CS 6170: Computational Topology, Spring 2019
Lecture 05
Topological Data Analysis for Data Scientists

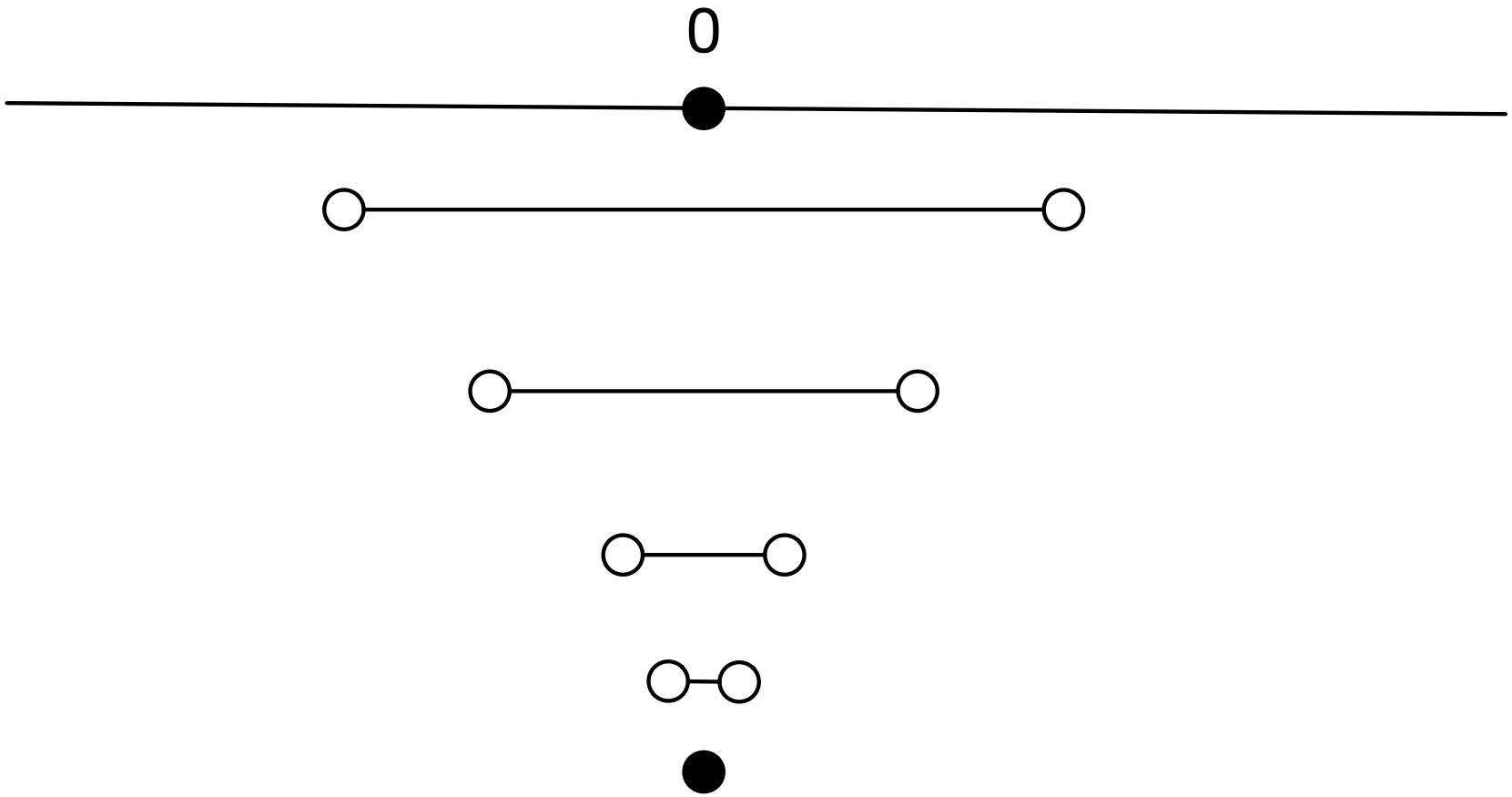
Dr. Bei Wang

School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah
www.sci.utah.edu/~beiwang
beiwang@sci.utah.edu

Jan 22, 2019

Topology requires a finite number of set intersections?

- Consider the topology on the real line with open intervals
- Imagine a set of intervals $[-\frac{1}{n}, +\frac{1}{n}]$ on the real line, where n goes to ∞.
- Now take the *infinite* number of intersections...
- 0 in the intersection is now a discrete point



Čech complexes and Vietoris-Rips complexes

Book Chapter A.III

Demo:

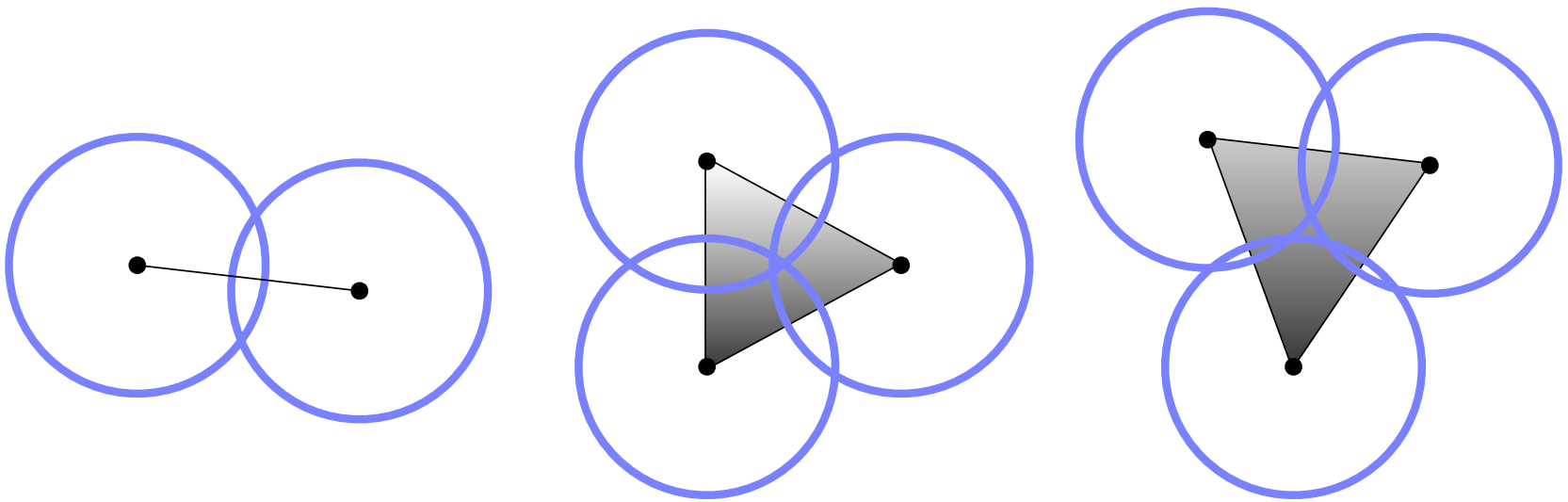
`http://www.sci.utah.edu/~tsodergren/prob_net_vis_working/`

