Cluster-based reduced-order modelling

From shear flows to engine tumble motion

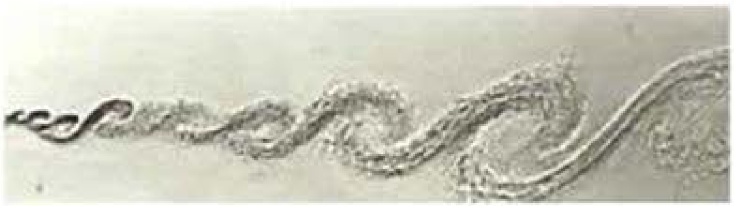
E. Kaiser[1,7] B. R. Noack[1] L. Cordier[1] A. Spohn[1] M. Segond[2]
M. Abel[2] G. Daviller[3] J. Östh[4] S. Krajnović[4] Y. Cao[1] J. Borée[1]
Lionel Thomas[1] Stéphane Guilain[5] R. K. Niven[6] L. N. Cattafesta[7]

[1]Institut P', CNRS – Université de Poitiers – ENSMA, France, [2]Ambrosys GmbH, Germany, [3]CERFACS, France
[4]Chalmers University, Sweden, [5]Renault s.a.s., France, [6]UNSW/ADFA, Australia, [7]Florida State University / FCAAP, USA

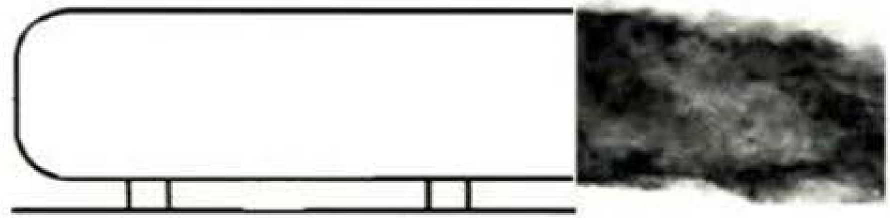
– supported by TUCOROM, SepaCoDe, Poitou-Charentes/France, PIRE, ADFA@UNSW –

SIAM conference CSE, 14–18 March 2015

Motivation



Brown & Roshko (1974)



Barros et al. (2014)

How can we identify physical mechanisms in an unsupervised manner directly from data?

- Analysis of physical mechanisms.
- Quick exploratory studies for optimization and parameter analysis.
- Robust models for real-time feedback control applications.

Cluster analysis for structure identification

- Part of pattern recognition/machine learning
  How to find a hidden structure or groups in data?
- Applications: Data compression, Tracking, 3D reconstruction, territorial behavior, etc.
- Goal: Organize data $\{\mathbf{x}^m\}_{m=1}^{M}$ into clusters such that there is a
  - high intra-cluster similarity and
  - low inter-cluster similarity
- What does *"similar"* mean? E.g.:
  $d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2$
- What are the points $\mathbf{c}_k$ that generate these subsets?

original

Original Image: Colorful Bird

$K = 3$

k-Means Clustering of 3 Colours

$K = 6$

k-Means Clustering of 6 Colours

K-means algorithm

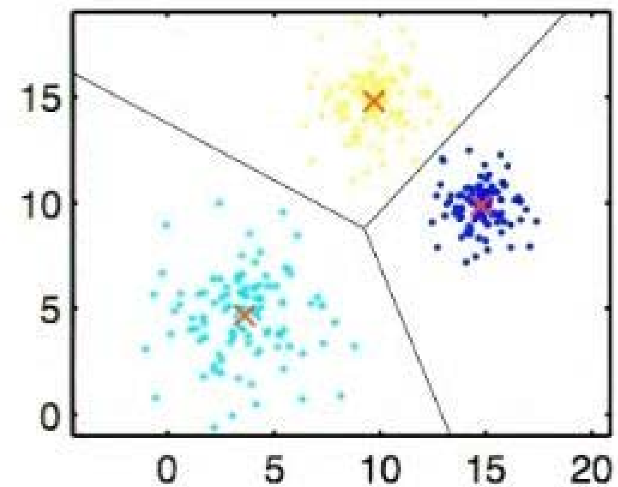
Given a data points/observations $\{\mathbf{x}^m\}_{m=1}^{M}$ where $\mathbf{x} = [x_1, \ldots, x_N]^T$

- **Initialization:** K cluster centers $\mathbf{c}_1, \ldots, \mathbf{c}_K$
- **Assign** each $\mathbf{x}^m$ to its closest cluster center $\mathbf{c}_k$
- **Update** each cluster center $\mathbf{c}_k$

$$\mathbf{c}_k = \frac{1}{n_k} \sum_{\mathbf{x}^m \in \mathcal{C}_k} \mathbf{x}^m$$

- **Iterate** until converged
- Minimization of total cluster variance

$$J(\mathbf{c}_1, \ldots, \mathbf{c}_k) = \sum_{k=}^{K} \sum ||\mathbf{x}^m - \mathbf{c}_k||^2$$

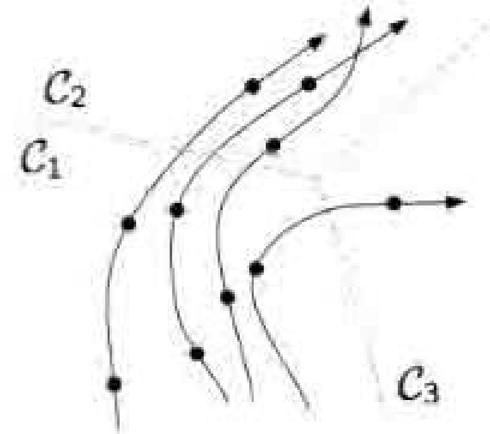


Dynamical model

Cluster transition probability matrix

$$P_{jk} = \frac{n_{jk}}{n_k}$$

$n_{jk} :=$ # of transitions from $\mathcal{C}_k$ to $\mathcal{C}_j$
$n_k :=$ # of data points in cluster $\mathcal{C}_k$



- $\mathbf{P}$ defines a *one-step time-homogeneous Markov chain* where

$$P_{jk} = \mathrm{Prob}(\mathbf{c}_j \mid \mathbf{c}_k)$$

- Evolution of probability distribution

$$\mathbf{p}_{m+1} = \mathbf{P}\,\mathbf{p}_m \quad \rightarrow \quad \mathbf{p}_m = \mathbf{P}^m\,\mathbf{p}_0$$

