Generative network modeling reveals quantitative definitions of bilateral symmetry exhibited by a whole insect brain connectome

Benjamin D. Pedigo[1*], Mike Powell[1], Eric W. Bridgeford[1], Michael Winding[2], Carey E. Priebe[1], Joshua T. Vogelstein[1]

Abstract. Comparing connectomes can help explain how neural connectivity is related to genetics, disease, development, learning, and behavior. However, making statistical inferences about the significance and nature of differences between two networks is an open problem, and such analysis has not been extensively applied to nanoscale connectomes. Here, we investigate this problem via a case study on the bilateral symmetry of a larval *Drosophila* brain connectome. We translate notions of "bilateral symmetry" to generative models of the network structure of the left and right hemispheres, allowing us to test and refine our understanding of symmetry. We find significant differences in connection probabilities both across the entire left and right networks and between specific cell types. By rescaling connection probabilities or removing certain edges based on weight, we also present adjusted definitions of bilateral symmetry exhibited by this connectome. This work shows how statistical inferences from networks can inform the study of connectomes, facilitating future comparisons of neural structures.

1 Introduction

Connectomes – maps of neural wiring – have become increasingly important in neuroscience, and are thought to be an important window into studying how connectivity relates to neural activity, evolution, disease, genetics, and learning [1–4]. However, many of these pursuits in connectomics depend on being able to compare networks. For instance, to understand how memory relates to connectivity, one would need to map a connectome which has learned something and one which has not, and then assess whether and how the two networks are different. To understand how a gene affects connectivity, one would need to map a connectome from an organism with a genetic mutation and one from a wild-type organism, and then assess whether and how the two networks are different. Authors have advocated for comparing connectomes across the phylogenetic tree of life [3] or disease states [2]. Several recent works have already started towards this goal of comparative connectomics. Gerhard et al. [5] compared the connections in the nerve cord (the insect equivalent of a spinal cord) of the L1 and L3 stages of the larval *Drosophila melanogaster* to understand how these connections change as the animal develops. Similarly, Witvliet et al. [6] collected connectomes from *Caenorhabditis elegans* at various life stages, and examined which connections were stable and which were dynamic across development. Cook et al. [7] generated connectomes for both a male and hermaphrodite *C. elegans* worm to understand which aspects of this organism's wiring diagram differ between the sexes. Valdes-Aleman et al. [8] made genetic perturbations to different individual *D. melanogaster* fly larva, and examined how these perturbations affected the connectivity of a local circuit in the organism's nerve cord. Viewed through the lens of the wiring diagrams alone (i.e., ignoring morphology, subcellular structures, etc.), these pursuits all amount to comparing two or more networks.

In addition to those described above, one comparison that has been prevalent in the connectomics literature is to assess the degree of left/right structural similarity of a nervous system. *Bilateria* is a group of animals which have a left/right structural symmetry. This clade is thought to have emerged around 550 million years ago [9], making it one of the oldest groups of animals. Most organisms studied in neuroscience (including *C. elegans*, *D. melanogaster*, mice, rats, monkeys, and humans) are all bilaterians. While functional asymmetries in the brain have been discovered, this axis of structural symmetry is generally thought to extend to the brain [10].

Connectomic studies have investigated this structural similarity in various ways. The degree of left/right symmetry in a single connectome has often been studied as a proxy or lower bound for the

[1]Johns Hopkins University (JHU), [2]University of Cambridge.
[*] corresponding author: bpedigo@jhu.edu.

amount of sterotypy that one could expect between connectomes of different individuals. Lu et al. [11] reconstructed the connectome of the axons projecting to the interscutularis muscle on the left and right side of two individual mice. They found that the branching patterns of axons between the left and right sides within one animal were no more similar than a comparison between the two animals, and also no more similar than two random branching patterns generated by a null model. In contrast, Schlegel et al. [12] found a striking similarity between the morphologies of neurons (as measured by NBLAST [13]) in the left and right hemispheres of the *D. melanogaster* antennal lobe, and a similar level of stereotypy between the antennal lobes of two different individuals. Cook et al. [7] used the observed level of left-right variability in a *C. elegans* hermaphrodite connectome as a proxy for the amount of variability in connectivity between individuals, assuming that one should expect the connectomes of the left and right to be the same up to developmental and experiential variability. Conversely, they also point out the fact that the ASEL neuron (on the left side) projects more strongly to neuron class AWC than the analogous version on the right, verifying this difference via fluorescent labeling in another animal. These studies highlight the complicated relationship between neuroscientists and bilateral symmetry: at times, we may assume that the left and right sides of a nervous system are in some sense the same in expectation, but at other times we find marked, reproducible differences between them. To date, no study (to our knowledge) has framed this question of bilateral symmetry of connectivity as a statistical hypothesis comparing two networks.

In this work, we compare the connectivity of the left and the right hemispheres of an insect connectome from the perspective of statistical hypothesis testing. Motivated by the discussion above, in this work we make three major contributions: 1) we formally state several notions of bilateral symmetry for connectomes as statistical hypotheses, 2) we present test procedures for each of these hypotheses of bilateral symmetry, and 3) we demonstrate the utility of these tests for understanding the significance and nature of bilateral symmetry/asymmetry in the brain of a *D. melanogaster* larva. In doing so, we provide a framework and methodology for any neuroscientist wishing to compare two networks, facilitating future work in comparative connectomics. We also provide Python implementations and documentation for the statistical tests for network comparison developed in this work.

2 Results

2.1 Connectome of a larval *Drosophila* brain

Recently, authors mapped a connectome of the brain of a *D. melanogaster* larva [14]. To understand how the neurons in this brain were connected to each other, the authors first imaged this brain using electron microscopy, and then manually reconstructed each neuron and its pre- and post-synaptic contacts. This synaptic wiring diagram consists of 3,013 neurons and over 544,000 synapses. We represent this connectome as a network, with nodes representing neurons and edges representing some number of synapses between them (Figure 1). Importantly, this work yielded a complete reconstruction of both the left and right hemispheres of the brain. In order to assess bilateral symmetry, we focused on the left-to-left and right-to-right (ipsilateral) induced subgraphs. While there are conceivable ways to define bilateral symmetry which include the contralateral connections, we did not consider them here in order to restrict our methods to the more widely applicable case of two-network-sample testing. More details on how we created the networks to compare here are available in Section 4.1. This process yielded a 1,504 neuron network for the left hemisphere, and a 1,503 neuron network for the right.

We sought to understand whether these two networks were significantly different according to some definition, in order to characterize whether this brain was bilaterally symmetric. As with any statistical hypothesis test, this required that we make some modeling assumptions about the nature of the networks being compared. We stress that our subsequent results should be interpreted in light of these models and what they do (and do not) tell us about these networks (see [15] for an excellent discussion of this point in network neuroscience, and see Section 3.2 for a discussion of alternative modeling assumptions). For all of our models, we treated the networks as directed (since we knew the direction

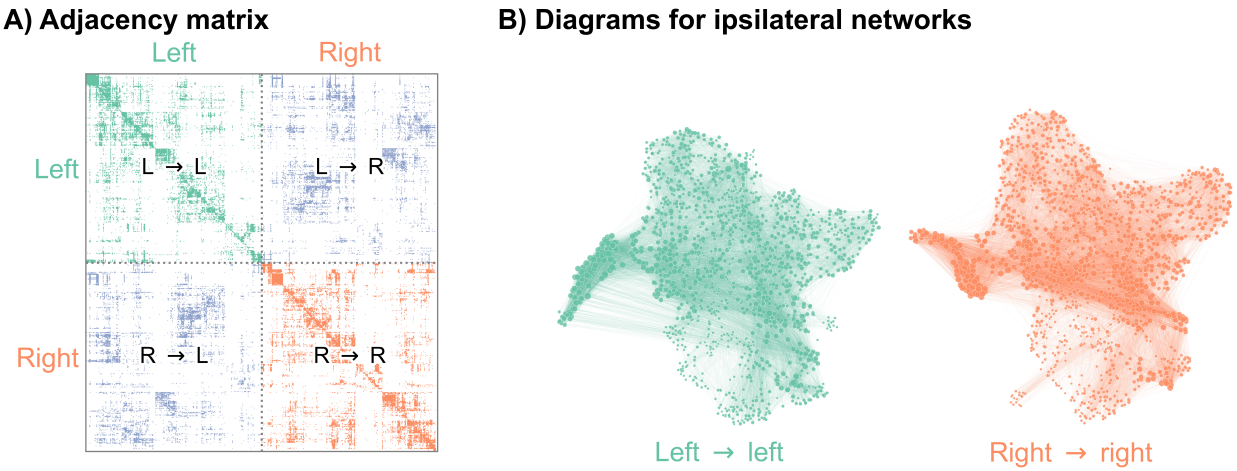


Figure 1: Visualizations of a larval *Drosophila* brain connectome from Winding et al. [14]. **A)** Adjacency matrix for the full brain connectome network, sorted by brain hemisphere. Note that we ignore the left → right and right → left (contralateral) subgraphs in this work. **B)** Network layouts for the left → left and right → right subgraphs.

of synapses), unweighted (creating an edge when there was one or more synapse between neurons unless otherwise specified), and loopless (since we ignored any observed self-loops). We made no assumptions about whether individual neurons in the left hemisphere correspond with individual neurons in the right hemisphere. Next, we detail a series of more specific models, what aspects of the networks they characterize, and how we construct a hypothesis test from each.

2.2 Density test Our first test of bilateral symmetry was based on perhaps the simplest network model, the **Erdos-Renyi (ER) model** [16, 17], which models each potential edge as independently generated with the same probability, p. Comparing two networks under the ER model amounts to simply comparing their densities (Figure 2A).

$$H_0 : p^{(L)} = p^{(R)} \text{ vs. } H_A : p^{(L)} \neq p^{(R)} \tag{2.1}$$

This comparison of probabilities can be tested using well-established statistical machinery for two-sample tests under the binomial distribution (see Section Section 4.2.2 for more details). We refer to this procedure as the **density test**.

Figure 2B shows the comparison of the network densities between the left and right hemisphere networks. The densities of the left and right are ~0.016 and ~0.017, respectively, making the density of the left ~0.93 that of the right. To determine whether this is a difference likely to be observed by chance under the ER model, we ran a two-sided Fisher's exact test, which tests whether the probabilities of two independent binomials are significantly different. This test yielded a p-value $<10^{-23}$, suggesting that we have strong evidence to reject this version of our hypothesis of bilateral symmetry. While the ratio of the estimated densities is only ~0.93, this extremely small p-value resulted from the large sample size for this comparison: there are 2,260,512 and 2,257,506 potential edges on the left and right, respectively.

