

A PHYSICS-BASED APPROACH TO UNSUPERVISED DISCOVERY OF COHERENT STRUCTURES IN SPATIOTEMPORAL SYSTEMS

Adam Rupe^{1,2}, James P. Crutchfield¹, Karthik Kashinath², Mr Prabhat²

Abstract—Given that observational and numerical climate data are being produced at ever more prodigious rates, increasingly sophisticated and automated analysis techniques have become essential. Deep learning is quickly becoming a standard approach for such analyses and, while great progress is being made, major challenges remain. Unlike commercial applications in which deep learning has led to surprising successes, scientific data is highly complex and typically unlabeled. Moreover, interpretability and detecting new mechanisms are key to scientific discovery. To enhance discovery we present a complementary physics-based, data-driven approach that exploits the causal nature of spatiotemporal data sets generated by local dynamics (e.g. hydrodynamic flows). We illustrate how novel patterns and coherent structures can be discovered in cellular automata and outline the path from them to climate data.

I. MOTIVATION

Incredibly complex and sophisticated models are currently employed to simulate the global climate system to facilitate our understanding of climate as well as increase our predictive power, most notably in regards to the effects of increased carbon levels. Our ability to simulate however has rapidly outpaced our ability to analyze the resulting data. Often the climate community resorts to rather simplistic data analyses, such as linear decomposition methods like EOF analyses [1], [2] or detecting (linear) trends in climate data time series [3]. Nonlinear and more sophisticated techniques are rarely brought to bear. Here we focus on one particular aspect of nonlinear dynamical systems analysis, the detection and discovery of coherent structures, such as cyclones and atmospheric rivers in climate data.

Coherent structures were introduced in the study of fluid dynamics and were initially defined as regions characterized by high levels of coherent vorticity, *i.e.*

regions where instantaneously space and phase correlated vorticity are high. The contours of coherent vorticity constitute an identifier to the structure's boundaries. However, pinning down this concept of coherent structures with rigorous and principled definitions or heuristics which produce consistent results across a wide class of physical systems is a challenging and open problem [4]. Climate practitioners are left with more ad hoc approaches [5], [6], [7] which can make it difficult to draw meaningful conclusions from analysis [8].

Deep learning attempts to sidestep this issue by learning how to identify coherent structures from labeled data [9]. However, we currently can not peer into the box to find out exactly what the defining characteristics a deep net uses to identify structures. Current state of the art achieves semi-supervised bounding box identification [10]. The ultimate goal would be unsupervised segmentation; that is, a pixel-level identification without reliance on labeled training data. It is not yet clear how to achieve this.

Like deep learning, our theory [11] approaches coherent structures from a rather different (and more general) perspective than the original context of Lagrangian coherence principles in fluid flows.

II. METHOD

Starting from basic physics principles, coherent structures can most generally be seen as *localized broken symmetries*. Two questions naturally arise; what are the symmetries which are broken and how can we identify such symmetry in a diverse range of spatiotemporal systems? Coherent structures can be found in a variety of systems with different physical properties. Convection cells in hydrodynamic systems and spiral waves in reaction-diffusion systems, for example. It is clear that the common thread is the underlying nonlinear dynamics of these systems [12], [13], [14].

Corresponding author: A Rupe, atrupe@ucdavis.edu ¹Complexity Sciences Center, Department of Physics, University of California Davis ²NERSC, Lawrence Berkeley National Laboratory

A framework known as *computational mechanics* [15], [16] has been developed to study pattern and structure in this dynamical context. The canonical object of computational mechanics is the ϵ -*machine* [17], a type of stochastic finite-state machine known as a hidden Markov model, which consists of a set of *causal states* and transitions between them. The causal states are constructed from the *causal equivalence relation*.

$$\overleftarrow{x}_i \sim_{\epsilon} \overleftarrow{x}_j \iff \Pr(\overrightarrow{X} | \overleftarrow{X} = \overleftarrow{x}_i) = \Pr(\overrightarrow{X} | \overleftarrow{X} = \overleftarrow{x}_j).$$

In words, two pasts \overleftarrow{x}_i and \overleftarrow{x}_j are *causally equivalent* if and only if they make the same prediction for the future \overrightarrow{X} ; that is, they have the same conditional distribution over the future. The causal states are the unique *minimal sufficient statistic* of the past to predict the future.

