

Maggot brain, mirror image? A statistical analysis of bilateral symmetry in an insect brain connectome

Benjamin D. Pedigo

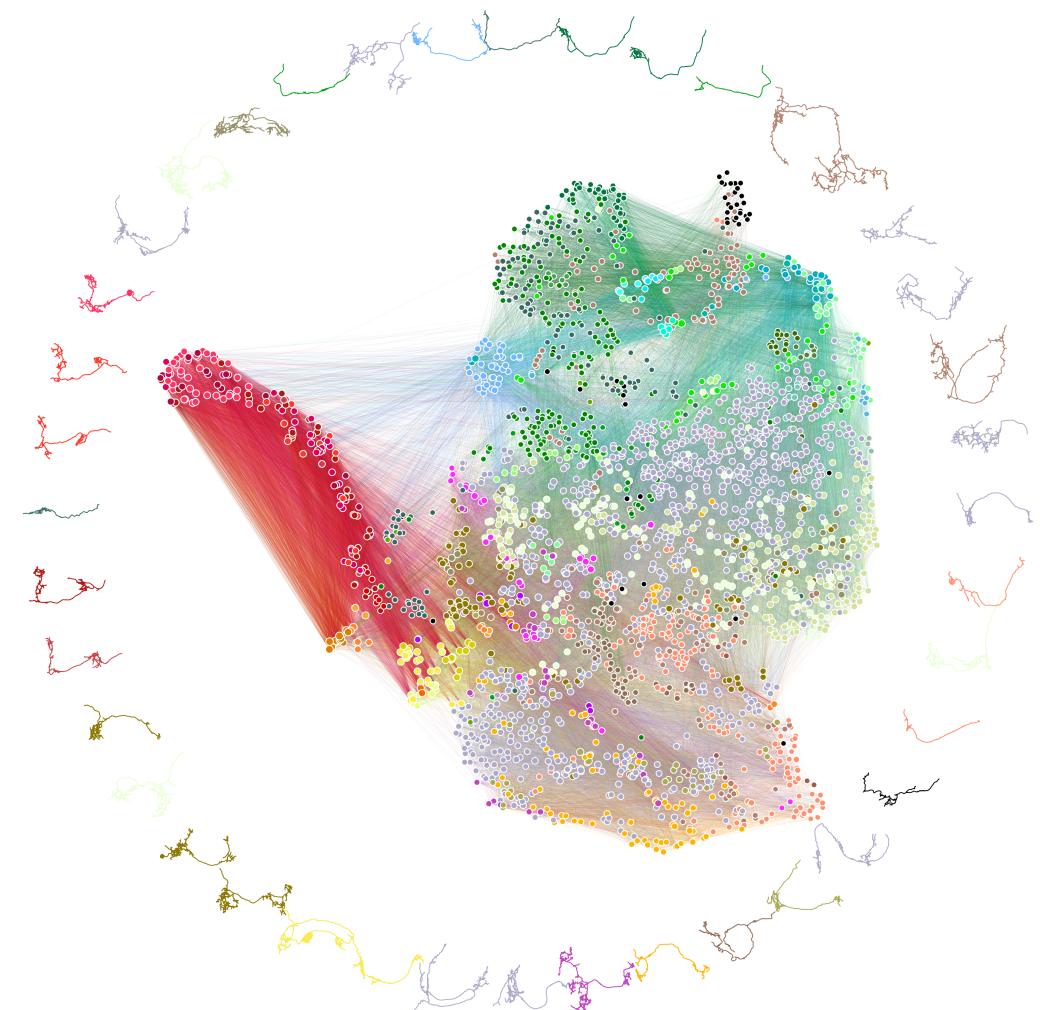
Johns Hopkins University

NeuroData lab

[@bdpedigo \(Github\)](#)

[@bpedigod \(Twitter\)](#)

bpedigo@jhu.edu



What is a connectome? (for *this* talk)

A **connectome** is a network model of brain structure consisting of nodes which represent individual neurons and edges which represent the presence of a synaptic connection between those neurons.

