Metric for attractor overlap

Article · December 2017

4 authors:

Rishabh Ishar
Helmholtz-Zentrum Dresden-Rossendorf
4 PUBLICATIONS   **20** CITATIONS

SEE PROFILE

Eurika Kaiser
University of Washington Seattle
47 PUBLICATIONS   **841** CITATIONS

SEE PROFILE

Marek Morzynski
Poznan University of Technology
115 PUBLICATIONS   **2,337** CITATIONS

SEE PROFILE

Bernd R. Noack
303 PUBLICATIONS   **7,524** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Modelling coherent structures in free and wall-bounded shear flows View project

VIRDIAMED View project

Metric for attractor overlap

Rishabh Ishar[1]†, **Eurika Kaiser**[2], **Marek Morzynski**[3] **and Bernd R. Noack**[4,5,6]

[1]Department of Mechanical Engineering, PEC University of Technology, Chandigarh 160012, India

[2]University of Washington, Department of Mechanical Engineering, Stevens Way, Box 352600, Seattle, WA 98195, United States

[3]Poznań University of Technology, Chair of Virtual Engineering, Jana Pawla II 24, PL 60-965 Poznań, Poland

[4]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 Cedex, France

[5]Institute for Turbulence-Noise-Vibration Interaction and Control, Harbin Institute of Technology, Shenzhen Campus, Shenzhen, Peoples Republic of China

[6]Institut für Strömungsmechanik und Technische Akustik (ISTA), Technische Universität Berlin, Müller-Breslau-Straße 8, D-10623 Berlin, Germany

We present the first general metric for attractor overlap (MAO) facilitating an unsupervised data comparison. The starting point is two or more attractors, i.e. ensembles of states representing different operating conditions. The proposed metric quantifies the overlap between these attractors in three steps. First, all states are clustered into discrete centroid-based bins. Second, the probability distribution of each operating condition is determined. Third, the pairwise similarity of attractors is assessed with a cluster-based pseudometric for the probability distributions. MAO is augmented by proximity maps for the snapshots, the clusters, and the attractors. We demonstrate MAO for seven different control laws applied to the fluidic pinball, i.e. the flow around three circular cylinders whose centers form an equilateral triangle pointing in the upstream direction. These seven different operating conditions comprise unforced shedding, boat tailing, base bleed, high- and low-frequency forcing as well as two opposing Magnus effects. MAO compares and classifies these actuated flows in agreement with physical intuition. Moreover, the overlap and difference of the attractors are identified via shared and unshared clusters. MAO has a large spectrum of potential applications ranging from a quantitative comparison between numerical simulations and experimental particle image velocimetry data to the analysis of simulations representing a myriad of different operating conditions.

1. Introduction

In this study, we propose arguably the first general metric between attractor data from different sources or from different operating conditions. With *attractor data*, we refer to an ensemble of statistically representative flow snapshots which allow for the computation of statistical moments and resolve coarse-grained coherent structures. More generally, we propose an unsupervised comparison methodology with little subjective bias. This methodology is exemplified for the *fluidic pinball*, i.e. the flow around three stationary rotating circular cylinders (Noack *et al.* 2016). Contrary to the single rotating cylinder,

† Email address for correspondence: rishabhishar.bemech14@pec.edu.in

the fluidic pinball exhibits rich spatial-temporal dynamics under different actuation laws at similarly low computational cost (Noack & Morzyński 2017).

Most fluid mechanics publications contain a comparison of flows from different sources, e.g. experiments versus simulations or from different operating conditions, e.g. optimal control versus the unforced benchmark. In an engineering application, this comparison is easily performed for the global order parameter of interest, e.g. drag for a car, lift for an airfoil, mixing for a combustor or far-field noise of a jet engine. These comparisons are simple at the cost of a limited assessment of the flow physics. For instance, Reynolds-averaged Navier-Stokes (RANS) simulations of cylinder wakes may predict well the drag coefficient. Yet, RANS predicts the von Kármán vortex street to dissipate far too quickly in downstream direction. A commonly used more refined comparison includes the statistical moments of the flow field, or at least transverse or streamwise velocity profiles. The comparison of statistical moments is straightforward using a corresponding Hilbert space norm. This comparison is more detailed than employing a single order parameter. Yet, it excludes spatial-temporal dynamics of coherent structures.

Coherent structures are often visible to the naked eye, as beautifully depicted for wakes over 500 years ago by Leonardo da Vinci. Their quantification has been subject of thousands of publications and many disputes. Vortices provide important dynamical insights into geometrically simple two- and three-dimensional flows and have been the cornerstone of early reduced-order modeling efforts starting with the famous Helmholtz vortex laws in 1869 (Lugt 1996). Data-driven vortex identifications have been proposed by Jeong & Hussain (1995) for snapshots and by Haller (2005) for flow histories. These frequently cited publications represent — pars pro toto — myriad of other feature analyses, e.g. Galilean invariant snapshot topology (Kasten *et al.* 2014).

An alternative approach is a reduced-order representation by expansions in terms of global modes. Proper orthogonal decomposition (Berkooz *et al.* 1993), dynamic mode decomposition (Schmid 2010; Rowley *et al.* 2009), stability eigenmodes (Theofilis 2011) or plain temporal Fourier expansions may serve as examples. These modes provide important insights into physical mechanisms of the coherent structure dynamics. Yet, one cannot expect the modes of different sources to coincide or even to be similar. In addition, the residual of low-order flow representations may easily exceed the resolved component for turbulence.

Both, the analysis of local features and expansions in terms of global modes pose significant challenges for an automated comparison of different attractor data. In this study, we follow the pioneering clustering approach of Burkardt *et al.* (2004) in fluid dynamics. The affiliation of snapshots into a given number of similarly sized clusters with representative centroids enables a straightforward comparison of different attractor data. Now, each attractor is characterized as frequency of snapshots in these clusters. The attractor overlap or similarity can be quantified using the centroid configuration and frequency distributions. The comparison methodology is augmented with powerful feature extraction from machine learning. The beauty and few biases of this approach will be elucidated in the following section.

As demonstrated for the proposed framework, we chose a direct numerical simulation of the fluidic pinball, i.e. the flow around three equal circular cylinders with centers on an equilateral triangle pointing in upstream direction (Noack *et al.* 2016). The term 'fluidic pinball' is owed to the possibility of moving fluid particles like balls in a conventional pinball machine by suitably rotating the cylinders. The pinball configuration includes most wake stabilization strategies with suitable rotation of the three cylinders. Examples include phasor control (Roussopoulos 1993), aerodynamic boat tailing (Geropp 1995; Geropp & Odenthal 2000; Barros *et al.* 2016), base bleed (Wood 1964; Bearman 1967),
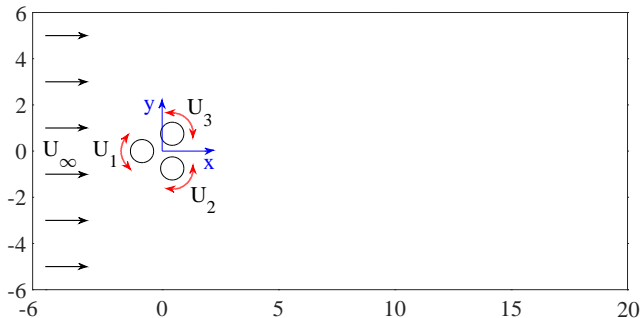
FIGURE 1. Configuration of the fluidic pinball.

high-frequency forcing (Thiria *et al.* 2006; Oxlade *et al.* 2015) and low-frequency forcing (Pastoor *et al.* 2008). Another possibility is to deflect the wake via the Magnus effect. In this study, we compare the unforced reference and six open-loop actuation mechanisms.

The manuscript is outlined as follows: In §2, the employed fluidic pinball simulation is described. The corresponding data includes converged data of the seven attractors. In §3, the comparison methodology for attractor data is outlined. The proposed approach is exemplified for all operating conditions of the fluidic pinball in §4. Section §5 summarizes this study and outlines future directions of research.

2. Fluidic pinball simulation

In this section, the computation of the employed flow data is described. In §2.1, the configuration of the fluidic pinball is introduced. The corresponding direct Navier-Stokes solver is described in §2.2. Finally, the fluidic pinball simulation of seven different open-loop actuations is detailed. These flow data are subjected to the proposed comparison methodology of the next section.

2.1. *Configuration*

In the following, the considered two-dimensional flow control configuration is described. Three equal circular cylinders with radius R are placed parallel to each other in a viscous incompressible uniform flow at speed U_∞ (see Fig. 1). The centers of the cylinders form an equilateral triangle with side length $3R$, symmetrically positioned with respect to the flow. The leftmost triangle vertex points upstream, while the rightmost side is orthogonal to the oncoming flow. Thus, the transverse extent of the three-cylinder configuration is given by $5R$.

