

Write-up on plans for structural and functional connectivity

Contents

1	Methods in literature for structural connectivity	2
2	Plan for analyzing structural connectivity	4
2.1	Data overview (individual subject)	4
2.2	Network construction	4
2.2.1	Global thresholding	5
2.2.2	Local thresholding	5
2.2.3	Sign test	6
3	Plan for analyzing functional connectivity	7
3.1	Marginal and first-order correlation-based network	7
3.2	Glasso-based network	8
3.3	Average over multiple subjects	9

1 Methods in literature for structural connectivity

As raw brain connectivity data are usually noisy, there may be numerous spurious connections. Thresholding is a widely-used method to suppress false positives and fix the density of the network (Fornito et al., 2016).

The most straightforward strategy of thresholding, known as global thresholding, is to apply a pre-specified cutoff value on all elements in the matrix. It can be further classified into two categories: weighted- and density-based thresholding (Fornito et al., 2016). In weighted-based thresholding, an edge between two nodes is established if and only if the corresponding entry in the matrix exceeds the threshold. While this method focuses solely on the weight of connections, it does not regulate the density of the resulting network. In contrast, density-based thresholding aims to select a fixed number of edges corresponding to the largest values in the matrix. This can be helpful if we desire a resulting network with a specific density in mind.

While global thresholding offers simplicity, it may be problematic if weights in a connectivity matrix follow a long-tailed distribution or exhibit multi-scale connection strength patterns across different nodes (Neal, 2022; Yeh et al., 2020). To overcome this issue, the idea of local thresholding can be employed.

Serrano et al. (2009) proposed the *disparity filter* approach, which involves the normalization of weights at every node through dividing each of them by the total weights associated with that particular node. In this case, the node-wise normalized weights represent relative connection strengths at each node. They are subsequently compared to a null model, which assumes those values to arise from a random assignment from a uniform distribution over $[0, 1]$. An edge is preserved if the corresponding normalized weight is tested to be significant at a predetermined level (more details in next section).

Following the normalization process, Foti et al. (2011) proposed the *locally adaptive network sparsification* (LANS) method, a nonparametric approach to select edges locally. The method calculates the empirical CDF for non-zero normalized edge weights at each node. Given a significance level α , an edge is locally significant and selected if the normalized weight surpasses the value yielded by empirical cumulative density at $1 - \alpha$. These methods account for connection strength variations at local level – two nodes with a weak connection measure might be determined as connected, as long as that is large enough compared to their connectivity strengths with other nodes.

Another potential drawback of global thresholding is that, without post-processing, it may yield disconnected network components, which is not scientifically reasonable in a brain connectivity network. To fix this issue, additional edges may be manually added to disconnected node to ensure that every node is connected to at least one other node. Alternatively, a minimum spanning tree

(MST) can be computed as a starting point, and then add more edges to it based on their weights (Alexander-Bloch et al., 2010).

There are even more sophisticated methods to construct brain connectivity networks. For instance, some wish to preserve brain modularity and community structures using Louvain algorithm (Newman & Girvan, 2004), and some based on path lengths between regions (GadElkarim et al., 2013). Given that a complicated process would introduce excessive uncertainties, and considering our aim to utilize the structural connectivity network as a reference for the functional one, I believe those methods are currently beyond the scope of our project.

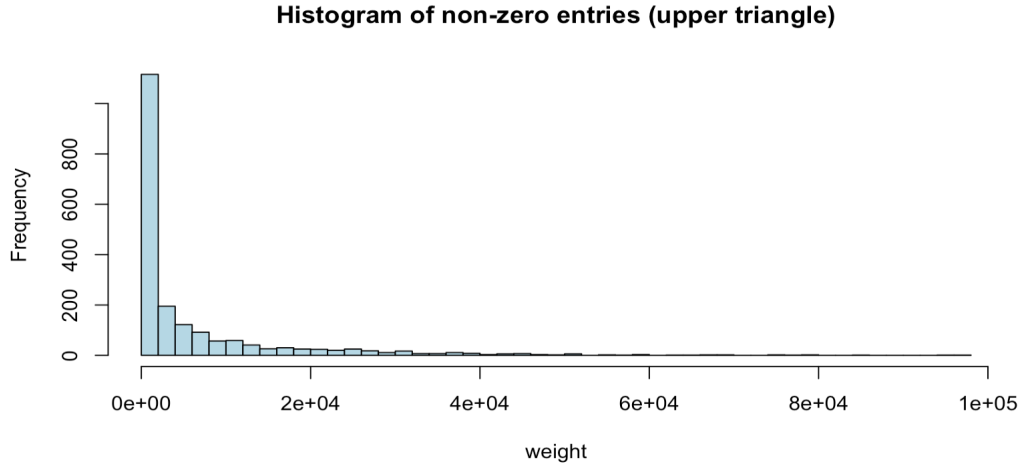
In addition to directly thresholding a weighted adjacency matrix, transformations and statistical tests may be performed on the data to guide the retention of statistically significant edges and the removal of insignificant ones. Gong et al. (2008) employed a non-parametric one-tailed sign test for each entry in the connectivity matrix, under the null hypothesis of no existing connection between the corresponding region pairs. After applying Bonferroni correction for multiple comparisons, they selected significant edges based on p-value of 0.05 to form a symmetric binarized connectivity matrix.

