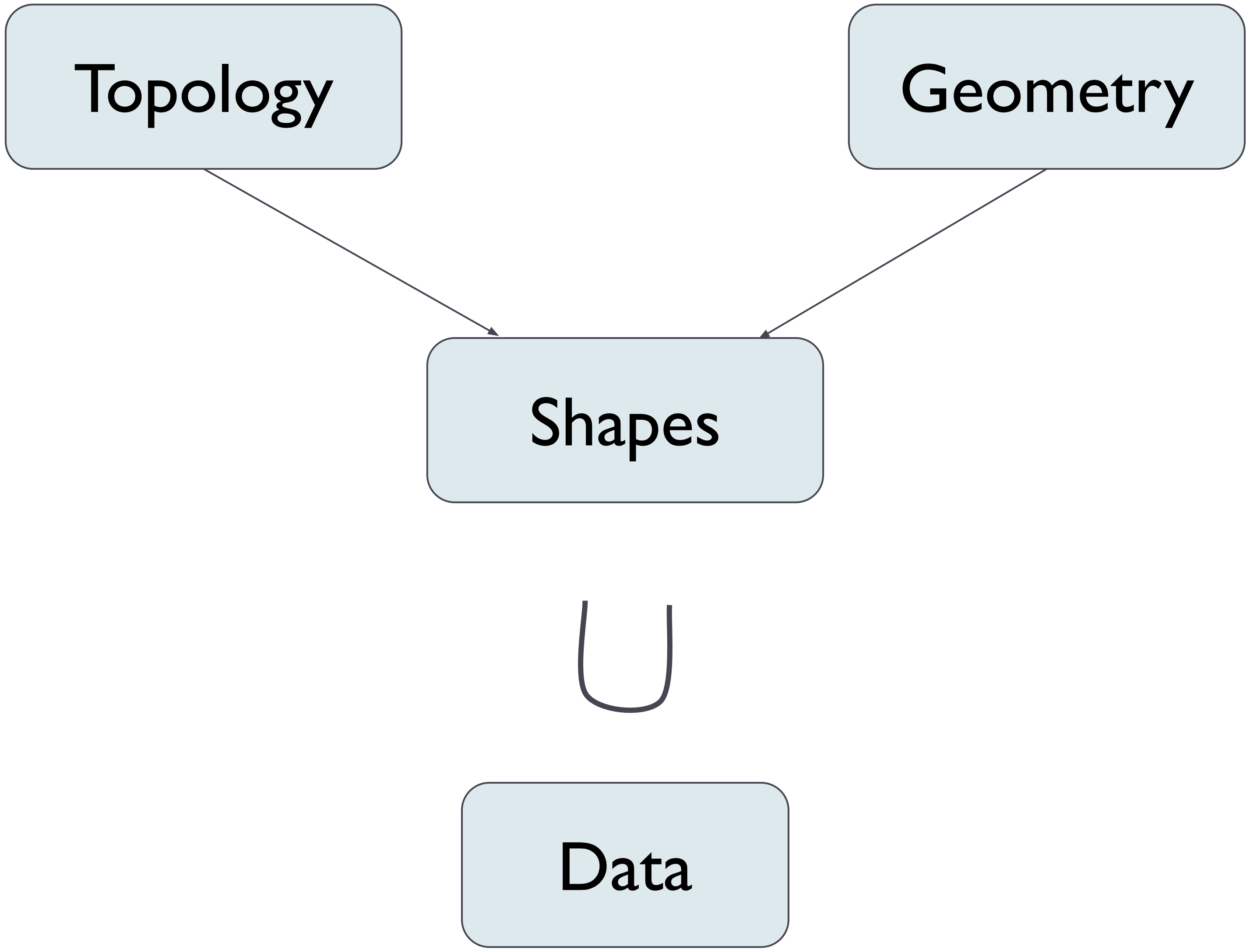
MATH412/COMPSCI434/MATH713
Fall 2025

Topological Data Analysis

Lecture 0: Introduction

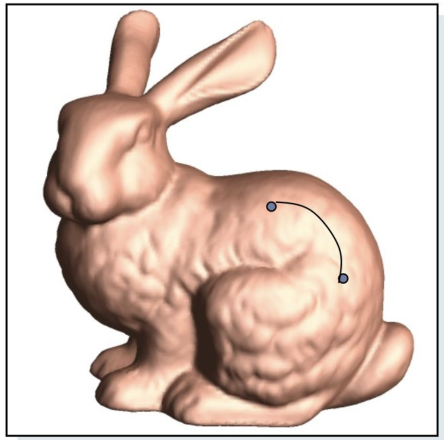
Instructor: Ling Zhou


Topology
Geometry
Shapes
∪
Data

Geometry = Geo (earth) + metry (measure)

- Distances and angles
  - area, volume, curvatures, etc
- Euclidean geometry (“flat” space)
- Riemannian geometry (“curved” space)
  - hyperbolic geometry
  - spherical geometry
- …

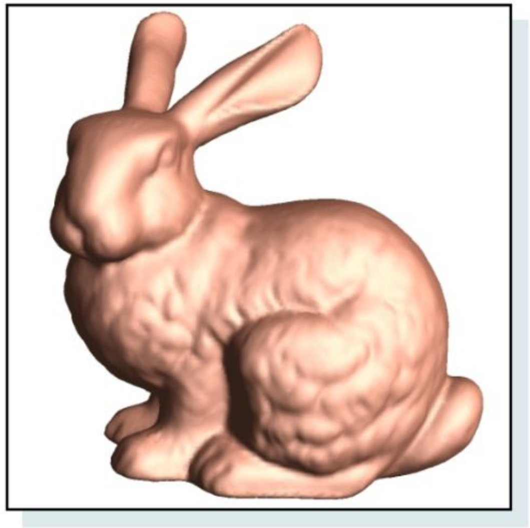
Using geometry, by unknown artist, 15th century



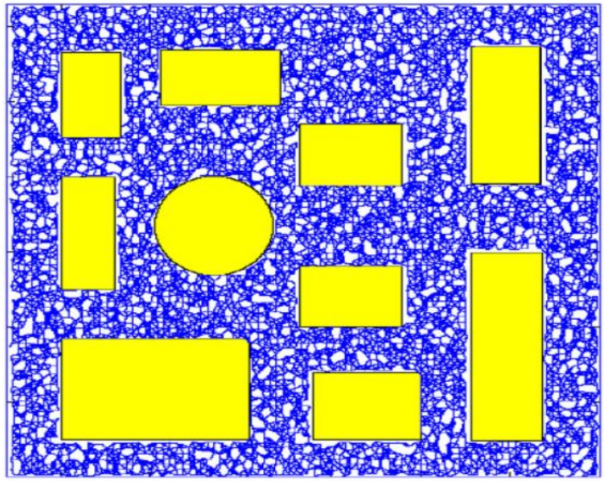
Topology

- Geometry can be too detailed:
  - Too local
  - Not always needed
  - Sometimes misleading
- Topology captures **coarse yet essential** info:
  - Connectivity
  - Holes, voids, etc

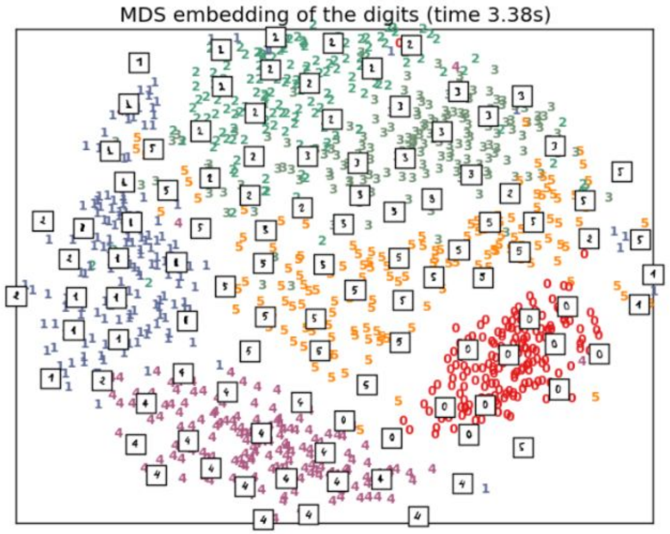
Data has shape!



