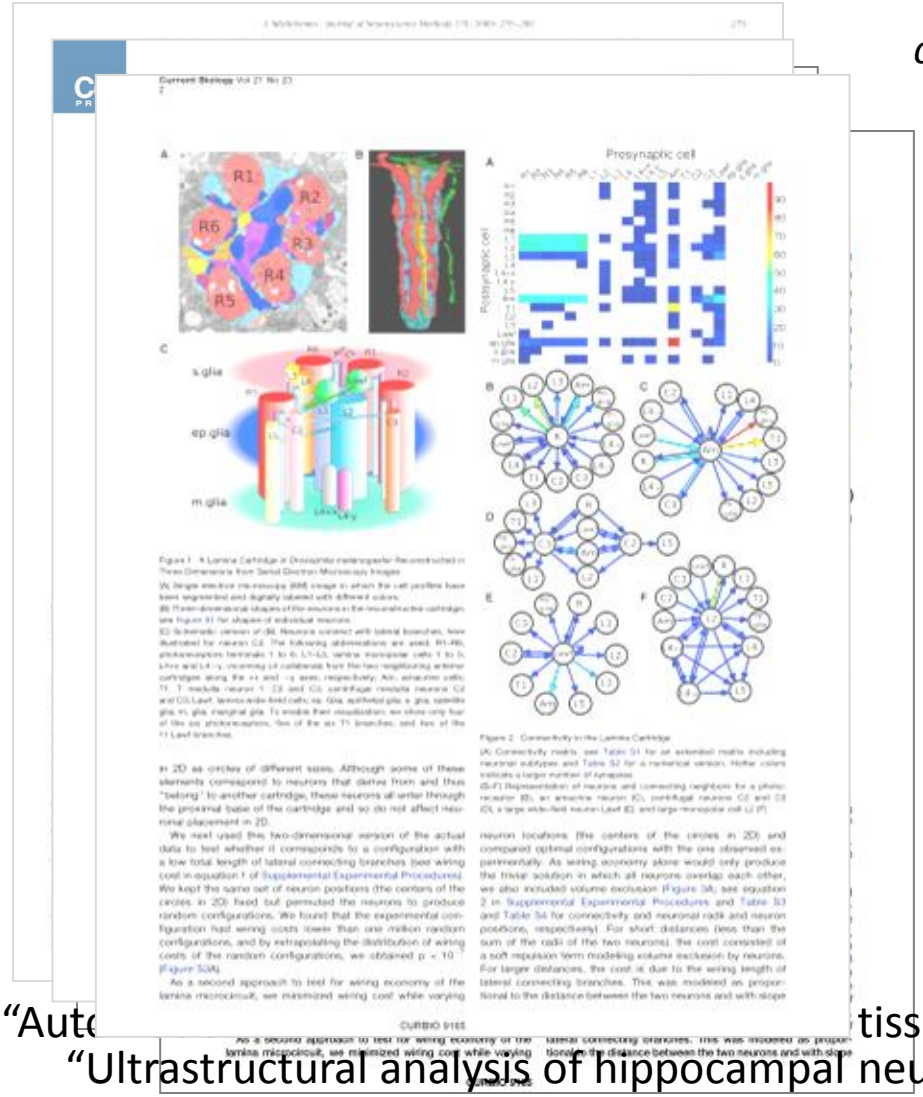
Reconstructing functional neural circuits with single cell resolution

Statistical methods for inferring neural network topology from large scale neural activity imaging data

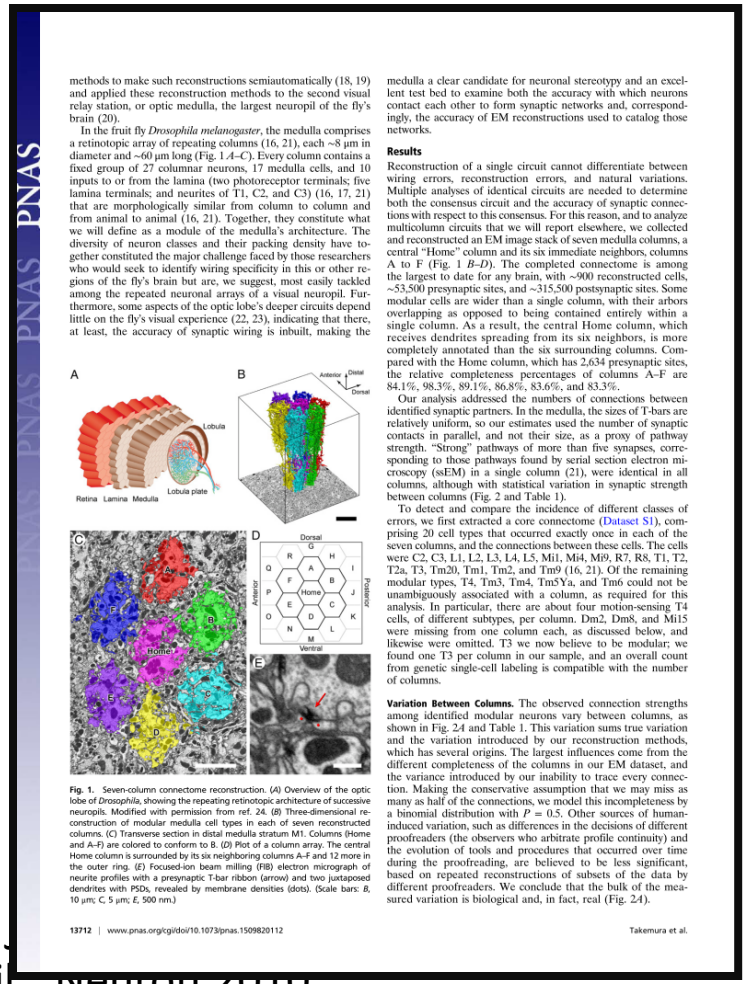
Yuriy Mishchenko

Prior encounter with Janelia 2007-2008: Analysis of serial EM data and reconstruction of dense volumes of cortical neuropil

Janelia 2007-2008: Serial EM reconstructions of dense neuropil …



"Synaptic circuits and their variations within different columns in the visual system of Drosophila", PNAS'2015



methods to make such reconstructions semiautomatically (18, 19) and applied these reconstruction methods to the second visual relay station, or optic medulla, the largest neuropil of the fly's brain (20).

In the fruit fly *Drosophila melanogaster*, the medulla comprises a retinotopic array of repeating columns (16, 21), each ~8 μm in diameter and ~60 μm long (Fig. 1 *A–C*). Every column contains a fixed group of 27 columnar neurons, 17 medulla cells, and 10 inputs to or from the lamina (two photoreceptor terminals; five lamina terminals; and neurites of T1, C2, and C3) (16, 17, 21) that are morphologically similar from column to column and from animal to animal (16, 21). Together, they constitute what we will define as a module of the medulla's architecture. The diversity of neuron classes and their packing density have together constituted the major challenge faced by those researchers who would seek to identify wiring specificity in this or other regions of the fly's brain but are, we suggest, most easily tackled among the repeated neuronal arrays of a visual neuropil. Furthermore, some aspects of the optic lobe's deeper circuits depend little on the fly's visual experience (22, 23), indicating that there, at least, the accuracy of synaptic wiring is inbuilt, making the medulla a clear candidate for neuronal stereotypy and an excellent test bed to examine both the accuracy with which neurons contact each other to form synaptic networks and, correspondingly, the accuracy of EM reconstructions used to catalog those networks.

Results

Reconstruction of a single circuit cannot differentiate between wiring errors, reconstruction errors, and natural variations. Multiple analyses of identical circuits are needed to determine both the consensus circuit and the accuracy of synaptic connections with respect to this consensus. For this reason, and to analyze multicolumn circuits that we will report elsewhere, we collected and reconstructed an EM image stack of seven medulla columns, a central "Home" column and its six immediate neighbors, columns A to F (Fig. 1 *B–D*). The completed connectome is among the largest to date for any brain, with ~900 reconstructed cells, ~53,500 presynaptic sites, and ~315,500 postsynaptic sites. Some modular cells are wider than a single column, with their arbors overlapping as opposed to being contained entirely within a single column. As a result, the central Home column, which receives dendrites spreading from its six neighbors, is more completely annotated than the six surrounding columns. Compared with the Home column, which has 2,634 presynaptic sites, the relative completeness percentages of columns A–F are 84.1%, 98.3%, 89.1%, 86.8%, 83.6%, and 83.3%.

Our analysis addressed the numbers of connections between identified synaptic partners. In the medulla, the sizes of T-bars are relatively uniform, so our estimates used the number of synaptic contacts in parallel, and not their size, as a proxy of pathway strength. "Strong" pathways of more than five synapses, corresponding to those pathways found by serial section electron microscopy (ssEM) in a single column (21), were identical in all columns, although with statistical variation in synaptic strength between columns (Fig. 2 and Table 1).

To detect and compare the incidence of different classes of errors, we first extracted a core connectome (Dataset S1), comprising 20 cell types that occurred exactly once in each of the seven columns, and the connections between these cells. The cells were C2, C3, L1, L2, L3, L4, L5, Mi1, Mi4, Mi9, R7, R8, T1, T2, T2a, T3, Tm20, Tm1, Tm2, and Tm9 (16, 21). Of the remaining modular types, T4, Tm3, Tm4, Tm5Ya, and Tm6 could not be unambiguously associated with a column, as required for this analysis. In particular, there are about four motion-sensing T4 cells, of different subtypes, per column. Dm2, Dm8, and Mi15 were missing from one column each, as discussed below, and likewise were omitted. T3 we now believe to be modular; we found one T3 per column in our sample, and an overall count from genetic single-cell labeling is compatible with the number of columns.

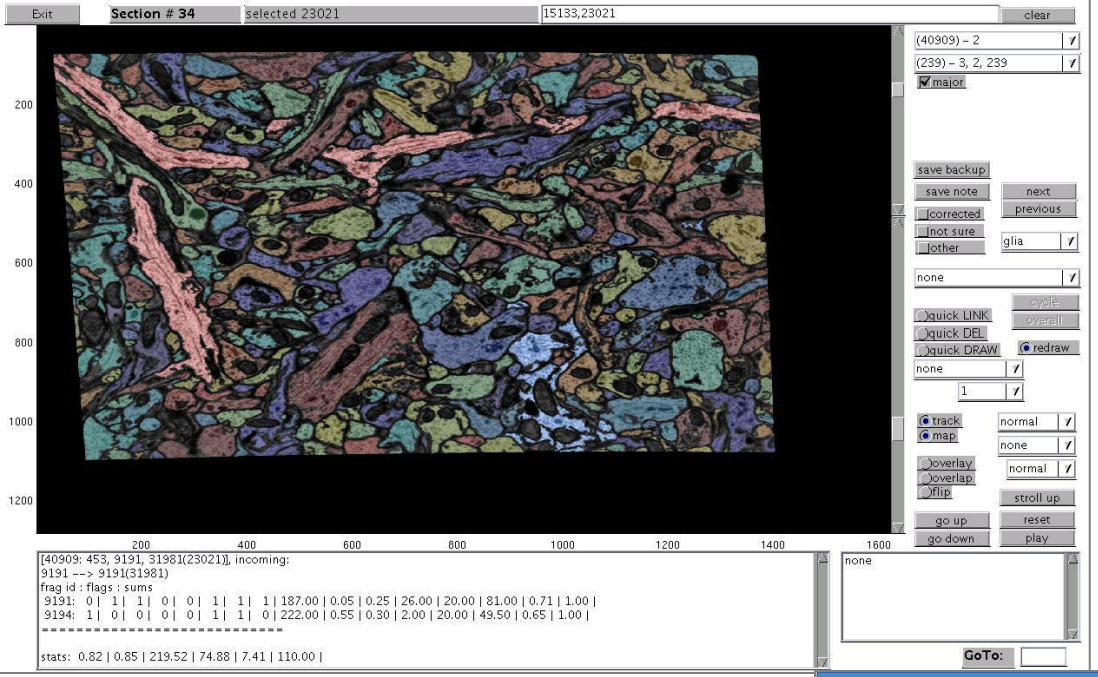
Variation Between Columns. The observed connection strengths among identified modular neurons vary between columns, as shown in Fig. 2*A* and Table 1. This variation sums true variation and the variation introduced by our reconstruction methods, which has several origins. The largest influences come from the different completeness of the columns in our EM dataset, and the variance introduced by our inability to trace every connection. Making the conservative assumption that we may miss as many as half of the connections, we model this incompleteness by a binomial distribution with $P = 0.5$. Other sources of human-induced variation, such as differences in the decisions of different proofreaders (the observers who arbitrate profile continuity) and the evolution of tools and procedures that occurred over time during the proofreading, are believed to be less significant, based on repeated reconstructions of subsets of the data by different proofreaders. We conclude that the bulk of the measured variation is biological and, in fact, real (Fig. 2*A*).

Fig. 1. Seven-column connectome reconstruction. (*A*) Overview of the optic lobe of *Drosophila*, showing the repeating retinotopic architecture of successive neuropils. Modified with permission from ref. 24. (*B*) Three-dimensional reconstruction of modular medulla cell types in each of seven reconstructed columns. (*C*) Transverse section in distal medulla stratum M1. Columns (Home and A–F) are colored to conform to B. (*D*) Plot of a column array. The central Home column is surrounded by its six neighboring columns A–F and 12 more in the outer ring. (*E*) Focused-ion beam milling (FIB) electron micrograph of neurite profiles with a presynaptic T-bar ribbon (arrow) and two juxtaposed dendrites with PSDs, revealed by membrane densities (dots). (Scale bars: *B*, 10 μm; *C*, 5 μm; *E*, 500 nm.)