For our application to coherent structures we use a straightforward spatiotemporal generalization known as the *local causal states* [18]. For systems which evolve under some local dynamic and information propagates through the system at a finite speed, it is quite natural to use *lightcones* as local notions of pasts and futures. Formally, the past lightcone of a spacetime point $x(\vec{r}, t)$ is the set of all points at previous times that could possibly influence it. That is,

$$\ell^-(\vec{r}, t) \equiv \{x(\vec{r}', t') \mid t' \leq t \text{ and } \|\vec{r}' - \vec{r}\| \leq c(t' - t)\}$$

where c is the finite speed of information propagation in the system. Similarly, the future lightcone is given as all the points at subsequent times that could possibly be influenced by $x(\vec{r}, t)$.

$$\ell^+(\vec{r}, t) \equiv \{x(\vec{r}', t') \mid t' > t \text{ and } \|\vec{r}' - \vec{r}\| < c(t - t')\}$$

The choice of lightcone representations for both local pasts and futures is ultimately a weak-causality argument; influence and information propagate locally through a spacetime site from its past lightcone to its future lightcone.

The generalization of the causal equivalence relation is straightforward. Two past lightcones are causally equivalent if they have the same conditional distribution over future lightcones.

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

This *local causal equivalence relation* over lightcones is designed around an intuitive notion of *optimal local prediction* [18]. At some site $x(\vec{r}, t)$ in spacetime, given knowledge of all past spacetime points which could possibly affect $x(\vec{r}, t)$, i.e. its past lightcone $\ell^-(\vec{r}, t)$, what might happen at all subsequent spacetime points which could be affected by $x(\vec{r}, t)$, i.e. its future

lightcone $\ell^+(\vec{r}, t)$? Local causal states are minimal sufficient statistics for optimal local prediction. Moreover, the particular local prediction done here uses lightcone shapes, which are associated with local causality in the system. Thus it is not direct causal relationships (e.g. learning equations of motion from data) that the local causal states are discovering. Rather, they are exploiting a kind of causality in the system (i.e. that the future follows the past and that information propagates at a finite speed) in order to discover spacetime structure.

Once local causal states have been inferred from data, each site in a representative spacetime field can be assigned its local causal state label in a process known as *causal filtering* [11]. This is how we achieve unsupervised image segmentation. Though it must be clearly stated that this is a *spacetime segmentation*, and not a general image segmentation algorithm, exactly because it works *only* in systems for which lightcones are well-defined.

Using the local causal states we can, in a general and principled manner, discover dynamical spatiotemporal symmetries in a system from data. These symmetry regions are known as *domains* and are defined as regions where the associated local causal state field, after causal filtering, has spacetime symmetry tilings. A *coherent structure* is then defined as a set of spatially localized, temporally persistent (in the Lagrangian sense) non-domain local causal states.

From prior work by Hanson and Crutchfield [19], [20], [21], the domains of 1-D cellular automata are well understood as dynamically invariant sets of homogeneous spatial configurations. There is strong empirical evidence [11] that the domains of cellular automata discovered by the local causal states are *exactly* the domains as described by Hanson and Crutchfield. Therefore the local causal states are discovering spatiotemporally symmetries which are externally well-defined. In turn there is a strong agreement between the description of coherent structures in cellular automata discovered by local causal states and the coherent structures as described by Hanson and Crutchfield.

III. TOWARDS CLIMATE

With consistent and readily interpretable results on cellular automata we are now working on generalizing to real-valued spatiotemporal systems, with specific emphasis on canonical fluid flows. Others have done preliminary work on this generalization, where an extra discretization (typically via clustering) step is needed during reconstruction [22], [23].

