Project DisCo: Physics-based discovery of coherent structures in spatiotemporal systems

Adam Rupe [1 2]
Karthik Kashinath [2], Nalini Kumar [3],
James P. Crutchfield [1], Ryan G. James [1], and Prabhat [2]

[1]Department of Physics
Complexity Sciences Center
University of California, Davis

[2]NERSC
Lawrence Berkeley National Laboratory

[3]Intel®

Big Data Summit 2018

Project DisCo

John Travolta in 'Saturday Night Fever' (Paramount)

- The Science Problem
- The Theory
- The Computation
- The Unsupervised Ladder
- The HPC Challenge

Project DisCo

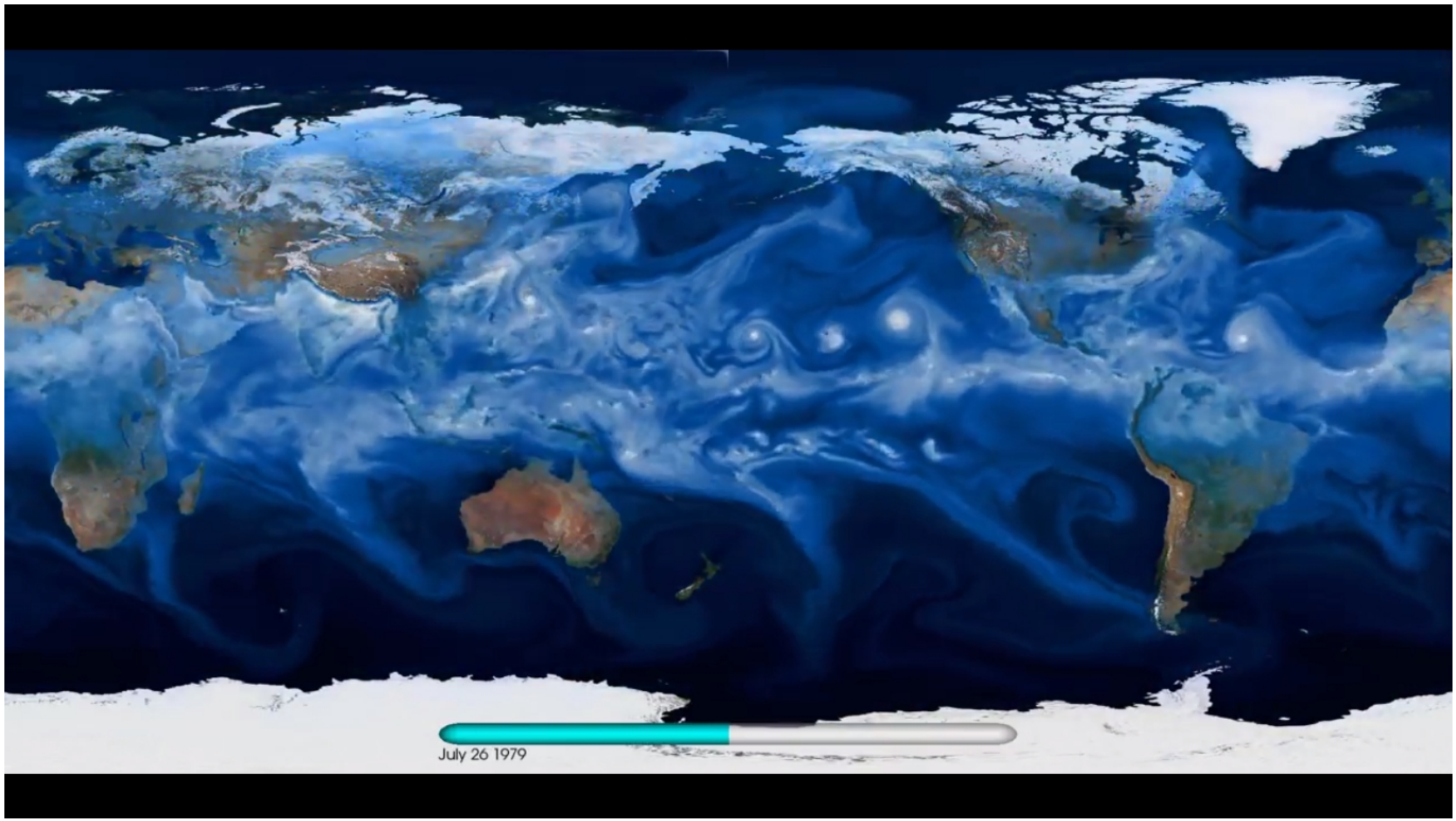
John Travolta in 'Saturday Night Fever' (Paramount)