To our knowledge, when neuroscientists have considered the question of bilateral symmetry, they have not meant such a simple comparison of network densities. In many ways, the ER model is too simple to be an interesting description of connectome structure. However, it is also striking that perhaps the simplest network comparison produced a significant difference between brain hemispheres for this brain. It is unclear whether this difference in densities is biological (e.g., a result of slightly differing rates of development for this individual), an artifact of how the data was collected (e.g., technological limitations causing slightly lower reconstruction rates on the left hemisphere), or something else entirely. Still, in addition to highlighting a simple departure from symmetry in this dataset, the density test result also provides important considerations for other tests. More complicated models of symmetry could

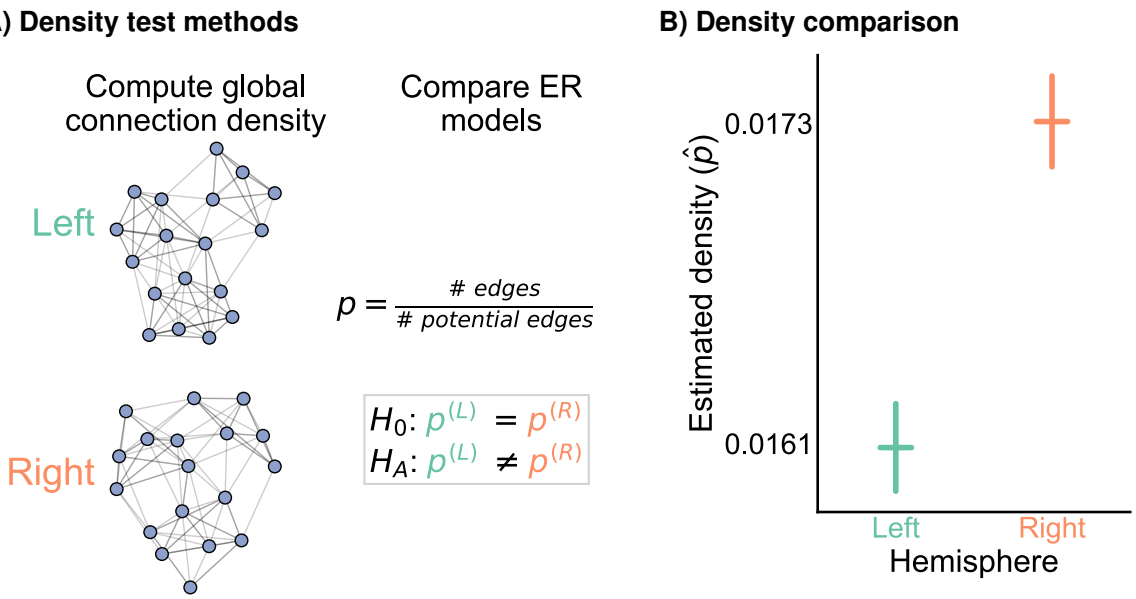


Figure 2: Comparison of left and right hemisphere networks via the density test. **A)** Diagram of the methods used for testing based on the network density. See Section 4.2.2 for more details. **B)** The estimated density $\hat{p}$ (probability of any edge averaged across the entire network) for the left hemisphere is ~0.016, while for the right it is ~0.017–this makes the left density ~0.93 that of the right. Vertical lines denote 95% confidence intervals for this estimated parameter $\hat{p}$. The p-value for testing the null hypothesis that these densities are the same is $<10^{-23}$ (two-sided Fisher's exact test), meaning very strong evidence to reject the null that the left and right hemisphere have the same density.

compare other network statistics—say, the clustering coefficients, the number of triangles, and so on. These statistics, as well as the model-based parameters we will consider in this paper, are strongly related to the network density [18, 19]. Thus, if the densities are different, it is likely that tests based on any of these other test statistics will also reject the null hypothesis of bilateral symmetry. Later, we describe methods for adjusting for a difference in density in other tests for bilateral symmetry.

2.3 Group connection test To understand whether this broad difference between the hemispheres can be localized to a specific set of connections, we next tested bilateral symmetry by making an assumption that the left and right hemispheres both come from a **stochastic block model (SBM)**. Under the SBM, each neuron is assigned to a group, and the probability of any potential edge is a function of the groups to which the source and target neurons belong. For instance, the probability of a connection from a neuron in group k to a neuron in group l is set by the parameter B_{kl}, where B is a $K \times K$ matrix of connection probabilities if there are K groups. Here, we used broad cell type categorizations from Winding et al. [14] to determine each neuron's group. Alternatively, there are many methods for estimating these assignments to groups for each neuron which we do not explore here (see Section 3.2 for discussion on this point). Under the SBM with a fixed group assignment for each node, testing for bilateral symmetry amounts to testing whether the group-to-group connection probability matrices, $B^{(L)}$ and $B^{(R)}$, are the same.

$$H_0 : B^{(L)} = B^{(R)} \text{ vs. } H_A : B^{(L)} \neq B^{(R)} \tag{2.2}$$

Rather than having to compare one probability as in Equation 2.1, we were interested in comparing all K^2 group-to-group connection probabilities between the SBM models for the left and right hemispheres. We developed a novel statistical hypothesis test for this comparison, which uses many Fisher's exact tests to compare each of the group-to-group connection probabilities, followed by appropriate correction for multiple comparisons (when examining the individual group-to-group connections) or combination of p-values (when assessing the overall null hypothesis in Equation 2.2). Details on the methodology used

here is provided in Section 4.2.3, and is shown as a schematic in Figure 3A. We refer to this procedure as the **group connection test**.

Figure 3B shows both of the estimated group-to-group probability matrices, $\hat{B}^{(L)}$ and $\hat{B}^{(R)}$. From a visual comparison of $\hat{B}^{(L)}$ and $\hat{B}^{(R)}$, the group-to-group connection probabilities appear qualitatively similar. Note also that some group-to-group connection probabilities are zero, making it nonsensical to do a comparison of probabilities. We highlight these elements in the matrices with explicit "0"s, and note that we did not run the corresponding test in these cases. Figure 3C shows the p-values from all 284 tests that were run to compare each element of these two matrices. After multiple comparisons correction, six tests produced p-values less than $\alpha = 0.05$, indicating that we could reject the null hypothesis that those specific connection probabilities are the same between the two hemispheres. We also combined all (uncorrected) p-values, yielding an overall p-value for the entire null hypothesis (Equation 2.2) of equivalence of group-to-group connection probabilities of $<10^{-7}$.

Taken together, these results suggest that while the group-to-group connections are roughly similar between the two hemispheres, they are not the same under this model. Notably, there are six group-to-group connections which were significantly different: Kenyon cells (KCs) $\rightarrow$ KCs, projection neurons (PNs) $\rightarrow$ lateral horn neurons (LHNs), somatosensory PNs $\rightarrow$ somatosensory PNs, others (uncategorized neurons) $\rightarrow$ LHNs, LHNs $\rightarrow$ others, and others $\rightarrow$ others. Knowing the neuron groups which are wired significantly differently between the two hemispheres highlights the interpretability of this test. If a neuroscientist wanted to study mechanisms which could cause bilateral asymmetries in the brain, these six group-to-group connections would be prime candidates for investigation.

However, in Section 2.2, we saw that the densities of the two networks are significantly different. p, the density of the network, can be thought of as a weighted average of the individual group-to-group connection probabilities, B. Should we then be surprised that if the density is different, the group-to-group connection probabilities are, too? Interestingly, for all the group-to-group connection probabilities which are different, the probability on the right hemisphere (which has the greater density) is higher (Figure 3D). We consider this phenomenon in the next section.

2.4 Density-adjusted group connection test

Next, we examined whether the group-to-group connection probabilities on the right are simply a “scaled-up” version of those on the left. Figure 3D showed that for all the individual connections which are significant, the connection probability on the right hemisphere is higher. This is consistent with the hypothesis stated above, which predicts that the connection probabilities in $B^{(R)}$ should be consistently higher than those in $B^{(L)}$.

We thus created a test for this notion of bilateral symmetry in group-to-group connections (up to a density adjustment):

$$H_0 : B^{(L)} = cB^{(R)} \text{ vs. } H_A : B^{(L)} \neq cB^{(R)} \tag{2.3}$$

To implement this hypothesis test, we first computed the density correcting constant (c in Equation 2.3), which is simply the ratio of the left to the right hemisphere densities, finding that $c \approx 0.93$. Then, we replaced each of the component tests in the group connection test with a modified version of Fisher’s exact test for non-unity probability ratios (see Section 4.2.4 for more details). We refer to this procedure as the **density-adjusted group connection test** (Figure 4A). The p-values for each of the component tests for the density-adjusted group connection test are shown in Figure 4B. After correction for multiple comparisons, there are two group-to-group connections which are significantly different (at significance level 0.05): KCs $\rightarrow$ convergence neurons (CNs) and KCs $\rightarrow$ KCs. Thus, all remaining significant differences between the hemispheres under this version of the SBM are associated with the Kenyon cells.

2.5 Removing Kenyon cells

Based on the results of Figure 4C, we sought to verify that the remaining differences in group-to-group connection probabilities after adjusting for a difference in density can be explained by asymmetry that is isolated to the Kenyon cells. To confirm this, we simply removed the

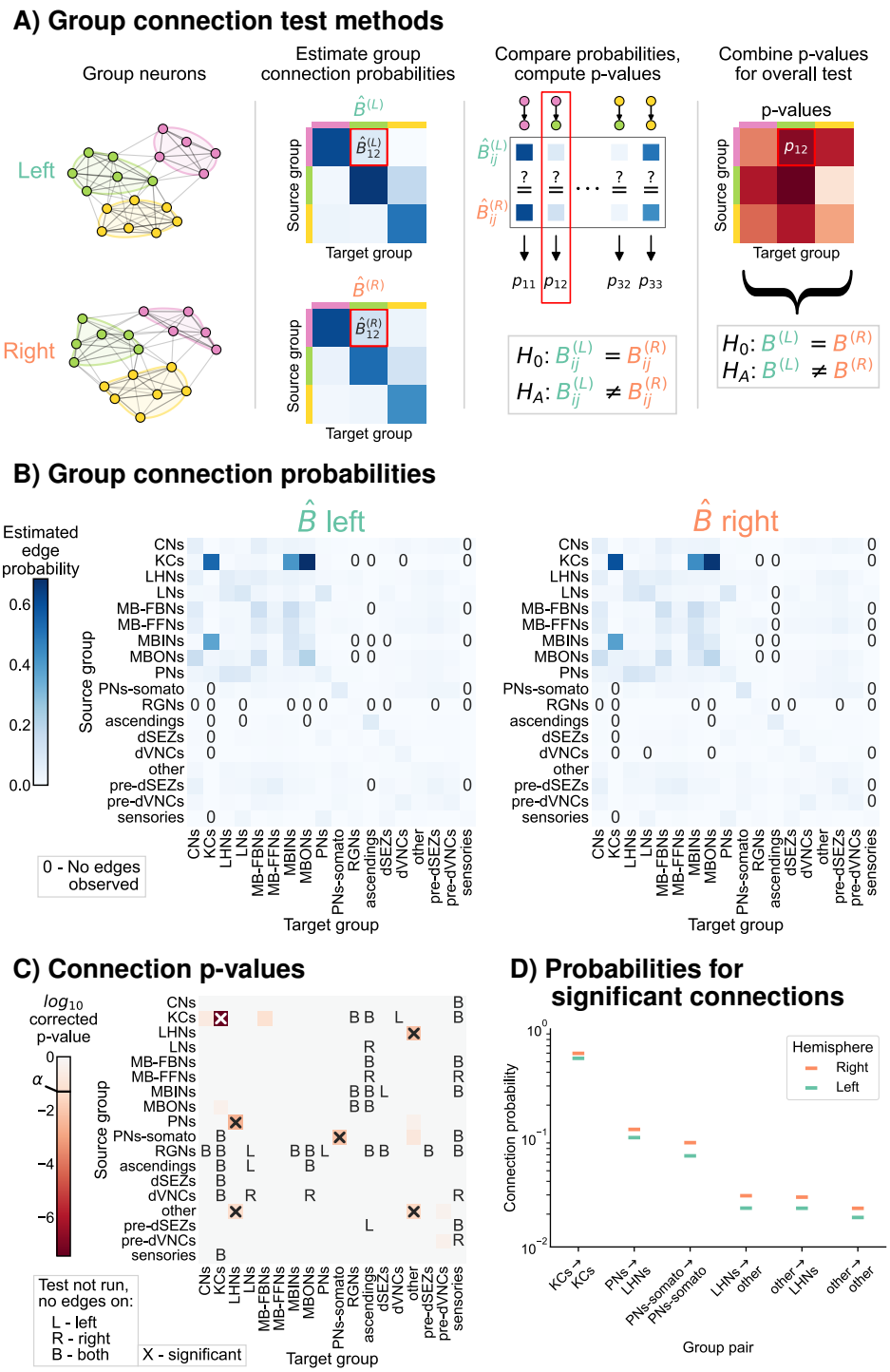


Figure 3: Comparison of left and right hemisphere networks via the group connection test. **A)** Description of methodology for the group connection test. See Section 4.2.3 for more details. **B)** Estimated group-to-group connection probabilities for both hemispheres. Note that they appear qualitatively similar. Estimated probabilities which are zero (no edge was present between that pair of groups) are indicated with a "0" in those cells. **C)** p-values (after multiple comparisons correction) for each hypothesis test between individual elements of the connection probability matrices. Each cell represents a test for whether a specific group-to-group connection probability is the same on the left and right sides. "X" denotes a significant p-value after multiple comparisons correction, with significance level $\alpha = 0.05$. "B" indicates that a test was not run since the estimated probability was zero on both hemispheres, "L" indicates this was the case on the left only, and "R" that it was the case on the right only. The individual (uncorrected) p-values were combined using Tippett's method, resulting in an overall p-value (for the null hypothesis that the two group connection probability *matrices* are the same) of $<10^{-7}$. **D)** Comparison of estimated group-to-group connection probabilities for the group pairs that are significantly different. In each case, the connection probability on the right hemisphere is higher.

