Connectivity, common input, and calcium

Liam Paninski

Department of Statistics
Center for Theoretical Neuroscience
Grossman Center for the Statistics of Mind
Columbia University
http://www.stat.columbia.edu/∼liam
liam@stat.columbia.edu
July 28, 2013

— with Y. Ahmadian, J. Huggins, S. Keshri, T. Machado, T. Paige, A. Pakman, D. Pfau, E. Pnevmatikakis, B. Shababo, C. Smith, J. Vogelstein

Support: NIH/NSF CRCNS, Sloan, NSF CAREER, DARPA, ARO, McKnight, Swartz.

A golden age of statistical neuroscience

Some notable recent developments:

- machine learning / statistics / optimization methods for extracting information from high-dimensional data in a computationally-tractable, systematic fashion
- computing (Moore's law, massive parallel computing)
- optical and optogenetic methods for recording from and perturbing neuronal populations, at multiple scales
- large-scale, high-density multielectrode recordings
- growing acceptance that many fundamental neuroscience questions are in fact statistics questions in disguise

A few grand challenges

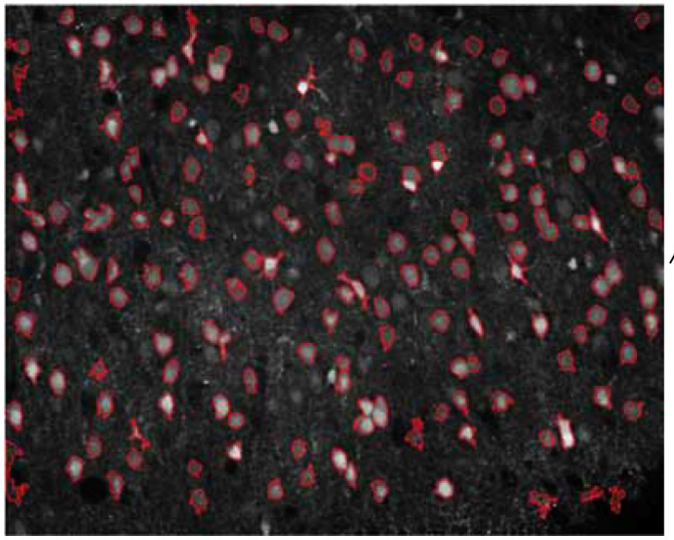
- Optimal decoding and dimensionality reduction of large-scale multineuronal spike train data
- Circuit inference from multineuronal spike train data
- Optimal control of spike timing in large neuronal populations
- Hierarchical nonlinear models for encoding information in neuronal populations
- Robust, expressive brain-machine interfaces; brain reading and writing
- Understanding dendritic computation and location-dependent synaptic plasticity via optical imaging (statistical spatiotemporal signal processing on trees)

A few grand challenges

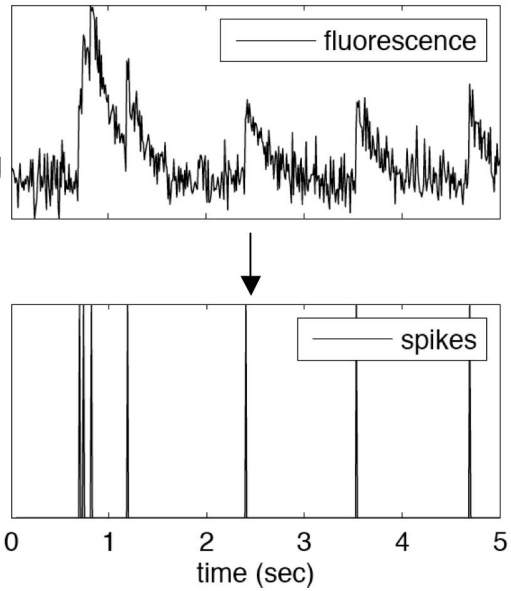
- Optimal decoding and dimensionality reduction of large-scale multineuronal spike train data
- Circuit inference from multineuronal spike train data
- Optimal control of spike timing in large neuronal populations
- Hierarchical nonlinear models for encoding information in neuronal populations
- Robust, expressive brain-machine interfaces; brain reading and writing
- Understanding dendritic computation and location-dependent synaptic plasticity via optical imaging (statistical spatiotemporal signal processing on trees)