This flow is described in a Cartesian coordinate system where the x-axis points in the direction of the flow, the z-axis is aligned with the cylinder axes, and the y-axis is orthogonal to both. The origin $\mathbf{0}$ of this coordinate system coincides with the geometric center of the cylinder triangle. The location is denoted by $\boldsymbol{x} = (x, y, z) = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z$, where $\boldsymbol{e}_{x,y,z}$ are unit vectors pointing in the direction of the corresponding axes. Analogously, the velocity reads $\boldsymbol{u} = (u, v, w) = u\boldsymbol{e}_x + v\boldsymbol{e}_y + w\boldsymbol{e}_z$. The pressure is denoted by p and the time by t. In the following, we assume a two-dimensional flow, i.e. no dependency of any flow quantity on z and vanishing spanwise velocity $w \equiv 0$.

The Newtonian fluid is characterized by a constant density ρ and kinematic viscosity ν. In the following, all quantities are assumed to be non-dimensionalized with cylinder diameter $D = 2R$, the velocity U_∞ and the fluid density ρ. The corresponding Reynolds

number is $Re_D = 100$ where $Re_D = U_\infty D/\nu$. The Reynolds number based on the transverse length $L = 5D$ is 2.5 times larger. The non-dimensionalization with respect to the diameter is more common for clusters of cylinders (Hu & Zhou 2008a,b; Bansal & Yarusevych 2017) and will be adopted in the following. With this non-dimensionalization, the cylinder axes are located at

$$\begin{array}{llll} x_F & = -\sqrt{3}/2, & y_F & = 0, \\ x_B & = \sqrt{3}/4, & y_B & = -3/4, \\ x_T & = \sqrt{3}/4, & y_T & = +3/4. \end{array} \tag{2.1}$$

Here, and in the following, the subscripts 'F', 'B' and 'T' refer to the front, bottom and top cylinder. An alternate reference is the subscripts 1, 2, 3 for the front, bottom, and top cylinder, respectively. The numbering is in mathematically positive orientation. Note that the choice of the origin and the non-dimensionalization with respect to the cylinder diameter give rise to a simple description of the cylinder centers.

The incompressibility condition reads

$$\nabla \cdot \boldsymbol{u} = 0, \tag{2.2}$$

where ∇ represents the Nabla operator. The evolution is described by the Navier-Stokes equations,

$$\partial_t \boldsymbol{u} + \nabla \cdot \boldsymbol{u} \otimes \boldsymbol{u} = -\nabla p + \frac{1}{Re_D} \triangle \boldsymbol{u}, \tag{2.3}$$

where ∂_t and $\triangle$ denote the partial derivative with respect to time t and the Laplace operator, respectively. The dot '·' and dyadic product sign '$\otimes$' refer to inner and outer tensor products.

Without forcing, the boundary conditions comprise a no-slip condition on the cylinder and a free-stream condition in the far field:

$$\boldsymbol{u} = \boldsymbol{0} \text{ on the cylinder and } \boldsymbol{u} = \boldsymbol{e}_x \text{ at infinity.} \tag{2.4}$$

A typical initial condition is the unstable steady Navier-Stokes solution $\boldsymbol{u}_s(\boldsymbol{x})$.

The forcing is exerted by rotation of the cylinders with circumferential velocities $b_1 = U_1 = U_F$, $b_2 = U_2 = U_B$ and $b_3 = U_3 = U_T$ for the front, bottom and top cylinder, respectively. The actuation command $\boldsymbol{b} = (b_1, b_2, b_3) = (U_1, U_2, U_3)$ is preferably used for control theory purposes (Brunton & Noack 2015; Duriez et al. 2016) while (U_F, U_B, U_T) are more natural for physical discussions. The actuation is conveniently expressed with the vector cross product '$\times$':

$$\boldsymbol{u} = 2U_i\, \boldsymbol{x} \times \boldsymbol{e}_z \text{ on the } i\text{th cylinder.} \tag{2.5}$$

The factor 2 counterbalances the non-dimensional radius $1/2$.

We like to refer to this configuration as *fluidic pinball* as the rotation speeds allow one to change the paths of the incoming fluid particles like flippers manipulate the ball of a conventional pinball machine. The front cylinder rotation may determine if the fluid particle passes by on the upper or lower side of the cylinder, while the top and bottom cylinder may guide the particle through the interior.

2.2. Direct Navier-Stokes solver

The chosen computational domain is bounded by the rectangle $[-6, 20] \times [-6, 6]$ and excludes the interior of the cylinders:

$$\Omega = \left\{ (x, y) \colon\, -6 \leqslant x \leqslant 20 \wedge |y| \leqslant 6 \wedge (x - x_i)^2 + (y - y_i)^2 \geqslant 1/4, i = 1, 2, 3 \right\}.$$
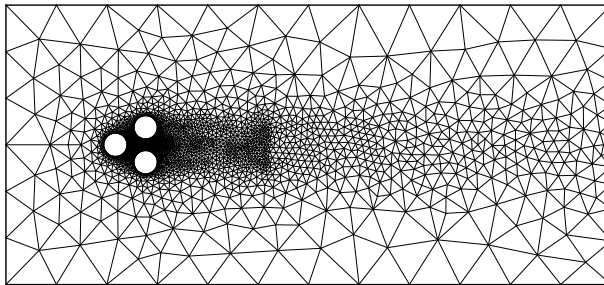
Figure 2. Grid of the fluidic pinball simulation.

This flow domain is discretized on an unstructured grid with 4225 triangles and 8633 vertices (see Fig. 2). This discretization optimizes the speed of the numerical simulation while keeping the accuracy at an acceptable level. Increasing the number of triangles by a factor 4 yields virtually indistinguishable results. The Navier-Stokes equation is numerically integrated with an implicit Finite-Element Method (Noack *et al.* 2003, 2016). The numerical integration is second-order accurate in space and third-order accurate in time.

This direct numerical simulation has a companion experiment at turbulent Reynolds numbers in the group of Prof. Martinuzzi at University of Calgary (Raibaudo *et al.* 2017).

2.3. *Attractor data*

We simulate seven different actuations in a single simulation starting at $t = 0$ with the unstable steady Navier-Stokes solution. Each phase is associated with one control law and lasts 100 convective time units. Figure 3 provides a preview of the simulation which is detailed in the following.

In the first phase, vortex shedding is kick-started with a motion of the front cylinder,

$$U_F = \begin{cases} 1/2 & t \leqslant 6.25 \\ -1/2 & 6.25 < t \leqslant 12.5 \\ 0 & \text{otherwise} \end{cases} \tag{2.6a}$$

$$U_B = -U_T = 0. \tag{2.6b}$$

The imposed period 12.5 corresponds to a Strouhal number of 0.2 based on the transverse width $5R$.

In the second phase, strong boat tailing with symmetric cylinder rotation of the upper and lower cylinder pushes the separation close to the x-axis and completely suppresses vortex shedding,

$$U_F = 0, \quad U_B = -U_T = 4. \tag{2.7}$$

In the third phase, base-bleed is enforced with opposite cylinder motion:

$$U_F = 0, \quad U_B = -U_T = -2. \tag{2.8}$$

This actuation widens the wake and pushes vortex formation downstream.

In the fourth phase, symmetric high-amplitude, high-frequency actuation energizes both shear layers:

$$U_F = U_B = U_T = 2 \cos \left(10\pi t / 12.5 \right). \tag{2.9}$$

The frequency corresponds roughly to five times the one of natural vortex shedding, following Thiria *et al.* (2006) for a single cylinder.
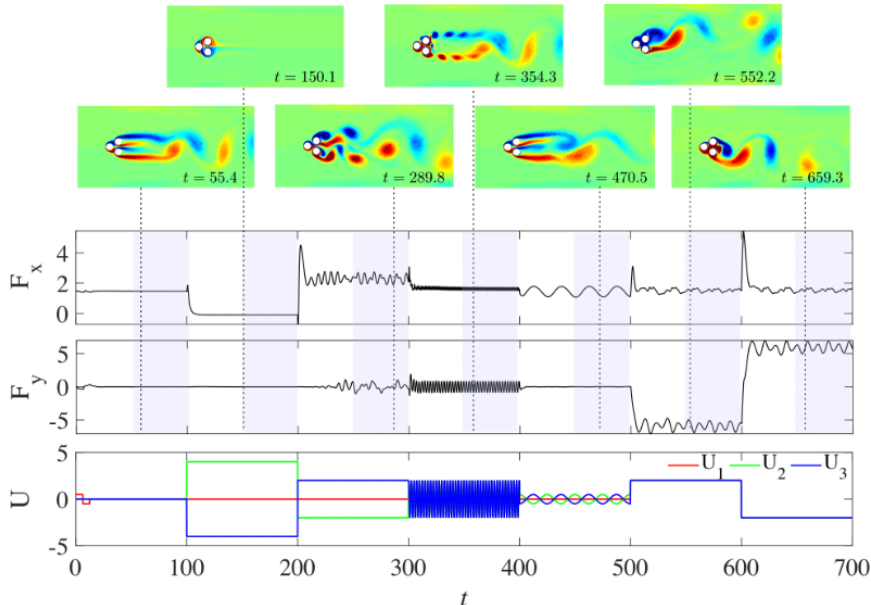
FIGURE 3. Operating conditions of the fluidic pinball simulation. Bottom: time series of the actuation commands for each cylinder over all phases together. Top: the flow of each phase is illustrated with a vorticity snapshot corresponding to maximum lift or minimum drag of the converged phase. Middle: the time series of the total drag F_x and total lift F_y are displayed.

