Challenges and opportunities in statistical neuroscience

Liam Paninski

Department of Statistics and Center for Theoretical Neuroscience
Columbia University
http://www.stat.columbia.edu/∼liam
liam@stat.columbia.edu

May 14, 2012

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

The coming statistical neuroscience decade

Some notable recent developments:

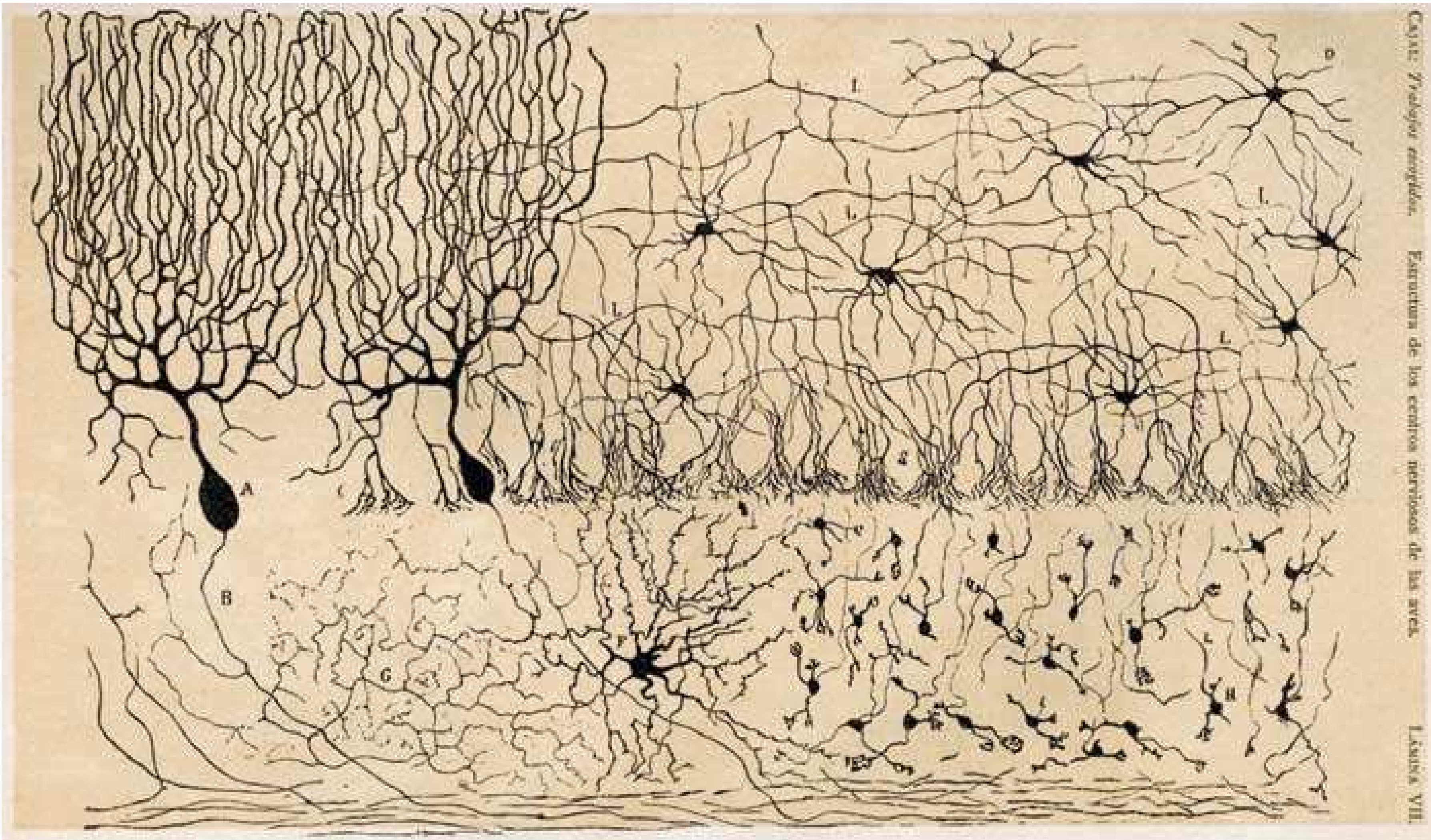
- machine learning / statistics methods for extracting information from high-dimensional data in a computationally-tractable, systematic fashion
- computing (Moore's law, massive parallel computing, GPUs)
- optical methods for recording and stimulating many genetically-targeted neurons simultaneously
- high-density multielectrode recordings (Litke's 512-electrode retinal readout system; Shepard's 65,536-electrode active array)

Some exciting open challenges

- inferring biophysical neuronal properties from noisy recordings
- reconstructing the full dendritic spatiotemporal voltage from noisy, subsampled observations
- estimating subthreshold voltage given superthreshold spike trains
- extracting spike timing from slow, noisy calcium imaging data
- reconstructing presynaptic conductance from postsynaptic voltage recordings
- inferring connectivity from large populations of spike trains
- decoding behaviorally-relevant information from spike trains
- optimal control of neural spike timing

— to solve these, we need to combine the two classical branches of computational neuroscience: dynamical systems and neural coding