In many studies, two-sample t-tests are used to determine whether significant differences exist in connectivity edges between different groups, such as a group of patients with neurological problems and a control group. This is different from our situation, as we do not distinguish between different subject groups and aim to test whether edges should present using non-negative values in the weighted matrix.

2 Plan for analyzing structural connectivity

2.1 Data overview (individual subject)

In our case, we have a weighted matrix of size 68×68 for one subject, which represents structural information of 68 nodes and $68 \times (68 - 1) / 2 = 2278$ node pairs (brain regions). Notably, 1958 (85.95%) of them correspond to a non-zero entry in the matrix, and the distribution of weights presents a highly right-skewed pattern. Below is a summary of 1958 positive values in the matrix, which shows the presence of extremely large values and the huge discrepancy between the median and mean.



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	80	1163	6171	6765	97495

2.2 Network construction

Due to the prevalence of non-zero entries in the matrix, binarizing the connectivity matrix based solely on the differentiation between zero and non-zero entries would result in an excessively dense network, likely containing too many false positives. Thus, we need to employ some method to sparsify it effectively.

We will discuss three approaches to construct the connectivity network in details, including global thresholding, local thresholding (disparity filter and LANS), and the sign test, as well as their feasibility in our scenario.

2.2.1 Global thresholding

With a purpose to keep m edges in the structural network, we can simply select pairs of regions which correspond to entries of m greatest connection strengths, set them to be 1 and the remaining to be 0.

In the current individual-subject data, suppose we would like to keep $2 \times 68 = 136$ edges in the network. The 136-th greatest value is 25557, and keeping values greater or equal to this value yields a network with exactly 136 edges. But this network contains 13 regions not connecting to any other region. If we would like a network structure with all nodes having a positive degree, we may need to manually add edges to those disconnected nodes.

In the case of multiple subjects, we can sum over all subjects' connection matrices and may do some transformations to prevent extremely large values. Then we select a predetermined number of edges corresponding to highest connection strengths. Alternatively, rather than sum raw values directly, we can first normalize each entry, for every subject, by the maximum value in the matrix, and then average across subjects for edge selection.

2.2.2 Local thresholding

For this idea, we first normalize connection strengths at each node. Let M denote a connection matrix and M_{ij} denote the i, j -th entry in M . For each node i , we perform the normalization for its connection with every other node j as $\frac{M_{ij}}{\sum_j M_{ij}} \in [0, 1]$. Let the normalized matrix be denoted as M^* , where $M_{ij}^* = \frac{M_{ij}}{\sum_j M_{ij}}$ and each row i sums to 1, i.e. $\sum_j M_{ij}^* = 1$. Now each entry (i, j) in M^* represents the relative strength at node i between node pair i, j .

Next, we can apply the disparity filter (Serrano et al., 2009). Suppose any specific node i has degree k , i.e. the i -th row in M^* has k non-zero entries. Denote the non-zero normalized weights as w_{i1}, \dots, w_{ik} . The null hypothesis states that the k normalized weights associated with node i come from random assignment of a uniform distribution on $[0, 1]$. Specifically, assume $Y_1, \dots, Y_{k-1} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}[0, 1]$, and define $W_1 = Y_{(1)}, W_2 = Y_{(2)} - Y_{(1)}, \dots, W_{k-1} = Y_{(k)} - Y_{(k-1)}, W_k = 1 - Y_{(k)}$, where $0 < Y_{(1)} < \dots < Y_{(k-1)} \leq 1$ are the order statistics. Note here W_1, \dots, W_k are invariant under permutation, and $W_j \sim \text{Beta}(1, k-1) \forall j = 1, \dots, k$. Then if this hypothesis is true, w_{i1}, \dots, w_{ik} would arise from the same distribution as W_1, \dots, W_k , which is $\text{Beta}(1, k-1)$. The pdf is $f(w) = (k-1)(1-w)^{k-2}, w \in (0, 1]$. Then the probability of observing a value greater than or equal to a value w is $p = 1 - \int_0^w (k-1)(1-w)^{k-2} dw$. With a predetermined significance level α , we declare an edge if and only if $p < \alpha$. Note the resulting normalized weighted adjacency matrix and connection network are not symmetric, and we can use the OR-rule to include edges – an edge

between i, j is preserved if either i, j -th entry or j, i -th entry is determined as significant. This algorithm is implemented in R package *backbone* (Neal, 2022), and I have managed to implement this process to get the same result.

