Spacetime Autoencoders Using Local Causal States

Adam Rupe* **and James P. Crutchfield†**
Complexity Sciences Center and Department of Physics and Astronomy, University of California Davis
One Shields Avenue, Davis CA. 95616

Abstract

Local causal-states are latent representations that capture organized pattern and structure in complex spatiotemporal systems. We expand their functionality, framing them as spacetime autoencoders. Previously, they were only considered as maps from observable spacetime fields to latent local causal-state fields. Here, we show that there is a stochastic decoding that maps back from the latent fields to observable fields. Furthermore, their Markovian properties define a stochastic dynamic in the latent space. Combined with stochastic decoding, this gives a new method for forecasting spacetime fields.

Introduction

Physics-based representation learning is a key emerging tool for analyzing high-dimensional nonlinear systems (Willard et al. 2020; Brunton, Noack, and Koumoutsakos 2020), and one with a long history in dynamical systems (Crutchfield and McNamara 1987). Nonlinear systems often require additional tools beyond computer simulation to extract actionable insight from the complex behaviors they exhibit. There are two main aspects to physics-based representation learning. The first is *dimensionality-reduction*: Learn a low-dimensional representation of a behavior—such as coherent structures in fluid flows (Holmes et al. 2012; Peacock and Haller 2013)—that provide a more human-interpretable accounting for the full system's behavior. The second is *generative modeling*: Provide a modeling alternative to numerical simulation of the equations of motion that can be applied directly to data. This is particularly helpful if the proper equations of motion are unknown. For high-dimensional systems the data-driven models are typically computationally less expensive than direct numerical simulation.

For learning representations of high-dimensional nonlinear dynamics, several approaches have been introduced. *Proper Orthogonal Decomposition* (POD)—rather similar to *Principle Component Analysis* (PCA)—is a canonical method for fluid flows (Holmes et al. 2012; Rowley and Dawson 2017). POD modes provide a low-dimensional latent representation that can give interpretable insights into a flow's large-scale organization. Flow equations of motion may also incorporate a set of POD modes, through Galerkin projection, giving a truncated set of ODEs as a generative model that is less expensive to simulate than the full Navier-Stokes PDE. Like PCA, the POD modes are linear latent representations. For that matter, PCA is equivalent to a linear autoencoder (Baldi and Hornik 1989). Representation linearity substantially restricts the complexity of flow structures that can be modeled appropriately.

Addressing this, the Perron-Frobenius and Koopman operators recently gained popularity as nonlinear generalizations for spectral (modal) analysis of high-dimensional dynamical systems (Froyland and Padberg 2009; Mezić 2013; Klus et al. 2019). As with nonlinear autoencoders, these operators' modes are typically used for nonlinear dimensionality-reduction.

Similarly, the Koopman operator (Alexander and Giannakis 2020) and autoencoders (Hernández et al. 2018) are now used for generative modeling which, in the dynamical systems setting, is a form of *predictive forecasting*. As an aside, reservoir computing was shown to be effective for predictive modeling (Pathak et al. 2018). It is somewhat analogous to Koopman operator approaches, as the system dynamics are learned in a higher-dimensional latent space.

Success with operator-approximation methods turns on a fortuitous matching of their chosen (or inherent) function-basis dictionary and a system's emergent structures. The reality, though, is that spatially-extended nonlinear systems generate a diverse set of complicated organizations—vortices, target patterns, dislocations, and the like. In short, these emergent structures are not easily or naturally modeled in terms of known or numerically-approximated spatially-global function bases. Recent work employs nonlinear autoencoders (Lusch, Kutz, and Brunton 2018; Mardt et al. 2018; Otto and Rowley 2019) as a means to bypass explicit dictionary choices. Issues still persist with these methods, however; finding the "best" finite-dimensional operator approximations for a given application remains an open problem.

Local causal states are yet another tool for physics-based

*atrupe@ucdavis.edu
†chaos@ucdavis.edu

representation learning. They start from a markedly different conceptualization of latent space, however, with the promise of learning the more complex emergent patterns generated by nonlinear pattern-forming systems. Recall that reducing dimension through spectral decomposition implicitly assumes algebraic representations that are spatially global and spatially coherent. In contrast, local causal states are spatially-local latent representations that are learned at each point in spacetime. As such, they are better adapted to capture structures that self-organize from local interactions governing the dynamics—structures that consist of many localized or "coherent" substructures (Rupe and Crutchfield 2018; Rupe et al. 2019). Crucially, since the representations are learned locally, the latent space shares the same coordinate geometry as the observable spacetime fields. This adds, among other benefits, a helpful visual interpretability. The following formally connects local causal states and spacetime autoencoders and presents preliminary results for predictive forecasting using them.

