

**DSC214**

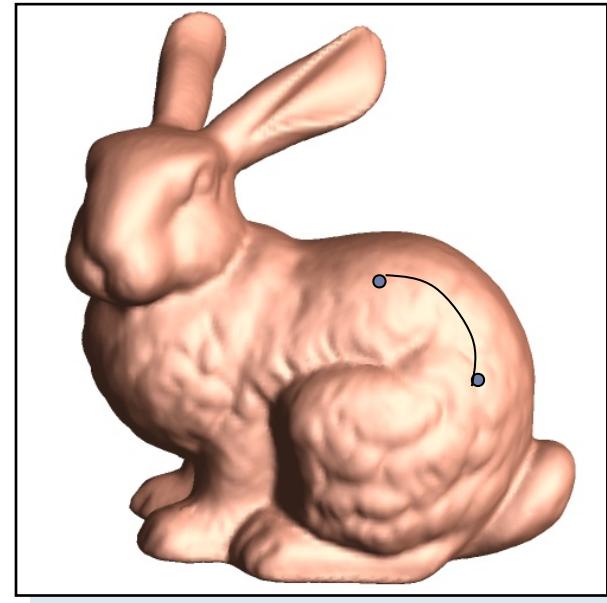
# **Topological Data Analysis**

## **Lecture 0: Introduction**

Instructor: Zhengchao Wan

# Geometry = Geo (earth) + metry (measure)

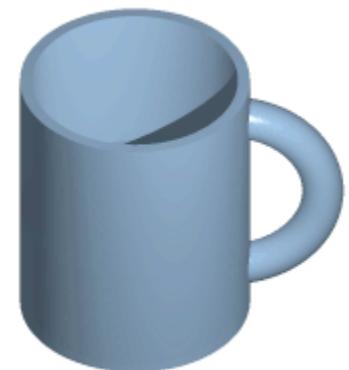
- ▶ Distances and angles
  - ▶ area, volume, curvatures, etc
- ▶ Euclidean geometry
- ▶ Riemannian geometry
  - ▶ Hyperbolic geometry
  - ▶ Spherical geometry
- ▶ ...



Using geometry, by  
unknown artist,  
15th century

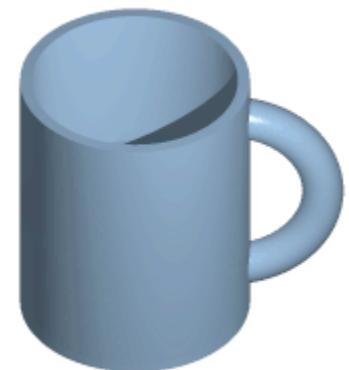
# Topology

- ▶ Detailed geometric information not sufficient
  - ▶ Too local
  - ▶ Or not necessary
  - ▶ Or may even be harmful
  - ▶ Wish to identify key information, “qualitative” structure
  
- ▶ Topology
  - ▶ Connectivity
  - ▶ Holes, Voids

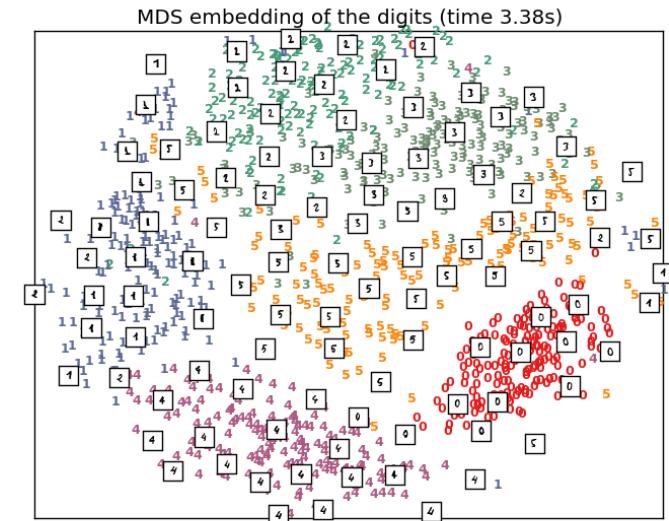
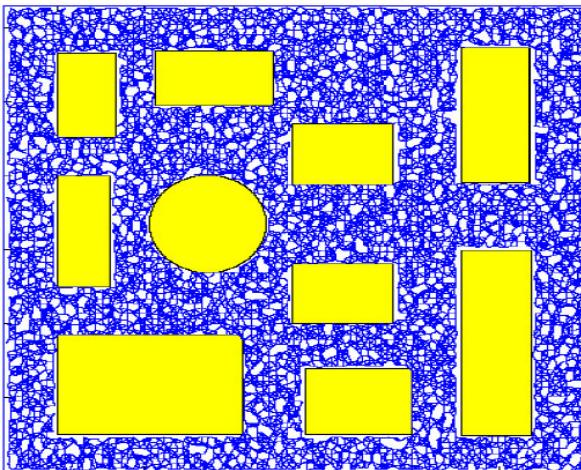
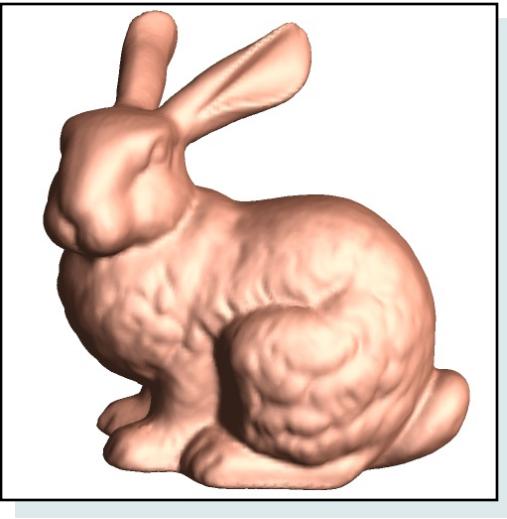


# Topology

- ▶ Detailed geometric information not sufficient
  - ▶ Too local
  - ▶ Or not necessary
  - ▶ Or may even be harmful
  - ▶ Wish to identify key information, “qualitative” structure
  
- ▶ Topology
  - ▶ Connectivity
  - ▶ Holes, Voids



# Data has shape!



Data has shape!

# Data has shape!

- ▶ **TGDA**: Topological and geometric data analysis
  - ▶ Develop effective methodologies and **algorithms** for data analysis through the **topological** and **geometric** lens.

# Data has shape!

- ▶ **TGDA:** Topological and geometric data analysis
  - ▶ Develop effective methodologies and **algorithms** for data analysis through the **topological** and **geometric** lens.
- ▶ This course: *Topological data analysis* (TDA)
  - ▶ Focus on topological concepts, structures, and algorithms for data analysis

# TDA - a brief introduction

- ▶ Aims at extracting and analyzing topological information from data sets
  - ▶ Coarse yet essential information
  - ▶ Characterization, feature identification
  - ▶ General, powerful tools for both space and functions defined on a space
  - ▶ Elegant mathematical understanding available
- ▶ Young and active (started around 2000)
- ▶ Interdisciplinary: computer science, mathematics, statistics, etc
- ▶ Many applications
  - ▶ Visualization, medical image processing, computational neuron science, computational biology, material science
  - ▶ TDA + machine learning

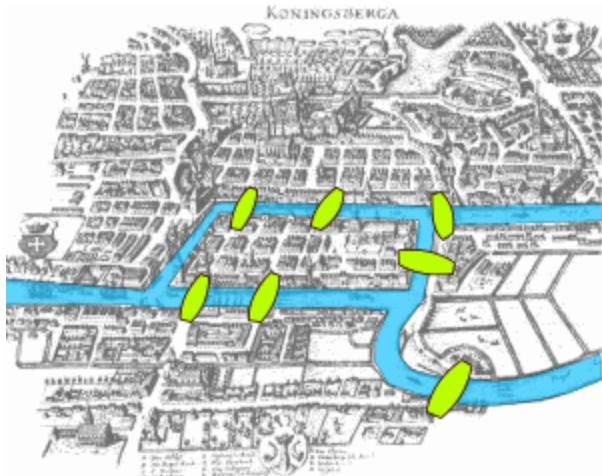
# Introduction to Topology

# History

- ▶ Seven Bridges of Königsberg (1736)

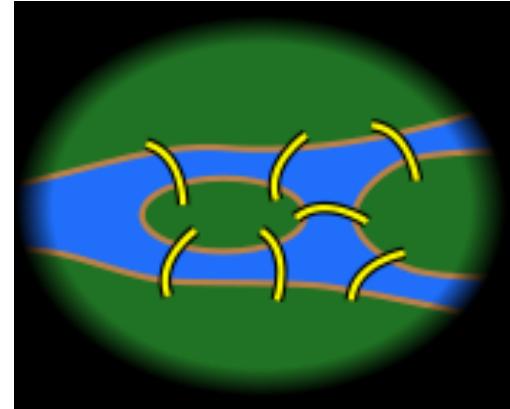
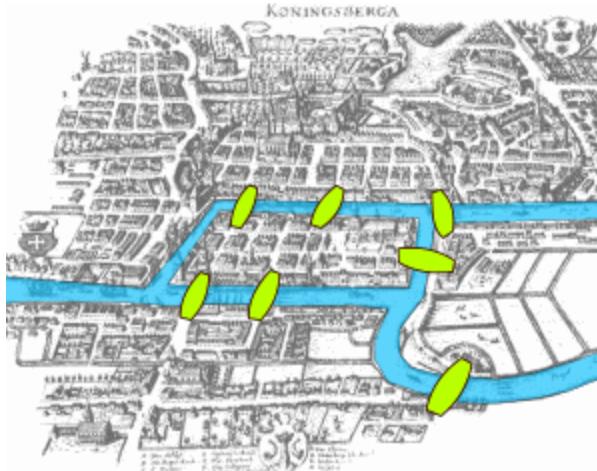
# History

- ▶ Seven Bridges of Königsberg (1736)



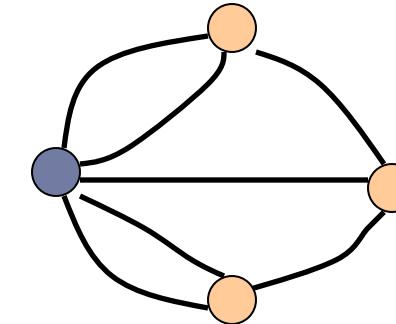
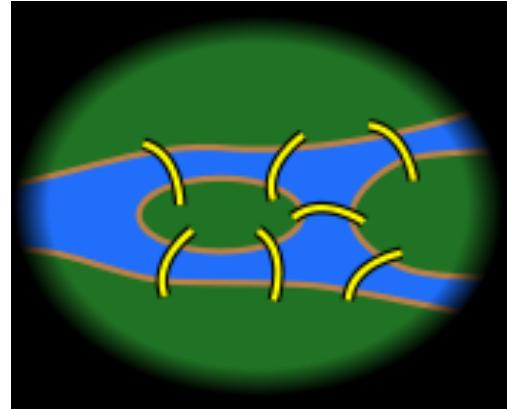
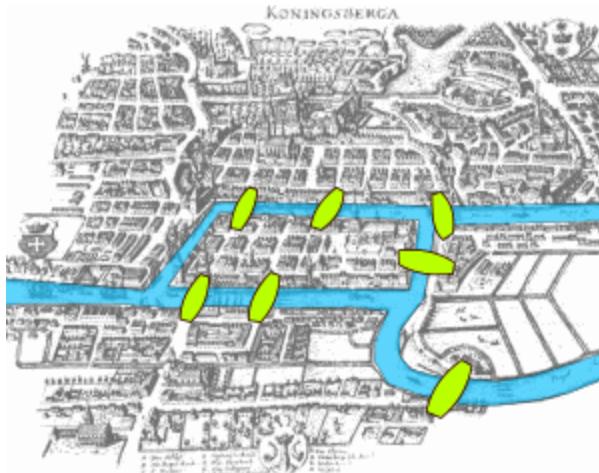
# History

- ▶ Seven Bridges of Königsberg (1736)



