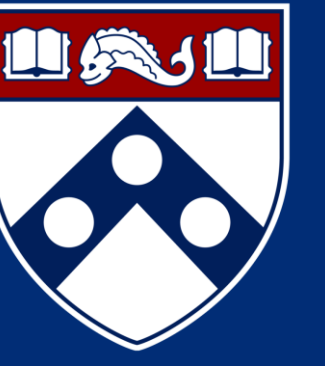# Geometric and Topological Analysis of Connectome Data

Iris Horng

Department of Mathematics, University of Pennsylvania, Philadelphia, PA

Abstract

- Topological Data Analysis is widely used to investigate the organization of brain circuits and abstract underlying geometric structures.
- Research presented in this study investigated a proper filtration that can be used to generate meaningful barcodes in order to represent brain networks across species (including humans). Filtration using curvature was also used to explore the neural network of C. Elegans.
- Analysis of the correlation matrices and barcodes established a pattern among brain networks.

Introduction

- Brain connectivity can be abstracted to a graph of nodes [3]
- This poses a limitation due to the need to represent the brain as a set of discrete vertices and edges
- **TDA** provides strong tools such as persistent homology in order to filter out noise to better understand the shape and abstract structure of high-dimensional data.
- **Ollivier-Ricci curvature** reveals defining features in graph structure: positively curved edges represent an edge within a cluster; negative curvature represents a bridge within clusters [5]

Background

Connectome data
- Functional connectomes: describe functional connectivity between regions in the brain [4]
- resting-state fMRI data can be represented as a correlation matrix
- Neuronal connectomes: represent neuron-to-neuron connections in a nervous system

Persistent Homology and Barcodes
- studies data that persists throughout a defined scale of measurement
- 0th dimension barcode: measures the number of connected components
- 1st dimension barcode: measures number of loops

Curvature
- Ollivier-Ricci curvature compares the "average distance between two sufficiently close points and the distance between two small balls around them" [2]
- Positive Ricci curvature: distance between two neighborhoods of endpoints is less than distance between the two endpoints
- Negative Ricci curvature: distance between two neighborhoods of endpoints is very large

(A) Computing the persistent homology of a point cloud. Adapted from [1]
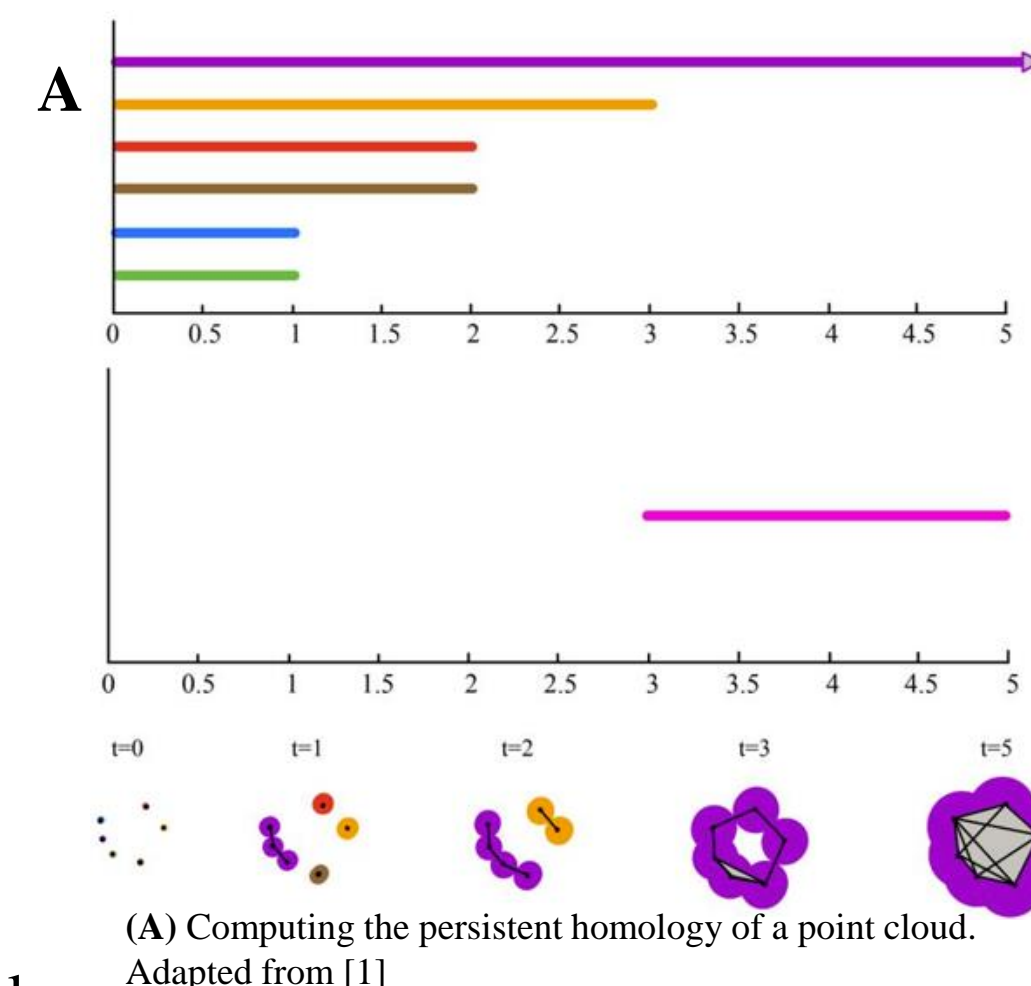
Motivating Question

