
Barycenters in the brain: An optimal transport approach to modeling connectivity

Max Dabagia^b and Eva Dyer^{b,#}

^b - School of Electrical and Computer Engineering

[#] - Coulter Department of Biomedical Engineering

Georgia Institute of Technology, Atlanta, GA

{maxdabagia, evadyer}@gatech.edu

Abstract

A growing body of evidence points to the distributed nature of information processing in the brain, with many distinct functional areas like the visual or auditory cortex being implicated in brain-wide computations. Thus, understanding how whole-brain circuits are organized and impacted by learning and disease is an important problem that has wide reaching implications. Here, we introduce a novel approach for learning a generative model for shape-based maps of the long-range connections in the brain. To do this, we define a collection of anchors drawn from our dataset of “projections”, which are spatial distributions representing the strength of long-range connections between regions. Our approach leverages Wasserstein barycenter regression to transform each observed projection into an interpolation of the anchor projections, and show that this representation preserves meaningful information about both the regions connected and the shape of the circuit traveled. It is also capable of generating novel samples at given coordinates by mapping into this barycenter space and back. We applied our method to sets of whole-brain viral tracing experiments in the Mouse Connectivity Atlas from the Allen Institute for Brain Science and demonstrate that our approach is capable of representing a variety of patterns of connectivity, including cortico-thalamic maps present in the somatosensory system.

1 Introduction

A growing body of evidence points to the distributed nature of information processing in the brain [1]. Many distinct areas are involved in brain-wide computations, whether it be in explicitly multi-sensory tasks (like language processing [2]) or even in tasks that at first glance appear uni-sensory (like object recognition [3]). Characterizing whole brain circuits that underlie these computations is critical, and revealing the underlying shape and diversity of connections between specific areas is still an active topic of research in the neuroscience community.

Methods for modeling the architecture of these whole-brain circuits (often called the “connectome”) have thus far used graphs and predictive linear models [4] to capture the strength and direction of projections between regions-of-interest. These graph-based connectivity models essentially model the whole brain as an intricate network, based upon the strength or proportion of connections from one area to another. However, when representing these rich volumetric data through a graph, information about the shape and geometry of the projection is unfortunately lost. There are many critical applications which rely on capturing the shape and morphology of projections, including discovery of new cell types [5], characterizing differences in connections between the same areas across individual subjects [6], and understanding the effects of aging and disease on neural computation [7]. Thus models that preserve spatial information and geometry are indispensable in developing maps of the