Vietoris-Rips Complex

Given a finite metric space (P, d), e.g., a point cloud $P \subset \mathbb{R}^n$:

Definition

The *Vietoris-Rips complex* of P and r consisting of all subsets of diameter at most $2r$. That is, $\sigma \in \mathcal{R}(r)$ iff $d(p, q) \leq 2r$ for all $p, q \in \sigma$.

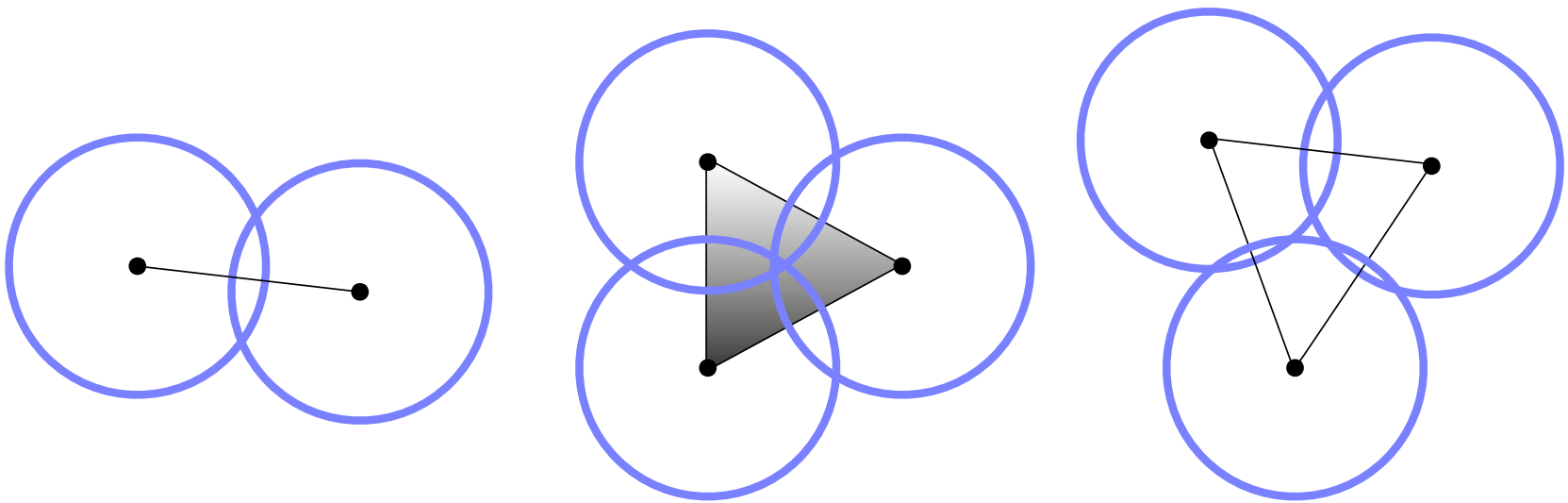


Čech Complex

Given a finite metric space (P, d), e.g., a point cloud $P \subset \mathbb{R}^n$:

