

<sup>1</sup> RESEARCH

<sup>2</sup> **Modelling the cell-type specific murine connectome**

<sup>3</sup> **Samson Koelle**<sup>1,2</sup>, **Jennifer Whitesell**<sup>1</sup>, **Karla Hirokawa**<sup>1</sup>, **Hongkui Zeng**<sup>1</sup>, **Marina Meila**<sup>2</sup>, **Julie Harris**<sup>1</sup>, **Stefan Mihalas**<sup>1</sup>

<sup>4</sup> <sup>1</sup>Allen Institute for Brain Science, Seattle, WA, USA

<sup>5</sup> <sup>2</sup>Department of Statistics, University of Washington, Seattle, WA, USA

<sup>6</sup> **Keywords:** [a series of capitalized words, separated with commas]

**ABSTRACT**

<sup>7</sup> The Allen Brain Connectivity Atlas consists of thousands of labelling experiments targeting  
<sup>8</sup> interrogating diverse structures and classes of projecting neurons. This paper describes the  
<sup>9</sup> conversion of these experiments into class-specific connectivity matrices representing the connection  
<sup>10</sup> between source and target structures. We introduce and validate a novel statistical model for creation  
<sup>11</sup> of connectivity matrices that combines spatial and categorical smoothing to share information  
<sup>12</sup> between similar neuron classes. We then illustrate overall and cell-type specific connectivity patterns  
<sup>13</sup> in the resultant connectivities.

**AUTHOR SUMMARY**

## INTRODUCTION

<sup>14</sup> The animal nervous system enables an extraordinary range of natural behaviors, and has inspired  
<sup>15</sup> much of modern artificial intelligence. Neural connectivities - axon-dendrite connections from one  
<sup>16</sup> region to another - form the architecture underlying this capability. These connectivities vary by  
<sup>17</sup> neuron type, as well as axonic source and dendritic target structure. Thus, characterization of the  
<sup>18</sup> relationship between neuron type and source and target structure is an important step to  
<sup>19</sup> understanding the nervous system.

<sup>20</sup> Viral tracing experiments - in which a viral vector expressing GFP is transduced into neural cells  
<sup>21</sup> through stereotaxic injection - are a useful tool for understanding these connections on the mesoscale  
<sup>22</sup> (???). The GFP protein moves from axon to dendrite through the process of anterograde projection, so  
<sup>23</sup> neurons 'downstream' of the injection site will also fluoresce. Two-photon tomography imaging can  
<sup>24</sup> then determine the location and strength of the fluorescent signals in two-dimensional slices. These  
<sup>25</sup> locations can then be mapped back into three-dimensional space, and the signal is partitioned into  
<sup>26</sup> the transduced source and merely transfected target regions.

<sup>27</sup> The conversion of such experiment-specific signals into an overall estimate of the connectivity  
<sup>28</sup> strength of two regions is accomplished by a statistical model. ? and ? describe two such methods.  
<sup>29</sup> Intuitively, both of these models provide some improvement over simply averaging the projection  
<sup>30</sup> signals of injections in a given region. is another. These models are evaluated based off of their ability  
<sup>31</sup> to predict held-out experiments in leave-one-out cross validation. A model that performs well in such  
<sup>32</sup> validation experiments is then assumed to generate the most accurate connectivity.

<sup>33</sup> Both ? and ? develop models for mostly wild-type mice using a standardized vector over all  
<sup>34</sup> experiments. However, recent work (?) has extended these datasets to include viral tracing  
<sup>35</sup> experiments inducing cell-type specific fluorescence. This is accomplished by injecting vectors with  
<sup>36</sup> Cre-recombinase triggered GFP promoters into transgenic mice with cell-type specific  
<sup>37</sup> Cre-recombinase expression Thus, the this paper extends the methodology of ? and ? to deal with the  
<sup>38</sup> diverse set of cre-lines described in ?.

<sup>39</sup> This extension relies on a to our knowledge novel estimator that takes into account both the spatial  
<sup>40</sup> position of the labelled source, as well as the categorical cre-label. This model outperforms the model  
<sup>41</sup> of ?, even for wild-type experiments.

<sup>42</sup> The resulting cell-type specific connectivity matrices form a multi-way *neural connection tensor* of  
<sup>43</sup> information about neural structure. We do not attempt an exhaustive analysis of this data, but do  
<sup>44</sup> demonstrate several basic phenomena. First, we verify several cell-type specific patterns found  
<sup>45</sup> elsewhere in the literature. Second, we discover cell-type specific signals in the neural connection  
<sup>46</sup> tensor. Finally, we decompose the overall (wild-type) connectivity matrix into factors representing  
<sup>47</sup> archetypal connective patterns.