In the fifth phase, symmetric low-amplitude, low-frequency forcing delays vortex shedding:

$$U_F = 0, \quad U_B = -U_T = \frac{1}{2} \cos\left(\pi t/12.5\right). \tag{2.10}$$

Low amplitudes were found to be most effective in stabilizing the wake in a parametric study at LIMSI (Rolland 2017).

In the sixth phase, a uniform rotation of all cylinders deflects the wake upwards via the Magnus effect,

$$U_F = U_B = U_T = 2. \tag{2.11}$$

In the seventh and last phase, the opposite Magnus effect is imposed,

$$U_F = U_B = U_T = -2. \tag{2.12}$$

Table 1 summarizes all control laws for later reference †.

The attractor data contains time-resolved snapshots from the latest 50 convective units of each phase. These snapshots are equidistantly sampled with a time step of 0.1, i.e. each phase is represented by 500 velocity fields. The first 50 convective time units correspond to 2.5 downwash times — enough for the transient dynamics to die out.

3. Comparison methodology for different attractors

In this section, we propose a cluster-based comparison methodology for attractor data. The first constitutive element is a standard metric for snapshots as described in §3.1. The second step and key technical enabler is the clustering of snapshot data based on this metric. The cluster analysis is described in §3.2. In the third step, the attractor is

† A visualization of the whole simulation can be found on `http://fluidicpinball.com`.

Phase	Time	Actuation	Control law
I	$t \leqslant 100$	unforced reference	$U_F = U_B = U_T = 0$ for $t > 12.5$, see (2.6)
II	$t \in (100, 200]$	boat tailing	$U_F = 0, U_B = -U_T = 4$
III	$t \in (200, 300]$	base bleed	$U_F = 0, U_B = -U_T = -2$
IV	$t \in (300, 400]$	high-frequency forcing	$U_F = U_B = U_T = 2\cos(10\pi t/12.5)$
V	$t \in (400, 500]$	low-frequency forcing	$U_F = 0, U_B = -U_T = 1/2\cos(\pi t/12.5)$
VI	$t \in (500, 600]$	positive Magnus effect	$U_F = U_B = U_T = 2$
VII	$t \in (600, 700]$	negative Magnus effect	$U_F = U_B = U_T = -2$

TABLE 1. Specifications of the different control phases of the fluidic pinball simulation.
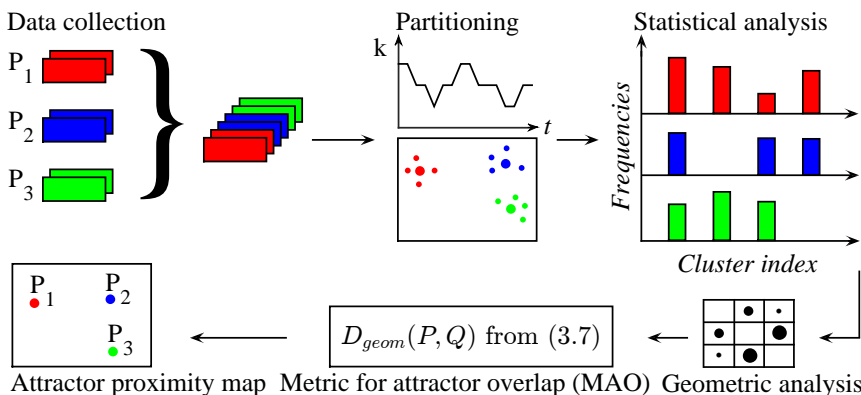

FIGURE 4. Schematic of the pipeline for the metric of attractor overlap and associated proximity maps.

represented by the frequency distribution of the populated clusters (§3.3). In the fourth and final step (§3.4), the metric of attractor overlap between two data ensembles is geometrically motivated and expressed in terms the respective probability distributions and the centroid distance matrix. In §3.5, proximity maps are introduced to visualize all snapshots, all cluster centroids, and all attractors. Figure 4 provides a schematic of the proposed methodology.

3.1. *Distance between snapshots*

Let $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$ be two incompressible velocity fields in the domain Ω. We define the distance between these fields as

$$D(\boldsymbol{u}, \boldsymbol{v}) = \int_{\Omega} d\boldsymbol{x} \, \|\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{v}(\boldsymbol{x})\|^2. \tag{3.1}$$

Here, $\|\cdot\|$ denotes the Euclidean norm. Note that this distance is based on the norm associated with the Hilbert space $\mathcal{L}^2(\Omega)$ of square-integrable functions. It fulfills all properties of a metric, like positive definiteness, commutativity, scaling with a real factor and triangle inequality. For the fluidic pinball data, the distance measure uses the whole computational domain Ω.

3.2. Cluster analysis

Next, the $M = 3500$ snapshots $\boldsymbol{u}^m(\boldsymbol{x})$, $m = 1, \ldots, M$, comprising all seven fluidic pinball phases are coarse-grained into $K = 50$ representative centroids $\boldsymbol{c}_k(\boldsymbol{x})$, $k = 1, \ldots, K$.

These centroids are chosen to minimize the total variance of the snapshots $\boldsymbol{u}^m$ with respect to the nearest centroid $\boldsymbol{c}_k$:

$$V = \sum_{k=1}^{K} \sum_{\boldsymbol{u}^m \in \mathcal{C}_k} D^2\left(\boldsymbol{u}^m, \boldsymbol{c}_k\right) \overset{!}{=} \min. \tag{3.2}$$

Each centroid $\boldsymbol{c}_k$ defines a cluster $\mathcal{C}_k$ containing all flow states $\boldsymbol{u}^m$ which are closer to $\boldsymbol{c}_k$ as compared to any other centroid $\boldsymbol{c}_j$, $j \neq k$. Thus, each snapshot $\boldsymbol{u}^m$ can be attributed to one cluster $\mathcal{C}_k$. This cluster affiliation is coded as characteristic function

$$T_k^m := \begin{cases} 1, & \text{if } \boldsymbol{u}^m \in \mathcal{C}_k, \\ 0, & \text{otherwise.} \end{cases} \tag{3.3}$$

The number of snapshots in the kth cluster reads

$$n_k = \sum_{m=1}^{M} T_k^m. \tag{3.4}$$

The centroids can also be expressed in terms of this characteristic function. It can be shown that they are the mean of all snapshots in the corresponding cluster,

$$\boldsymbol{c}_k = \frac{1}{n_k} \sum_{\boldsymbol{u}^m \in \mathcal{C}_k} \boldsymbol{u}^m = \frac{1}{n_k} \sum_{m=1}^{M} T_k^m \boldsymbol{u}^m. \tag{3.5}$$

Numerically, the optimization problem (3.2) for the centroids is solved using k-means clustering (Steinhaus 1956; MacQueen 1967; Loyd 1982) and k-means++ (Arthur & Vassilvitskii 2007) for the initialization. As k-means shows a dependency on the initial conditions, we run the corresponding MATLAB routine for 1000 initial conditions and select the one having the smallest variance. The iteration stops when convergence is reached, i.e. the characteristic function (3.3) does not change. The number of iterations is limited to $10,000$ iterations.

For practical reasons, we perform a lossless POD decomposition prior to the clustering of M snapshots into the mean flow $\boldsymbol{u}_0$ and $M - 1$ modes $\boldsymbol{u}_i$, $i = 1, \ldots, M - 1$ (see §A). Let $\boldsymbol{u} = \boldsymbol{u}_0 + \sum_{m=1}^{M-1} a_i \boldsymbol{u}_i$ be one snapshot and $\boldsymbol{v} = \boldsymbol{u}_0 + \sum_{m=1}^{M-1} b_i \boldsymbol{u}_i$ another one. Then, $D^2(\boldsymbol{u}, \boldsymbol{v}) = \sum_{m=1}^{M-1} (a_i - b_i)^2$. The right-hand side requires $3M$ floating point operations while the computational load of the integral (3.1) scales with the number of grid points and is much larger even for the employed coarse grid.

