeling feature vectors is more robust to mode deformation, moving geometry, and varying operating condition. This procedure is shown schematically in figure 1.

The manuscript is organized as follows: §2 provides an overview of the flow configuration considered in this work, namely the incompressible, two-dimensional flow past a circular cylinder at $Re = 100$. §3 describes the proposed gray-box modeling procedure, including modeling in feature space and obtaining a generalized modal expansion if full-state data is available. In §4, numerical results for the gray-box modeling procedure are presented and analyzed for the cylinder flow. §5 provides some leads for application of the present approach to higher-dimensional systems. It also highlights some connections
with previous works. Finally, §6 summarizes our key findings and provides the reader with possible future directions to extend this work.

2. Flow configuration

The flow configuration considered in the present work is the two-dimensional incompressible viscous flow past a circular cylinder at $Re = 100$, based on the free-stream velocity $U_\infty$, the cylinder diameter $D$ and the kinematic viscosity $\nu$. This Reynolds number is well above the critical Reynolds number ($Re_c = 48$) for the onset of the two-dimensional vortex shedding (Zebib 1987; Schumm et al. 1994) and below the critical Reynolds number ($Re_c = 188$) for the onset of three-dimensional instabilities (Zhang et al. 1995; Barkley & Henderson 1996). In the fluid dynamics community, a large body of literature exists in which this particular setup has been chosen to illustrate modal decomposition (Bagheri 2013) and model identification techniques (Noack et al. 2003; Sengupta et al. 2015; Brunton et al. 2016b; Rowley & Dawson 2016). This setup is thus a particularly compelling test case to illustrate our model identification strategy, as well as to draw connections and quantify its performance against other well-established techniques.

The dynamics of the flow are governed by the incompressible Navier-Stokes equations

\[
\frac{\partial u}{\partial t} + (u \cdot \nabla) u = -\nabla p + \frac{1}{Re} \nabla^2 u
\]

\[
\nabla \cdot u = 0,
\]

where $u = (u, v)^T$ and $p$ are the velocity and pressure fields, respectively. The center of the cylinder has been chosen as the origin of the reference frame $x = (x, y)$, where $x$ denotes the streamwise coordinate and $y$ denotes the spanwise coordinate. This study considers the same computational domain as in Noack et al. (2003), extending from $x = -5$ to $x = 15$ in the streamwise direction, and from $y = -5$ to $y = 5$ in the spanwise direction. A uniform velocity profile is prescribed at the inflow, a classical stress-free boundary condition is used at the outflow, and free-slip boundary conditions are used on the lateral boundaries of the computational domain. Based on the spectral element solver Nek 5000 (Fischer et al. 2008), the domain is discretized by 1832 seventh-order spectral elements. Finally, the time-integration of the diffusive terms relies on a Backward Differentiation of order 3 (BDF3), while the convective terms are advanced in time based on a third-order accurate extrapolation.

Two out of the three direct numerical simulations used in this work have been initialized with the following initial condition

\[
u(x, 0) = u_s(x) + 0.001\epsilon(x),
\]

where $u_s(x)$ is the linearly unstable steady solution of the Navier-Stokes equations and $\epsilon(x)$ is a zero-mean and unit-variance random white-noise velocity field. Each simulation is run for 150 convective time units, providing $M = 1200$ equidistantly sampled velocity snapshots $u^{(m)}(x) = u(x, t_m)$, $m = 1, \ldots, M$, and synchronous measurements of the lift and drag coefficients, $C_L(t_m)$ and $C_D(t_m)$. This time-span covers the entire unforced transient phase, from the steady solution to the fully developed von Kármán vortex street. Note that only one of these two simulations is used to identify the low-order model in §4.1, while the second one serves only cross-validation purposes.

Figure 2 depicts a typical evolution of the lift coefficient $C_L$, while figure 3 shows snapshots of the vorticity field at different instants of time. For $t \leq 50$, the flow is governed by linear dynamics. Consequently, the vorticity field of the perturbation
Figure 2. (a) Evolution of the lift coefficient $C_L$ as a function of time for the two-dimensional cylinder flow at $Re = 100$. (b) Trajectory of the system in the phase-plane ($C_L$, $dC_L/dt$). In both figures, the orange dots indicate the instants of time for which the corresponding vorticity field is shown in figure 3.

Figure 3. Snapshots of the vorticity field at the different time instants highlighted in figure 2. Note that the vorticity field of the linearly unstable baseflow, $u_s(x)$, has been subtracted in order to highlight the vorticity induced by the perturbation.

$\mathbf{v}(x,t) = \mathbf{u}(x,t) - \mathbf{u}_s(x)$, shown in figure 3(a), can be well approximated by the leading instability mode. For $50 \leq t \leq 80$, the perturbation grows to an extent that nonlinear effects cause the perturbation to distort. Eventually, for $t \geq 80$, the flow settles onto a periodic limit cycle corresponding to the classical von Kármán vortex street. Given the evolution of the lift coefficient depicted in figure 2 and the associated snapshots shown in figure 3, the aim of the present work is to propose a new reduced-order modeling strategy able to accurately reproduce such dynamics and flow structures. Note that this setup mimics an experiment in Hosseini et al. (2016) where the pressure difference between the top and bottom side of a cylinder is a surrogate quantity for the lift. This difference had been recorded in a time-resolved manner while non-time-resolved PIV flow snapshots were taken.

3. Sparse sensor-based modeling

Here, we discuss the core mathematical and algorithmic framework used to identify nonlinear reduced-order models and build a full-state estimator from data. The identification procedure relies on the sparse identification of nonlinear dynamics (SINDy) method (Brunton et al. 2016b) briefly summarized in §3.2. As a second step, we introduce the reader to two different full-state estimators in §3.3 based on stochastic estimation and nearest-neighbors algorithms.
3.1. From sensor signals to feature space

Performing direct numerical simulations allows one to gather space-time resolved data at every single point within the computational domain considered. In contrast, experimentalists typically have access only to a limited number of time-resolved sensor measurements $s$. The aim of the present work is thus to illustrate how one can leverage recent advances in system identification and machine learning to construct reduced-order models directly from these limited sensor measurements. For that purpose, experimental conditions are mimicked with direct numerical simulations and we consider a single sensor measurement given by

$$s(t) := C_L(t),$$

(3.1)

where $C_L$ is the lift coefficient. Note that in general $s$ may be a vector of measurements, including for instance the lift and drag coefficients, pressure measurements on an immersed body, or point velocity field measurements at selected locations. However, in the present study, the lift coefficient is sufficient to characterize the flow. Given the sensor measurements $s$, our aim is to identify a low-order model that allows us to predict the evolution of our system. Raw signals may unfortunately be ill-behaved for system identification purposes and may need to be augmented, or lifted, to include functions of the sensor measurements. We consider the augmented state $a$ to be a feature vector given by

$$a = g(s).$$

(3.2)