Local Causal States

The local causal states are part of the *computational mechanics* framework (Crutchfield 2012), which learns nonparametric models of dynamical systems in an unsupervised fashion using the *causal equivalence relation*:

$$\mathrm{p}_t \sim_\epsilon \mathrm{p}_{t'} \iff \Pr(\mathrm{Future}_t|\mathrm{p}_t) = \Pr(\mathrm{Future}_{t'}|\mathrm{p}_{t'}) \,.$$

The induced equivalence classes over pasts $\{\mathrm{p}_t\}$ are a system's *causal states*—the unique minimal sufficient statistic of the past for optimally predicting the future.

For spatiotemporal systems, *lightcones* are local features that represent pasts and futures; see Fig. 1. Lightcones capture the history and propagation of local interactions in the system through space and time. A lightcone *configuration* is an assignment of observable values to the lightcone templates shown in Fig. 1. Two past lightcone configurations ℓ_i^- and ℓ_j^- are causally equivalent if they have the same conditional distribution over future lightcones L^+:

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

The resulting equivalence classes are the system's *local causal states* (Shalizi 2003). They are the unique minimal sufficient statistic of past lightcones for optimally predicting future lightcones.

Encoding

The lightcone equivalence relation can be recast as a function that generates the causal equivalence classes—the ϵ-*function* $\epsilon : \ell^- \mapsto \xi$ maps from past lightcone configurations ℓ^- to local causal states ξ. This local mapping from observables to their corresponding latent local causal-state representation is central to using and interpreting our method. Specifically, $\epsilon\big(\ell^-\big)$ is applied in parallel to all points $(\vec{r}, t)$ in a spacetime field X, mapping the entire field to its associated *local causal state field* $S = \epsilon(X)$. Every feature $X(\vec{r}, t)$ is mapped to its latent variable (local causal state) via its past lightcone $\xi = S(\vec{r}, t) = \epsilon\big(\ell^-(\vec{r}, t)\big)$. One result is that the global latent spacetime field S maintains X's

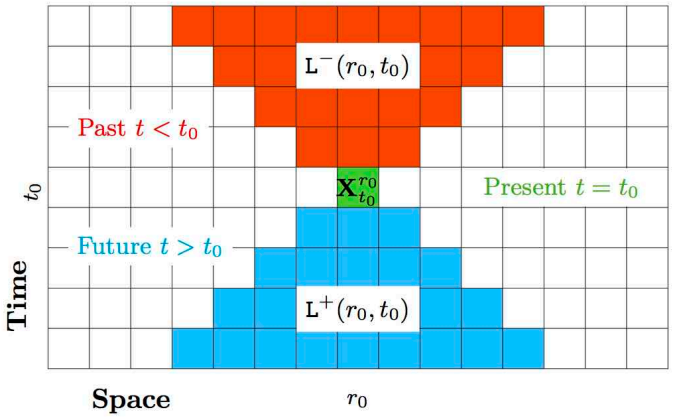

Figure 1: Lightcones: past lightcone $\mathtt{L}^-(r_0, t_0)$ (red) and future lightcone $\mathtt{L}^+(r_0, t_0)$ (blue) shown for spacetime point $X(r_0, t_0)$ (green) for a system with radius $R = 1$ interactions, such as the map lattice example shown below.

coordinate geometry such that $S(\vec{r}, t)$ is the local latent variable corresponding to the local observable $X(\vec{r}, t)$. (This is markedly unlike neural network autoencoders.) The shared geometry of the observable space X and the latent space S facilitates extracting physical features of X from special (e.g., algebraic) properties in the corresponding spacetime region in S.

For real-valued systems, such as the map lattice analyzed shortly, local causal state inference requires a discretization to empirically estimate $\Pr(\mathrm{L}^+|\ell^-)$ (Goerg and Shalizi 2012). Rather than discretize observable space, we discretize in the lightcone feature space using K-Means to cluster lightcones using distance metric:

$$\mathrm{D}_{\mathrm{lc}}(\mathbf{a}, \mathbf{b}) \equiv \sqrt{(a_1 - b_1)^2 + \ldots + \mathrm{e}^{-\tau d(n)}(a_n - b_n)^2}, \quad (1)$$

where $\mathbf{a}$ and $\mathbf{b}$ are flattened lightcone vectors, $d(n)$ is the temporal depth of the lightcone vector at index n, and τ is the temporal decay rate ($1/\tau$ can be thought of as a coherence time).