3.3. Frequency analysis

The clustering enables to characterize each operating condition by a corresponding probability distribution. Let $l = 1, \ldots, L$ be the index of the L operating conditions (here: $L = 7$). Let N_k^l be the number of snapshots of the lth attractor data in the kth cluster. Let $N^l = \sum_{k=1}^{K} N_k^l$ be the total number of snapshots of the lth attractor data. Then, the probability that a snapshot belonging to the lth operating condition lies in the kth cluster reads

$$P_k^l = \frac{N_k^l}{N^l}, \quad k = 1, \ldots, K. \tag{3.6}$$

In particular, the probability distribution is determined here as relative frequencies of cluster visits and fulfills the non-negativity, $P_k^l \geqslant 0$, and normalization condition, $\sum_{k=1}^{K} P_k^l = 1$.

3.4. *Metric of attractor overlap*

In this section, we define a measure for the attractor similarity based on the centroid configuration. The geometric configuration of the centroids is characterized by the distances between all centroids in terms of the Hilbert space norm:

$$D(\boldsymbol{c}_i, \boldsymbol{c}_j) = \|\boldsymbol{c}_i - \boldsymbol{c}_j\|_\Omega, \quad i, j = 1, \ldots, K$$

Let $\boldsymbol{P}$ and $\boldsymbol{Q}$ denote the probability distributions of attractor data $\mathcal{A}$ and $\mathcal{B}$, respectively. By definition, the support of $\boldsymbol{P}$ contains all cluster indices i with $P_i > 0$.

The distance of centroid $\boldsymbol{c}_i$ to the centroids $\boldsymbol{c}_j$ of attractor $\mathcal{A}$ is defined by

$$D(\boldsymbol{c}_i, \mathcal{A}) = \min_{j=1,\ldots,K} \{D(\boldsymbol{c}_i, \boldsymbol{c}_j) \colon P_j > 0\}.$$

Analogously, we define

$$D(\boldsymbol{c}_i, \mathcal{B}) = \min_{j=1,\ldots,K} \{D(\boldsymbol{c}_i, \boldsymbol{c}_j) \colon Q_j > 0\}.$$

as the distance of cluster $\boldsymbol{c}_i$ to attractor $\mathcal{B}$. The attractor similarity is quantified as the average distance of the centroids of attractor $\mathcal{A}$ and $\mathcal{B}$:

$$D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2} \sum_{i=1}^{K} [P_i \, D(\boldsymbol{c}_i, \mathcal{A}) + Q_i \, D(\boldsymbol{c}_i, \mathcal{B})]. \tag{3.7}$$

We refer to this quantity as *Metric of Attractor Overlap (MAO)* for following reasons. MAO has the properties of a *pseudometric* for $\boldsymbol{P}$ and $\boldsymbol{Q}$, as the defining properties can be shown for probability distributions $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{R}$:

(1) Positive semidefiniteness: $D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q}) \geqslant 0$ for all $\boldsymbol{P}$ and

$$\boldsymbol{P} = \boldsymbol{Q} \quad \Rightarrow \quad D_{geom}(\boldsymbol{P}, \boldsymbol{Q}) = 0.$$

(2) Symmetry:

$$D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q}) = D_{\text{geom}}(\boldsymbol{Q}, \boldsymbol{P}).$$

(3) Triangle inequality:

$$D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{R}) \leqslant D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q}) + D_{\text{geom}}(\boldsymbol{Q}, \boldsymbol{R}).$$

The pseudometric $D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q})$ vanishes if and only $\boldsymbol{P}$ and $\boldsymbol{Q}$ have the same support, irrespective of the non-zero values of the probability distributions. Hence, the pseudometric characterizes the geometric overlap and not the ergodic measures of the attractors. The pseudometric becomes a *metric* for the maximum cluster resolution $K = M$ of non-identical snapshots. In this case, each snapshot represents a centroid and the probability distributions associated with each cluster are unit vectors.

Another extreme case is that attractors $\mathcal{A}$ and $\mathcal{B}$ have no shared centroids, $\mathcal{A} \cap \mathcal{B} = \emptyset$. For concreteness, let us assume that both attractors arise from periodic dynamics defining limit cycles in the same plane with the same origin but with radii 10 and 11. In this case, the distance approaches unity for sufficiently large amount of snapshots. The metric is purely geometric and blind to angular frequencies and even to non-uniformity of the ergodic measure.

As a side note, a commonly used metric for probability distributions is the Jensen-Shannon distance (see §B). This distance measures probability differences on shared

clusters but is blind to the geometric distances of disjoint clusters. In the previous example of concentric co-planar circles, the metric is the same for all non-identical radii.

The cluster-based metric (3.7) is a coarse-grained approximation of the snapshot-based metric

$$D_{\text{geom}}(\mathcal{A}, \mathcal{B}) = \frac{1}{2} \sum_{i=1}^{M} \left[\chi(\boldsymbol{u}^m, \mathcal{A}) D(\boldsymbol{u}^m, \mathcal{A}) + \chi(\boldsymbol{u}^m, \mathcal{B}) D(\boldsymbol{u}^m, \mathcal{B}) \right] \tag{3.8}$$

Here, χ represents the characteristic function $\chi(\boldsymbol{u}^m, \mathcal{A}) = 1$ if $\boldsymbol{u}^m \in \mathcal{A}$ and 0 otherwise—with an analogous definition for $\mathcal{B}$. Note that $D_{\text{geom}}(\boldsymbol{P}, \boldsymbol{Q}) = D_{\text{geom}}(\mathcal{A}, \mathcal{B})$ for the maximal resolution $K = M$ and the generic case of non-identical snapshots. In the following, we consider the cluster-based pseudometric as an approximation of the snapshot metric and will refer to it as metric for reasons of simplicity.

The main computational load for the metric comes from the distance matrix of the snapshots or centroids. The computational savings of the cluster-based measure (3.7) with respect to snapshot-based metric (3.8) scale with $(K/M)^2$. For the current study with $M = 3500$ and $K = 50$ this translates to the quite dramatic factor of $\approx 2/10000$. This saving does not include the operations for clustering. These can be large for a field-based clustering of §3.2 and small for an alternative clustering, e.g. based on few aerodynamic force components.

A lossless POD reduces the computational load of the distance matrix and of the clustering to a tiny fraction of the original cost. The main cost of POD originates from the correlation matrix which requires a similar amount of operations as the distance matrix. The advantage of POD preprocessing is the many post-processing options at negligible cost.

3.5. *Proximity map*

The attractors may be visualized in a two-dimensional proximity map which preserves the metric of attractor overlap as good as possible. This task is performed with multidimensional scaling (Cox & Cox 2000). Let $\boldsymbol{P}^l$ be the cluster population in terms of the probability distributions of the lth phase. Let

$$D_{ln} := D_{\text{geom}} \left(\boldsymbol{P}^l, \boldsymbol{P}^n \right), \quad l, n = 1, \ldots, L \tag{3.9}$$

be the distance matrix between the attractor data. Let $\boldsymbol{\gamma}^l \in \mathbb{R}^2$ be a two-dimensional feature vector associated with the lth operating condition. The goal of a proximity map is to find a mapping $\boldsymbol{P} \mapsto \boldsymbol{\gamma}$ so that the pointwise distances in the feature plane are preserved as well as possible:

$$\sum_{l,n=1}^{L} \left(D_{ln} - \left\| \boldsymbol{\gamma}^l - \boldsymbol{\gamma}^n \right\| \right)^2 \overset{!}{=} \min \tag{3.10}$$

At this point, the feature plane is indeterminate with respect to translation, rotation, and reflection. However, one can request that the center of the feature vector is at the origin, removing the translative degree of freedom. Moreover, the variances of the first feature coordinate can be maximized, removing the rotational degree of freedom. Now, only the indeterminacy with respect to reflection is left, like for POD modes.

The proximity map can easily be generalized for higher-dimensional feature spaces. Moreover, the proximity map requires only a distance matrix and can thus also be used for snapshots or centroids using the snapshot metric (3.1). Proximity maps have been presented for mixing layer and Ahmed body wake data (Kaiser *et al.* 2014) and for ensembles of control laws (Kaiser *et al.* 2017*b*; Duriez *et al.* 2016; Kaiser *et al.* 2017*a*).
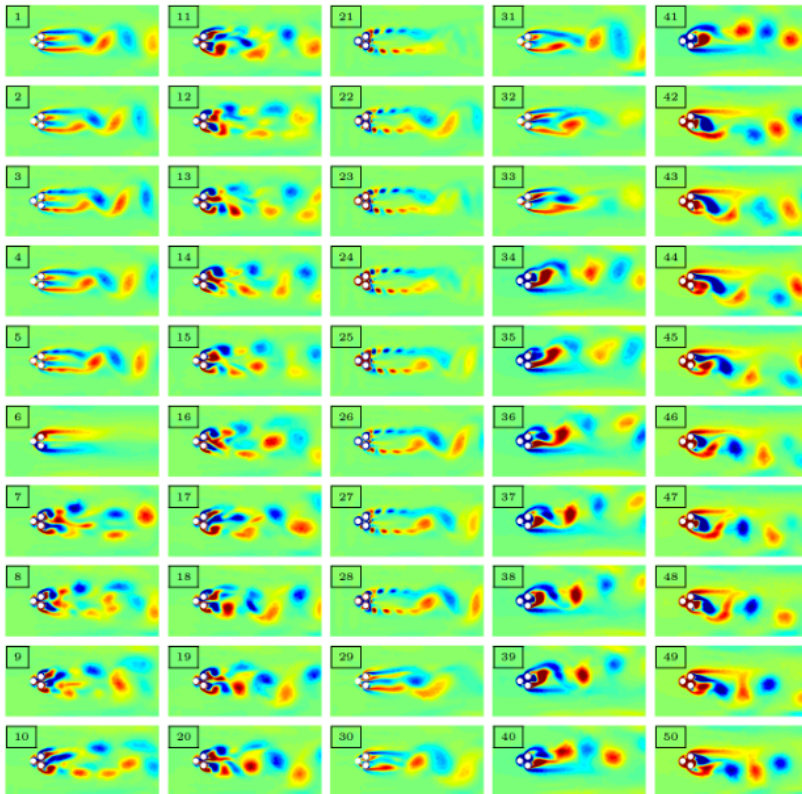
FIGURE 5. Cluster centroids of the fluidic pinball simulation. The centroids are sorted in order of appearance neglecting the first 50 time units (transients) of each phase. The visualization depicts the vorticity distribution: green, red and blue represent vanishing, positive and negative values, respectively.