There are many choices for the mapping $g$ to enrich the sensor measurements and improve models. If the sensors are sufficient to define the state of the system, then $g$ may be the identity map. If the sensors consist of high-dimensional snapshots, then $g$ may extract POD mode coefficients. It may also be possible to augment the measurements with delay-embedding (Takens 1981; Juang & Pappa 1985; Brunton et al. 2017) in order to obtain a sufficiently high-dimensional feature vector $a$ as to fully characterize the dynamics of the system. More generally, choosing a good transformation $g$ is an important open problem, with connections to representation theory and the Koopman operator perspective on dynamical systems (Mezić 2005, 2013; Williams et al. 2015; Brunton et al. 2016a, 2017; Arbabi & Mezić 2016). In the present study, we choose $g$ to augment the sensor measurement of the lift coefficient with its time derivative, along with a proper scaling:

$$a_1(t) := s(t),$$

$$a_2(t) := \frac{1}{\omega_\infty} \frac{ds}{dt}(t),$$

(3.3)

where $\omega_\infty$ is the post-transient angular shedding frequency estimated from the power spectral density of the lift signal depicted in figure 2. This choice has been guided by physical considerations: the lift can be measured and allows for the characterization of the state of the flow. It is also well-known that the flow around a two-dimensional cylinder at $Re = 100$ only necessitates three degrees of freedom (Noack et al. 2003), or features, to be approximately described: the shedding amplitude, shedding phase and the degree of base flow deformation. In this work, we will show that the degree of base flow deformation can be described by the drag coefficient, which is then modeled as an algebraic equation of the feature vector $a$. In comparison, in Hosseini et al. (2016), the analogues of $a_1$ and $a_2$ were defined as the cosine and sine Morlet transforms of the pressure difference history. Please refer to §5.1 for a discussion on appropriate sensor selection and for the definition of the mapping function $g$ in flows where the number of degrees of freedom required to model the dynamics of the system may not be known a priori.
3.2. Sparse Identification of Nonlinear Dynamics (SINDy)

Identifying reduced-order models from data is a central challenge in mathematical physics, with a rich history of developments in fluid dynamics. The form of the dynamics is typically either constrained via prior knowledge, as in the Galerkin projection, or a particular model structure is chosen heuristically, and parameters are optimized to match the data. Simultaneous identification of the model structure and parameters from data is considerably more challenging, as there are combinatorially many possible model structures. The sparse identification of nonlinear dynamics (SINDy) architecture (Brunton et al. 2016b) bypasses the intractable combinatorial search through all possible model structures, leveraging the fact that many systems may be modeled by dynamics \( f \) that are sparse in the space of possible right-hand side functions:

\[
\frac{da}{dt} = f(a),
\]

where \( a \) is the same state vector as in §3.1. It is then possible to solve for the relevant terms that are active in the dynamics using either a convex \( \ell_1 \)-regularized regression (Tibshirani 1996) or a sequentially thresholded least-squares method (Brunton et al. 2016b). Both algorithms penalize the number of terms in the dynamics and scale favorably to large problems.

First, time-series data is collected and formed into a data matrix:

\[
A = \begin{bmatrix} a(t_1) & a(t_2) & \cdots & a(t_M) \end{bmatrix}^T
\]

(3.5)

where ‘\( T \)’ denotes the matrix transpose. A similar matrix of derivatives is formed:

\[
\dot{A} = \begin{bmatrix} \frac{da}{dt}(t_1) & \frac{da}{dt}(t_2) & \cdots & \frac{da}{dt}(t_M) \end{bmatrix}^T.
\]

(3.6)

In practice, this may be computed directly from the data in \( A \). However, for noisy data, the total-variation regularized derivative (Chartrand 2011) tends to provide numerically robust derivatives. Based on the data in \( A \), a library of candidate nonlinear functions \( \Theta(A) \) is constructed:

\[
\Theta(A) = \begin{bmatrix} 1 & A & P_2(A) & \cdots & P_d(A) & \cdots & \sin(A) & \cdots \end{bmatrix}.
\]

(3.7)

Here, \( P_d(A) \) denotes a matrix with column vectors given by all possible time-series of \( d \)-th degree polynomials in the state \( a \). Any basis functions may be used to generate the library \( \Theta \), although polynomials work well for fluids. The dynamical system in Eq. (3.4) may now be represented in terms of the data matrices in Eqs. (3.6) and (3.7) as

\[
\dot{A} = \Theta(A)\Xi.
\]

(3.8)

Each column \( \Xi_k \) in \( \Xi \) is a vector of coefficients determining the active terms in the \( k \)-th row equation in Eq. (3.4). A parsimonious model will provide an accurate model fit in Eq. (3.8) with as few terms as possible in \( \Xi \). Such a model may be identified using a convex \( \ell_1 \)-regularized sparse regression:

\[
\Xi_k = \arg\min_{\Xi_k} \|\dot{A}_k - \Theta(A)\Xi_k\|_2 + \lambda\|\Xi_k\|_1.
\]

(3.9)

Here, \( \dot{A}_k \) is the \( k \)-th column of \( \dot{A} \). Sparse regression, such as the LASSO (Tibshirani 1996) or the sequential thresholded least-squares algorithm used in SINDy, improves the numerical robustness of this identification for noisy overdetermined problems, in contrast to earlier methods (Wang et al. 2011) that used compressed sensing (Donoho...
Sparse reduced-order modeling

Once identified, the sparse vectors $\Xi_k$ may be synthesized into a nonlinear dynamical system model:

$$\frac{da_k}{dt} = \Theta(a)\Xi_k,$$

(3.10)

where $a_k$ is the $k$-th element of $a$ and $\Theta(a)$ is a row vector of symbolic functions of $a$, as opposed to the data matrix $\Theta(A)$. Identifying the most parsimonious nonlinear model by applying sparse regression in the library $\Theta$ is a convex procedure. The alternative approach, which involves regression onto every possible sparse nonlinear structure, constitutes an intractable brute-force procedure. SINDy thus bypasses this combinatorial search with modern convex optimization and machine learning. The SINDy algorithm is closely related to NARMAX models (Billings 2013) and fast function extraction (FFX) (McConaghy 2011).

A major benefit of the SINDy architecture is its ability to identify parsimonious models that contain only the required nonlinear terms, resulting in interpretable models that avoid overfitting. In the optimization above, the sparsifying parameter $\lambda$ may be varied from $\lambda = 0$ (i.e., least-squares) to $\lambda \to \infty$ (i.e., trivial dynamics $da/dt = 0$), sweeping out a Pareto front. To identify the most parsimonious model that best balances model complexity with accuracy, Mangan et al. (2017) proposed an efficient methodology to rank candidate models on the Pareto front using the Akaike information criterion (AIC) (Akaike 1974) or the Bayes information criterion (BIC) (Schwarz et al. 1978).

