Transfer Entropy reconstruction of neuronal networks from calcium imaging data

Javier G. Orlandi
Dept. ECM, University of Barcelona
June 17, 2014

Universitat de Barcelona

CIDNET14
16-20 June 2014

reconstruction

networks

data

Why now?

BIG DATA

We are recording more data than ever before

Probably more than we can analyse

Economics
Climate research
High energy physics

Generating huge amounts of data
Some of it publicly available
Leading to new patterns and emerging behaviours

reconstruction

networks

data

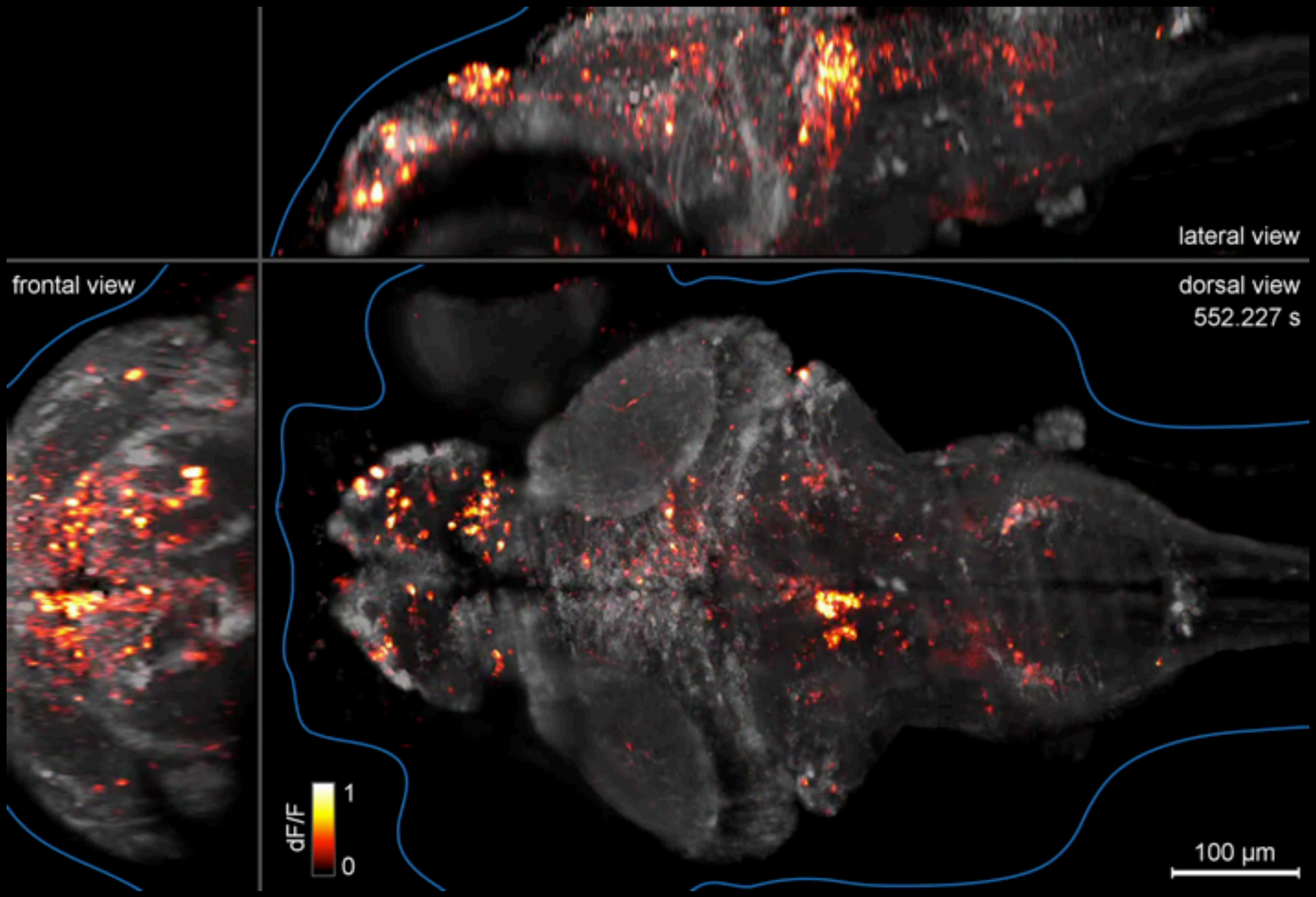
2013

THE BLUE BRAIN PROJECT EPFL

neuroscience is in everyone’s mind

The Human Connectome Project

Whole-brain functional imaging



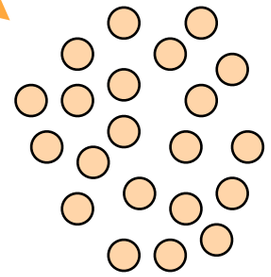
Ahrens et al, Nat Met **2013**

reconstruction of neuronal networks from imaging data

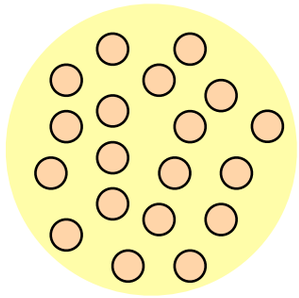
We still need to build (more) models and tools to infer connectivity in neuroscience
Well controlled experimental setup: Neuronal Cultures

