

Background  
oooooooo

Community Detection  
oooooooooooo

Kernel Regression  
oooooooooooo

Collaborative Work  
ooo

# Statistical Methods for Network Association and Discovery in Neuroimaging Data

Alexandria Jensen

UC Denver Anschutz Medical Campus - Department of Biostatistics and Informatics

31 January 2022



Background  
oooooooo

Community Detection  
oooooooooooo

Kernel Regression  
oooooooooooo

Collaborative Work  
ooo

# Outline

Background

Community Detection

Kernel Regression

Collaborative Work

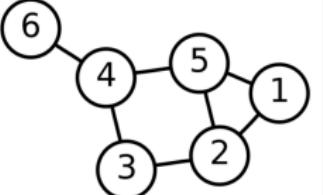
# About Me

- BS in Mathematics and BA in Economics from Colorado State University (2011)
- MS in Biostatistics from CU Anschutz (2018)
- PhD candidate in the Department of Biostatistics & Informatics at CU Anschutz
- Graduate research assistant in CIDA, working in the Colorado Aging Brain Lab as well as the Translational Outcomes in Radiology Research (TORR) Lab.



# A Brief Review of Graphs

A graph  $\mathcal{G}$  is a mathematical structure used to model pairwise relationships between objects, composed of a nonempty, finite set  $V$  of nodes connected by a set  $E \subseteq \{\{x,y\} | x,y \in V \text{ and } x \neq y\}$  called edges.

Labelled Graph	Adjacency Matrix	Degree Matrix
	$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

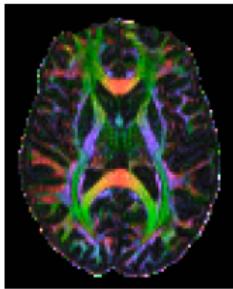
# Graphy Theory and the Human Brain

- Applied to structural and functional brain connectivity by Ed Bullmore and Olaf Sporns in their seminal 2009 *Nature* publication
- Thousands of articles have been published using graph theory to understand brain connectivity
  - Network organization of the brain.
  - Neuropsychiatric disorders as disconnectivity syndromes.

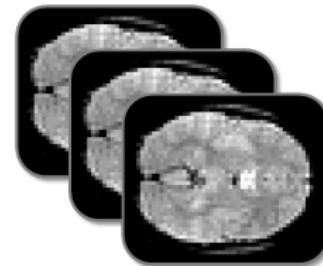
# Building Brain Graphs

In its most general form, a brain graph is composed of (i) a set of nodes characterizing anatomical, functional, or computational units and (ii) a set of edges representing pairwise relationships between nodes.

When using MRI as the basis for a brain graph, either a structural or functional MR sequence can be used.



(a) Structural MR Image



(b) Functional MR Image

# Modular Organization of the Brain

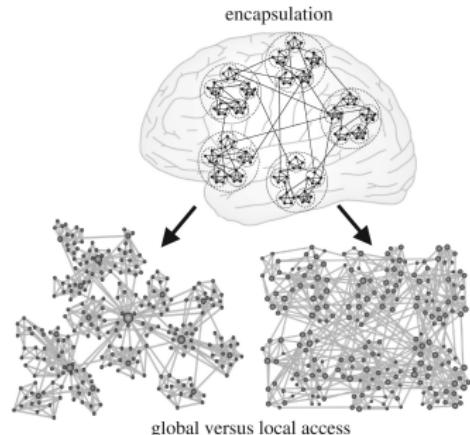


Modular (or community) organization of the brain is thought to be a method of minimizing energy consumption in the development and maintenance of connections.

Modularity lies on the mesoscale and the composition of these modules underpins the brain's communication patterns, promoting a well-balanced and efficient mechanism for integration and segregation.

# Hierarchical Organization of the Brain

The brain has an intrinsic hierarchical structure - from neurons to entire brain regions and systems. Specialized cognitive sub-modules exist within large communities and can dynamically change their coupling with each other based on the cognitive process.



