# System-level matching of structural and functional connectomes in the human brain

# Supplementary Information

Yusuf Osmanlıoğlu[1,*], Birkan Tunç[2,4], Drew Parker[1], Mark A. Elliott[3], Graham L. Baum[4], Rastko Ciric[4], Theodore D. Satterthwaite[4], Raquel E. Gur[4], Ruben C. Gur[4], Ragini Verma[1]

[1] Center for Biomedical Image Computing and Analytics, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA.
[2] Center for Autism Research, Children's Hospital of Philadelphia, Philadelphia, PA, 19104, USA.
[3] Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA.
[4] Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA.

# SI.1. Details of structure-function relationship analysis with partial correlation based functional connectomes

### Generating Partial Functional Connectomes

Considering that pairwise relationships obtained in full functional connectome can be affected by the functional activations of other regions, we also considered partial correlation-based functional connectomes in our structure-function analysis. Partial correlation removes the effect of tertiary regions in calculating the functional connections between region pairs. Since the number of available time points in our data was less than the number of ROIs in parcellations, it was necessary to estimate the partial functional connectome matrices via regularization approach, which is known to retain network properties even when the number of ROIs are a magnitude of order larger than the number of time points especially when L1 regularization is used [Schmittmann et al. 2015, Wang et al. 2016]. Consequently, we utilized L1 regularization at multiple resolutions to obtain connectomes using the parcor package of R (Krämer et al. 2009).

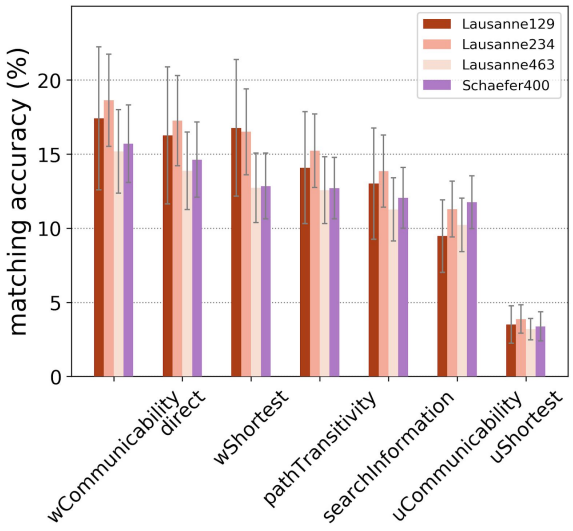


**Figure SI.1. Matching various structure-based connectomes with partial positive functional connectome, we observed that the ordering of matching accuracies at connectome level was highly consistent with that of structure-full function matching (Fig.SI.4.a).**

### Matching structure-based connectomes with partial functional connectome

The traffic patterns that we considered in our analysis describe various degrees of indirect functional connectivity, spanning a spectrum from direct structural connections on one end, which ignores any indirect connectivity, and weighted communicability on the other end, which accounts for indirect interactions that can be captured through all possible pathways. Despite this variability in the traffic patterns, the Pearson's correlation-based full functional connectome captures direct and indirect connectivity between regions jointly. Thus, it is desirable to investigate the similarity between structure-based connectomes and partial correlation-based functional connectomes which capture only direct functional connectivity between regions.

We observed that the matching slightly decreased for weighted communicablity while it increased for direct communication over most of the parcellations. We also noted that the ordering of traffic patterns are mostly preserved when contrasted with that of structure-full function matching (Fig. SI.1). Consistent with results of full functional connectomes, weighted communicability outperformed all other traffic patterns. As in the previous experiment, WC was followed by direct communication and the derivatives of shortest path.