- Convergence to unique, stationary distribution

$$\mathbf{p}_\infty = \lim_{m\to\infty} \mathbf{P}^m\,\mathbf{p}_0$$

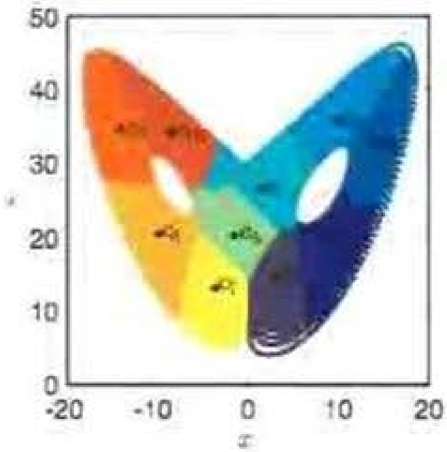
CROM applied to Lorenz attractor data

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM

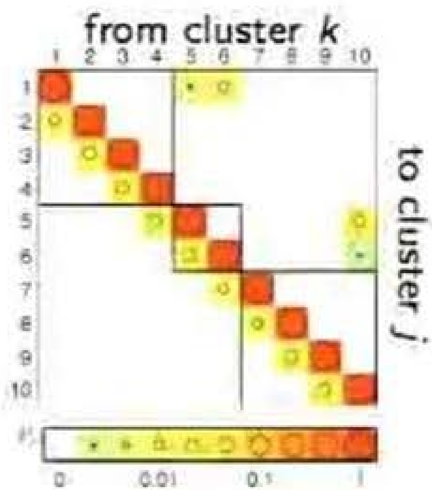
Lorenz system (*Lorenz, 1963*) with $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$.

Data set: $\mathbf{x}^m := \mathbf{x}(t_m) = [x(t_m), y(t_m), z(t_m)]^T$

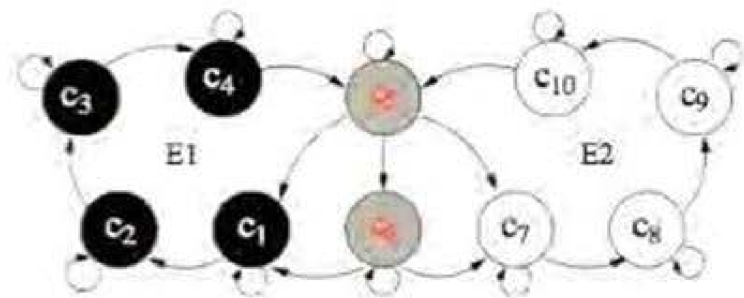
Kinematics



Dynamics



Graphical representation



- Coarse-grained state space
- Homogeneous partitioning around fixed points → phase averaging
- Identification of osc … nsition clusters

Interpretation as generalization of Ulam's method

Dynamical system:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x})$$

Liouville equation:

$$\frac{\partial \rho(\mathbf{x},t)}{\partial t} = -\nabla \cdot (\mathbf{f}\rho(\mathbf{x},t)) = \mathcal{L}(\rho(\mathbf{x},t))$$

Perron-Frobenius operator:

$$\mathcal{P} = \exp(\mathcal{L}t)$$

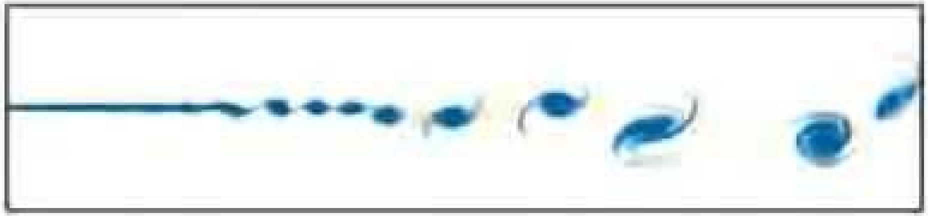
Ulam-Galerkin method (*Li, 1976, J. Approx. Theory*):

$$P_{jk} = \frac{\mathrm{card}(\{\mathbf{x}^m | \mathbf{x}^m \in B_k \text{ and } \mathbf{f}(\mathbf{x}^m) \in B_j\})}{\mathrm{card}(\{\mathbf{x}^m \in B_k\})}$$

CROM:

- Low dimension
- Other distance metrics

Spatially developing mixing layer

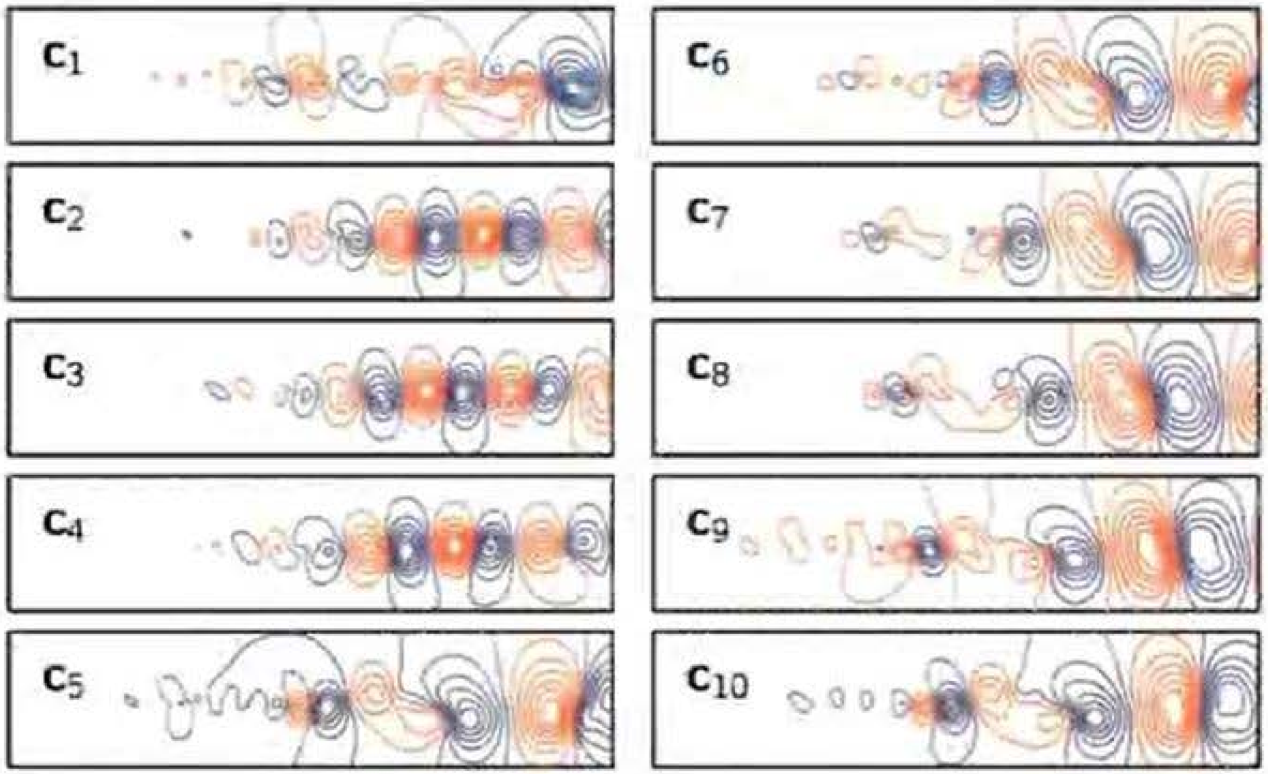


- 2D mixing layer simulation with Kelvin-Helmholtz vortices undergoing vortex pairing
- $Re = \Delta U \delta_\omega / \nu = 500$; $Ma = 0.3$; $r = U_1/U_2 = 3$
- Finite-difference Navier-Stokes solver (*Daviller, PhD thesis (2010), Cavalieri et al. (2011)*)
- Data compression using proper orthogonal decomposition (used for all following examples)

CROM of the mixing layer

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM

Kinematics: Cluster centroids

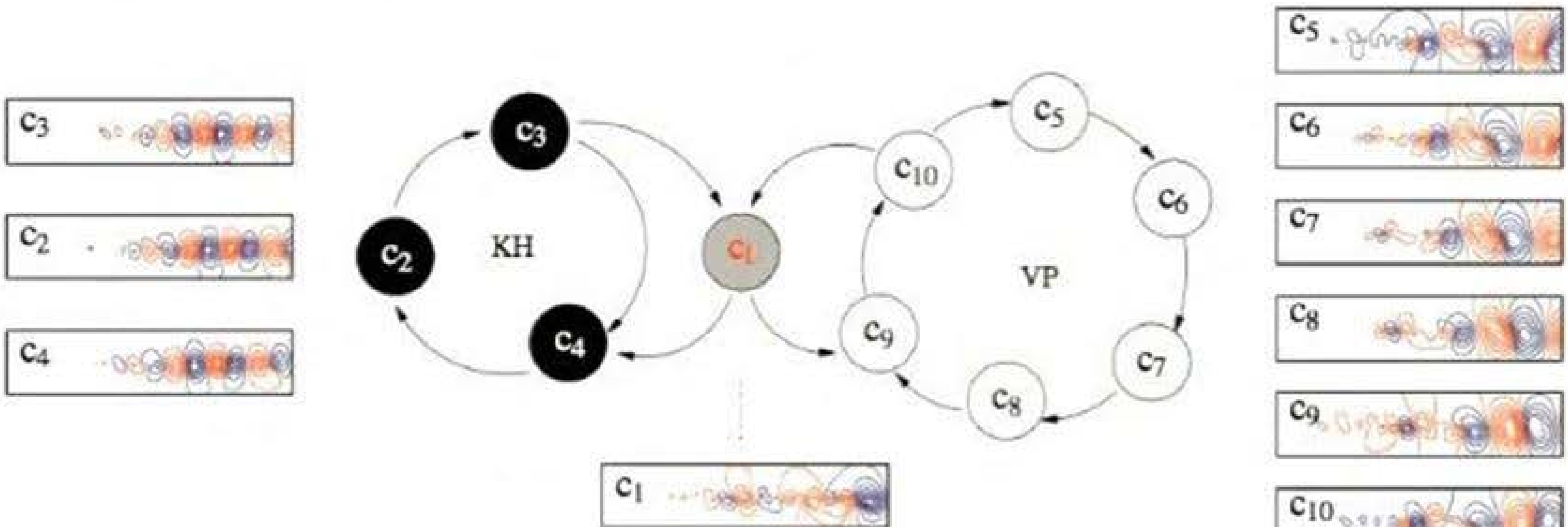


- Most clusters are "phase bins"

CROM of the mixing layer

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM

Dynamics: Graph of the transition probability matrix



- Identification of two shedding regimes.
- Flipper cluster acts as a switch between both shedding regimes.

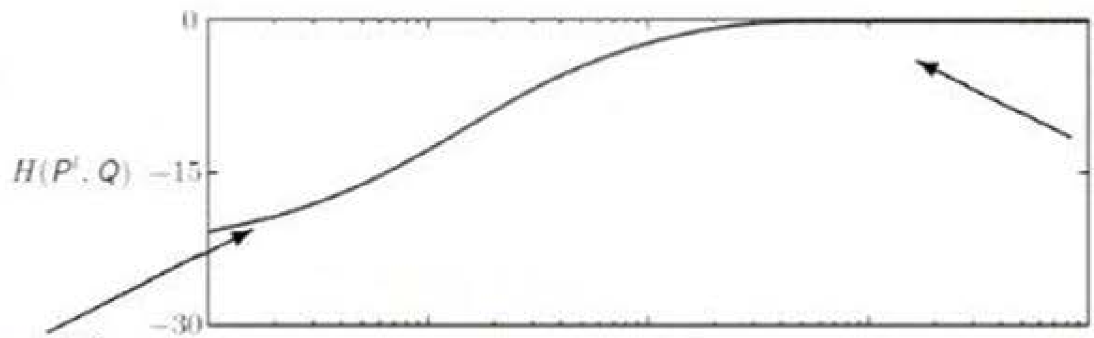
Attractor properties of the mixing layer

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Morzyński, Östh, Krajnović, Niven, 2014, AIP Conf. Proc. (MaxEnt2013)