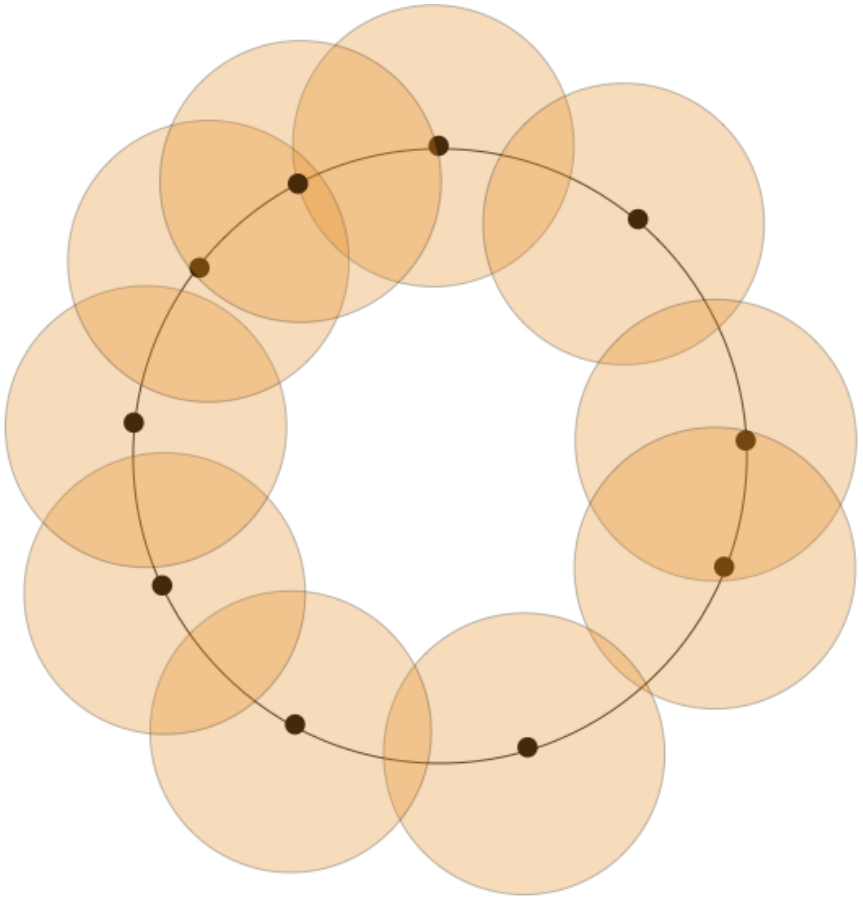
Definition

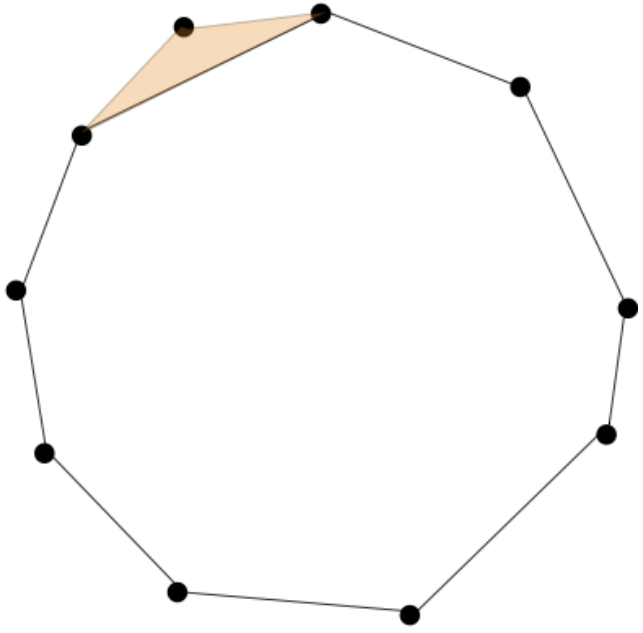
The *Čech complex* of P and r consisting of all subsets whose corresponding sets of r-balls have nonempty intersections. That is, $\sigma \in \mathcal{C}(r)$ iff $\bigcap_{x \in \sigma} B_x(r) \neq 0$.



Sensor Networks

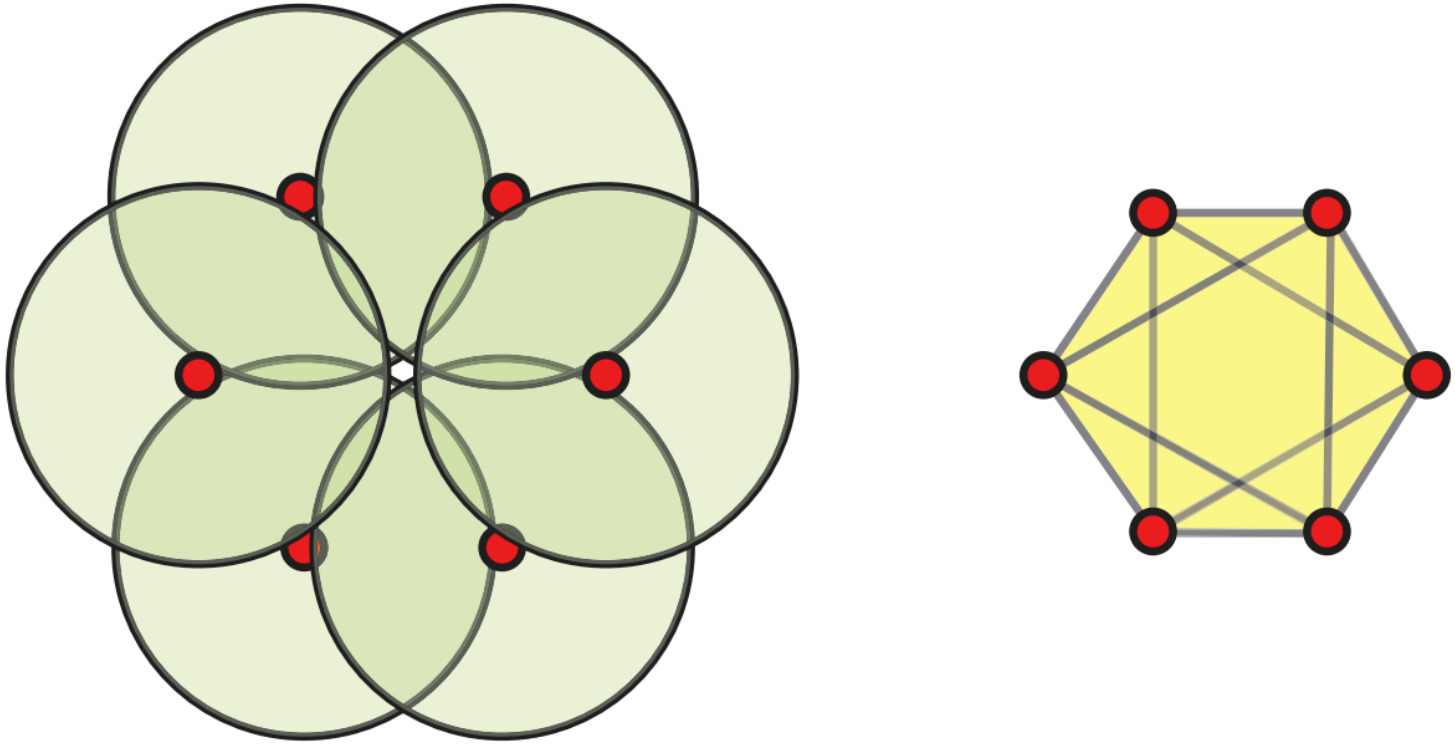
Sensor coverage with Čech complex





https://en.wikipedia.org/wiki/Cech_complex

Rips complex detects "phantom" topological features



Left: Union of balls (Čech). Right: Rips.
Image courtesy of de Silva and Ghrist (2007a).