- The Science Problem
- The Theory
- The Computation
- The Unsupervised Ladder
- The HPC Challenge

on track for SC'19 Gordon Bell submission

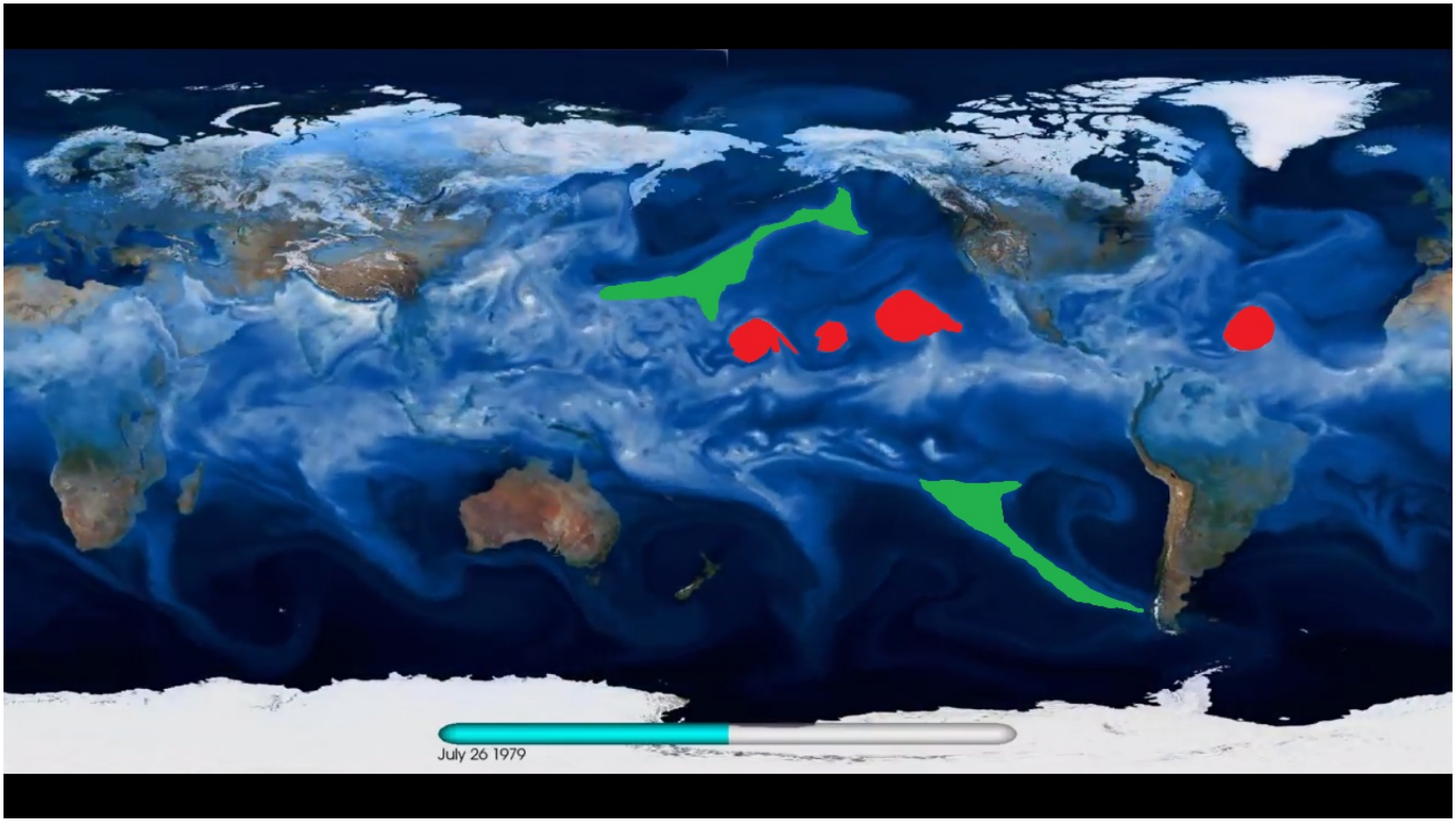
The Science Problem

Unsupervised detection (segmentation) of spatiotemporal structures in climate



The Science Problem

Unsupervised detection (segmentation) of spatiotemporal structures in climate



The Theory: Computational Mechanics

Think of this as a *physics-based* machine learning technique

Segmentation achieved through use of *local causal states*

The Theory: Computational Mechanics

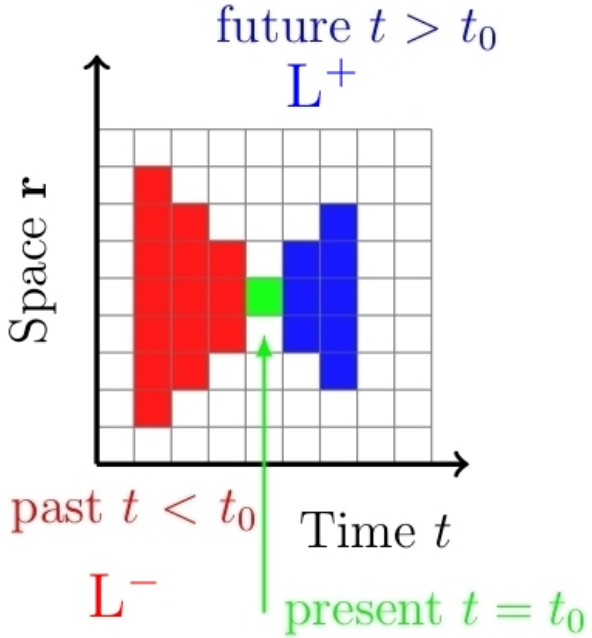
Think of this as a *physics-based* machine learning technique

