Machine Learning
for Fluid Mechanics
Features and Autoencoders



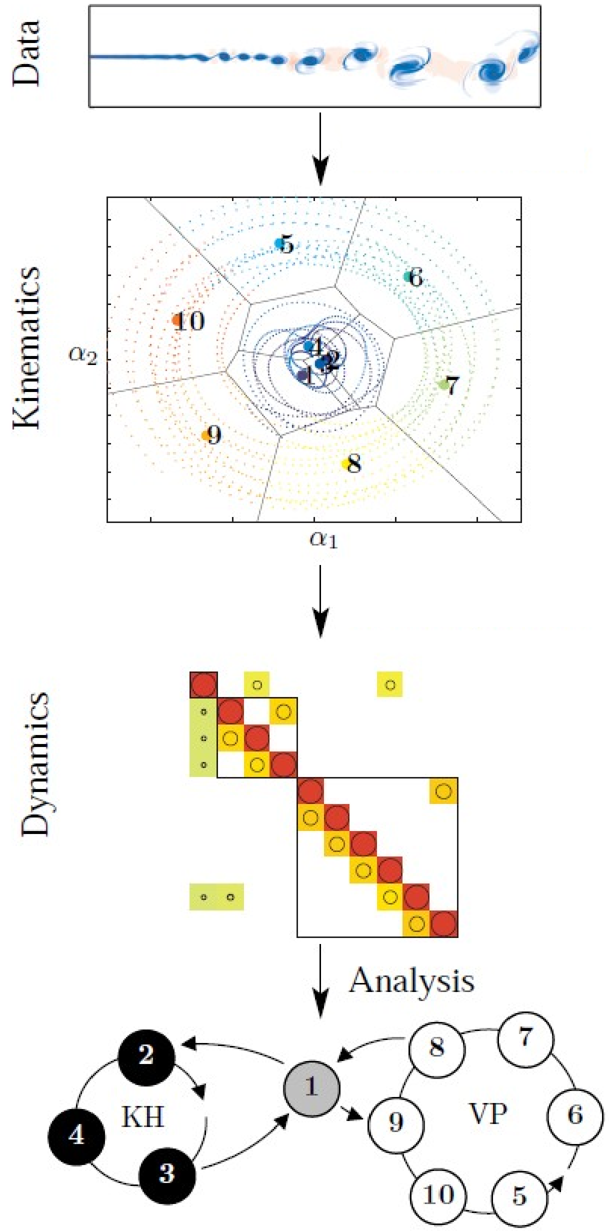
Bernd Noack

HIT, China & TU Berlin

TU Berlin, Germany, 2020-03-03

Overview

How does Cluster-based ROM work?



Time-resolved snapshot data

Cluster analysis

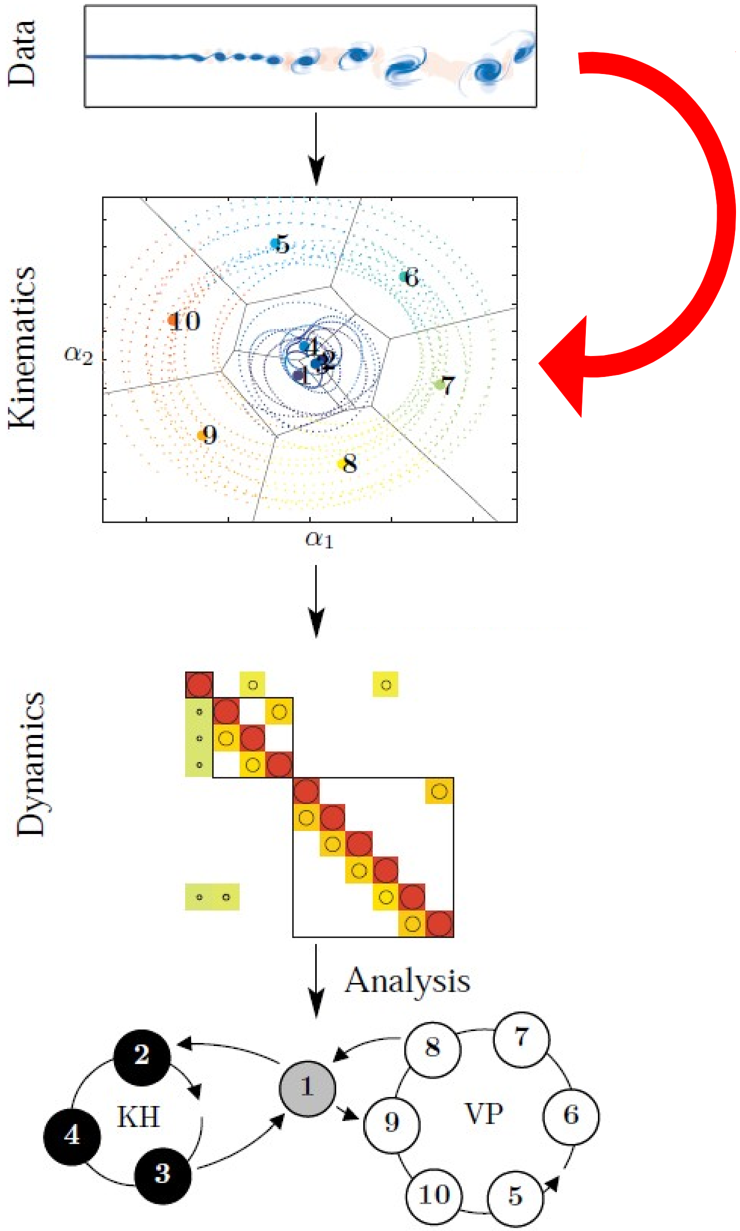
Low number of representative states

Markov process

Dynamic model for the cluster transitions

Further analysis of transition processes:
Cluster sociology (cliques, hubs),
dynamically important clusters, etc.

How does CROM work?



Time-resolved snapshot data

Cluster analysis

Low number of representative states

Markov process

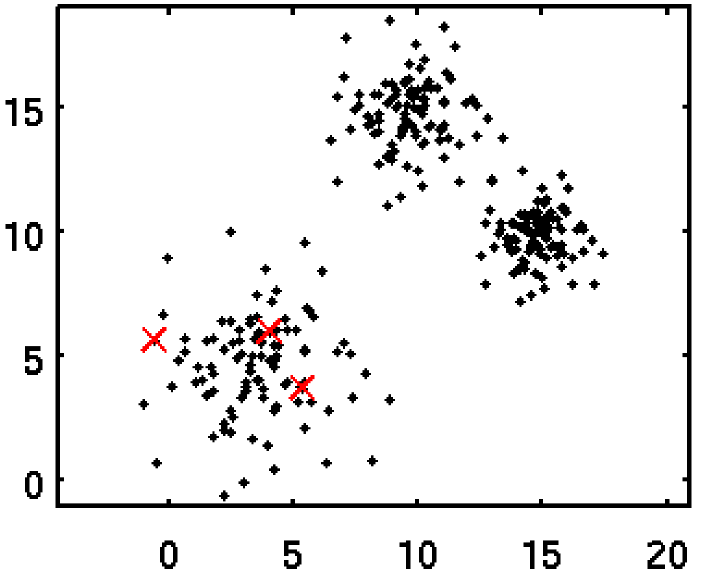
Dynamic model for the cluster transitions

Further analysis of transition processes:
Cluster sociology (cliques, hubs),
dynamically important clusters, etc.

Cluster analysis

K-means algorithm (see: *Loyd, 1957 & 1982)*:

1. Initialize K clusters (randomly, Kmeans++)



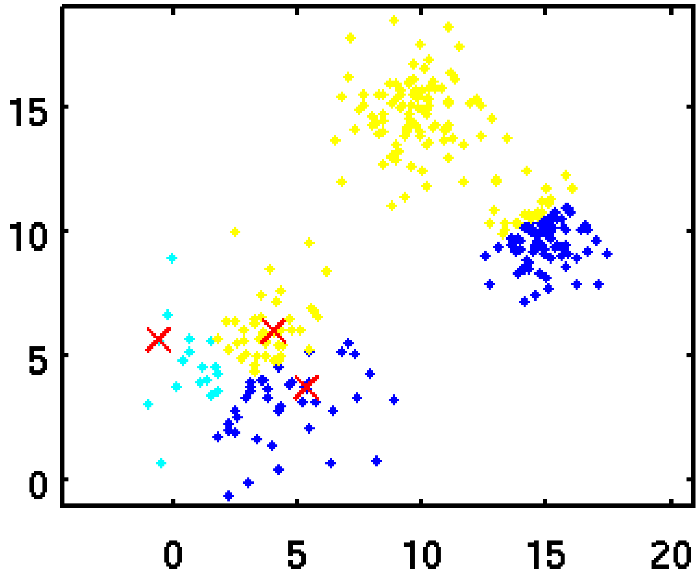
Cluster analysis

K-means algorithm (see: *Loyd, 1957 & 1982)*:

1. Initialize K clusters (randomly, Kmeans++)

   Do

2. Assign objects to closest cluster



Cluster analysis

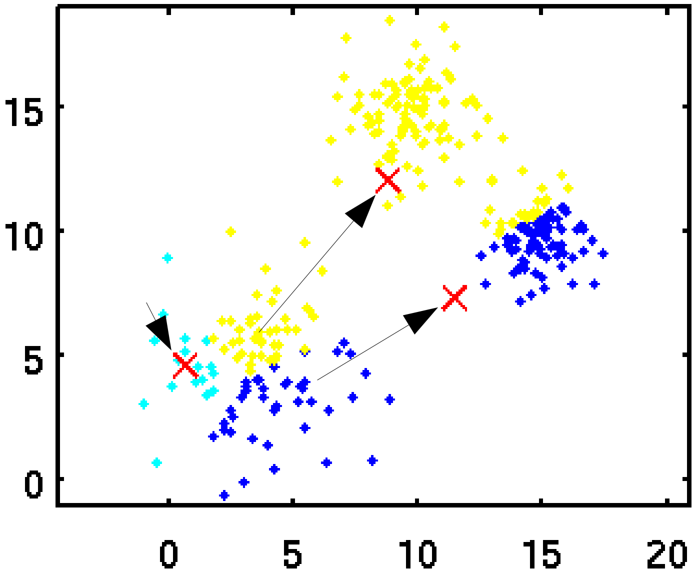
K-means algorithm (see: *Loyd, 1957 & 1982)*:

1. *Initialize K clusters (randomly, Kmeans++)*

   Do

2. *Assign objects to closest cluster*
3. *Compute new mean of objects in clusters*

   until clusters are converged.



Cluster analysis

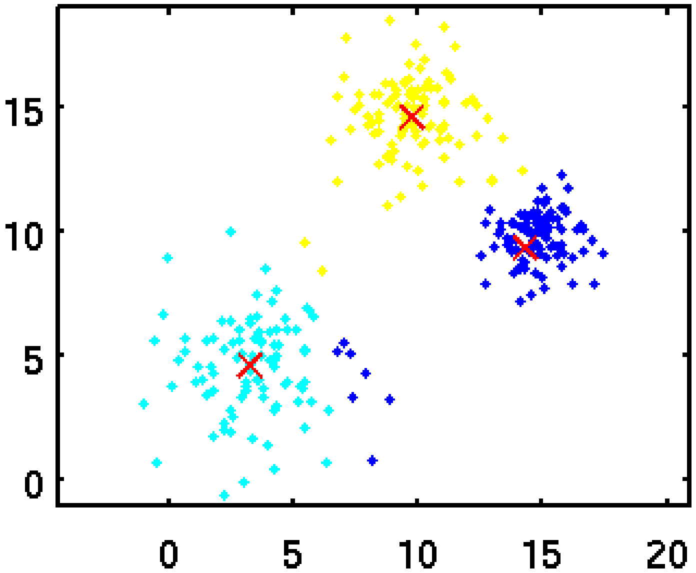
K-means algorithm (see: *Loyd, 1957 & 1982)*:

1. *Initialize K clusters (randomly, Kmeans++)*

   Do

2. *Assign objects to closest cluster*
3. *Compute new mean of objects in clusters*

   until clusters are converged.



Cluster analysis

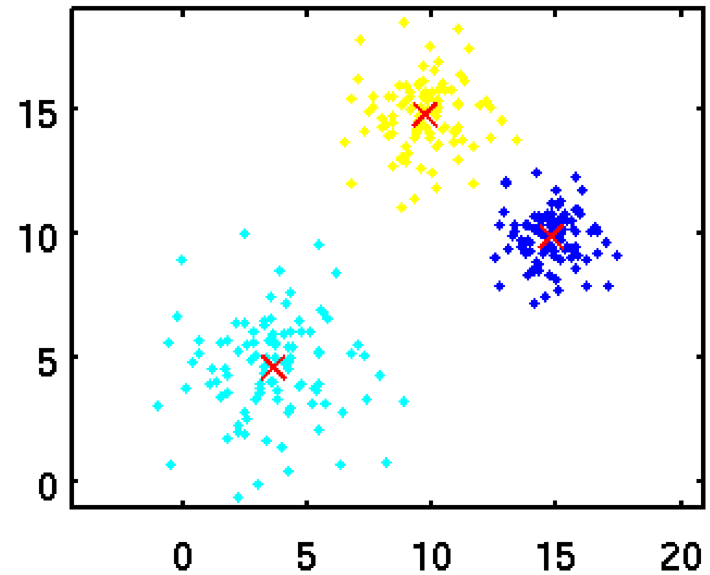
K-means algorithm (see: *Loyd, 1957 & 1982)*:

1. *Initialize K clusters (randomly, Kmeans++)*

   Do

2. *Assign objects to closest cluster*
3. *Compute new mean of objects in clusters*

   until clusters are converged.



Cluster analysis

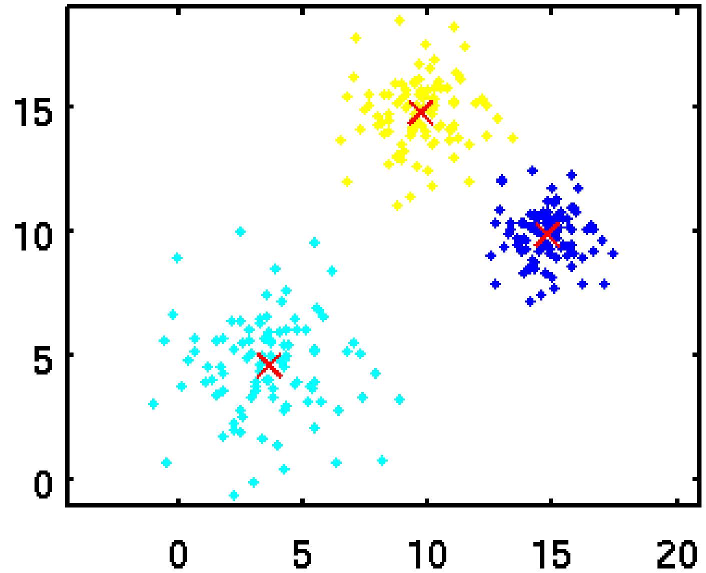
K-means algorithm (see: *Loyd, 1957 & 1982*):

1. *Initialize K clusters (randomly, Kmeans++)*

   Do

2. *Assign objects to closest cluster*
3. *Compute new mean of objects in clusters*

   until clusters are converged.



Optimization problem:

$$J(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K) = \sum_{k=1}^{K} \sum_{\boldsymbol{v}^m \in \mathscr{C}_k} \|\boldsymbol{v}^m - \boldsymbol{c}_k\|_{\Omega}^2$$

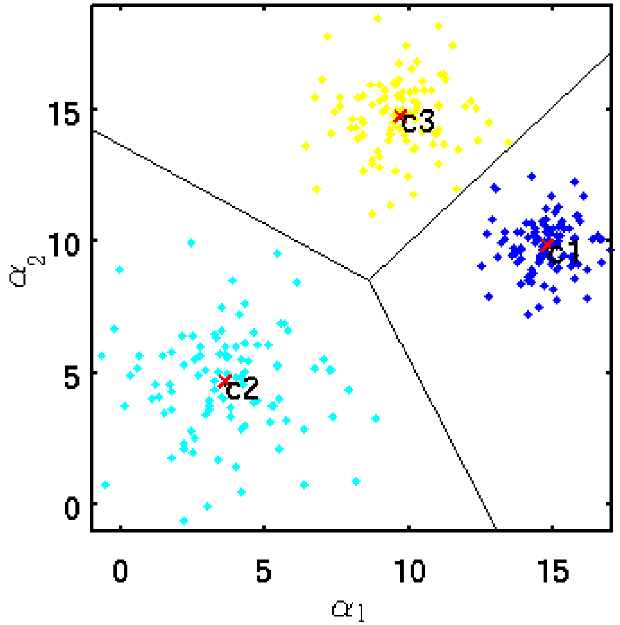
$$\boldsymbol{c}_1^{opt}, \ldots, \boldsymbol{c}_K^{opt} = \underset{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K}{\operatorname{argmin}} J(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_K)$$

Cluster analysis

Issues:

- Cluster number needs to be defined in advance (tests for good choice)
- Converges to local optimum
- Clustering effected by initial choice of centroids
- Definition of distance

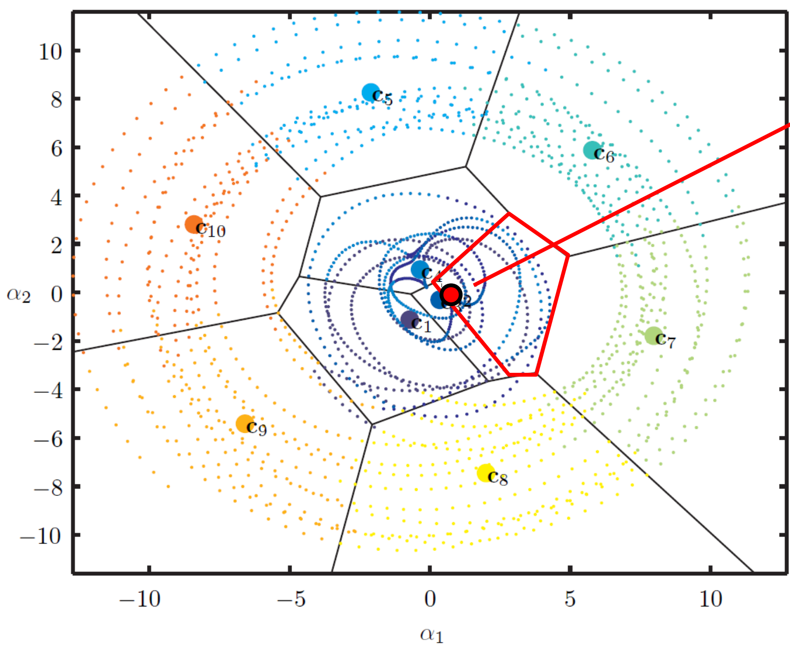
Voronoi diagram:



Clustering of mixing layer data

K-means applied to POD coefficient vectors for K=10, M=2000 snapshots

- Compression of all snapshots into a few modes
- Uncertainty due to cluster size

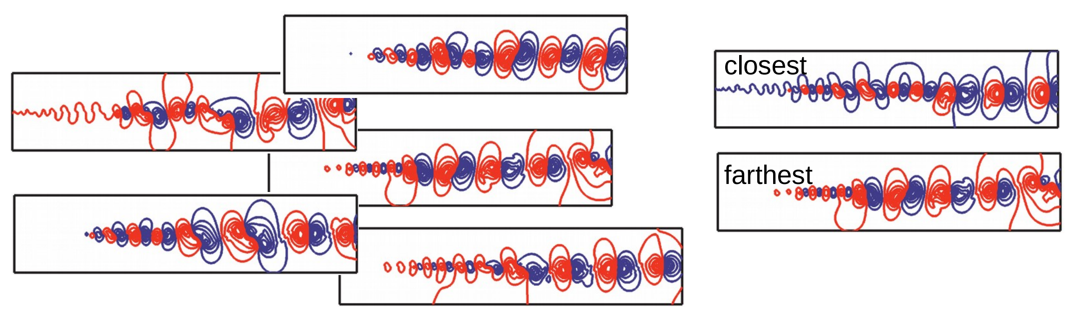


Voronoi diagram

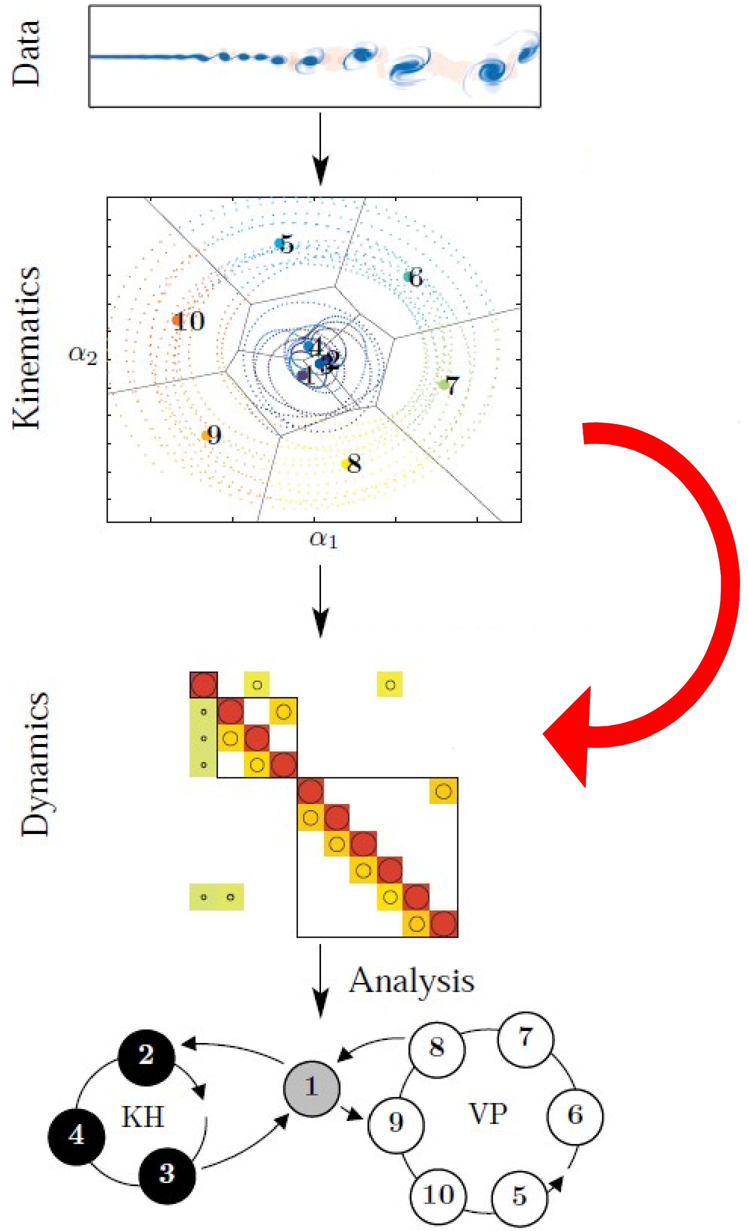
Cluster centroid c2:

$$c_k = \frac{1}{n_k} \sum_{v^m \in \mathscr{C}_k} v^m$$

Some snapshots in this cluster:



How does CROM work?



Time-resolved snapshot data

Cluster analysis

Low number of representative states

Markov process

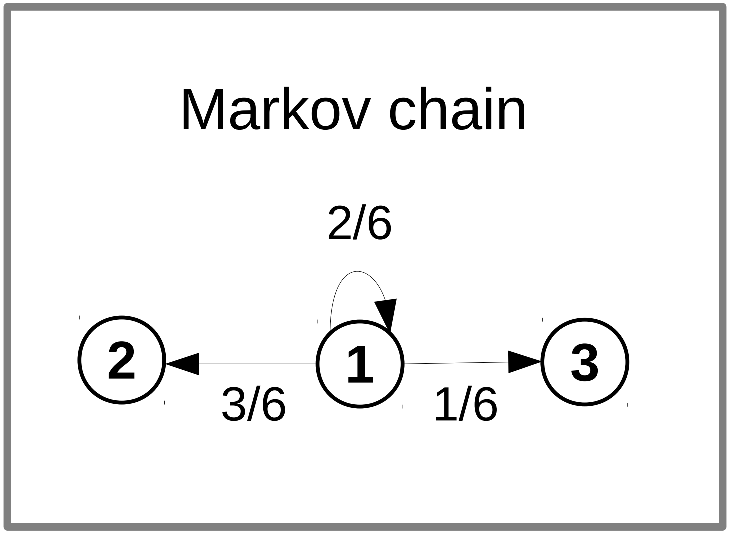
Dynamic model for the cluster transitions

Further analysis of transition processes:
Cluster sociology (cliques, hubs),
dynamically important clusters, etc.

Transition matrix model

- Discrete states with probabilities $\boldsymbol{p}^l = \left[p_1^l, \ldots, p_K^l\right]^T$ $\quad \sum_{k=1}^{K} p_k^l = 1$
- Transition probabilities to move from one state to another depends on current state

$$\boldsymbol{p}^{l+1} = P\,\boldsymbol{p}^l$$

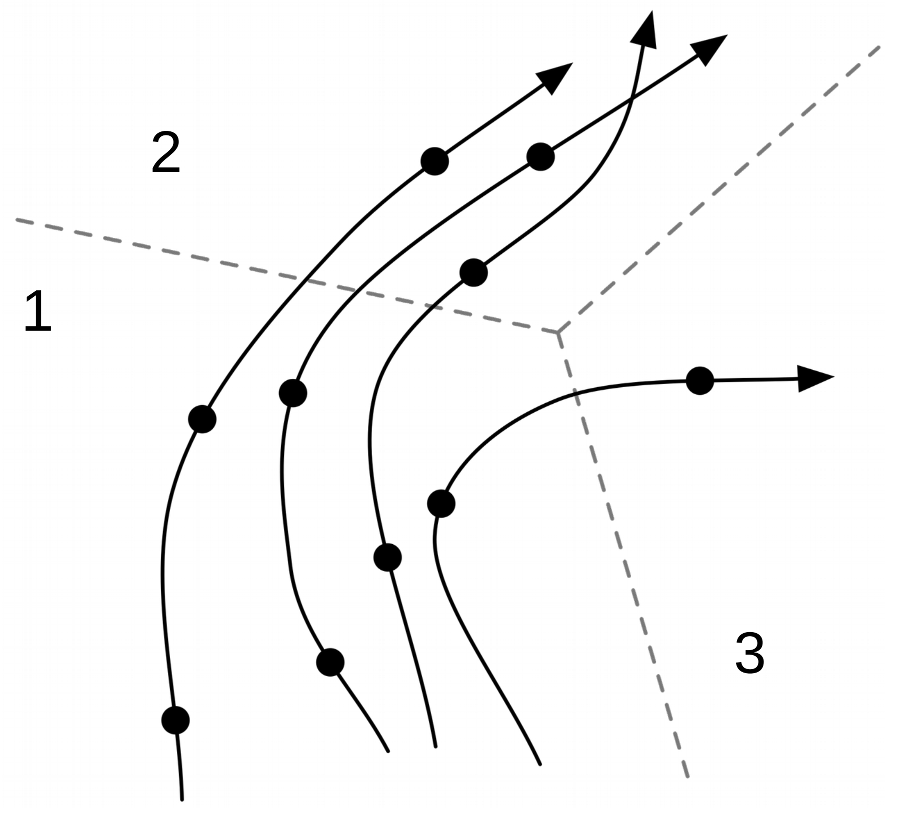


Transition matrix model

- Transition matrix: $P_{jk} = \frac{n_{jk}}{n_k}$

- Example:

$P_{21} = ?$

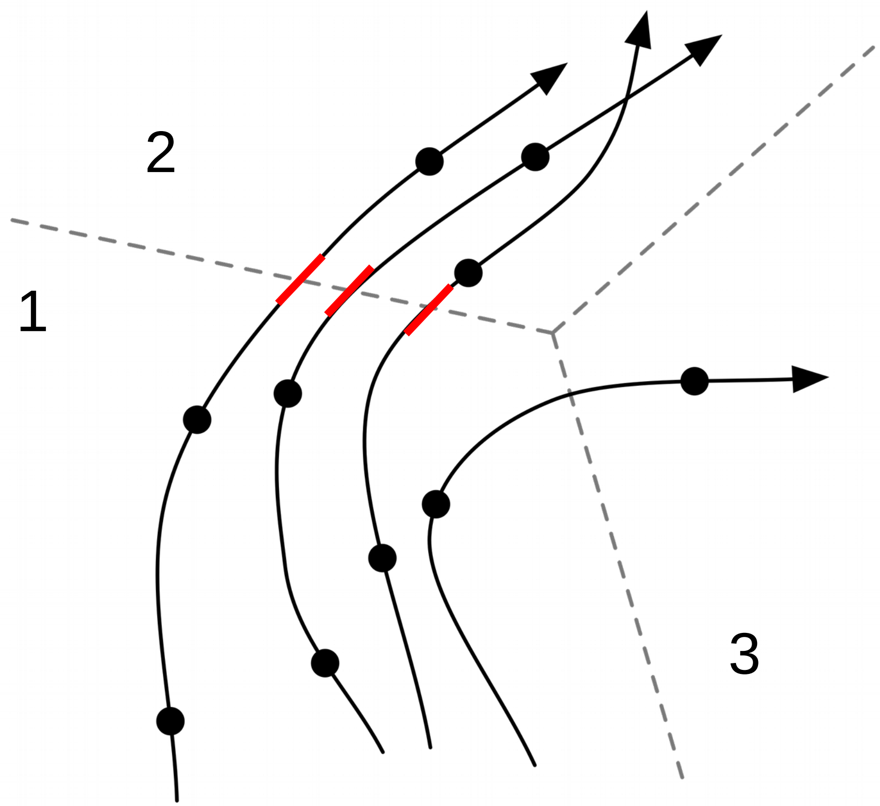


Transition matrix model

- Transition matrix: $P_{jk} = \frac{n_{jk}}{n_k}$

- Example:

$$P_{21} = \frac{3}{?}$$

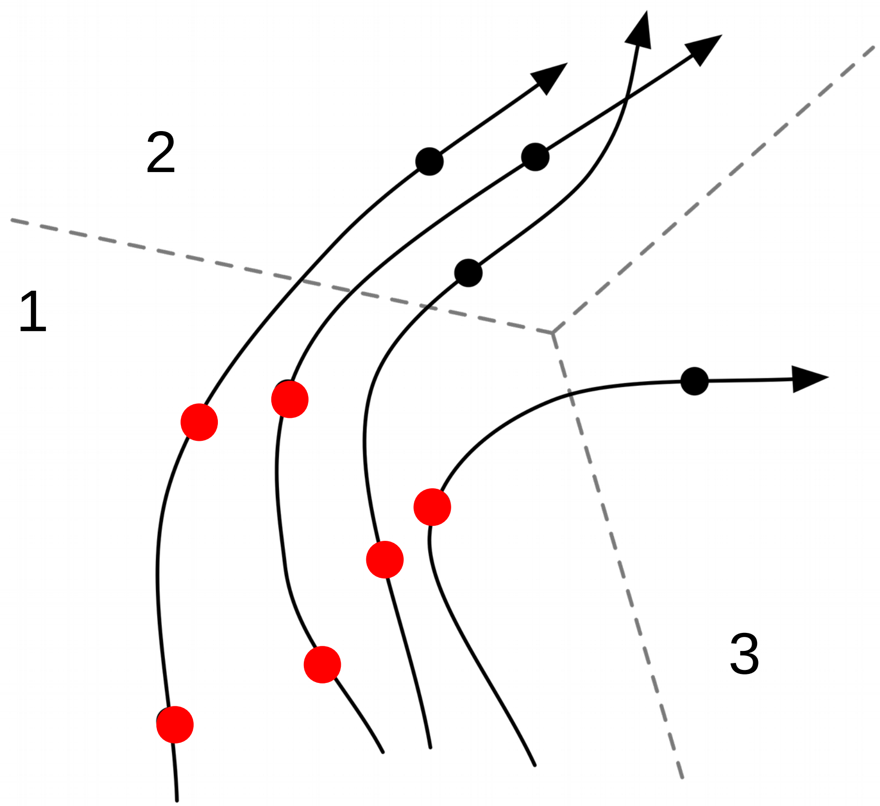


Transition matrix model

- Transition matrix: $P_{jk} = \frac{n_{jk}}{n_k}$

- Example:

$$P_{21} = \frac{3}{6}$$



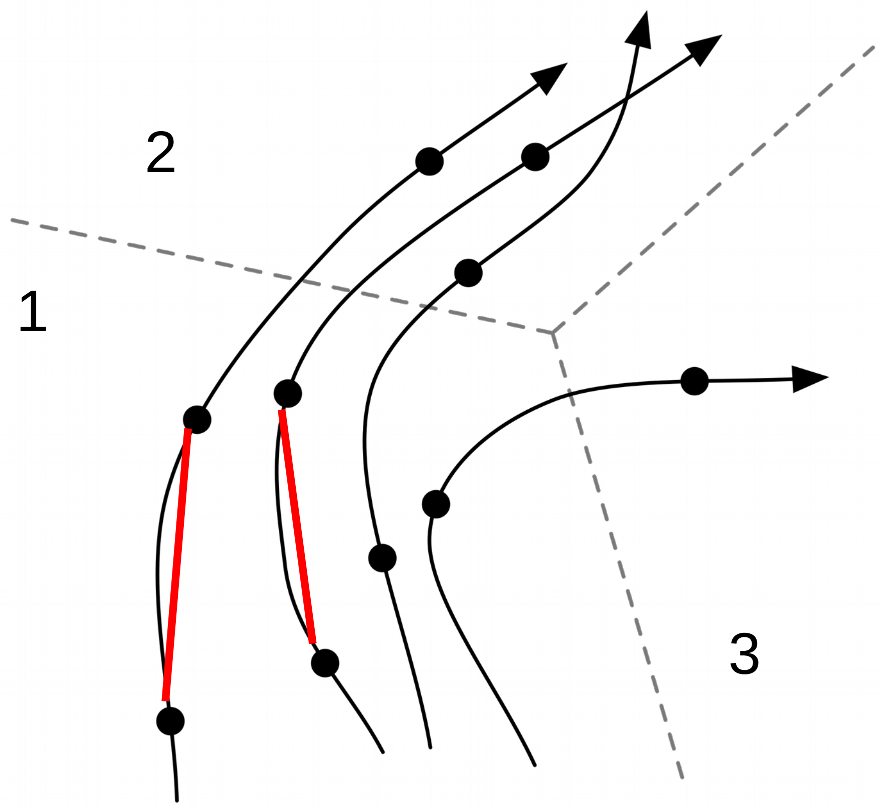
Transition matrix model

- Transition matrix: $P_{jk} = \dfrac{n_{jk}}{n_k}$

- Example:

$$P_{21} = \frac{3}{6}$$

$$P_{11} = \frac{2}{6}$$



Transition matrix model

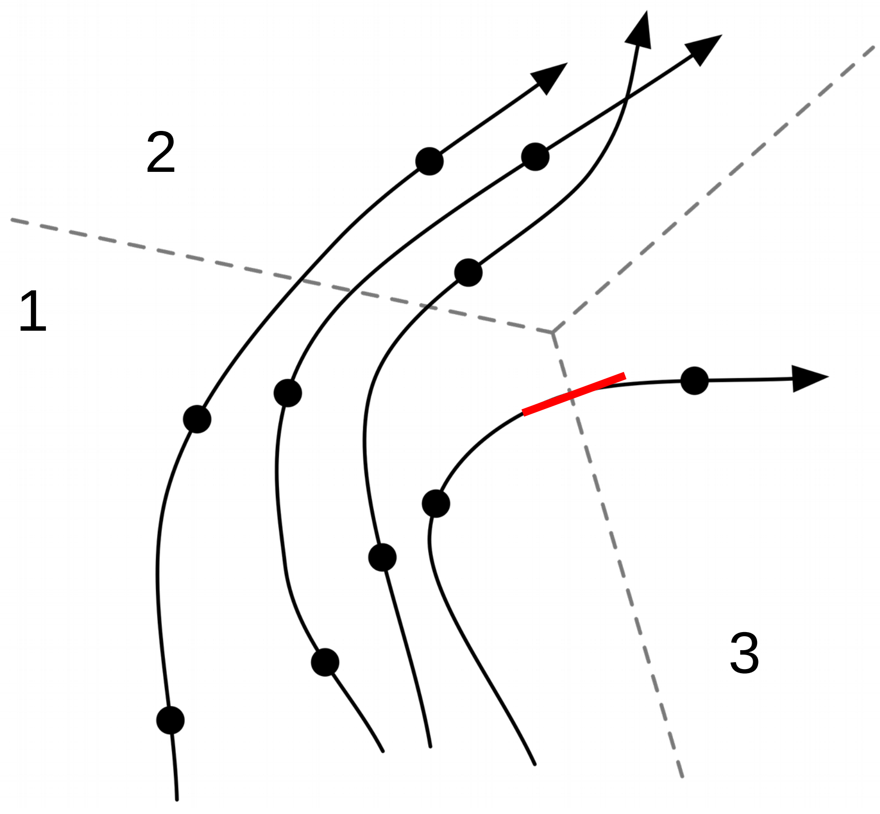
- Transition matrix: $P_{jk} = \dfrac{n_{jk}}{n_k}$