3D surface



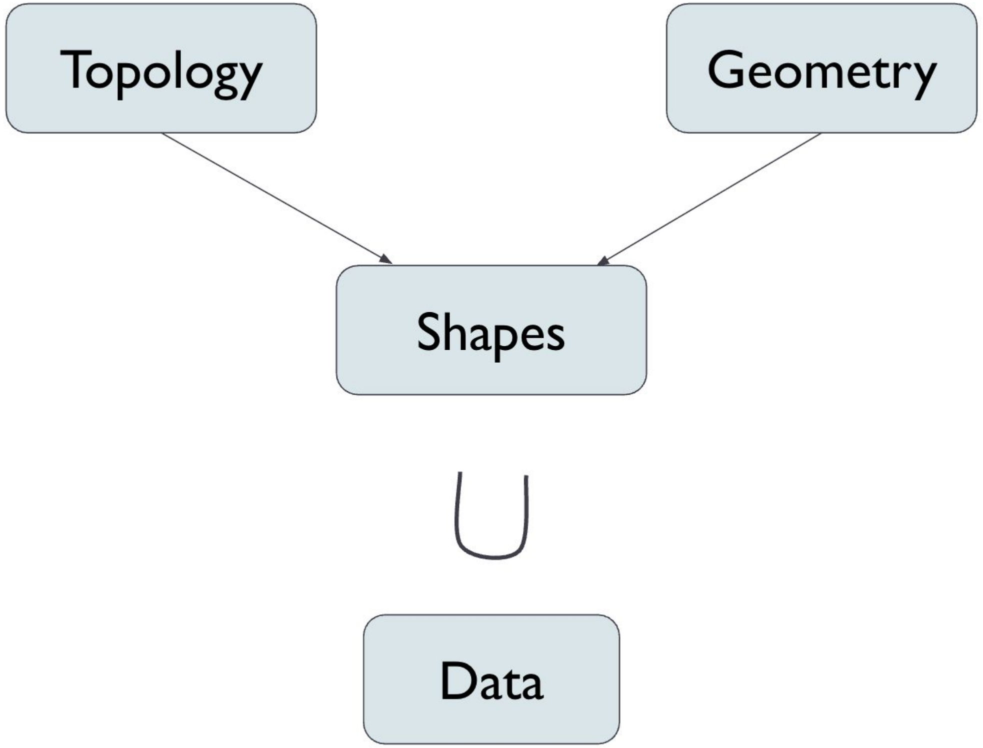
Holes in data



Clusters in data

TDA: Topological data analysis

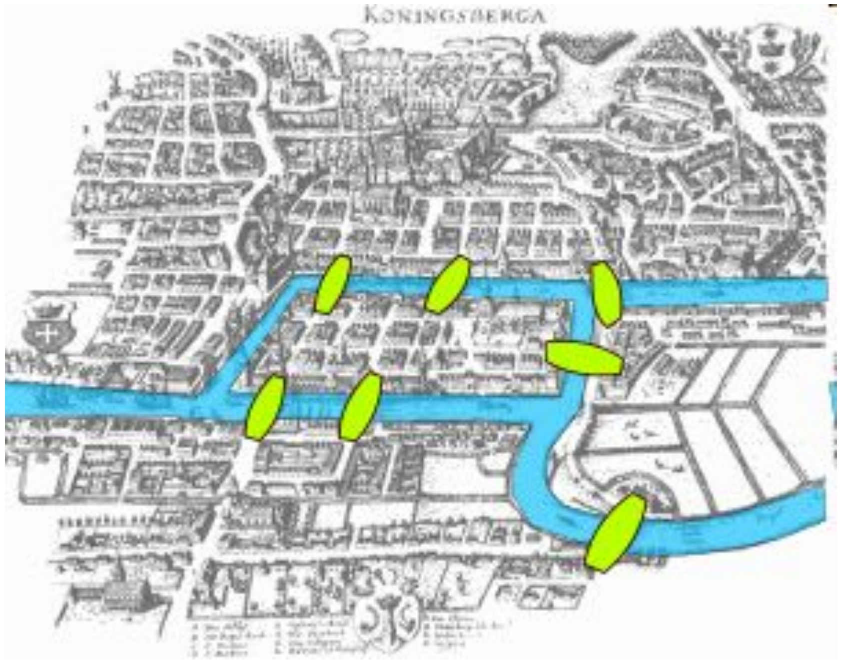
- Uses topology to capture global structure: connectivity, holes, clusters
- Built on geometry — distances and neighborhoods guide the computation
- Provides algorithms for extracting "shape" from data



Introduction to Topology

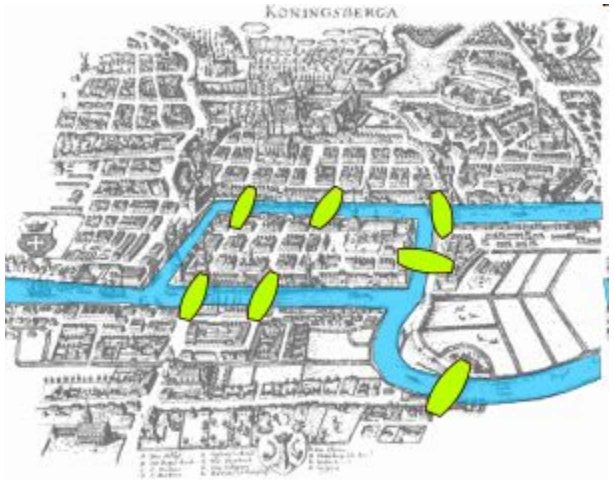
Introduction

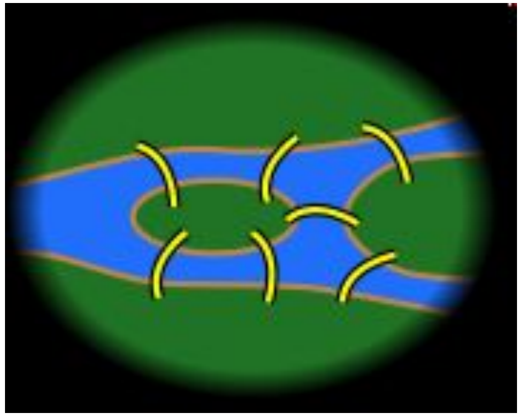
- What is topology
- Why should we be interested in it
- What to expect from this course

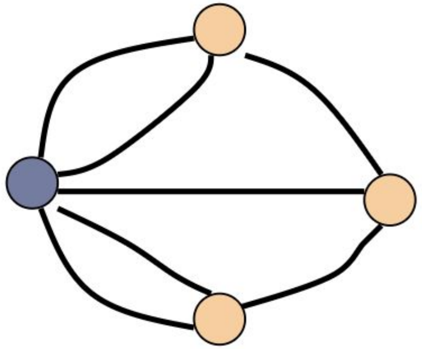


History

- Seven Bridges of Königsberg (1736)







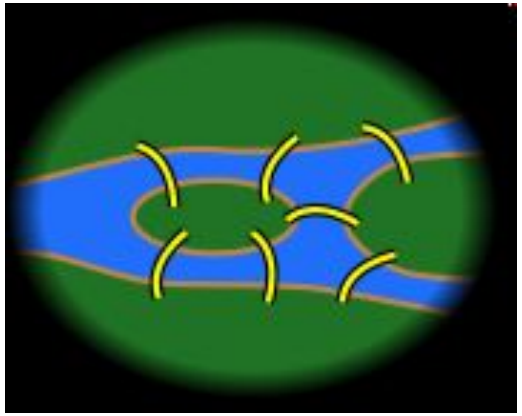
Euler cycle problem

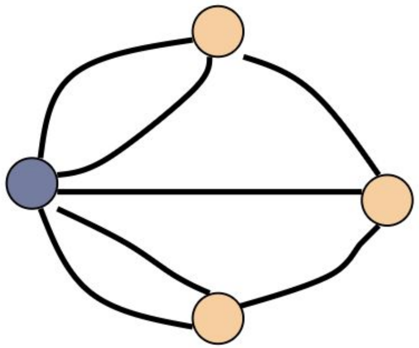
Abstraction of connectivity

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

Fundamental Questions

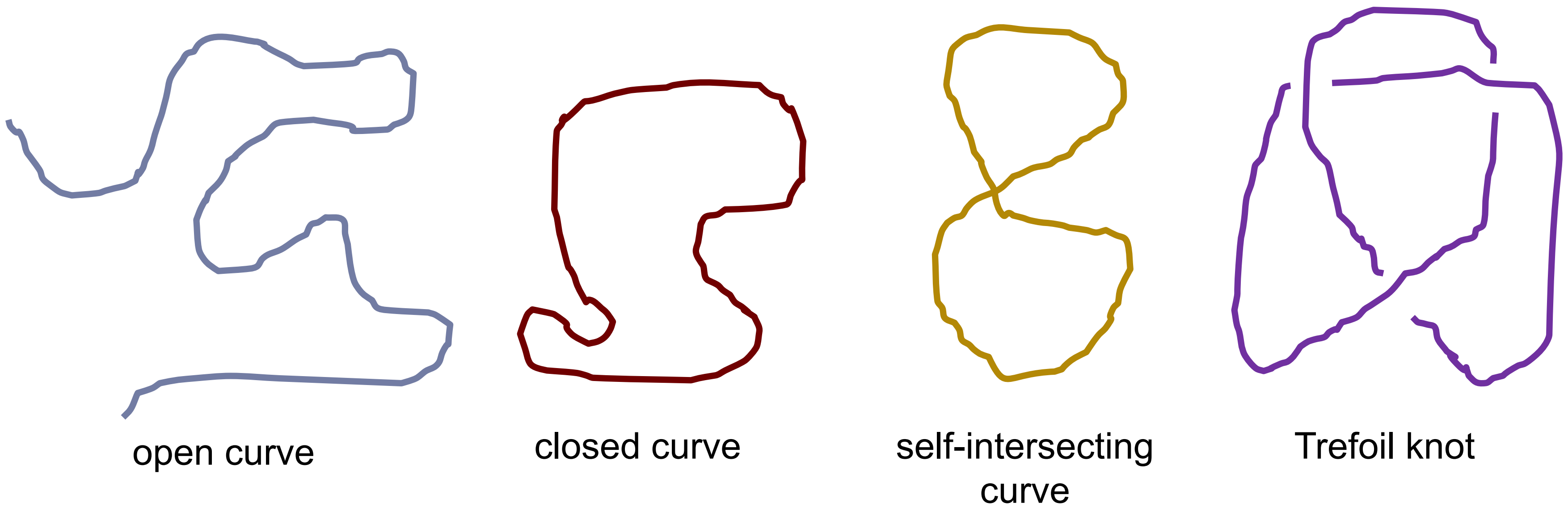
- How do we know when two spaces are the "same"?
- Why things are the "same" after deformation?