Circuit inference via calcium imaging

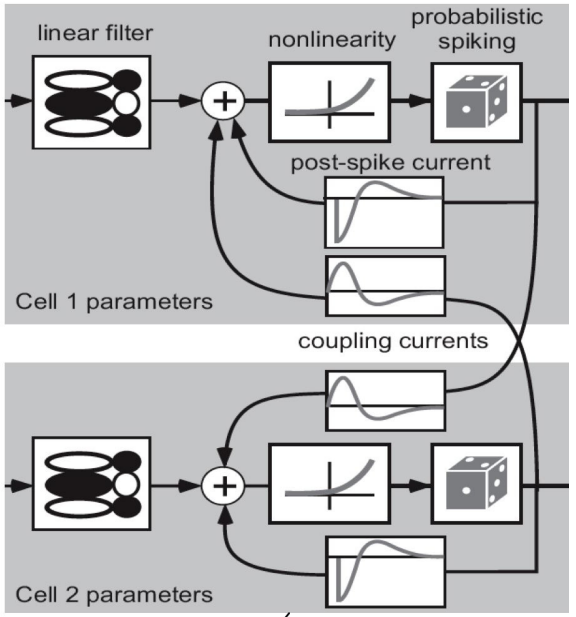
Record large-scale calcium movie



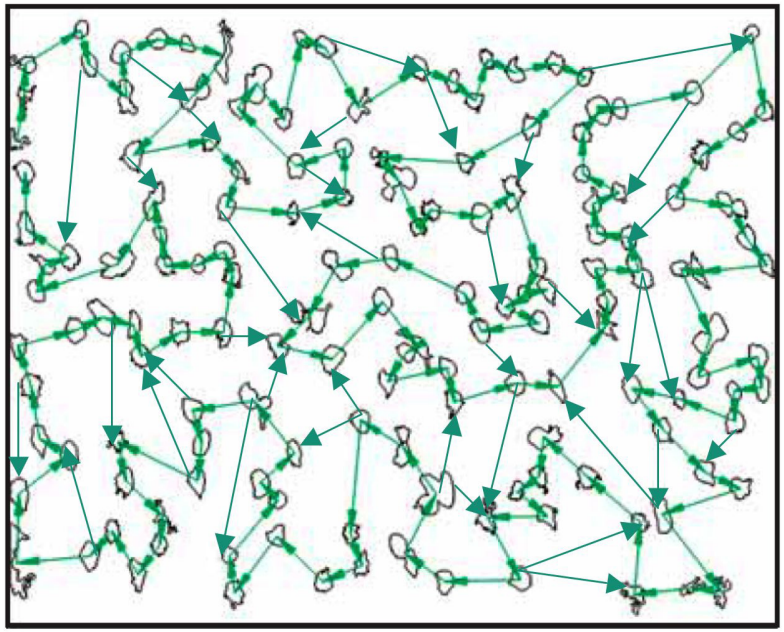
Aim 1: Extract spike times



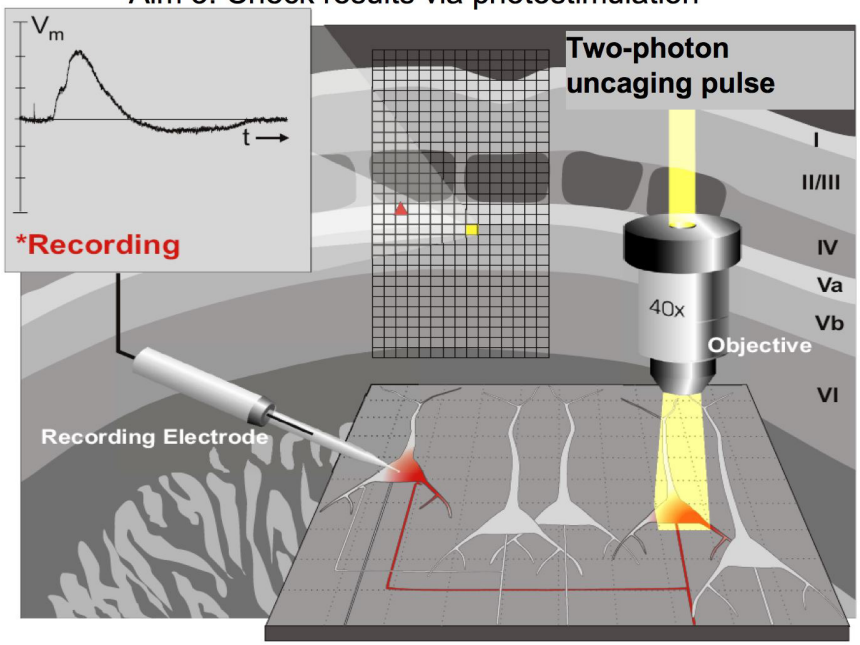
Aim 2: Estimate network model



Inferred network model

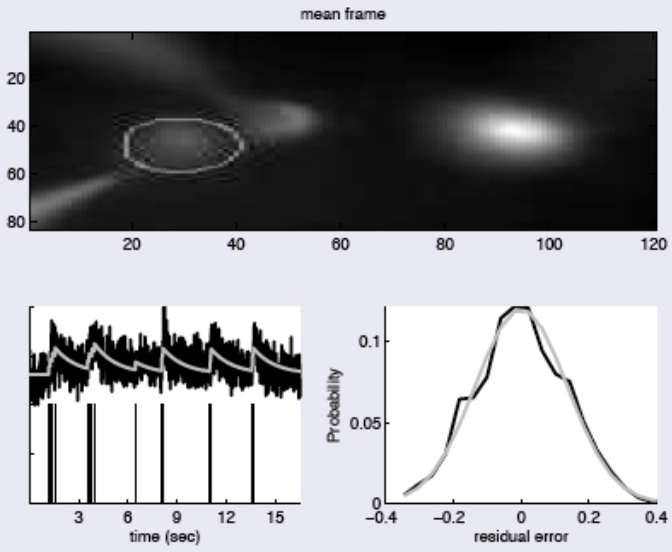


Aim 3: Check results via photostimulation

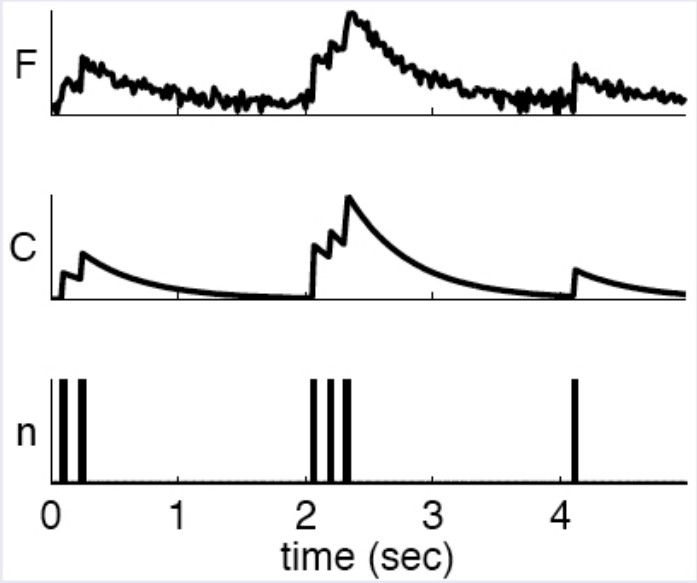


Aim 1: Model-based estimation of spike rates

data



schematic



equations

$$F_t = \alpha C_t + \beta + \sigma\varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0,1)$$

$$C_t = -(1 - \Delta/\tau)C_{t-1} + n_t$$

$$n_t \sim \text{poisson}(\lambda\Delta)$$

Note: each component here can be generalized easily.

Fast maximum a posteriori (MAP) estimation

Start by writing out the posterior:

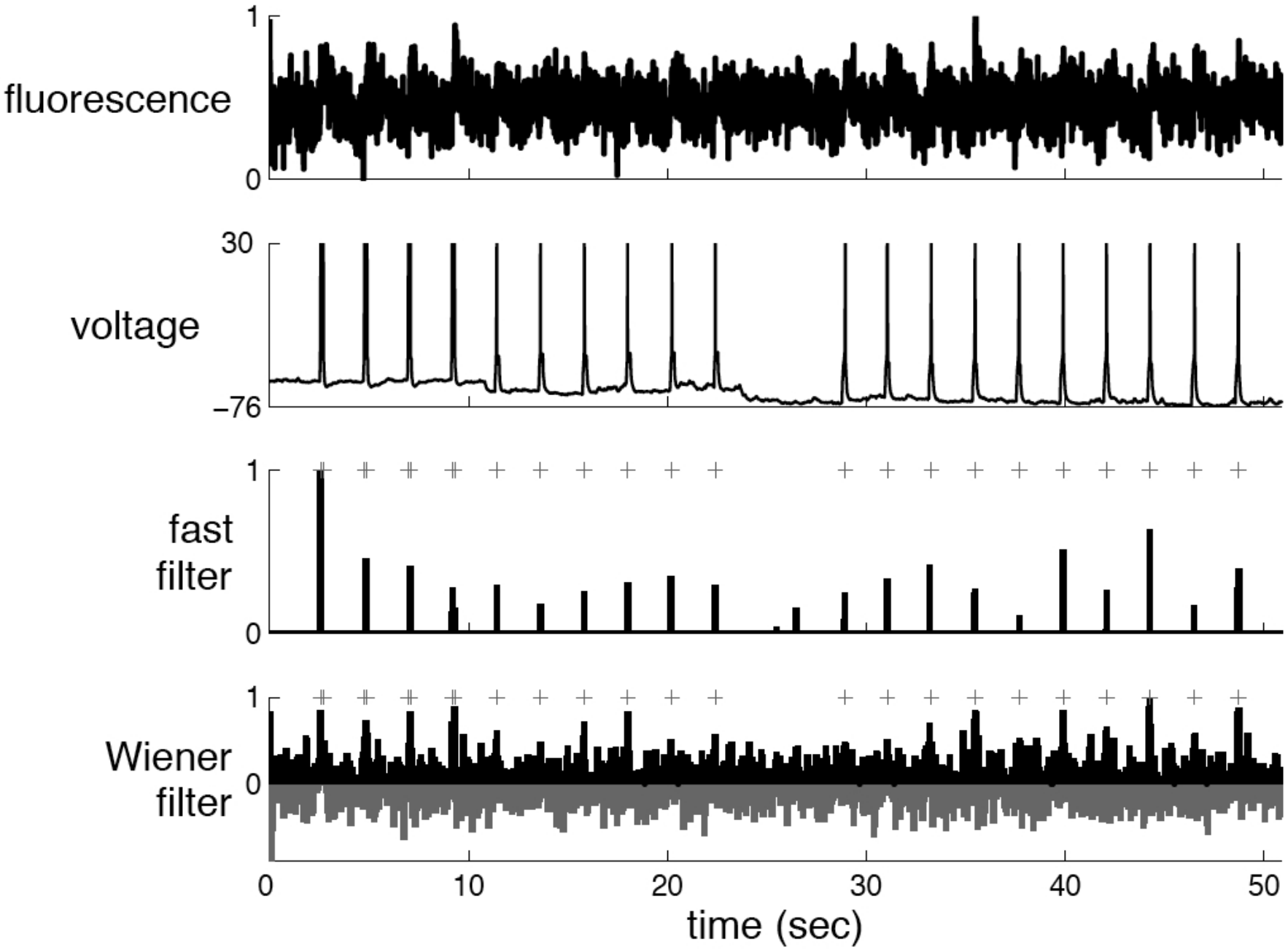
$$\begin{aligned} \log p(C|F) &= \log p(C) + \log p(F|C) + const. \\ &= \sum_t \log p(C_{t+1}|C_t) + \sum_t \log p(F_t|C_t) + const. \end{aligned}$$

Three basic observations:

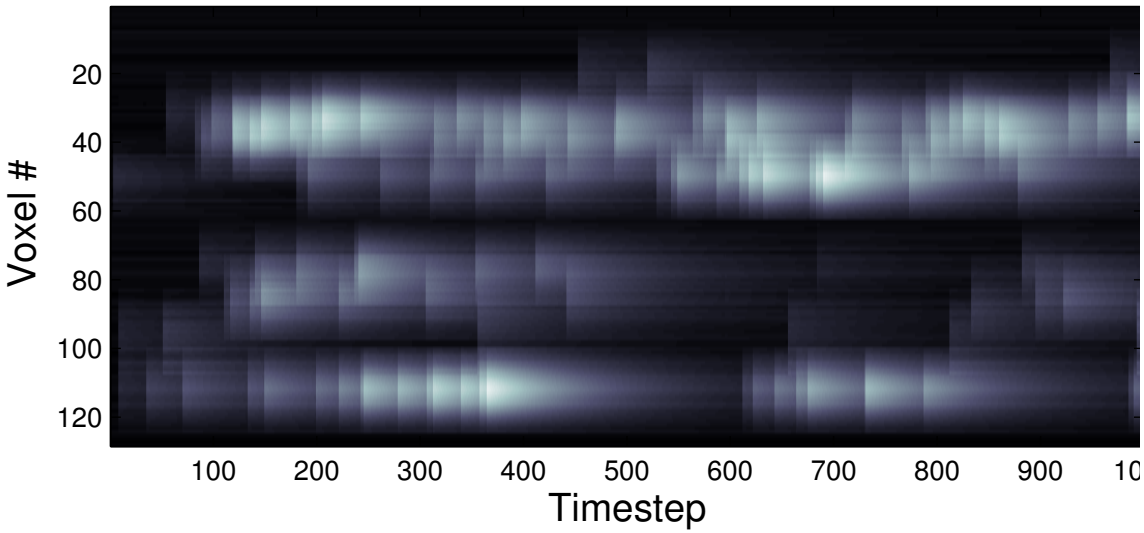
- If $\log p(C_{t+1}|C_t)$ and $\log p(F_t|C_t)$ are concave, then so is $\log p(C|F)$.
- Hessian H of $\log p(C|F)$ is tridiagonal: $\log p(F_t|C_t)$ contributes a diag term, and $\log p(C_{t+1}|C_t)$ contributes a tridiag term (Paninski et al., 2010).
- C is a linear function of n.