- Example:

$$P_{21} = \frac{3}{6}$$

$$P_{11} = \frac{2}{6}$$

$$P_{31} = \frac{1}{6}$$



Transition matrix model

- Transition matrix: $$P_{jk} = \frac{n_{jk}}{n_k}$$

- Example:

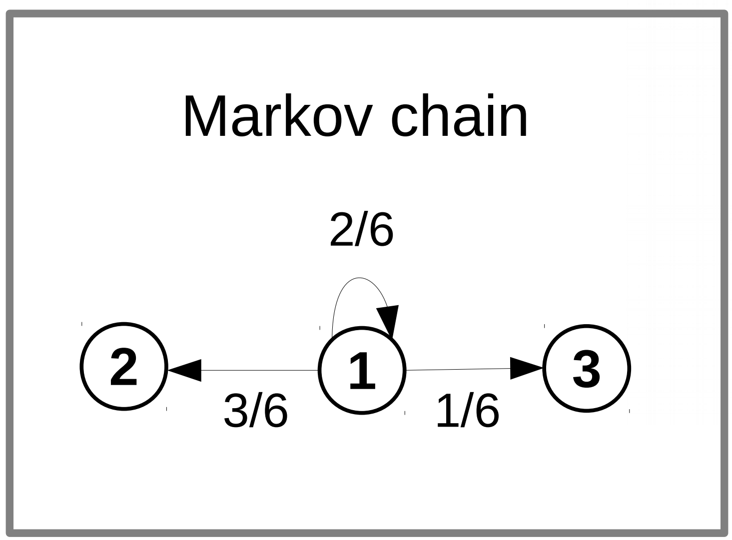
$$P_{21} = \frac{3}{6}$$

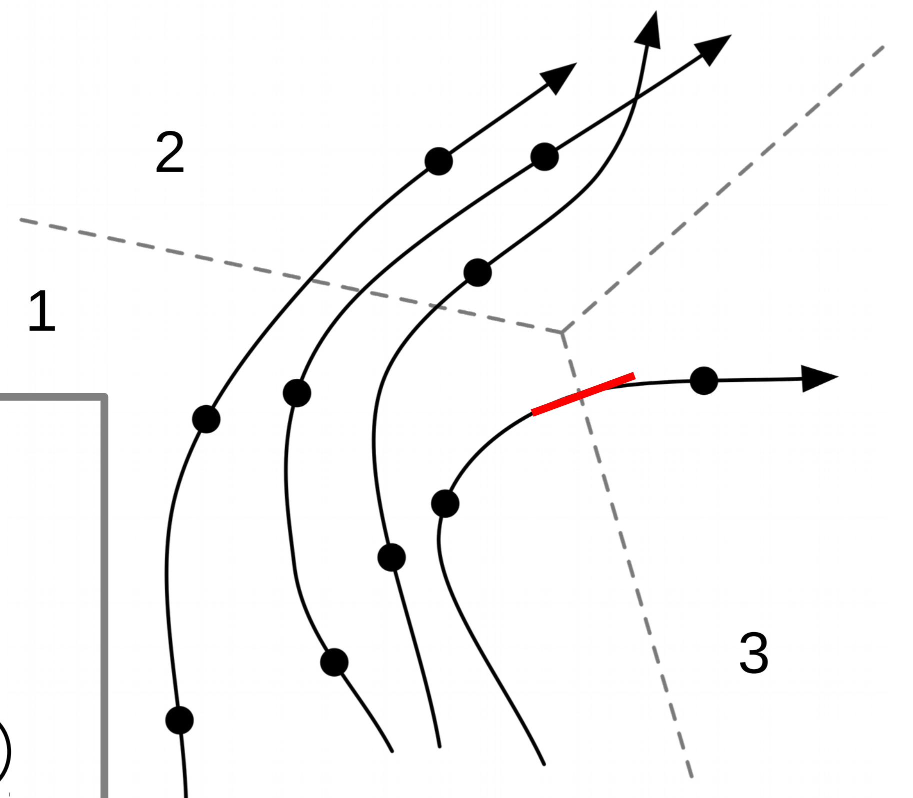
$$P_{11} = \frac{2}{6}$$

$$P_{31} = \frac{1}{6}$$

$$P_{j2} = 0$$

$$P_{j3} = 0$$





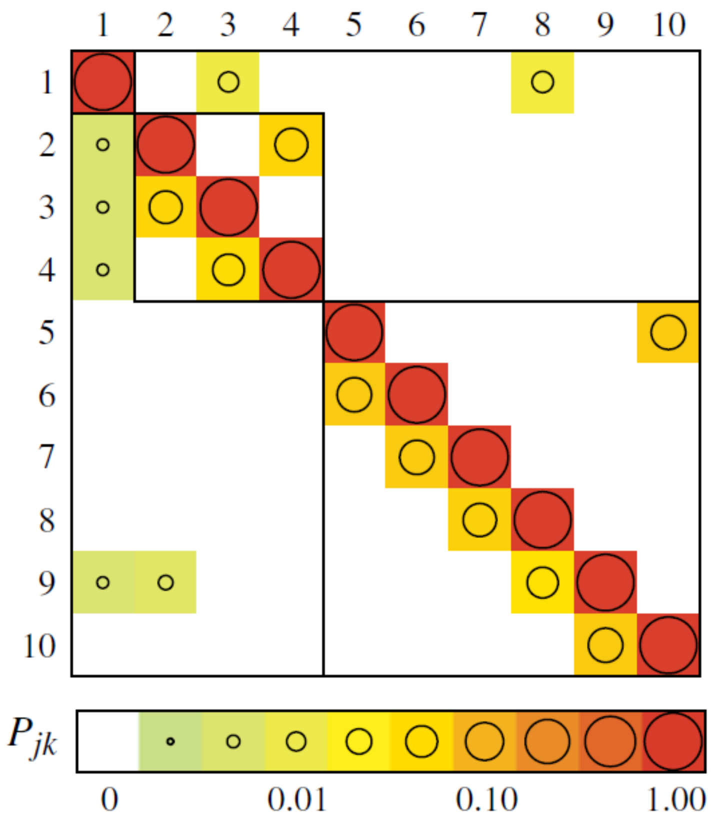
CROM results of a mixing layer

Cluster transition matrix

Transition probability

$$P_{jk} = Prob(c_j | c_k)$$

to move to c_j if the current state is c_k



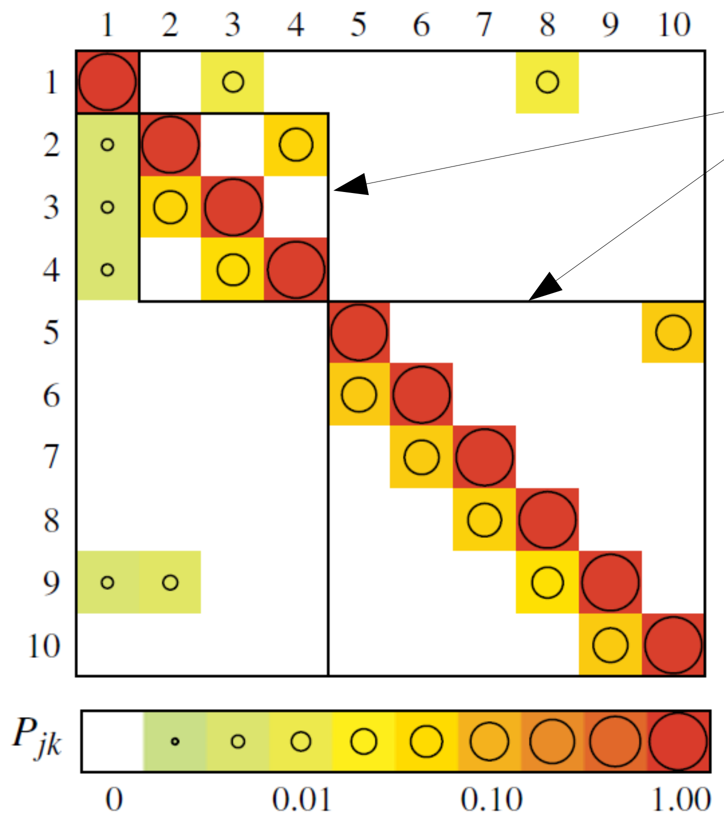
CROM results of a mixing layer

Cluster transition matrix

Transition probability

$$P_{jk} = Prob(c_j | c_k)$$

to move to c_j if the current state is c_k



Two cluster groups with oscillatory behavior

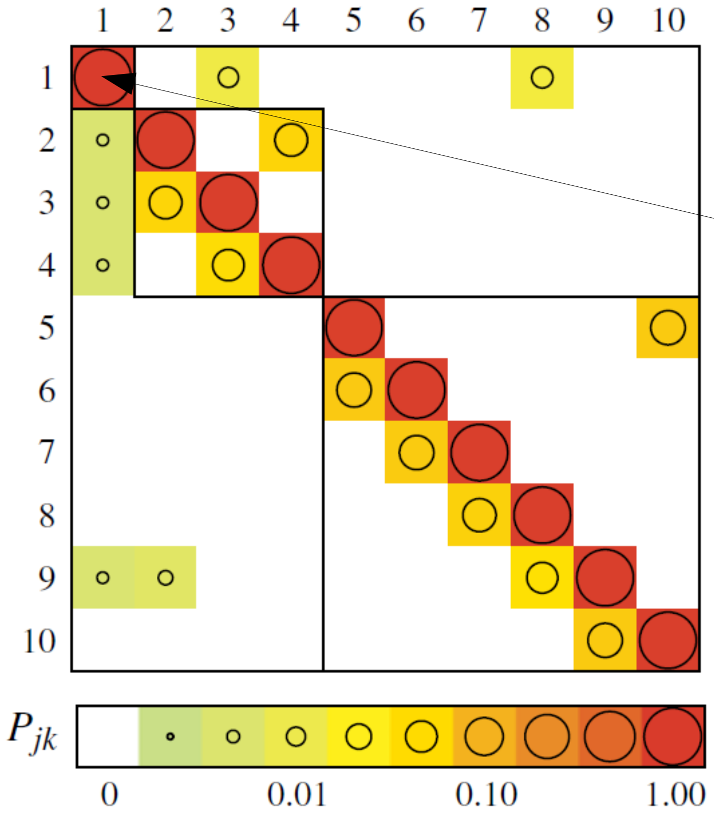
CROM results of a mixing layer

Cluster transition matrix

Transition probability

$$P_{jk} = Prob\left(c_j \middle| c_k\right)$$

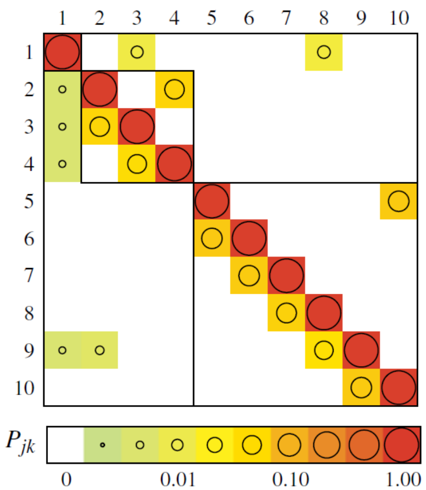
to move to c_j if the current state is c_k



Two cluster groups with oscillatory behavior.

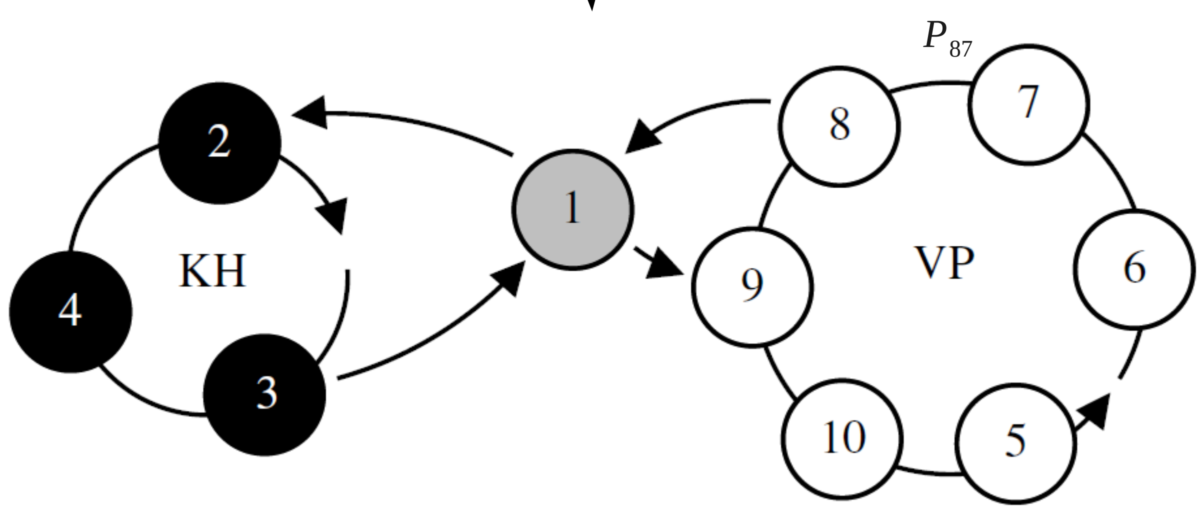
Flipper cluster c1 connects The two groups.

CROM of a mixing layer



Cluster transition matrix

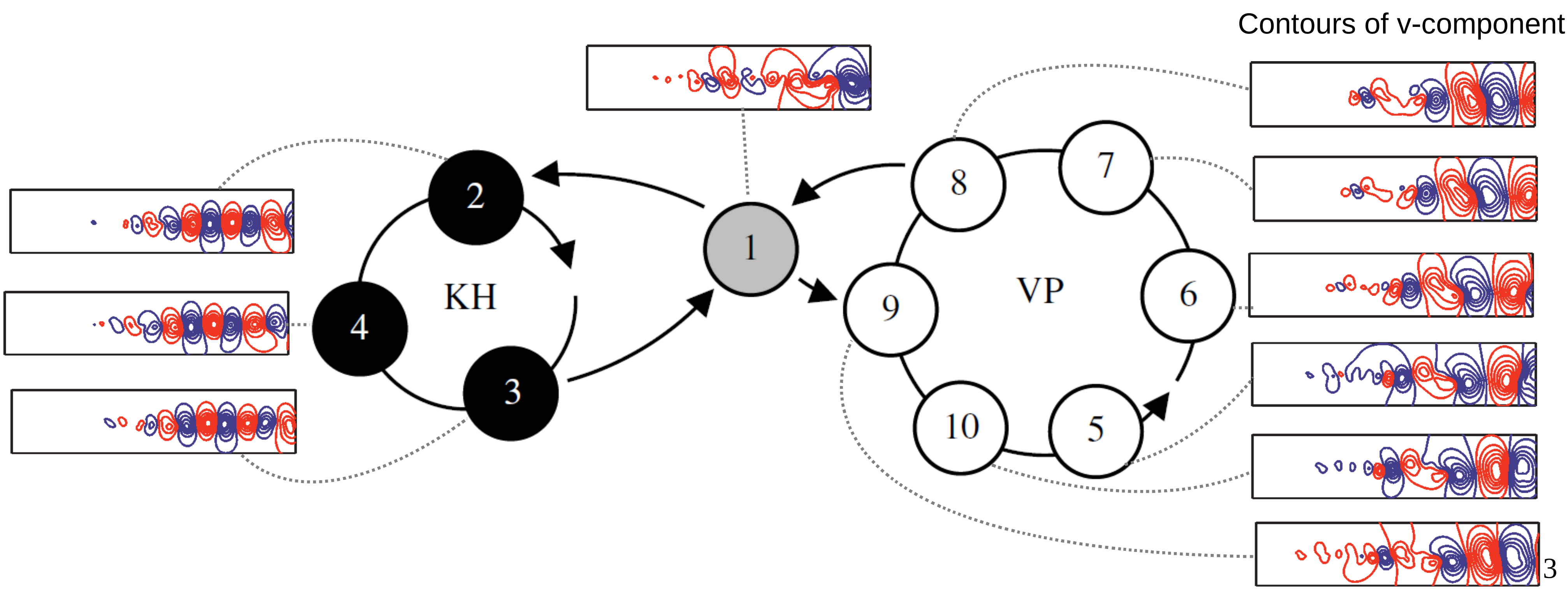
Simplified cluster transitions



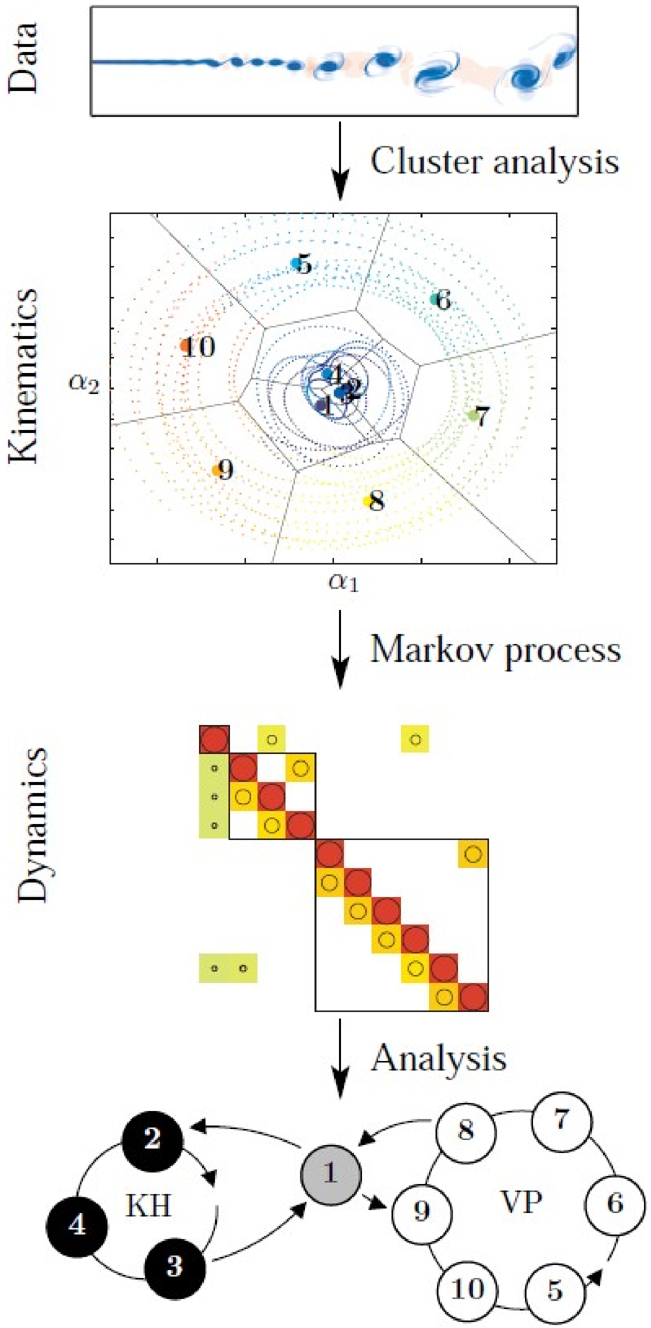
Markov chain:
Weighted directed graph

CROM of a mixing layer

- Most clusters are 'phase bins'
- Centroids are aligned with the dynamical evolution of the flow
- Identification of two shedding regimes
- Flipper cluster acts as a switch between both shedding regimes