## METHODS

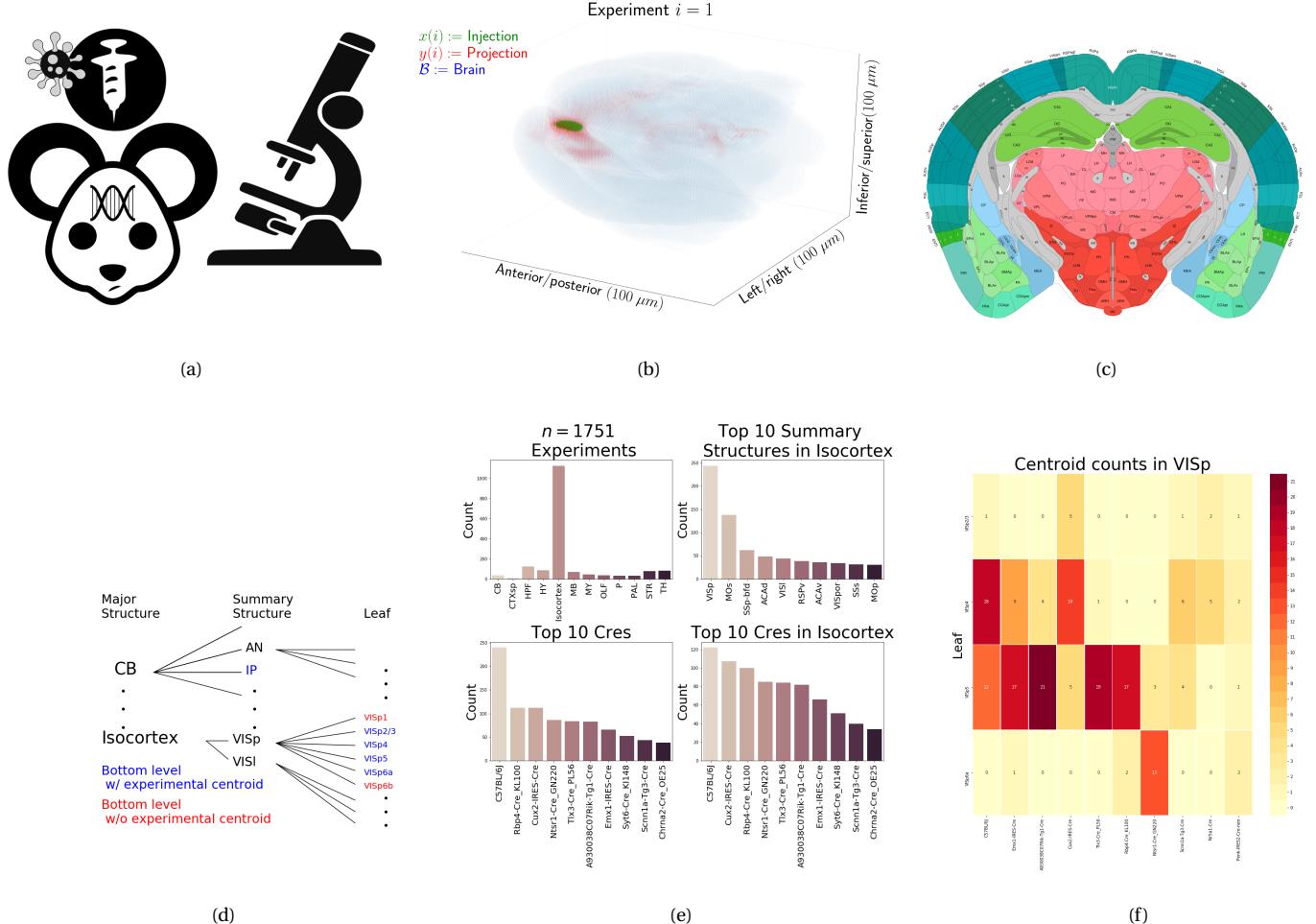


Figure 1: a) Background on histology. a) Within the brain (blue), injection (green) and projection (red) areas are determined via histological analysis and alignment to the Allen Common Coordinate Framework (CCF). b) An example of the segmentation of projection and injection for a single experiment. c) Example of structural segmentation within a horizontal plane. d) Explanation of nested structural ontology highlighting lowest-level and data-relevant structures. e) Abundances of celines and structural injections. f) Co-occurrence of layer-specific centroids and cline within VISp

48 Our main result is the creation of cell-type specific connectivity matrices using a model trained on  
49 murine viral-tracing experiments. This section first describes the data used to generate the model, the  
50 model itself, evaluation of the model, and the use of the model in creation of the connectivity  
51 matrices. The model we ultimately propose is selected based off it's preferable performance in  
52 creation of connectivity matrices with greater accuracy than alternative approaches. We accompany  
53 these matrices with exploratory analyses of the resulting connectivities that illustrate their key  
54 features.