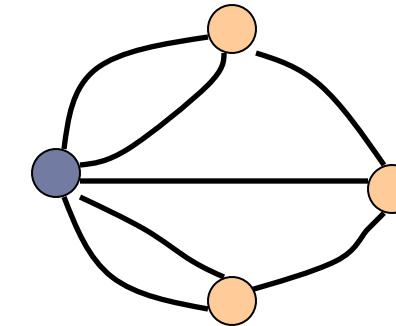
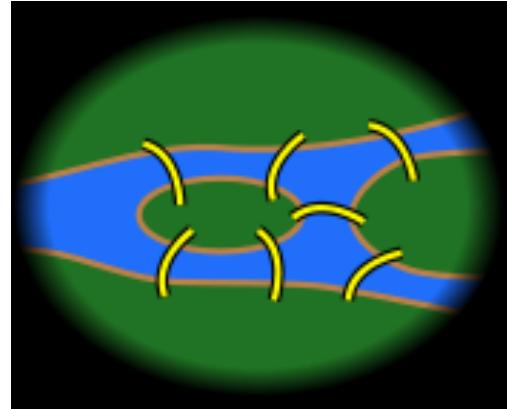
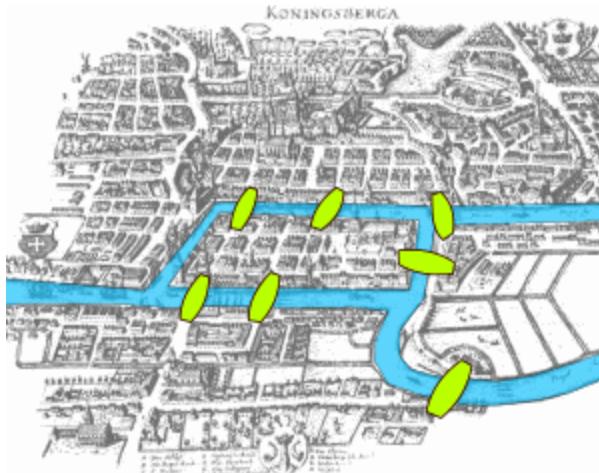
# History

- ▶ Seven Bridges of Königsberg (1736)



# History

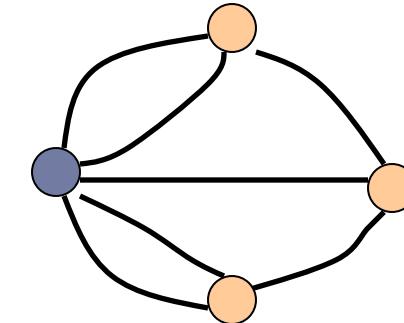
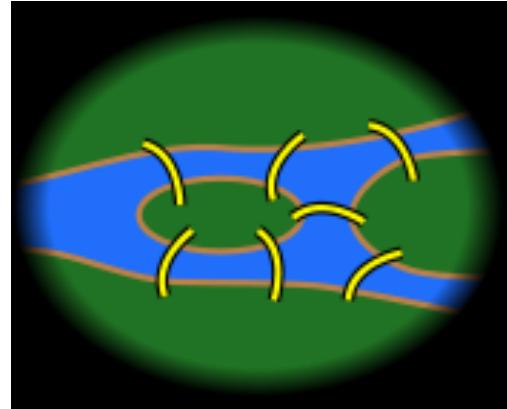
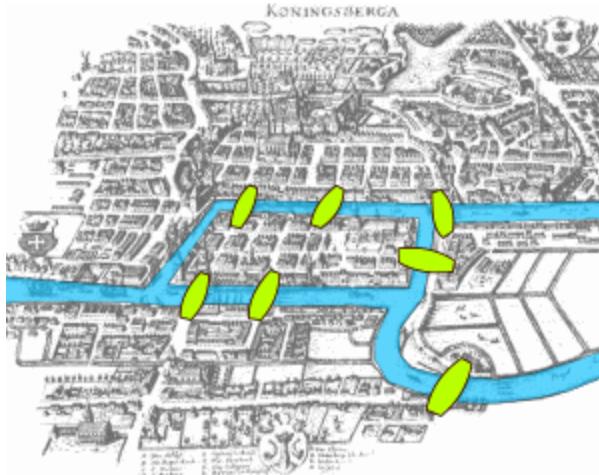
- ▶ Seven Bridges of Königsberg (1736)



Euler cycle problem

# History

- ▶ Seven Bridges of Königsberg (1736)

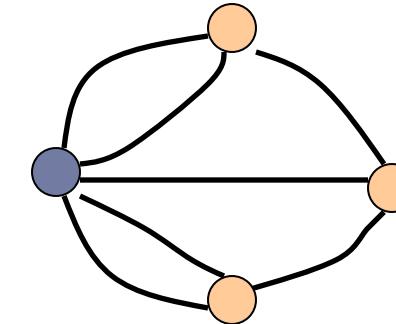
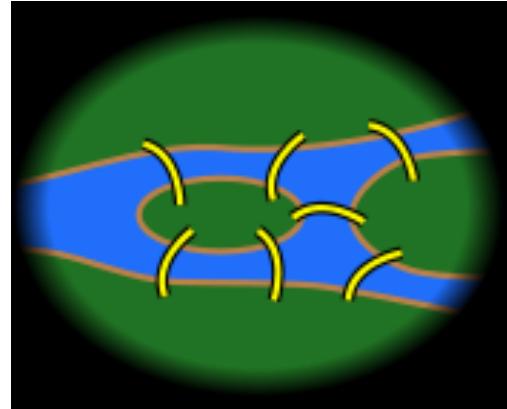
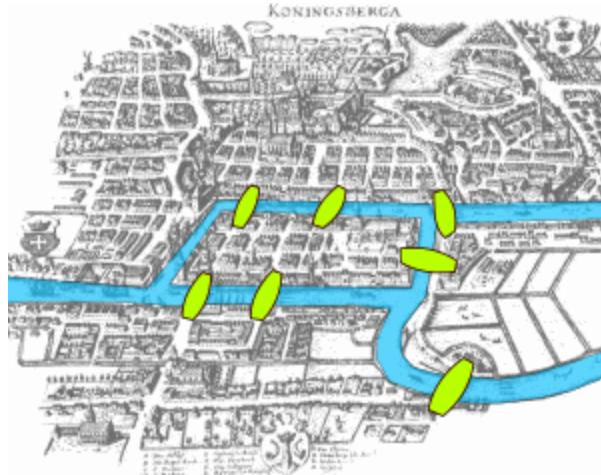


Euler cycle problem

Abstraction of connectivity

# History

- ▶ Seven Bridges of Königsberg (1736)



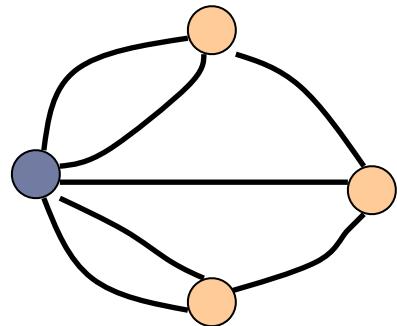
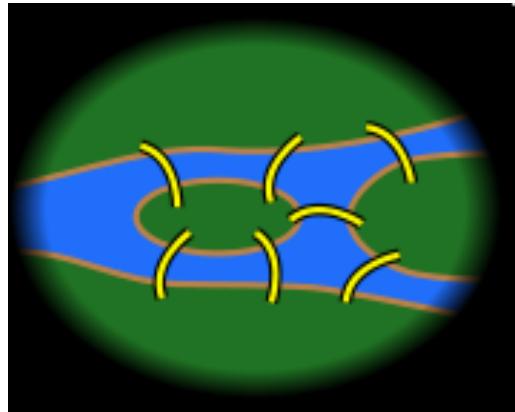
Euler cycle problem

Abstraction of connectivity

Topology: “ distinguish qualitative geometry from the ordinary geometry in which quantitative relations chiefly are treated ”

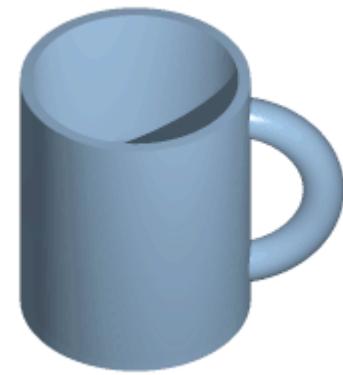
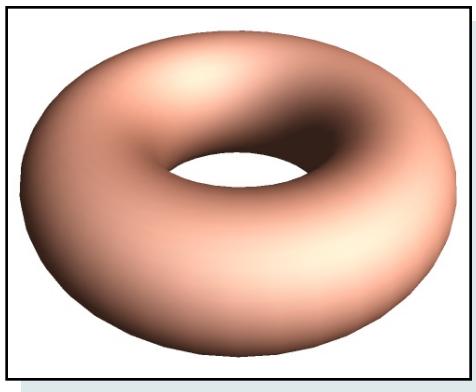
# Fundamental Questions

- ▶ How do you know when the two spaces are the “same”?
- ▶ Why things are the “same” after deformation?



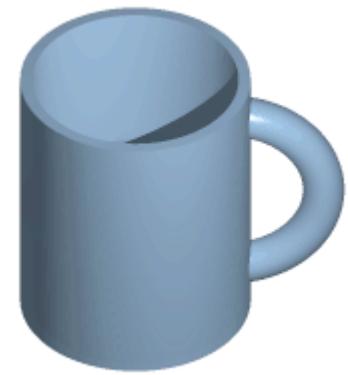
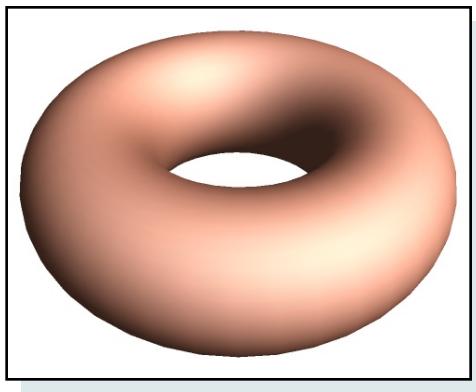
# Homeomorphism

- ▶ Intuitively, two spaces have the same topology if one can continuously deform one to the other without breaking, gluing, and inserting new things



# Homeomorphism

- ▶ Intuitively, two spaces have the same topology if one can continuously deform one to the other without breaking, gluing, and inserting new things

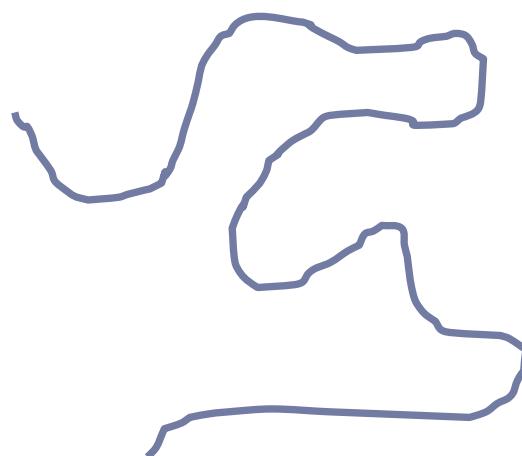


# Homeomorphism

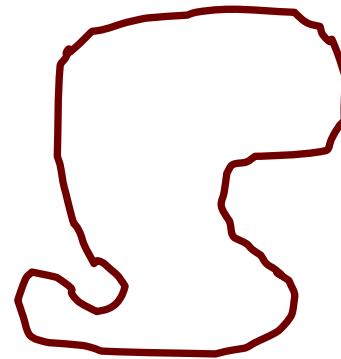
- ▶ Intuitively, two spaces have the same topology if one can continuously deform one to the other without breaking, gluing, and inserting new things