The embedding of nonlinear dynamics in terms of a linear regression problem in (3.9) makes the SINDy method highly extensible. Recent extensions to SINDy enable the identification of nonlinear differential equations with rational function nonlinearities by reformulating the problem as an implicit differential equation and solving for the active terms by finding the sparsest vector in the null space of an augmented library containing functions of the state and derivative terms (Mangan et al. 2016). SINDy has also been generalized to identify partial differential equations from data (Rudy et al. 2017; Schaeffer 2017), to include integral terms (Schaeffer & McCalla 2017), and it has been extended to include inputs and control (Brunton et al. 2016c; Kaiser et al. 2017).

3.3. Full-state estimation

The proposed methodology enables the identification of a low-order model that reproduces the system dynamics recorded by a few sensors. Although it may provide useful insights into the physics, it does not allow for a straightforward full-state estimation of the system considered. To accommodate this estimation, one needs to define a function

$$h(x, a(t)) \approx u(x, t)$$

(3.11)

mapping the state of the system from the low-dimensional feature space to the high-dimensional physical space. In the most general case, $h$ is a nonlinear mapping function. Note, however, that $h(x, 0)$ physically corresponds to the reference state $u_s(x)$ at $a = 0$. While this reference state is classically chosen as the mean flow, in the present work, it is chosen as the linearly unstable steady solution $u_s(x)$ of the Navier-Stokes equations. In the following, two different strategies to approximate $h(x, a)$ from data will be presented.

3.3.1. Local linear mapping

Let us consider, for the sake of simplicity and without loss of generality, the feature vector $a = [a_1 \ a_2]^T$. Given different transient evolutions of $a(t)$, and having stored the associated velocity field snapshots, the nonlinear mapping $h(x, a(t)) \approx u(x, t)$ can be approximated by a local linear mapping. If non-time-resolved snapshots are available,
Figure 4. In both figures, the blue line depicts the trajectory of the testing dataset for which we reconstruct the flow field. (a) Delaunay triangulation of the state plane. In addition to the two transient trajectories started from the fixed point, a third trajectory with an initial condition above the limit cycle has been used to obtain this triangulation. (b) Close-up view in the vicinity of the query point (•). The corresponding flow field can then be estimated as a weighted average of the flow fields associated to each vertex of the triangle highlighted in orange.

then only the corresponding synchronous instances of \( \mathbf{a} \) are considered. In the rest of this work, \( \mathbf{a}^\bullet(t) \) will denote time-evolutions of the feature vector obtained from a direct numerical simulation, while \( \mathbf{a}^\circ(t) \) will denote the evolution predicted by the low-dimensional models identified using the SINDy architecture outlined in §3.2. A Delaunay triangulation of the phase plane of the low-dimensional system can then be obtained from the transient evolutions of \( \mathbf{a}_1^\circ(t) \) and \( \mathbf{a}_2^\circ(t) \) in the training dataset. An example triangulation is illustrated in figure 4(a). Estimating the flow field associated with a point \( [\mathbf{a}_1^\circ \mathbf{a}_2^\circ]^T \) then amounts to a two-step procedure:

(i) Given the Delaunay triangulation of the phase plane, identify in which triangle the point \( [\mathbf{a}_1^\circ \mathbf{a}_2^\circ]^T \) is contained. See figure 4 for an illustration.

(ii) Based on the vertices of this triangle, the flow field associated to \( [\mathbf{a}_1^\circ \mathbf{a}_2^\circ]^T \) can then be estimated as a weighted average of the flow fields associated to each vertex. In the present work, these weights are chosen such that the query point \( [\mathbf{a}_1^\circ \mathbf{a}_2^\circ]^T \) is the barycentre of the corresponding triangle.

Although it may be memory-intensive, since numerous snapshots need to be stored, it will be shown in §4.2 that this local linear mapping procedure allows for an unprecedented accuracy when reconstructing the flow field. Further, it is possible to reduce the memory by compressing the snapshots. In such a case, the output of this local linear mapping are the coefficients of the POD modes needed to reconstruct the high-dimensional flow field.

3.3.2. Feature-based modal expansion

If full-state snapshots of the flow field are available, then it is possible to construct a set of generalized feature-based modes that make the sparse models physically interpretable, providing spatial structures associated with feature variables or specific interaction terms identified in the dynamics. These modes are defined as the spatial structures most correlated with either the feature vector \( \mathbf{a} \) or with specific nonlinear terms in the sparse model (3.4). They are obtained by stochastic estimation, a means of approximating a conditional average using a knowledge of unconditional statistics. As shown in Adrian & Moin (1988), a stochastic estimate of the conditional average can be obtained by means
of a Taylor series expansion
\[
\hat{u}(x,t) = \langle u(x,t)|a(t) \rangle \\
\approx a_i(t)u_{ij}(x) + a_i(t)a_j(t)u_{ij}(x) + \cdots
\tag{3.12}
\]
where \(a_i(t)\) is the \(i\)-th entry of the feature vector \(a(t)\), \(\langle \cdot \rangle\) is the expected value, and \(\hat{u}(x,t)\) is the estimate of the conditional average, in the present case the estimated velocity field. The first and second order feature-based modes \(u_{ij}(x)\) and \(u_{ij}(x)\), respectively, are determined by minimizing the mean square error of the estimate \((\langle \hat{u}(x,t) - u(x,t) \rangle^2)\).

In practice, these feature-based modes are obtained by solving a least-mean square problem. Consider a sequence of baseflow-subtracted full-state snapshots\(\dagger\)
\[Q(x) = [u^{(1)}(x)\ u^{(2)}(x)\ u^{(k_f)}(x)].\tag{3.13}\]
Note that these snapshots do not need to be collected at the same sampling period as the measurements \(s\) but only at a multiple of this sampling period. As such, they are collected at times \(t_{k_1}, \ldots, t_{k_p}\) that correspond to a subset of the resolved measurement times with which \(s\) and \(a\) are collected. Recall that the columns of \(A^T\) are a time-resolved sequence of feature vectors \(a(t_m)\), with \(m = 1, \ldots, M\). The columns of \(A^T\) associated with the non-time-resolved snapshots in \(Q\) are given by
\[A^T_Q = [a(t_{k_1})\ a(t_{k_2})\ \cdots\ a(t_{k_p})].\tag{3.14}\]
Considering only the first-order Taylor expansion of the conditional average, the snapshots sequence may thus be approximated by
\[Q = U(x)A^T_Q(t) + R(x,t),\tag{3.15}\]
where \(R\) is the truncation residual (i.e. the higher-order terms in the Taylor series expansion of the conditional average). The columns of \(U(x)\) are the first-order feature-based modes that are most correlated with the terms in the feature vector \(a(t)\). These modes are found via least-squares regression:
\[U = Q(A^T_Q)\dagger,\tag{3.16}\]
where \((A^T_Q)\dagger\) is the Moore-Penrose pseudo-inverse of \(A^T_Q\). More generally, it is possible to compute modes that are most correlated with the dynamic interaction terms in the sparse model (3.8). Let \(\gamma_1, \ldots, \gamma_q\) denote the indices of the rows in \(\Xi\) with non-zero entries, i.e. corresponding to active terms in the sparse dynamics. The corresponding terms in the dynamics may be extracted via:
\[\alpha = (\Theta(a)[e_{\gamma_1}\ e_{\gamma_2}\ \cdots\ e_{\gamma_q}])^T,\tag{3.17}\]
where \(e_{\gamma_j}\) is a column vector consisting entirely of zeros, except for a one in the \(\gamma_j\)-th row; i.e., \(e_{\gamma_j}\) is the \(\gamma_j\)-th column of the identity matrix. For example, we consider in §4.2 a vector of nonlinear terms \(\alpha\) given by \(\alpha = [a_1\ a_2\ a_1^2 + a_2^2\ 2a_1a_2\ a_1^2 - a_2^2]^T\). This choice corresponds to the different terms active in the dynamics of the low-order model identified in §4.2. The five components comprise the first and second harmonics as well