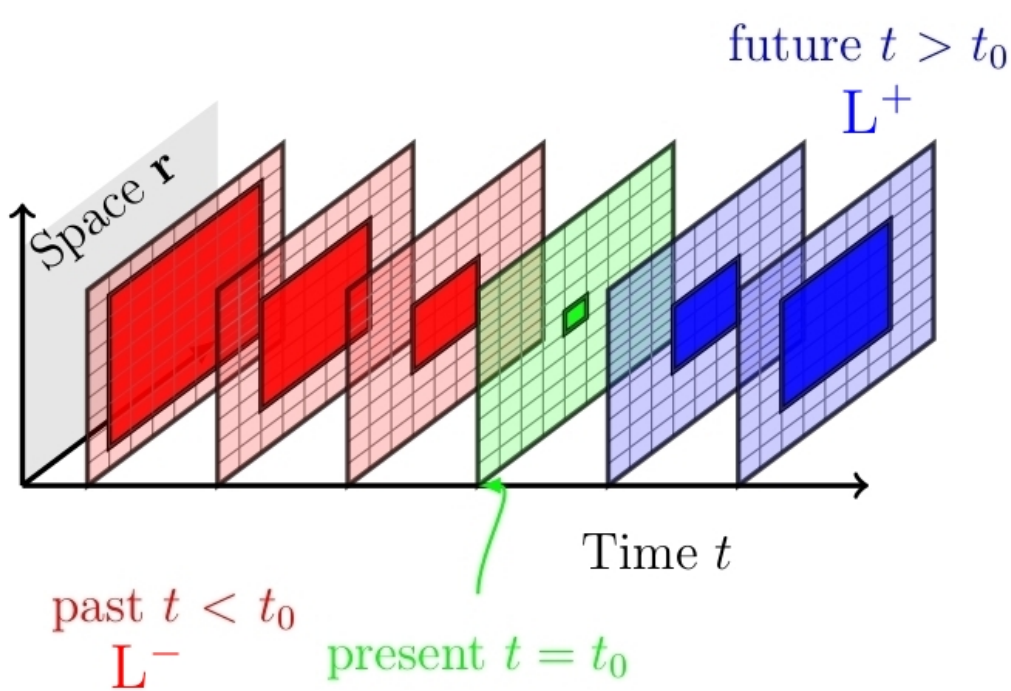
Segmentation achieved through use of *local causal states*

Causal states defined through the *causal equivalence relation*

$$\text{past}_i \sim_\epsilon \text{past}_j \iff \Pr(\text{Future}|\text{past}_i) = \Pr(\text{Future}|\text{past}_j)$$

Lightcones used as local notions of past and future





The Computation

Goal - reconstruction of *causal equivalence relation* from data

$$\ell_i^- \sim_\epsilon \ell_j^- \iff \Pr\left(\mathtt{L}^+|\ell_i^-\right) = \Pr\left(\mathtt{L}^+|\ell_j^-\right)$$

The Computation

Goal - reconstruction of *causal equivalence relation* from data

$$\ell_i^- \sim_\epsilon \ell_j^- \iff \Pr\left(\mathtt{L}^+|\ell_i^-\right) = \Pr\left(\mathtt{L}^+|\ell_j^-\right)$$

- Training
  1. Reconstruct morph$(\ell_i^-) = \Pr\left(\mathtt{L}^+|\ell_i^-\right)$
     - extract $\left(\ell^-, \ell^+\right)$ pairs from sample fields
     - for real-valued fields, need to cluster space of lightcones

The Computation

Goal - reconstruction of *causal equivalence relation* from data

$$\ell_i^- \sim_\epsilon \ell_j^- \iff \Pr\left(\mathtt{L}^+|\ell_i^-\right) = \Pr\left(\mathtt{L}^+|\ell_j^-\right)$$

- Training
  1. Reconstruct morph$(\ell_i^-) = \Pr\left(\mathtt{L}^+|\ell_i^-\right)$
     - extract $\left(\ell^-, \ell^+\right)$ pairs from sample fields
     - for real-valued fields, need to cluster space of lightcones
  2. Cluster together pasts with same morph - resulting clusters are *local causal states*
     - gives ϵ-map; $\epsilon(\ell_i^-) = \xi_{\ell_i^-} = \{\ell_j^- \ : \ \ell_j^- \sim_\epsilon \ell_i^-\}$

The Computation

Goal - reconstruction of *causal equivalence relation* from data

$$\ell_i^- \sim_\epsilon \ell_j^- \iff \Pr\left(\mathtt{L}^+|\ell_i^-\right) = \Pr\left(\mathtt{L}^+|\ell_j^-\right)$$

- Training
  1. Reconstruct $\text{morph}(\ell_i^-) = \Pr\left(\mathtt{L}^+|\ell_i^-\right)$
     - extract $\left(\ell^-, \ell^+\right)$ pairs from sample fields
     - for real-valued fields, need to cluster space of lightcones
  2. Cluster together pasts with same morph - resulting clusters are *local causal states*
     - gives ϵ-map; $\epsilon(\ell_i^-) = \xi_{\ell_i^-} = \{\ell_j^- \ : \ \ell_j^- \sim_\epsilon \ell_i^-\}$
- Inference
  1. Use ϵ-map to perform *causal filtering*; $\mathbf{x} \to \mathcal{S} = \epsilon(\mathbf{x})$
     - *Segmentation* semantics from structural properties of $\mathcal{S}$

