Characterizing Hierarchical Computation in Primary Visual Cortex (V1)

Daniel A Butts[1,2], Felix Bartsch[1], Matthew R Whiteway[2], Bruce G. Cumming[3]

(1) Program in Neuroscience and Cognitive Science and Dept of Biology, (2) Applied Mathematics and Scientific Computation Program, University of Maryland, College Park, MD USA
(3) Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, MD USA

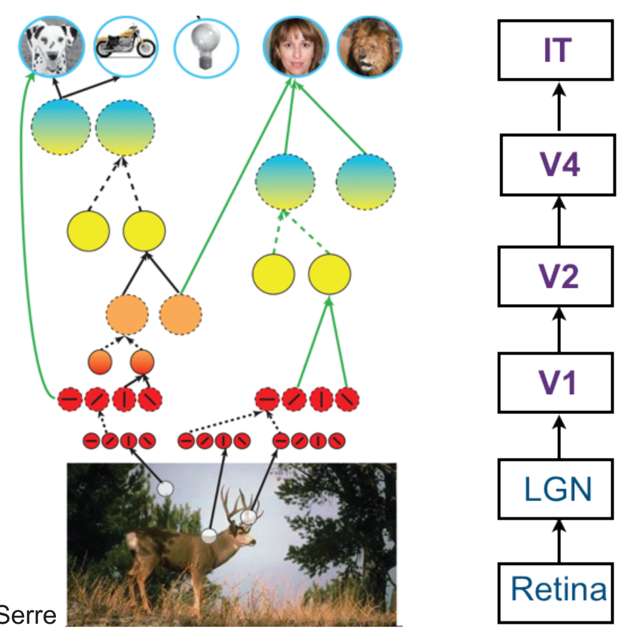
Introduction

The ability to perform complex visual tasks such as object recognition is thought to require many successive stages of nonlinear computations, a hypothesis supported by the hierarchical structure of visual areas within the primate cortex. This assumption underlies the modern machine learning approach of deep neural networks (DNNs), which can be trained to perform such visual tasks with similar performance to humans. However, while such conceptions of visual system function serve as a useful qualitative description, it is unclear how to validate such descriptions using neural data, and – additionally – how to understand neural activity in the context of hierarchical computation.

1. What is the role of individual neurons in system-wide computation?

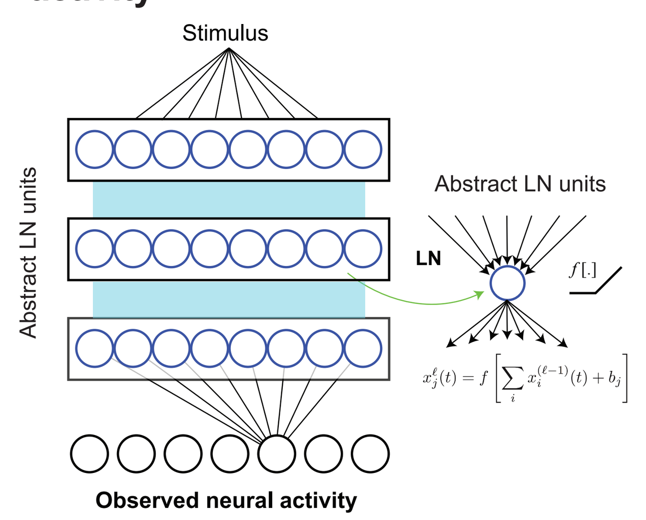
Classical viewpoint: feature detection
Machine learning: need to address on systems-level