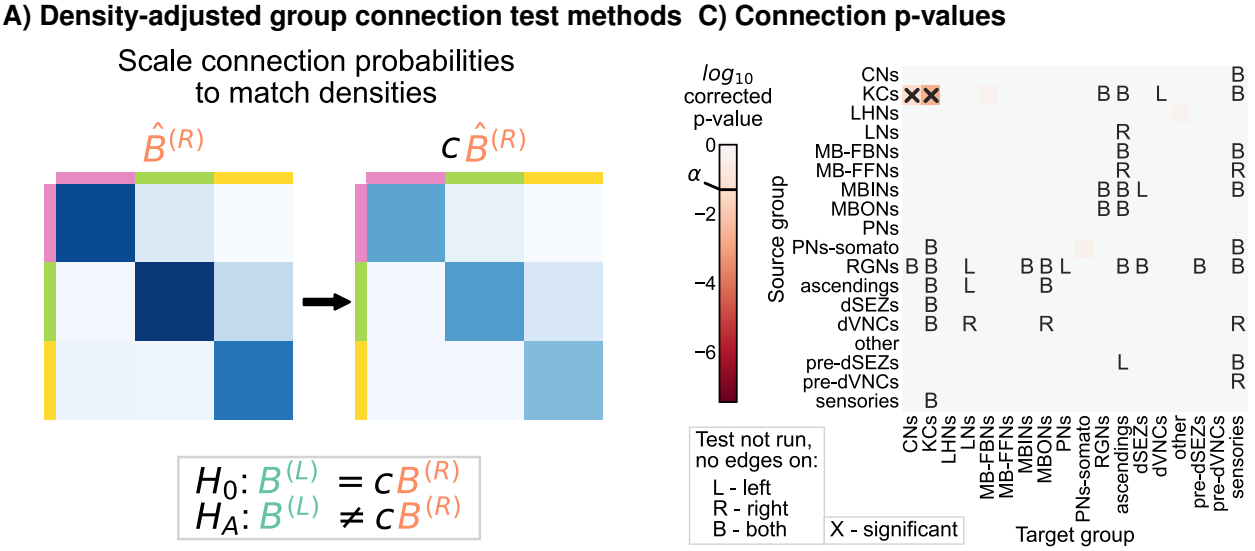


Figure 4: Comparison of left and right hemisphere networks via the density-adjusted group connection test. **A)** Description of methodology for adjusting for a density difference between the two stochastic block models. See Section 4.2.3 for more details. The adjustment factor (ratio of the left to the right density), c, is ~0.93. **B)** P-values for each group-to-group comparison after adjusting for a global density difference. P-values are shown after correcting for multiple comparisons. Note that there are two significant p-values, and both are in group connections incident to Kenyon cells. These individual (uncorrected) p-values were combined using Tippett's method, resulting in an overall p-value (for the null hypothesis that the two group connection probability *matrices* are the same after correcting for the density difference) of $<10^{-2}$.

Kenyon cells (i.e., all Kenyon cell nodes and edges to or from those nodes) from both the left and right hemisphere networks, and then re-ran each of the tests for bilateral symmetry presented here (Figure 5A). We observed significant differences between the left and right hemispheres for the density and group connection tests when excluding Kenyon cells, yielding p-values of $<10^{-27}$ and $<10^{-2}$, respectively (Figure 5B and C). However, for the density-adjusted group connection test, the p-value was ~0.51, indicating that we no longer rejected bilateral symmetry under this definition when the Kenyon cells are excluded from the analysis (Figure 5D). This sequence of results suggests that the difference between the left and right hemispheres (at least in terms of the high-level network statistics studied here) can be explained as the combination of a global effect (the difference in density) and a cell-type-specific effect (the difference in Kenyon cell projection probabilities).

It is noteworthy that the Kenyon cells were the specific cell type where we detected asymmetry after correcting for the density difference. Kenyon cells are involved in associative learning in *Drosophila* and other insects [20–22]. Other studies have suggested that certain connections (specifically from antennal lobe projection neurons to Kenyon cells) are random [22, 23]. The marked lack of symmetry we observed specifically in the Kenyon cells in the current study could be the result of these features, which suggest their uniquely non-stereotyped patterns of connectivity in this nervous system.

2.6 Edge weight thresholds Next, we sought to examine how the definition of an edge used to construct our binary network affects the degree of symmetry under each of the definitions considered here. For the networks considered in the previous sections, we considered an edge to exist if one or more synapses from neurons i to j were in the dataset. To understand how our analysis might change based on this assumption, we considered two types of edge weight threshold schemes for creating a binary network before testing: the first based simply on a threshold on the number of synapses, and the second based on a threshold of the proportion of a downstream neuron's input (Figure 6A). By varying the threshold in both schemes, we were able to evaluate many hypotheses about bilateral symmetry, where higher thresholds meant that we only considered the symmetry present in strong

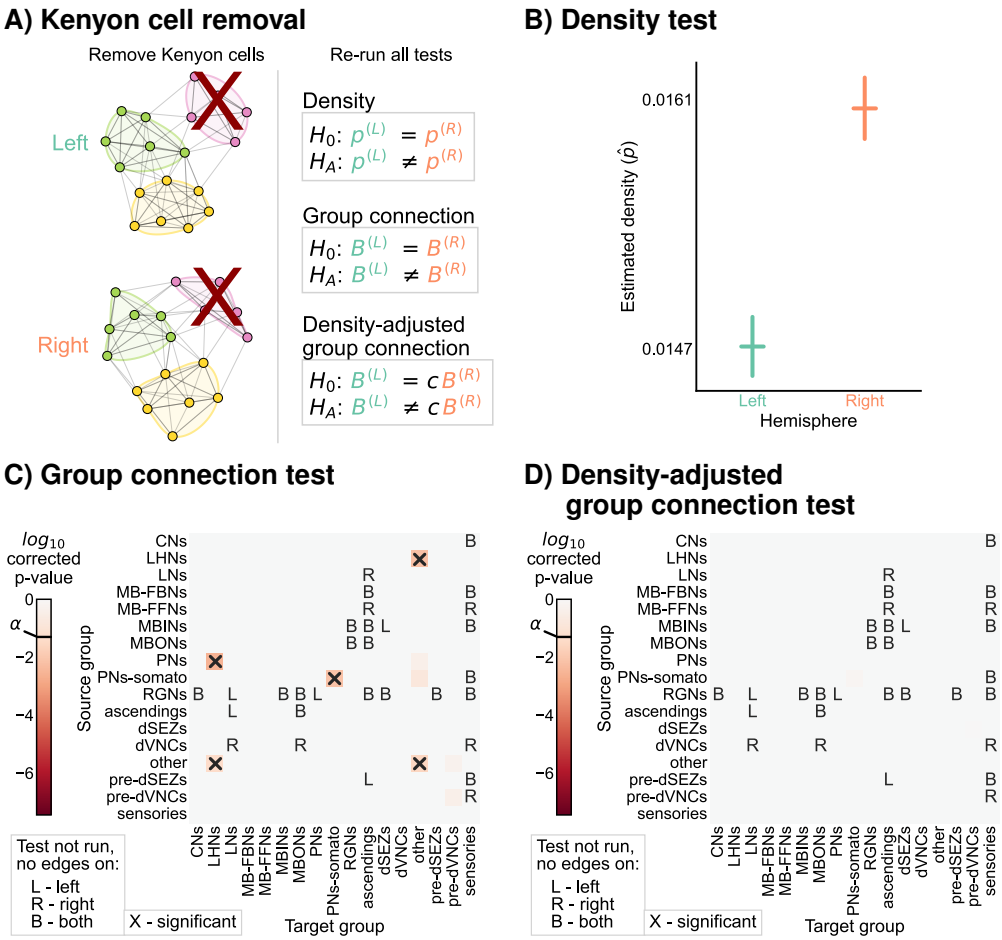


Figure 5: Comparison of left and right hemisphere networks when not including Kenyon cells. **A)** Diagram of the methods used, indicating that Kenyon cells (and any incident edges) were simply removed from the network, and all previously mentioned tests were run again. **B)** Comparison of network densities, as in Figure 2B. The p-value for this comparison is $<10^{-27}$, indicating very strong evidence to reject the null that the two networks share the same density. **C)** Comparison of group-to-group connection probabilities, as in Figure 3C. P-values are shown for each group-to-group connection comparison (after multiple comparison correction). The (uncorrected) p-values were combined to yield an overall p-value of $<10^{-2}$, showing evidence that the group connection probabilities are not the same even after removing Kenyon cells. **D)** Comparison of group-to-group connection probabilities after density adjustment, as in Figure 4C. P-values are shown for each group-to-group connection comparison (after multiple comparison correction). Note that there are no longer any significantly different connections. The (uncorrected) p-values were combined to yield an overall p-value of ~0.51. After removing Kenyon cells, there is no longer evidence to reject the null that the group connection probabilities are the same.

Test method	**Model**	H_0 **(vs.** $H_A \neq$**)**	**KCs**	**p-value**
Density test	ER	$p^{(L)} = p^{(R)}$	+	$<10^{-23}$
Group connection test	SBM	$B^{(L)} = B^{(R)}$	+	$<10^{-7}$
Density-adjusted group connection test	DA-SBM	$B^{(L)} = cB^{(R)}$	+	$<10^{-2}$
Density test	ER	$p^{(L)} = p^{(R)}$	-	$<10^{-27}$
Group connection test	SBM	$B^{(L)} = B^{(R)}$	-	$<10^{-2}$
Density-adjusted group connection test	DA-SBM	$B^{(L)} = cB^{(R)}$	-	~0.51

Table 1: Summary of tests, models, hypotheses, whether Kenyon cells (KCs) were included, and the resulting p-values for each evaluation of bilateral symmetry.