Many connectomics questions require comparison

- Understand wiring substrate of learning/memory
- Understand links between genetics or disease and connectivity
- Understand how different neural architectures lead to different computational abilities

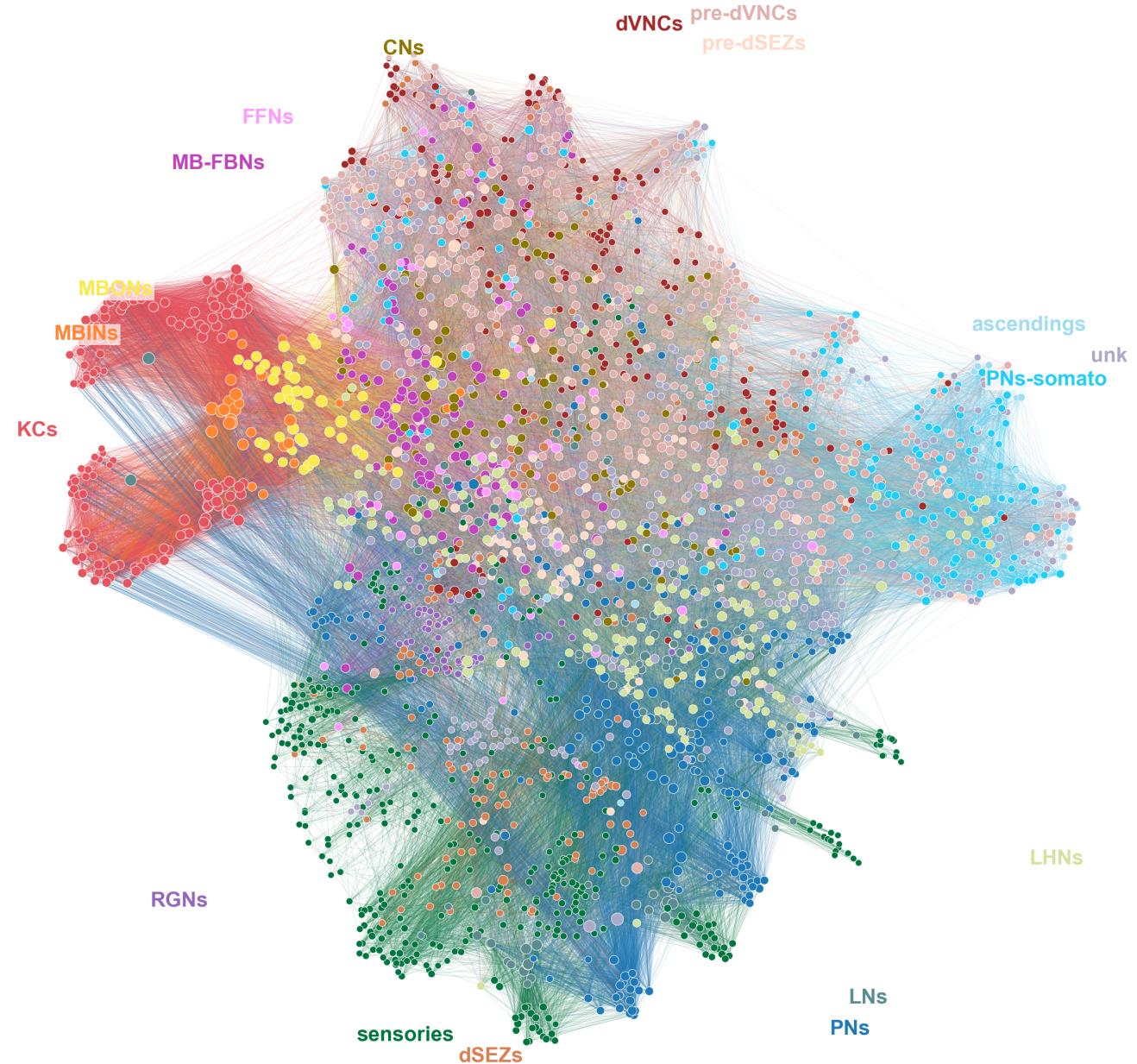
"Understanding statistical regularities and learning which variations are stochastic and which are secondary to an animal's life history will help define the substrate upon which individuality rests and *require comparisons between circuit maps within and between animals.*" [1] (emphasis added)

Larval *Drosophila* brain connectome

- ~2500 brain neurons +
~500 sensory neurons
- ~544K synapses
- Both hemispheres of
the brain reconstructed