The Unsupervised Ladder

No labeled data – no error metric to optimize

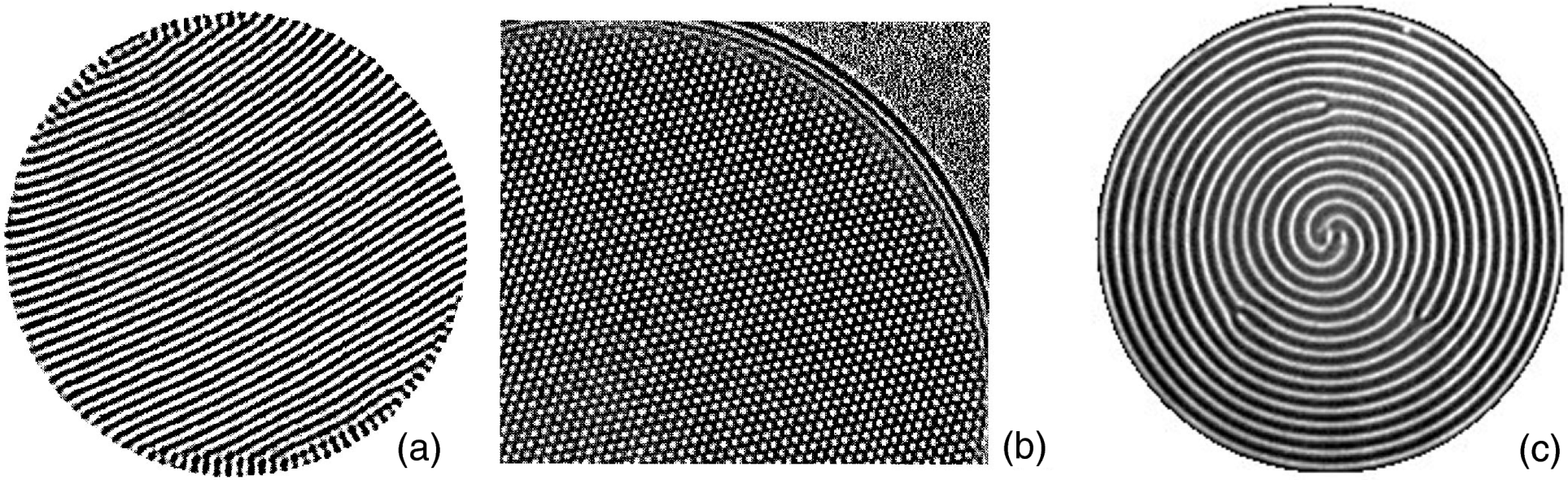
The Unsupervised Ladder

No labeled data – no error metric to optimize

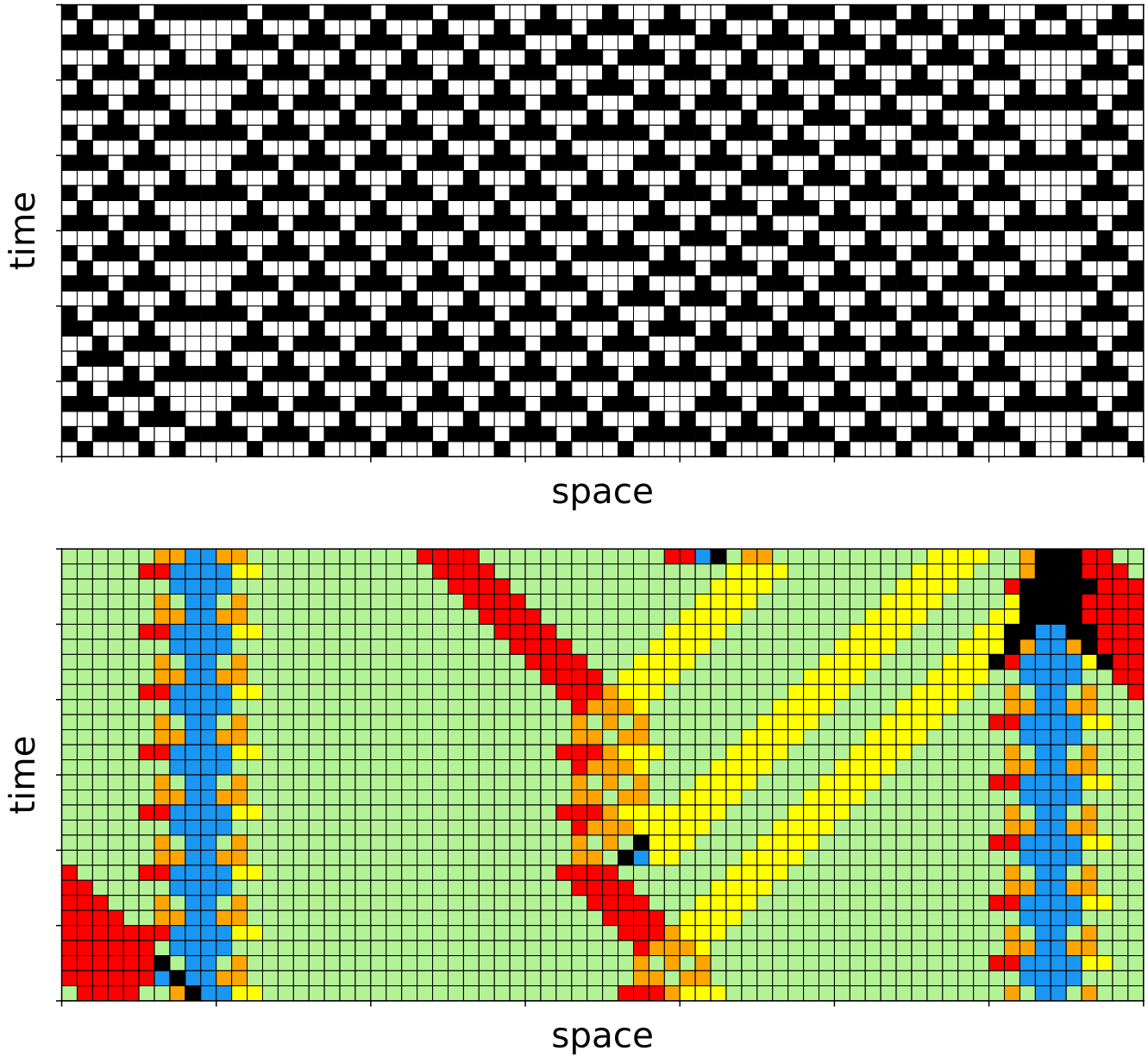
Built from physical theory

- validation established using physical principles

