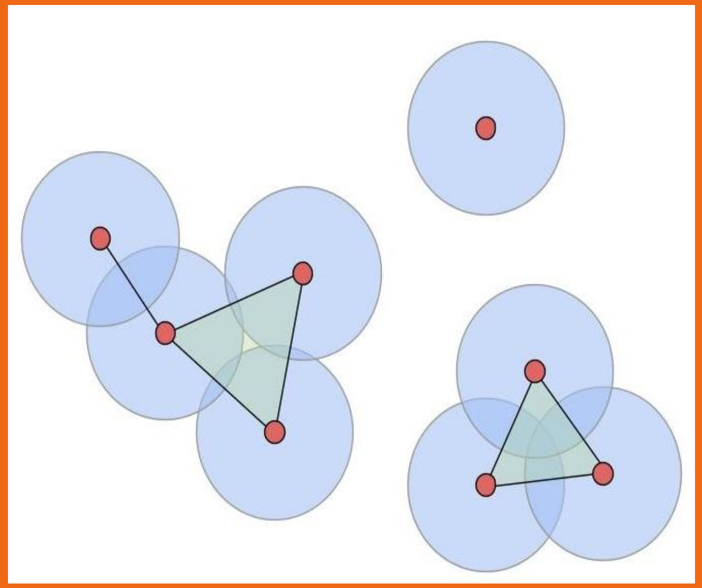
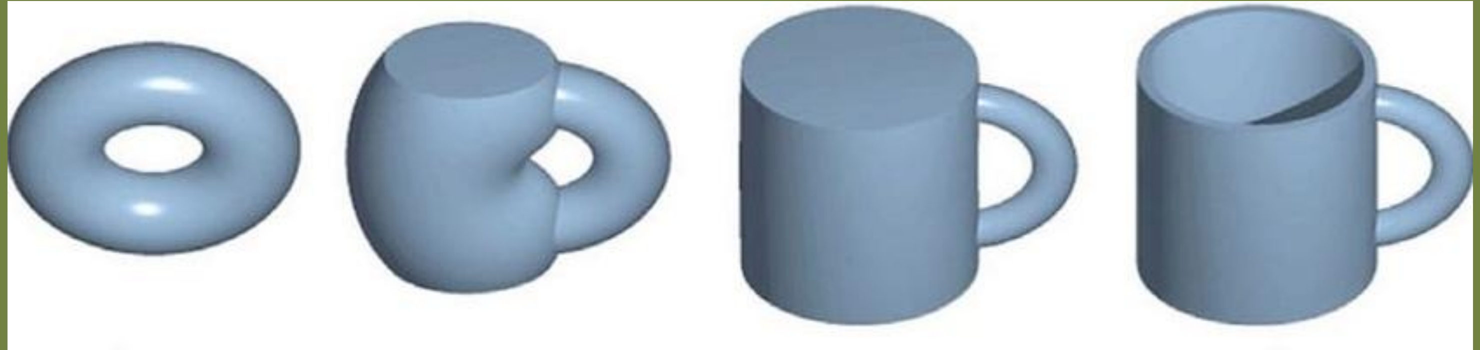
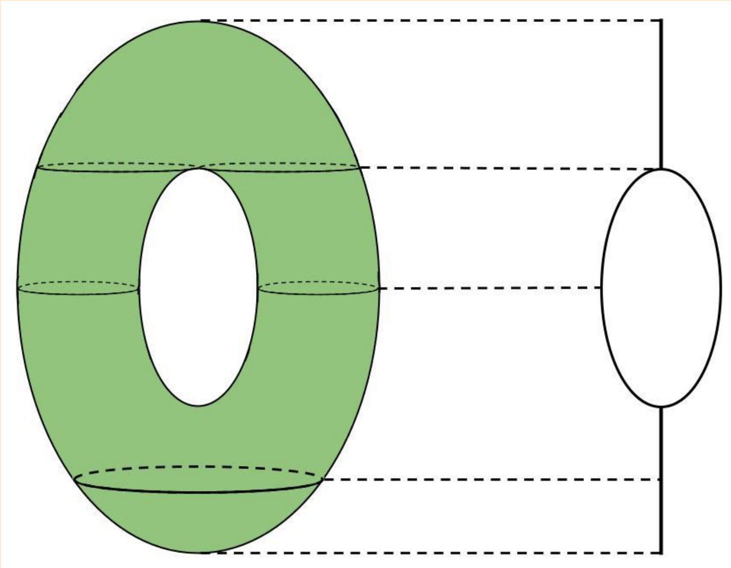
Topological Data Analysis

An *Approachable* Introduction

Emma Schmidt







Overview

Topology

Homology & Homotopy

Complexes

Filtration & Persistence

Why Care About TDA?

And a brief history of the field

History:

- Built on algebraic topology and computational geometry
- Big paper: "Topological Persistence and Simplification," 2002, Edelsbrunner, Letscher, and Zomorodian
- Introduced filtration, persistence, & algorithms

Today:

- Datasets are becoming more pervasive, more complex, and larger in size
- TDA: great for visualization and capturing high-dimensional behavior
- Lots of applications

01 Topology

TDA In a word:

Looks at patterns in data and see what other tools can't see, encodes information into lower-dimensional visualizations

What is Topology?

- The way shapes move and deform
- "Rubber sheet geometry"
- Continuous deformation -> preserves topological behavior

Topological Spaces

- Object of study -> topological space
- Metric space: use some metric to relate points to each other
- Typical metric of choice: distance

