Topology in Machine Learning

Henry Adams
University of Florida

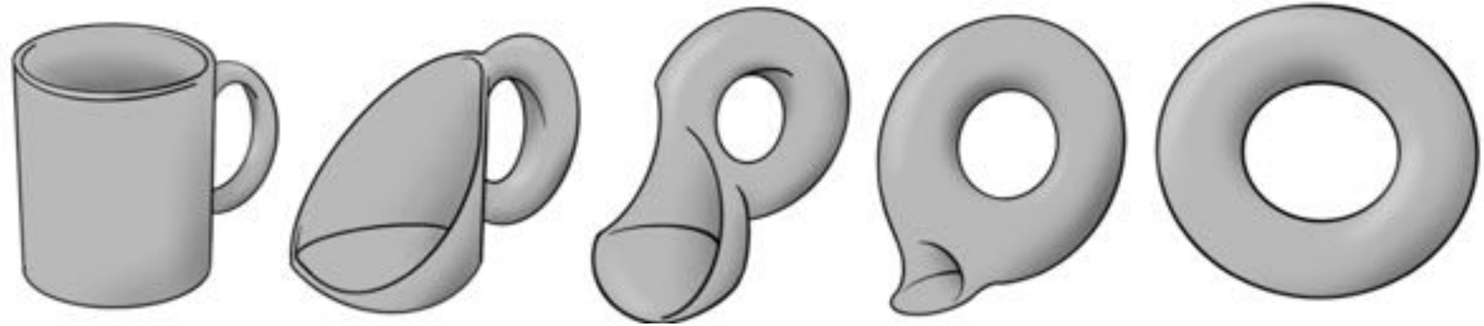
AATRN, www.aatrn.net , 1-2 live talks per week

YouTube : 6,000 subscribers, 22 hours watched per day

Topology in Machine Learning

- How to vectorize geometry ?
- Introduction to persistent homology
- Applications in materials science, computer vision, and explainable machine learning

What is the difference between geometry and topology?

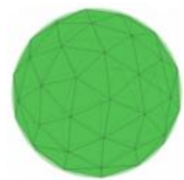


Topology ignores some geometrical properties (distances, curvatures) but preserves holes.

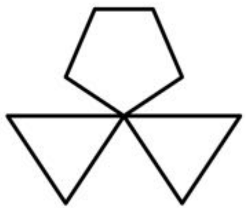
Topology is computable.

Homology (counts holes)

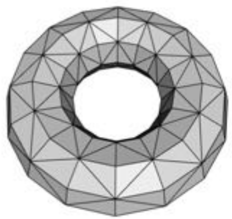
H_0: rank 6
H_1: rank 0



H_0: rank 1
H_1: rank 0
H_2: rank 1

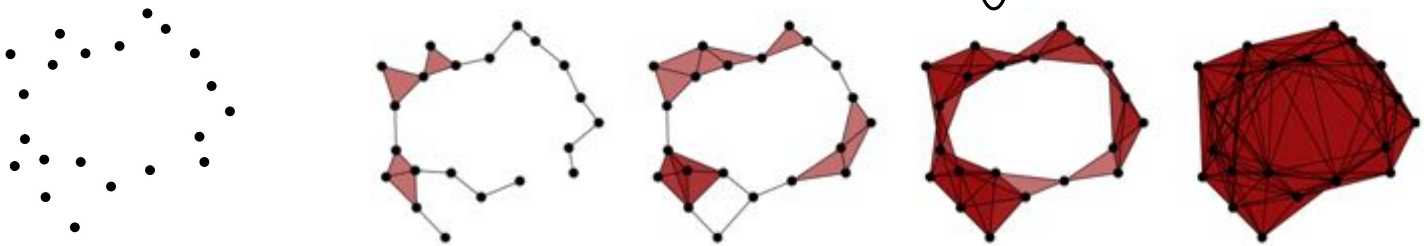


H_0: rank 1
H_1: rank 3

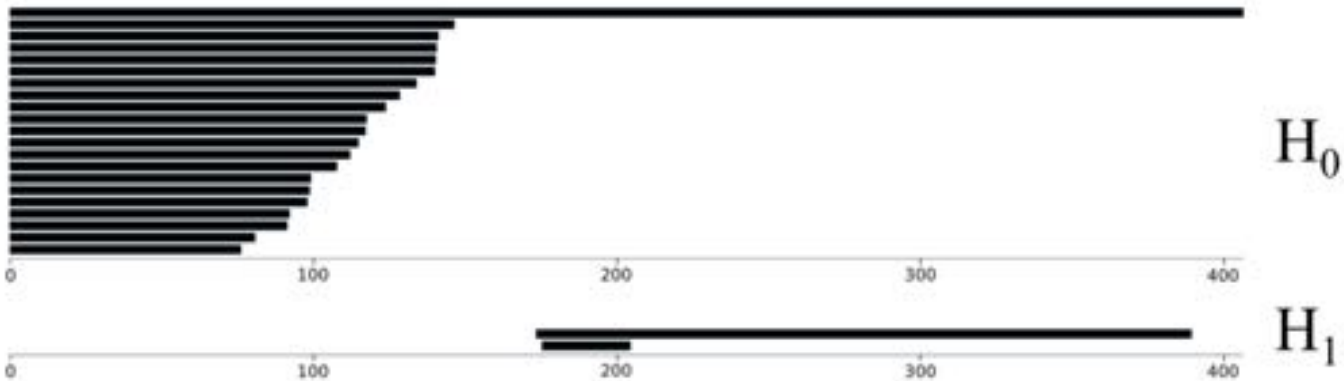


H_0: rank 1
H_1: rank 2
H_2: rank 1

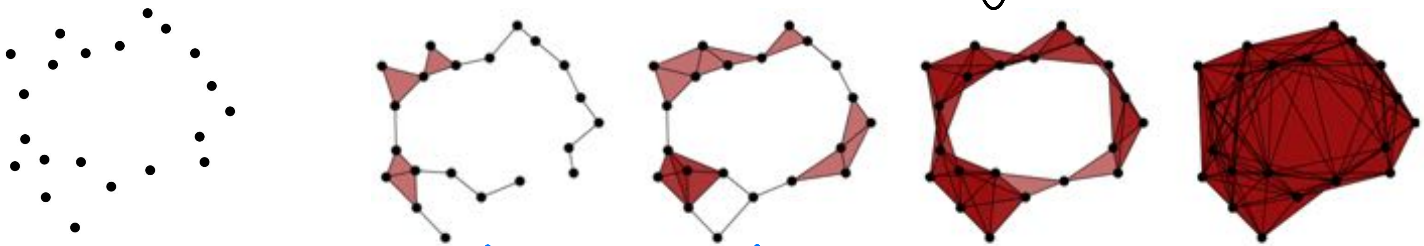
Persistent Homology
(tracks holes as a space grows)



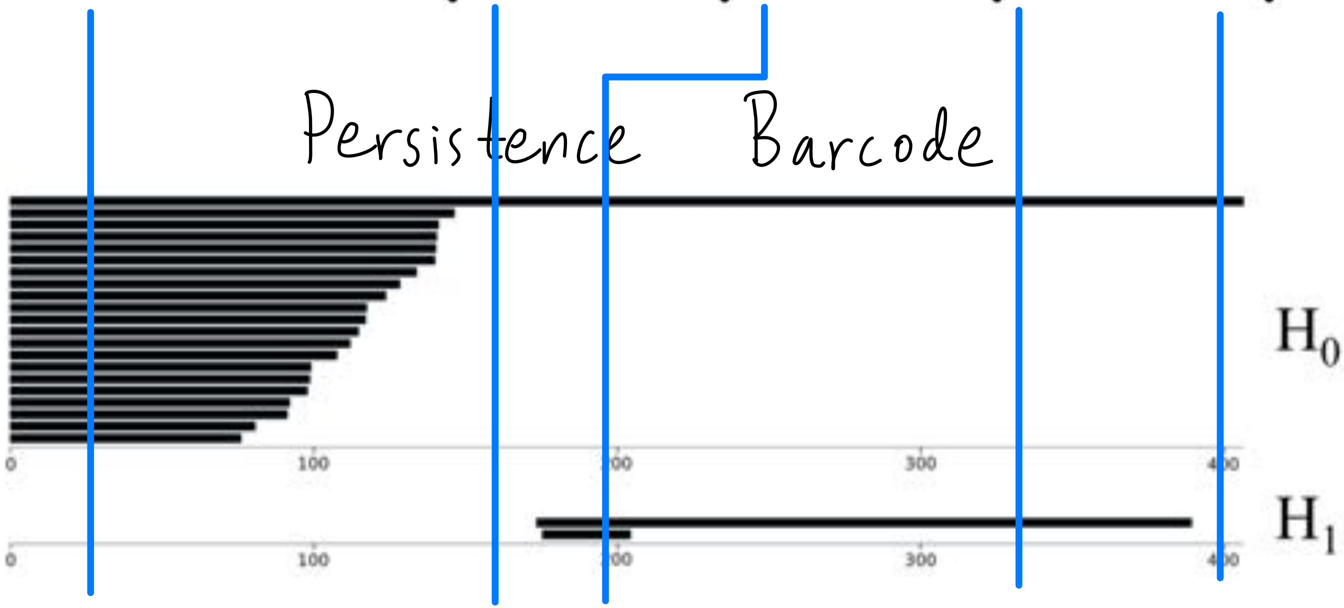
Persistence Barcode

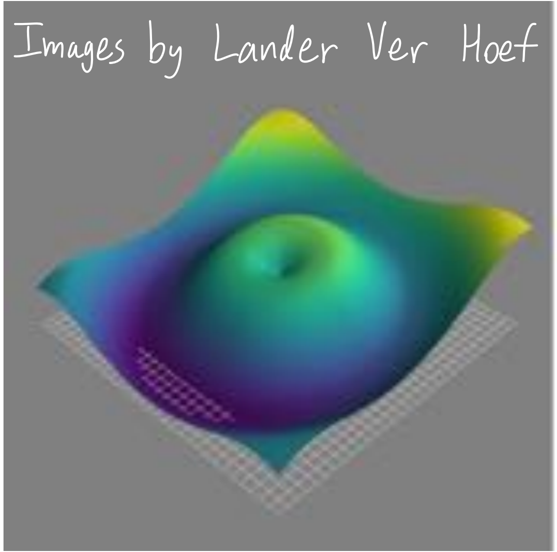


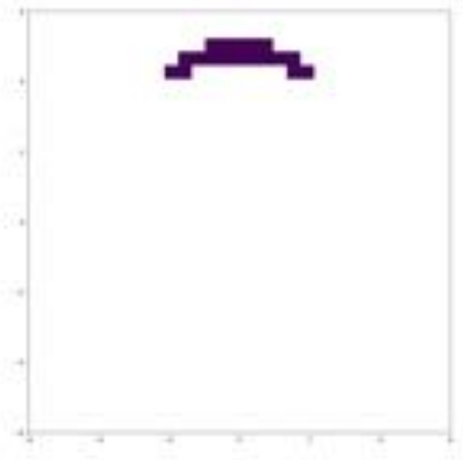
Persistent Homology
(tracks holes as a space grows)