2. How can neural activity be interpreted in the context of such a model?

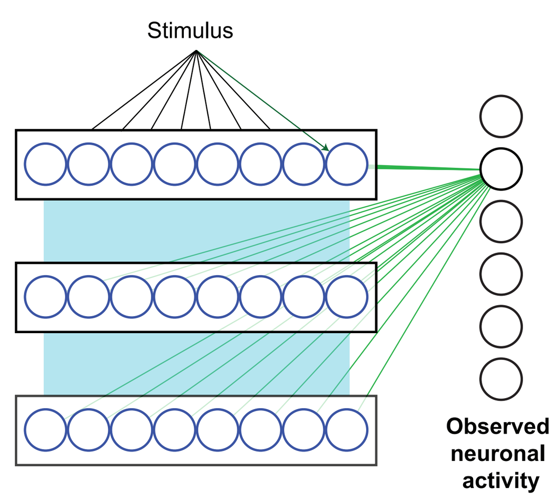


Approach

Part 1: Machine learning to predict observed neural activity



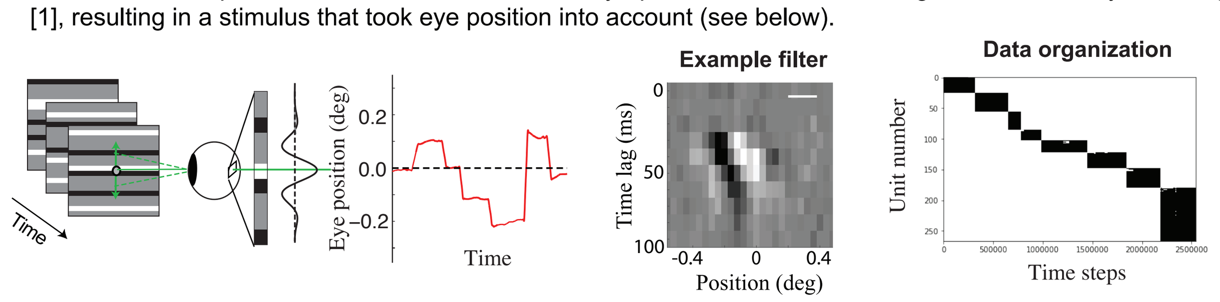
Part 2: "Computational scaffold" to infer neural function within hierarchal computation



Methods

Electrophysiology: Recordings were made in primary visual cortex (V1) of awake two macaques using 24-electrode linear arrays (50 μm spacing), and (in one case) a Utah array, as previously described [1]. Animals performed a simple fixation task to obtain a liquid reward upon completion of each 4-second trial. Here we pooled data from eight experiments, resulting in 48 well isolated single units (SUs) and 219 multiunits (MUs).

Stimuli: Uncorrelated random bar patterns ('1D ternary noise'), aligned as close as possible to the preferred orientation, were presented at 100 Hz refresh rate. Eye-position was tracked using "model-based eye-tracking" [1], resulting in a stimulus that took eye position into account (see below).



Machine learning: Convolutional neural networks (CNNs) of various configurations were fit using custom software using Google's TensorFlow package. Parameters were fit using stochastic gradient descent (the "adam" optimizer to maximize the regularization-penalized population Poisson log-likelihood (per spike), given by:

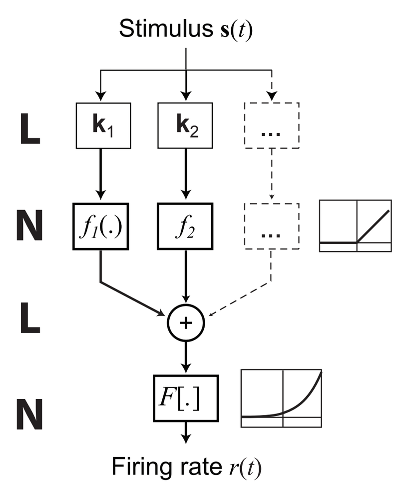
$$LL^*_{pop} = \sum_i \frac{1}{N^{(i)}_{spk}} \sum_t d_i(t) \left[R^{(i)}_{obs}(t) \log_2 r_i(t) - r_i(t) \right] - \text{(regularization penalties)}$$

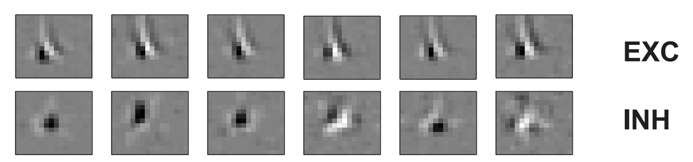
where $r_i(t)$ is the model predicted firing rate, and $d_i(t)$=1 for all time points where there is recorded data for neuron i, and zero otherwise. The network itself was build from LN units with a rectified linear (relu) nonlinearity. All weights were constrained to be positive, but 1/4 of units in each level were made to be "inhibitory" by multiplying their output by -1.