55 ***Mice***

56 (SK's comment:**Experiments involving mice were approved by the Institutional Animal Care and**  
57 **Use Committees of the Allen Institute for Brain Science in accordance with NIH guidelines.**)

58 ***Data***

59 Our dataset  $\mathcal{D}$  consists of  $n = 1751$  experiments from the Allen Mouse Brain Connectivity Atlas. Figure  
60 2 describes the key features of the dataset. Each experiment is performed by injecting a GFP-labelled  
61 transgene cassette with a potentially cre-specific promoter into a particular location in a cre-driver  
62 mouse. The resultant fluorescent signal is imaged, and aligned into the Allen Common Coordinate  
63 Framework (CCF), a three-dimensional idealized model of the brain that is consistent between  
64 animals.

65 Within the dataset generating the connectivity model reported in this paper, certain structure and  
66 cre-line combinations ( $S, V$ ) appear frequently, while others appear not at all. Since users of the  
67 connectivity matrices may be interested in particular combinations, or interested in the amount of  
68 data used to generate a particular connectivity estimate, we exhaustively present this information  
69 about all experiments in Appendix ??.

70 ***Data processing***

71 We discretize the fluorescent intensity photographically determined through histological image  
72 analysis at the  $100 \mu\text{m}$  voxel level. Thus, the fluorescence is represented as a tensor  $\mathcal{F} \in \mathbb{R}^B$  where  
73  $B \subset [1 : 132] \times [1 : 80] \times [1 : 104]$  corresponds to the subset of the voxelized  $(1.32 \times 0.8 \times 1.04)$  cm  
74 rectangular space occupied by the standard mouse brain. This fluorescence is segmented into

*75 injection and projection areas corresponding to areas of transduction and transduction/transfection,*  
*76 respectively. For a given experiment, we denote these as  $x(i)$  and  $y(i)$ , respectively. An example of*  
*77 such segmentation areas is given in Figure ??.* In order to relate the regularly discretized 3D space  $B$   
*78 with biologically informative structures such as the cortex, we also apply several levels of*  
*79 regionalization, as shown in Figure 2.* Mathematically, we refer to the regionalization map as  
*80  $r : \mathcal{B} \times \mathbb{R} \rightarrow \mathcal{R} \times \mathbb{R}$ .* Given a vector  $a$ , we also define a normalization map  $n : a \mapsto \frac{a}{\sum_{t \in T} a_t}$ . A detailed  
*81 mathematical description of these data preprocessing steps is given in Appendix ??.*

*82 **Connectivity***

*83 Our goal is the estimation of structural connectivity from one structure to another. However, at an*  
*84 essential level, cell-class specific neural connectivity is representable as a function*  
*85  $f : \mathcal{V} \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^+$*  giving the connection of a particular cell-class from a source position to a target  
*86 position. As mentioned in the previous section, our injection and projection intensities are*  
*87 discretized into  $100\mu\text{m}$  cubic voxels. These voxels are contained within structures, so mathematically*  
*88 we can write a region  $R = \{r\}$ . We generically denote source regions  $S = \{s\}$  and target regions  $T = \{t\}$ .*

There are several notions of structural connectivity worth considering based on normalization with respect to the sizes of the source and/or target regions. Given a set of source regions  $\mathcal{S} = \{S\}$ , target regions  $\mathcal{T} = \{T\}$ , and cre-lines  $\mathcal{V}$  we shall estimate the following tensors:

*connectivity strength  $\mathcal{C} \in \mathcal{V} \times \mathcal{S} \times \mathcal{T} \times \mathbb{R}_{\geq 0}$  with  $\mathcal{C}(V, S, T) = \sum_{s \in S} \sum_{t \in T} f(v, s, t)$*

*normalized connectivity strength  $\mathcal{C}^S \in \mathcal{V} \times \mathcal{S} \times \mathcal{T} \times \mathbb{R}_{\geq 0}$  with  $\mathcal{C}^S(V, S, T) = \frac{1}{|S|} \sum_{s \in S} \sum_{t \in T} f(v, s, t)$*

*normalized projection density  $\mathcal{C}^D \in \mathcal{V} \times \mathcal{S} \times \mathcal{T} \times \mathbb{R}_{\geq 0}$  with  $\mathcal{C}^D(V, S, T) = \frac{1}{|S||T|} \sum_{s \in S} \sum_{t \in T} f(v, s, t)$ .*

*89 Our goal is to estimate  $\mathcal{C}(V, S, T)$  with data  $\mathcal{D}$ . We call this estimator  $\hat{\mathcal{C}}$ .*

*90 **Modelling connectivity***

Construction of such an estimator raises the important questions of 1) what data to use for estimating which connectivity, 2) how to featurize the dataset, 3) what statistical estimator to use, and 4) how to reconstruct the connectivity using the chosen estimator. Mathematically, we represent these

considerations as

$$\widehat{\mathcal{C}}(V, S, T) = e^*(\widehat{e}(e_*(\mathcal{J}(\mathcal{D}))). \quad (1)$$

This makes explicit the data featurization  $e_*$ , statistical estimator  $\widehat{e}$ , and any potential subsequent transformation  $e^*$  such as averaging over the source region, as well as the fact that different data  $\mathcal{D}$  may be used to estimate different connectivities. For example, a simple model would be to take the mean regionalized projection of all the experiments with injection centroid in a given structure. Table 1 reviews estimators used for this data-type, and explain the intuition behind our new cell-class specific estimator. Additional information is given in Appendix ??