Analysis of model convergence to its asymptotic state via relative entropy (negative of Kullback-Leibler function D):

$$H(\mathbf{P}^m, \mathbf{Q}) = -D(\mathbf{P}, \mathbf{Q}) := -\sum_{j=1}^{K} \sum_{k=1}^{K} P_{jk}^m \ln \frac{P_{jk}^m}{Q_{jk}}$$

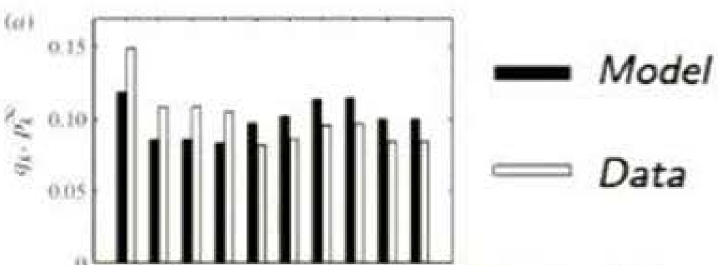
Transients:



All information on initial conditions is lost

Backwards integration possible

Attractor:

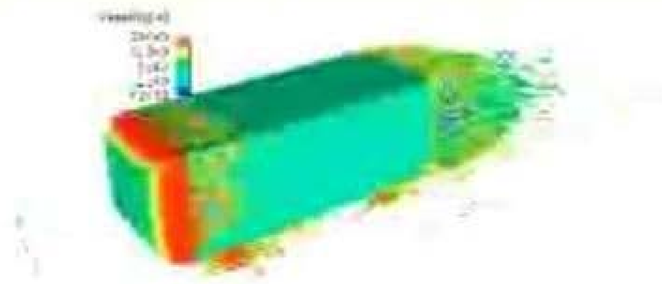


Probability distributions: Model $\mathbf{p}_\infty$ *compared to data* $\mathbf{q}$

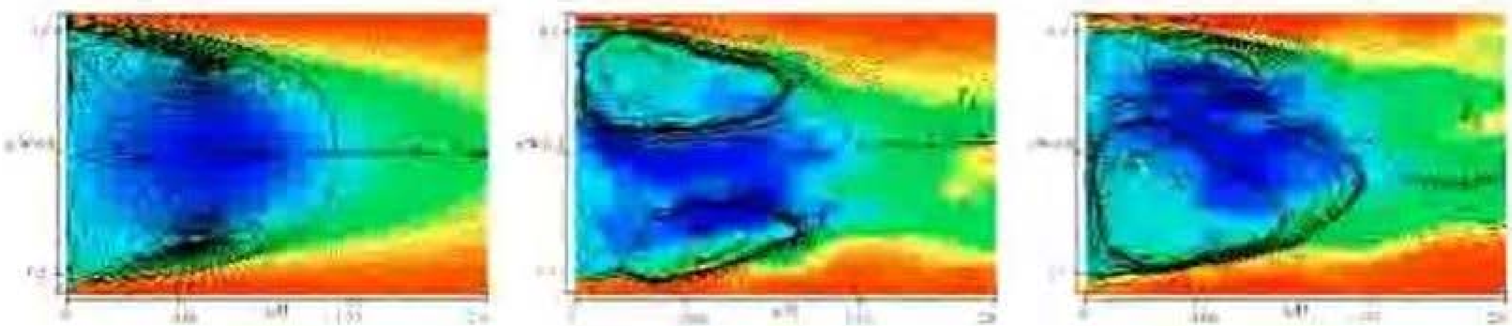
Turbulent wake of a bluff body

Östh, Noack, Krajnović, Barros & Boreé, 2014, JFM

- 3D incompressible turbulent flow around a bluff body (finite-volume LES)
- $Re = U_\infty H/\nu = 3x10^5$
- Data compression using POD