# Representing Brain Connectivity

Alterations to functional and structural connectivity have been observed across the lifespan and in the presence of neurological and psychiatric disorders.

Crucially important to take into account both the modular and hierarchical aspects to brain connectivity to answer relevant scientific questions, including

- (a) network variation among subjects
- (b) network changes during aging
- (c) differences between networks as a consequence of psychological and neurological disorders

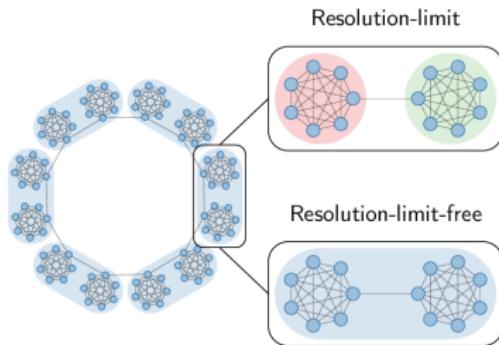
# Research Aim

To address the limitations many community detection algorithms have when applying them to neuroimaging data. It is more accurate to construct a hierarchical tree of communities, where the number and size of these communities can vary. As well, allowing for multimodal data sources to be incorporated can help with the robust creation of these hierarchical community structures.

# Popular Methods: Modularity Maximization

Newman and Girvan (2004):  
partitions network nodes into  
communities in such a way that  
nodes within communities are  
more likely to be grouped  
together than expected in a  
random network null model.

Failure to resolve small  
communities, known as the  
resolution limit, is a limitation of  
these methods.



Background  
oooooooo

Community Detection  
○○●oooooooooooo

Kernel Regression  
oooooooooooo

Collaborative Work  
ooo

# Other Hierarchical Methods

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION  
2020, AHEAD-OF-PRINT, 1-18  
<https://doi.org/10.1080/01621459.2020.1833888>



## Hierarchical Community Detection by Recursive Partitioning

Tianxi Li<sup>a</sup>, Lihua Lei<sup>b</sup>, Sharmodeep Bhattacharyya<sup>c</sup>, Koen Van den Berge<sup>d,e</sup>, Purnamrita Sarkar<sup>f</sup>, Peter J. Bickel<sup>d</sup>, and Elizaveta Levina<sup>g</sup>

PLOS ONE

RESEARCH ARTICLE

## Multi-scale detection of hierarchical community architecture in structural and functional brain networks

Arian Ashourvan<sup>1,2</sup>, Qawi K. Telesford<sup>1,2</sup>, Timothy Verstynen<sup>3</sup>, Jean M. Vettel<sup>1,2,4</sup>, Danielle S. Bassett<sup>1,5,6,7,8\*</sup>

# The Spinglass Algorithm

First introduced by Reichardt and Bornholdt in 2006, relies on the analogy between the Potts model and community detection, where the optimal community configuration is akin to the ground state of the Potts model.

$$\begin{aligned} \mathcal{H}(\{C\}) = & - \sum_{i \neq j} a_{ij} \underbrace{W_{ij} \delta(C_i, C_j)}_{\text{internal links}} + \sum_{i \neq j} b_{ij} \underbrace{(1 - W_{ij}) \delta(C_i, C_j)}_{\text{internal non-links}} \\ & + \sum_{i \neq j} c_{ij} \underbrace{W_{ij} (1 - \delta(C_i, C_j))}_{\text{external links}} - \sum_{i \neq j} d_{ij} \underbrace{(1 - W_{ij}) (1 - \delta(C_i, C_j))}_{\text{external non-links}} \end{aligned}$$

# Rewriting the Spinglass Hamiltonian

The algorithm can be simplified to only consider internal links and non-links, with  $a_{ij} = 1 - \gamma p_{ij}$  and  $b_{ij} = \gamma p_{ij}$ , where  $p_{ij}$  is the probability that a non-zero edge exists between nodes  $i$  and  $j$  normalized to the total weight of the graph.

$$\mathcal{H}(\{C\}) = - \sum_{i \neq j} (W_{ij} - \gamma p_{ij}) \delta(C_i, C_j)$$

Girvan-Newman modularity can be written as a special case of the Potts spinglass model by choice of  $p_{ij}$  and normalization of the Hamiltonian.

