High resolution neural connectivity from incomplete tracing data using nonnegative spline regression

Kameron Decker Harris[1] (kamdh@uw.edu), Stefan Mihalas[2], Eric Shea-Brown[1]

[1]Applied Mathematics, University of Washington, Seattle, WA USA; [2]Allen Institute for Brain Science, Seattle, WA USA

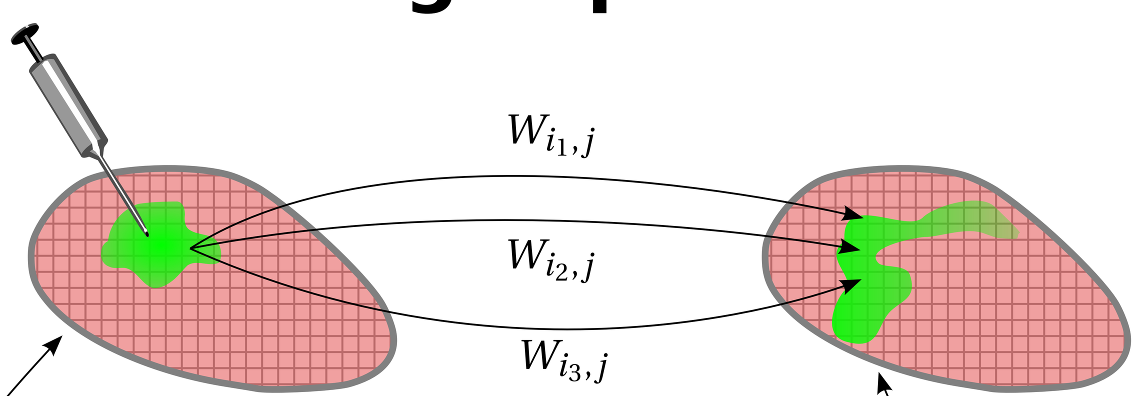
Code, movies, and more!
github.com/kharris/high-res-connectivity-nips-2016

Motivation

Scientific: Understand interplay of brain **network structure** and **information representation** (coding). Examine functional correlations such as spatial maps in relation to structural connectivity.

Mathematical: Develop tractable methods for analyzing high-dimensional structural connectivity data from next-generation experiments.