# Homeomorphism

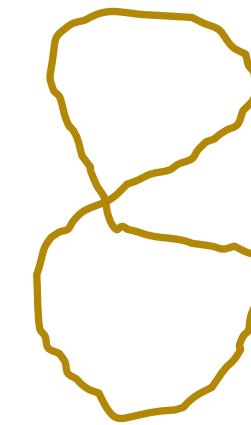
- ▶ Intuitively, two spaces have the same topology if one can continuously deform one to the other without breaking, gluing, and inserting new things



open curve



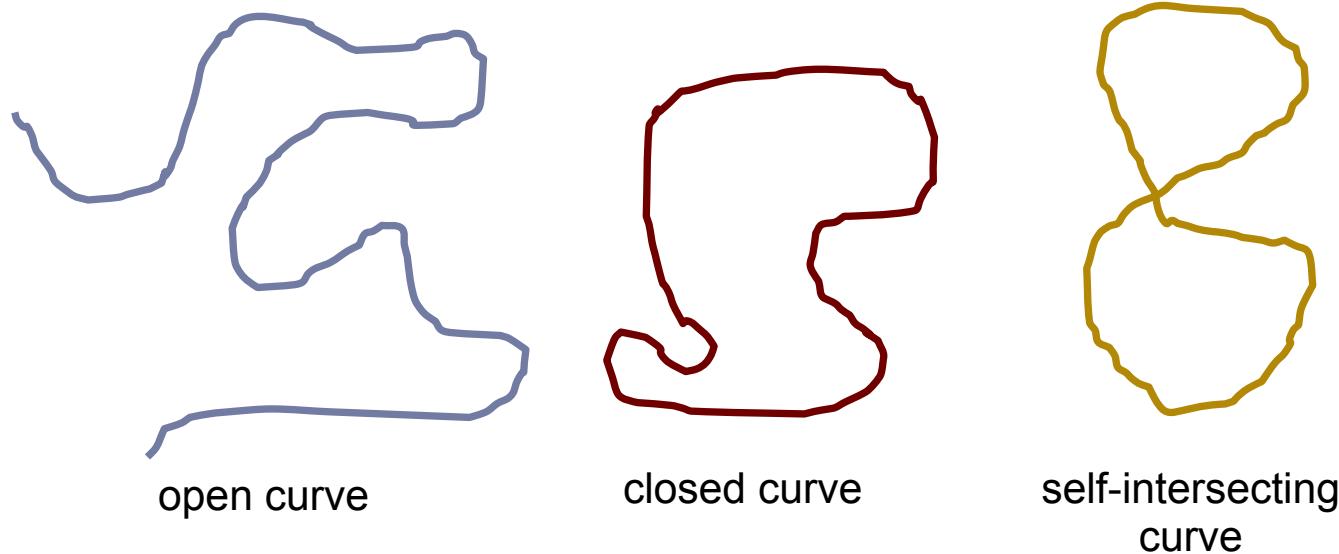
closed curve



self-intersecting  
curve

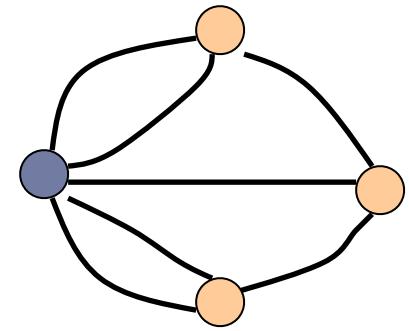
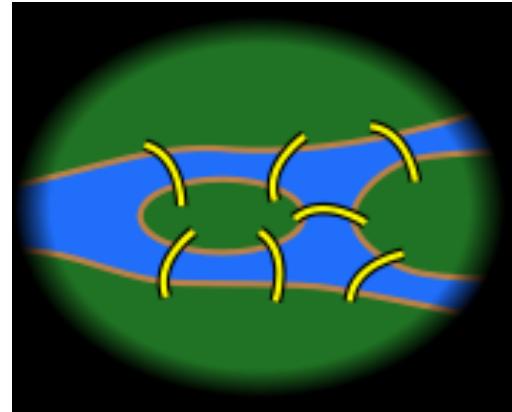
# Homeomorphism

- ▶ Intuitively, two spaces have the same topology if one can continuously deform one to the other without breaking, gluing, and inserting new things

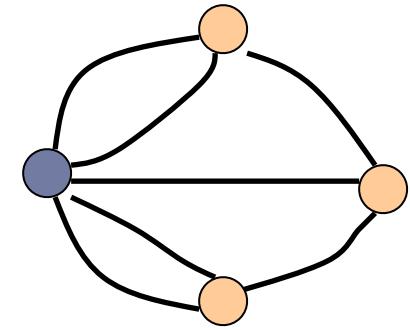
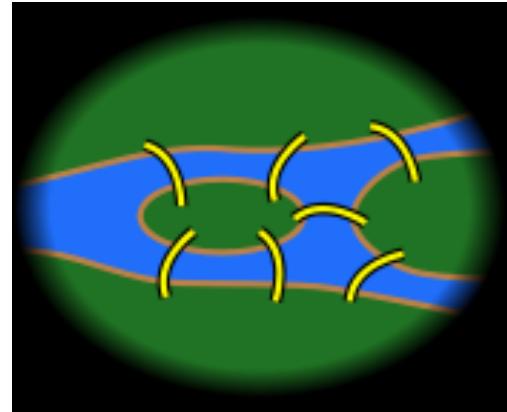
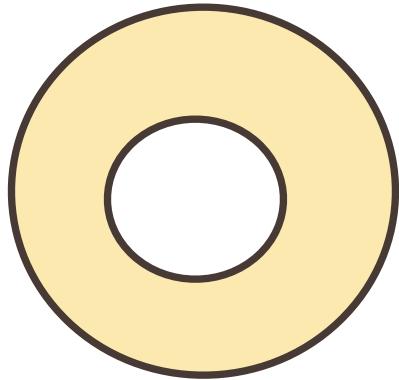
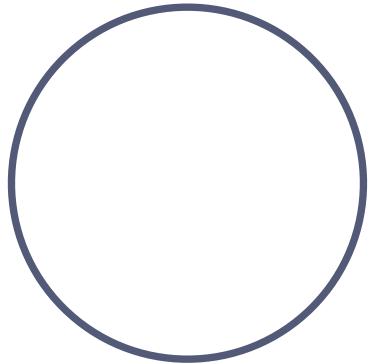


Two spaces with the same topology are ***homeomorphic***

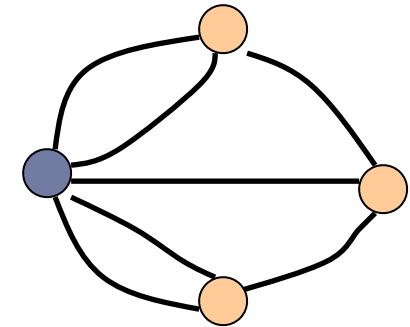
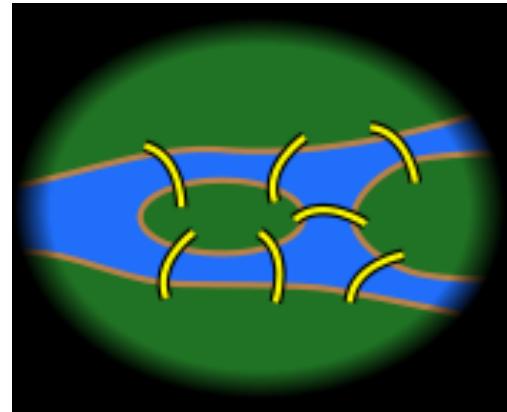
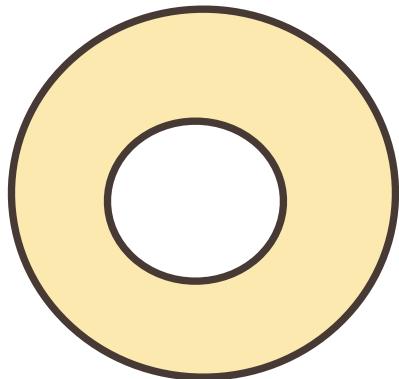
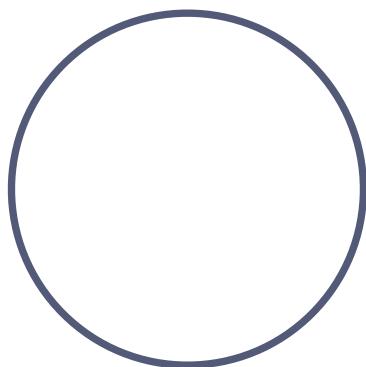
# Homotopy Equivalent - A Relaxation of Homeomorphism



# Homotopy Equivalent - A Relaxation of Homeomorphism



# Homotopy Equivalent - A Relaxation of Homeomorphism



Homotopy equivalent - “same” in a topologist’s perspective



- ▶ All these are so abstract!! They belong to **Point Set Topology**

- ▶ All these are so abstract!! They belong to **Point Set Topology**

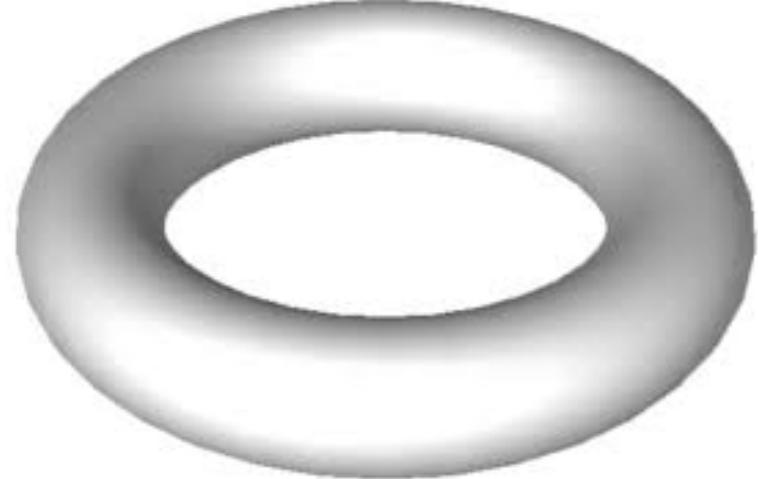
- ▶ All these are so abstract!! They belong to **Point Set Topology**

- ▶ All these are so abstract!! They belong to **Point Set Topology**
- ▶ How can we use these notions in practice?

- ▶ All these are so abstract!! They belong to **Point Set Topology**
- ▶ How can we use these notions in practice?
  - ▶ Algebraic topology

# Homology - Algebraic Tools for Determining Homotopy

- ▶ Homology can count # connected components, # holes, # voids.



- ▶ The above two spaces are not homotopy equivalent

# Topological Quantities

# Topological Quantities

- ▶ Homeomorphism → homotopy equivalence
  - ▶ Different levels of similarity between spaces

# Topological Quantities

- ▶ Homeomorphism → homotopy equivalence
  - ▶ Different levels of similarity between spaces
- ▶ Quantities invariant under them (topological quantities)

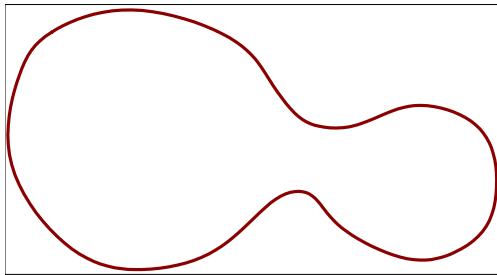
# Topological Quantities

- ▶ Homeomorphism → homotopy equivalence
  - ▶ Different levels of similarity between spaces
- ▶ Quantities invariant under them (topological quantities)  
    ⇒ *(Essential) features*
  - ▶ Powerful for feature identification and characterization
  - ▶ This course will give
    - ▶ definitions, intuition, and their computation
    - ▶ also examples of applications

TDA - Principle

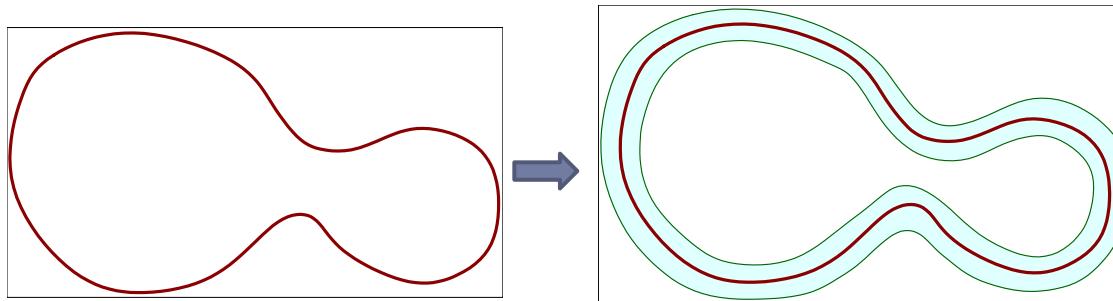
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



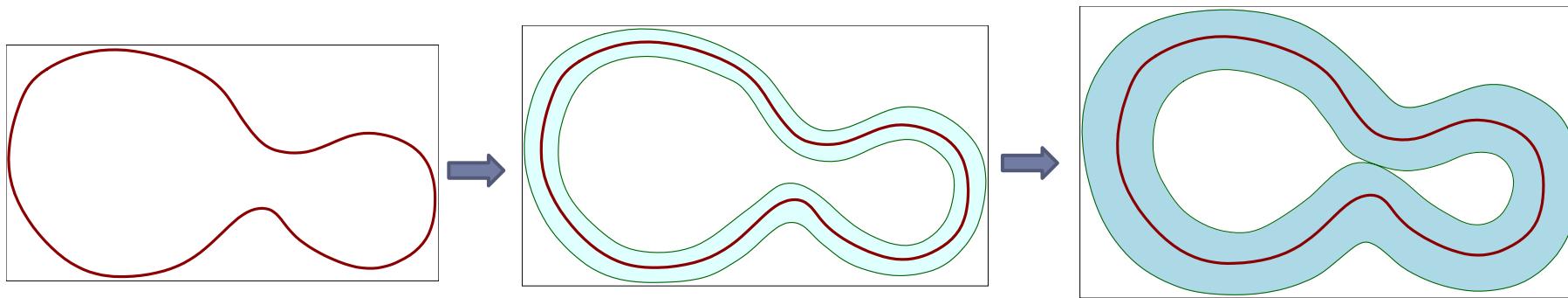
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