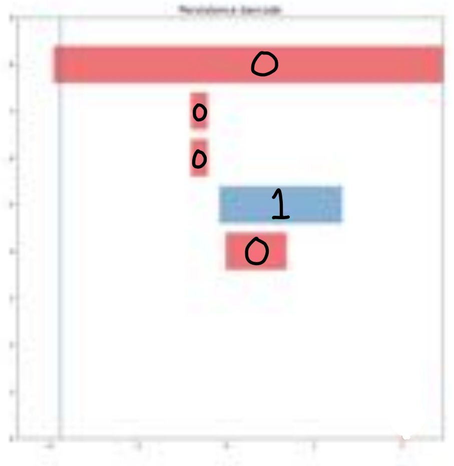


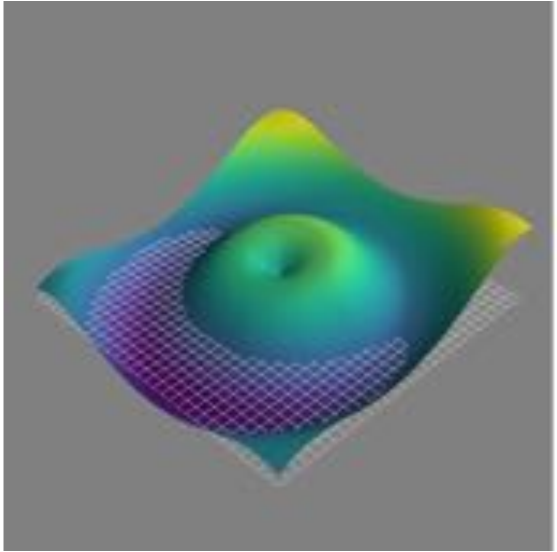
Persistence Barcode

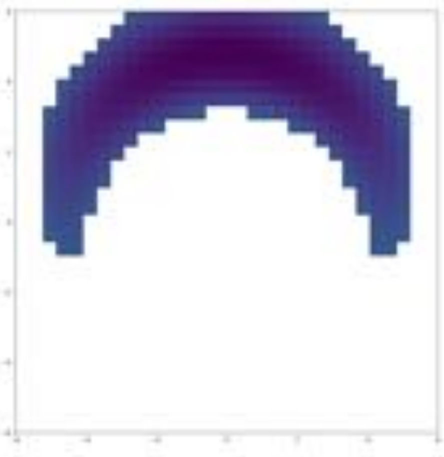


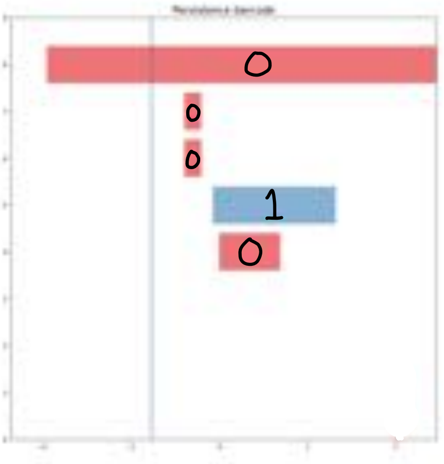

Images by Lander Ver Hoef

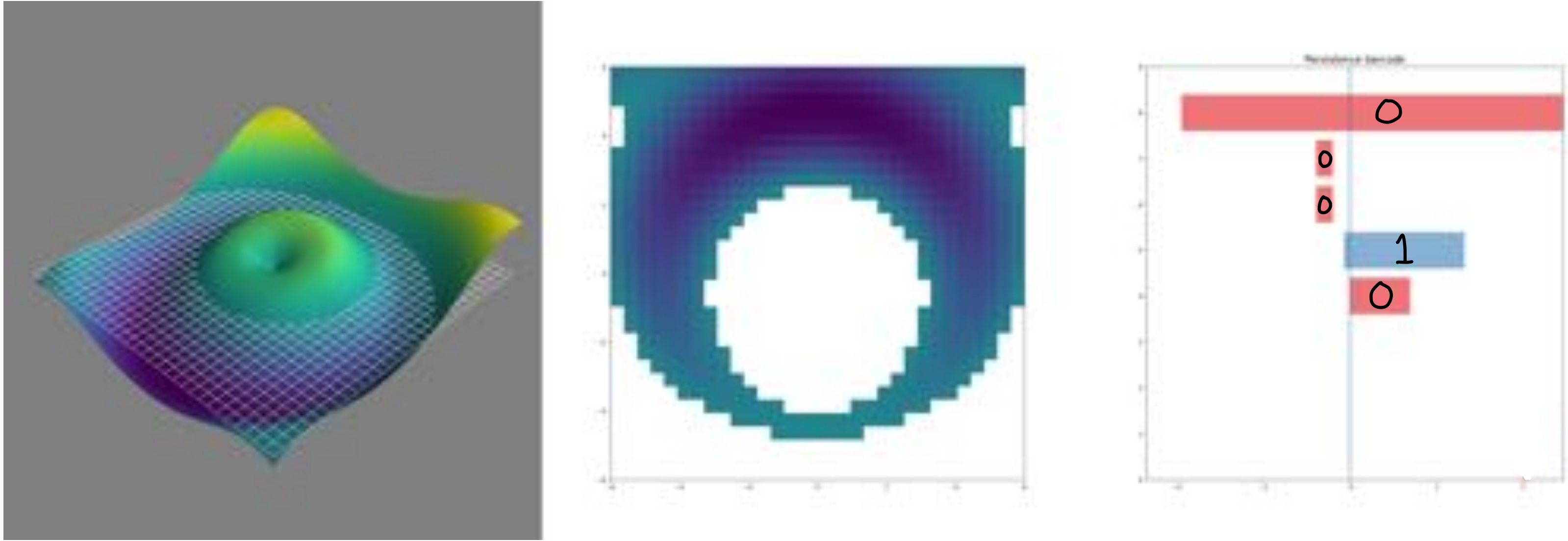


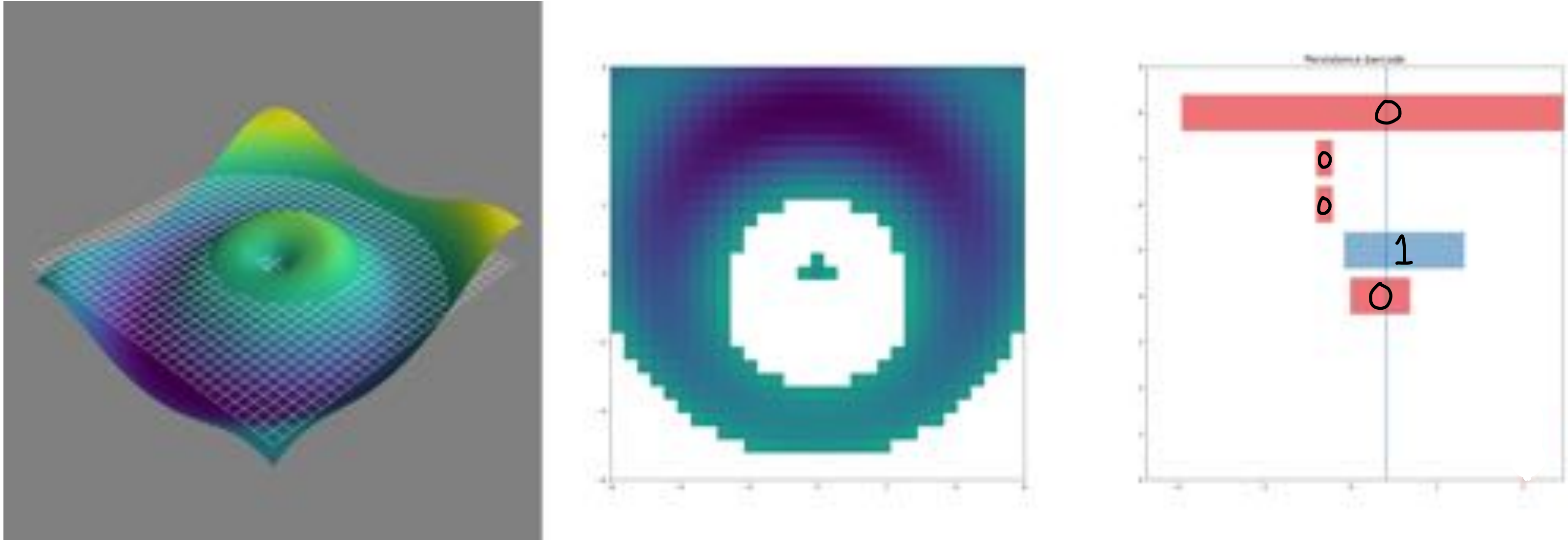

0
0
0
1
0

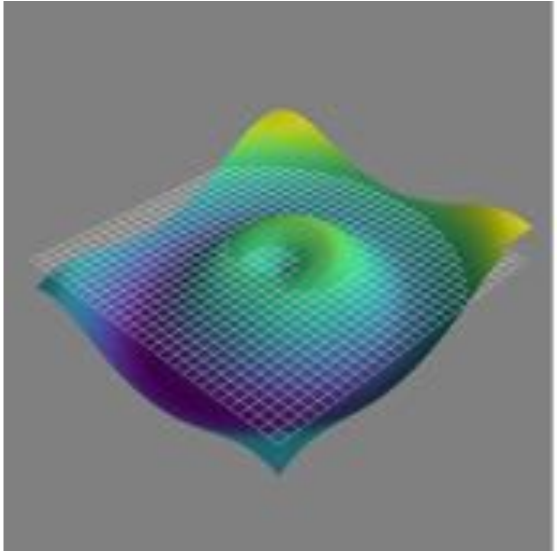


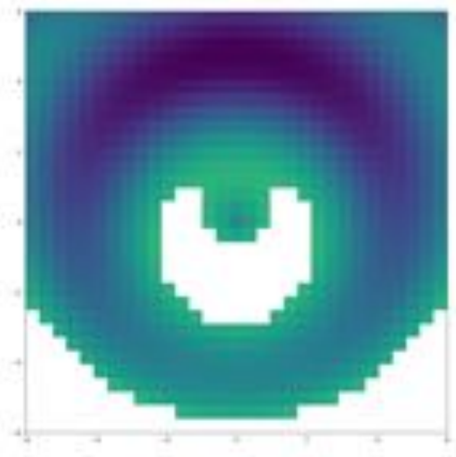


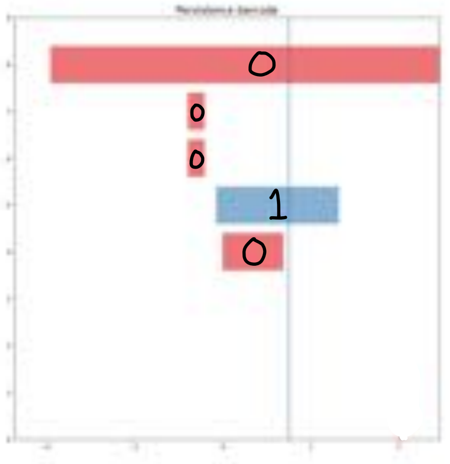

0
0
0
1
0

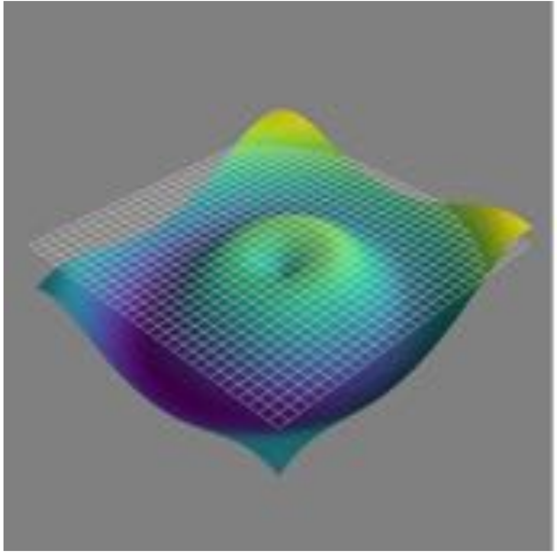