Using Rips complex to approximate Čech complex

Lemma (Edelsbrunner and Harer (2010), page 62)

Let P be a finite set of points in some Euclidean space and letting $r \geq 0$, then

$$\mathcal{R}(r) \subseteq \mathcal{C}(\sqrt{2}r).$$

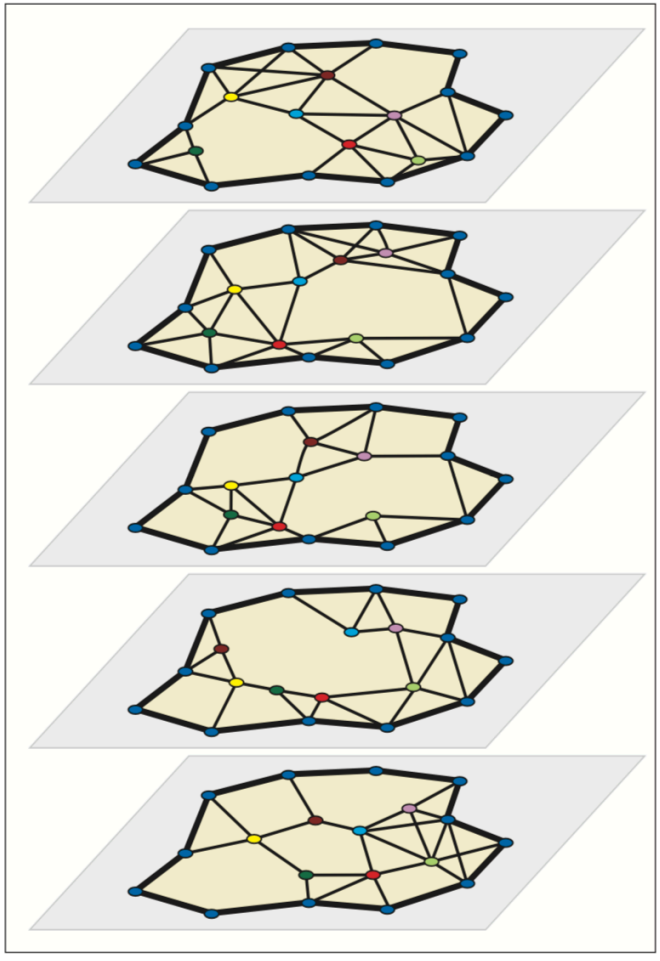
Theorem (de Silva and Ghrist (2007a), Theorem 2.5)

For a point set P in $\mathbb{R}^n$,

$$\mathcal{C}(r) \subset \mathcal{R}(r) \subset \mathcal{C}(\theta r),$$

whenever $\theta \geq \sqrt{\frac{2n}{n+1}}$.

Time-Varying sensor networks



A time-sequence of network graphs for a mobile network. Does this network admit a wandering hole? de Silva and Ghrist (2007b)

Topological Data Analysis for Brain Networks

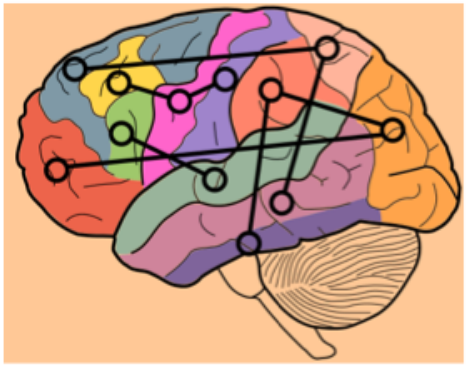
Correlating Brain Network Topology with Autism Severity
Wong et al. (2016)

Big Picture

Goal: Quantify the relationship between brain functional networks and behavioral measures.

Our Contribution: Use topological features based on persistent homology.

Result: Combining correlations with topological features gives better prediction of autism severity than using correlations alone.



functional network

behavior

Motivation

About Autism Spectrum Disorders (ASD):

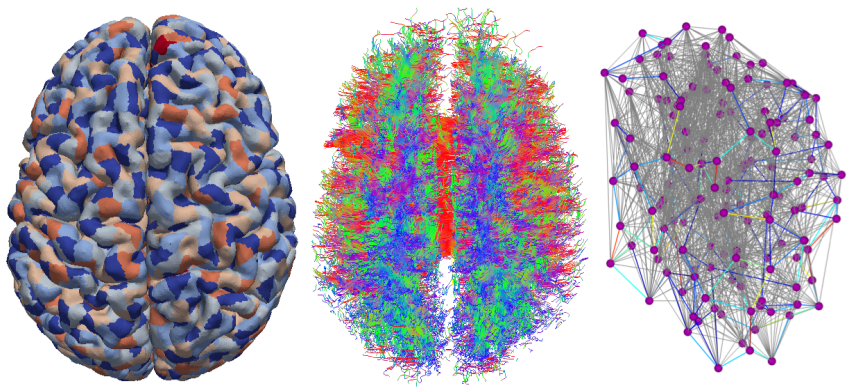
- No cure, causes unknown
- Diagnosis:
  - No systematic method
  - ADOS (Autism Diagnostic Observation Schedule)

Correlate functional brain network to ADOS scores

- Early diagnosis
- Treatment tracking

What is a Brain Network?

- Represents brain regions and pairwise associations
- Computation of Correlation Matrices:
  - Resting state functional MRI (R-fMRI)
  - Preprocessing
  - Define regions of interest (ROIs)
  - Estimate time series signals
  - Compute pairwise associations - Pearson Correlation



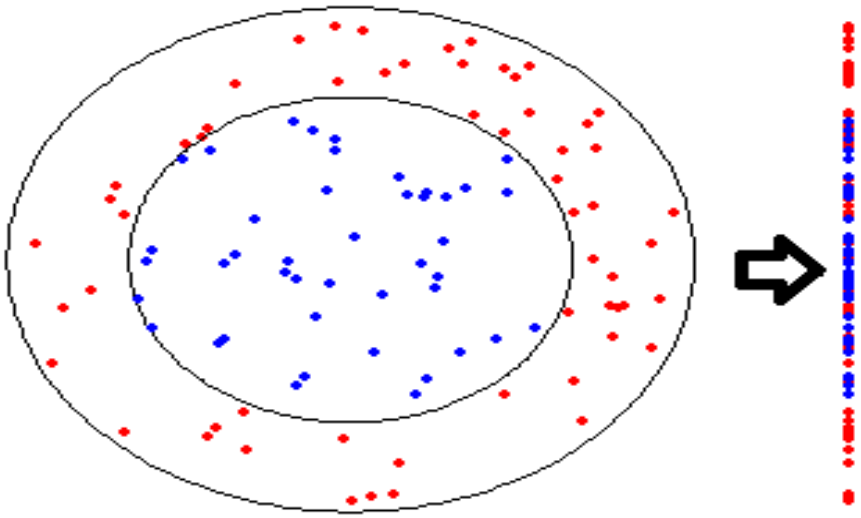
Why Topology?

How to use this data?