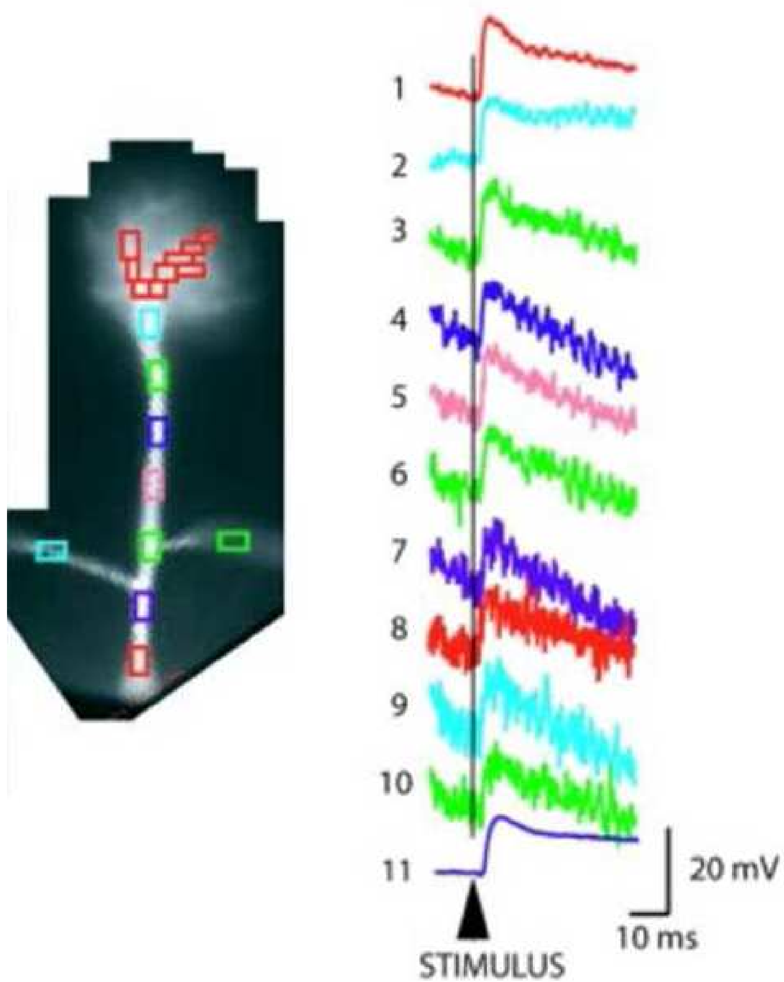
1. Basic goal: understanding dendrites



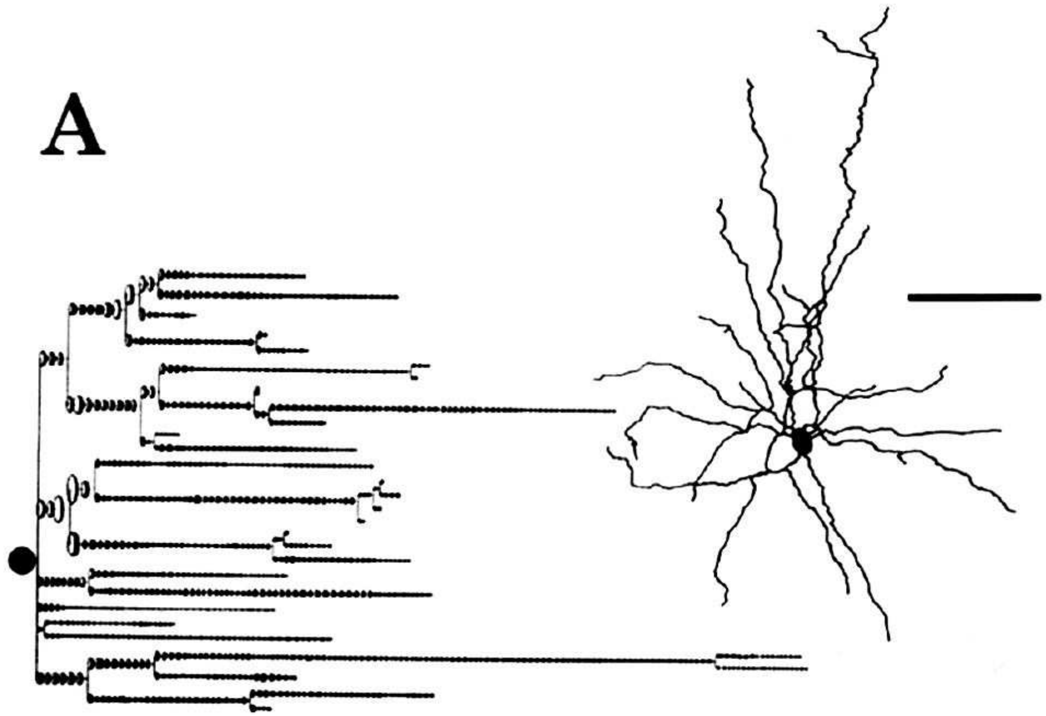
Ramon y Cajal, 1888.

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.



Basic paradigm: compartmental models



- write neuronal dynamics in terms of equivalent nonlinear, time-varying RC circuits
- leads to a coupled system of stochastic differential equations

Inference of spatiotemporal neuronal state given noisy observations

State-space approach: q_t = state of neuron at time t.

We want $p(q_t|Y_{1:t}) \propto p(q_t, Y_{1:t})$. Markov assumption:

$$p(Q, Y) = p(Q)p(Y|Q) = p(q_1)\left(\prod_{t=2}^{T} p(q_t|q_{t-1})\right)\left(\prod_{t=1}^{T} p(y_t|q_t)\right)$$

To compute $p(q_t, Y_{1:t})$, just recurse

$$p(q_t, Y_{1:t}) = p(y_t|q_t)\int_{q_{t-1}} p(q_t|q_{t-1})p(q_{t-1}, Y_{1:t-1})dq_{t-1}.$$

Linear-Gaussian case: requires $O(\dim(q)^3 T)$ time; in principle, just matrix algebra (Kalman filter). Approximate solutions in more general case via sequential Monte Carlo (Huys and Paninski, 2009).

Major challenge: $\dim(q)$ can be $\approx 10^4$ or greater.

Low-rank approximations

Key fact: current experimental methods provide just a few low-SNR observations per time step.

Basic idea: if dynamics are approximately linear and time-invariant, we can approximate Kalman covariance $C_t = cov(q_t|Y_{1:t})$ as a perturbation of the marginal covariance $C_0 + U_t D_t U_t^T$, with $C_0 = \lim_{t\to\infty} cov(q_t)$.

C_0 is the solution to a Lyapunov equation. It turns out that we can solve linear equations involving C_0 in $O(\dim(q))$ time via Gaussian belief propagation, using the fact that the dendrite is a tree.

The necessary recursions — i.e., updating U_t, D_t and the Kalman mean $E(q_t|Y_{1:t})$ — involve linear manipulations of C_0, using

$$\begin{aligned} C_t &= [(AC_{t-1}A^T + Q)^{-1} + B_t]^{-1} \\ C_0 + U_t D_t U_t^T &= \left([A(C_0 + U_{t-1}D_{t-1}U_{t-1}^T)A^T + Q]^{-1} + B_t\right)^{-1}, \end{aligned}$$

and can be done in $O(\dim(q))$ time (Paninski, 2010). Generalizable to many other state-space models (Pnevmatikakis and Paninski, 2011).

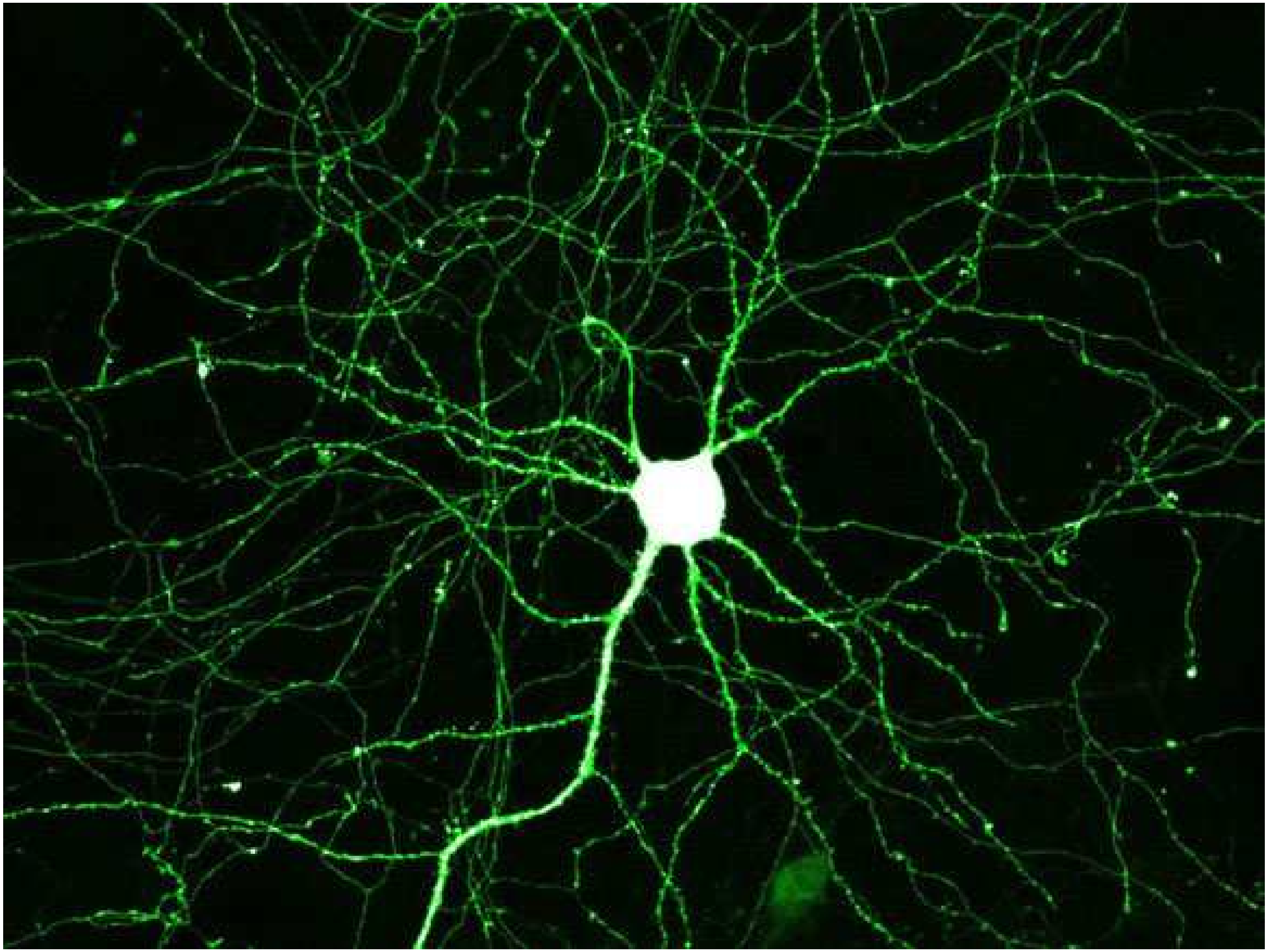
Example: inferring voltage from subsampled observations