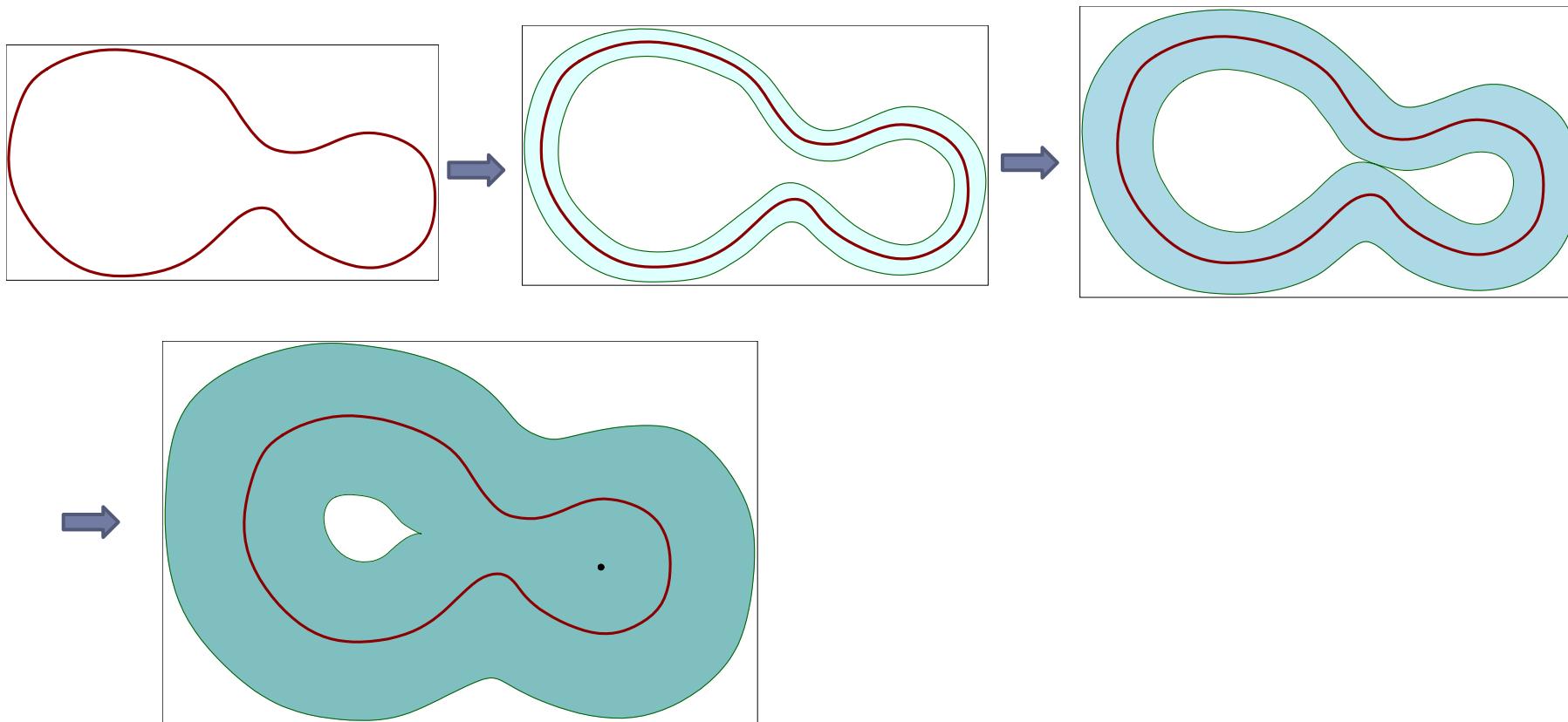
# TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



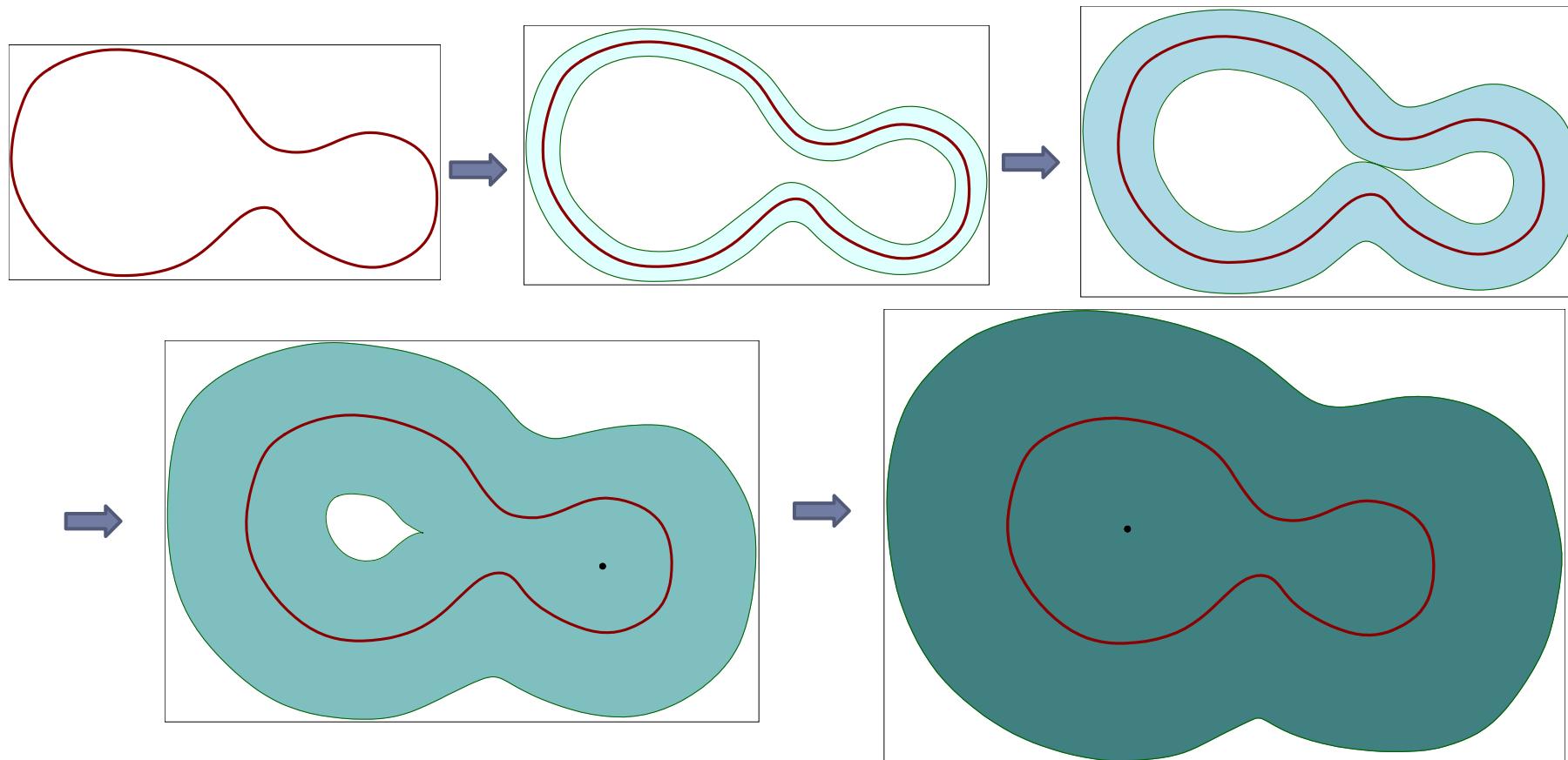
# TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



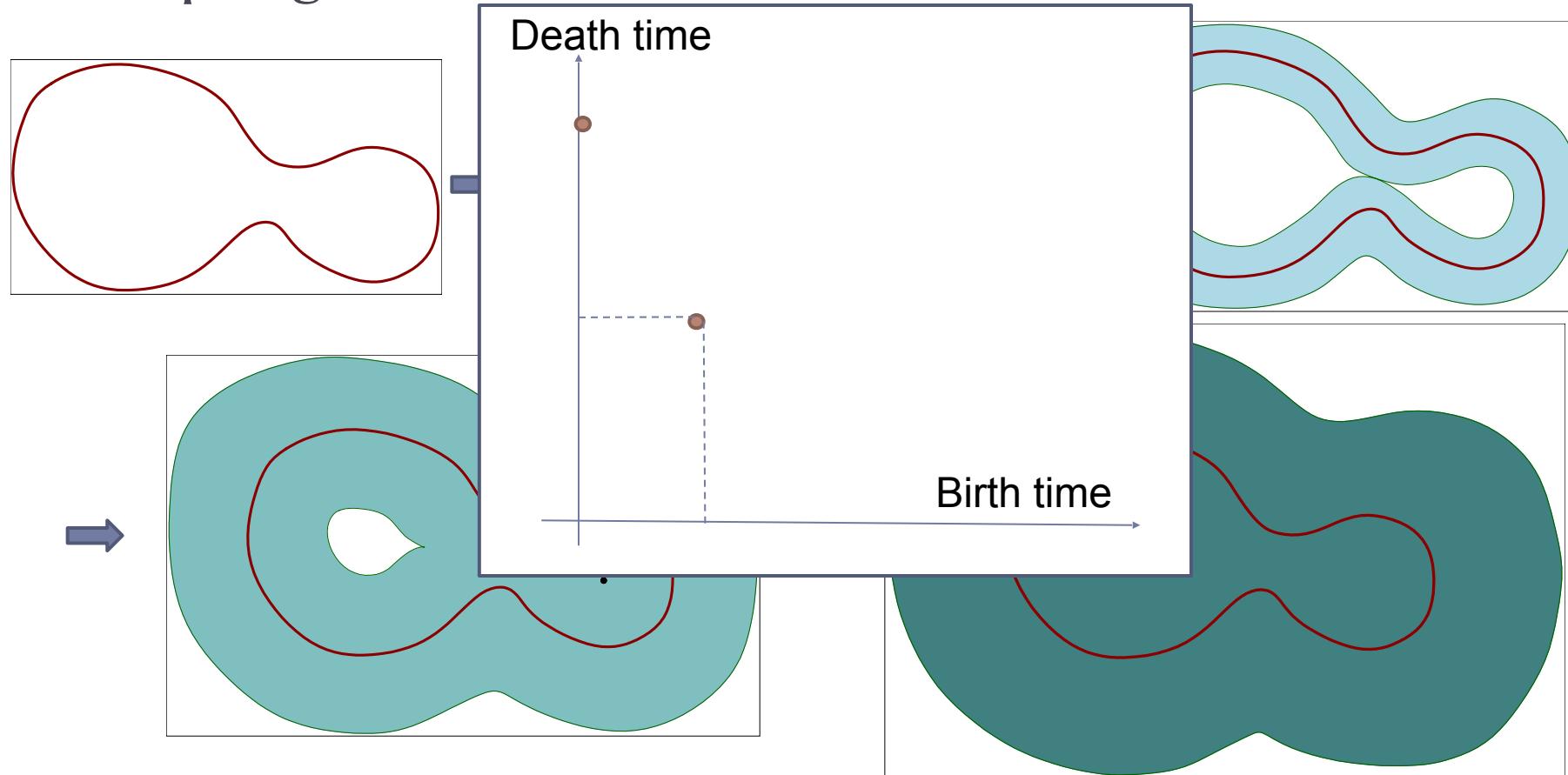
# TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