- Graph and graph theoretic measures (e.g. small worldness)
  - Require binary associations (thresholding)
- Correlations as features
  - High dimensionality, not enough samples
- Dimensionality reduction: PCA, random projections
  - May lose structures in higher dimensions

Why Topology

Projection - may lose structures in higher dimensions



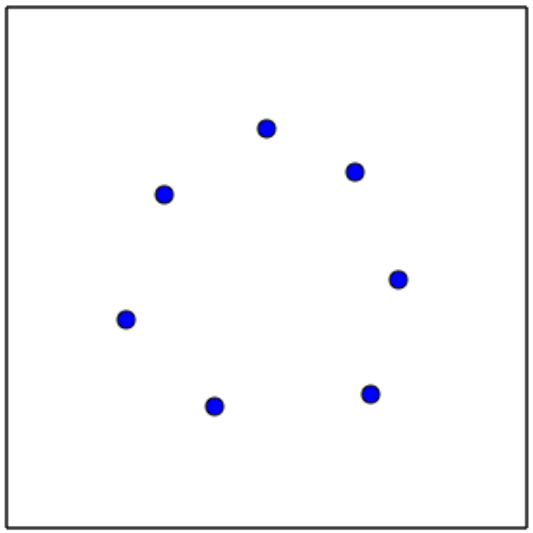
Topology captures structure

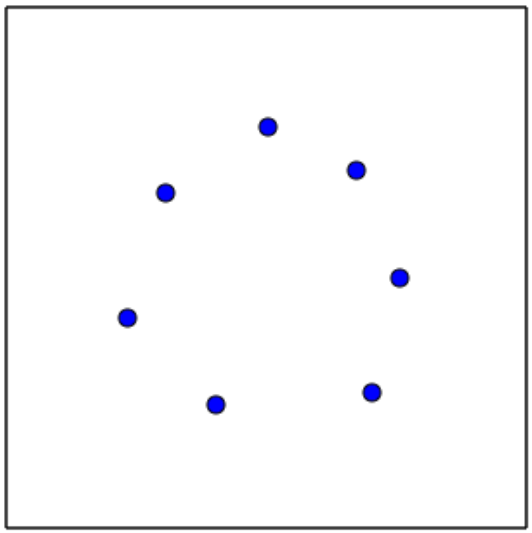
- In higher dimensions
- Across all continuous thresholds

Persistent Homology

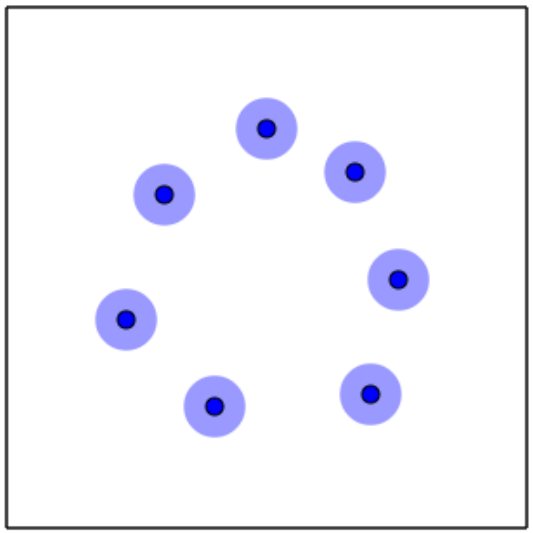
- What are topological features? Homological features:
  - Dim 0 - Connected Components
  - Dim 1 - Tunnels / Loops
  - Dim 2 - Voids
- How to compute them (in a nutshell)?
  - Begin with point cloud
  - Grow balls of diameter t around each point
  - Track features of the union of balls as t increases