(Loading low-rank-speckle.mp4)

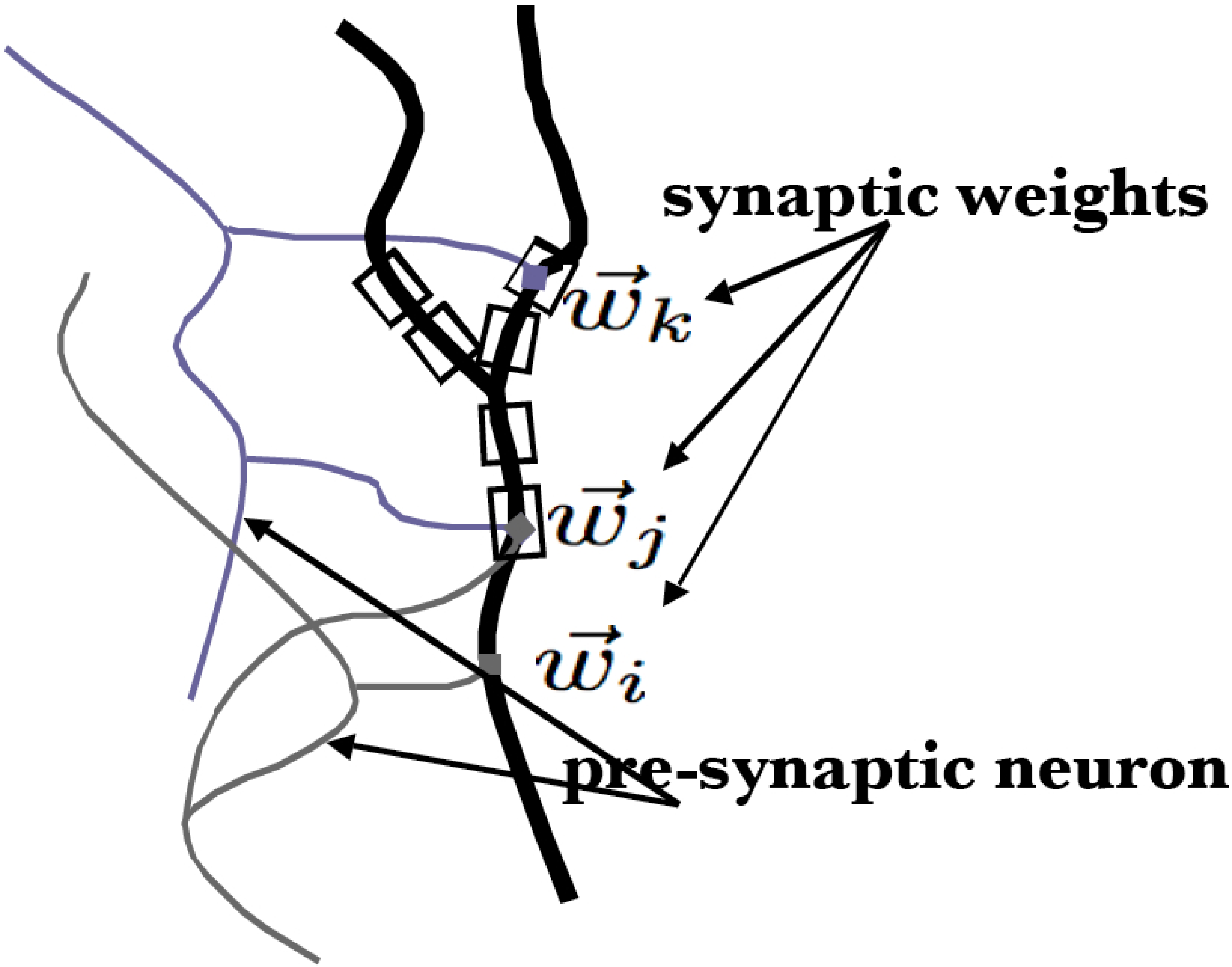
Applications

- Optimal experimental design: which parts of the neuron should we image? (Submodular optimization; Krause and Guestrin, 2007)
- Estimation of biophysical parameters (e.g., membrane channel densities, axial resistance, etc.): reduces to a simple nonnegative regression problem once $V(x,t)$ is known (Huys et al., 2006)
- Detecting location and weights of synaptic input (Pakman et al., 2012)

Application: synaptic locations/weights



Application: synaptic locations/weights



Application: synaptic locations/weights

Including known terms:

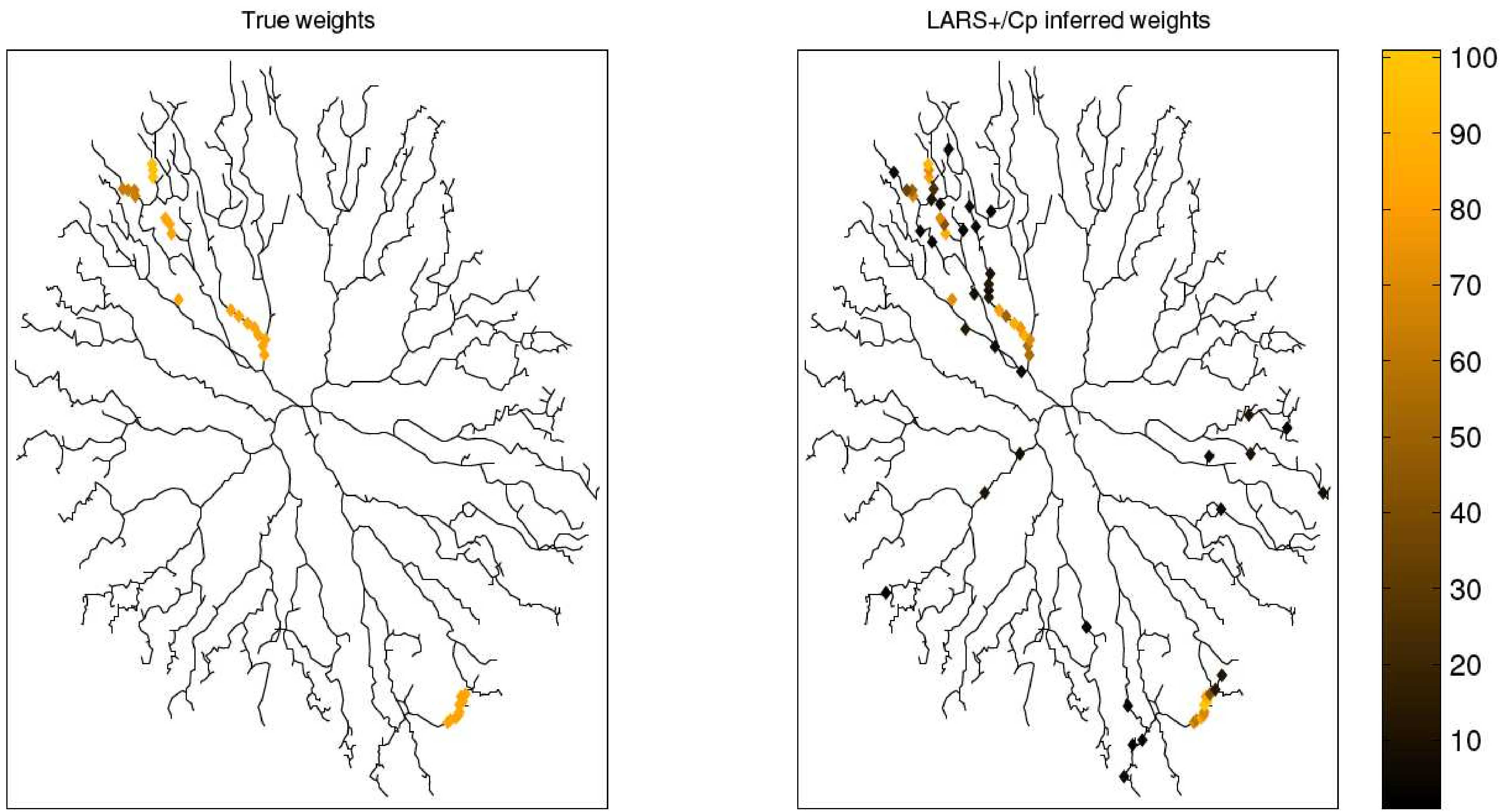
$$d\vec{V}/dt = A\vec{V}(t) + W\vec{U}(t) + \vec{\epsilon}(t);$$

$U_j(t)$ = known input terms.