"Auto… tissue", …

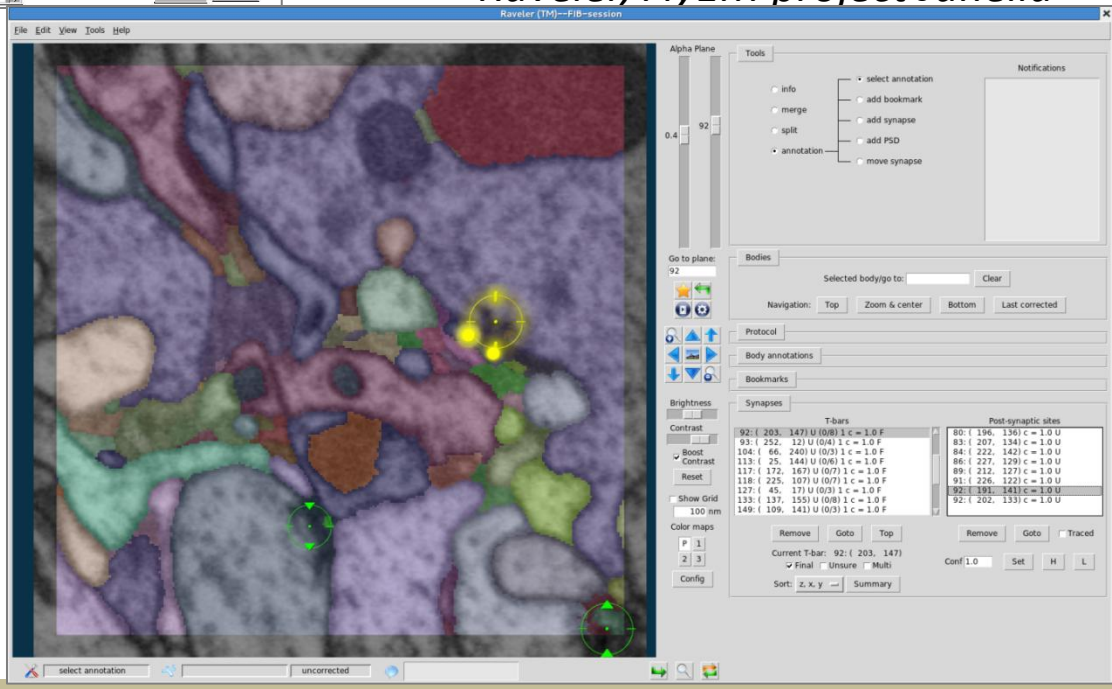
"Ultrastructural analysis of hippocampal neuropil", Neuron 2010

"Wiring economy and volume exclusion determines neuronal placements", Curr Biol'2011



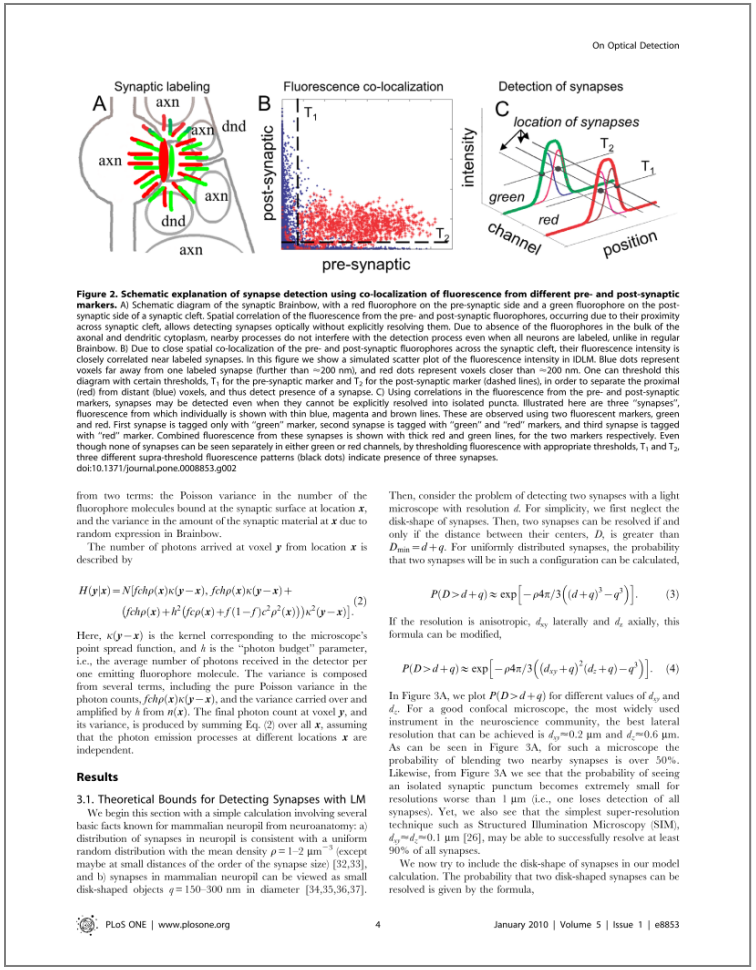
Raveler, FlyEM project Janelia

The old ProofReading Tool GUI in Matlab



On Optical Detection of Densely Labeled Synapses …

synaptic Brainbow



On Optical Detection

Figure 2. Schematic explanation of synapse detection using co-localization of fluorescence from different pre- and post-synaptic markers. A) Schematic diagram of the synaptic Brainbow, with a red fluorophore on the pre-synaptic side and a green fluorophore on the post-synaptic side of a synaptic cleft. Spatial correlation of the fluorescence from the pre- and post-synaptic fluorophores, occurring due to their proximity across synaptic cleft, allows detecting synapses optically without explicitly resolving them. Due to absence of the fluorophores in the bulk of the axonal and dendritic cytoplasm, nearby processes do not interfere with the detection process even when all neurons are labeled, unlike in regular Brainbow. B) Due to close spatial co-localization of the pre- and post-synaptic fluorophores across the synaptic cleft, their fluorescence intensity is closely correlated near labeled synapses. In this figure we show a simulated scatter plot of the fluorescence intensity in IDLM. Blue dots represent voxels far away from one labeled synapse (further than ~200 nm), and red dots represent voxels closer than ~200 nm. One can threshold this diagram with certain thresholds, T_1 for the pre-synaptic marker and T_2 for the post-synaptic marker (dashed lines), in order to separate the proximal (red) from distant (blue) voxels, and thus detect presence of a synapse. C) Using correlations in the fluorescence from the pre- and post-synaptic markers, synapses may be detected even when they cannot be explicitly resolved into isolated puncta. Illustrated here are three "synapses", fluorescence from which individually is shown with thin blue, magenta and brown lines. These are observed using two fluorescent markers, green and red. First synapse is tagged only with "green" marker, second synapse is tagged with "green" and "red" markers, and third synapse is tagged with "red" marker. Combined fluorescence from these synapses is shown with thick red and green lines, for the two markers respectively. Even though none of synapses can be seen separately in either green or red channels, by thresholding fluorescence with appropriate thresholds, T_1 and T_2, three different supra-threshold fluorescence patterns (black dots) indicate presence of three synapses.
doi:10.1371/journal.pone.0008853.g002

from two terms: the Poisson variance in the number of the fluorophore molecules bound at the synaptic surface at location $\boldsymbol{x}$, and the variance in the amount of the synaptic material at $\boldsymbol{x}$ due to random expression in Brainbow.

The number of photons arrived at voxel $\boldsymbol{y}$ from location $\boldsymbol{x}$ is described by

$$H(\boldsymbol{y}|\boldsymbol{x}) = N[fch\rho(\boldsymbol{x})\kappa(\boldsymbol{y}-\boldsymbol{x}),\ fch\rho(\boldsymbol{x})\kappa(\boldsymbol{y}-\boldsymbol{x}) + (fch\rho(\boldsymbol{x}) + h^2(fc\rho(\boldsymbol{x}) + f(1-f)c^2\rho^2(\boldsymbol{x})))\kappa^2(\boldsymbol{y}-\boldsymbol{x})]. \quad (2)$$

Here, $\kappa(\boldsymbol{y}-\boldsymbol{x})$ is the kernel corresponding to the microscope's point spread function, and h is the "photon budget" parameter, i.e., the average number of photons received in the detector per one emitting fluorophore molecule. The variance is composed from several terms, including the pure Poisson variance in the photon counts, $fch\rho(\boldsymbol{x})\kappa(\boldsymbol{y}-\boldsymbol{x})$, and the variance carried over and amplified by h from $n(\boldsymbol{x})$. The final photon count at voxel $\boldsymbol{y}$, and its variance, is produced by summing Eq. (2) over all $\boldsymbol{x}$, assuming that the photon emission processes at different locations $\boldsymbol{x}$ are independent.

Results

3.1. Theoretical Bounds for Detecting Synapses with LM

We begin this section with a simple calculation involving several basic facts known for mammalian neuropil from neuroanatomy: a) distribution of synapses in neuropil is consistent with a uniform random distribution with the mean density $\rho = 1–2\ \mu m^{-3}$ (except maybe at small distances of the order of the synapse size) [32,33], and b) synapses in mammalian neuropil can be viewed as small disk-shaped objects $q \approx 150–300$ nm in diameter [34,35,36,37].

Then, consider the problem of detecting two synapses with a light microscope with resolution d. For simplicity, we first neglect the disk-shape of synapses. Then, two synapses can be resolved if and only if the distance between their centers, D, is greater than $D_{min} = d + q$. For uniformly distributed synapses, the probability that two synapses will be in such a configuration can be calculated,

$$P(D > d+q) \approx \exp\left[-\rho 4\pi/3\left((d+q)^3 - q^3\right)\right]. \quad (3)$$

If the resolution is anisotropic, d_{xy} laterally and d_z axially, this formula can be modified,

$$P(D > d+q) \approx \exp\left[-\rho 4\pi/3\left((d_{xy}+q)^2(d_z+q) - q^3\right)\right]. \quad (4)$$

In Figure 3A, we plot $P(D > d+q)$ for different values of d_{xy} and d_z. For a good confocal microscope, the most widely used instrument in the neuroscience community, the best lateral resolution that can be achieved is $d_{xy} \approx 0.2$ μm and $d_z \approx 0.6$ μm. As can be seen in Figure 3A, for such a microscope the probability of blending two nearby synapses is over 50%. Likewise, from Figure 3A we see that the probability of seeing an isolated synaptic punctum becomes extremely small for resolutions worse than 1 μm (i.e., one loses detection of all synapses). Yet, we also see that the simplest super-resolution technique such as Structured Illumination Microscopy (SIM), $d_{xy} \approx d_z \approx 0.1$ μm [26], may be able to successfully resolve at least 90% of all synapses.

We now try to include the disk-shape of synapses in our model calculation. The probability that two disk-shaped synapses can be resolved is given by the formula,

"On optical detection of densely labeled synapses", PLoS ONE'2010

genetic BOINC

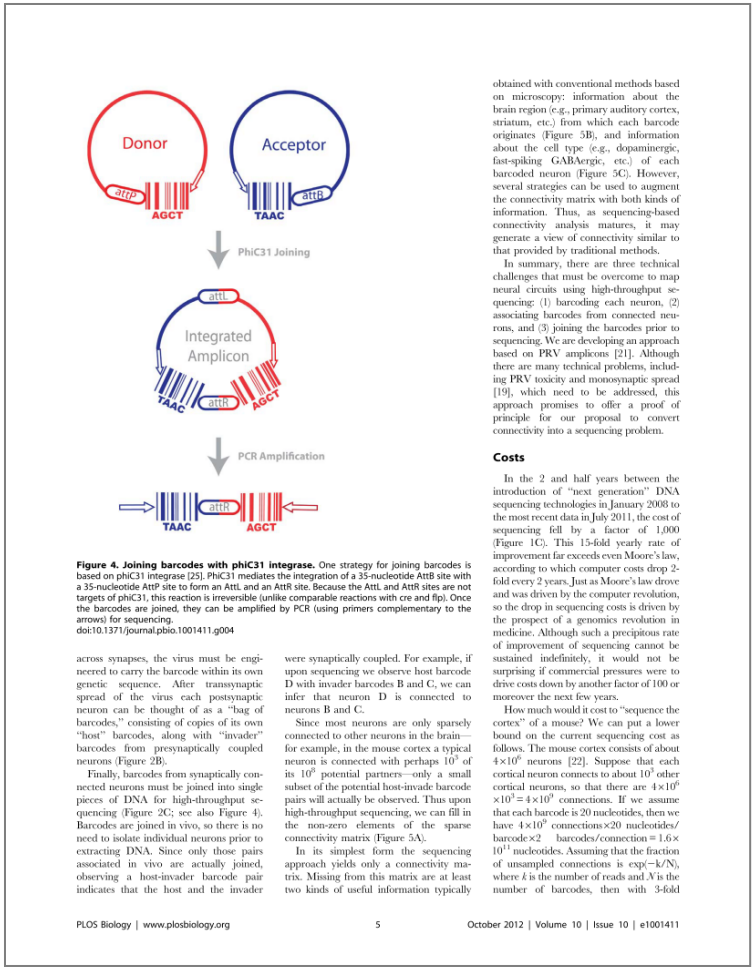


Figure 4. Joining barcodes with phiC31 integrase. One strategy for joining barcodes is based on phiC31 integrase [25]. PhiC31 mediates the integration of a 35-nucleotide AttB site with a 35-nucleotide AttP site to form an AttL and an AttR site. Because the AttL and AttR sites are not targets of phiC31, this reaction is irreversible (unlike comparable reactions with cre and flp). Once the barcodes are joined, they can be amplified by PCR (using primers complementary to the arrows) for sequencing.
doi:10.1371/journal.pbio.1001411.g004

across synapses, the virus must be engineered to carry the barcode within its own genetic sequence. After transsynaptic spread of the virus each postsynaptic neuron can be thought of as a "bag of barcodes," consisting of copies of its own "host" barcodes, along with "invader" barcodes from presynaptically coupled neurons (Figure 2B).

Finally, barcodes from synaptically connected neurons must be joined into single pieces of DNA for high-throughput sequencing (Figure 2C; see also Figure 4). Barcodes are joined in vivo, so there is no need to isolate individual neurons prior to extracting DNA. Since only those pairs associated in vivo are actually joined, observing a host-invader barcode pair indicates that the host and the invader were synaptically coupled. For example, if upon sequencing we observe host barcode D with invader barcodes B and C, we can infer that neuron D is connected to neurons B and C.

Since most neurons are only sparsely connected to other neurons in the brain—for example, in the mouse cortex a typical neuron is connected with perhaps 10^3 of its 10^8 potential partners—only a small subset of the potential host-invade barcode pairs will actually be observed. Thus upon high-throughput sequencing, we can fill in the non-zero elements of the sparse connectivity matrix (Figure 5A).

In its simplest form the sequencing approach yields only a connectivity matrix. Missing from this matrix are at least two kinds of useful information typically obtained with conventional methods based on microscopy: information about the brain region (e.g., primary auditory cortex, striatum, etc.) from which each barcode originates (Figure 3B), and information about the cell type (e.g., dopaminergic, fast-spiking GABAergic, etc.) of each barcoded neuron (Figure 5C). However, several strategies can be used to augment the connectivity matrix with both kinds of information. Thus, as sequencing-based connectivity analysis matures, it may generate a view of connectivity similar to that provided by traditional methods.

In summary, there are three technical challenges that must be overcome to map neural circuits using high-throughput sequencing: (1) barcoding each neuron, (2) associating barcodes from connected neurons, and (3) joining the barcodes prior to sequencing. We are developing an approach based on PRV amplicons [21]. Although there are many technical problems, including PRV toxicity and monosynaptic spread [19], which need to be addressed, this approach promises to offer a proof of principle for our proposal to convert connectivity into a sequencing problem.

Costs

In the 2 and half years between the introduction of "next generation" DNA sequencing technologies in January 2008 to the most recent data in July 2011, the cost of sequencing fell by a factor of 1,000 (Figure 1C). This 15-fold yearly rate of improvement far exceeds even Moore's law, according to which computer costs drop 2-fold every 2 years. Just as Moore's law drove and was driven by the computer revolution, so the drop in sequencing costs is driven by the prospect of a genomics revolution in medicine. Although such a precipitous rate of improvement of sequencing cannot be sustained indefinitely, it would not be surprising if commercial pressures were to drive costs down by another factor of 100 or moreover the next few years.