Newton's method: iteratively solve $HC_{dir} = \nabla$. Tridiagonal solver requires $O(T)$ time. Can include nonneg constraint $n_t \geq 0$ (Koyama and Paninski, 2010) — real-time deconvolution (Vogelstein et al., 2010).

Example: nonnegative MAP filtering



Multineuronal case: spatiotemporal demixing



Model:

$$
\begin{aligned}
Y &= C + \epsilon \\
C(x,t) &= \sum_{i=1}^{r} s_i(x) f_i(t) \\
f_i(t+dt) &= \left(1 - \frac{dt}{\tau_i}\right) f_i(t) + n_i(t)
\end{aligned}
$$

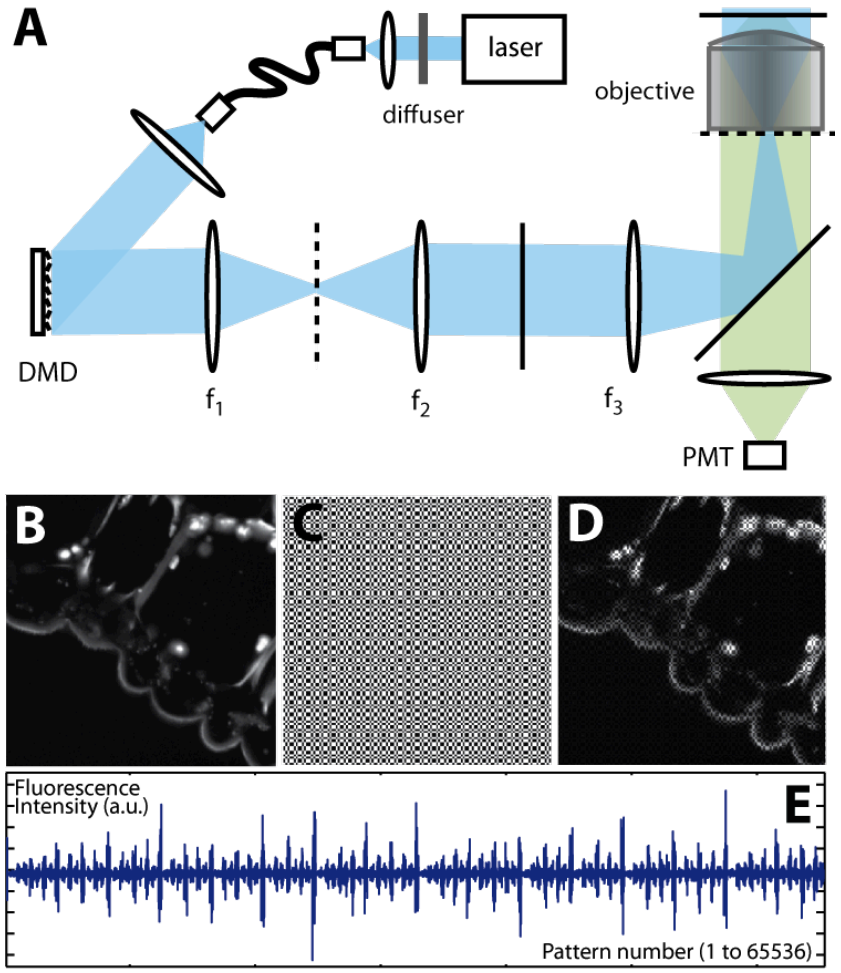
Goal: infer low-rank matrix C from noisy Y. Rank r = number of visible neurons

Additional structure: jumps in $f_i(t)$ are non-negative

Rank-penalized convex optimization with nonnegativity constraints to infer C, followed by iterative matrix factorization under nonnegativity constraints to infer $s_i(x)$, $f_i(t)$ (Pnevmatikakis et al, 2013). Examples: Machado, Lacefield

Compressed sensing imaging

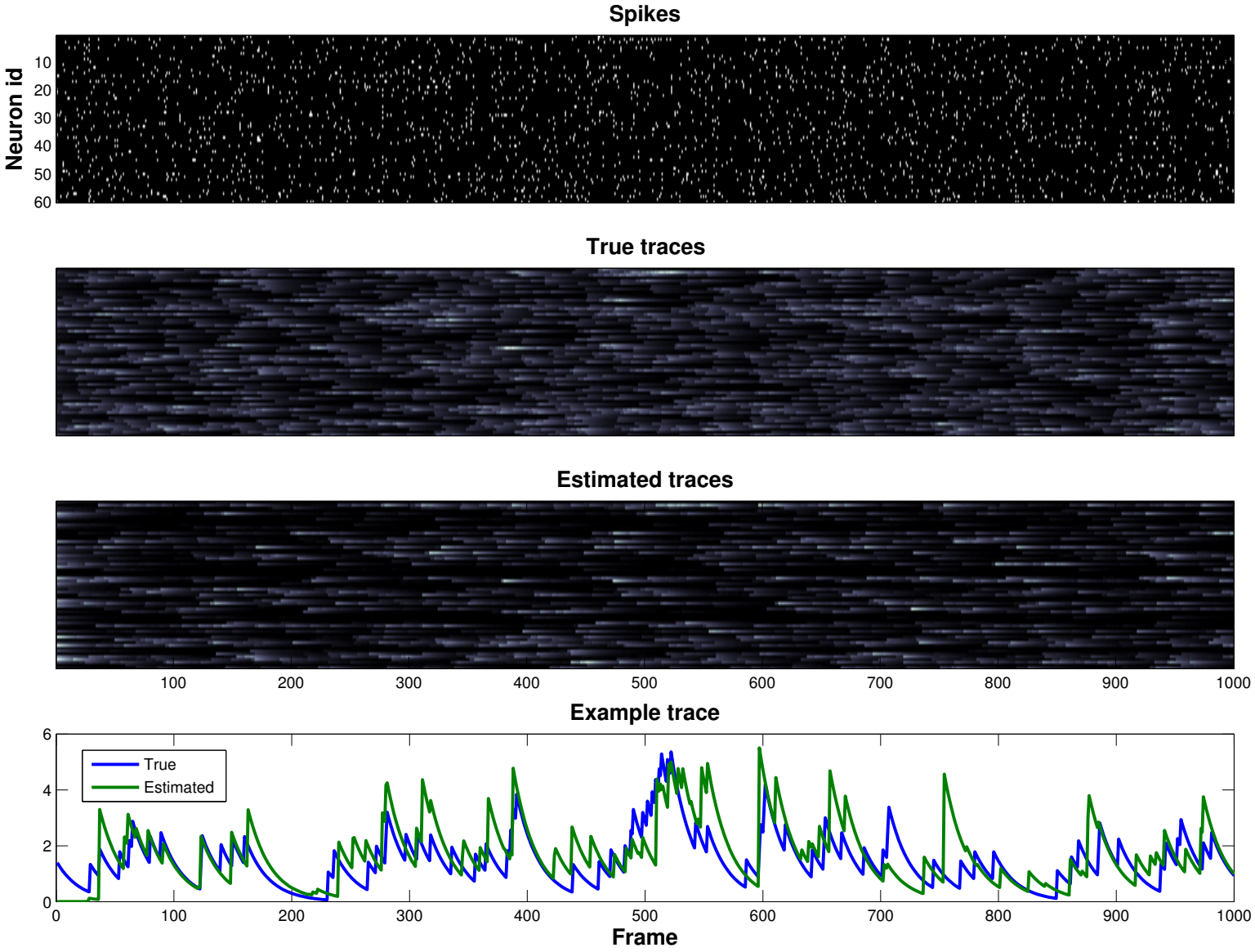
Idea: instead of raster scans, take randomized projections in each frame.



(from Studer et al, 2011)