Can we find a proper filtration that will give us meaningful barcodes, which can be used to identify patterns in connectome connectivity?

Methodology

Brain Connectome Data
- Functional MRI data for 18 subjects from NeuroData was analyzed (undirected graphs).
- data was processed into symmetric correlation matrices with 0's on the diagonal.
- Algorithm for filtration: all nodes (0-simplices) appear at $\varepsilon=0$; all edges (1-simplices) appear at the values listed in the matrix; all filled-in triangles (solid 2-simplices) appear once the third side of the triangle is added

C. Elegans
- Data from NeuroData (directed graphs)
- data was processed into appropriate non-symmetric adjacency matrices A of curvature values where no edge is represented by value 2.
- Algorithm for filtration: all nodes appear at $\varepsilon=0$; edges appear at the maximum of $A(i,j)$ and $A(j,i)$; taking a representative value for each edge as the minimum of the curvature in one direction and the curvature in the opposite direction, filled-in triangles appear at the maximum of the three representative values of the edges
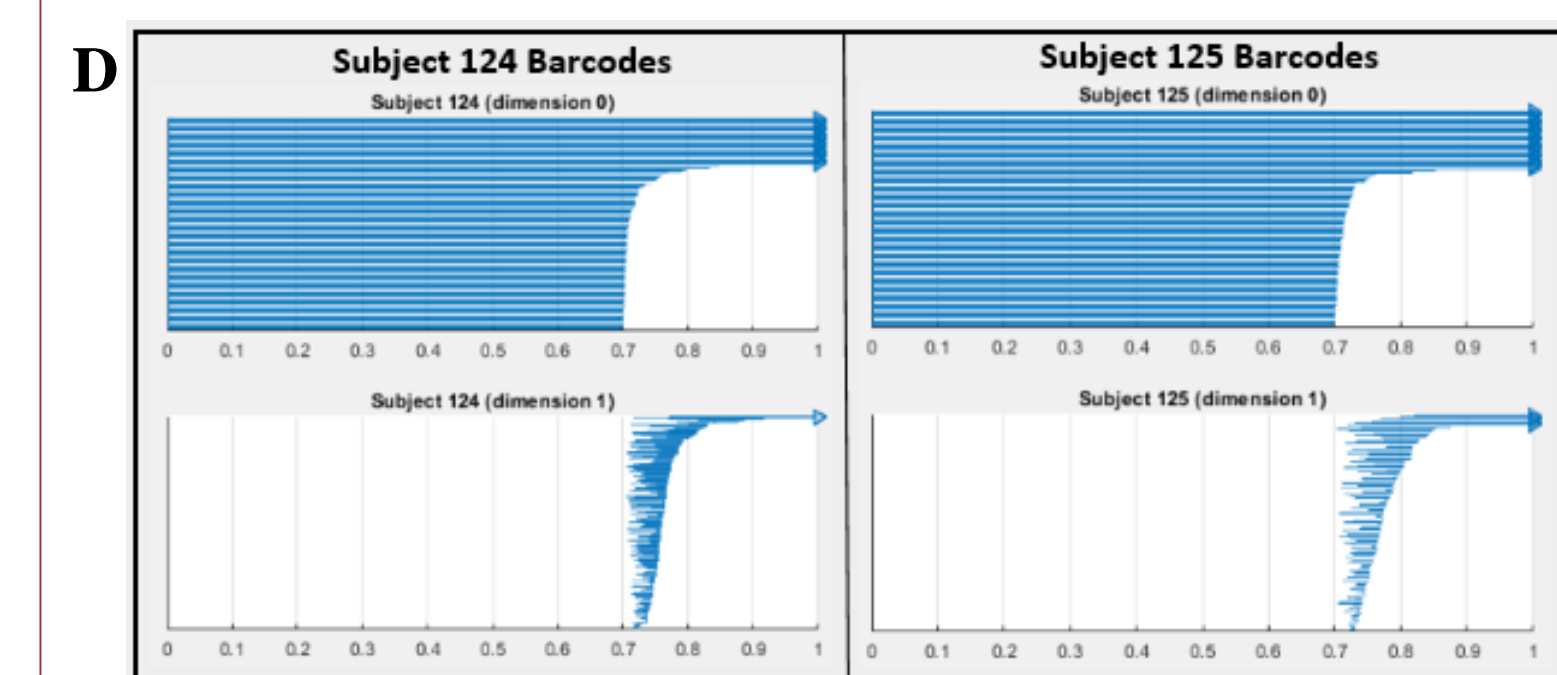
Results
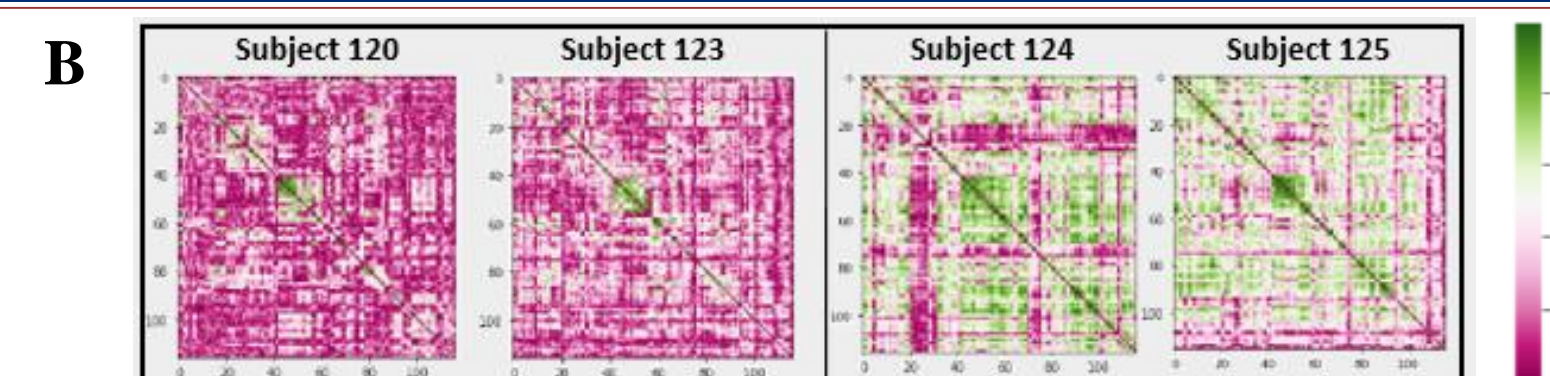
(B) Heat maps of the correlation matrices show brain connectivity for the functional brain networks