Model	$e^*$	$\widehat{e}$	$e_*$	Training Data
(?)	$\widehat{e}(S)$	NNLS(X,Y)	$X = r(x(I)), Y = r(y(I))$	$I = I_M$
(?)	$\sum_{s \in S} \widehat{e}(s)$	NW(X,Y)	$X = c(x(I)), Y = r(y(I))$	$I = I_M$
Cre-NW	$\sum_{s \in S} \widehat{e}(s)$	NW(X,Y)	$X = c(x(I)), Y = n(r(y(I)))$	$I = I_S \cap I_V$
Expected-loss	$\sum_{s \in S} \widehat{e}(s)$	EL <sub>S</sub> (X, Y, V)	$X = c(x(I)), Y = n(r(y(I))), V = v(I)$	$I = I_S$

Table 1: Estimation of  $\mathcal{C}$  using connectivity data. The regionalization, estimation, and featurization steps are denoted by  $e^*$ ,  $\widehat{e}$ , and  $e_*$ , respectively. The training data used to fit the model is given by  $I$ . We generically denote the set of experiments used to train a particular model as  $I$ , and experiments from particular major brain divisions, summary structures, and leafs as  $I_M$ ,  $I_U$ , and  $I_L$ , respectively.

Our new methodological contributions in this area - the Cre-NW and Expected-loss models - have several differences from the previous methods. Both the ? non-negative least squares and ? Nadaraya-Watson take into account  $s$  and  $t$ , but not  $v$ . Since our goal is creation of cre-specific connectivities, our new estimators specifically account for this information. The cre-specific Nadaraya-Watson estimator only uses experiments from a particular cre-line to predict cell-class connectivity, while the Lxpected Loss estimator shares information between cre-lines. We also normalize projections by total intensity to account for differences in the cre-driven expression of eGFP via the various transgene promoters.

The final and perhaps most important modelling question is how to select optimum functions from within and between these estimator classes. Examining Equation 1, we can see that in 3D coordinates,  $\hat{f}(v, s, t) = \hat{e}(e^*(\mathcal{J}(\mathcal{D})))$ , with  $e^*$  a deterministic step that is determined without input by the data. Thus, we can evaluate our model by its ability to predict held-out experiments in leave-one-out cross validation. In order to compare between methods, we necessarily restrict to the smallest set of evaluation experiments suggested by any of our models - that is, those with combinations of cre-line and structure-leaf that are present at least twice. We use  $l_2$ -loss and weighted  $l_2$ -loss to evaluate these predictions:

$$\text{l2-loss } l(\hat{f}) = \frac{1}{|I_M|} \sum_{i \in I_M} \|r(y(i)) - \hat{f}(c(i))\|_2^2$$

$$\text{weighted l2-loss } l(\hat{f}) = \frac{1}{|\{S, V\}|} \sum_{s, v \in \{S, V\}} \frac{1}{|I_{s, v}|} \sum_{i \in I_{s, v}} l(r(y(i)), \hat{f}(\mathbb{D} \setminus i))$$

- <sup>105</sup> Due to our normalizing preprocessing step, the former is equivalent to the normalized loss in ?, and  
<sup>106</sup> also to cosine loss, while the latter weights all region cell-class combinations equally.

<sup>107</sup> As a final modelling step, we establish a lower limit of detection. This is covered in Appendix.

<sup>108</sup> **Connectivity analyses**

- <sup>109</sup> We quantify and illustrate some of the interesting neuronal processes underlying our estimated  
<sup>110</sup> connectome. First, we cluster projection pattern by cell-class and source structure. This shows that  
<sup>111</sup> cell-class has a dominating effect on projection in certain regions. Second, we extend the  
<sup>112</sup> characterization of ? on structural differences in short-range projections. These are primarily  
<sup>113</sup> assumed to be due to diffusion, and the diffusion-rate helps to characterize the basic structural  
<sup>114</sup> anatomy. Third, since the overall wild-type connectome results from the combination of underlying  
<sup>115</sup> cell-classes, we apply non-negative matrix factorization (NMF) to decompose the observed  
<sup>116</sup> long-range connectivity into *connectivity archetypes* that linearly combine to reproduce the observed  
<sup>117</sup> connectivity. These methods identify structures with both known and plausible biological meaning,  
<sup>118</sup> and simplistically exemplify useful posthoc analyses for data of this type. Technical details of these  
<sup>119</sup> approaches are given in Appendix.