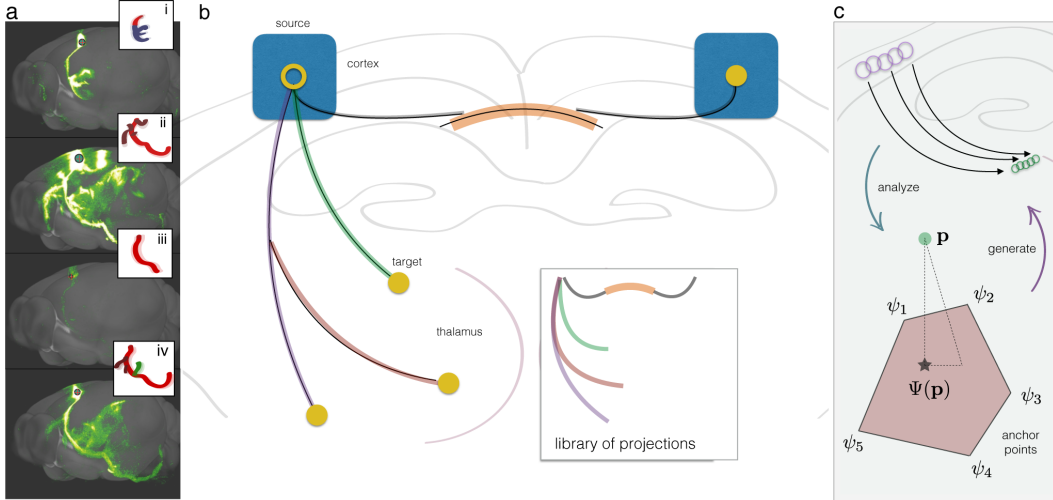


Figure 1: *Overview of the model and its application to modeling whole-brain circuits.* On the left, in (a) we provide examples of images of cell-type specific projections from the Mouse Connectivity Atlas (green), with different projections highlighted in each example. In (b), we show the idea behind our spatial projectome model, where a dataset and circuit is defined through a spatial mixture of projections from a library of possible projections (defined for a source area). By combining these projections into a spatial mixture, this generates a “projectome”. To model geometric variability in a particular projection (part) in the model (c), we construct a local generator which finds a map between an the image space to a set of coordinates that describe the projection in terms of a set of anchors (other experiments drawn from the dataset). Below, we depict how the selected anchors are used to represent other projections in terms of their coefficients in Wasserstein barycenter space, after being mapped onto the simplex which has the anchors as vertices.

connections in the brain that push our understanding of the connectome beyond a purely functional view of which regions are connected to how and why the pathways diverge.

Here, we introduce a new model for high-resolution maps of connectivity (called projectomes [8]) drawn directly from real images that can be used to reconstruct 3D connection profiles with spatial detail. We call this model “generative” because it attempts to represent the full range of variation for connections between given areas, such that plausible new examples could potentially be created. Our formulation uses a generalization of the concept of center of mass arising in the study of optimal transport. The simple physical problem of finding a center of mass (in physics, a barycenter) can be framed as finding an interpolation between the positions of several points, such that the sum of the distance to each point weighted by its mass is minimized. The *Wasserstein barycenter* [9] extends this concept to a collection of histograms instead of points, and replaces Euclidean distance with the Wasserstein distance, which is the total work required to transform one histogram into another. We construct a barycenter starting from a collection of shapes or “anchors” which can be incrementally deformed into each other to capture the range of variability seen in each pathway of interest. We then reconstruct a projection by finding a weighted interpolation between anchors that most resembles the target distribution.

Additionally, we show that when using our generative OT-based modeling approach on connectivity data, the representations provided purely by the weights best approximating a shape can be used to recover important signatures of variability across profiles, as well as succinctly describing the shape as a combination of the anchors. We applied this method to map cortico-thalamic circuits present in the somatosensory system, as revealed by whole mouse brain cell-type specific viral tracing datasets (Ntsr and Syt6 cre-lines) [10]. This result thus demonstrates the potential applications of this generative modeling approach for shape-based modeling of projectome datasets.

Experimental results suggest that the flexible generative model provided by Wasserstein barycenters is a useful geometric model for whole-brain connectivity profiles. Our results thus open up new opportunities to apply convolutional and other Wasserstein barycenter methods to high-resolution connectomics datasets. These methods will provide new strategies for understanding disease and other pathologies that impact whole-brain connectivity and function.

1.1 Contributions

The contributions of this paper are as follows:

- *A novel formulation for modeling mesoscale connectivity data:* We present a new abstraction for modeling variability in high-resolution whole-brain datasets like those recently introduced by the Allen Institute (Mouse Connectivity Atlas) [11]. This generative approach can be used for modeling cell type-specific projections, an important challenge in modern neuroanatomy.
- *A new method for Wasserstein barycenter-based generative modeling:* We couple barycenter coordinate regression with subset selection of representative anchors that span different parts of a brain area. Combining these methods provides a simple yet effective strategy for defining a geometric generative model.
- *Demonstration that our barycenter-based generative model provides insight into whole-brain circuits:* Optimal transport and barycenter regression play an essential role in the performance of the method, which can both compress existing samples and generate new ones in meaningful ways.

1.2 Background and Related Work

Reminders on Optimal Transport: Over the space of normalized histograms with N bins, which we denote as Σ_N , we let C_{ij} be the cost of moving a unit of mass between the i th and j th bins. In Euclidean space we take this cost to be ℓ^2 distance, such that $C_{ij} \triangleq \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$. We define the entropy-regularized Wasserstein distance [12] between two histograms as:

$$\mathcal{W}_{2,\gamma}^2(\mathbf{p}, \mathbf{q}) \triangleq \inf_{\mathbf{P} \in \mathcal{B}_N} \sum_{i,j} C_{ij} \mathbf{P}_{ij} - \gamma H(\mathbf{P})$$