Tracing experiments

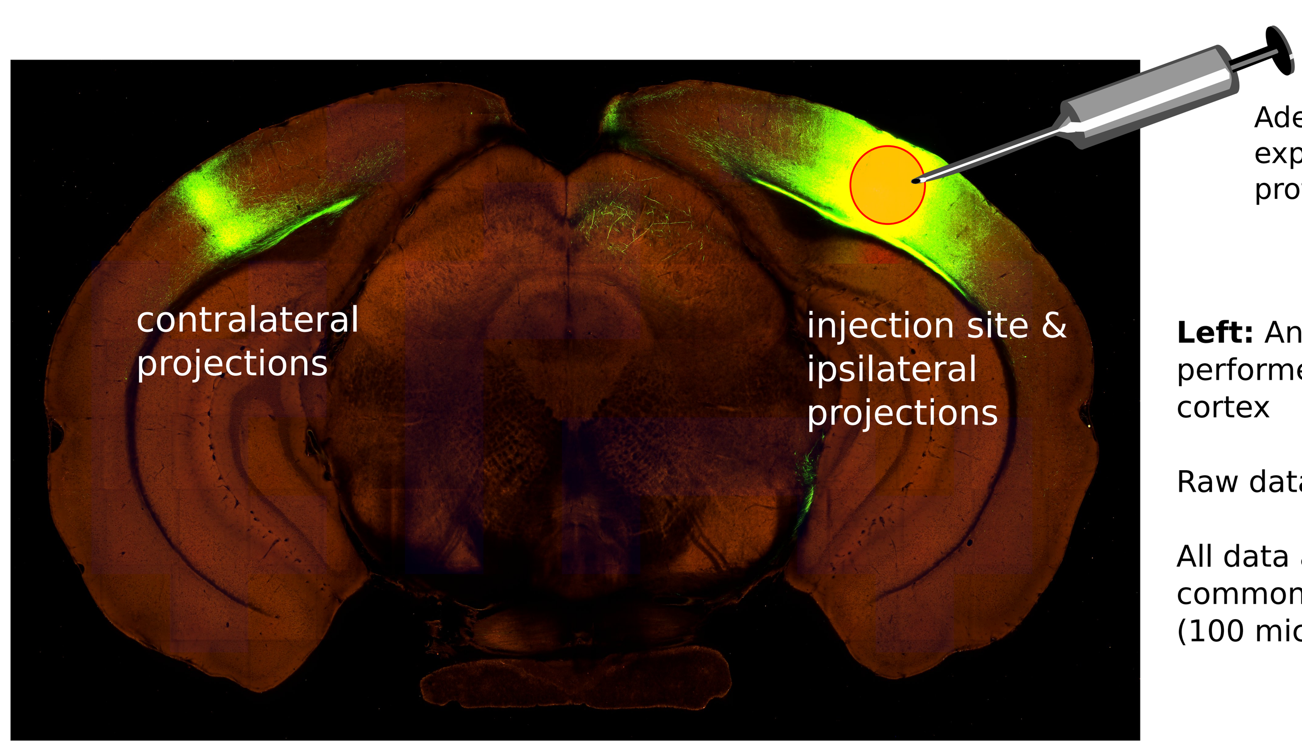


x: source expression (n_x x 1), n_x is # voxels in source regions

y: target expression (n_y x 1), n_y is # voxels in target regions

$$\mathbf{y} \approx W\mathbf{x}$$

Goal: fit unknown (n_y x n_x) weight matrix

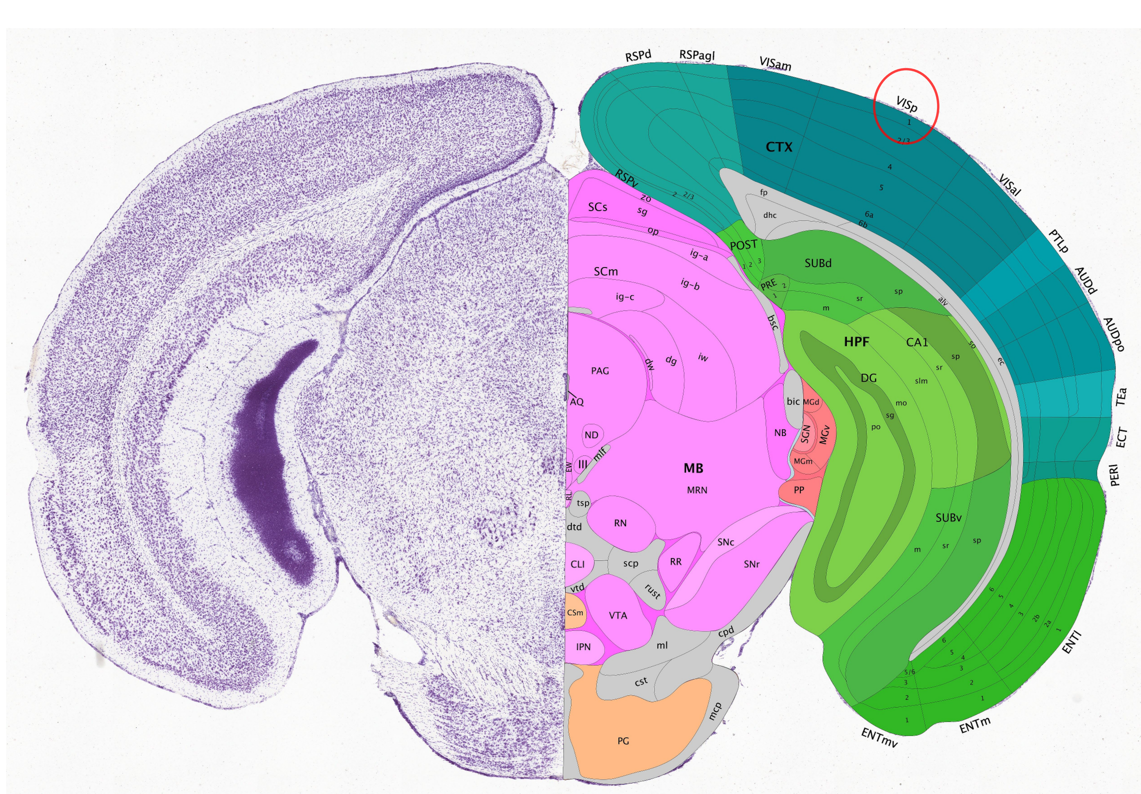


Adeno-associated virus expressing green fluorescent protein

Left: An example of an injection performed in primary visual cortex

Raw data (images) 2 PB

All data are aligned onto a common coordinate system of (100 micron)^3 cubic voxels



