

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

Benjamin D. Pedigo

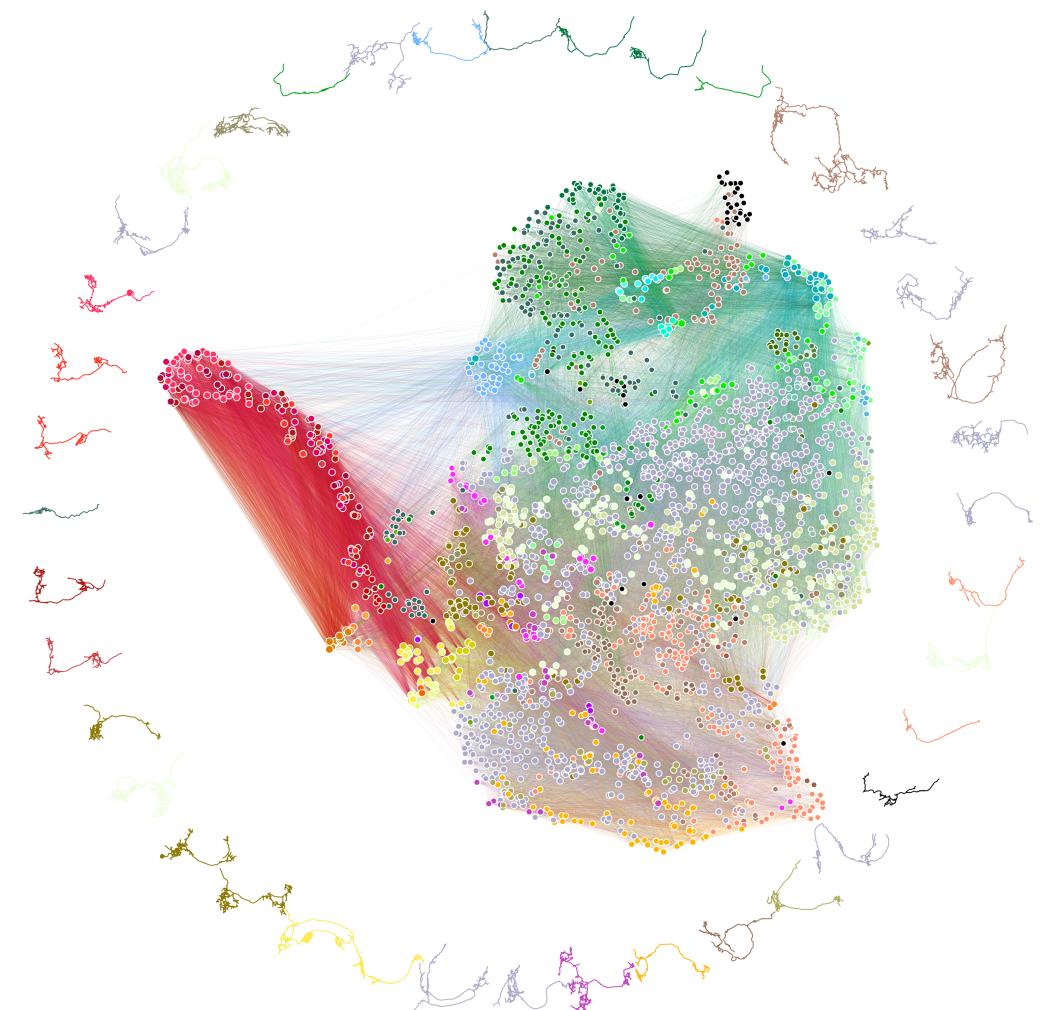
*Johns Hopkins University*

*NeuroData lab*

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

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

[bpedigo@jhu.edu](mailto:bpedigo@jhu.edu)



# Many connectomics questions require comparison

For instance,

- Understand connectomes across evolution [1]
- Understand connectomes across development [2]
- Understand links between genetics and connectivity [3]

[1] Bartsotti + Correia et al. *Curr. Op. Neurobiology* (2021)