Conclusions



- Data-driven approach to extract physical mechanisms in an unsupervised manner
- Some tuning parameters:
  - cluster number
  - time step of data
  - distance metric
- Cluster sociology
- Linear model taking into account nonlinear actuation dynamics

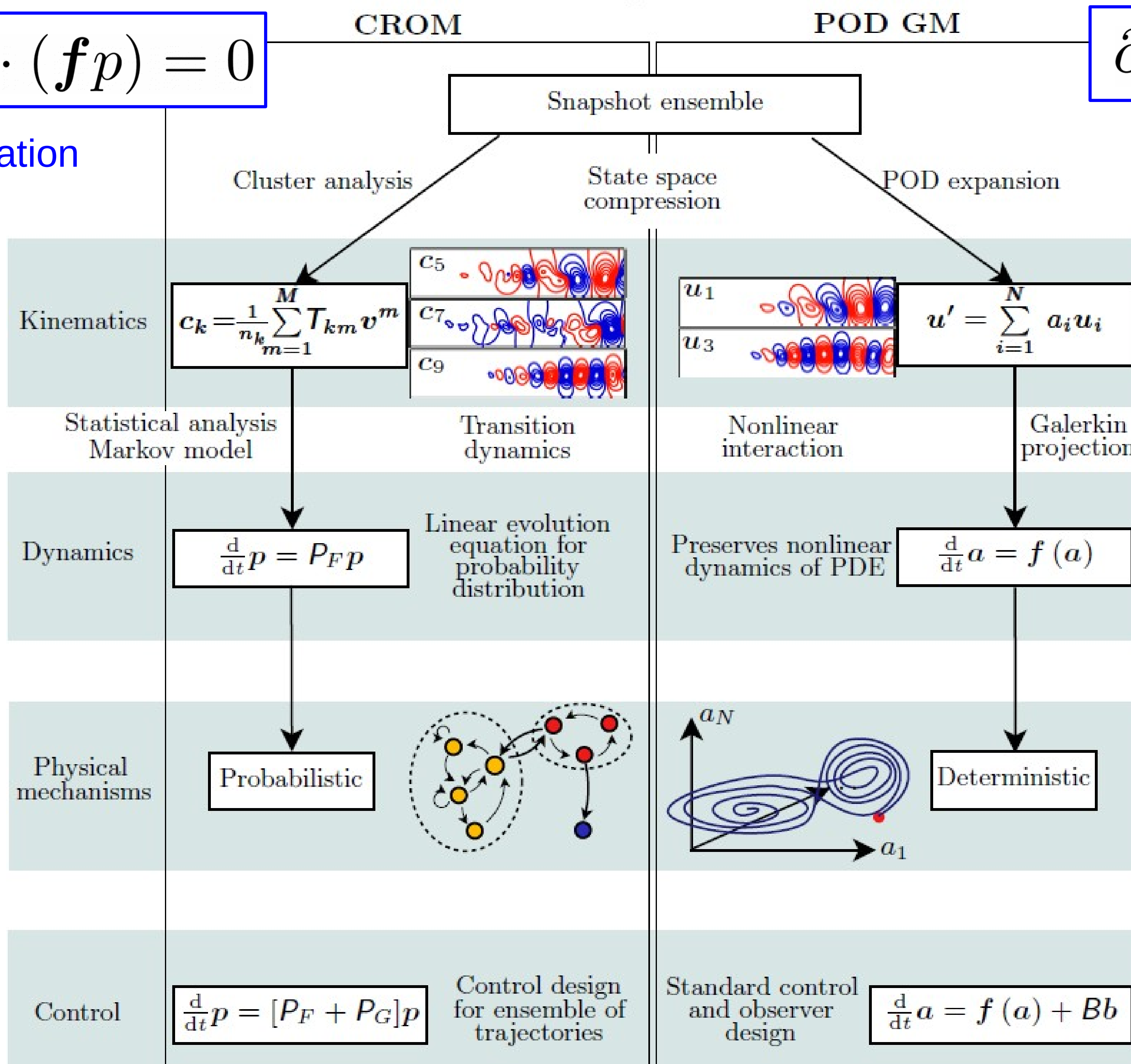
CROM vs. POD Galerkin models

$$\partial_t p + \nabla \cdot (\boldsymbol{f} p) = 0$$

Liouville equation

$$\partial_t \boldsymbol{u} = \boldsymbol{F}(\boldsymbol{u})$$

Navier-Stokes Eq.



Machine Learning
for Fluid Mechanics
Applications $\mapsto$Autoencoders



Bernd Noack

HIT, China & TU Berlin

Overview

Overview

A. Recursive DMD

. . . . *Combining the advantages of POD+DMD*

B. Feature-based manifold modeling

. . . . *The case for manifolds as opposed to POD*

C. Metric of attractor overlap

. . . . *Comparing attractor data*

Alternatives to POD

(1) **Dynamic Mode Decomposition (DMD)**

Rowley+ 2009 JFM, Schmid 2010 JFM

ID of stability modes near steady solution;
Fourier modes on attractor

(2) **Recursive DMD (rDMD)** Noack+ 2016 JFM

Fourier-like modes with low residual

(3) **Extended POD (EPOD)** Hoarau 2006 PF

Linearly links flow to sensor data

(4) **Spectral POD (SPOD)** Sieber 2016 JFM

Interpolation between POD and Fourier modes

(5) **Spectral POD (SPOD)** II Towne+ 2018 JFM

Space and time-dependent modes

(6) **Convective POD (CPOD)** 2018 AFCC Schulze

POD-like modes for convection

... **Many more for many more purposes**

POD—Recursive algorithm

Step 0 **Snapshot data**

$$\boldsymbol{u}^m(\boldsymbol{x}) := \boldsymbol{u}(\boldsymbol{x}, t^m),\ t^m = m\Delta t,\ m = 1, \ldots, M$$

Step 1 **Compute mean**

$$\boldsymbol{u}_0(\boldsymbol{x}) := \langle \boldsymbol{u}^m(\boldsymbol{x}) \rangle_M := \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{u}^m(\boldsymbol{x})$$

Step 2 **Compute fluctuation**

$$\boldsymbol{v}^m(\boldsymbol{x}) := \boldsymbol{u}^m(\boldsymbol{x}) - \boldsymbol{u}_0(\boldsymbol{x})$$

Step 3 **Compute POD mode** $\boldsymbol{u}_1(\boldsymbol{x})$

—most energetic direction

$$\left\langle \left| (\boldsymbol{v}^m(\boldsymbol{x}), \boldsymbol{u}_1(\boldsymbol{x}))_\Omega \right|^2 \right\rangle_M \stackrel{!}{=} \max$$

Step 4 **Remove orthogonal component** w.r.t $\boldsymbol{u}_1(\boldsymbol{x})$

$\boldsymbol{v}^m(\boldsymbol{x}) \leftarrow \boldsymbol{v}^m(\boldsymbol{x}) - a_1^m \boldsymbol{u}_1(\boldsymbol{x})$ where $a_1^m := (\boldsymbol{v}^m, \boldsymbol{u}_1)_\Omega$

Step 5 **Repeat Step 3 & 4** for the ith POD mode

$\boldsymbol{u}_i(\boldsymbol{x})$

$i = 1, \ldots, N$

DMD—Basic idea

Rowley+ 2009 JFM, Schmid 2010 JFM

Step 0 **Snapshot data**—Now time-resolved, i.e. with small Δt

$\boldsymbol{u}^m(\boldsymbol{x}) := \boldsymbol{u}(\boldsymbol{x}, t^m),\ t^m = m\Delta t,\ m = 1, \ldots, M$

Step 1 **Identify linear map $\boldsymbol{A}$**

$\boldsymbol{u}^{m+1} = \boldsymbol{A}\boldsymbol{u}^m, m = 1, \ldots, M-1$

Step 2 **Compute eigenmodes of $\boldsymbol{A}$**

—These are the DMD modes!

$\lambda_i \boldsymbol{u}_i = \boldsymbol{A}\boldsymbol{u}_i, i = 1, \ldots, M-1$

Comment 1 **Near steady solution/linear growth (CFD)**

$\Rightarrow$ **Some DMD modes are stability modes**

Comment 2 **On attractor / in experiment**

—statistically representative data

$\Rightarrow$ **All DMD modes become Fourier modes**

Recursive DMD (rDMD)—Algorithm

Noack, Stankiewicz, Morzyński & Schmid 2016 JFM

Step 0 **Snapshot data**—Now time resolved, small Δt

$$\boldsymbol{u}^m(\boldsymbol{x}) := \boldsymbol{u}(\boldsymbol{x}, t^m),\ t^m = m\Delta t,\ m = 1, \ldots, M$$

Step 1 **Compute mean**

$$\boldsymbol{u}_0(\boldsymbol{x}) := \langle \boldsymbol{u}^m(\boldsymbol{x}) \rangle_M := \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{u}^m(\boldsymbol{x})$$

Step 2 **Compute fluctuation**

$$\boldsymbol{v}^m(\boldsymbol{x}) := \boldsymbol{u}^m(\boldsymbol{x}) - \boldsymbol{u}_0(\boldsymbol{x})$$

Step 3 **Compute dominant DMD mode** $\boldsymbol{u}_1(\boldsymbol{x})$

From all DMD modes, take the one which resolves the largest fluctuation level.

Step 4 **Remove orthogonal component** w.r.t $\boldsymbol{u}_1(\boldsymbol{x})$

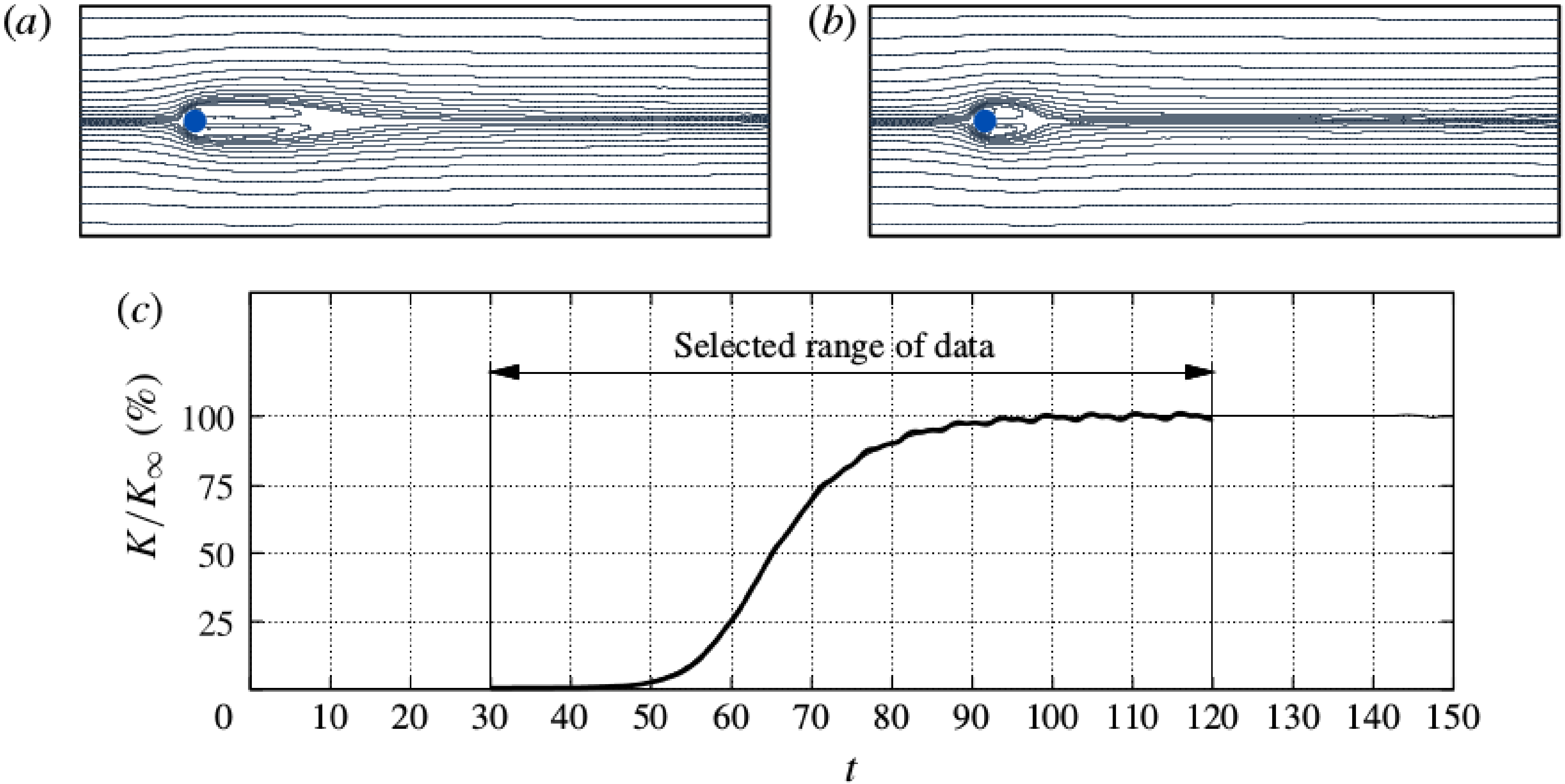
$\boldsymbol{v}^m(\boldsymbol{x}) \leftarrow \boldsymbol{v}^m(\boldsymbol{x}) - a_1^m \boldsymbol{u}_1(\boldsymbol{x})$ where $a_1^m := (\boldsymbol{v}^m, \boldsymbol{u}_1)_\Omega$

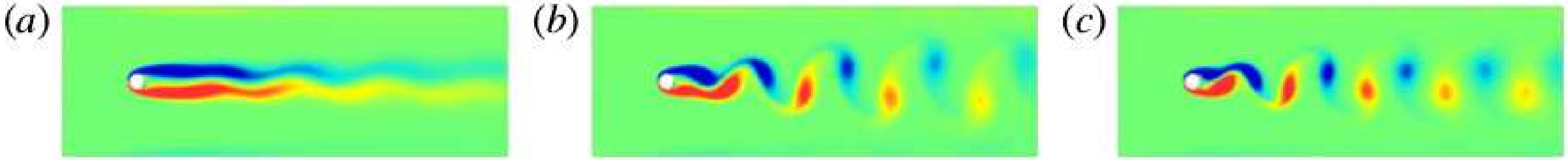
Step 5 **Repeat Step 3, 4** for ith rDMD mode $\boldsymbol{u}_i(\boldsymbol{x})$

$i = 2, \ldots, N$

Recursive DMD—Wake transient

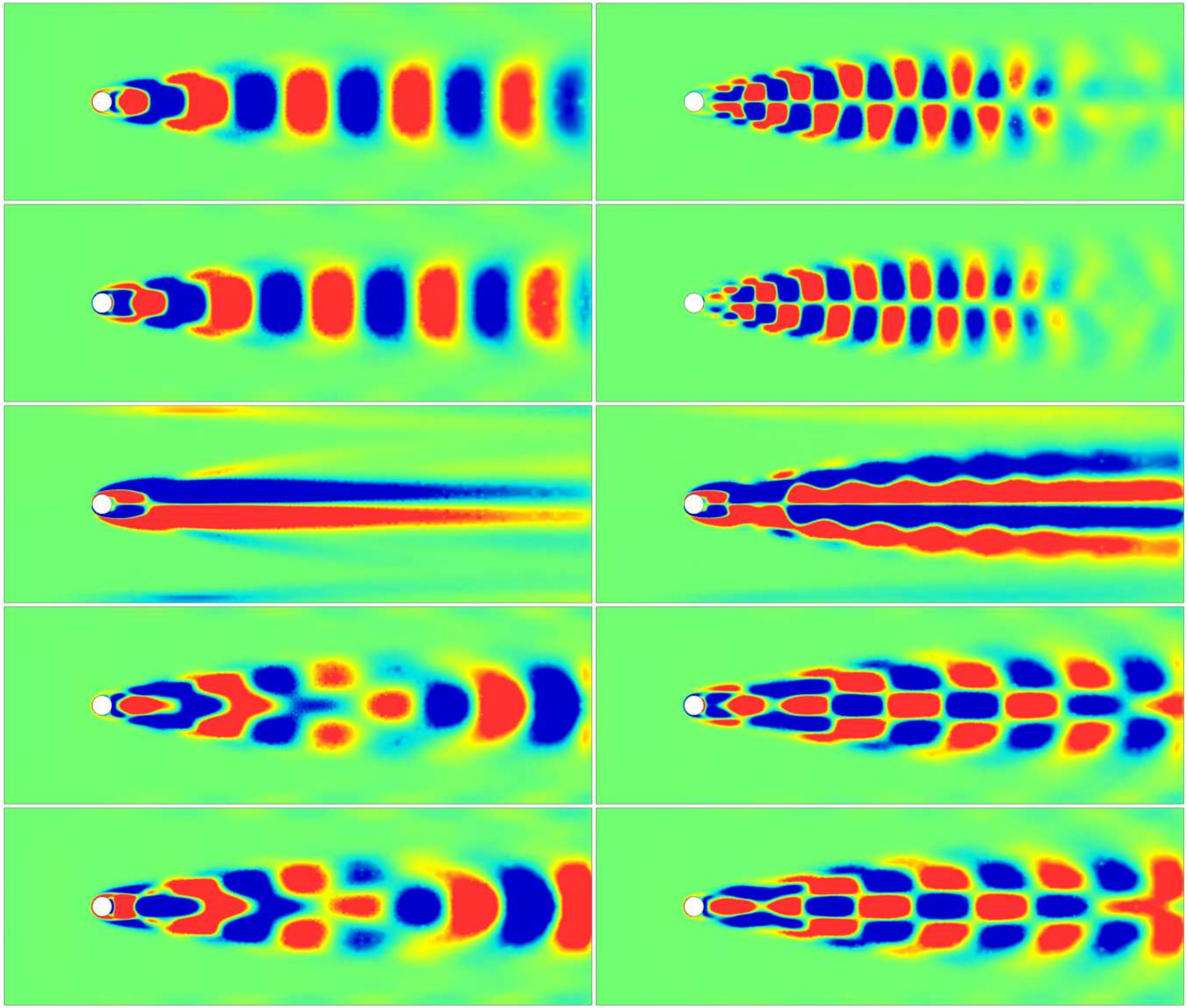
Noack, Stankiewitz, Morzyński & Schmid (2016) JFM