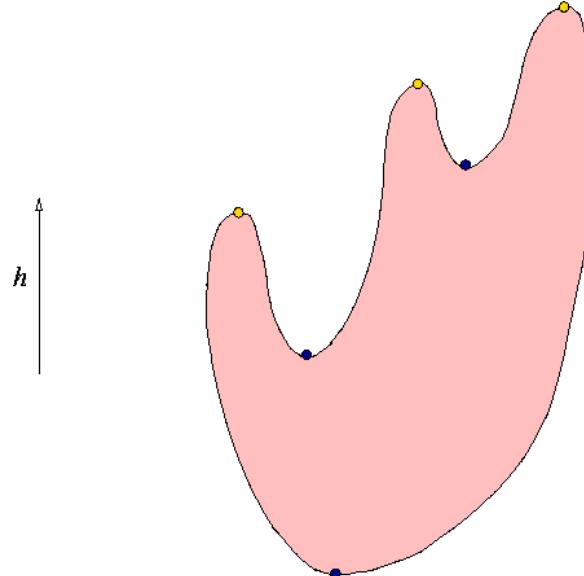
# TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



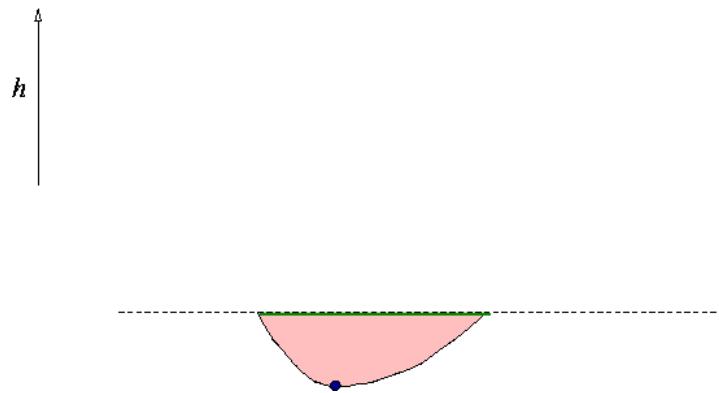
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



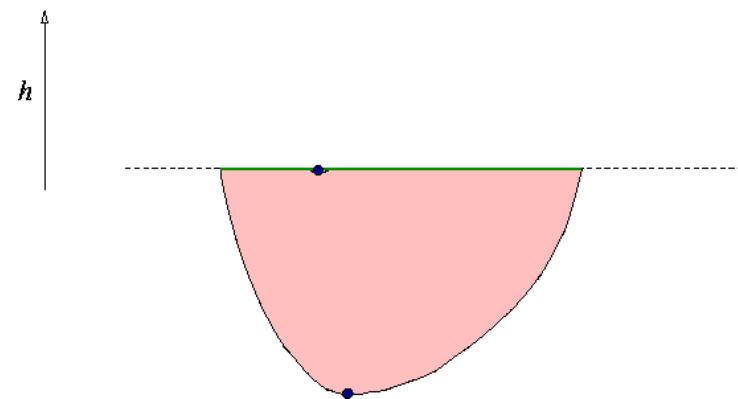
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



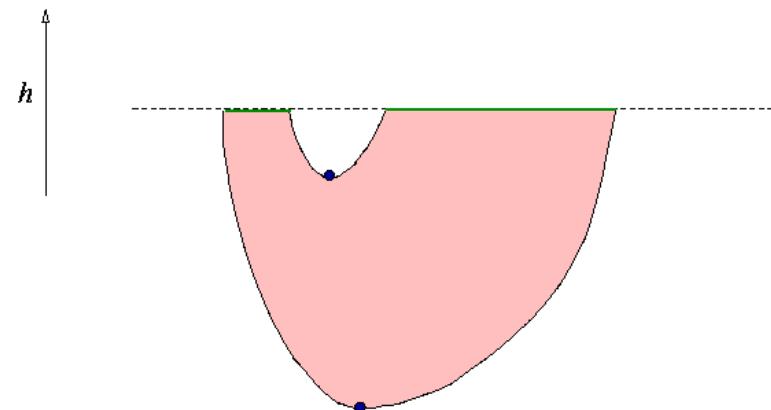
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



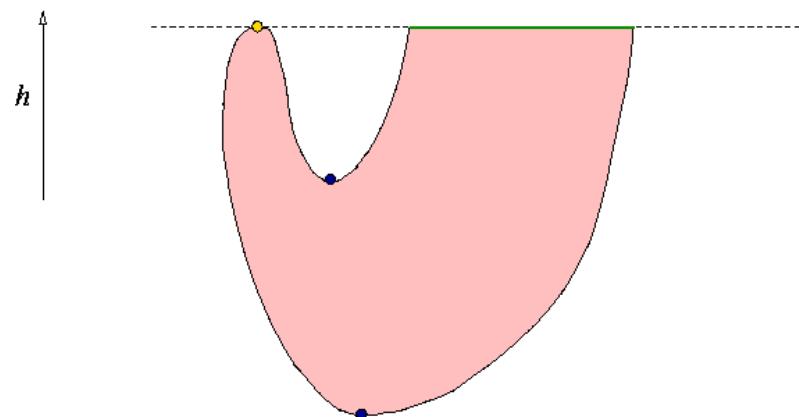
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



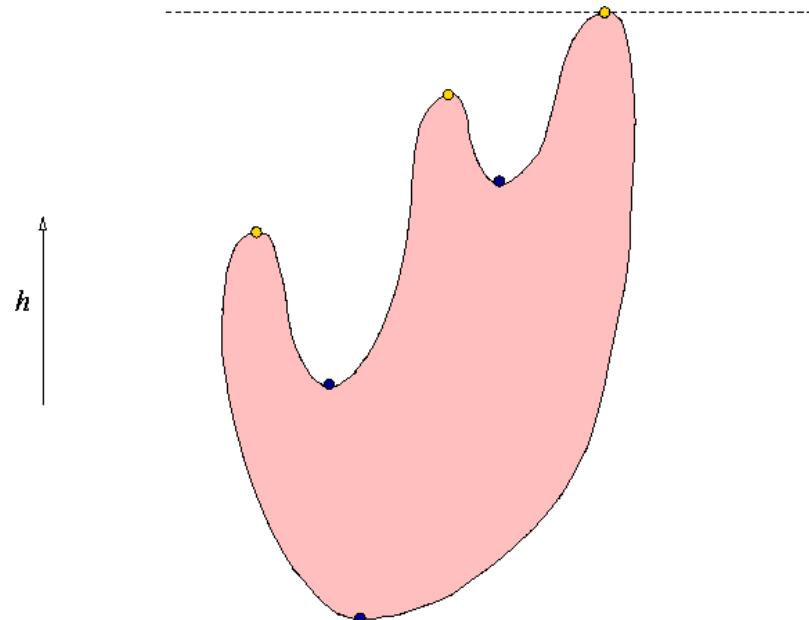
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



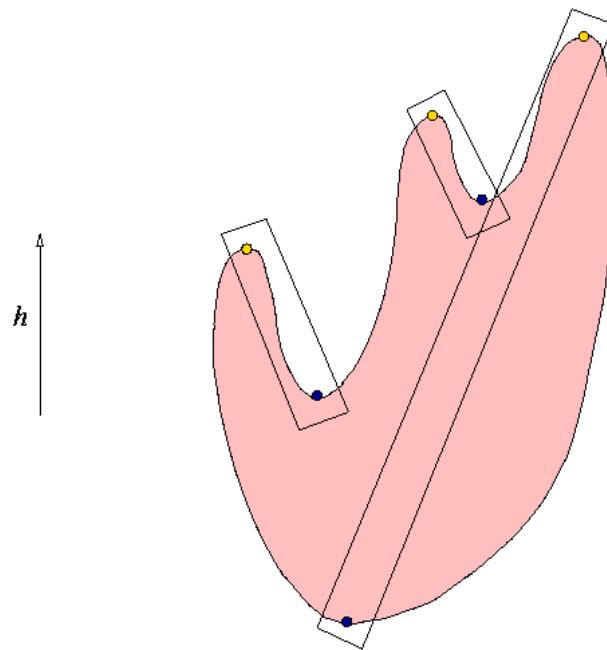
# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



# TDA Principle - Tracking Multiscale Topology

- ▶ Incorporate geometry, functions or maps of a space to create and capture multiscale topological features



# Topics

- ▶ Basics in Topology
- ▶ Common complexes (*aka. how we model space of interests*)
- ▶ (Simplicial) homology (*aka. how do we quantify topological features*)
  
- ▶ Persistent homology (*aka. powerful modern extension of homological features*)
- ▶ Analysis of point cloud data (PCD) and graph data
  
- ▶ Analysis of functions on data
- ▶ Discrete Morse theory (*higher order skeletal structure behind data*) and applications
- ▶ TDA and machine learning
- ▶ Spectral approaches to TDA

# Topics

- ▶ Basics in Topology
  - ▶ Common complexes (*aka. how we model space of interests*)
  - ▶ (Simplicial) homology (*aka. how do we quantify topological features*)
  
  - ▶ Persistent homology (*aka. powerful modern extension of homological features*)
  - ▶ Analysis of point cloud data (PCD) and graph data
  
  - ▶ Analysis of functions on data
  - ▶ Discrete Morse theory (*higher order skeletal structure behind data*) and applications
  - ▶ TDA and machine learning
  - ▶ Spectral methods
- Will focus on concepts, definitions, algorithms, also intuition why they work, and how they can be used.

# Applications of TDA

# Motivating Examples I

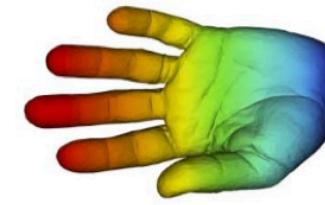
- ▶ Shape of data
- ▶ Topological summary

# Motivating Examples I

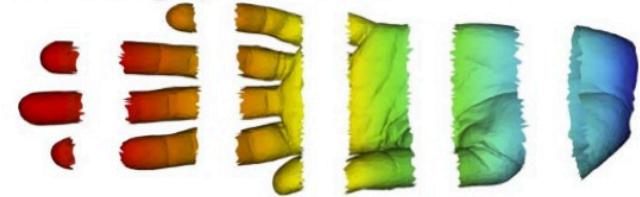
- ▶ Shape of data
- ▶ Topological summary