Two past lightcones are considered γ^--equivalent if they are placed into the same cluster C^- by the distance-based clustering:

$$\gamma^-(\ell_i^-) = \gamma^-(\ell_j^-) \iff \ell_i^- \in C^- \text{ and } \ell_j^- \in C^- \,. \quad (2)$$

Similarly for future lightcones.

This allows us to build empirical distributions over lightcone-clusters from simple counting. Two past lightcone clusters are considered ψ-equivalent if they have (approximately) the same empirical predictive distributions:

$$\psi(C_i^-) = \psi(C_j^-) \iff \Pr(C^+|C_i^-) \approx \Pr(C^+|C_j^-) \,. \quad (3)$$

This approximates the ϵ-function as:

$$\epsilon(\ell^-) \approx \psi\big(\gamma^-(\ell^-)\big) \,. \quad (4)$$

Decoding

Previously, when performing coherent-structure segmentation the ϵ-function *encoded* observable spacetime fields X

to a corresponding latent local causal state field S (Rupe and Crutchfield 2018; Rupe et al. 2019). Such dimensionality-reduction identifies coherent structures pointwise in observable spacetime fields. We now introduce, for the first time, the ϵ^{-1}-function—a stochastic *decoding* that maps from latent spacetime fields S to *reconstructed* observable fields $\overline{X}$.

Each local causal state ξ in S is defined by its predictive distribution $\Pr(\mathtt{L}^+|\xi)$, since every past lightcone configuration ℓ_i^- in ξ has, by definition, the same predictive distribution: $\Pr(\mathtt{L}^+|\ell_i^-) = \Pr(\mathtt{L}^+|\xi)$, for all $\ell_i^- \in \xi$. The decoding $\overline{X} = \epsilon^{-1}(S)$ is performed by sampling the distributions $\Pr(\mathtt{L}^+|\xi)$ for each ξ in S as follows. For each spacetime coordinate $(\mathbf{r}, t)$:

1. Retrieve the local causal state $\xi = S(\mathbf{r}, t)$;
2. Sample a future cluster C^+ from $\Pr(C^+|\xi)$;
3. Retrieve the centroid $\overline{\ell^+}$ of C^+; and then
4. Place $\overline{\ell^+}$ in $\overline{X}$, with base at $(\mathbf{r}, t)$.

Since we use K-Means to cluster both past and future lightcones, we take the centroid $\overline{\ell^+}$ as a representative real-valued future lightcone for cluster C^+. Due to the spacetime extent of nontrivial future lightcones, each point $\overline{X}(\mathbf{r}, t) \in \overline{X}$ makes predictions from several local causal states. In fact, for future-lightcone depth h^+ a prediction will be made from each point in its h^+-depth past lightcone. The ultimate prediction for $\overline{X}(\mathbf{r}, t)$ averages these predictions, with the same time-exponential weighting used in the lightcone metric of Eq. (1). That is, predictions made further out in time are discounted exponentially compared to more recent predictions.

Combining the ϵ-function encoding with the ϵ^{-1}-function decoding, the local causal states form a *spacetime autoencoder*. As shown in the top portion of Fig. 2, the ϵ-function encodes an observable spacetime field X to the compressed latent field S and ϵ^{-1} decodes to a reconstructed observable field $\overline{X} = \epsilon^{-1}(S)$. Said another way, the identity $I \approx \epsilon \circ \epsilon^{-1}$ is learned through the causal state bottleneck S.

There are several points to emphasize. First, unlike neural network autoencoders, the local causal states are nonparametric models. And so, rather than using the encoding and decoding together to train parameters as in neural network autoencoders, the ϵ-map and its inverse are learned directly by approximating the local causal equivalence relation from data. Second, as already noted, a crucial distinction is that the ϵ-map encoding is done locally so that the latent space and observable space share spacetime coordinate geometry. The latent space of neural network autoencoders, in contrast, does not share geometry with its inputs due to how their bottleneck is created. For local causal states approximated from real-valued spacetime data, the bottleneck comes from having a finite number of latent local causal states and the latter are determined by the inherent structural dynamics. However, this alone is a rather weak notion of autoencoder. Last, along these lines, accounting for temporal evolution is critical for capturing pattern and structure that is spontaneously generated by complex spatiotemporal systems. Accounting for dynamics leads to a more powerful view of a causal state autoencoder.