See [Michael Windings's talk](#), 11 AM (EST) Dec 2nd

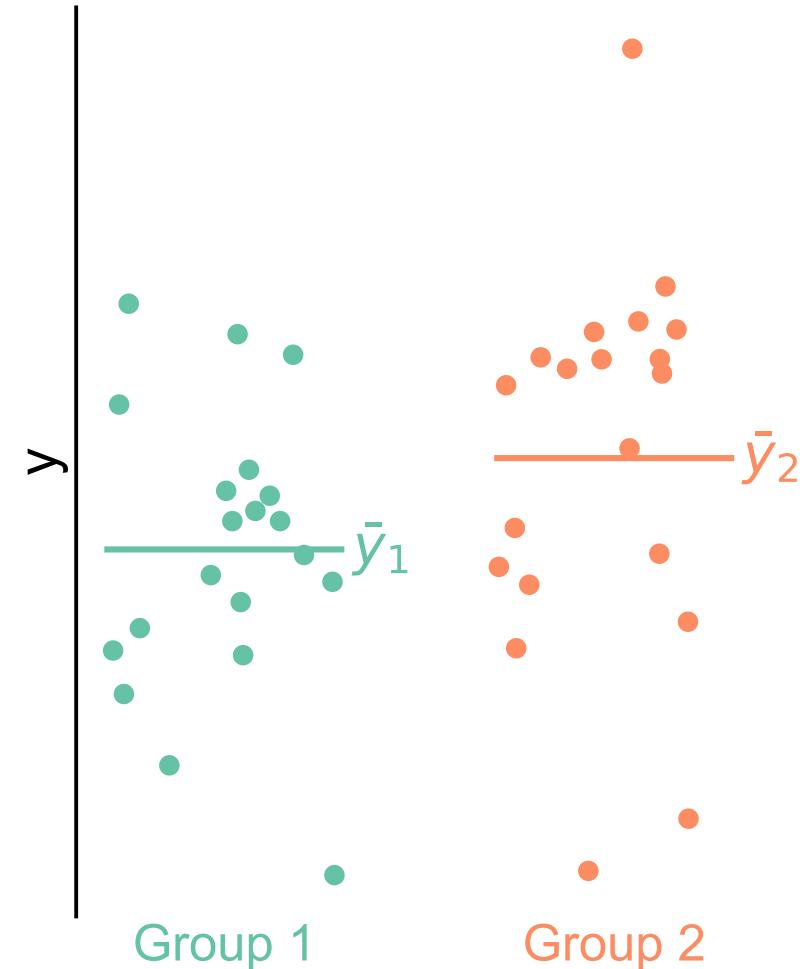
Winding et al. "The complete connectome of an insect brain." In prep (2021)



Are the left and the right sides of this brain *the same?*

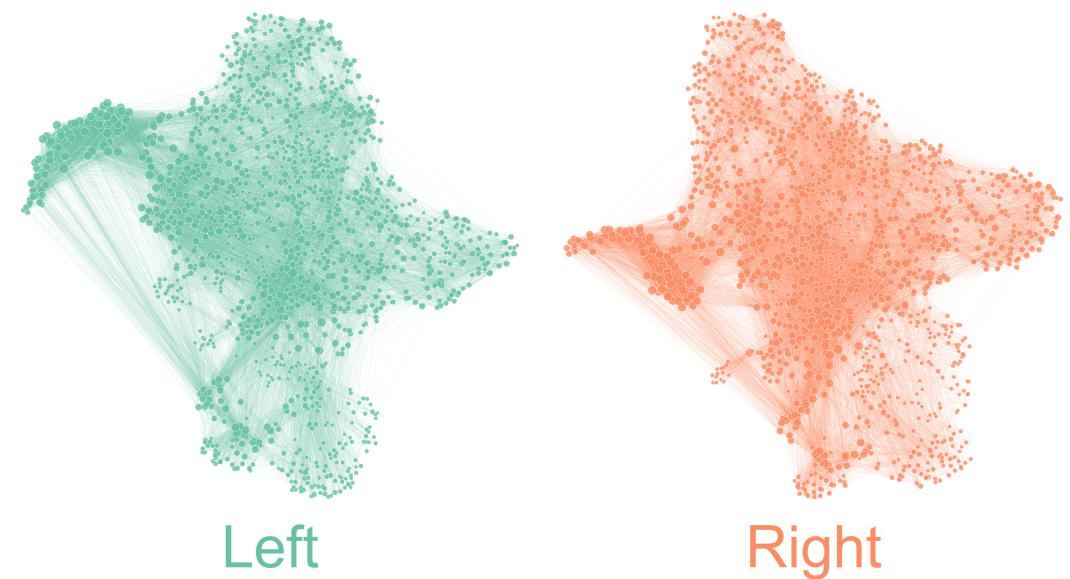
Are these populations the same?