Neuronal Cultures

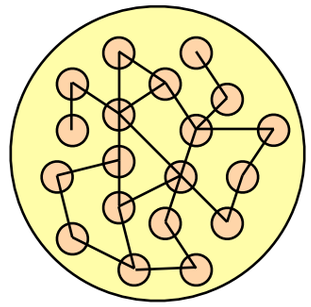
cortical tissue
17-19 days embryo



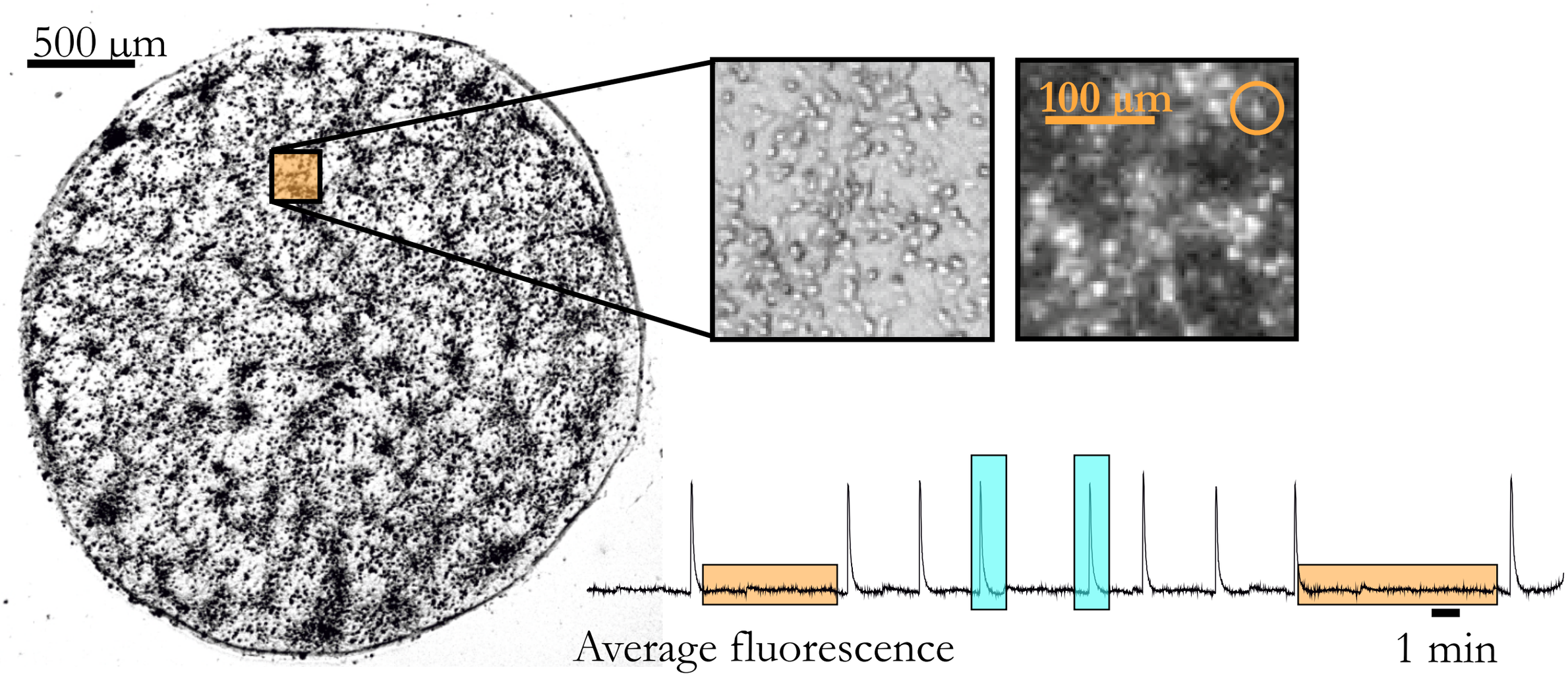
neuron
dissociation



plating
(day 0)

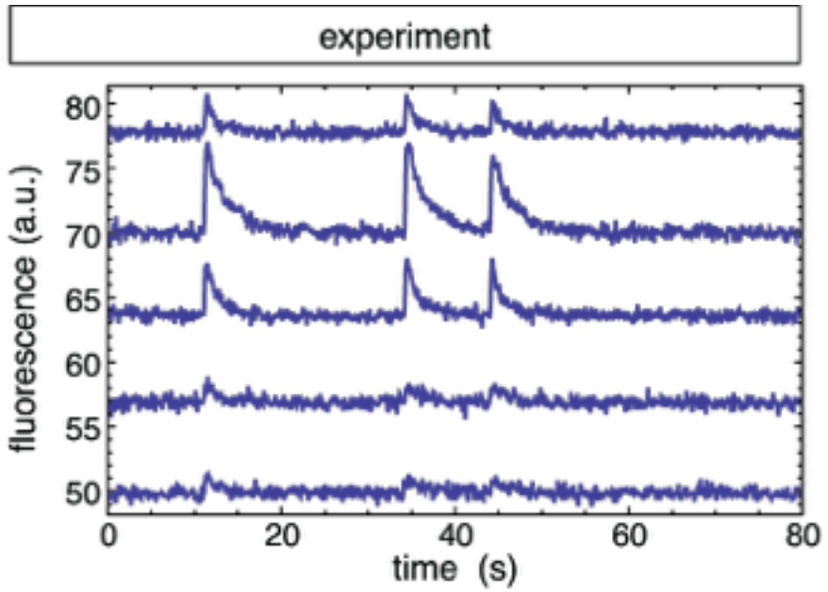


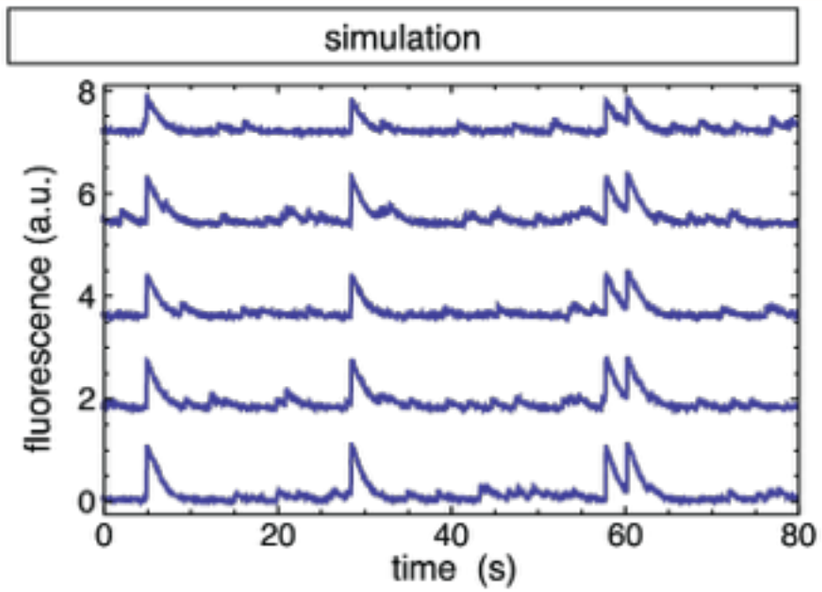
“mature” culture
(day 14)