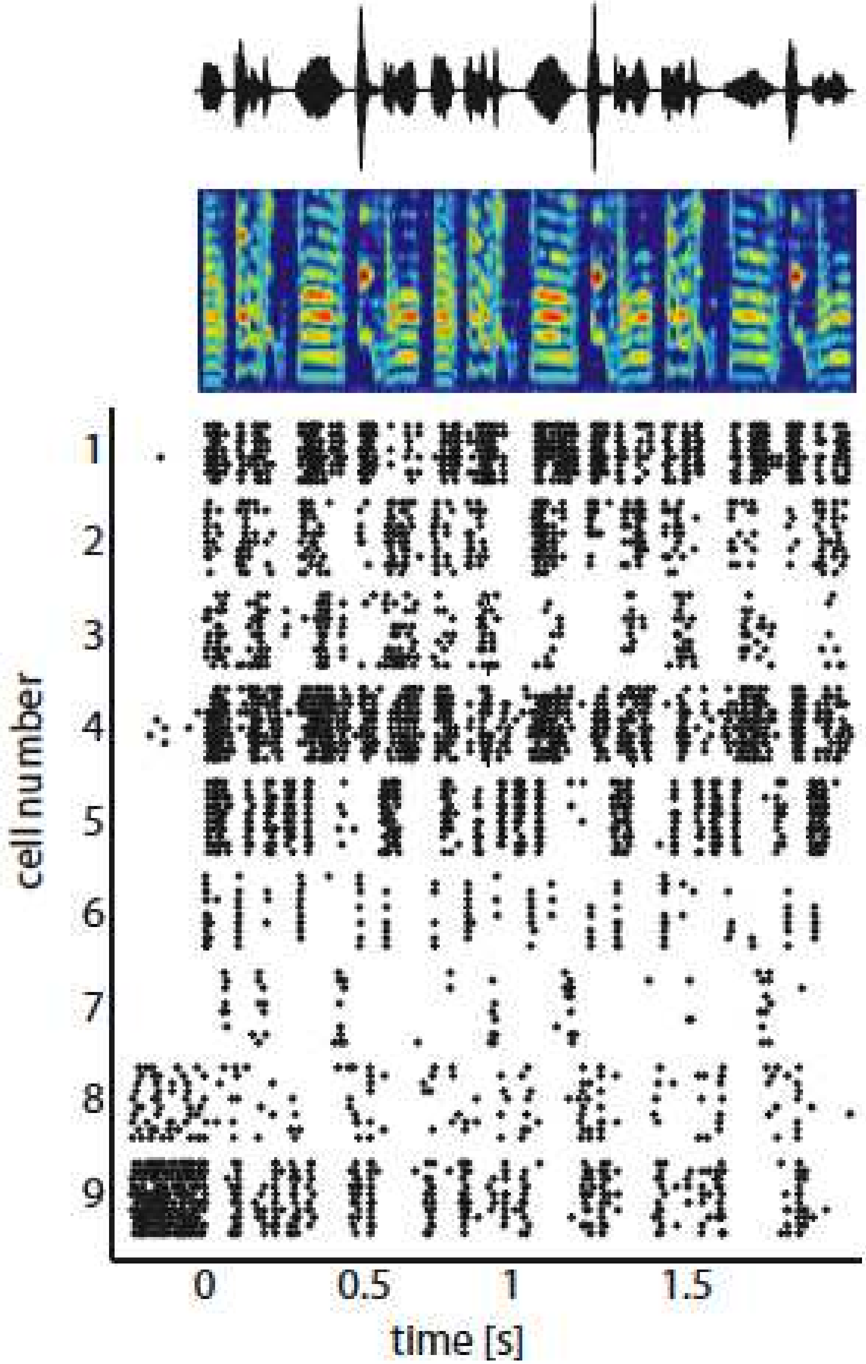
Example: $U(t)$ are known presynaptic spike times, and we want to detect which compartments are connected (i.e., infer the weight matrix W).

Loglikelihood is quadratic; L_1-penalized loglikelihood can be optimized efficiently with LARS-like approach. Total computation time is $O(NTk)$: N = # compartments, T = # timesteps, k = # nonzero weights.

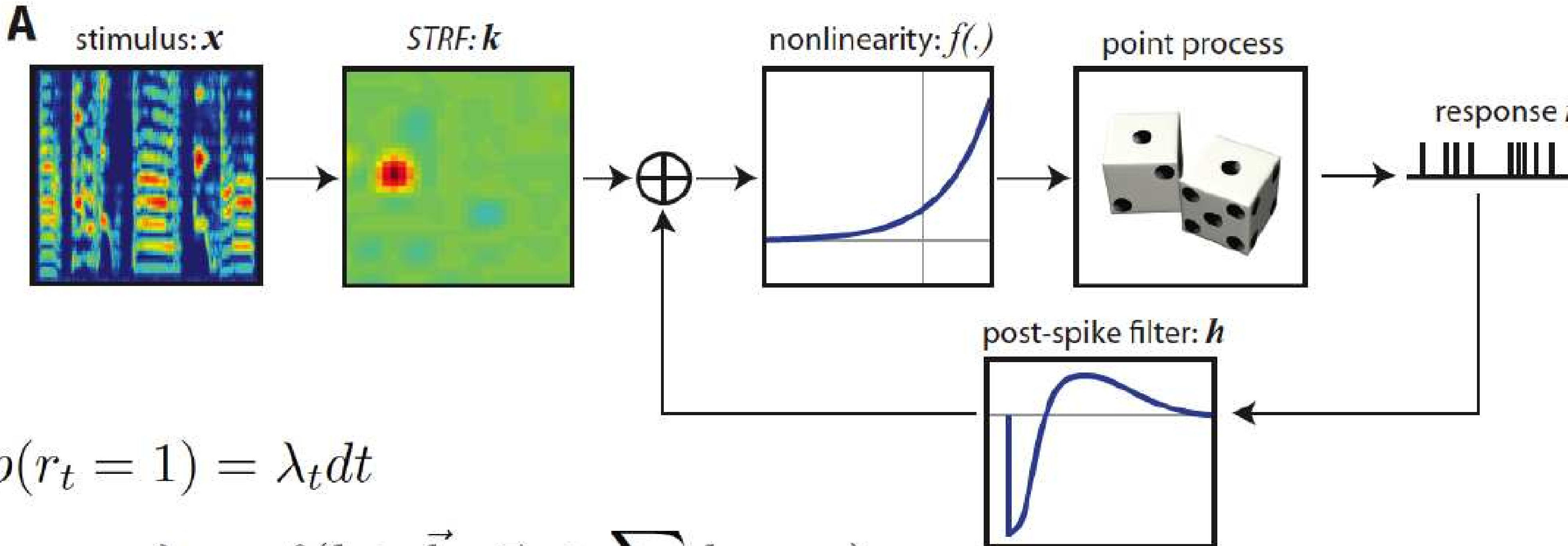
Application: synaptic locations/weights



Part 2: optimal decoding of spike train data



Semiparametric GLM



$$p(r_t = 1) = \lambda_t dt$$

$$\lambda_t = f(b + \vec{k} \cdot \vec{x}_t + \sum_j h_j r_{t-j})$$

Parameters $(\vec{k}, h)$ estimated by L_1-penalized maximum likelihood (concave); f estimated by log-spline (Calabrese, Woolley et al. 2009). Currently the best predictive model of these spike trains.