Homeomorphism

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



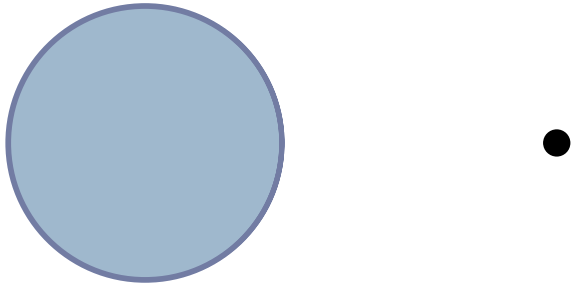
Examples of non-homeomorphic shapes

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

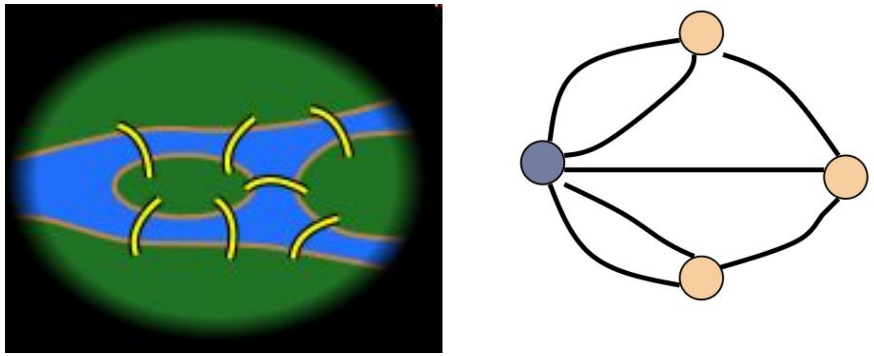
Homeomorphism

Example of homeomorphic shapes

Homotopy Equivalent - A Relaxation of Homeomorphism



- Not homeomorphic
- Homotopy equivalent



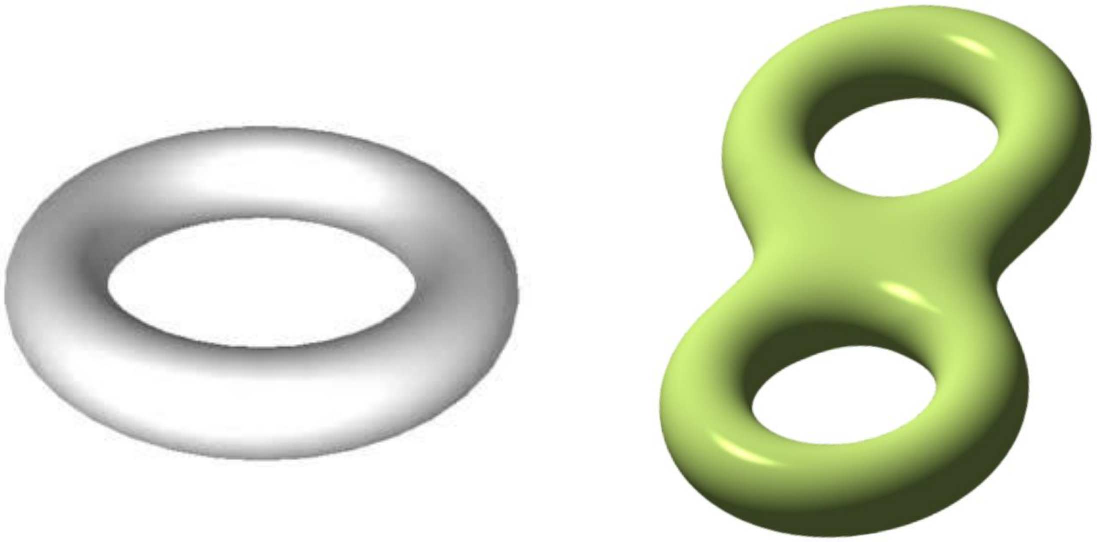
- Not homeomorphic
- Homotopy equivalent

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

- Homeomorphism and homotopy equivalence:
  - Too abstract to be verified in practice
  - Belong to Point Set Topology

- What do we use in practice?
  - Algebraic Topology

Homologous - A Relaxation of Homotopy Equivalence



Example of not homeomorphic, not homotopy equivalent & not homologous

Homologous - “same” number of holes in all dimensions

Topological Quantities

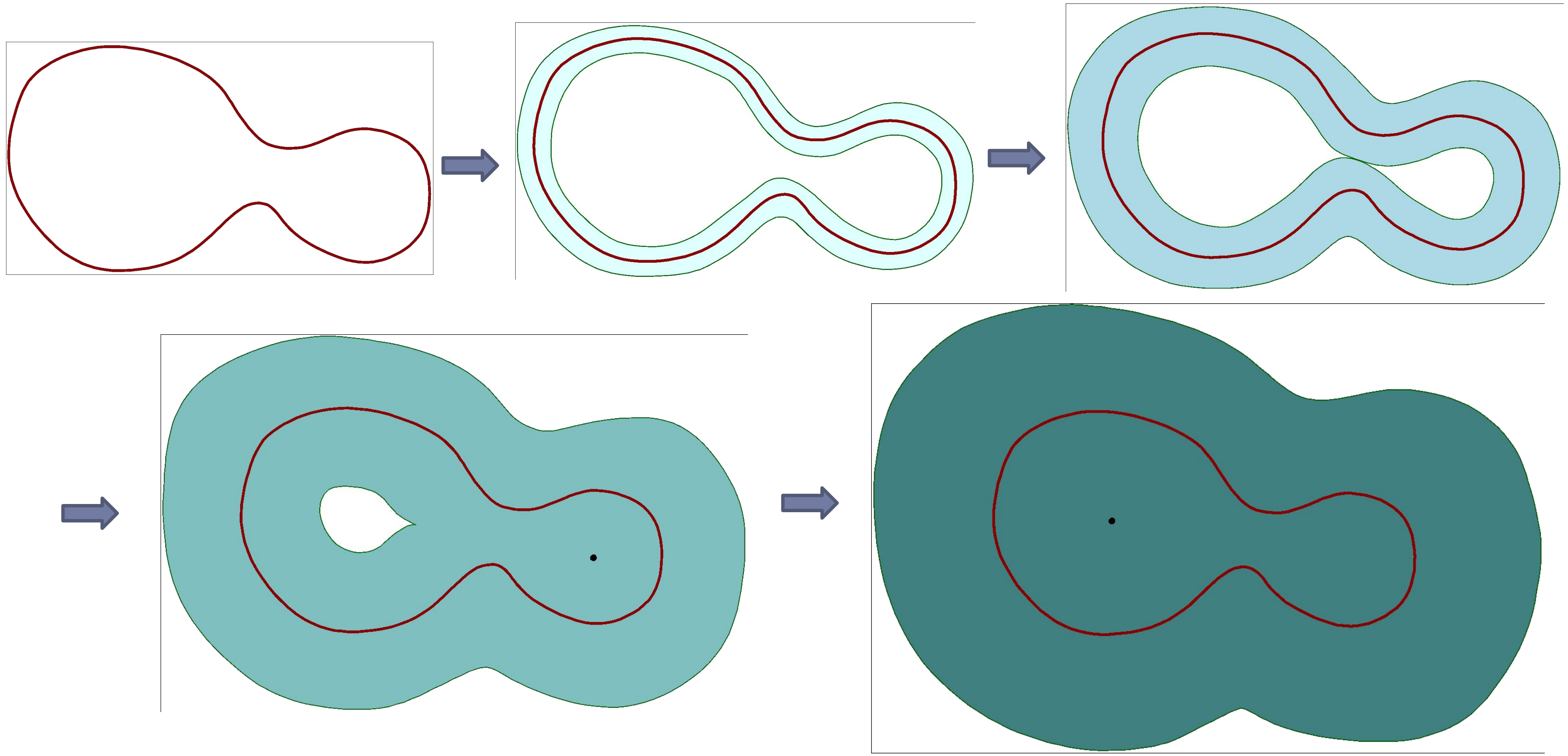
- Levels of "sameness"
  - Homeomorphism: exact topological equivalence (stretching without cutting)
  - Homotopy equivalence: coarser, allows collapsing parts
  - Homologous: even coarser, counts components, loops, voids
- What matters
  - Quantitative Invariants under these transformations = essential features
  - These are what topology captures and what make it useful for data

- Course goals
  - Learn definitions and build intuition for these invariants
  - Learn how to compute them
  - See applications in data analysis