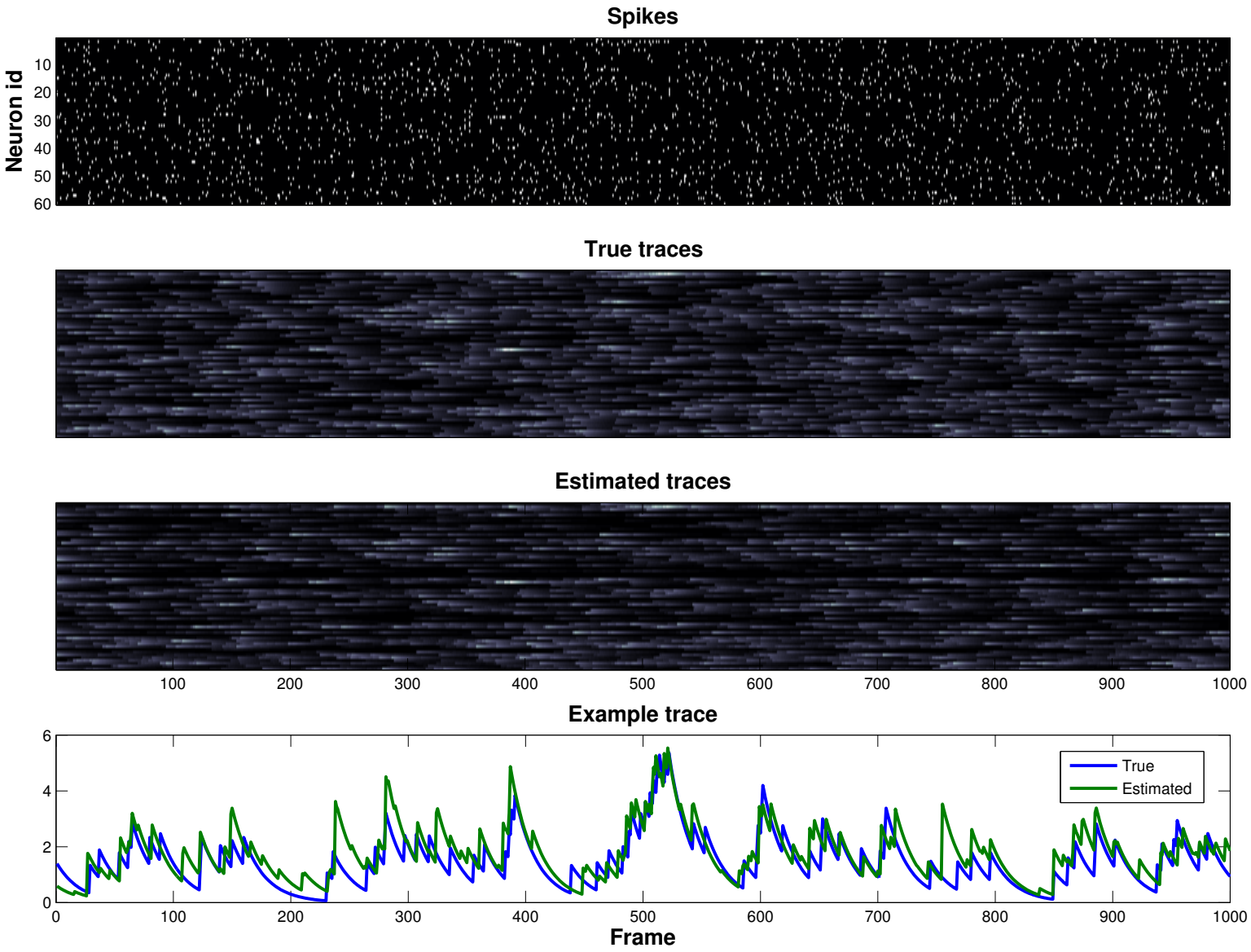
Estimating C given randomized projections Y can still be cast as a convex optimization.

Compressed sensing imaging



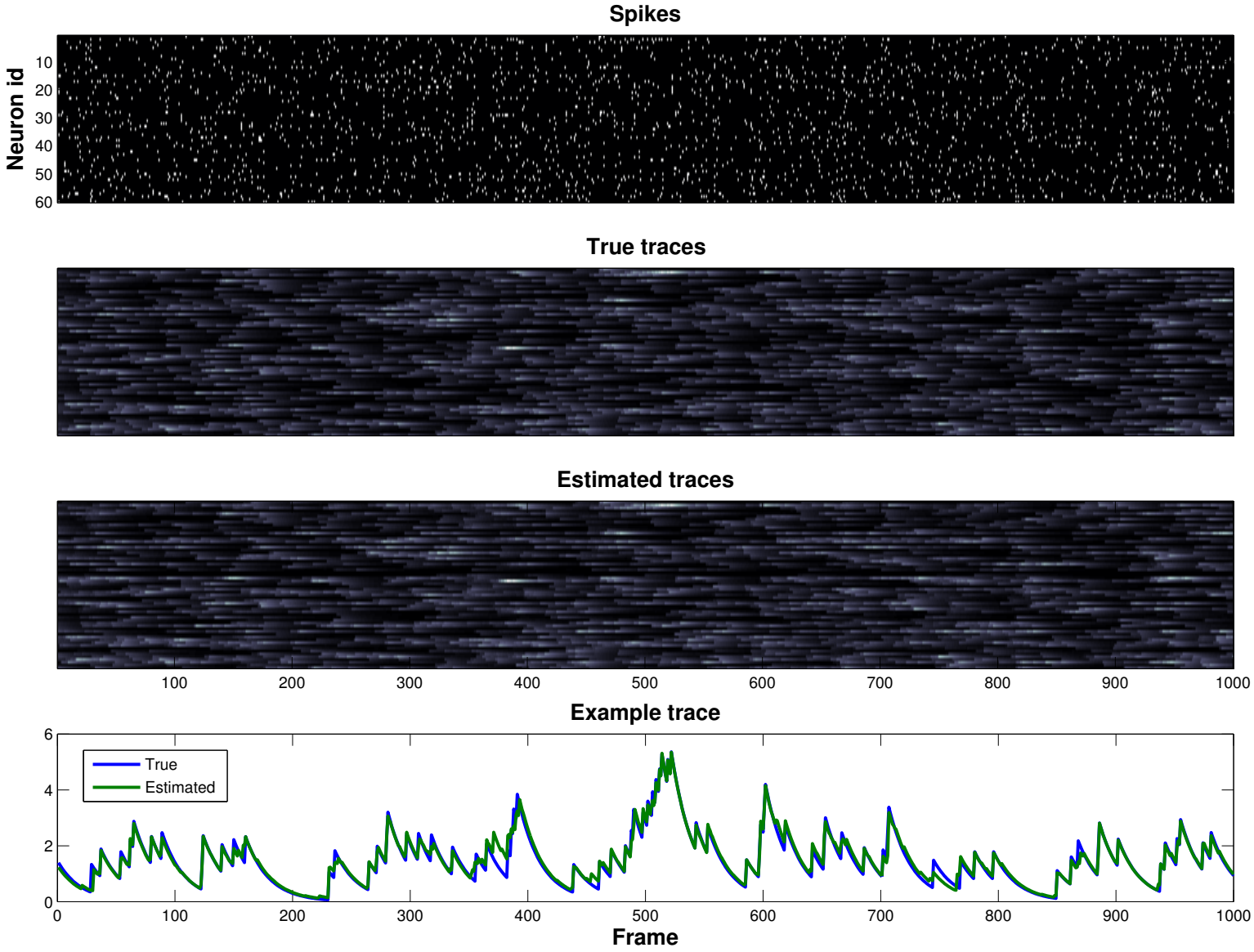
2 measurements per timestep (30x undersampling); Pnevmatikakis et al (2013)

Compressed sensing imaging



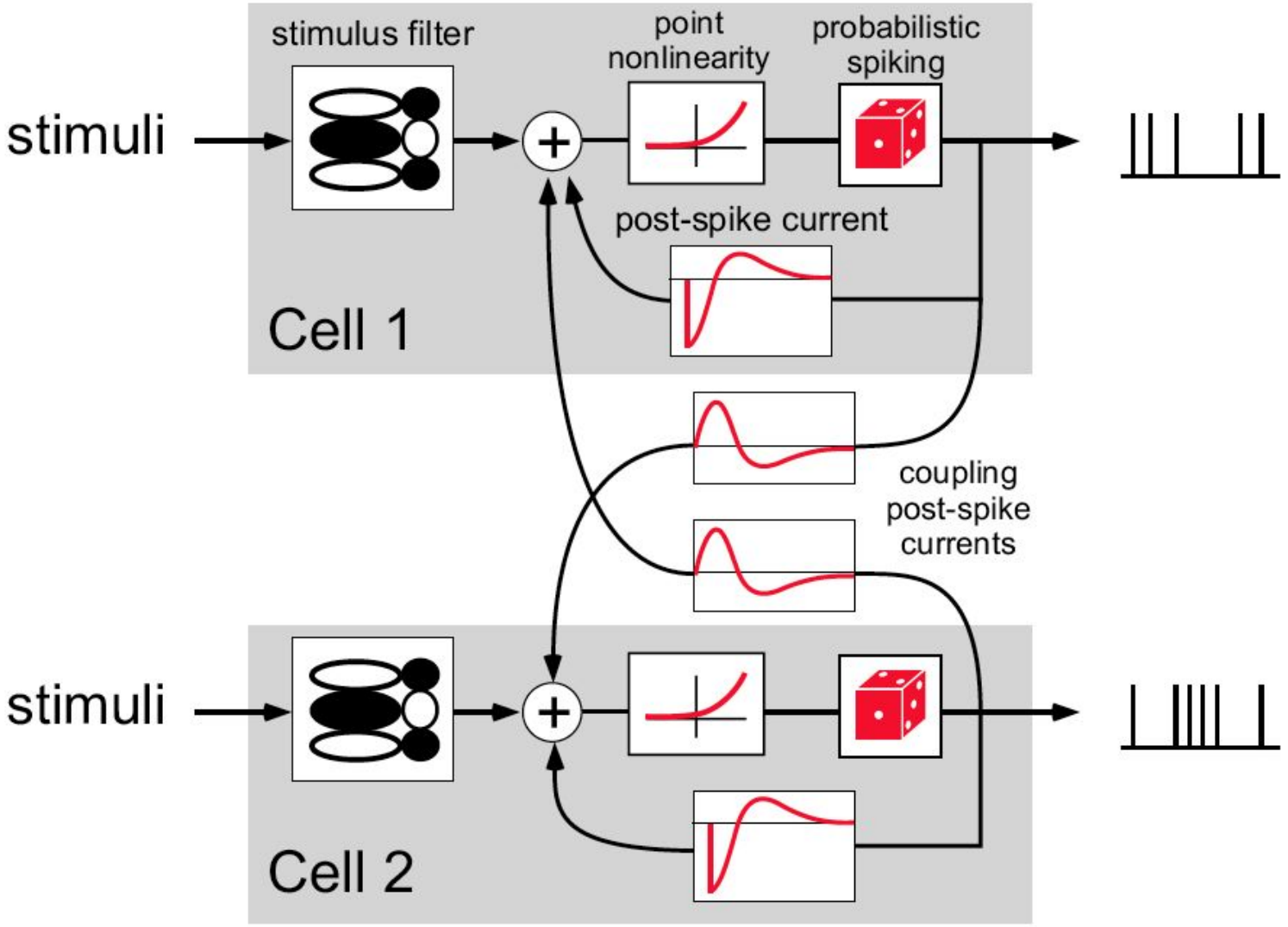
4 measurements per timestep (15x undersampling); Pnevmatikakis et al (2013)

Compressed sensing imaging



8 measurements per timestep (7.5x undersampling); Pnevmatikakis et al (2013)

Aim 2: estimating network connectivity



Given the spike times in the network, L_1-penalized concave loglikelihood optimization is easy (Paninski, 2004; Pillow et al., 2008). Fast, efficient methods from generalized linear model, compressed sensing literature.

