

Hypothesis testing for connectome comparisons: a statistical analysis of bilateral symmetry in an insect brain connectome

Benjamin D. Pedigo

(he/him) - *NeuroData lab*

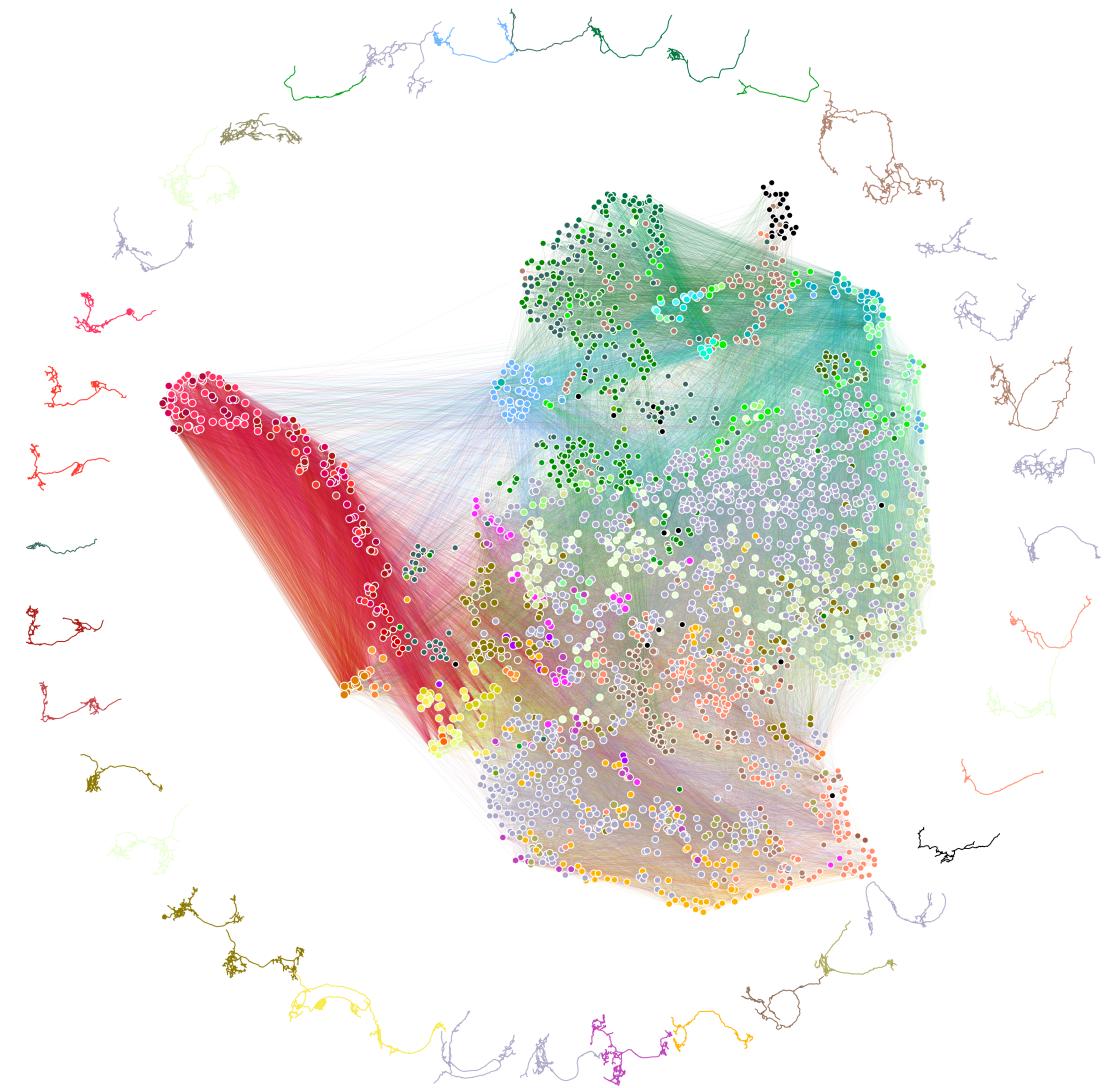
Johns Hopkins University - Biomedical Engineering

 bpedigo@jhu.edu

 [@bdpedigo \(Github\)](https://github.com/bdpedigo)

 [@bpedigod \(Twitter\)](https://twitter.com/bpedigod)

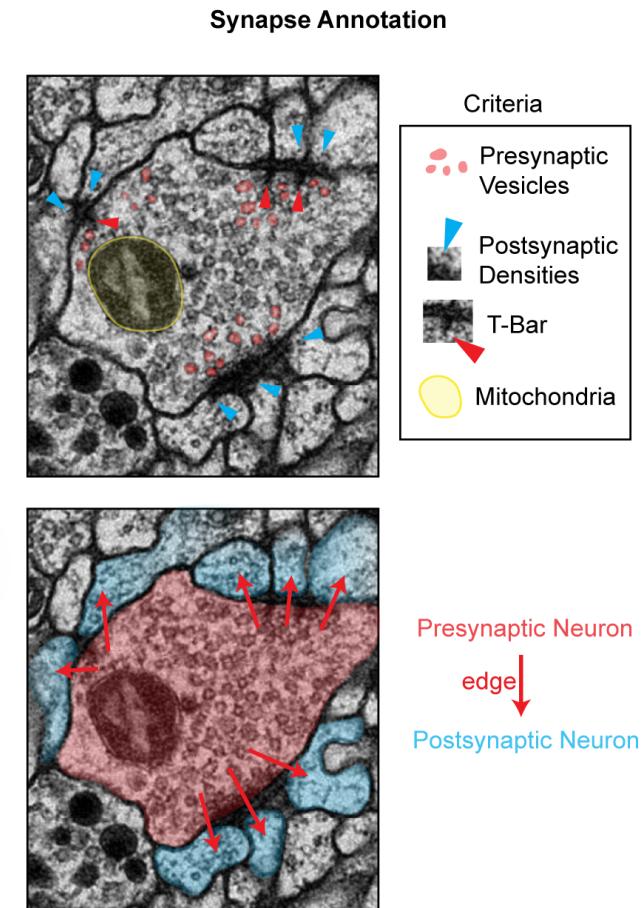
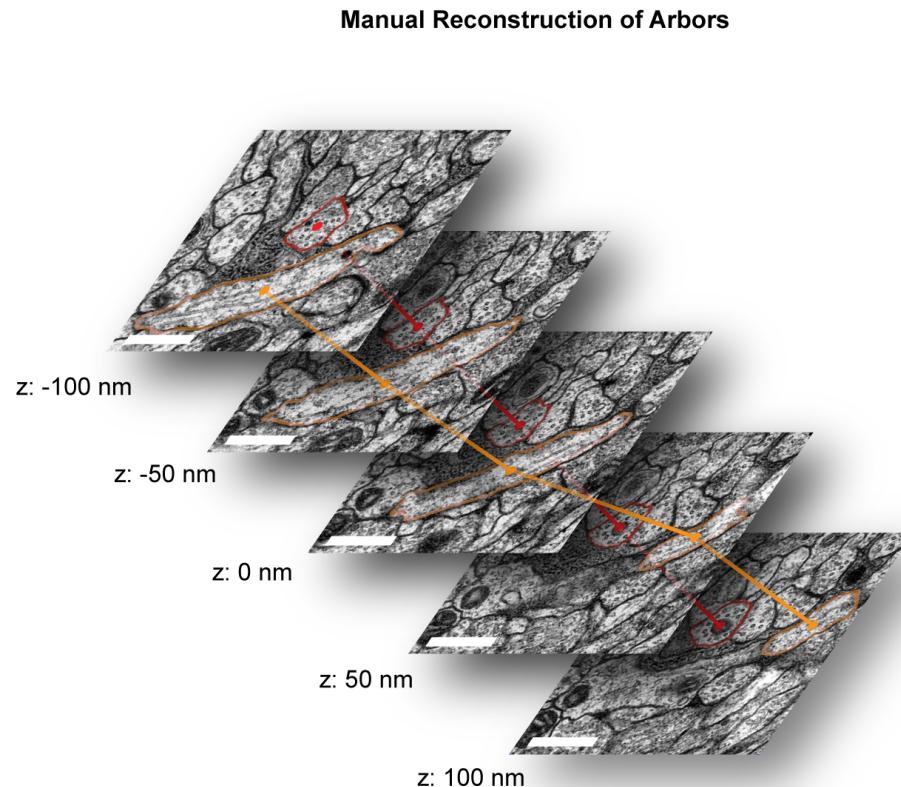
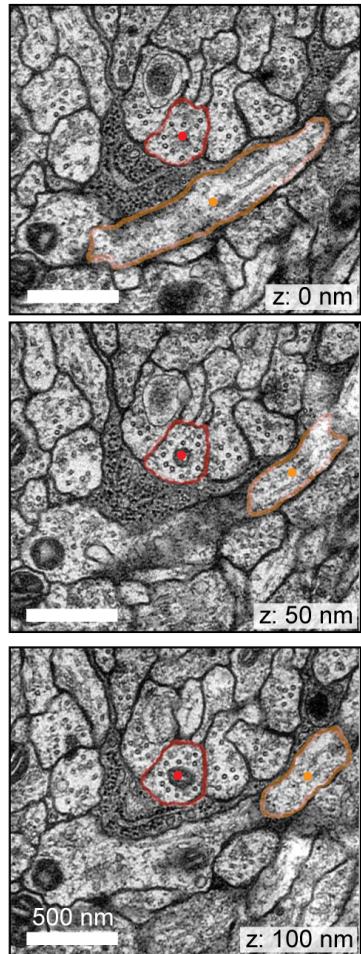
 <https://bpedigo.github.io/>



Outline

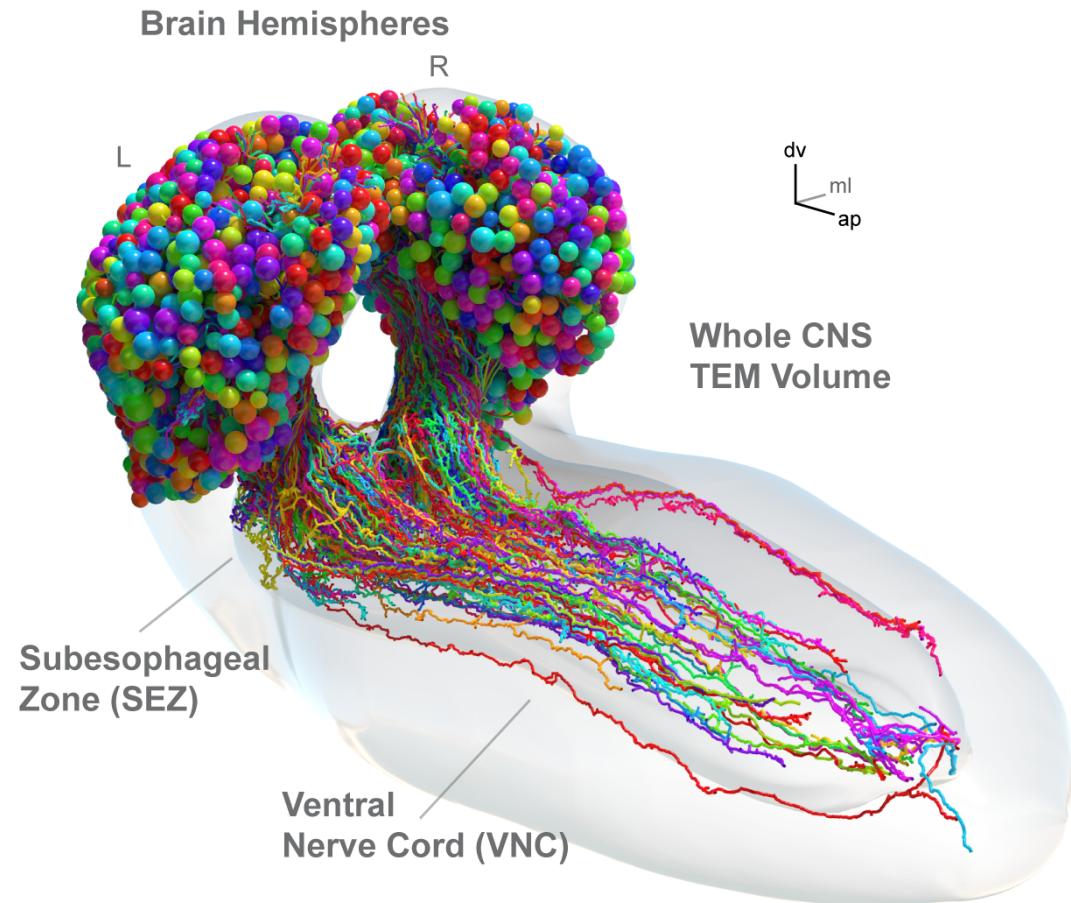
- What is electron microscopy connectomics
- *Drosophila* larva brain connectome
- Why we should care about comparing connectomes
- Bilateral symmetry
- Extensions and other tools

Electron microscopy connectomics

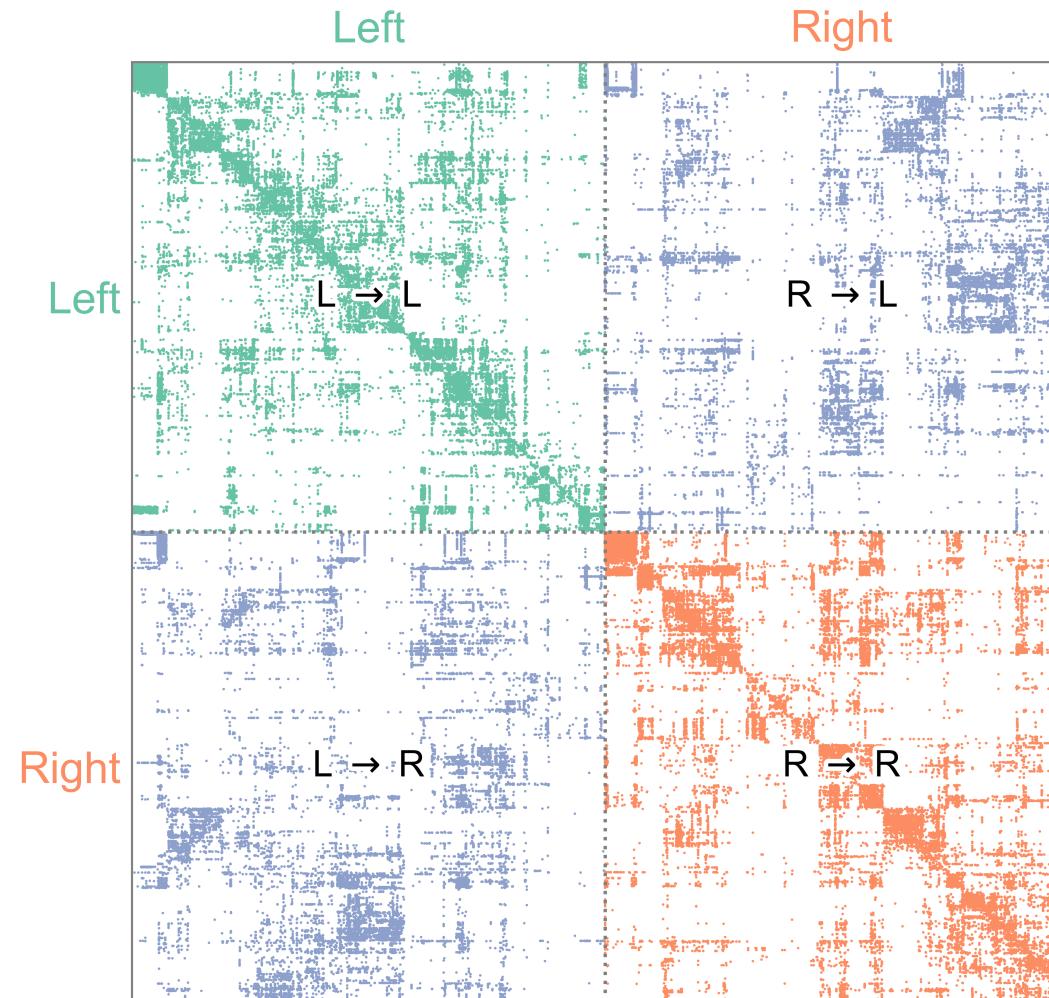


Drosophila larva (AKA a maggot) brain connectome

- Collaboration with Marta Zlatic/Albert Cardona's groups - led by Michael Winding
- First whole-brain, single-cell connectome of any insect
- ~3000 neurons, ~550K synapses
- Both hemispheres of the brain reconstructed