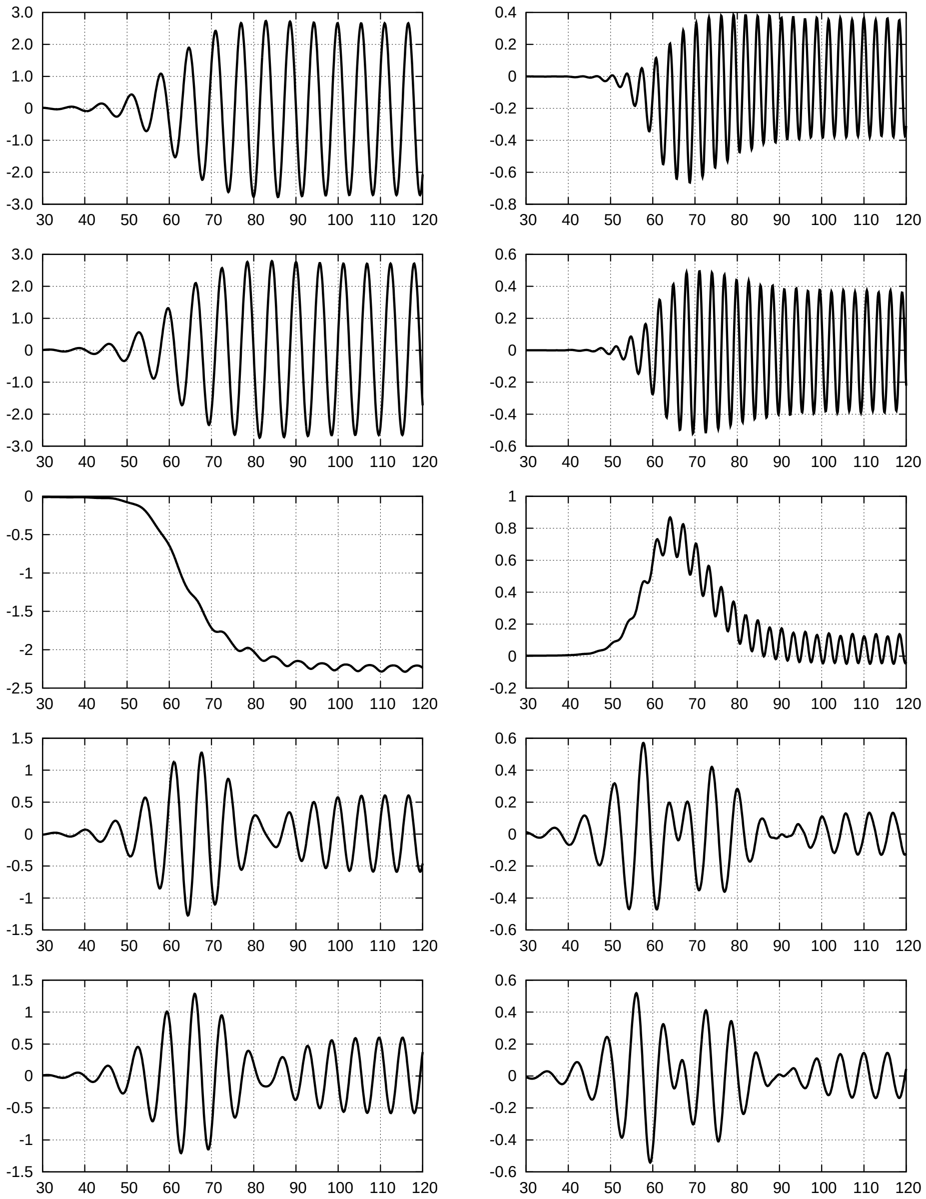
POD modes

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



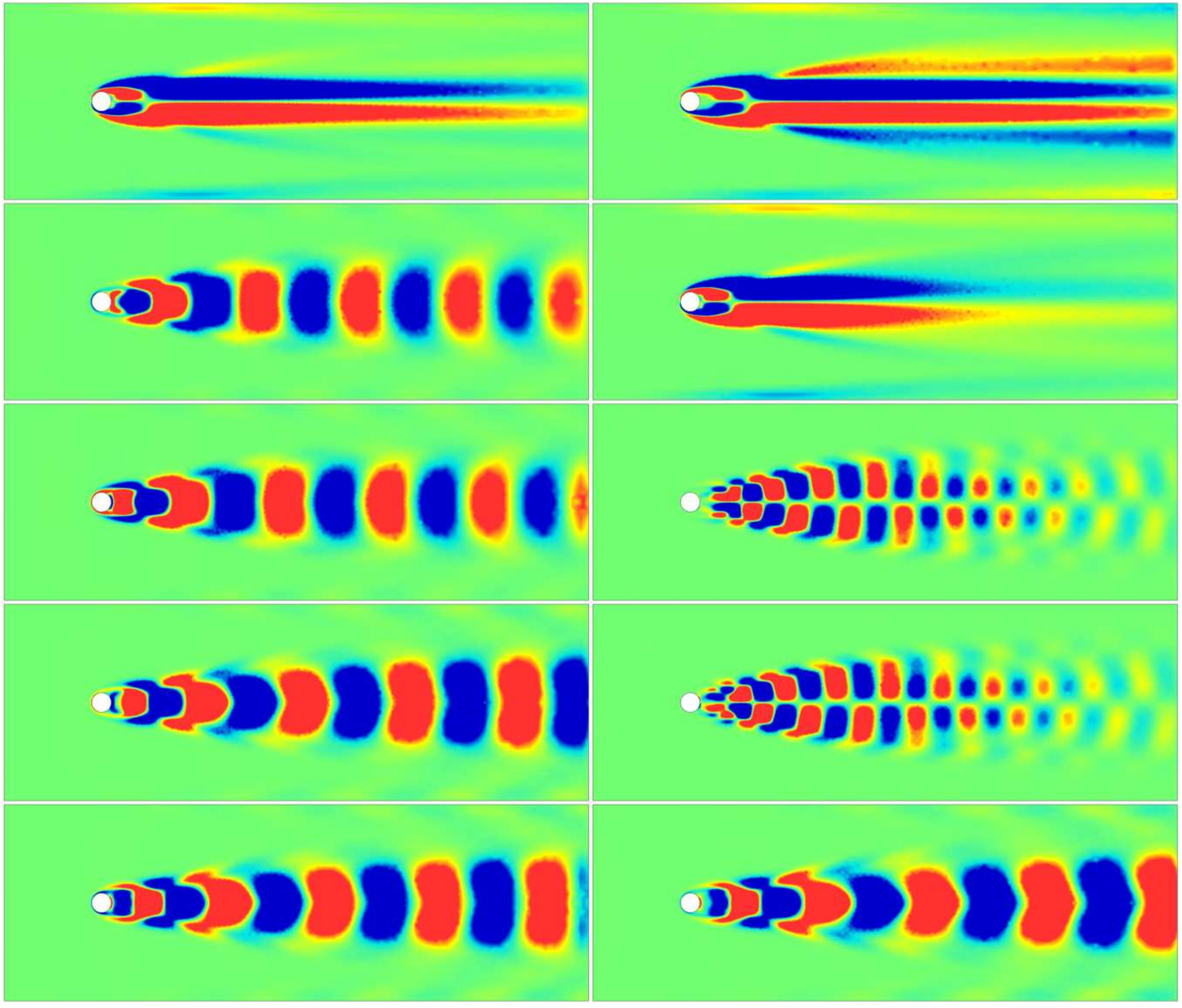
POD amplitudes

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



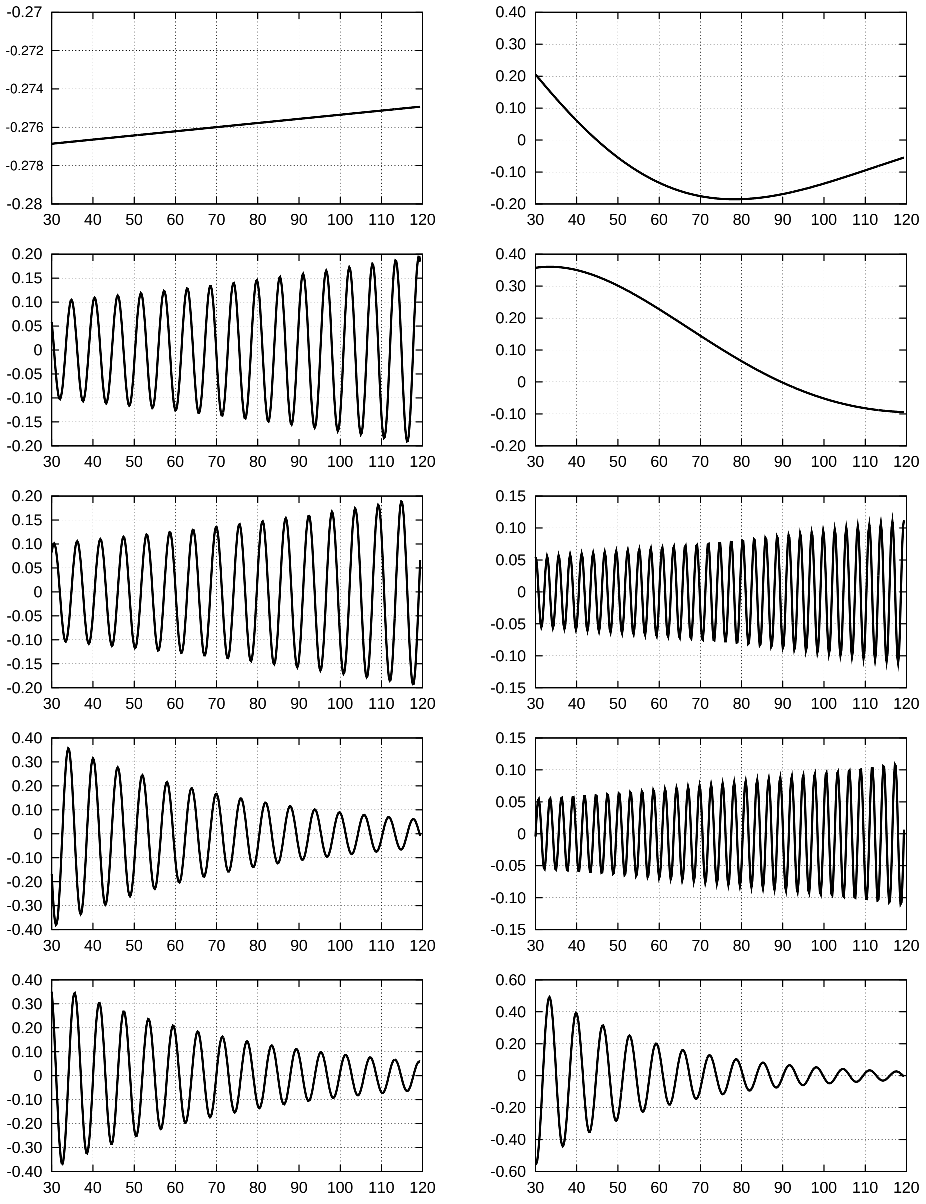
DMD modes

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



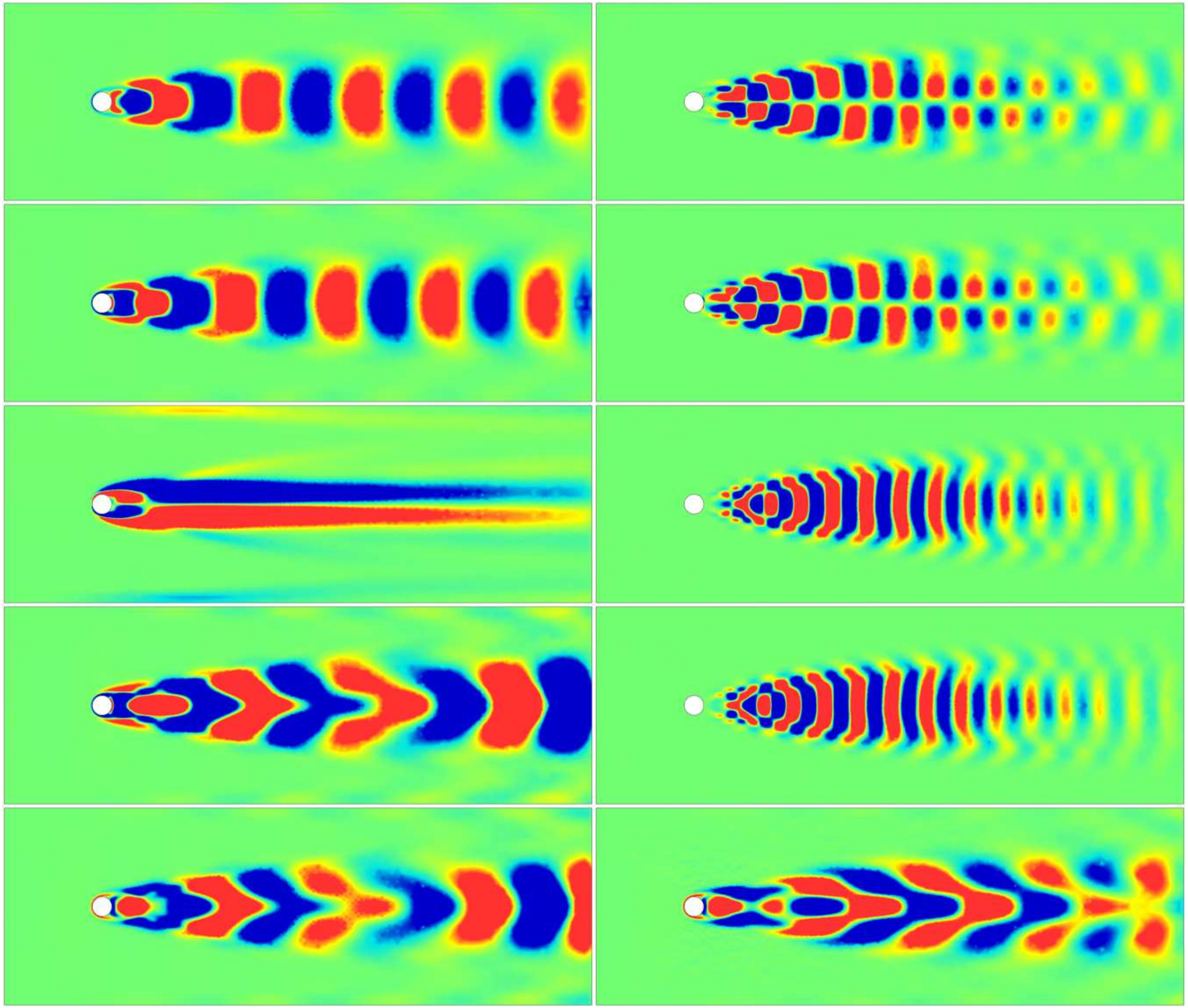
DMD amplitudes

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



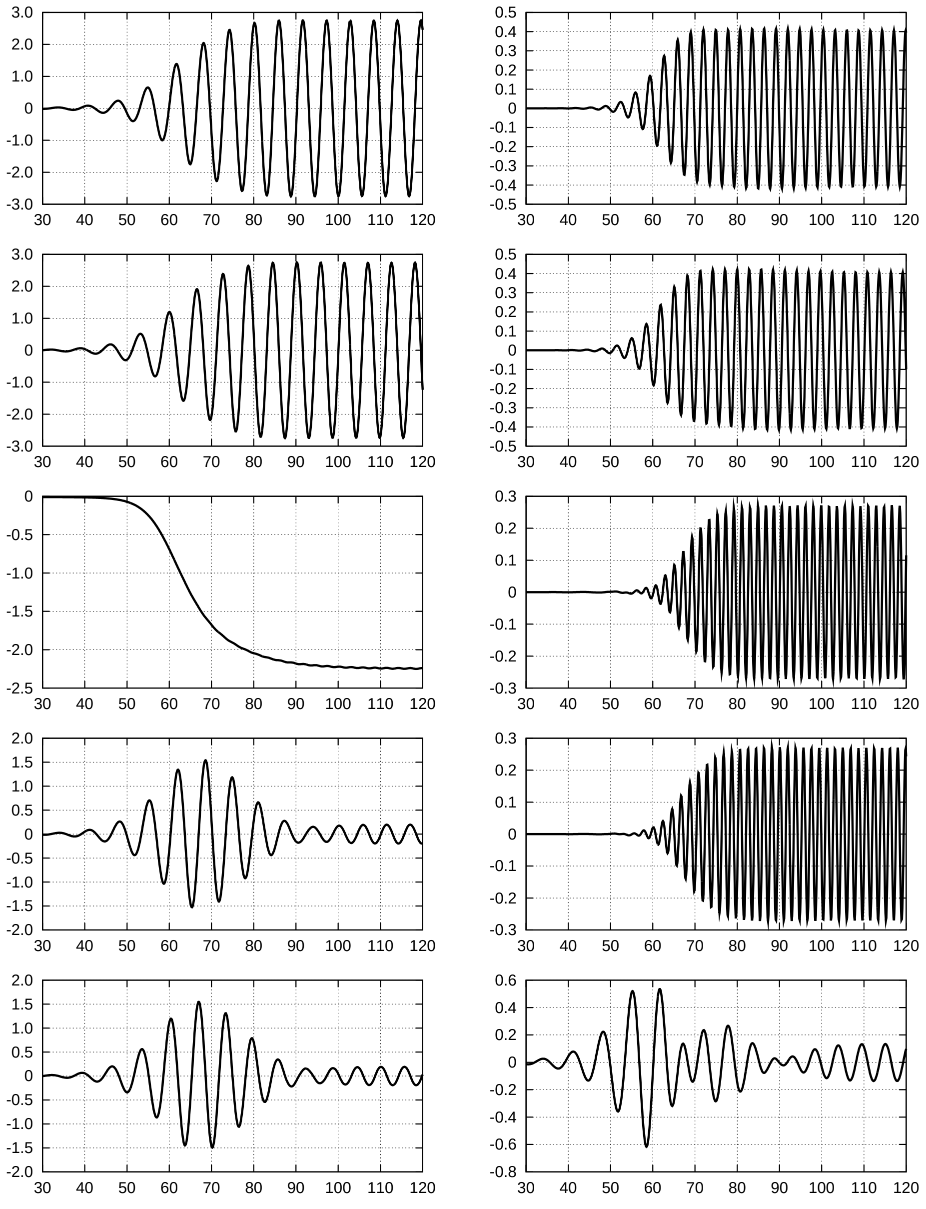
Recursive DMD—Modes

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



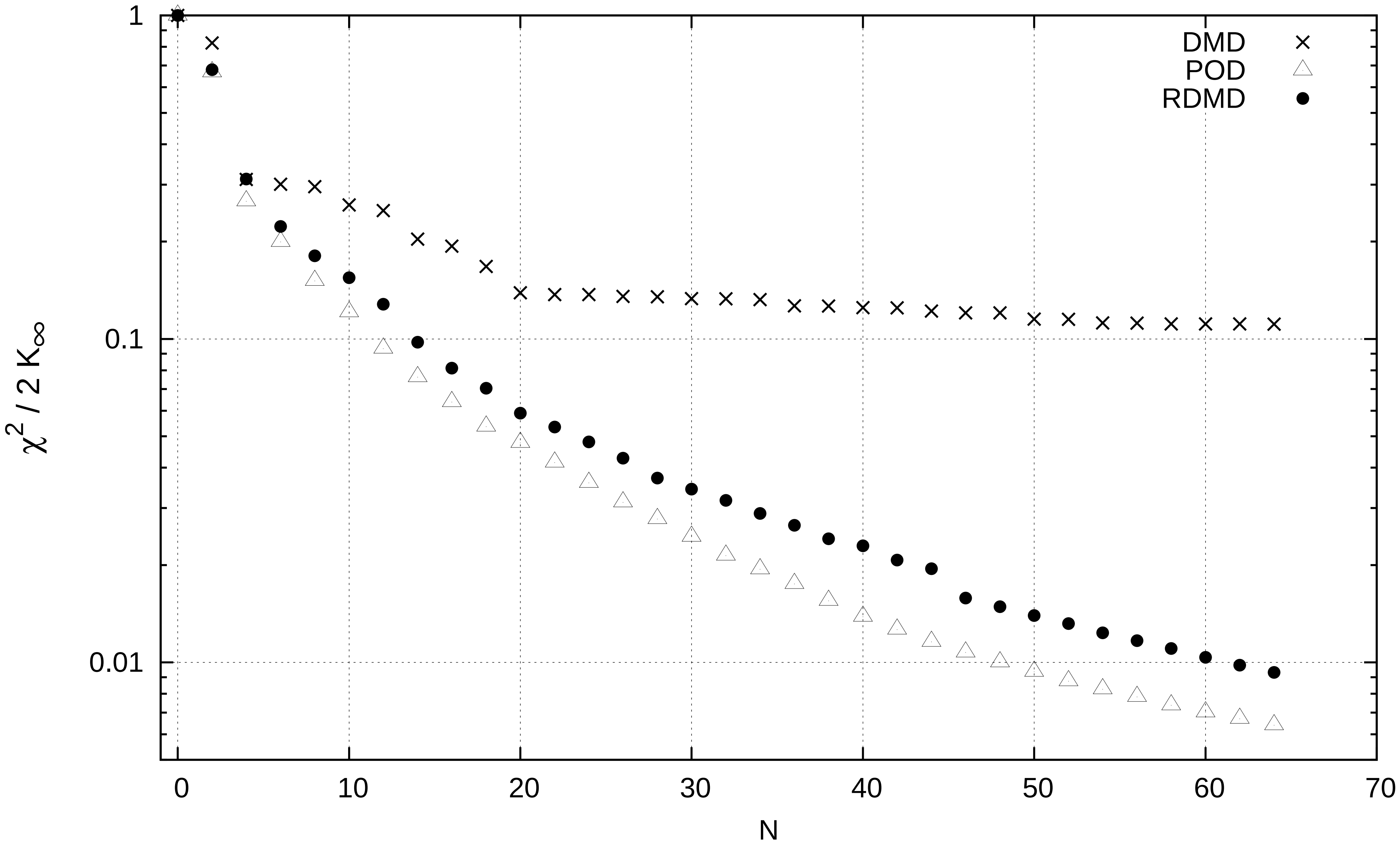
Recursive DMD—Amplitudes

≡ *Noack, Stankiewitz, Morzyński & Schmid (2016) JFM*



Recursive DMD—Residuals

Noack, Stankiewitz, Morzyński & Schmid (2016) JFM



Overview

A. Recursive DMD

. . . . .Combining the advantages of POD+DMD

B. Feature-based manifold modeling

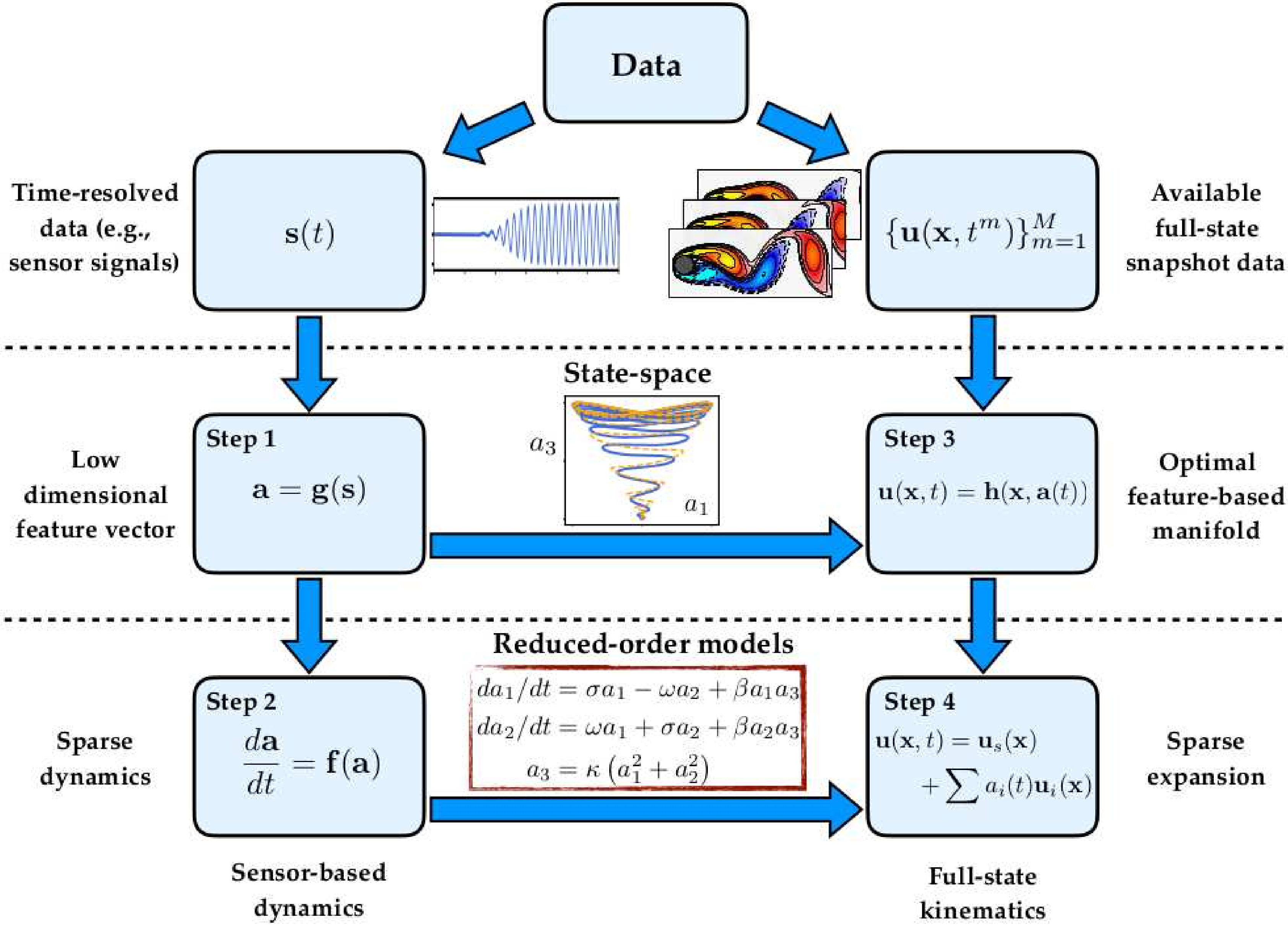
. . . . .The case for manifolds as opposed to POD

C. Metric of attractor overlap

. . . . . .Comparing attractor data

Feature-based Manifold Model (FeMM)

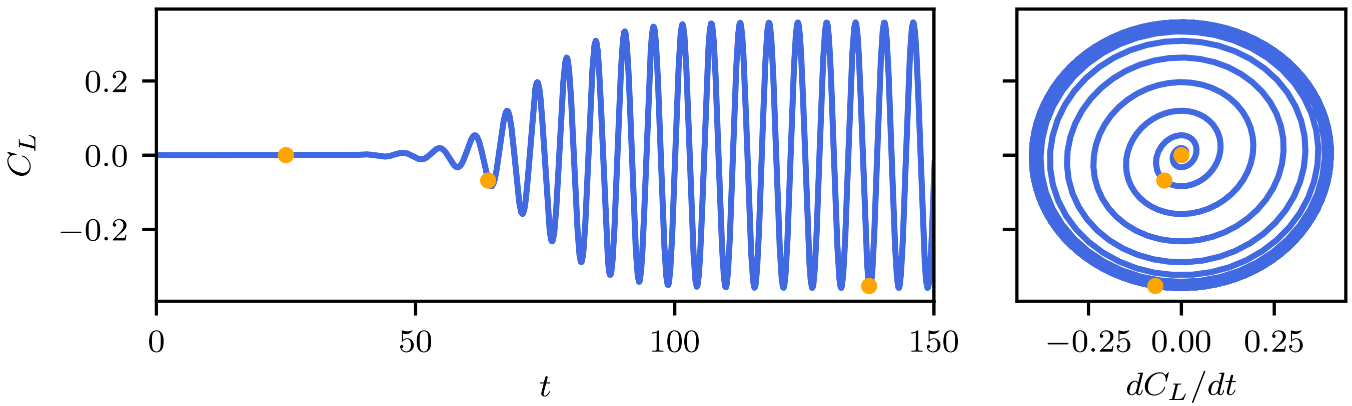