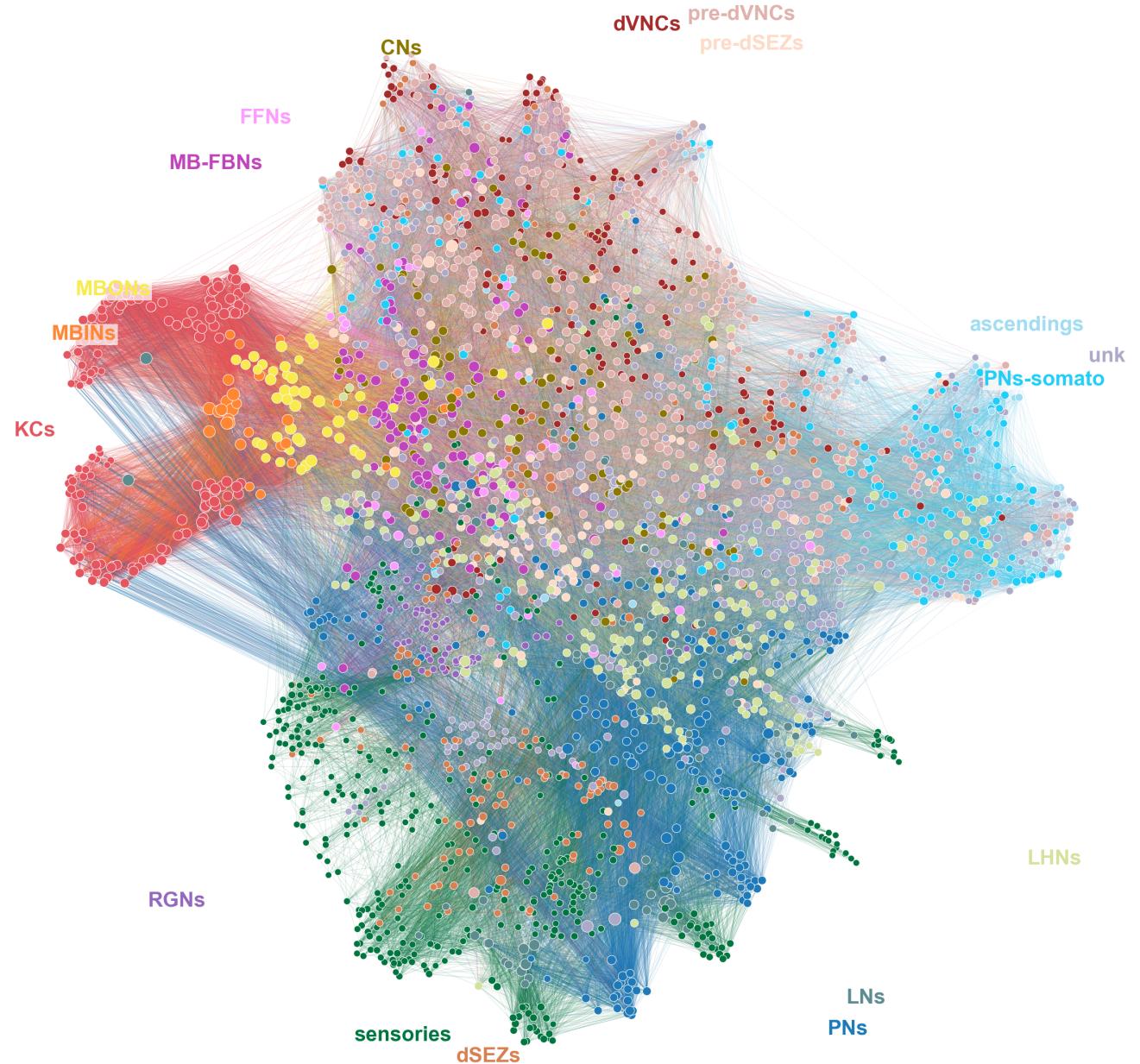
[2] Witvliet et al. *Nature* (2021)

[3] Valdes-Aleman et al. *Neuron* (2021)

# Larval *Drosophila* brain connectome

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

- ~3000 neurons,
- ~544K synapses
- Both hemispheres of the brain reconstructed

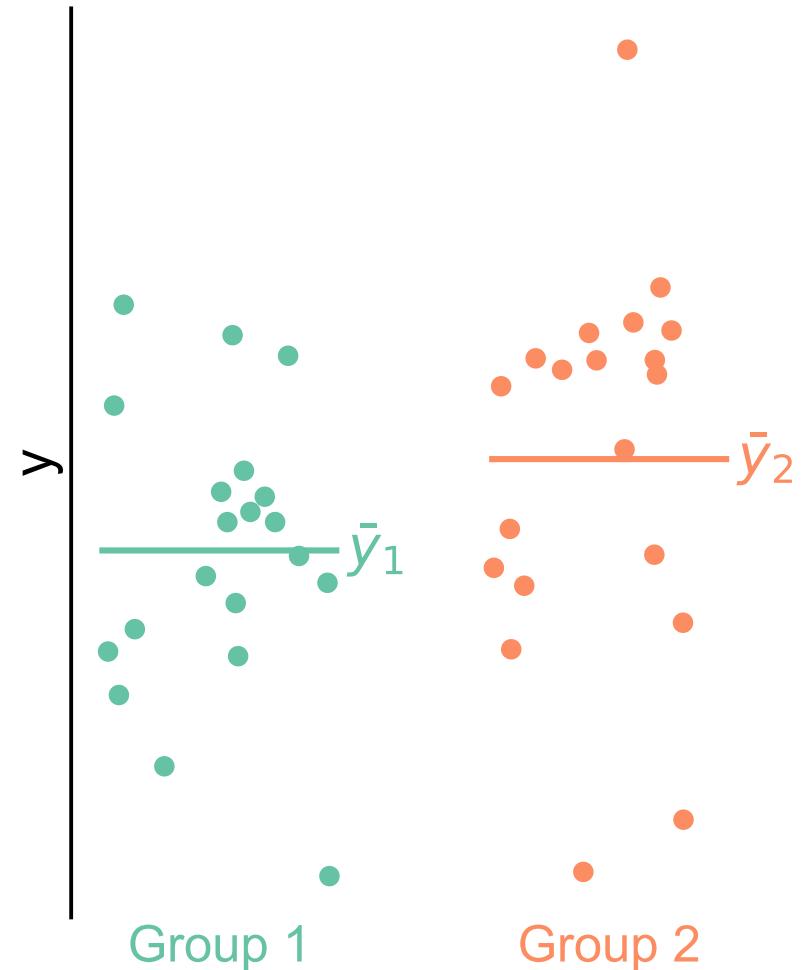


Winding et al. "The complete connectome of an insect brain." In prep (2021)