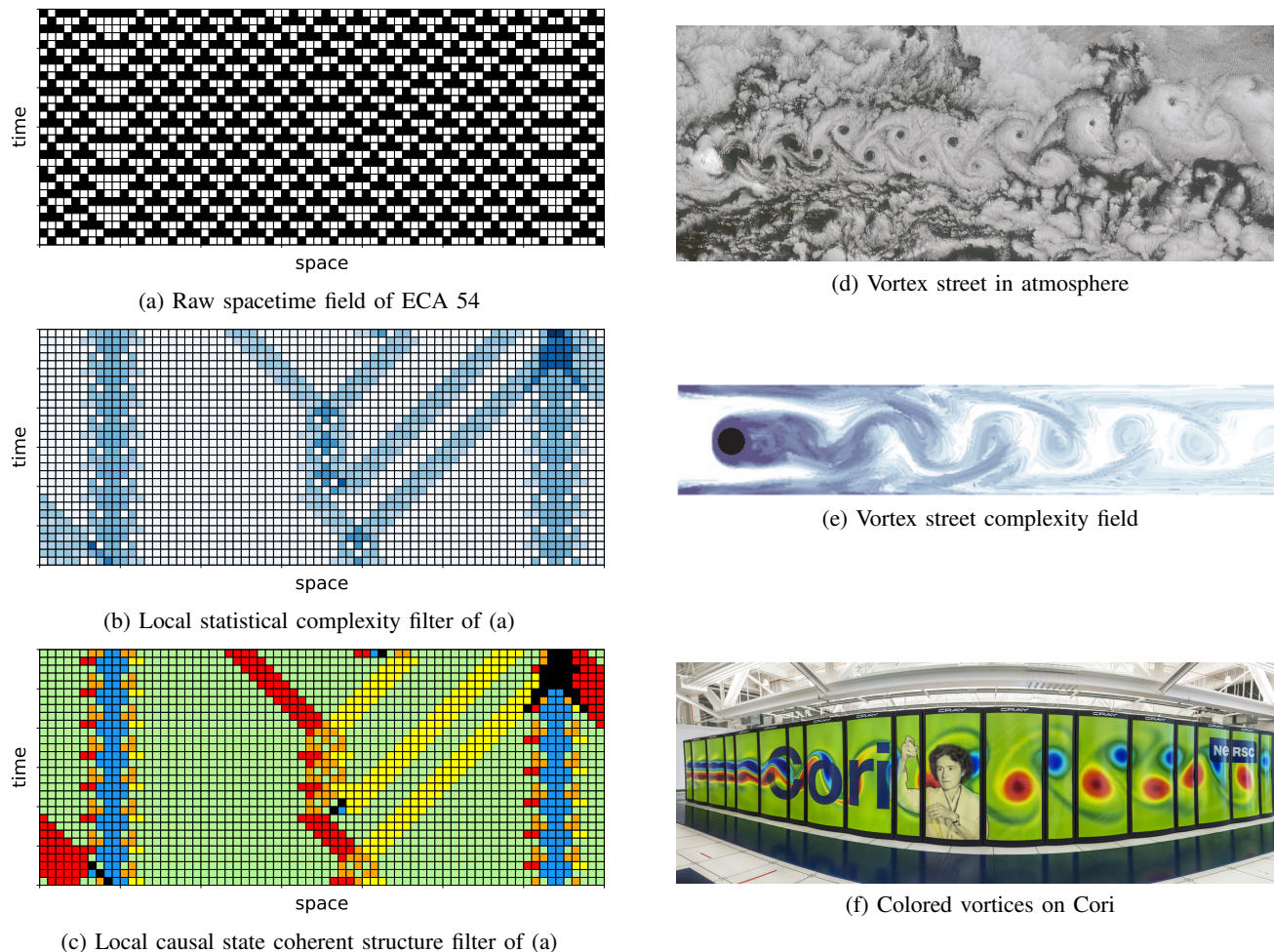


Fig. 1. Visualization of results on 1D cellular automata (fully-discrete spatiotemporal models) and projected analogous results for fluid systems. CA results for elementary cellular automaton rule 54 are given in (a)-(c). The raw spacetime field is shown in (a) and a corresponding local statistical complexity field in (b). From the local statistical complexity filter, which is a qualitative information-theoretic “rare event” filter, it is clear there are coherent structures on top of a background domain, but the four different structures can not be explicitly distinguished and identified. Thus a more detailed coherent structure filter using our unsupervised local causal state segmentation analysis is given in (c). Here states participating in the domain spacetime symmetry tiling are colored green, and other non-domain states which satisfy our definition for a coherent structure are colored according to the structure(s) they belong to. Interaction states not associated with domain or a coherent structure are in black. An outline of analogous results for vortex shedding is shown in (d)-(f). (d) A vortex street in the cloud layer over the arctic (Source: <https://photojournal.jpl.nasa.gov/catalog/PIA03448>). (e) The local statistical complexity of the vorticity field for a canonical vortex street simulation, taken from [22], analogous to the qualitative structure filter of (b). (f) Closer to the more detailed and principled coherent structure filter of (c) are the colored vortices displayed on the cover of the NERSC Cori HPC system. We emphasize the analogy is not that learning about coherent structures in CAs will give insight into fluid and climate structures. Rather, it is to illustrate how we foresee our approach will discover coherent structures in fluids and climate, in much the same way we can currently discover structures in CAs.