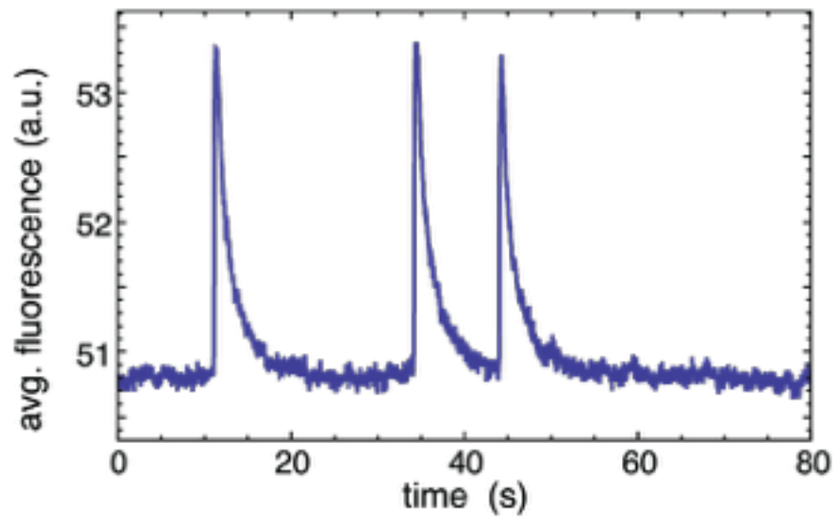


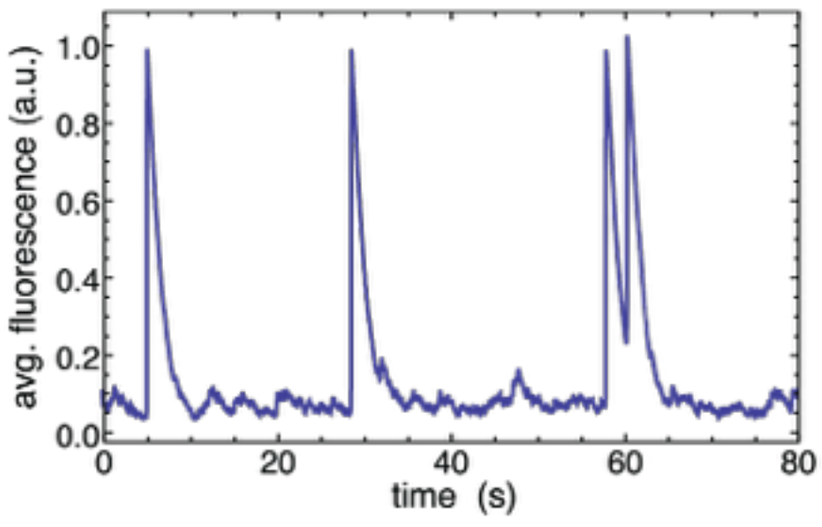
Orlandi et al, Nat Phys **2013**

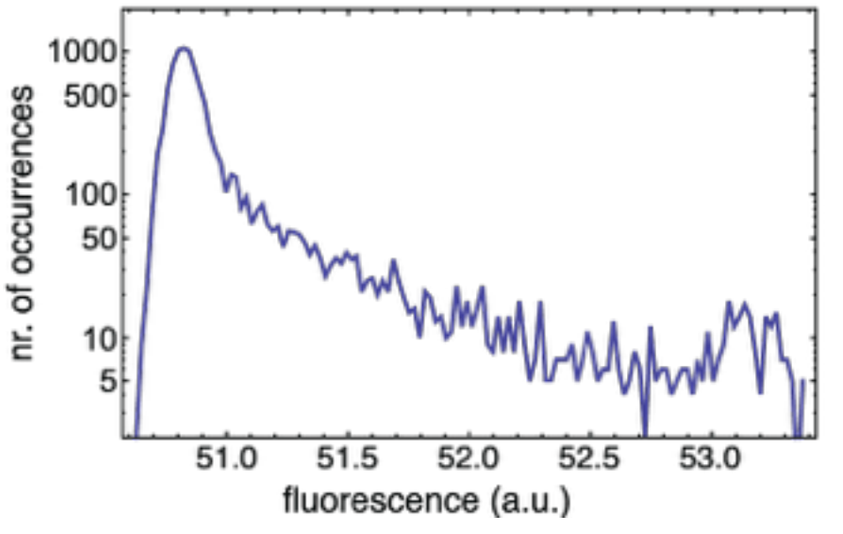
Population and single cell activity

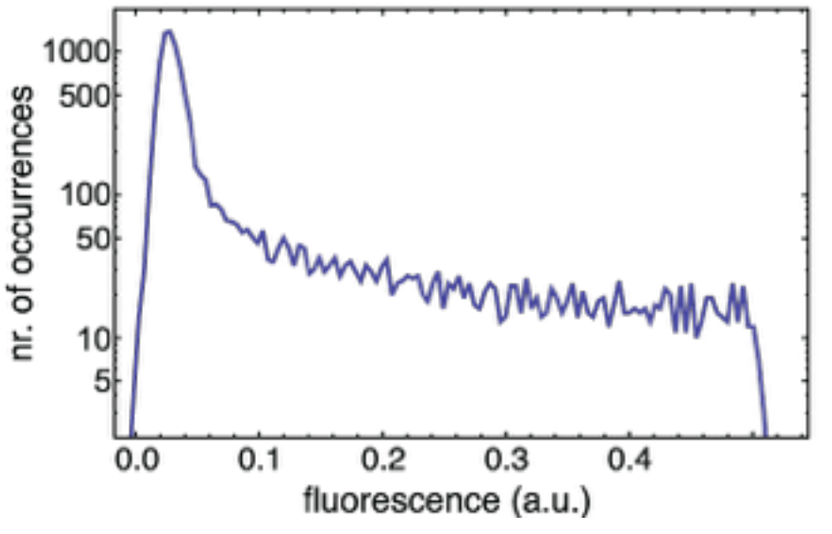












Model

Integrate and fire neurons (only excitatory)

Short term synaptic depression

Clustered and/or local connectivity

Calcium dynamics

Light Scattering effects

Stetter et al, PLOS Comput Biol **2012**

Transfer Entropy

$$TE_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log \frac{p(i_{n+1}|i_n^{(k)}, j_n^{(l)})}{p(i_{n+1}|i_n^{(k)})}$$