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

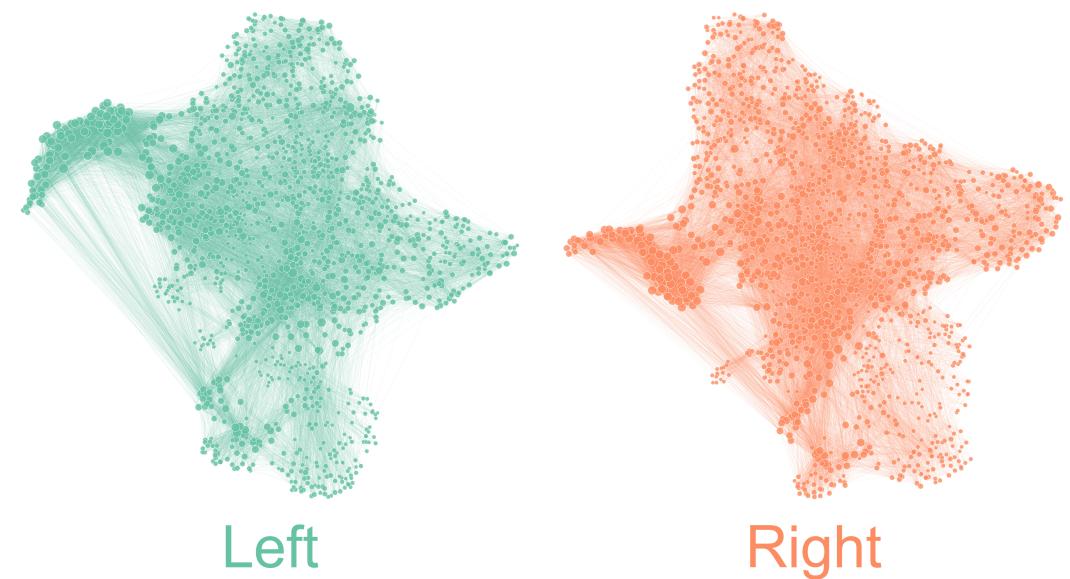
# Are these populations the same?

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



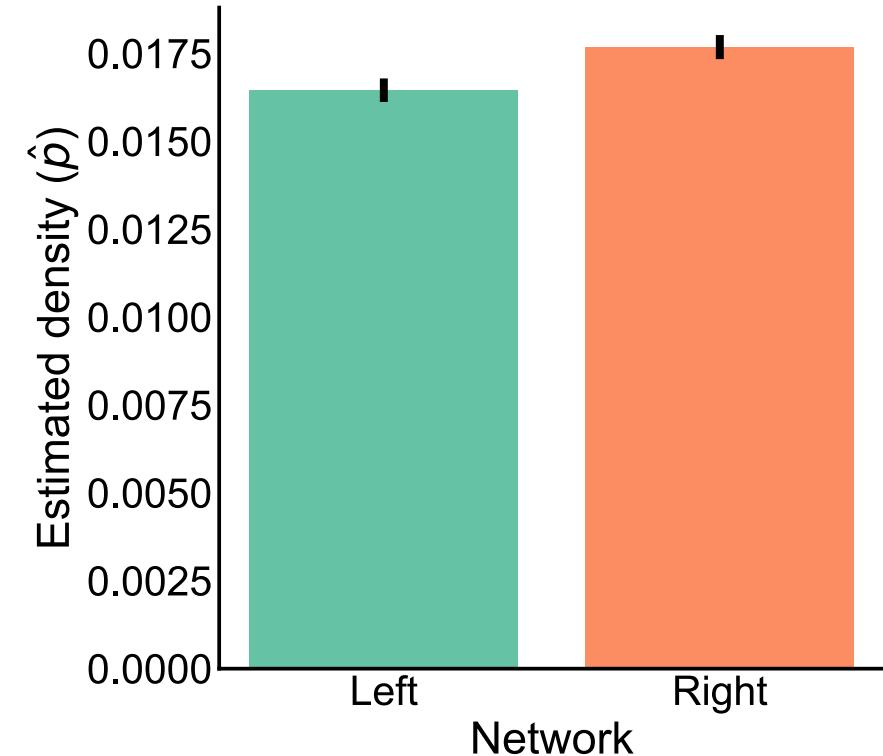
# Are these two *networks* the same?

