



Combining Multiple Atlases to Estimate Data-Driven Mappings Between Functional Connectomes Using Optimal Transport

Javid Dadashkarimi¹(✉), Amin Karbasi², and Dustin Scheinost³

¹ Department of Computer Science, Yale University, New Haven, USA
javid.dadashkarimi@yale.edu

² Department of Electrical Engineering, Yale University, New Haven, USA
amin.karbasi@yale.edu

³ Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, USA
dustin.scheinost@yale.edu

Abstract. Connectomics is a popular approach for understanding the brain with neuroimaging data. Yet, a connectome generated from one atlas is different in size, topology, and scale compared to a connectome generated from another atlas. These differences hinder interpreting, generalizing, and combining connectomes and downstream results from different atlases. Recently, it was proposed that a mapping between atlases can be estimated such that connectomes from one atlas (*i.e.*, source atlas) can be reconstructed into a connectome from a different atlas (*i.e.*, target atlas) without re-processing the data. This approach used optimal transport to estimate the mapping between one source atlas and one target atlas. Yet, restricting the optimal transport problem to only a single source atlases ignores additional information when multiple source atlases are available, which is likely. Here, we propose a novel optimal transport based solution to combine information from multiple source atlases to better estimate connectomes for the target atlas. Reconstructed connectomes based on multiple source atlases are more similar to their “gold-standard” counterparts and better at predicting IQ than reconstructed connectomes based on a single source mapping. Importantly, these results hold for a wide-range of different atlases. Overall, our approach promises to increase the generalization of connectome-based results across different atlases.

Keywords: Optimal transport · Functional connectome · fMRI

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-16431-6_37.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
L. Wang et al. (Eds.): MICCAI 2022, LNCS 13431, pp. 386–395, 2022.
https://doi.org/10.1007/978-3-031-16431-6_37

1 Introduction

A connectome—a matrix describing the connectivity between any pair of brain regions—is a popular approach used to model the brain as a graph-like structure. They are created by parcellating the brain into distinct regions using an atlas (*i.e.*, the nodes of a graph) and estimating the connections between these regions (*i.e.*, the edges of a graph). As different atlases divide the brain into a different number of regions of varying size and topology, connectomes created from different atlases are not directly comparable. In other words, connectome-based results generated from one atlas cannot be directly compared to connectome-based results generated from a different atlas. This fact hinders not only replication and generalization efforts, but also simply comparing the results from two independent studies that use different atlases. For example, large-scale projects—like the Human Connectome Project (HCP), the Adolescent Brain Cognitive Development (ABCD) study [5], and the UK Biobank [25]—share fully processed connectomes to increase the wider-use of the data, while reducing redundant processing efforts [15]. Yet, several atlases, but no gold standards, exist [2]. As such, released connectomes for each project are based on different atlases, which prevents these datasets being combined without reprocessing data from thousands of participants. Being able to map between these connectomes—without need for raw data—would facilitate existing connectomes to be easily reused in a wide-range of analyses while eliminating wasted and duplicate processing efforts.

To facilitate this mapping, it was shown that an existing connectome could be transformed into a connectome from a different atlas without needing the raw functional magnetic imaging (fMRI) data [8]. This method used optimal transport, or the mathematics of converting a probability distribution from one set to another, to find a spatial mapping between a pair of atlases. This mapping could then be applied the timeseries fMRI data parcellated with the first atlas (source atlas), then creating connectome based on the second atlas (target atlas). While these previous mappings were based on only a single source atlas, most large-scale projects, release data processed data from 2–4 atlases. As such, richer information than that provided by a single source atlas is available and ignored in the current approach. We propose to combine information from multiple source atlases to jointly estimate mappings to the target atlas. Using 6 different atlases, we show that our approach results in significant improvements in the quality of reconstructed connectomes and their performance in downstream analyses.