We're just going to consider this to be a network



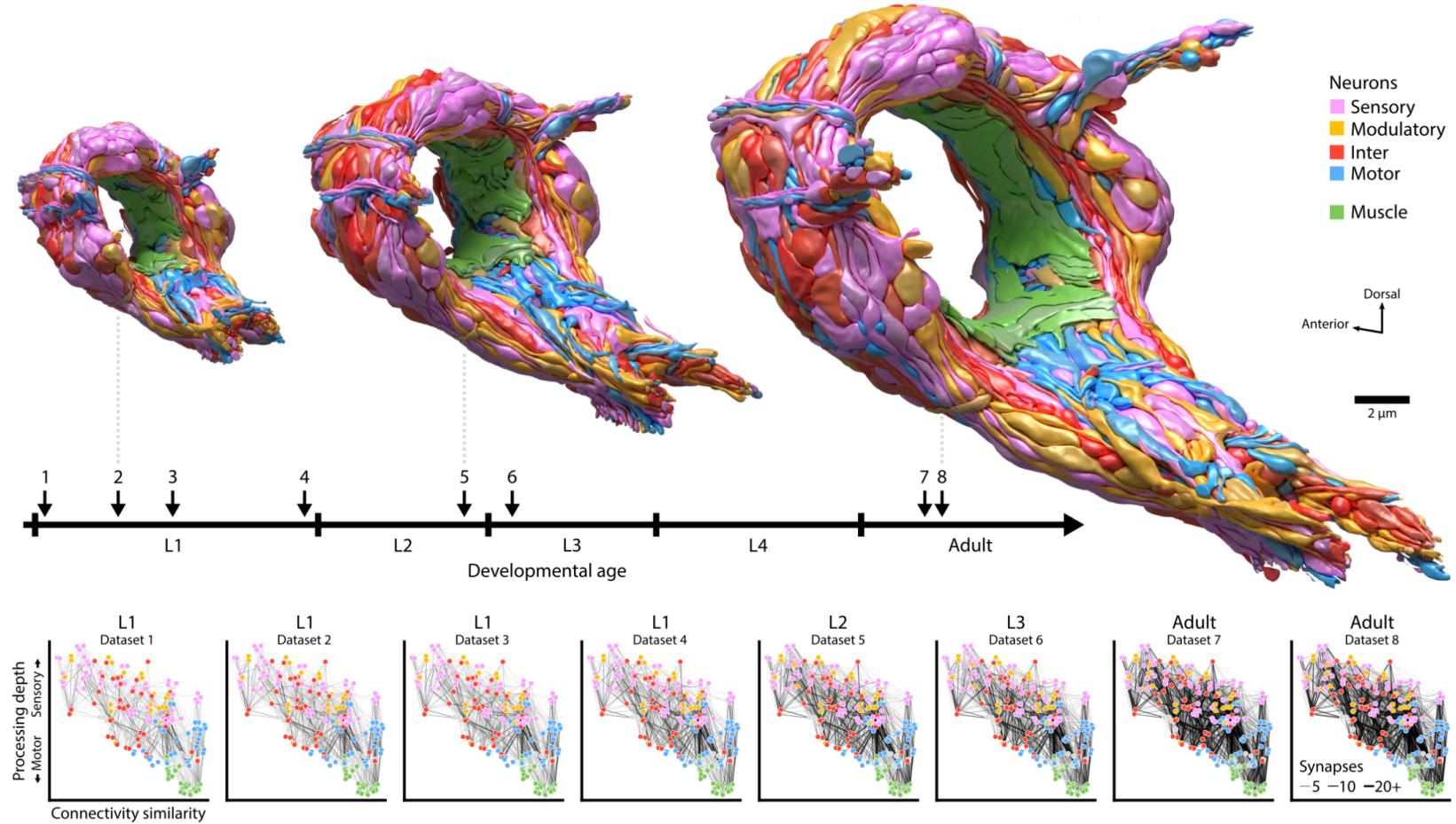
Why bilateral symmetry?

"We examined the connectivity of members of left–right homologous neuron pairs ... to assess the amount of natural variability in connectivity. ... Differences between individual worms will be expected to be at least this large."

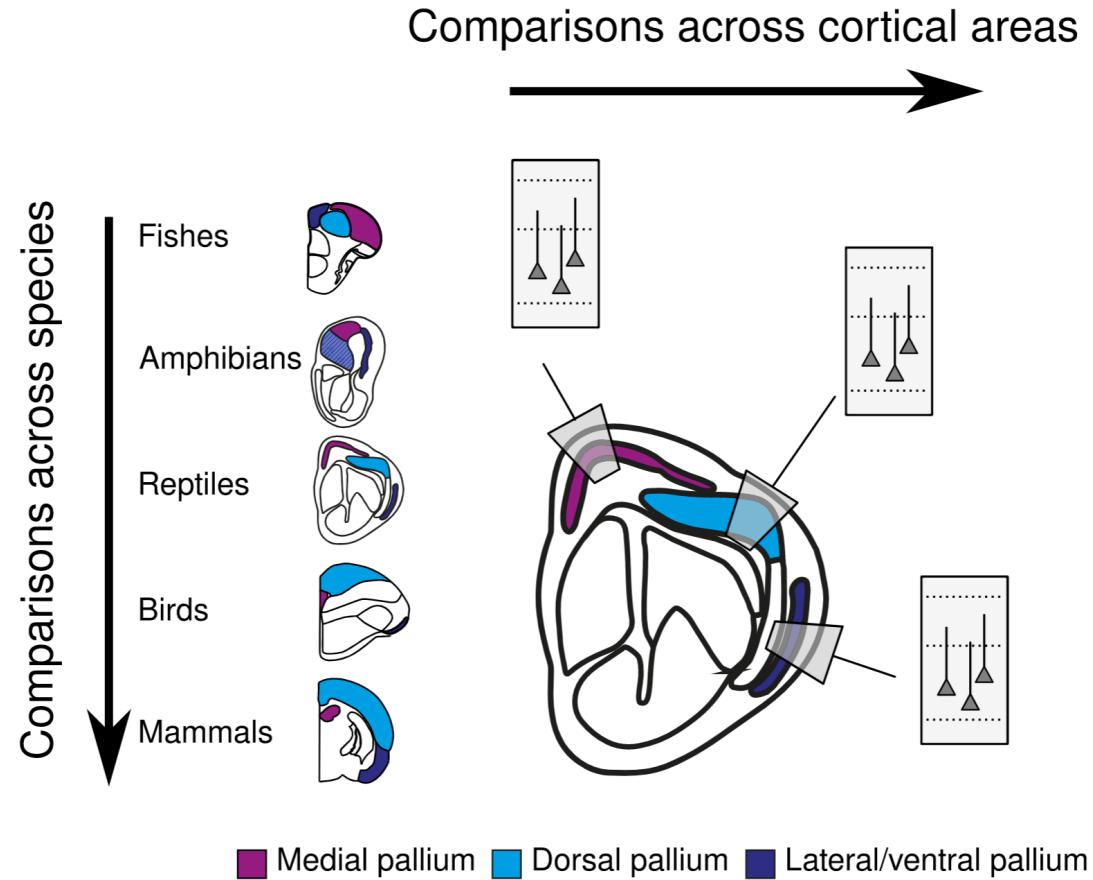
"... the gustatory neuron ASEL (that is, the left neuron of the pair) has greater chemical connectivity than ASER (that is, the right neuron of the pair) to the olfactory neuron class AWC."

Many connectomics questions require comparison

Connectomes across development



Connectomes across evolution, cortex

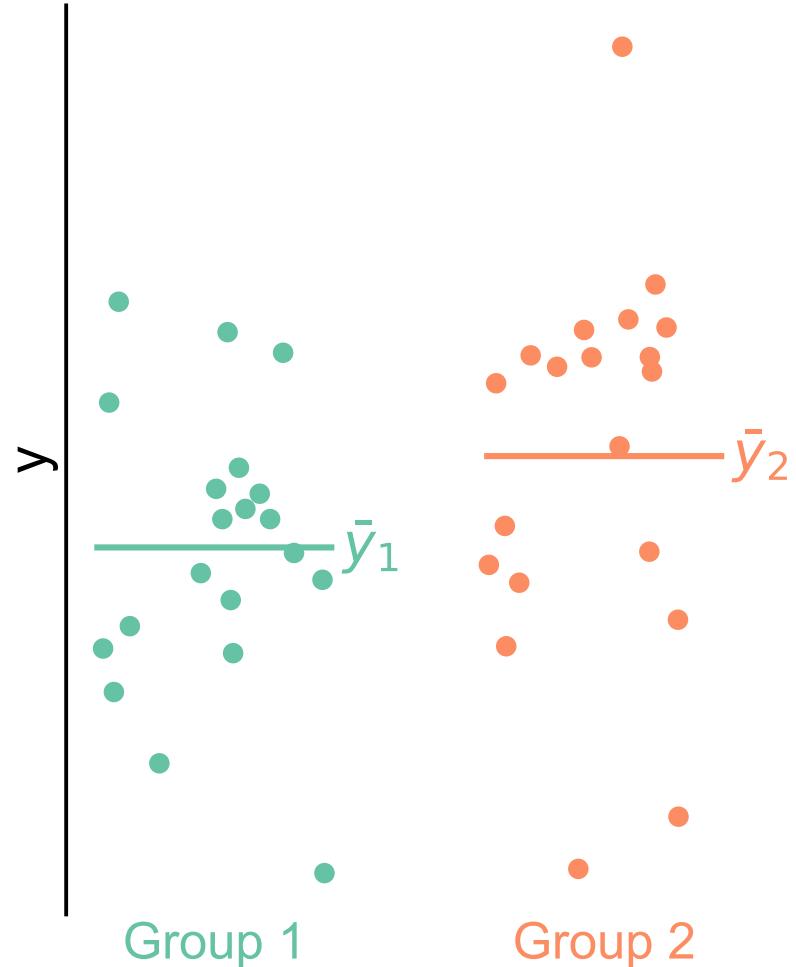


So, studying bilateral symmetry here lets us

- Try to formalize what we even mean by this property, and make claims about what we find in this connectome, and
- Test out methods for comparing networks for these future pursuits

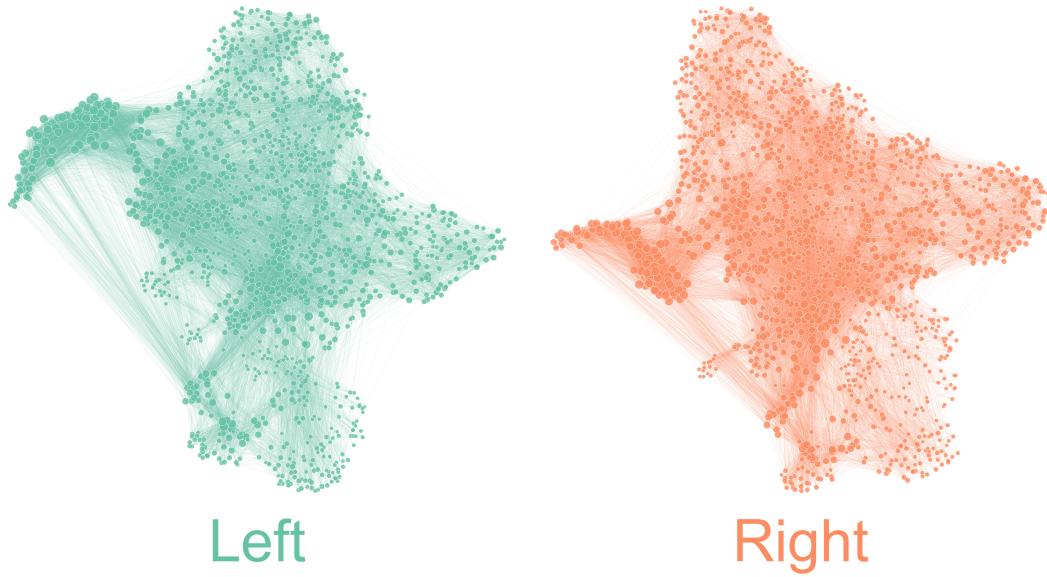
Are the **left** and **right** sides of this connectome
different?

Are these populations different?



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

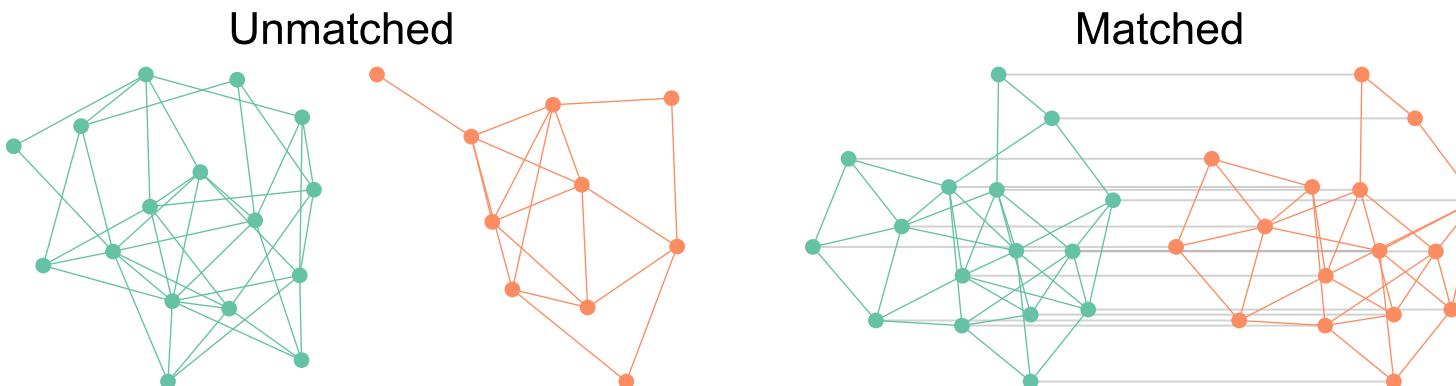
Are these two *networks* different?