- Want a two-network-sample test!
- For simplicity (for now), consider networks to be *directed, unweighted*.
- For simplicity (for now), consider the **left  $\rightarrow$  left** and **right  $\rightarrow$  right** (ipsilateral) connections only.
- $A^{(L)} \sim F^{(L)}, A^{(R)} \sim F^{(R)}$
- $H_0 : F^{(L)} = F^{(R)}$   
 $H_A : F^{(L)} \neq F^{(R)}$



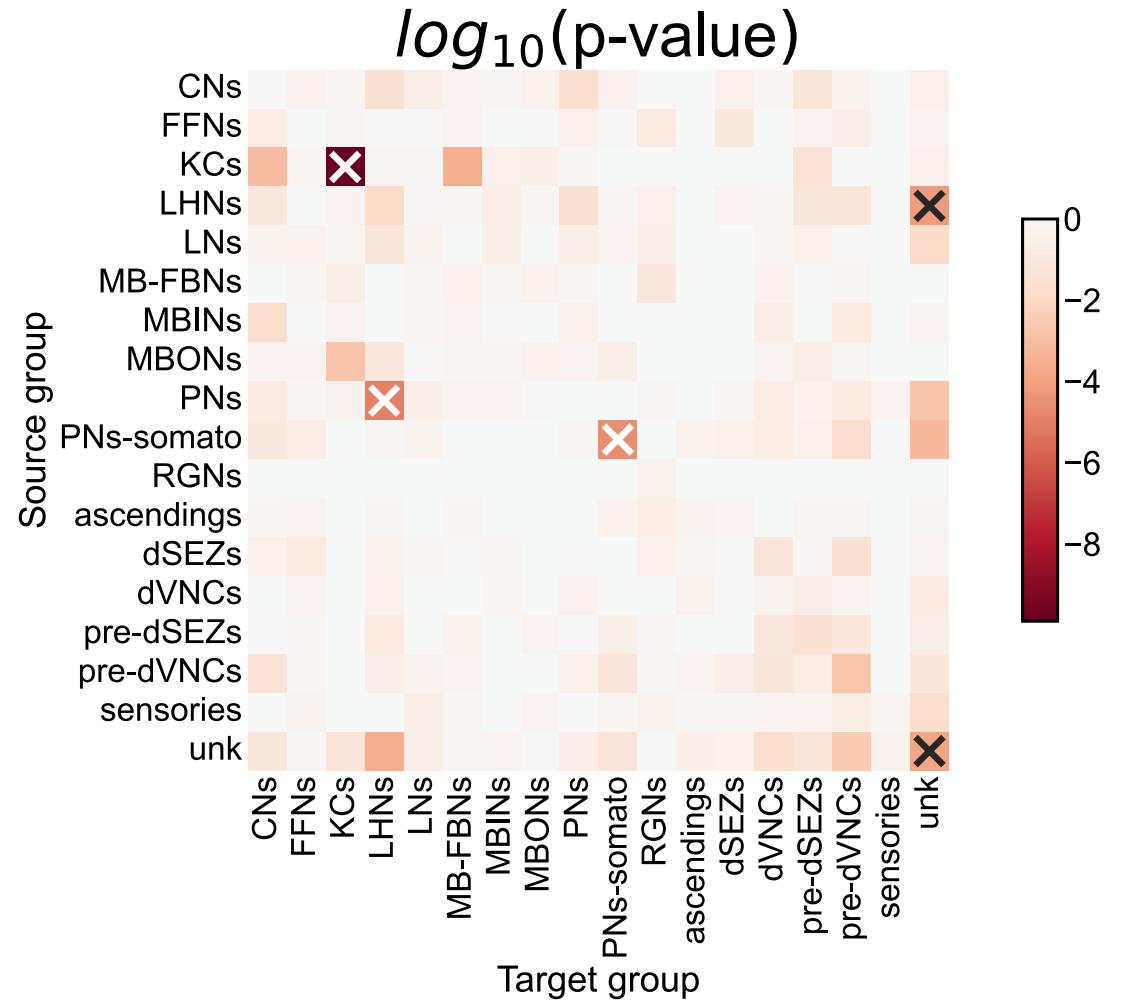
# Density-based testing: Erdos-Renyi (ER) model

- Connections independent, same connection probability  $p$  for all edges
- $P[i \rightarrow j] = p$
- Compare probabilities:  
 $H_0 : p^{(L)} = p^{(R)}$   
 $H_A : p^{(L)} \neq p^{(R)}$
- **p-value**  $< 10^{-23}$
- Is this a difference we care about?



