The LICORS Cabinet: Nonparametric Light Cone Methods for Spatio-Temporal Modeling

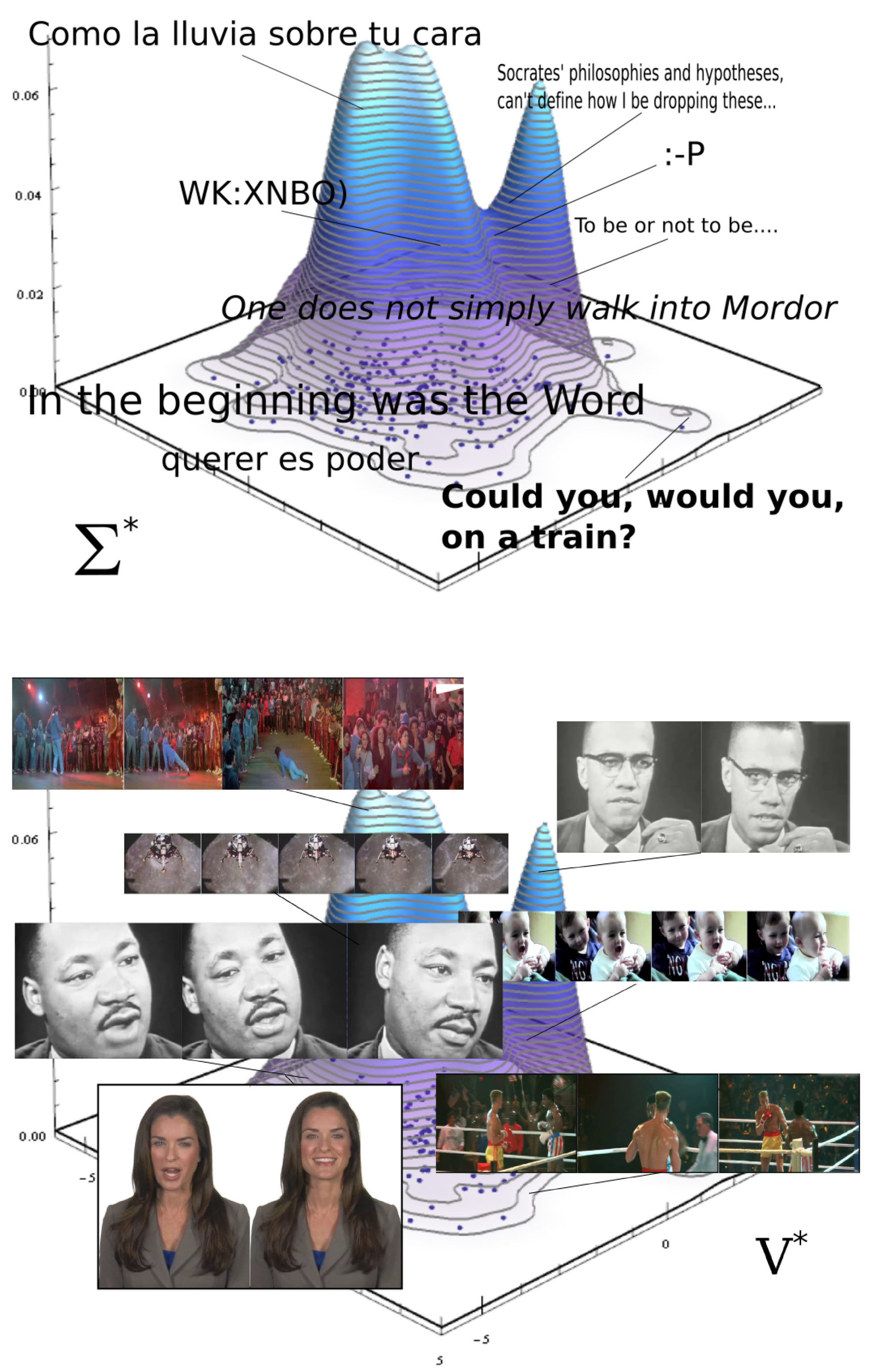
George D. Montañez and Cosma Rohilla Shalizi

Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA

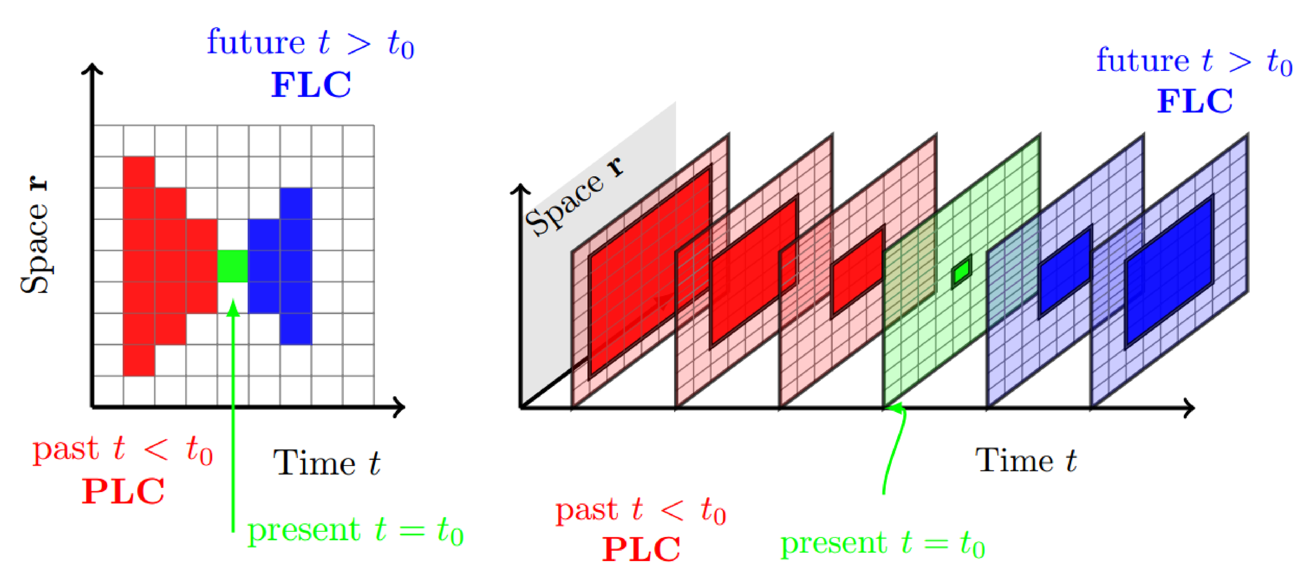
Introduction

- Spatio-temporal data is intrinsically high dimensional.
- Must exploit structure.
- Split the global field into many lower-dimensional “light cones”.
- Use light cone decompositions for predictive state reconstruction.
- Allow for tractable inference of spatio-temporal data.