where $\mathbf{p}, \mathbf{q} \in \Sigma_N$, \mathbf{P} is a valid transport plan (in practice, a joint probability distribution) between them, and γ is a constant specifying the strength of the regularization of the entropy of \mathbf{P} , defined as $H(\mathbf{P}) = -\sum_{i,j} \mathbf{P}_{ij} \log \mathbf{P}_{ij}$. We define a distance kernel $\mathbf{K}_\gamma \triangleq e^{-C/\gamma}$, and then note that the cost-minimizing transport plan with the entropy penalty takes the form $\mathbf{P}^* = \text{diag}(\mathbf{u}) \mathbf{K}_\gamma \text{diag}(\mathbf{v})$ for some $\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^N$. These vectors are readily computed at each iteration t by Sinkhorn's algorithm [12]:

$$\mathbf{u}^t = \frac{\mathbf{p}}{\mathbf{K}_\gamma \mathbf{v}^{t-1}}, \quad \mathbf{v}^t = \frac{\mathbf{q}}{\mathbf{K}_\gamma^\top \mathbf{u}^t}.$$

We interpret this as iterative Bregman projections of the kernel \mathbf{K}_γ onto the constraint sets $\mathbf{P} \mathbb{1} = \mathbf{p}$ and $\mathbf{P}^\top \mathbb{1} = \mathbf{q}$ until convergence [13].

The space requirement of the distance kernel scales exponentially in the dimension and would be prohibitively large for most domains of practical interest (e.g. images), so following [14] we note that the kernel is translation-invariant, and under a Euclidean distance cost its multiplicative application is actually equivalent to convolution with a Gaussian of variance $\sigma^2 = \gamma/2$. This eliminates the need to store the distance kernel entirely.

Wasserstein Barycenters: After establishing Wasserstein distance as a distance over histogram space Σ_N , we can introduce the concept of a Wasserstein barycenter of a set of histograms $\Psi = \{\psi_i \in \Sigma_N\}_{i=1}^S$. Much like the familiar physical notion of center of mass, the barycenter corresponds to a single histogram $\beta \in \Sigma_N$ for which the weighted Wasserstein distance to each distribution in Ψ is minimized. We define β given a set of weights $\alpha \in \Sigma_S$ as:

$$\beta_\alpha = \arg \min_{\beta \in \Sigma_N} \sum_i \alpha_i W_{2,\gamma}(\beta, \psi_i).$$

For uniform weights α this is the Wasserstein mean of Ψ , but we observe that taking α to be any point in Σ_S defines a simplex of barycenters in Σ_N , with vertices (or anchors) $\beta_{\delta_i} = \psi_i$. A set of weights α then defines the coordinates of a weighted barycenter β_α within this space. As demonstrated in

[13], Sinkhorn’s algorithm can be extended to simultaneously estimate the barycenter and compute its distance to the histograms of Ψ , as follows:

$$\beta^t = \prod_{i=1}^S (\mathbf{v}_i^{t-1} \odot \mathbf{K}_\gamma^\top \mathbf{u}_i^t)^{\alpha_i}$$

$$\mathbf{u}_i^t = \frac{\psi_i}{\mathbf{K}_\gamma \mathbf{v}^t}, \quad \mathbf{v}_i^t = \frac{\beta^t}{\mathbf{K}_\gamma^\top \mathbf{u}^t}$$

In practice this method converges rapidly, producing reasonable approximations of the barycenter in fewer than ten iterations.

Related Work: Methods derived from optimal transport have previously been applied to neuroscience datasets in solving the source imaging problem, which arises in processing MEG and EEG data [15]. Much as in our model, optimal transport is attractive in this situation because it respects the notion of distance inherent to the physical space of the brain. An example closer to the work we present here is the application of Wasserstein barycenters as a means of combining maps of the cortical surface of different subjects into a single atlas [16]. Wasserstein barycenters prove capable of generating much sharper combinations of individual maps than Euclidean averaging.

2 Method

Mesoscale Connectomics Dataset: As a demonstration of the capabilities of our method, we focus our attention to a subset of the Mouse Connectivity Atlas (MCA) from the Allen Institute for Brain Science [11], which uses fluorescent viral tracing in transgenic mice to resolve long-range projections across the brain (**Figure 1a**). This figure highlights some of the experiments that are encountered in practice. In (i), we show an example of a projection pattern obtained from a highly localized cre line, where only cells in Layer 6 of cortex are visualized (infected). In (ii), we show a complex projection pattern in a wild type animal, where all cell types are highlighted and multiple projections are visible. In (iii), a sparse projection with some “missing data” due to either segmentation errors or low intensity and non-uniform signal. Finally, in (iv) we show an experiment with intermediate expression where a small (sparse) number of projections are visualized simultaneously.