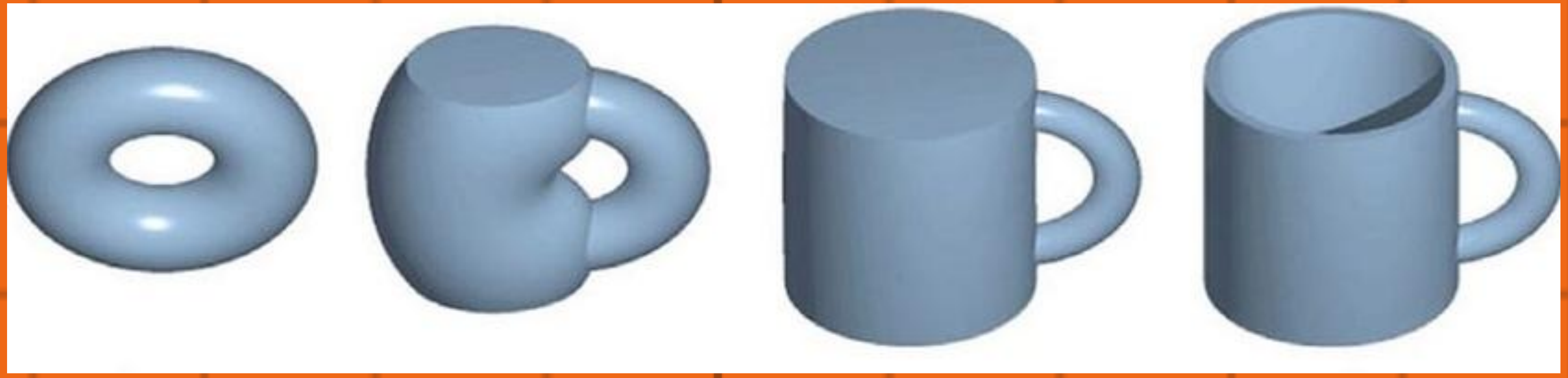
In a word:
Look at number of holes in an object, group shapes by comparing hole structure

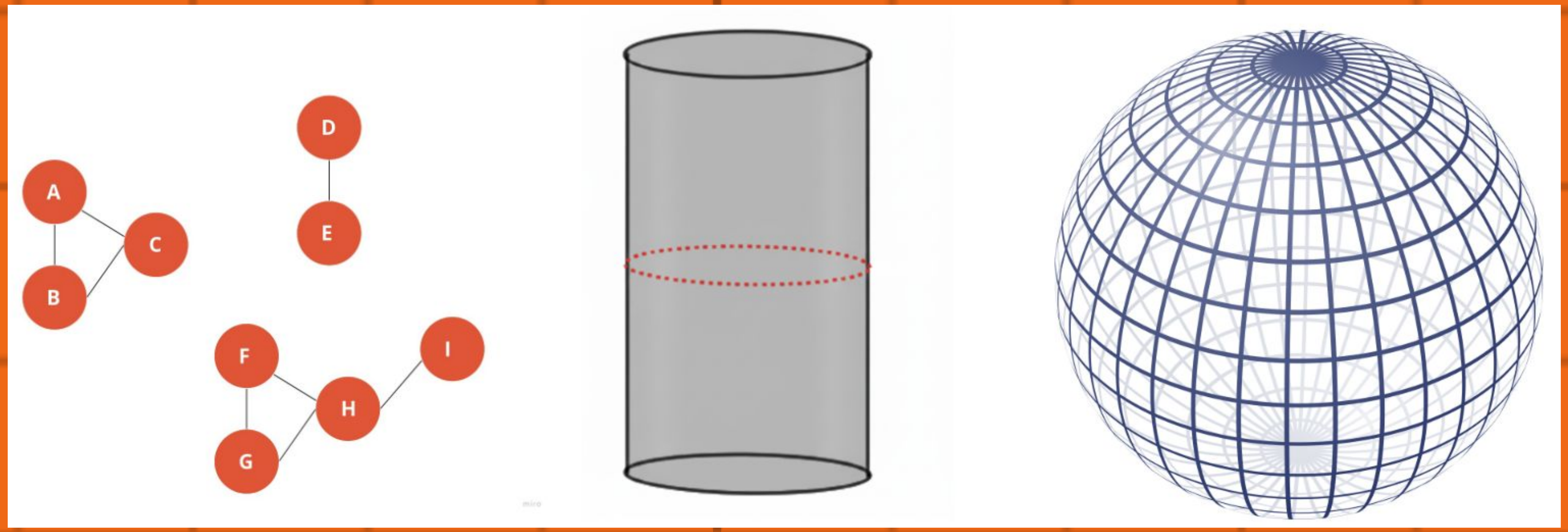
Homology

02

Dimensions of homology:

- **0-dim**: connected components
- **1-dim**: loops
- **2-dim**: voids
- **3-dim+**: hard to visualize





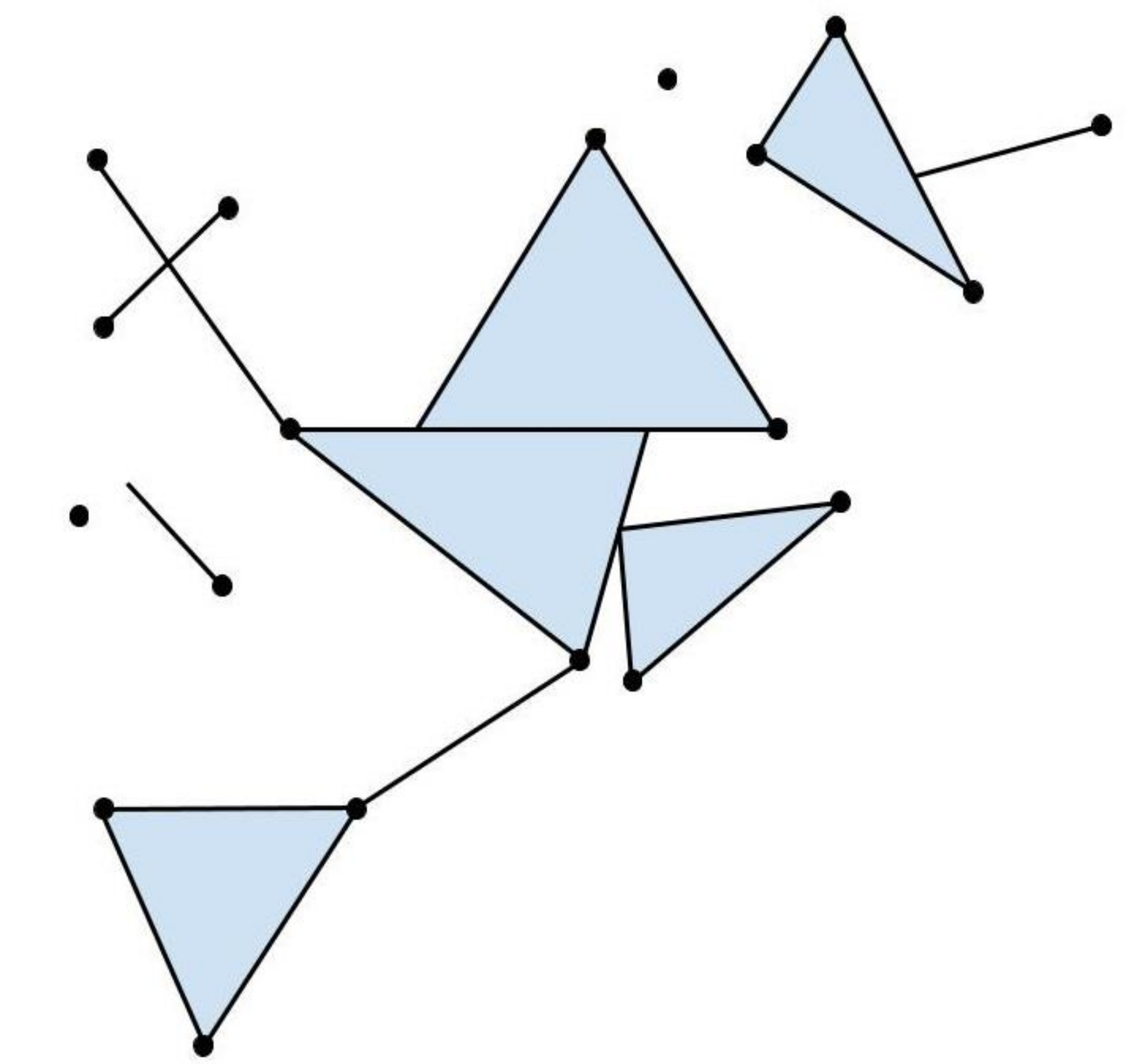
Homotopy:

- Studies the *way* things deform, tracks entire path
- Requires smooth step-by-step deformation
- Harder to compute, less used

03 Simplicial Complexes

In a word:

Graph containing vertices (points), edges (lines), and faces (solid triangles)



Simplex: generalization of shape with respect to dimension

Rules for building a legal simplicial complex:

- Only use simplices
- Connect things together by (full) simplices
- Every simplex must be connected to all of its lower-dimensional simplices

Cĕch & Vietoris-Rips Complexes

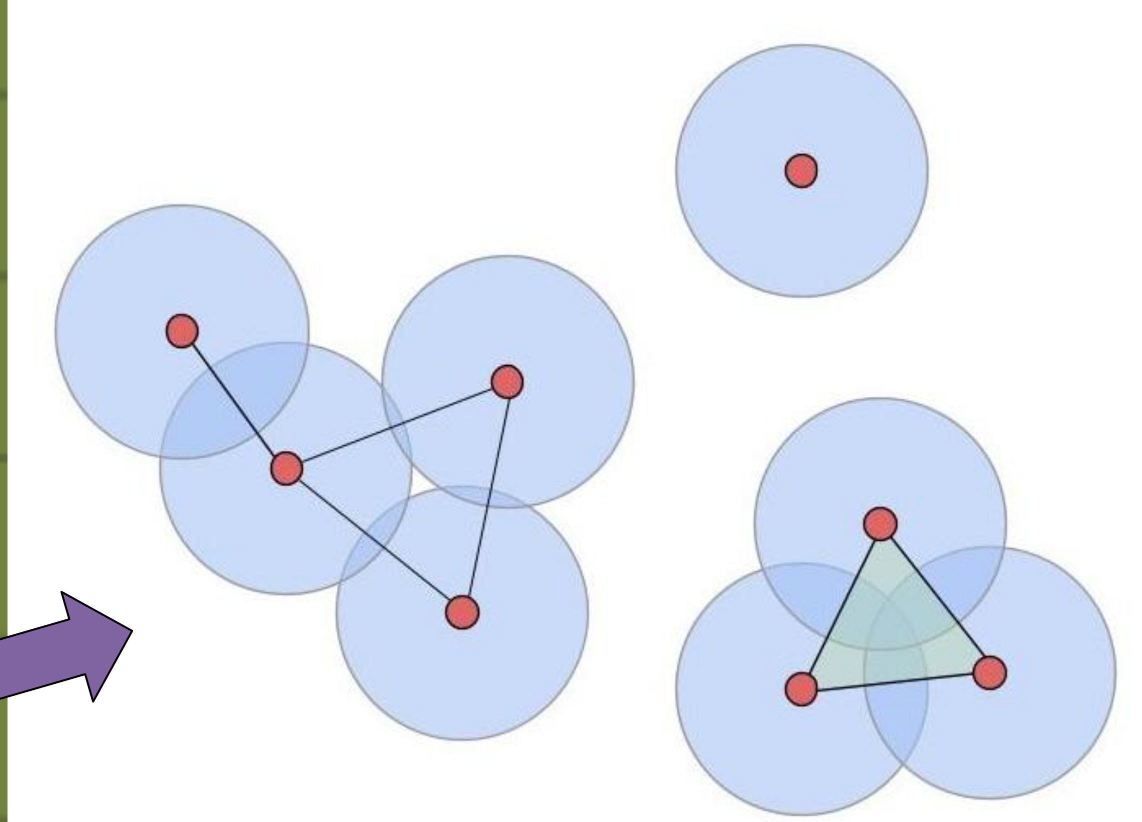
Types of simplicial complexes

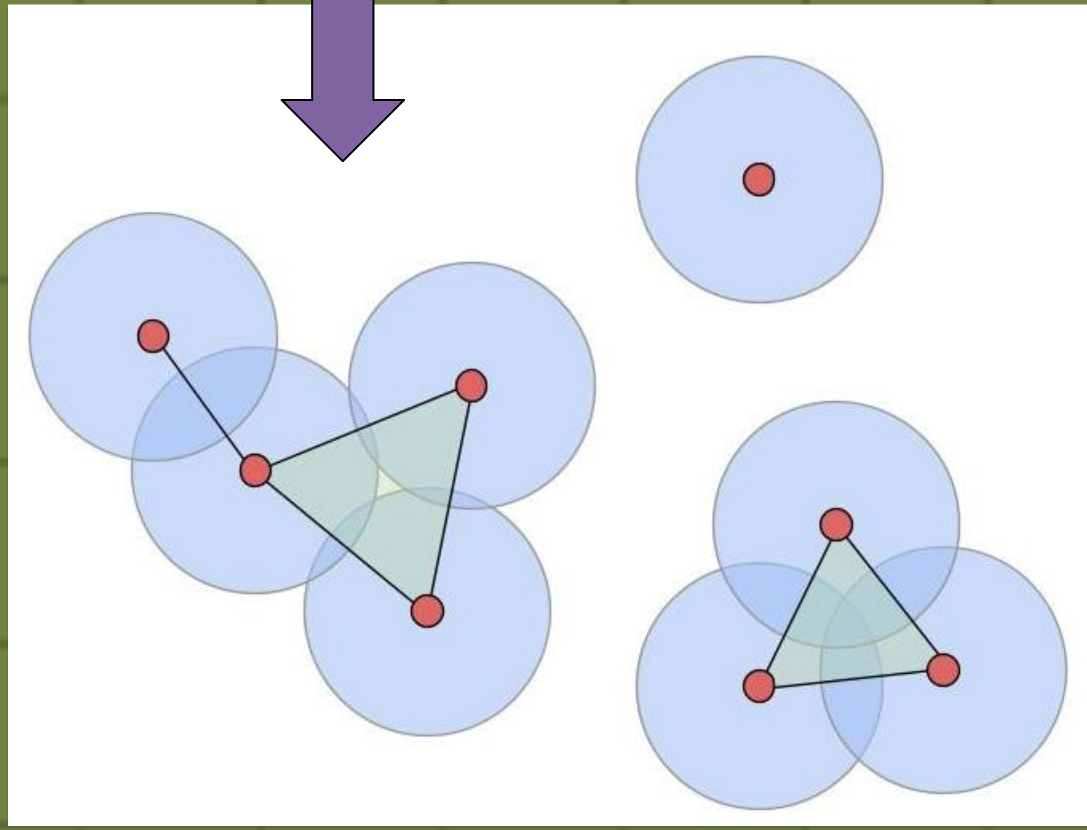
Vietoris-Rips Complex

- Start with points, draw small circles around each
- If two circles intersect, draw edge between centers
- If three circles intersect pairwise, insert shaded triangle

Cĕch Complex

- Start with same process as Vietoris-Rips complex
- For faces: only draw face if all three circles intersect at some common point (pairwise intersection not enough)





In a word:

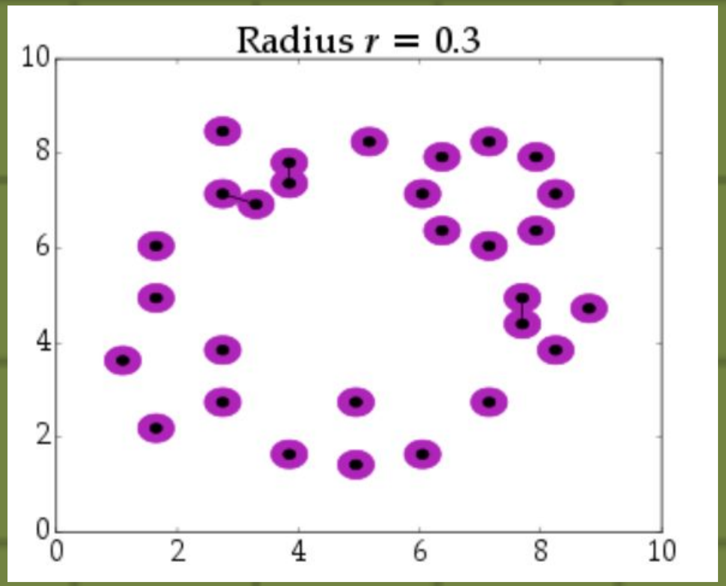
Look at an object step-by-step as some parameter value changes; creates nested sequence