# Semi-Supervised Spinglass

Extending on the work of Reichardt and Bornholdt, in 2012 Eaton and Mansbach incorporated external guidance into the community search, creating a semi-supervised spinglass algorithm.

$$\mathcal{H}'(C) = \mathcal{H}(C) + \mu \sum_{i \neq j} (u_{ij} - (\bar{u}_{ij}) \delta(C_i, C_j))$$

where  $u_{ij}$  is the penalty for violating guidance that nodes  $v_i$  and  $v_j$  belong to the same community,  $\bar{u}_{ij}$  is the penalty for violating guidance that  $v_i$  and  $v_j$  belong to different communities, and  $\mu \geq 0$  controls the balance between inherent community structure and external guidance.

# Hierarchical, Multimodal Spinglass

Adopting Eaton and Mansbach's semi-supervised methodology for multimodal community detection in the human brain, our modified Hamiltonian takes the following form:

$$\mathcal{H}'(C) = - \sum_{i \neq j} M_{ij} \delta(C_i, C_j) - \alpha \sum_{i \neq j} S_{ij} \delta(C_i, C_j),$$

where  $M_{ij}$  is the modularity matrix associated with functional connectivity,  $S_{ij}$  is the structural connectivity matrix, and  $\alpha \geq 0$  controls the balance between the community structure determined solely by functional connectivity and being informed by the structural connectivity via white matter fiber tract information.

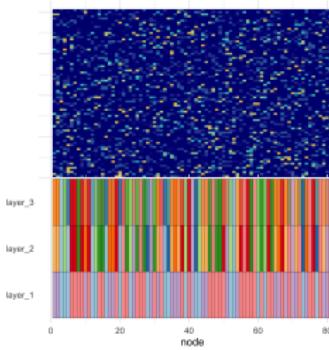
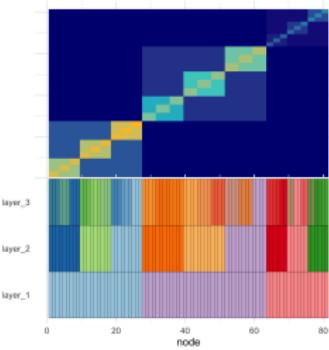
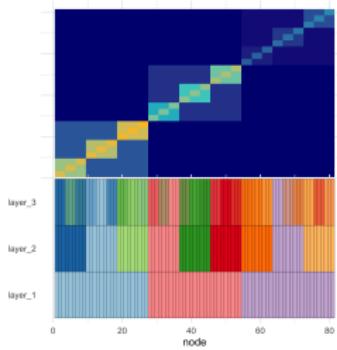
Background  
oooooooo

Community Detection  
oooooooo●oooo

Kernel Regression  
oooooooooooo

Collaborative Work  
ooo

# Hierarchical Community Structure Recovery



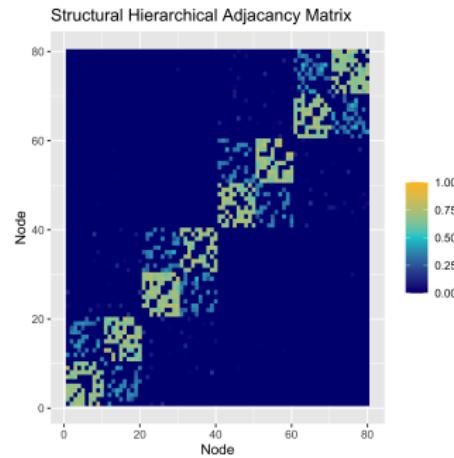
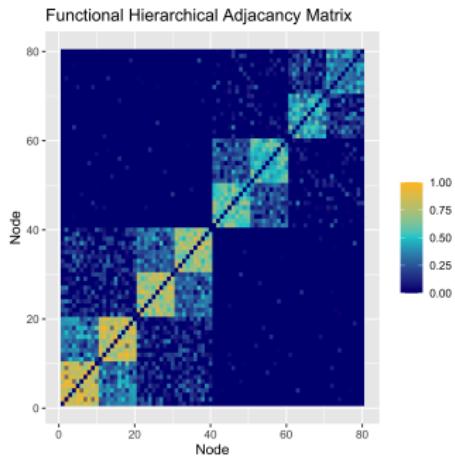
Background  
○○○○○○○