Barcodes for the brain connectome data were generated using correlation as the filtration values
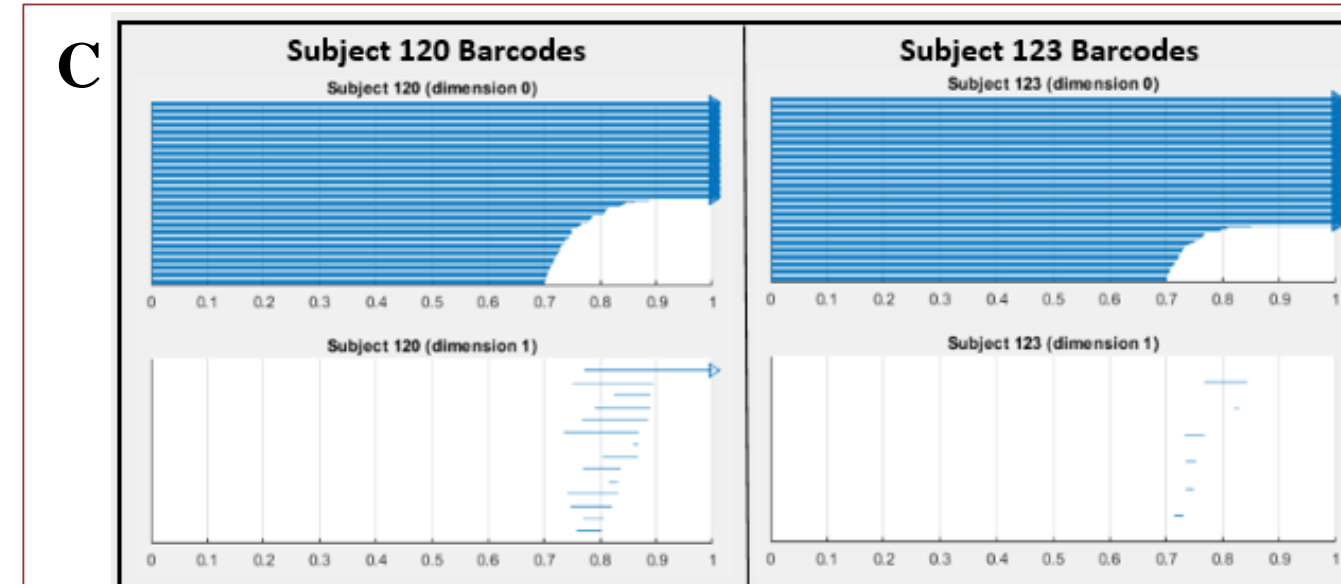
(C) The 0th and 1st dimension barcodes for Subjects 120 and 123 show 14 loops and 6 loops respectively
- Less connected → fewer loops

Acknowledgments

Results

(D) The 0th and 1st dimension barcodes for Subjects 124 and 125 show 173 loops and 107 loops respectively
- More connected → more loops