4. Comparison of fluidic pinball attractors with different control laws

Here, the comparison of fluidic pinball attractors is performed following the proposed methodology of the previous section. Firstly (§4.1 and §4.2), the cluster and frequency analysis are performed. Sections §4.3 and §4.4 present the metric of attractor overlap and associated proximity maps. It should be noted that the analysis is only based on converged post-transient data while the shown temporal dynamics includes also the actuation transients.

4.1. *Clustering*

Figure 5 illustrates 50 centroids distilled from 7×500 post-transient snapshots. These centroids can be considered as the 'piano keys' for the analysis.

The cluster affiliation as a function of time is shown in Fig. 6. Clusters 1–5 describe unforced vortex shedding of phase I. We emphasize that the centroids are only meant to resolve converged attractor data and that the presentation of the transient dynamics shall only indicate the transient times. The stabilized boat tailing of phase II is described by a single centroid $k = 6$. Base bleed (phase III) is resolved by the new centroids 7–20, indicating that base bleed leads to significantly different vortex shedding structures. High-frequency (phase IV) and low-frequency forcing (phase V) can be resolved with the centroids of unforced vortex shedding and of base-bleed actuation. Low-frequency forcing
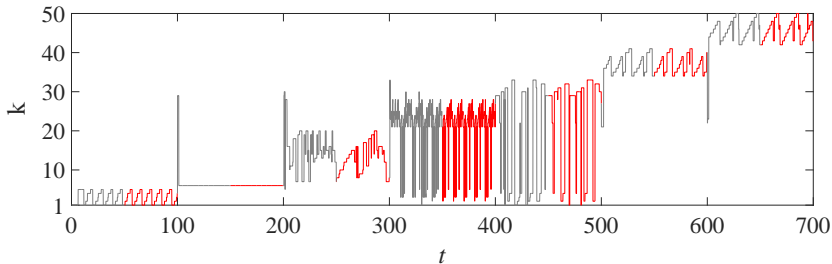
FIGURE 6. Cluster affiliation of the fluidic pinball simulation in terms of time. The applicable control laws are described in table 1. By construction, new cluster indices are a monotonously increasing function of time, neglecting the initial actuation transients.
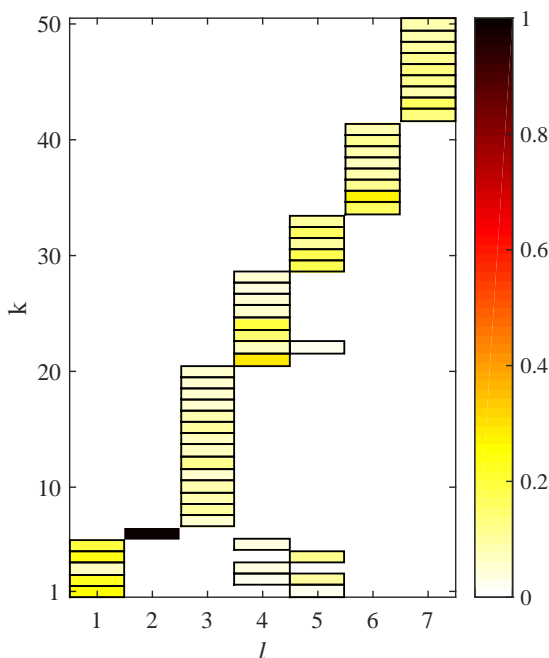

FIGURE 7. Population probabilities of the clusters for the seven fluidic pinball phases (table 1). The abscissa denotes the phase l and the ordinate the cluster index k. Each box represents a non-vanishing probability. The yellow to red tones indicate the population probability or relative frequency P_k^l from (3.6). Probabilities below 1% are kept white.

leads to a wider wake, like base bleed, and has more overlap with base-bleed dynamics. The new centroids 34–41 resolve the positive Magnus effect (phase VI) while centroids 42–50 the negative Magnus effect (phase VII). As physical intuition suggests, the strong deflected wake has no overlap with forced states which are symmetric or statistically symmetric.

4.2. *Frequency analysis*

Figure 7 displays the population of each cluster by the seven different dynamics. Probabilities below 1% are kept white. The probability distributions for each attractor are far from being uniform. This behavior is different from single attractor cluster analyses,
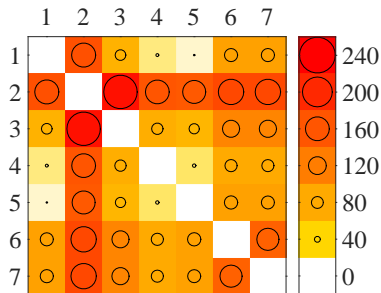
FIGURE 8. Metric of attractor overlap (3.7) for the seven control laws of the fluidic pinball (table 1). The value of $D_{\text{geom}}\left(\boldsymbol{P}^l, \boldsymbol{P}^n\right)$ is shown in the lth row and nth column as color code from white to red and by the size of the circle (see the caption on the right).

where all clusters are generally populated (Kaiser *et al.* 2014). In contrast, considering the whole simulation results in the total probability $P_k = P_k^1 + \ldots + P_k^7$ which is nonzero for any cluster k. It should be born in mind that one cluster may be populated by different attractor dynamics.

4.3. *Metric of attractor overlap*

The metric of attractor overlap for the seven operating conditions is displayed in figure 8. The positive semidefiniteness gives rise the vanishing elements in the diagonal and non-negative values elsewhere. The symmetry condition leads to the corresponding symmetry of the matrix. Natural vortex shedding ($l = 1$) is seen to be close to high- and low-frequency forcing ($l = 4, 5$). This closeness is corroborated by the probability distributions 7: the three states share joint clusters. In contrast, the wake stabilized with boat tailing shares no centroids with the other six operating conditions and has a large MAO value with them. Similarly, the positive and negative Magnus effect ($l = 6, 7$) have a large distance to each other, as expected from the opposite deflections of the wake (see figure 3) and the empty overlap of both attractors (no shared clusters in figure 7).

4.4. *Proximity map*

Figure 9 beautifully displays all centroids and all snapshots in a proximity map following §3.5. The boat-tailed and base-bleed dynamics are on the leftmost and rightmost sides near the γ_1–axis, respectively. The positive and negative Magnus effects are depicted at the bottom and top sides near the γ_2–axis. The remaining flow states populate the neighborhood near the origin. Referring to figures 3 and 6, γ_1 is seen to correspond to the drag and γ_2 to the lift. It should be noted that the automatically determined feature coordinates are strongly linked to aerodynamic forces although the forces did not enter the multidimensional scaling analysis.