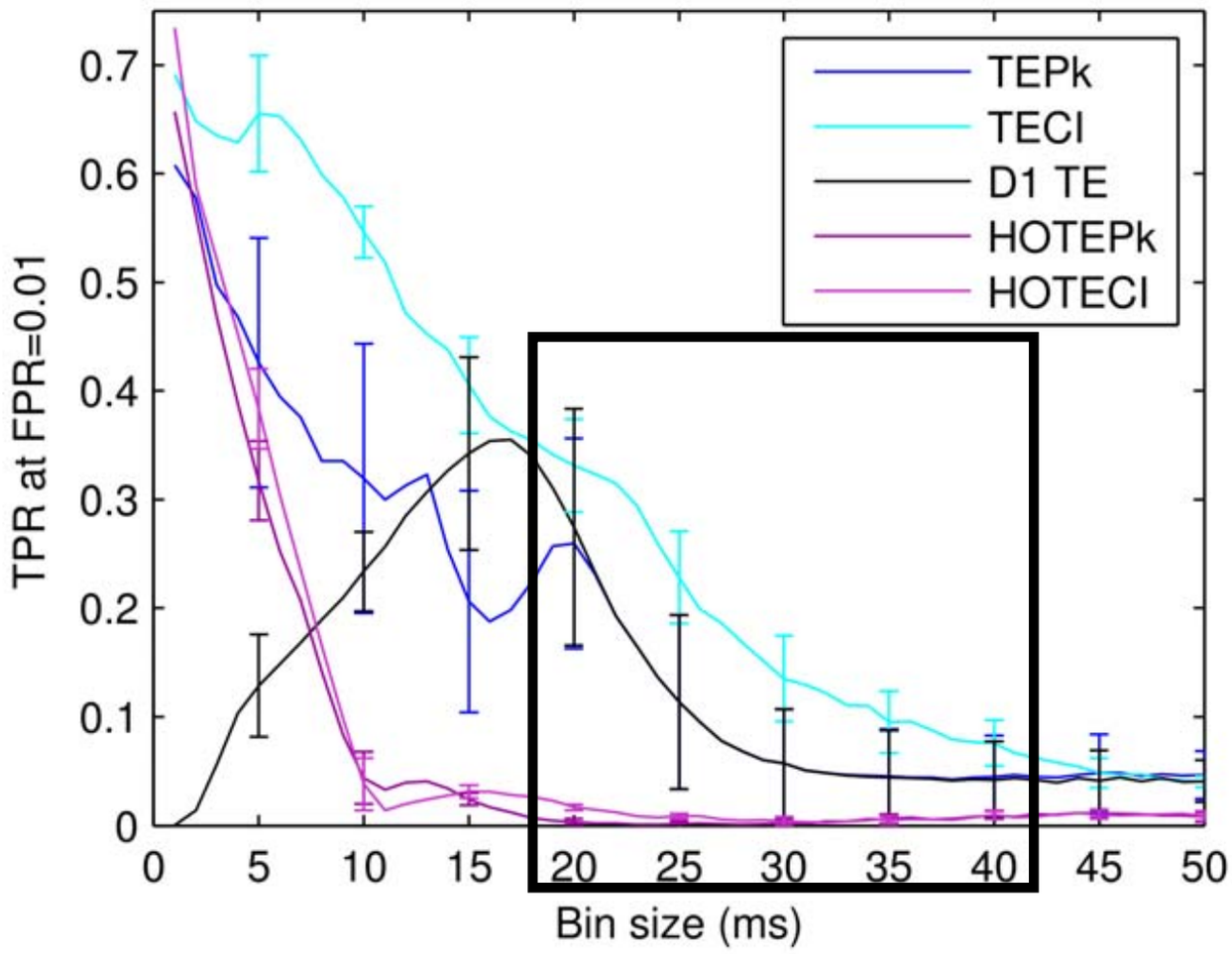
Schreiber, PRL **2000**

From J. Beggs' talk yesterday…

Non-bursting regime
Spiking data
Bin size dependence

Now

Bursts
Fluorescence data
Low temporal resolution



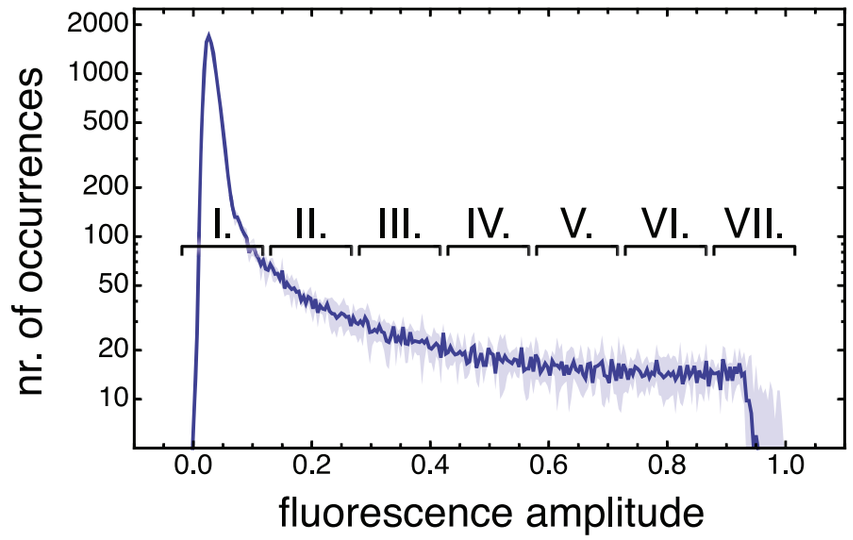
Ito et al, PLOS ONE **2011**

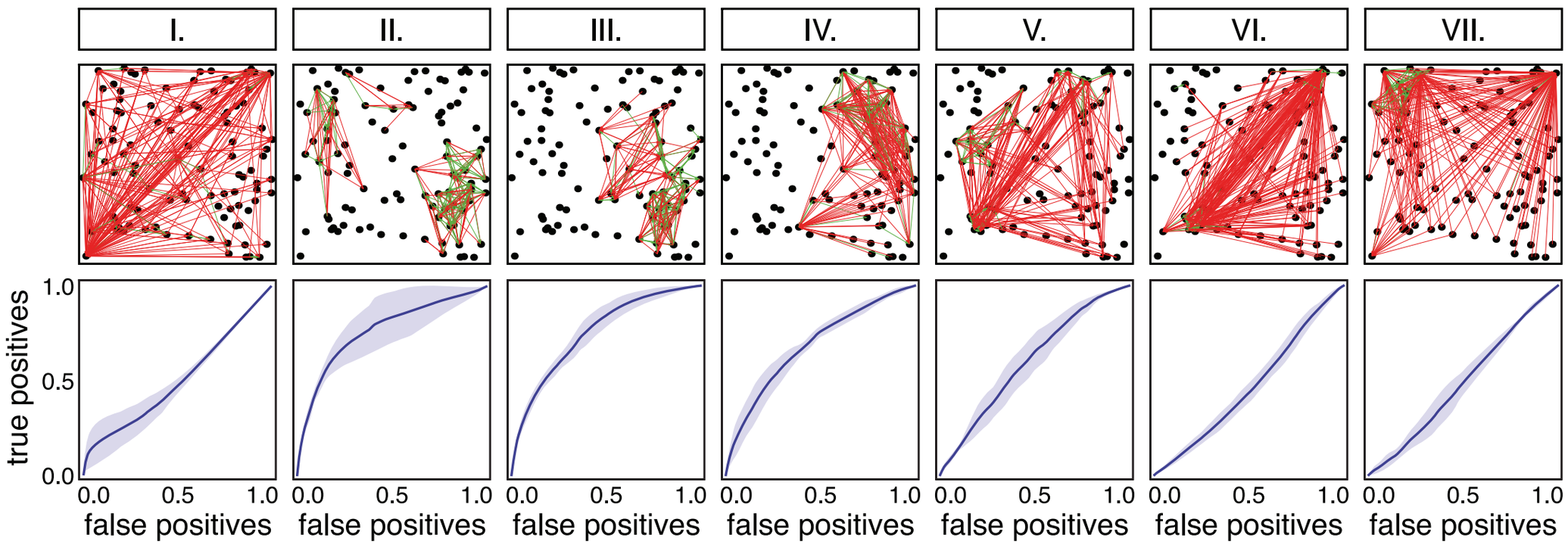
Generalizing Transfer Entropy

$$GTE_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, {j_{n+1}}^{(l)} | g_{n+1} \in S) \log \frac{p(i_{n+1} | i_n^{(k)}, {j_{n+1}}^{(l)}, g_{n+1} \in S)}{p(i_{n+1} | i_n^{(k)}, g_{n+1} \in S)}$$