**Traffic patterns are limited in explaining complicated indirect functional interactions**

Studies involving fMRI data most commonly utilize Pearson's correlation in estimating functional connectivity from BOLD timeseries (Leming et al. 2019). Edges of the connectomes obtained using this approach represent the sum of direct and indirect functional interactions between regions, without providing sufficient information to differentiate the two interaction types. The partial correlation-based approach to functional connectivity estimation, on the other hand, removes the effect of tertiary regions in estimating the functional connections between region pairs, providing a measure of direct functional interaction among brain regions (Smith et al. 2011). Due to the disparity between two functional connectivity types, there is likely to be a difference in their agreement with structural connectivity. In matching structure-based connectomes with partial functional connectomes, we expected direct structural connections to achieve a higher matching accuracy relative to other traffic patterns, as both direct structural and partial functional connectomes consider direct connectivity between regions. We further expected the agreement between structure and the full functional connectome to be stronger (relative to the agreement between structure and the partial functional connectome) for weighted communicability, as it incorporate indirect connectivity between regions similar to full functional connectomes.

Although we observed an increased matching accuracy between the direct and partial functional connectivity, its structure-partial function agreement was lower relative to weighted communicability. This might indicate that partial functional connectomes still contain indirect connectivity information, which can be attributed to L1 regularization-based estimation method failing to completely remove the effect of tertiary regions in pairwise connections (Smith et al. 2011). In line with our second expectation, structure-function agreement was stronger in full functional connectomes relative to partial functional connectome for the weighted communicability. It is also noteworthy that, partial functional connectivity obtained via L1 regularization can capture the network information in par with full functional connectivity despite number of time series being much less than the number of ROIs in larger resolutions, supporting previous findings (Schmittmann et al. 2015).

# SI.2. Details of communication patterns used in the study

In this section, we provide a detailed explanation of the communication patterns contrasted in the study.

- Direct connections (DC): This is the standard weighted structural connectome which assumes that the communication occurs only between directly connected regions.
- Unweighted shortest path (USP): Commonly considered as the communication pattern of the brain (Honey et al. 2009), this traffic pattern assumes that the communication between region pairs occurs through minimum number of intermediary regions. After binarizing the weighted direct connectomes by representing nonzero values as ones, these connectomes are calculated by the Floyd-Warshall all pairs shortest path algorithm.
- Weighted shortest path (WSP): Being the weighted version of the USP, this scenario assumes that the communication occurs through the strongest path, i.e., the path with maximal number of connecting fibers. Since the longest path problem is intractable, it is calculated by first taking the reciprocal of the edge weights of the direct connectome, then calculating the shortest path between all node pairs, and finally reporting the reciprocal of calculated values as the weighted shortest path.
- Search information (SI): An extension to WSP, search information (Goñi et al. 2014) quantifies the accessibility of the shortest path between two nodes within the network by measuring the amount of knowledge needed to access the path. For any two nodes S,T with $\Omega_{S\to T} = \{S, a, b, ..., z, T\}$ denoting the sequence of nodes constituting the shortest path between them, and $w_{Si}$ denoting the weighted edges in the shortest path, probability of taking the shortest path from s to t can be quantified as $P(\Omega_{S\to T}) = \frac{w_{Sa}}{w_S} \cdot \frac{w_{ab}}{w_a} \cdot ... \cdot \frac{w_{zT}}{w_z}$ where $w_i$ represents the strength of the node i. Then, information needed to access the shortest path is calculated as $SI(\Omega_{S\to T}) = \frac{-log_2(P(\Omega_{S\to T}))-log_2(P(\Omega_{T\to S}))}{2}$. The main assumption in here is that the intermediate nodes with several neighbors that are located on the shortest path are likely to affect communication negatively, as the signals can be routed incorrectly to one of the several neighbors that are not on the (shortest) path.
- Path transitivity (PT): Being another extension to WSP, this traffic pattern (Goñi et al. 2014) quantifies the density of local detours available on the shortest path between two nodes, assuming that having alternate reroutes at nodes over the shortest path will facilitate communication. Calculation of PT involves a quantifier called matching index, measuring the similarity of incoming and outgoing edges of two nodes as follows:

$$m_{ij} = \frac{\sum_{k\neq i,j} (w_{ik}+w_{jk})\Theta(w_{ik})\Theta(w_{jk})}{\sum_{k\neq j} w_{ik} + \sum_{k\neq i} w_{jk}},$$

where $\Theta$ is an indicator function taking value 1 for edges with nonzero weights and zero, otherwise. Extending this definition to the edges on the shortest path between a source node S and destination node T, path transitivity is then calculated as

$$PT(\Omega_{S\to T}) = \frac{2}{|\Omega_{S\to T}|(|\Omega_{S\to T}|-1)} \sum_{i>j\in\Omega_{S\to T}} m_{ij}.$$

Note that PT and SI are comparable in that PT considers whether the connections surrounding the shortest path can improve the connectivity, while SI quantifies how much dispersion is introduced to the shortest path by its surrounding connections.

- Unweighted communicability (UC): This traffic pattern considers communication as a diffusion process as it assumes that the communication between regions occur through multiple pathways simultaneously (Estrada and Hatano 2008). This unweighted version considers the strength of connection between two nodes to be proportional to the number of possible walks between them. A walk over a graph is defined as a sequence of vertices establishing a consecutive link between two nodes where loops can possibly exist. In contrast to paths, which allow traversing a node or edge at most once, the potential for loops in walks supports the idea that reentry loops can enhance a signal that needs to travel a longer distance (Goñi et al. 2014). Given a binarized connectivity matrix A, UC between nodes *s,t* is calculated by

$$UC(s,t) = \sum_{n=0}^{\infty} \frac{[A^n]_{st}}{n!}$$

where $A^n$ constitutes all walks of length $n$ in the network. Note that due to the factorial in the denominator, higher powers of the exponentiation have a diminishing return, implying that longer walks will contribute less to the sum.

- Weighted communicability (WC): This is the weighted version of UC, where exponents of weighted DC are calculated and the values in the resulting matrix are normalized by the multiplication of the strength of the source and destination nodes of each path. More specifically, given the weighted connectivity matrix $W$ and the diagonal matrix $S$ containing the node strengths (the sum of all edges directly connected to each node) at its diagonal entries, WC between nodes *s,t* is calculated by

$$WC(s,t) = \sum_{n=0}^{\infty} \frac{\left[(S^{-1/2}WS^{-1/2})^n\right]_{st}}{n!}.$$

We note that WC is an approximation, while UC provides an exact number of walks in the network.

# SI.3. Comparison of structure-function matching accuracy for weighted communicability with various number of hops

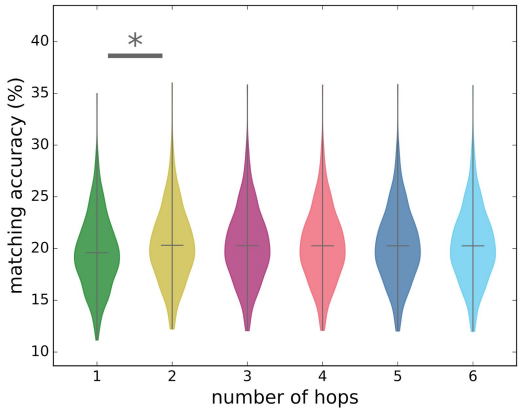


**Fig. SI.2. Comparison of structure-function matching scores for various hop counts in the calculation of the weighted communicability model.**

Equation used in calculation of communicability implies that, longer walks contribute less to the total sum due to the factorial term in the denominator. Since calculating number of longer walks is computationally more expensive despite contributing little to the accuracy of total number of walks, it is desirable to find a feasible cut off value for $n$. Using the structure-function matching of brain as our test bed, we explored the parameter space of walk length $n$ taking integer values in [1,6] interval. For each value of $n$, we calculated weighted communicability of the structural connectome and matched it with positive functional connectome. We then contrasted the connectome level matching accuracy of these six experiments. We observed that increasing the number of hops from 1 to 2 hops, improves the structure-function matching accuracy from 19.6% to 20.3% ($p<10^{-3}$) (Fig. SI.2). Increasing number of hops beyond 2 in the calculation of weighted communicability, did not provide any statistically significant improvement over the structure-function matching score. Thus, we used $n$=2 in our experiments for calculation of communicability-based connectomes, which account for walks of up to 2 hops.