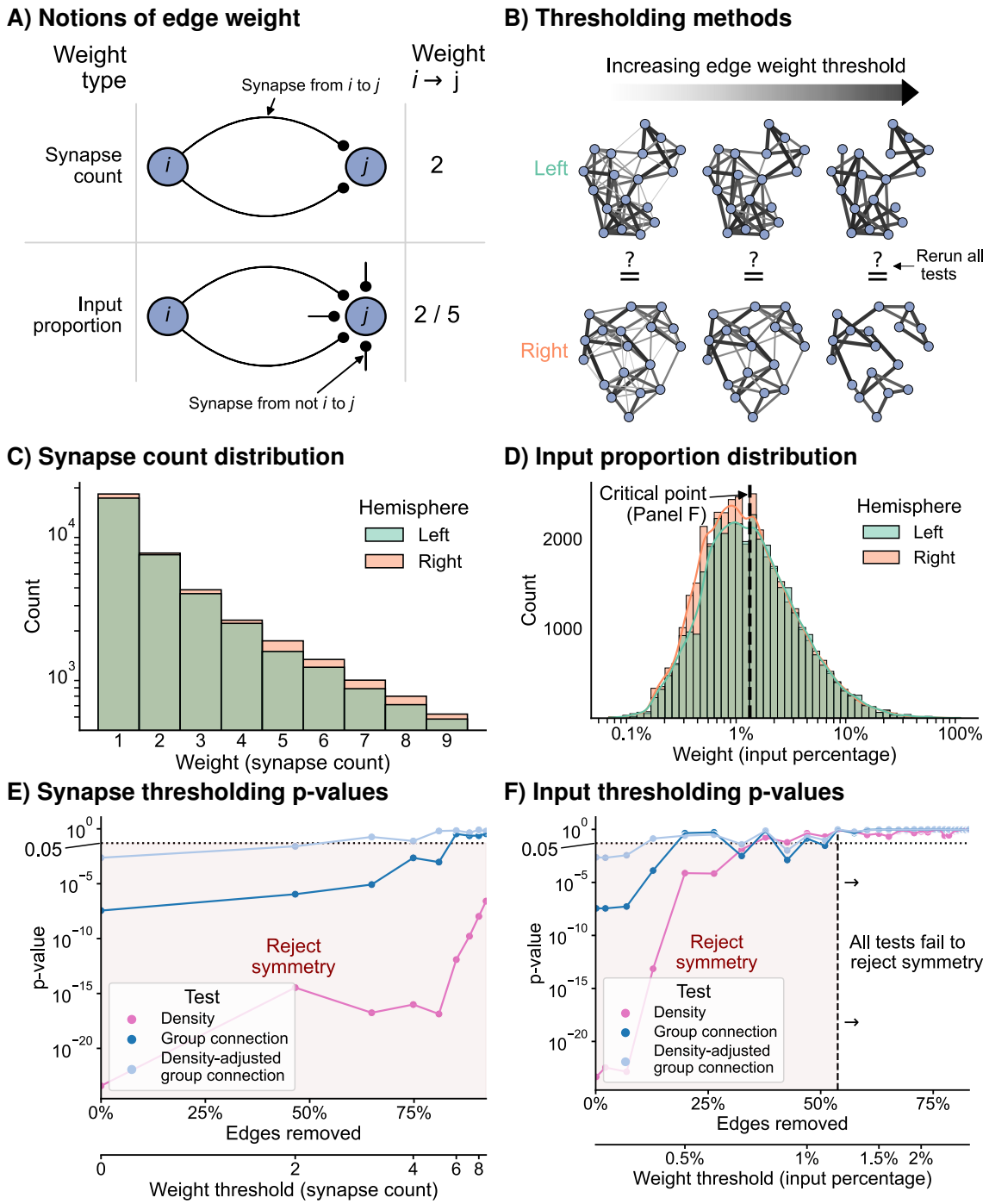


Figure 6: The effect of edge weight threshold on the significance level for each of the tests of bilateral symmetry. Diagrams of **A)** two notions of edge weight, and **B)** application of edge weight thresholds to examine bilateral symmetry. See Section 4.3 for more explanation. **C)** Distribution of synapse count edge weights. The right hemisphere consistently has more edges in each synapse count bin. **D)** Distribution of input percentage edge weights. The right hemisphere has more edges in the lower ($<1\%$) portion of this distribution, but the hemispheres match well for high edge weights. **E)** p-values for each test after synapse count thresholding, plotted as a function of the percentage of edges which are removed from the networks, as well as the corresponding weight threshold (lower x-axis). The p-values for all tests generally increased as a function of synapse count threshold, but the density test never reached a p-value >0.05 over this range of thresholds. **F)** p-values for each test after input percentage thresholding, plotted as a function of the percentage of edges which were removed from the networks, as well as the corresponding weight threshold (lower x-axis). Note that all tests yielded insignificant (>0.05) p-values after a threshold of around 1.25% input proportion. Compared to the results in E), thresholding based on input percentage reached insignificant p-values faster as a function of the total amount of edges removed for all tests.

edges (Figure 6B).

Before running the tests for each of these notions of symmetry, we first examined the distributions of edge weights to get a sense for how we should expect these tests to perform. Figure 6C and D display the distribution (total count) for the synapse count or input proportion edge weights, respectively. The right hemisphere has more connections than the left for all synapse count values (Figure 6C), hinting that the density of the right hemisphere will be slightly higher for any potential edge weight threshold using this definition. Conversely, the distributions of weights as an input percentage shows a different trend. For edge weights less than ~1.25%, the right appears to have more edges, but past this threshold, the counts of edges between left and right appear comparable (Figure 6D).

Figure 6E and F show the effect of varying these thresholds on the p-values from each of our tests of bilateral symmetry. We observed that for either thresholding scheme (synapse count or input proportion), the p-value for each test generally increased as a function of the threshold—in other words, the left and right hemisphere networks became less significantly different (under the definitions of "different" we have presented here) as the edge weight threshold increased. Previous works have shown that higher-weight edges are more likely to have that corresponding edge present on the other side of the nervous system [5, 24]. Here, we provide evidence that considering networks formed from only strong edges also decreases asymmetry at a broad, network-wide level.

To make these two thresholding schemes more comparable, we also examined these results as a function of the proportion of edges from the original network which that threshold removed (Figure 6E and F, lower x-axis). We found that when thresholding based on synapse counts, the majority (~60%) of the edges of the networks need to be removed for any test (in this case the density-adjusted group connection test) to yield non-significant p-values. Conversely, for the thresholds based on input proportion, the density-adjusted group connection test yielded a p-value greater than 0.05 after removing only the weakest ~20$\%$ of edges. Strikingly, we observed that when considering only the strongest ~60% of edges in terms of input proportion, even the density test had a high p-value (>0.05), while for the synapse-based thresholds we examined, this never occurred.

These findings are consistent with previous work in connectomics which has hinted at the importance of input proportion as a meaningful "edge weight." Gerhard et al. [5] compared the connectivity of select neurons in the nerve cord between L1 and L3 stages of the larva. They observed that while the number of synapses from the mdIV cell type onto various nerve cord local neurons can grow ~3-10 fold from L1 to L3, the proportion of that downstream neuron's input stays relatively conserved. Based on this finding, the authors suggested that perhaps the nervous system evolved to keep this parameter constant as the organism develops. An analysis of wiring in the olfactory system of the adult *Drosophila* suggested a similar interpretation after examining a projection neuron cell type with an asymmetric number of neurons on the two sides of the brain [25]. Here, we provide further evidence based on the entire brain of the *Drosophila* larva that while the left and right hemispheres may appear significantly different when considering all observed connections, the networks formed by only the strongest edges (especially in terms of input proportion) are not significantly different between the hemispheres when viewed through the lens of the models considered in this work.

3 Discussion

3.1 Summary We began with what was at its face a very simple question: is the connectivity on the left and the right side of this brain "different?" We then described several ways that one could mathematically formalize notions of "different" from the perspective of network model parameters: difference in density of connections across the entire network (Section 2.2), difference in group connection probabilities (Section 2.3), or difference in group connection probabilities while adjusting for a difference in density (Section 2.4). We proposed a test procedure corresponding with each of these notions, relying on well established statistical techniques for evaluating contingency tables and combining p-values to construct our tests. The results of these different test procedures varied markedly (Table 1). Specif-

ically, we saw that the network densities were significantly different between the hemispheres. The group connection test also detected a difference, highlighting six group-to-group connections which had significantly differing connection probabilities when comparing the hemispheres. However, when we added an adjustment to the group connection comparison to account for the difference in network density, this test had only two significant group connections, and both were projections from the Kenyon cells. Thus, the asymmetry observed (at least when viewed through the lens of these high-level network statistics) between the hemispheres can be thought of as a global density difference in addition to a cell type-specific effect shown in the Kenyon cells. We confirmed this finding by simply removing the Kenyon cells, and showing that the density-adjusted group connection test no longer rejected (Section 2.5). Finally, we examined whether the left and right hemisphere networks would become less dissimilar when only high-edge-weight edges were considered (Section 2.6). We found that whether thresholding based on number of synapses or the proportion of input to the post-synaptic neuron, p-values generally increased for each test (i.e., less significant asymmetry was detected) as the edge weight threshold grew. However, we observed that thresholds based on neuron input proportion could achieve symmetry while removing fewer (only 20% for some tests) edges. These results are consistent with the idea that the nervous system evolved to preserve a relative balance of inputs to individual neurons, which has been suggested by previous studies on specific subcircuits in the larval and adult *Drosophila* nervous system [5, 25, 26].

3.2 Limitations As with any statistical inference, our conclusions are valid under particular model assumptions. Therefore, it is important to highlight the assumptions which motivated each of our tests in order to understand what each p-value means (and what it does not). We highlight several of these assumptions below, and comment on alternative assumptions that one could make in each case.

What model? First, while we motivated the tests presented here by assuming that some statistical model produced the connectivity of the left and the right hemispheres, these models do not literally describe the process which generated these networks. However, without knowledge of how genes and development give rise to the connectome, we know of no more correct model for how this connectome was generated [1, 6, 27] (and even this would still be just a model). Without an agreed upon definition of bilateral symmetry, we chose to start from the simplest definition of what one *could* mean by bilateral symmetry. From this simplest network model, we iteratively added complexity to the definition of bilateral symmetry until we found the simplest model for which the *Drosophila* larva connectome displayed no significant asymmetry.

However, many other network models could have been applied to examine different definitions of bilateral symmetry. Tests based on the random dot product graph model [28–30] would allow us to compare connection probabilities between hemispheres without assuming that neurons belong to a finite number of groups. Bravo-Hermsdorff et al. [31] showed that a two-network-sample test could be constructed from subgraph counts, which they argue characterize a network's "texture" rather than its "backbone" as studied in this work. We also did not use network models that incorporate edge weights, as two-network-sample tests for this case are even less developed than for the unweighted case. Further, a variety of neuroscience-specific network models (such as those which incorporate spatial information) have been proposed [15]. Nevertheless, we note that even if one is concerned with these more elaborate notions of symmetry, they are still related to the simple models studied here. For instance, the network density would affect a network's latent positions under the random dot product graph model, as well as the count of any possible subgraph. Thus, even if one prefers a different definition of bilateral symmetry, the definitions presented here were worth testing.

What is a cell type? Second, even if these networks were generated from SBMs, alternative groupings of neurons could have been used. We used broad cell type categorizations from previous literature [14] to partition our network into groups. However, we could have used a coarser partition, categorizing

neurons as sensory, interneuron, and descending/output. Conversely, we could have used a finer partition, splitting the cell types used here into subgroups (such as whether a sensory neuron receives odor or visual information). Thus, the results presented for any group connection test need to be interpreted in terms of the specific cell type groupings used.

Further, a rich literature exists on *inferring* the partition for an SBM from the observed connectivity [32–37] - this is one perspective for clustering neurons based on their observed connectivity, much like clustering procedures are used to predict meaningful groups of neurons based on morphology, activity, or gene expression. Applying these techniques to a connectome would yield alternative groupings of neurons (as in Winding et al. [14]) to use for a group connection test, which again, could change its conclusions. However, this approach requires further study, as it introduces a new source of uncertainty since more model parameters are estimated from the data.

What about neuron pairs? Third, we assumed that the two networks we observed were *unmatched* – that is, the tests we applied did not use any pairing of individual neurons between hemispheres. In *Drosophila*, this 1-to-1 neuron correspondence is known to exist for most neurons, particularly in the larva. GAL-4 lines are able to reliably label bilateral neuron pairs on the basis of their gene expression [38, 39]. These neurons tend to be similar in terms of their morphology and their connectivity [5, 12, 14, 24, 39–41]. Methods which use this pairing (e.g., [28, 42, 43], as well as tests based on correlated ER and SBM models) would be able to evaluate symmetry in light of edge correspondences between the two networks, and could have higher power at detecting certain asymmetries. However, these methods assume that the matching of nodes is *perfect* and *complete*–if even one neuron pairing is a mistake, or if even one neuron does not have a partner in the opposite hemisphere, then these tests could be invalid or inapplicable. We note that graph matching techniques could estimate a correspondence between nodes for all neurons [14, 40, 44–46]; however, the statistical consequences of first learning this (likely imperfect) alignment prior to using a method which assumes the alignment is known and exact have not been thoroughly studied, so we did not explore it further here.