J.C. Loiseau, B.R. Noack & S.L. Brunton 2018 JFM



2D Cylinder wake transient ($Re = 100$)

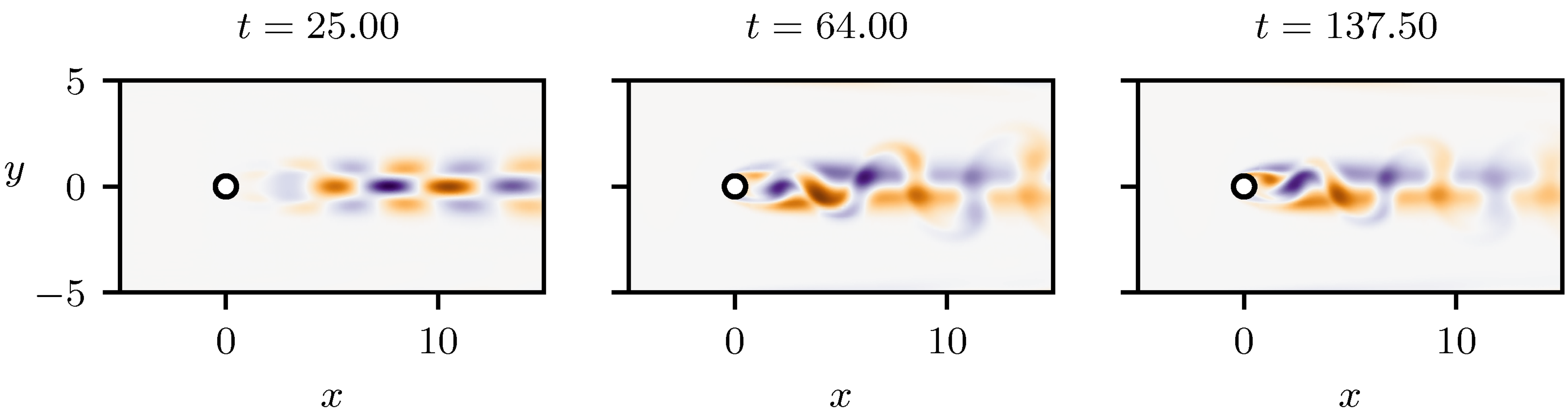
J.C. Loiseau, B.R. Noack & S.L. Brunton 2018 JFM

(1) Sensor signal $s = c_L$



(a) (b)

(2) Velocity field $\boldsymbol{u}(\boldsymbol{x}, t)$ (Visualization of vorticity fluctuation)

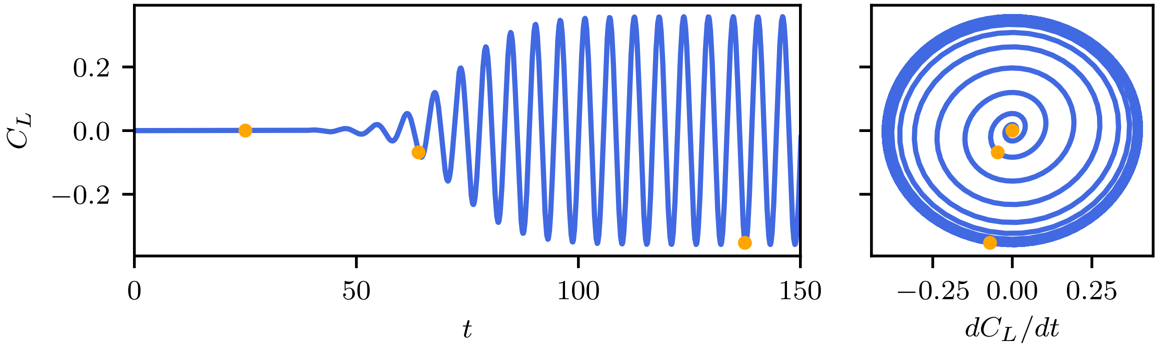


(a) (b) (c)

Feature-based Manifold Model (FeMM)

J.C. Loiseau, B.R. Noack & S.L. Brunton 2018 JFM

2D Cylinder wake transient steady$\mapsto$periodic



(a) (b)

(1) Feature vector $\boldsymbol{a} = \begin{bmatrix} c_L \\ (dc_L/dt)/\omega_\infty \end{bmatrix}$

(2) SINDy Brunton et al. 2016 PNAS . . . . *Black-box model*

$$\frac{\mathrm{d}a_1}{\mathrm{d}t} = 1.12 a_2$$

$$\frac{\mathrm{d}a_2}{\mathrm{d}t} = -1.12 a_1 + 0.2(1 - a_1^2 - a_2^2) a_2.$$

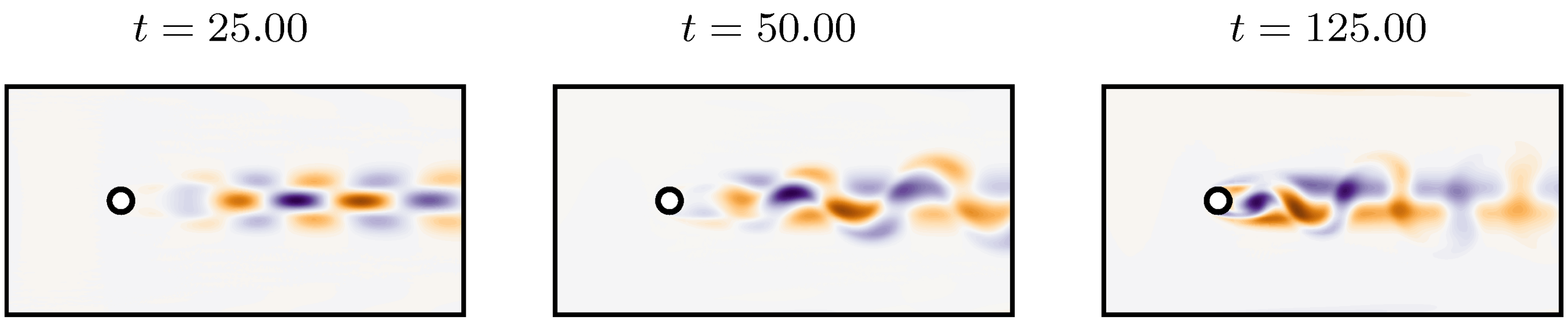
(3) Local mapping $\boldsymbol{a} \mapsto \boldsymbol{u}(\boldsymbol{x})$ . . . . *Gray-box model with (2)*

(4) EXTRA $\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_s(\boldsymbol{x}) + \sum_i a_i(t) \boldsymbol{u}_i(\boldsymbol{x})$ *Galerkin model*