TDA Principle

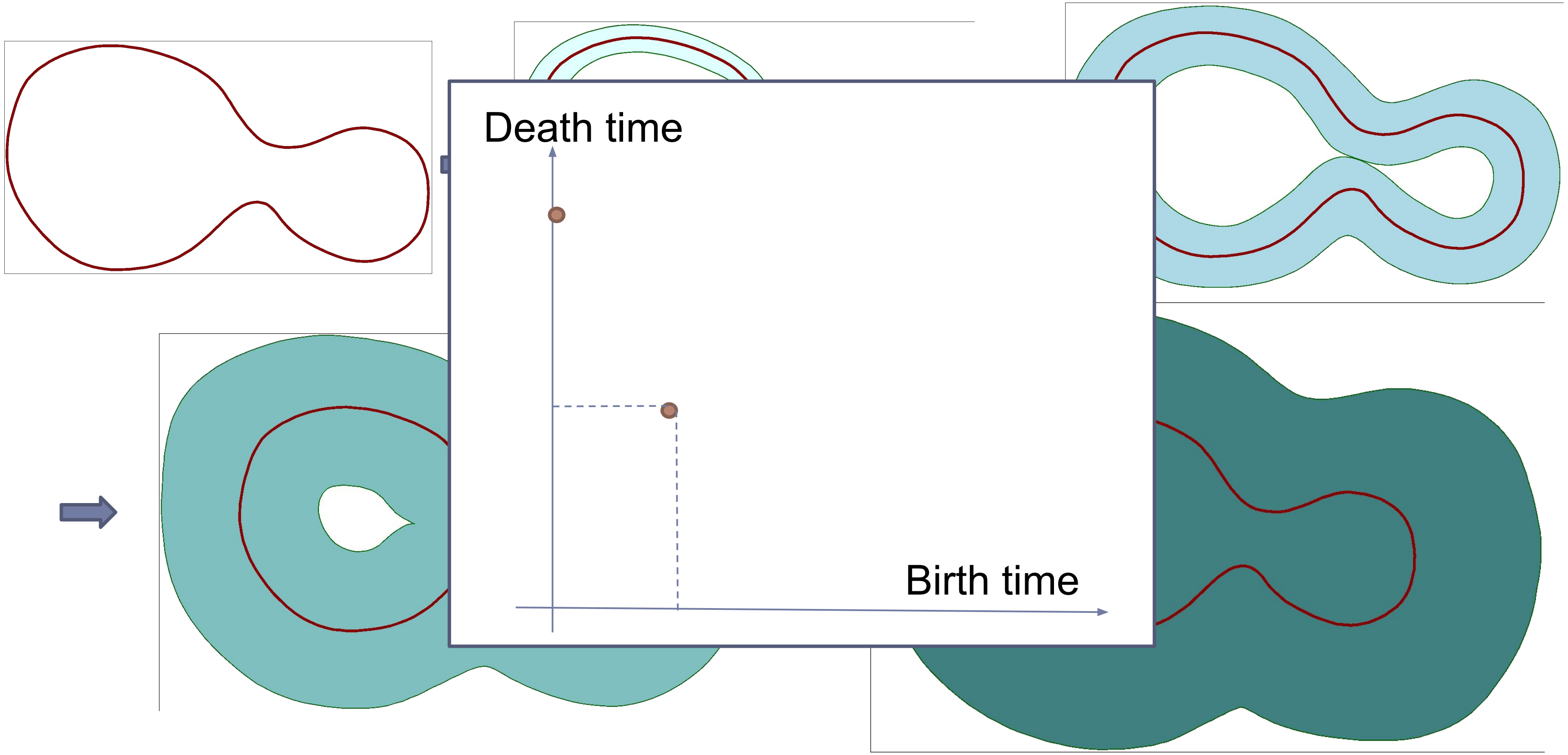
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



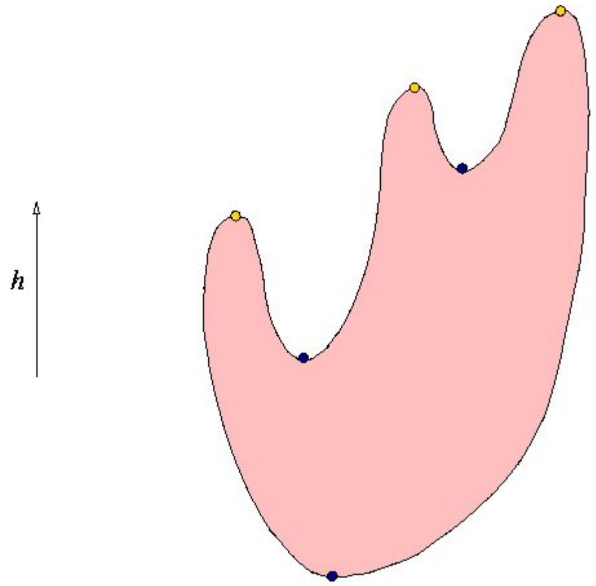
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



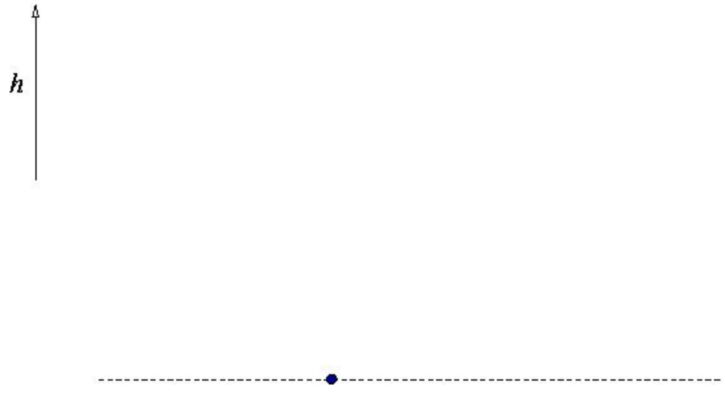
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



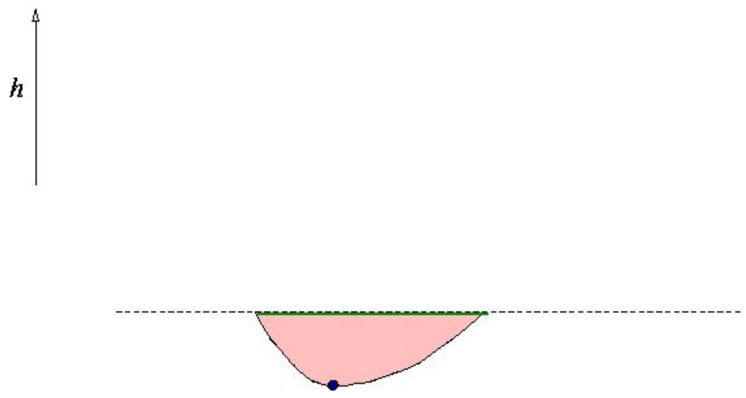
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



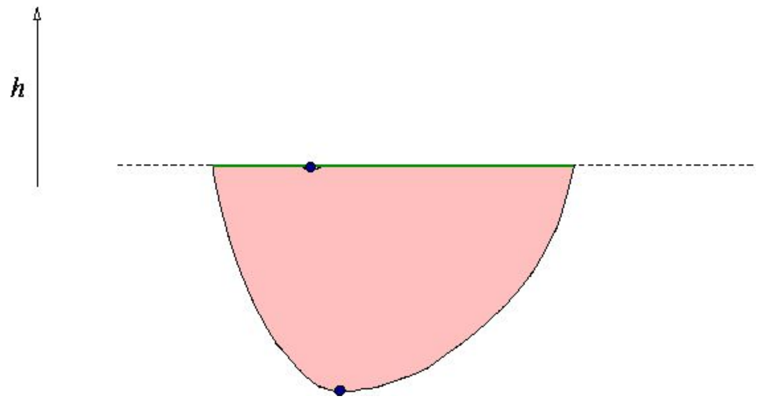
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - 'lifetime' of features



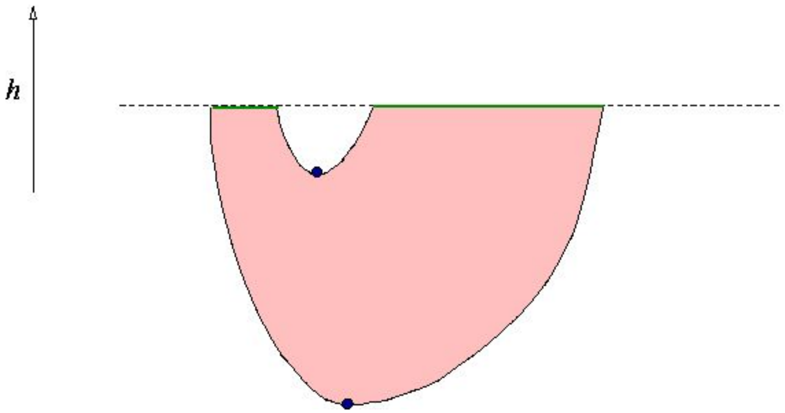
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