Monte Carlo EM approach

...But we only have noisy calcium observations; true spike times are hidden variables. Thus an EM approach is once again natural.

- E step: sample spike train responses n from $p(n|F, \theta)$
- M step: given sampled spike trains, perform L_1-penalized likelihood optimization to update parameters θ.

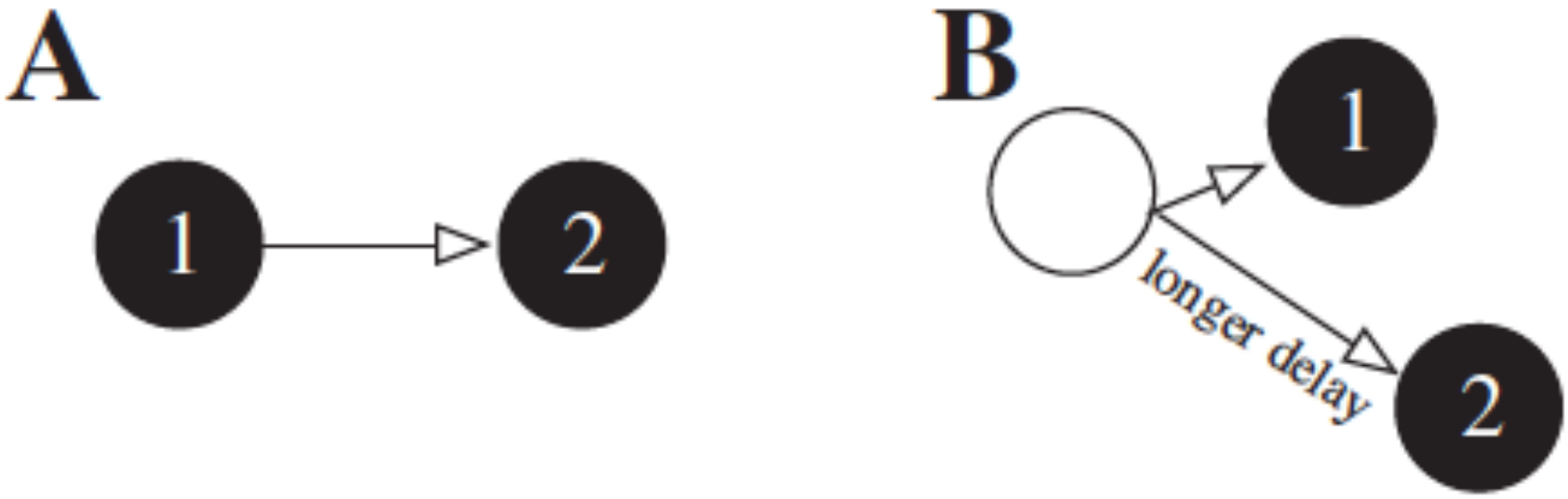
Both steps are highly parallelizable. Can also exploit many sources of prior information about cell type, proximity, anatomical likelihood of connectivity, etc.

Good news: MAP connections are inferred with the correct sign from realistic simulated network data, in just a couple minutes of compute time, if we observe the full network.

Bad news: poor results unless we observe a large fraction of the network.

The dreaded common input problem

How to distinguish direct connectivity from common input?



(from Nykamp '07)

Previous work (e.g., Vidne et al, 2012) modeled common input terms explicitly as latent variables; works well given enough a priori information, but not a general solution.

A "shotgun" solution to the common input problem

Idea: don't observe the same subset of cells throughout the experiment.

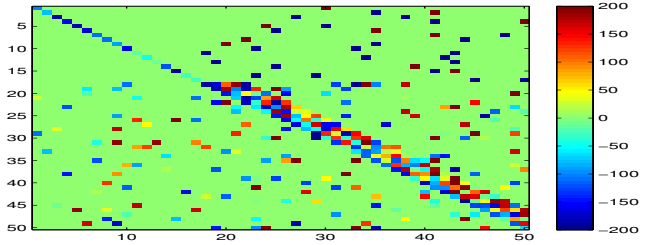
Instead, observe as many different subsets as possible.

Hard with multi-electrode arrays; easy with imaging approaches.

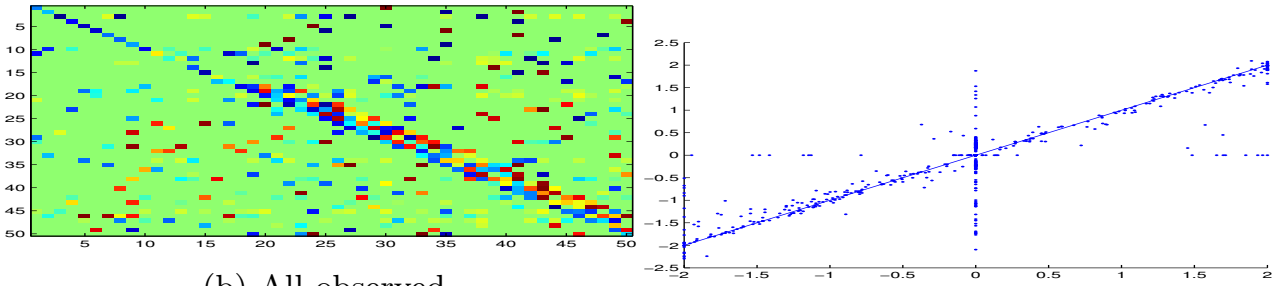
Statistics problem: how to patch together all of the estimated subnetworks?

Solution: same EM approach discussed above.

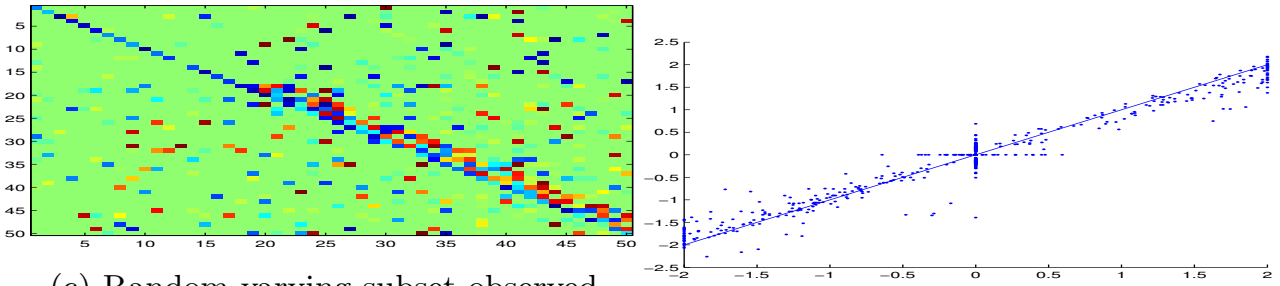
A "shotgun" solution



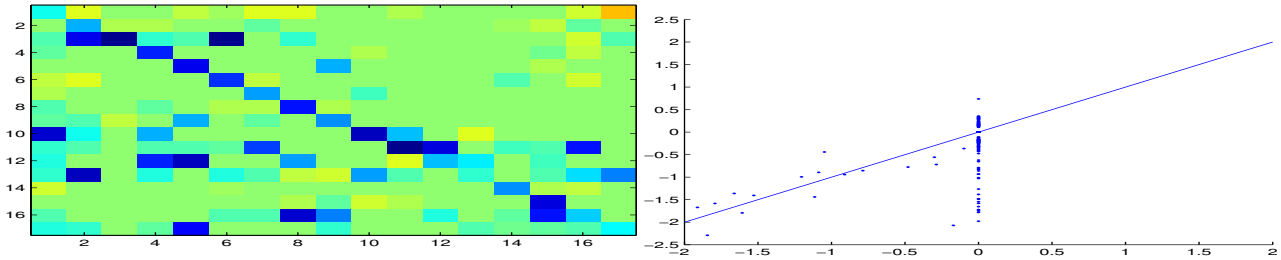
(a) True Weights



(b) All observed



(c) Random varying subset observed



(d) Fixed subset observed

only 20% of network observed at any timestep (Keshri et al, 2013)

Aim 3: Optimal control of spike timing

To test our results, we want to perturb the network at will. How can we make a neuron fire exactly when we want it to?