In our experiments, we focus on projections from somatosensory cortex to thalamus. This region is of particular interest because areas within this region are well-understood to map to tactile stimulation of specific body parts, with barrel fields (whiskers), mouth and nose, and trunk and lower limbs, each having characteristic shapes and locations that are reasonably consistent across subjects.

Barycenter Coordinate Regression: Our first task to calculate the barycenter coordinates of a histogram. Consider $\Psi = \{\psi_i \in \Sigma_N\}_{i=1}^S$, our set of anchor histograms. With slight abuse of notation, we define $\Psi: \Sigma_N \mapsto \Sigma_S$ as the barycenter transform operator, which associates a histogram \mathbf{p} with its nearest set of coordinates on the barycenter simplex of the anchor histograms. Nearness is defined according to the choice of loss function (typically ℓ^2 or Wasserstein) as follows:

$$\Psi(\mathbf{p}) = \arg \min_{\alpha \in \Sigma_S} \mathcal{L}(\beta_\alpha, \mathbf{p})$$

As proposed in [17], we can calculate a gradient for this function by recursive differentiation of the Wasserstein barycenter algorithm. We then compute the coordinate coefficients by supplying the gradient to a descent-based optimizer, such as L-BFGS [18].

Anchor Selection: Given this general algorithm for expressing one histogram in terms of a set of others, we must choose a small collection of samples from the dataset to be our anchors, which are the vertices of the barycenter simplex used to represent the others. We found that this is a delicate process; a simplex that yields unique coordinates can’t have too much redundancy across anchors, but too much disparity and it becomes impossible to accurately reconstruct examples after the forward mapping. Heuristically, we restricted our anchors to one from each region of interest, but within that region selected projections close to the mean and avoided outliers.

Generative Model: We can efficiently generate new samples using this technique as well. An injection is simulated by placing a Gaussian mass at the injection coordinates of interest. We then