broken symmetry

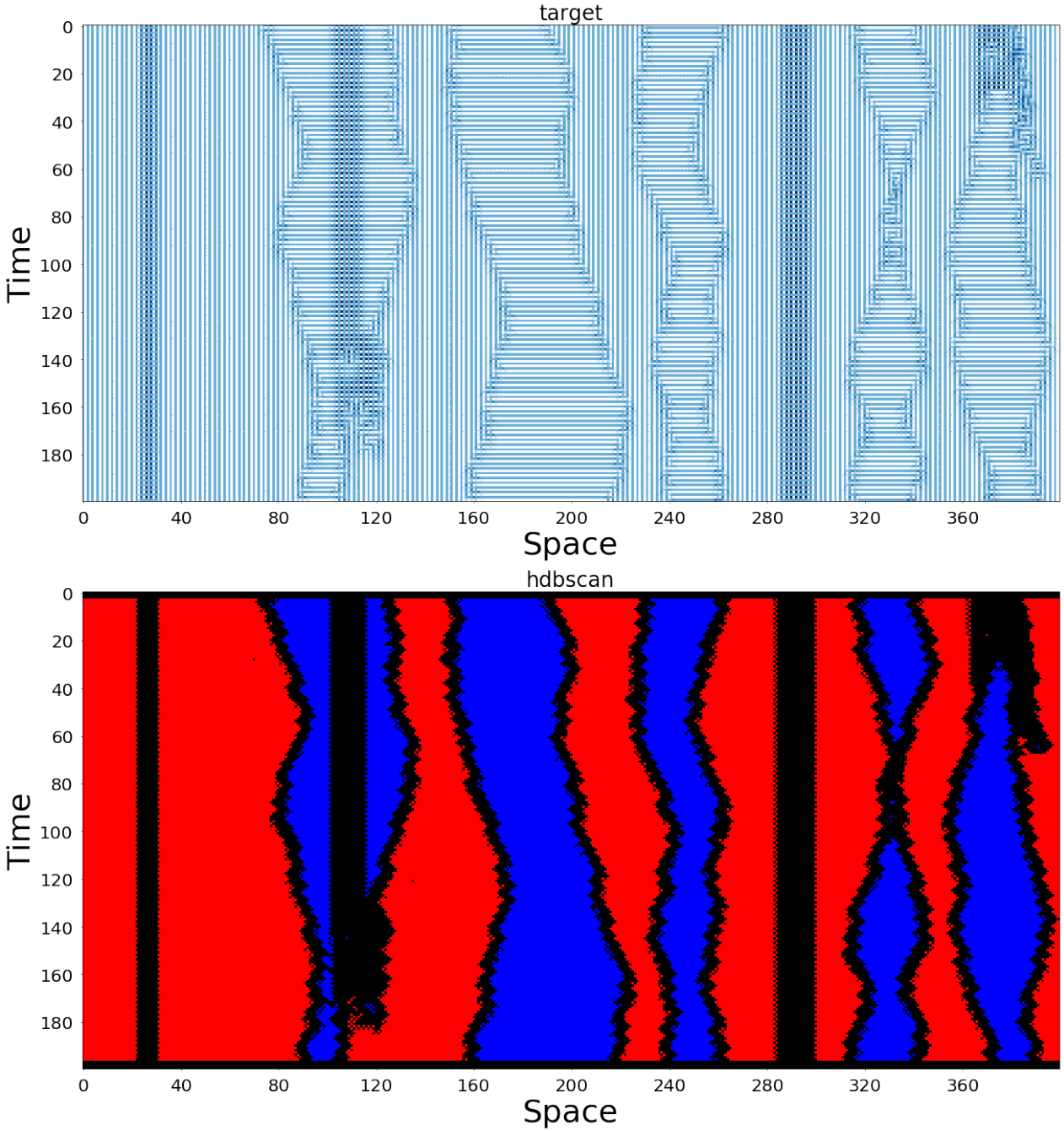


Step 1 – Cellular Automata

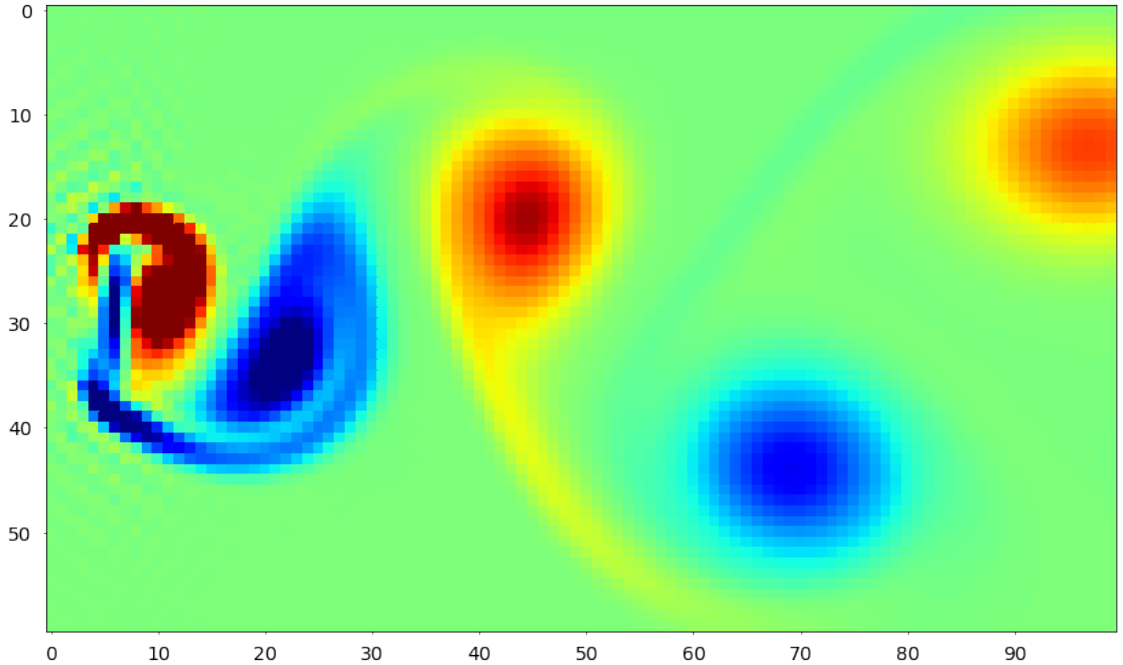


A. Rupe and J.P. Crutchfield (2018). Local Causal States and Discrete Coherent Structures. arXiv preprint arXiv:1801.00515.