# SI.4. Regions with highest accurate matching scores

| Matching Accuracy | Region Name | Func. System | Matching Accuracy | Region Name | Func. System |
|---|---|---|---|---|---|
| 57.22 | Left_lateraloccipital-4 | Visual | 36.18 | Left_precuneus-5 | Dorsal |
| 44.99 | Right_superiorfrontal-2 | DMN | 35.97 | Right_precuneus-1 | Dorsal |
| 41.41 | Right_precentral-3 | Somato. | 35.66 | Left_cuneus-2 | Visual |
| 41.37 | Right_rostralanteriorcingulate | DMN | 35.52 | Left_superiorfrontal-1 | DMN |
| 40.00 | Left_cuneus-1 | Visual | 35.45 | Right_cuneus | Visual |
| 39.80 | Left_isthmuscingulate | DMN | 35.44 | Right_postcentral-4 | Somato. |
| 39.19 | Left_superiorparietal-1 | Dorsal | 35.31 | Left_fusiform-2 | Visual |
| 39.00 | Right_postcentral-3 | Somato. | 35.20 | Right_superiortemporal-3 | Default |
| 38.60 | Left_lateraloccipital-2 | Visual | 35.10 | Left_pericalcarine-2 | Visual |
| 38.22 | Left_parsopercularis-2 | Fronto. | 34.97 | Left_rostralanteriorcingulate-1 | DMN |
| 38.13 | Left_postcentral-5 | Somato. | 34.97 | Left_lateraloccipital-1 | Visual |
| 38.05 | Left_postcentral-3 | Somato. | 34.57 | Right_precentral-1 | Somato. |
| 36.55 | Right_lateraloccipital-5 | Visual | 33.91 | Left_lateraloccipital-5 | Visual |
| 36.52 | Right_paracentral-1 | Somato. | 33.86 | Right_isthmuscingulate | DMN |
| 36.40 | Left_medialorbitofrontal-1 | DMN | 33.79 | Left_precuneus-3 | DMN |

**Table SI.1. Top 30 regions with highest accurate structure-positive function matching score along with the functional systems that they belong to. Reported results are due to structure-function matching using weighted communicability as the traffic pattern.**

As stated before, we considered a node in structure-based graph getting matched with its counterpart in the function-based graph as an accurate match. Consequently, the diagonal entries of the average structure-function matching matrix indicates the probability of accurate matchings for each node of the brain. In Table SI.1, we list the top 30 nodes that has highest probability of accurately matching its counterpart. Analyzing the table, we observe two patterns in the results:

- Regions belonging to visual, DMN, dorsal, and motor systems has highest matching accuracy at the node level.
- Regions with high matching accuracy are of the same type or sub divisions of the same region in contralateral hemispheres. Specifically, we observe that lateral occipital, isthmus cingulate, medial orbitofrontal, precuneus, superio-parietal, rostral anterior cingulate, superio-frontal, and fusiform have high matching accuracy in left and right hemisphere.