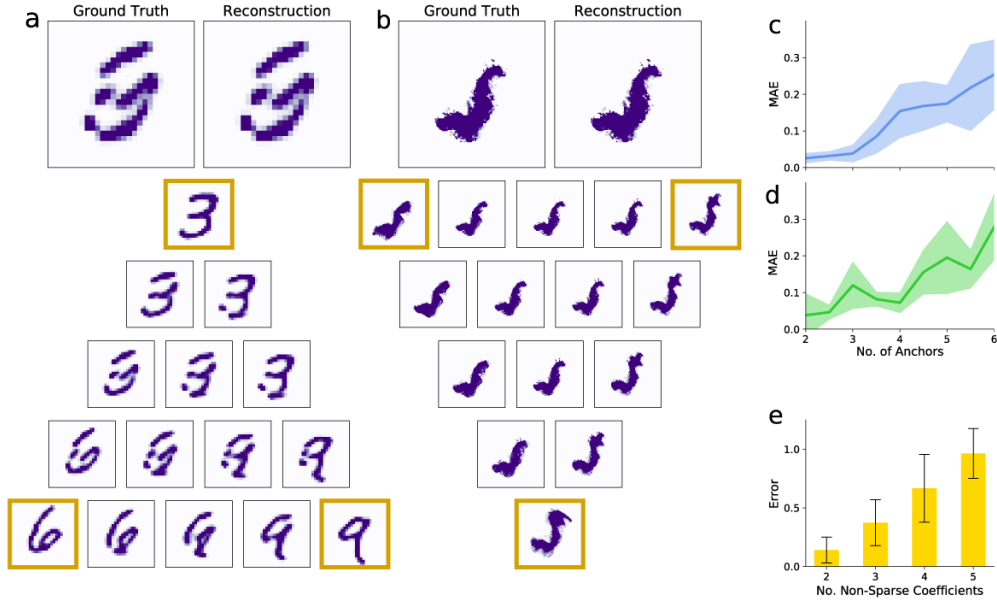


Figure 2: Results of barycenter regression and interpolation over a simplex of anchors. Results of barycenter regression to on simplices of anchors from MNIST (a) and of our data (b) with synthetic weights, to ensure we can recover them reliably with our method. Below, we demonstrate the expressive power of this method by interpolating between the anchors of the respective simplices. On the right, we compare the error of recovered weights on the MNIST simplex in two different situations: (c) anchors are held constant and weights are randomly sampled, and in (d) we hold weights constant and randomly select anchors from the digits 0 to 9. We report absolute error across 25 trials for each number of anchors. In (e) we compare the results of regression with varying degrees of sparsity of random coefficients for seven anchors, chosen from our connectomics dataset. As anticipated, redundancy in such a collection means coefficients rapidly become unrecoverable, since multiple anchors can yield equally faithful reconstructions with similar weights.

find the k nearest neighbors in our dataset of this location in injection space, and find the barycentric coordinates of this mass on their simplex. The new sample is then generated by mapping back to histogram space, through interpolation between the k neighbors.

3 Results

To first confirm that we can recover the correct model coefficients from observed data, we applied our method for barycenter regression to synthetic examples where the true coefficients are known and weighted barycenters are generated from a known collection of anchor histograms. We can do this both for digits in MNIST and also for our connectomics dataset of interest (Figure 2a, b). In both cases, we sample coefficients from a gamma distribution with shape $k = 1$ and unit mean and normalize to a valid probability distribution, apply barycenter regression to recover the underlying coefficients, and then measure the error in recovery (Figure 2c, d, e).

Our results demonstrate that it is possible to accurately recover the correct coefficients under certain assumptions about the makeup of the collection of anchors. The set cannot contain anchors that are nearly identical, or else the associated coefficients of similar anchors will be distributed across them; conversely, the anchors cannot be *too* disparate, because the barycenter of several anchors with limited overlapping support and dramatically different contours tends toward an amorphous mass. As expected, when redundant histograms are incorporated into the set, correctly recovering the coefficients becomes increasingly difficult (Figure 2c, d, e). This spreading effect is similar in least-squares regression across correlated predictors, and thus suggests that a carefully selected set of anchors is necessary to build more meaningful and interpretable representations of data using Wasserstein barycenter regression.

Next, we sought to understand how the coefficients in the expansion could be used to understand information about the shape and projection patterns of mesoscale connectomes. Therefore, we selected a set of representative anchors from the somatosensory dataset and applied barycenter regression to the remaining projections, and confirmed that the coefficients obtained are indeed