Community Detection  
○○○○○○○●○○○

Kernel Regression  
○○○○○○○○○○○○

Collaborative Work  
○○○

# Modifications of the BTSBM Functions



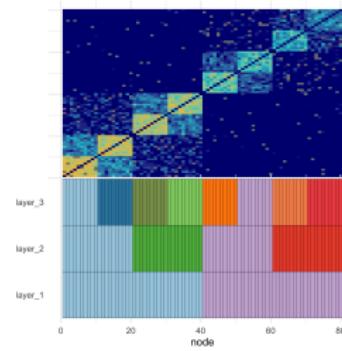
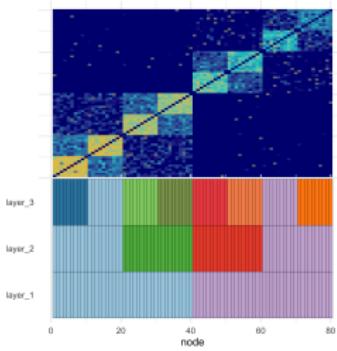
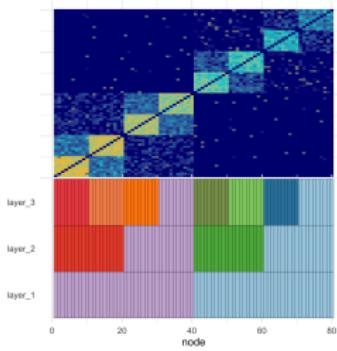
Background  
oooooooo

Community Detection  
oooooooooooo●○○

Kernel Regression  
oooooooooooo

Collaborative Work  
○○○

# More Simulation Results



# Multimodal MRI Dataset

- Deidentified multimodal imaging data from Carnegie Mellon University made publicly available.
  - Streamline count matrices derived from diffusion tractography ( $n = 60$ )
  - fMRI BOLD time series ( $n = 53$ )
  - Limited demographics data
- Regions of interest derived from an upsampled version of the 90-region AAL atlas, resulting in 600 regions (average area 268 voxels with standard deviation of 35 voxels).

Background  
oooooooo

Community Detection  
oooooooooooo●

Kernel Regression  
oooooooooooo

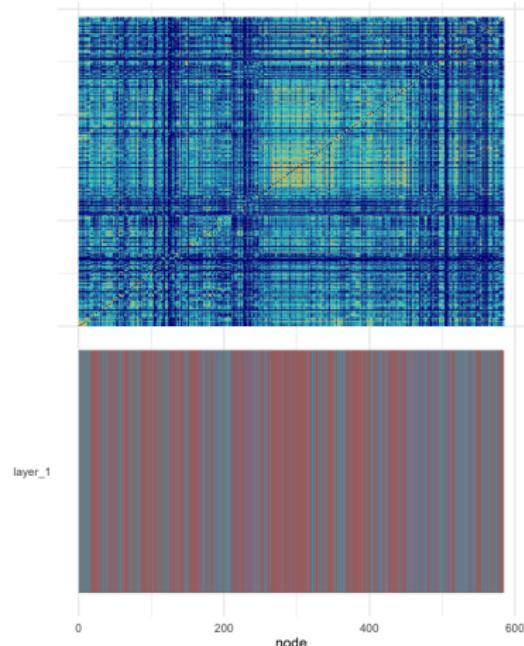
Collaborative Work  
ooo

# CMU Data Analysis

Preliminary data for functional connectivity only is available.  
More results are currently being processed!

