Smooth network inference from neural tracing data

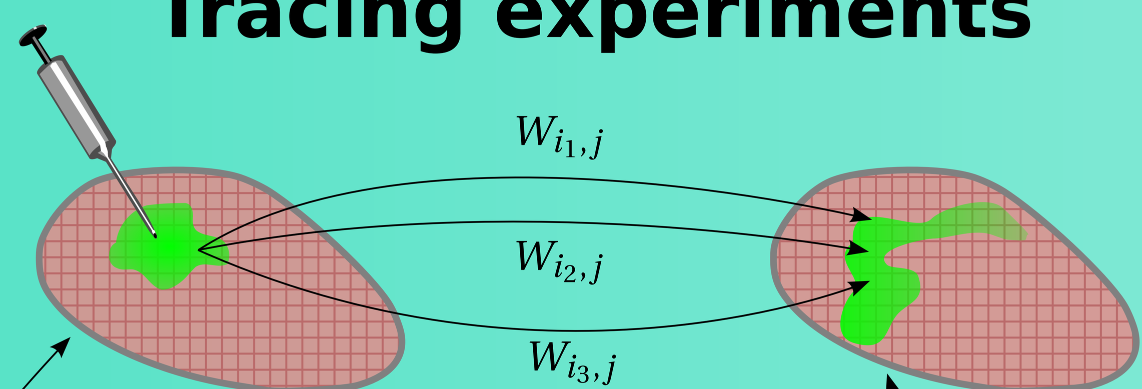
Kameron Decker Harris[1] (kamdh@uw.edu), Nicholas Cain[2], Stefan Mihalas[1,2], Eric Shea-Brown[1]
[1]Applied Mathematics, University of Washington, Seattle, WA USA; [2]Allen Institute for Brain Science, Seattle, WA USA

Motivation

Broadly, understand interplay of brain **structure** and **information representation**

Improve the resolution of the connectivity weight matrix to voxel scale: underconstrained w/o regularization