- Want a two-network-sample test!
- $A^{(L)} \sim F^{(L)}$, $A^{(R)} \sim F^{(R)}$
- $H_0 : F^{(L)} = F^{(R)}$
 $H_A : F^{(L)} \neq F^{(R)}$

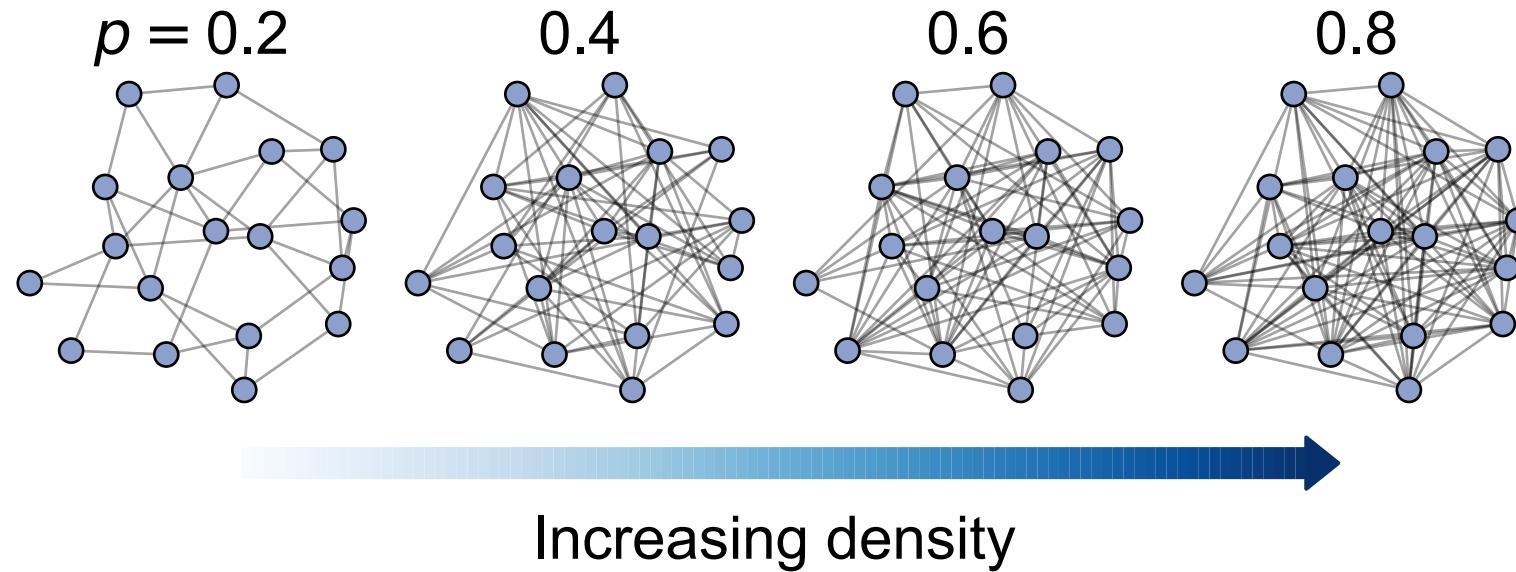
Assumptions

- We know the direction of synapses, so network is *directed*.
- For simplicity (for now), consider networks to be *unweighted*.
- For simplicity (for now), consider the **left** → **left** and **right** → **right** (*ipsilateral*) connections only.
- Not going to assume any nodes are matched



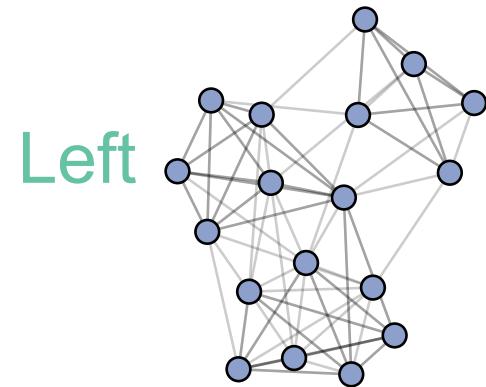
Erdos-Renyi model

- All edges are independent
- All edges generated with the same probability, p



Density-based testing

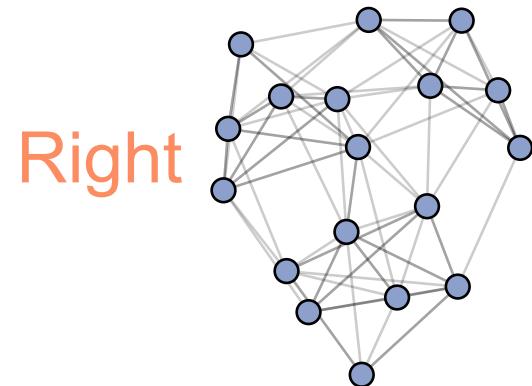
Compute global
connection density



Left

Compare ER
models

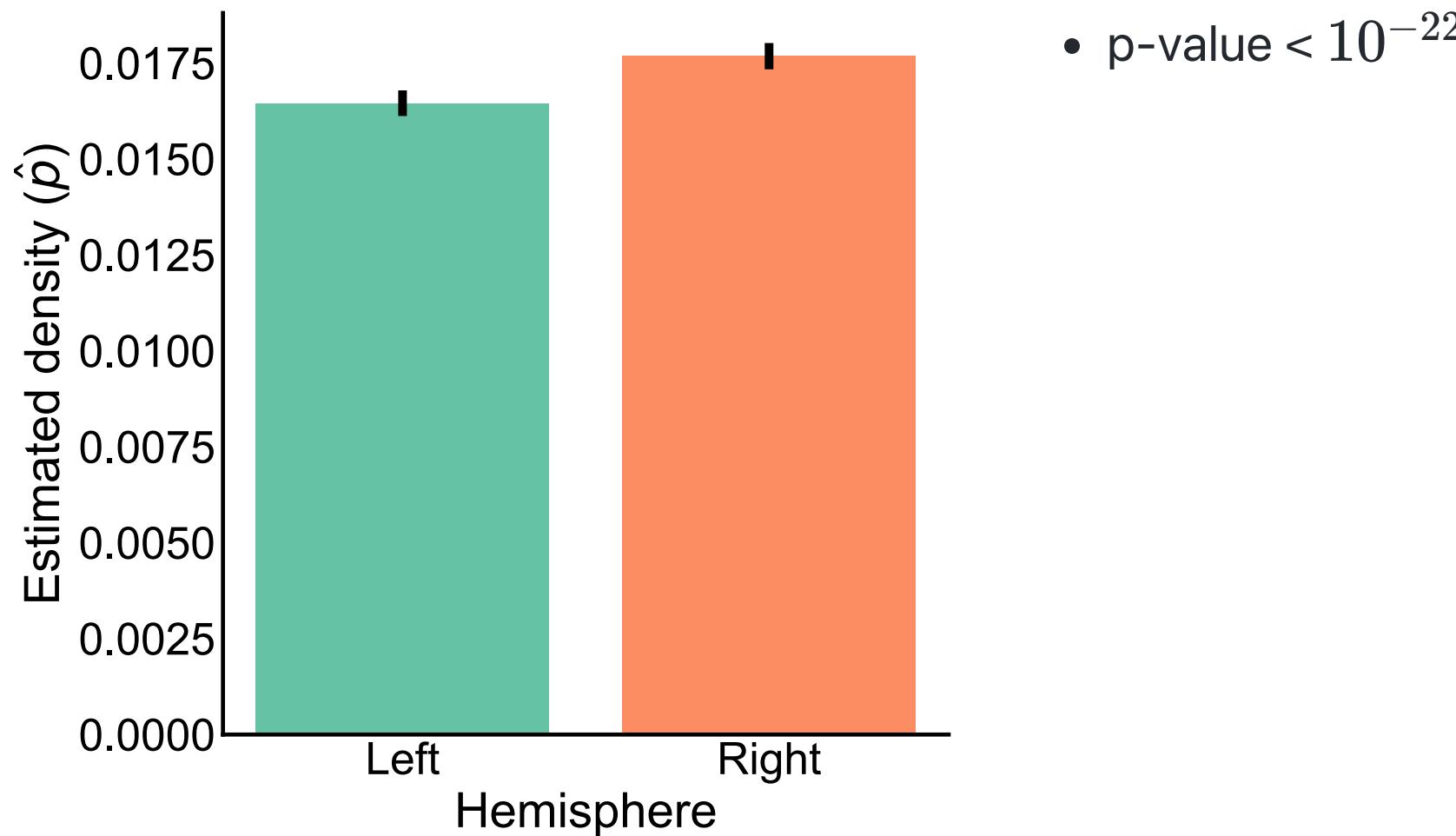
$$p = \frac{\# \text{ edges}}{\# \text{ potential edges}}$$



Right

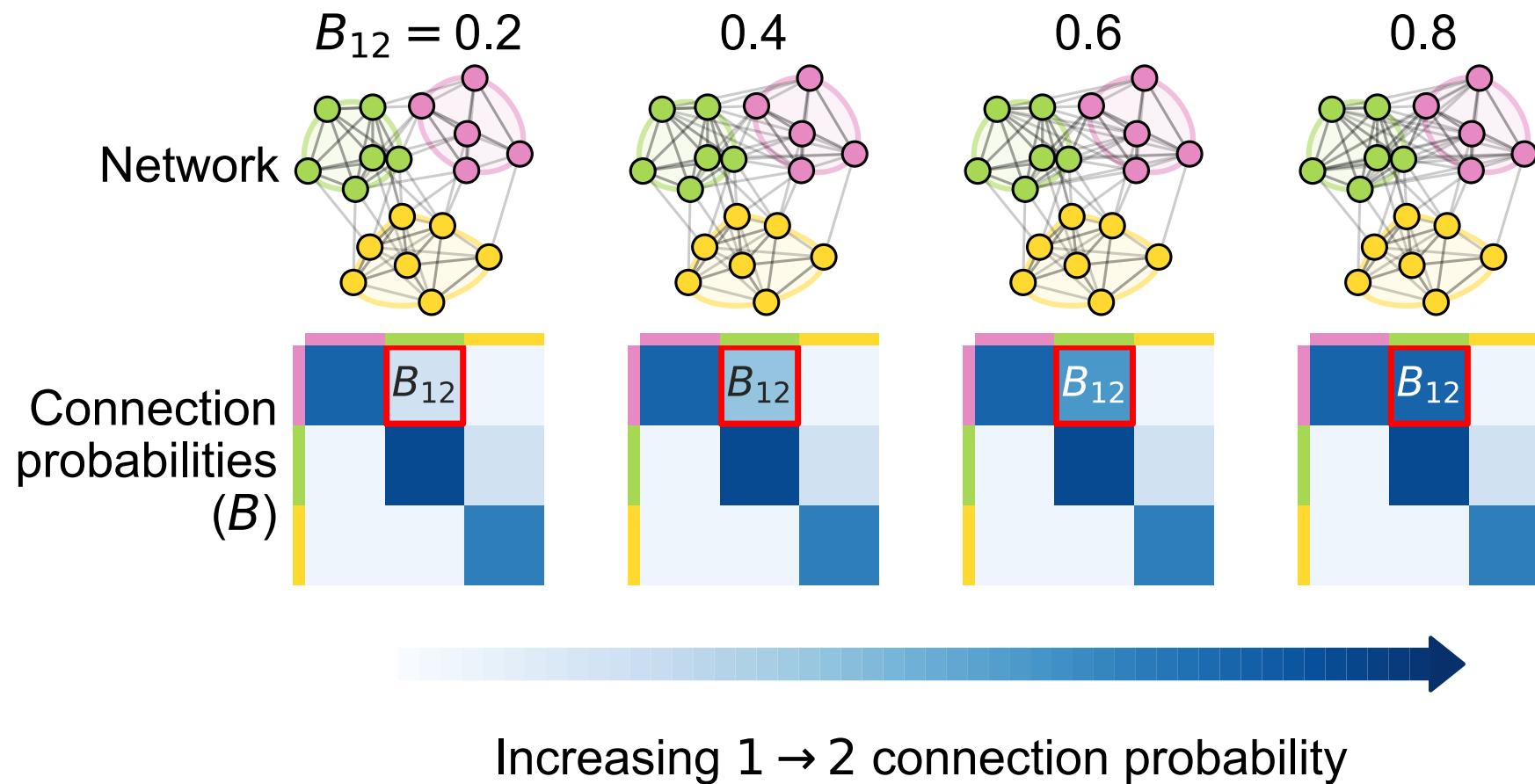
$$\begin{aligned} H_0: p^{(L)} &= p^{(R)} \\ H_A: p^{(L)} &\neq p^{(R)} \end{aligned}$$