Modeling Space of All Strings and All Videos



Light Cone Geometry: Local Structure



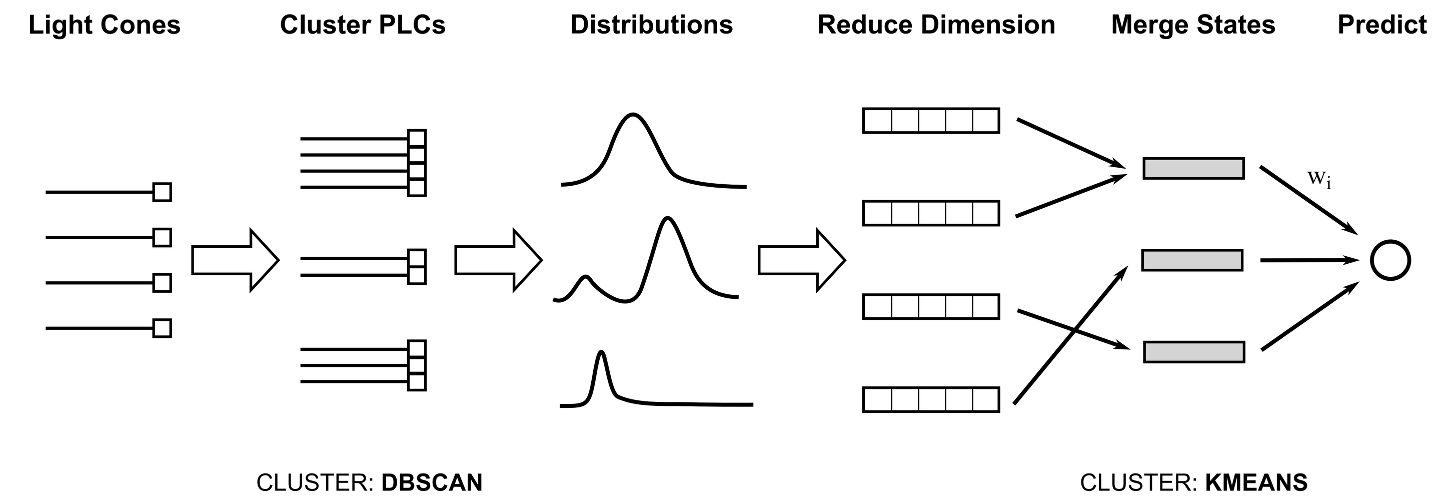
(a) $(1+1)D$ field (b) $(2+1)D$ field

(Taken from Goerg and Shalizi, “Mixed LICORS: A Nonparametric Algorithm for Predictive State Reconstruction,” 2013)

Light Cone Methods

- Mixed LICORS (Goerg and Shalizi, 2013)
- Moonshine
- One Hundred Proof
- Light Cone Linear Regression