\(\dagger\) Note that the columns of the matrix \(Q\) correspond to a time-sequence of full-state snapshots, which is the mathematical convention in snapshot proper orthogonal decomposition and dynamic mode decomposition (Kutz et al. 2016). However, the matrix \(A\) uses transposed notation to be consistent with the original SINDy paper (Brunton et al. 2016b), with rows corresponding to a time-sequence of transposed feature vectors \(a^T\). Thus, the columns of \(A^T\) are a time-sequence of the features \(a\), similar to \(Q\).
as the baseflow drift. It is now possible to obtain generalized modes:

$$U = Q ([\alpha(t_{k_1}) \quad \alpha(t_{k_2}) \quad \cdots \quad \alpha(t_{k_p})])^\dagger.$$  

Thus, each mode $u_i(x)$ is a spatial field corresponding to a specific interaction term in the dynamical system, given by a component of $\alpha$.

Compared to the local linear mapping presented in §3.3.1, such a feature-based expansion has a low memory footprint, although it will typically be less accurate. However, even if the feature-based modes are not used for full-state reconstruction, they imbue the sparse model with physical interpretability. The modal representation above may be thought of as closely related to the proper orthogonal decomposition or dynamic mode decomposition, except generalized to identify modes that are most correlated with the features in $a$ or the dynamic interaction terms in $\alpha$.

4. Results

A detailed description of the application of the present methodology to the two-dimensional cylinder flow at $Re = 100$ is given in this section. As a starting point, our attention is focused in §4.1 on the identification of a low-dimensional descriptor system describing the dynamics of the flow solely based on the instantaneous lift and drag coefficients. As a second step, the performances of different estimators to reconstruct the full-state of the system are discussed in §4.2.

4.1. Sensor-based dynamics

In this section, a low-dimensional model of the transient and post-transient laminar cylinder wake is presented. First, a dynamical model capturing the dynamics of the lift coefficient is identified. Then, this low-order model is supplemented with a nonlinear algebraic measurement equation in order to infer the evolution of the drag coefficient. Note that prior to system identification, the lift signal is low-pass filtered as to simplify the identification process. Figure 5 depicts the power spectral density of the raw signal and its low-pass filtered counterpart. As can be seen, the low-pass filter suppresses the second low-amplitude harmonic. Nonetheless, plotting both signals in the time-domain shows no discernible differences (not shown). In practice, such a procedure is similar to the Krylov-Bogulobov assumption used by Tadmor et al. (2010) to simplify the low-order model derived from a Galerkin projection of the Navier-Stokes equations onto the span of the leading POD modes. The influence of this low-amplitude harmonic on the identified model is discussed in Appendix A.
It is well known that the two-dimensional cylinder flow behaves as a self-excited, self-limiting and nearly harmonic nonlinear oscillator. This behavior is clearly visible in the time evolution of the instantaneous lift coefficient depicted in figure 2. As such, the dynamics can be described by a nonlinear second-order ordinary differential equation, or by a set of two coupled first-order ODEs. The feature vector \( \mathbf{a}(t) \) characterizing the state of the system in the feature space is defined as

\[
\mathbf{a} = [\hat{a}_1 \quad \hat{a}_2]^T,
\]

with \( \hat{a} = \frac{a}{a_{\text{max}}} \), \( a_{\text{max}} \) being the maximum absolute value of \( a \) once the system evolves onto the periodic limit cycle. Such normalization ensures that

\[-1 \leq a_{1,2} \leq 1,
\]
a condition which greatly simplifies the sparse optimization problem involved in the identification procedure. Although this mapping function has been defined analytically in the present work, similar features could be identified using delay coordinates, as in the singular spectrum analysis (SSA) in meteorology and ecology (Colebrook 1978; Barnett & Hasselmann 1979; Weare & Nasstrom 1982; Ghil et al. 2002) or the eigensystem realization algorithm (ERA) from system identification and control theory (Juang & Pappa 1985). These delay coordinates have recently been shown to provide an embedding of nonlinear dynamics into a new coordinate system where the dynamics are approximately linear, establishing a connection to Koopman operator theory (Brunton et al. 2017; Arbabi & Mezić 2016). Alternatively, the Hilbert transform can also be used to embed the dynamics. In addition, a number of techniques in nonlinear time series analysis have been developed to determine the dimension of the feature vector \( \mathbf{a} \) needed to correctly reproduce the dynamics, e.g. the false nearest neighbors algorithm proposed by Kantz & Schreiber (2004).

4.1.1. Determining the structure of the model

As a starting point, let us first determine the optimal structure of the low-order model. Based on the transient evolution of the lift coefficient \( s^\ast(t) \) depicted in figure 2, the corresponding time series of the feature vector \( \mathbf{a}^\ast(t) \) has been computed. The Sparse Identification of Nonlinear Dynamics algorithm is used to identify the equations governing the dynamics of \( \mathbf{a} \). Four libraries of polynomial functions are considered, with the maximum polynomial degree ranging from \( d = 1 \) (i.e., linear) up to \( d = 7 \) (i.e., septic). For \( d = 3 \) (i.e., cubic), this pool of functions reads

\[
\Theta(a_1, a_2) = [1 \quad a_1 \quad a_2 \quad a_1^2 \quad a_1 a_2 \quad a_2^2 \quad a_1^3 \quad a_1^2 a_2 \quad a_1 a_2^2 \quad a_2^3].
\]

Such a pool of functions, which can easily be enriched if needed, is a natural choice for the identification of nonlinear oscillators.

Given our different polynomial libraries \( \Theta(a_1, a_2) \), the procedure proposed by Mangan et al. (2017) based on the Akaike Information Criterion (AIC) is used to select the model structure that best balances the tradeoff between accuracy and model complexity. Models with lower scores are more parsimonious, as in Mangan et al. (2017).