Conditioning on the global population (state selection)

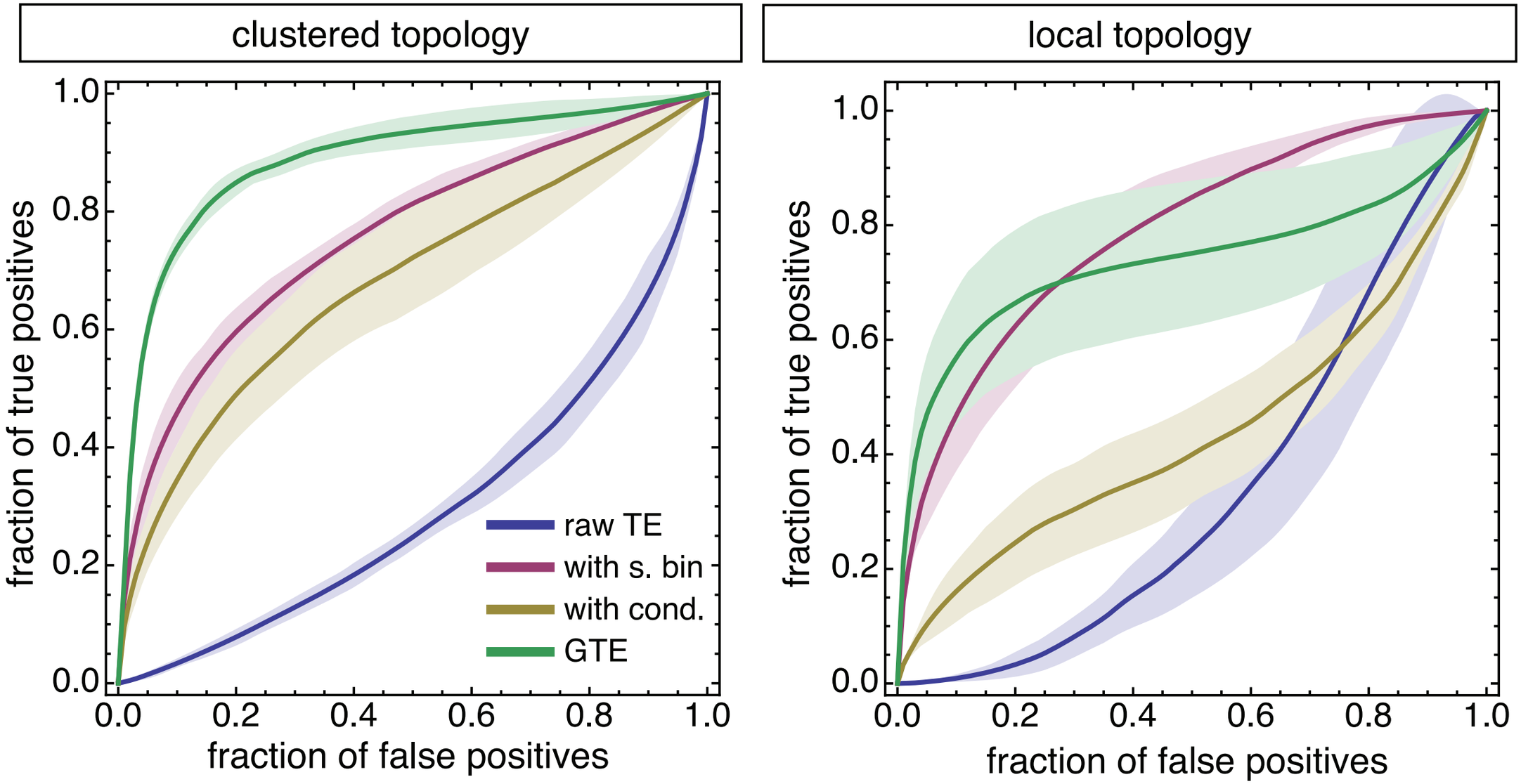
Same-bin (instantaneous) interactions





Stetter et al, PLOS Comput Biol **2012**

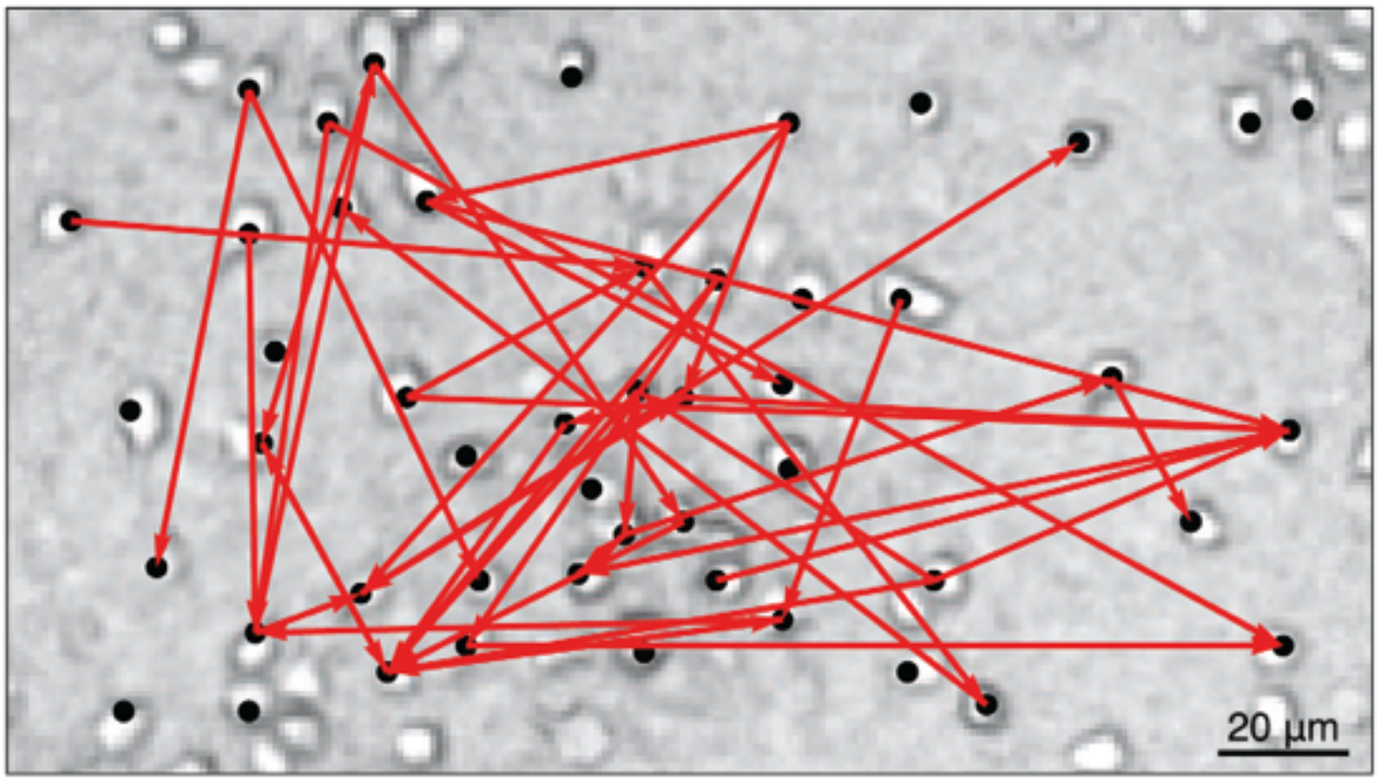
Improvements on raw TE

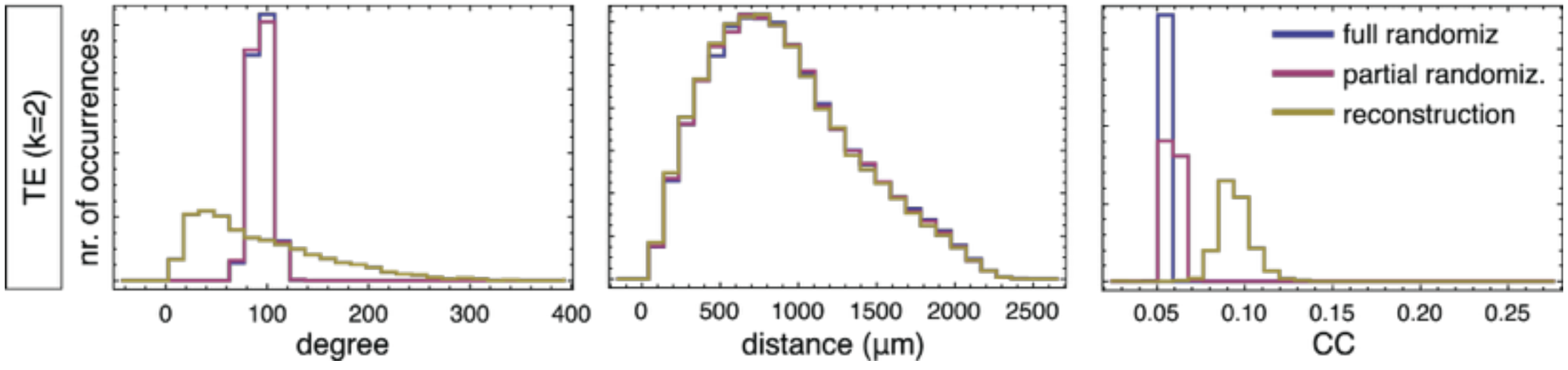