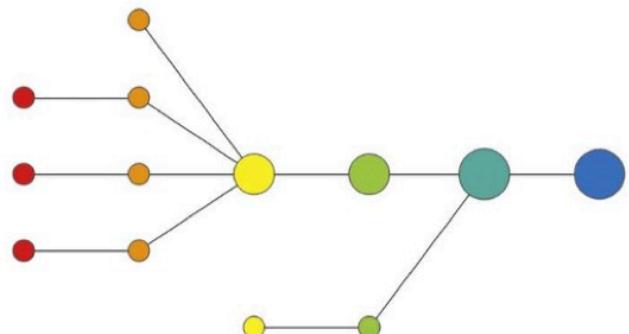
B Coloring by filter value



C Binning by filter value

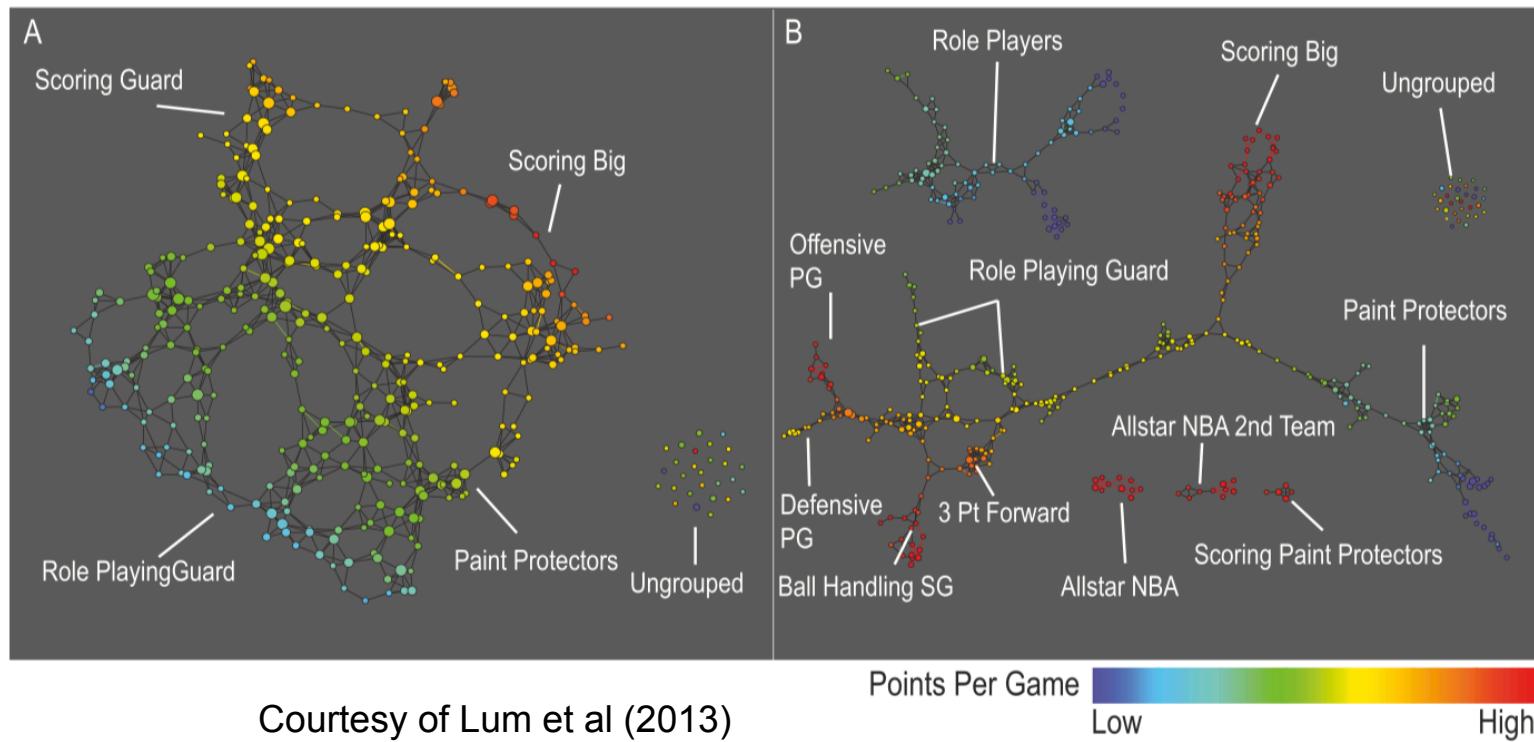


D Clustering and network construction



# Motivating Examples I

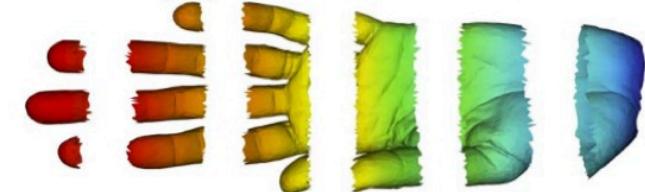
- ▶ Shape of data
- ▶ Topological summary



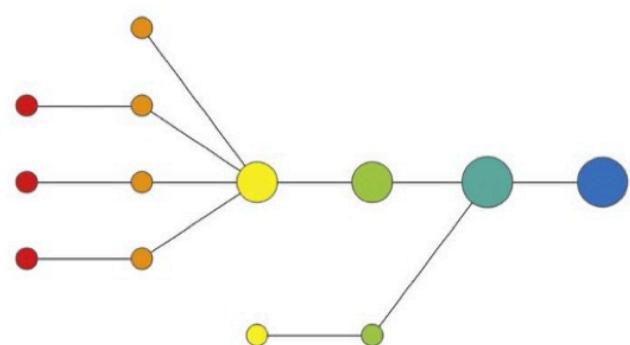
B Coloring by filter value



C Binning by filter value



D Clustering and network construction

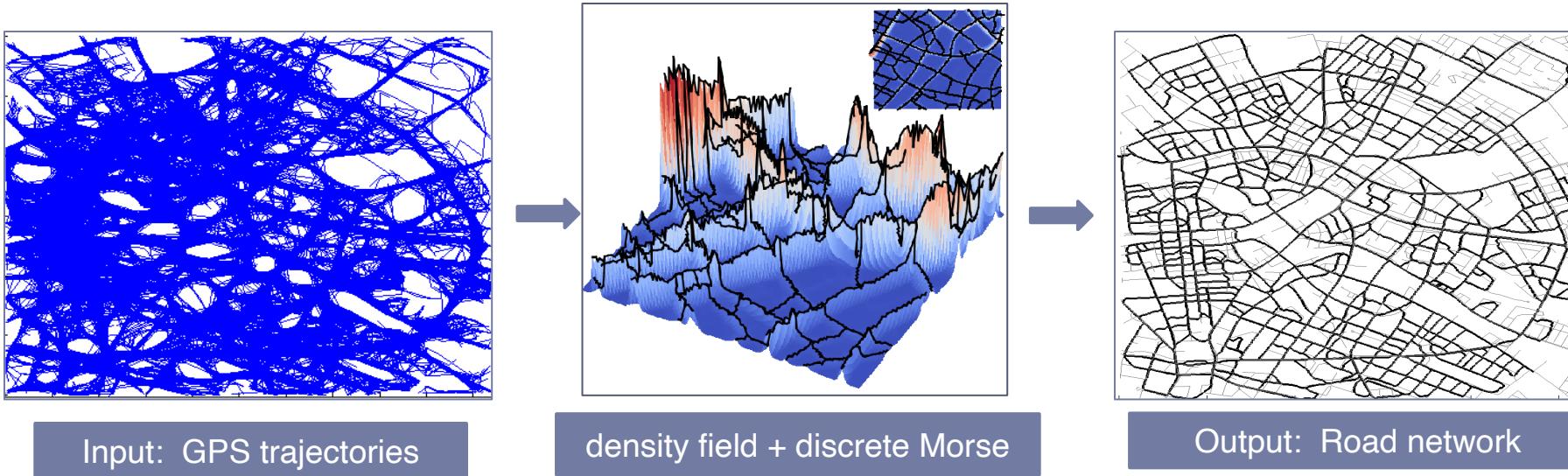


# Motivating Examples II

- ▶ Graph reconstruction
  - ▶ Road network reconstruction, neuron skeletonization, etc

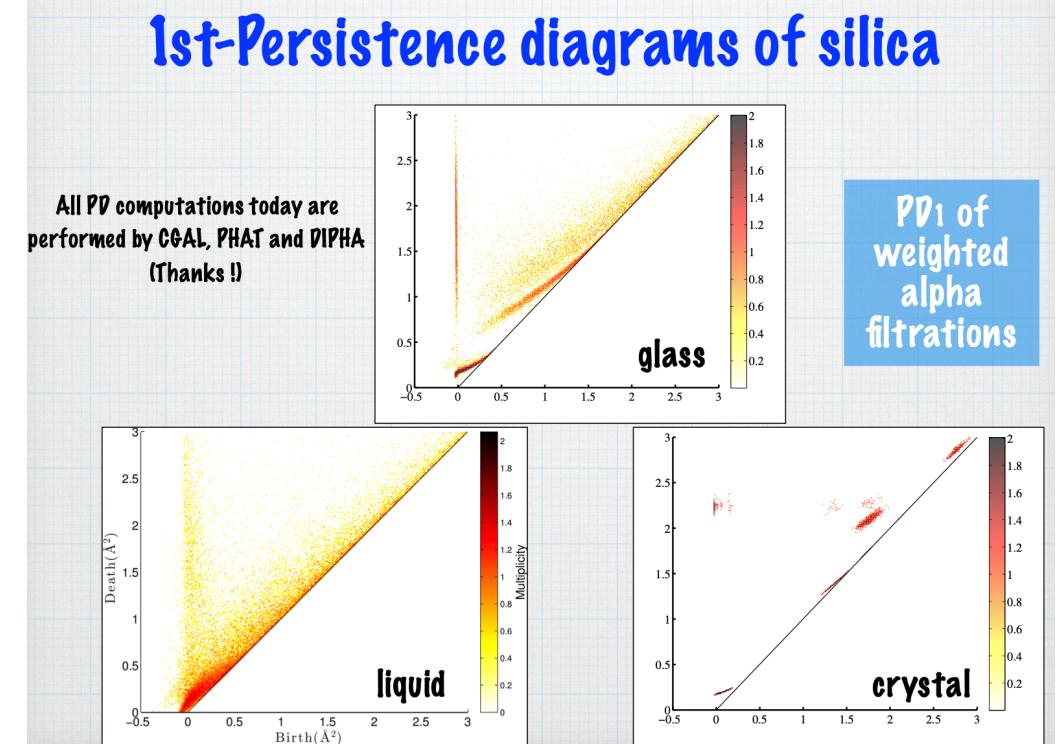
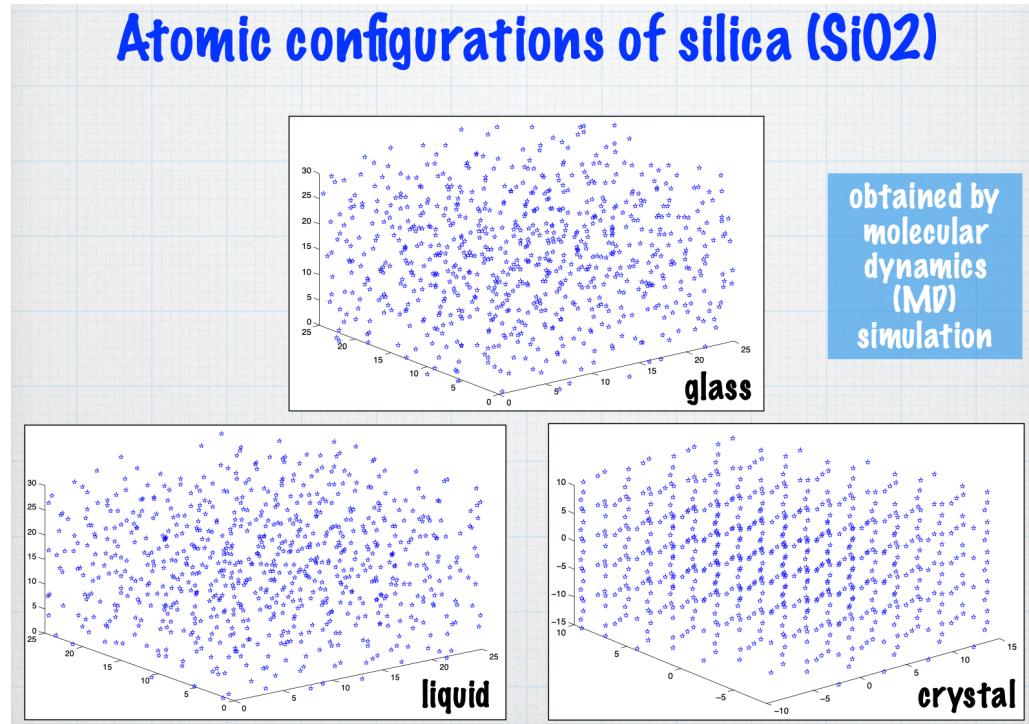
# Motivating Examples II

- ▶ Graph reconstruction
  - ▶ Road network reconstruction, neuron skeletonization, etc