Latent Space Dynamics

In fact, a stochastic dynamic can be defined over the local causal states. This combined with ϵ^{-1} decoding gives spacetime forecasting: infer the local causal states and their dynamics up to the present time, evolve the states forward in time, then decode to a forecasted observable field. This view of the local causal states as *predictive spacetime autoencoders* is much more useful than the weak notion above. It synthesizes the two main aspects of representation learning in physics—dimensionality-reduction and generative modeling. For nonlinear dynamical systems the generative modeling of interest is predictive forecasting, which generates sequential states and captures how they are correlated and organized in time.

The local causal state field exhibits *Markov shielding* (Shalizi 2003). The ϵ-function uniquely determines a local causal state $\xi = S(\mathbf{r}, t)$ from a full past lightcone $\ell^-(\mathbf{r}, t)$ of observables. If the local causal states are known for each point in $\ell^-(\mathbf{r}, t)$, only the states in the immediately preceding time step are required to determine ξ. That is, $S(\mathbf{r}, t)$ is independent of the local causal states in $\ell^-(\mathbf{r}, t)$, given the local causal states in its depth-1 past lightcone—the *local neighborhood* $\eta\big(S(\mathbf{r}, t-1)\big)$ of $S(\mathbf{r}, t-1)$.

Markov shielding was derived in the setting of stochastic field theories, where the observable field X is *not* from a deterministic dynamical system. The stochasticity inherent in the system, along with Markov shielding, implies the temporal dynamics over the local causal states is a *stochastic cellular automaton* (SCA), where $S(\mathbf{r}, t)$ is given by a stochastic function ϕ of the local neighborhood $\eta\big(S(\mathbf{r}, t-1)\big)$.

While our ultimate interests lie in deterministic systems (e.g., those governed by partial differential equations), the local causal state dynamics is still an SCA in this case. For systems with finite-range local interactions, which *define* a system's lightcones, a site value $X(\mathbf{r}, t)$ in X is uniquely determined by the past lightcone of $X(\mathbf{r}, t)$. In fact, for deterministic, non-delay dynamics $X(\mathbf{r}, t)$ is uniquely determined by $\eta\big(X(\mathbf{r}, t-1)\big)$. In this way, Markov shielding in the local causal states is inherited from the dynamics of the observable field. Since the local causal states are compressed (local) representations, the local causal states alone are not sufficient for deterministic evolution and, in general, the information loss in the bottleneck implies the dynamic Φ over the local causal states must be stochastic. Note that this argument is agnostic to the functional form of the observable dynamics (i.e., the equations of motion). In this way, the local causal states and their SCA dynamic Φ provide a universal probabilistic model for spatially-extended dynamical systems.

Consider a finite set Ξ of symbols (e.g., local causal state labels). A radius-R *deterministic* cellular automata (CA) over Ξ is specified by a *local update rule* $\phi : \eta \mapsto \xi$ that is a deterministic function of radius-R neighborhoods η. In $1+1$-dimensions (one space, one time), the neighborhoods are tuples of symbols from Ξ. For example, in a radius-1 CA in $1+1$-dimensions: $\xi(\mathbf{r}, t+1) = \phi\big(\eta(\xi(\mathbf{r}, t))\big) = \phi\big(\xi(\mathbf{r}-1, t), \xi(\mathbf{r}, t)\xi(\mathbf{r}+1, t)\big)$. The global CA dynamic Φ, that evolves spatial configurations over a time step, ap-