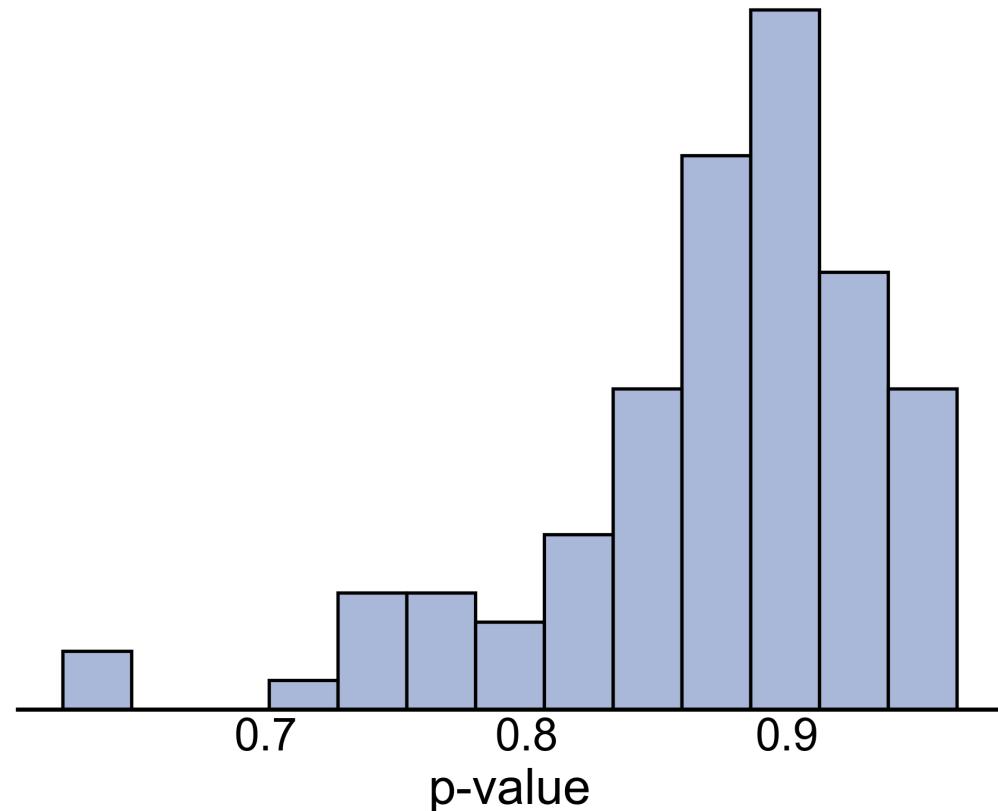
# Group-based testing: stochastic block model (SBM)

- Connection probability set by the source node's group ( $\tau_i$ ) and target node's group ( $\tau_j$ )
- $P[i \rightarrow j] = B_{\tau_i, \tau_j}$
- Compare group-to-group connection probabilities:  
 $H_0 : B^{(L)} = B^{(R)}$   
 $H_A : B^{(L)} \neq B^{(R)}$
- p-value  $< 10^{-4}$