Alternatively, after obtaining M^* , we may choose to apply the distribution-free method LANS (Foti et al., 2011). Again suppose any node i has degree k and normalized weights w_{i1}, \dots, w_{ik} . Then the empirical CDF is calculated as $\hat{F}(w) = \frac{1}{k} \sum_{j=1}^k \mathbb{1}(w_{ij} \leq w)$, representing the (empirical) probability of observing a weight less than or equal to w . Given a significance level α , we select an edge between node i, j if and only if $1 - \hat{F}(w_{ij}) < \alpha$.

Note that if we aim to select edges, then we may adjust α based on the density of the obtained network. However, if we are interested in making inference, we will need to correct for the multiple comparison issue, e.g. using Bonferroni correction.

Applying these two methods on the single-subject data and controlling for the number of edges in the network to be fewer than 136, all nodes in the two obtained networks have positive degrees. This demonstrates that the local thresholding method can account for heterogeneity of connection strength distributions at different nodes.

For multiple subjects, as discussed in global thresholding case, we can either sum over all subjects' connection matrices, with some transformations to prevent extremely large values, or sum the values normalized by the maximum in one's matrix. Then determine a threshold α and apply local thresholding methods described above.

2.2.3 Sign test

This method only works for multiple subjects. Suppose we have connection matrices for N subjects M_1, \dots, M_N , and let $M_{i,jk}$ denote the j, k -th entry in i -th subject's connection matrix. For every entry (j, k) , we collect data from every subject to get $M_{1,jk}, \dots, M_{N,jk}$ and denote the number of non-zero values among them as y_{jk} . Then use y_{jk} as a test statistic and calculate the p -value $= P(Y \geq y_{jk})$ with $Y \sim \text{Bin}(N, q)$ for some prespecified probability q . If and only if the resultant p -value for j, k -th entry is below a threshold α , we draw an edge between node i and j . Since this test involves the issue of multiple comparison, we can apply Bonferroni correction, which leads the threshold to become α divided by number of node pairs (in our scenario $\frac{\alpha}{2278}$). We may adjust different values of q and α here to obtain a network with a desired density and/or structure.

3 Plan for analyzing functional connectivity

Functional connectivity networks are constructed using time-series fMRI data, which includes 4 sessions and 1200 time points each.

We will use two methods to construct functional connectivity networks – method based on marginal & first-order correlation, and graphical lasso.

We first consider analyzing each session for any single subject, which can provide insight into the consistency of functional connectivity across multiple sessions. Then we employ an aggregated representation of functional connectivity through averaging the correlation matrices from all sessions to construct a connectivity network encapsulating their common patterns.

In consideration of the autocorrelation in time series data, we will apply those methods on both of: (1) all of 1200 time points in any session, (2) a subset of 1200 observations, corresponding to indexes in the form of $6n + 1, n \in \mathbb{N}, 6n + 1 \leq 1200$. Using our data, the subsample contains 200 observations.

3.1 Marginal and first-order correlation-based network

For each of the 4 sessions, we can perform the following procedure:

0. Initialize a complete graph G .
1. Compute the marginal correlation matrix with i, j -th entry $\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$, where Σ is the covariance matrix.
2. Perform Fisher transformation and calculate p-values for every pair of nodes. For the analysis using all 1200 time points in an fMRI session, the test statistic is calculated as $\frac{\text{atanh}(\rho_{ij})}{1/\sqrt{1197}}$, and $\frac{\text{atanh}(\rho_{ij})}{1/\sqrt{197}}$ for the subsample of size 200. The p -values are obtained through comparing the test statistics to $N(0, 1)$. Given a significance level α , remove edges in G corresponding to a $p\text{-value} > \alpha$. We may use $\alpha = 0.05$ here.
3. Compute first-order correlation matrix for pairs of nodes remaining connected in G . Specifically, for every node pair (i, j) , compute the first-order correlation conditional on node $k \neq i, j$ to obtain $\rho_{ij|k} = \frac{\rho_{ij} - \rho_{ik}\rho_{jk}}{\sqrt{(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)}}$. This process gives 66 correlations for every node pair. Then for each (i, j) , find the minimum of 66 correlations, denoted as ρ_{ij}^* .
4. Perform Fisher transformation on ρ_{ij}^* 's and get the p-values. Given a desired number of m edges for the connectivity network, select m smallest p -values to be connected in G and the others to be disconnected. We can use $m = 2 \times 68 = 136$ here. Thereby we obtain a functional connectivity

network with exactly m edges.