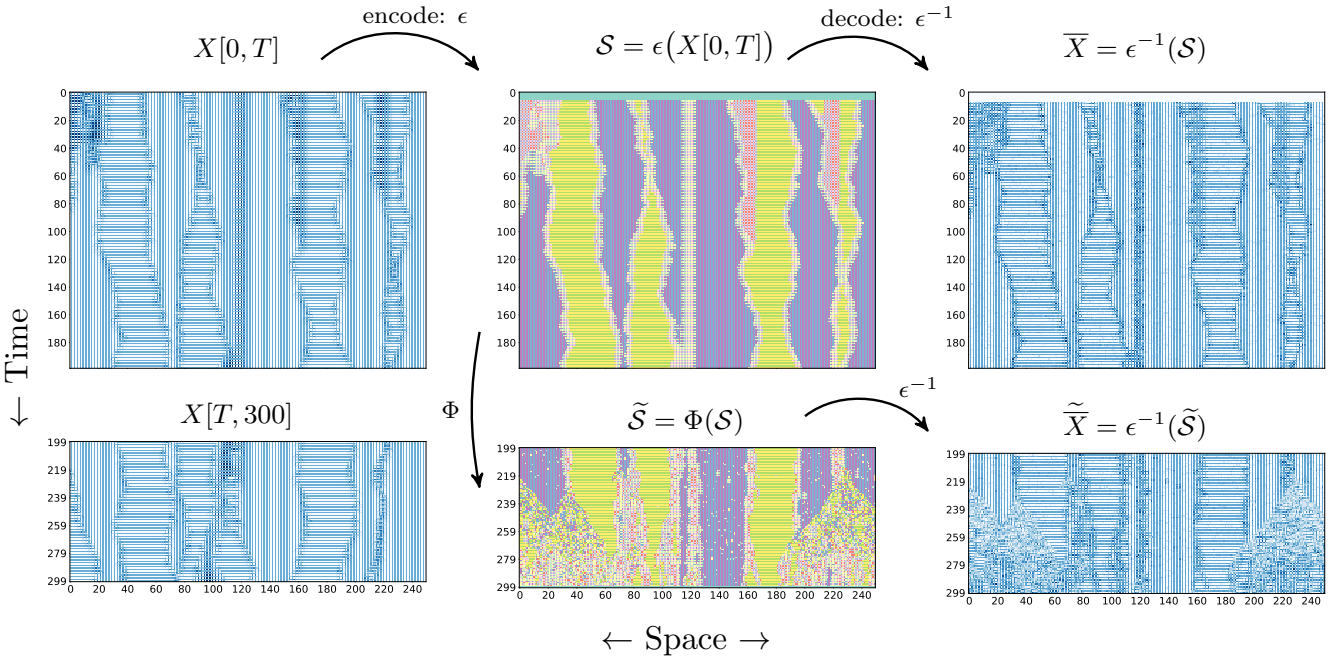

Figure 2: Spacetime local causal-state autoencoder: (Top-left) Observable spacetime field $X[0,T]$ of the circle map lattice used for training. A 250-site portion of the $20,000$ spatial sites are shown (horizontal), which are evolved for an initial 200 time steps (vertical) for training. (Top-middle) Encoded latent local causal state field $S = \epsilon\big(X[0,T]\big)$. Each unique color represents a unique local causal state. This compressed latent space provides a compressed representation of $X[0,T]$ for dimensionality-reduction. (Top-right) Reconstructed observable field $\overline{X} = \epsilon^{-1}(S)$, decoded from S. (Bottom-middle) Forecasted local causal state field, evolved forward in time after inference, starting at $T = 199$, using $\widetilde{S} = \Phi(S)$ as described above. (Bottom-right) Forecasted observable field $\widetilde{\overline{X}}$, decoded from the evolved local causal state field $\widetilde{\overline{X}} = \epsilon^{-1}(\widetilde{S})$. (Bottom-left) Ground-truth observable field $X[T, 300]$, evolved using the circle map lattice equation of motion.

plies the local update ϕ synchronously and in parallel across a configuration. For the more general stochastic CAs, the local update ϕ is still a function of local neighborhoods η. However, instead of outputting a single symbol ξ, it outputs a *probability mass function* (PMF) over symbols from $\mathbf{\Xi}$. Deterministic CAs then are the special case when the PMFs assign unity mass to and only one symbol.

For a given observable spacetime field X, once the associated local causal state field $S = \epsilon(X)$ has been inferred, the state dynamic Φ can be estimated from S, again by simple counting. Empty histograms are initialized for all possible neighborhoods η of local causal states. For a particular neighborhood η_i found at $S(\mathbf{r}, t)$, the histogram of $\phi(\eta_i)$ is incremented at ξ for $\xi = S(\mathbf{r}, t+1)$.

Since real-valued systems require approximation schemes and models are always inferred from finite data, using Φ estimated in this way to evolve the local causal states forward in time may yield a neighborhood η not seen in the original inferred field S. The local dynamic ϕ outputs an empty histogram in this case. In practice, to circumvent this issue we keep and update a separate PMF over local causal states that is the spatial distribution over the local causal states.