How much would it cost to "sequence the cortex" of a mouse? We can put a lower bound on the current sequencing cost as follows. The mouse cortex consists of about 4×10^6 neurons [22]. Suppose that each cortical neuron connects to about 10^3 other cortical neurons, so that there are $4\times10^6 \times 10^3 = 4\times10^9$ connections. If we assume that each barcode is 20 nucleotides, then we have 4×10^9 connections×20 nucleotides/barcode×2 barcodes/connection = 1.6×10^{11} nucleotides. Assuming that the fraction of unsampled connections is $\exp(-k/N)$, where k is the number of reads and N is the number of barcodes, then with 3-fold

"Sequencing the connectome", PLoS Biology'2012

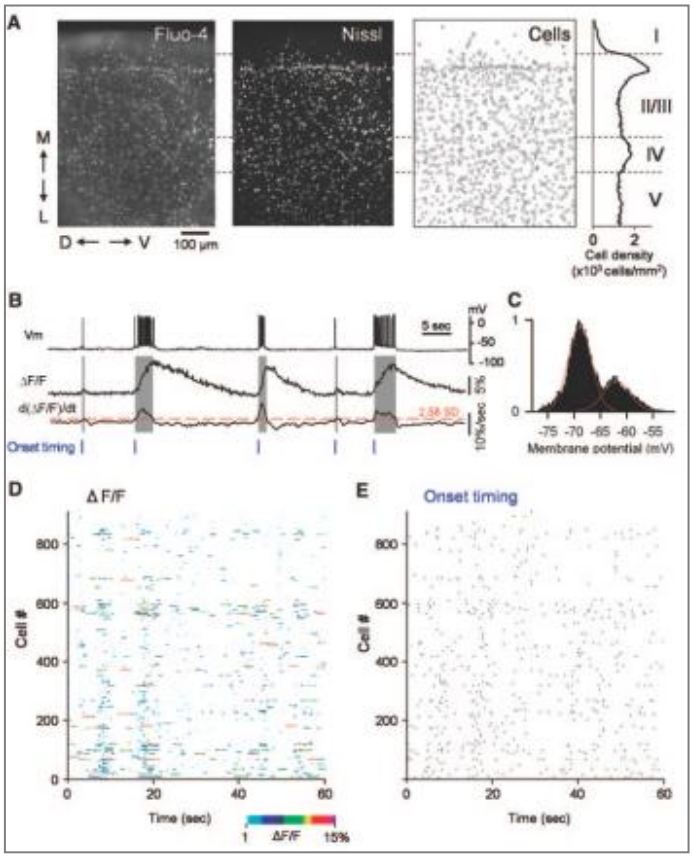
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
CENTER FOR THEORETICAL NEUROSCIENCE
columbia university college of physicians and surgeons | 1051 riverside dr, new york ny 10032

- Statistical estimation of neural circuits from large-scale calcium imaging data (Columbia University, CTN)

Liam Paninski, Columbia Univ.

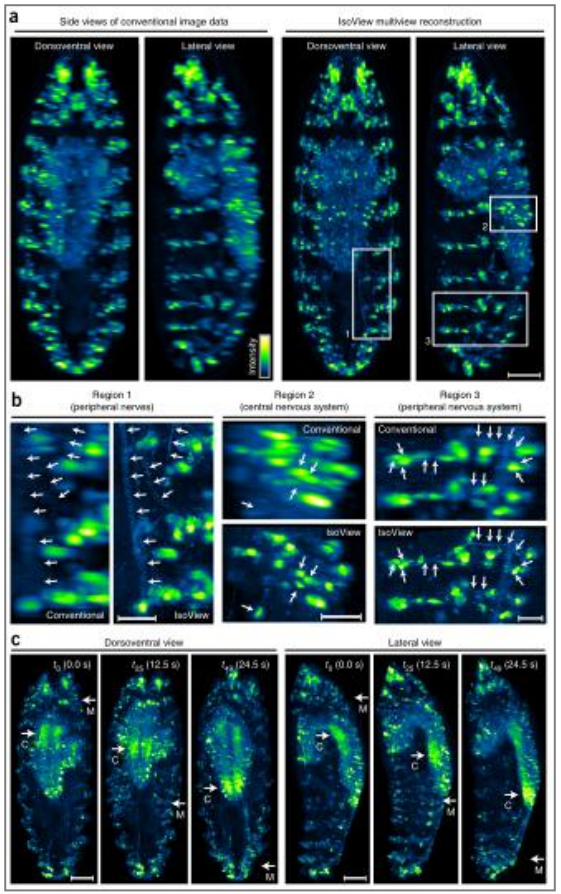
Dramatic progress in population calcium imaging …

2005



Ikegaya et al., Science'2004

2015



Chhetri et al, Nat Methods'2015

"Compared to the rapid advances in experimental methods, computational analysis of imaging data remains in its infancy. Currently used methods are *ad hoc*, slow, poorly documented, and differ across labs, implying that hard won experimental data are underutilized. A lack of standardization hinders reproducibility and comparison across studies… Nearly complete automation and modern computational methods … will have to supplant the semi-manual methods in use today to fully exploit the richness of these datasets."

Peron, Chen & Svoboda " Comprehensive imaging of cortical networks", Curr Opin Neurobiol'2015.

In 2004-2010 Chichilnisky, Simoncelli, Pillow and Liam Paninski made significant progress in the applications of statistical models of neuronal activity to the analysis of real biological neurons, demonstrating that a certain class of such models (GLM) can be extremely successful in describing the behavior of real Ganglion cells in retina

My question: How can this framework be applied to the problem of reconstructing the connectivity of neural networks from large-scale calcium neural activity imaging data?

The Generalized Linear Model

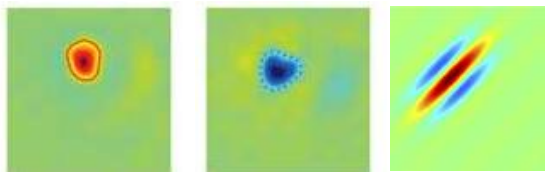
GLM is a statistical model of neuronal spiking

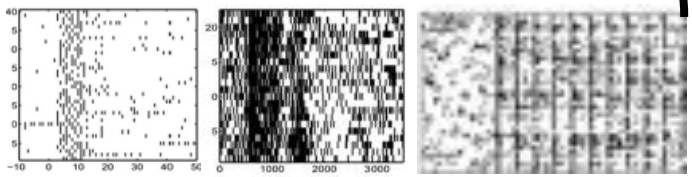
$$P\{s(t)=1\} = f\left(k \cdot x(t) + \sum_{t'<t} h(t-t')s(t')\right)$$

Probability of spiking at time *t*

Stimulus term

Spike-history term





$$P\{s(t)=1\} = f\left(k \cdot x(t) + \sum_{t'<t} h(t-t')s(t')\right)$$

Two reasons for the success of the GLM in the prior Chichilnisky et al's work:

- The rich repertoire of neuronal behaviors that can be captured by the GLMs
- The ease with which the model parameters can be fit to describe the real neurons

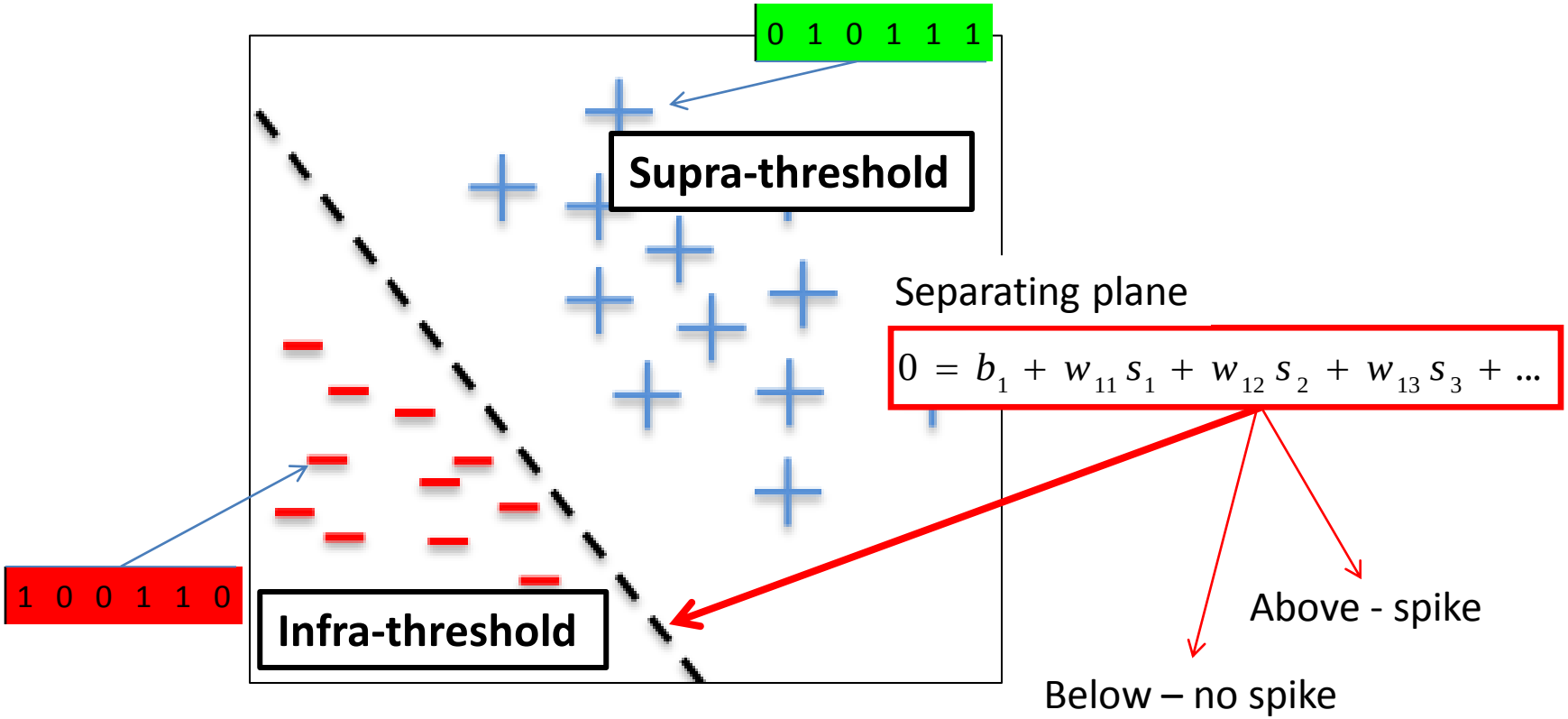
Estimating GLM for real neurons

Contrast patterns of inputs associated with the target neuron's producing a spike vs. such not producing a spike:

Target neuron	Other neuron-1	Other neuron-2	Other neuron-3	Other neuron-4	Other neuron-5	Other neuron-6
0	1	0	0	1	1	0
1	**0**	**0**	**1**	**1**	**1**	**0**
1	**0**	**1**	**0**	**1**	**1**	**1**
0	1	1	0	0	0	1
0	1	0	1	1	0	0
1	**0**	**0**	**1**	**1**	**0**	**0**
0	0	1	0	0	1	1

High-D patterns of inputs

"Target-spike" and "target-no spike" patterns in a high-D configuration space of input patterns:



A principled approach to finding the separating plane is provided by the Maximum Likelihood Estimation (MLE, we use, others are available)

- Find the GLM that maximizes the chances of having observed the neural activity that was actually observed given generative GL model

$$\text{loglik} = \sum_{i=1}^{i=N_n} \sum_{t=1}^{t=T} \left\{ \bar{s}_i(t) \log f\left(J_i(t)\right) - f\left(J_i(t)\right)\Delta t \right\}$$

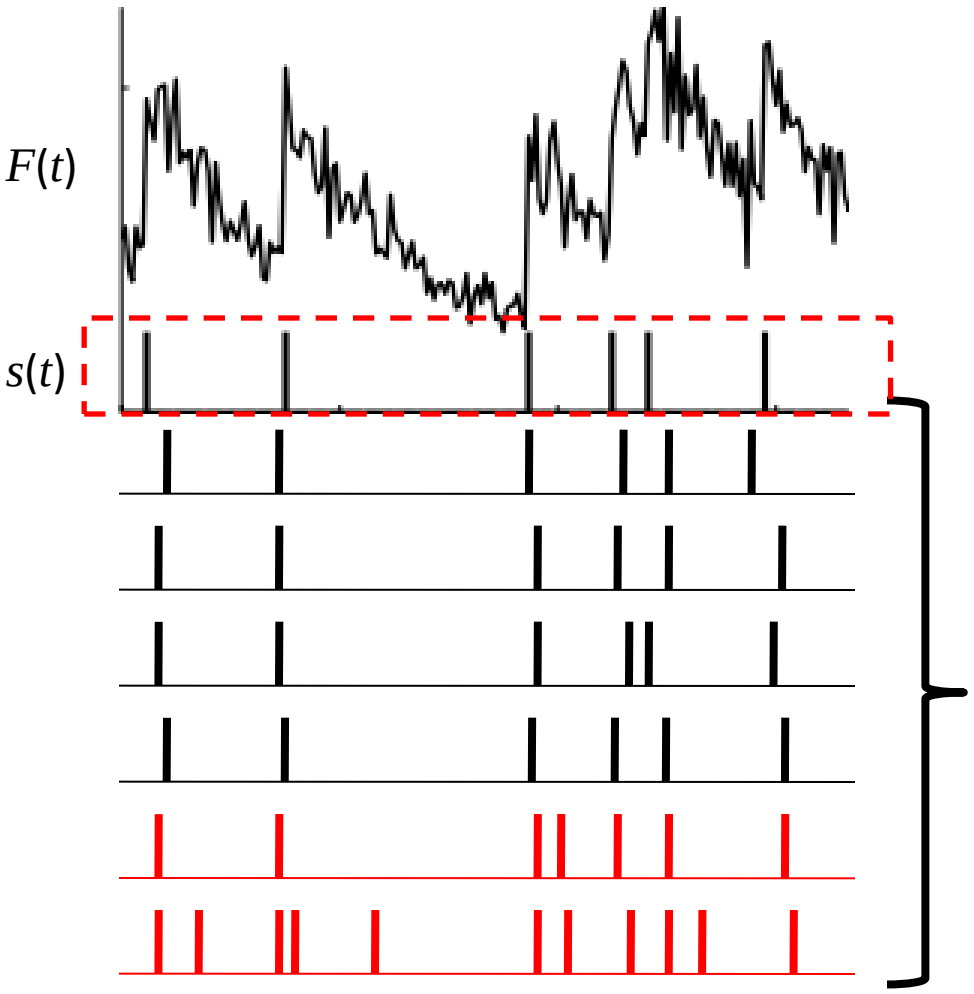
$$J_i(t) = \hat{b}_i + \hat{k}_i \cdot \bar{x}(t) + \sum_{t'<t} \hat{h}_i(t-t')\bar{s}_i(t') + \sum_{j\neq i} \sum_{t'<t} \hat{w}_{ij}(t-t')\bar{s}_j(t')$$

- In fact, not a difficult problem, the solution for several hundreds to thousands of neurons can be produced on a laptop with Matlab in matter of hours

- Calcium imaging data → new layer of complexity

The problem with calcium imaging data is that the Ca fluorescence traces, $F_i(t)$, do not really fix the underlying spike trains, $s_i(t)$.

Example:



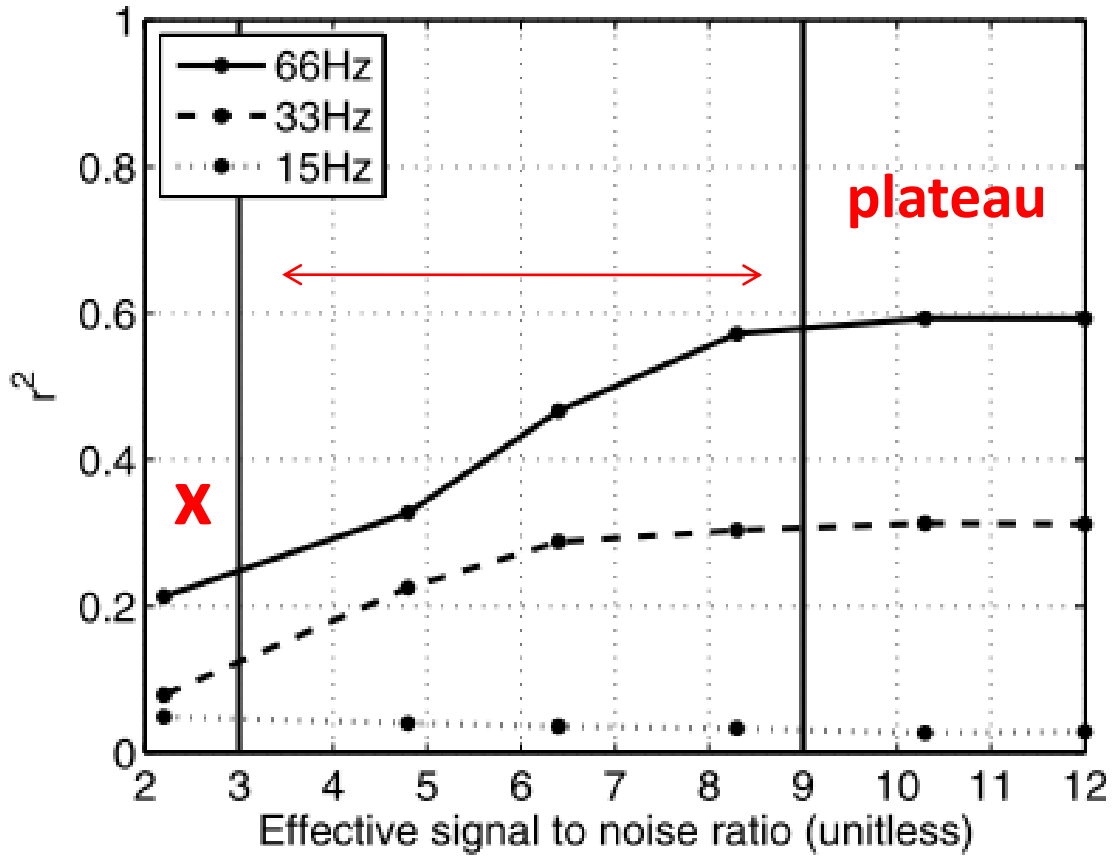
AVERAGE OVER ALL !