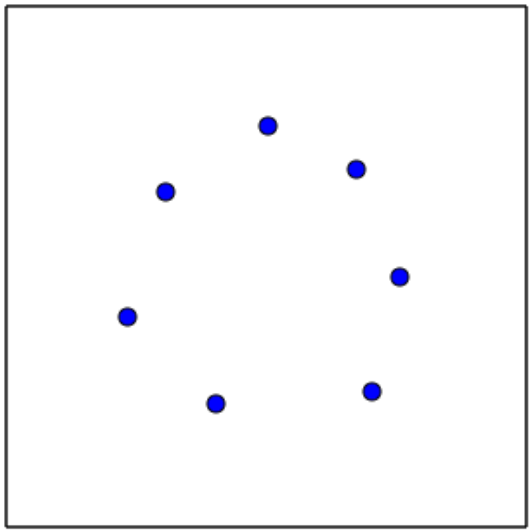
Persistent Homology



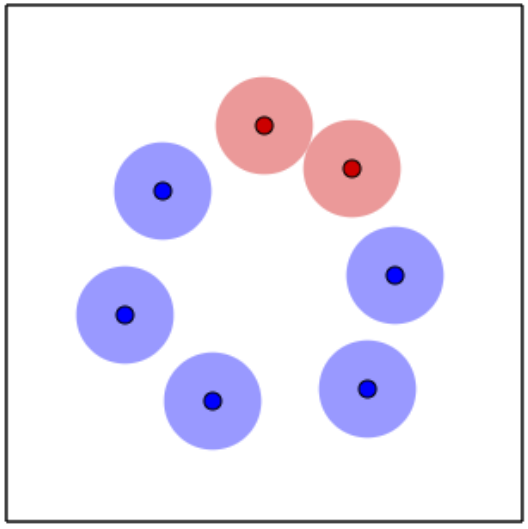


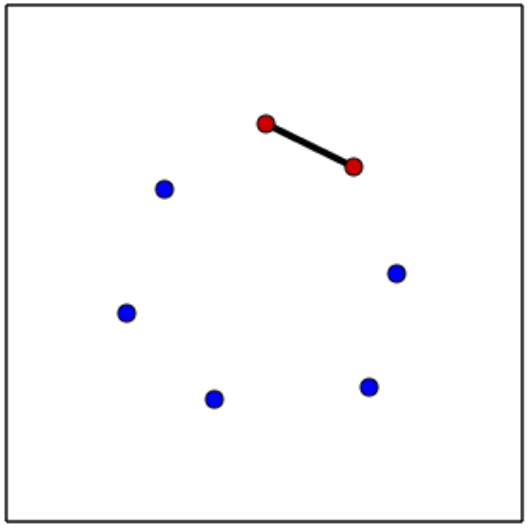
Persistent Homology



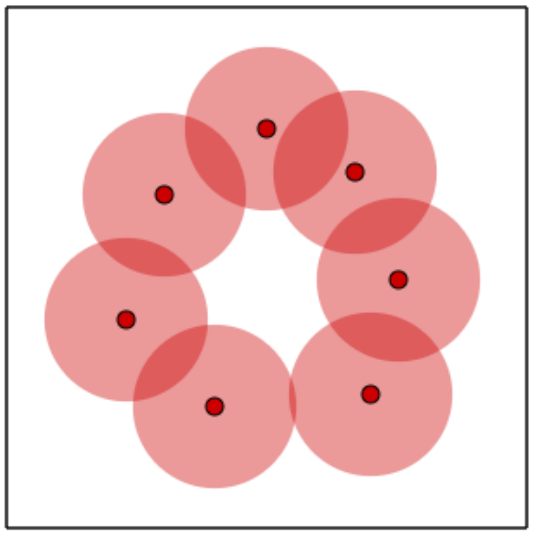


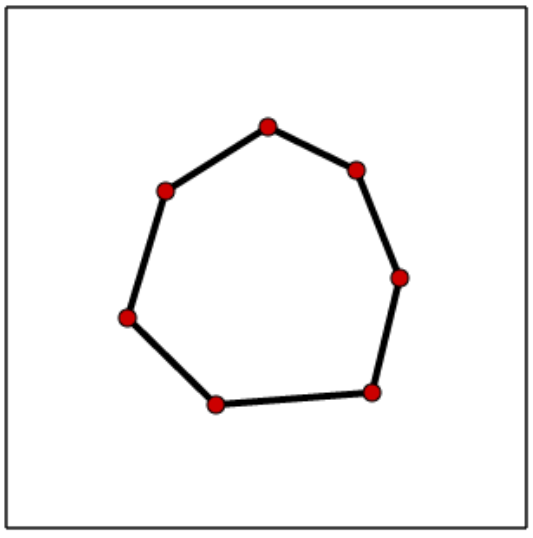
Persistent Homology



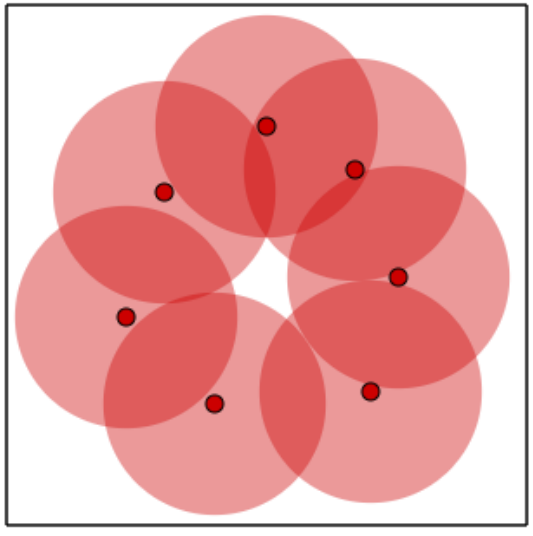


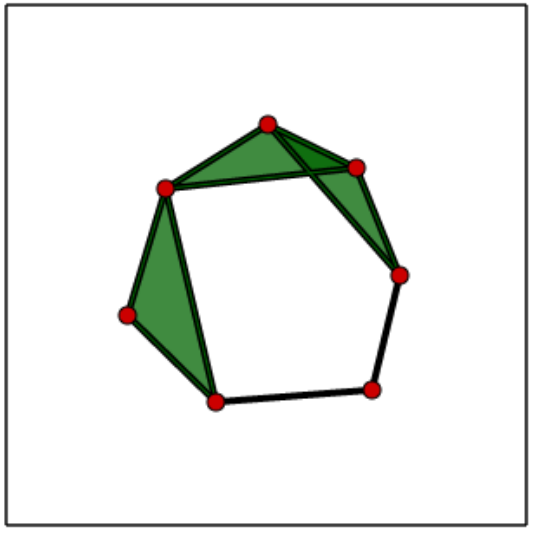