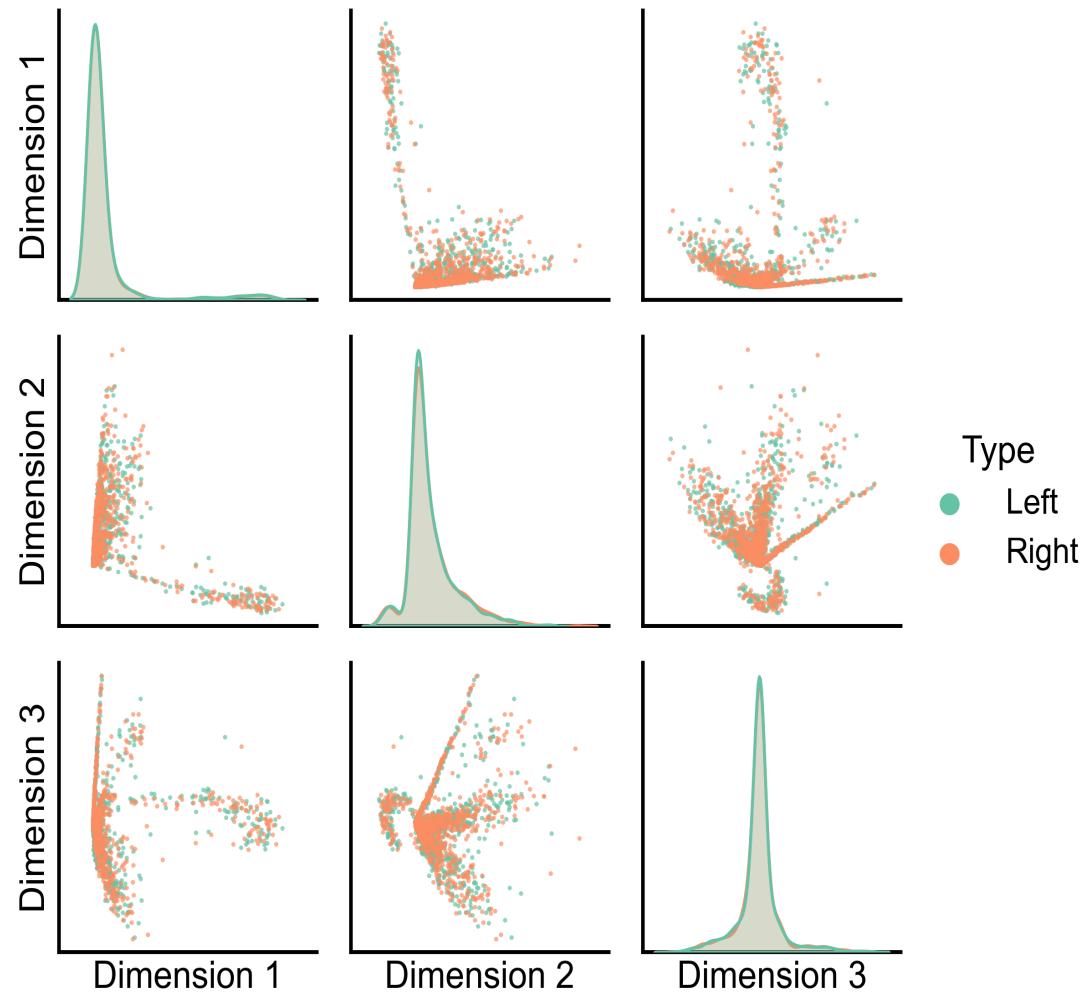
# Adjusting for a difference in density

- 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**  $> 0.6$



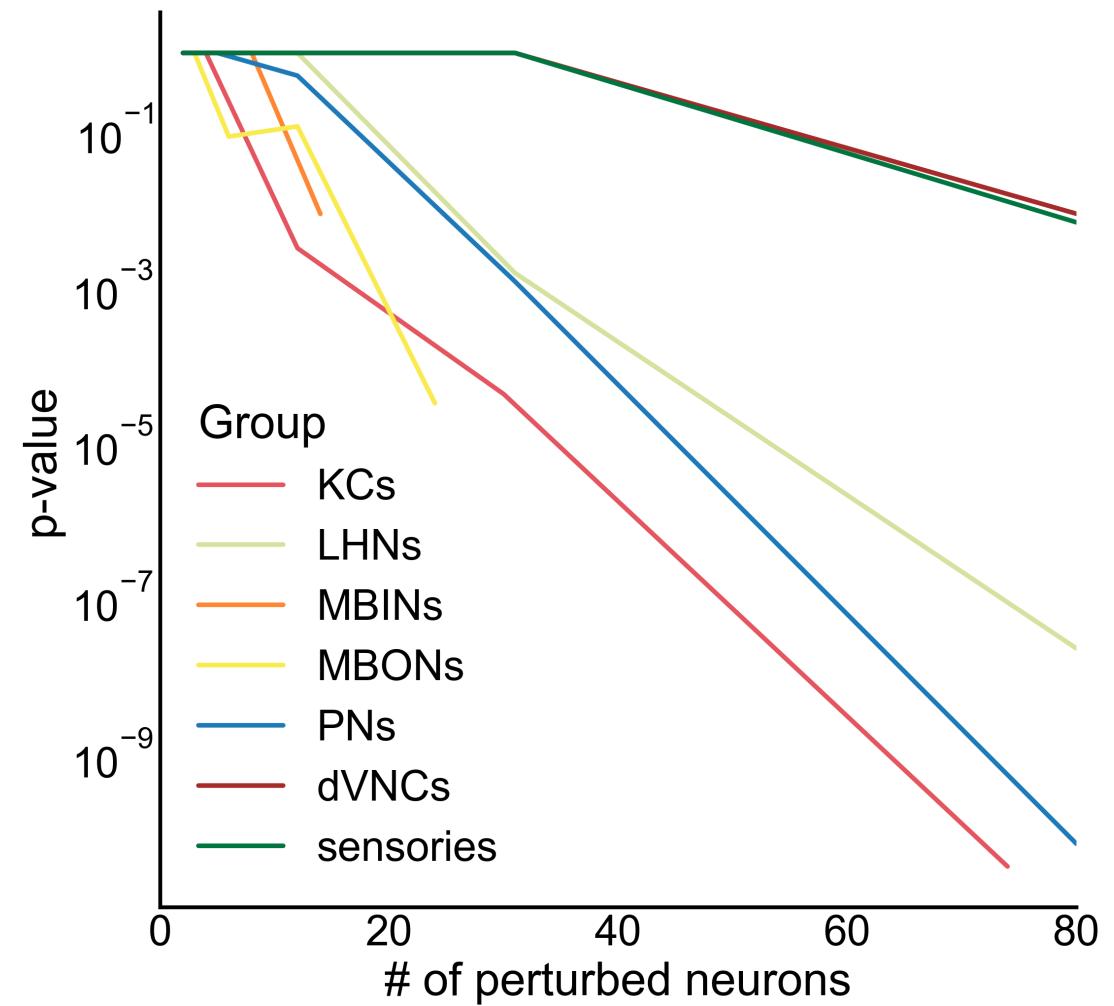
# More flexibility: Random dot product graph (RDPG)

- Probability from dot product of source node's latent vector, target node's latent vector.
- $P[i \rightarrow j] = \langle \mathbf{x}_i, \mathbf{y}_j \rangle$
- $\mathbf{x}_i^{(L)} \sim F^{(L)}, \mathbf{x}_i^{(R)} \sim F^{(R)}$  in  $d$  dimensions
- Compare distributions of latent vectors:  
 $H_0 : F^{(L)} = F^{(R)}$   
 $H_A : F^{(L)} \neq F^{(R)}$
- **p-value**  $\approx 1$



# Can we detect differences when we know they exist?

- Make two copies of right hemisphere network
- Apply some perturbation to one of them:
  - Ex: Shuffle edges incident to some number of nodes in some group
- Rerun the RDPG-based test for symmetry



# To sum up...

Model	$H_0$ (vs. $H_A \neq$ )	p-value	Interpretation
ER	$p^{(L)} = p^{(R)}$	$< 10^{-23}$	Reject densities the same
SBM	$B^{(L)} = B^{(R)}$	$< 10^{-4}$	Reject cell type connection probabilities the same
SBM	$B^{(L)} = cB^{(R)}$	$\approx 0.7$	Don't reject the above after density adjustment
RDPG	$F^{(L)} = F^{(R)}$	$\approx 1$	Don't reject latent distributions the same

The answer to this very simple question totally depends on how you frame it!