We detect a difference in density

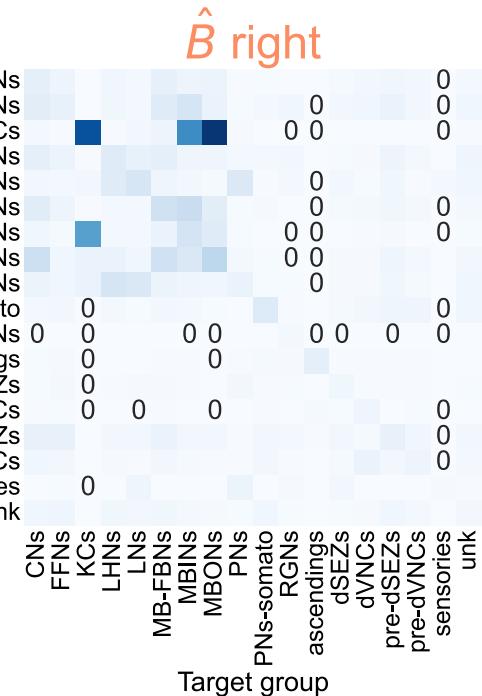
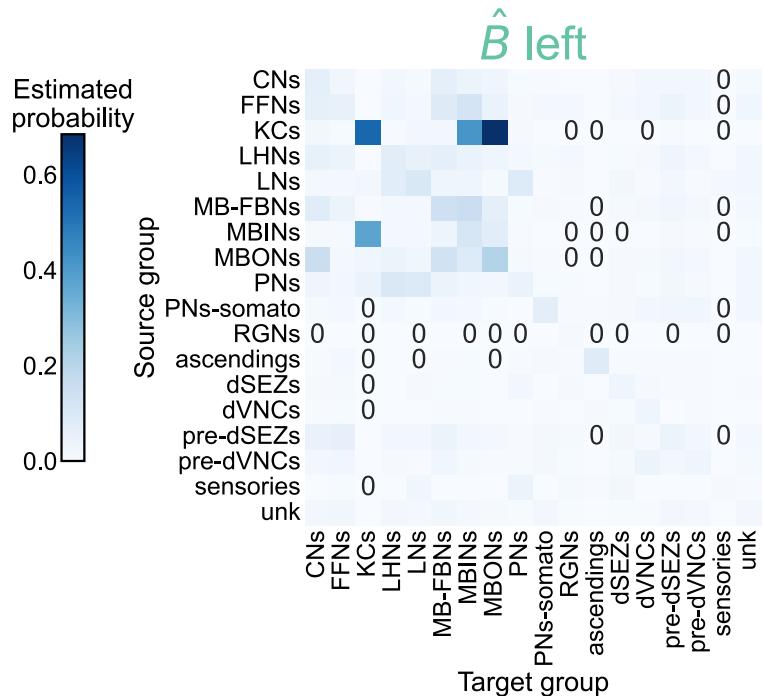
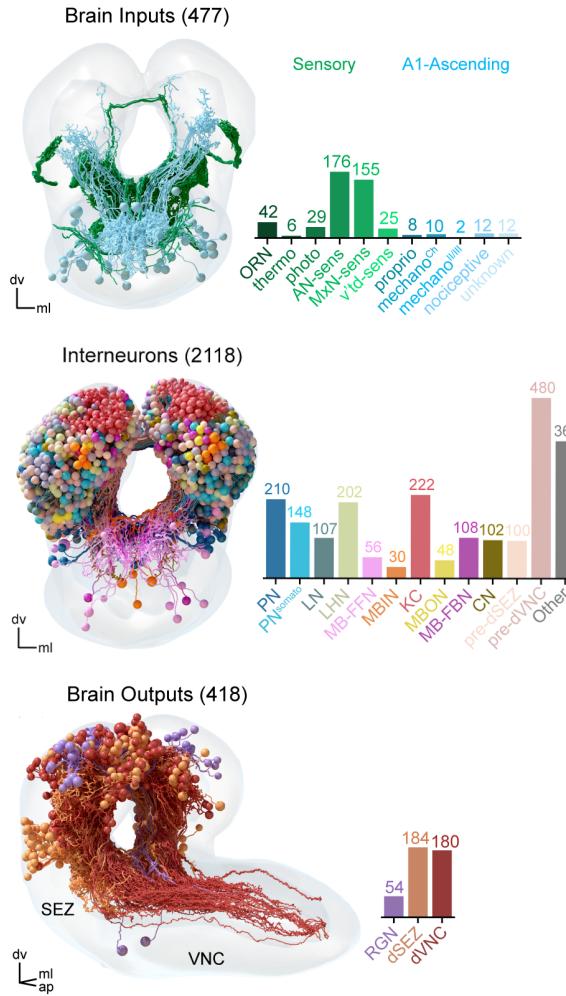


Stochastic block model

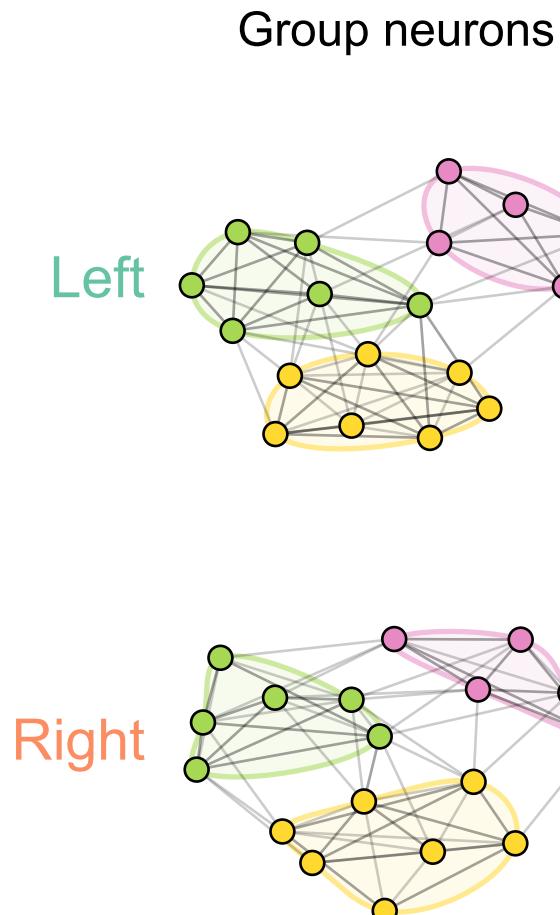
- Edge probabilities are a function of a neuron's group



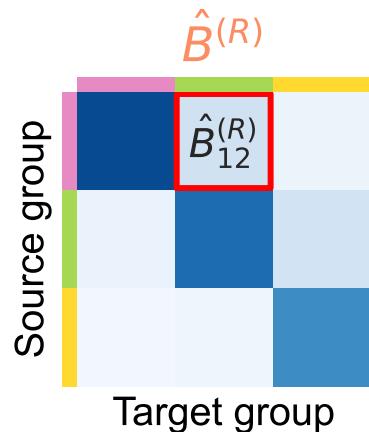
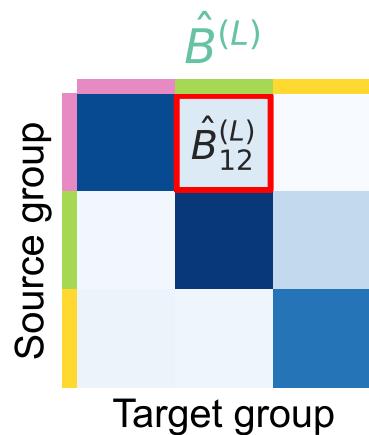
Connection probabilities between groups



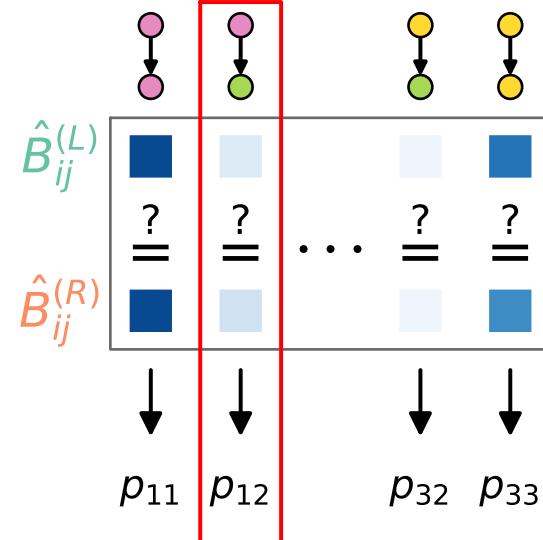
Group-based testing



Estimate group-to-group connection probabilities

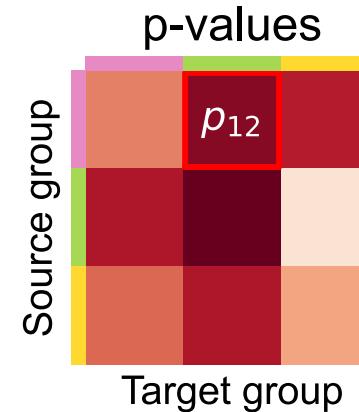


Compare estimated probabilities



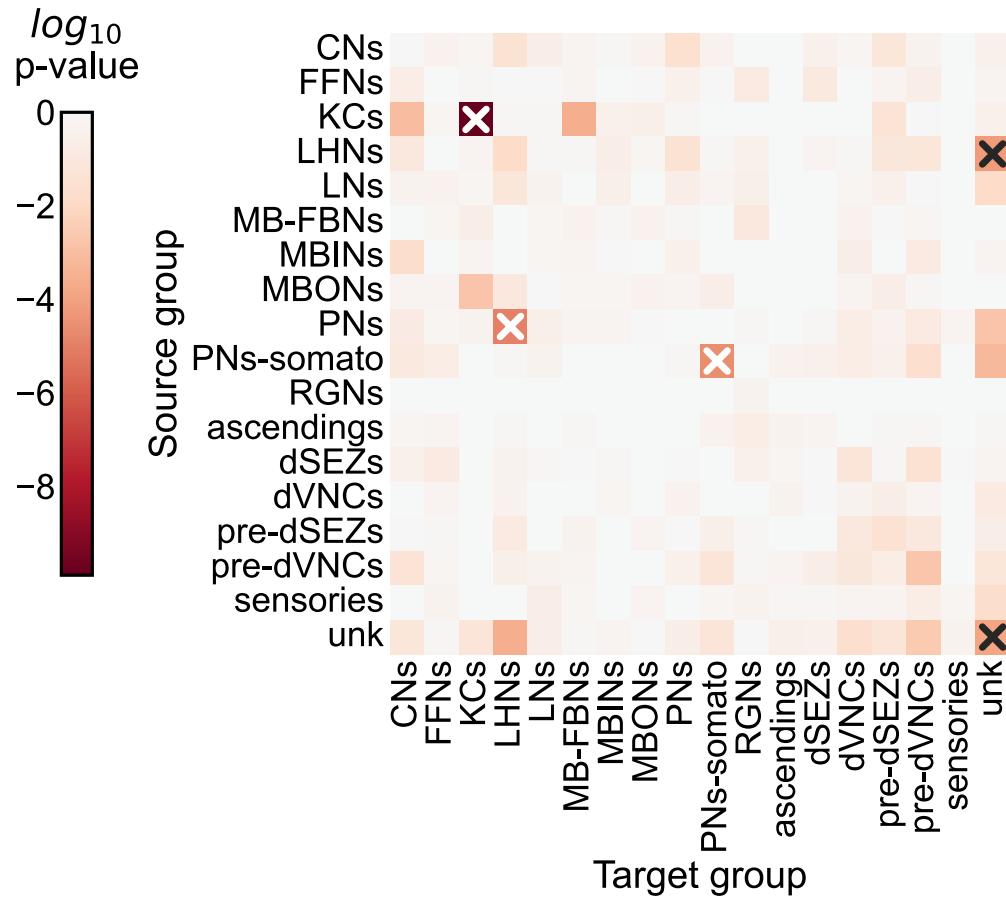
$$H_0: \hat{B}_{ij}^{(L)} = \hat{B}_{ij}^{(R)}$$
$$H_A: \hat{B}_{ij}^{(L)} \neq \hat{B}_{ij}^{(R)}$$

Combine p-values for overall test



$$H_0: \hat{B}_{ij}^{(L)} = \hat{B}_{ij}^{(R)}$$
$$H_A: \hat{B}_{ij}^{(L)} \neq \hat{B}_{ij}^{(R)}$$

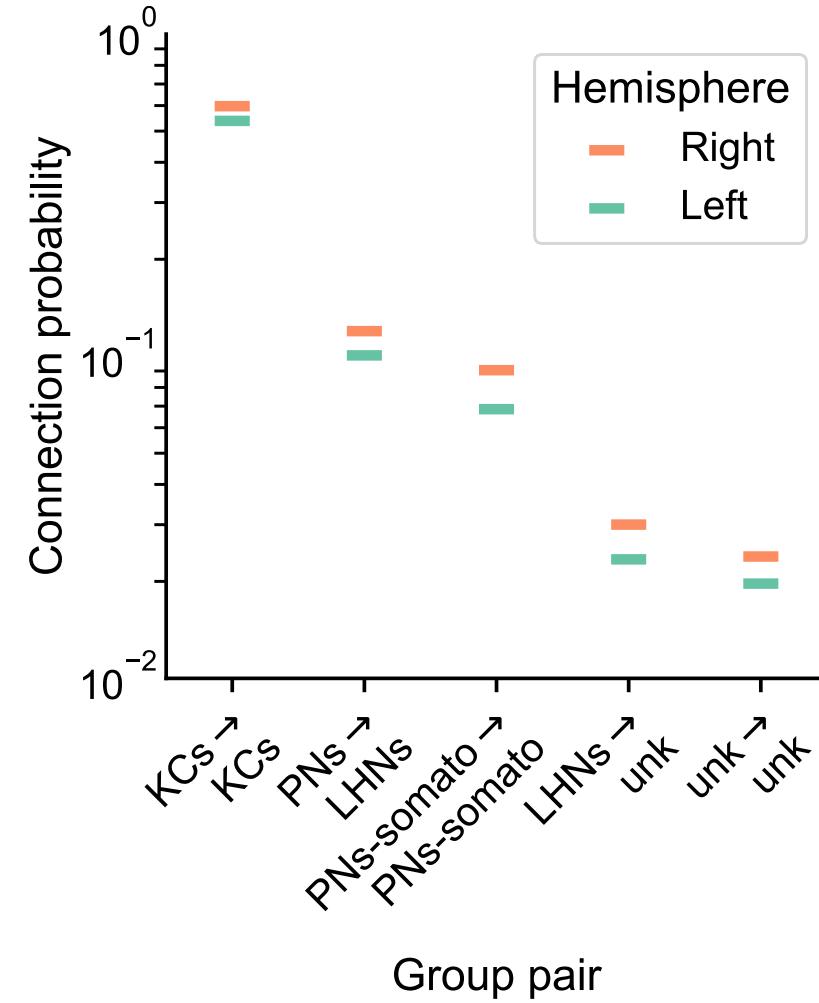
We detect a difference in group-to-group connection probabilities



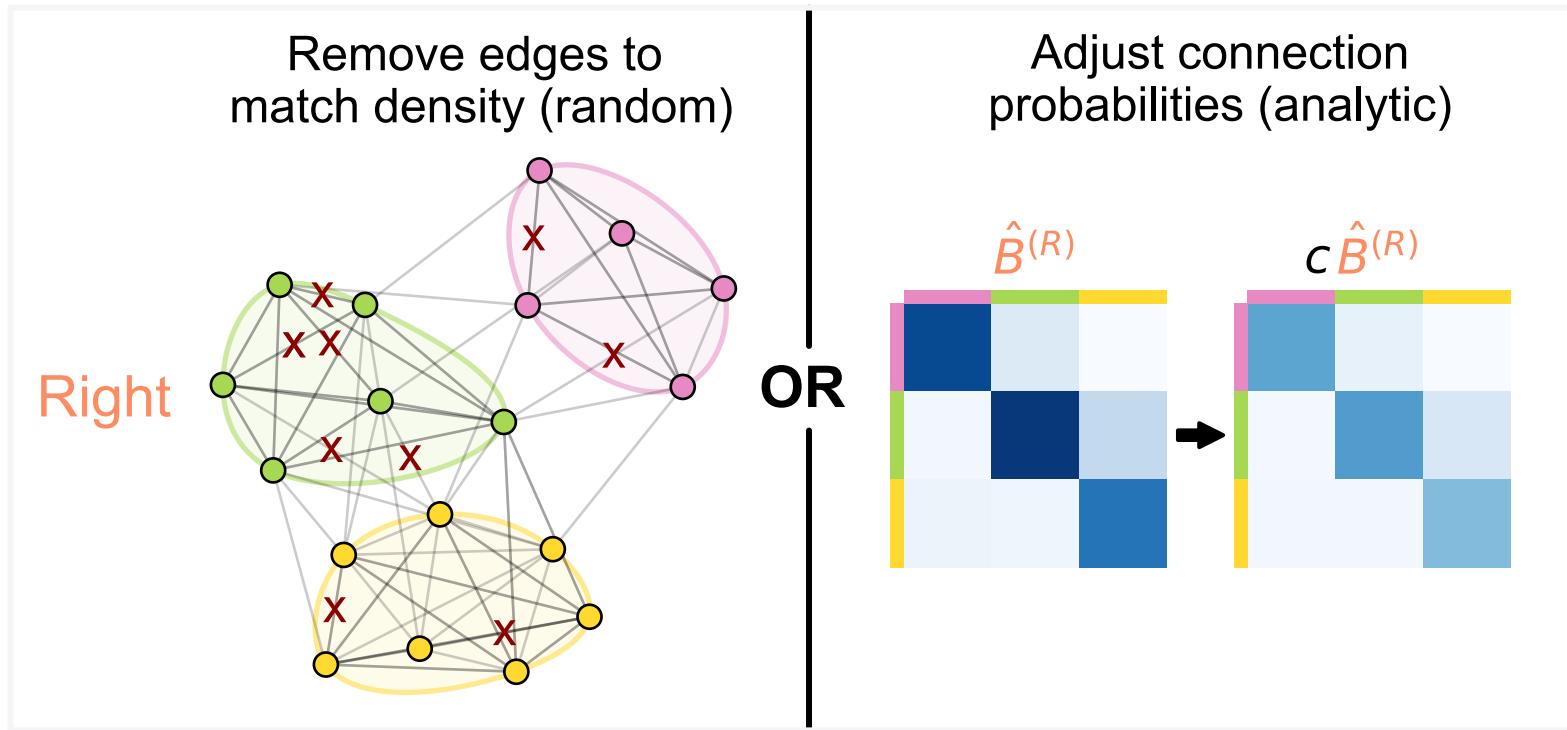
- After multiple comparison, find 5 group-to-group connections which are significantly different
- Combine (uncorrected) p-values (like a meta-analysis), leads to p-value for overall test of $< 10^{-7}$

Should we be surprised?