Tracing experiments

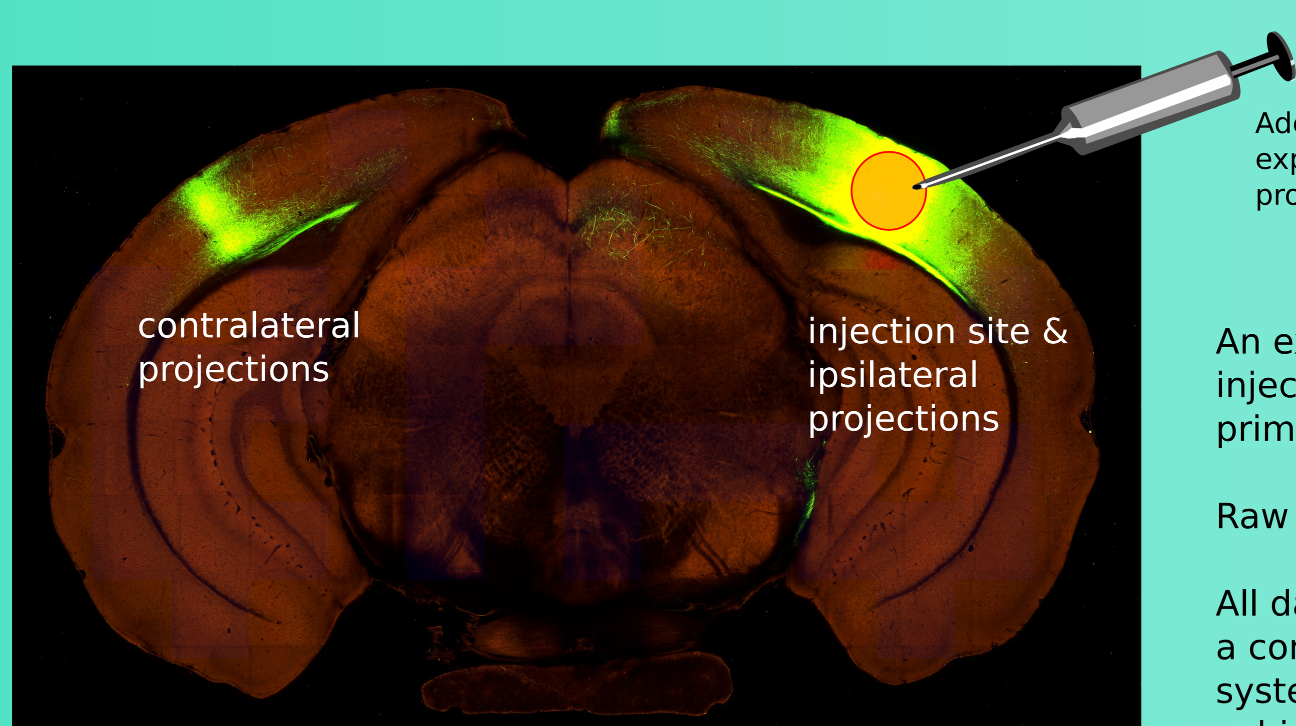


x: source expression (n_x x 1), n_x is # voxels in source regions

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

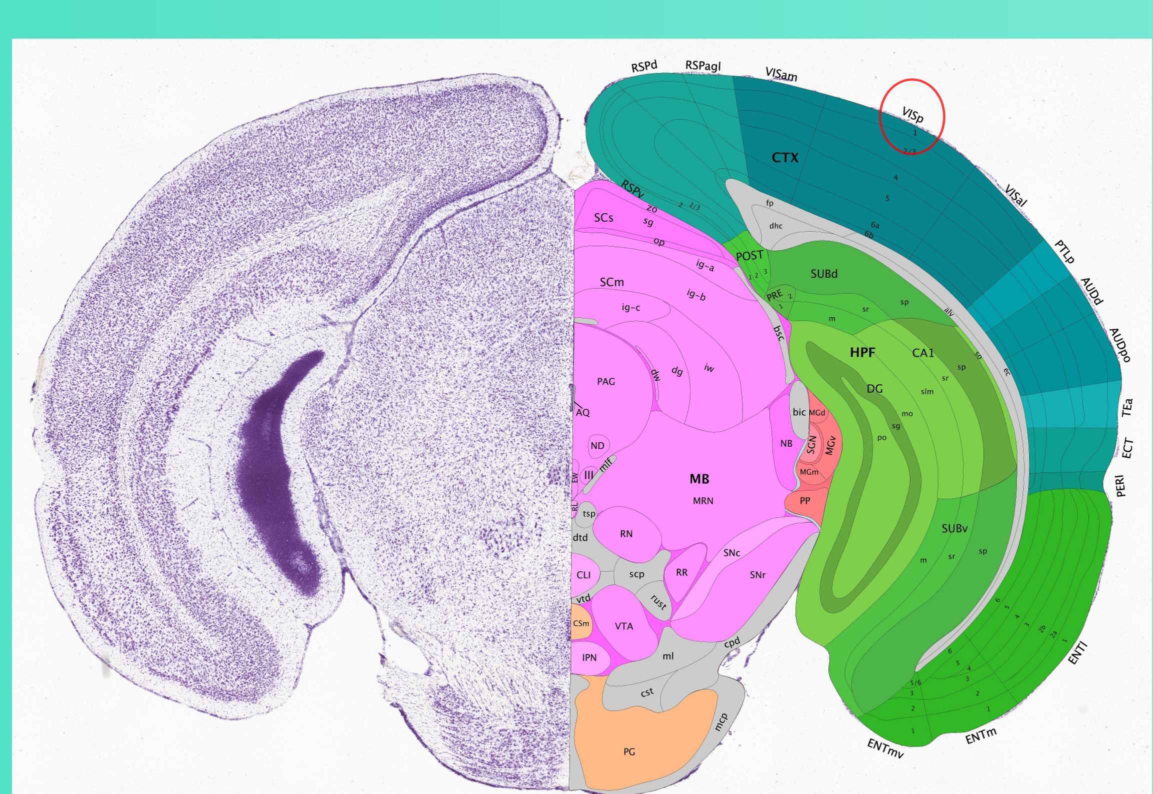


Adeno-associated virus expressing green fluorescent protein

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, which are organized in a heirarchy

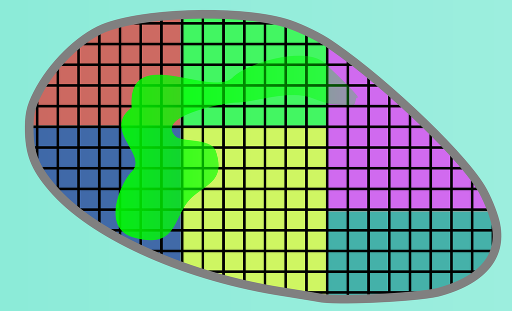
Other information:
- 100-1000 neurons/voxel
- 5 x 10^5 voxels
- each injection covers 30-500 voxels, 240 avg
- projections identifiable at voxel scale
- approx 400 injections in current version of data (many more planned)

Above images: Allen Institute for Brain Science

Acknowledgements: ESB and KDH acknowledge support from NSF Grant #1122106 and a Simons Fellowship in Mathematics. KDH was also supported by a Boeing fellowship. We wish to thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement and support. Thanks also to the computational neuroscience community at UW, including Joel Zylberberg, for suggestions and discussions.