These groups have also used local causal states for coherent structure detection, including real-valued applications like fluids and even climate [22]. However, they have all relied on the “local statistical complexity” [24], which is the point-wise entropy over local causal states. At best this is simply a qualitative filtering tool which aides in visual recognition of structures and at worst can give both false positive and false negative misidentification. We are the first to give a principled

and rigorous method for coherent structure discovery and description using the local causal states, and are working to generalize this more detailed analysis to real-valued systems. In doing so we hope to move beyond the scope of data visualization these prior groups were working in, and facilitate novel scientific discovery, particularly in climate science.

On the theory side, we must confirm our methods on known fluid structures. As the theory is founded in

basic dynamical principles it is likely to apply without much modification in fluid systems. We will also begin to explore whether our methods can facilitate additional mechanistic insight beyond structure discovery. For example, whether there are any links between the local causal state analysis and thermodynamic considerations.

On the implementation side, the computational costs of local causal state reconstruction in more complex systems will require fully-distributed execution on large HPC machines. This will certainly be the case for TB scale climate data sets we ultimately are interested in. As our primary objective is automated coherent structure discovery, moving from canonical fluid flows to large-scale climate data will largely be a matter of computational scaling. With access to HPC experts from the Intel Big Data Center and the NERSC Cori system at Lawrence Berkeley National Laboratory we feel well-positioned to tackle these computational challenges.

ACKNOWLEDGMENTS

Adam Rupe and Jim Crutchfield would like to acknowledge Intel for supporting the IPCC at UC Davis. Prabhat and Karthik Kashinath were supported by the Intel Big Data Center. This research is based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contract W911NF-13-1-0390, and used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