- We fully implemented the solution for this problem as NETFIT+OOPSI package for Matlab (*details in "A Bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data", Annals of Applied Statistics'2011*)
- Successfully tested the reconstruction of neural networks' structure in simulated cortical neural networks for up to 1000 neurons

- Calculations ran on Columbia University's STAT computing cluster – 256 cores Intel Xeon L5430 2.66GHz
- Typical solution time – 1 hour per 1 neuron
- Computation cost is not too high – can be easily handled by Amazon AWS or NFS's HPC infrastructures
- Hypothetically, 100,000 neurons → 100,000 compute-hours solution time – a below average ran-time of many physics/weather HPC simulation projects

Use that solution to look at how the calcium imaging inference is affected by different parameters of the experimental calcium imaging setups

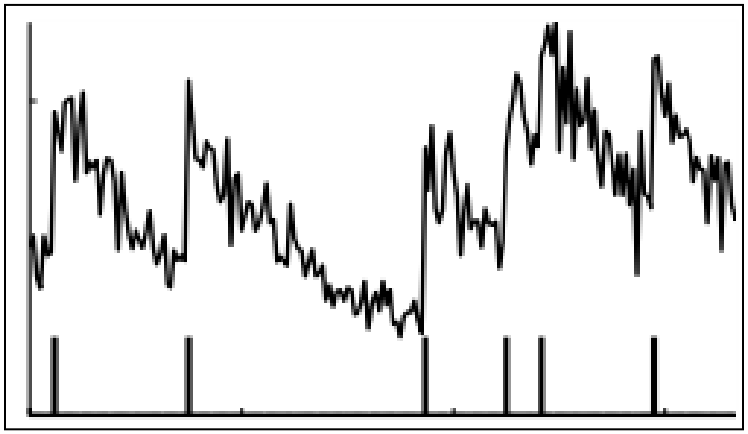
Signal-to-Noise Ratio



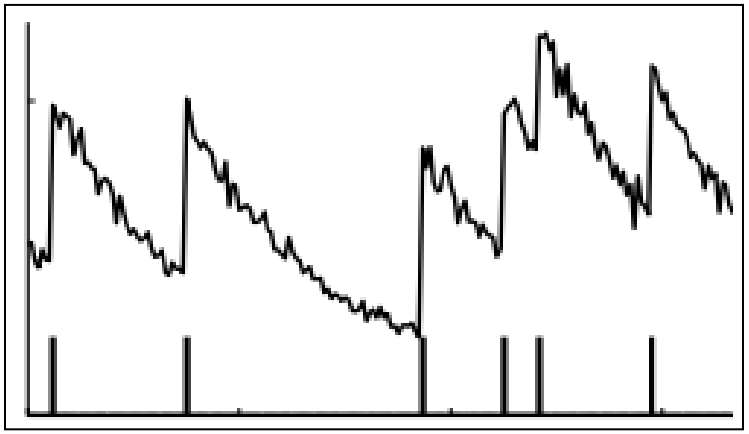
SNR here is:

$$SNR=\Delta F(spike)/STD[\Delta F(nospike)]$$

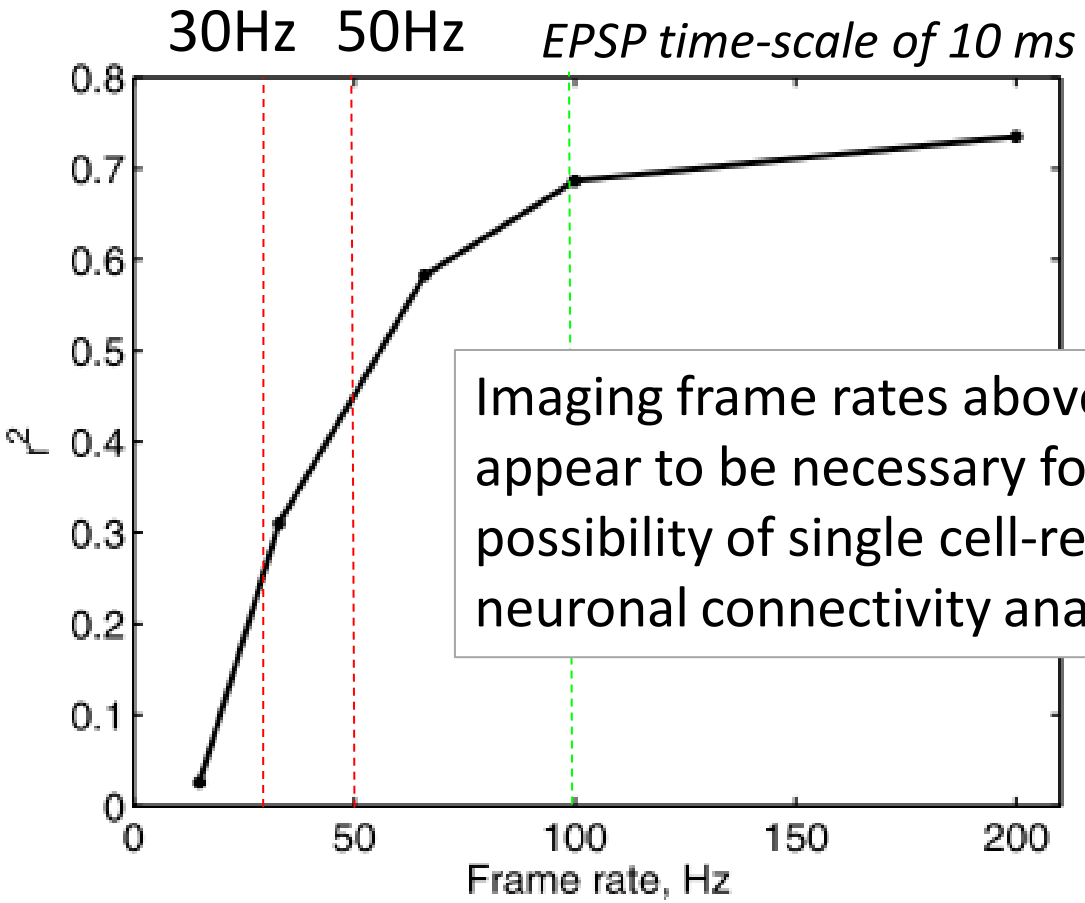
SNR=3



SNR=9

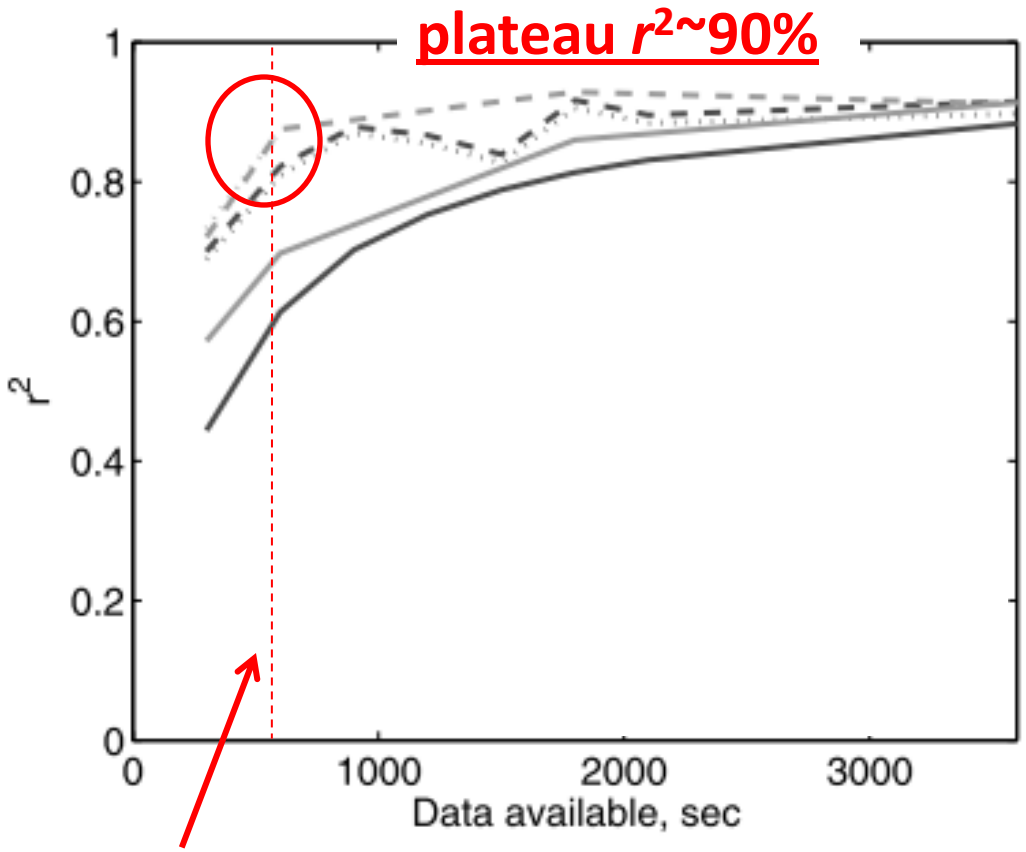


Frame Rate



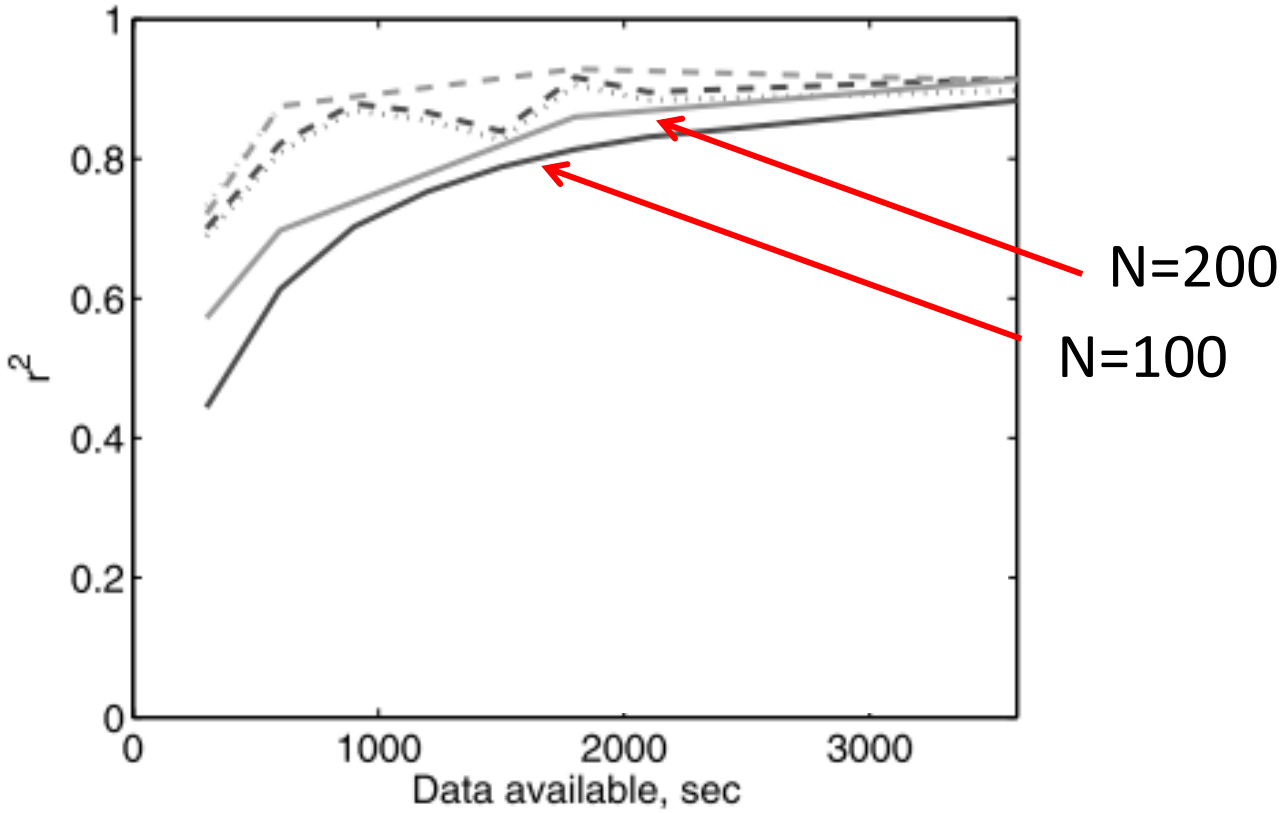
Imaging frame rates above 30 Hz appear to be necessary for the possibility of single cell-resolution neuronal connectivity analysis

Imaging Time Requirements

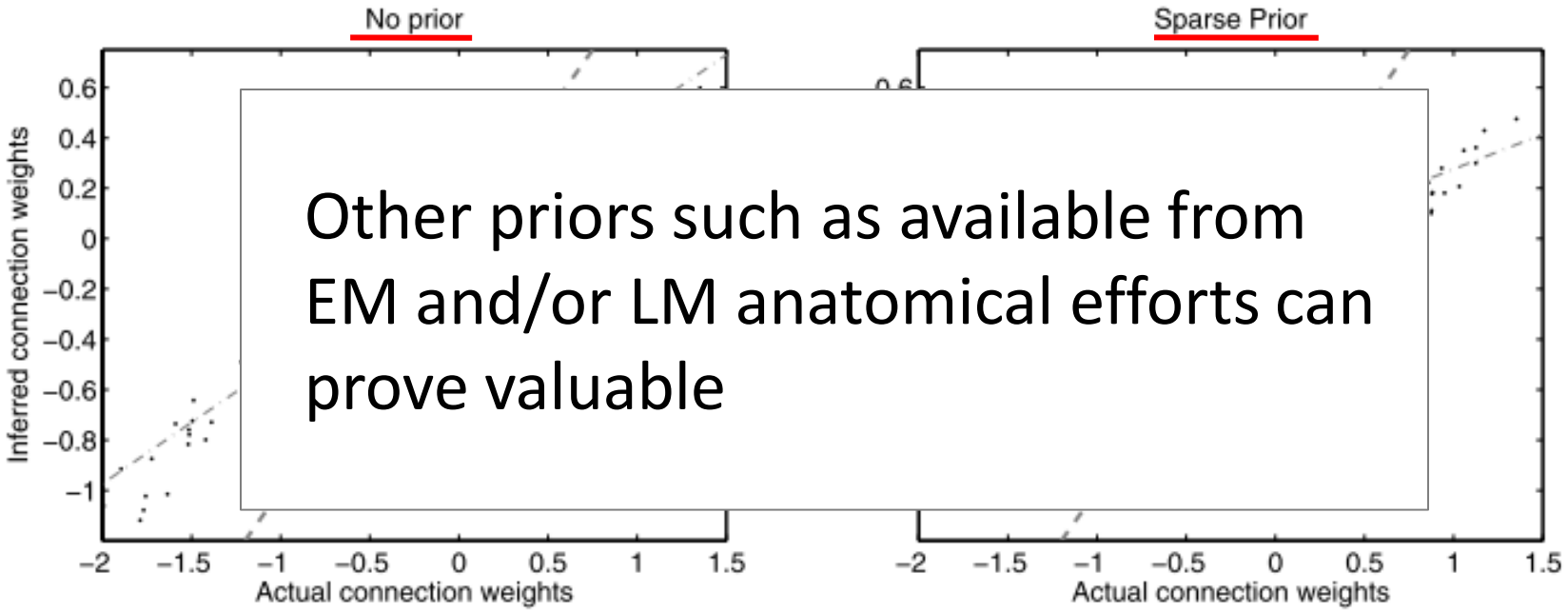


600 seconds at <r>=5 Hz ~ 3000 spikes/neuron

Larger neuronal circuits do not require longer imaging times



Prior information can help dramatically!