Previous work: regional model

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



colors indicate regions

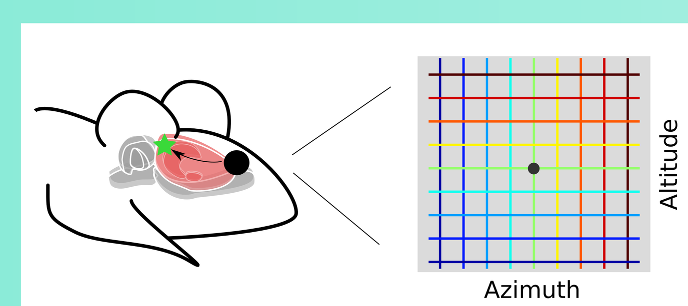
Integrate source and target expression over regions to produce regional vectors of expression $\mathbf{x}^{(r)}$ and $\mathbf{y}^{(r)}$
(n_r x 1) (n_r x 1)

Then fit the (n_r x n_r) weight matrix $W^{(r)}$ via least squares by solving

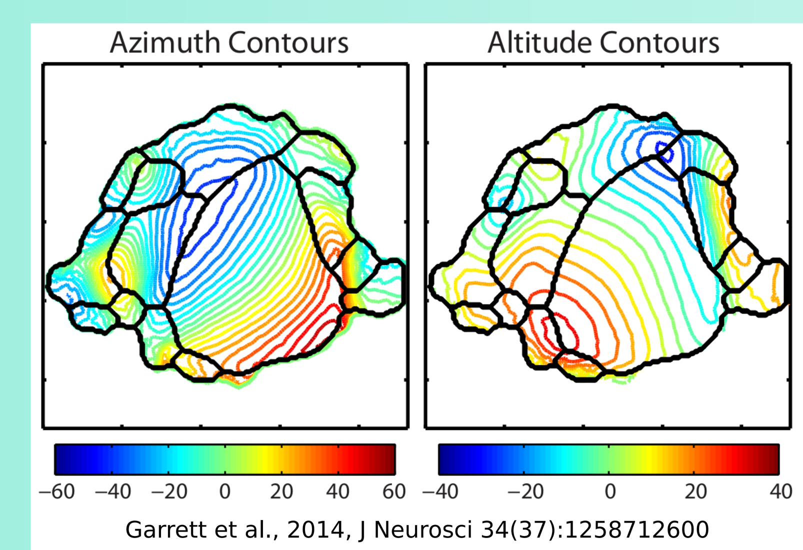
$$\min_W \left\| Y^{(r)} - W^{(r)} X^{(r)} \right\|_F^2$$

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

Evidence for smoothness



What the mouse sees



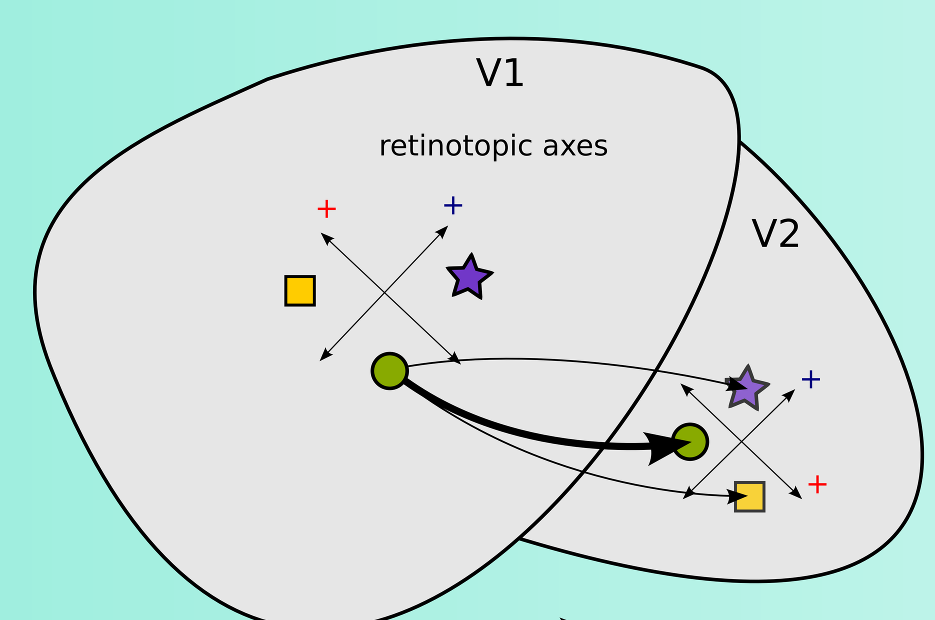
... and it's cortical representation

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.