3.3 Outlook We presented the first statistical comparison of bilateral networks in a neuron-level brain connectome. While we focused on the larval *Drosophila* brain connectome, these techniques could be applied to future connectomes to evaluate bilateral symmetry in other individuals or organisms. More generally, we presented several notions that can be used to compare two networks, a particularly relevant problem in the current age of connectomics. Human (macroscale) connectomics has seen an explosion in the number of network samples that can be obtained, allowing for different approaches for comparing connectomes across populations, from simple comparisons of edges [47] to low-rank and sparse regressions across networks [48]. However, nanoscale connectomics is still technologically limited in its acquisition rate, often to only one or at best a few (< 10, e.g., [6]) individuals for a given experiment. Nevertheless, we wish to make valid inferences and comparisons between these connectomes [1–4]. The framework for two-network-sample testing presented here will facilitate these kinds of comparisons. To make these comparisons more practical to neuroscientists, we demonstrated the importance of adjustments to simple null hypotheses—as we saw, even a difference in something as simple as a network density can be related to other network comparisons. For example, take the problem of comparing the connectome of the larval and adult *Drosophila*. Since the adult *Drosophila* brain has orders of magnitude more nodes [14, 49, 50], the density of this network is likely to be smaller than that of the larva. Therefore, we may want to consider a more subtle question—are the connectomes of the adult and larva different (and if so, how) after adjusting for this difference in density? These kinds of biologically-motivated adjustments to out-of-the-box statistical hypotheses will be key to drawing valid inferences from connectomes which are also relevant to meaningful questions in neuroscience.

4 Methods

4.1 Network construction Here, we explain how we generated networks for the bilateral symmetry comparison. We started from a network of all neurons in the brain and sensory neurons which project into it for a larval *Drosophila* [14]. As in Winding et al. [14], we removed neurons which were considered partially differentiated. From this network, we selected only the left-to-left (ipsilateral) induced subgraph, and likewise for the right-to-right. We ignored a pair of neurons which had no left/right designation, as their cell bodies lie on the midline [14]. To ensure we had fully connected networks on either hemisphere, we took the largest weakly connected component of neurons on the left, and likewise on the right.

With this selection for our nodes of interest, we then choose our set of edges to be:

- *Unweighted*: we only considered the presence or absence of a connection, creating a binary network. For most analyses except where explicitly indicated, this meant we considered an edge to exist if there was at least one synapse from the source to the target neuron. For this connectome, four edge types are available: axo-axonic, axo-dendritic, dendro-dendritic, and dendro-axonic. We made no distinction between these four edge types when constructing the binary networks. One could consider notions of bilateral symmetry for a weighted network, but we focused on the unweighted case for simplicity and the fact that most network models are for binary networks. We studied the effect of varying the edge weight requirement (i.e., the threshold) for an edge to exist in Section 2.6.
- *Directed*: we allow for a distinction between edges which go from neuron i (presynaptic) to neuron j (postsynaptic) and the reverse.
- Loopless: we remove any edges which go from neuron i to neuron i, as the theory on network testing typically makes this assumption. We note that while ~18% of neurons have a connection to themselves, these self-loops comprise only ~0.7% of edges.

When comparing two networks, methods may make differing assumptions about the nature of the two networks being compared. One of the most important is whether the method assumes a correspondence between nodes [51]. Some methods (**matched** comparisons, also called known node-correspondence) require that the two networks being compared have the same number of nodes, and that for each node in network 1, there is a known node in network 2 which corresponds to it. Other methods (**unmatched** comparisons, also called unknown node-correspondence) do not have this requirement. To make an analogy to the classical statistical literature on two-sample testing, this distinction is similar to that between an unpaired (unmatched) and a paired (matched) t-test. We focused on the unmatched case in this work, where we say nothing about whether any neurons on the left correspond with any specific neurons on the right.

4.2 Two-network-sample testing Here, we describe in more detail the methods used to evaluate bilateral symmetry, each of which is based on some generative statistical model for the network. For each model, we formally define the model, describe how its parameters can be estimated from observed data, and then explain the test procedure motivated by the model. A more thorough review of these models can be found in Chung et al. [52].

4.2.1 Independent edge random networks Many statistical network models fall under the umbrella of independent edge random networks, sometimes called the Inhomogeneous Erdos-Renyi model. Under this model, the elements of the network's adjacency matrix A are sampled independently from a Bernoulli distribution:

$$A_{ij} \sim Bernoulli(P_{ij})$$

If n is the number of nodes, the matrix P is a $n \times n$ matrix of probabilities with elements in $[0, 1]$. Depending on how the matrix P is constructed, we can create different models. We next describe several of these choices. Note that for each model, we assume that there are no loops, or in other

words the diagonal of the matrix P will always be set to zero.

4.2.2 Erdos-Renyi model and density testing

Perhaps the simplest model of a network is the Erdos-Renyi (ER) model. 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. Thus, for all $(i, j), i \neq j$, with i and j both running from $1...n$, the probability of the edge (i, j) occurring is:

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

where p is the global connection probability.

Thus, for this model, the only parameter of interest is the global connection probability, p. This is sometimes also referred to as the **network density**. For a directed, loopless network, with n nodes, there are $n(n-1)$ unique potential edges (since we ignore the n elements on the diagonal of the adjacency matrix). If the observed network A has m total edges, then the estimated density is simply

$$\hat{p} = \frac{m}{n(n-1)}.$$

In order to compare two networks $A^{(L)}$ and $A^{(R)}$ under this model, we simply need to compute these estimated network densities ($\hat{p}^{(L)}$ and $\hat{p}^{(R)}$), and then run a statistical test to see if these densities are significantly different. Under this model, the total number of edges m comes from a $Binomial(n(n-1), p)$ distribution. This is because the number of edges is the sum of independent Bernoulli trials with the same probability. If $m^{(L)}$ is the number of edges on the left hemisphere, and $m^{(R)}$ is the number of edges on the right, then we have:

$$m^{(L)} \sim Binomial(n^{(L)}(n^{(L)} - 1), p^{(L)})$$

and independently,

$$m^{(R)} \sim Binomial(n^{(R)}(n^{(R)} - 1), p^{(R)})$$

To compare the two networks, we are interested in a comparison of $p^{(L)}$ vs. $p^{(R)}$. Formally, we are testing:

$$H_0 : p^{(L)} = p^{(R)}, \quad H_a : p^{(L)} \neq p^{(R)}.$$

Fortunately, the problem of testing for equal proportions under the binomial is well studied. In our case, we use Fisher's exact test [53] to run this test for the null and alternative hypotheses above.

4.2.3 Stochastic block model and group connection testing

A **stochastic block model (SBM)** is a popular statistical model of networks [54]. Put simply, this model treats the probability of an edge occurring between node i and node j as purely a function of the communities or groups that node i and j belong to. This model is parameterized by:

- An assignment of each node in the network to a group. Note that this assignment can be considered to be deterministic or random, depending on the specific framing of the model one wants to use. Here we are assuming τ is a fixed vector of assignments. We represent this non-random assignment of neuron to group by an n-length vector τ. If there are K groups, τ has elements in $\{1...K\}$. If the i-th element of τ is equal to k, then that means that neuron i is assigned to group k.
- A set of group-to-group connection probabilities. We represent these probabilities by the matrix $B \in [0, 1]^{K \times K}$, where the element (k, l) of this matrix represents the probability of an edge from a neuron in group k to one in group l.

Thus, the probability of any specific edge (i, j) can be found by looking up the appropriate element of B:

$$P[A_{ij} = 1] = P_{ij} = B_{\tau_i, \tau_j}$$

In our case, we assume τ is known—in the case where it is not, or one simply wishes to estimate an alternative partition of the network, many methods exist for estimating τ. But with τ known, estimating B becomes simple, amounting to doing K^2 subgraph density estimates. Specifically, let $m(k,l)$ be the number of edges from nodes in group k to nodes in group l. We then compute the density of this subgraph for each (k,l) pair (ignoring self-loops):

$$\hat{B}_{k,l} = \begin{cases} \frac{m(k,l)}{n_k n_l}, & \text{if } k \neq l \\ \frac{m(k,l)}{n_k(n_k-1)}, & \text{if } k = l \end{cases}$$

where n_k is the number of nodes in group k, and likewise for n_l.

Assuming the SBM, we are interested in comparing the group-to-group connection probability matrices, B, for the left and right hemispheres. The null hypothesis of bilateral symmetry becomes

$$H_0 : B^{(L)} = B^{(R)}, \quad H_A : B^{(L)} \neq B^{(R)} \tag{4.1}$$

Rather than having to compare one proportion as in Section 4.2.2, we are now interested in comparing all K^2 probabilities between the SBM models for the left and right hemispheres. The hypothesis test above can be decomposed into K^2 hypotheses. $B^{(L)}$ and $B^{(R)}$ are both $K \times K$ matrices, where each element B_{kl} represents the probability of a connection from a neuron in group k to one in group l. We also know that group k for the left network corresponds with group k for the right. In other words, the *groups* are matched. Thus, we are interested in testing, for k, l both running from $1...K$:

$$H_0 : B^{(L)}_{kl} = B^{(R)}_{kl}, \quad H_A : B^{(L)}_{kl} \neq B^{(R)}_{kl} \tag{4.2}$$

Now, we are left with K^2 p-values from Equation 4.2, each of which bears upon the overall null hypothesis in Equation 4.1. We therefore require some method of combining these p-values into one, or otherwise making a decision about the hypothesis in Equation 4.1. Many methods for combining p-values have been proposed. This problem of combining p-values can itself be viewed as a hypothesis testing problem. Denoting the (k,l)th p-value from Equation 4.2 as p_{kl}, we are testing

$$H_0 : p_{kl} \sim Uniform(0,1)$$

versus the alternative hypothesis that at least one of the p-values is distributed according to some non-uniform, non-increasing density with support $[0,1]$ [55, 56]. Birnbaum [55] showed that no method of combining these p-values can be optimal in general to all alternatives, so we are left with a decision to make (with no universally preferred answer) about which methods to use to combine p-values [56]. Here, we select Tippett's method [56, 57] due to its ubiquity, simplicity, and power against various alternatives to bilateral symmetry under a simulation described in Section 7.1. We note that for future works, specific classes of alternatives may motivate different methods for combining p-values, as described in Heard and Rubin-Delanchy [56].

We also examined the p-values from each of the individual tests after Holm-Bonferroni correction to correct for multiple comparisons. As in Section 4.2.2, we used Fisher's exact test [53] to perform each of the individual hypothesis tests in Equation 4.2. Note also that in some cases, an element of $B^{(L)}$ and/or $B^{(R)}$ could be 0; in each of these cases, we did not run that specific comparison between elements, as the notion of testing for proportions being the same becomes nonsensical. We indicated these tests in Figure 3C, Figure 4C and Figure 5C-D, and note that these tests were not included when computing the number of comparisons for the Holm-Bonferroni correction.