I. Machine learning to predict neural activity

Nonlinear models of [single] V1 neurons

LNLN cascade models like the Nonlinear Input Model (NIM) [2] and others [3,4] can identify many stimulus features that a given neuron is sensitive to.





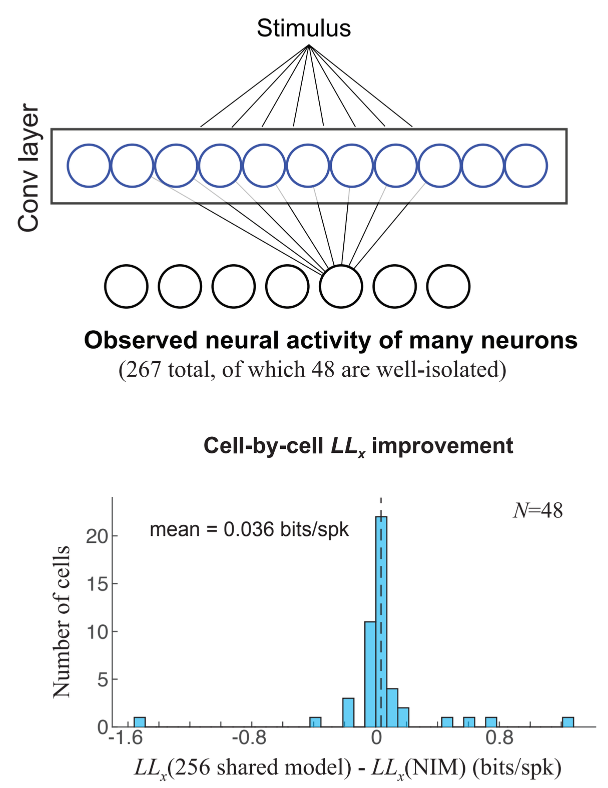
=> Defines a "feature space" combined with a how features are nonlinearly combined.

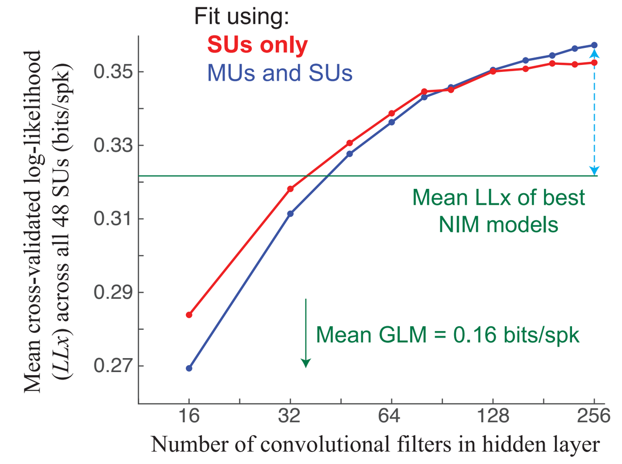
LNLN cascades can in principle represent any high-dimensional nonlinear function $r=f(s)$ with enough LN subunits, but...

... in practice, number of subunits that can be fit is data-limited.

Population fitting via shared computation

Machine learning frameworks to fit neural data. See related work in V1 [5-7] and retina [8].