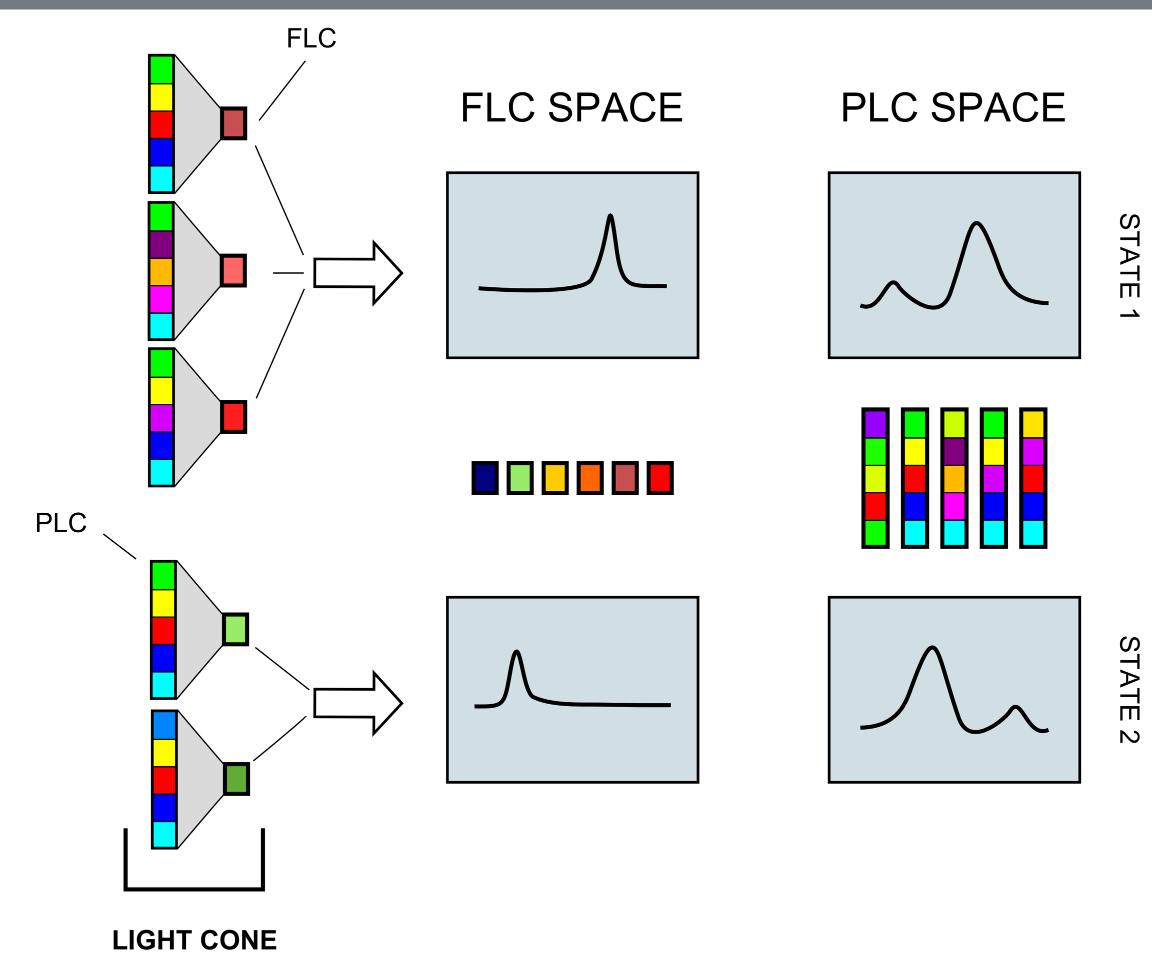
Moonshine



Algorithm 1 Moonshine

Decompose spatio-temporal process into light cone (PLC, FLC) observation tuples.
Cluster PLCs using density based clustering.
Compute cluster-conditioned density estimates for $2K+1$ random points.
if number of clusters $>$ maximum number **then**
  Merge clusters in the space of reduced dimension.
end if
Map original light cones to final clusters.

One Hundred Proof



Algorithm 2 One Hundred Proof

Decompose spatio-temporal process into light cone (PLC, FLC) observation pairs.
Cluster FLCs using k-means++ clustering.
Map original light cone pairs to final clusters.

Light Cone Linear Regression

Light cone linear regression uses the same light cone decomposition as the LICORS, Moonshine and OHP methods, but learns a regression rule directly from past light cones to future light cone values.

Experiments

- Forecasting Task 1: Electrostatic Potentials
  - Model electrostatic potential changes in organic electronic materials.
- Forecasting Task 2: Human Speaker Video
  - Predict the next frame of a full-resolution video from a recording of a human speaker.

For all light cone methods, the same set of light cones were extracted from the data, with $h_p = 1$, $h_f = 0$, and $c = 1$.

Measured mean-squared-error (MSE), correlation (Pearson ρ) with the ground truth, and average per pixel log-likelihood.