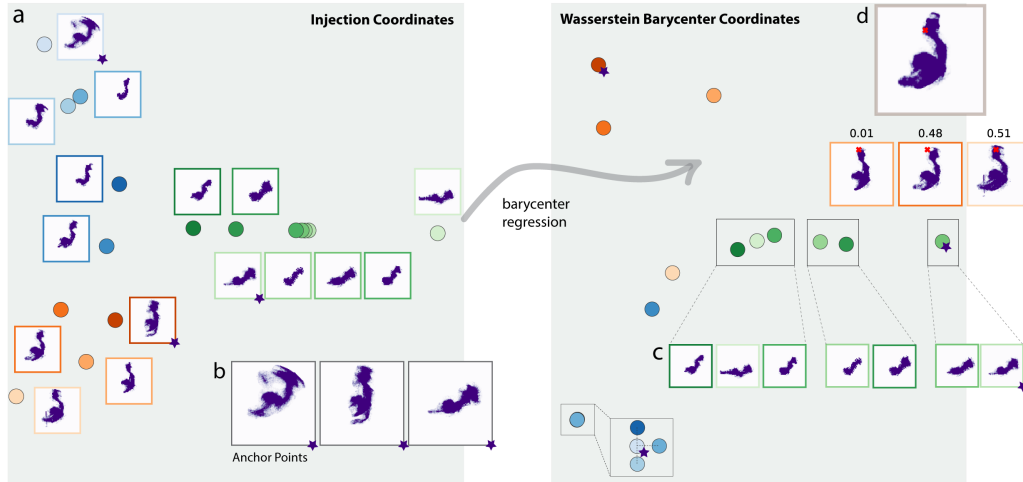


Figure 3: *Barycenter Coordinate Space*. We compare the representation of the dataset in injection coordinates (a) versus the coordinates of projection onto Wasserstein barycenter coordinate (WBC) space (c), with anchors displayed in (b). The barycenter transform displays several desirable properties; for one, the mapping clearly preserves information about individual morphology, since it allows reconstruction of samples with reasonable fidelity. Additionally, region prediction of a sample based solely its highest coefficient in WBC gave correct predictions for all but one sample, compared to 3 misclassified by injection coordinate prediction. Finally, we generate a new sample (d) by projecting a simulated mass placed at the prospective injection site (red 'X') onto the barycenter simplex of its 3 nearest neighbors in injection space and reconstructing the result.

meaningful (Figure 3a, b, c), through visualizing coefficients after Isomap embeddings and also through nearest neighbor-based predictions of subregion (mouth nose, barrel fields, trunk). We find that barycenter coordinate transformations preserve information about shape and subregion, such that with appropriate choice of anchors the coefficients can be used to predict the structure of injection better than the injection coordinates themselves. Whereas coefficients in injection space are tightly clustered with little space between cortical regions (Figure 3a), barycenter coordinates provide useful and descriptive of variability both inside and across cortical subregions (Figure 3a, c). The flexibility of this shape-based matching approach thus provides more precise classification of source regions from their observed shapes.

Through synthesizing a weighted barycenter with a specific set of coefficients, it is possible to generate a new instance of a connectivity profile. However, it is rarely obvious which combinations of coefficients will yield a desired sample. We propose a solution to this problem via barycenter regression: After placing a Gaussian mass at the desired injection site, we project the resulting image onto the barycenter simplex and examine the results. (Figure 3d). In practice this yields high-quality new mesoscale brain datasets at positions that haven't been previously observed. Thus, our highly data-driven geometric approach for generative modeling of connectomes provides intuitive representations of data and a useful way to model modifications to morphology and shape differences.

4 Discussion

This paper introduces a new approach for modeling the shape and geometry of whole-brain connectivity patterns, which goes beyond graphs to produce a simplified yet meaningful and geometrically-constrained representation.

Optimal transport, and in particular the Wasserstein barycenter formulation that we introduced here, provide a useful way to build local part-based generators to describe neural circuits and build projectomes. However, our analysis on both synthetic datasets (MNIST and Allen data) demonstrate that the selection of anchors to use to form the barycenter representation of new test data points (experiments) is extremely important and will impact both the quality of representations, and also the stability of the method. Our current selection procedures, which use either biological or purely heuristic priors to select landmarks, could be improved further by learning optimal sets of connectivity datasets that provide meaningful barycenter representations.

This paper demonstrates a novel application of OT to model connectivity and shape of whole-brain neural circuits. Thus, this will provide new opportunities to use OT in neuroscience. In particular, our generative model for mesoscale connectivity datasets provides a useful tool for understanding when circuits deviate from their typical patterns of variability. In disease the shape and patterns of connections are known to be altered [19], and accordingly, we hope to use this approach to model disease and in other applications of comparative connectomics.