MAP stimulus decoding

It is reasonable to estimate the song X that led to a response R via the MAP

$$\hat{X} = \arg\max_X p(X|R).$$

(Note that X is very high-dimensional!) For this model, we have:

$$\begin{aligned} \log p(X|R) &= \log p(X) + \log p(R|X) + const. \\ &= \log p(X) + \sum_t \log p(r_t|X, R_{\ldots,t-1}) + const. \end{aligned}$$

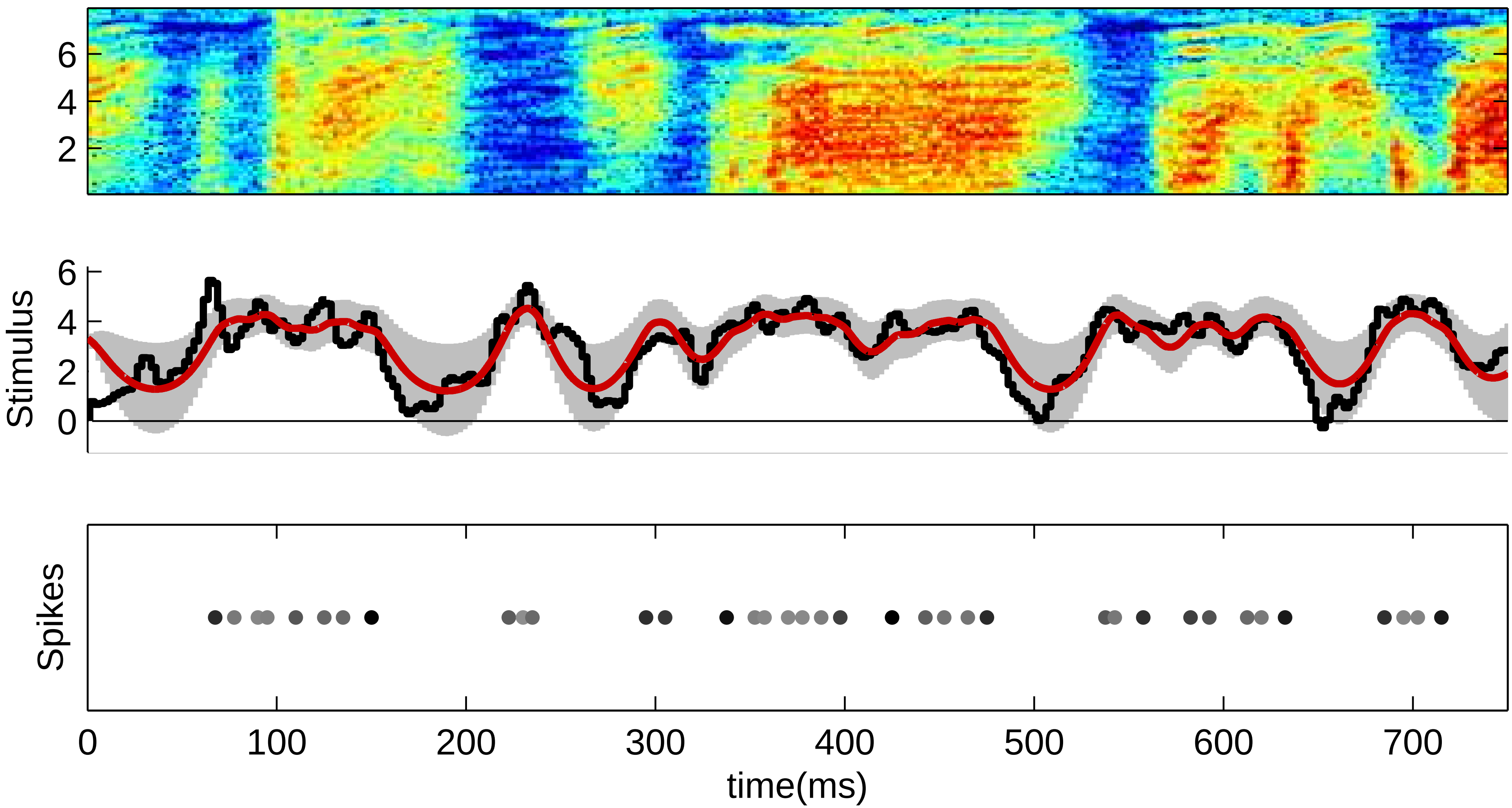
Two basic observations:

- If $\log p(X)$ is concave, then so is $\log p(X|R)$, since each $\log p(r_t|X, Y_{\ldots,t-1})$ is.
- Hessian H of $\log p(R|X)$ w.r.t. X is banded: each $p(r_t|X, R_{\ldots,t-1})$ depends on X locally in time, and therefore contributes a banded term.

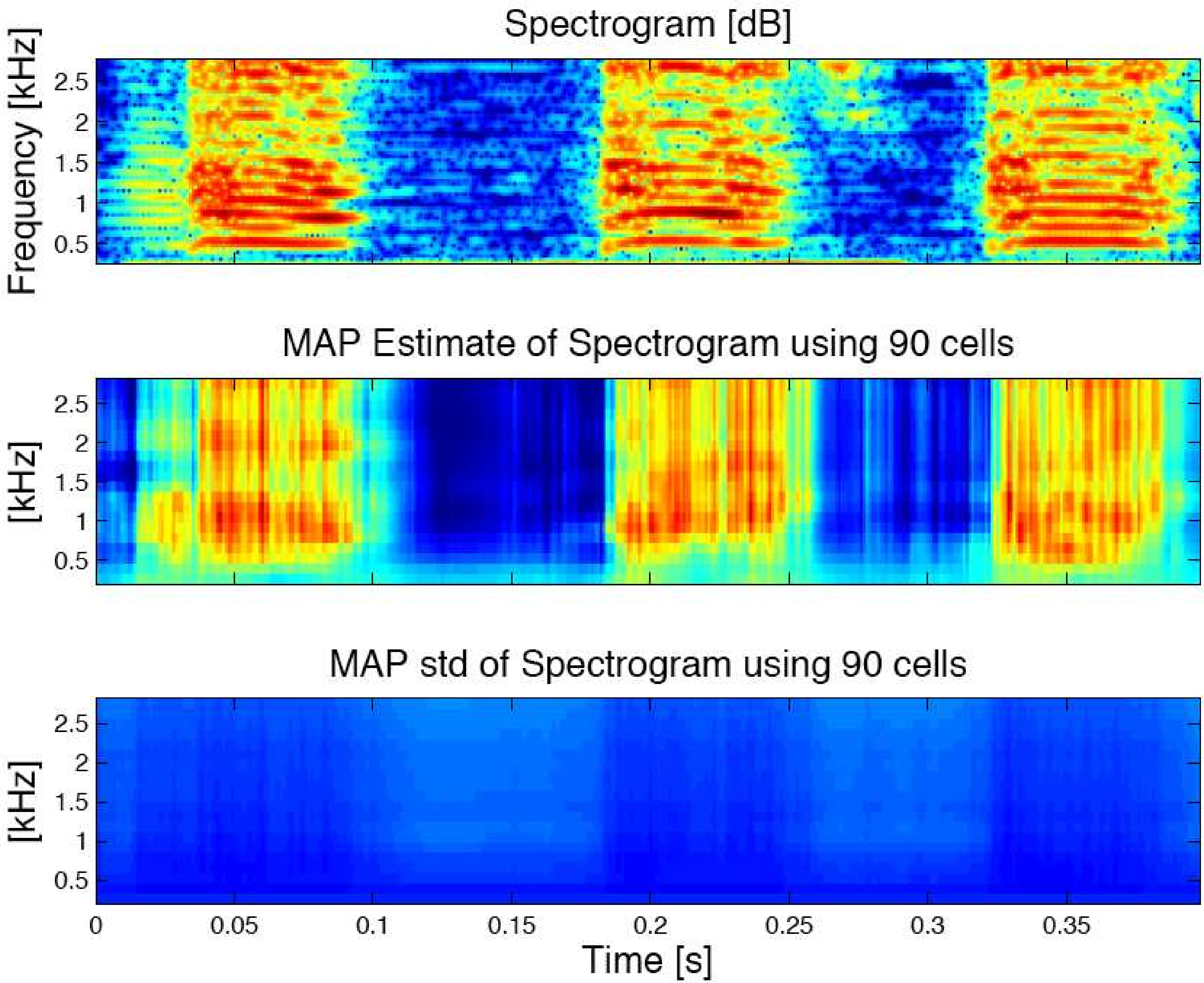
Newton's method iteratively solves $HX_{dir} = \nabla$. Solving banded systems requires $O(T)$ time, so computing MAP requires $O(T)$ time if log-prior is concave with a banded Hessian.