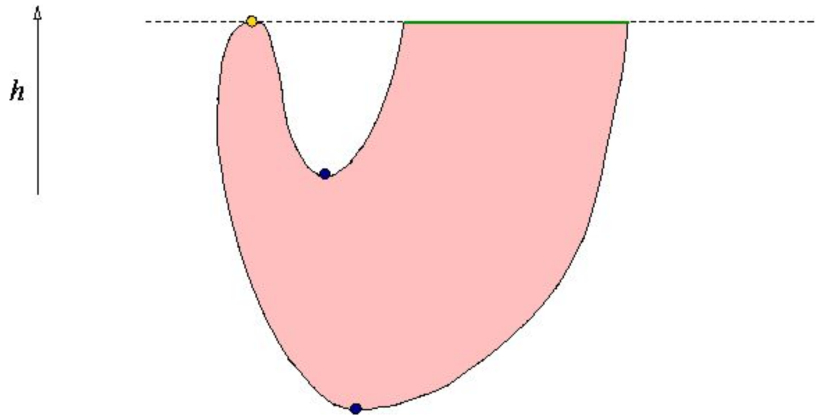
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



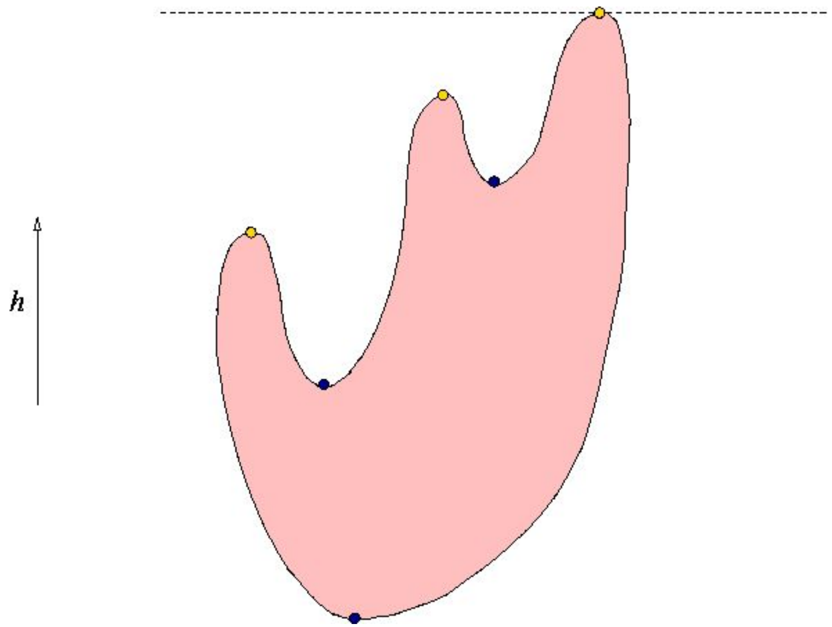
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



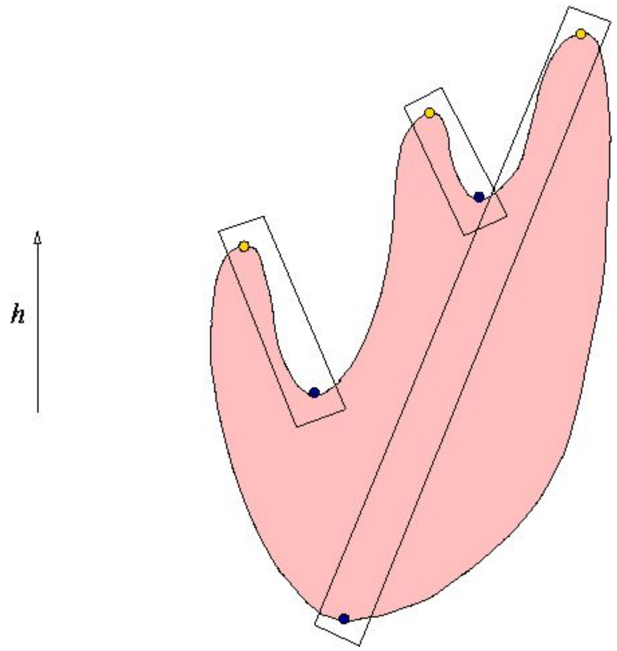
TDA Principle - Tracking Multiscale Topology

- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features



TDA Principle - Tracking Multiscale Topology

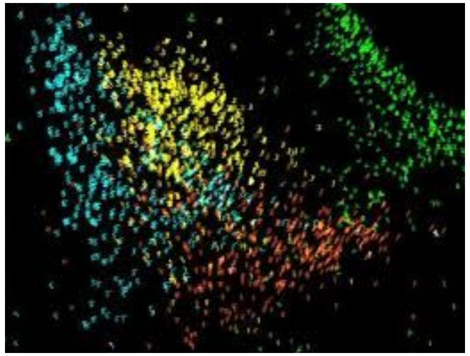
- Incorporate geometry, also functions or maps of a space to capture
  - existence of features
  - ‘lifetime’ of features

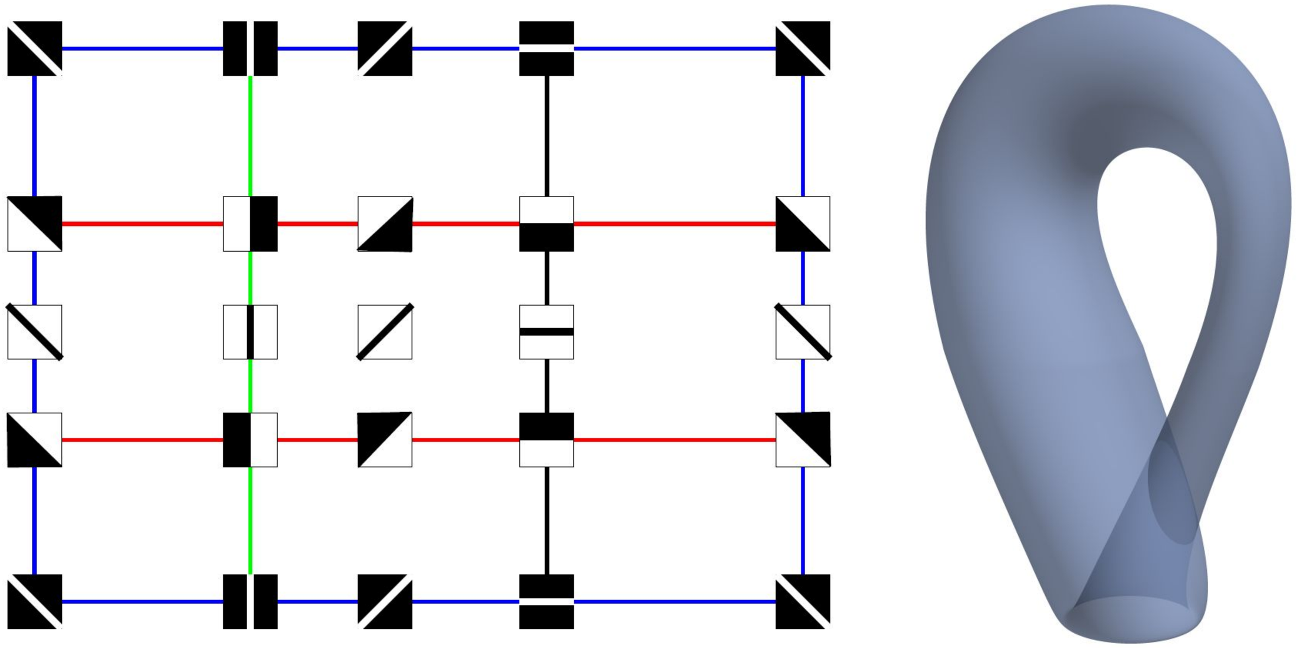


Examples of TDA Applications

Motivating Examples I

- Computer Vision
  - Clustering
  - Shape space

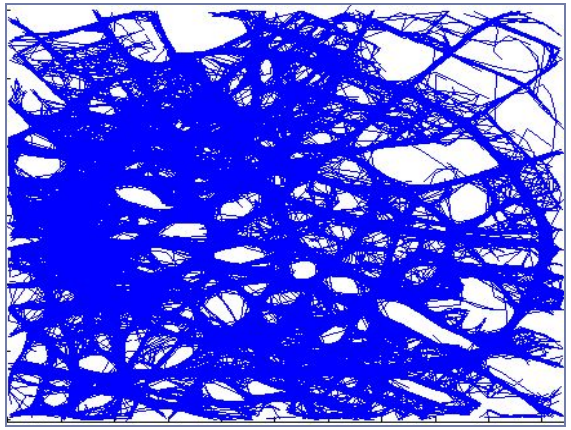




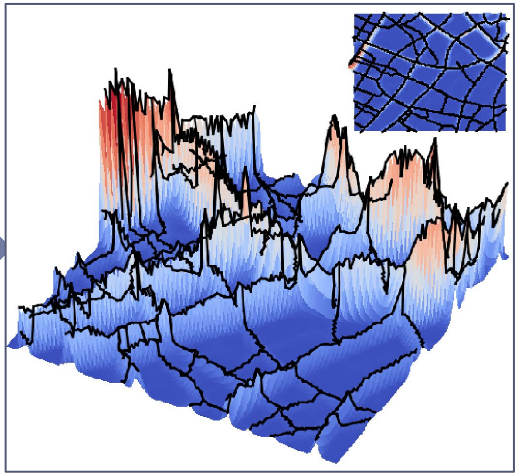
Courtesy of Carlsson et al, *On the local behavior of spaces of natural images*