The reference atlas labels voxels belonging to known regions (**left**), which are organized in a heirarchy

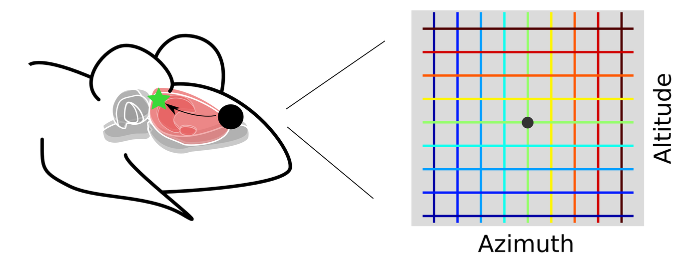
Data summary

1) Mesoscale: 100-1000 neurons/voxel, averages over neurons. Each injection covers 30-500 voxels, 240 avg.
2) Still, projections are identifiable at voxel scale and all regions are targeted by at least 1 injection.
3) Connectome fitting is very underconstrained: 5 x 10^5 voxels and only ~1000 injections.

Above images: Allen Institute for Brain Science

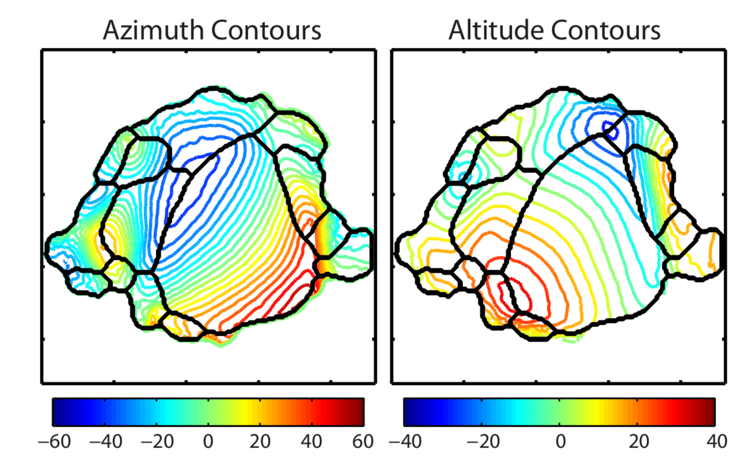
Acknowledgements: We acknowledge the support of the UW NIH Training Grant in Big Data for Neuroscience and Genetics (KDH), Boeing Scholarship (KDH), NSF Grant DMS-1122106 and 1514743 (ESB & KDH), and a Simons Fellowship in Mathematics (ESB). We thank Liam Paninski for helpful insights at the outset of this project. We wish to thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement, and support. This work was facilitated though the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington.

Voxel model strategy: assume smoothness



What the mouse sees

... and it's cortical representation

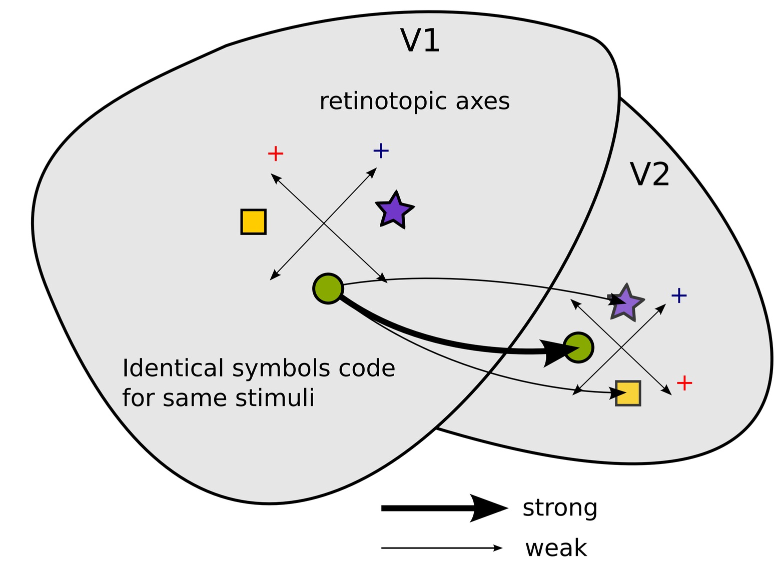


Garrett et al., 2014, J Neurosci 34(37):1258712600

Retinotopy (map representation of visual field) in primary visual cortex is maintained from V1 into deeper areas analagous to V2, etc.