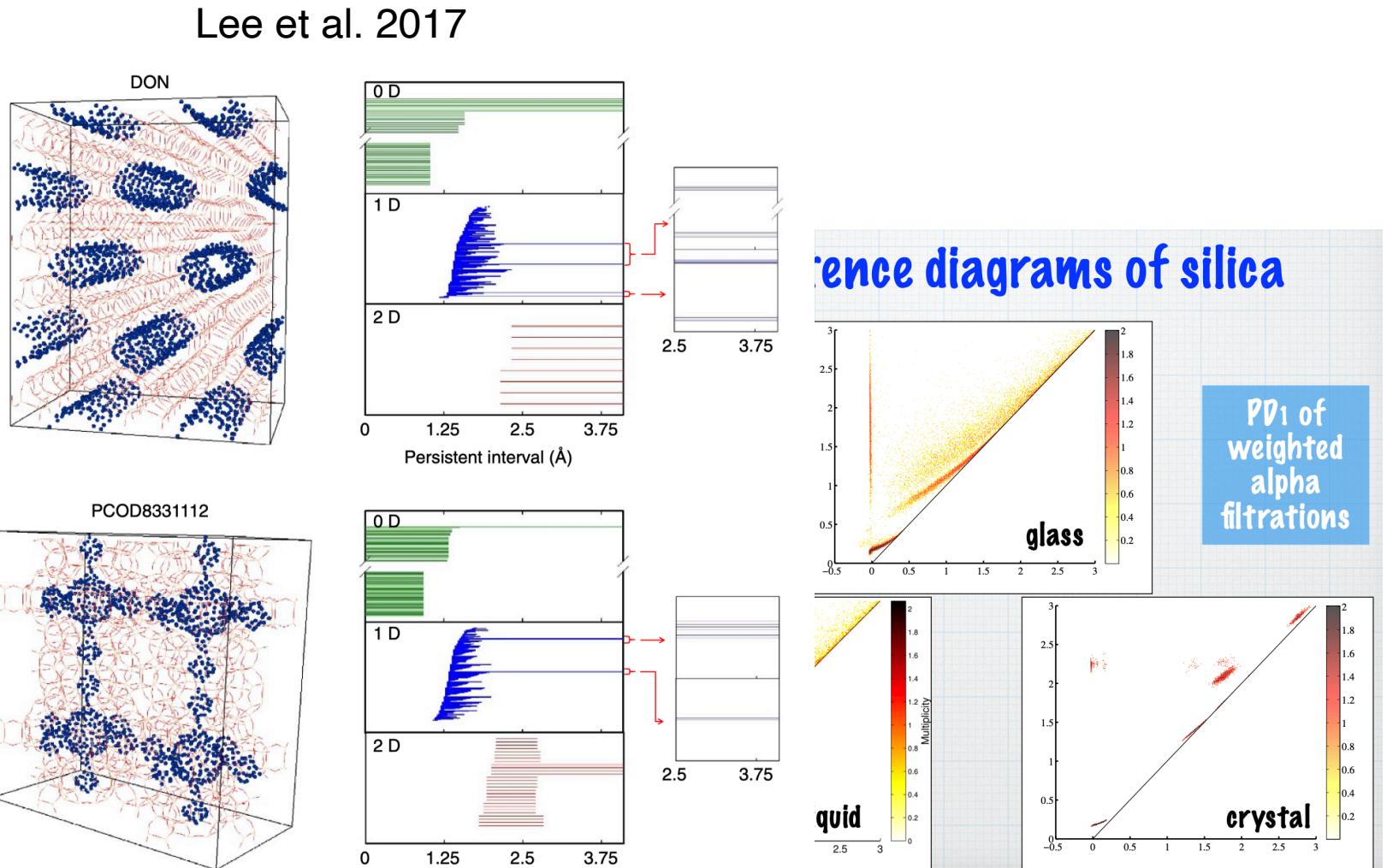
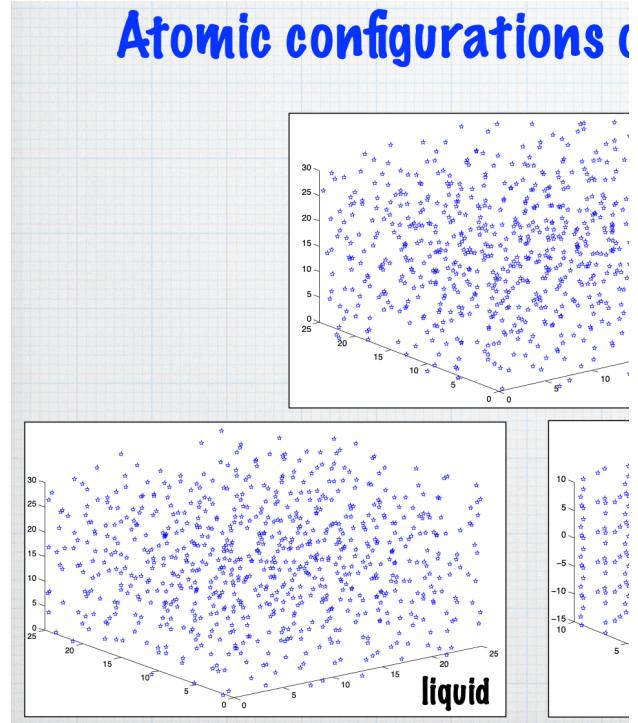
# Motivating Examples III

## ► Material Science



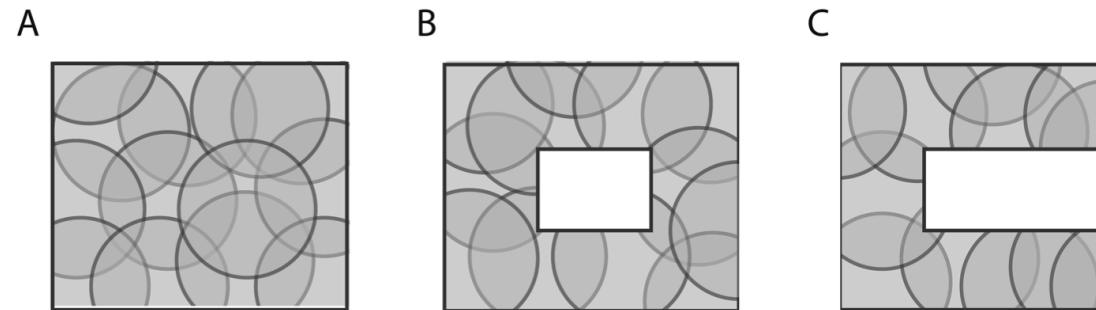
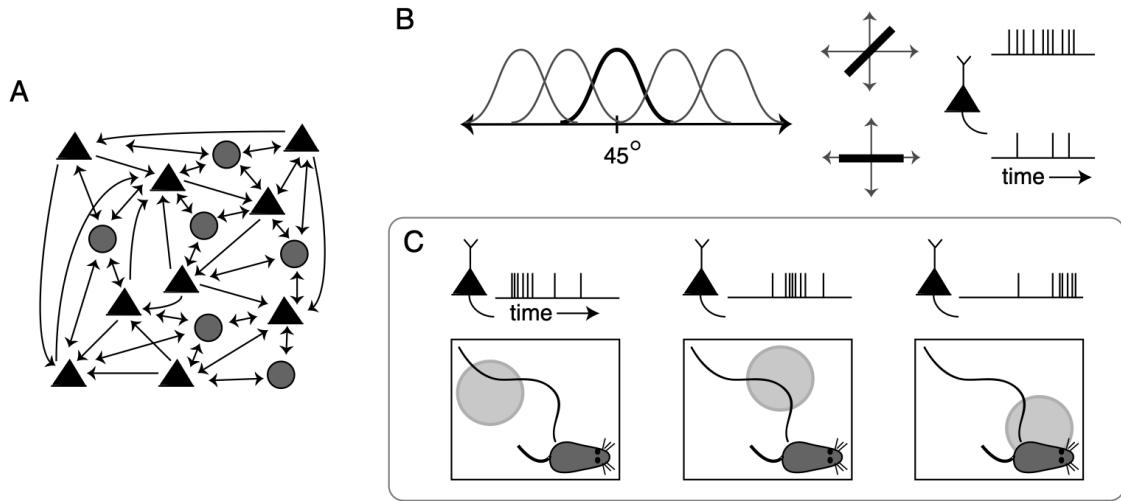
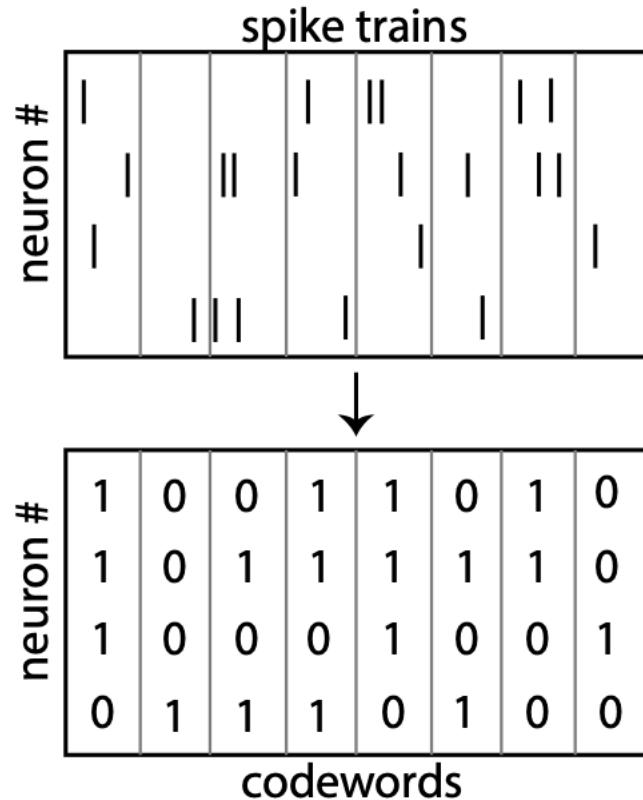
# Motivating Examples III

## ► Material Science



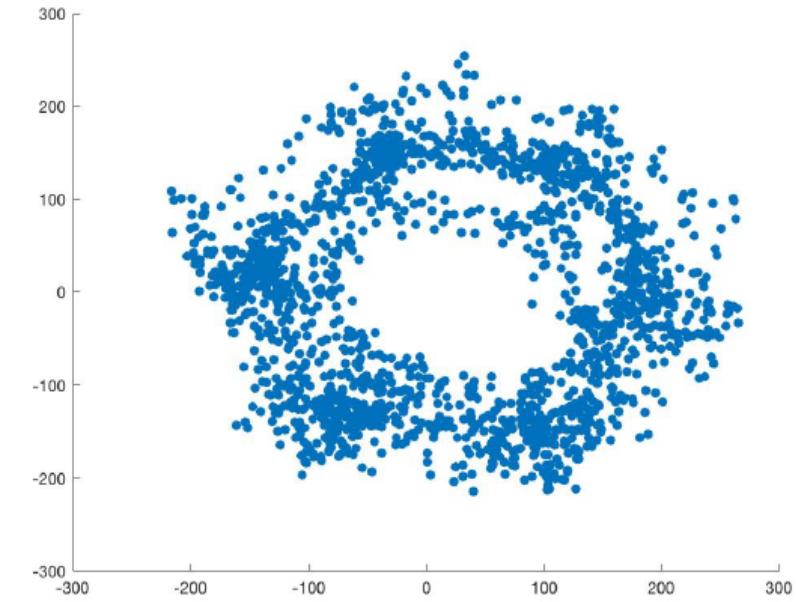
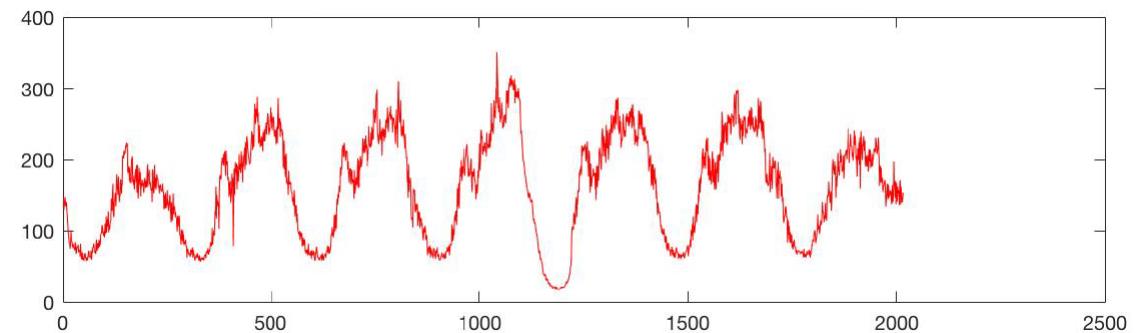
# Motivating Examples IV