- Already saw that even the overall densities were different
- For all significant comparisons, probabilities on the right hemisphere were higher
- Maybe the right is just a "scaled up" version of the left?
 - $H_0 : B^{(L)} = cB^{(R)}$ where c is a density-adjusting constant, $\frac{p^{(L)}}{p^{(R)}}$



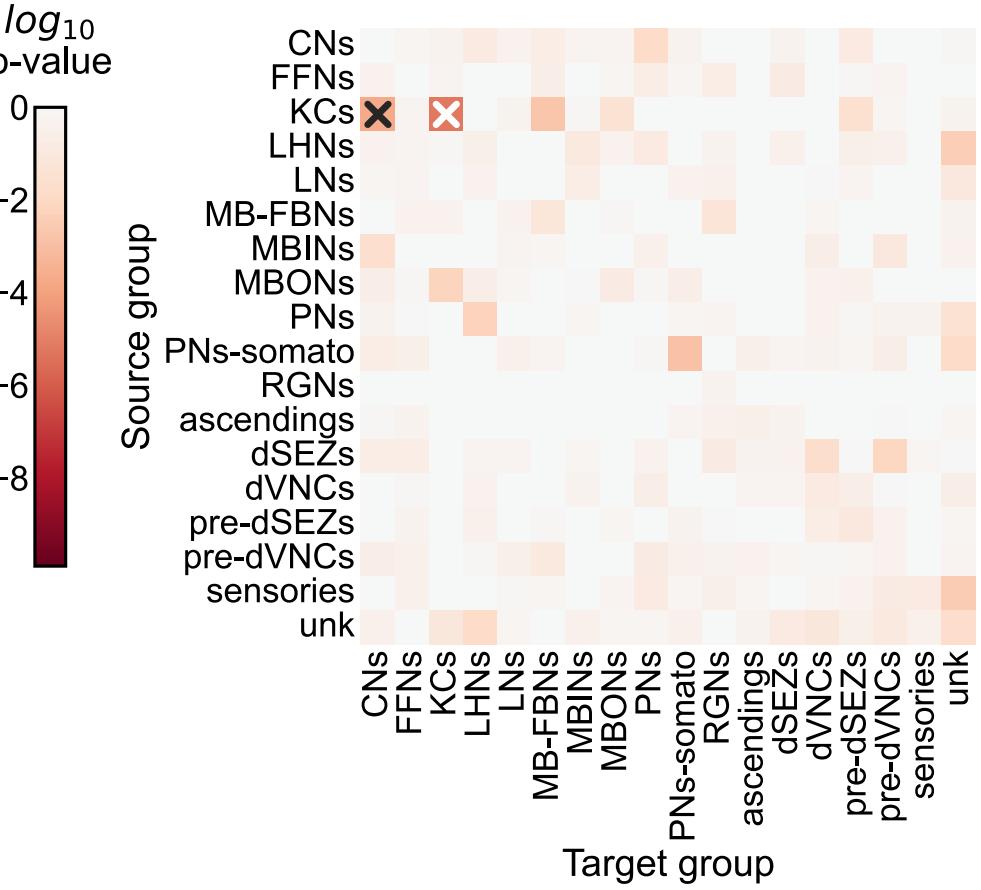
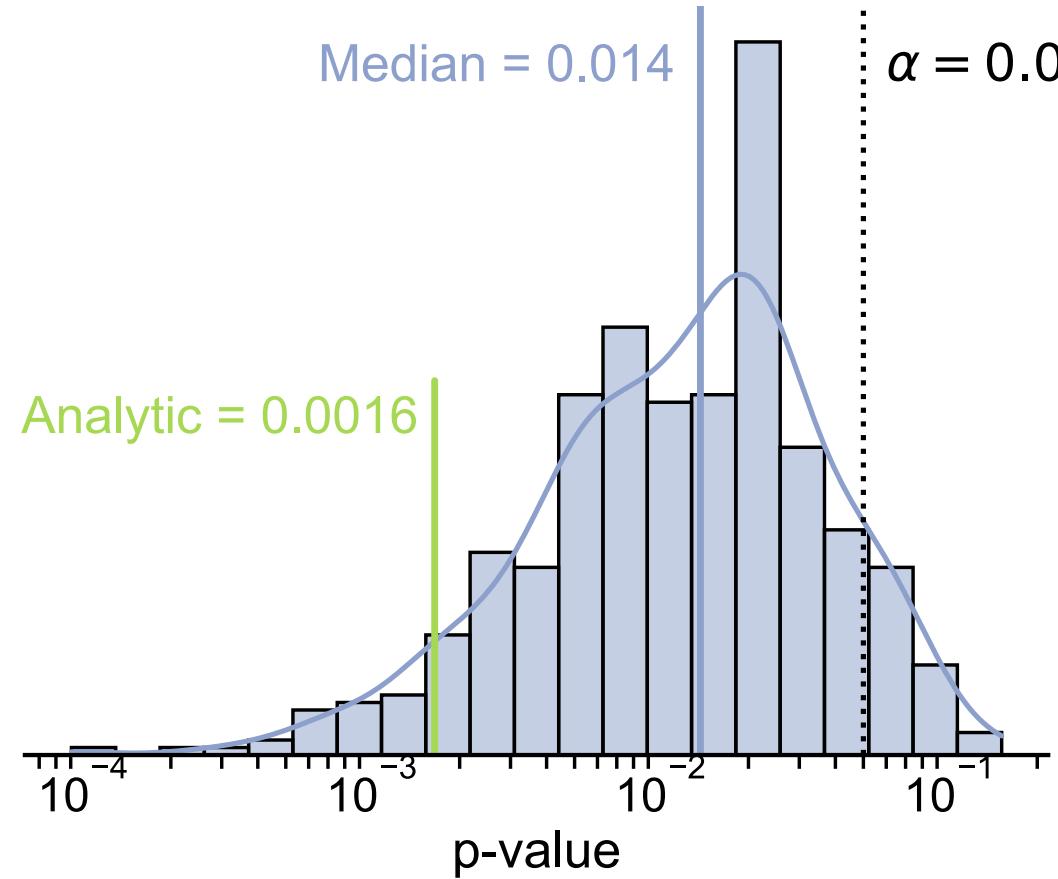
Adjusting for a difference in density



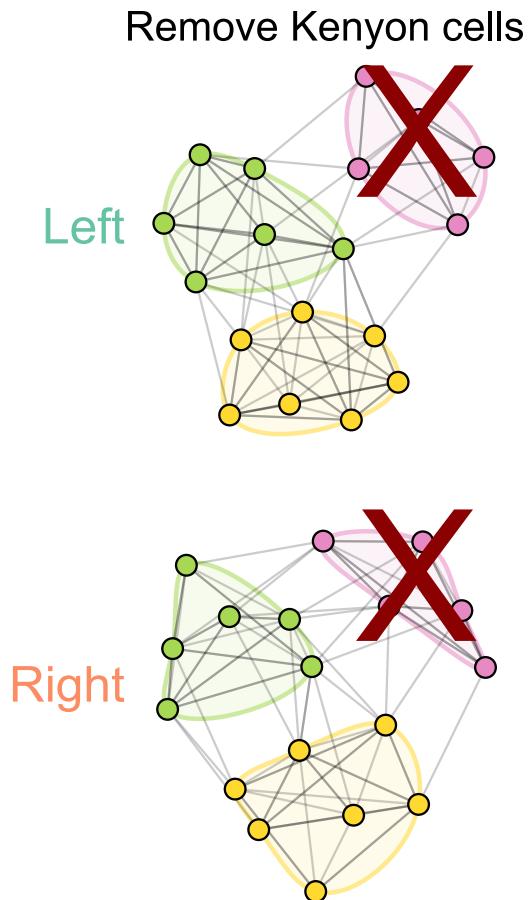
Rerun SBM testing

$$H_0: B^{(L)} = c B^{(R)}$$
$$H_A: B^{(L)} \neq c B^{(R)}$$

Even with density adjustment, we detect a difference



So the Kenyon cells (KCs) are the only group where we detect remaining differences...



Re-run all tests

ER

$$\begin{aligned} H_0: p^{(L)} &= p^{(R)} \\ H_A: p^{(L)} &\neq p^{(R)} \end{aligned}$$

SBM

$$\begin{aligned} H_0: B^{(L)} &= B^{(R)} \\ H_A: B^{(L)} &\neq B^{(R)} \end{aligned}$$

aSBM

$$\begin{aligned} H_0: B^{(L)} &= c B^{(R)} \\ H_A: B^{(L)} &\neq c B^{(R)} \end{aligned}$$

- ER test: $p < 10^{-26}$
- SBM test: $p \approx 0.003$
- Adjusted SBM test: $p \approx 0.43$

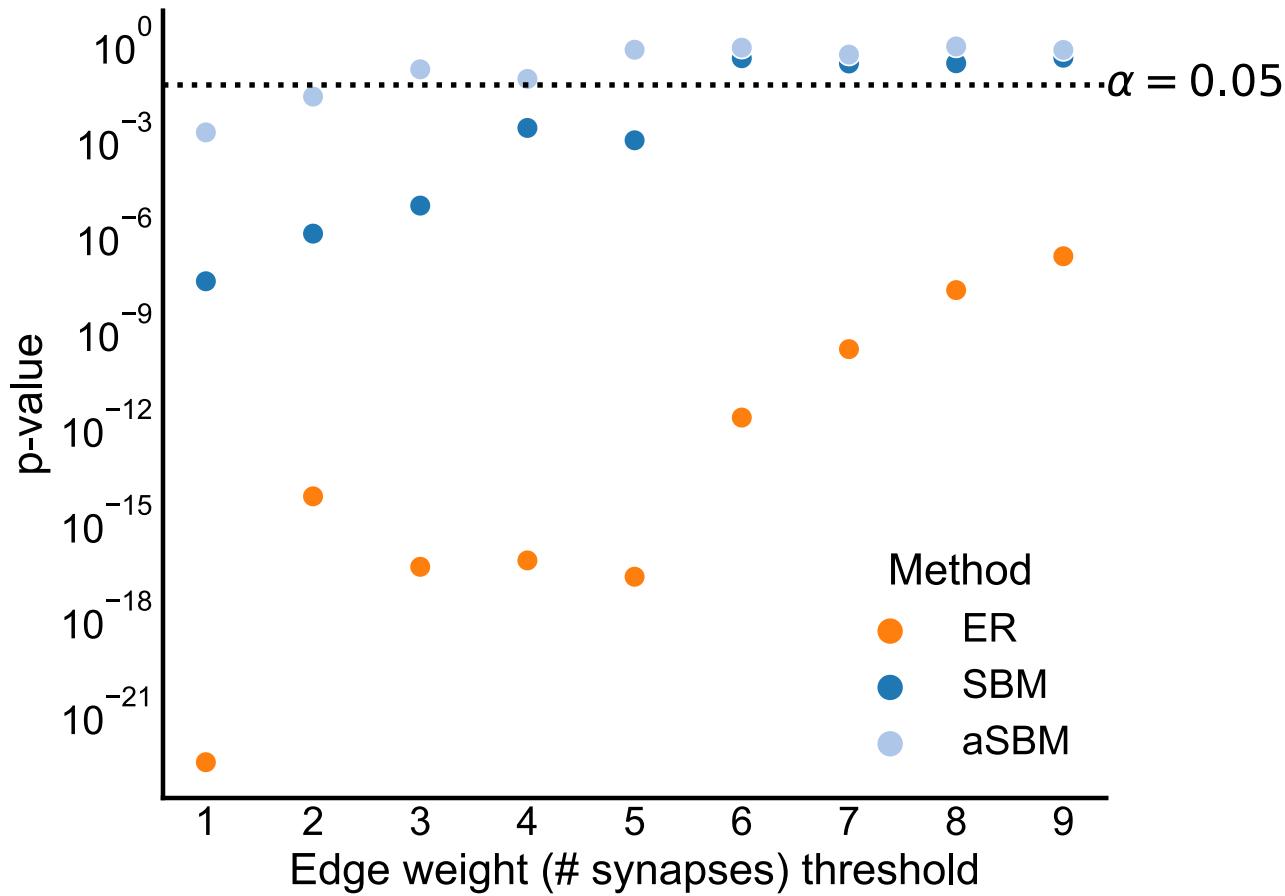
To sum up...

| Model | H_0 (vs. $H_A \neq$) | KCs | p-value | Interpretation |
|-------|-------------------------|-----|-----------------|--|
| ER | $p^{(L)} = p^{(R)}$ | + | $< 10^{-23}$ | Reject densities the same |
| SBM | $B^{(L)} = B^{(R)}$ | + | $< 10^{-7}$ | Reject group connection probabilities the same |
| aSBM | $B^{(L)} = cB^{(R)}$ | + | ≈ 0.002 | Reject above even after accounting for density |
| ER | $p^{(L)} = p^{(R)}$ | - | $< 10^{-26}$ | Reject densities the same (w/o KCs) |
| SBM | $B^{(L)} = B^{(R)}$ | - | ≈ 0.003 | Reject group connection probabilities the same (w/o KCs) |
| aSBM | $B^{(L)} = cB^{(R)}$ | - | ≈ 0.43 | Don't reject above after density adjustment (w/o KCs) |