4.2.4 Density-adjusted group connection testing

In Section 2.4, we considered the null hypothesis that the left hemisphere connection probabilities under the SBM are a scaled version of those on the right:

$$H_0 : B^{(L)} = cB^{(R)} \text{ vs. } H_A : B^{(L)} \neq cB^{(R)}. \tag{4.3}$$

The scale for this comparison is the ratio of the densities between the left and the right hemisphere networks:

$$c = \frac{p^{(L)}}{p^{(R)}}. \tag{4.4}$$

Analogous to the group connection testing in Equation 4.2, this means that the individual group connection hypotheses become

$$H_0 : B_{kl}^{(L)} = cB_{kl}^{(R)}, \quad H_A : B_{kl}^{(L)} \neq cB_{kl}^{(R)}. \tag{4.5}$$

The parameter c can be viewed as a probability ratio:

$$B_{kl}^{(L)} \stackrel{?}{=} cB_{kl}^{(R)}$$

$$\frac{B_{kl}^{(L)}}{B_{kl}^{(R)}} \stackrel{?}{=} c$$

In essence, we wish to test whether a probability ratio matches a prespecified hypothesized value, c. To test Equation 4.2, we were able to use Fisher's exact test [53], which aims to determine whether two proportions are significantly different in a 2×2 contingency table. By conditioning on the marginal observations of the table, it was shown that the distribution of one of the cells follows a hypergeometric distribution under the null that the two proportions are the same. Fisher's exact test compares the observed count to this hypergeometric distribution under the null to compute its p-value. Similarly, it was shown that the distribution of one of the cells conditioned on the marginals *and* a specific odds ratio follows Fisher's non-central hypergeometric distribution under the null hypothesis [58, 59]. This leads to a test analogous to Fisher's exact test, but for a potentially non-unity probability ratio, c. Note that this test reduces exactly to Fisher's exact test when the probability ratio $c = 1$. We used this non-unity Fisher's exact test in the individual group connection tests, with all other machinery (e.g., for combining p-values or correcting for multiple comparisons) remaining the same as in Section 4.2.3.

4.3 Edge weight thresholds To examine the effect of which edges are used to define the left and right networks on the p-values from each test, we tested various edge weight thresholds used to define our binary networks for comparison. Given a set of edges (i.e., (i, j) pairs) with corresponding weights w_{ij}, a thresholding $\mathcal{E}(t)$ simply selects the subset of those edges for which w_{ij} is greater than or equal to some threshold, t.

$$\mathcal{E}(t) = \{(i, j) : w_{ij} >= t\}$$

Let s_{ij} be the observed number of synapses from neuron i to neuron j. We considered two thresholding schemes: the first was to simply use the number of synapses from neuron i to j as the edge weight and the second was to consider the edge weight from neuron i to j to be the number of synapses from i to j divided by the total number of observed synapses onto neuron j. We stress that the number of synapses onto neuron j is not necessarily equal to the weighted degree of neuron j. This is simply because we consider all annotated postsynaptic contacts onto neuron j, and some number of those contacts may not be connected to another neuron in the current networks considered here. We denote the number of synapses onto neuron j as D_j. To summarize:

- **Synapse number threshold:** $w_{ij} = s_{ij}$
- **Input proportion threshold:** $w_{ij} = \frac{s_{ij}}{D_j}$

Given either definition of the weighting scheme, we formed a series of networks by varying the edge weight threshold, t. We stress that edge weights were used *only* for the purposes of defining the edges to consider for our (binary) networks—the edge weights themselves were not used in the

statistical tests. We then re-ran the density, group connection, and density-adjusted group connection tests for each network. The p-values for these tests are plotted against the weight thresholds and the proportion of edges removed in Figure 6E and F for the synapse number and input proportion thresholds, respectively.

5 Code

Analyses relied on graspologic [60], NumPy [61], SciPy [62], Pandas [63], and NetworkX [64]. Plotting was performed using matplotlib [65] and Seaborn [66]. All code used for this paper can be found at github.com/neurodata/bilateral-connectome and viewed as a JupyterBook [67] at docs.neurodata.io/bilateral-connectome. The data used in these analyses will be made public and included in the source code repository above upon the release of the data in Winding et al. [14].

6 Acknowledgements

B.D.P. was supported by the NSF Graduate Research Fellowship (Grant no. DGE1746891). J.T.V. was supported by the NSF CAREER Award (Grant no. 1942963). J.T.V. was supported by the NSF NeuroNex Award (Grant no. 2014862). J.T.V and C.E.P. were supported by the NIH BRAIN Initiative (Grant no. 1RF1MH123233-01).

References

[1] Joshua T Vogelstein, Eric W Bridgeford, Benjamin D Pedigo, Jaewon Chung, Keith Levin, Brett Mensh, and Carey E Priebe. Connectal coding: discovering the structures linking cognitive phenotypes to individual histories. Current Opinion in Neurobiology, 55:199–212, April 2019. ISSN 0959-4388. doi: 10.1016/j.conb.2019.04.005. URL http://www.sciencedirect.com/science/article/pii/S0959438818301430.

[2] Larry F. Abbott, Davi D. Bock, Edward M. Callaway, Winfried Denk, Catherine Dulac, Adrienne L. Fairhall, Ila Fiete, Kristen M. Harris, Moritz Helmstaedter, Viren Jain, Narayanan Kasthuri, Yann LeCun, Jeff W. Lichtman, Peter B. Littlewood, Liqun Luo, John H.R. Maunsell, R. Clay Reid, Bruce R. Rosen, Gerald M. Rubin, Terrence J. Sejnowski, H. Sebastian Seung, Karel Svoboda, David W. Tank, Doris Tsao, and David C. Van Essen. The Mind of a Mouse. Cell, 182(6):1372–1376, 2020. ISSN 10974172. doi: 10.1016/j.cell.2020.08.010. URL https://doi.org/10.1016/j.cell.2020.08.010. Publisher: Elsevier Inc.

[3] Elizabeth Barsotti, Ana Correia, and Albert Cardona. Neural architectures in the light of comparative connectomics. Current Opinion in Neurobiology, 71:139–149, 2021. ISSN 18736882. doi: 10.1016/j.conb.2021.10.006. URL https://doi.org/10.1016/j.conb.2021.10.006. Publisher: MRC Laboratory of Molecular Biology.

[4] Dana Shani Galili, Gregory S. X. E. Jefferis, and Marta Costa. Connectomics and the neural basis of behaviour. Current Opinion in Insect Science, page 100968, September 2022. ISSN 2214-5745. doi: 10.1016/j.cois.2022.100968. URL https://www.sciencedirect.com/science/article/pii/S2214574522001031.

[5] Stephan Gerhard, Ingrid Andrade, Richard D Fetter, Albert Cardona, and Casey M Schneider-Mizell. Conserved neural circuit structure across Drosophila larval. eLife, 6:e29089:1–17, 2017. ISSN 2050-084X. ISBN: 2050-084X (Electronic) 2050-084X (Linking).

[6] Daniel Witvliet, Ben Mulcahy, James K. Mitchell, Yaron Meirovitch, Daniel R. Berger, Yuelong Wu, Yufang Liu, Wan Xian Koh, Rajeev Parvathala, Douglas Holmyard, Richard L. Schalek, Nir Shavit, Andrew D. Chisholm, Jeff W. Lichtman, Aravinthan D. T. Samuel, and Mei Zhen. Connectomes across development reveal principles of brain maturation. Nature, 596(7871):257–261, August 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03778-8. URL https://www.nature.com/articles/s41586-021-03778-8. Number: 7871 Publisher: Nature Publishing Group.

[7] Steven J. Cook, Travis A. Jarrell, Christopher A. Brittin, Yi Wang, Adam E. Bloniarz, Maksim A. Yakovlev, Ken C. Q. Nguyen, Leo T.-H. Tang, Emily A. Bayer, Janet S. Duerr, Hannes E. Bülow, Oliver Hobert, David H. Hall, and Scott W. Emmons. Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature, 571(7763):63–71, July 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1352-7. URL https://www.nature.com/articles/s41586-019-1352-7. Number: 7763 Publisher: Nature Publishing Group.

[8] Javier Valdes-Aleman, Richard D. Fetter, Emily C. Sales, Emily L. Heckman, Lalanti Venkatasubramanian, Chris Q. Doe, Matthias Landgraf, Albert Cardona, and Marta Zlatic. Comparative Connectomics Reveals How Partner Identity, Location, and Activity Specify Synaptic Connectivity in Drosophila. Neuron, 109(1):105–122.e7, 2021. ISSN 10974199. doi: 10.1016/j.neuron.2020.10.004. URL https://doi.org/10.1016/j.neuron.2020.10.004. Publisher: Elsevier Inc.

[9] Mikhail A. Fedonkin and Benjamin M. Waggoner. The late Precambrian fossil Kimberella is a mollusc-like bilaterian organism. Nature, 388(6645):868–871, 1997. ISSN 00280836. doi: 10.1038/42242.

[10] Oliver Hobert. Development of left/right asymmetry in the Caenorhabditis elegans nervous system: From zygote to postmitotic neuron. Genesis, 52(6):528–543, 2014. ISSN 1526968X. doi: 10.1002/dvg.22747. ISBN: 1526-954X.

[11] Ju Lu, Juan Carlos Tapia, Olivia L. White, and Jeff W. Lichtman. The interscutularis muscle connectome. PLoS Biology, 7(2):0265–0277, 2009. ISSN 15457885. doi: 10.1371/journal.pbio.1000032.

[12] Philipp Schlegel, Alexander Shakeel Bates, Tomke Stürner, Sridhar R Jagannathan, Nikolas Drummond, Joseph Hsu, Laia Serratosa Capdevila, Alexandre Javier, Elizabeth C Marin, Asa Barth-Maron, Imaan FM Tamimi, Feng Li, Gerald M Rubin, Stephen M Plaza, Marta Costa, and Gregory S X E Jefferis. Information flow, cell types and stereotypy in a full olfactory connectome. eLife, 10:e66018, May 2021. ISSN 2050-084X. doi: 10.7554/eLife.66018. URL https://doi.org/10.7554/eLife.66018. Publisher: eLife Sciences Publications, Ltd.

[13] Marta Costa, James D. Manton, Aaron D. Ostrovsky, Steffen Prohaska, and Gregory S.X.E. Jefferis. NBLAST: Rapid, Sensitive Comparison of Neuronal Structure and Construction of Neuron Family Databases. Neuron, 91(2):293–311, 2016. ISSN 10974199. doi: 10.1016/j.neuron.2016.06.012. URL http://dx.doi.org/10.1016/j.neuron.2016.06.012. Publisher: Elsevier Inc.

[14] Michael Winding, Benjamin D. Pedigo, Christopher L. Barnes, Heather G. Patsolic, Youngser Park, Tom Kazimiers, Akira Fushiki, Ingrid V. Andrade, Feng Li, Javier Valdes-Aleman, Avinash Khandelwal, Nadine Randel, Elizabeth Barsotti, Ana Correia, Richard D. Fetter, Volker Hartenstein, Carey E. Priebe, Joshua T. Vogelstein, Albert Cardona, and Marta Zlatic. The connectome of an insect brain, November 2022. URL https://www.biorxiv.org/content/10.1101/2022.11.28.516756v1. Pages: 2022.11.28.516756 Section: New Results.

[15] František Váša and Bratislav Mišić. Null models in network neuroscience. Nature Reviews Neuroscience, May 2022. ISSN 1471-003X, 1471-0048. doi: 10.1038/s41583-022-00601-9. URL https://www.nature.com/articles/s41583-022-00601-9.

[16] E. N. Gilbert. Random Graphs. The Annals of Mathematical Statistics, 30(4):1141–1144, 1959. URL http://www.jstor.org/stable/2237458.

[17] P. Erdős and A. Rényi. On the Evolution of Random Graphs. In Publication of the Mathematical Institute of the Hungarian Academy of Sciences, pages 17–61, 1960.

[18] Laura E. Suárez, Yossi Yovel, Martijn P. van den Heuvel, Olaf Sporns, Yaniv Assaf, Guillaume Lajoie, and Bratislav Misic. A connectomics-based taxonomy of mammals, March 2022. URL https://www.biorxiv.org/content/10.1101/2022.03.11.483995v1. Pages: 2022.03.11.483995 Section: New Results.

[19] Hang Chen, Utkarsh Soni, Yafeng Lu, Vahan Huroyan, Ross Maciejewski, and Stephen Kobourov. Same Stats, Different Graphs: Exploring the Space of Graphs in Terms of Graph Properties. IEEE Transactions on Visualization and Computer Graphics, 27(3):2056–2072, March 2021. ISSN 1077-2626, 1941-0506, 2160-9306. doi: 10.1109/TVCG.2019.2946558. URL https://ieeexplore.ieee.org/document/8863985/.