REFERENCES

- [1] NCAR, "Climate data analysis tools and methods." <https://climatedataguide.ucar.edu/climate-data-tools-and-analysis>, Accessed: 2017-07-25.
- [2] M. Ghil, M. Allen, M. Dettinger, K. Ide, D. Kondrashov, M. Mann, A. Robertson, A. Saunders, Y. Tian, and F. V. P. Yiou, "Advanced spectral methods for climatic time series," *Reviews of Geophysics*, vol. 40, no. 1, 2002.
- [3] D. Shea, "Climate data analysis tools and methods - trend analysis." <https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/trend-analysis>. Accessed: 2017-07-25.
- [4] A. Hadjighasem, M. Farazmand, D. Blazeovski, G. Froyland, and G. Haller, "A critical comparison of lagrangian methods for coherent structure detection," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 27, no. 5, p. 053104, 2017.
- [5] F. Vitart, J. Anderson, and W. Stern, "Simulation of interannual variability of tropical storm frequency in an ensemble of gcm integrations," *Journal of Climate*, vol. 10, no. 4, pp. 745–760, 1997.
- [6] K. Walsh and I. Watterson, "Tropical cyclone-like vortices in a limited area model: Comparison with observed climatology," *Journal of climate*, vol. 10, no. 9, pp. 2240–2259, 1997.
- [7] Prabhat, S. Byna, V. Vishwanath, E. Dart, M. Wehner, W. D. Collins, *et al.*, "Teca: Petascale pattern recognition for climate science," in *International Conference on Computer Analysis of Images and Patterns*, pp. 426–436, Springer, 2015.
- [8] D. Faranda and D. DeFrance, "A wavelet-based approach to detect climate change on the coherent and turbulent component of the atmospheric circulation," *Earth System Dynamics*, vol. 7, no. 2, pp. 517–523, 2016.
- [9] Y. Liu, E. Racah, Prabhat, J. Correa, A. Khosrowshahi, D. Lavers, K. Kunkel, M. Wehner, and W. Collins, "Application of deep convolutional neural networks for detecting extreme weather in climate datasets," *arXiv preprint arXiv:1605.01156*, 2016.
- [10] E. Racah, C. Beckham, T. Maharaj, Prabhat, and C. Pal, "Semi-supervised detection of extreme weather events in large climate datasets," *arXiv preprint arXiv:1612.02095*, 2016.
- [11] A. Rupe and J. P. Crutchfield, "Local causal states and discrete coherent structures." <http://csc.ucdavis.edu/~cmg/compmech/pubs/dcs.htm>, 2017.
- [12] M. Cross and H. Greenside, *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, 2009.
- [13] R. Hoyle, *Pattern Formation: An Introduction to Methods*. New York: Cambridge University Press, 2006.
- [14] M. Golubitsky and I. Stewart, *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*, vol. 200. Birkhäuser, 2003.
- [15] J. P. Crutchfield and K. Young, "Inferring statistical complexity," *Phys. Rev. Lett.*, vol. 63, pp. 105–108, 1989.
- [16] J. P. Crutchfield, "Between order and chaos," *Nature Physics*, vol. 8, no. January, pp. 17–24, 2012.
- [17] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: Pattern and prediction, structure and simplicity," *J. Stat. Phys.*, vol. 104, pp. 817–879, 2001.
- [18] C. Shalizi, "Optimal nonlinear prediction of random fields on networks," *Discrete Mathematics & Theoretical Computer Science*, 2003.
- [19] J. E. Hanson and J. P. Crutchfield, "The attractor-basin portrait of a cellular automaton," *J. Stat. Phys.*, vol. 66, pp. 1415 – 1462, 1992.
- [20] J. P. Crutchfield, "Discovering coherent structures in nonlinear spatial systems," in *Nonlinear Ocean Waves* (A. Brandt, S. Ramberg, and M. Shlesinger, eds.), (Singapore), pp. 190–216, World Scientific, 1992. also appears in *Complexity in Physics and Technology*, R. Vilela-Mendes, editor, World Scientific, Singapore (1992).
- [21] J. E. Hanson and J. P. Crutchfield, "Computational mechanics of cellular automata: An example," *Physica D*, vol. 103, pp. 169–189, 1997.
- [22] H. Jänicke, A. Wiebel, G. Scheuermann, and W. Kollmann, "Multifield visualization using local statistical complexity," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1384–1391, 2007.
- [23] G. Goerg and C. Shalizi, "Mixed licors: A nonparametric algorithm for predictive state reconstruction," in *Artificial Intelligence and Statistics*, pp. 289–297, 2013.
- [24] C. Shalizi, R. Haslinger, J.-B. Rouquier, K. Klinkner, and C. Moore, "Automatic filters for the detection of coherent structure in spatiotemporal systems," *Physical Review E*, vol. 73, no. 3, p. 036104, 2006.