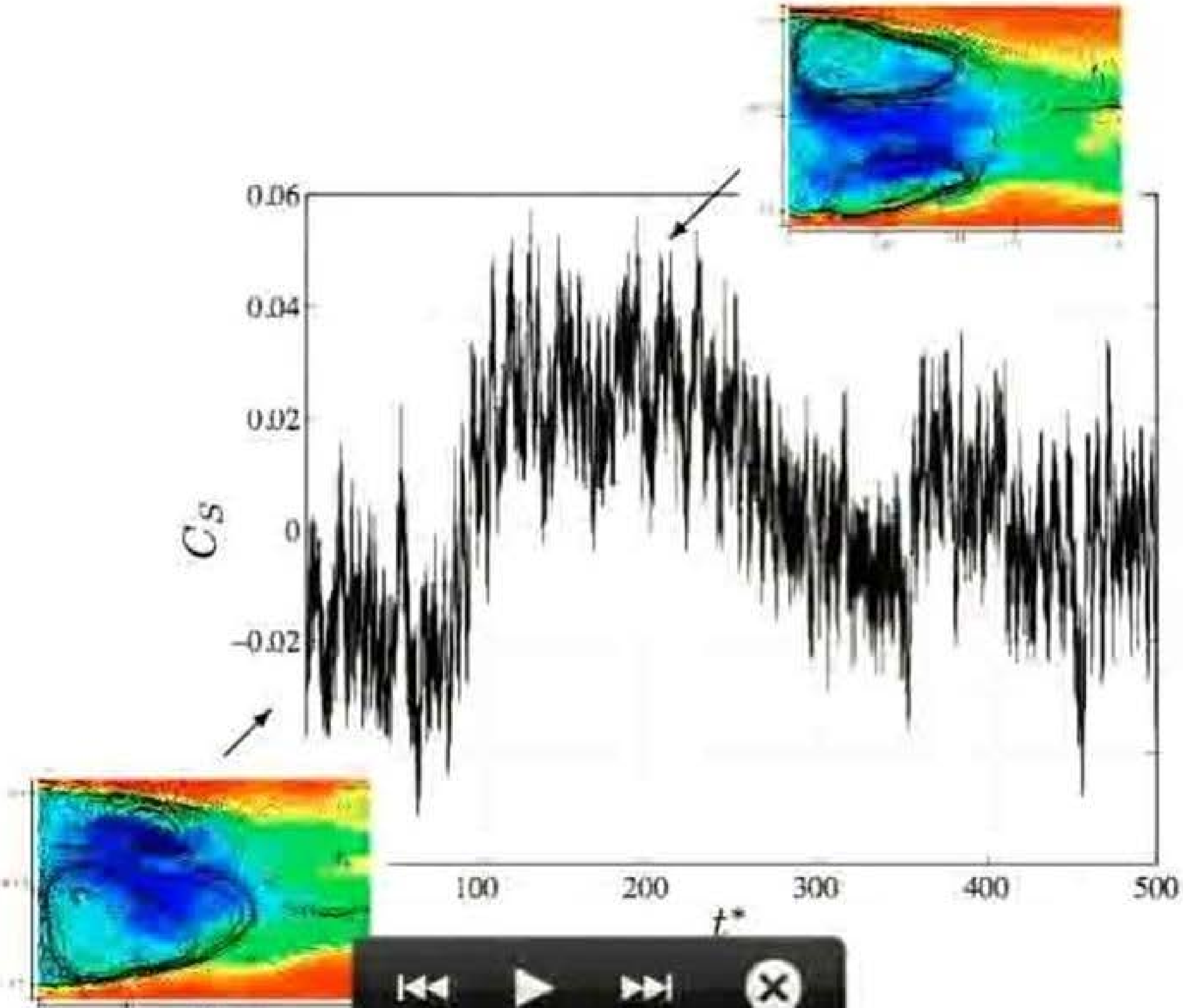


Bi-modal behaviour of the wake (*Grandmange et al. 2013*):
Flow changes between two asymmetric states over time scales of order $T_s \approx 100\, J/U_\infty \Longrightarrow$ Challenge for ROM.



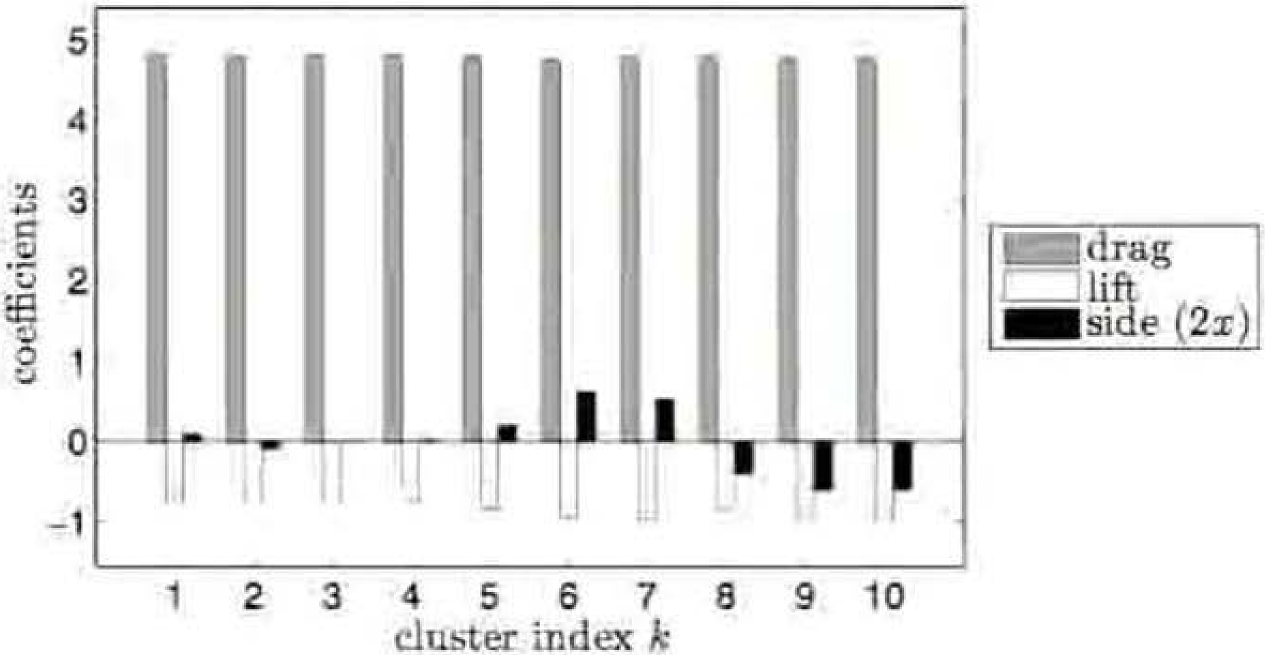
Side force of the bluff body

Östh, Noack, Krajnović, Barros & Boreé, 2014, JFM



Mean cluster forces of the bluff body wake

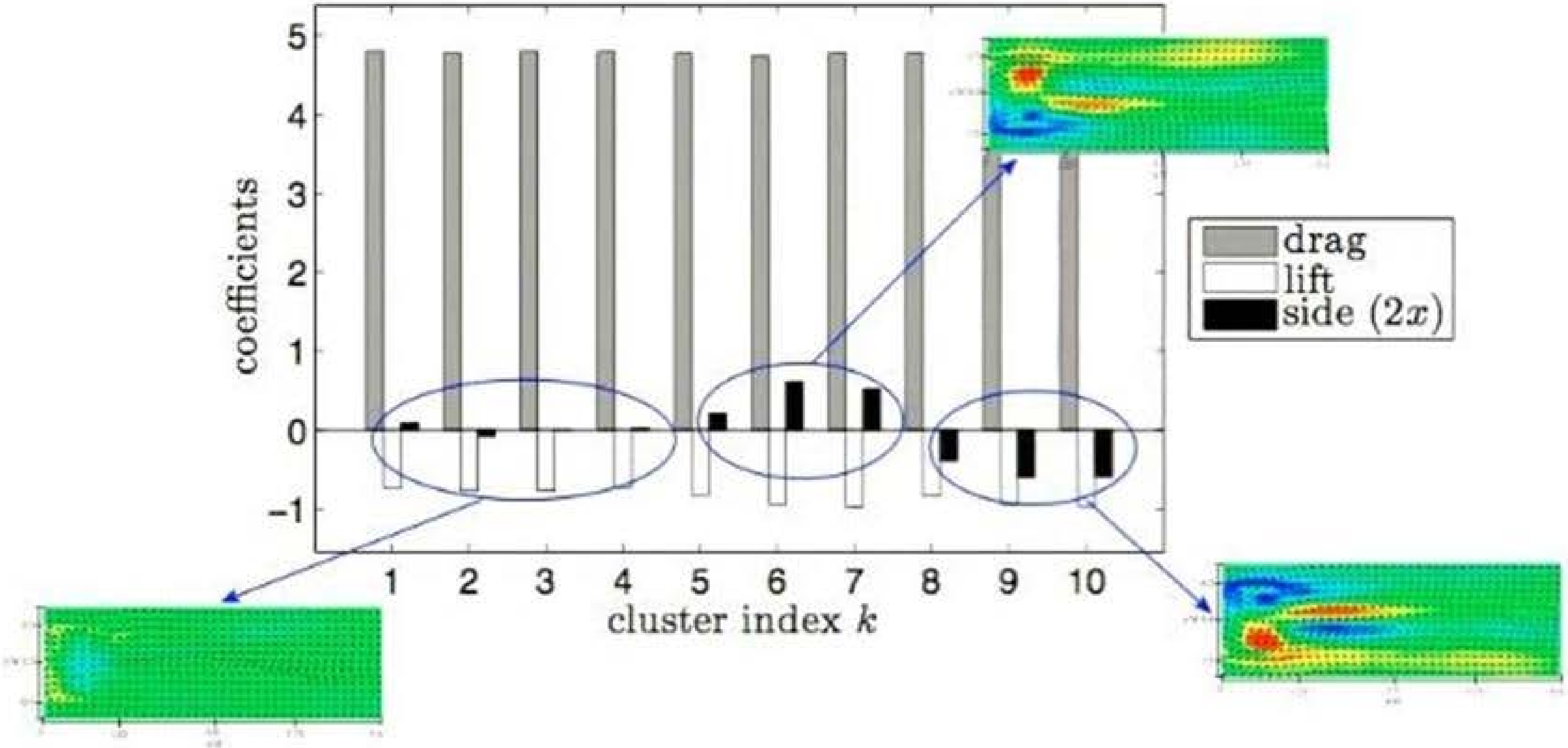
Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM



- Clusters are associated with different mean forces
- Desirable state space regions

Mean cluster forces of the bluff body wake

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM

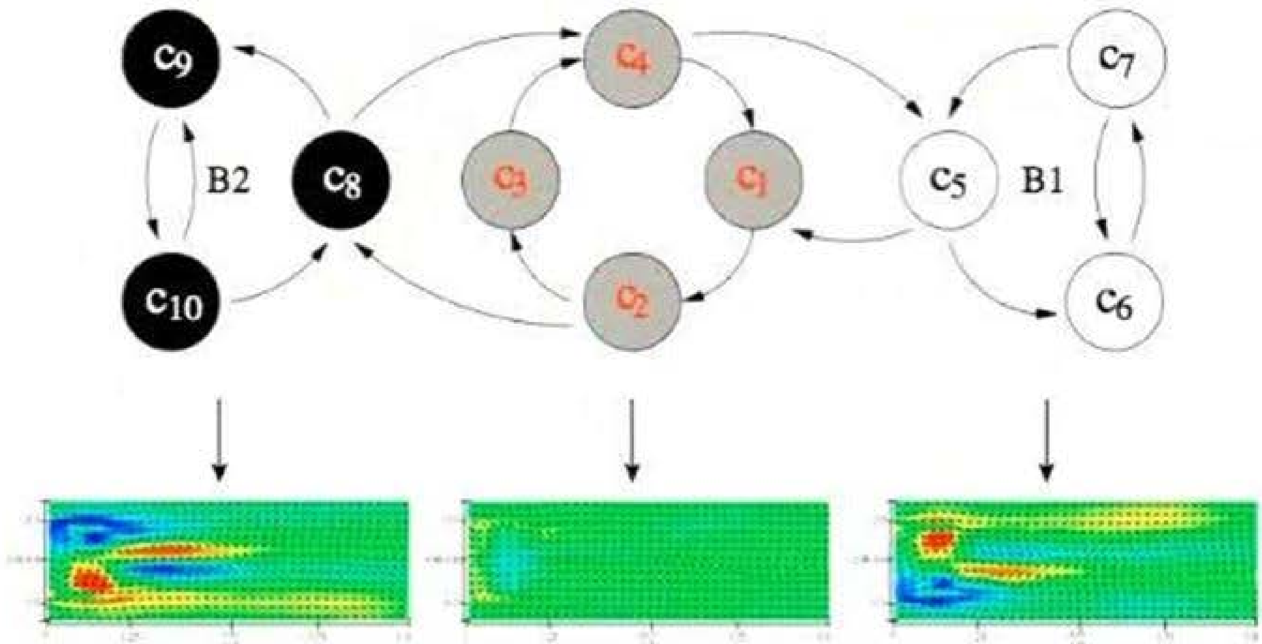


- Clusters are associated with different mean forces
- Desirable state spac

CROM of the bluff body wake

Kaiser, Noack, Cordier, Spohn, Segond, Abel, Daviller, Östh, Krajnović, Niven, 2014, JFM

Cluster dynamics



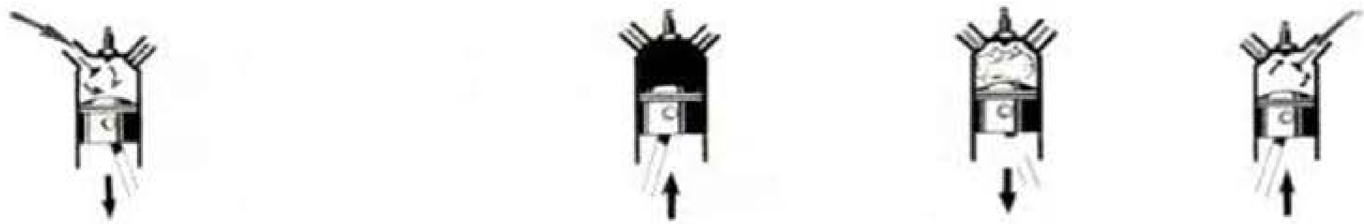
- Bi-modal behavior and transition processes are distilled
- Identification of branching clusters ($\mathcal{C}_5$ and $\mathcal{C}_8$)