The proximity map for the attractors based on MAO is depicted in figure 10. Here, the feature coordinate δ_1 correlates with the drag, δ_2 with the average lift—like in 9. Even the numerical values of the feature coordinates associated with the same operating conditions are similar. Note that both proximity maps are based on the same velocity field norm and the attractor proximity map is an aggregate of the snapshots and centroids. This quantitative similarity indicates the robustness of the multidimensional scaling for the construction of proximity maps.
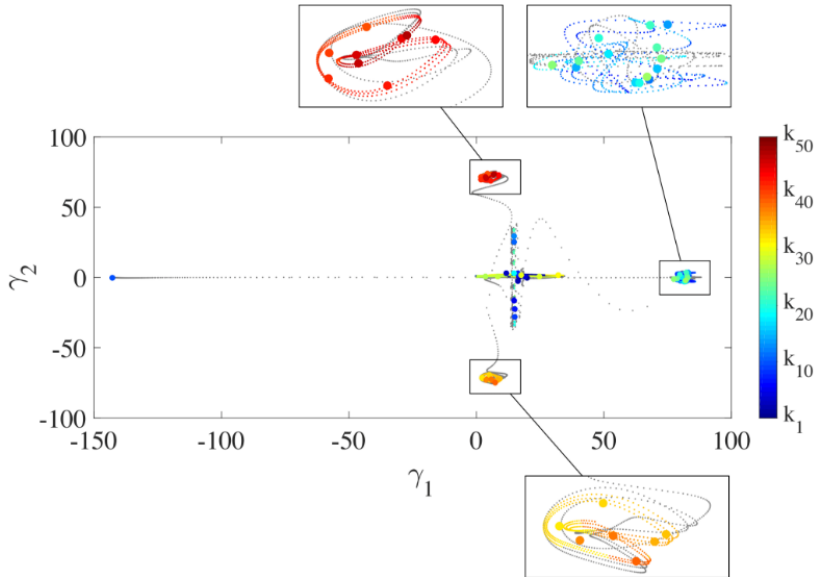
FIGURE 9. Proximity map of the centroids (large circles) together with snapshots (small circles). Snapshots from transients are displayed as gray dots, while the data for the analysis is color-coded by cluster affiliation.
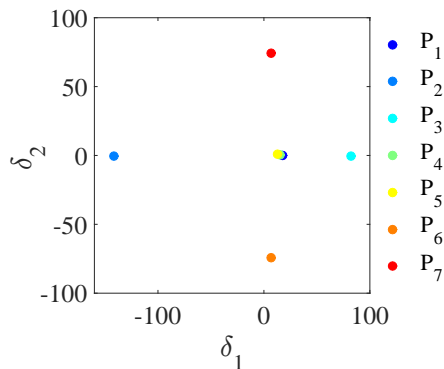

FIGURE 10. Attractor proximity map based on the geometric MAO (3.7). The points $P_1 \dots P_7$ correspond to the seven control laws in table 1.

5. Conclusions and outlook

We have proposed a novel quantitative measure for the similarity of attractor data—augmented by proximity maps for the individual snapshot to the whole attractor. A critical enabler is the coarse-graining of the snapshot data from all attractors into discrete clusters. This coarse-graining allows one to associate discrete frequency distributions with each attractor. The proposed metric of attractor overlap (MAO) incorporates the cluster geometry and the probability distributions.

Our proposed comparison methodology has five discriminating features. Firstly, the difference between two attractors is quantified in a single, non-negative number encapsulating all resolved coherent structures. Secondly, the cluster frequency distribution of each attractor identifies shared and disjoint flow states as a manageable number of centroids. Thirdly, the metric of attractor overlap enables the automatic creation of a proximity map from a myriad of attractors. Neighboring feature points represent attractors with

similar flow states, largely separated points indicate attractors with different physics. Fourthly, the approach is compatible with statistical post-processing: The centroids and their probabilities define the mean flow and — in principle — any higher-order moment, like Reynolds stress or fluctuation energy. Even POD modes can be derived from the centroids. Finally, the comparison methodology can be automatized and has little subjective bias. One does not need to decide in advance on a cost function, such as drag power, a frequency filter, such as low-pass filter, important flow features, such as vortices, just to name a few.

The metric of attractor overlap is only based on three decisions. Firstly, one needs to choose the flow state space, e.g. the velocity field in a domain Ω. Such a decision is unavoidable for any metric. Secondly, a metric for two flow states is needed. A $L^2(\Omega)$ Hilbert space norm appears as a natural candidate. Thirdly, the number of clusters or, synonymously, the typical size of a cluster needs to be chosen. The metric is found to quickly converge with increasing number of clusters. The terminal value is defined only by the snapshots and their relative distances.

The comparison methodology has been applied to the fluidic pinball, the flow around a cluster of three rotating cylinders. The considered attractors include the unforced reference, aerodynamic boat tailing, base bleed, symmetric high-frequency forcing. symmetric low frequency forcing and a Magnus effect deflecting the wake to both sides. The metric of attractor overlap confirms physical inspection of the seven different configurations. Firstly, boat tailing with complete wake suppression differs strongly from all other states. Secondly, both Magnus effects also strongly differ from all other states. Thirdly, unforced vortex shedding and wakes manipulated by base bleed as well as high-frequency and low-frequency forcing have a significant overlap of shared clusters, i.e. are similar.

Particularly noteworthy are the proximity maps of the snapshots and the cluster-based attractors. The trajectory through all seven configurations displays the dynamics very clearly. Surprisingly. the first and second feature coordinates resemble the drag and lift respectively. The features of the snapshots show the reduced drag by boat tailing (left) and the increased drag by base bleed (right) and significant averaged lift by both Magnus forcing (top and bottom). The proximity map of the attractor visually elucidates the discussed neighborhood relations in the first three feature coordinates.

Evidently, the presented MAO comparison encourages numerous other applications, like a comparison between computed and experimental velocity fields and a comparison between flow behavior under different actuations, e.g. periodic forcing or machine learning control studies (Duriez *et al.* 2016). The analysis may also be applied to sensor data, e.g. a hot-wire rig.

Future research may significantly extend the MAO methodology. The employed Hilbert space norm is a good initial choice but highly sensitive to small mode deformations (Noack 2016). A small change in wavenumber gives rise to unphysically large differences in the norm. Force-related feature vectors and manifold learning may be one remedy (Loiseau *et al.* 2017). Rigorous physics-based criteria for the cluster numbers need to be advanced. The comparison may also be targeted towards single- or multi-objective goals by generalizing the snapshot metric. So far, the comparison only addressed ergodic properties of the attractor, and the snapshot data only needed to be statistically representative and not time-resolved. The temporal dynamics of different attractors may be compared using time-resolved snapshots and corresponding Markov models (Kaiser *et al.* 2014).

Summarizing, a rational automated comparison method has been proposed which holds significant promise for future data assessments. The authors are actively pushing this direction.

Appendix A. Proper orthogonal decomposition versus clustering

The cluster analysis is compared with the proper orthogonal decomposition (POD) from the same data in the same observation domain. Each snapshot is expanded in terms of the mean flow $\boldsymbol{u}_0(\boldsymbol{x})$ of all post-transient attractor data and N POD modes $\boldsymbol{u}_i(\boldsymbol{x})$, $i = 1, \ldots, N$ and their corresponding amplitudes $a_i(t)$, $i = 1, \ldots, N$:

$$\boldsymbol{u}(\boldsymbol{x}, t) \approx \boldsymbol{u}_0(\boldsymbol{x}) + \sum_{i=1}^{N} a_i(t)\, \boldsymbol{u}_i(\boldsymbol{x}). \tag{A 1}$$

Figures 11 and 12 display the first ten POD modes and their amplitudes, respectively. Mode 1 resolves a base-flow deformation, like a shift mode (Noack *et al.* 2003). Mode 5 represents a symmetric near-field modulation. The other modes have more oscillatory structures. We shall not pause to hypothesize about the physical meaning of these modes. Modes with cleaner frequency content yet near the optimal residual could, for instance, be constructed with recursive DMD (Noack *et al.* 2016). The cluster analysis resulting in the snapshot cluster affiliation (Fig. 6) and centroids (Fig. 5) is physically easier to interpret than the POD.

It may be noted that POD also allows for another metric of attractor overlap. Let $p(\boldsymbol{a})$ be the probability distribution associated with one attractor and $q(\boldsymbol{a})$ the distribution of another one. Then, the continuous version of the Jensen-Shannon distance (B 4) can be applied. An advantage is that the snapshot distance (3.1) is reflected in the construction of the metric. A disadvantage is the need for approximations to construct a continuous probability distribution from a finite number of snapshots.