- Known as two-sample testing
- $Y^{(1)} \sim F^{(1)}$, $Y^{(2)} \sim F^{(2)}$
- $H_0 : F^{(1)} = F^{(2)}$
vs.
 $H_A : F^{(1)} \neq F^{(2)}$



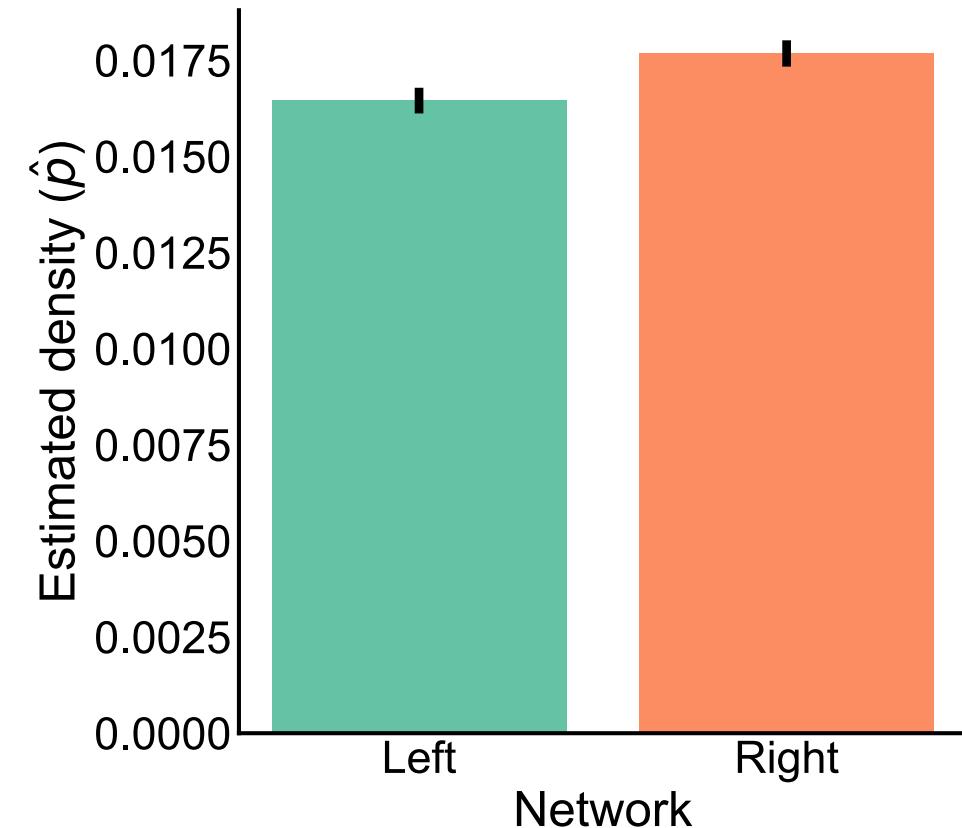
Are these two *networks* the same?

- Left side network: $A^{(L)} \sim F^{(L)}$
- Right side network: $A^{(R)} \sim F^{(R)}$
- $H_0 : F^{(L)} = F^{(R)}$
vs.
 $H_A : F^{(L)} \neq F^{(R)}$