Assume bounded inputs; otherwise problem is trivial.

Start with a simple model:

$$\begin{aligned} \lambda_t &= f(V_t + h_t) \\ V_{t+dt} &= V_t + dt\,(-gV_t + aI_t) + \sqrt{dt}\sigma\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1). \end{aligned}$$

Now we can just optimize the likelihood of the desired spike train, as a function of the input I_t, with I_t bounded.

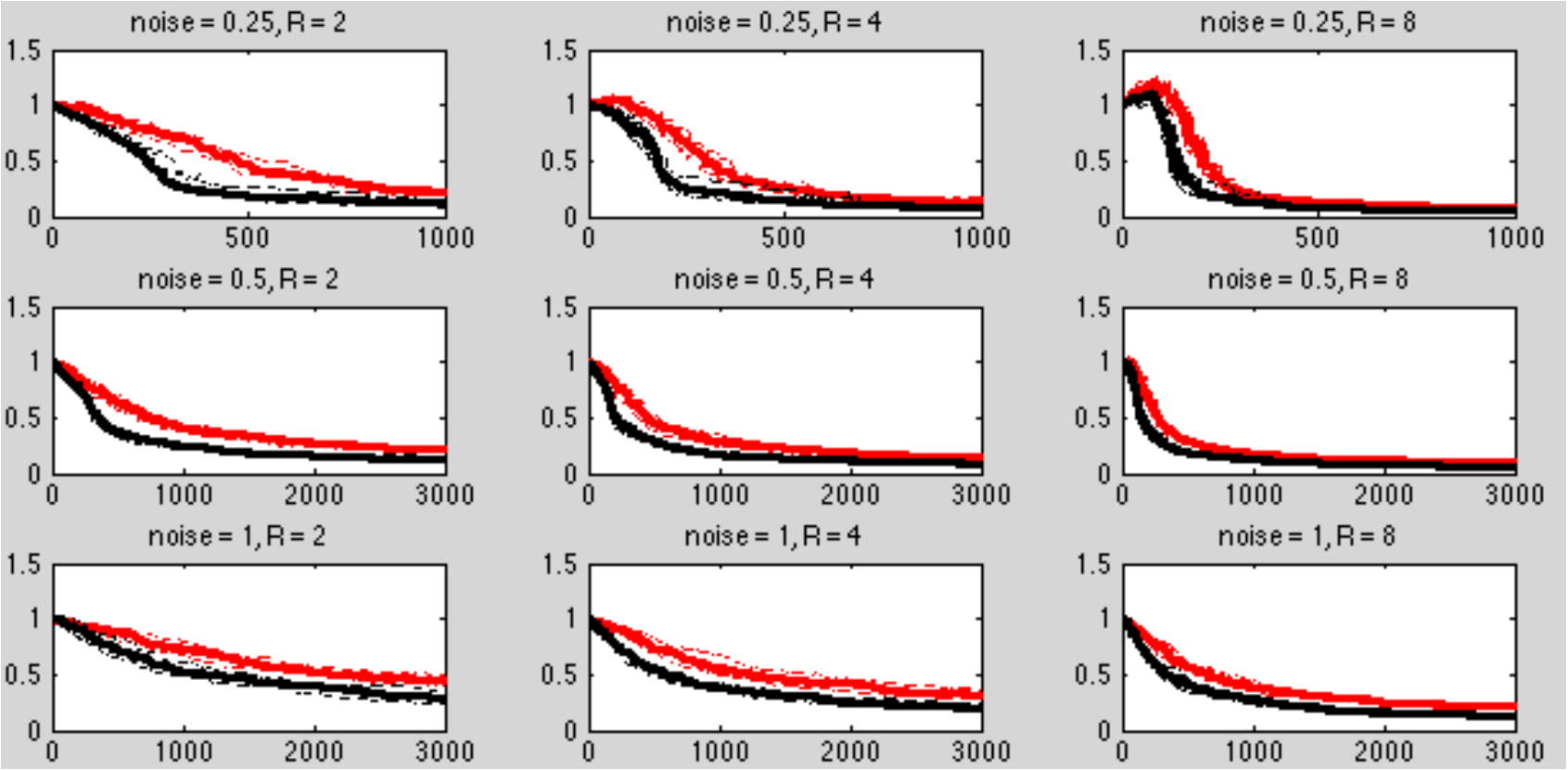
Concave objective function over convex set of possible inputs I_t + Hessian is tridiagonal $\implies$ $O(T)$ optimization.

— again, can be done in real time (Ahmadian et al., 2011).

Applications

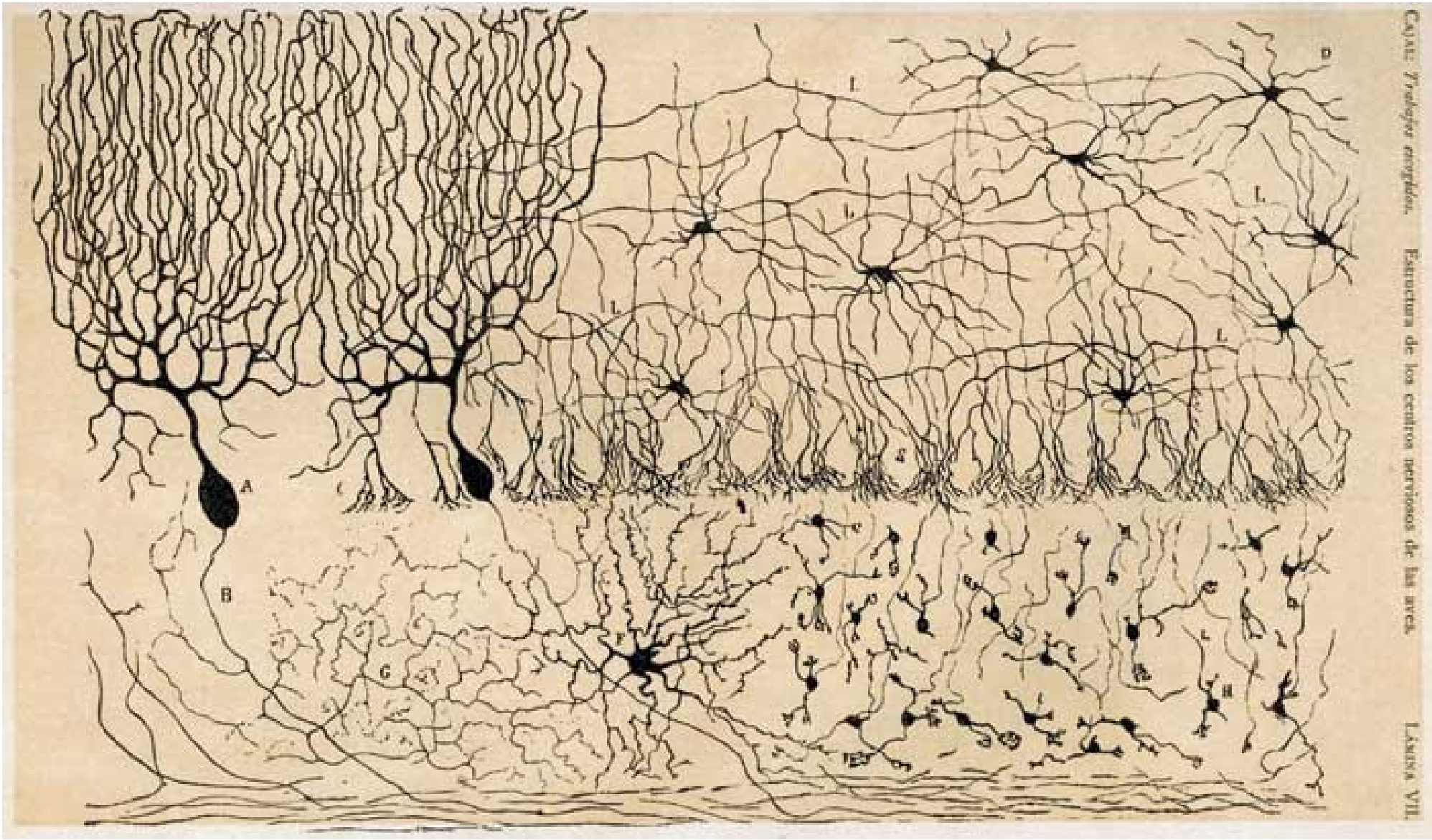
- sensory prosthetics, e.g. retinal prosthetics

- online adaptive experimental design: choose stimuli which provide as much information about network as possible.