2 Background

Optimal Transport: The optimal transport problem solves how to transport resources from one location α to another β while minimizing the cost C to do so [12, 14, 18, 26]. Using a probabilistic approach in which the amount of mass located at x_i potentially dispatches to several points in target [17], admissible solutions are defined by a coupling matrix $\mathcal{T} \in \mathbb{R}_+^{n \times m}$ indicating the amount of mass being transferred from location x_i to y_j by $\mathcal{T}_{i,j}$:

$$U(a, b) = \{\mathcal{T} \in \mathbb{R}_+^{n \times m} : \mathcal{T} \mathbb{1}_m = a, \mathcal{T}^T \mathbb{1}_n = b\}, \quad (1)$$

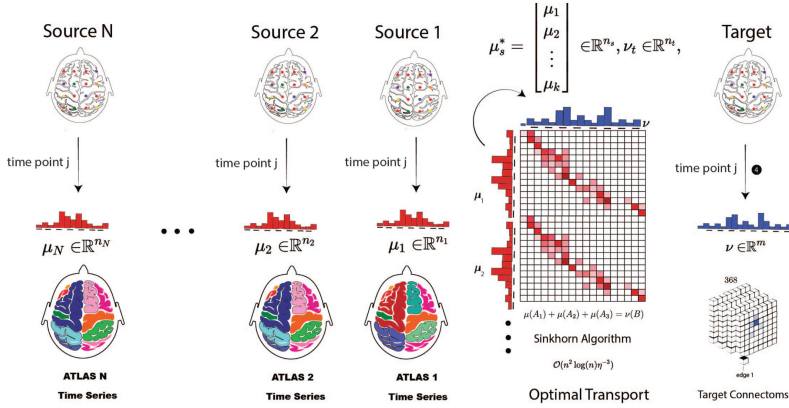


Fig. 1. Our all-way optimal transport algorithm combines timeseries data from multiple source atlases to jointly estimate a single mapping from these atlases to the target atlas. Once this mapping is found, it can be applied to independent data to reconstruct connectomes based on the target data based only timeseries data from the source atlases.

for vectors of all 1 shown with $\mathbb{1}$. An optimum solution is obtained by solving the following problem for a given “ground metric” matrix $C \in \mathbb{R}^{n \times m}$ [21]:

$$L_c(a, b) = \min_{T \in U(a, b)} \langle C, T \rangle = \sum_{i, j} C_{i, j} T_{i, j}. \quad (2)$$

which is a linear problem and is not guarantee to have a unique solution [19], but always there exists an optimal solution (see proof in [3, 4]). Unlike, the KL divergence, optimal transport is one of the few methods that provides a well-defined distance metric when the support of the distributions is different.

Single-Source Optimal Transport: The single-source optimal transport algorithm from Dadashkarimi et al. [8], first, transforming timeseries data from one atlas (labeled the source atlas) into timeseries from an unavailable atlas (labeled the target atlas). Next, the corresponding functional connectomes can be estimated using standard approaches (e.g., full or partial correlation). Formally, it is assumed that we have training timeseries data consisting of T time-points from the same individuals but from two different atlases (atlas $\mathcal{P}_n$ with n regions and atlas $\mathcal{P}_m$ with m regions). Additionally, let $\mu_t \in \mathbb{R}^n$ and $\nu_t \in \mathbb{R}^m$ to be the vectorized brain activity at single timepoint t based on atlases $\mathcal{P}_n$ and $\mathcal{P}_m$, respectively. For a fixed cost matrix $C \in \mathbb{R}^{n \times m}$, which measures the pairwise distance between regions in $\mathcal{P}_m$ and $\mathcal{P}_n$, this approach aims to find a mapping $T \in \mathbb{R}^{n \times m}$ that minimizes transportation cost between μ_t and ν_t :

$$L_c(\mu_t, \nu_t) = \min_T C^T T \text{ s.t., } AT = \begin{bmatrix} \mu_t \\ \nu_t \end{bmatrix}, \quad (3)$$

in which $\underline{\mathcal{T}} \in \mathbb{R}^{nm}$ is vectorized version of $\mathcal{T}$ such that the $i + n(j-1)$'s element of $\mathcal{T}$ is equal to $\mathcal{T}_{ij}$ and A is defined as:

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & n \end{matrix} \\ \begin{matrix} m \\ n \end{matrix} & \left(\begin{array}{cccc} \left(\begin{smallmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{smallmatrix} \right) & \left(\begin{smallmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{smallmatrix} \right) & \dots & \left(\begin{smallmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{smallmatrix} \right) \\ \left(\begin{smallmatrix} 1 & 1 & \dots & 1 \end{smallmatrix} \right) & \left(\begin{smallmatrix} 1 & 1 & \dots & 1 \end{smallmatrix} \right) & \dots & \left(\begin{smallmatrix} 1 & 1 & \dots & 1 \end{smallmatrix} \right) \\ \vdots & & & \\ \left(\begin{smallmatrix} 1 & 1 & \dots & 1 \end{smallmatrix} \right) & \dots & \dots & \left(\begin{smallmatrix} 1 & 1 & \dots & 1 \end{smallmatrix} \right) \end{array} \right). \end{matrix} \quad (4)$$

$\mathcal{T}$ represents the optimal way of transforming the brain activity data from n regions into m regions. Thus, by applying $\mathcal{T}$ to every timepoint from the time-series data of the source atlas, the timeseries data of the target atlas and corresponding connectomes can be estimated. The cost matrix C was based on the similarity of pairs of timeseries from the different atlases:

$$C = 1 - \begin{pmatrix} \rho(U_{1,\cdot}, N_{1,\cdot}) & \dots & \rho(U_{1,\cdot}, N_{n,\cdot}) \\ \vdots & \ddots & \vdots \\ \rho(U_{m,\cdot}, N_{1,\cdot}) & \dots & \rho(U_{m,\cdot}, N_{n,\cdot}) \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (5)$$

where U_x and N_x are timeseries from $\mathcal{S}_m$ and $\mathcal{S}_n$ and $\rho(U_x, N_y)$ is correlation between them.

3 All-Way Optimal Transport

A key drawback of the single-source optimal transport is that it relies on a single pair of source and target atlases (i.e., one source atlas and one target atlas), which ignores additional information when multiple source atlases exist. To overcome this weakness, we designed a new approach, called all-way optimal transport, that uses a varying number of source atlases to better reconstruct the target atlas. All-way optimal transport combines information from multiple source atlases by using a larger cost matrix generated from stacking the set of region centers in each source atlas (see Fig. 1). In general, assume we have paired time-series, from the same person, but from k different source atlases with a total of n_s regions (where $n_s = n_1 + n_2 + \dots + n_k$ from source atlas $\mathcal{S}_{n_1}$ with n_1 regions, $\mathcal{S}_{n_2}$ with n_2 regions, ..., $\mathcal{S}_{n_k}$ with n_k regions) and a target atlas $\mathcal{S}_m$ with m regions, let's define $\mu_t \in \mathbb{R}^{n_s}$ and $\nu_t \in \mathbb{R}^m$ to be the distribution of brain activity at single time point t based on atlases $\mathcal{S}_s$ and $\mathcal{S}_m$:

$$\mu_s^* = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix} \in \mathbb{R}^{n_s}, \nu_t \in \mathbb{R}^{n_t}, C^* = \begin{pmatrix} C_{1,1} & \dots & C_{1,m} \\ \vdots & \ddots & \vdots \\ C_{n_s,1} & \dots & C_{n_s,m} \end{pmatrix} \in \mathbb{R}^{n_s \times m}, \quad (6)$$