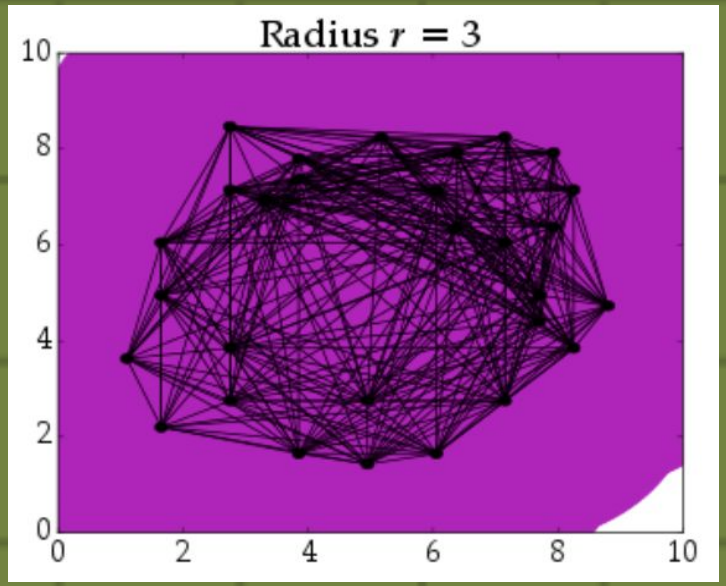
Filtrations

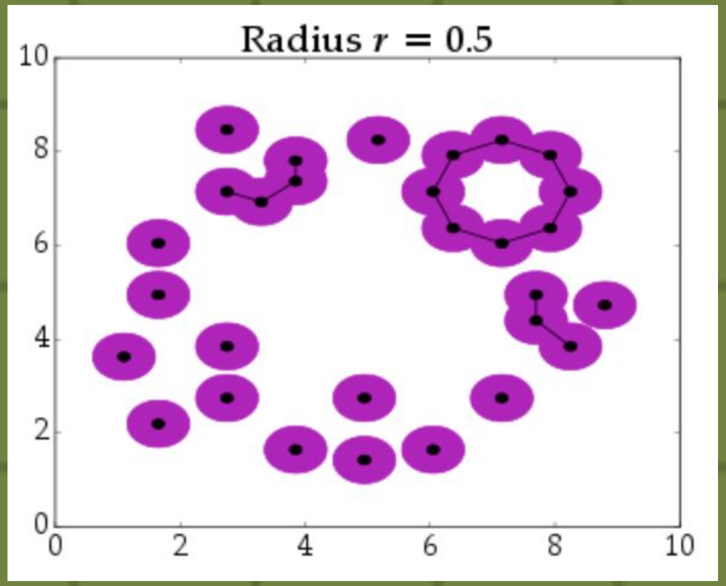
04

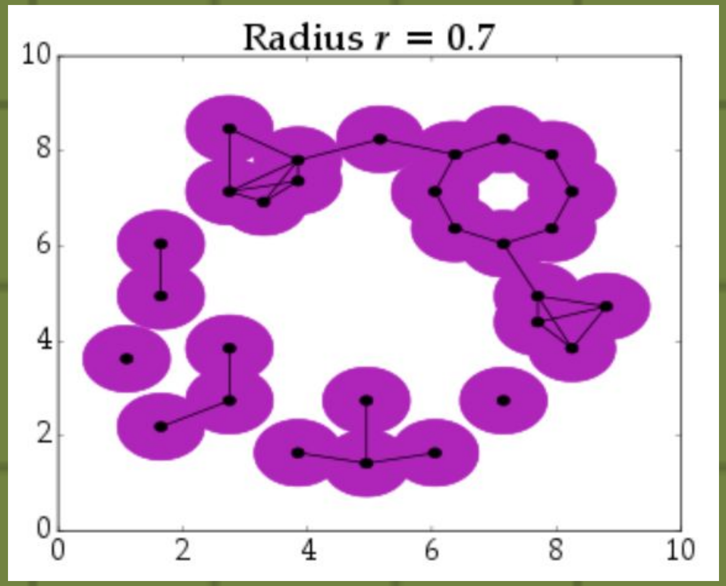
Simplicial Complex Filtration:

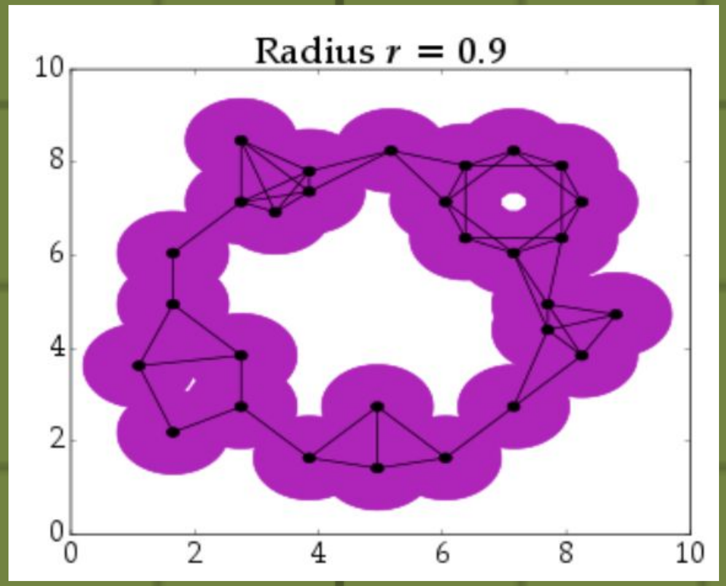
- Look at what *persists* across parameter changes
- For simplicial complex: how does the complex change as radius changes?