0
0
0
1
0

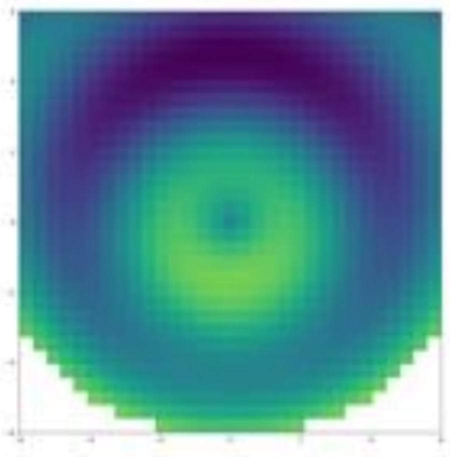

0
0
0
1
0

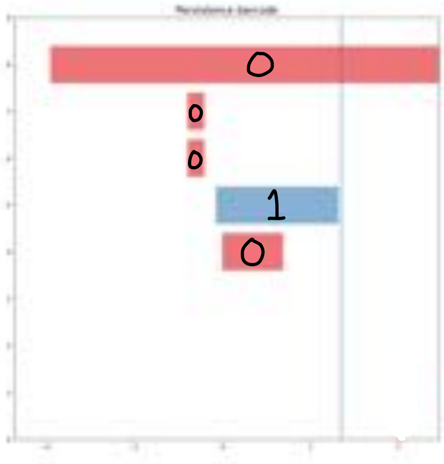




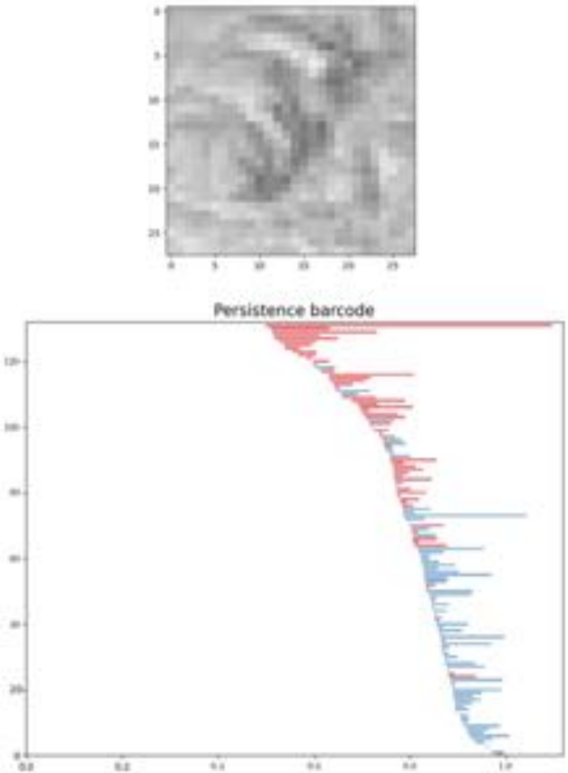

0
0
0
1
0



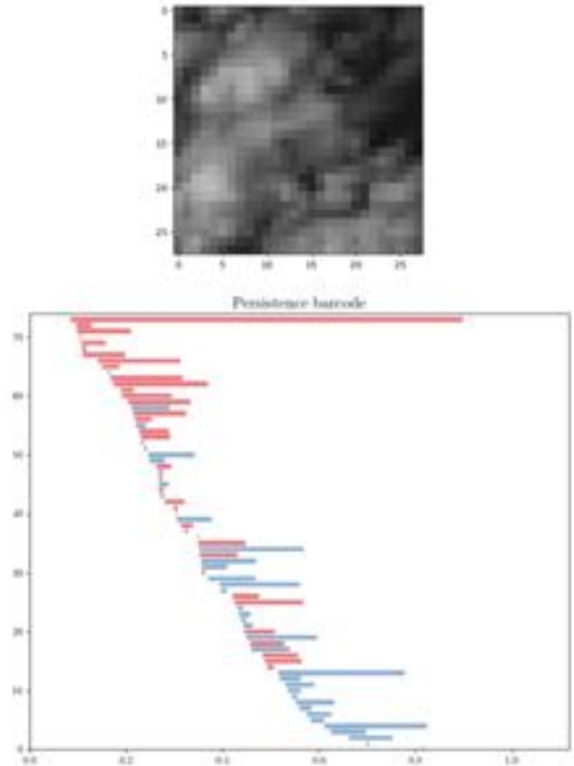



0
0
0
1
0

Convective

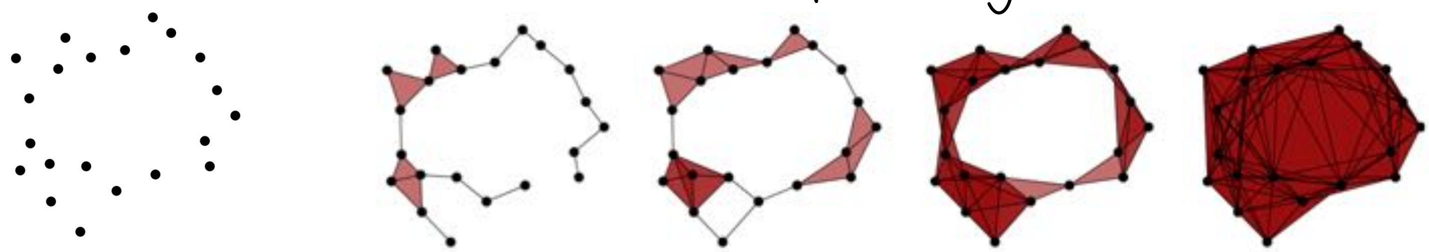


Non-Convective

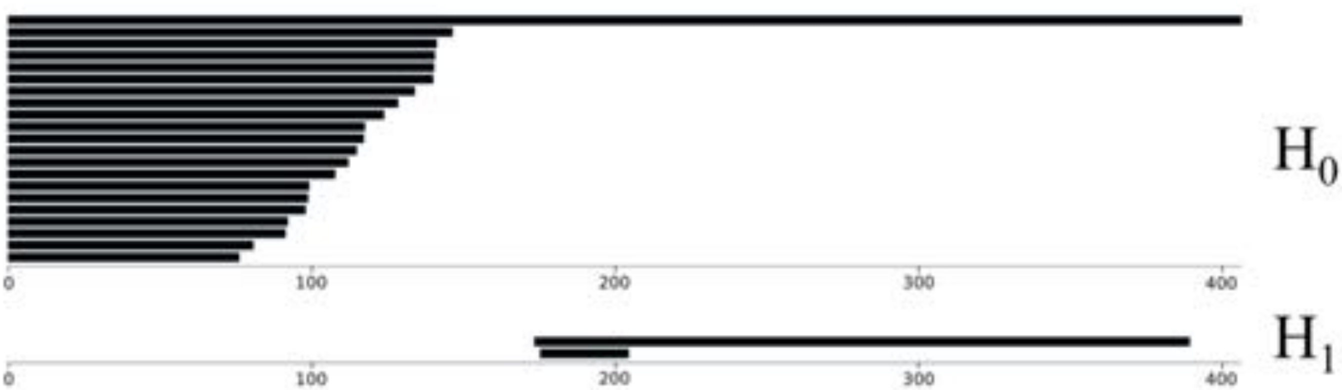


Ver Hoef, Lee, Adams, King, Ebert-Uphoff

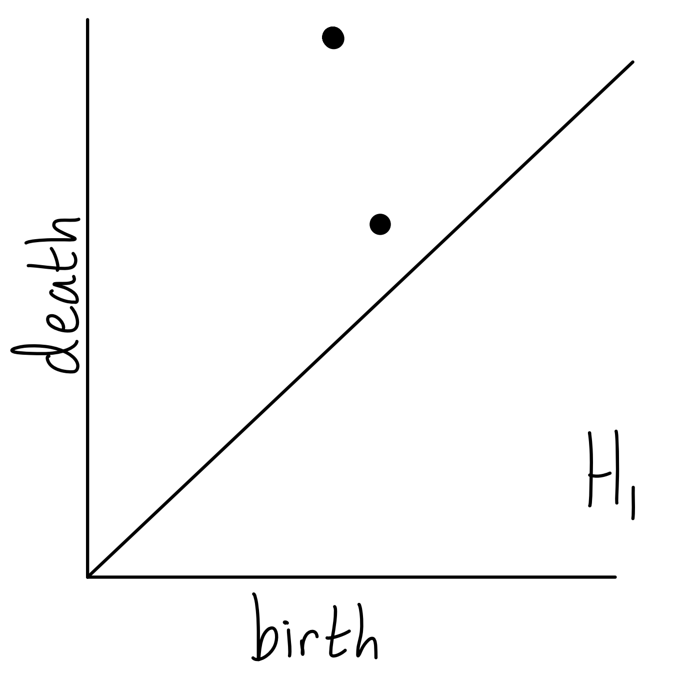
Persistent Homology
(tracks holes as a space grows)