Single subject consensus communities show that the optimal number of spins is 3-4.

The normalized mutual information (NMI) between multiple runs of the algorithm on the same subject is  $\approx 0.65$ .



# Research Aim

Kernels provide an efficient and generic approach to deal with structured data. Through their incorporation into a semiparametric regression framework, the goal is to address the question of whether there is an association between tree structures and a phenotype, while controlling for other covariates.

# Distance-Based Kernels

$$k_d(x, y) = \exp \left\{ -\frac{d^2(x, y)}{\rho} \right\}$$

- $d(x, y)$  is a distance function
- $\rho$  is an unknown bandwidth/scaling parameter

$k_d(i, j)$  is a measure of similarity between two subjects  $i$  and  $j$  in terms of some multi-dimensional variable set,  $Z$

- A kernel can be incorporated into a regression framework to test to what extent variation in  $Z$  can explain variation in a binary outcome variable,  $Y$

# Resistance Perturbation Distance Formula

Effective resistance:

$$R = \text{diag}(L^\dagger)1^T + 1\text{diag}(L^\dagger)^T - 2L^\dagger, \text{ where } \text{diag}(L^\dagger) = \begin{bmatrix} L_{11}^\dagger \\ \vdots \\ L_{nn}^\dagger \end{bmatrix}$$

Resistance perturbation distance for  $1 \leq p < \infty$

$$\begin{aligned} d_{rp(p)}(G^{(1)}, G^{(2)}) &= \| R^{(1)} - R^{(2)} \|_p \\ &= \left[ \sum_{i,j \in V} |R_{ij}^{(1)} - R_{ij}^{(2)}|^p \right]^{\frac{1}{p}} \end{aligned}$$

# Kernel-Based Score Test: Framework

Assume  $\text{logit} [\Pr(y_i = 1)] = X_i \beta_i + g(Z_i) + \epsilon_i$  for the logistic framework

- $g(\cdot)$  is a centered, smooth function in a reproducing kernel Hilbert space
- $Z_i$  is the vectorized form of a distance or similarity matrix
- A differentially-expressed multi-dimensional variable set  $Z$  would lead to rejecting the null hypothesis  $H_0 : g(\cdot) = 0$

# Kernel-Based Score Test: Framework

Model corresponds to a penalized quasi-likelihood:

$$\ell[\beta, g(\cdot)] = \sum_{i=1}^n [Y_i(X_i^T \beta + g(Z_i)) - \log(1 + \exp\{X_i^T \beta + g(Z_i)\})] - \frac{\lambda}{2} \|g\|_{\mathcal{H}_k}^2,$$

where  $\lambda$  is a parameter that controls the trade off between goodness-of-fit and model complexity and  $\rho$  controls the smoothness of  $g(\cdot)$

The penalized likelihood can also be approached from a GzLMM perspective as logistic regression is a special case and a slight modification to the penalized log likelihood allows for testing of the null hypothesis  $H_0 : \tau = 1/\lambda = 0$  without explicit specification of the basis functions of  $g(\cdot)$

## Kernel-Based Score Test: Test Statistic

However, under the null hypothesis, the kernel matrix disappears, making  $\rho$  a special type of nuisance parameter that is inestimable under the null hypothesis

- Davies (*Biometrika* 1977, 1987) studied this issue and created a score statistic that treated the nuisance parameter like a parameter-indexed Gaussian process
- Maximization of the penalized likelihood in terms of the parameters using a Newton-Raphson algorithm

The score test for  $H_0 : \tau = 1/\lambda = 0$  can be written as

$$S(\rho) = \frac{Q_\tau(\hat{\beta}_0, \rho) - \mu_Q}{\sigma_Q}$$

## Kernel-Based Score Test: P-Value

As we do not know the distribution of this test statistic, we will apply Davies' results to get an upper bound for the score test's p-value