Appendix B. Jensen-Shannon distance

Probability distributions can easily be compared using information measures. Let $\boldsymbol{Q} = [Q_1, \ldots, Q_K]$ be a reference probability distribution and $\boldsymbol{P} = [P_1, \ldots, P_K]$ a new measured one. Then, the information gained from $\boldsymbol{P}$ with respect to the reference $\boldsymbol{Q}$ is quantified by the *Kullback-Leibler divergence* (Kullback & Leibler 1951; Kullback 1959),
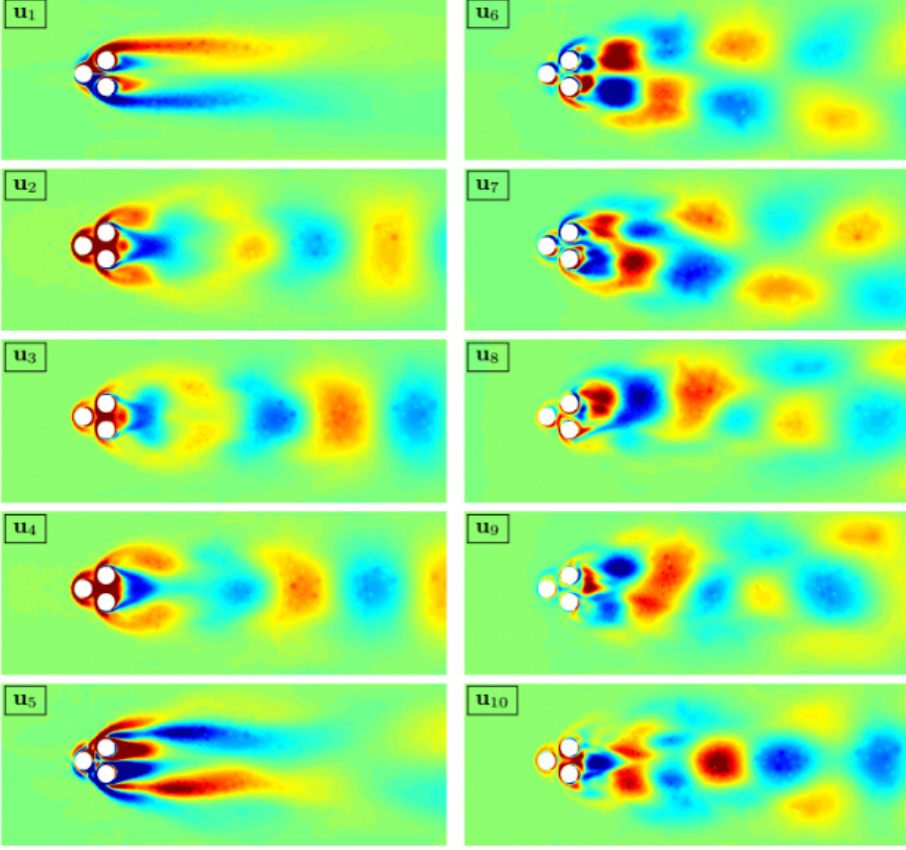
FIGURE 11. POD modes $\boldsymbol{u}_i(\boldsymbol{x})$, $i = 1, \ldots, 10$.

also called *relative entropy*,

$$D_{KL}(\boldsymbol{P}\|\boldsymbol{Q}) = \sum_{k=1}^{K} P_k \ln \left[\frac{P_k}{Q_k} \right]. \tag{B 1}$$

Identical probability distributions $\boldsymbol{P} = \boldsymbol{Q}$ give rise to a vanishing Kullback-Leibler divergence. Different distributions yield a positive value. If $P_k = 0$, the term $P_k \ln P_k / Q_k$ is interpreted as zero because $\lim_{x\to 0} x \ln x = 0$. Strictly speaking, the Kullback-Leibler divergence is not defined in case there exists a k for which $Q_k = 0$ and $P_k > 0$. One could follow a common practice of smoothing using the *absolute discounting method* by replacing Q_k by a small value, e.g. $\epsilon = 0.001$.

The Kullback-Leibler divergence is not symmetric, i.e. $D_{KL}(\boldsymbol{P}\|\boldsymbol{Q}) \not\equiv D_{KL}(\boldsymbol{Q}\|\boldsymbol{P})$. Hence, it cannot serve as a metric. The probability distributions $\boldsymbol{P}$ and $\boldsymbol{Q}$ may be compared with the *Jensen-Shannon divergence* which is defined as symmetrized and smoothened Kullback-Leibler divergence:

$$JSD(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2} \left[D_{KL}(\boldsymbol{P}\|\boldsymbol{M}) + D_{KL}(\boldsymbol{Q}\|\boldsymbol{M}) \right], \quad \text{where} \quad \boldsymbol{M} = \frac{1}{2} \left[\boldsymbol{P} + \boldsymbol{Q} \right]. \tag{B 2}$$
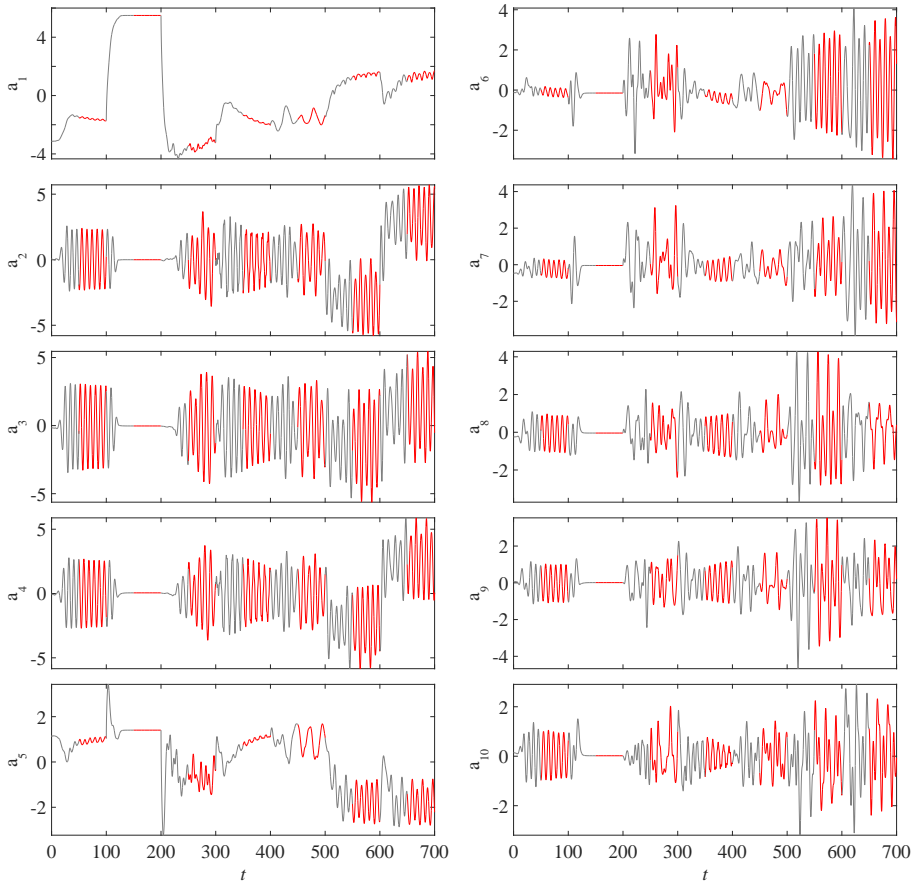
FIGURE 12. POD amplitudes $a_i(t)$, $i = 1, \ldots, 10$. Snapshot sequences after the decay of the transients are employed for the analysis (colored in red).

Alternatively,

$$JSD(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2} \sum_{k=1}^{K} \left(P_k \ln \left[\frac{2P_k}{P_k + Q_k} \right] + Q_k \ln \left[\frac{2Q_k}{P_k + Q_k} \right] \right). \qquad (\text{B 3})$$

The summation term is interpreted as zero if $P_k = Q_k = 0$ or if the numerator of the logarithm argument vanishes. Note that we don't need the ϵ threshold of the Kullback-Leibler entropy anymore.

The square root of the Jensen-Shannon divergence has all properties of a metric (Endres & Schindelin 2003) and is referred to as *Jensen-Shannon distance*,

$$D_{JS}(\boldsymbol{P}, \boldsymbol{Q}) = \sqrt{JSD(\boldsymbol{P}, \boldsymbol{Q})}. \qquad (\text{B 4})$$

This distance (B 4) defines an *entropic* metric of attractor overlap between two attractor data represented by the cluster distributions $\boldsymbol{P}$ and $\boldsymbol{Q}$.

Figure 13 displays the Kullback-Leibler and Jensen-Shannon divergences between the seven fluidic pinball attractors. The diagonal entries vanish by definition. Unforced vortex shedding is seen to be similar to high- and low-frequency forcing, as was indicated already by the many shared clusters in Fig. 6. The difference between unforced shedding, base-bleed dynamics, and both Magnus effects is larger as they share no joint clusters. The
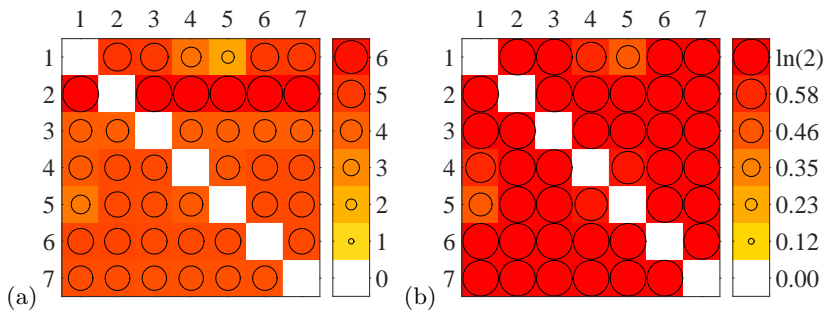
FIGURE 13. Distances between attractors: (a) Kullback-Leibler divergence and (b) Jensen-Shannon divergence based on the cluster probability distribution for each of the seven control phases. The row and column of $D_{KL}(P\|Q)$ corresponds to the first ($\boldsymbol{P}$) and second argument ($\boldsymbol{Q}$). The value of the divergence is color-coded and indicated by the size of the circle (see the corresponding caption).

stabilized boat tailing with a single cluster is seen to be very different from all other phases with no cluster overlap. The Kullback-Leibler divergence is, as expected, not symmetric. For instance. the divergence of boat-tailing with respect to base-bleed is larger than the other way round. It should be noted that the Jensen-Shannon distance is typically ln(2) (or unity if the base 2 logarithm is used) corresponding to no shared clusters.