# SI.5. Pearson's correlation as a measure of structure-function similarity

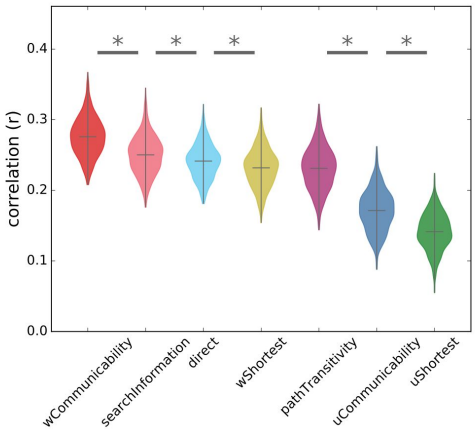


**Figure SI.3. Using pearson's correlation as a similarity metric, we calculated the edgewise correlation between the structural and positive full functional connectomes. Results shown here are calculated over Lausanne 234 parcellation.**

Pearson's correlation has been widely used in measuring similarity between structural and functional connectivity in brain (Hagmann et al., 2008, Honey et al., 2009, Honey et al., 2010, Zimmermann et al., 2016). In order to validate our graph matching-based approach as a similarity measure capturing the structure-function relationship in brain, we calculated edgewise correlation between structural and positive functional connectomes. In our experiment, we utilized structural connectomes derived from all seven traffic patterns and contrasted them with full correlation-based functional connectomes. We observed that the ordering of traffic patterns are highly consistent with our previous results (Fig. SI.3). Weighted communicability achieved the highest correlation, which is followed by direct connections and shortest path-based traffic patterns (i.e., search information, weighted shortest path, and path transitivity). Unweighted versions of the traffic patterns, again, achieved the lowest accuracy scores. Repeating the experiment over other parcellations, we observed a similar ordering of traffic patterns, where weighted communicability consistently achieved top performance in capturing the structure-function similarity.

# SI.6. Stability of structure-function matching accuracy across parcellations

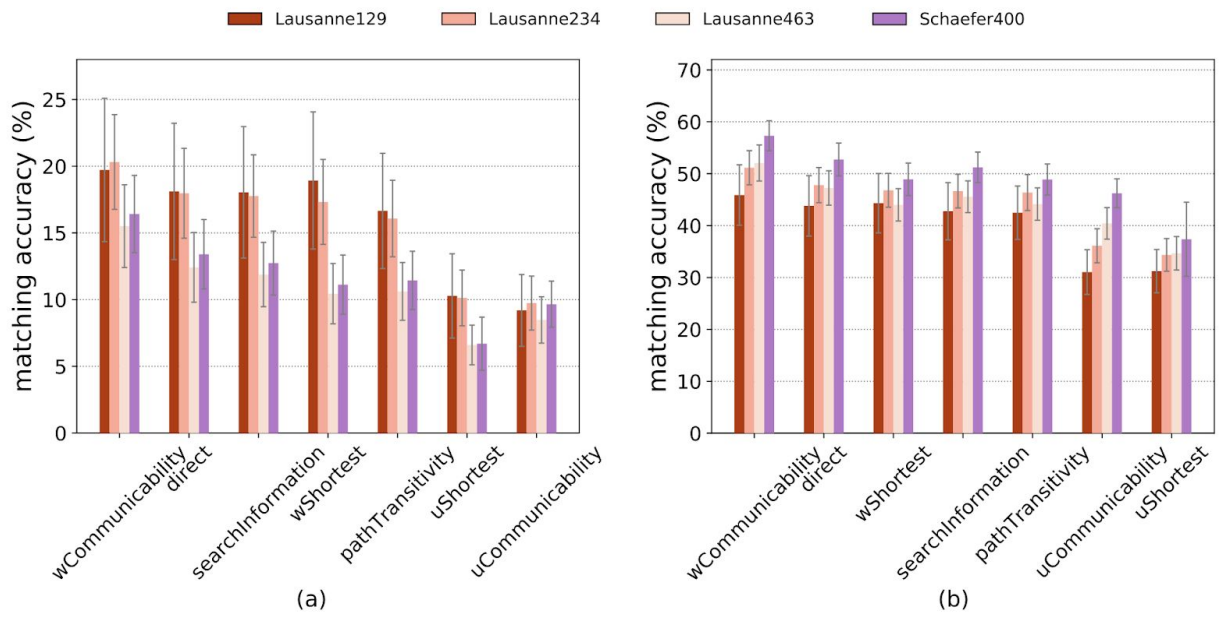


**Figure SI.4. Structure-function matching accuracies for various traffic patterns across parcellations at (a) connectome level and (b) system level.**