## RESULTS

120 Our results include evaluation of model fit, the cre-specific connectivity matrices themselves, and  
121 retrospective analyses of these matrices for patterns related to cre-type and source and target regions.

122 ***Model evaluation***

123 Table contains the sizes of these evaluation sets in each major structure. This information may be  
124 cross-referenced visually with the figures in Our two-stage model generally performs better than the  
125 cre-line specific NW estimator.

	Estimator	EL	NW	Average	NW	NW-wt
Smoothing		SS	Cre-SS	Cre-SS	SS	M
Target		SS	SS	SS	SS	SS
Structure	# Eval exps					
CB	10	0.044	0.081	0.081	0.058	0.439
CTXsp	2	0.497	0.497	0.497	0.497	0.000
HPF	79	0.122	0.140	0.143	0.155	0.471
HY	41	0.241	0.266	0.269	0.244	1.019
Isocortex	838	0.173	0.195	0.202	0.234	0.404
MB	23	0.151	0.151	0.166	0.139	0.759
MY	7	0.186	0.233	0.233	0.184	0.452
OLF	17	0.069	0.095	0.100	0.073	0.110
P	8	0.236	0.239	0.239	0.264	0.984
PAL	11	0.190	0.198	0.198	0.260	1.401
STR	45	0.084	0.088	0.089	0.097	0.265
TH	29	0.351	0.678	0.678	0.365	1.088

Table 2: Weighted losses with summary structure targets.

126 **Connectivities**

127 Our main result is the estimation of matrices  $\hat{\mathcal{C}}_v$ , representing connections of source structures to  
128 target structures for particular cre-lines  $v$ . We exhibit several characteristics of interest, and confirm  
129 the detection of several well-established connectivities within our tensor. Many additional interesting  
130 biological processes are visible within this matrix - more than we can report in this paper - and it is  
131 our expectation that these will be identified by users of our results. The connectivity tensor and code  
132 to reproduce it are available at ([SK's comment:footnote](#)).

133 The connectivity matrix for wild-type connectivities from leaf sources to summary structure targets  
134 is illustrated in Figure ???. The clear intraareal connectivities mirror previous estimates in ? and ? and  
135 descriptive depictions of individual experiments in ?. Compared with ?, our more discretized source  
136 smoothing and greater number of experiments leads to a significantly more discretized connectivity  
137 matrix. This is generally expected - for example, different cortical layers have more substantially  
138 different connectivities.

139 The cell-type specific connectivities that we provide also conform to well-known behaviors.  
140 Examples from the visual processing and motor control regions of the cortex are given in Figure ?? for  
141 both wild type and several cre-lines. Rbp4-Cre and Ntsr1-Cre target layers 5 and 6, respectively. As in  
142 ?, layer 5 projects to anterior basolateral amygdala (BLA) and capsular central amygdala (CEA), while  
143 layer 6 does not.



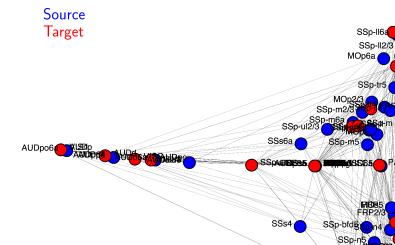
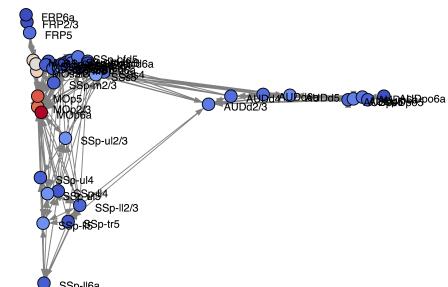
(a)

	# Ipsilateral Leaf Targets	Top Entropy	Bottom Sparsity	Bottom Entropy	Top Sparsity
Isocortex	51	CP	BAC	BAC	ENT1
OLF	11	TMv	III	III	NaN
HPF	15	IG	EPv	PA	NaN
CTXsp	7	TT	FC	APr	TT
STR	14	RPA	ISN	PVR	TU
PAL	9	PG	ACVII	GR	MG
TH	44	NOD	DN	SSp-ll	SCm
HY	44	CLA	SH	LSc	DG
MB	39	NDB	SubG	SGN	SUB
P	26	MT	Acs5	SOC	NDB
MY	43	RT	NaN	OV	EPd
CB	18	ECT	AOB	MOB	GU

(b)

	# Ipsilateral Leaf Targets	Top Entropy	Bottom Sparsity	Bottom Entropy	Top Sparsity
Isocortex	51	CP	BAC	BAC	ENT1
OLF	11	TMv	III	III	NaN
HPF	15	IG	EPv	PA	NaN
CTXsp	7	TT	FC	APr	TT
STR	14	RPA	ISN	PYR	TU
PAL	9	PG	ACVII	GR	MG
TH	44	NOD	DN	SSp-ll	SCm
HY	44	CLA	SH	LSc	DG
MB	39	NDB	SubG	SGN	SUB
P	26	MT	Acs5	SOC	NDB
MY	43	RT	NaN	OV	EPd
CB	18	ECT	AOB	MOB	GU