- Birth: a feature appears or is created
- Death: a feature disappears or is destroyed

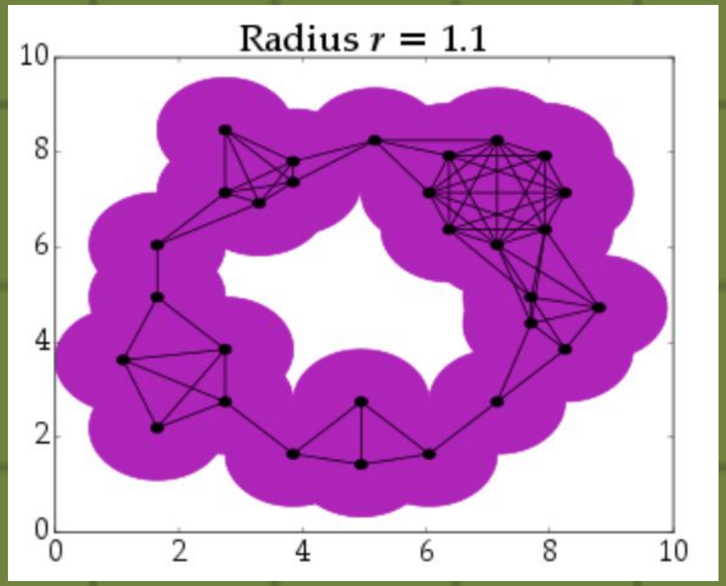


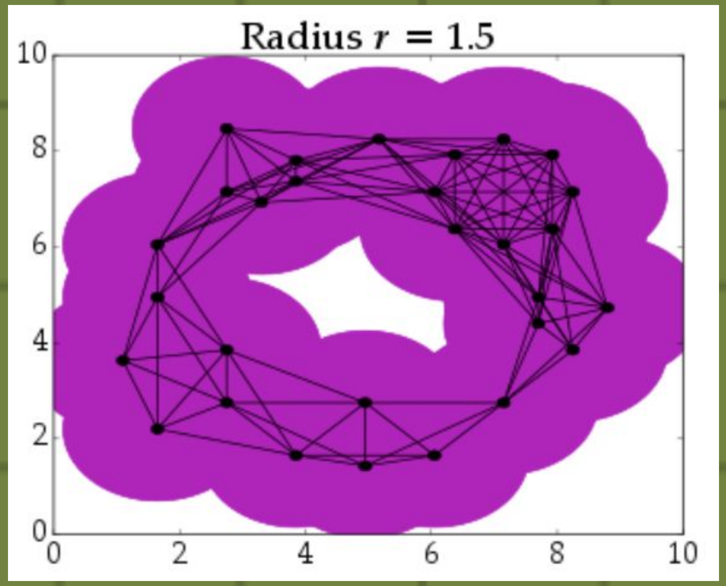












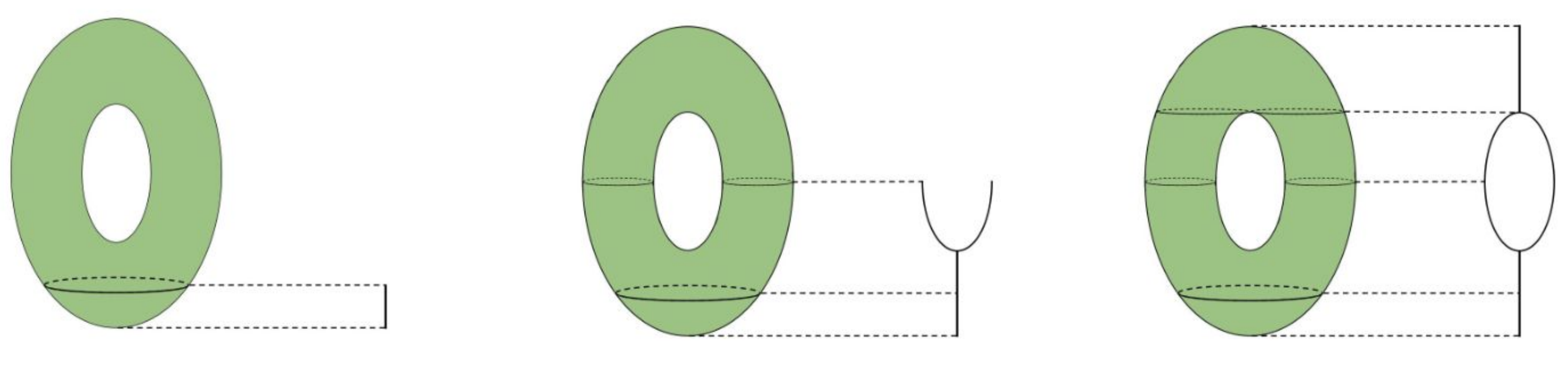
Reeb Graphs & 3D Shape Filtration

Reeb graph:

- Skeleton of 3D shape
- Level set: all things captured at specific height
- Sublevel set: going from bottom to top

3D Shape Filtration:

- How Reeb graph changes as you scan from bottom to top
- Torus ex: one connected component



05

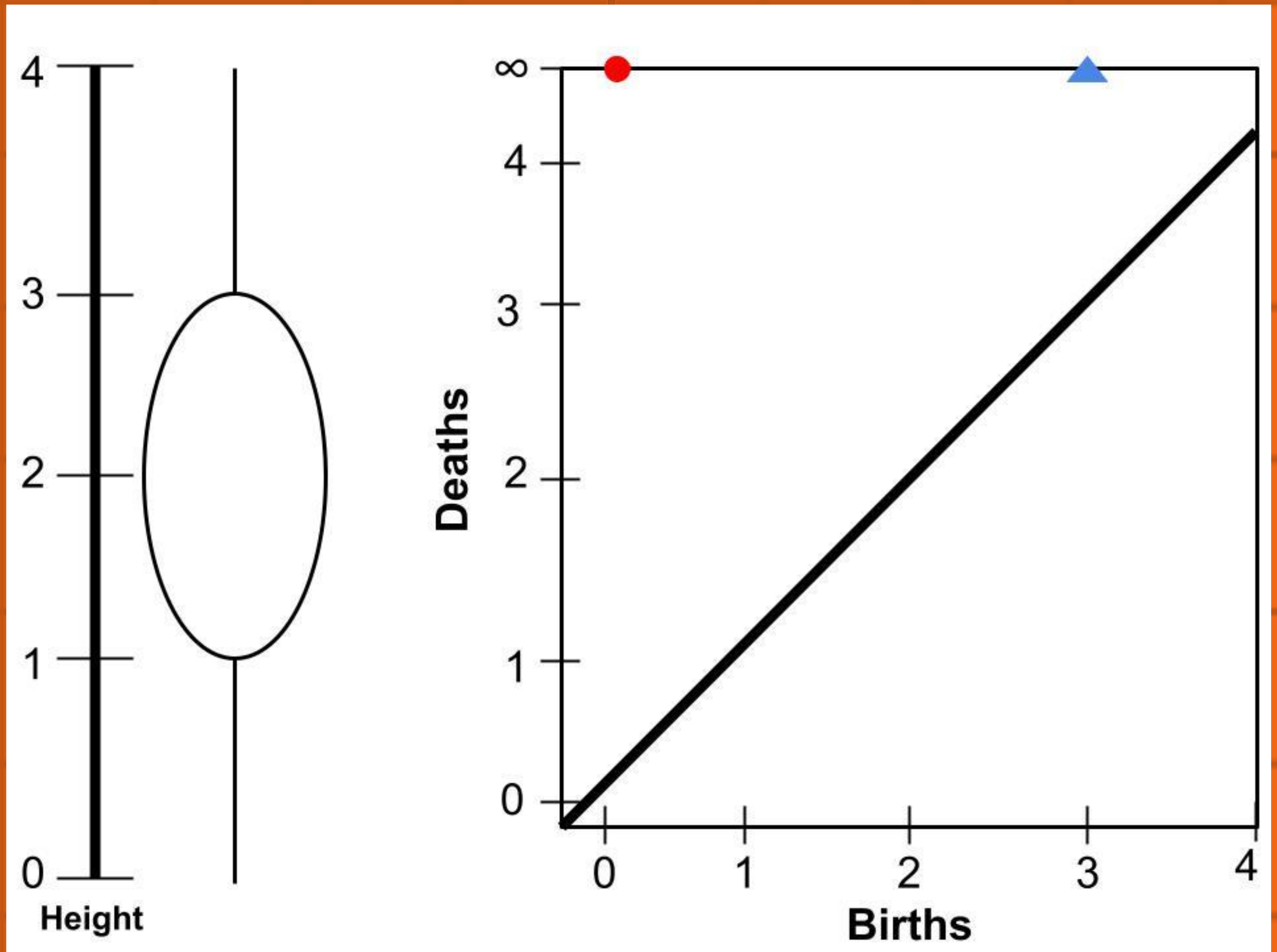
Persistent Homology & Persistence Diagrams

In a word:

Graph of births and deaths of H0 and H1 features throughout filtration changes

Persistent homology:

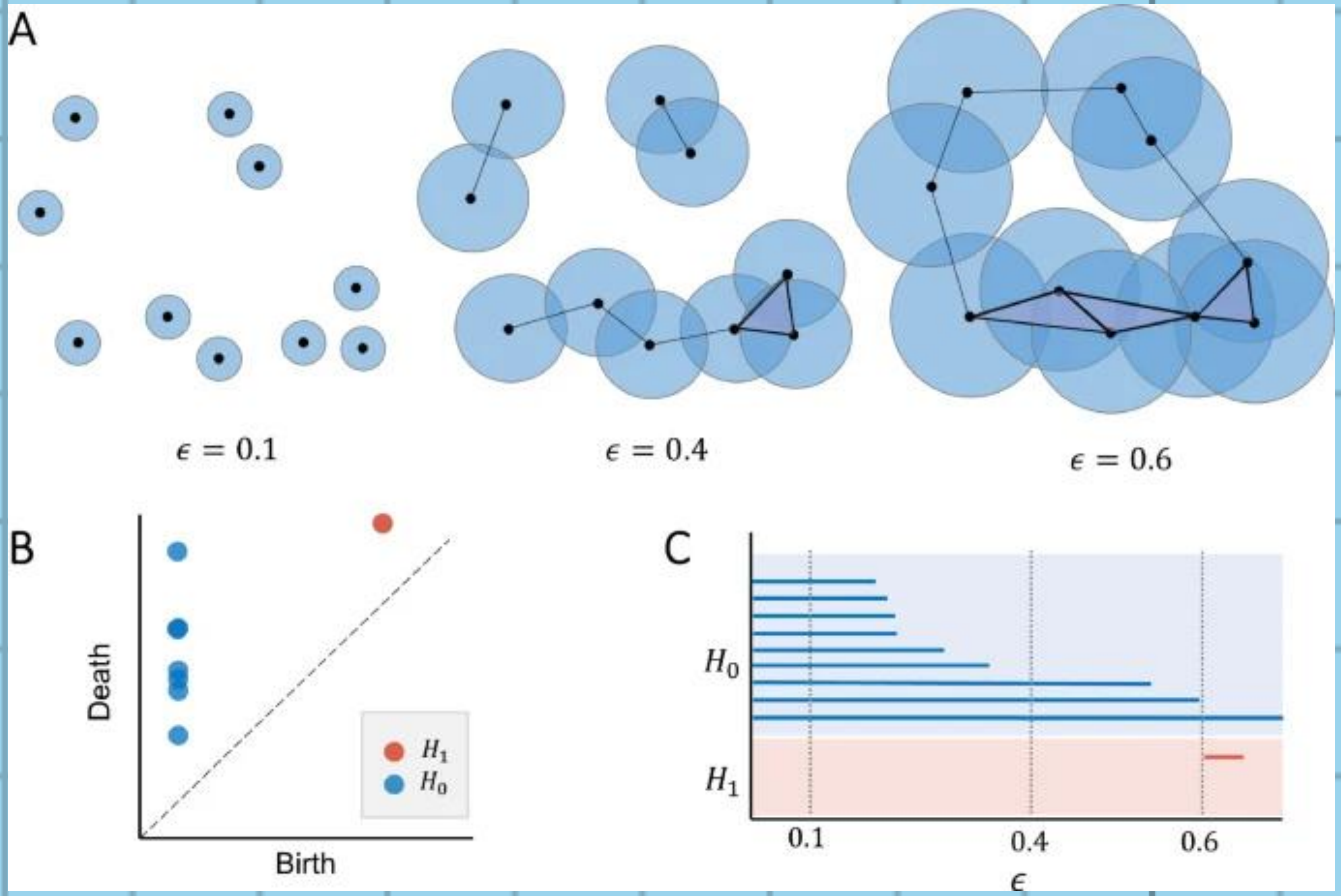
- TDA tool to measure how long each feature survives
- Generally concerned with H0 and H1
- Short life: possibly noise, unimportant
- Long life: important feature / behavior



Persistence diagrams (PD):

- Births on x-axis, deaths on y-axis
- Diagonal line y=x, birth=death
- Close to diagonal: short lifespan, noise
- Far from diagonal: long lifespan, important feature
- Lives forever: height at ∞

Ex: Vietoris-Rips Complex, Persistence Diagram, & Barcode



Barcodes:

- Similar to Persistence Diagrams
- Horizontal bars encode lifespan (birth to death)

Example:

- Vietoris-Rips complex
- Persistence diagram & barcode for H0 and H1 elements

Additional Resources

How to learn more about TDA

Articles:

- Elizabeth Munch, "A User's Guide to Topological Data Analysis," 2017
- Frédéric Chazal and Bertrand
- Michel, "An Introduction to TDA...for Data Scientists," 2021

Textbooks:

- Robert Ghrist, *Elementary Applied Topology*, 2014
- Tamal Krishna Dey and Yusu Wang, *Computational Topology for Data Analysis*, 2022

Other:

- TTK (Topology ToolKit)
- Blog post: Ben Windsor, "An Introduction to Topological Data Analysis," 2020
- Web article: Alexander Del Toro Barba, "Topological Data Analysis with Persistent Homology," 2024

Continuing Work

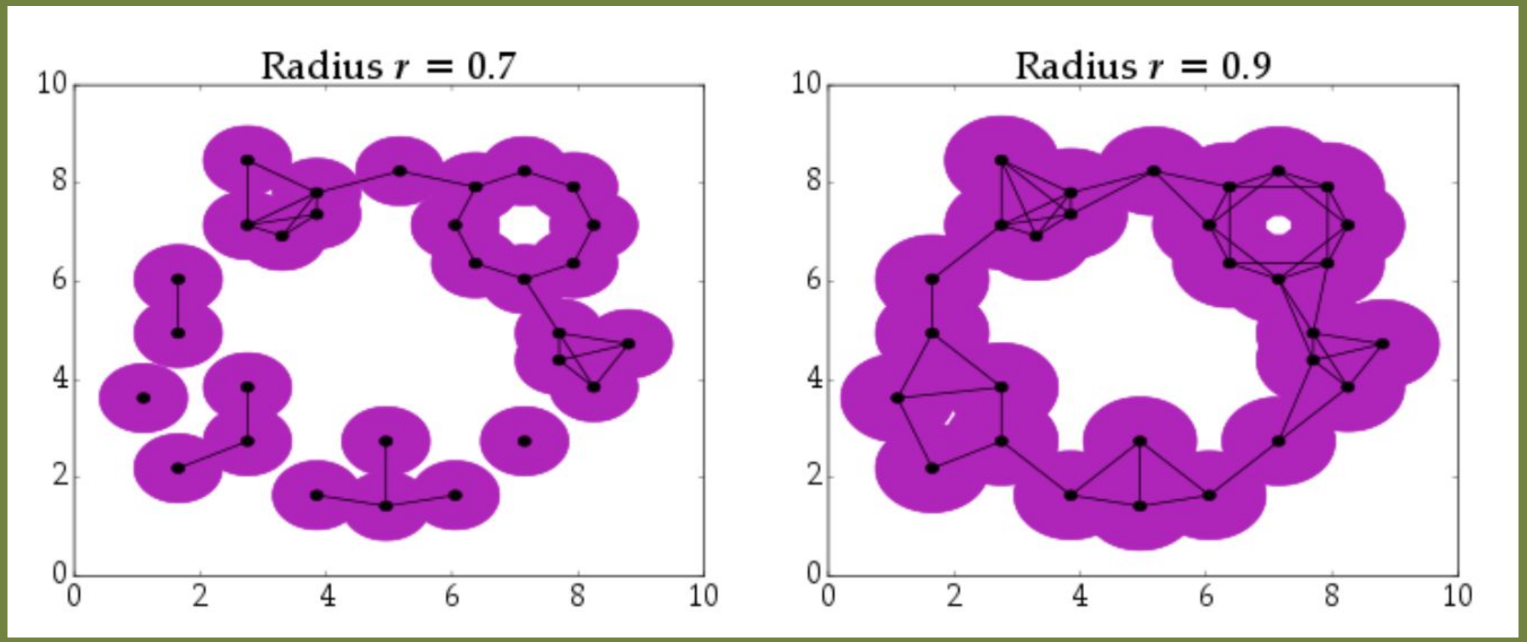
Extensions & expansions:

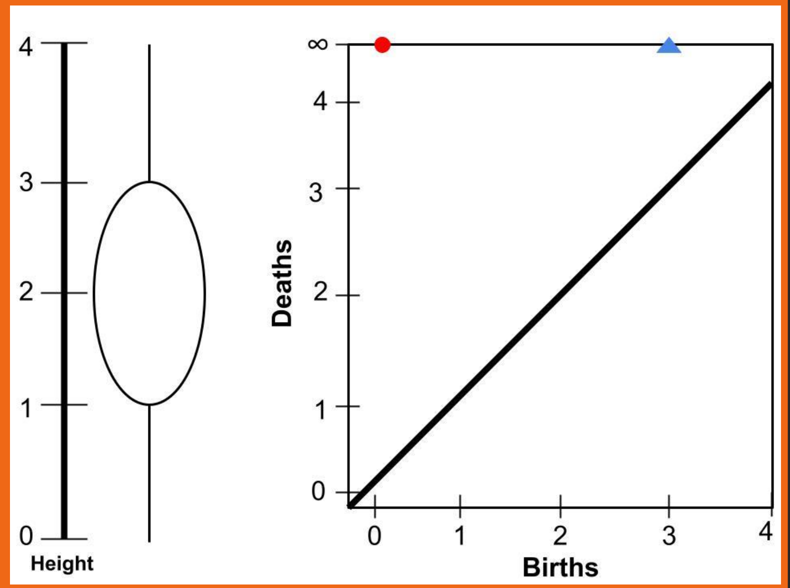
- CODING! Application to computer science / data science for beginners
- Morse theory
- Chain complexes
- Comparing PDs (different distance notions, vineyards)
- Relative homology, extended persistence, bipersistence

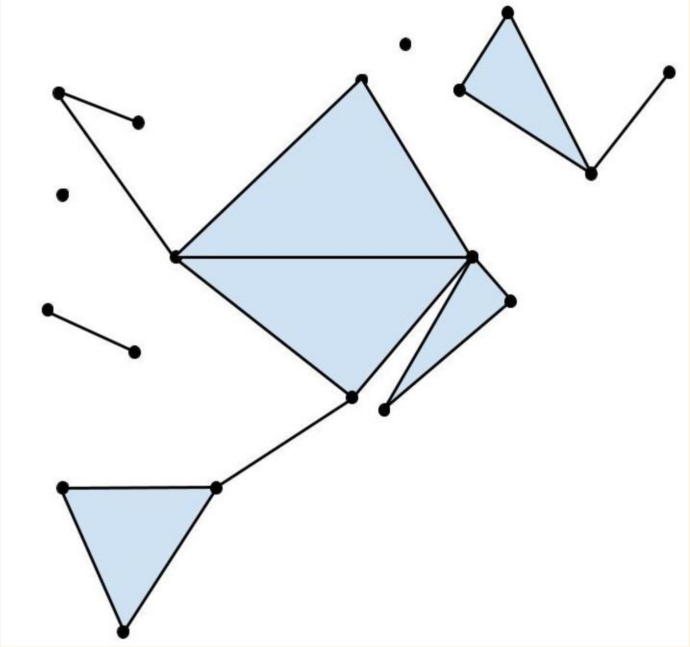
Reader feedback:

- Feedback on accessibility, readability, & understanding
- Variety of backgrounds
- If you're interested in being a reader, email me at eschmi23@nd.edu

THANK







YOU!

References

- H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological Persistence and Simplification. *Discrete and Computational Geometry,* 28:511–533, 2002. https://pub.ista.ac.at/~edels/Papers/2002-04-TopologicalPersistence.pdf .
- A. Muratov. Why a Coffee Cup Is the Same as a Donut (Mathematically), 2025. https://medium.com/@andrew-muratov/why-a-coffee-cup-is-the-same-as-a-donut-mathematically-8964a82d4612.
- E. Munch. A User's Guide to Topological Data Analysis. Journal of Learning Analytics, 4(2):47–61, 2017. https://learning-analytics.info/index.php/JLA/article/view/5196/6061.
- F. Rottach, S. Schieferdecker, and C. Eickhoff. The Topology of Molecular Representations and its Influence on Machine Learning Performance. Journal of Cheminformatics, 17(109), 2025. https://link.springer.com/article/10.1186/s13321-025-01045-w.
- G. Carlsson. Topology and Data. Bulletin of the American Mathematical Society, 46:255–308, 2009. https://www.researchgate.net/publication/243073634_Topology_and_Data.
- What is Topology? University of Waterloo, 2015. University of Waterloo: Pure Mathematics. https://uwaterloo.ca/pure-mathematics/about-pure-math/what-is-pure-math/what-is-topology.
- F. Chazal and B. Michel. An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists. Frontiers in Artificial Intelligence, 4, 2021. https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2021.667963/full.
- R. Ghrist. Elementary Applied Topology. Createspace, 2014. https://www2.math.upenn.edu/~ghrist/notes.html.
- T. K. Dey and Y. Wang. Computational Topology for Data Analysis. Cambridge University Press, 2022. https://www.cs.purdue.edu/homes/tamaldey/book/CTDAbook/CTDAbook.pdf.
- Topology ToolKit: https://topology-tool-kit.github.io/
- B. Windsor. An Introduction to Topological Data Analysis, 2020. Github Blog. https://benwindsorcode.github.io/TDA-Introduction/.
- A. D. T. Barba. Topological Data Analysis with Persistent Homology. Medium, 2024. Blog post. https://medium.com/@deltorobarba/quantum-topological-data-analysis-the-most-powerful-quantum-machine-learning-algorithm-part-1-c6d055f2a4de.