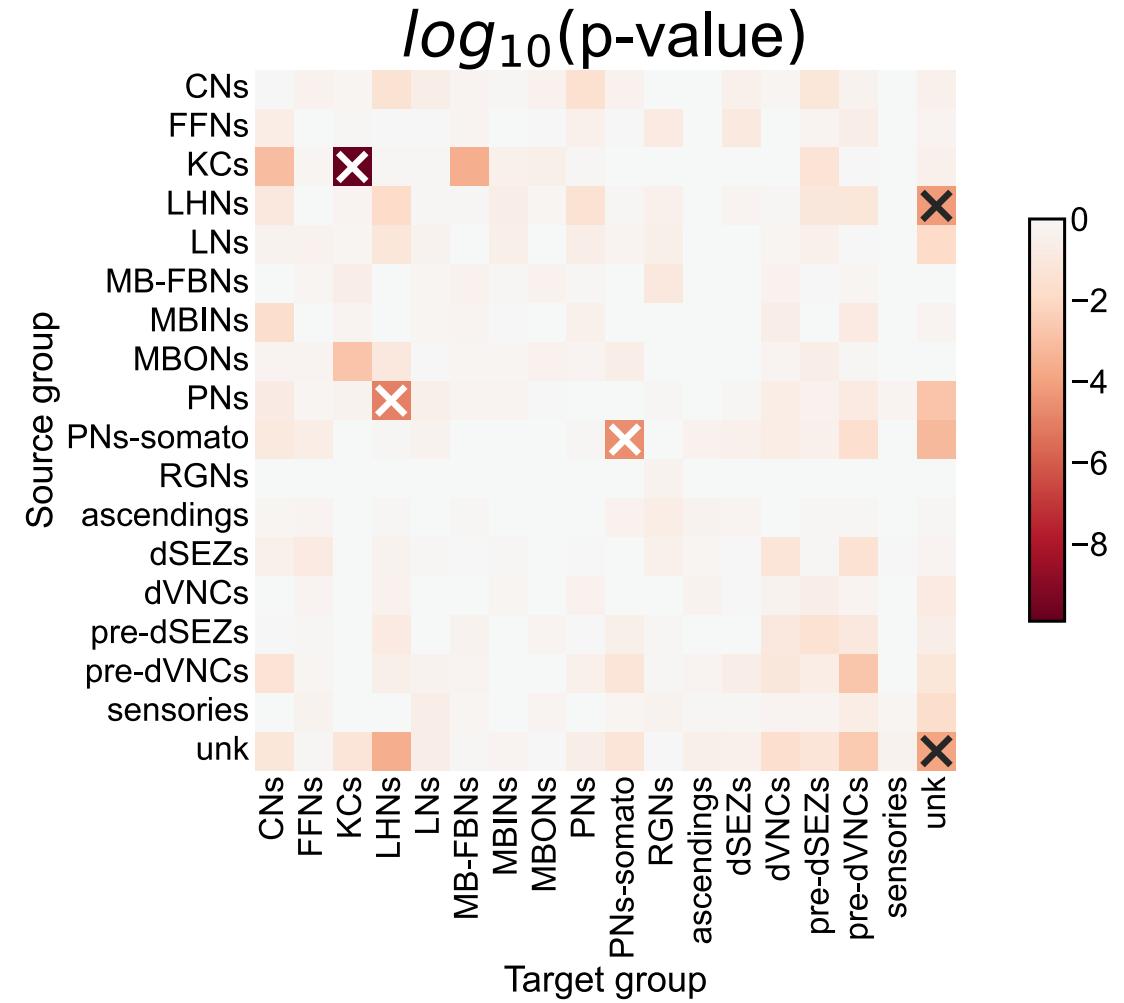
The simplest thing: Erdos-Renyi (ER) model

- Each edge is sampled independently, same connection probability p for all edges
- $A_{ij} \sim Bernoulli(p)$
- Compare $\hat{p}^{(L)}$ vs $\hat{p}^{(R)}$ (binomial test)
- p-value $< 10^{-23}$