Extensions (and other tools)

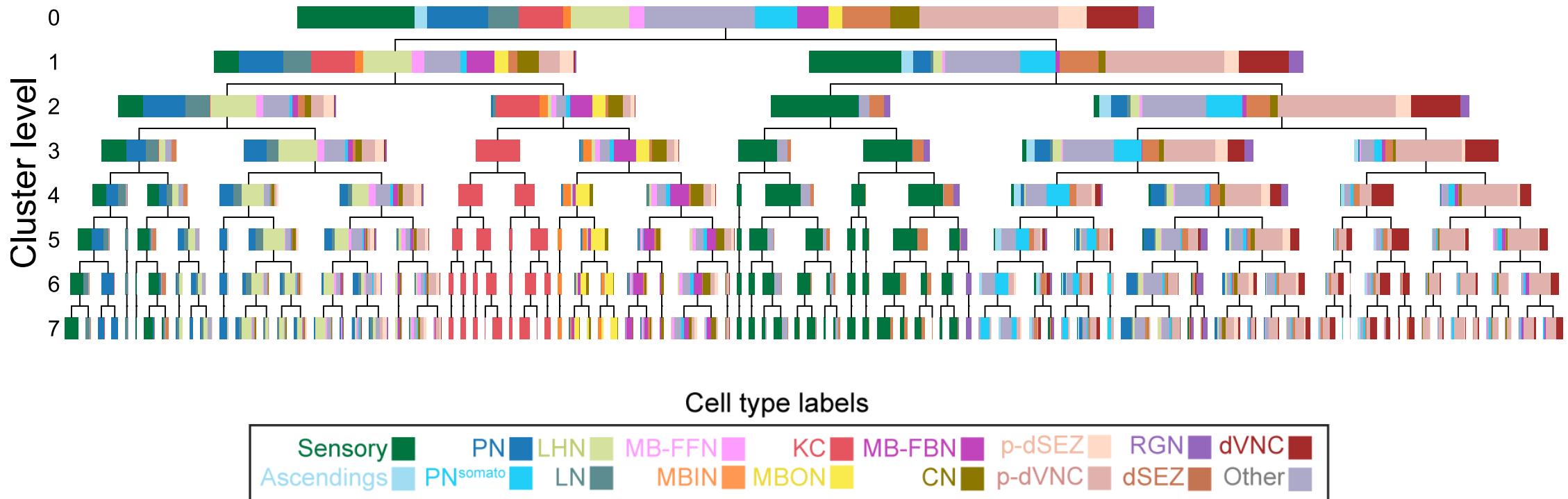
But you threw out all of the edge weights!

Thresholding at higher synapse counts reduces asymmetry



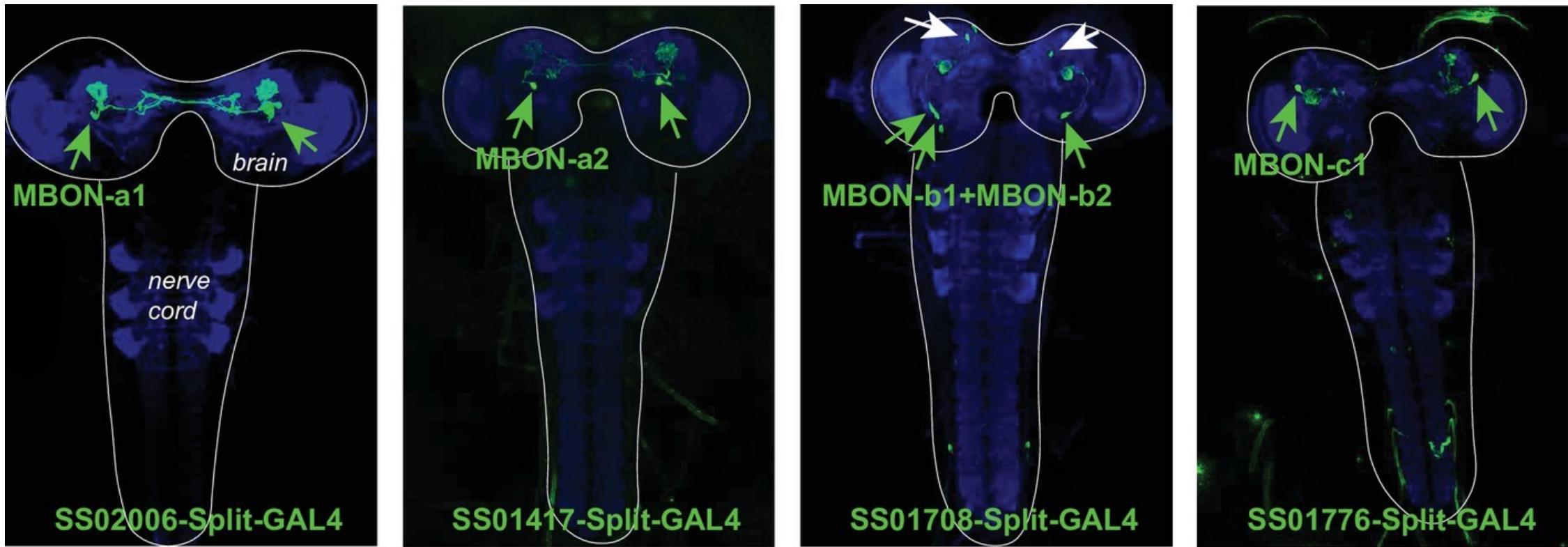
What do we consider to be a "cell type"?

Hierarchical clustering of neurons based on observed connectivity

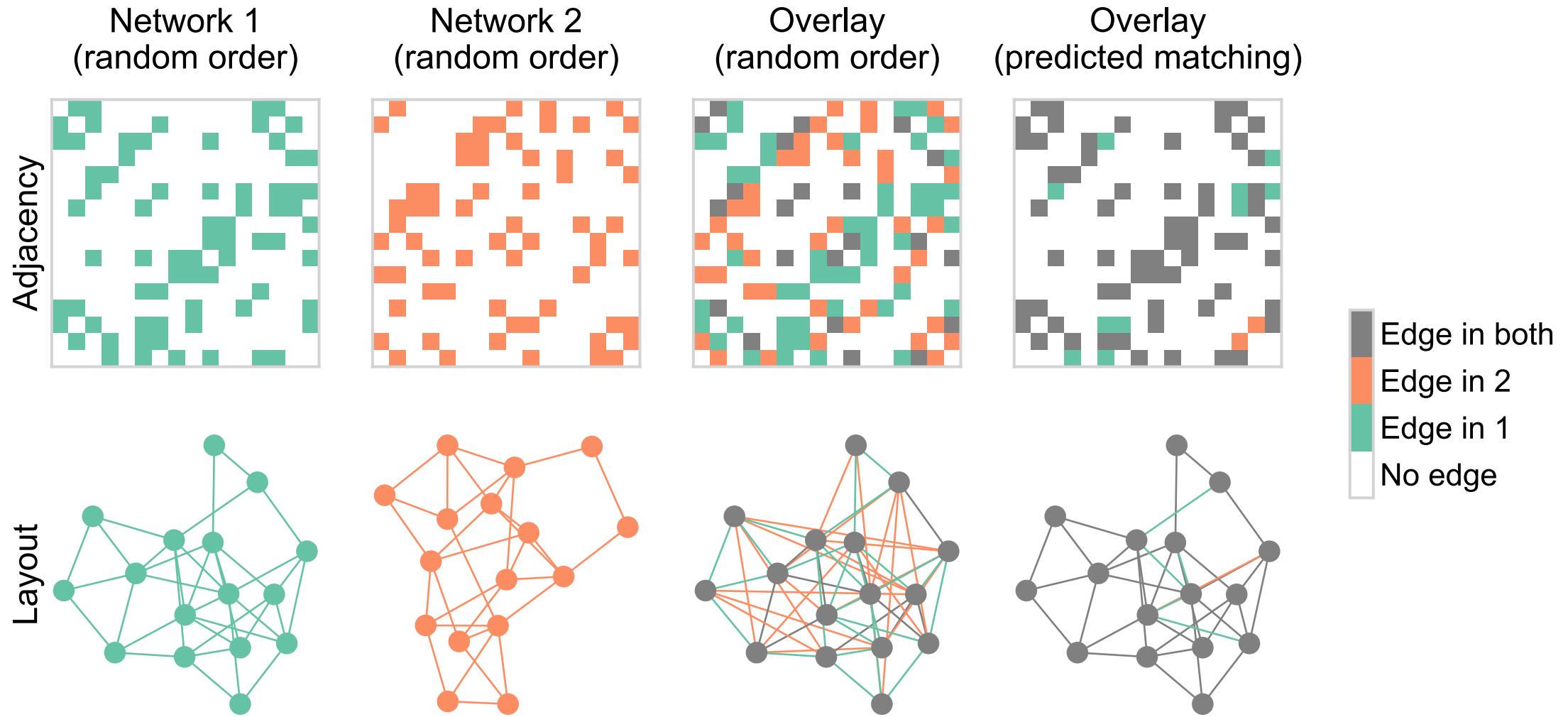


Are nodes/edges matched across hemispheres?

Bilateral neuron pairs



Graph matching



Graph matching predicts single-neuron pairs between hemispheres

- ~86% of predicted pairs are confirmed by a human annotator

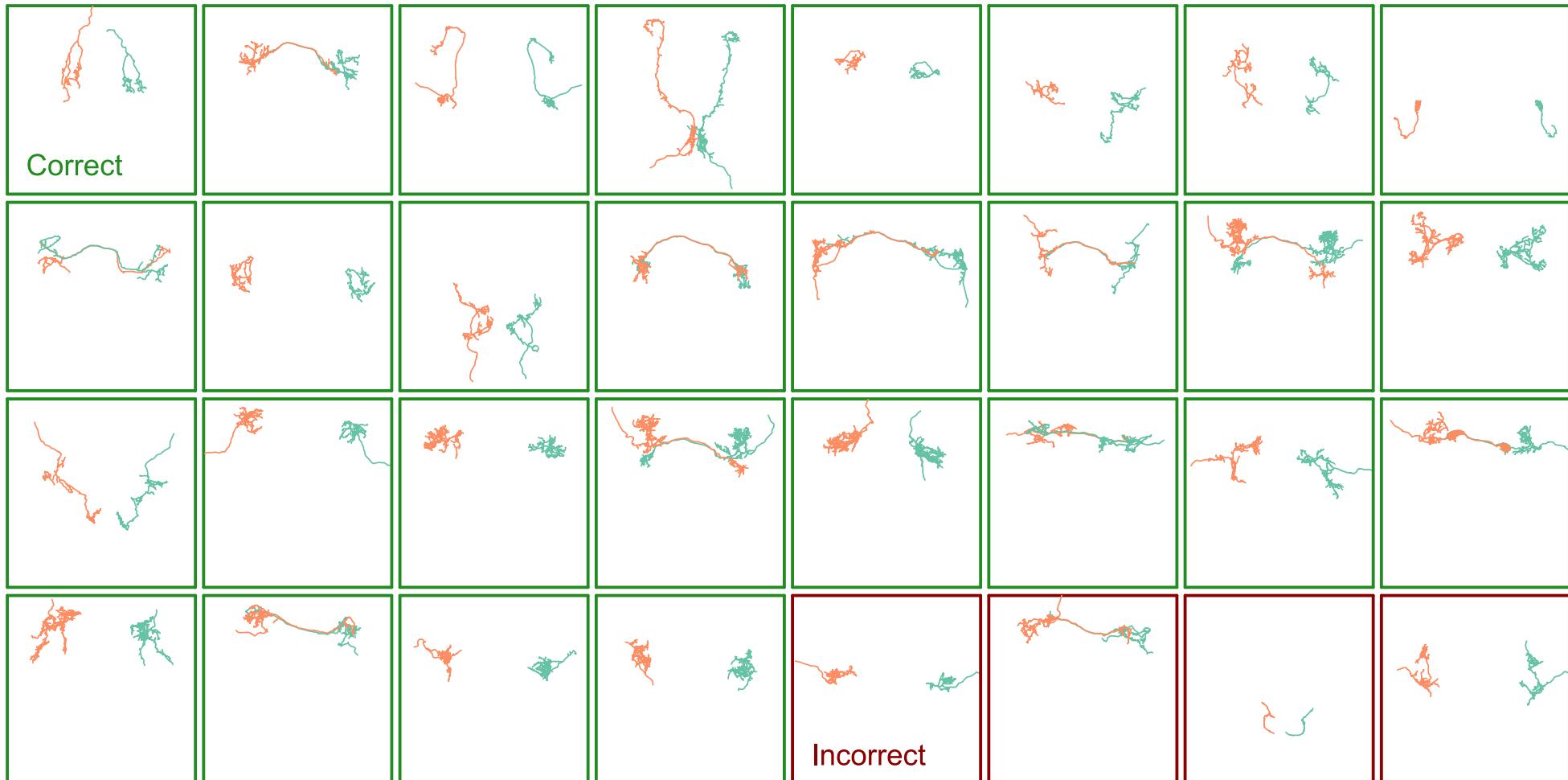
Left adjacency



Right adjacency



Predicted pairs are morphologically similar

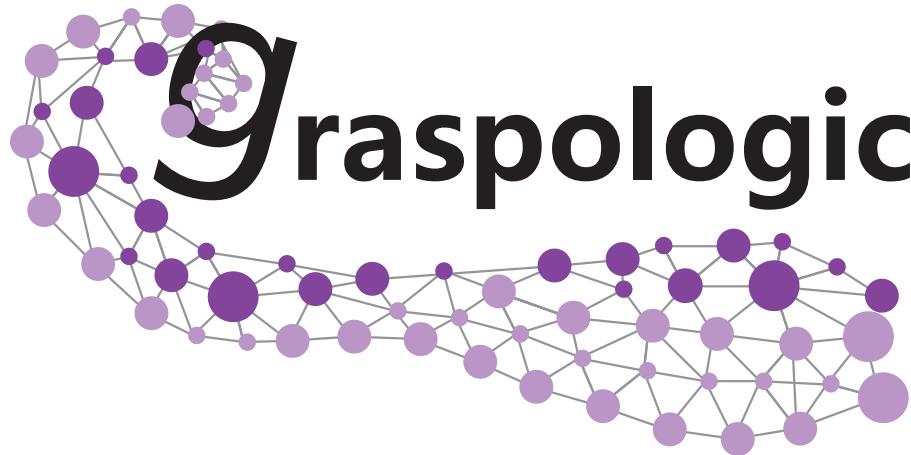


In summary...