The Jensen-Shannon distance is a natural metric for probability distributions. Yet, it is blind to the geometric location of the centroids. Let us assume the first attractor data populates only cluster 1, $P_i = \delta_{i,1}$, and the second data occupies only cluster 2, $Q_i = \delta_{i,2}$. In this case, $JSD(\boldsymbol{P},\boldsymbol{Q}) = \ln 2$, regardless of whether the centroids are very close or very far from each other. This property makes the Jensen-Shannon distance strongly dependent on the number of clusters or, equivalently, on the typical size of the clusters. For the minimum number of clusters, $K = 1$, we have obtained the trivial result $P_1 = Q_1 = 1$ and $JSD = 0$. For the maximum number of clusters, each snapshot represents one centroid and defines one cluster. In the generic case of different snapshots, the Jensen-Shannon distance is also $\ln 2$ independent of the geometric location of the snapshots. Only a few of the seven fluidic pinball phases share joint clusters and the proximity maps based on (B 4) provides limited physical insight. The metric may be far more meaningful in other cases in which most attractors have significant overlap.

REFERENCES

ARTHUR, D. & VASSILVITSKII, S. 2007 k-means++: The advantages of careful seeding. In *Proc. of the 18th annual ACM-SIAM symposium on Discrete algorithms*, pp. 1027–1035. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

BANSAL, M. S. & YARUSEVYCH, S. 2017 Experimental study of flow through a cluster of three equally spaced cylinders. *Exp. Thermal Fluid Sci.* **80**, 203–217.

BARROS, D., BORÉE, J., NOACK, B. R., SPOHN, A. & RUIZ, T. 2016 Bluff body drag manipulation using pulsed jets and Coanda effect. *J. Fluid Mech.* **805**, 442–459.

BEARMAN, PW 1967 The effect of base bleed on the flow behind a two-dimensional model with a blunt trailing edge. *Aeronautical Quarterly* **18** (03), 207–224.

BERKOOZ, G., HOLMES, P. & LUMLEY, J. L. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Ann. Rev. Fluid Mech.* **25**, 539–575.

BRUNTON, S. L. & NOACK, B. R. 2015 Closed-loop turbulence control: Progress and challenges. *Appl. Mech. Rev.* **67** (5), 050801:01–48.

BURKARDT, J., GUNZBURGER, M.D. & LEE, H.-C. 2004 Centroidal Voronoi Tessellation-Based Reduced-Order Modeling of Complex Systems. *Tech. Rep.*. Florida State University.

COX, T. F. & COX, M. A. A. 2000 *Multidimensional Scaling*, 2nd edn., *Monographs on Statistics and Applied Probability*, vol. 88. Chapman and Hall.

DURIEZ, T., BRUNTON, S. L. & NOACK, B. R. 2016 *Machine Learning Control — Taming Nonlinear Dynamics and Turbulence*, *Fluid Mechanics and Its Applications*, vol. 116. Springer-Verlag.

ENDRES, D. M. & SCHINDELIN, J. E. 2003 A new metric for probability distributions. *IEEE Transactions on Information Theory* **49**, 1858–1860.

GEROPP, D. 1995 Process and device for reducing the drag in the rear region of a vehicle, for example, a road or rail vehicle or the like. United States Patent **US 5407245 A**.

GEROPP, D. & ODENTHAL, H.-J. 2000 Drag reduction of motor vehicles by active flow control using the Coanda effect. *Exp. Fluids* **28** (1), 74–85.

HALLER, G. 2005 An objective definition of a vortex. *J. Fluid Mech.* **525**, 1–26.

HU, J. & ZHOU, Y. 2008a Flow structure behind two staggered circular cylinders. Part 1. Downstream evolution and classification. *J. Fluid Mech.* **607**, 51–80.

HU, J. & ZHOU, Y. 2008b Flow structure behind two staggered circular cylinders. Part 2: Heat and momentum transport. *J. Fluid Mech.* **607**, 81–107.

JEONG, J. & HUSSAIN, F. 1995 On the identification of a vortex. *J. Fluid Mech.* **285**, 69–94.

KAISER, E., LI, R. & NOACK, B. R. 2017a On the control landscape topology. In *The 20th World Congress of the International Federation of Automatic Control (IFAC)*, pp. 1–4. Toulouse, France.

KAISER, E., NOACK, B. R., CORDIER, L., SPOHN, A., SEGOND, M., ABEL, M. W., DAVILLER, G., ÖSTH, J., KRAJNOVIĆ, S. & NIVEN, R. K. 2014 Cluster-based reduced-order modelling of a mixing layer. *J. Fluid Mech.* **754**, 365–414.

KAISER, E., NOACK, B. R., SPOHN, A., CATTAFESTA, L. N. & MORZYŃSKI, M. 2017b Cluster-based control of a separating flow over a smoothly contoured ramp. *Theor. Comput. Fluid Dyn.* **31**, 579–593.

KASTEN, J., REININGHAUS, J., HOTZ, I., HEGE, H.-C., NOACK, B. R., DAVILLER, G., COMTE, P. & MORZYŃSKI, M. 2014 Acceleration feature points of unsteady shear flows. *Archives of Mechanics* **68**, 55–80.

KULLBACK, S. 1959 *Information Theory and Statistics*, 1st edn. New York: John Wiley.

KULLBACK, S. & LEIBLER, R. A. 1951 On information and sufficiency. *Annals Math. Stat.* **22**, 79–86.

LOISEAU, J.-CH., NOACK, B. R. & BRUNTON, S. L. 2017 Sparse reduced-order modeling: Sensor-based dynamics to full-state estimation. *J. Fluid Mech.* **(submitted, see arXiv:1706.03531 [physics.fui-dyn])**, 1–28.

LOYD, S. 1982 Least squares quantization in pcm. *IEEE Trans. Inform. Theory* **28** (2), 129–137.

LUGT, H.J. 1996 *Introduction to Vortex Theory*. Potomac (Maryland, U.S.A.): Vortex Flow Press.

MACQUEEN, J. 1967 Some methods for classification and analysis of multivariate observations. *Proc. of the Fifth Berkeley Symp. On Math. Stat. and Prob.* **1**, 281–297.

NOACK, B. R. 2016 From snapshots to modal expansions – bridging low residuals and pure frequencies. *J. Fluid Mech. – Focus in Fluids* **802**, 1–4.

NOACK, B. R., AFANASIEV, K., MORZYŃSKI, M., TADMOR, G. & THIELE, F. 2003 A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.* **497**, 335–363.

NOACK, B. R. & MORZYŃSKI, M. 2017 The fluidic pinball — a toolkit for multiple-input multiple-output flow control (version 1.0). *Tech. Rep.* 02/2017. Chair of Virtual Engineering, Poznan University of Technology, Poland.

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

OXLADE, A. R., MORRISON, J. F., QUBAIN, A. & RIGAS, G. 2015 High-frequency forcing of a turbulent axisymmetric wake. *J. Fluid Mech.* **770**, 305–318.

PASTOOR, M., HENNING, L., NOACK, B. R., KING, R. & TADMOR, G. 2008 Feedback shear layer control for bluff body drag reduction. *J. Fluid Mech.* **608**, 161–196.

RAIBAUDO, C., ZHONG, P., MARTINUZZI, R. J. & NOACK, B. R. 2017 Closed-loop control

of a triangular bluff body using rotating cylinders. In *The 20th World Congress of the International Federation of Automatic Control (IFAC)*, pp. 1–6. Toulouse, France.

ROLLAND, R. 2017 Fluidic pinball — a control study. Master's thesis, LIMSI and ENSAM, Paris, France.

ROUSSOPOULOS, K. 1993 Feedback control of vortex shedding at low Reynolds numbers. *J. Fluid Mech.* **248**, 267–296.

ROWLEY, C. W., MEZIĆ, I., BAGHERI, S., SCHLATTER, P. & HENNINGSON, D.S. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **645**, 115–127.

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

STEINHAUS, H. 1956 Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci.* **4** (12), 801–804.

THEOFILIS, V. 2011 Global linear instability. *Ann. Rev. Fluid Mech.* **43**, 319–352.

THIRIA, B., GOUJON-DURAND, S. & WESFREID, J. E. 2006 The wake of a cylinder performing rotary oscillations. *J. Fluid Mech.* **560**, 123–147.

WOOD, C. J. 1964 The effect of base bleed on a periodic wake. *Journal of the Royal Aeronautical Society* **68** (643), 477–482.