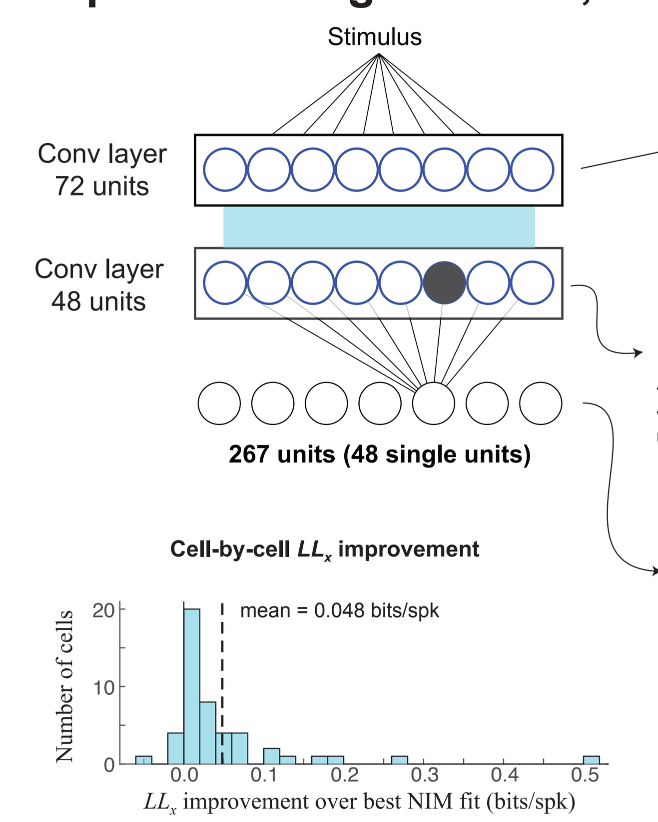


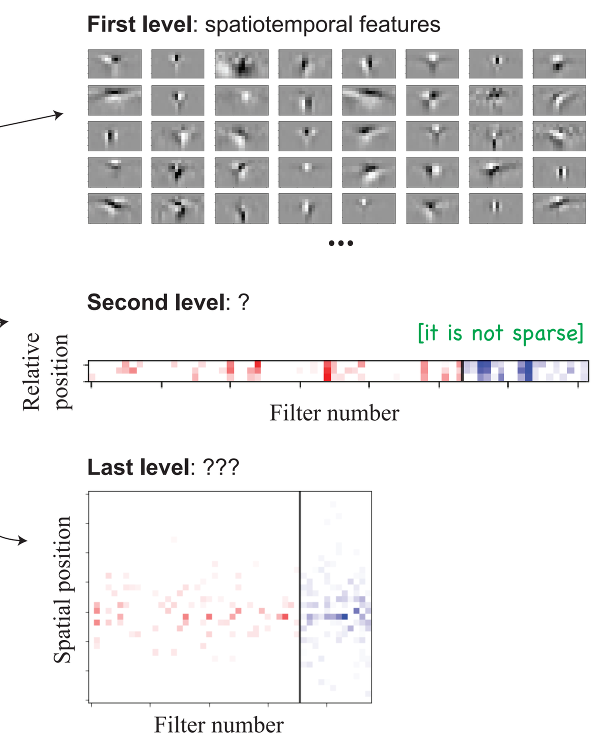
=> LNLN models of V1 are data-limited

=> adding more recorded units can improve all fits via shared computation

Methods to incorporate large amounts of recording into a single population model

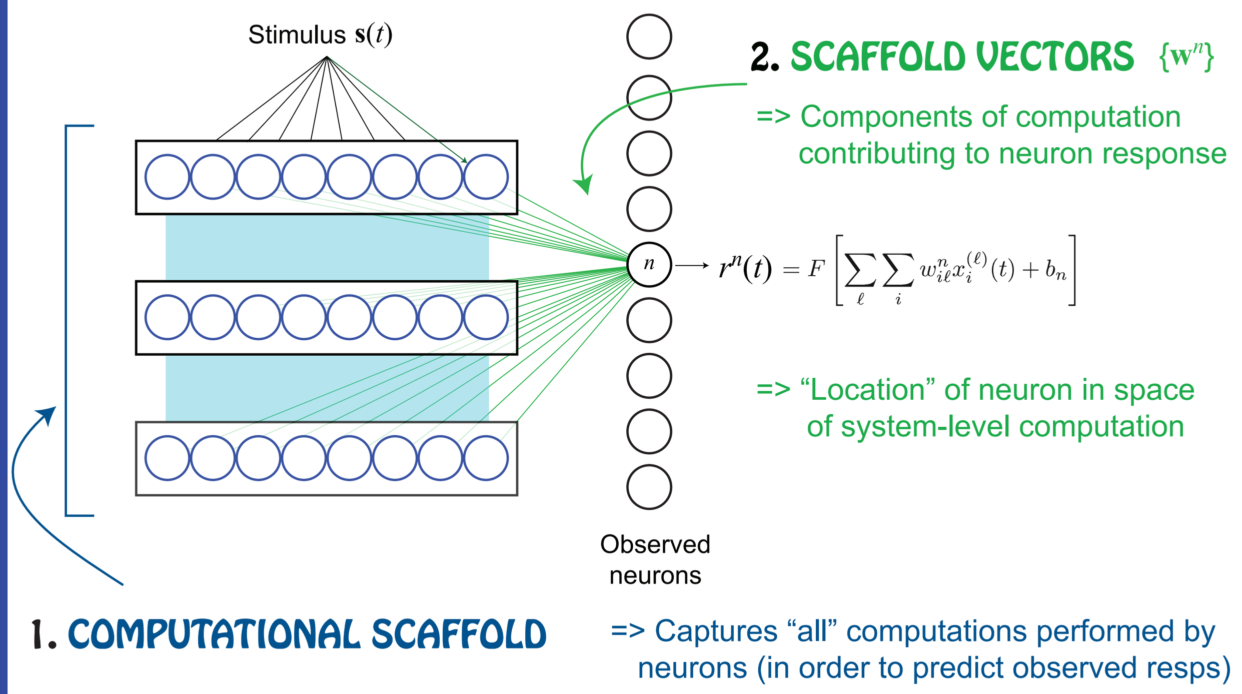
Deeper models give better, less interpretable, predictions