Shababo, Paige et al (2013)

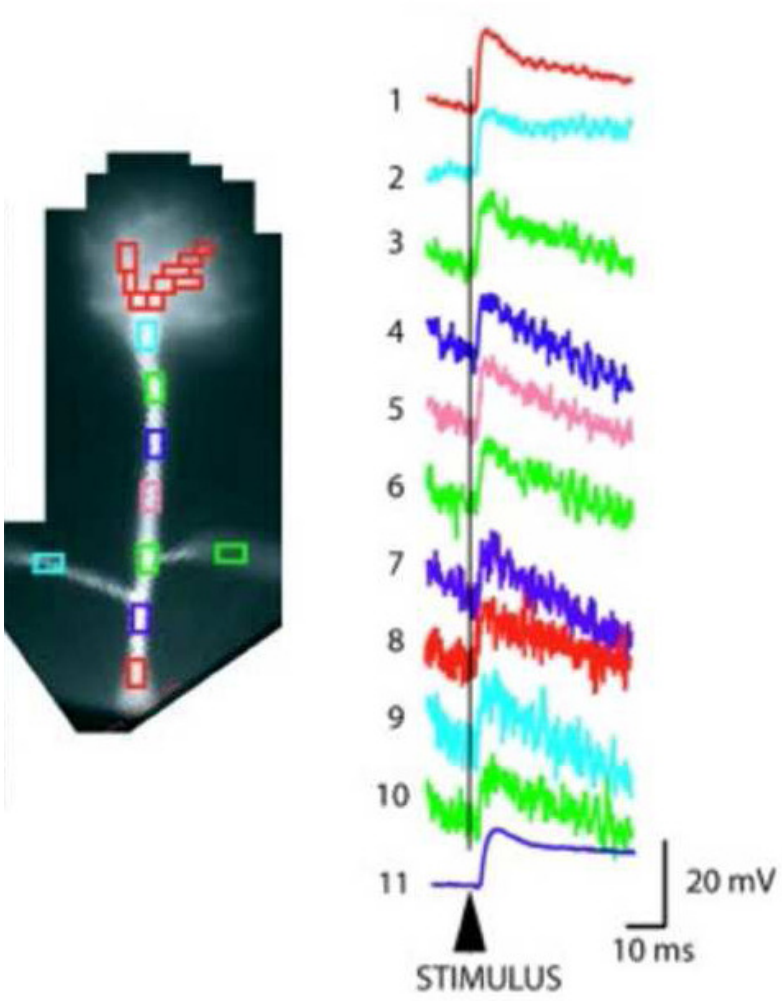
Aim 4: Connectivity at the dendritic scale



Ramon y Cajal, 1888.

The filtering problem

Spatiotemporal imaging data opens an exciting window on the computations performed by single neurons, but we have to deal with noise and intermittent observations.



Inference of spatiotemporal neuronal state given noisy observations

Variable of interest, q_t, evolves according to a noisy differential equation (e.g., cable equation):

$$dq/dt = f(q) + \epsilon_t.$$

Make noisy observations:

$$y(t) = g(q_t) + \eta_t.$$

We want to infer $E(q_t|Y)$: optimal estimate given observations. We also want errorbars: quantify how much we actually know about q_t.

If $f(.)$ and $g(.)$ are linear, and ϵ_t and η_t are Gaussian, then solution is classical: Kalman filter.

Extensions to nonlinear dynamics, non-Gaussian observations: hidden Markov ("state-space") model, particle filtering (Huys and Paninski, 2009)

Basic idea: Kalman filter

Dynamics and observation equations:

$$d\vec{V}/dt = A\vec{V} + \vec{\epsilon}_t$$

$$\vec{y}_t = B_t\vec{V} + \vec{\eta}_t$$

$V_i(t)$ = voltage at compartment i

A = cable dynamics matrix: includes leak terms ($A_{ii} = -g_l$) and intercompartmental terms ($A_{ij} = 0$ unless compartments are adjacent)

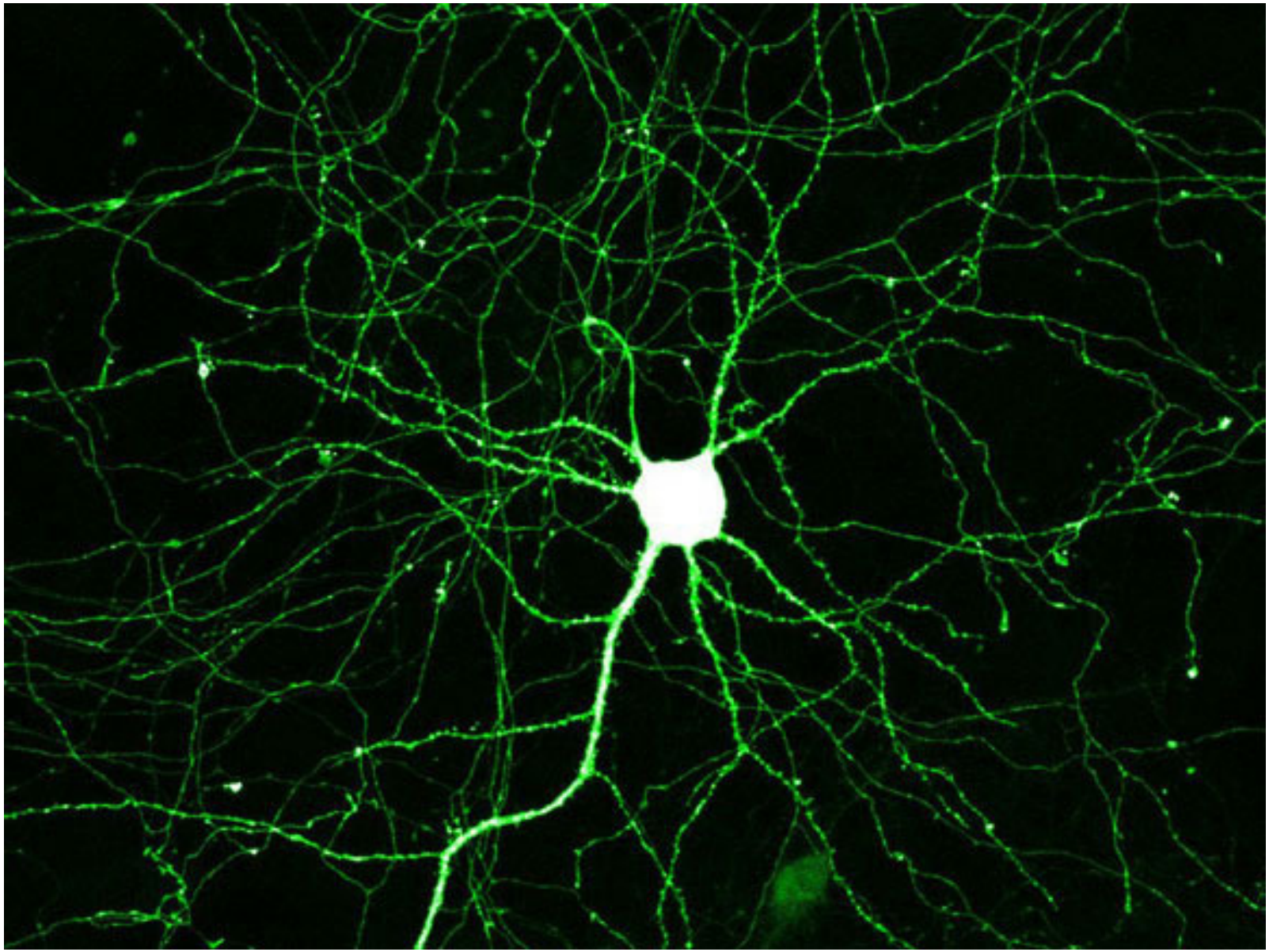
B_t = observation matrix: point-spread function of microscope