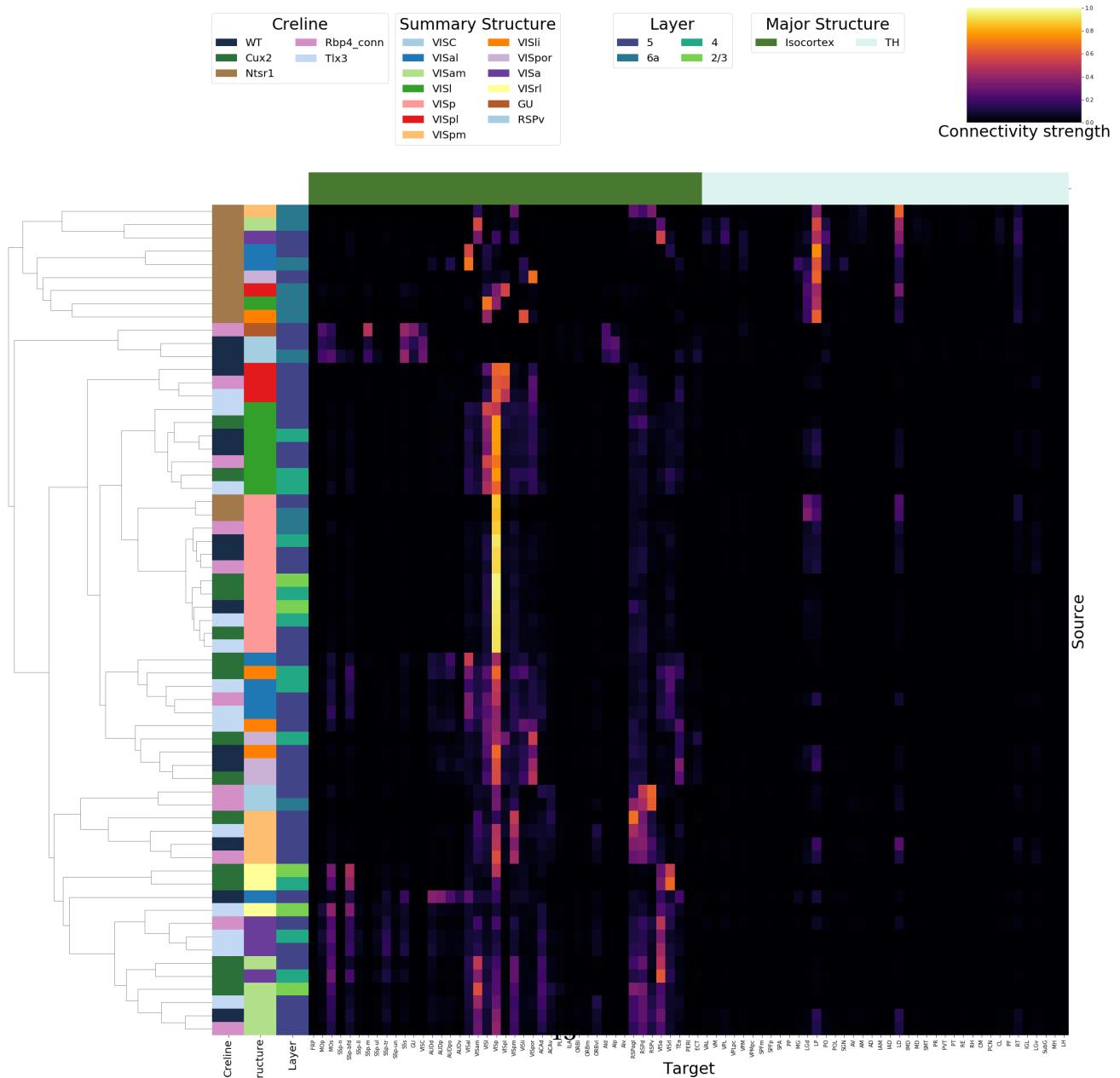
(d)

Source  
Target

(e)



(a)



144 **Connectivity Analyses**

145 The connectivity matrix represents a collection of relatively few biological processes. For example,  
146 certain cell-types and layers have a characteristic connectivity pattern, and structures tend to connect  
147 most strongly to the most proximal areas. We elucidate these patterns through two types of analyses.  
148 First, we demonstrate cell-type specific connectivity patterns by hierarchical clustering of  
149 connectivities from multiple cre-lines, and showing that cre-line is a key factor driving the observed  
150 behavior. Then, we perform a different unsupervised analysis - non-negative matrix factorization - of  
151 distal wild-type connectivities, to estimate underlying overall connectivity patterns.