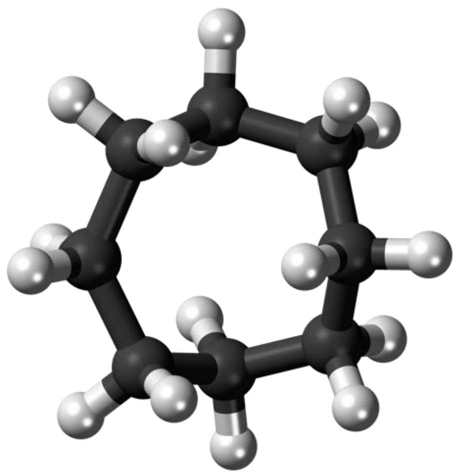
Persistence Barcode

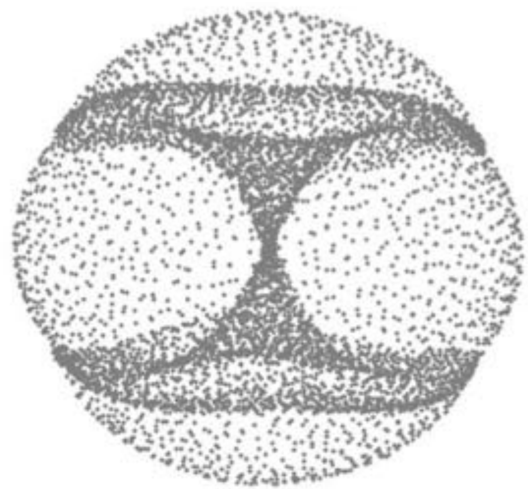


Persistence Diagram



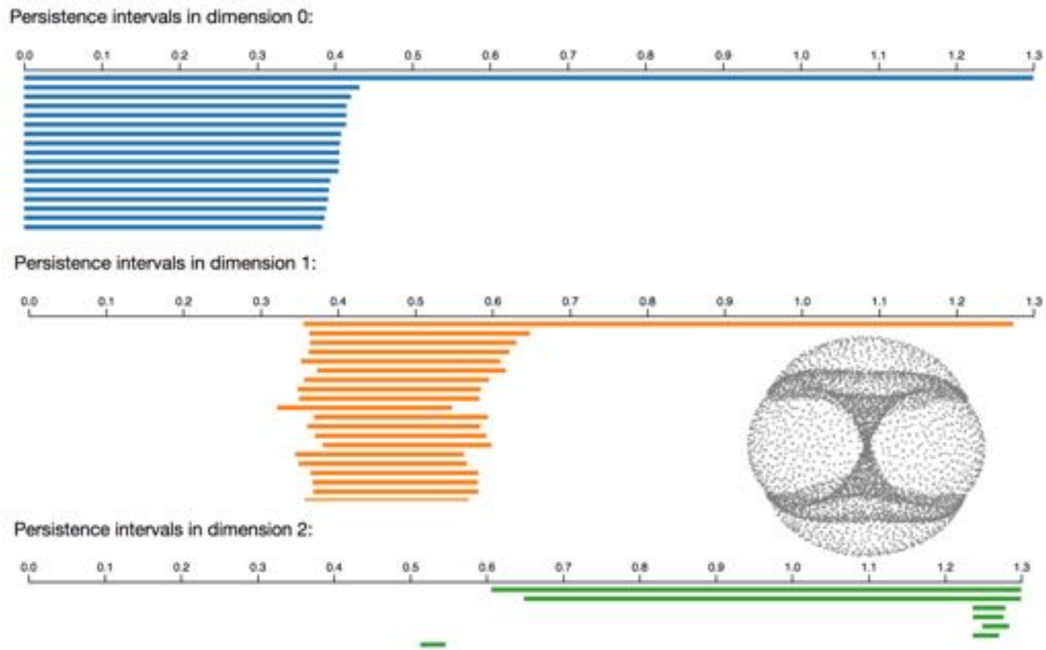
Global topology





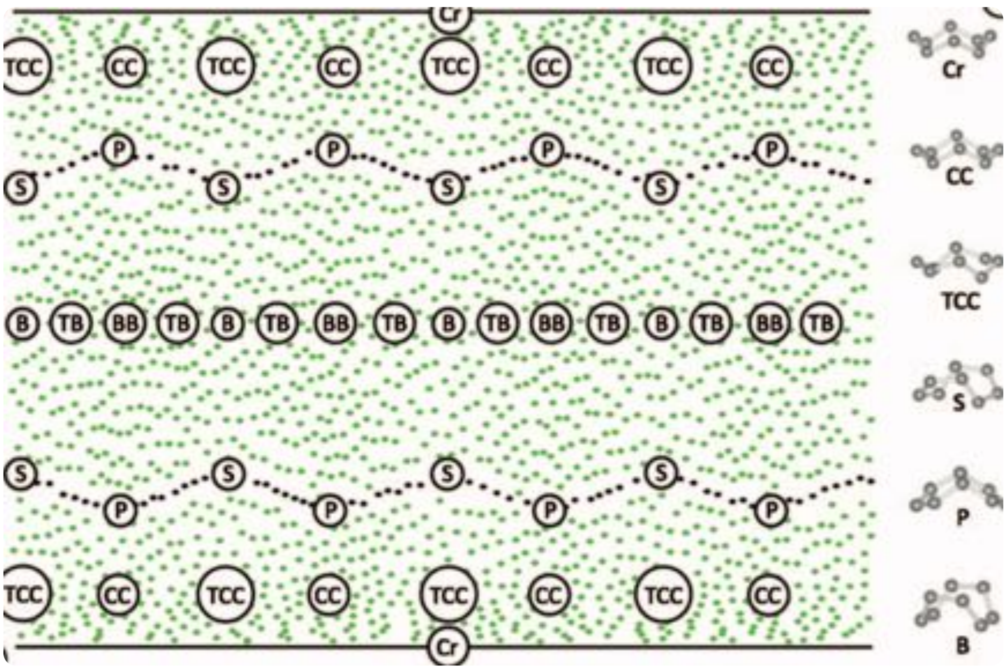
Topology of cyclo-octane energy landscape
Martin, Thompson, Coutsias, Watson, 2010

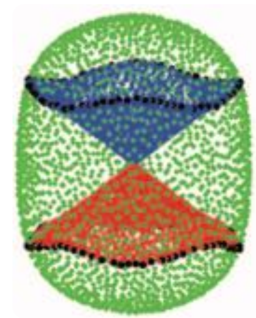
Global topology



Topology of cyclo-octane energy landscape
Martin, Thompson, Coutsias, Watson, 2010

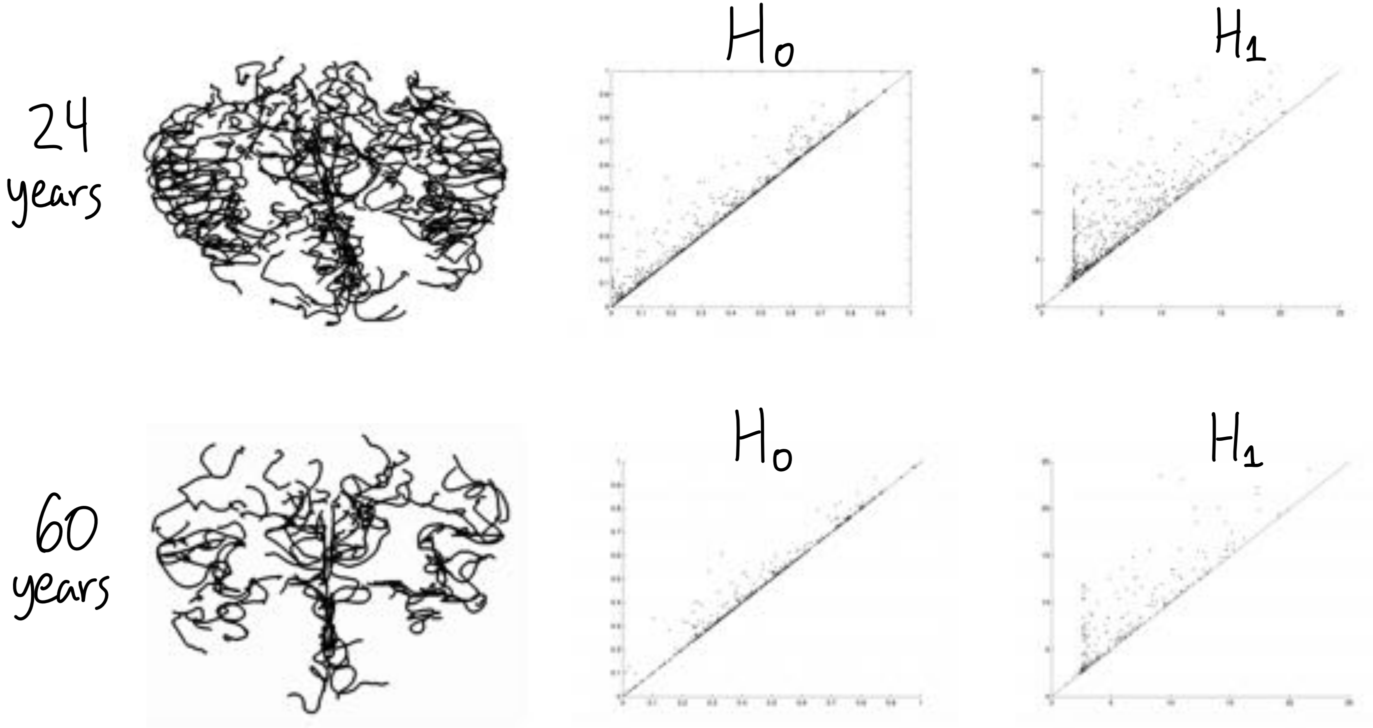
Global topology





Topology of cyclo-octane energy landscape
Martin, Thompson, Coutsias, Watson, 2010