Persistent Homology



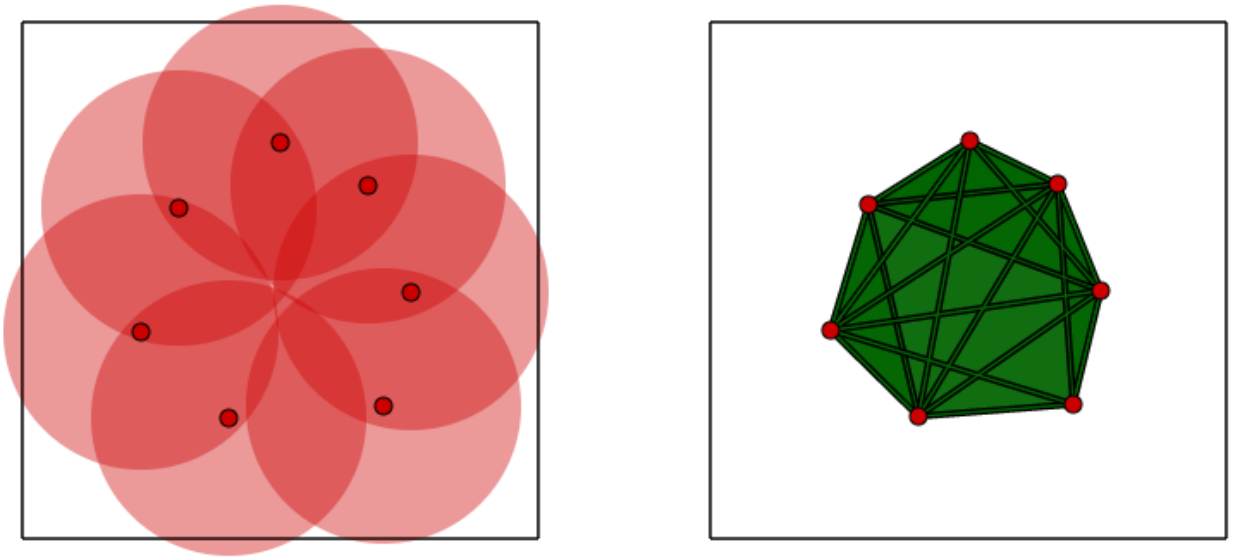


Persistent Homology





Persistent Homology



Persistence Diagrams

Persistent homological features - encoded as barcodes or persistent diagrams

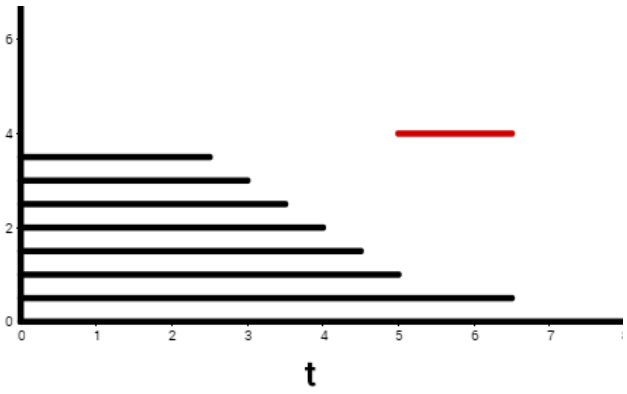


Figure: Barcode

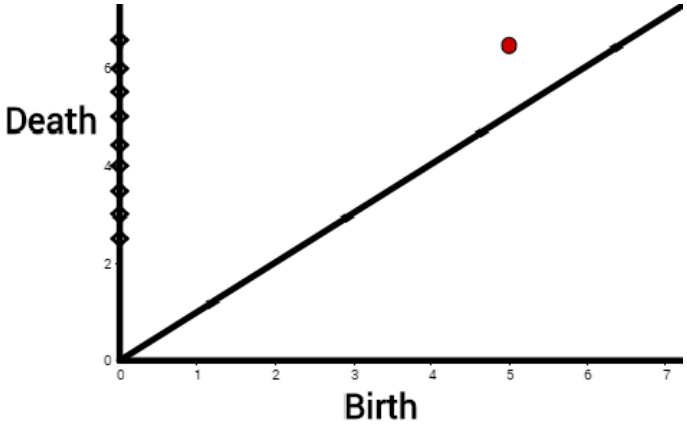
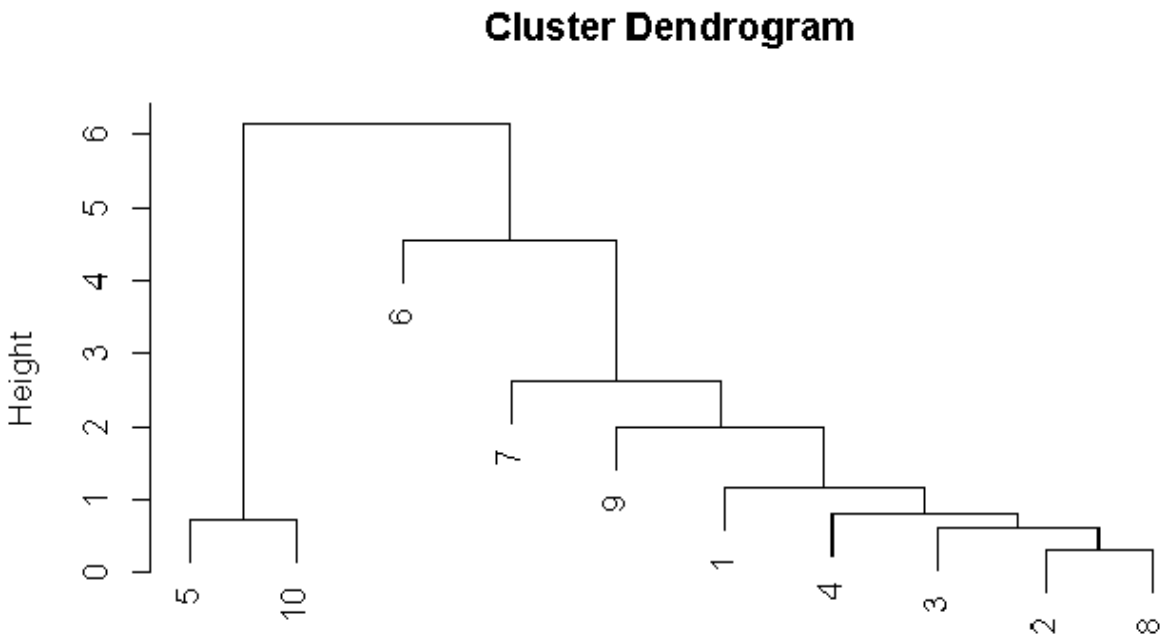