152 Figure 3 shows a collection of connectivity strengths generated using cre-specific models for  
153 wild-type, Cux2, Ntsr1, Rbp4, and Tlx3 cre-lines from visual signal processing leafs in the cortex to  
154 cortical and thalamic nucleii. Heirarchical clustering is applied to sort the different source/cre  
155 combinations by the similarity of their connectivities to summary-structure targets. This analysis  
156 shows that Ntsr1 cre-lines tend to target thalamic nucleii, in particular LP and LD ?. However, with  
157 this exception, for the other plotted cre-lines, connectivity tends to cluster by source structure. That  
158 the tendency for structures to connect to themselves is quite strong emphasizes the special nature of  
159 the Ntsr1-Thalamic connection in this analysis.

160 The overall wild-type connectivity strength matrix also displays an underlying modellable  
161 structure. As discussed in ?, one of the most basic processes underlying the observed connectivity is  
162 the tendency of each source region to predominantly project to proximal regions. The heatmap in ??a)  
163 shows intraregion distances clearly contains an overall pattern reminiscent of the connectivity matrix  
164 in ???. This relationship is plotted in ?? b), showing that there exists substantial variability that would  
165 be impossible to model with low-error in a univariate model, even using the diffusion model  
166 suggested in ?. These connections are biologically meaningful, but also unsurprising, and their  
167 relative strength biases learned latent coordinate representations away from long-range structures.  
168 For this reason, we establish a  $1500\mu m$  'distal' threshold within which to exclude connections for our  
169 analysis. We then apply non-negative matrix factorization (NMF) to decompose the remaining  
170 censored matrix into a relatively small number of distinct projection signals, and apply an

171 unsupervised cross-validation method to select the optimum number of signals ([SK's](#)

172 [comment:Percent error... show reconstruction? log scale?](#)).

## DISCUSSION

173 Flattening  $\mathcal{C}$  prior to unsupervised analysis is not necessarily recommended, but provides an easy  
174 solution for this problem.

175 With respect to the model, a Wasserstein-based measure of injection similarity per structure would  
176 combine both the physical simplicity of the centroid model while also incorporating structural  
177 knowledge.

178 The Nadaraya-Watson weighting procedure introduced here is, to our knowledge, novel. In  
179 particular, our method of utilizing the expected loss to weight points differs from the minimization  
180 task of fitting data to weighted sums of neighbors (?). We make a key assumption: that the additional  
181 statistical accuracy of including more samples makes up for the fact that their expected accuracy is  
182 lower. Note that this assumption can be easily violated, if, for example, the data is distributed on a  
183 circle without error, and only nearest neighbors are most predictive.

184 Model averaging based off of cross-validation has been implemented in ?, but we note that our  
185 approach makes use of a non-parametric estimator, rather than an optimization method for selecting  
186 the weights. ([SK's comment:CITE METHOD THAT SELECTS WEIGHTS IN KERNEL \(has catchy  
name\)](#))  
187

## ACKNOWLEDGMENTS

188 The Funder and award ID information you input at submission will be introduced by the publisher  
189 under a Funding Information head during production. Please use this space for any additional  
190 acknowledgements and verbiage required by your funders.

## SUPPORTING INFORMATION

191 This is an optional section. Please use this space to provide information about any supporting  
192 information referred to in your manuscript.

## COMPETING INTERESTS

<sup>193</sup> This is an optional section. If you declared a conflict of interest when you submitted your manuscript,  
<sup>194</sup> please use this space to provide details about this conflict.

## TECHNICAL TERMS

<sup>195</sup> All NETN article types require Technical Terms.

<sup>196</sup> Identify approximately 10 key terms that are mentioned in your article and whose usage and  
<sup>197</sup> definition may not be familiar across the broad readership of the journal. Provide brief (20-word or  
<sup>198</sup> less) definitions for each term, avoiding in these definitions the use of jargon, or highly technical or  
<sup>199</sup> specialized language. When the article is typeset, the Technical Terms will appear in the margins at or  
<sup>200</sup> near their first mention in the text.

<sup>201</sup> In your manuscript, bold the first occurrence of each **Technical Term** and then provide a list of the  
<sup>202</sup> terms and their definitions at the end of the manuscript after the references.

<sup>203</sup> **Technical Term** a key term that is mentioned in an NETN article and whose usage and definition  
<sup>204</sup> may not be familiar across the broad readership of the journal.