— similar speedups available in constrained cases (Paninski et al., 2010), and in MCMC setting (e.g., fast hybrid Monte Carlo methods (Ahmadian et al., 2010b)).

Application: fast optimal decoding

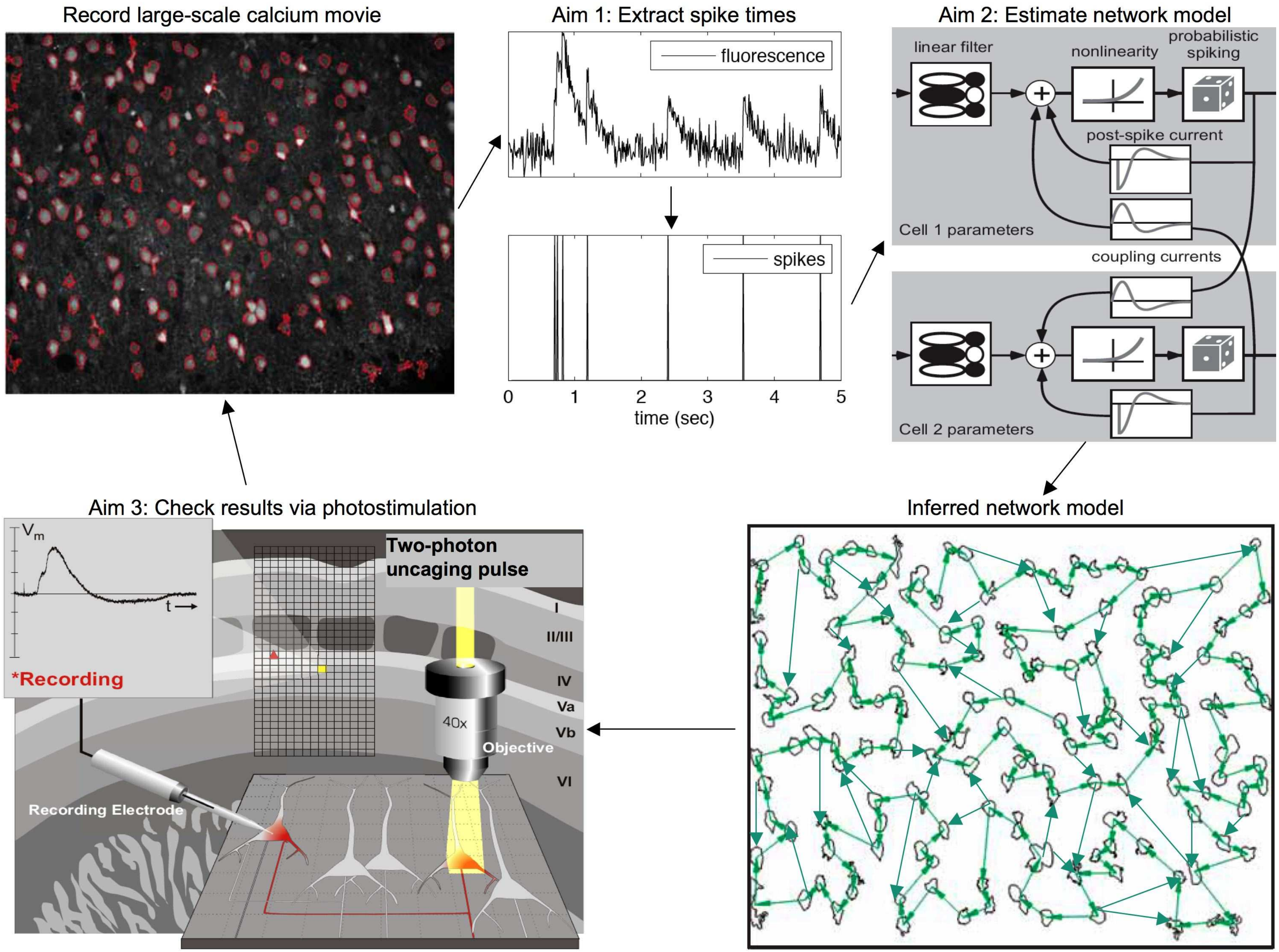


Decoding a full song

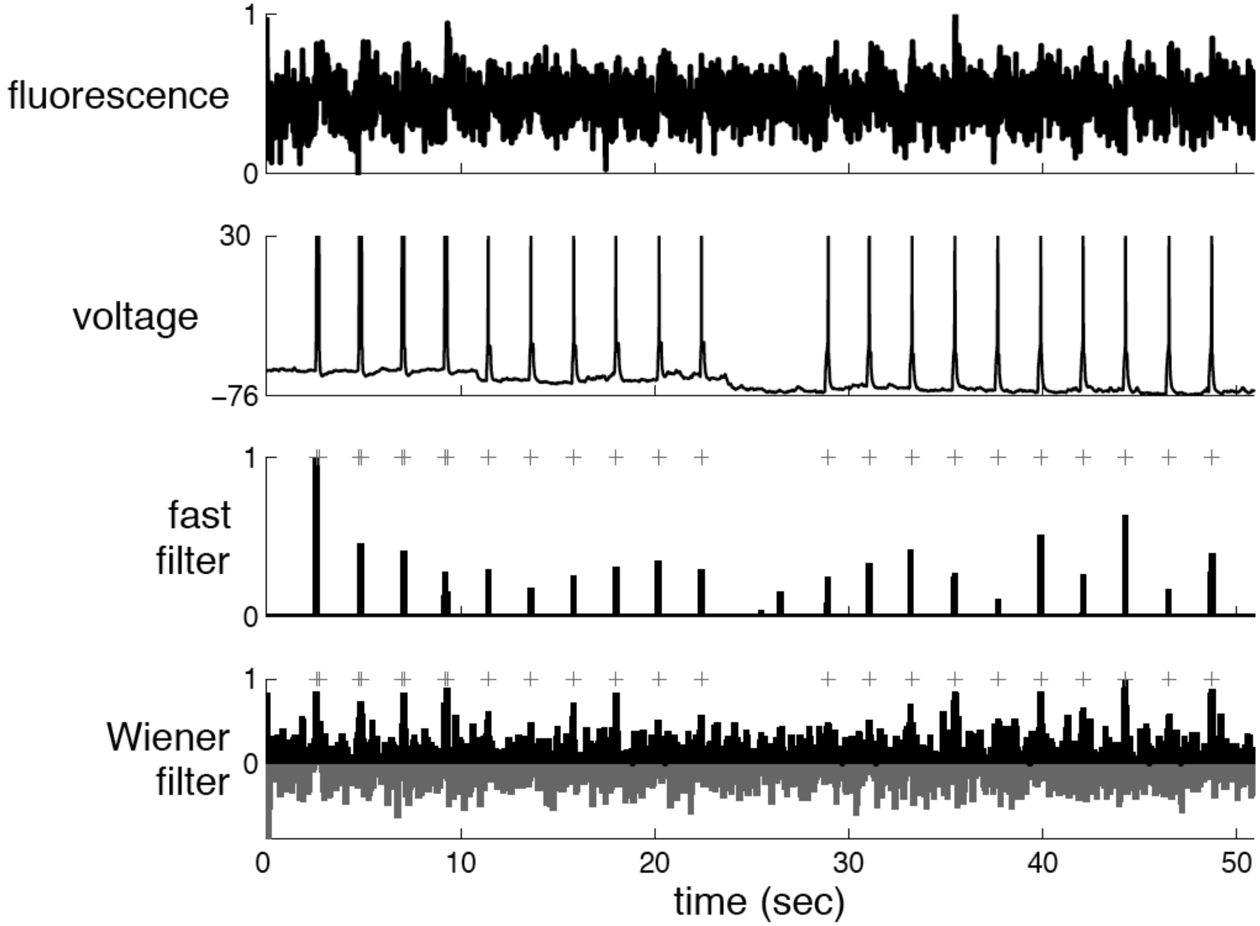


(Ramirez et al., 2011)