At each t we estimate a histogram over the local causal states according to the spatial configuration $S(t)$ of the local causal state field. If an empty histogram is encountered during the evolution $S(t+1) = \Phi\big(S(t)\big)$, a local causal state for that point is chosen from the spatial PMF instead. This heuristic leaves room for variation, and we use a running-estimate spatial distribution since the distribution over local causal states generally is not temporally stationary.

Forecasting Spacetime Fields

To demonstrate that local causal states are spacetime autoencoders, along with their predictive forecasting, the following presents preliminary results for a nonlinear map lattice system based on the *circle map*. While the choice is somewhat arbitrary, we selected the circle map lattice due to the complexity of the self-organized patterns and structures it generates. Map lattices are also markedly simpler to simulate than PDEs, such as the Kuramoto-Sivashinsky equation popular in data-driven forecasting explorations (Pathak et al. 2018; Otto and Rowley 2019), that require elaborate numerical integration schemes.

A one-dimensional *map lattice* (Crutchfield and Kaneko 1987) is a spatially-extended dynamical system that evolves configurations on a discrete spatial lattice $\mathbb{Z}$ in discrete time

steps according to the local dynamics:

$$x(\mathbf{r}, t+1) = (1-\alpha) f\big(x(\mathbf{r}, t)\big) + \frac{\alpha}{2}\left[f\big(x(\mathbf{r}+1, t)\big) + f\big(x(\mathbf{r}-1, t)\big)\right], \quad (5)$$

where $\mathbf{r}$ is the spatial index, t is the time index, α is the coupling strength, and f is an iterated map of the unit interval: $x(t+1) = f\big(x(t)\big),\ x \in [0,1]$. We use the circle map:

$$f(x) = x + \omega - \frac{K}{2\pi}\sin(2\pi x) \bmod 1\ ,$$

where ω is a phase shift and K is the strength of the nonlinearity. Following (Gravner and Johnson 2018), Fig. 2 uses $\omega = 0.5$, $K = 1.0$, and $\alpha = 1.0$.

The map lattice's complex behaviors seen there in the observable spacetime field X in the left panels arise from two competing background *domains*. One is spatial period-2 (vertical stripes) and the other is temporal period-2 (horizontal stripes). Interfaces between these domains (also know as *dislocations* in the statistical mechanics literature) diffuse through space over time and sometimes pairwise annihilate upon collision. Due to pairwise annihilation, the macroscopic behavior of domains and the interactions of their interfaces is not stationary in time.

When evolved from random initial conditions, as done here, there are initially many domain interfaces that quickly annihilate. These pairwise interactions decrease over time until the system's spacetime behavior is mostly comprised of domain regions. We are interested here in intermediate times that have a balance between dislocation dynamics and (meta-)stable domain regions. To this end, we let the map lattice evolve for 300 time steps before starting local causal state inference. Note that time $t = 0$ in Fig. 2 starts after this initial transient time. We use a lattice size of $N = 20,000$ (only a portion of which is shown in Fig. 2) and periodic boundary conditions for our map lattice.

After the initial 300-step transient time, the map lattice is evolved for another $T = 200$ time steps to produce the training observable field $X[0,T]$ that is used for local causal state inference. For inference we use past lightcone depth $h^- = 6$, future lightcone depth $h^+ = 1$, and propagation speed $c = 1$ (determined by the radius $R = 1$ local interactions of the map lattice). In the lightcone clustering step, i.e. γ-equivalence, we use $K = 10$ for past lightcone clustering and $K = 40$ for future lightcone clustering. For ψ-equivalence, we use hierarchical agglomerative clustering, using a χ^2 test with $\alpha = 0.05$ for distribution comparison.

After inference, the training observable field $X[0,T]$ is encoded using the approximated ϵ-map, $\epsilon(\ell^-) \approx \psi\big(\gamma^-(\ell^-)\big)$, to produce the associated local causal state field $S = \epsilon\big(X[0,T]\big)$, shown in the top-middle panel of Fig. 2. Each unique local causal state is assigned an arbitrary integer label during inference, and these integer labels are then assigned arbitrary colors for the latent field visualizations in Fig. 2. Each unique color seen in S identifies a unique local causal state at that point.

As the two domains in X are period-2, each is described by two local causal states in $S = \epsilon(X)$. The result is that only four local causal states capture the majority of the observed spacetime behavior, with additional local causal states capturing particular interactions at the domain interfaces. From the dimensionality-reduction perspective, these are the structures that one would like to capture and in terms of which we would then re-express the system's evolution. Note that at the "microscopic" level, in terms of the full spatial lattice, the dynamics of the system are deterministic, following Eq. (5), while the evolution of the reduced "macroscopic" description is probabilistic, following the stochastic diffusion and annihilation of domain interfaces.