In order to investigate the stability of structure-function matching accuracy across different parcellations, we carried out the matching experiment using two atlases at 4 different resolutions. Specifically, we used Lausanne atlas with 129, 234, and 463 ROIs (Hagmann et al., 2008) as well as Schaefer atlas with 400 ROIs (Schaefer et al., 2017) for generating both structural and functional connectomes.

Measuring the similarity at connectome level, we observed a highly consistent ordering among the traffic patterns across the parcellations (Fig. SI.4). In all four cases, we observed that weighted communicability achieves the highest accuracy with a relatively large margin, which is followed by direct connection and weighted shortest path-based traffic patterns with small difference among each other. We also observe that unweighted versions of the traffic patterns consistently achieve lowest accuracy scores (Fig. SI.4.a). We further observed that, matching accuracy is reduced with increasing number of ROIs in the connectomes. This can be attributed to the fact the the matching algorithm needs to select the best match among increased number of nodes in connectomes with higher resolution, which will have possibly similar connectivity signature.

Measuring the similarity at system level, we again observe a highly consistent ordering of the traffic patterns with weighted communicability achieving highest structure-function matching accuracy score (Fig. SI.4.b). We note that, system level matching accuracy improves with increasing number of ROIs in connectomes in contrast to the inverse behavior we observed at the connectome level matching accuracy. Considering the two results together, it implies that the mismatches observed in connectomes with higher resolution occur within the same large scale systems. It is also noteworthy that, Schaefer 400 parcellation achieves the highest matching accuracy at system level for all of the traffic patterns, despite having a lower resolution relative to Lausanne 463 parcellation. This can be attributed to the fact that, Schaefer

atlas is obtained from a much larger sample spanning a wider age range relative to Lausanne atlas, possibly resulting in a more accurate delineation of functional systems that overlaps with age range of samples that we investigate in our experiments.

# References (Supplementary)


Estrada, E., Hatano, N., 2008. Communicability in complex networks. Phys Rev. E 77(3):036111.

Goñi, J., van den Heuvel, M.P., Avena-Koenigsberger, A., de Mendizabal, N.V., Betzel, R.F., Griffa, A., Hagmann, P., Corominas-Murtra, B., Thiran, J.P. and Sporns, O., 2014. Resting-brain functional connectivity predicted by analytic measures of network communication. Proc Natl Acad Sci USA. 111(2), 833-838.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J. and Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS Biology. 6(7), e159.

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R. and Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc Natl Acad Sci USA. 106(6), 2035-2040.

Honey, C.J., Thivierge, J.P., Sporns, O., 2010. Can structure predict function in the human brain? Neuroimage. 52(3):766–776.

Krämer, N., Schäfer, J. and Boulesteix, A.L., 2009. Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC* bioinformatics, *10*(1), p.384.

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.N., Holmes, A.J., Eickhoff, S.B. and Yeo, B.T., 2017. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb Cortex*. 28(9):1-20.

Schmittmann, V.D., Jahfari, S., Borsboom, D., Savi, A.O. and Waldorp, L.J., 2015. Making large-scale networks from fMRI data. *PloS one*, *10*(9), p.e0129074.

Wang, Y., Kang, J., Kemmer, P.B. and Guo, Y., 2016. An efficient and reliable statistical method for estimating functional connectivity in large scale brain networks using partial correlation. Frontiers in neuroscience, 10, p.123.

Zimmermann, J., Ritter, P., Shen, K., Rothmeier, S., Schirner, M., McIntosh, A.R., 2016. Structural architecture supports functional organization in the human aging brain at a regionwise and network level. Hum brain Mapp. 37(7), 2645-2661.