Part 3: circuit inference



Challenge: slow, noisy calcium data



First-order model:

$$C_{t+dt} = C_t - dtC_t/\tau + r_t;\ r_t > 0;\ y_t = C_t + \epsilon_t$$

— $\tau \approx 100$ ms; nonnegative deconvolution problem. Can be solved by $O(T)$ relaxed constrained interior-point optimization (Vogelstein et al., 2010) or sequential Monte Carlo (Vogelstein et al., 2009).

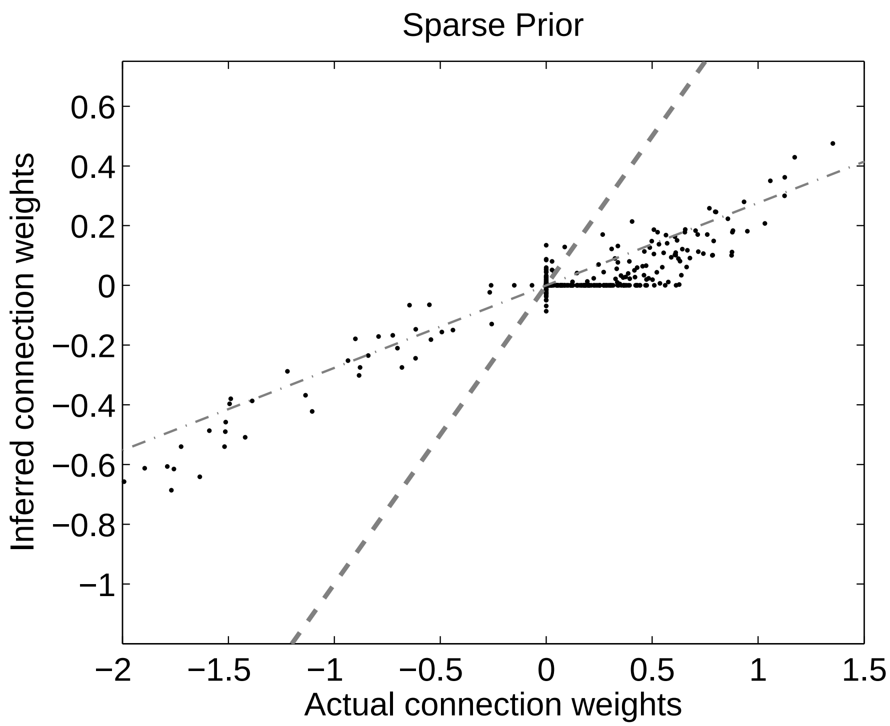
Monte Carlo EM approach

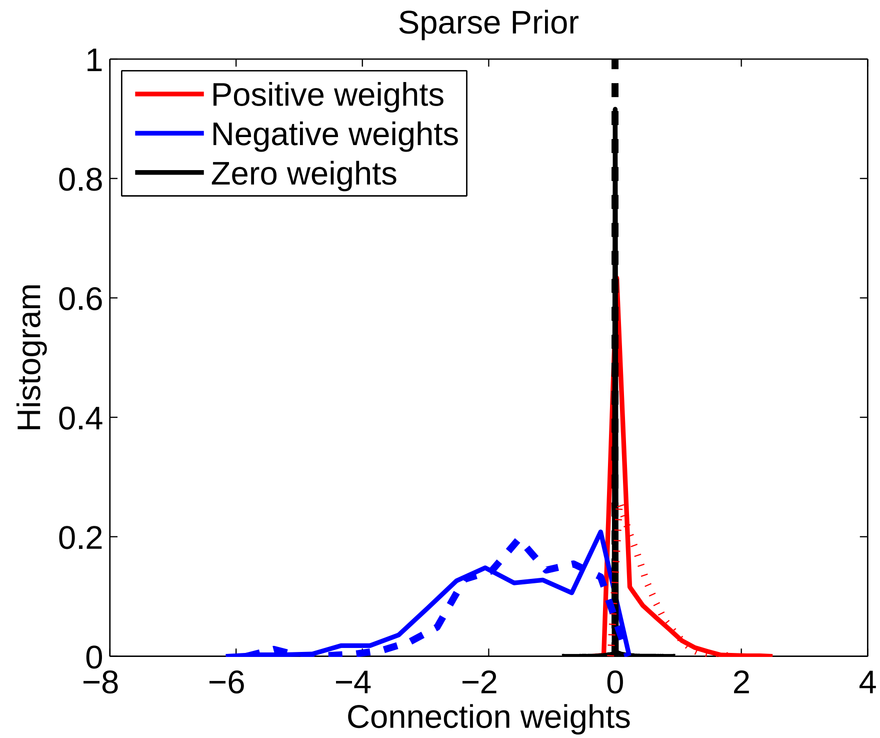
Given the spike times in the network, L_1-penalized likelihood optimization is easy. But we only have noisy calcium observations Y; true spike times are hidden variables. Thus an EM approach is natural.

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

E step is hard part here. Use the fact that neurons interact fairly weakly; thus we need to sample from a collection of weakly-interacting Markov chains (Mishchenko and Paninski, 2010).

Simulated circuit inference





— Connections are inferred with the correct sign in conductance-based integrate-and-fire networks with biologically plausible connectivity matrices (Mishchencko et al., 2009).

Good news: connections are inferred with the correct sign. Exact offline methods are slow; fast approximate methods can be implemented online (Machado et al., 2010).

Optimal control of spike timing

Optimal experimental design and neural prosthetics applications require us 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.