Other priors such as available from EM and/or LM anatomical efforts can prove valuable

R^2=0.65

R^2=0.85

What is extracted?

- The matrix of parameters w of the statistical model of neuronal population activity below;

$$P\{s_i(t)=1\} = f\left(b_i + \sum_{t'<t} h_i(t-t')s_i(t') + \underbrace{\sum_{j\neq i}\sum_{t'<t} w_{ij}(t-t')s_j(t')}\right)$$

The long winding road of the concept of neural connectivity …

Synaptic connectivity

Structural connectivity

Neuronal connectivity

Functional connectivity

Effective connectivity

Effective connectivity is defined as the parameter of a statistical generative model (typically, a network-type model) of a neuronal population's activity

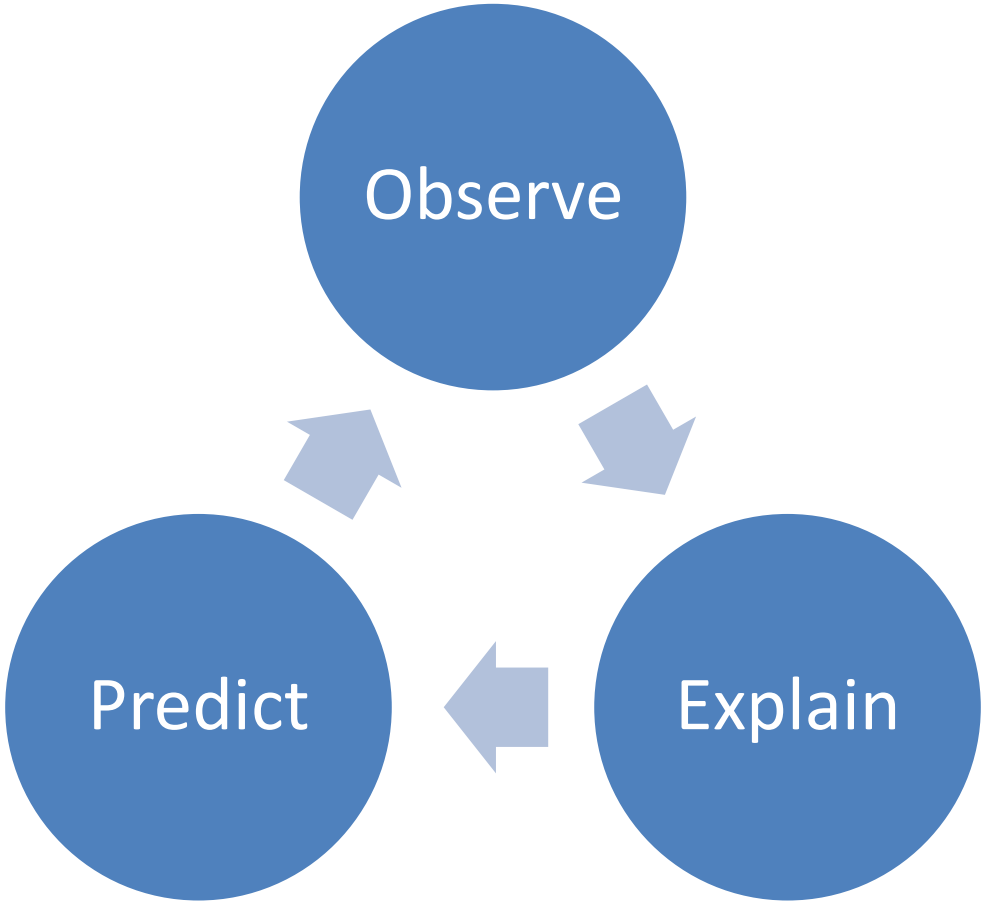
$$s(t) \sim P(s(t); s(t<t'), \underline{\boldsymbol{W}})$$

Meant to say that it is **NOT** synaptic connectivity, **NOT** structural connectivity, and **NOT** functional connectivity

$$s(t) \sim P(s(t); s(t<t'), W)$$

Why?

- Effective connectivity is the parameter of a model of neuronal population activity that is *generative* , *causal* and *predictive* at that

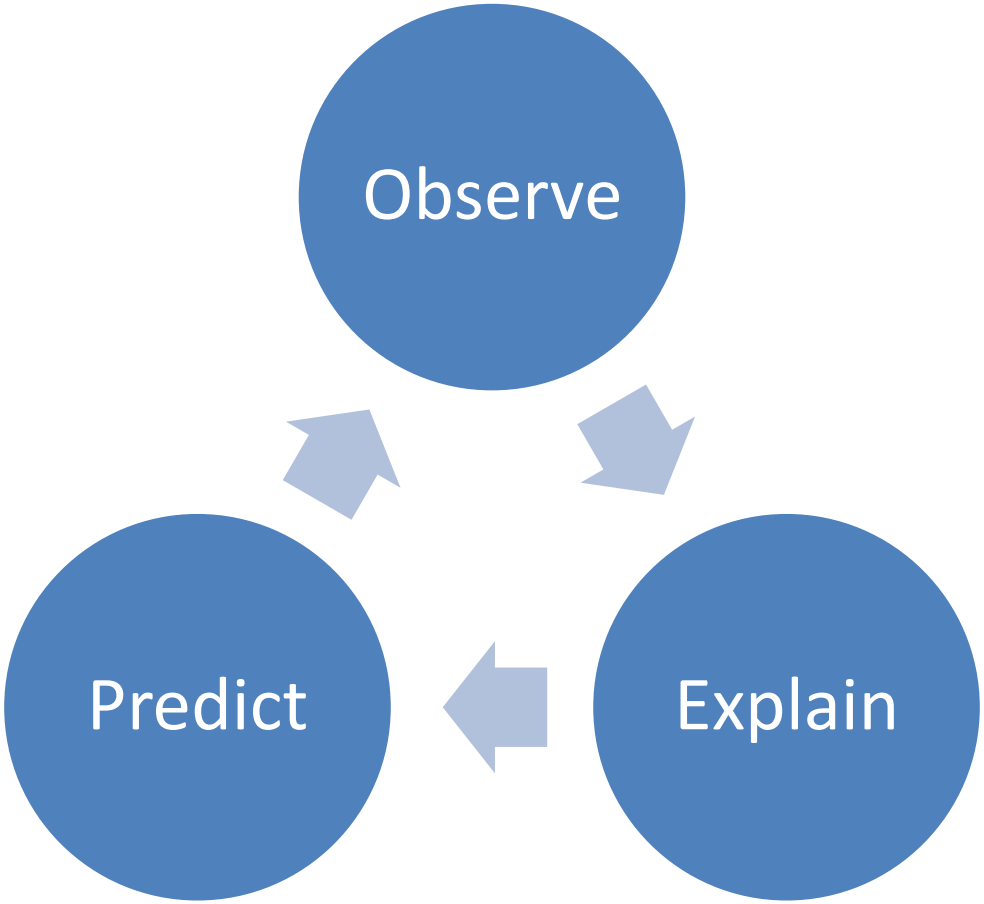

Observe
Explain
Predict

Another way to look at it:

- Allows to simulate examples of neural activity equivalent to such observed in a real neuronal population;
- Allows to contrast different models against real neuronal populations and vice versa;
- Can explain how different activity patterns emerge in neuronal population, also causally;
- Can predict the response if something in neuronal population changes;
- Provides quantitative way to check if that prediction was correct

= Allows one to ask ‘what if’ questions, make testable predictions for them, and test such predictions (quantitatively)

Scientific Method's Cycle:

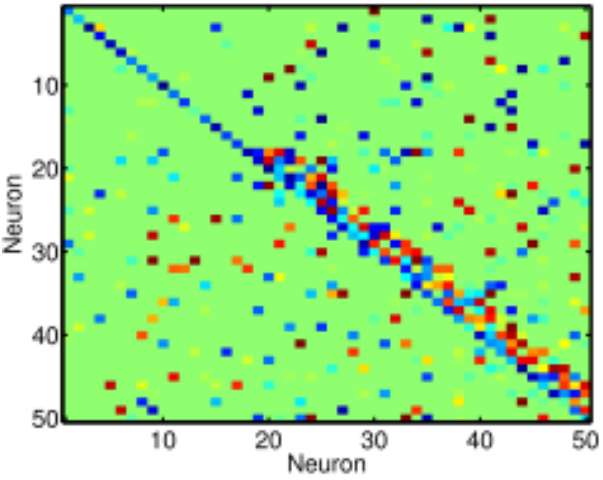


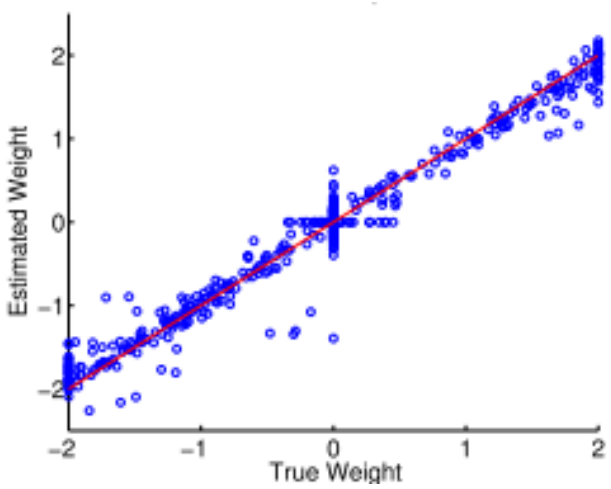
Sparse imaging of neuronal activity

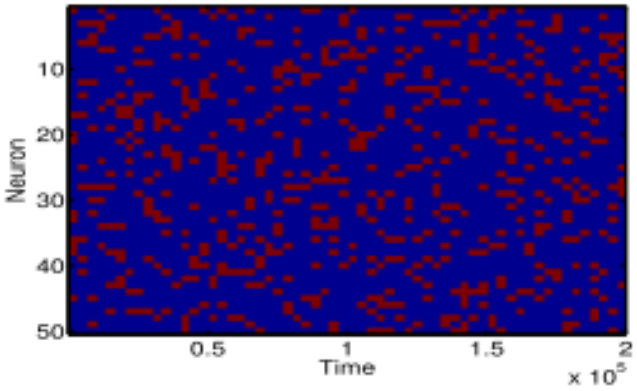
- Originally suggested for the problem of hidden inputs, but can be also used to improve the Frame-Rate and the SNR in large-scale population calcium imaging
- The question: If we cannot observe the activity of complete neuronal circuit, what shall we do?

“Shotgun” proposition

- Look at a small number of random neurons at a time
- Do this long enough so that all neurons had been looked at
- Attempt to recover the *complete connectivity matrix* from such “shredded” observation







Keshri et al, "A shotgun sampling solution for the common input problem in neural connectivity inference" arXiv'2013

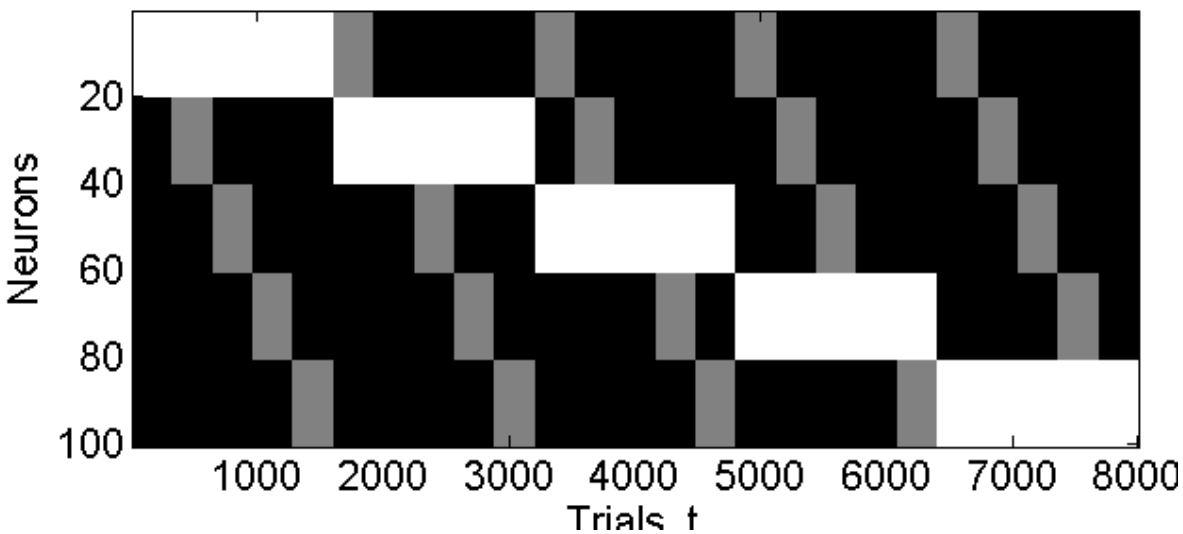
"Consistent estimation of complete neuronal connectivity in large neuronal populations…" (under review in *J Comp Neurosci*)

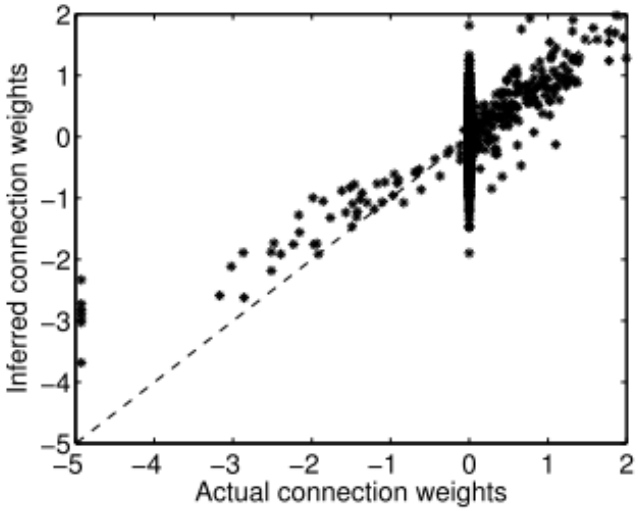
- Determine precisely under what conditions the effective connectivity matrix of a complete neuronal population can be recovered from such partial observations
- Develop a numerical algorithm for solving the associated connectivity estimation problem
- Test in simulations