Figure 6 depicts the distribution of the different models identified in the complexity vs. AIC plane. Note that the relative \( \text{AIC}_c \) score depicted is given by \( \text{AIC}_c = \text{AIC} - \text{AIC}_{\text{min}} \). It should be emphasized that the model characterized by \( \text{AIC}_c = 0 \) is not necessarily the best model possible, but only the best one among the different models tested. In the present case, the model that best balances accuracy and complexity is given by a cubic
polynomial:
\[ \frac{da_1}{dt} = 1.12a_2 \]
\[ \frac{da_2}{dt} = -1.12a_1 + 0.2(1 - a_1^2 - a_2^2)a_2. \] (4.3)

Surprisingly, the last three terms involved in the equations governing the dynamics of \( a_2 \) are identical. Moreover, the first coefficient identified corresponds to the circular frequency of the natural vortex shedding.

Figure 7 depicts a comparison of the time trace of \( a_1 \) obtained from DNS and predicted by the low-order model identified for the training dataset (figure 7a and 7b) and the cross-validation dataset (figure 7c and 7d). Note that the latter has not been used in the identification process. Although it accurately captures the dynamics of the system close to the periodic orbit, the identified low-order model underestimates the growth rate of the fixed point’s linear instability. This underestimation leads to a delayed nonlinear saturation of the oscillation amplitude. This point will be investigated in the next section.

Finally, note that models including nonlinearities up to the 7\(^{th}\) order reproduce very accurately the testing and cross-validation time traces. These models may, however, be prone to overfitting, as highlighted by the AIC score.

4.1.2. Correcting the growth rate

As shown in the previous section, the cubic model identified underestimates the growth rate of the fixed point’s linear instability, despite having the lowest AIC score, balancing accuracy and complexity. Figure 8 depicts the evolution of the identified growth rate as a
function of the maximum polynomial degree of the model (blue dots) and compares it to the one obtained from a linear stability analysis of the corresponding high-dimensional linearized Navier-Stokes operator (dashed line). It appears that SINDy identifies the correct growth rate for libraries of polynomial functions with order $N \geq 7$. Despite the good agreement between the phase plane trajectories predicted by the cubic model and the DNS observations (see figure 7b and 7d), it appears that higher-order terms are necessary to accurately identify the instability growth rate, even though it may increase the risk of overfitting.

Although the need for high-order terms may appear counter-intuitive, this is relatively common in regression analysis with classical polynomials. To understand this, let us rewrite our regression model for $a_2$ as

$$\frac{da_2}{dt} = \Theta_1 \xi_1 + \Theta_2 \xi_2$$

(4.4)

where $\Theta_1$ contains the linear terms (i.e. $\Theta_1(a_1, a_2) = [a_1 \ a_2]$), $\Theta_2(a_1, a_2)$ the higher-order terms, while $\xi_1$ and $\xi_2$ are the corresponding vectors of coefficients. Neglecting the
sparsity constraint for the sake of simplicity, the least-squares problem becomes
\[
\begin{bmatrix}
\Theta_1^T \Theta_1 & \Theta_1^T \Theta_2 \\
\Theta_2^T \Theta_1 & \Theta_2^T \Theta_2
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2
\end{bmatrix}
= \begin{bmatrix}
\Theta_1^T \hat{a}_2 \\
\Theta_2^T \hat{a}_2
\end{bmatrix}. \tag{4.5}
\]
Thus, the coefficients \(\xi_1\) of the linear terms are given by
\[
\xi_1 = (\Theta_1^T \Theta_1)^{-1} \Theta_1^T (\hat{a}_2 - \Theta_2 \xi_2). \tag{4.6}
\]
Note that here \(\hat{a}_2\) is used to denote the time-derivative of \(a_2\) for the sake of clarity. The term \((\Theta_1^T \Theta_1)^{-1} \Theta_1^T \hat{a}_2\) is the least-squares solution when our pool of admissible functions is limited to \(\Theta_1(a_1, a_2)\). The second term, \(- (\Theta_1^T \Theta_1)^{-1} \Theta_1^T \Theta_2 \xi_2\), is the correction of \(\xi_1\) due to the inclusion of higher-order polynomials in the regression problem. This correction is proportional to the cross-correlation matrix \(\Theta_1^T \Theta_2\) and is related to the non-orthogonality properties of the set of functions forming our total library \(\Theta(a_1, a_2)\).

Note that that one needs not identify the complete high-order model in order to estimate the correct growth rate of the instability. Given the Frisch-Waugh-Lovell theorem, one needs only solve a partial regression problem for that purpose. Having estimated the rate of the instability, one can then enforce this particular value directly during the identification of the cubic model, as in Loiseau & Brunton (2016). The resulting optimization problem then reads
\[
\minimize_{\xi} \| \Theta(a_1, a_2) \xi - \hat{a}_2 \|^2_2 + \lambda \| \xi \|_1
\text{subject to } C \xi = d \tag{4.7}
\]
where \(C\) and \(d\) are the linear constraints enforcing the estimated growth rate of the instability. We then follow the same procedure as in the previous section to identify the optimal constrained cubic model. The resulting model reads
\[
\frac{d}{dt} \begin{bmatrix}
a_1 \\
a_2
\end{bmatrix} = \begin{bmatrix}
0 & 1.12 \\
-1.12 & 0.3(1 - a_1^2 - a_2^2)
\end{bmatrix} \begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}. \tag{4.8}
\]
As shown in figure 9, the time evolution of \(a_1^2(t)\) predicted by this low-dimensional model is in excellent agreement with DNS \((a_1^\bullet(t))\) on both the training and testing datasets. It is remarkable that, despite its apparent simplicity, this two-degrees-of-freedom model captures all of the key features of the cylinder flow, namely
- It admits only one linearly unstable fixed point given by \(a = 0\) and one attracting limit cycle characterized by \(\|a\| = 1\). The corresponding circular frequency \(\omega^\circ = 1.12\) in the nonlinearly saturated state is identical to the frequency observed in DNS \((\omega^\bullet = 1.12\)).
- It explicitly highlights the quadratic dependency of the instantaneous growth rate \(2\sigma(a) = 0.3(1 - a_1^2 - a_2^2)\). Such quadratic dependencies are consistent with our current understanding of the nonlinear saturation process of globally unstable flows; see Mantić-Lugo et al. (2014) for more details.
- Once the amplitude of the oscillation has saturated to \(\|a\| = 1\), the system reduces to a simple harmonic oscillator. A similar structure could be derived based on a Galerkin projection of the Navier-Stokes equations onto the span of the first two POD modes and using the marginally stable mean flow as the reference state.

From a physical point of view, this low-order system describes the dynamics of the original high-dimensional system when constrained to the low-dimensional manifold structuring its phase phase (Noack et al. 2003).