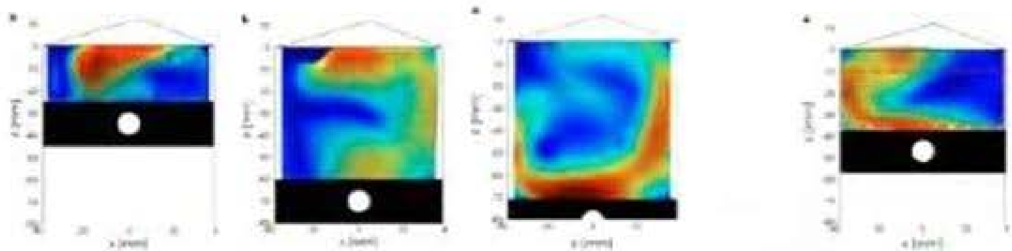
Engine tumble motion of an IC engine

Voisine et al. 2011, Exp. Fluids

Four-stroke internal combustion engine

Intake Compression Ignition Exhaust



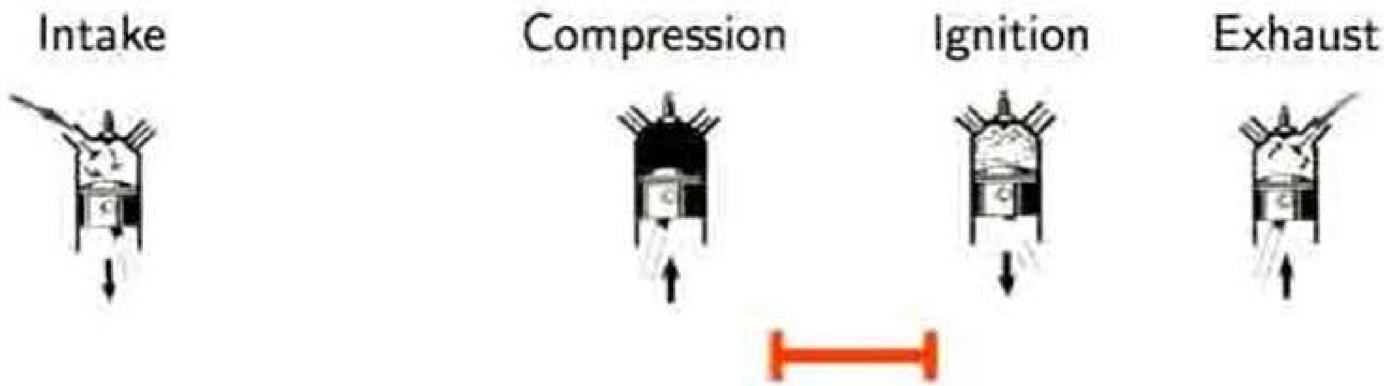


Mid-compression
(270 CAD)

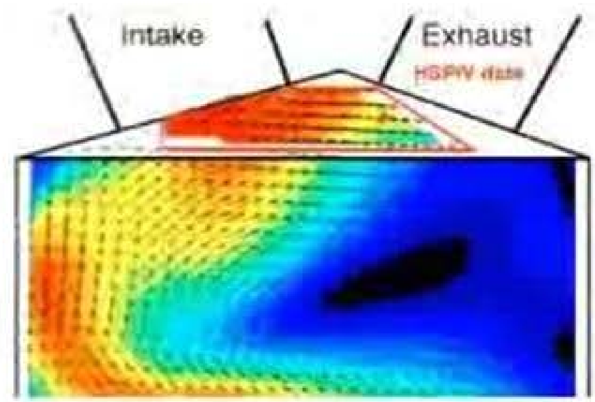
- Tumble breakdown from large-scale to small-scale structures
- Large-scale cycle-to-cycle variability effects engine efficiency and emissions

Data set of IC engine

Voisine et al. 2011, Exp. Fluids



- HSPIV in symmetry plane of cylinder head/ pent-roof chamber
- From mid-compression to end of compression phase ($\theta = 270 \ldots 354\,CAD$, $\Delta\theta = 4\,CAD$)
- $N = 161$ consecutive engine cycles are measured with $M = 22$ snapshots/cycle