Recall that the ϵ^{-1}-function is also learned during inference, and it is then used to create the reconstruction $\overline{X} = \epsilon^{-1}(S)$ of $X[0,T]$, shown in the top-right panel of Fig. 2. The reconstructed observable field $\overline{X}$ qualitatively reproduces the macroscopic behavior observed in $X[0,T]$, even though the microscopic details are noisy due to the stochasticity of the decoding. Given that the latent space S is comprised of just 10 local causal states, the reconstructed observable field $\overline{X} = \epsilon^{-1}(S)$ is surprisingly faithful to the original observable field $X[0,T]$.

Once the local causal state field S has been inferred, the stochastic latent space dynamic Φ can be estimated, as described above. Using the estimated dynamic, the latent space is evolved forward in time to produce a forecasted local causal state field $\widetilde{S} = \Phi(S)$. In the bottom-middle of Fig. 2 the local causal state field is predicted forward in time another 100 time steps. Note that in the original inferred latent space field $S = \epsilon\big(X[0,T]\big)$ no local causal states are placed in regions that do not have full past or future lightcones of the chosen depths. These points are known collectively as the *margin* of the latent space field S. Due to the periodic boundaries in space, there are only time margins. Margin points are assigned a unique label in S, rather than just being omitted, so that S is the same shape as the observable field it is trained on. Since we use future lightcone depth $h^+ = 1$ here, there are no local causal states assigned at time step 199. Thus, as can be seen in the bottom panels of Fig. 2, the forecast starts at $T = 199$.

There are several points to note for the forecasted local causal state field $\widetilde{S} = \Phi(S)$ in the bottom-middle panel. We again emphasize that the following analysis is enabled and enhanced by the visual interpretability afforded by the shared coordinate geometry of the observable and latent spaces. From visual inspection, we can see that the most stable and well-predicted regions are the periodic domains. By relying on the previous-time spatial distribution when null PMFs are encountered applying Φ, the domain regions are stable to incorrect predictions, to some degree. We can see this, for instance, in the predicted domain region around site $r = 140$. There is noise present in the prediction of this domain, as there are local causal states present other than the two states that identify this domain in the same region in the inferred field $S = \epsilon\big(X[0,T]\big)$ just before the forecast starts. The domain persists in the prediction, despite this noise.

However, resilience to incorrect predictions does not hold generally. As can be seen, for instance on the right-hand side of $\widetilde{S}$ after $t \approx 220$, prediction errors can sometimes cause

cascading failures in the forecast, which typically travel around the propagation speed of $c = 1$. Once a cascading failure starts, it is generally not recoverable due to instability in the inferred dynamic and the forecast breaks down.

Again using the ϵ^{-1}-function, we decode $\widetilde{S}$ to produce a forecasted observable field $\widetilde{\overline{X}} = \epsilon^{-1}(\widetilde{S})$. As with the reconstructed field $\overline{X}$, the small-scale details are noisy, but the large-scale behavior predicted by the local causal state field $\widetilde{S}$ is reproduced in the observable forecast. Comparing to the ground-truth observable field $X[T, 300]$, provided by evolving the map lattice forward another 100 time steps after creating the training field, we see that the domain regions are forecasted well. Several interfaces remain stable for the full duration without breaking down into a cascading failure, such as those near $r = 140$.

Ideally, the forecast should not produce the kind of cascading failures seen in $\widetilde{S}$ and $\widetilde{\overline{X}}$. These are likely due to empty PMFs encountered when applying the estimated latent space dynamic Φ. Clearly, there is room for improvement in the algorithms, heuristics, and training protocols. These preliminary results nonetheless demonstrate the twofold benefit of the local causal states for physics-based representation learning: dimensionality-reduction and generative modeling—predictive forecasting, in this case. However the method is improved, we suspect there will ultimately be a trade-off between these two aspects of local causal states. Allowing for more local causal states to be learned during inference (the hyperparameter choice of K in K-Means cluster over past lightcones sets an upper bound to the possible number of states that can be learned) will generally increase forecasting performance, but at the same time detract from with visual interpretability of the latent space for dimensionality-reduction.

Python source code and a Jupiter notebook that produces the results shown in Fig. 2 is available at https://github.com/adamrupe/spacetime_autoencoders.