Results: Prediction Task 1

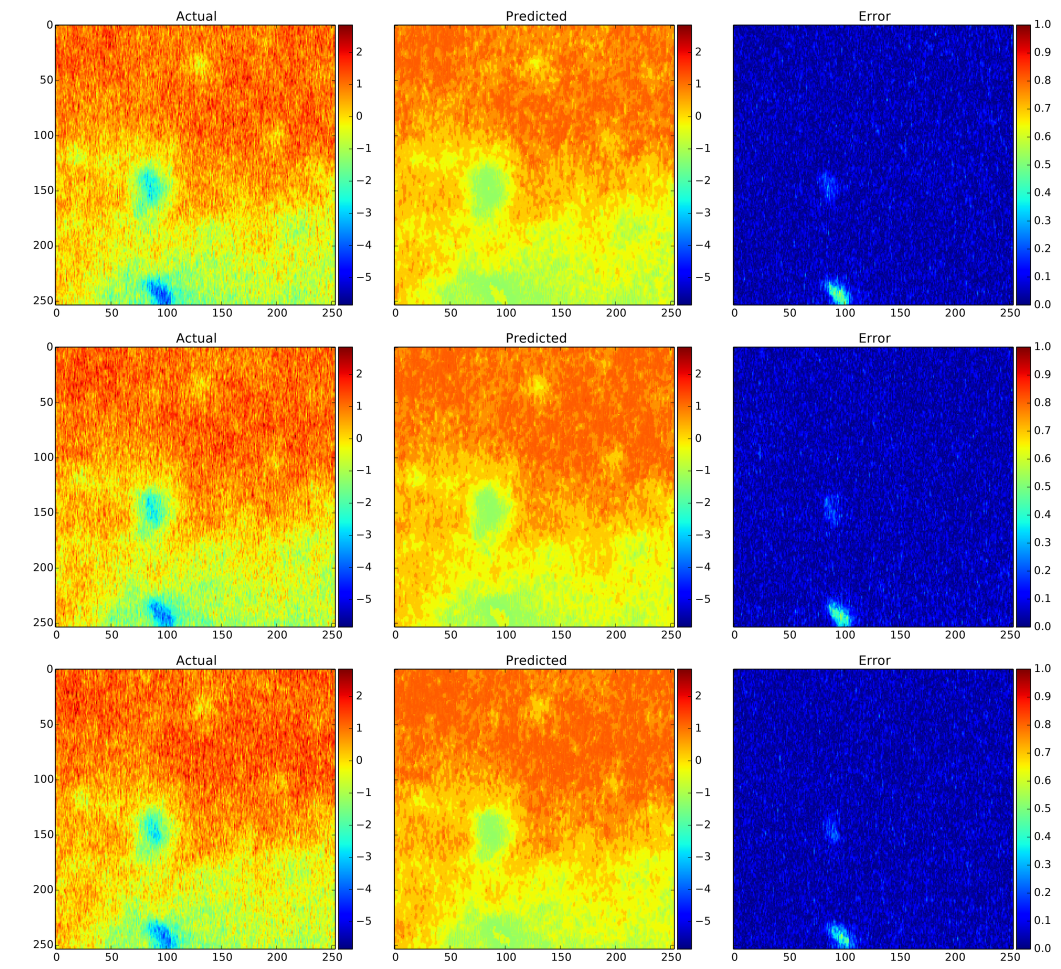


Figure: Predicting electrostatic potentials with Moonshine.

Method	K_{max}	MSE	95% CI	Avg. LL	95% CI	Perplexity	Pearson ρ	95% CI
Future-like-the-Past	·	0.778	[0.777, 0.780]	·	·	·	0.615	[0.614, 0.616]
KNN Regression	·	0.852	[0.851, 0.853]	·	·	·	0.506	[0.505, 0.506]
Light Cone Linear Regression	·	0.607	[0.606, 0.608]	·	·	·	0.628	[0.627, 0.628]
Mixed LICORS	100	**0.569**	**[0.567, 0.571]**	-1.034	[-1.110, -0.964]	2.052	**0.663**	**[0.661, 0.665]**
Moonshine	100	**0.570**	**[0.569, 0.572]**	**-0.672**	**[-0.727, -0.617]**	**1.593**	0.656	[0.655, 0.657]
One Hundred Proof	100	0.592	[0.591, 0.593]	-1.724	[-2.127, -1.321]	3.303	0.641	[0.640, 0.642]
Mixed LICORS	10	**0.566**	**[0.565, 0.567]**	-1.022	[-1.096, -0.947]	2.030	**0.668**	**[0.667, 0.669]**
Moonshine	10	0.609	[0.605, 0.613]	-0.722	**[-0.767, -0.678]**	1.650	0.625	[0.622, 0.628]
One Hundred Proof	10	0.597	[0.595, 0.598]	**-0.682**	**[-0.757, -0.608]**	**1.605**	0.648	[0.646, 0.649]