Stetter et al, PLOS Comput Biol **2012**

After many more (statistical) tests… check with the experiments

Reconstruction on experimental data

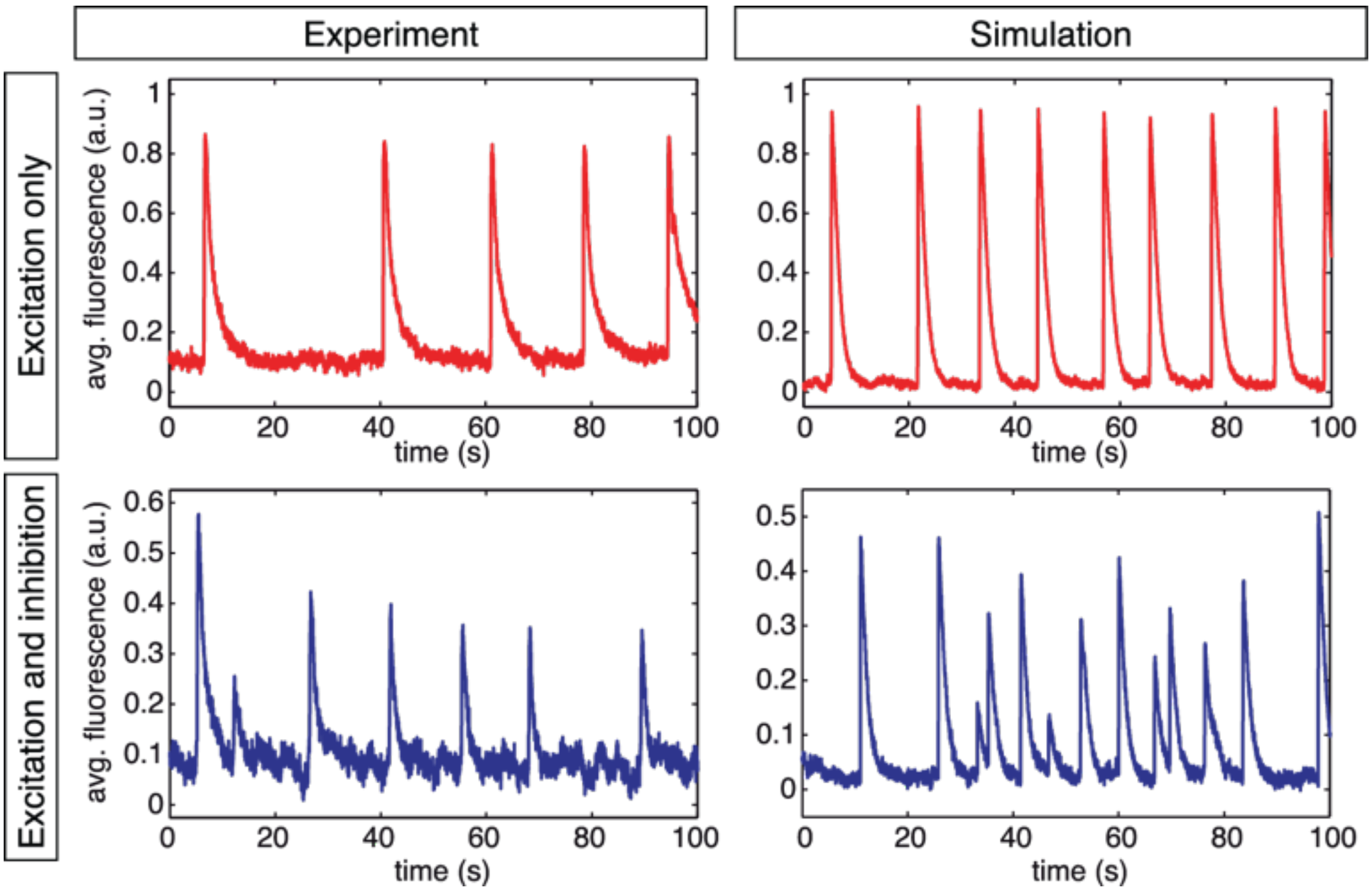




Stetter et al, PLOS Comput Biol **2012**

We are still missing the "truth"