- Tests are sensitive to some alternatives and not others
- Difference you might not care about (e.g. density) need to be explicitly accounted for

# Future work

- Many other tests (e.g. compare subgraph counts)
- Studying the sets of alternatives each test is/is not sensitive to
- Roadmap for future principled comparisons of connectome networks!

# graspologic:

[github.com/microsoft/graspologic](https://github.com/microsoft/graspologic)



# This work:

[github.com/neurodata/bilateral-connectome](https://github.com/neurodata/bilateral-connectome)

The screenshot shows a documentation page for 'Bilateral Connectome' under the 'NEURODATA' brand. The page includes a search bar, navigation links for 'Abstract', 'PRELIMINARIES', and other sections like 'Introduction', 'Outline', and 'Unmatched vs. matched networks'. A sidebar on the right is titled 'The Erdos-Renyi (ER) model' and contains a 'Math' section with text and equations. The equations include  $P[A_{ij} = 1] = p_{ij} = p$  and the statement 'Where  $p$  is the global connection probability.'

**NEURODATA**

Bilateral Connectome

Search this book...

Abstract

PRELIMINARIES

Introduction

Outline

Unmatched vs. matched networks

Larval *Drosophila melanogaster* brain connectome

**i Math**

Let  $n$  be the number of nodes. We say that for all  $(i, j)$ ,  $i \neq j$ , with  $i$  and  $j$  both running from  $1 \dots n$ , the probability of the edge  $(i, j)$  occurring is:

$$P[A_{ij} = 1] = p_{ij} = p$$

Where  $p$  is the global connection probability.

Each element of the adjacency matrix  $A$  is then

# 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**

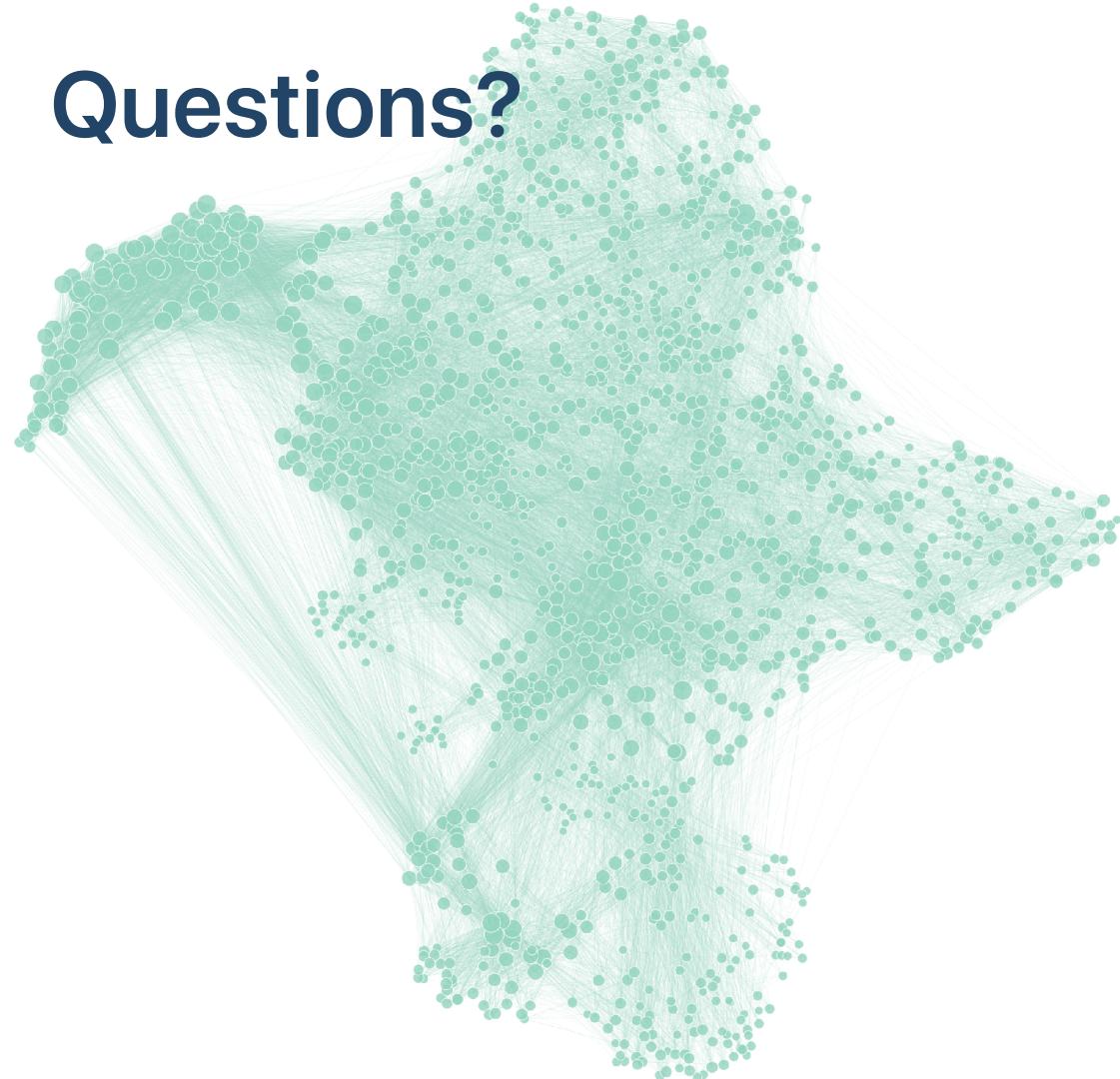
Michael Winding, Albert Cardona, Marta Zlatic, Chris Barnes