Results: Prediction Task 2

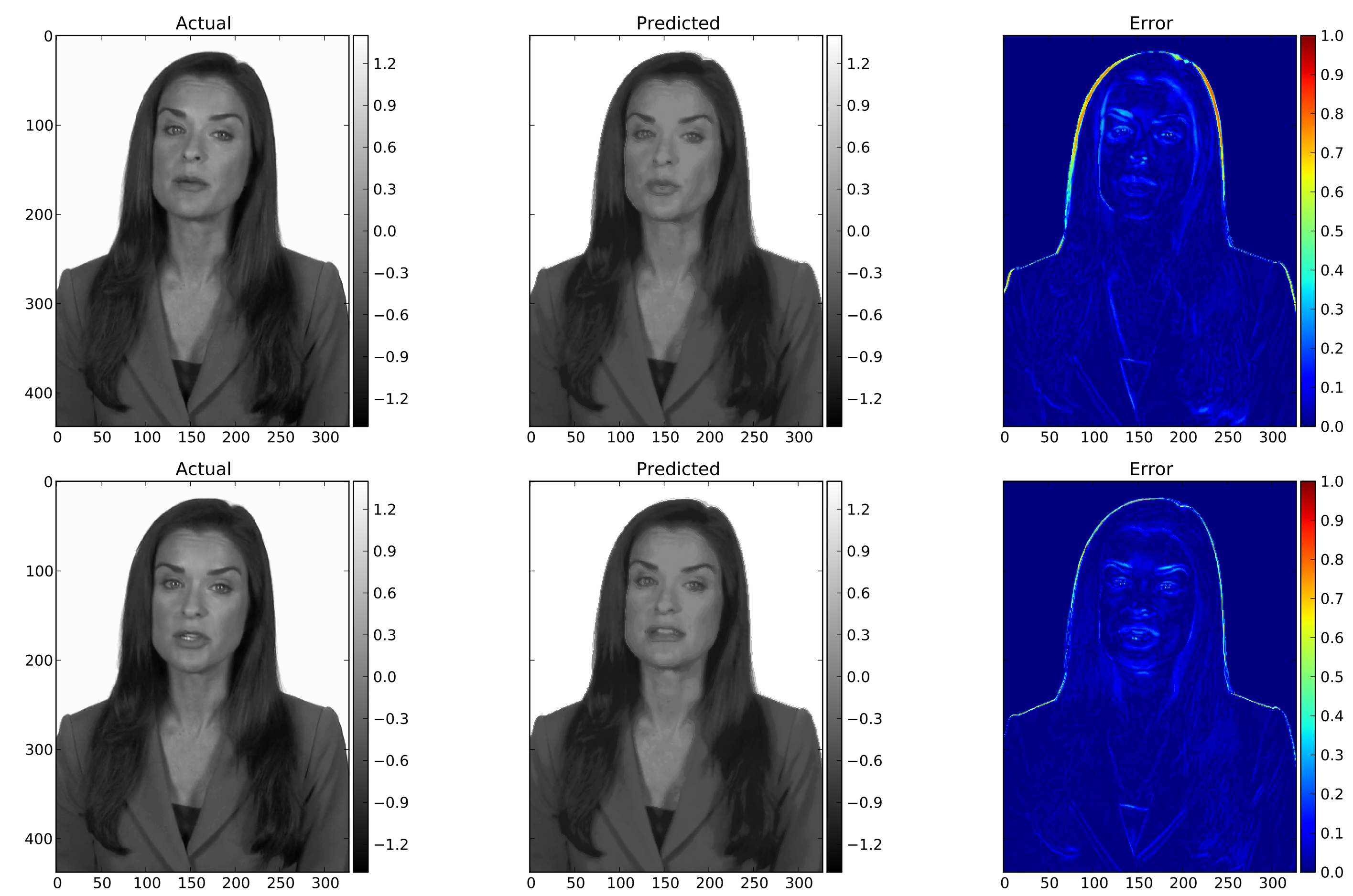


Figure: Predicting video with Mixed LICORS light cone system.

Method	K_{max}	MSE	95% CI	Avg. LL	95% CI	Perplexity	Pearson ρ	95% CI
Future-like-the-Past	·	0.031	[0.031, 0.031]	·	·	·	0.984	[0.984, 0.984]
KNN Regression	·	0.033	[0.033, 0.033]	·	·	·	0.984	[0.984, 0.984]
Light Cone Linear Regression	·	**0.028**	**[0.028, 0.028]**	·	·	·	**0.986**	**[0.986, 0.0986]**
Mixed LICORS	100	0.038	[0.038, 0.038]	0.102	[0.099, 0.105]	0.932	0.981	[0.981, 0.981]
Moonshine	100	0.039	[0.039, 0.039]	**0.925**	**[0.874, 0.976]**	**0.527**	0.981	[0.981, 0.981]
One Hundred Proof	100	1.060	[0.460, 1.659]	-6.48	[-8.025, -4.948]	89.641	0.911	[0.871, 0.952]

Conclusions

- Decompose global field into local structures (light cones).
- Use kernel density estimators on clusters of light cones.
- Can model and forecast complex spatio-temporal data.