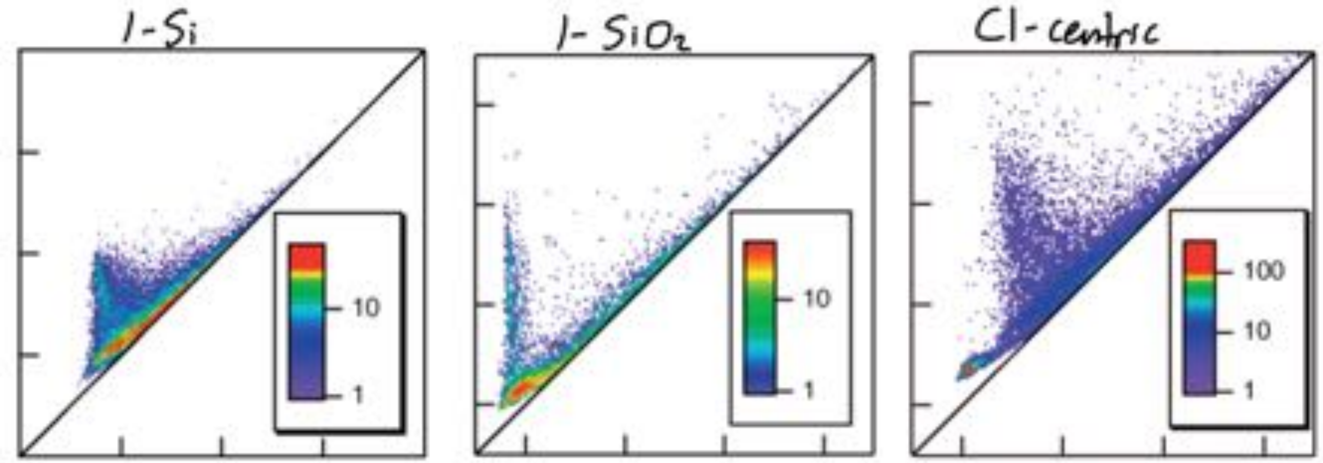
Persistent homology measures topology and geometry



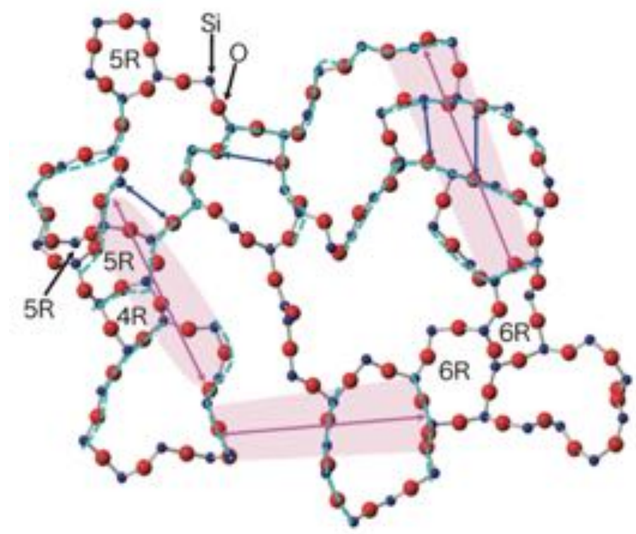
Persistent homology analysis of brain artery trees
Bendich, Marron, Miller, Pieloch, Skwerer, 2014

Topology in Machine Learning

- How to vectorize geometry ?
- Introduction to persistent homology
- Applications in materials science, computer vision, and explainable machine learning



"Understanding diffraction patterns of glassy, liquid, and amorphous material via persistent homology analysis" by Onodera, Kohara, Tahara, Masuno, Inoue, Shiga, Hirata, Tsuchiya, Hiraoka, Obayashi, Ohara, Mizuno, Sokata, 2019



"Persistence diagrams with linear machine learning models" by Obayashi, Hiraoka, Kimura

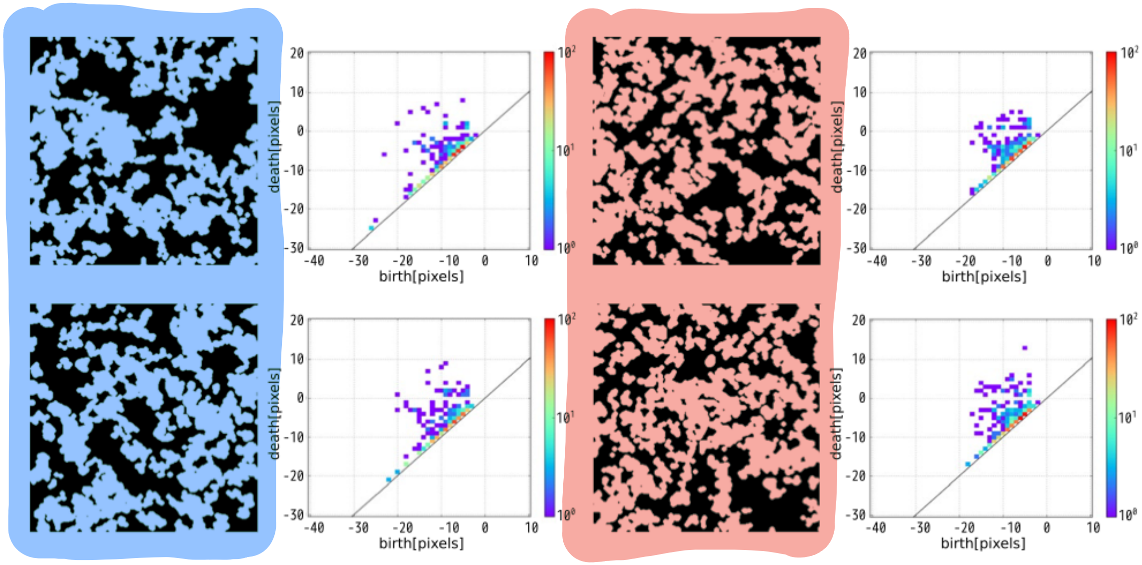


Fig. 4 Input binary images and their 0th persistence diagrams. The left and right two images are sampled from the parameter pairs (**A**) and (**B**), respectively

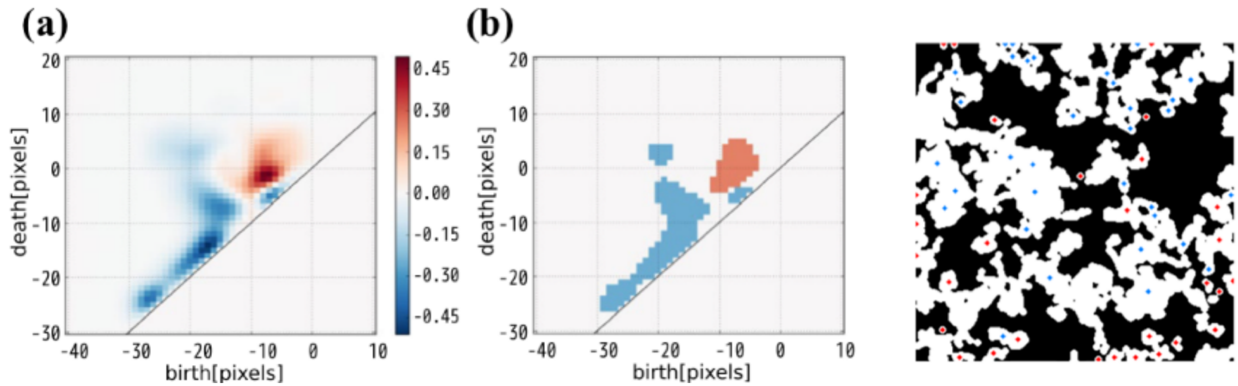
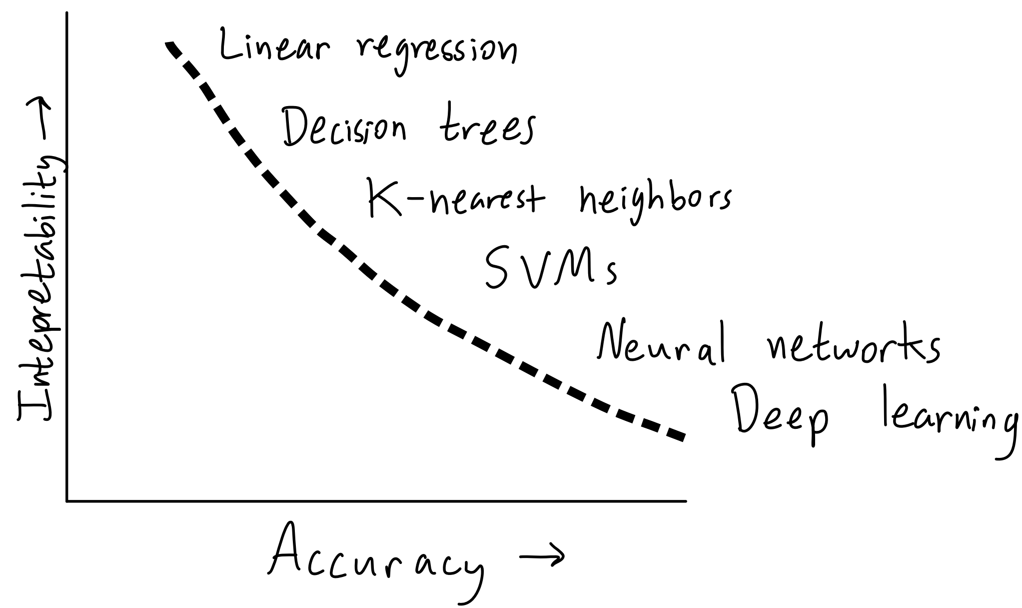
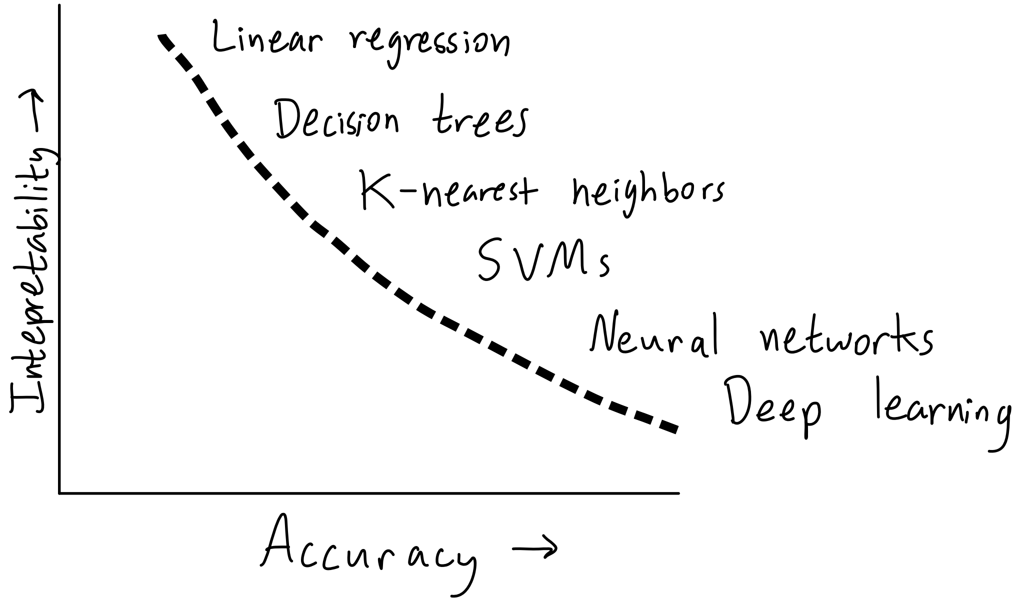
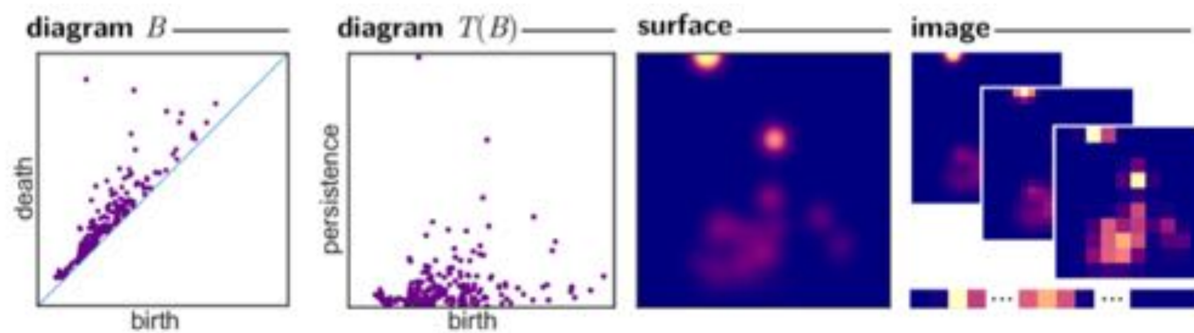