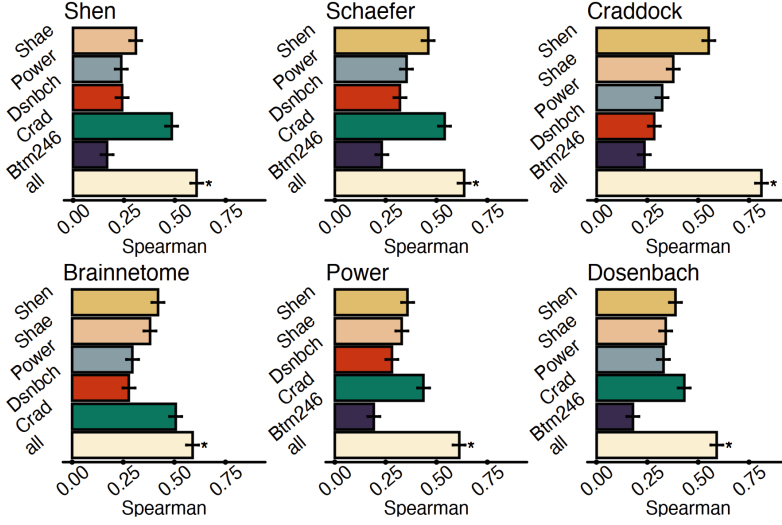


Fig. 2. Using multiple source atlases improves the similarity of reconstructed connectomes. The correlation between the reconstructed connectomes and connectomes generated directly with the target atlases are shown for each pair of source and target atlas as well reconstructed connectomes using all source atlases. For each target atlas, using all source atlases produces higher quality reconstructed connectomes. Error bars are generated from 100 iterations of randomly splitting the data into training and testing data. In all cases, all-way optimal transport resulted in significantly more similar connectomes (indicated by *).

and $C_{i,j}$ is based the similarity of pairs of timeseries from nodes i and j from different atlases. Next, we want to minimize distance between μ_s^* and ν_t as:

$$L_c(\mu_t^*, \nu_t^*) = \min_{\mathcal{T}} C^T \mathcal{T} \text{ s.t., } A\mathcal{T} = \begin{bmatrix} \mu_t^* \\ \nu_t^* \end{bmatrix}, \quad (7)$$

4 Implementation

Solving the large linear program in Eq. 7 is computationally hard [9]. As such for both all-way and single source optimal transport, we used the entropy regularization, which gives an approximation solution with complexity of $\mathcal{O}(n^2 \log(n) \eta^{-3})$ for $\epsilon = \frac{4 \log(n)}{\eta}$ [19], and instead solve the following:

$$L_c(\mu_t^*, \nu_t^*) = \min_{\mathcal{T}} C^T \mathcal{T} - \epsilon H(\mathcal{T}) \text{ s.t., } A\mathcal{T} = \begin{bmatrix} \mu_t^* \\ \nu_t^* \end{bmatrix}. \quad (8)$$

Specifically, we use the Sinkhorn algorithm—an iterative solution for Eq. 8 [1]—to find $\mathcal{T}$ as implemented in the Python Optimal Transport toolbox [11].

5 Results

Datasets: To evaluate our approach, we used data from the Human Connectome Project (HCP) [27], starting with the minimally preprocessed data [13]. First, data with a maximum frame-to-frame displacement of 0.15 mm or greater were excluded, resulting in a sample of 515 resting-state scans. Analyses were restricted only to the LR phase encoding, which consisted of 1200 individual time points. Further preprocessing steps were performed using BioImage Suite [16]. These included regressing 24 motion parameters, regressing the mean white matter, CSF, and grey matter time series, removing the linear trend, and low-pass filtering. After processing, Shen (268 nodes) [24], Schaefer (400 nodes) [22], Craddock (200 nodes) [7], Brainnetome (246 nodes) [10], Power (264 nodes) [20], and Dosenbach (160 nodes) [6] atlases were applied to the preprocessed to create mean timeseries for each node. Connectomes were generated by calculating the Pearson’s correlation between each pair of these mean timeseries and then taking the fisher transform of these correlations.