- Large values of  $Q_\tau(\hat{\beta}_0, \rho)$  would result in a rejection of  $H_0$

$$W \exp \left\{ -\frac{1}{2} M^2 \right\}$$

Upper bound of the p-value is  $\Phi(-M) + \frac{W}{\sqrt{8\pi}}$

- $M$  is the maximum of  $S(\rho)$  over all of the searched space of  $\rho$
- $W = |S(\rho_1) - S(L)| + |S(\rho_2) - S(\rho_1)| + \dots + |S(U) - S(\rho_m)|$
- $L = 0.1 \min_{i \neq j} \sum_{l=1}^p (z_{il} - z_{jl})^2$  and  $U = 100 \max_{i \neq j} \sum_{l=1}^p (z_{il} - z_{jl})^2$

# COBRE-I Dataset

Center for Biomedical Research Excellence (COBRE-I) made raw anatomical and functional MR data publicly available on 147 subjects through the 1000 Functional Connectomes Project

- 72 subjects had diagnosed schizophrenia and 75 were healthy controls
- Phenotypic data on sex (F/M), age (18-65), handedness (LH/RH/Ambidextrous), and diagnostic information (for cases only) was provided for every subject

# Full COBRE-I Dataset Analysis

- 69 schizophrenic and 72 control patients
- Four scenarios were tested to look at the impact of inclusion of demographic variables in the regression and the treatment of negative correlations within the connectivity matrices

Regression		Score	Test	P-Value
Semiparametric	Zero Negative Correlation			0.53
	Keeping Negative Correlations			0.61
Fully Non-Parametric	Zero Negative Correlation			0.02
	Keeping Negative Correlations			0.48

# Comparison Analyses

As a comparison, graph theoretic measures were calculated for each COBRE-I subject's connectivity matrix and included in a conventional logistic regression framework, with group as the outcome of interest and age, sex, and handedness included as additional covariates of interest.

Graph Theoretic Measure	Wald Test P-Value
Global Efficiency	0.06
Rich Club Coefficient	0.35

# Current Work

In the process of writing the code for five different tree kernels:

- Tree edit distance
- Subtree kernel
- Subset tree kernel
- LabelGram kernel
- Tree Fisher kernel

Simulation studies will be conducted to understand each kernel regression's power and type I error, as well as the number of simulated trials that give power  $< 70\%$  or type I error  $> 0.05$ .

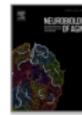
Real data analysis will be conducted using the Autism Brain Imaging Data Exchange (ABIDE) dataset.

# Colorado Aging Brain Lab

- Created a pre-processing pipeline for cross-sectional analysis of T1 scans, gathering ROI-based grey matter volumes.
- Wrote a Shiny app that communicates with RedCap databases to produce realtime information on the study cohort.
- Currently working on the analysis for a project that is investigating the association between neuritic orientation and dispersion indices and AD blood biomarkers.



Neurobiology of Aging  
Volume 103, July 2021, Pages 68-77



---

Astrogliosis and episodic memory in late life:  
higher GFAP is related to worse memory and  
white matter microstructure in healthy aging and  
Alzheimer's disease

# Translational Outcomes in Radiology Research Lab

- GERRAF: Ethnoracial disparities in access to transjugular intrahepatic portosystemic shunt (TIPS) creation and hospital survival for acute variceal bleeding in the United States.
- Compliance with imaging guidelines for workup of transient ischemic attack (TIA) within emergency departments.
- Submitted grants to AHA and NIH related to inferior vena cava filter (IVCF) retrieval.

# Thank You

**Members of my dissertation committee:** Dr. Debasish Ghosh, Dr. Katerina Kechris, Dr. Julia Wrobel, Dr. Peter DeWitt, and Dr. Brianne Bettcher

**CIDA:** Dr. Nichole Carlson, Dr. Brianne Bettcher, Dr. Ashesh Thaker, Dr. Premal Trivedi

**And especially everyone who came to listen to my seminar today!**