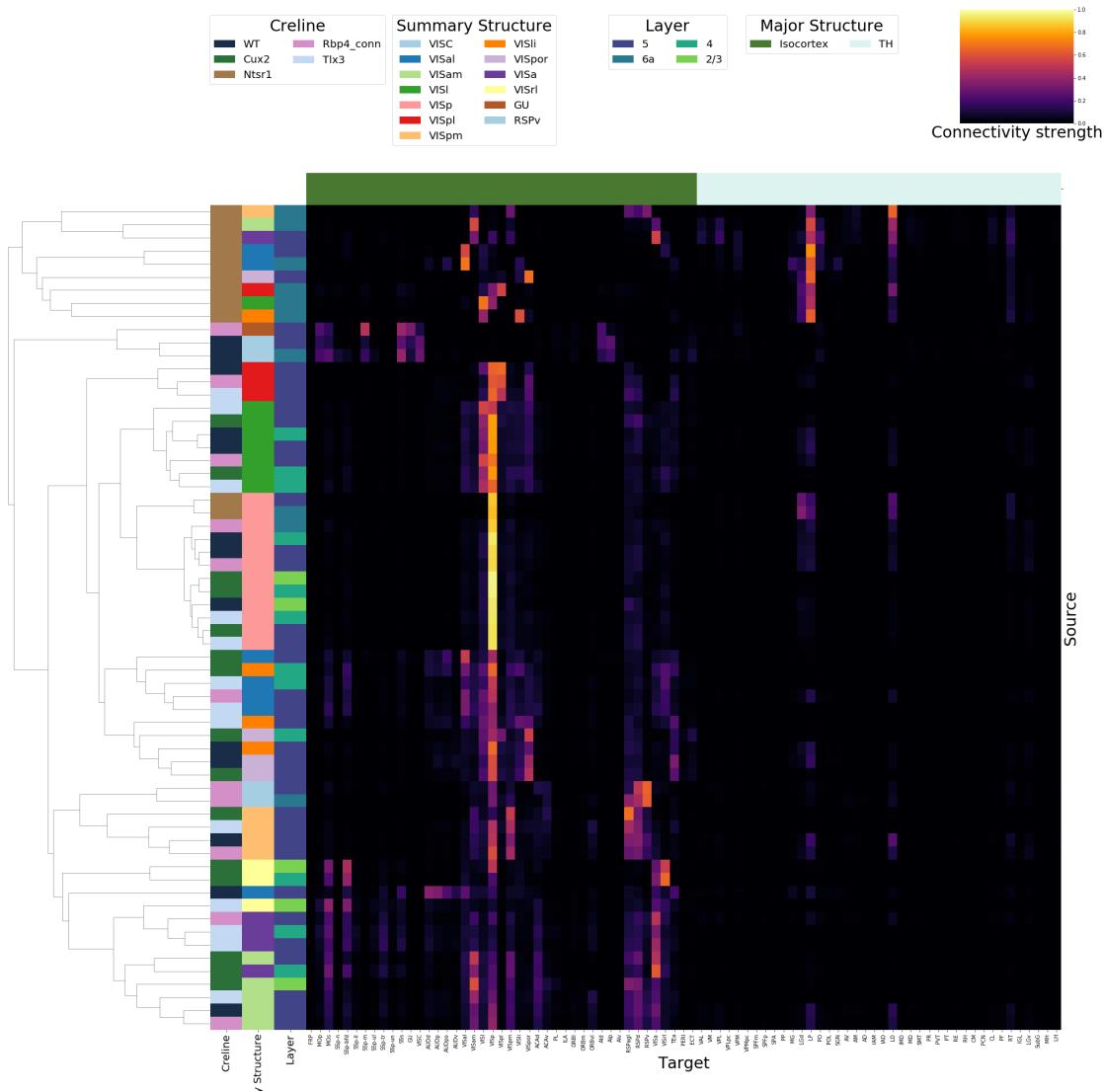
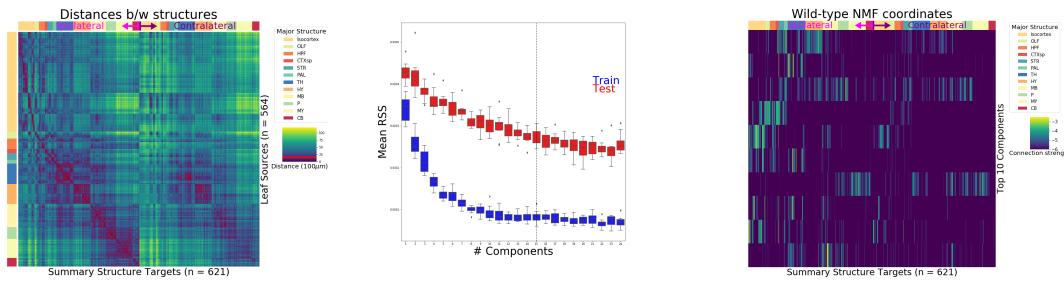


Figure 3: Heirarchical clustering of connectivity strengths from visual signal processing cell-types to cortical and thalymic targets. Cre-line, summary structure, and layer are labelled on the sources. Note that sources/cre combinations are only included if there is at least one experiment of that cre-line in that particular leaf.



(a)

(b)

(c)