References

- [1] E. Amico, A. Arenas, and J. Goñi, “Centralized and distributed cognitive task processing in the human connectome,” *Network neuroscience (Cambridge, Mass.)*, vol. 3, no. 2, pp. 455–474, 2019. DOI: 10.1162.[Online].AAvailable: <https://www.ncbi.nlm.nih.gov/pubmed/30793091>.
- [2] I. Blank, Z. Balewski, K. Mahowald, and E. Fedorenko, “Syntactic processing is distributed across the language system,” *NeuroImage*, vol. 127, pp. 307–323, 2016. DOI: <https://doi.org/10.1016/j.neuroimage.2015.11.069>. [Online].AAvailable: <http://www.sciencedirect.com/science/article/pii/S1053811915011064>.
- [3] J. M. Pakan, V. Francioni, and N. L. Rochefort, “Action and learning shape the activity of neuronal circuits in the visual cortex,” *Current opinion in neurobiology*, vol. 52, pp. 88–97, 2018, ISSN: 1873-6882. DOI: 10.1016.[Online].AAvailable: <https://www.ncbi.nlm.nih.gov/pubmed/29727859>.
- [4] S. Henriksen, R. Pang, and M. Wronkiewicz, “A simple generative model of the mouse mesoscale connectome,” *eLife*, vol. 5, e12366–e12366, 2016. DOI: 10.7554/eLife.12366.[Online].AAvailable: <https://www.ncbi.nlm.nih.gov/pubmed/26978793>.
- [5] H. Zeng and J. R. Sanes, “Neuronal cell-type classification: Challenges, opportunities and the path forward,” *Nature Reviews Neuroscience*, vol. 18, no. 9, p. 530, 2017.
- [6] F.-C. Yeh, J. M. Vettel, A. Singh, B. Poczos, S. T. Grafton, K. I. Erickson, W.-Y. I. Tseng, and T. D. Verstynen, “Quantifying differences and similarities in whole-brain white matter architecture using local connectome fingerprints,” *PLoS computational biology*, vol. 12, no. 11, e1005203–e1005203, 2016, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005203.[Online].AAvailable: <https://www.ncbi.nlm.nih.gov/pubmed/27846212>.
- [7] J. S. Damoiseaux, “Effects of aging on functional and structural brain connectivity,” *NeuroImage*, vol. 160, pp. 32–40, 2017, Functional Architecture of the Brain, ISSN: 1053-8119. DOI: <https://doi.org/10.1016/j.neuroimage.2017.01.077>. [Online].AAvailable: <http://www.sciencedirect.com/science/article/pii/S1053811917301015>.
- [8] N. Kasthuri and J. W. Lichtman, “The rise of the ‘projectome’,” *Nature Methods*, vol. 4, no. 4, p. 307, 2007.
- [9] M. Aguehand G. Carlier, “Barycenters in the wasserstein space,” *SIAM J. Math. Analysis*, vol. 43, pp. 904–924, Jan. 2011. DOI: 10.1137/100805741.
- [10] J. A. Harris, S. Mihalas, K. E. Hirokawa, J. D. Whitesell, J. Knox, A. Bernard, P. Bohn, S. Caldejon, L. Casal, A. Cho, *et al.*, “The organization of intracortical connections by layer and cell class in the mouse brain,” *BioRxiv*, p. 292961, 2018.
- [11] S. W. Oh, J. A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan, A. M. Henry, *et al.*, “A mesoscale connectome of the mouse brain,” *Nature*, vol. 508, no. 7495, p. 207, 2014.
- [12] M. Cuturi, “Sinkhorn distances: Lightspeed computation of optimal transportation distances,” *Advances in Neural Information Processing Systems*, vol. 26, Jun. 2013.
- [13] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, “Iterative Bregman Projections for Regularized Transportation Problems,” *arXiv e-prints*, arXiv:1412.5154, arXiv:1412.5154, 2014. arXiv: 1412.5154.
- [14] J. Solomon, F. De Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas, “Convolutional wasserstein distances,” *ACM Transactions on Graphics*, vol. 34, Jul. 2015. DOI: 10.1145/2766963.

- [15] H. Janati, T. Bazeille, B. Thirion, M. Cuturi, and A. Gramfort, "Group level MEG/EEG source imaging via optimal transport: minimum Wasserstein estimates," *arXiv e-prints*, arXiv:1902.04812, arXiv:1902.04812, 2019. arXiv: 1902.04812 [stat.ML].
- [16] Z. Chen, Z. Wu, L. Sun, F. Wang, L. Wang, F. Zhao, W. Lin, J. H. Gilmore, D. Shen, and G. Li, "Construction of 4d neonatal cortical surface atlases using wasserstein distance," pp. 995–998, 2019. DOI: 10.1109.
- [17] N. Bonneel, G. Peyré, and M. Cuturi, "Wasserstein barycentric coordinates: Histogram regression using optimal transport," *ACM Transactions on Graphics*, vol. 35, Apr. 2016. DOI: 10.1145/2897824.2925918.
- [18] R. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, Feb. 2003. DOI: 10.1137/0916069.
- [19] J. W. Prescott, A. Guidon, P. M. Doraiswamy, K. Roy Choudhury, C. Liu, and J. R. Petrella, "The alzheimer structural connectome: Changes in cortical network topology with increased amyloid plaque burden," *Radiology*, vol. 273, no. 1, pp. 175–184, 2014, ISSN: 1527-1315. DOI: 10.1148/radiol.14132593.