## **Microsoft Research**

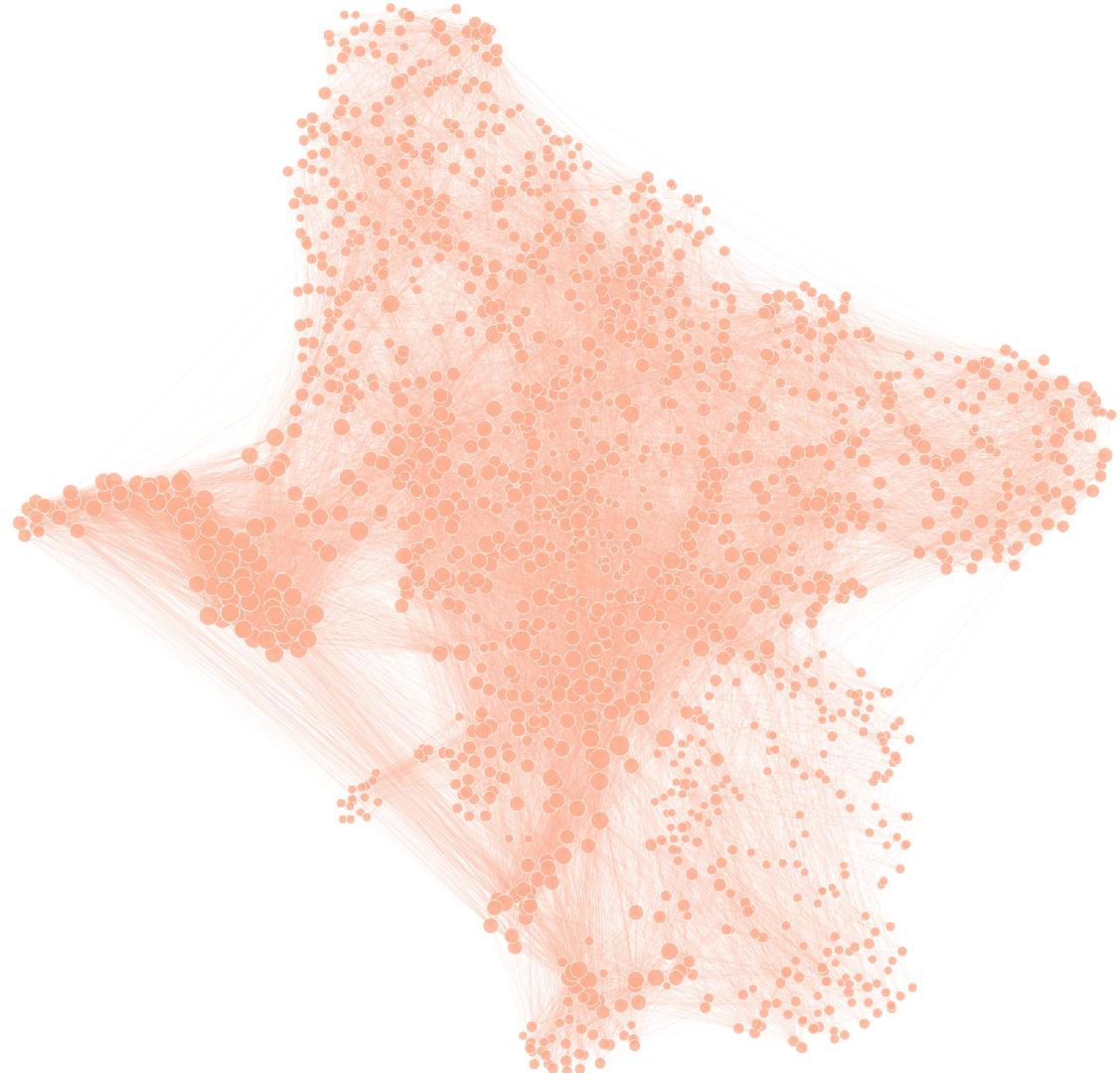
Hayden Helm, Dax Pryce, Nick Caurvina, Bryan Tower, Patrick Bourke, Jonathan McLean, Carolyn Buractaon, Amber Hoak

## **NMC organizers!**

# Questions?



Left



Right