4.1.3. Predicting the drag evolution

The dynamical system identified in the previous section provides valuable information and accurately predicts the dynamics of its high-dimensional counterpart. Being based
solely on quantities derived from the instantaneous lift coefficient, the model does not enable direct estimation of the corresponding drag force. However, estimating the instantaneous drag force might be of critical importance in aerodynamic applications. Defining the drag coefficient as an additional degree of freedom of our system, two different possibilities arise:

(i) extend the identified system with a (nonlinear) algebraic measurement equation, also identified using sparse regression; or,

(ii) identify a new dynamical system comprised of three coupled ODEs.

We have investigated both of these options. Given the simplicity and accuracy of the dynamical system introduced previously, the former approach is preferred over the latter. Using the drag coefficient of the baseflow-subtracted velocity field as our additional degree of freedom $a_3$, and normalizing it such that its mean value is $\bar{a}_3 = 1$ on the limit cycle, the identified measurement equation reads

$$a_3^2 = 1.08a_1^2 + 0.92a_2^2 - 0.18a_1a_2.$$  

(4.9)

Note that it can be re-written as

$$a_3^2 = (a_1^2 + a_2^2) + 0.8(a_1^2 - a_2^2) - 0.18a_1a_2.$$  

(4.10)

The first term highlights that the amplitude of the drag force is directly proportional to the amplitude of the lift oscillation. The second and third terms then describe the contribution of the second harmonic of the lift on the drag, hence characterizing the instantaneous oscillation of the drag force around its mean value. Combining this measurement equation with the dynamical system (4.8) identified in §4.1.2 results in a low-dimensional descriptor system governing the evolution of both the instantaneous lift and drag coefficients.

Equation (4.10) describes a distorted cone and is notably different from the mean-field paraboloid of the 3-dimensional POD-Galerkin model by Noack \textit{et al.} (2003). This difference can be explained by the choice of the state space $a$. The oscillation of von Kármán vortex shedding is characterized by $a_1$ and $a_2$—albeit from different physical mechanisms. During the transient, the vortex shedding moves upstream from the stagnation point of the unstable steady solution to the immediate vicinity of the cylinder. The global POD mode amplitudes $a_1$, $a_2$ resolve the vortex shedding even when it is far downstream. However, the lift-based feature coordinates only respond to the vortex shedding in the final stage when the vortices are close to the cylinder. In contrast, the
J.-Ch. Loiseau, B. R. Noack and S. L. Brunton

Figure 9. (a) and (c) Comparison of the time-evolution of $a_1$ obtained from direct numerical simulation and predicted by the identified low-dimensional dynamical system. (b) and (d) Trajectories of the true system and the model on the phase plane $(a_1, a_2)$. For (a) and (b), the initial condition is close to the linearly unstable fixed point $u_b$, given by $a = 0$. (c) and (d) use the cross-validation dataset for which the initial condition used in the direct numerical simulation has been chosen to lie outside of the limit cycle and, for physical reasons, was also constrained to start close to the parabolic manifold structuring the phase space of the system.

Drag variation and the shift-mode amplitude of the 3-dimensional POD-Galerkin model are linearly related. Thus, $a_3$ of the Galerkin expansion and of the force-related ROM resolve the same physics. Hence, the fluctuation amplitude $\sqrt{a_1^2 + a_2^2}$ increases much faster with $a_3$ for the POD-based model than for the force-based system.

Figures 10(a) and 10(c) compare the predicted time-evolution of $a_3^\circ(t)$ against the evolution of $a_3^\bullet(t)$ obtained from direct numerical simulation, whereas figures 10(b) and 10(d) depict the associated trajectories projected onto the $(a_1, a_3)$ phase plane. Along with the dynamics of $a_1$ and $a_2$ being correctly captured by the dynamical system (4.8), the nonlinear algebraic measurement equation (4.10) correctly infers the evolution of $a_3$. It is clear from these results that the low-dimensional descriptor system identified in the present work is one of the simplest and yet most accurate and physically interpretable low-order models available in the literature to reproduce the dynamics of the cylinder flow at $Re = 100$.

4.2. Full-state estimation

The descriptor system identified in the previous section provides valuable information and accurately predicts the dynamics of our sensor. However, it does not allow us to directly infer the corresponding flow field. In order to reconstruct the flow field, we must
supplement our low-dimensional model with a full-state estimator

$$u(x, t) = h(a).$$

Formally, this full-state estimator $h$ is a nonlinear mapping from the low-dimensional feature-space to the high-dimensional physical space. In the rest of this section, four different estimators will be employed:

(i) a first estimator based on the Locally Linear Mapping (LLM) using the full-state snapshots from the training dataset to reconstruct those of the testing dataset,

(ii) a second estimator based on LLM outputting the estimated amplitudes of the first five POD modes,

(iii) a third estimator based on LLM outputting the estimated amplitudes of the first twenty-five POD modes,

(iv) a fourth estimator using the feature-based modal expansion described in §3.3.2.

Note that for the fourth estimator, we employ generalized feature modes. Given a sparse nonlinear model, this generalized feature-based modal decomposition is based on modal structures that are most correlated with specific terms in the dynamics. We define a new feature vector $\alpha$ containing all of the nonzero terms identified in the right hand side of

Figure 10. (a) and (c) Comparison of the time-evolution of $a_3$ (i.e. the drag coefficient) obtained from direct numerical simulation and predicted by the identified low-dimensional dynamical system. (b) and (d) Trajectories of the true system and the model on the phase plane $(a_1, a_3)$. For (a) and (b), the initial condition is close to the linearly unstable fixed point $u_b$, given by $a = 0$. (c) and (d) use the cross-validation dataset for which the initial condition used in the direct numerical simulation is the same as described in figure 9.


Figure 11. Figure (a) depicts the vorticity field of the linearly unstable base flow, while the dashed line highlights the spatial extent of the reversed flow region. Figures (b) to (e) show the vorticity field of the feature modes associated to the extended feature vector $\alpha$.

the sparse model:

$$\alpha \triangleq \begin{bmatrix} a_1 \\ a_2 \\ a_1^2 + a_2^2 \\ 2a_1a_2 \\ a_1^2 - a_2^2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix}. \quad (4.11)$$

The feature-based modal expansion considered hereafter then reads

$$u(x, t) = u_0(x) + \sum_{i=1}^{5} \alpha_i(t)u_i(x) + r(x, t) \quad (4.12)$$

where $u_0(x)$ is the linearly unstable steady solution to the Navier-Stokes equations and $r(x, t)$ is the residual. The different feature modes $u_i(x)$ have been computed following the stochastic estimation procedure described in §3.3. The associated vorticity fields are shown in figure 11, while the time-evolution of the different basis coefficients $\alpha_i(t)$ is depicted in figure 12(a). Figure 12(b) shows the cross correlation matrix of these signals. Given its diagonal structure, it is clear that the different basis coefficients $\alpha_i(t)$ are uncorrelated one to another. For the cylinder flow, these feature modes are very similar to the classical POD modes. A key advantage over POD modes is that the present modes are directly interpretable as being the coherent structures most correlated with our different measurements and the sparse nonlinear interaction terms in the model.