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

Above items taken together suggest connections will vary smoothly in space



Smoothness regularized model

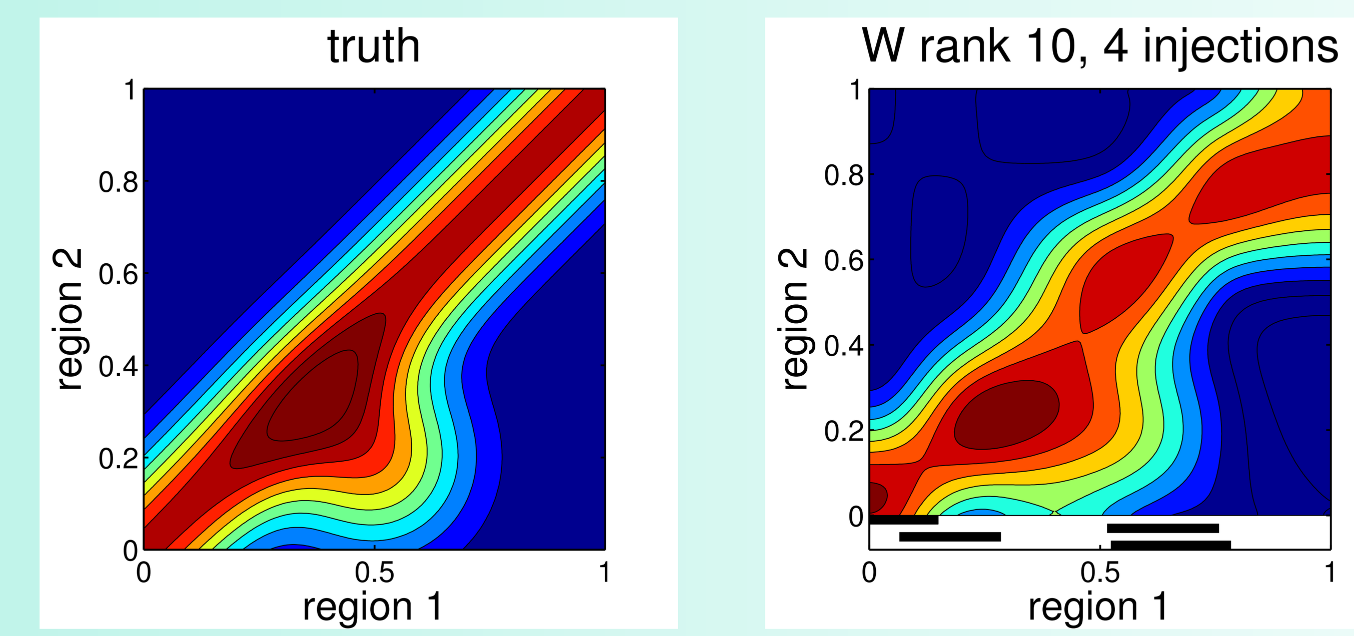
Find the voxel-resolution connection matrix W that balances goodness of fit and smoothness:

$$\min_W \|Y - WX\|_F^2 + \lambda \left\| W L_x^T + L_y W \right\|_F^2 \quad *$$

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

Laplacian matrices (smoothness terms), Neumann boundary conditions

The choice of a Laplacian penalty in this regression is analagous to so-called "thin-plate splines" for curve fitting or interpolation



Reconstruction of a smooth matrix in a test problem, λ=2

Method must be scalable to work with O(10^5) voxels in dataset

Result: With n_{inj} small (relative to n_x, n_y) a low rank (≈3 x n_{inj}) solution $W=UV^T$ works well in our 1-dimensional test problem

Require: $U_0 \in \mathbb{R}^{n_y \times r}$, $\epsilon > 0$ (and X, Y, L_x, L_y, λ to evaluate J)
Ensure: $W = USV^T$ rank r solution to the regularized problem
$k \leftarrow 0$
repeat
$k \leftarrow k+1$
$\tilde{V}_k = \operatorname{argmin}_V J(U_{k-1}, V)$
$[V_k, S_k] = \operatorname{svd}(\tilde{V}_k)$
$\tilde{U}_k = \operatorname{argmin}_U J(U, V_k)$
$U_k = \operatorname{orth}(\tilde{U}_k)$
until $\|U_k - U_{k-1}\|_F < \epsilon$ **and** $\|V_k - V_{k-1}\|_F < \epsilon$
return $W = U_k S_k V_k^T$

Solve a generalized Sylvester equation of the form AXB + CXD + XE + FX = G for X (J is cost function *)

Algorithm: Alternating least-squares low-rank solution

Challenges and future work

- Algorithmic bottleneck is Sylvester equation solve
- Regularization parameter & goodness of fit w/ cross-validation
- Does this work equally well in higher dimensions?
- How much injection coverage is needed?
- Hope to recover **retinotopy correlation from connectivity**
- Fit entire mouse visual system... fit entire mouse brain