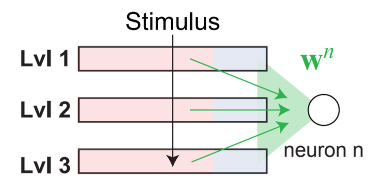
What does this neuron respond to?

II. Computational Scaffold Network



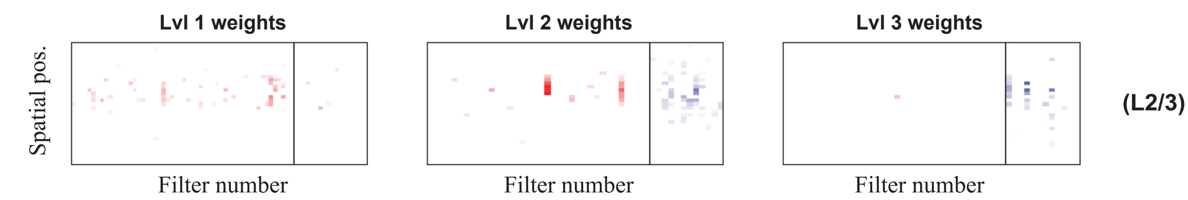
**will just approximate computation, and topology of scaffold vectors should be preserved with different scaffold structures and nonlinearities (can show)

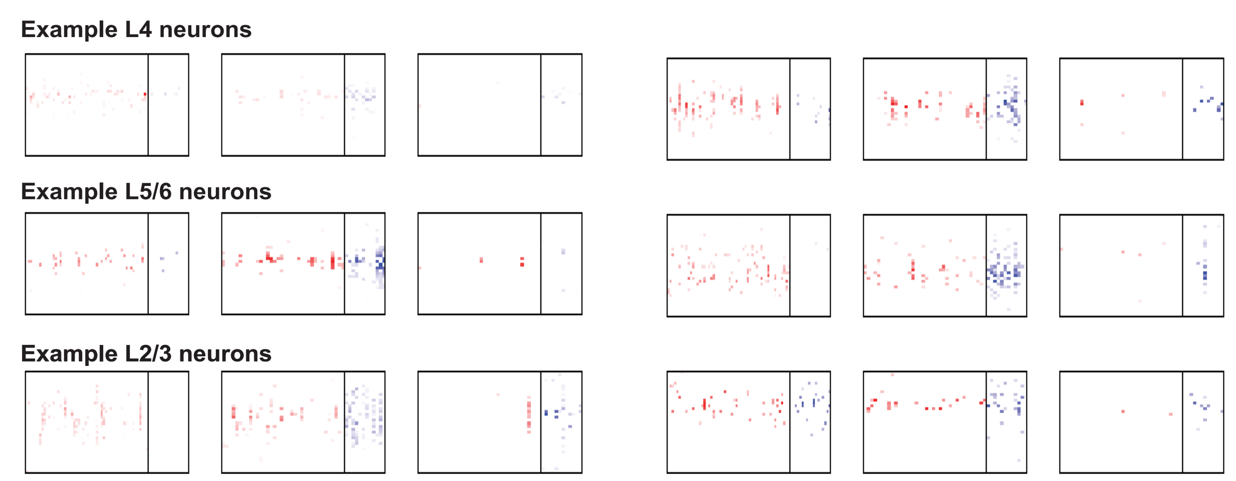
The "structure" of V1 neuron computation