Assumptions:
- like connects strongest to like
- connection strength decreases as similarity decreases
- retinotopy (hence similarity) is smooth

Taken together, suggests connections vary smoothly in space



Smoothness regularized model

Find the voxel-resolution connection matrix W that balances goodness of fit and smoothness:

$$W^* = \arg\min_{W \succeq 0} \underbrace{\|P_\Omega(WX - Y)\|_F^2}_{\text{Goodness of fit (loss) term}} + \underbrace{\lambda \frac{n_{\text{inj}}}{n_x} \left\|L_y W + W L_x^T\right\|_F^2}_{\text{Smoothness (regularization) term}}$$

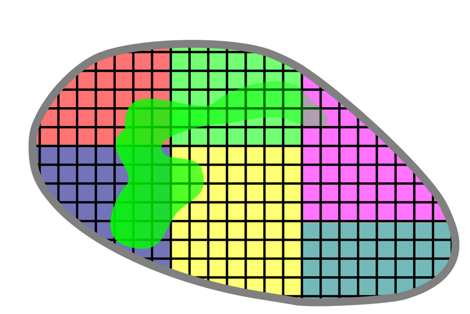
Data matrices:
$Y = [\mathbf{y}_1, \ldots, \mathbf{y}_{n_{\text{inj}}}]$
$X = [\mathbf{x}_1, \ldots, \mathbf{x}_{n_{\text{inj}}}]$

P_Ω: Loss is summed only over voxels outside injection site, as in "matrix completion"

L_x, L_y: finite difference Laplacian matrices. Regularization is analogous to so-called "thin-plate splines" radial basis functions for regression or interpolation

Previous work: regional model

Oh et al., 2014, Nature 508(7495):207-214



colors indicate regions

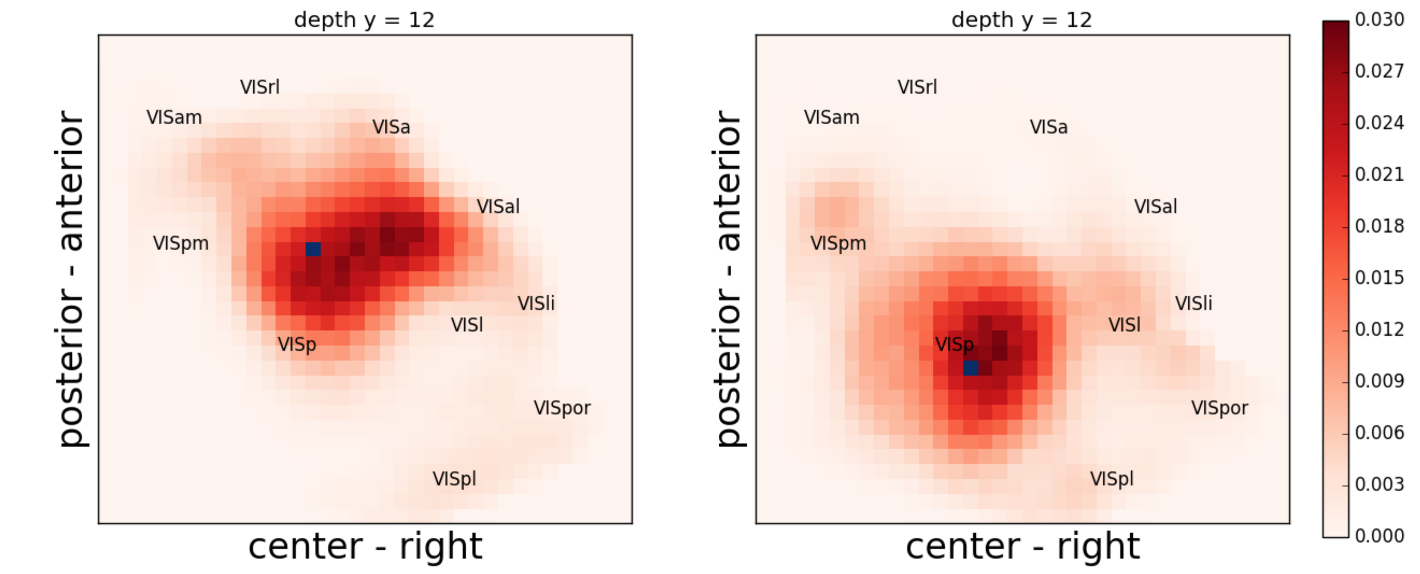
Fit the weight matrix $W^{(r)}$ via non-negative least squares