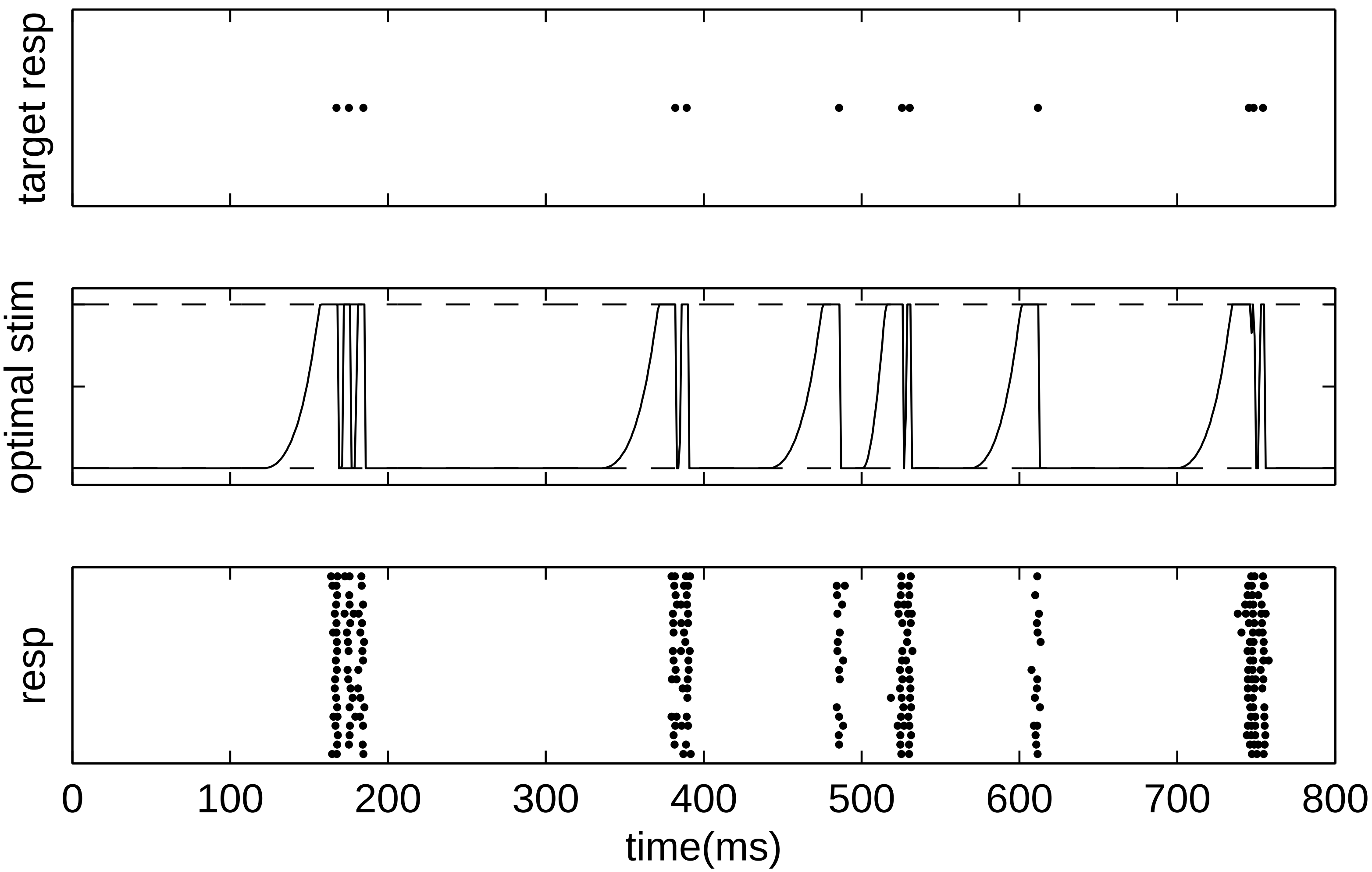
Start with a simple model:

$$\lambda_t \;=\; f(\vec{k} * I_t + h_t).$$

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

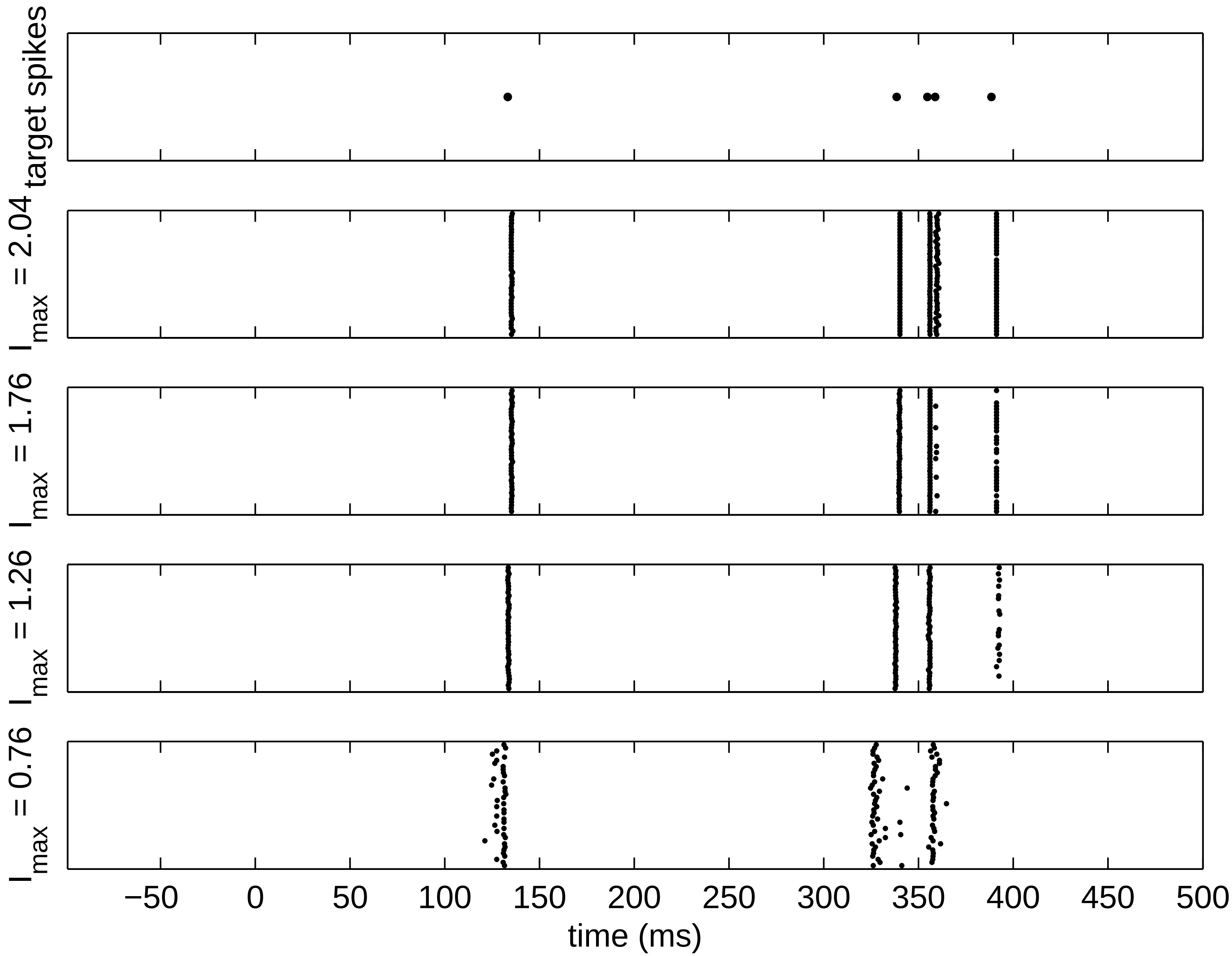
Concave objective function over convex set of possible inputs I_t + Hessian is banded $\implies$ $O(T)$ optimization.

Optimal electrical control of spike timing



Extension to optical stimulation methods is straightforward (Ahmadian et al., 2010a).

Example: intracellular control of spike timing



(Ahmadian et al., 2010a)

Conclusions

- GLM and state-space approaches provide flexible, powerful methods for answering key questions in neuroscience
- Close relationships between encoding, decoding, and experimental design (Paninski et al., 2007)
- Log-concavity, banded matrix methods make computations very tractable
- Experimental methods progressing rapidly; many new challenges and opportunities for breakthroughs based on statistical ideas

References

Ahmadian, Y., Packer, A., Yuste, R., and Paninski, L. (2010a). Fast optimal control of spike trains. *Under review.*

Ahmadian, Y., Pillow, J., and Paninski, L. (2010b). Efficient Markov Chain Monte Carlo methods for decoding population spike trains. *In press, Neural Computation.*

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., Ahrens, M., and Paninski, L. (2006). Efficient estimation of detailed single-neuron models. *Journal of Neurophysiology*, 96:872–890.

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.

Mishchenko, Y. and Paninski, L. (2010). Efficient methods for sampling spike trains in networks of coupled neurons. *In preparation.*

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.

Paninski, L., Pillow, J., and Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. In Cisek, P., Drew, T., and Kalaska, J., editors, *Computational Neuroscience: Progress in Brain Research*. Elsevier.

Pnevmatikakis, E. and Paninski, L. (2011). Fast interior-point inference in high-dimensional sparse, penalized state-space models. *Submitted.*

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.*, In press.

Vogelstein, J., Watson, B., Packer, A., Jedynak, B., Yuste, R., and Paninski, L. (2009). Model-based optimal inference of spike times and calcium dynamics given noisy and intermittent calcium-fluorescence imaging. *Biophysical Journal*, 97:636–655.