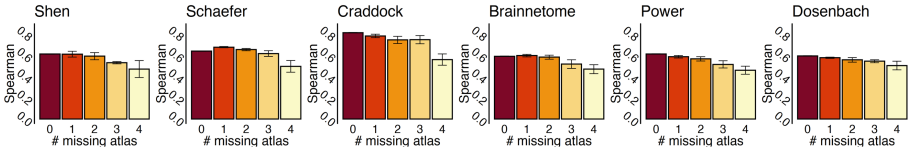


Fig. 3. Bar plots exhibit correlation of estimated connectomes and original connectomes based on $n - k$ samplings of available atlases (*i.e.*, n indicates the number of all available atlases to be transported) for each target atlas for all-way optimal transport. Strong correlations can be observed with less than the maximum number of source atlases.

Similarity Between Reconstructed and Original Connectomes: To validate our approach, we assessed the similarity of connectomes reconstructed using the proposed optimal transport algorithms and the original connectomes generated directly from the raw data. First, We partitioned our sample into 80% for optimal ‘parameter estimation’. These optimal parameters were then applied on 20% remaining data for measuring the efficacy of the method. Therefore, we estimated $\mathcal{T}$ using all 1200 time points and 412 participants for each source-target atlas pairs (for single-source optimal transport) as well as using all available source atlases to a single target atlas (for all-way optimal transport).

Next, in the left out partition, we applied the estimated $\mathcal{T}$ to reconstruct the target atlases. Finally, the reconstructed connectomes were compared to the “gold-standard” connectomes (*i.e.*, connectomes generated directly from an atlas) using correlation. Results from all-way were compared to results from the single-source optimal transport algorithm.

As shown in Fig. 2, we observed strong correlation between the reconstructed connectomes and their original counterparts when using the all-way optimal

transport algorithm. In every case, these algorithms produce significantly more similar connectomes than the previous single-source optimal transport algorithm (all $\rho's > 0.50$; $p < 0.01$). For most atlases, explained variance is more than tripled using multiple source atlases compare to using a single source atlas.

Effect of Number of Source Atlases: We investigated the impact of using a smaller number of source atlases by only including k random source atlases when creating connectome for the target atlas. This process was repeated with 100 iterations over a range of $k = 2-6$. As shown in Fig. 3, while similarity between reconstructed and original connectomes increases as the number of source atlases increases, strong correlations (e.g., $\rho > 0.6$) can be observed with as little as two or three source atlases, suggesting that a small number of atlases may be sufficient for most applications. Overall, improvements in similarity level off after combining a few atlases, suggesting that adding a greater number of atlases than tested here will have diminished returns.

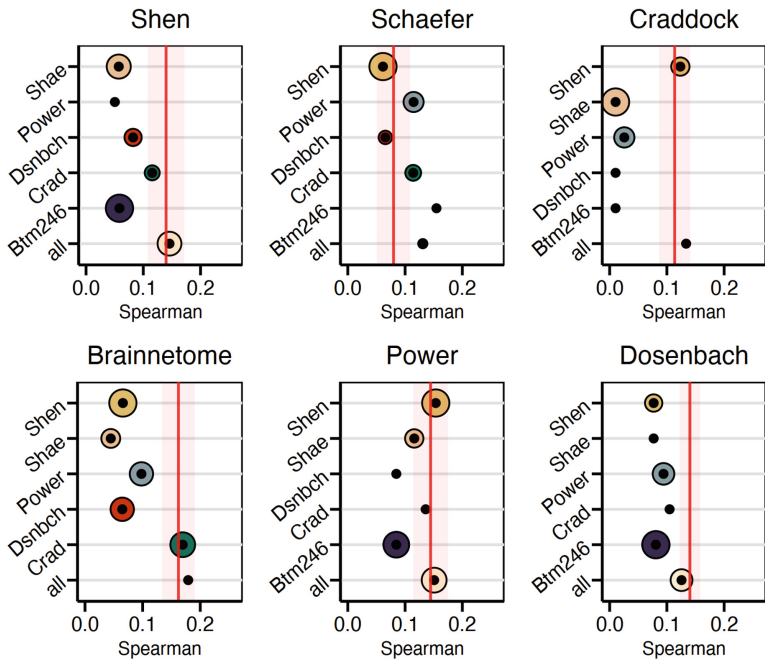


Fig. 4. The reconstructed connectomes using multiple source atlases retain significantly more individual differences than using a single source atlas and predicted IQ as well or better than the original connectome (red line). Size of circle represents the variance of prediction of 100 iteration of 10-fold cross-validation. (Color figure online)