Key conclusions: simply assembling the observations of all input-output neuronal activity pairs, e.g. $(s_i(t),s_j(t\text{-}1))$, is sufficient to estimate the effective connectivity matrix of a complete neuronal population in great many cases

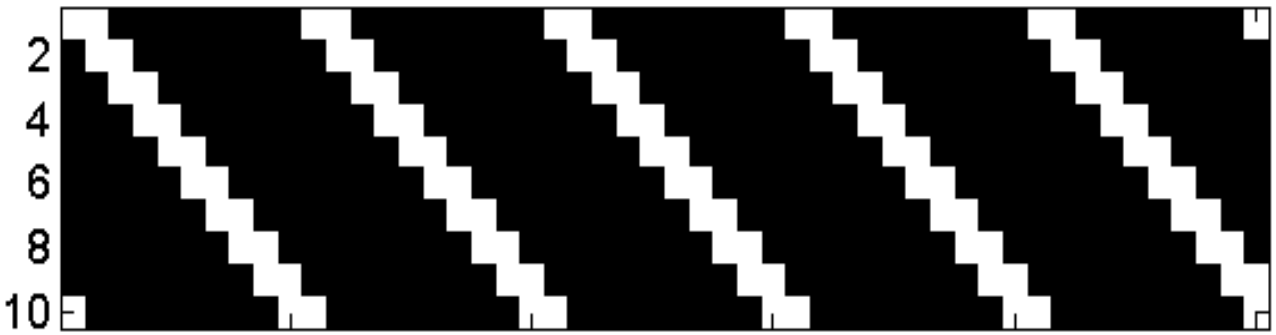
Experimentally difficult randomized "shotgun" scanning of neuronal population is not strictly necessary, much simpler and equally effective imaging protocols can be designed

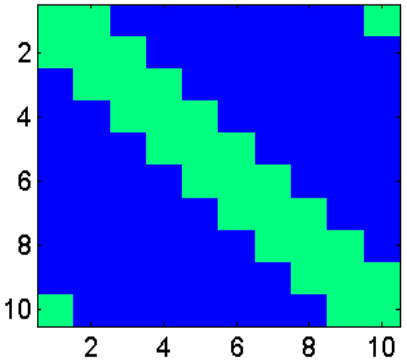
Alternative sparse imaging strategy – double-block scanning



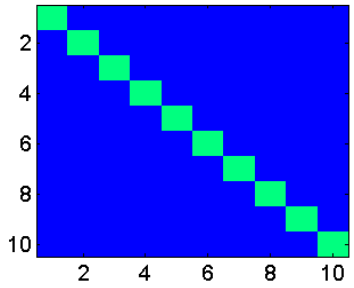


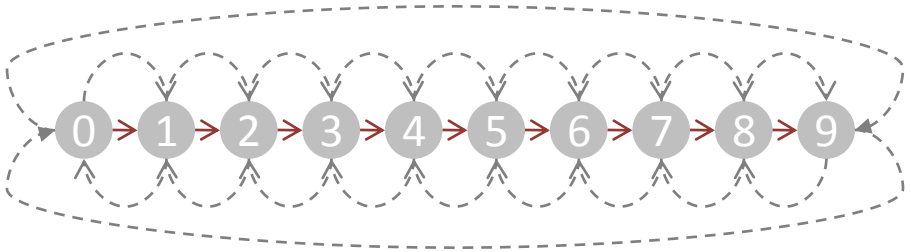
Single-block scanning does not provide necessary information to extract the connectivity of a neuronal population





TRUTH:





Suggestions:

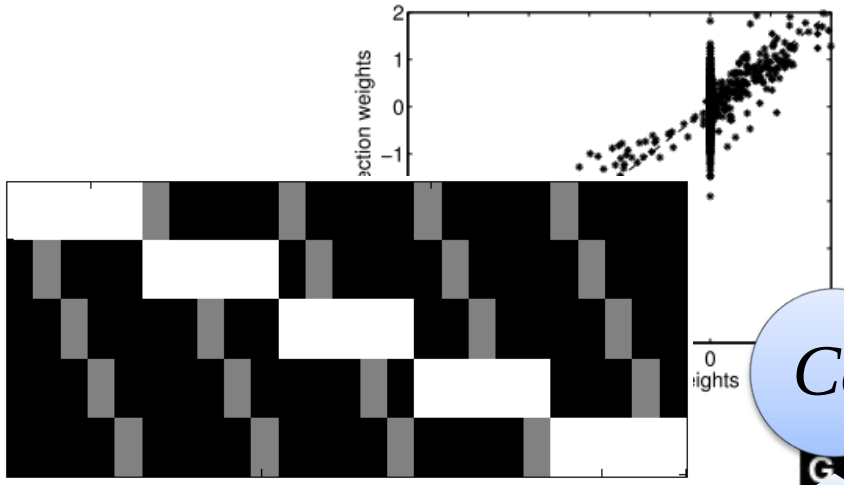
- Organize the whole-brain imaging in double-block scanning manner, microscopy can dwell at single input-output blocks extended times
- Can increase the SNR and the frame-rate of imaging (5Hz→ 30Hz)
- Can piece the observation information together computationally, nontrivial but definitely possible and the proof as well as the proof-of-principle are now available

Outlook

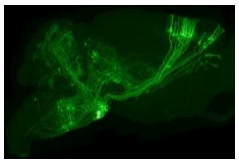
- Recent advances at Janelia related to whole-brain imaging look promising together with the described analytical framework for the goal of reconstructing whole-brain functional connectomes

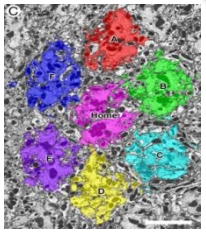
- Critical to combine theoretical and experimental effort: effective integration of theoretical and experimental work cannot be achieved in isolation

- Janelia's extensive LM and EM mapping projects may be valuable as set priors

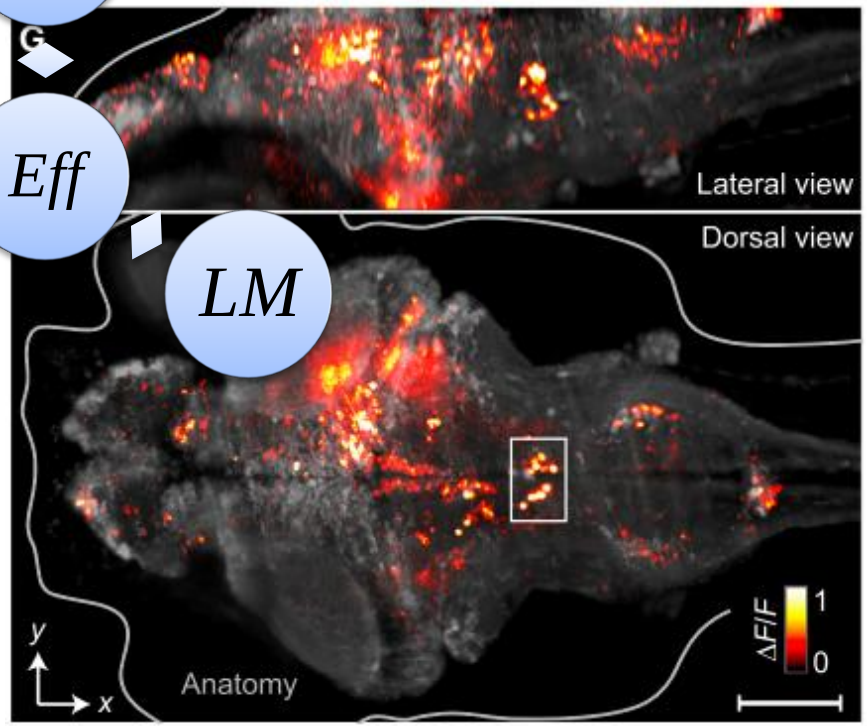


The theory of effective connectivity estimation from sparse neural activity data





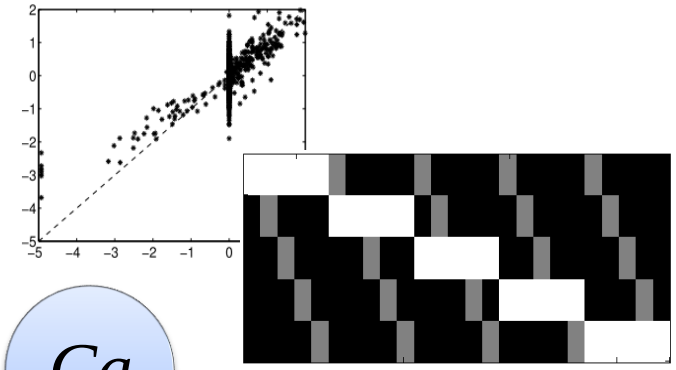
The whole-brain neural activity imaging using emerging LM technologies

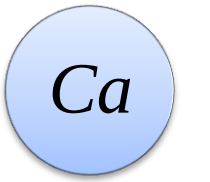


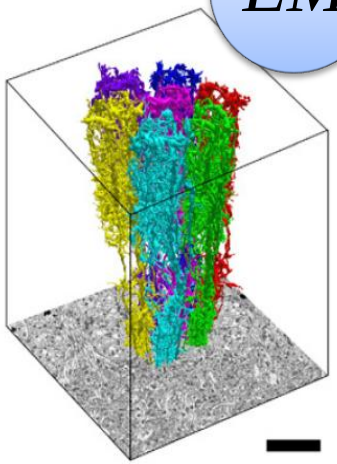
"Visualizing Whole-Brain Activity and Development", Neuron'2015

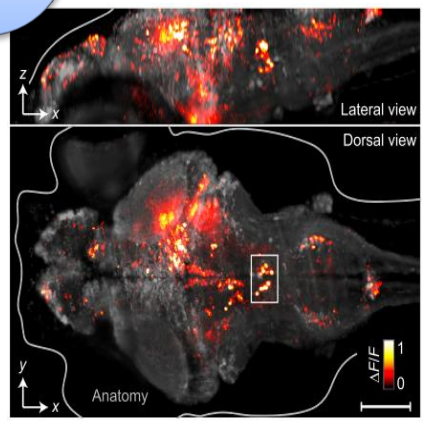
Acknowledgements

- Liam Paninski, Prof. Dr.
- Columbia University, Dept. of Statistics and Center for Theoretical Neuroscience
- Toros University, Dept. of Computer and Software Engineering
- BAGEP Young Investigator Scholarship Award, The Science Academy, Turkey
- TUBITAK ARDEB 1001 Grant Number 113E611
- Janelia Research Campus, HHMI









EXTRAS

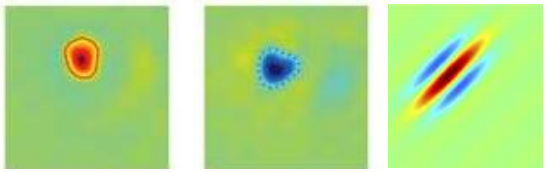
GLM model of neuronal spiking

Rate function

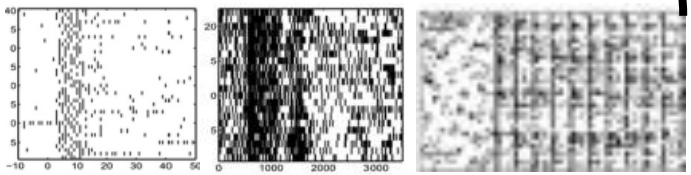
$$P\{s(t)=1\} = f\left(k \cdot x(t) + \sum_{t'<t} h(t-t')s(t')\right)$$

Probability of spiking at time t

Stimulus term

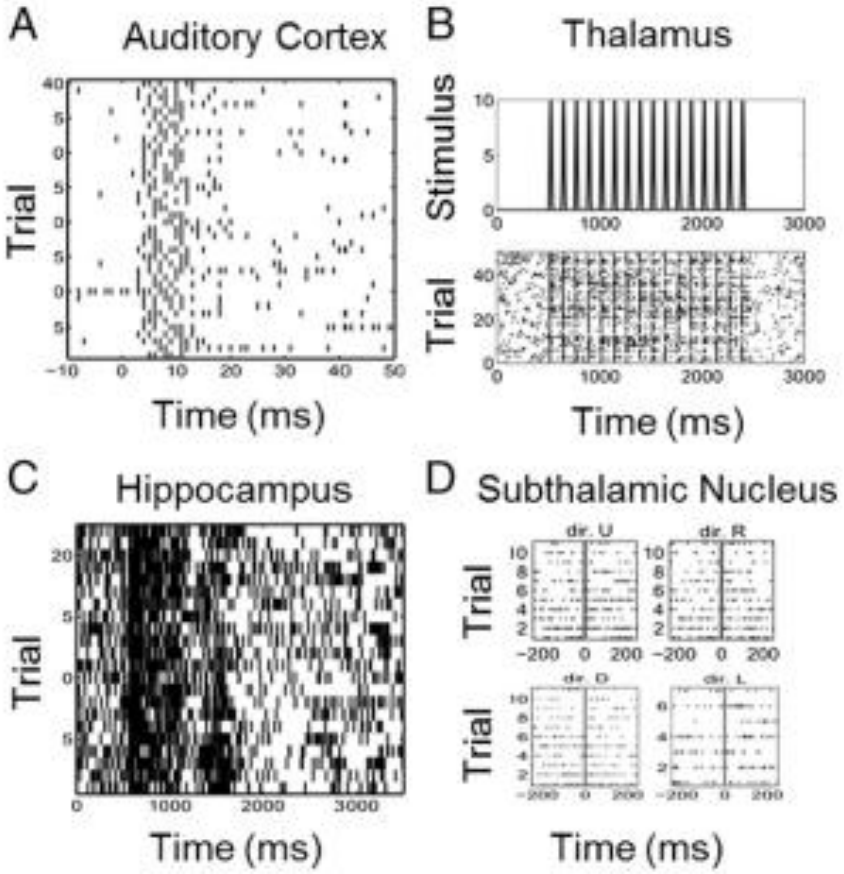


Spike-history term

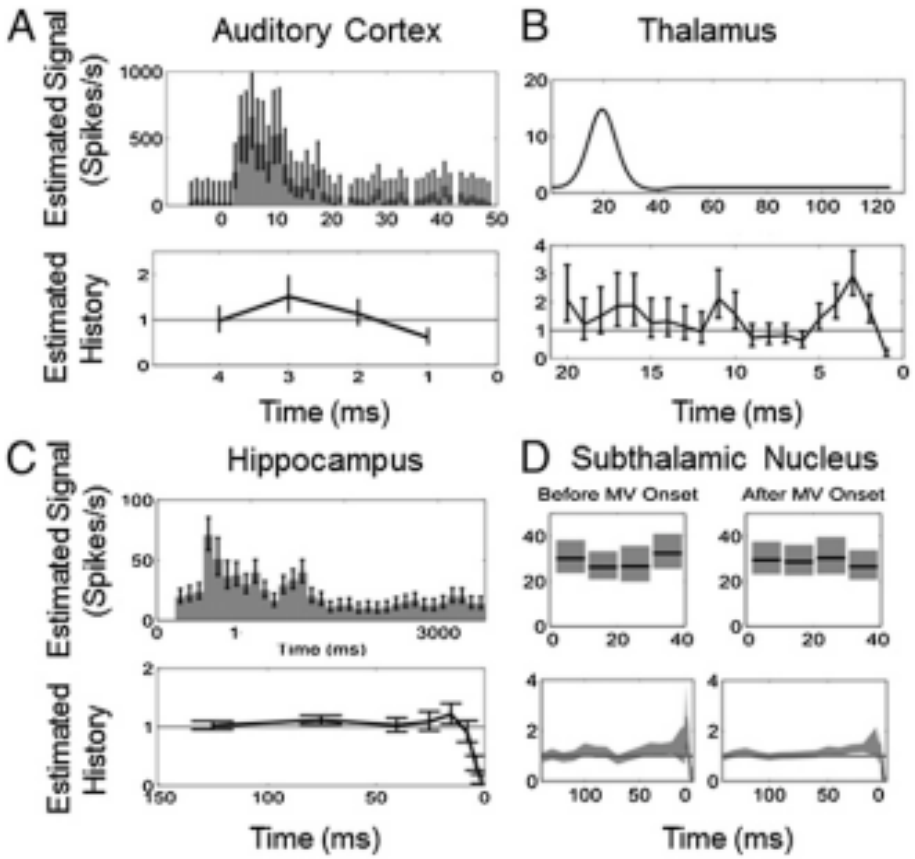


Key addition to earlier models is the spike-history h-term that allows the model to capture and reproduce a variety of complex neural behaviors such as refractory periods, bursting, periodic spiking, etc.