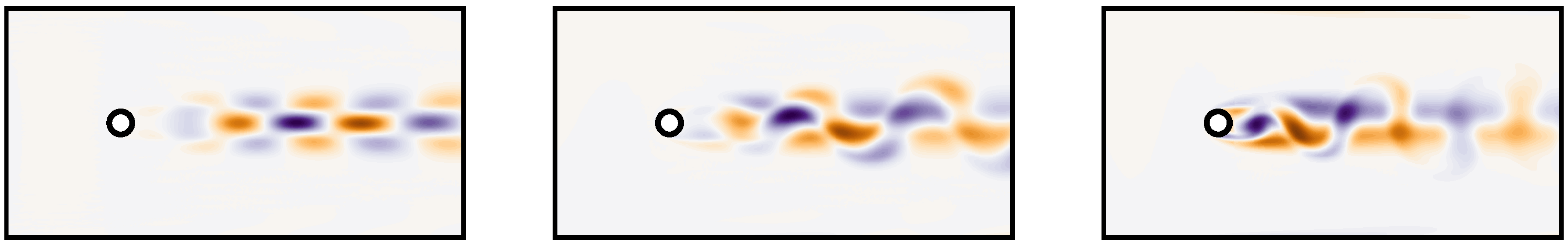
Feature-based Manifold Model II

J.C. Loiseau, B.R. Noack & S.L. Brunton 2018 JFM

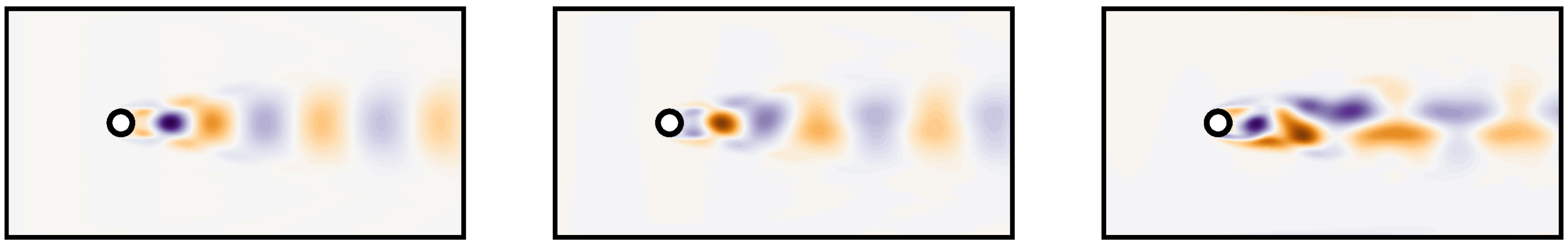
2D Cylinder wake transient steady$\mapsto$periodic



(a) Direct numerical simulation



(b) Local linear mapping

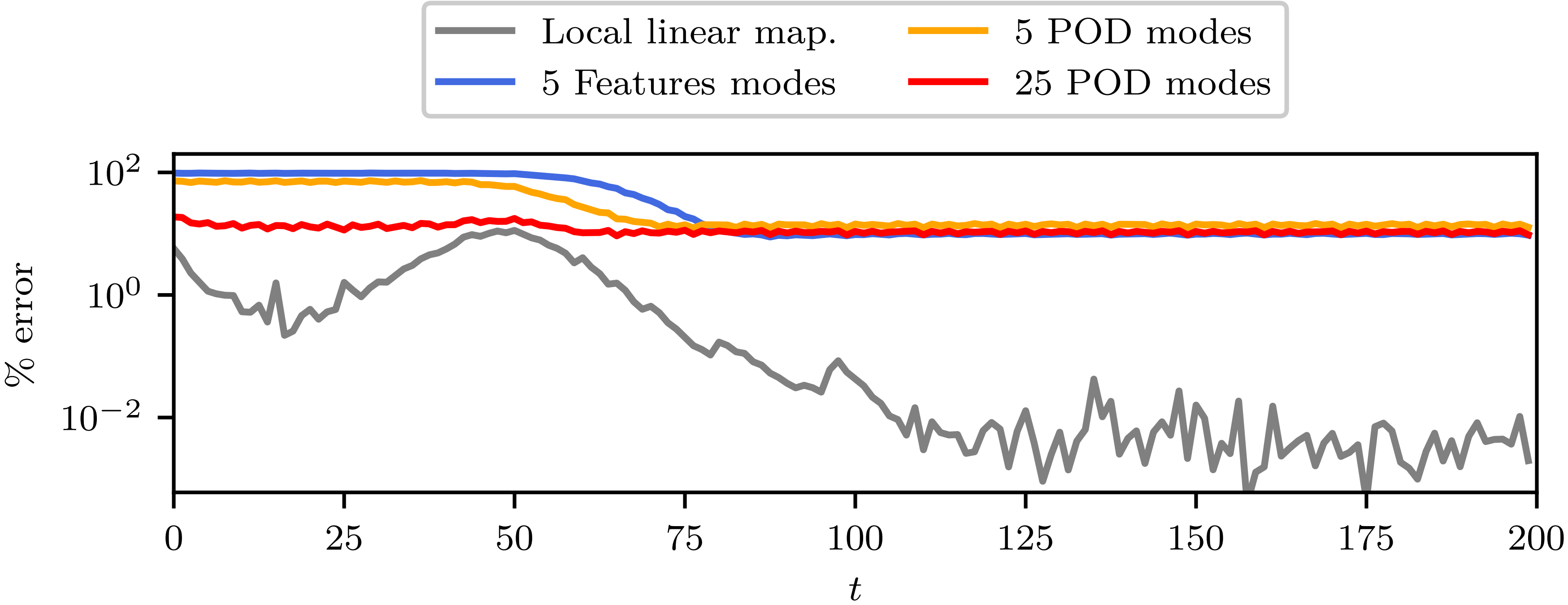


(c) Galerkin expansion

Feature-based Manifold Model vs POD

J.C. Loiseau, B.R. Noack & S.L. Brunton 2018 JFM

2D Cylinder wake transient steady$\mapsto$periodic



The 25 POD mode expansion is orders of magnitude worse than the 2-dimensional feature-based manifold!

Overview

A. Recursive DMD

. . . . *Combining the advantages of POD+DMD*

B. Feature-based manifold modeling

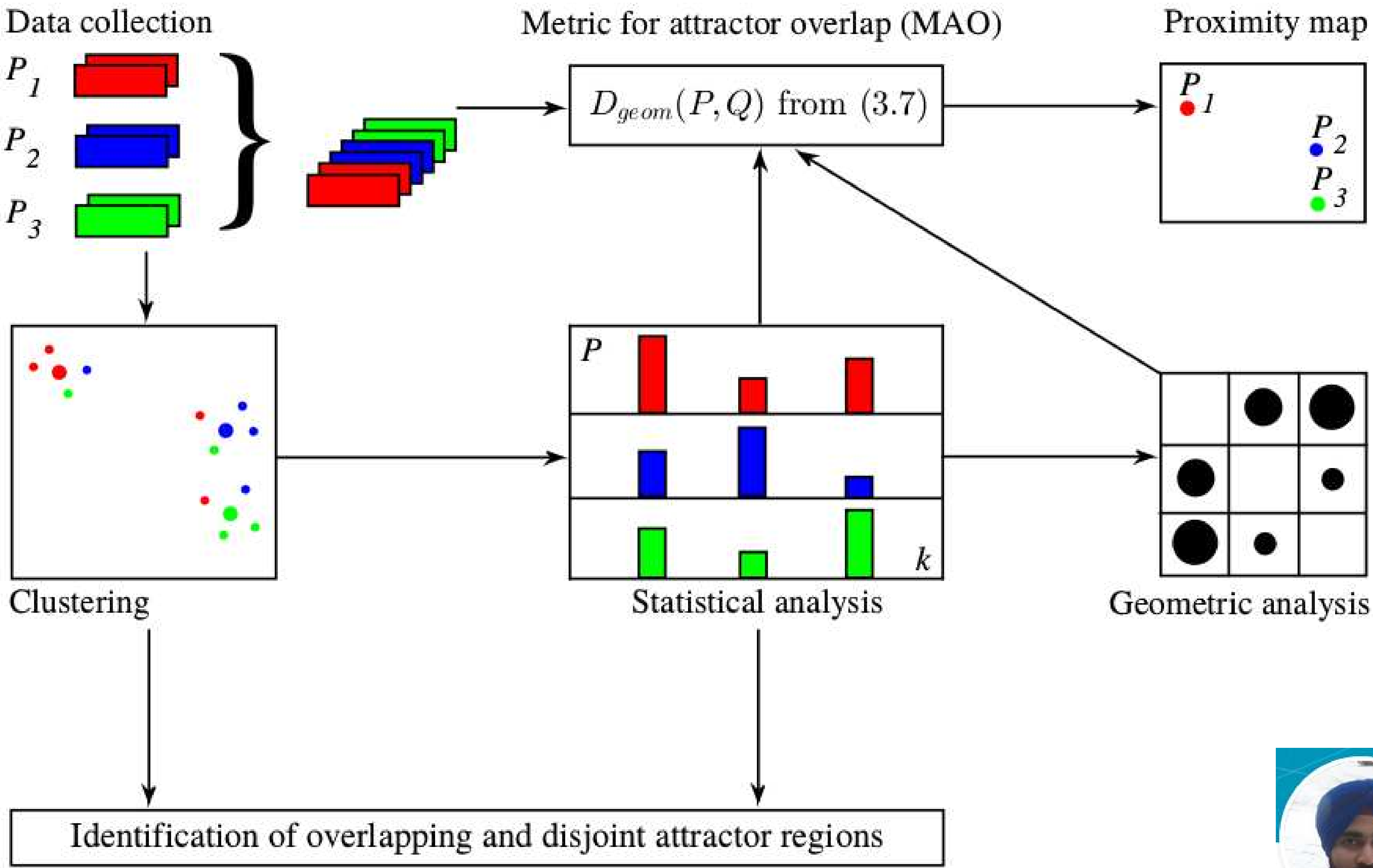
. . . . *The case for manifolds as opposed to POD*

C. Metric of attractor overlap

. . . . *Comparing attractor data*

Metric of attractor overlap (MAO)

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2018 JFM (in revision)



Metric of attractor overlap (MAO)

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2018 JFM (in revision)

(0) Snapshots of attractor $\mathcal{A}, \mathcal{B}$: $\boldsymbol{u}^m, m = 1, \ldots, M$

$$\chi_A^m = \begin{cases} 1 & \text{for } \boldsymbol{u}^m \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases} \qquad M_A := \sum_{m=1}^{M} \chi_A^m \qquad \text{analog var. for } \mathcal{B}$$

(1) Snapshot-based metric:

$$D(\mathcal{A}, \mathcal{B}) = \frac{1}{M_A} \sum \chi_A^m \, D\left(\boldsymbol{u}^m, \mathcal{B}\right) + \frac{1}{M_B} \sum \chi_B^m \, D\left(\boldsymbol{u}^m, \mathcal{A}\right)$$

(2) Clustering: M snapshots $\boldsymbol{u}^m$ in K centroids $\boldsymbol{c}_k$;

$\xi_k^m = 1 \leftrightarrow \boldsymbol{u}^m$ belongs to $\boldsymbol{c}_k$; otherwise $\xi_k^m = 0$.

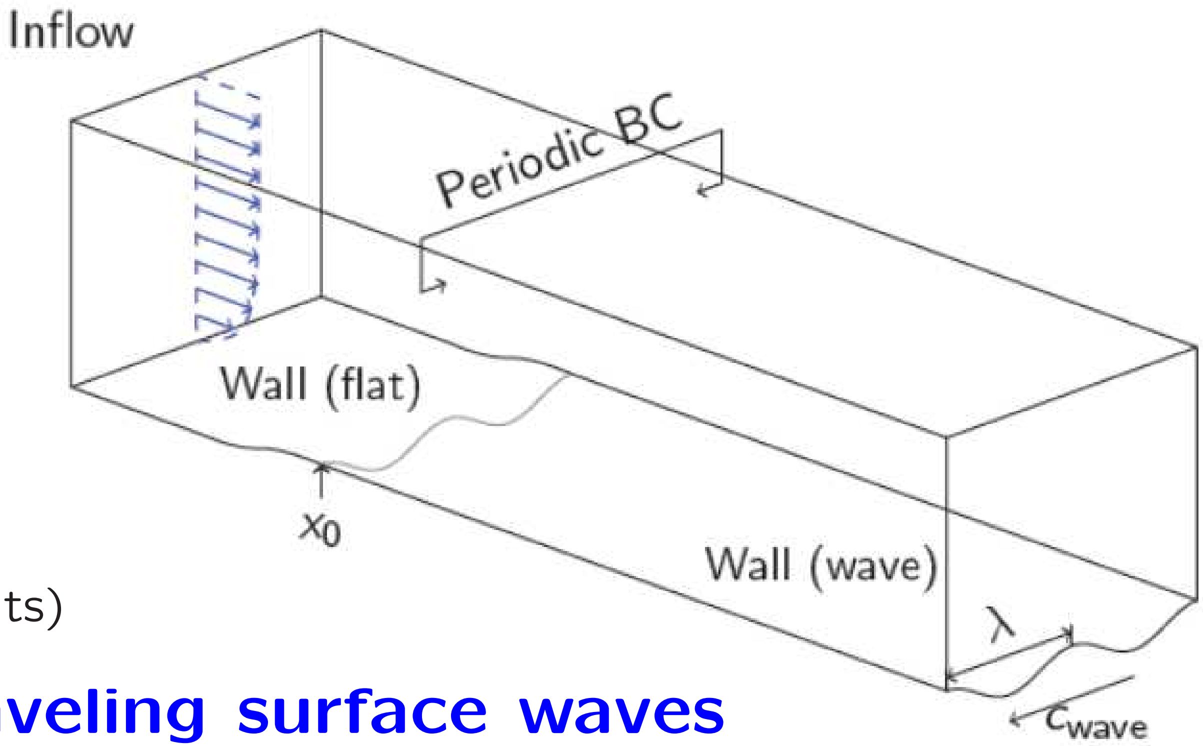
(3) Overlap clusters $\mathcal{A} \cap \mathcal{B}$: $\xi_k^m = \chi_A^m = \chi_B^m = 1$

(4) Disjoint clusters $\mathcal{A}$, $\mathcal{B}$: $\xi_k^m = 1$ and $\chi_A^m + \chi_B^m = 1$

(5) Proximity map for many attractos: $\mathcal{A}^l \mapsto \boldsymbol{\gamma}^l \in R^2$.

Drag reduction LES data—Made in Aachen

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2019 JFM



$Re_\theta = 1000$

$Ma = 0.2$

Analysis domain

- $x/\delta_\theta \in [50, 100]$,
- $y/\delta_\theta \in [0, 15]$
- $z/\delta_\theta \in [0, 21]$ (1000+ units)

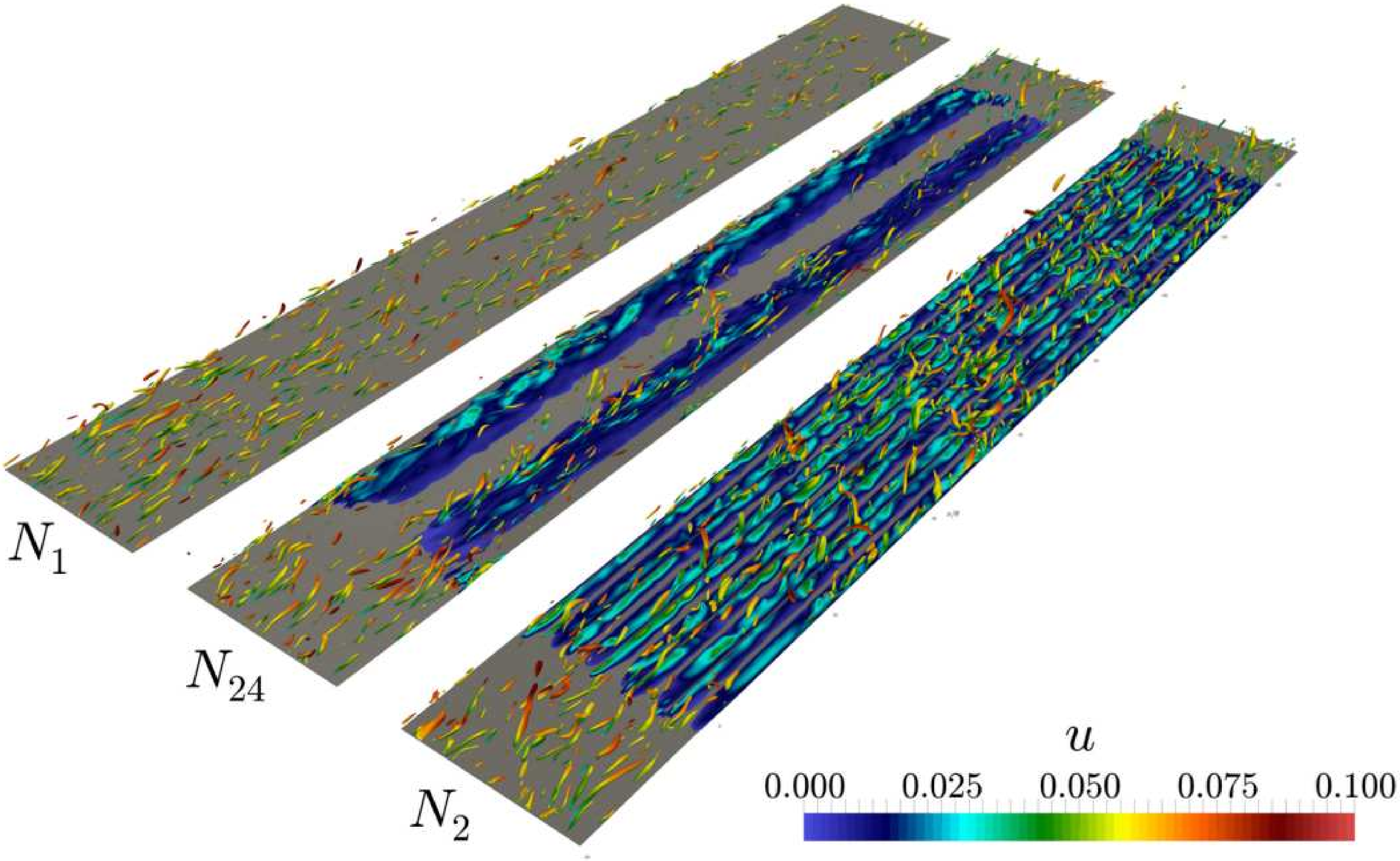
80 data sets with traveling surface waves

$$y = A \cos 2\pi \left[\frac{t}{T} - \frac{z}{\lambda_z}\right]$$

- $A^+ \in [0, 78]$
- $T^+ \in [20, 120]$
- $\lambda^+ \in [200, 3000]$

LESdrag reduction study

≡ *R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2019 JFM*



Contour plot of the λ_2 -criterion (Jeong & Hussain 1995), coloured by the instantaneous streamwise velocity, for three turbulent boundary layer flows; non-actuated reference case N_1 , actuated highest drag reduction case N_{24} , actuated lowest drag reduction case N_2.