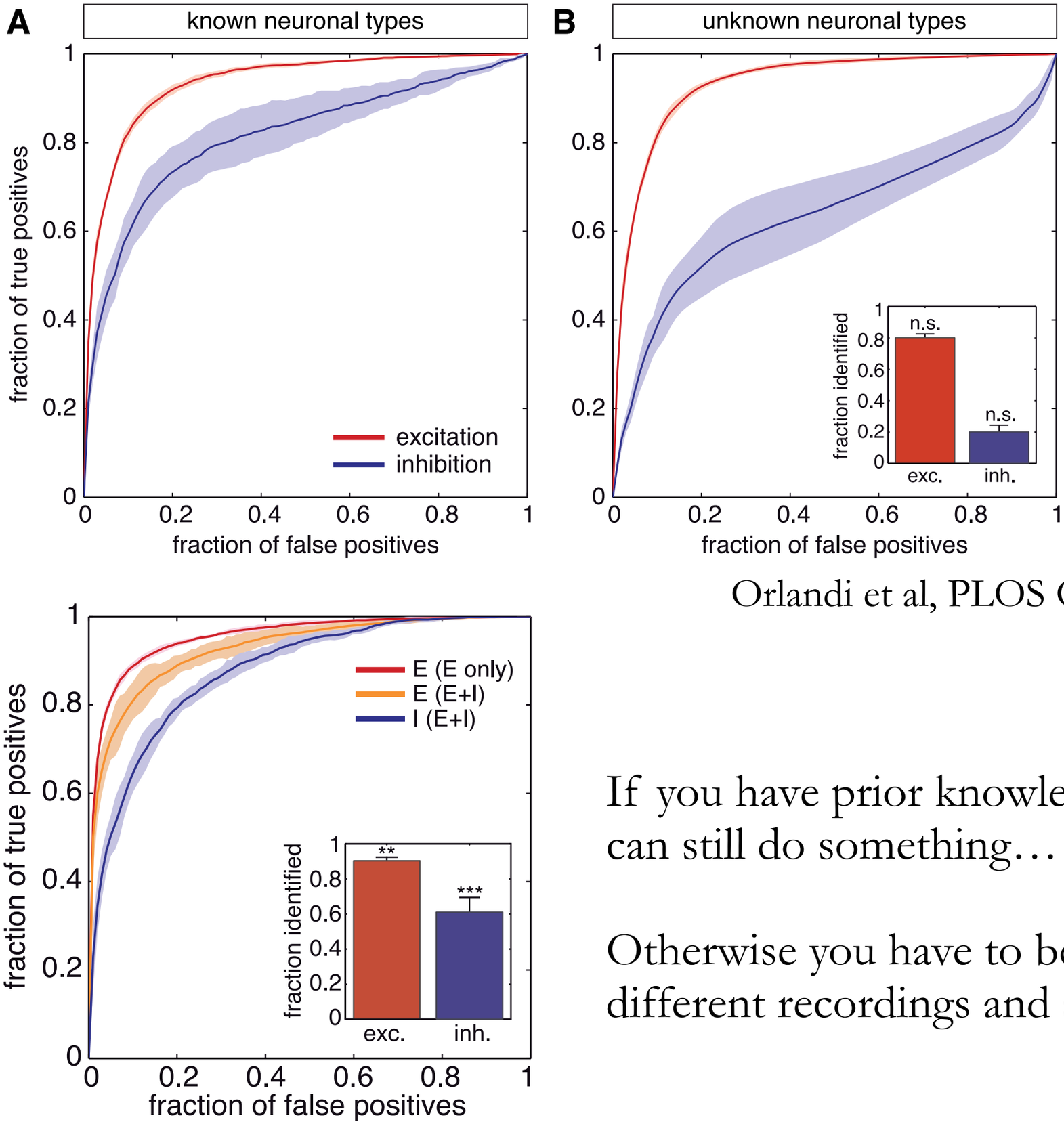
What about inhibition?



Orlandi et al, PLOS ONE **2014**

I should talk about this…
I won't (very short)

What about inhibition?



Orlandi et al, PLOS ONE **2014**

If you have prior knowledge of neuronal type you can still do something…

Otherwise you have to be creative (combine different recordings and stimulation protocols)

Yesterday....

From M. Eichler's talk…

Be very wary about using single methods…

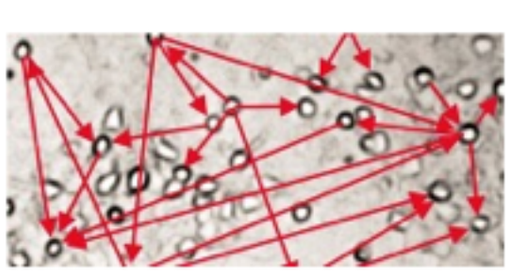
What to use in practice?
For causal inference… all and more

A different approach

Challenges in Machine Learning (collaborative competitions)

https://www.kaggle.com/c/connectomics/

kaggle | Customer Solutions | Competitions | Community | Javier G. Orlandi | Logout



CONNECTOMICS

Finished

Wednesday, February 5, 2014 — $3,000 • 144 teams — Monday, May 5, 2014

Dashboard

- Home
  - Data
  - Make a submission
- Information
  - Description
  - Evaluation
  - Rules
  - Prizes
  - Timeline
- Forum

Competition Details » Get the Data » Make a submission

Reconstruct the wiring between neurons from fluorescence imaging of neural activity

Understanding the brain structure and some of its disease alterations is key to research on the treatment of epilepsy, Alzheimer's disease, and other neuropathologies, as well as understanding the general function of the brain and its learning capabilities. The brain contains nearly 100 billion neurons with an average

First connectomics challenge

https://www.kaggle.com/c/connectomics/data

Deep CNN for Time Series Correlation Measurement
28 days ago

Fast Matlab code to get a score of .93985
30 days ago

144 teams

185 players

1459 entries

What's in the data?

Validation and test data include time series of activities of neurons "extracted" from simulated calcium fluorescence imaging data. The neurons are arranged on a flat surface simulating a neural culture and you get the coordinates of each neuron. Th model takes into account light scattering effects. [Learn more...]

For "training" data, the network connectivity is also provided.

What am I predicting?

The network connectivity for the validation (valid) and test data.

How do I get started?