To average over 4 sessions:

1. Average the marginal correlation matrices across all sessions and perform Fisher transformation to obtain p -values.
2. As in step 2 for any single session, keep significant edges and remove insignificant ones.
3. Then follow step 3 and 4 above to compute first-order correlations using the averaged marginal correlation matrix and select m edges corresponding to the most significant set of first-order correlations.

3.2 Glasso-based network

The graphical lasso aims to solve the optimization problem

$$\hat{\Theta} = \arg \min_{\Theta \geq 0} \left(tr(S\Theta) - \log \det(\Theta) + \rho \sum_{j \neq k} |\Theta_{jk}| \right),$$

where S is the sample covariance and ρ is the penalizing parameter. As $\hat{\Theta}$ is sparse, we can use it to select edges in the connectivity network. This method is implemented in R package *glasso*.

For each of the 4 sessions, perform the following procedure:

0. Initialize a complete graph G .
1. Fit *glasso* with *approx*=F (not using Meinhausen-Buhlmann (2006) approximation), adjust parameter ρ until exactly m distinct pairs of nodes remain in the matrix (set diagonal elements to 0). In other words, leave exactly m non-zero off-diagonal entries in the upper triangular. Note ρ values for data from different sessions may differ to make the resulting networks have the same density. We can use $m = 136$ here as in the previous method.
2. Let edges corresponding to non-zero entries be connected in G and others disconnected.

For the network averaging over 4 sessions:

0. Initialize a complete graph G .
1. Take the average of covariance matrices from all 4 sessions.
2. Fit *glasso* with *approx*=F (not using Meinhausen-Buhlmann (2006) approximation) using the averaged covariance matrix. Adjust parameter ρ until the resulting sparse matrix contains exactly m distinct pairs of nodes (diagonal elements = 0).
3. Let edges corresponding to non-zero entries be connected in G and others disconnected.

3.3 Average over multiple subjects

To construct a network at the population level, we can compute the marginal and first-order correlation matrices for every individual subject and take the average across all subjects. This averaged correlation matrix provides an estimate of connectivity patterns representative of the entire population. Methods described previously can then be applied to the averaged correlations.

Since the objective is to create one connectivity representation for multiple subjects rather than assessing the behavior of individual subjects, it is unnecessary to repeat the procedure for every single session. Instead, we can aggregate the information from every subject's correlations and construct a network based on them.

References

- Fornito, A., Zalesky, A., & Bullmore, E. T. (2016). Fundamentals of Brain Network Analysis. In Elsevier eBooks. <https://doi.org/10.1016/c2012-0-06036-x>
- Neal, Z.P. (2022). backbone: An R package to extract network backbones. PLoS ONE, 17, e0269137. <https://doi.org/10.1371/journal.pone.0269137>
- Serrano, M. Á., Boguñá, M., & Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. Proceedings of the National Academy of Sciences of the United States of America, 106(16), 6483–6488. <https://doi.org/10.1073/pnas.0808904106>
- Foti, N. J., Hughes, J. M., & Rockmore, D. N. (2011). Nonparametric sparsification of complex multiscale networks. PLOS ONE, 6(2), e16431. <https://doi.org/10.1371/journal.pone.0016431>
- Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., & Bullmore, E. T. (2010). Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. Frontiers in systems neuroscience, 4, 147. <https://doi.org/10.3389/fnsys.2010.00147>
- Newman, M. G., & Girvan, M. (2004). Finding and evaluating community structure in networks. Physical Review E, 69(2). <https://doi.org/10.1103/physreve.69.026113>
- GadElkarim, J. J., Ajilore, O., Schonfeld, D., Zhan, L., Thompson, P. M., Feusner, J. D., Kumar, A., Altshuler, L. L., & Leow, A. D. (2013). Investigating brain community structure abnormalities in bipolar disorder using path length associated community estimation. Human Brain Mapping, 35(5), 2253–2264. <https://doi.org/10.1002/hbm.22324>
- Yeh, C., Jones, D. K., Liang, X., Descoteaux, M., Connelly, A. (2020). Mapping Structural Connectivity Using Diffusion MRI: Challenges and opportunities. Journal of Magnetic Resonance Imaging, 53(6), 1666–1682. <https://doi.org/10.1002/jmri.27188>
- Gaolang Gong, Yong He, Luis Concha, Catherine Lebel, Donald W. Gross, Alan C. Evans, Christian Beaulieu, Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography, Cerebral Cortex, Volume 19, Issue 3, March 2009, Pages 524–536, <https://doi.org/10.1093/cercor/bhn102>