Figure 13 compares the evolution of the relative estimation error for the different estimators considered. Different sampling frequencies for the PIV-like snapshots forming the training dataset are used in figure 13(b) and (c). The flow estimator based on the local linear mapping (LLM) largely outperforms the other estimators. Its accurate performance, on average two to three orders of magnitude more accurate than the other estimators considered, results from the fact that LLM leverages all of the information contained in the different snapshots matrices used whereas the different modal expansions only provide low-rank approximations of these same matrices. In all cases, it has to be noted that increasing the sampling period $T_{s(PIV)}$ between two consecutive PIV-like snapshots when creating the training dataset hardly influences the performance of the different estimators. Moreover, it can be observed that the performance of the feature-based modal expansion is on a par with the LLM procedure outputting the amplitudes
of the first five POD modes. Finally, increasing the number of POD modes considered is only beneficial during the early stage of transition while leaving the accuracy of the POD-based estimator almost unchanged once the system has reached a statistically stationary state.

Figures 14(b) and (c) depict the estimated vorticity field, with the base flow subtracted, at different instants in time and compare them with the true vorticity field obtained from direct numerical simulation in figure 14(a). As expected, the vorticity fields estimated by the local linear mapping (with full-state PIV-like snapshots) are in much better agreement, from a physical and kinematic point of view, than the ones obtained by the other low-dimensional estimators. This is especially pronounced during the period of exponential growth of the linear instability and at the onset of nonlinear saturation during which the local linear mapping correctly captures the deformation and distortion of the flow structure. In contrast, low-rank modal expansions are notorious for their inability to capture such mode deformation and/or changes in operating conditions. Given that the POD modes and generalized feature modes used in this work essentially approximate the flow structure once the system has reached the periodic limit cycle, it is thus expected that they provide only a very crude estimation of the flow structures when the system evolves in the vicinity of the linearly unstable base flow. This inability of a modal expansion to easily address mode deformation is one of its key limitations and is the principal reason why the local linear mapping strategy using full-state PIV-like snapshots should be preferred. Combining the descriptor system identified in §4.1 with the local linear flow estimator finally allows us to construct a two-degrees-of-freedom reduced-order model of the cylinder flow, having an unprecedented accuracy.

5. Discussion

In this work, the instantaneous lift force measurement had been used as our single sensor. Here, the problem of sensor selection for more complex flow configurations is discussed in §5.1. Moreover, possible connections with the works of Tu et al. (2013) and Iñigo et al. (2014) are highlighted in §5.2.
5.1. Optimal sensor selection

It has been implicitly assumed so far that it is possible to choose a single sensor measurement whose time series contains sufficient information to fully characterize the dynamics of the flow under consideration. While using the instantaneous lift coefficient is a natural choice for the two-dimensional cylinder flow, it may not be as straightforward for more complex flow configurations to determine which quantity to monitor, nor where to place the required sensors as to obtain the most informative measurements. With the recent advances in machine learning and sparse sensing, a considerable body of work has
Sparse reduced-order modeling

Figure 14. Comparison of the vorticity fields at different instants of time obtained (a) from direct numerical simulation, (b) from the local linear mapping procedure and (c) from the feature-based modal expansion. Note that the vorticity field of the linearly unstable baseflow has been subtracted in order to highlight the perturbative vorticity field.

emerged in the literature about near-optimal sensor placement for estimation or reconstruction tasks of high-dimensional systems. One can cite for instance the recent work of Manohar et al. (2017) on data-driven sparse sensor placement. Given prior velocity field snapshots, it is possible to use a combination of POD/SVD and QR factorization with pivots of the collected snapshots matrix. The POD analysis allows the user to determine the rank of the snapshot matrix (i.e. the dimension of the low-dimensional model to be considered), while the pivots of the QR factorization then provide the spatial location of the most informative sensors with which to optimally reconstruct the high-dimensional state of the system (Drmac & Gugercin 2016; Manohar et al. 2017). Figure 15 depicts the locations of the first five near-optimal streamwise velocity sensors identified by this approach for the present flow configuration. Three of these sensors are located along the \( y = 0 \) line, monitoring the mean flow distortion. The remaining two sensors are located approximately at \( y \simeq \pm 0.5 \) along a diagonal, capturing the amplitude and phase of the vortex shedding. These velocity measurements may then form the initial measurement vector \( s \) required for the low-dimensional modeling methodology proposed herein. For more details details about data-driven sensor placement, interested readers are referred to Manohar et al. (2017) and numerous references therein.

5.2. Comparisons with previous works

The low-dimensional modeling strategy proposed herein shares some connections with the recent works of Tu et al. (2013) and Iiigo et al. (2014). All three approaches are summarized as block-diagrams in figure 16. For the sake of simplicity, the block-diagrams depicted in figure 16 all use the present notation where:

- \( s \) denotes the input sensor measurements vector,
- \( a \) denotes the state vector of the low-dimensional model,
- \( y \) denotes the output vector.

In Iiigo et al. (2014) and Tu et al. (2013), this output \( y \) corresponds to the coefficients...
of POD modes used to reconstruct the velocity snapshots. In the present case, $y$ can denote the coefficients of the POD modes, those of the feature-based modes, or directly the reconstructed high-dimensional snapshot, depending on the particular choice made by the user. All three approaches thus have the same input and same output. They essentially differ on how one maps the input vector $s$ to the output vector $y$. Iñigo et al. (2014) use the sensor measurement vector $s$ as the input to a Multiple-Input-Multiple-Output (MIMO) Linear Time Invariant (LTI) dynamical system obtained using linear subspace identification techniques. The vector $y$ of POD coefficients used to reconstruct the high-dimensional velocity snapshot then comes as the output of this MIMO LTI system. Tu et al. (2013) first define a mapping $g$ from the sensor measurement vector $s$ to the state $a$ of the low-dimensional system using a multiple-time-delay modified linear stochastic estimator (MTD-mLSE). Next, a discrete-time LTI system for $a$ is identified using a combination of linear stochastic estimation and prior knowledge obtained from DMD analysis. The state vector $a$ of this low-dimensional system is chosen as the amplitude of the leading POD modes, and the output vector $y$ is simply $a$ itself.

Similar to Tu et al. (2013), we first define a function $g$ mapping the sensor measurement vector $s$ to the state $a$ of the low-dimensional model to be identified. In its present formulation, this mapping has been defined analytically, although one could have formally used the same technique as in Tu et al. (2013), i.e. defining a mapping from time-delayed lift measurements to the leading POD coefficients. As a second step, a (possibly) nonlinear continuous-time model governing the dynamics of the state $a$ is identified using SINDy. Finally, the predicted state $a$ is mapped to the output $y$. Depending on the choice of the user, this output $y$ can denote the coefficients of a modal expansion of the velocity field (e.g. POD modes or feature-based modes) as in Iñigo et al. (2014) and Tu et al. (2013), or directly the reconstructed high-dimensional velocity snapshot obtained from the Locally Linear Mapping procedure.