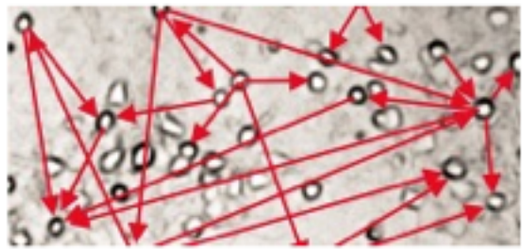
We provide a sample submission corresponding to the Correlation benchmark, sample code and tutorial material. See also our frequently asked questions.

File types

- **"fluorescence" files (type F):** These are the time series of neural activities obtained from fluorescence signals. The neurons are in columns and the rows are time ordered samples. The signals are sampled at 20ms intervals.
- **"networkPosition" files (type P):** Each row represents a neuron. First column = X position; second column = Y position. The neurons span a $1mm^2$ square area.

AUC-based scoring

https://www.kaggle.com/c/connectomics/leaderboard/private



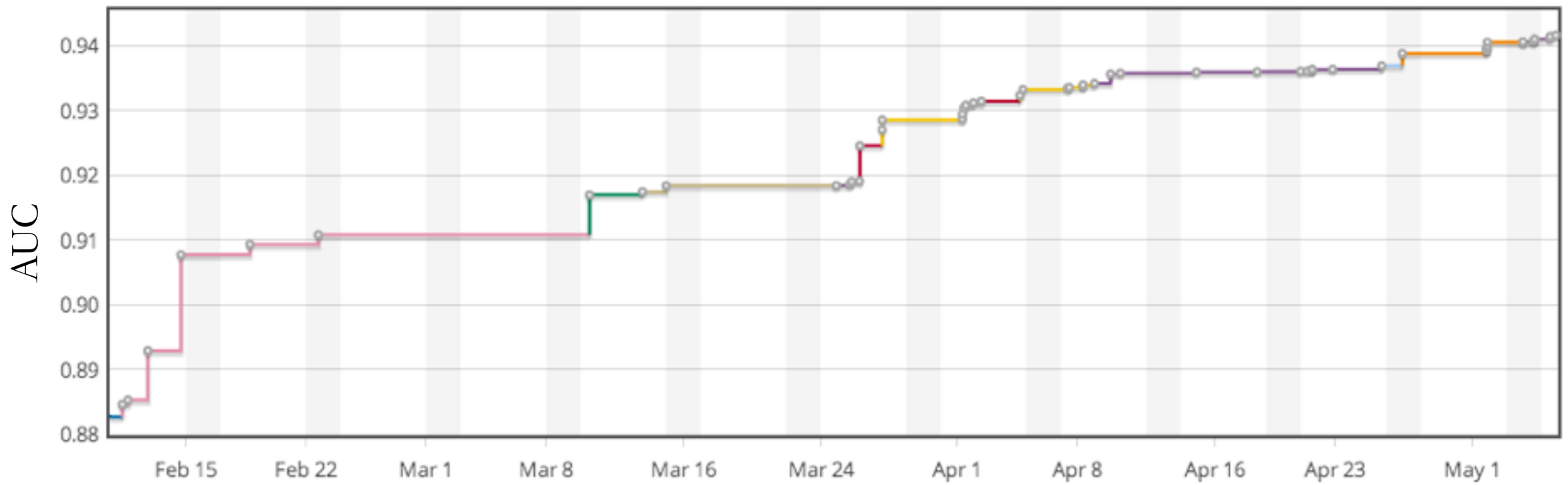
CONNECTOMICS

Finished

Wednesday, February 5, 2014 — $3,000 • 144 teams — Monday, May 5, 2014

Dashboard

Private Leaderboard - CONNECTOMICS



Rankings

#	Δ1w	Team Name ‡ model uploaded * in the money	Score	Entries	Last Submission UTC (Best – Last Submission)
1	↑6	AAAGV ‡ *	0.94161	145	Mon, 05 May 2014 20:50:39
2	↑3	Matthias Ossadnik *	0.94102	71	Mon, 05 May 2014 19:27:34 (-0h)
3	↓2	Ildefons Magrans ‡ *	0.94063	21	Sun, 04 May 2014 13:47:26 (-0.3h)
4	↓2	Lukasz 8000	0.93956	101	Mon, 05 May 2014 15:54:48 (-18.1h)
5	↓2	Lejlot & Rafal	0.93826	97	Mon, 05 May 2014 15:39:39 (-20.9h)
6	—	Sium	0.93711	28	Mon, 05 May 2014 21:45:26 (-0.2h)
7	↓3	Alexander N & vopern	0.93666	7	Sat, 03 May 2014 23:46:07
8	—	gaucho81	0.93385	43	Sun, 04 May 2014 21:48:37 (-30.2h)
9	↑8	killertom	0.93011	13	Mon, 05 May 2014 12:50:08
10	new	dhanson	0.92885	6	Mon, 05 May 2014 21:02:20
11	↓2	DJMN	0.92609	20	Mon, 05 May 2014 19:44:13
12	↓2	Gideon & Alex	0.92420	48	Mon, 05 May 2014 16:16:57 (-19.7h)
13	↓1	Sandro	0.92039	18	Mon, 05 May 2014 23:15:48 (-8.7h)
14	↑5	Selfish Gene	0.92039	20	Sun, 04 May 2014 16:38:44
15	↑7	Nitai Dean	0.91945	3	Mon, 05 May 2014 12:43:11

Post-challenge

Analysis of methods used by the participants:
Deep convolutional neural networks
State selection
Multivariate logistic regression of inferred spike trains
Inverse covariance matrix
Random forests and gradient boosting machines
Network deconvolution
…

Checking for robustness of the methods

Code from the winning teams will be publicly available with open sources licenses (and all the data associated with the challenge) http://www.kaggle.com/c/connectomics

Acknowledgments (and references)

Universitat de Barcelona


Jordi


Sara


Jaume


Olav


Demian


Theo

Max Planck Institute for Dynamics and Self-Organization

Isabelle Guyon, ChaLearn, California, USA
Alexander Statnikov, New York University, Langome Medical Center, New-York, USA
Bisakha Ray, New York University, Langome Medical Center, New-York, USA
Mehreen Saeed, Department of Computer Science, FAST, National University of Computer & Emerging Sciences, Lahore Campus, Pakistan

Model-Free Reconstruction of Excitatory Neuronal Connectivity from Calcium Imaging Signals

O. Stetter, D. Battaglia, J. Soriano, T. Geisel - PLOS Comput Biol 8(8): e1002653 (2012)

Transfer Entropy Reconstruction and Labeling of Neuronal Connections from Simulated Calcium Imaging

J. G. Orlandi, O. Stetter, J. Soriano, T. Geisel, D. Battaglia - PLOS ONE 9(6): e98842 (2014)