Excitation and Inhibition

- Each layer is composed of 3/4 excitatory and 1/4 inhibitory units
- All weights (other than to the stimulus) have pos. constraints
- Yields equivalent fits, but gives more structural information

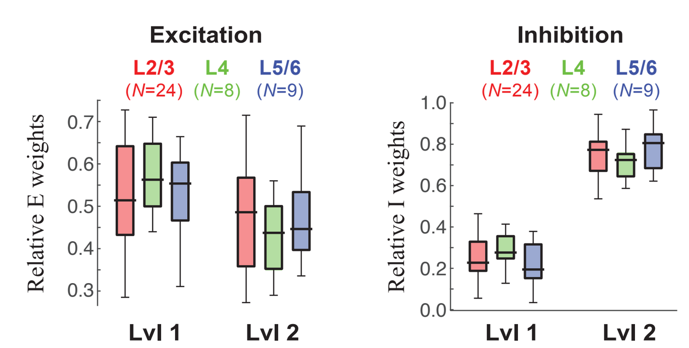




=> Computation of most V1 neurons is spread across levels of scaffold

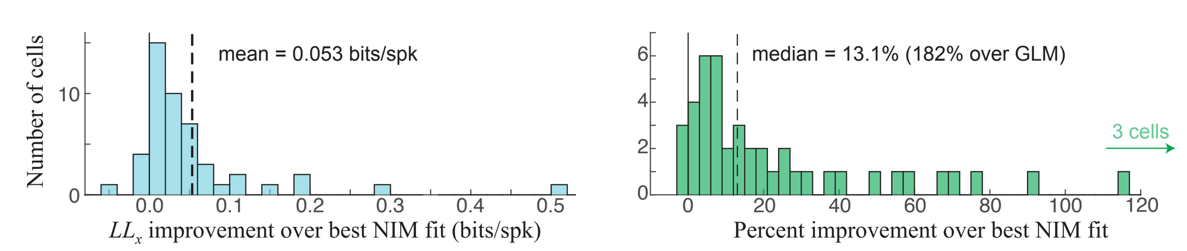
Trends across layers

The expectation was that layer-4 neurons would connect preferentially to earlier levels than other V1 neurons. There is only a weak trend (with limited data), which is best seen with a two-layer scaffold.



Scaffold network maintains predictive power

(and further improves)



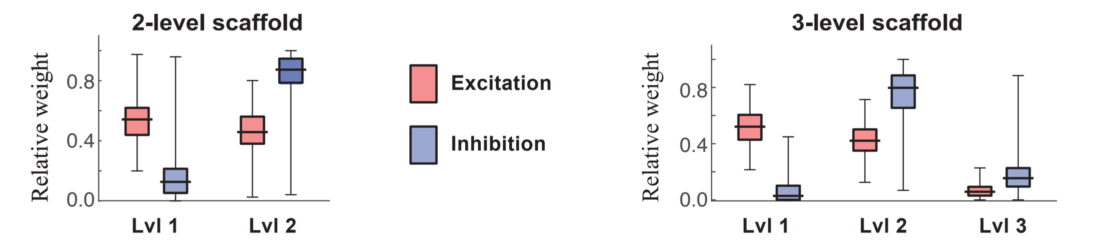
=> Substantial improvement over the best V1 models (and just the beginning)

Conclusions

1. Machine learning approaches offer the opportunity to capture nonlinear computations performed by V1 neurons significantly better than current models, but what they produce is difficult to interpret on its own.
2. The **computational scaffold** network offers a new conception of neural function in the context of hierarchical computation, as an alternative to descriptions based on feature detection.
3. The scaffold network reveals complex structure of V1 neuron computation across nearly all neurons. Such models significantly outperform less complex models.
4. Putative inhibitory inputs are derived from deeper levels of the scaffold, suggesting it is more computationally complex (and likely not easily captured by simpler models.
5. The model predicts a plausible array of size-tuning, which derives from inhibition in deeper layers.