Motivating Examples II

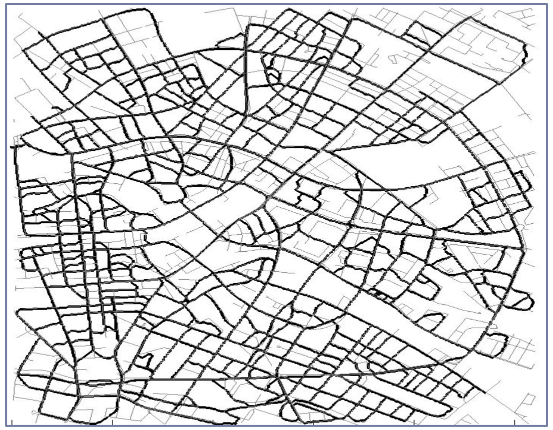
- Graph reconstruction
  - Road network reconstruction, neuron skeletonization, etc



Input: GPS trajectories



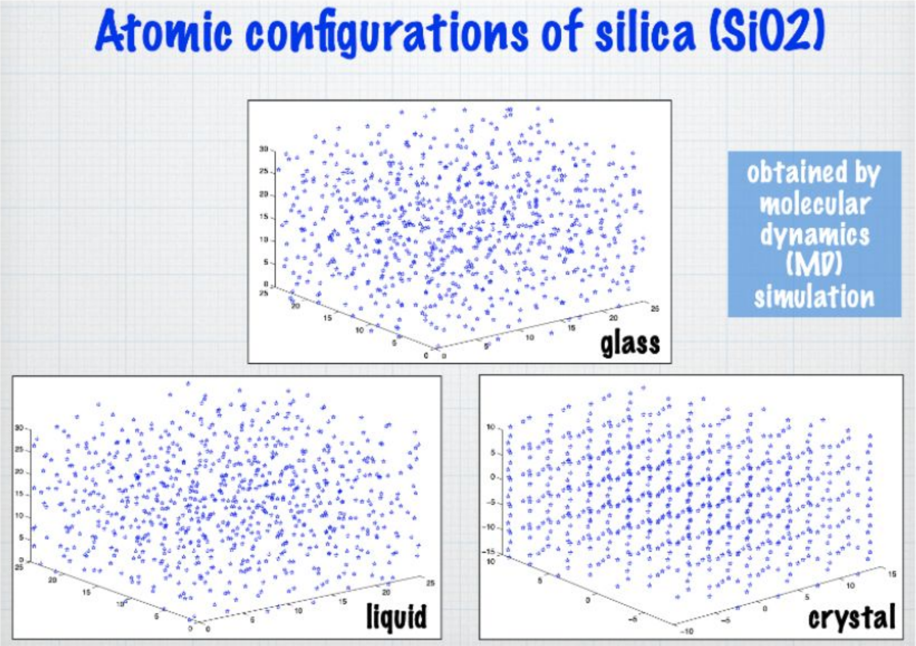
density field + discrete Morse

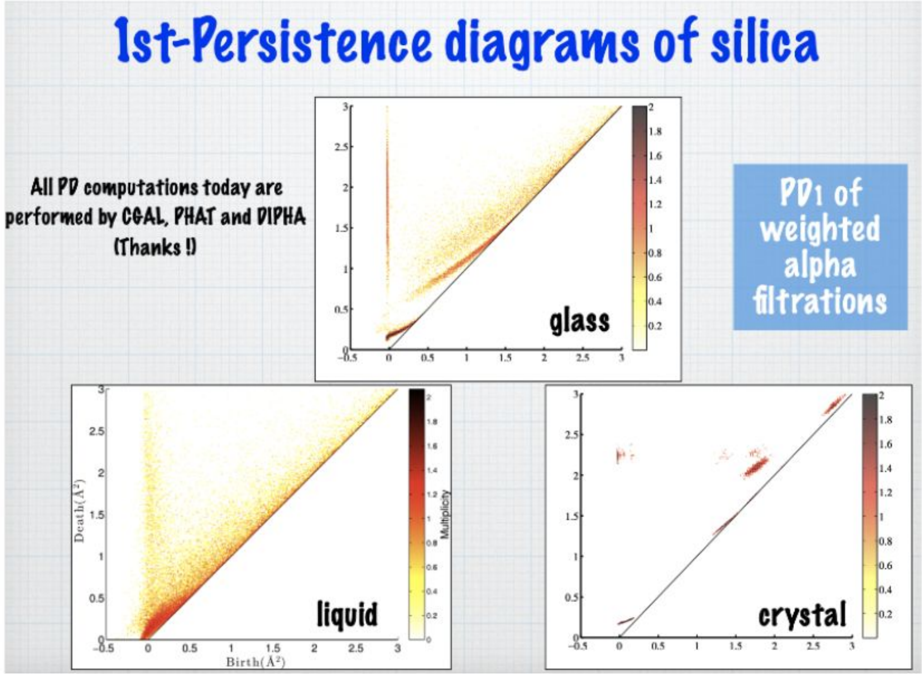


Output: Road network

Motivating Examples III

- Material Science

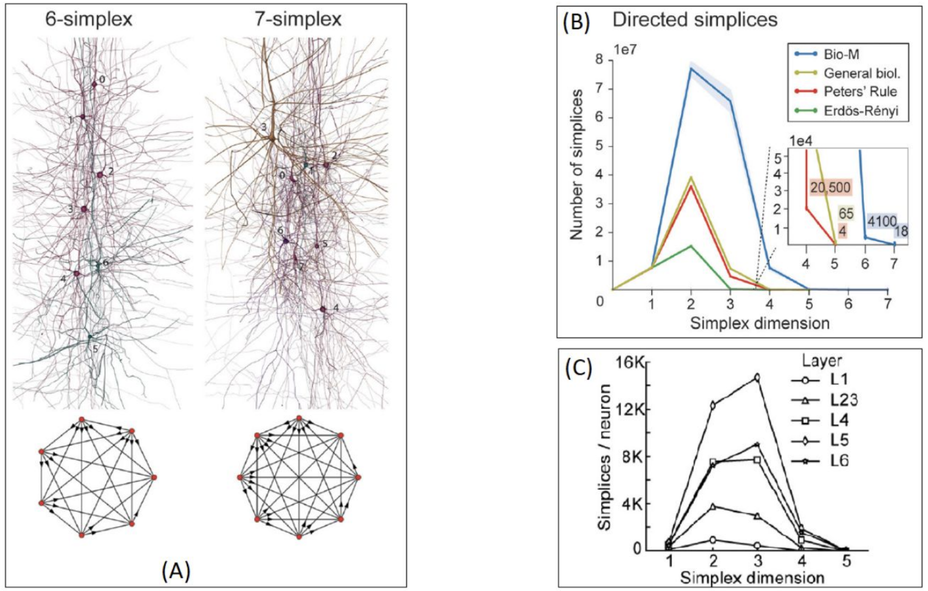




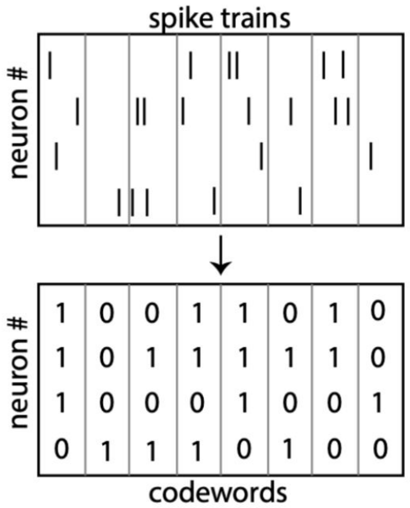
Nakamura et al (2015), *Persistent homology and many-body atomic structure for medium-range order in the glass*

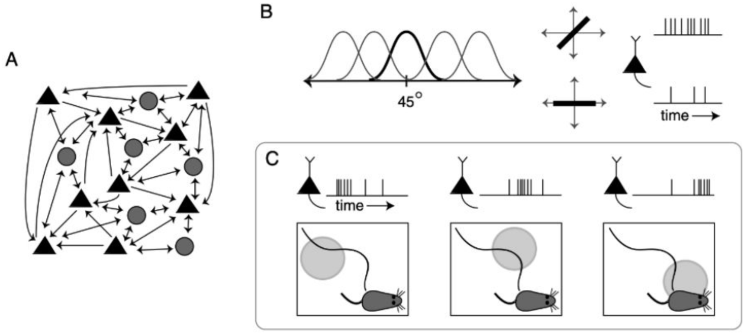
Motivating Examples IV

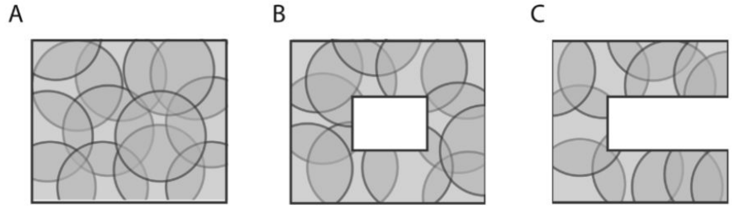
- Neural Science
  - Shape of neurons
  - Cofiring behavior of neurons