Different patterns of neural activity in stimulus-response paradigm

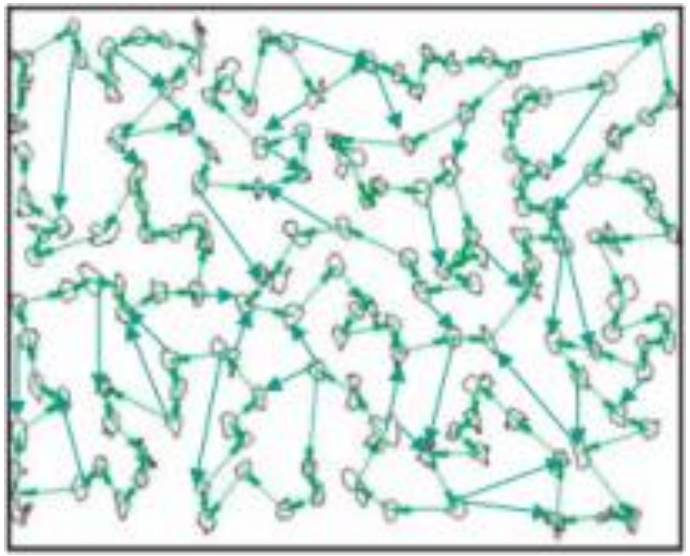


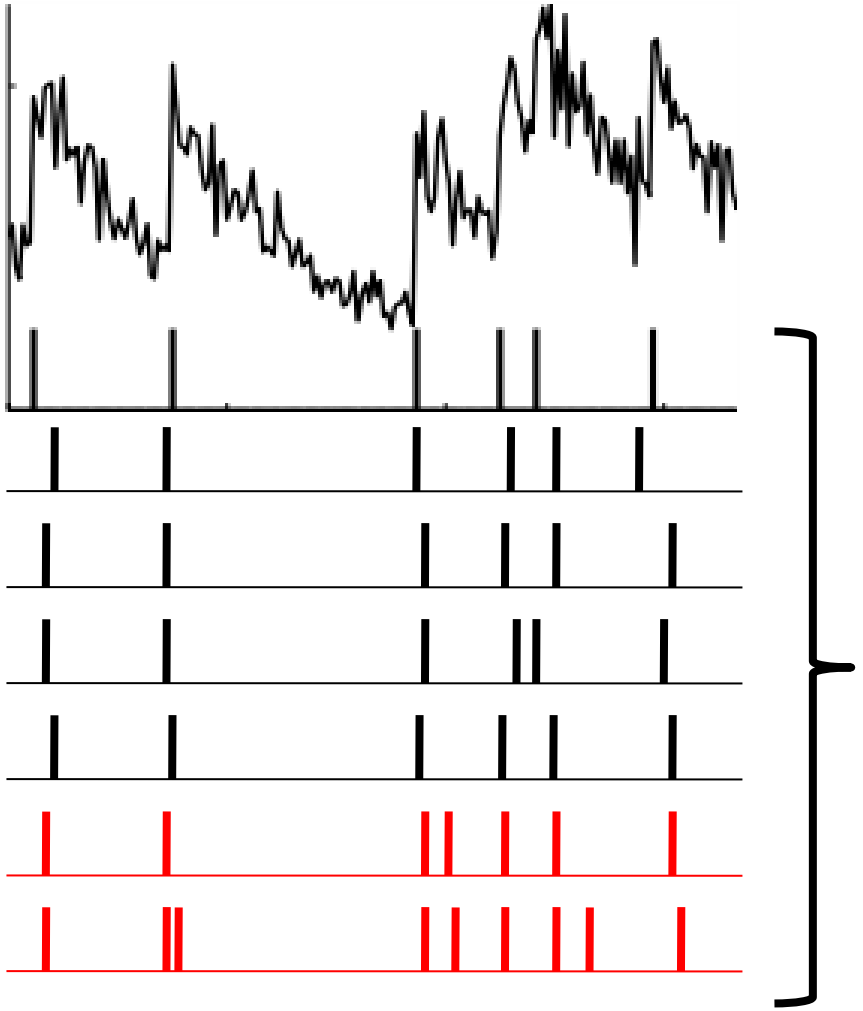
Corresponding GLM models, k- and h-terms



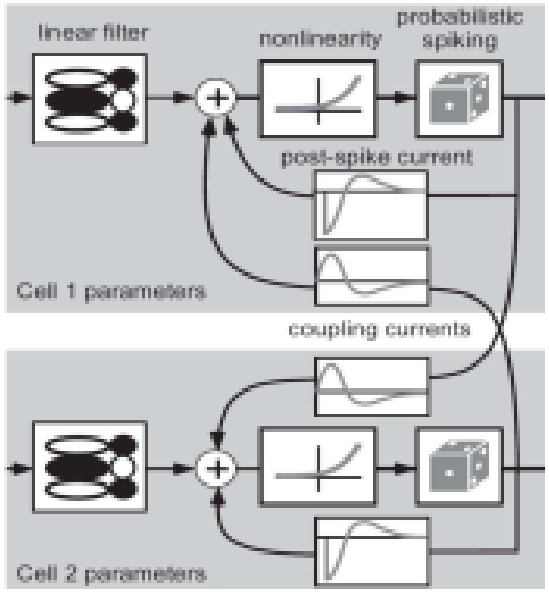
Czanner et al. "Measuring the signal-to-noise ratio of a neuron", PNAS'2015

Inferring neural connectivity from Ca imaging data





Neural population activity model - system of coupled GLM Markov processes:



$$s_1(t) \sim P[f(J_1(t))]$$

$$J_1(t) = b_1 + k_1 \cdot x(t) + \sum_{\tau>0} \sum_j w_{1j}(\tau) s_j(t-\tau)$$

$$s_2(t) \sim P[f(J_2(t))]$$

$$J_2(t) = b_2 + k_2 \cdot x(t) + \sum_{\tau>0} \sum_j w_{2j}(\tau) s_j(t-\tau)$$

$$s_3(t) \sim P[f(J_3(t))]$$

$$J_3(t) = b_3 + k_3 \cdot x(t) + \sum_{\tau>0} \sum_j w_{3j}(\tau) s_j(t-\tau)$$

…

… lots of neurons – lots of parameters :(

Fluorescence model:

$$C_i(t) \sim N[C_b^i + \exp(-\Delta t / \tau_i^c)(C_i(t) - C_b^i) + A_i s_i(t), (\sigma_i^c)^2 \Delta t]$$

$$F_i(t) = N[\alpha_i S(C_i(t)) + \beta_i, (\sigma_i^f)^2 + \gamma_i S(C_i(t))]$$

- Autoregressive model for calcium concentrations
- Ca jumps on each spike, then exponentially decays to baseline
- Fluorescence is related to instantaneous Ca concentration via a saturating fluorescence function $S(.)$

Model estimation (Bayesian MAP)

MAP estimator

$$w^{MAP} = \arg\max P(w, \theta \mid F)$$

$$\theta^{MAP} = \arg\max P(w, \theta \mid F)$$

Want this (a-posterior)

$$P(w, \theta \mid F) \propto P(F \mid w, \theta) P(w) P(\theta)$$

From this (generative)

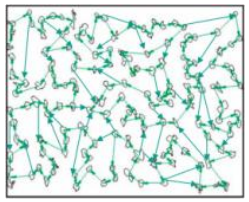
$$P(F \mid w, \theta) = \int dCdsP \quad (F \mid C) P(C \mid s, \theta) P(s \mid w)$$

Solution plan (also a rigorous statistical method for solving this kind of problems known as the Expectation Maximization algorithm)

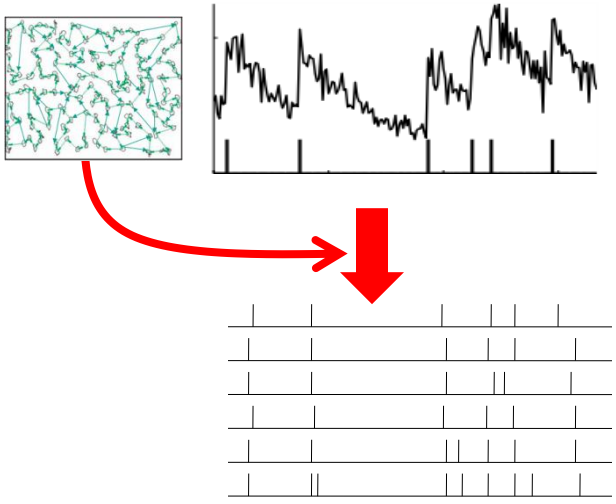
- Use a model of the neural network and individual cells' fluorescence to produce a large number of "plausible" spike trains consistent with the observed Ca fluorescence under that model
- Calculate the average likelihood of the actual observation over these spike trains
- Re-fit the neural network model by maximizing the average observation's likelihood
- Repeat

REPEAT IN A LOOP

Re-fit the network model and the Ca fluorescence model to maximize Q^*



Produce a large number of *possible* spike trains under the observed Ca fluorescence traces for all neurons



$$Q^* = L(F, s) + L(s)$$

Calculate the average likelihood of the spike trains and the Ca fluorescence observations

How this works in practice

– Obtain a sample from

$$\{C^{k}, s^{k}\} \sim P[C, s \mid F, w^{(l)}, \theta^{(l)}]$$

– Define $Q(w,\theta; w^{(l)},\theta^{(l)})$ by (sum k over that sample)

$$Q\left(w, \theta; w^{(l)}, \theta^{(l)}\right) = \sum_{k} \log \; P[F, C^{k}, s^{k} \mid w, \theta]$$

– Compute new parameters

$$(w^{(l+1)}, \theta^{(l+1)}) = \arg \max \; Q\left(w, \theta; w^{(l)}, \theta^{(l)}\right) + \log \; P\left(w, \theta\right)$$

The mathematics of EM algorithm

– Given an estimate for network and fluorescence model parameters (w, θ), calculate

$$Q\left(w, \theta; w^{(l)}, \theta^{(l)}\right) = E_{P[C, s|F, w^{(l)}, \theta^{(l)}]} \log P[F, C, s \mid w, \theta]$$

$$= \int dCds \; P[C, s \mid F, w^{(l)}, \theta^{(l)}] \log P[F, C, s \mid w, \theta]$$

– Obtain new estimate for network and fluorescence parameters (w, θ)

$$(w^{(l+1)}, \theta^{(l+1)}) = \arg \max Q\left(w, \theta; w^{(l)}, \theta^{(l)}\right) + \ln P(w, \theta)$$

- Each pass is guaranteed to increase

$$\delta \log P(w, \theta \mid F) \geq \delta Q(w', \theta' \mid w, \theta) + D(w', \theta' \| w, \theta)$$

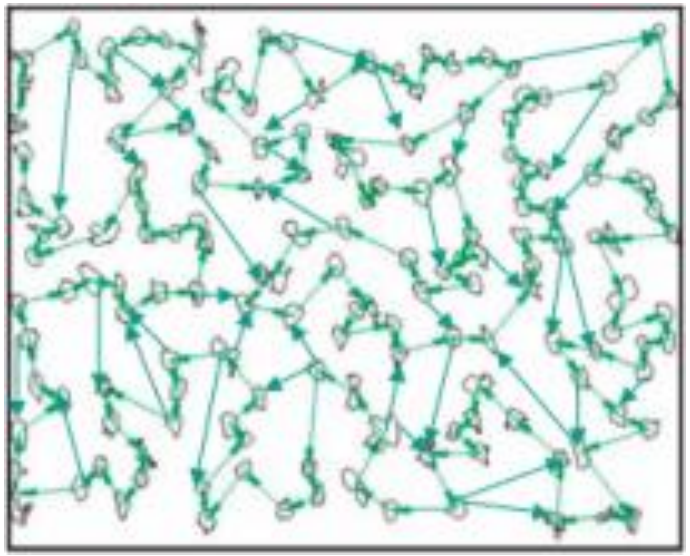
- Thus this is at worst locally, at best globally MAP estimate

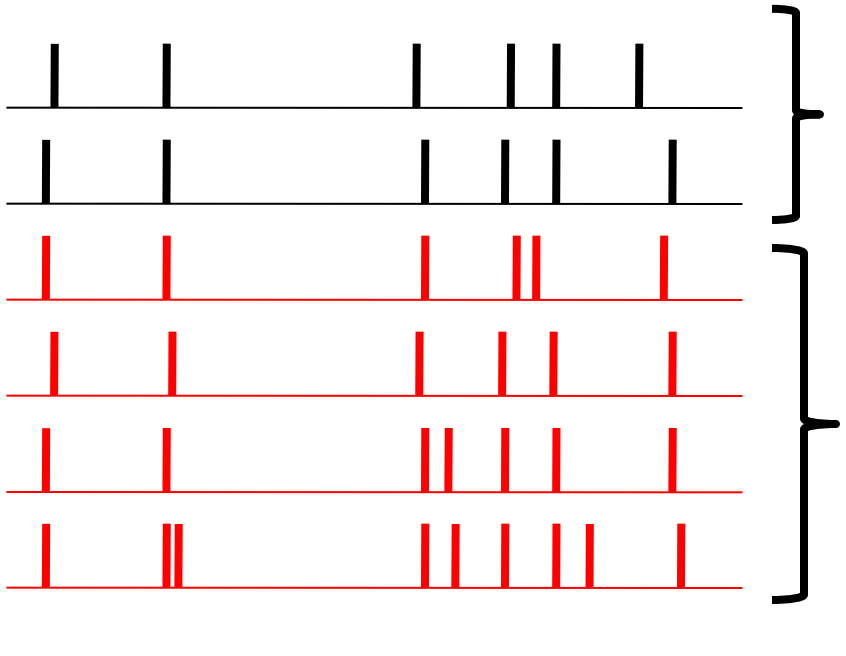
Estimation of effective connectivity from sparse neural activity samples

- The connectivity estimation problem from patchy observations of neural activity is in essence similar to that from Ca fluorescence observations
- In that case, either not all neurons are assumed to be linked to the fluorescence observations or the observed neurons take place of the Ca fluorescence as “observables” and the hidden neurons are averaged over, as before

To make life simpler, assume that observed neurons' spike trains are directly seen, that is, no Ca signal deconvolution is needed or the spikes had been extracted using a different method prior to this

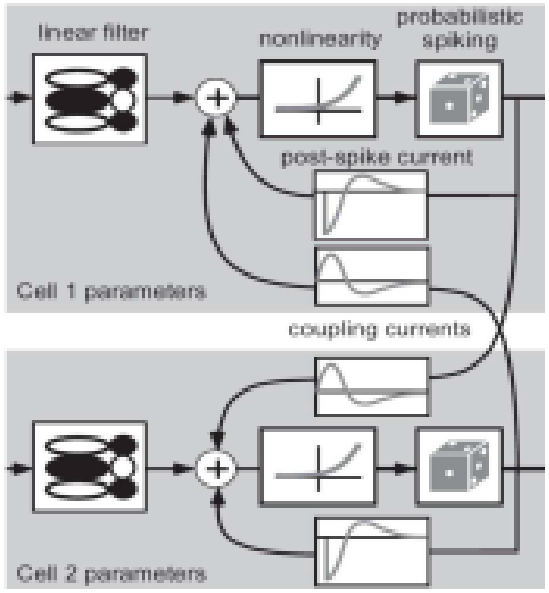
Inferring neural connectivity from sparse imaging





AVERAGE OVER !