Fig. 5 **a** The reconstructed persistence diagram from the learned vector w. The blue (resp. red) area contributes to the class 0 (resp. 1). **b** A thresholding of (**a**). **c** 1–4 The birth positions of the generators in blue and red areas in (**b**) are plotted with the same color (color figure online)



Linear regression
Decision trees
K-nearest neighbors
SVMs
Neural networks
Deep learning
Intepretability →
Accuracy →

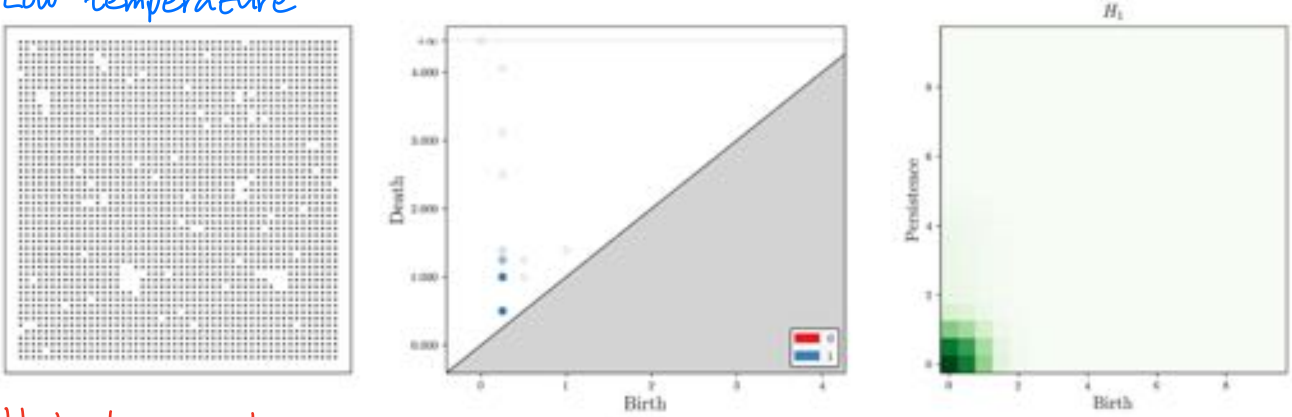




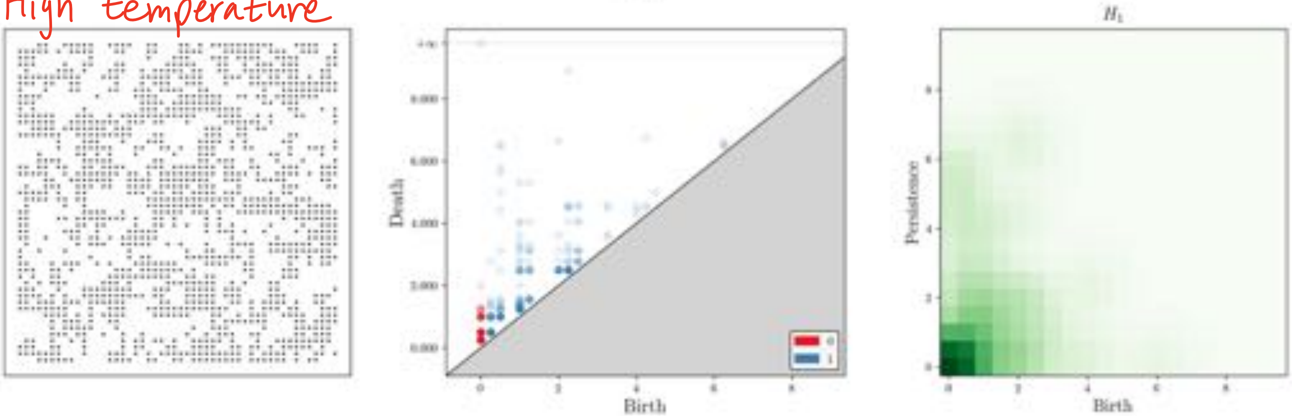
"Persistence images: A stable vector representation of persistent homology" by Adams, Chepushtanova, Emerson, Hanson, Kirby, Motta, Neville, Peterson, Shipman, Ziegelmeier, 2017

"Quantitative and interpretable order parameters for phase transitions from persistent homology" by Cole, Loges, Shiu, 2021

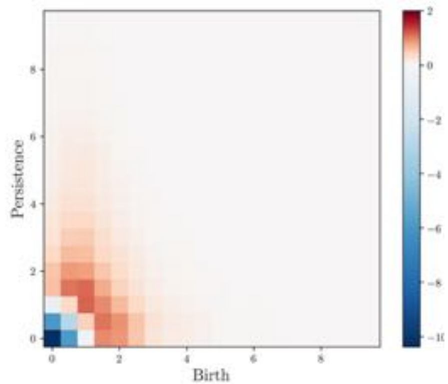
Low temperature

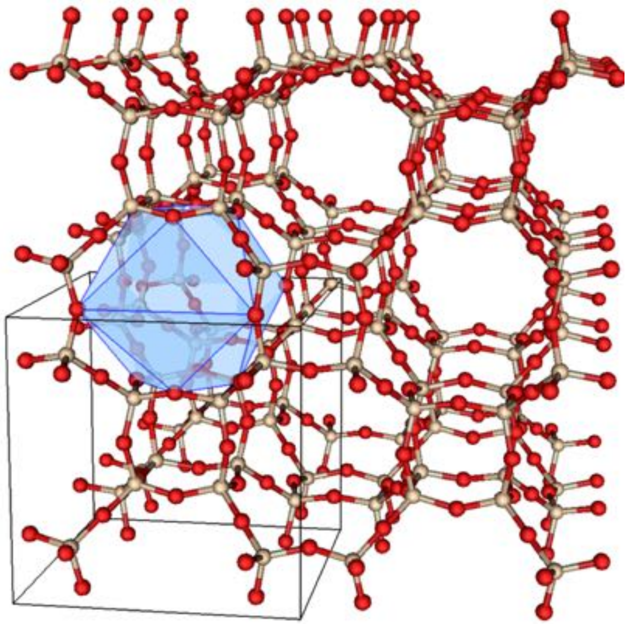


High temperature

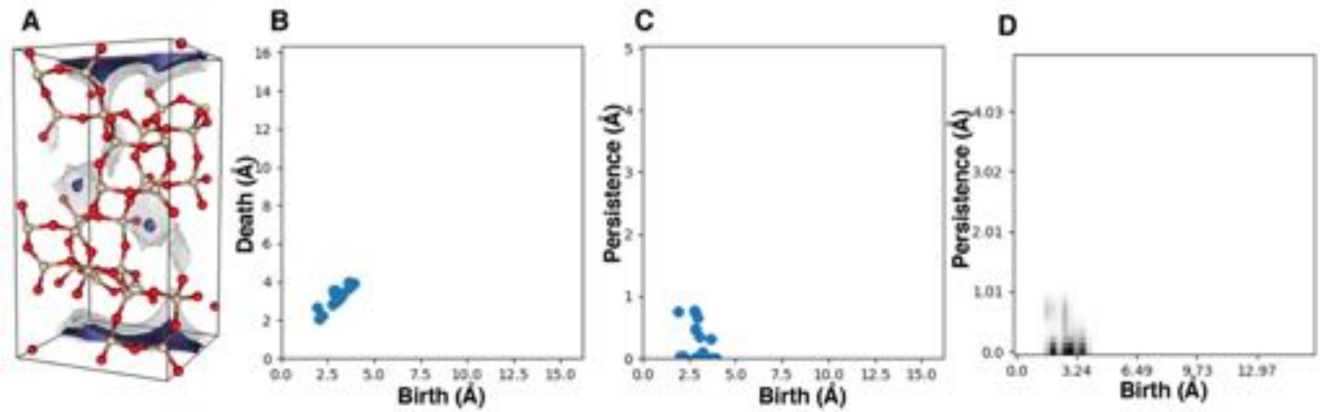


Statistical physics
Ising model
Phase transitions





"Topological descriptors help predict guest absorption in nanoporous materials" by Krishnapriyan, Haranczyk, Morozov, 2020