Acknowledgments AR and JPC acknowledge Intel's support for the IPCC at UC Davis and thank the Telluride Science Research Center for its hospitality during visits. 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-18-1-0028. Part of this research was performed while AR was visiting the Institute for Pure and Applied Mathematics (IPAM), which is supported by the National Science Foundation (Grant No. DMS-1440415).

References

Alexander, R., and Giannakis, D. 2020. Operator-theoretic framework for forecasting nonlinear time series with kernel analog techniques. *Physica D: Nonlin. Phenom.* 132520.

Baldi, P., and Hornik, K. 1989. Neural networks and principal component analysis: Learning from examples without local minima. *Neural Net.* 2(1):53–58.

Brunton, S. L.; Noack, B. R.; and Koumoutsakos, P. 2020. Machine learning for fluid mechanics. *Ann. Rev. Fluid Mech.* 52:477–508.

Crutchfield, J. P., and Kaneko, K. 1987. Phenomenology of spatio-temporal chaos. In Bai-lin, H., ed., *Directions in Chaos*. Singapore: World Scientific Publishers. 272.

Crutchfield, J. P., and McNamara, B. S. 1987. Equations of motion from a data series. *Complex Systems* 1:417 – 452.

Crutchfield, J. P. 2012. Between order and chaos. *Nature Physics* 8(January):17–24.

Froyland, G., and Padberg, K. 2009. Almost-invariant sets and invariant manifolds — connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D: Nonlin. Phenom.* 238(16):1507 – 1523.

Goerg, G. M., and Shalizi, C. R. 2012. LICORS: Light cone reconstruction of states for non-parametric forecasting of spatio-temporal systems. *arXiv:1206.2398*.

Gravner, J., and Johnson, K. 2018. Coupled map lattices as musical instruments. *Computer Music J.* 42(02):22–34.

Hernández, C. X.; Wayment-Steele, H. K.; Sultan, M. M.; Husic, B. E.; and Pande, V. S. 2018. Variational encoding of complex dynamics. *Phys. Rev. E* 97(6):062412.

Holmes, P.; Lumley, J.; Berkooz, G.; and Rowley, C. 2012. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge, United Kingdom: Cambridge University Press.

Klus, S.; Husic, B. E.; Mollenhauer, M.; and Noé, F. 2019. Kernel methods for detecting coherent structures in dynamical data. *Chaos: Interdisc. J. Nonlin. Sci.* 29(12):123112.

Lusch, B.; Kutz, J. N.; and Brunton, S. L. 2018. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Comm.* 9(1):1–10.

Mardt, A.; Pasquali, L.; Wu, H.; and Noé, F. 2018. VAMPnets for deep learning of molecular kinetics. *Nature Comm.* 9(1):1–11.

Mezić, I. 2013. Analysis of fluid flows via spectral properties of the Koopman operator. *Ann. Rev. Fluid Mech.* 45:357–378.

Otto, S. E., and Rowley, C. W. 2019. Linearly recurrent autoencoder networks for learning dynamics. *SIAM J. App. Dynamical Sys.* 18(1):558–593.

Pathak, J.; Hunt, B.; Girvan, M.; Lu, Z.; and Ott, E. 2018. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Phys. Rev. Let.* 120(2):024102.

Peacock, T., and Haller, G. 2013. Lagrangian coherent structures: The hidden skeleton of fluid flows. *Physics Today* 66(2):41–47.

Rowley, C. W., and Dawson, S. T. 2017. Model reduction for flow analysis and control. *Annual Review of Fluid Mechanics* 49:387–417.

Rupe, A., and Crutchfield, J. P. 2018. Local causal states and discrete coherent structures. *Chaos* 28(7):1–22.

Rupe, A.; Kumar, N.; Epifanov, V.; Kashinath, K.; Pavlyk, O.; Schlimbach, F.; Patwary, M.; Maidanov, S.; Lee, V.;

Prabhat; and Crutchfield, J. P. 2019. Disco: Physics-based unsupervised discovery of coherent structures in spatiotemporal systems. In *IEEE/ACM Workshop on Mach. Learn. High Perf. Comput. Envt.*, 75–87. IEEE.

Shalizi, C. 2003. Optimal nonlinear prediction of random fields on networks. *Discrete Math. Theo. Comp. Sci.*

Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; and Kumar, V. 2020. Integrating physics-based modeling with machine learning: A survey. *arXiv preprint arXiv:2003.04919*.