IQ Prediction: To further evaluate the reconstructed connectomes, we show that reconstructed connectomes can be used to predict fluid intelligence using connectome-based predictive modeling (CPM) [23]. We partitioned the HCP dataset into three groupings: g_1 , which consisted of 25% of the participants; g_2 , which consisted of 50% of the participants; and, g_3 , which consisted of the final 25% of the participants. In g_1 , $\mathcal{T}$'s for each algorithm were estimated as above. We then applied $\mathcal{T}$ on g_2 and g_3 to reconstruct connectomes. Finally, for each set of connectomes, we trained a CPM model of fluid intelligence using g_2 and tested this model in g_3 . Spearman correlation between observed and predicted values was used to evaluate prediction performance. This procedure was repeated with 100 random splitting of the data into the three groups. In all cases, connectomes reconstructed using all of the source atlases performed as well in prediction as the original connectomes (Fig. 4).

Parameter Sensitivity: We investigated the sensitivity of all-way optimal transport to the free parameters: frame size, training size, and entropy regularization (see Eq. 8). We observe stable correlations with original connectomes using different frame sizes, emphasizing that our cost matrix captures the geometry between the different atlases well. Also, all-way optimal transport is trainable with limited amount of data (see Fig. 5). Finally, increasing entropy regularization ϵ overly penalizes the mapping and degrades the quality of connectomes.

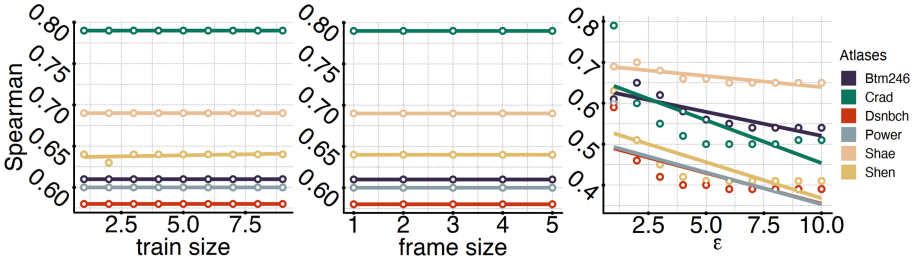


Fig. 5. Parameter sensitivity of frame size ($\times 50$), training data ($\times 100$), and entropy regularization ϵ for different target atlases using all-way optimal transport.

6 Discussion and Conclusions

Here, we significantly improve upon previous efforts to enable fMRI data, previously processed with one atlas, to be mapped to a connectome generated from a different atlas, without the need for further preprocessing. To accomplish this, we proposed and validate two algorithms that combine information from multiple source atlases to better estimate connectomes for the target atlas. All-ways optimal transport directly estimates a single mapping between multiple source atlases and the target atlas. In contrast, stacking optimal transport combines previously estimated mappings between a single source and target atlas, allowing

these previously estimated mappings to be reused. Reconstructed connectomes from both algorithms are more similar to their “gold-standard” counterparts and better at predicting IQ than reconstructed connectomes based on a single source mapping. Importantly, these results hold for a wide-range of different atlases. Future work includes generalizing our framework to other functional timeseries data—*e.g.*, electroencephalography (EEG) and functional near infrared spectroscopy (fNIRS). Overall, our approach is a promising avenue to increase the generalization of connectome-based results across different atlases.

Acknowledgements. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; U54 MH091657) and funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Amin Karbasi is partially supported by NSF (IIS-1845032), ONR (N00014-19-1-2406), and Tata.

Compliance with Ethical Standards. This research study was conducted retrospectively using human subject data made available in open access by the Human Connectome Project. Approval was granted by local IRB. Yale Human Research Protection Program (HIC #2000023326) on May 3, 2018.