- Studied statistical ways of framing "bilateral symmetry", proposing a test procedure for each
- All tests found the left and the right hemispheres significantly different, unless ignoring Kenyon cells and adjust for the difference in density
- Provided a foundation for future principled comparisons of connectomes
- Mentioned several other tools/analyses which could alter the definition of symmetry
 - Edge weights
 - Inferring neuron groups
 - Graph matching to find pairs

graspologic:

github.com/microsoft/graspologic



downloads 107k



Stars 242

contributors 46

License MIT

This work:

github.com/neurodata/bilateral-connectome

The Erdos-Renyi (ER) model

The [Erdos-Renyi \(ER\) model](#) is one of the simplest network models. This model treats the probability of each potential edge in the network occurring to be the same. In other words, all edges between any two nodes are equally likely.

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

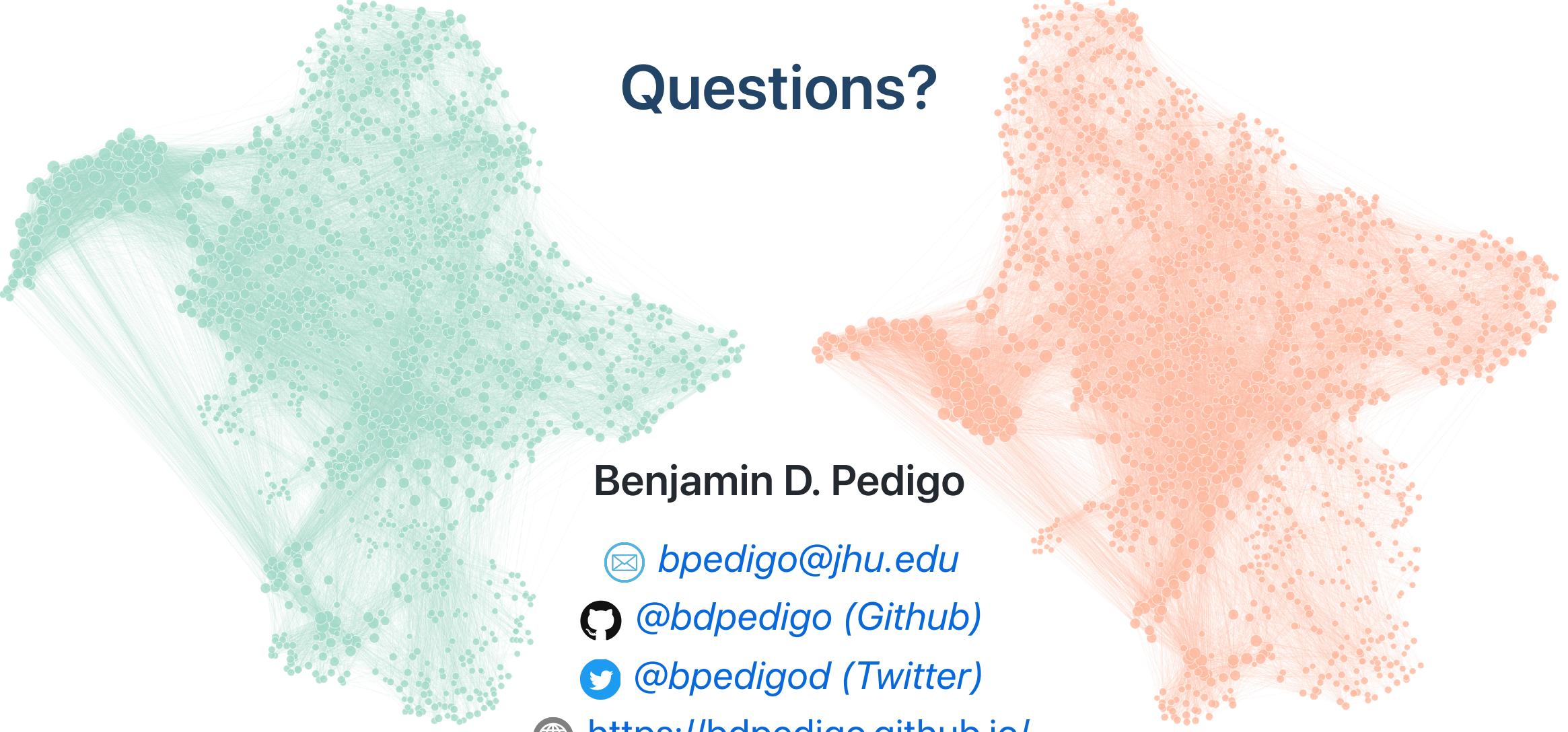
Mike Powell, Eric Bridgeford, Carey Priebe, Joshua Vogelstein, Kareef Ullah, Diane Lee, Sambit Panda, Jaewon Chung, Ali Saad-Eldin, NeuroData lab

University of Cambridge / MRC Laboratory of Molecular Biology

Michael Winding, Albert Cardona, Marta Zlatic, Chris Barnes

Funding





Questions?

Benjamin D. Pedigo

✉ bpedigo@jhu.edu

/github/ [@bdpedigo \(Github\)](https://github.com/bdpedigo)

/twitter/ [@bpedigod \(Twitter\)](https://twitter.com/bpedigod)

/website/ <https://bpedigo.github.io/>

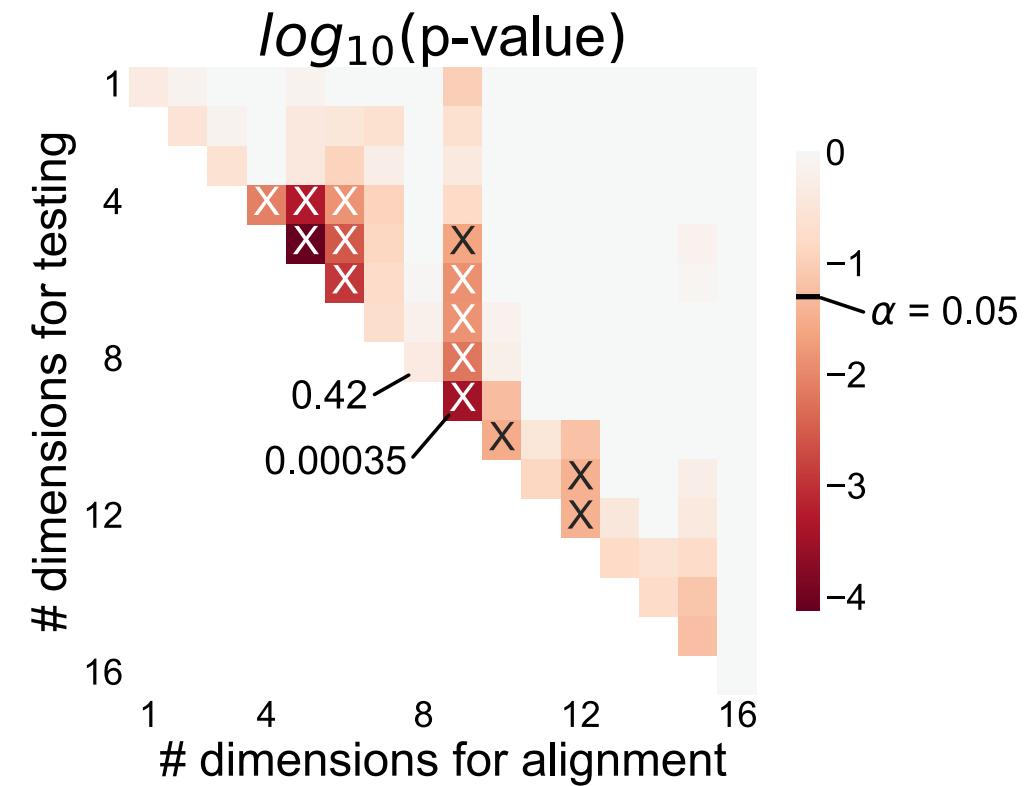
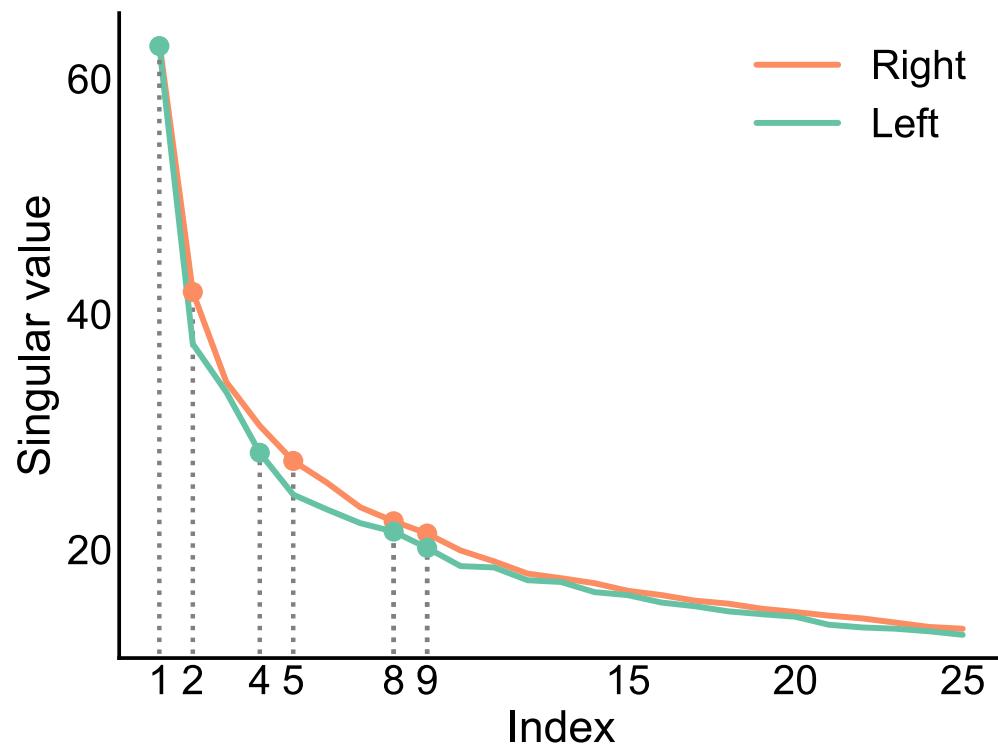
Left

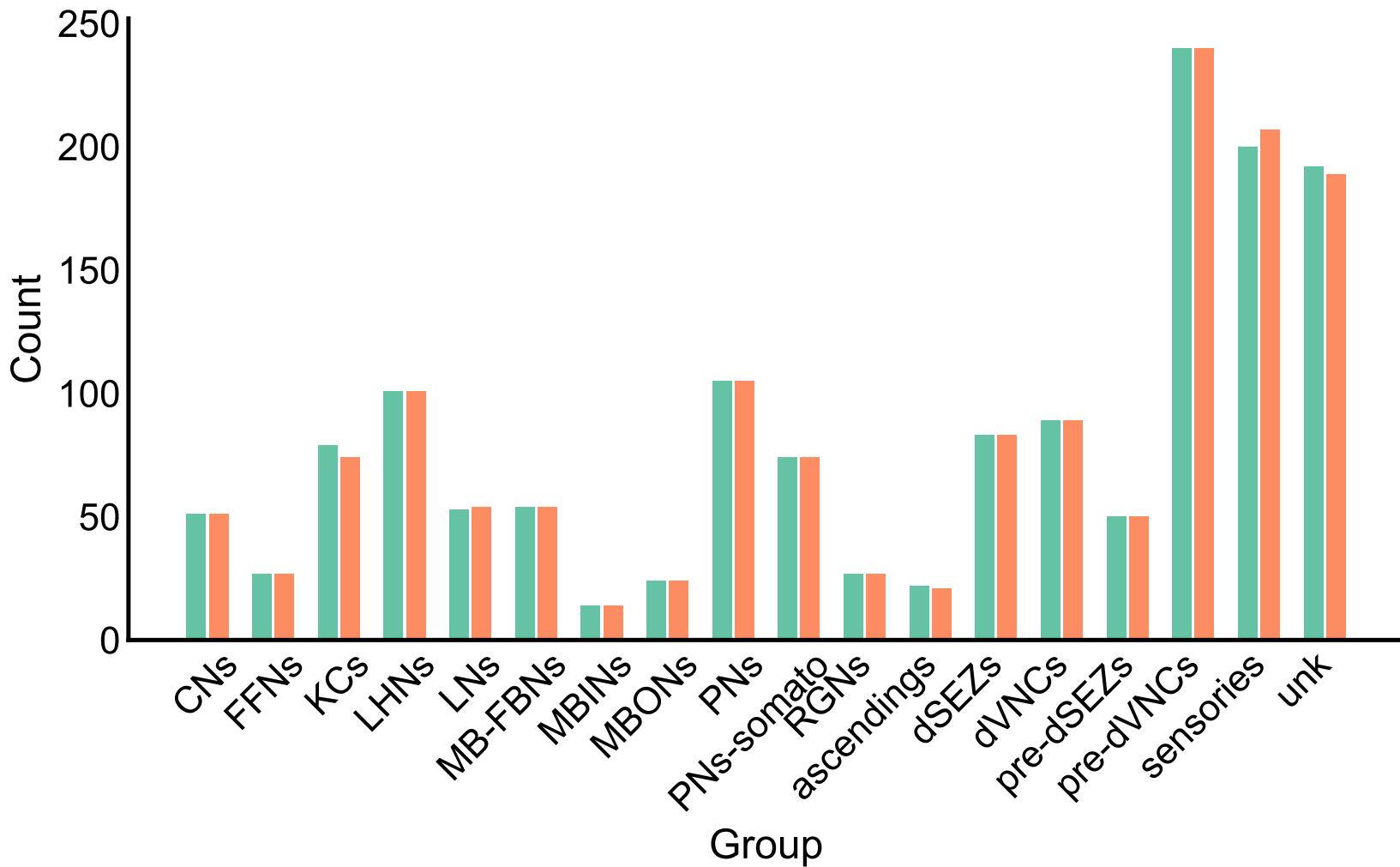
Right

Appendix

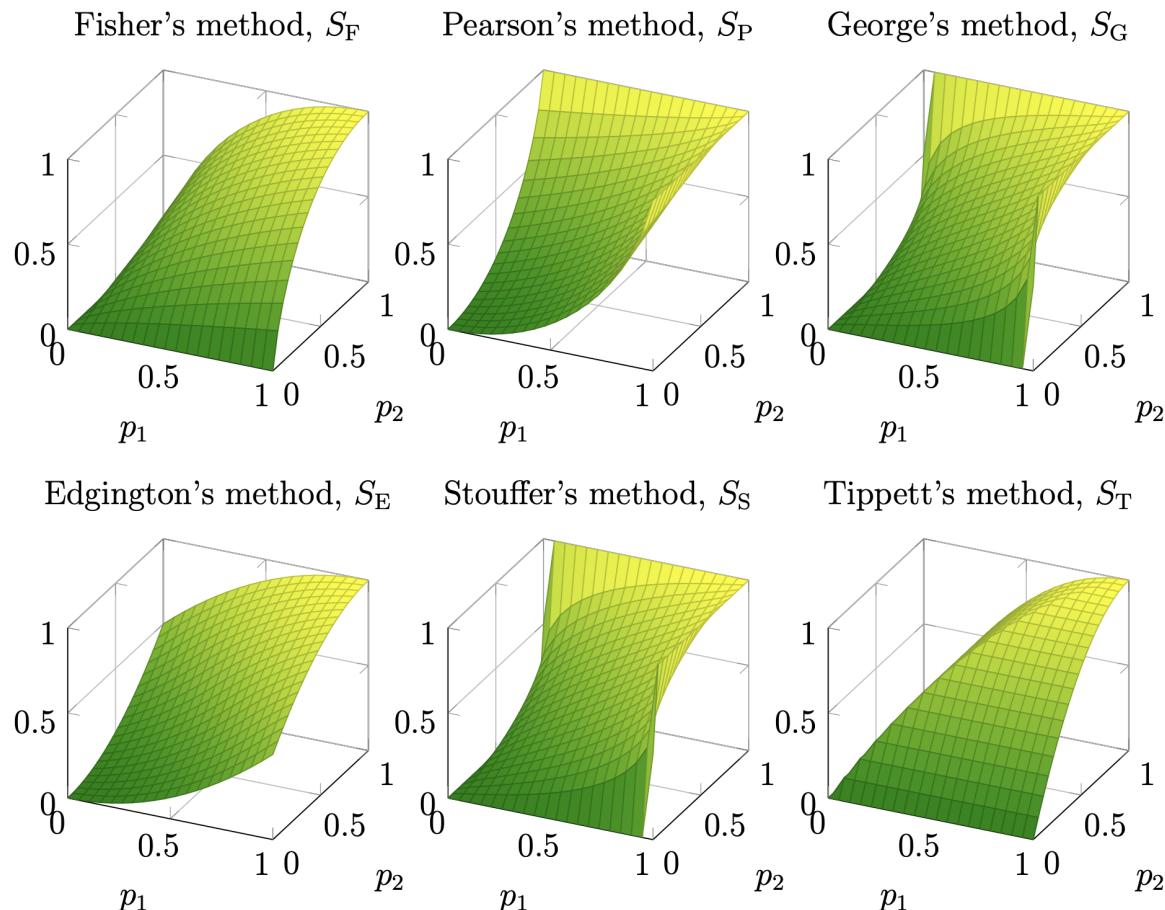
There are so many other models!