Michael W. Reimann et al. (2017), *Cliques of neurons bound into cavities provide a missing link between structure and function*



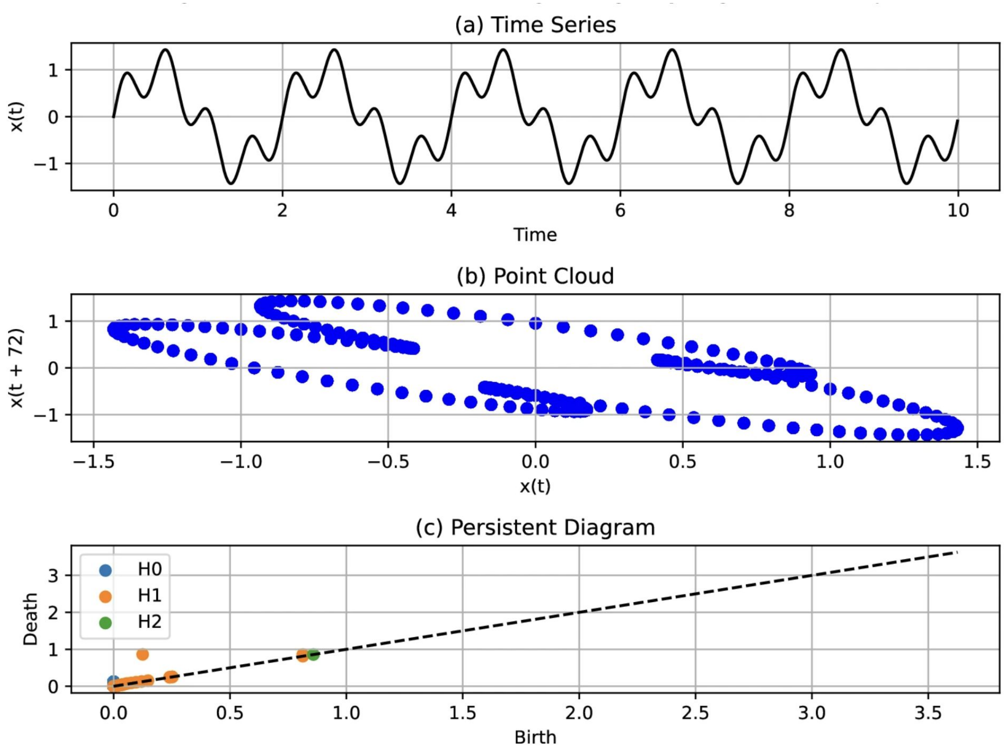


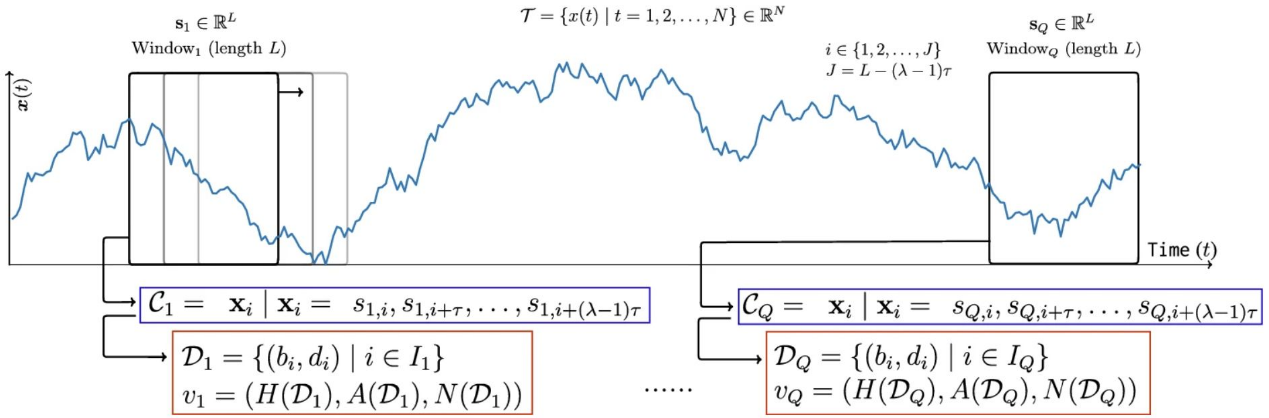


Carina Curto (2017), *What can topology tell us about the neural code?*

Motivating Examples V

- Time Series Data

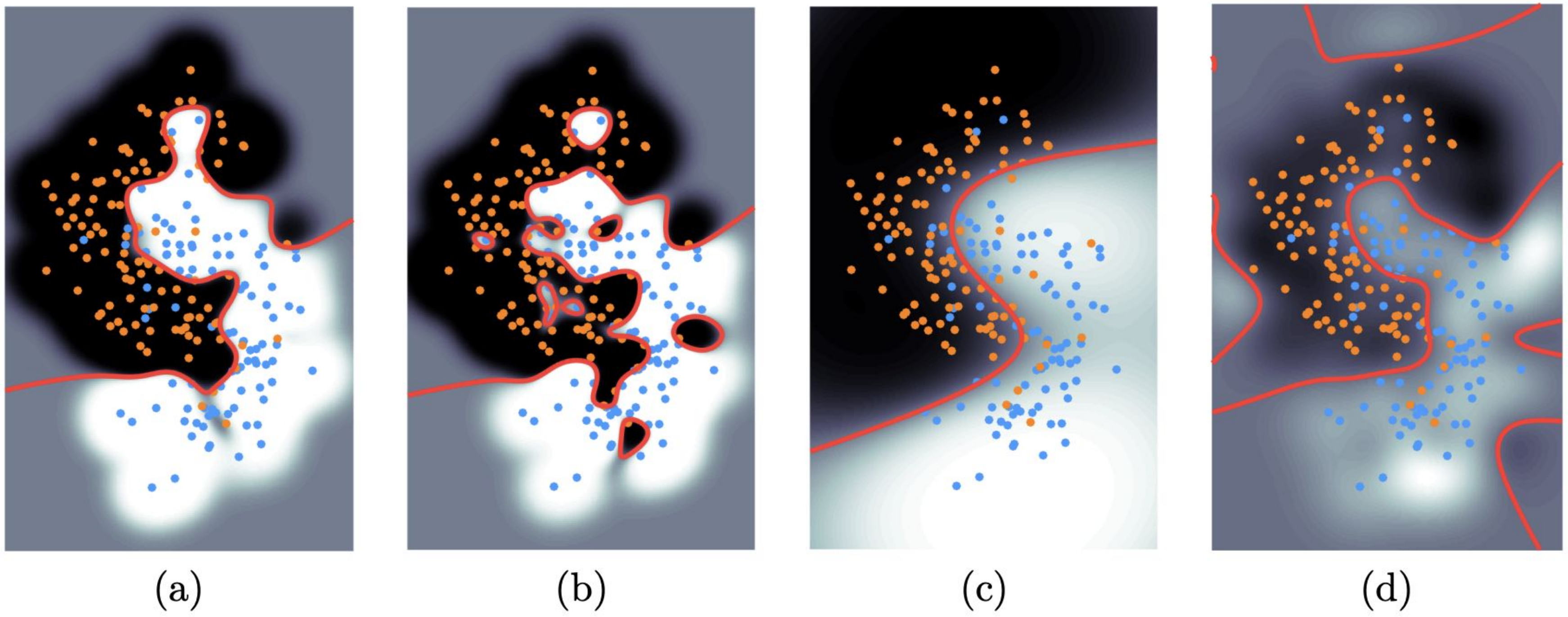




Luiz Carlos de Jesus Jr et al. (2025), *Enhancing financial time series forecasting through topological data analysis*

Motivating Examples VI

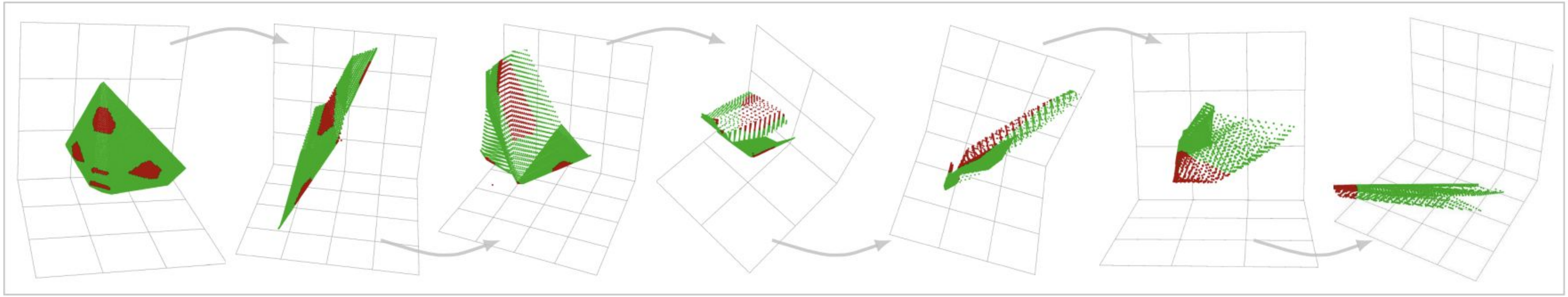
- Topological enhanced classifiers



Chao Chen et al. (2019), *A Topological Regularizer for Classifiers via Persistent Homology*