LES drag reduction study

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2019 JFM

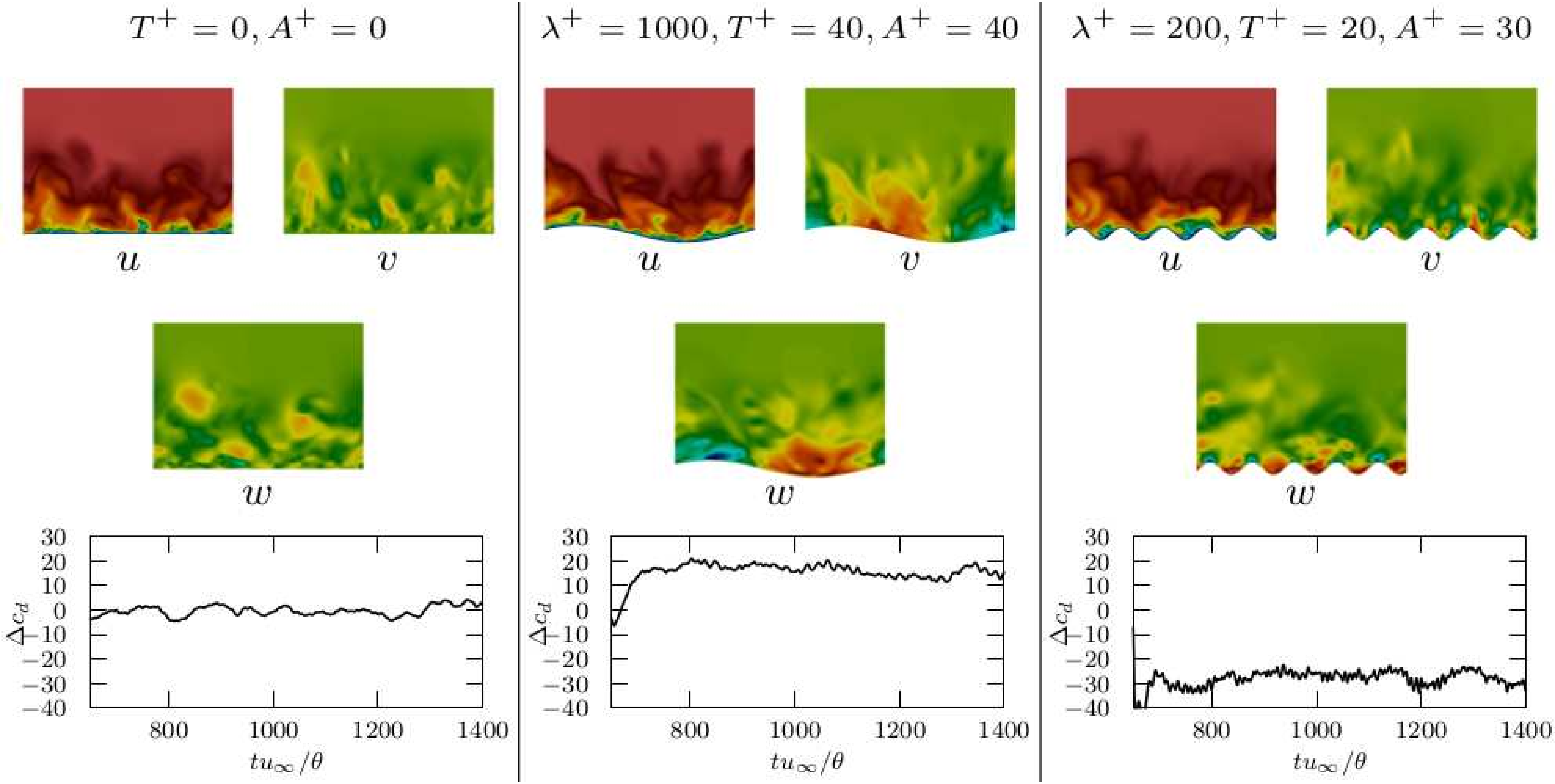
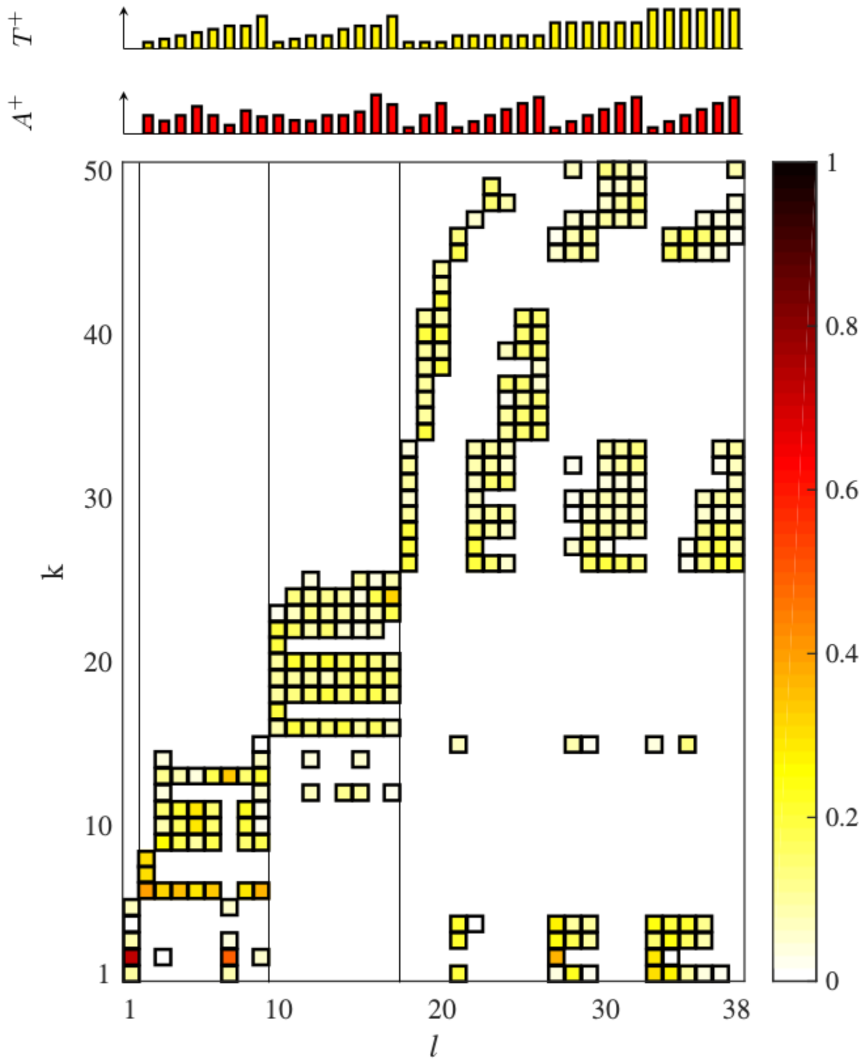


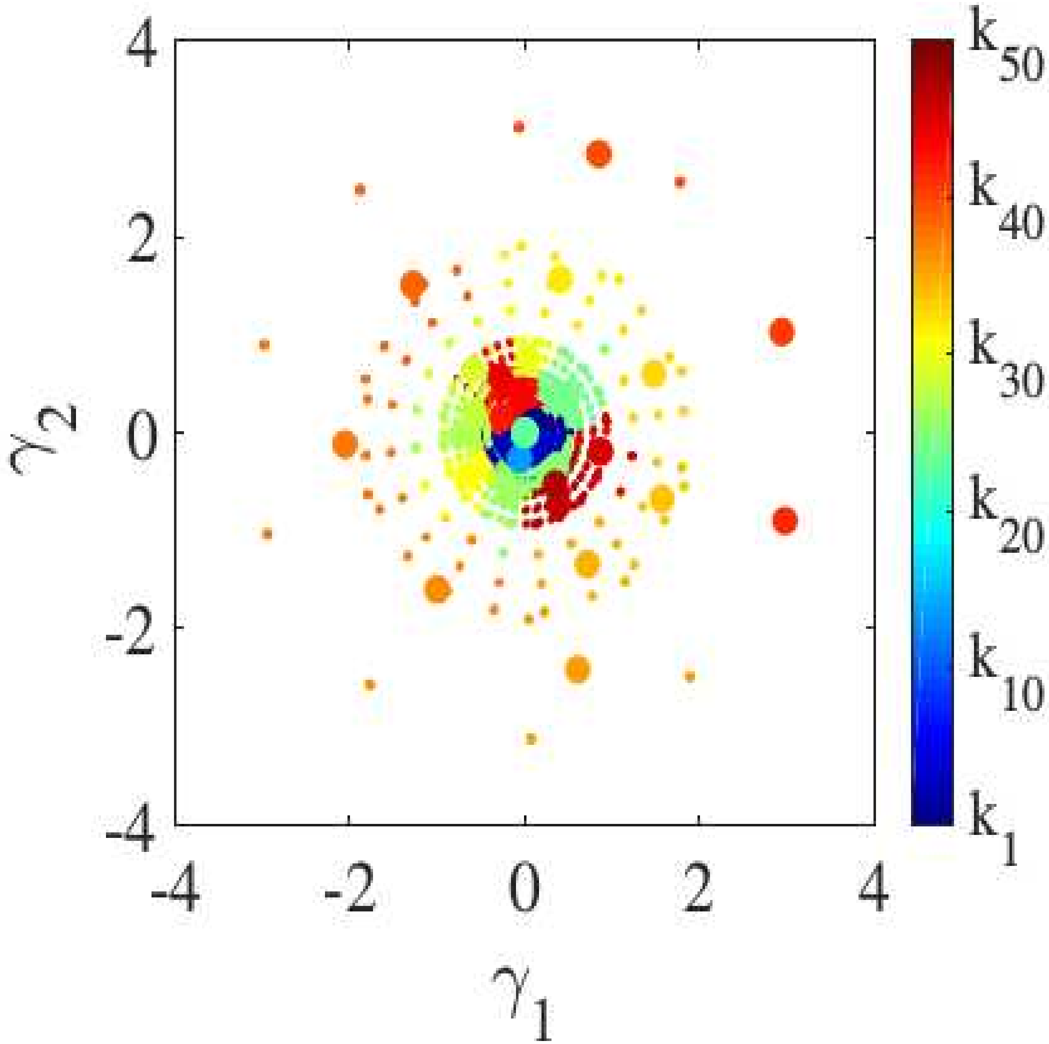
Illustration of the turbulent boundary layer flow: (left) non-actuated reference case N_1 , (center) actuated case with highest drag reduction N_{24} , and (right) actuated case with lowest drag reduction N_2; (top) contour plots of the instantaneous Cartesian velocity components u,v, and w in a $y-z$ plane at $x/\theta \approx 65$; (bottom) time evolution of the instantaneous drag reduction rate Δc_d.

TBL Drag reduction — Clustering

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2018 JFM (in revision)

50 clusters from 38 actuation cases

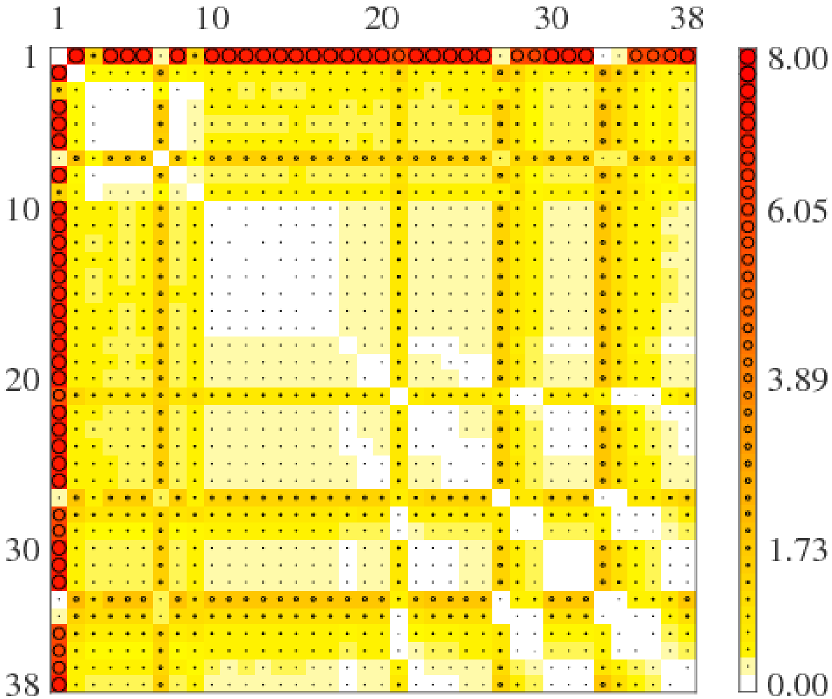


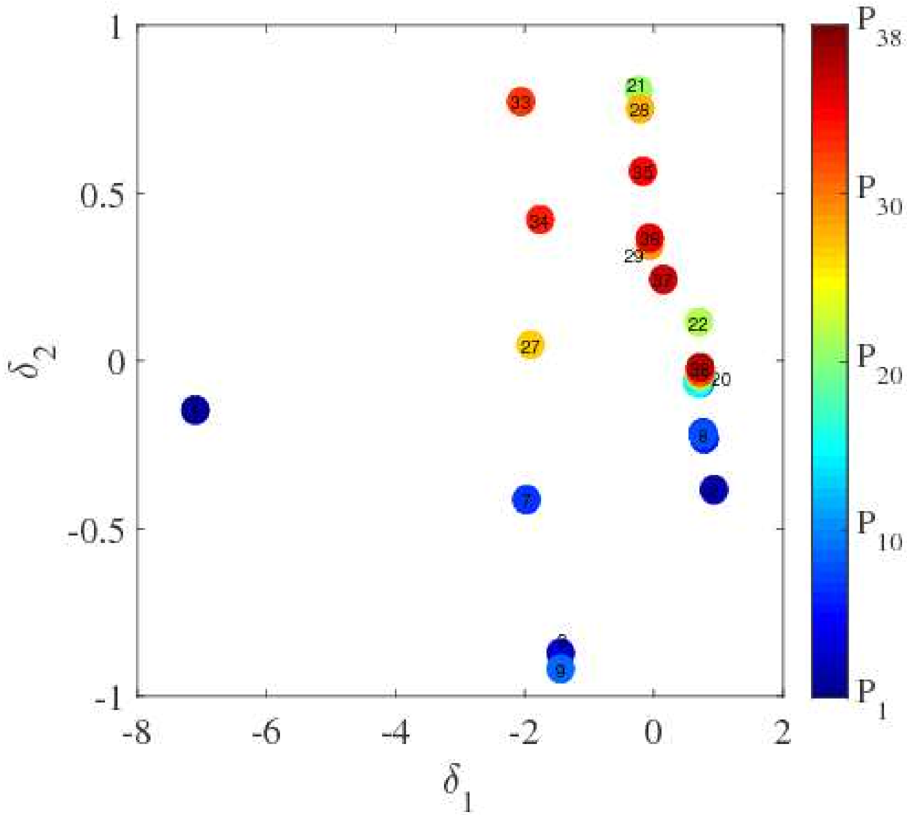


TBL drag reduction — Metric of Attractor Overlap

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2019 JFM

Metric of attractor overlap $D(A, B)$
$\sim$ average distance between attractors A and B

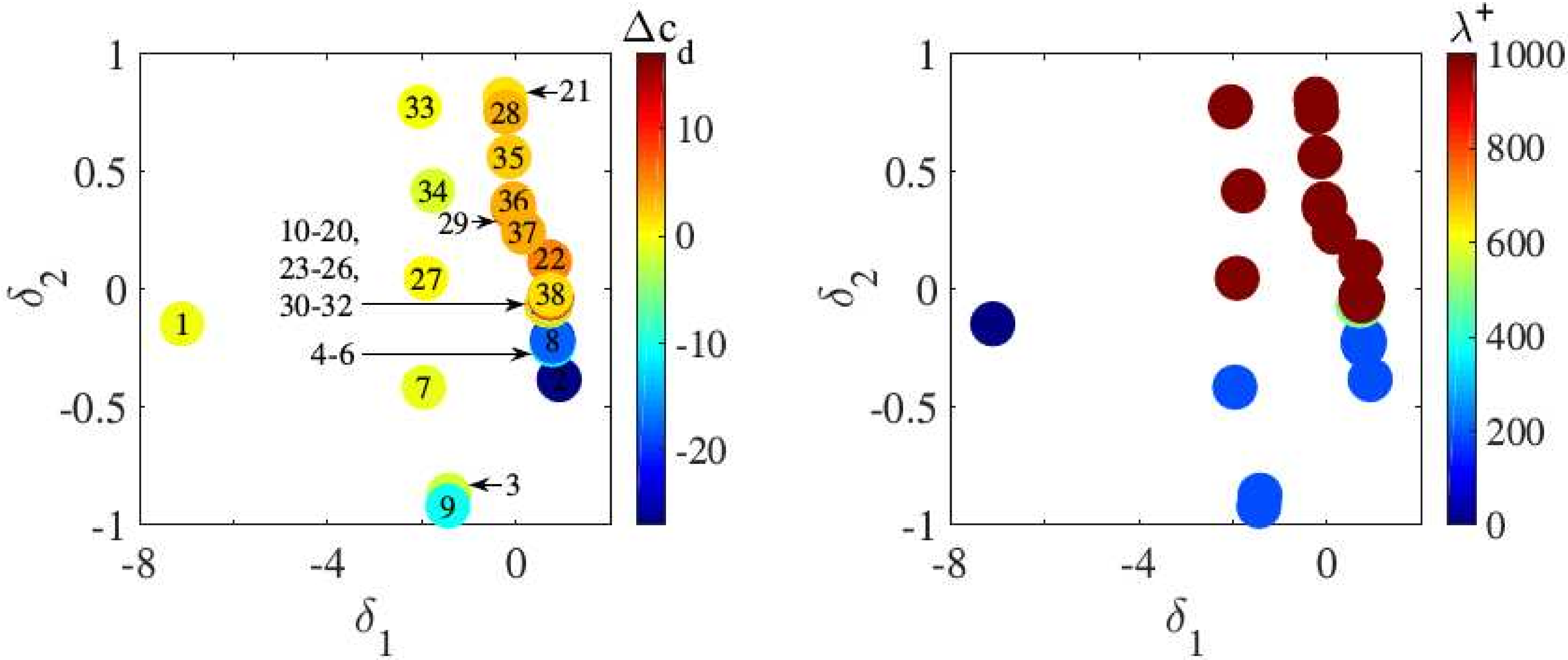




TBL drag reduction — Metric of Attractor Overlap

R. Ishar, E. Kaiser, M. Morzynski, M. Albers, P. Meysonnat, W. Schröder 2019 JFM

Interpretation of attractor proximity map



Attractor proximity map of the turbulent boundary layer simulations: (a) color-coded with drag reduction Δc_D; (b) color-coded with wavelength λ^+.