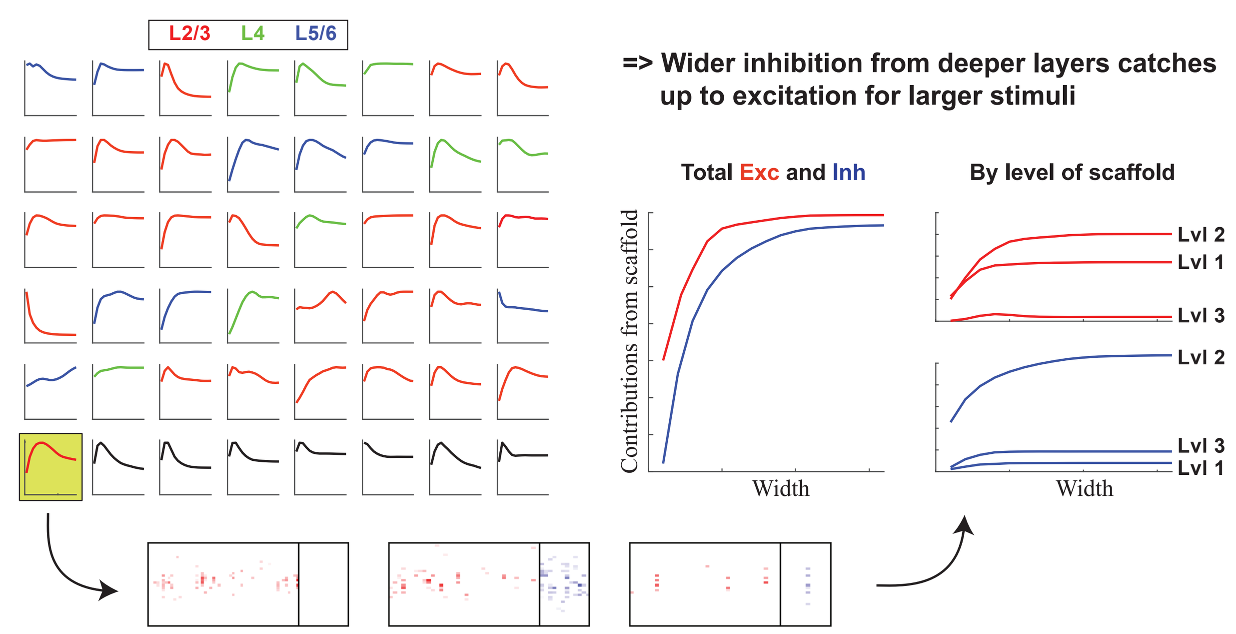
Properties of [putative] inhibition

Inhibitory connections (from the scaffold) to V1 neurons robustly derives from deeper levels, across different scaffold configurations.



Size tuning

Deep inhibition should have a wider spread, suggesting it could be result in size-tuning. Although this was not experimentally tested, we generated model responses to see if the fit models would predict size tuning, using random bar stimulation apertured at different widths, relative to each cell's RF location.



References

http://neurotheory.umd.edu

[1] McFarland JM, Bondy AG, Cumming BG, Butts DA (2014) Nat Commun 5:4605.
[2] McFarland JM, Cui Y, Butts DA (2013) PLoS Comput Biol 9:e1003143.
[3] Park M, Pillow JW (2013) Adv Neural Info Proc Sys (NIPS).
[4] Vintch B, Movshon JA, Simoncelli EP (2015) J Neurosci 35:14829–14841.
[5] Antolik J, Hofer SB, Bednar JA, Mrsic-Flogel TD (2016) PLoS Comput Biol 12:e1004927.
[4] Klindt DA, Ecker AS, Euler T, Bethge M (2017) Adv Neural Info Proc Sys (NIPS): 3509–19.
[6] Kindel WF, Christensen ED, Zylberberg J (2017) arXiv:1706.06208[q–bio.NC].
[7] Ikezoe K, Amano M, Nishimoto S, Fujita I (2018) NeuroImage 180:312–323.
[8] Maheswaranathan N, McIntosh LT, Kastner DB, ..., Ganguli S, Baccus SA (2018) bioRxiv:1–14.

Acknowledgements

This work is supported by the National Science Foundataion IIS-1350990 (DAB), and the National Institutes of Health EY025403 (FB, DAB) and the NEI intramural program (BGC).