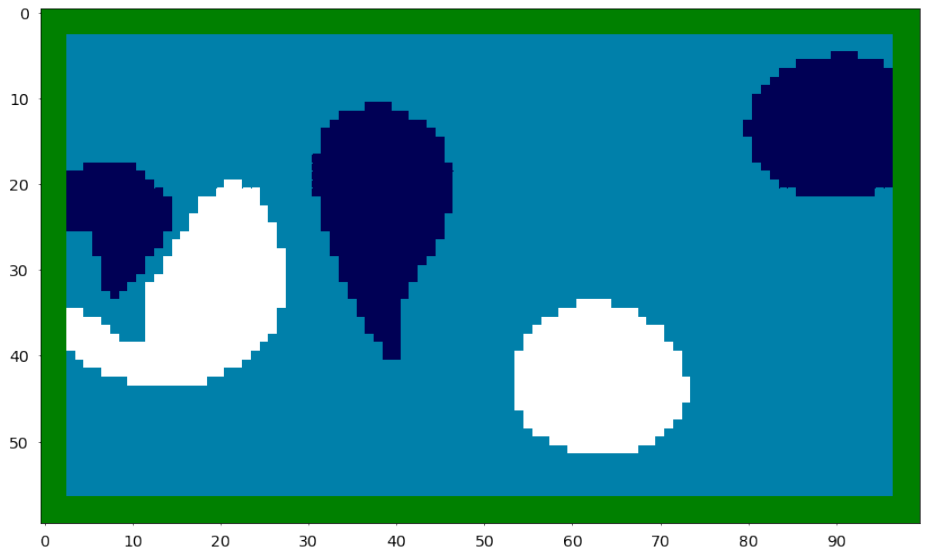
Step 2 – Coupled Map Lattices



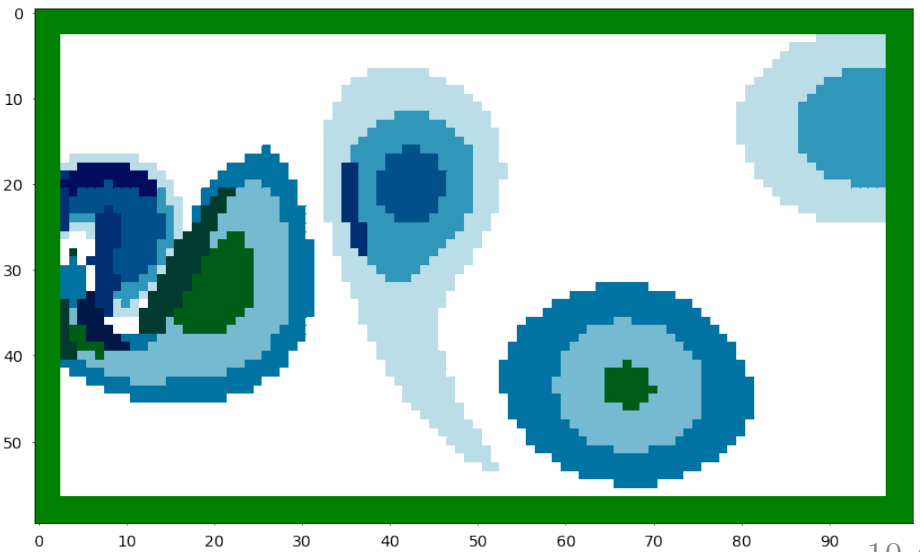
Step 3 – Vortex Shedding



K = 3

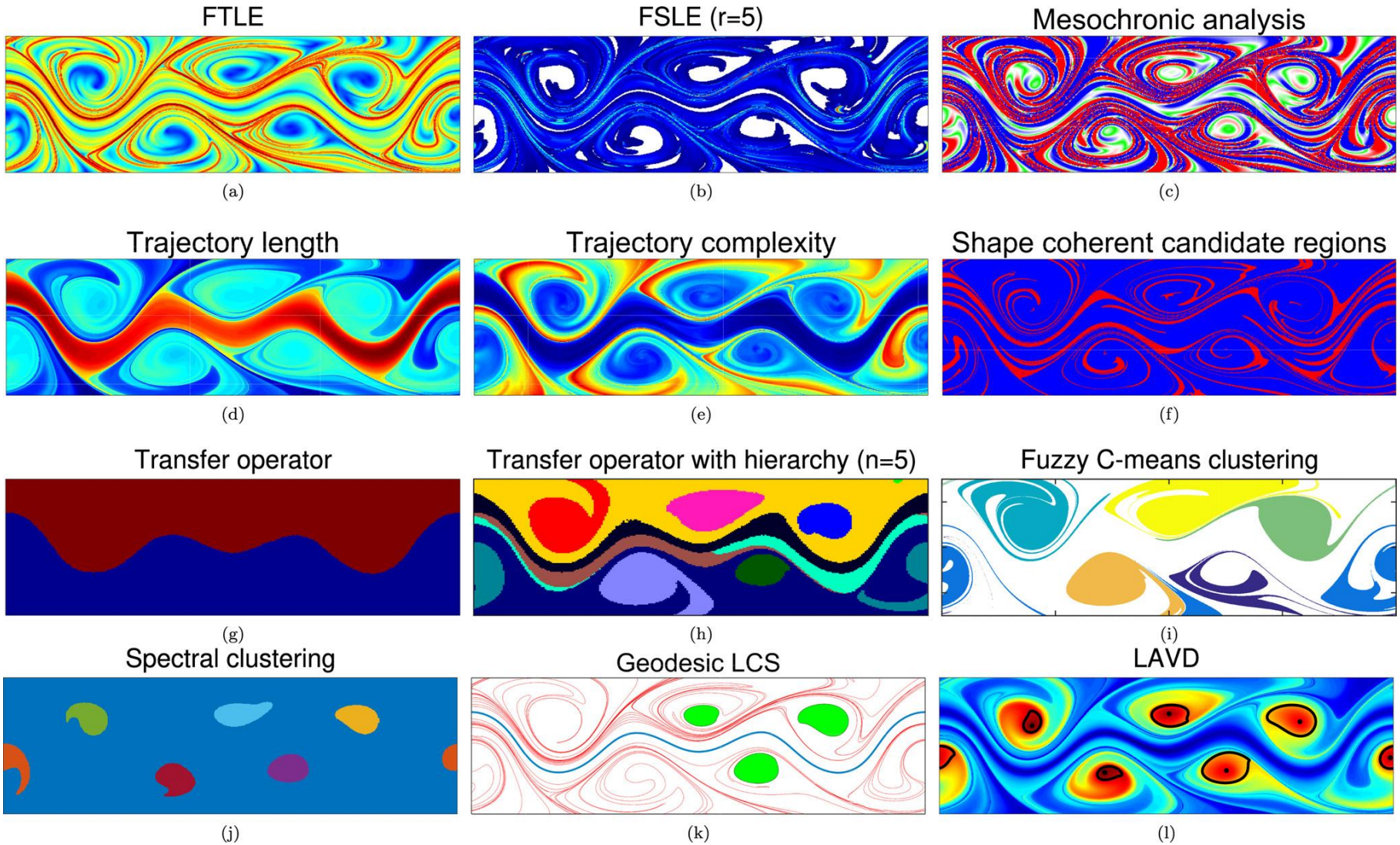


K = 11



Step 4 – Bickley Jet

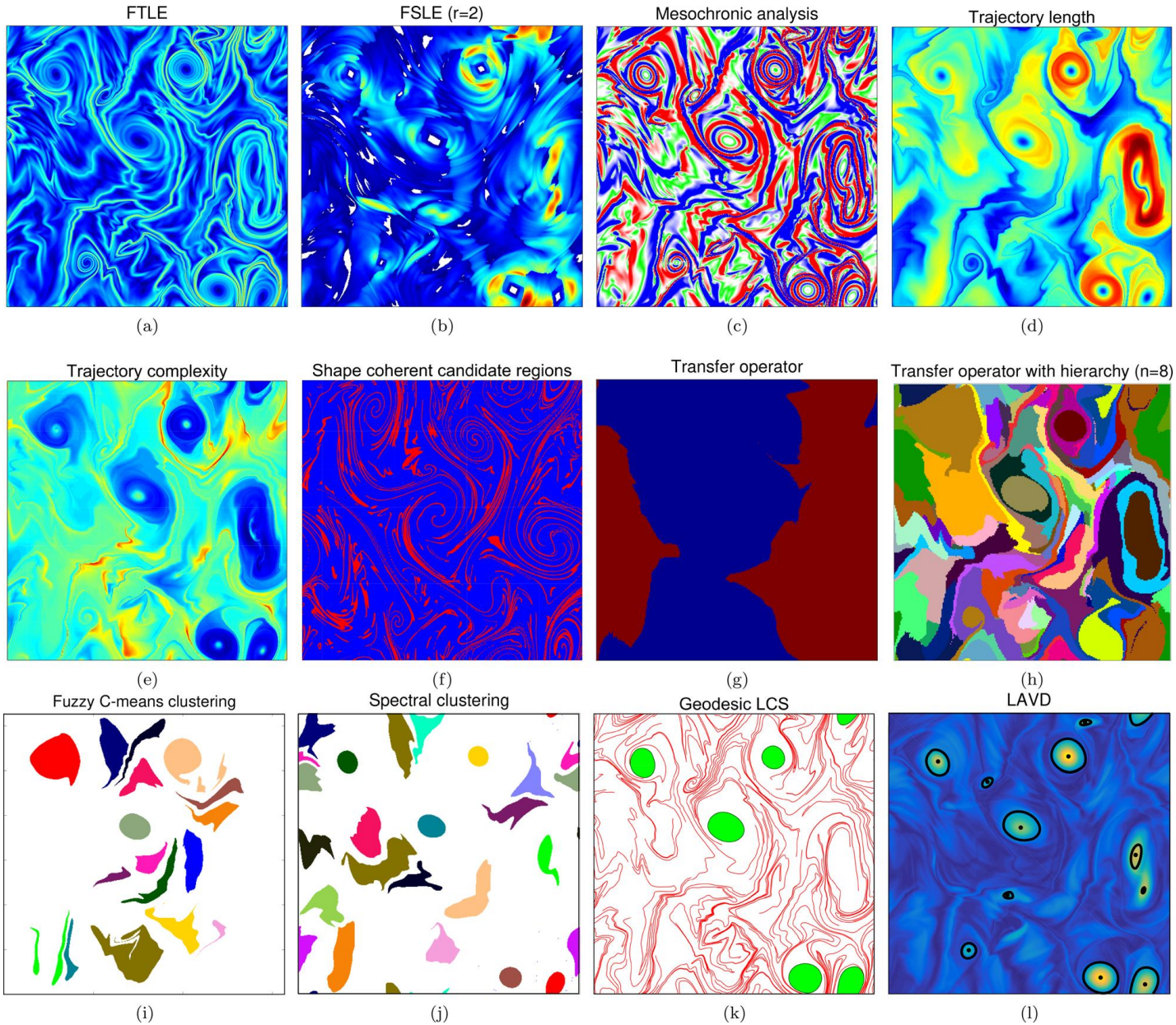
Candidate data set for SC'19 GB submission



Hadjighasem, A., Farazmand, M., Blazevski, D., Froyland, G., & Haller, G. (2017). A critical comparison of Lagrangian methods for coherent structure detection. Chaos: An Interdisciplinary Journal of Nonlinear Science, 27(5), 053104.

Step 5 – Transitional / Turbulent Flow

Candidate data set for SC'19 GB submission



The HPC Challenge

- Clustering very high-dimensional lightcone data
  - kmeans vs dbscan
- Multi-node clustering in Python
- Efficient use of memory

Outcomes

Completed Papers:

- A. Rupe, J.P. Crutchfield, K. Kashinath, and Prabhat (2017). A Physics-Based Approach to Unsupervised Discovery of Coherent Structures in Spatiotemporal Systems, Lyubchich, V., N. C. Oza, A. Rhines, and E. Szekely, eds., 2017: Proceedings of the 7th International Workshop on Climate Informatics: CI 2017. NCAR Technical Note NCAR/TN-536+PROC, 132 pp, doi:10.5065/D6222SH7. arXiv:1709.03184 [physics.flu-dyn]
- A. Rupe and J.P. Crutchfield (2018). Local Causal States and Discrete Coherent Structures. arXiv preprint arXiv:1801.00515.
  - *Accepted for publication* in Chaos: An Interdisciplinary Journal of Nonlinear Science
- A. Rupe and J.P. Crutchfield (2018). Spacetime Symmetries, Invariant Sets, and Additive Sub-Dynamics of Cellular Automata. Preparing for submission.

Planned Manuscripts:

- A. Rupe and J.P. Crutchfield (2018). Spacetime Computational Mechanics.
- Structural Semantics of Local Causal States, Part 1: Contamination.
- Structural Semantics of Local Causal States, Part 2: Coherence Detection.
- Local Causal States and Lagrangian Coherent Structures.

Outcomes

Presentations:

- Seminar – Center for Nonlinear Dynamics, UT Austin, 2016
- Talk – APS Far West Section, 2016
- Poster – AGU Fall Meeting, 2016
- Talk – Dynamics Days, 2017
- Talk – 7th Annual UC Davis Math Conference, 2017
- Poster – 7th International Workshop on Climate Informatics, NCAR, 2017
- Poster – Intel HPC Developer Conference, 2017
- Talk – 70th Annual Meeting of the APS Division of Fluid Dynamics, 2018
- Poster – Dynamics Days, 2018
- Poster – Intel AI Dev
- Talk – Information Engines at the Frontiers of Nanoscale Thermodynamics, Telluride Science Research Center

Thank You!