[20] Martin Heisenberg. Mushroom body memoir: From maps to models. Nature Reviews Neuroscience, 4(4):266–275, 2003. ISSN 14710048. doi: 10.1038/nrn1074.

[21] Yoshinori Aso, Divya Sitaraman, Toshiharu Ichinose, Karla R. Kaun, Katrin Vogt, Ghislain Belliart-Guérin, Pierre Yves Plaçais, Alice A. Robie, Nobuhiro Yamagata, Christopher Schnaitmann, William J. Rowell, Rebecca M. Johnston, Teri T.B. Ngo, Nan Chen, Wyatt Korff, Michael N. Nitabach, Ulrike Heberlein, Thomas Preat, Kristin M. Branson, Hiromu Tanimoto, and Gerald M. Rubin. Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. eLife, 3(3):e04580, 2014. ISSN 2050084X. doi: 10.7554/eLife.04580.

[22] Katharina Eichler, Feng Li, Ashok Litwin-Kumar, Youngser Park, Ingrid Andrade, Casey M. Schneider-Mizell, Timo Saumweber, Annina Huser, Claire Eschbach, Bertram Gerber, Richard D. Fetter, James W. Truman, Carey E. Priebe, L. F. Abbott, Andreas S. Thum, Marta Zlatic, and Albert Cardona. The complete connectome of a learning and memory centre in an insect brain. Nature, 548(7666):175–182, 2017. ISSN 14764687. doi: 10.1038/nature23455. URL http://dx.doi.org/10.1038/nature23455. Publisher: Nature Publishing Group.

[23] Sophie J.C. Caron, Vanessa Ruta, L. F. Abbott, and Richard Axel. Random convergence of olfactory inputs in the Drosophila mushroom body. Nature, 497(7447):113–117, 2013. ISSN 00280836. doi: 10.1038/nature12063. Publisher: Nature Publishing Group.

[24] Tomoko Ohyama, Casey M. Schneider-Mizell, Richard D. Fetter, Javier Valdes Aleman, Romain Franconville, Marta Rivera-Alba, Brett D. Mensh, Kristin M. Branson, Julie H. Simpson, James W. Truman, Albert Cardona, and Marta Zlatic. A multilevel multimodal circuit enhances action selection in Drosophila. Nature, 520(7549):633–639, 2015. ISSN 14764687. doi: 10.1038/nature14297.

[25] William F. Tobin, Rachel I. Wilson, and Wei Chung Allen Lee. Wiring variations that enable and constrain neural computation in a sensory microcircuit. eLife, 6:1–26, 2017. ISSN 2050084X. doi: 10.7554/eLife.24838.

[26] Matthew E Berck, Avinash Khandelwal, Lindsey Claus, Luis Hernandez-Nunez, Guangwei Si, Christopher J Tabone, Feng Li, James W Truman, Rick D Fetter, Matthieu Louis, Aravinthan DT Samuel, and Albert Cardona. The wiring diagram of a glomerular olfactory system. eLife, 5:1–21, 2016. doi: 10.7554/elife.14859.

[27] Dániel L. Barabási and Albert-László Barabási. A Genetic Model of the Connectome. Neuron, 105(3):435–445.e5, February 2020. ISSN 0896-6273. doi: 10.1016/j.neuron.2019.10.031. URL https://www.sciencedirect.com/science/article/pii/S0896627319309262.

[28] Minh Tang, Avanti Athreya, Daniel L. Sussman, Vince Lyzinski, Youngser Park, and Carey E. Priebe. A Semiparametric Two-Sample Hypothesis Testing Problem for Random Graphs. Journal of Computational and Graphical Statistics, 26(2):344–354, 2017. ISSN 15372715. doi: 10.1080/10618600.2016.1193505. URL https://doi.org/10.1080/10618600.2016.1193505. arXiv: 1409.2344 Publisher: Taylor & Francis.

[29] Avanti Athreya, Donniell E. Fishkind, Minh Tang, Carey E. Priebe, Youngser Park, Joshua T. Vogelstein, Keith Levin, Vince Lyzinski, Yichen Qin, and Daniel L. Sussman. Statistical Inference on Random Dot Product Graphs: a Survey. Journal of Machine Learning Research, 18(226):1–92, 2018. ISSN 1533-7928. URL http://jmlr.org/papers/v18/17-448.html.

[30] Jaewon Chung, Bijan Varjavand, Jesús Arroyo-Relión, Anton Alyakin, Joshua Agterberg, Minh Tang, Carey E. Priebe, and Joshua T. Vogelstein. Valid two-sample graph testing via optimal transport Procrustes and multiscale graph correlation with applications in connectomics. Stat, 11 (1):e429, 2022. ISSN 2049-1573. doi: 10.1002/sta4.429. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sta4.429. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sta4.429.

[31] Gecia Bravo-Hermsdorff, Lee M. Gunderson, Pierre-André Maugis, and Carey E. Priebe. A principled (and practical) test for network comparison, July 2021. URL http://arxiv.org/abs/2107.11403. arXiv:2107.11403 [cs, math, stat].

[32] Clement Lee and Darren J. Wilkinson. A review of stochastic block models and extensions for graph clustering. Applied Network Science, 4(1), 2019. ISSN 23648228. doi: 10.1007/s41109-019-0232-2. arXiv: 1903.00114 Publisher: Applied Network Science.

[33] Tiago P. Peixoto. Efficient Monte Carlo and greedy heuristic for the inference of stochastic block models. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 89(1):1–8, 2014. ISSN 15393755. doi: 10.1103/PhysRevE.89.012804. arXiv: 1310.4378.

[34] Tiago P. Peixoto. Nonparametric Bayesian inference of the microcanonical stochastic block model. Physical Review E, 95(1), 2017. ISSN 24700053. doi: 10.1103/PhysRevE.95.012317. arXiv: 1610.02703.

[35] Karl Rohe, Sourav Chatterjee, and Bin Yu. Spectral clustering and the high-dimensional stochastic blockmodel. Annals of Statistics, 39(4):1878–1915, 2011. ISSN 00905364. doi: 10.1214/11-AOS887.

[36] Daniel L. Sussman, Minh Tang, Donniell E. Fishkind, and Carey E. Priebe. A consistent adjacency spectral embedding for stochastic blockmodel graphs. Journal of the American Statistical Association, 107(499):1119–1128, 2012. ISSN 01621459. doi: 10.1080/01621459.2012.699795. arXiv: 1108.2228.

[37] Thorben Funke and Till Becker. Stochastic block models: A comparison of variants and inference methods. PLOS ONE, 14(4):e0215296, April 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.

0215296. URL https://dx.plos.org/10.1371/journal.pone.0215296.

[38] Arnim Jenett, Gerald M. Rubin, Teri-T. B. Ngo, David Shepherd, Christine Murphy, Heather Dionne, Barret D. Pfeiffer, Amanda Cavallaro, Donald Hall, Jennifer Jeter, Nirmala Iyer, Dona Fetter, Joanna H. Hausenfluck, Hanchuan Peng, Eric T. Trautman, Robert R. Svirskas, Eugene W. Myers, Zbigniew R. Iwinski, Yoshinori Aso, Gina M. DePasquale, Adrianne Enos, Phuson Hulamm, Shing Chun Benny Lam, Hsing-Hsi Li, Todd R. Laverty, Fuhui Long, Lei Qu, Sean D. Murphy, Konrad Rokicki, Todd Safford, Kshiti Shaw, Julie H. Simpson, Allison Sowell, Susana Tae, Yang Yu, and Christopher T. Zugates. A GAL4-Driver Line Resource for Drosophila Neurobiology. Cell Reports, 2(4):991–1001, October 2012. ISSN 2211-1247. doi: 10.1016/j.celrep.2012.09.011. URL https://www.cell.com/cell-reports/abstract/S2211-1247(12)00292-6. Publisher: Elsevier.

[39] Claire Eschbach, Akira Fushiki, Michael Winding, Casey M. Schneider-Mizell, Mei Shao, Rebecca Arruda, Katharina Eichler, Javier Valdes-Aleman, Tomoko Ohyama, Andreas S. Thum, Bertram Gerber, Richard D. Fetter, James W. Truman, Ashok Litwin-Kumar, Albert Cardona, and Marta Zlatic. Recurrent architecture for adaptive regulation of learning in the insect brain. Nature Neuroscience, 23(4):544–555, April 2020. ISSN 1546-1726. doi: 10.1038/s41593-020-0607-9. URL https://www.nature.com/articles/s41593-020-0607-9. Number: 4 Publisher: Nature Publishing Group.

[40] Benjamin D. Pedigo, Michael Winding, Carey E. Priebe, and Joshua T. Vogelstein. Bisected graph matching improves automated pairing of bilaterally homologous neurons from connectomes. Network Neuroscience, pages 1–29, October 2022. ISSN 2472-1751. doi: 10.1162/netn_a_00287. URL https://doi.org/10.1162/netn_a_00287.

[41] Casey M. Schneider-Mizell, Stephan Gerhard, Mark Longair, Tom Kazimiers, Feng Li, Maarten F. Zwart, Andrew Champion, Frank M. Midgley, Richard D. Fetter, Stephan Saalfeld, and Albert Cardona. Quantitative neuroanatomy for connectomics in Drosophila. eLife, 5(MARCH2016):1–36, 2016. ISSN 2050084X. doi: 10.7554/eLife.12059.

[42] Debarghya Ghoshdastidar and Ulrike Von Luxburg. Practical methods for graph two-sample testing. Advances in Neural Information Processing Systems, 2018-Decem:3015–3024, 2018. ISSN 10495258. arXiv: 1811.12752.

[43] Somnath Bhadra, Kaustav Chakraborty, Srijan Sengupta, and Soumendra Lahiri. A Bootstrap-based Inference Framework for Testing Similarity of Paired Networks, November 2019. URL http://arxiv.org/abs/1911.06869. arXiv:1911.06869 [stat].

[44] Donniell E. Fishkind, Sancar Adali, Heather G. Patsolic, Lingyao Meng, Digvijay Singh, Vince Lyzinski, and Carey E. Priebe. Seeded graph matching. Pattern Recognition, 87:203–215, 2019. ISSN 00313203. doi: 10.1016/j.patcog.2018.09.014. URL https://doi.org/10.1016/j.patcog.2018.09.014. arXiv: 1209.0367 Publisher: Elsevier Ltd.

[45] Joshua T. Vogelstein, John M. Conroy, Vince Lyzinski, Louis J. Podrazik, Steven G. Kratzer, Eric T. Harley, Donniell E. Fishkind, R. Jacob Vogelstein, and Carey E. Priebe. Fast Approximate Quadratic programming for graph matching. PLoS ONE, 10(4):1–17, 2015. ISSN 19326203. doi: 10.1371/journal.pone.0121002.

[46] Ali Saad-Eldin, Benjamin D. Pedigo, Carey E. Priebe, and Joshua T. Vogelstein. Graph matching via optimal transport. arXiv: 2111.05366, 2021. URL http://arxiv.org/abs/2111.05366.

[47] Madhura Ingalhalikar, Alex Smith, Drew Parker, Theodore D. Satterthwaite, Mark A. Elliott, Kosha Ruparel, Hakon Hakonarson, Raquel E. Gur, Ruben C. Gur, and Ragini Verma. Sex differences in the structural connectome of the human brain. Proceedings of the National Academy of Sciences, 111(2):823–828, January 2014. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1316909110. URL https://pnas.org/doi/full/10.1073/pnas.1316909110.

[48] Cedric Huchuan Xia, Zongming Ma, Zaixu Cui, Danilo Bzdok, Bertrand Thirion, Danielle S. Bassett, Theodore D. Satterthwaite, Russell T. Shinohara, and Daniela M. Witten. Multi-scale network regression for brain-phenotype associations. Human Brain Mapping, 41(10):2553–2566, 2020.