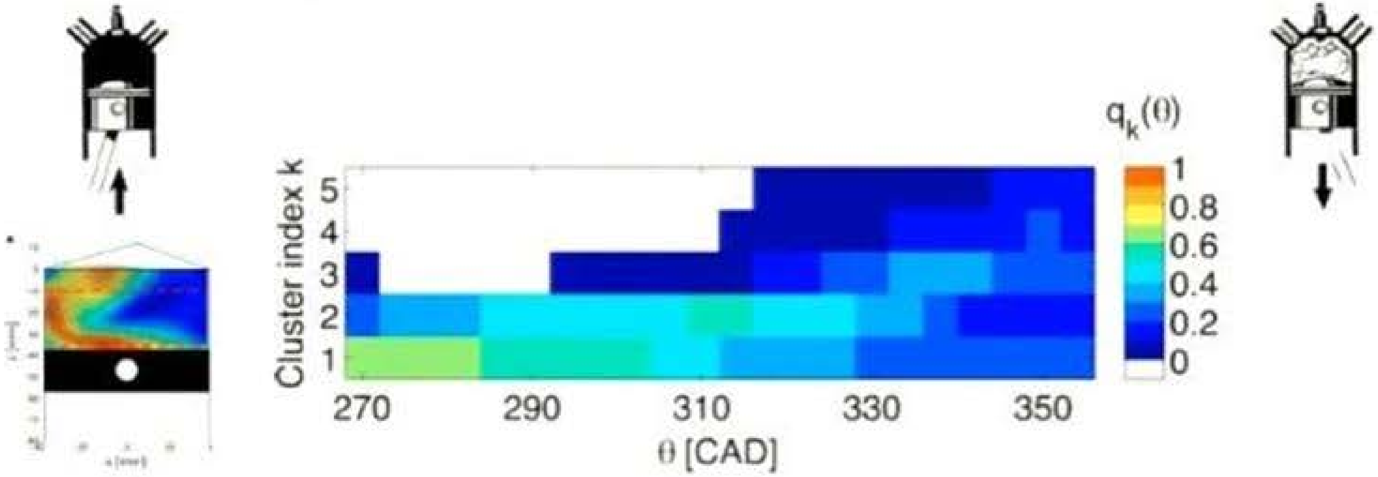


Evolution of cluster probability distribution

Cao, Kaiser, Boree, Noack, Thomas, Guilain, 2014, *Exp. Fluids*

Mid-Compression

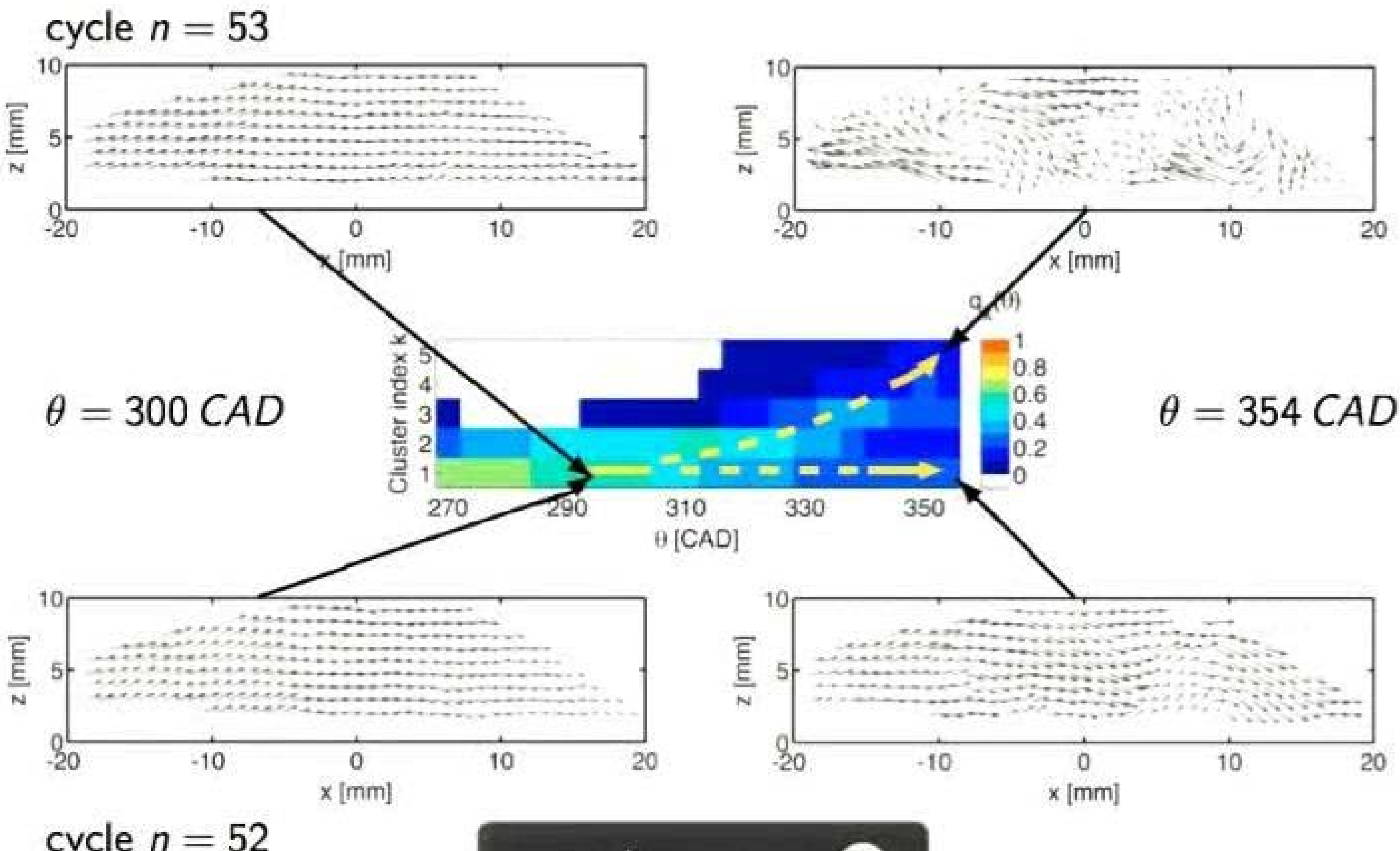
Ignition



- Cluster probability $q_k(\theta) = \frac{n_k(\theta)}{N}$
  - $N = 161$: total # cycles
  - $n_k(\theta)$: # snapshots at $\theta \in \mathcal{C}_k$
- Spreading of flow fields during tumble breakdown
- Almost uniform PDF at ignition

Large variation of flow patterns at ignition

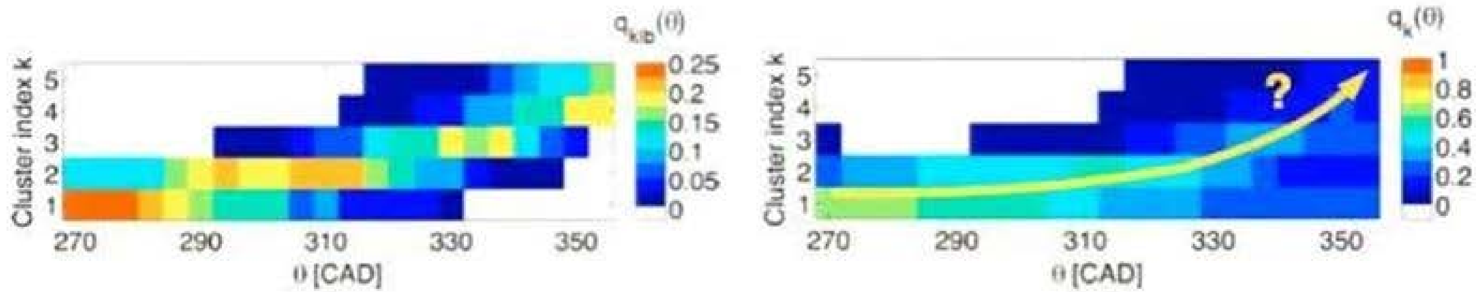
Cao, Kaiser, Boreé, Noack, Thomas, Guilain, 2014 Exp. Fluids



A stable and potentially cleaner engine

Cao, Kaiser, Boreé, Noack, Thomas, Guilain, 2014, Exp. Fluids

- Identification of important clusters: $\mathcal{C}_4$ and $\mathcal{C}_5$ are the most desirable states for the end of compression phase.



- Evolution of probability distribution estimated from trajectories ending in $\mathcal{C}_4/\mathcal{C}_5$: Gradual destabilisation of the large-scale flow structures
- Goal: Completion of tumble breakdown in each cycle

Conclusions and outlook

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