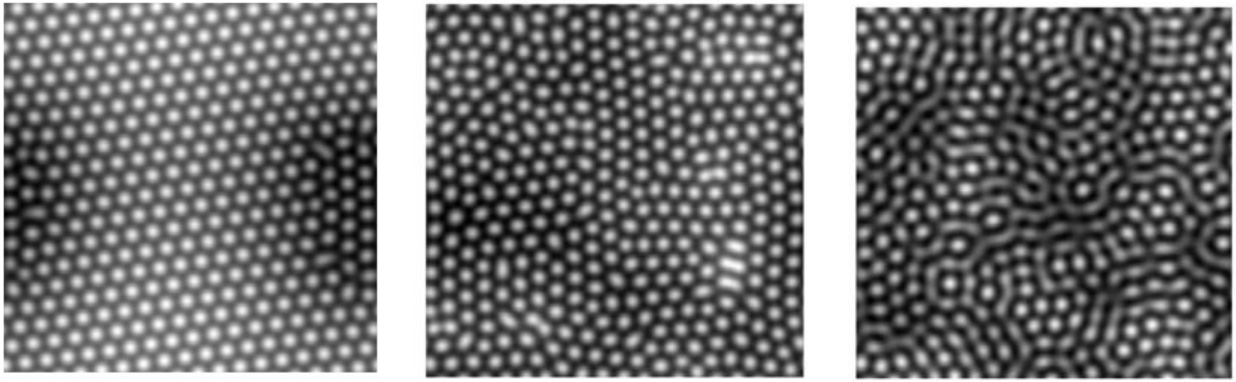


Methane absorption:

Accessible surface area

Largest cavity diameter

Local geometry

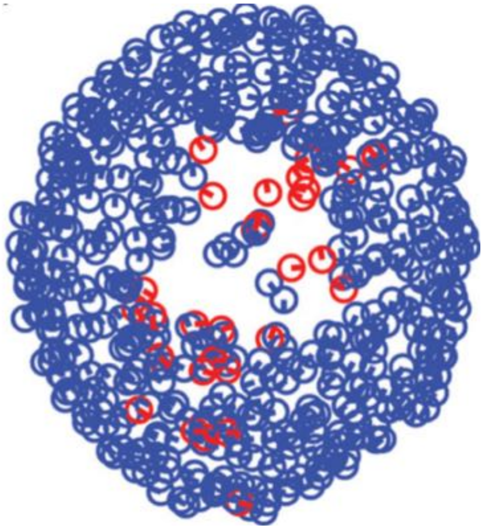


Measures of order for nearly hexagonal lattices
Motta, Neville, Shipman, Pearson, Bradley, 2018

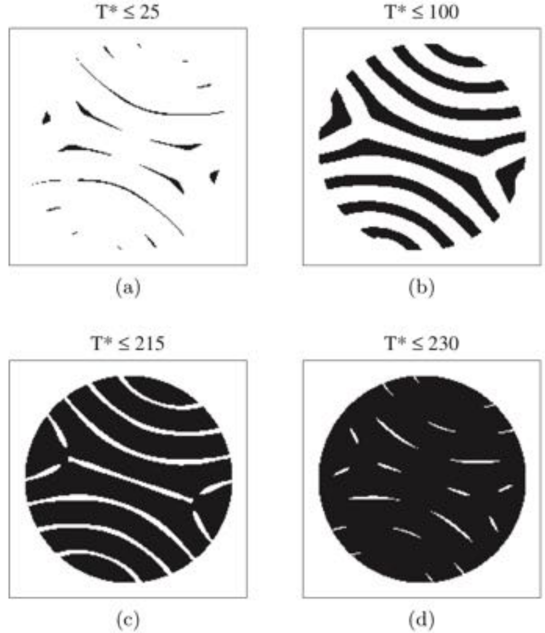
Collective motion, self-organization

Topological data analysis of biological aggregation models
Topaz, Ziegelmeier, Halverson, 2015



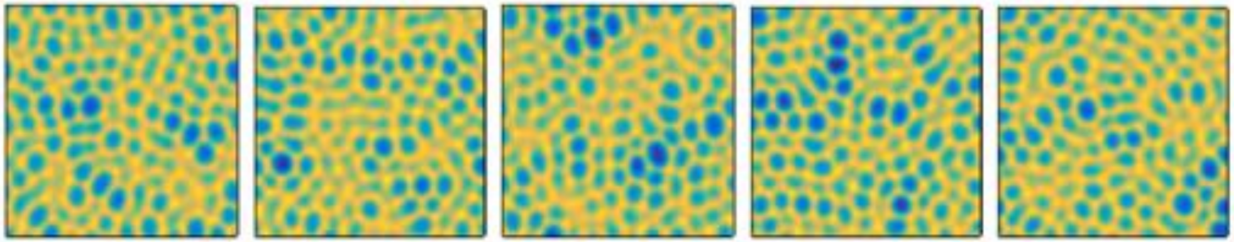
Collective motion, self-organization

Topological data analysis of biological aggregation models
Topaz, Ziegelmeier, Halverson, 2015



Analysis of Kolmogorov flow and Rayleigh-Bénard convection using persistent homology

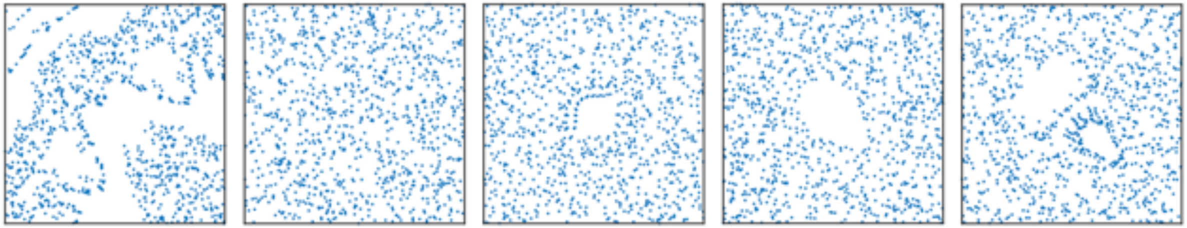
Kramár, Levanger, Tithof, Suri, Xu, Paul, Schatz, Mischaikow, 2016



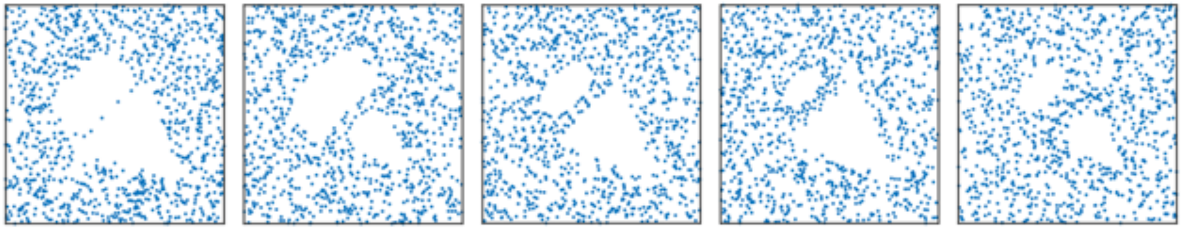
Answer: (from left) $r = 1.75, 2, 1.75, 2, 2$.

Persistence images: A stable vector representation of persistent homology. Adams, Chepushtanova, Emerson, Hanson, Kirby, Motta, Neville, Peterson, Shipman, Ziegelmeier, 2017

Different parameters:

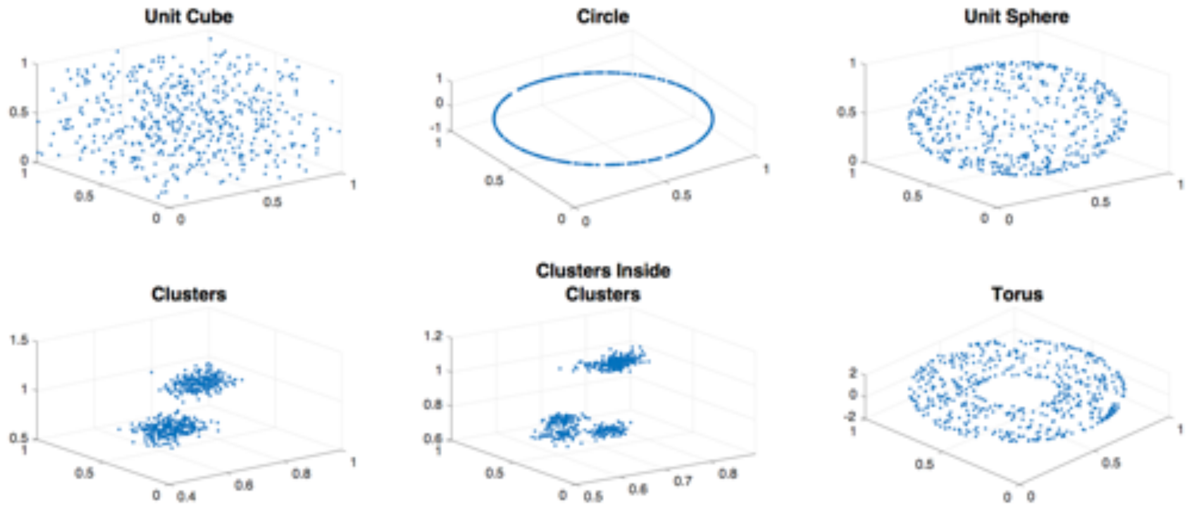


Same parameter:



Persistence images: A stable vector representation of persistent homology. Adams, Chepushtanova, Emerson, Hanson, Kirby, Motta, Neville, Peterson, Shipman, Ziegelmeier, 2017

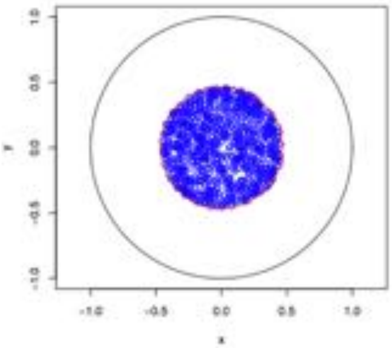
Local geometry



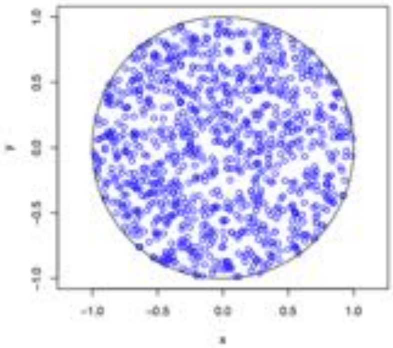
Persistence images: A stable vector representation of persistent homology. Adams, Chepushtanova, Emerson, Hanson, Kirby, Motta, Neville, Peterson, Shipman, Ziegelmeier, 2017