Finding: For the matrices with more **low correlation** values, the heatmap is more **pink**, and the barcodes show **fewer loops**. For the matrices with more **high correlation** values, there's more **green** in the heatmap, and the barcodes show there are **more loops**.
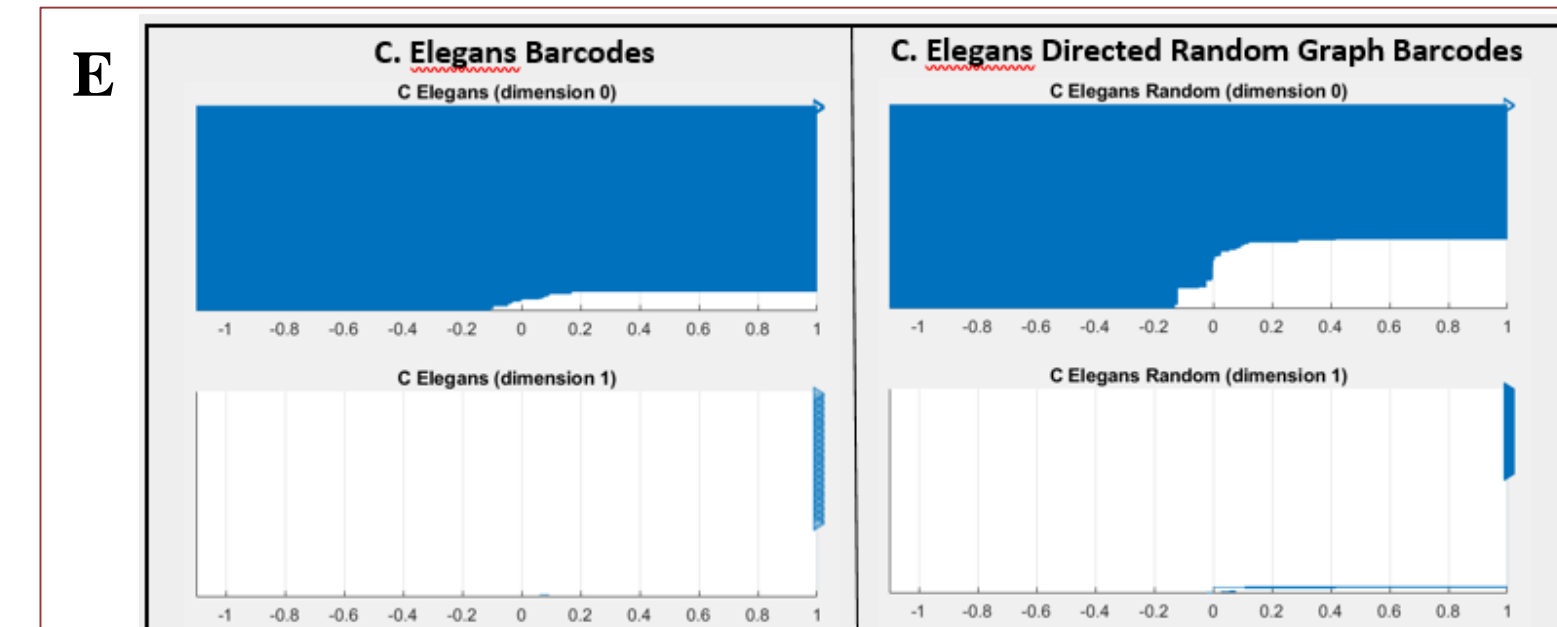
(E) C. Elegans Barcodes were generated using **curvature** as the filtration values

- indicates no loops in the branch-like network
- C. Elegans network (left) differs from a directed random graph (preserves weight, degree, and strength distributions but randomizes the edges) of the C. Elegans network (right)

Future Work

- Computing barcodes for higher dimensions may reveal further features in the topological shape of the data
- Advancing the **clique complex algorithm** to account for higher-dimensional simplices would be ideal for analyzing a high-dimensional complex network like the brain
- Determining which features or regions of the brain correspond to the persisting bars would allow us to form interpretations regarding what that reveals about the brain

References

[1] Brain Connectivity Toolbox. Google Sites: Sign-in. (n.d.). Retrieved August 19, 2022, from https://sites.google.com/site/bctnet/
[2] Eidi, M., & Jost, J. (2020). Ollivier Ricci curvature of directed hypergraphs. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-68619-6
[3] Fornito, A., Zalesky, A., & Zhu, X. shou. (2016). Fundamentals of Brain Network Analysis. Elsevier.
[4] Functional connectivity and the human brain functional connectome - florey. (n.d.). Retrieved July 30, 2022, from https://florey.edu.au/science-research/research-projects/functional-connectivity-and-the-human-brain-functional-connectome
[5] NeuroData. Networks. (n.d.). Retrieved July 29, 2022, from https://neurodata.io/project/connectomes/
[5] Ni, C.-C., Lin, Y.-Y., Luo, F., & Gao, J. (2019). Community detection on networks with Ricci flow. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-46380-9
[6] Tausz, A., Vejdemo-Johansson, M., & Adams, H. (2014). JavaPlex: A research software package for persistent (co)homology https://doi.org/10.1038/s41598-019-46380-9

Notes: Javaplex [6] was used to generate barcodes and [5] was used to compute curvatures