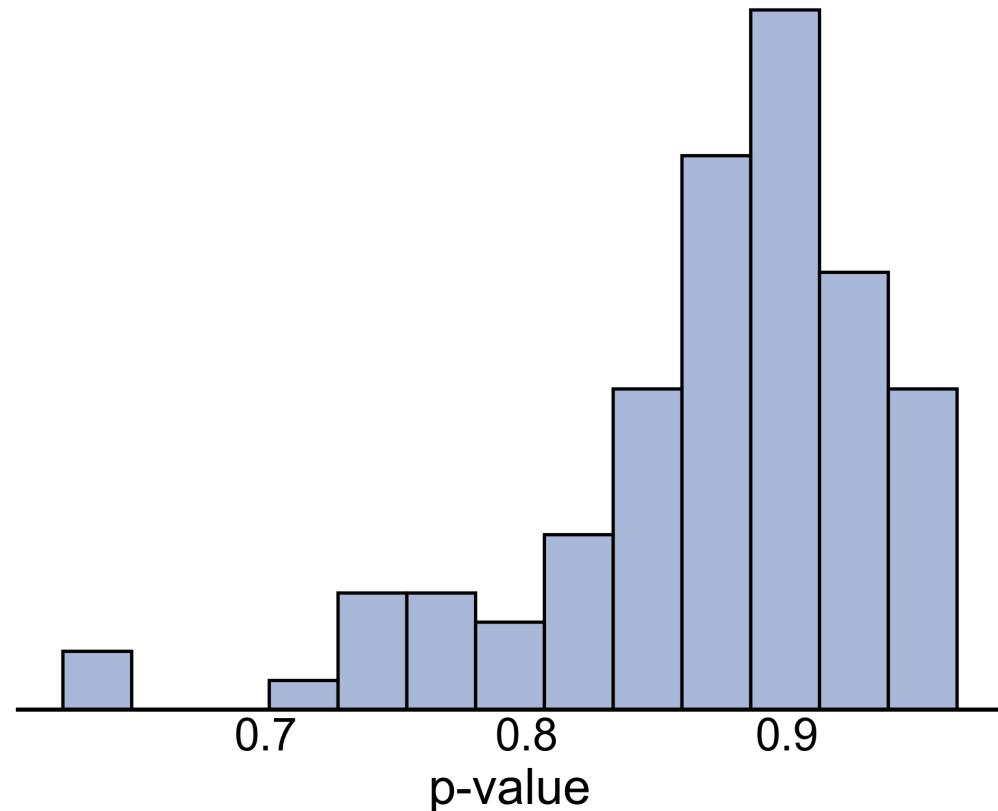
Testing under the stochastic block model (SBM)

- Connections independent, with probability set by the **source node's group** and **target node's group**
- $A_{ij} \sim \text{Bernoulli}(B_{\tau_i, \tau_j})$
- Compare group-to-group connection probabilities:
 $H_0 : B^{(L)} = B^{(R)}$
vs.
 $H_A : B^{(L)} \neq B^{(R)}$
- **p-value** $< 10^{-4}$



Adjusting for a difference in density

- Rejecting $B^{(L)} = B^{(R)}$ can be explained by the difference in density?
- New null hypothesis:
 $H_0 : B^{(L)} = cB^{(R)}$
where c is a density-adjusting constant, $\frac{p^{(L)}}{p^{(R)}}$
- Randomly subsample edges from denser network, rerun test
- p-values > 0.6



More flexibility: Random dot product graph (RDPG)

- Connections independent, probability set by dot product of source node's latent vector and target node's latent vector .
- $A_{ij} \sim \text{Bernoulli}(\langle \mathbf{x}_i, \mathbf{y}_j \rangle)$
- Vectors from distributions $\mathbf{x}_i^{(L)} \sim F^{(L)}$, $\mathbf{x}_i^{(R)} \sim F^{(R)}$
- Compare distributions of latent vectors:
 $H_0 : F^{(L)} = F^{(R)}$
 $H_A : F^{(L)} \neq F^{(R)}$

RDPG results

How sensitive are they?

Summary

- Even for a simple network question like "are the left and right the same," there are many different ways to write down what we mean as a statistical hypothesis
 - Each yields a different test procedure
 - Each test is sensitive to varying alternatives
- We may care about some differences and not others (e.g. density) and any test will need to adjust for that
- These techniques apply anytime one wants to compare connectomes/networks

graspologic:

github.com/microsoft/graspologic



This work:

github.com/neurodata/bilateral-connectome

The screenshot shows a page from a book titled "Bilateral Connectome" by NEURODATA. The title of the page is "The Erdos-Renyi (ER) model". The text explains that the Erdos-Renyi (ER) model is one of the simplest network models, where all edges between any two nodes are equally likely. A "Math" section provides the formula $P[A_{ij} = 1] = p_{ij} = p$ for the probability of an edge occurring. The page also includes links to "Abstract", "PRELIMINARIES", and other chapters like "Introduction" and "Outline".

Acknowledgements

Johns Hopkins University

Joshua Vogelstein, Carey Priebe, Mike Powell, Eric Bridgeford, Kareef Ullah, Diane Lee, Sambit Panda, Jaewon Chung, Ali Saad-Eldin

University of Cambridge / Laboratory of Molecular Biology

Michael Winding, Albert Cardona, Marta Zlatic, Chris Barnes

Microsoft Research

Hayden Helm, Dax Pryce, Nick Caurvina, Bryan Tower, Patrick Bourke, Jonathan McLean, Carolyn Buractaon, Amber Hoak

Questions?