Even this case is challenging, since $d = \dim(\vec{V})$ is very large

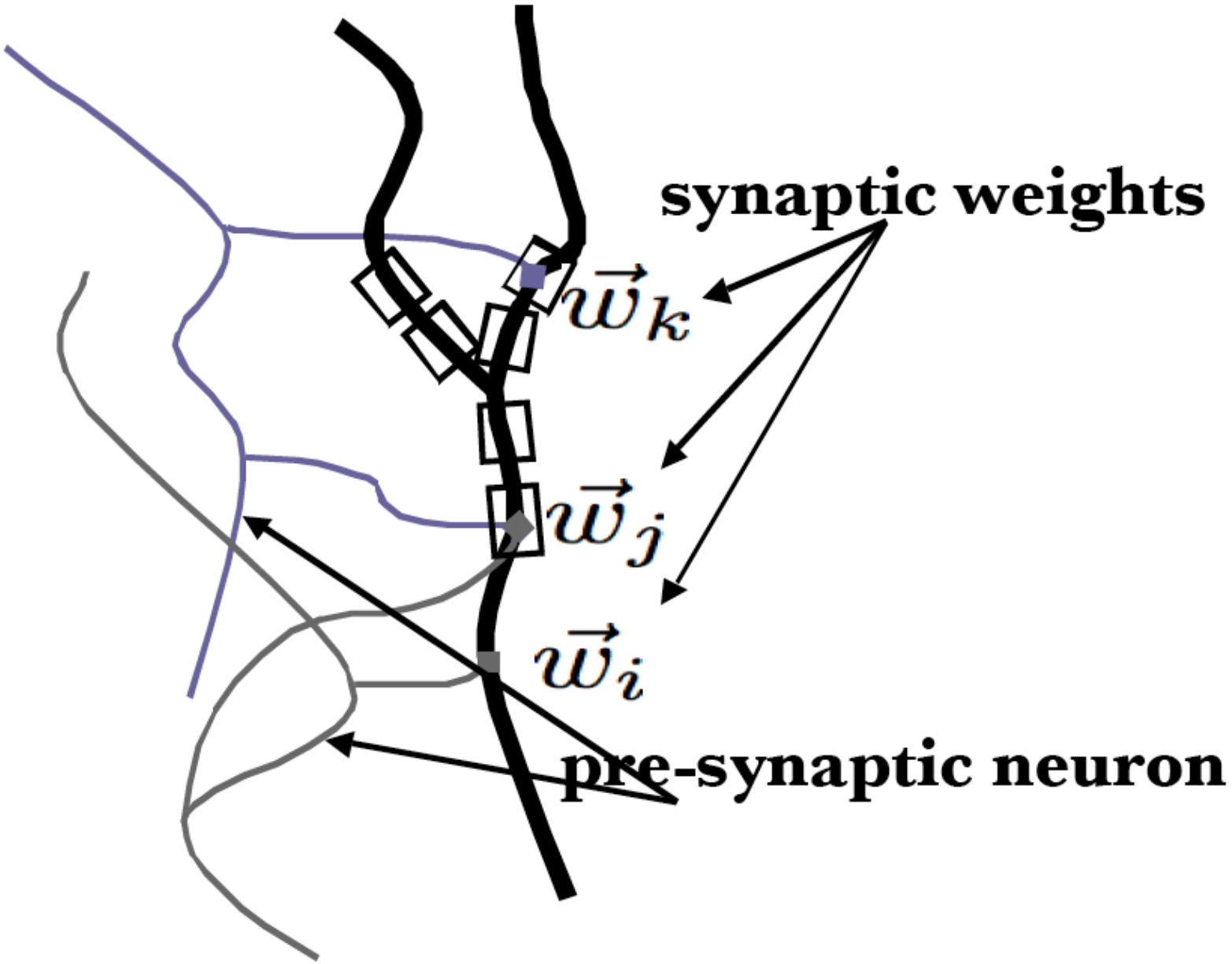
Standard Kalman filter: $O(d^3)$ computation per timestep (matrix inversion)

(Paninski, 2010): methods for Kalman filtering in just $O(d)$ time: take advantage of sparse tree structure.

Application: synaptic locations/weights



Application: synaptic locations/weights



Application: synaptic locations/weights

Including known terms:

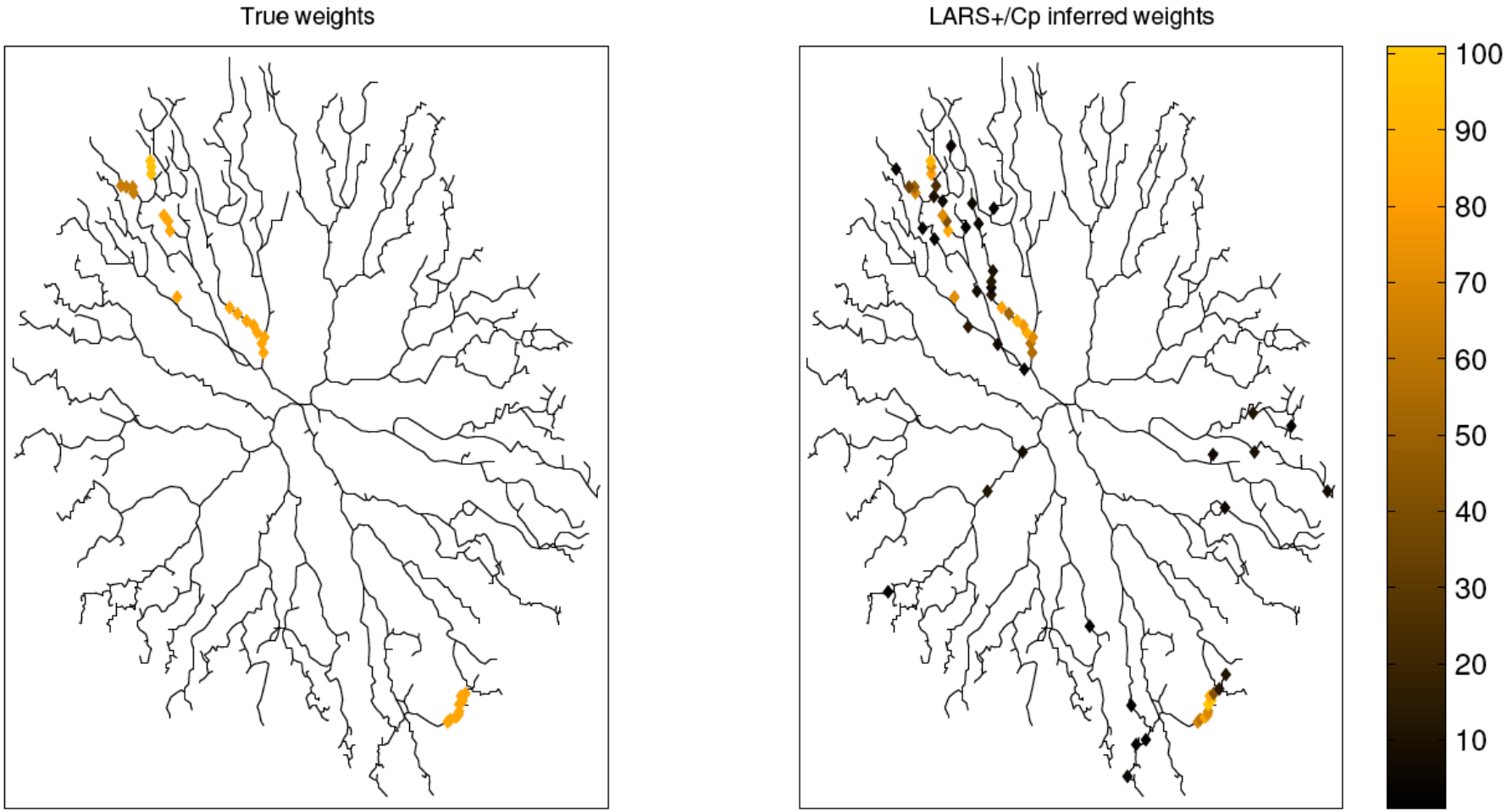
$$d\vec{V}/dt = A\vec{V}(t) + W\vec{U}(t) + \vec{\epsilon}(t);$$

$U(t)$ are known presynaptic spike times, and we want to detect which compartments are connected (i.e., infer the weight matrix W).

Loglikelihood is quadratic; W is a sparse vector. L_1-penalized loglikelihood can be optimized efficiently with homotopy (LARS) approach.

Total computation time: $O(dTk)$; $d = \#$ compartments, $T = \#$ timesteps, $k = \#$ nonzero weights.

Example: real neural geometry



700 timesteps observed; 40 compartments (of > 2000) observed per timestep

Note: random access scanning essential here: results are poor if we observe the same compartments at each timestep (Pakman, Huggins et al 2013).

Conclusions

- Modern statistical approaches provide flexible, powerful methods for answering key questions in neuroscience
- Close relationships between biophysics, statistical modeling, and experimental design
- Modern optimization methods make computations very tractable; suitable for closed-loop experiments
- Experimental methods progressing rapidly; many new challenges and opportunities for breakthroughs based on statistical ideas

References

Ahmadian, Y., Packer, A. M., Yuste, R., and Paninski, L. (2011). Designing optimal stimuli to control neuronal spike timing. *Journal of Neurophysiology*, 106(2):1038–1053.

Djurisic, M., Antic, S., Chen, W. R., and Zecevic, D. (2004). Voltage imaging from dendrites of mitral cells: EPSP attenuation and spike trigger zones. *J. Neurosci.*, 24(30):6703–6714.

Huys, Q. and Paninski, L. (2009). Model-based smoothing of, and parameter estimation from, noisy biophysical recordings. *PLOS Computational Biology*, 5:e1000379.

Knopfel, T., Diez-Garcia, J., and Akemann, W. (2006). Optical probing of neuronal circuit dynamics: genetically encoded versus classical fluorescent sensors. *Trends in Neurosciences*, 29:160–166.

Koyama, S. and Paninski, L. (2010). Efficient computation of the map path and parameter estimation in integrate-and-fire and more general state-space models. *Journal of Computational Neuroscience*, 29:89–105.

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, 15:243–262.

Paninski, L. (2010). Fast Kalman filtering on quasilinear dendritic trees. *Journal of Computational Neuroscience*, 28:211–28.

Paninski, L., Ahmadian, Y., Ferreira, D., Koyama, S., Rahnama, K., Vidne, M., Vogelstein, J., and Wu, W. (2010). A new look at state-space models for neural data. *Journal of Computational Neuroscience*, 29:107–126.

Pillow, J., Shlens, J., Paninski, L., Sher, A., Litke, A., Chichilnisky, E., and Simoncelli, E. (2008). Spatiotemporal correlations and visual signaling in a complete neuronal population. *Nature*, 454:995–999.

Vogelstein, J., Packer, A., Machado, T., Sippy, T., Babadi, B., Yuste, R., and Paninski, L. (2010). Fast non-negative deconvolution for spike train inference from population calcium imaging. *J. Neurophys.*, 104:3691–704.