ISSN 10970193. doi: 10.1002/hbm.24982.

[49] Joshua I. Raji and Christopher J. Potter. The number of neurons in Drosophila and mosquito brains. PLoS ONE, 16(5 May):1–11, 2021. ISSN 19326203. doi: 10.1371/journal.pone.0250381. URL http://dx.doi.org/10.1371/journal.pone.0250381. ISBN: 1111111111.

[50] Alexander S. Bates, Philipp Schlegel, Ruairi J.V. Roberts, Nikolas Drummond, Imaan F.M. Tamimi, Robert Turnbull, Xincheng Zhao, Elizabeth C. Marin, Patricia D. Popovici, Serene Dhawan, Arian Jamasb, Alexandre Javier, Laia Serratosa Capdevila, Feng Li, Gerald M. Rubin, Scott Waddell, Davi D. Bock, Marta Costa, and Gregory S.X.E. Jefferis. Complete Connectomic Reconstruction of Olfactory Projection Neurons in the Fly Brain. Current Biology, 30(16):3183–3199.e6, 2020. ISSN 18790445. doi: 10.1016/j.cub.2020.06.042. URL https://doi.org/10.1016/j.cub.2020.06.042. Publisher: Elsevier Ltd.

[51] Mattia Tantardini, Francesca Ieva, Lucia Tajoli, and Carlo Piccardi. Comparing methods for comparing networks. Scientific Reports, 9(1):1–19, 2019. ISSN 20452322. doi: 10.1038/s41598-019-53708-y.

[52] Jaewon Chung, Eric Bridgeford, Jesús Arroyo, Benjamin D. Pedigo, Ali Saad-Eldin, Vivek Gopalakrishnan, Liang Xiang, Carey E. Priebe, and Joshua T. Vogelstein. Statistical connectomics. Annual Review of Statistics and Its Application, 8:463–492, 2021. ISSN 2326831X. doi: 10.1146/annurev-statistics-042720-023234.

[53] Alan Agresti. Categorical Data Analysis. Wiley, third edition, January 2013. ISBN 978-0-470-46363-5.

[54] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. Social Networks, 5(2):109–137, June 1983. ISSN 0378-8733. doi: 10.1016/0378-8733(83)90021-7. URL http://www.sciencedirect.com/science/article/pii/0378873383900217.

[55] Allan Birnbaum. Combining Independent Tests of Significance. Journal of the American Statistical Association, 49(267):559–574, 1954. ISSN 1537274X. doi: 10.1080/01621459.1954.10483521.

[56] N. A. Heard and P. Rubin-Delanchy. Choosing between methods of combining p -values. Biometrika, 105(1):239–246, 2018. ISSN 14643510. doi: 10.1093/biomet/asx076. arXiv: 1707.06897.

[57] Leonard Henry Caleb Tippett. The Methods Of Statistics. Williams and Norgate LTD, third edition, 1931. URL http://archive.org/details/in.ernet.dli.2015.189563.

[58] Charles William Dunnett and Michael Gent. Significance Testing to Establish Equivalence between Treatments, with Special Reference to Data in the Form of 2 x 2 Tables. Biometrics, 33(4):593, 1977. ISSN 0006341X. doi: 10.2307/2529457.

[59] Ivan S.F. Chan. Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods. Statistical Methods in Medical Research, 12(1):37–58, 2003. ISSN 09622802. doi: 10.1191/0962280203sm314ra.

[60] Jaewon Chung, Benjamin D. Pedigo, Eric W. Bridgeford, Bijan K. Varjavand, Hayden S. Helm, and Joshua T. Vogelstein. GraSPy: Graph Statistics in Python. arXiv:1904.05329 [cs, stat], March 2019. URL http://arxiv.org/abs/1904.05329. arXiv: 1904.05329.

[61] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. Nature, 585(7825):357–362, September 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2649-2. URL https://www.nature.com/articles/s41586-020-2649-2. Number: 7825 Publisher: Nature Publishing Group.

[62] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der

Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, and Paul van Mulbregt. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods, 17(3):261–272, March 2020. ISSN 1548-7105. doi: 10.1038/s41592-019-0686-2. URL https://www.nature.com/articles/s41592-019-0686-2. Number: 3 Publisher: Nature Publishing Group.

[63] Wes McKinney. Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference (Scipy 2010), pages 56–61, 2010. doi: 10.25080/Majora-92bf1922-00a. URL https://conference.scipy.org/proceedings/scipy2010/mckinney.html.

[64] Aric A Hagberg, Daniel A Schult, and Pieter J Swart. Exploring Network Structure, Dynamics, and Function using NetworkX. Proceedings of Seventh Python in Science Conference (SciPy 2008), page 5, 2008.

[65] John D. Hunter. Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, 9 (3):90–95, May 2007. ISSN 1558-366X. doi: 10.1109/MCSE.2007.55. Conference Name: Computing in Science & Engineering.

[66] Michael L. Waskom. seaborn: statistical data visualization. Journal of Open Source Software, 6(60):3021, April 2021. ISSN 2475-9066. doi: 10.21105/joss.03021. URL https://joss.theoj.org/papers/10.21105/joss.03021.

[67] Executable Books Community. Jupyter Book, February 2020. URL https://zenodo.org/record/4539666.

7 Supplemental Information

7.1 Power and validity of group connection test under various alternatives

In Section 4.2.3, we considered the group connection test, where the goal was to test

$$H_0 : B^{(L)} = B^{(R)} \text{ vs. } H_A : B^{(L)} \neq B^{(R)}. \tag{7.1}$$

We saw that this set of hypotheses could be decomposed into K^2 (where K is the number of groups) different hypotheses

$$H_0 : B^{(L)}_{kl} = B^{(R)}_{kl}, \quad H_A : B^{(L)}_{kl} \neq B^{(R)}_{kl}, \tag{7.2}$$

yielding a p-value for the (k, l)th test, p_{kl}. We now consider the problem of trying to combine these p-values into one which bears on the overall hypotheses in Equation 7.1. We proposed using Tippett's method for combining p-values [57], and we now demonstrate the utility of this method against various alternatives.

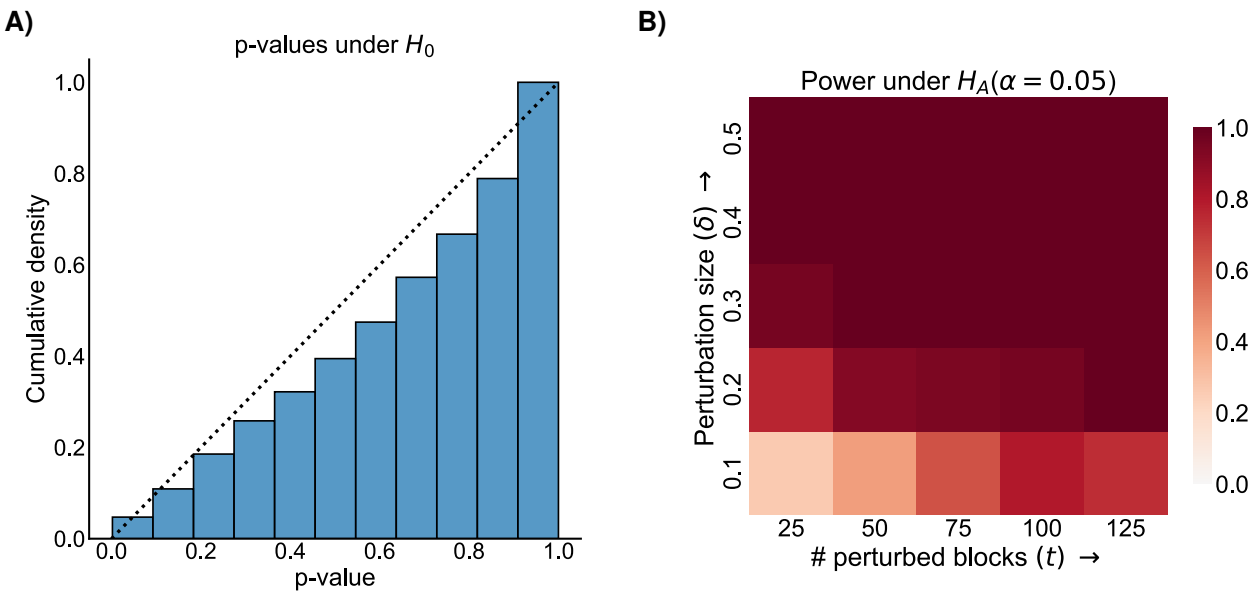
To do so, we performed the following simulation experiment. First, we consider two hypothetical group connection matrices, $B^{(1)}$ and $B^{(2)}$. We set $B^{(1)} = \hat{B}^{(L)}$. We also consider the matrix M, which is a $K \times K$ matrix denoting the number of possible edges in each block of an SBM. Here, we again set $M = \hat{M}^{(L)}$, in other words, we use the number of potential edges for each block observed for the left hemisphere network. To analyze the sensitivity of Tippett's method to different alternatives, we conducted the following simulation: Let t be the *number of probabilities to perturb*. Let δ represent the *strength of the perturbation*. We performed experiments using $\delta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\} \times t \in \{0, 25, 50, 75, 100, 125\}$ (note that if $\delta = 0$ or $t = 0$, then we are under the null hypothesis in Equation 7.1). For each (δ, t), we ran 50 replicates of the simulation below:

1. Randomly select t probabilities without replacement from the elements of B.
2. For each of the selected elements, set $B^{(2)}_{kl} = TN(B^{(1)}_{kl}, \delta B^{(1)}_{kl})$ where TN is a truncated normal distribution with support $[0, 1]$.
3. For each of the unselected elements, set $B^{(2)}_{kl} = B^{(1)}_{kl}$.
4. For each block (k, l), sample the number of edges in that block for network 1:

$$m^{(1)}_{kl} \sim Binomial(M_{kl}, B^{(1)}_{kl}).$$

   Sample the number of edges in each block similarly for network 2, but using $B^{(2)}$.
5. For each block (k, l), compare $m^{(1)}_{kl}$ and $m^{(2)}_{kl}$ using Fisher's exact test as in Section 4.2.3. This yields a set of p-values $\mathcal{P} = \{p_{1,1}, p_{1,2} \ldots p_{(K-1),K}, p_{K,K}\}$ for each comparison.
6. Apply Tippett's method to combine the p-values $\mathcal{P}$ into one p-value for the overall hypotheses.

We observed that the p-values obtained from Tippett's method were valid – they controlled the probability of Type I error for any significance level (Supplemental Figure 1A). Further, we observed that Tippett's method was also powerful against differing alternatives to the null hypothesis (Supplemental Figure 1B). Tippett's method had a power of 1 against the alternative $(t = 25, \delta = 0.5)$, meaning a small number of large perturbations. It also had a power of ~0.8 against the alternative $(t = 125, \delta = 0.1)$, in other words, a large number of small perturbations. Thus, we concluded that Tippett's method is a reasonable choice of method for combining p-values for our group connection test.



Supplemental Figure 1: Demonstration that the group connection test is both valid and powerful against a range of alternatives in a synthetic simulation based on the observed data. See Section 7.1 for more details on the simulation. **A)** Cumulative distribution of p-values from Tippett's method for combining p-values under the null, where the two group connection matrices $B^{(1)}$ and $B^{(2)}$ are the same. Note that the distribution of these p-values is sub-$Uniform(0, 1)$ (i.e., below the dashed line indicating the cumulative distribution of a $Uniform(0, 1)$ random variable, meaning that the test is valid and the Type-I error is properly controlled for any level α. **B)** Power (probability of correctly rejecting the null hypothesis when it is false) as a function of the number of perturbed blocks (t) and the strength of each perturbation (δ). Note that the test is powerful against both a small number of strong perturbations and a large number of small perturbations, indicating its general applicability.