## ► Neural Science



# Motivating Examples V

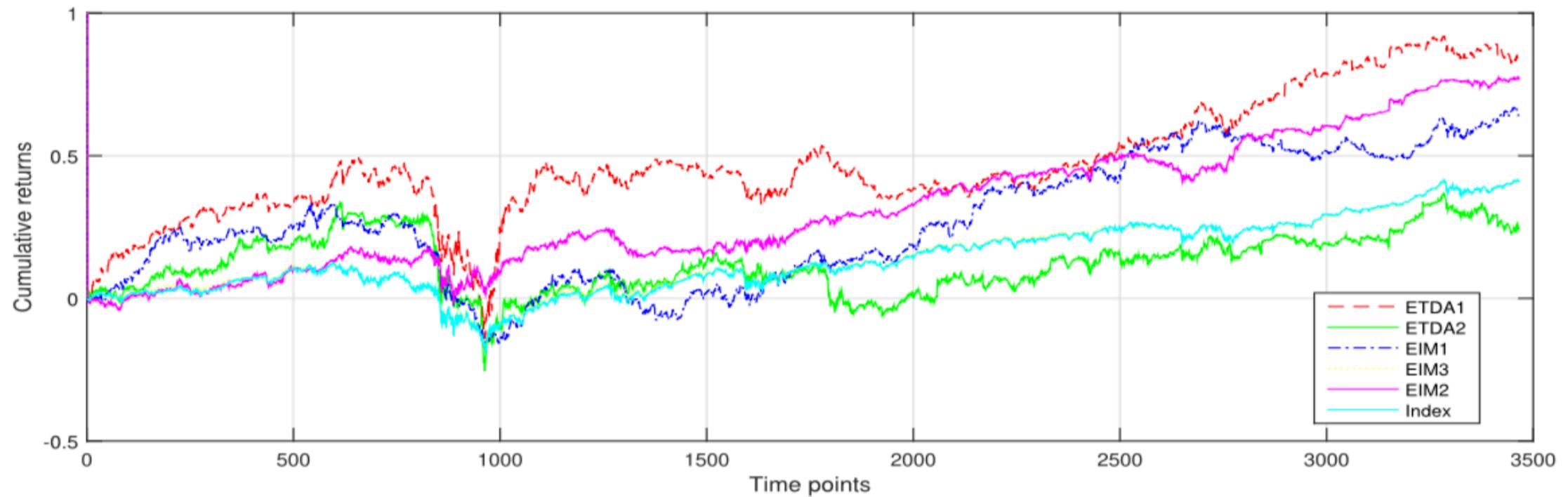
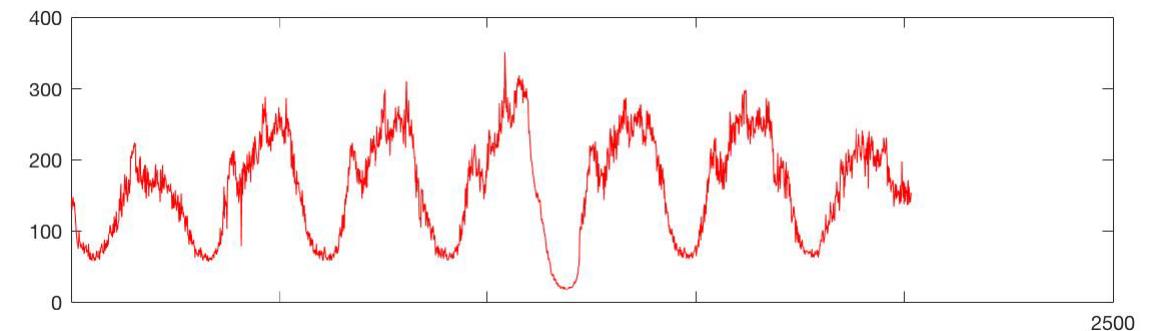
- ▶ Time series analysis
  - ▶ Finance - investment decisions



(a) Week 1

# Motivating Examples V

- ▶ Time series analysis
  - ▶ Finance - investment decisions

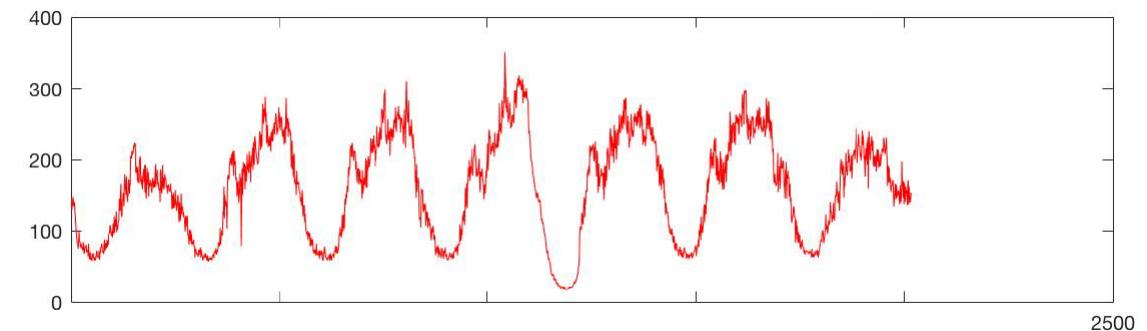


(d) Dow Jones Index

Goel et al. (2020)

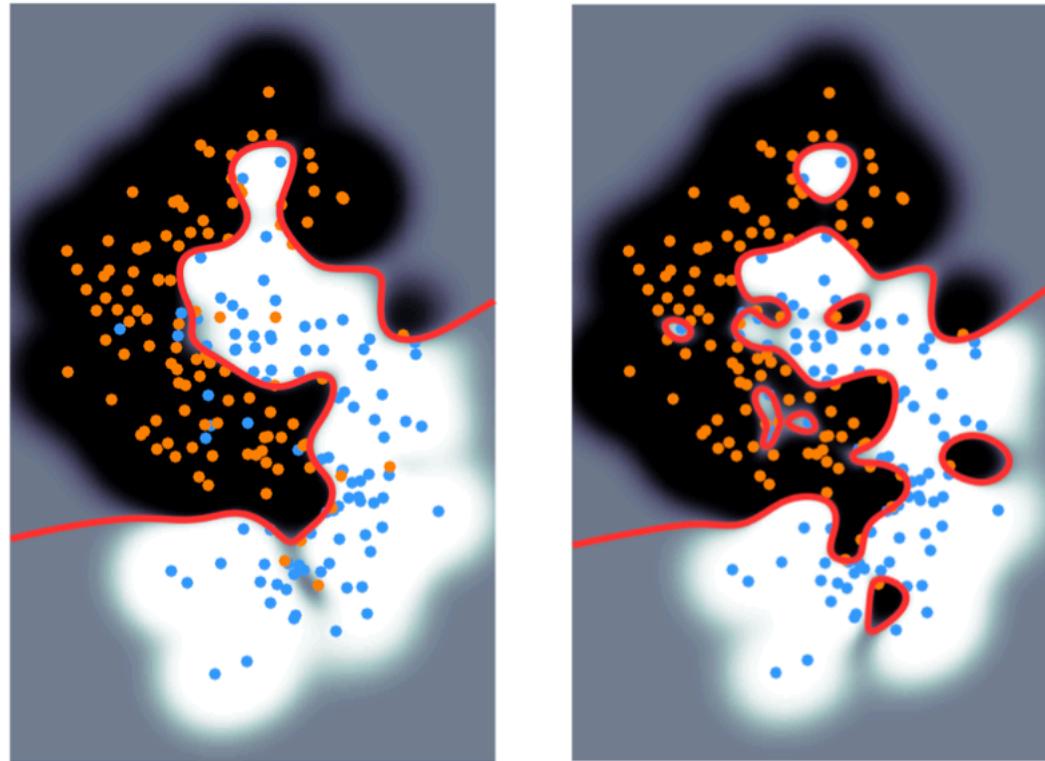
# Motivating Examples V

- ▶ Time series analysis
  - ▶ Finance - investment decisions



# Motivating Examples VI

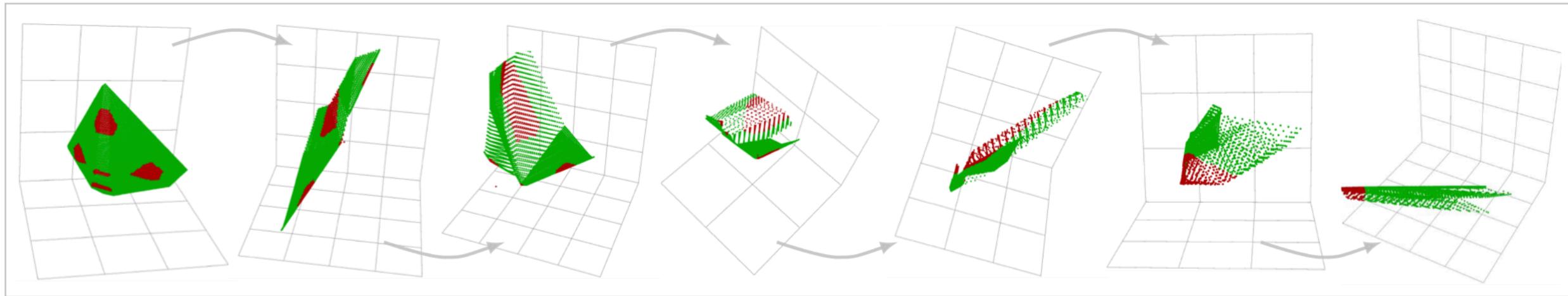
- ▶ Topological enhanced classifiers



Courtesy of Chen et al., A Topological Regularizer for Classifiers via Persistent Homology

# Motivating Examples VII

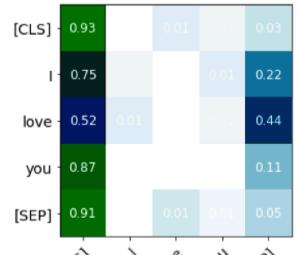
## ► Topology of neural networks



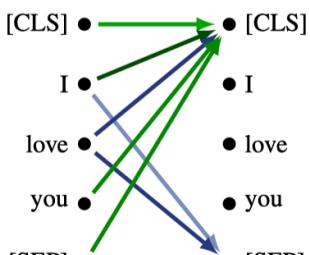
Courtesy of Naitzat et al., Topology of Deep Neural Networks

# Motivating Examples VIII

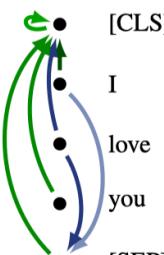
## ▶ Large Language Models



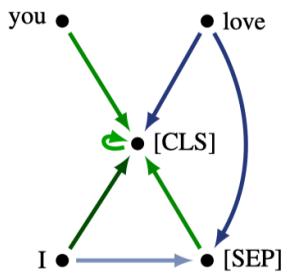
(a)



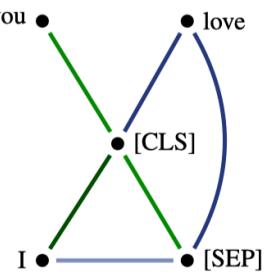
(b)



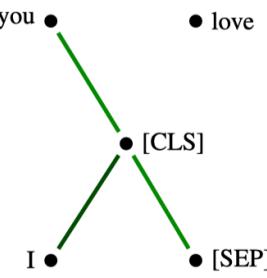
(c)



(d)



(e)



(f)

### Artificial Text Detection via Examining the Topology of Attention Maps

Laida Kushnareva<sup>1\*</sup>, Daniil Cherniavskii<sup>2\*</sup>, Vladislav Mikhailov<sup>3\*</sup>,  
Ekaterina Artemova<sup>1,4</sup>, Serguei Barannikov<sup>2,5</sup>, Alexander Bernstein<sup>2</sup>,  
Irina Piontkovskaya<sup>1</sup>, Dmitri Piontkovski<sup>4</sup>, Evgeny Burnaev<sup>2</sup>

<sup>1</sup>Huawei Noah's Ark lab, <sup>2</sup>Skolkovo Institute of Science and Technology,

<sup>3</sup>SberDevices, <sup>4</sup>HSE University, <sup>5</sup>CNRS, IMJ

# A Nice DataBase for TDA - Applications

## ► Zotero Groups: TDA-Applications

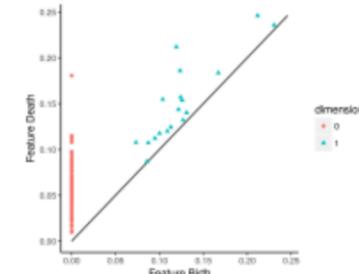
[Home](#) > [Groups](#) > TDA-Applications

### TDA-Applications

Group Library

#### Recently Added Items

Title	Added By	Date Modified
<a href="#">Chatter Classification in Turning using Machine Learning and...</a>	<a href="#">BarbaraGiunti</a>	<a href="#">3/29/2023, 04:31:38</a>
<a href="#">A Novel Quality Clustering Methodology on Fab-Wide Wafer Map...</a>	<a href="#">BarbaraGiunti</a>	<a href="#">2/4/2023, 05:44:16</a>
<a href="#">Dynamic State Analysis of a Driven Magnetic Pendulum Using O...</a>	<a href="#">BarbaraGiunti</a>	<a href="#">2/4/2023, 05:12:07</a>
<a href="#">Topological data analysis for true step detection in periodi...</a>	<a href="#">BarbaraGiunti</a>	<a href="#">2/4/2023, 04:55:32</a>
<a href="#">Shape Terra: mechanical feature recognition based on a persi...</a>	<a href="#">BarbaraGiunti</a>	<a href="#">2/4/2023, 04:40:14</a>



A database for applications of TDA outside maths. The scope of this database is to provide an as exhaustive as possible collection of applications of TDA to real data. Therefore, works pertaining (improving of) algorithms, new mathematical techniques or other improvements of the existing methods but not containing applications to real data sets will not be added.

For further questions, please contact the owner.

**FIN**