Latent distribution test (random dot product graph)





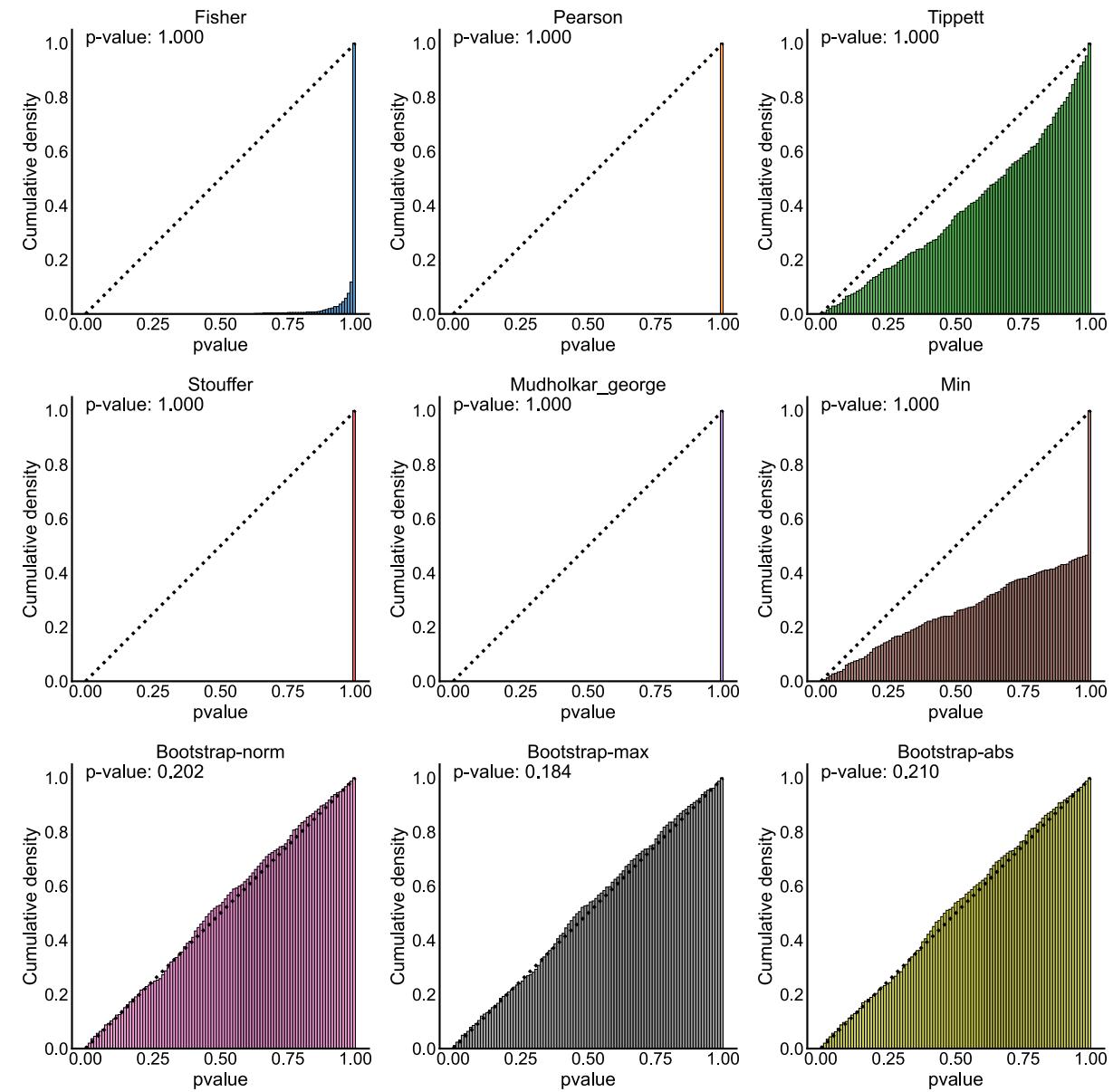
Combining p-values: nobody's perfect



Combining p-values: don't trust SciPy until 1.9.0

```
    elif method == 'tippett':  
        statistic = np.min(pvalues)  
        pval = distributions.beta.sf(statistic, 1, len(pvalues))  
        pval = distributions.beta.cdf(statistic, 1, len(pvalues))
```

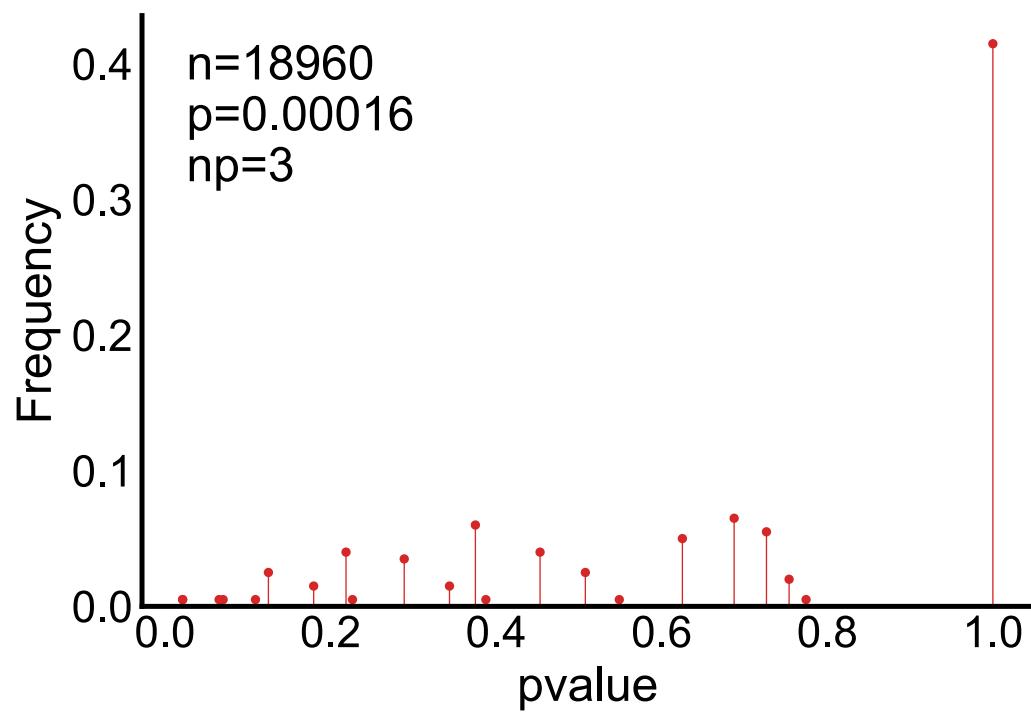
Distribution under the null for combining p-values



Combining p-values: be careful with discreteness

THE COMBINATION OF PROBABILITIES ARISING FROM DATA
IN DISCRETE DISTRIBUTIONS

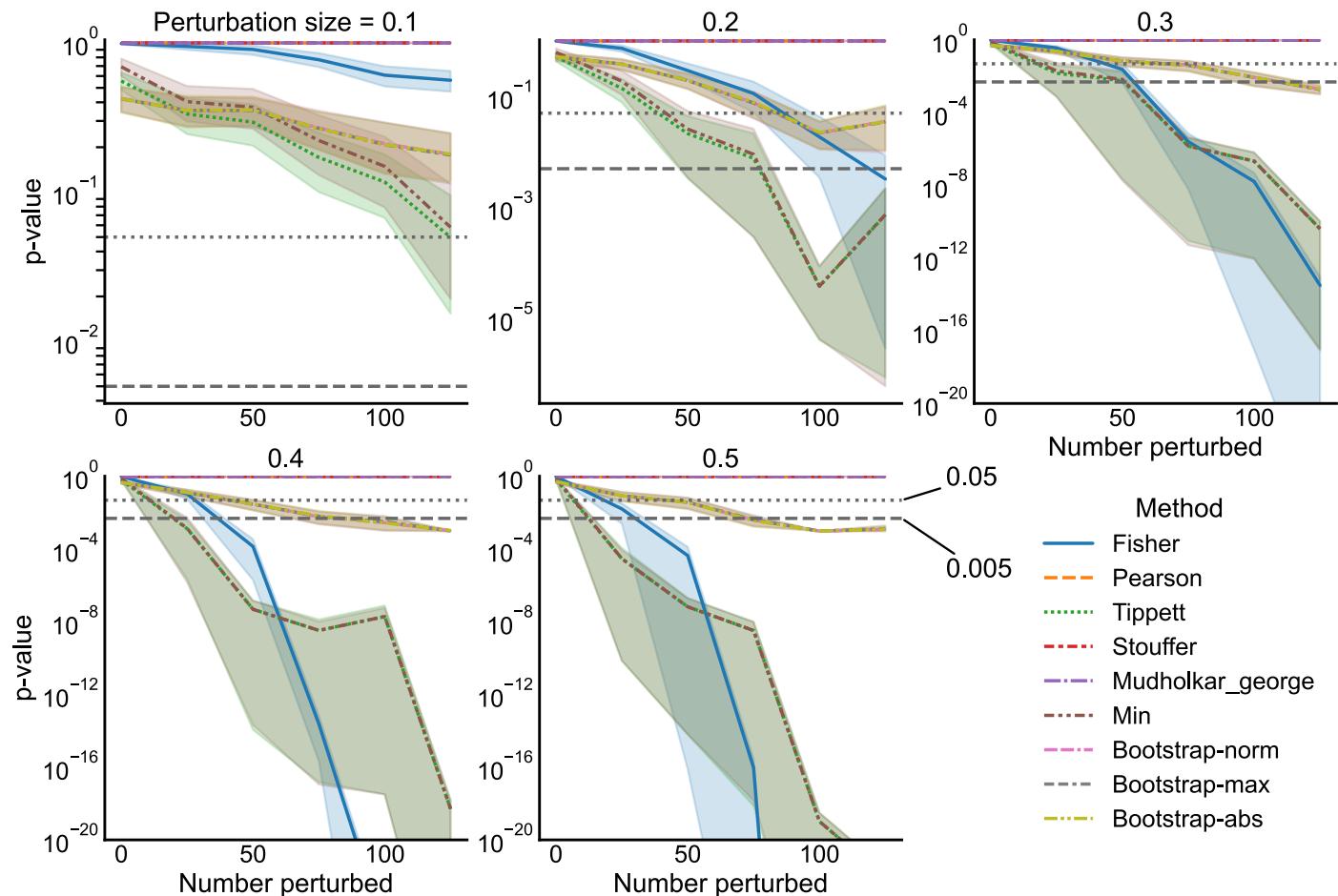
BY H. O. LANCASTER, *Rockefeller Fellow in Medicine*



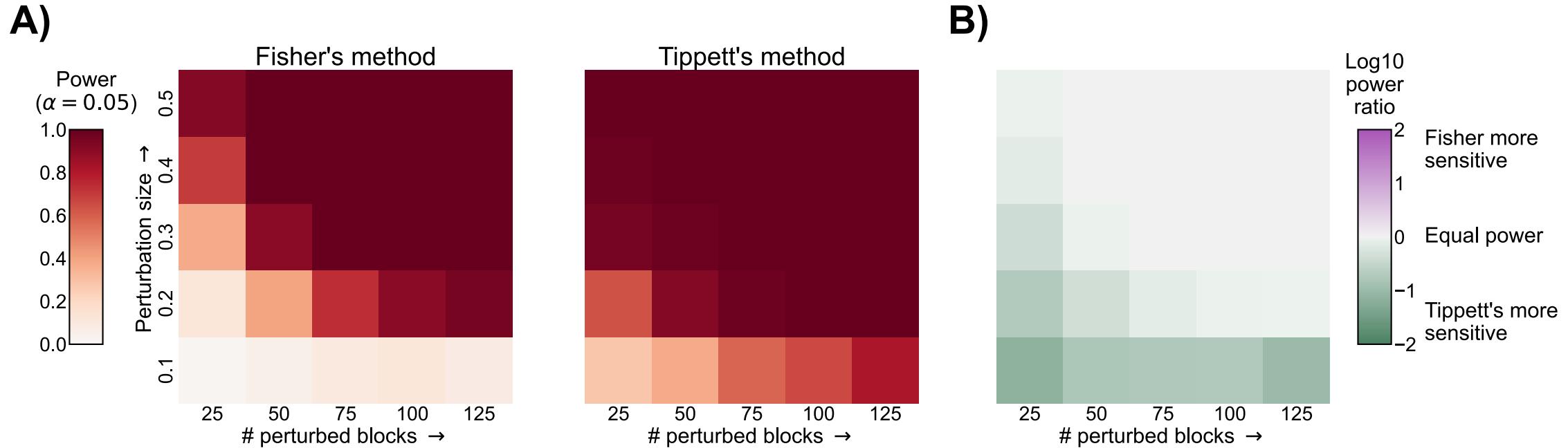
← We are trying to approximate this null distribution with something continuous $Uniform(0, 1)$

Power for combining p-values

- We perturb:
 - Some # of them (x-axis)
 - By some amount (panels)



Relative power (Fisher's vs Tippett's)



Plotting connection probabilities