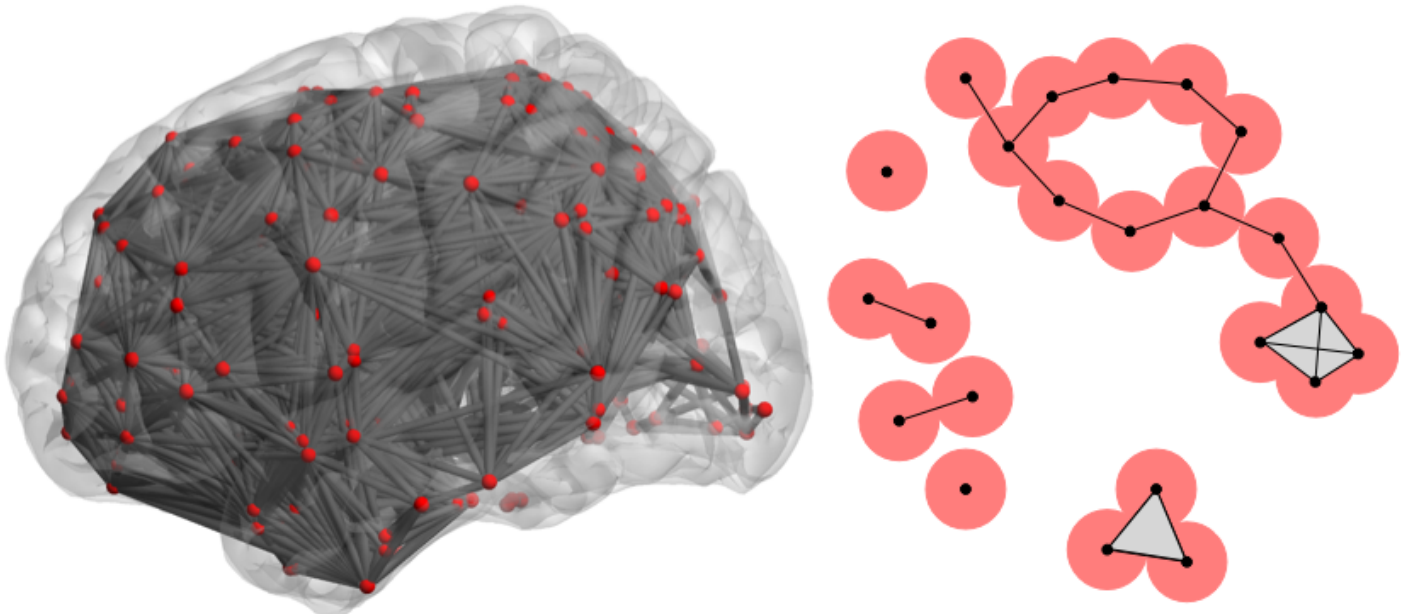
Figure: Persistence Diagram

Interpretation of Connected Components

- Dim 0 features - hierarchical clustering



Computing Topological Features for Brain Networks



Partial Least Squares (PLS) Regression

A dimensionality reduction technique that finds two sets of latent dimensions from datasets X and Y such that their **projections** on the latent dimensions are **maximally co-varying**.

- X - features from brain imaging: correlations, topological features (zero mean)
- Y - clinical measure of behavior: ADOS scores (zero mean)

PLS models the relations between X and Y by means of **score vectors**.

PLS Regression

- n - number of data points
- X - predictor/regressor $(n \times N)$, Y - response $(n \times M)$
- PLS - decompose X, Y such that:

$$X = TP^T + E$$
$$Y = UQ^T + F$$

Where

- T, U - latent variables/score vectors $(n \times p)$, factor matrices
- P $(N \times p)$, Q $(M \times p)$ - orthogonal loading matrices
- E $(n \times N)$, F $(n \times M)$ - residuals/errors

- T, U are chosen such that projections of X, Y, that is, T and U, are maximally co-varying.

PLS Regression: the Algorithm

Iterative NIPALS algorithm Wold (1975) (nonlinear iterative partial least squares)

- Find first latent dimension

  i.e. find vectors w, c such that

  $$t = Xw, \quad u = Yc$$

  have maximal covariance
- Deflate previous latent dimensions from X, Y and repeat

Kernel PLS

Kernel form of NIPALS algorithm (kPLS)

1. Initialize random vector u
2. Repeat until convergence
   a. $t = Ku/\|Ku\|$
   b. $c = Y^T t$
   c. $u = Yc/\|Yc\|$
3. Deflate $K = (I - tt^T)K(I - tt^T)$
4. Repeat to compute subsequent latent dimensions

Data

- 87 Subjects: 30 Control, 57 ASD
- ADOS scores: 0 to 21
- 264 ROIs (Power regions)
- 264×264 correlation matrix.
- 34,716 distinct pairwise correlations per subject.

Experiments

- Given: Correlation matrices
- Map to metric space

$$d(x, y) = \sqrt{1 - \mathsf{Cor}(x, y)}$$

- Compute persistence diagrams
- Define inner product of persistence diagrams Reininghaus et al. (2015) (i.e. kernel): Given two persistence diagrams F, G

$$k_\sigma(F, G) = \frac{1}{8\pi\sigma} \sum_{p \in F} \sum_{q \in G} e^{-\frac{\|p-q\|^2}{8\sigma}} - e^{-\frac{\|p-\bar{q}\|^2}{8\sigma}}$$

where for every $q = (x, y) \in G$, $\bar{q} = (y, x)$

Experiments

Performed experiments with 3 kernels:

1. K^{Cor} - Euclidean dot product of vectorized correlations
2. $K^{TDA} = w_0 K^{TDA_0} + (1 - w_0) K^{TDA_1}$
   - K^{TDA_0} - using only Dim 0 features
   - K^{TDA_1} - using only Dim 1 features
3. $K^{TDA+Cor} = w_0 K^{TDA_0} + w_1 K^{TDA_1} + (1 - w_0 - w_1) K^{Cor}$

Baseline predictor - mean ADOS score

Experiments

- Leave one out cross validation over parameters
  - σ_0, σ_1 - $(\log_{10} \sigma)$ from -8.0 to 6.0 by 0.2
  - w_0, w_1 - from 0.0 to 1.0 by 0.05
- k^{TDA} parameters: $\sigma_0 = -6.6$, $\sigma_1 = 1.8$, $w_1 = 0.95$
- $k^{TDA+Cor}$ parameters: $\sigma_0 = -7.8$, $\sigma_1 = 2.8$, $w_0 = 0.1$, $w_1 = 0.4$
- Compute RMSE
- Permutation test for significance

Results and Conclusions

	RMSE	ADOS mean	K^{TDA}	K^{cor}
ADOS mean	6.4302	-	-	-
K^{TDA}	6.3553	0.316	-	-
K^{cor}	6.0371	0.055	0.095	-
$K^{\mathrm{TDA+cor}}$	**6.0156**	**0.048**	0.075	0.288

Table: ADOS prediction results. Columns 2 to 4 are p-values for the permutation test of improvement of row method over column method.

Result Highlights:

- Baseline RMSE: 6.4302
- $K^{TDA+Cor}$:
  - Only method statistically significant over baseline
  - Permutation test p-value: 0.048
  - RMSE: 6.0156

Conclusion

- Augmenting correlations with topological features gives a **better** prediction of autism severity than using correlations alone
- (Hopefully) topological features derived from R-fMRI have the **potential** to explain the connection between functional brain networks and autism severity

Future Work

- Alternatives to correlation
- Different distance metric
- Different kernel
- Multi-site data
- Classification (combine with TDA features)

References I

de Silva, V. and Ghrist, R. (2007a). Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358.

de Silva, V. and Ghrist, R. (2007b). Homological sensor networks. *NOTICES OF THE AMS*, 54(1).

Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: An Introduction*. American Mathematical Society, Providence, RI, USA.

Reininghaus, J., Huber, S., Bauer, U., and Kwitt, R. (2015). A stable multi-scale kernel for topological machine learning. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*.

Wold, H. (1975). Soft modeling by latent variables; the nonlinear iterative partial least squares approach. *Perspectives in Probability and Statistics, Papers in Honour of M.S. Bartlett*, pages 520–540.

Wong, E., Palande, S., Wang, B., Zielinski, B., Anderson, J., and Fletcher, P. T. (2016). Kernel partial least squares regression for relating functional brain network topology to clinical measures of behavior. *International Symposium on Biomedical Imaging (ISBI)*.