References

1. Altschuler, J., Weed, J., Rigollet, P.: Near-linear time approximation algorithms for optimal transport via sinkhorn iteration. arXiv preprint [arXiv:1705.09634](https://arxiv.org/abs/1705.09634) (2017)
2. Arslan, S., Ktena, S.I., Makropoulos, A., Robinson, E.C., Rueckert, D., Parisot, S.: Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage* **170**, 5–30 (2018). <https://doi.org/10.1016/j.neuroimage.2017.04.014>, <https://www.sciencedirect.com/science/article/pii/S1053811917303026>, segmenting the Brain
3. Bertsimas, D., Tsitsiklis, J.: Introduction to linear optimization, Athena scientific (1997). <http://athenasc.com/linoptbook.html>
4. Birkhoff, G.: Tres observaciones sobre el algebra lineal. *Univ. Nac. Tucuman, Ser. A* **5**, 147–154 (1946)
5. Casey, B., et al.: The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018)
6. Cohen, A.L., et al.: Defining functional areas in individual human brains using resting functional connectivity MRI. *Neuroimage* **41**(1), 45–57 (2008)
7. Craddock, R.C., James, G.A., Holtzheimer, P.E., III., Hu, X.P., Mayberg, H.S.: A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapping* **33**(8), 1914–1928 (2012)
8. Dadashkarimi, J., Karbasi, A., Scheinost, D.: Data-driven mapping between functional connectomes using optimal transport. In: de Bruijne, M., et al. (eds.) MIC-CAI 2021. LNCS, vol. 12903, pp. 293–302. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87199-4_28
9. Dantzig, G.B.: Reminiscences about the origins of linear programming. In: *Mathematical Programming The State of the Art*, pp. 78–86. Springer, Heidelberg (1983). https://doi.org/10.1007/978-3-642-68874-4_4

10. Fan, L., et al.: The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* **26**(8), 3508–3526 (2016)
11. Flamary, R., Courty, N.: Pot python optimal transport library (2017). <https://pythonot.github.io/>
12. Gangbo, W., McCann, R.J.: The geometry of optimal transportation. *Acta Math.* **177**(2), 113–161 (1996)
13. Glasser, M.F., et al.: The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013)
14. Hitchcock, F.L.: The distribution of a product from several sources to numerous localities. *J. Math. Phys.* **20**(1–4), 224–230 (1941)
15. Horien, C., et al.: A hitchhiker’s guide to working with large, open-source neuroimaging datasets. *Nat. Hum. Behav.* 1–9 (2020)
16. Joshi, A., et al.: Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics* **9**(1), 69–84 (2011)
17. Kantorovich, L.: On the transfer of masses In: *Doklady Akademii Nauk*, vol. 37, pp. 227–229 (1942). (in russian)
18. Koopmans, T.C.: Optimum utilization of the transportation system. *Econometrica J. Econ. Soc.* **17**, 136–146 (1949)
19. Peyré, G., Cuturi, M., et al.: Computational optimal transport: With applications to data science. *Found. Trends® Mach. Learn.* **11**(5–6), 355–607 (2019)
20. Power, J.D., et al.: Functional network organization of the human brain. *Neuron* **72**(4), 665–678 (2011)
21. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover’s distance as a metric for image retrieval. *Int. J. Comput. Vis.* **40**(2), 99–121 (2000)
22. Schaefer, A., et al.: Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**(9), 3095–3114 (2018)
23. Shen, X., et al.: Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat. Protoc.* **12**(3), 506 (2017)
24. Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T.: Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* **82**, 403–415 (2013)
25. Sudlow, C., et al.: Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**(3), e1001779 (2015)
26. Tolstoi, A.: Methods of finding the minimal total kilometrage in cargo transportation planning in space. *Trans. Press Natl. Commissariat Transp.* **1**, 23–55 (1930)
27. Van Essen, D.C., et al.: The WU-Minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013)