It is clear that the approach described herein and the one proposed by Tu et al. (2013) are closely related. The key difference is that Tu et al. (2013) obtain a low-dimensional model using linear stochastic estimation and prior knowledge from DMD analysis (thus assuming that the state $a$ has linear dynamics), while our low-dimensional model has been obtained using sparse nonlinear system identification. In the limit of truly linear dynamics and for the same definition of the mapping $a = g(s)$, these two approaches should thus simply be discrete-/continuous-time counterpart of each other. The same observation holds once the system reaches its limit cycle as the model identified in the present work reduces in this case to a simple harmonic oscillator. The connection with Iñigo et al. (2014) is not as straightforward, although the two approaches should yield similar results if the dynamics of the state $a$ are indeed linear.
6. Conclusions

This work develops a new reduced-order modeling procedure for unsteady fluid flows that yields accurate nonlinear models and insight into relevant flow structures. This procedure identifies sparse nonlinear models, not on the full fluid state, but from time-resolved sensor measurements that may be realistically obtained in experiments. The sparsity of the model prevents overfitting and uncovers key nonlinear interaction terms. Although models are data-driven, they are interpretable, and it is also possible to incorporate partial prior knowledge of the physics or constraints to improve the models. If snapshots are also available, for example from non-time-resolved PIV, it is possible to estimate the full-state from the sparse model using local linear mapping: the full-state is interpolated between the most similar historical flow fields, based on the dynamics. It is also possible to construct a generalized modal decomposition that identifies coherent structures most correlated with each interaction term in the sparse nonlinear model.

This methodology is illustrated using the canonical two-dimensional cylinder flow at $Re = 100$. Despite its simplicity, this flow configuration is a prototypical example capturing the key physics of bluff body flows. Even though this study uses data from direct numerical simulations, the overall strategy is generally applicable to a real flow experiment with minor modifications. Despite their simplicity, the identified models do not suffer the same drawbacks as reduced-order models obtained from a Galerkin projection procedure, namely over-estimation of the duration of transients and energy overshoots at the onset of nonlinear saturation. Instead, the identified sparse models provide simple explanations for the nonlinear saturation process of globally unstable flows. Moreover, the models are based on sensor measurements, which may include lift, drag, or pressure measurements that are physically linked to the geometry. Working in these intrinsic coordinates has the potential to overcome many of the limitations of classical modal-based projection methods, including mode deformation due to moving geometry and varying parameters.

The effectiveness of the reduced-order models identified and the modularity of the
methodology proposed in the present work suggest a number of exciting future directions. There is significant potential for these methods to be applied broadly to obtain interpretable reduced-order models for a range of flow configurations in simulations and experiments. For example, these sensor-based models may be applied to develop nonlinear unsteady aerodynamic models, generalizing previous linear and linear parameter varying models (Brunton et al. 2013, 2014; Hemati et al. 2016).

A key motivation in this work is its extension to flow control. Given a feature vector $\mathbf{a}$ and actuators characterized by a control law $\mathbf{b}(t)$, one could use SINDy with control (SINDYc) (Brunton et al. 2016c; Kaiser et al. 2017) in order to identify low-order models

$$\frac{d\mathbf{a}}{dt} = f(\mathbf{a}, \mathbf{b})$$

that incorporate the influence of the actuation $\mathbf{b}$ on the dynamics of $\mathbf{a}$. Combining such an approach with Machine Learning Control (Duriez et al. 2016) may result in interpretable models of entirely new flow behaviors and previously unobserved flow physics that are discovered in the controlled flow. The identified models can then serve as a low-dimensional representation of the actual system in order to facilitate the computation of nonlinear optimal feedback control laws. This is an area of active research by the authors. In the near future, the authors aim to apply the methodology introduced in the present work to the optimal control of experimental flows.

Finally, there are a number of methodological extensions that may improve the performance of this sparse modeling framework. First, it will be important to demonstrate that these methods scale favorably to systems with higher-dimensional attractors. Because the algorithms are based on simple regression and sparse optimization, they should remain computationally tractable. Next, it may be possible to reduce the memory requirements of the local linear mapping by building local modal libraries in different dynamic regimes (e.g., linear instability, saturated limit cycle, etc.). The storage requirements may further be reduced using compression techniques and sparse sampling. It may also be possible to incorporate the accuracy of the generalized modal decomposition reconstruction into the cost function in the SINDy regression, so that nonlinear features are selected based on their dynamic relevance and their ability to reconstruct the full state.

Acknowledgments

SLB acknowledges generous funding support from the Defense Advanced Research Projects Agency (DARPA HR0011-16-C-0016) and from the Air Force Office of Scientific Research (AFOSR FA9550-18-1-0200). SLB would like to thank Nathan Kutz, Josh Proctor, Eurika Kaiser, Niall Mangan, and Sam Rudy for discussions related to sparse model identification. SLB would also like to thank Scott Dawson for valuable discussions related to nonlinear modeling in terms of aerodynamic force coefficients.

BRN acknowledges funding from the German Science Foundation (DFG grant SE 2504/2-1). This work is also supported by internal funds of LIMSI-CNRS and a public grant overseen by the French National Research Agency (ANR) as part of the “Investissement dAvenir” program, through the “iCODE Institute project” funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

Appendix A. Influence of the low-amplitude harmonics

As explained in §4.1, the lift signal recorded from DNS has been low-pass filtered prior to the identification of the low-dimensional model as to remove the low-amplitude
Sparse reduced-order modeling

second harmonics component of the signal. Although not mandatory, this preprocessing step enables the identification of a simpler model without altering its predictive and explanatory capabilities. In this appendix, the same procedure as described in §4.1 is used directly on the raw unfiltered lift signal. Including the correction of the growth rate, the identified cubic model reads

\[
\frac{da_1}{dt} = 1.12a_2
\]

\[
\frac{da_2}{dt} = -1.12a_1 + 0.3a_2 - 0.3a_2^3 - 0.34a_1^2a_2. \tag{A 1}
\]

This model can be rewritten as

\[
\frac{da_1}{dt} = 1.12a_2
\]

\[
\frac{da_2}{dt} = -1.12a_1 + 0.3(1 - a_1^2 - a_2^2)a_2 - 0.04a_1^2a_2. \tag{A 2}
\]

Comparing model (A 2) and its low-pass filtered counterpart (4.8) presented in §4.1, it is clear that keeping the low-amplitude second harmonics in the signal used for the identification gives rise to a correction term which is an order of magnitude smaller than the other terms involved. Simulating both models gives almost identical results. Following Ockham’s razor, the model identified in §4.1 is thus preferred over the one presented in this appendix due to its simpler formulation.

REFERENCES


Graham, W. R., Peraire, J. & Tang, K. Y. 1999 Optimal control of vortex shedding using...
Sparse reduced-order modeling