$$\min_W \left\| Y^{(r)} - W^{(r)} X^{(r)} \right\|_F^2$$

$y^{(r)}_{i,j} = \int_{\text{region } i}$ target signal of experiment j
$x^{(r)}_{i,j} = \int_{\text{region } i}$ source signal of experiment j

This is the same as choosing a voxel scale W where W_{ij} is constrained to be constant for all voxels *i* in region A and *j* in region B, for all regions A and B

Voxel approach begins to reveal spatial connectivity maps



Results of 2 "virtual" injections into VISp, visualizing columns of W

Spatial connectivities are more predictive in cross-validation

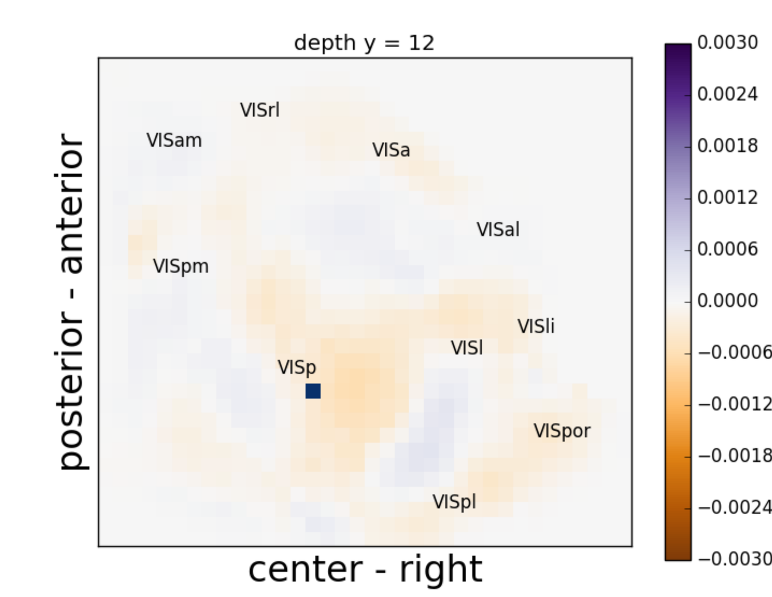
Model	Voxel MSE_{rel}	Regional MSE_{rel}
Regional	107% (70%)	48% (6.8%)
Voxel	**33% (10%)**	**16% (2.3%)**

$$\text{MSE}_{\text{rel}} = \frac{2\|P_\Omega(WX - Y)\|_F^2}{\|P_\Omega(WX)\|_F^2 + \|P_\Omega(Y)\|_F^2}$$

Low rank approximation enables inference that scales

Non-convex, low-rank version of previous spatial model:

$$(U^*, V^*) = \arg\min_{U,V \succeq 0} \|P_\Omega(UV^T X - Y)\|_F^2 + \lambda \frac{n_{\text{inj}}}{n_x} \left\|L_y UV^T + UV^T L_x^T\right\|_F^2$$



Left: Residuals between the full-rank and a rank-160 nonnegative approximation are an order of magnitude less than connectivity.

This corresponds to 23-fold compression of the 7497 x 7497 matrix, with error:

$$\frac{\|U^* V^{*T} - W^*_{\text{full}}\|_F}{\|W^*_{\text{full}}\|_F} = 13\%$$

Factorized matrix provides NMF-like "clustering" of voxels by projections.

- Whole brain: number of voxels is 5 × 10^5
- Same 23-fold compressibility means rank 10^4
- This is still a problem in O(10^9) unknowns
- Compresses W matrix from 1.9 TB to 75 GB

Challenges and future work

- Basis function representations: better than low rank?
- Compare *Cre* cell type-specific connectivity
- Compare spatial maps of retinotopy to connectivity
- Fit entire mouse brain, requires massive parallelization