Motivating Examples VII

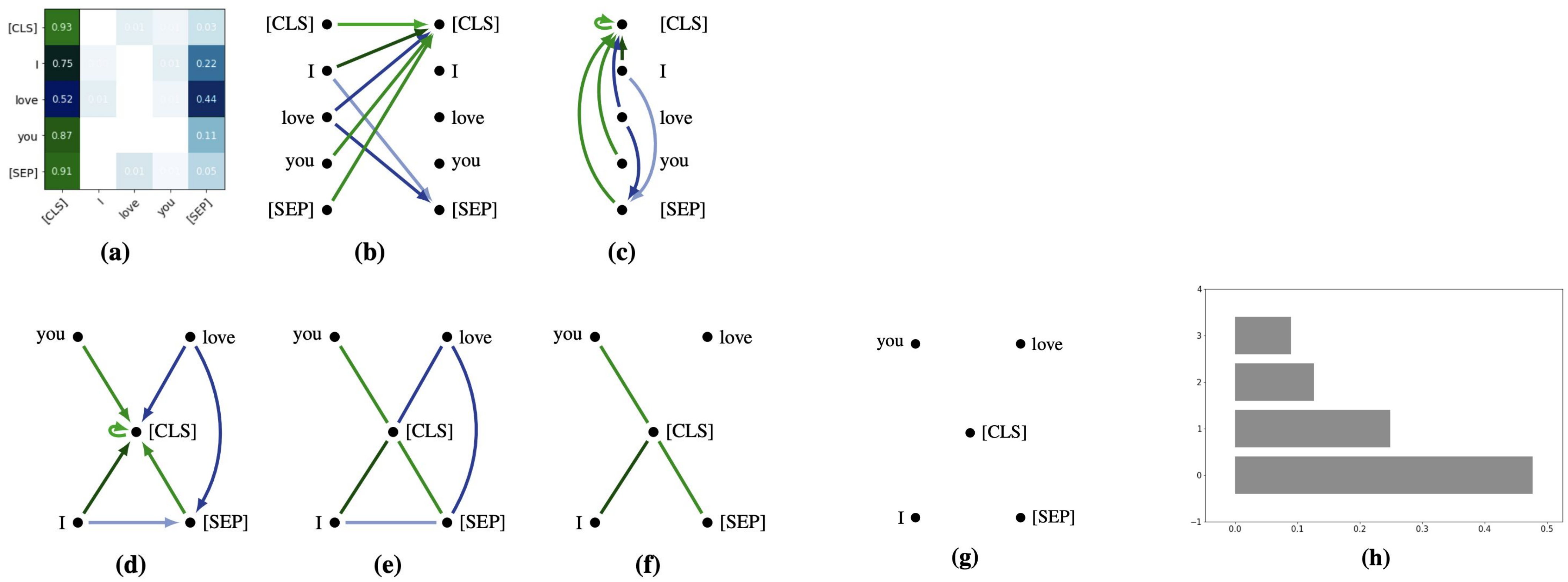
- Topology of neural networks



Gregory Naitzat et al. (2020), *Topology of Deep Neural Networks*

Motivating Examples VIII

- Large Language Models



Laida Kushnareva et al. (2021), *Artificial Text Detection via Examining the Topology of Attention Maps*

Motivating Examples IX

- Persistence-based feature vectorization + machine learning
  - Persistence images, kernels for persistence diagrams, etc

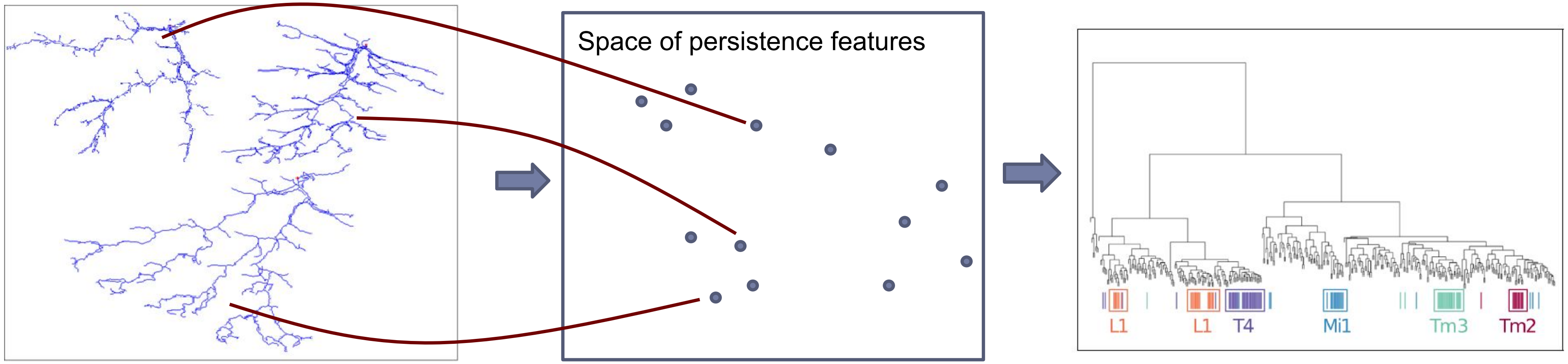
Input data:
A collection of point sets, graphs, 2D/3D images, arbitrary compacts, etc

→

Input data:
A collection of persistence-based feature representations (e.g, vectors)

→

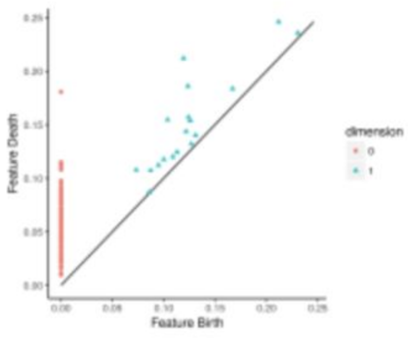
Further data analysis:
Feed to data analysis pipelines, say SVM, CNN, etc



A Nice DataBase for TDA - Applications

Zotero Groups: TDA-Applications

Title	Creator	Date
Persistence-based Hough Transform for Line Detec...	Ferner et al.	2025-04-18
A topological analysis of the space of recipes	Escolar et al.	2025-03-01
Advancing Precision Medicine: Algebraic Topology ...	Levenson et al.	2024-06
A novel multi-task machine learning classifier for rar...	Siva et al.	2024-05-09
Induction Motor Eccentricity Fault Detection and ...	Wang et al.	2024
Diverse 3D cellular patterns underlie the developme...	Mody et al.	2023-12-06
A functional data-driven approach to monitor and ...	Sansana et al.	2023-12-01
A novel approach for wafer defect pattern classifi...	Ko and Koo	2023-11-30
Pattern characterization using topological data an...	Chumley et al.	2023-09-01
Relational persistent homology for multispecies dat...	Stolz et al.	2023-08-11
Topological Singularity Detection at Multiple Scales	Rohrscheidt and Rieck	2023-07-03
Topological data analysis in medical imaging: curren...	Singh et al.	2023-04-01
A Primer on Topological Data Analysis to Support I...	Hoef et al.	2023-02-24
A data-driven workflow for evaporation performa...	Zhang et al.	2023-02-01
Skyler	Bokor Bleile	2023-01-17



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.

Summary

- In general, topology
  - Coarser yet essential information
  - Characterization, feature identification
  - General, powerful tools for both space and functions defined on a space
  - Elegant mathematical understanding available

- Indeed, topological ideas / methods have been used in many applications fields, including:
  - Graphics, visualization, medical image processing, computer vision, computational neuron science, computational biology, material science
  - TDA + machine learning

- However
  - Difficult mathematical language
  - Previously, lacks computational methods
  - Also, we need modern perspective / extensions to make them more suitable for data analysis

About the Course

This Course

- Introduce foundation of topological data analysis
- Goal:
  - Understand basic language in computational topology
  - Appreciate the power of topological methods
  - Potentially apply topological methods to your research
- Course materials:
  - Canvas Course page
  - Textbook: *Computational Topology for Data Analysis*, by Tamal Krishna Dey and Yusu Wang. Available for free online.
  - Instructor Contact: *ling.zhou@duke.edu*

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
- Ad

Will focus on concepts, definitions, algorithms, also intuition why they work, and how they can be used.

Grading policy

- Grading
  - Homework: 18%
  - Course project / survey: 82%
    - Proposal 10%
    - Report 32%
    - Presentation 40%

Important schedule

- Week 3-5:
  - Meet with me to discuss background and potentially design projects.
- Week 6-8:
  - Finalize project topics, with a proposal.
- Week 9-14:
  - Finish final report and prepare for the presentation
- Week 15:
  - Final presentation week: in-class + Q&A