Local geometry

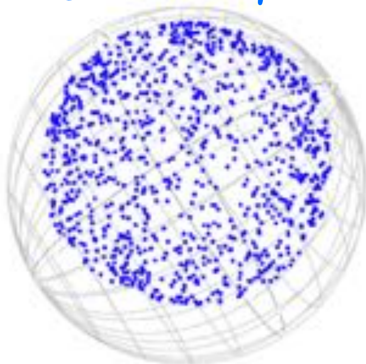
Hyperbolic disk



Flat disk

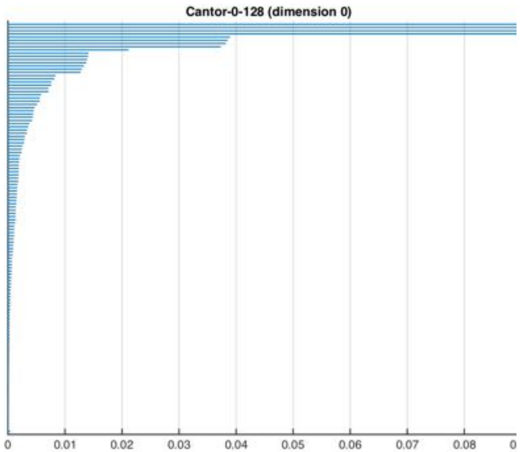


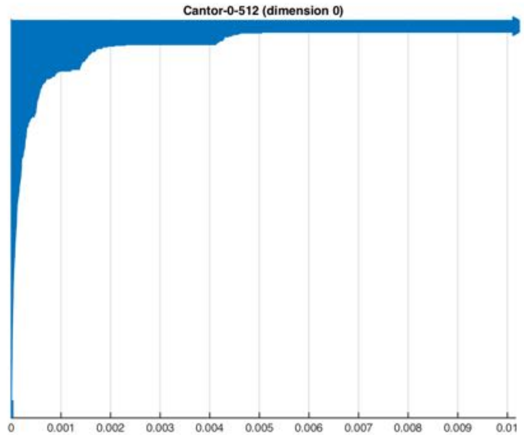
Disk on sphere



Persistent homology detects curvature
Bubenik, Hull, Patel, Whittle, 2019

Local geometry

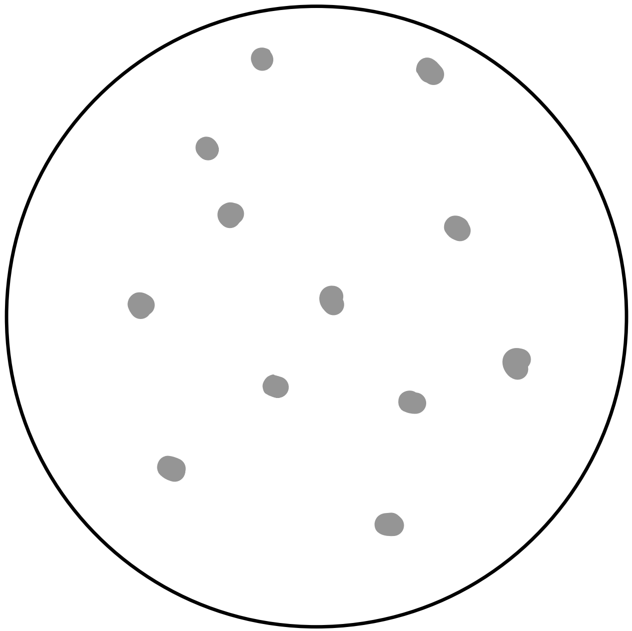




A fractal dimension for measures via persistent homology
Adams, Aminian, Farnell, Kirby, Peterson, Mirth, Neville, Shonkwiler, 2020

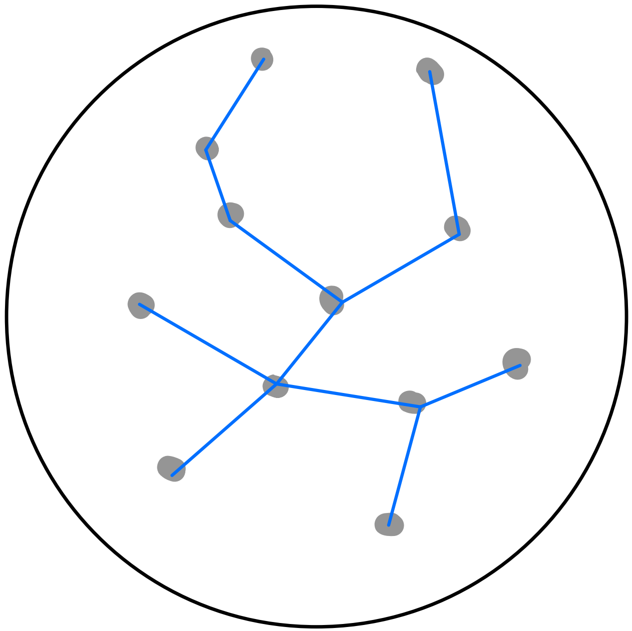
See also work by Robins and MacPherson & Schweinhart

Local geometry



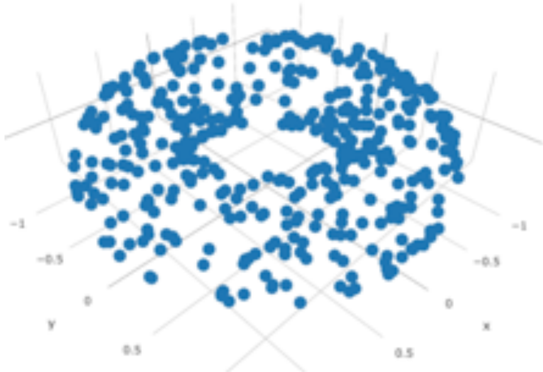
A fractal dimension for measures via persistent homology
Adams, Aminian, Farnell, Kirby, Peterson, Mirth, Neville, Shonkwiler, 2020

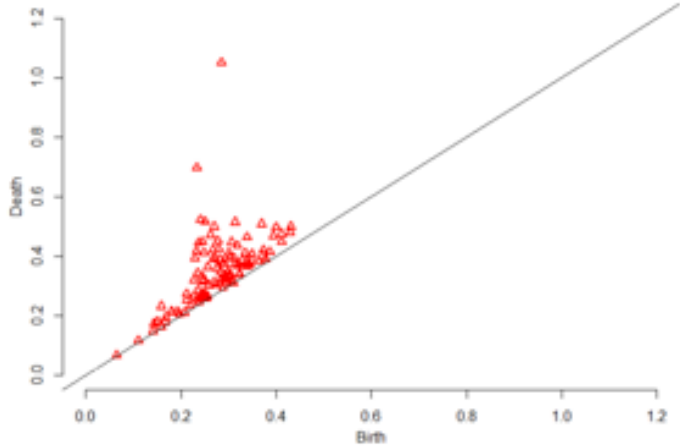
Local geometry



A fractal dimension for measures via persistent homology
Adams, Aminian, Farnell, Kirby, Peterson, Mirth, Neville, Shonkwiler, 2020

Local geometry





On the choice of weight functions for linear representations of persistence diagrams
Divol and Polonik, 2019

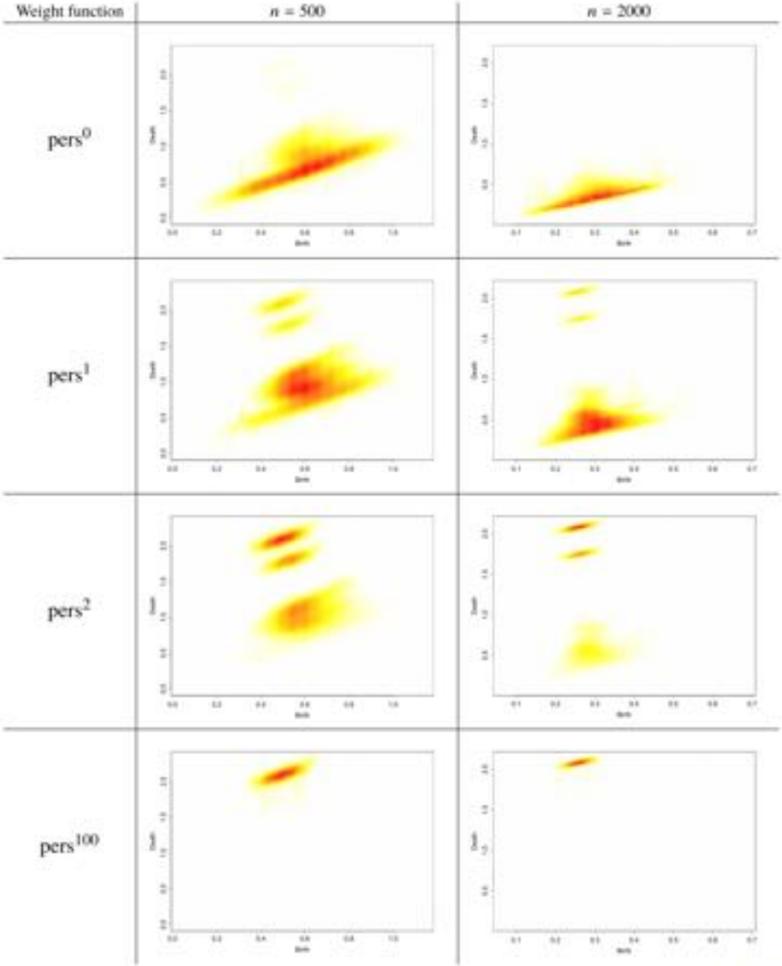


Fig. 2 For $n = 500$ or 2000 points uniformly sampled on the torus, persistence images (Adams et al. 2017) for different weight functions are displayed. For $\alpha < 2$, the mass of the topological noise is far larger than the mass of the true signal, the latter being comprised by the two points with high-persistence. For $\alpha = 2$, the two points with high-persistence are clearly distinguishable. For $\alpha = 100$, the noise has also disappeared, but so has one of the point with high-persistence

On the choice of weight functions for linear representations of persistence diagrams
Divol and Polonik, 2019

Topology in Machine Learning

- How to vectorize geometry ?
- Introduction to persistent homology
- Applications in materials science, computer vision, and explainable machine learning