Neural population activity model (same as before):



$$s_1(t) \sim P[f(J_1(t))]$$

$$J_1(t) = b_1 + k_1 \cdot x(t) + \sum_{\tau>0} \sum_{j} w_{1j}(\tau) s_j(t-\tau)$$

$$s_2(t) \sim P[f(J_2(t))]$$

$$J_2(t) = b_2 + k_2 \cdot x(t) + \sum_{\tau>0} \sum_{j} w_{2j}(\tau) s_j(t-\tau)$$

$$s_3(t) \sim P[f(J_3(t))]$$

$$J_3(t) = b_3 + k_3 \cdot x(t) + \sum_{\tau>0} \sum_{j} w_{3j}(\tau) s_j(t-\tau)$$

…

Model estimation (Bayesian MAP)

MAP estimator

$$w^{MAP} = \arg\max P(w \mid s^{observed})$$

Want this (a-posterior)

$$P(w \mid s^{observed}) \propto P(s^{observed} \mid w) P(w)$$

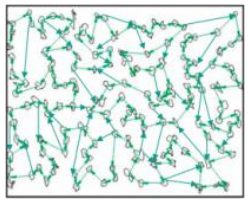
From this (generative)

$$P(s^{observed} \mid w) = \int ds^{hidden} P(s^{all} \mid w)$$

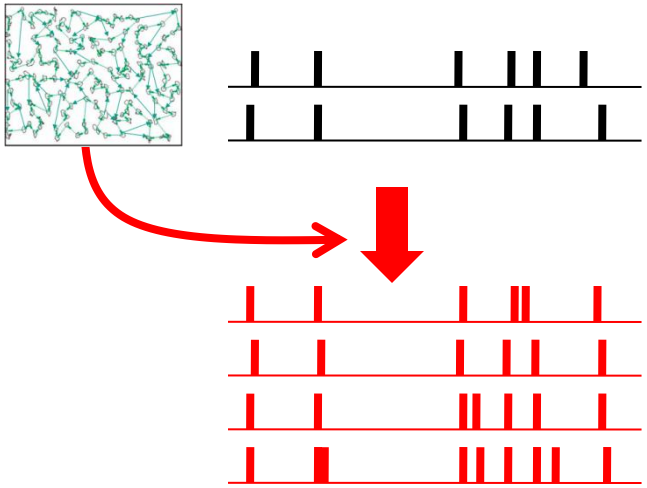
- Solution plan
  - Use a model of the neural network to produce a large number of “plausible” spike trains of hidden neurons, consistent with that of the observed neurons
  - Calculate the average likelihood of the actual observation over these spike trains
  - Re-fit the neural network model by maximizing the average observation likelihood

REPEAT IN A LOOP

Re-fit the network model to maximize Q^*



Produce a large number of *possible* spike trains of missing neurons



$$Q^* = L(s^{observed}, s^{hidden})$$

Calculate the likelihood of the actually seen activity (in the observed population), average over missing spike trains

How this works in practice

- Obtain a sample from

$$s^{hidden} \sim P[s^{hidden} \mid s^{observed}, w^{(l)}]$$

- Define $Q(w; w^{(l)})$ by (k-sum over that sample)

$$Q\left(w; w^{(l)}\right) = \sum_{k} \log P[s^{observed}, s^{hidden} \mid w]$$

- Compute new parameters

$$w^{(l+1)} = \arg\max Q\left(w; w^{(l)}\right) + \log P\left(w\right)$$

- Difficulties compared to Ca imaging case:
  - The vector of hidden neurons' activities is extremely high dim – a lot of missing neurons at a lot of intermediate time steps
  - The hidden activities may be poorly constrained by the observed neural activities – the distribution of the hidden activity samples is wide/dispersed
  - The set of neurons that are hidden changes from time to time

- The sampling step for missing activity is doable, although computationally challenging and demanding on the resources
- We can run the EM algorithm after that as before, and can reuse a lot of pieces from our Ca imaging analysis

In the “shotgun”-type neural activity imaging, the full observations’ log-likelihood becomes

$$\log \; P(s^{observed} \mid w) = \sum_{X} \log \; P(s^{X} \mid w)$$

where X are the different subsets of neurons included in the imaging at different times throughout experiment

For a single fixed segment of imaging, for a neural population

$$\log \; P(s^{observed} \mid w) \equiv \log \; P(s^{X} \mid w)$$

where the subset of the observed neurons X is fixed and constant at all times

Having the set $s^{observed}$ fixed makes the observations log-likelihood allow different models of hidden neurons' connectivity that can produce the same *observations*, mathematically this means that the MAP estimation gets multiple optima

$$\log P(s^{observed} \mid w) \equiv \log P(s^{X} \mid w) \rightarrow \text{multiple maxima}$$

However, new alternative plausible "explanation" models are added by incomplete observation, the true model is not erased from the data – it is still one of and among the multiple MAP optima (this is an important point!)

$$\log P(s^{observed} \mid w) \equiv \log P(s^{X} \mid w) \rightarrow \text{multiple maxima}$$

The above statement is another way of putting “the hidden inputs” or “the missing neurons” problem – unobserved neural populations do not break functional inference, *they make it ambiguous* (!)

$$\log P(s^{observed} \mid w) \equiv \log P(s^{X} \mid w) \rightarrow \text{multiple maxima}$$

For the “shotgun”-type neuronal activity imaging to be able to fully specify the complete connectivity matrix w, it is necessary that the collection of the imaged subsets of neurons X in

$$\log \; P(s^{observed} \mid w) = \sum_{X} \log \; P(s^{X} \mid w)$$

removes those multiple maxima from the log-likelihood function, leaving only the true maximum

The precise statement this becomes is:

- The collection of the subsets of neurons $\{ X \}$ covered by sparse neural activity imaging should be such that the set of marginal distributions

$$\left\{ \log \; P (s^{X} \mid w) \right\}$$

is distinct for each neural activity model w

In that case, the ambiguous likelihood maxima are removed from the observations' log-likelihood log $P(s^{observed}|w)$ and the neural population model w becomes again recoverable !

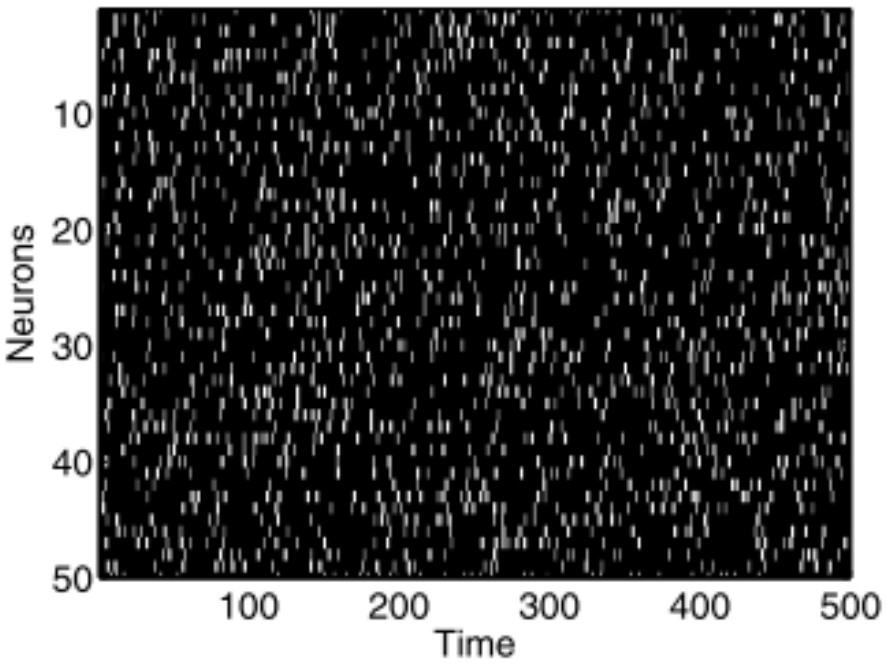
For several special cases, including the exponential spiking GLM used by Chichilnisky et al. in retina works, I explicitly show in my paper that the necessary set $\{ X \}$ is the set of all input-output plus same-time pair-wise neural activities, that is, $\{ P(s_i(t),s_j(t)|w) \}$ and $\{ P(s_i(t+1),s_j(t)|w) \}$

More generally, based on arguments known as the *implicit function theorem*, it appears that the set of N^2 pair-wise input-output neural activity distributions $\{ P(s_i(t+1),s_j(t)) \}$ is sufficient for recovering the unique connectivity in most network-type Markov models of neural population activity, as parameterized by the matrix of N^2 connection weights $\{ w_{ij} \}$

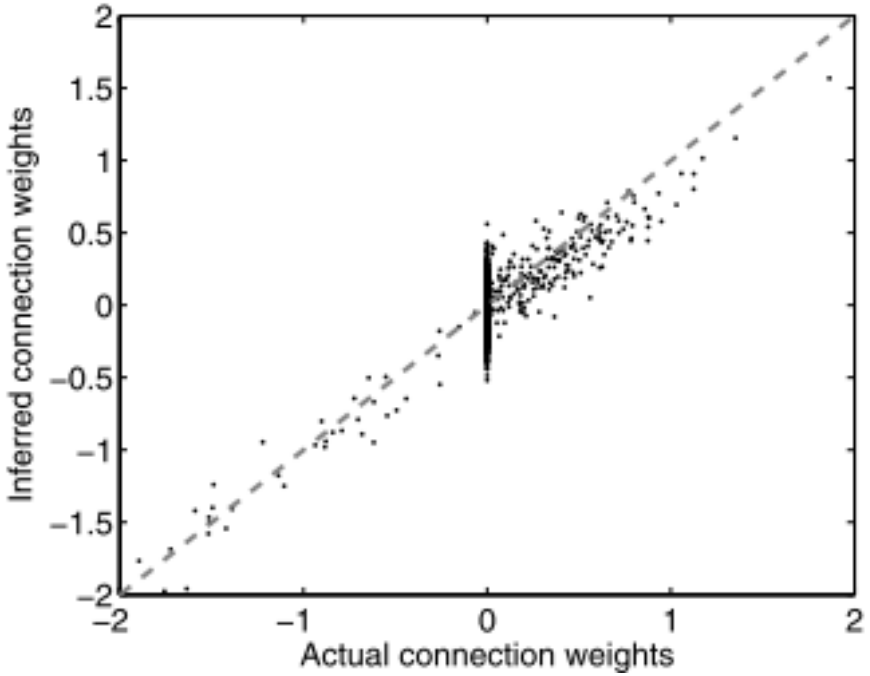
The bottom line for experimentalists: It is not necessary to produce continuous movies of entire neural populations' activity; observing the pairs of neurons in input-output configurations, randomly or deterministically no matter, will suffice for extracting the complete connectivity matrix of neural population

R^2=0.85

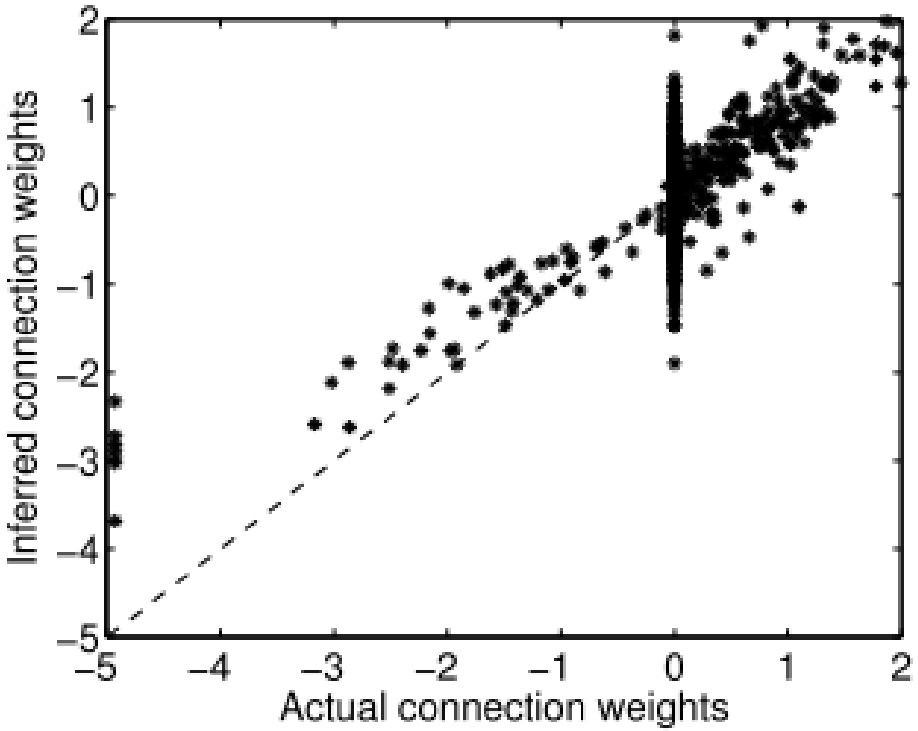
A snippet of neural activity raster



A scatter plot of inferred vs. true model connection weights



Neural connectivity matrix extracted from sparse double-block scanning data in a model cortical neural network



Supercomputing facilities in numbers

NERSC Computational Systems

System Name	System Type	CPU		Computational Pool						Node Interconnect	Scratch Disk
		Type	Speed (GHz)	Nodes	SMP Size	Total Cores	Flops per Core (Gflops/sec)	Peak Performance (Tflops/sec)	Aggregate Memory	Avg. Memory/core	
Edison	Cray XC30	Intel Ivy Bridge	2.4	5,576	24	133,824	19.2	2569.4	357 TB	2.67	Aries	7.56 PB (local) + 3.9 PB (global)
Hopper	Cray XE6	Opteron	2.1	6,384	24	153,216	8.4	1287.0	211.5 TB	1.41 GB	Gemini	2.2 PB (local) + 3.9 PB (global)
PDSF[1)]	Linux Cluster	Opteron, Xeon	2.0, 2.27, 2.33, 2.67	232	8, 12, 16	2,632	8.0, 9.08, 9.32, 10.68	17.6	9.5 TB	4 GB	Ethernet / InfiniBand	34.9 TB for batch nodes and 184 GB for interactive nodes
Genpool[2)]	Various vendor systems	Nehalem, Opteron	2.27, 2.67	547	8, 24, 32, 80	4,680	9.08, 10.68	42.8	33.7 TB	7.36 GB	Ethernet	3.9 PB (global)

Supercomputing facilities in numbers

Magerit/Blue Brain

Arquitectura POWER

Tiempo de espera medio

# Procesadores:	3.920 cores
Memoria:	7.840 GB
Potencia:	103,50 TFLOPS Rpeak
	72,03 TFLOPS Rmax

Arquitectura Intel

Tiempo de espera bajo

# Procesadores:	656 cores
Memoria:	2.624 GB
Potencia:	13,64 TFLOPS Rpeak
	12,69 TFLOPS Rmax