Statistical methods for understanding neural computation

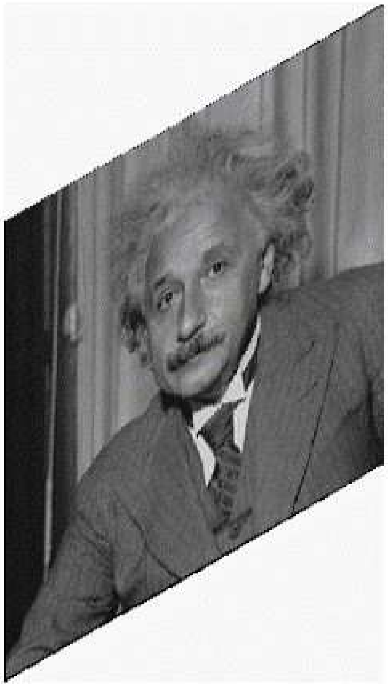
Liam Paninski

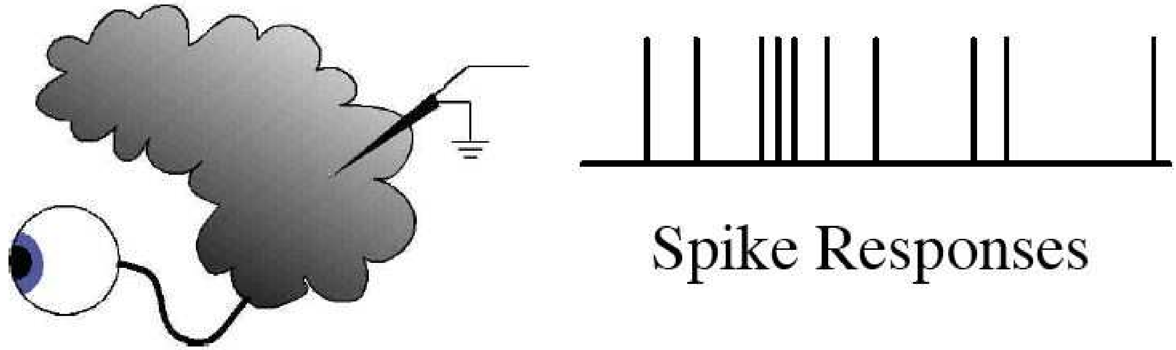
Department of Statistics and Center for Theoretical Neuroscience
Columbia University
http://www.stat.columbia.edu/∼liam
liam@stat.columbia.edu

October 29, 2009

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

The neural code





Basic goal: infer input-output relationship between

- External observables x (sensory stimuli, motor responses...)
- Neural responses r (spike trains, population activity...)

— many basic mechanisms are well-understood up to some unknown parameters: infer $p(r|x, \theta)$.

Several levels of neural data analysis

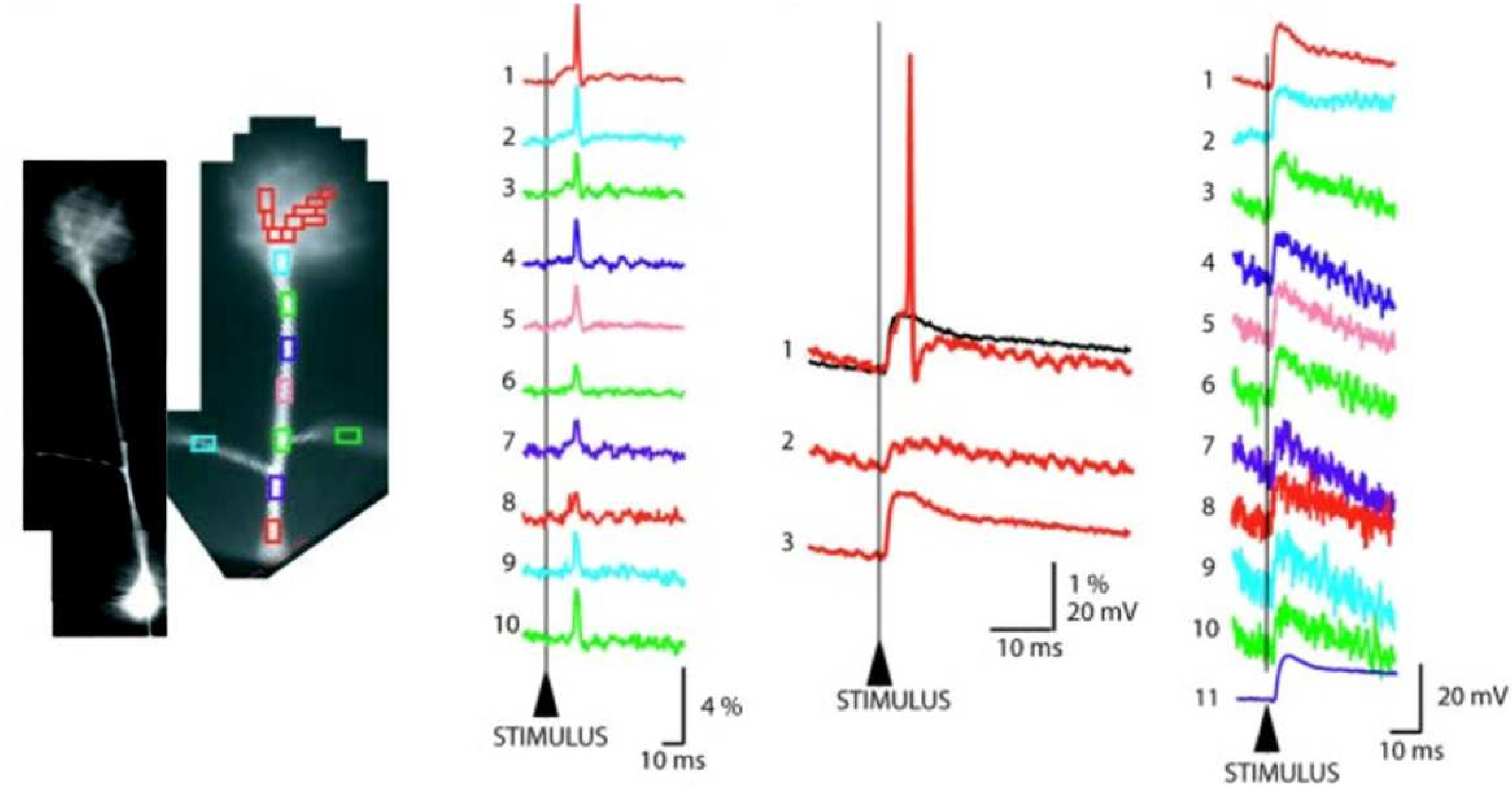
- "Subcellular" level: measurements of intracellular voltage or ionic concentrations (intracellular "patch" electrodes, two-photon imaging)
- "Circuit" level: electrical activity of single neurons or small groups of isolated neurons (multi-electrode recordings, calcium-sensitive microscopy)
- "Systems" level: blood flow or other indirect measurements of electrical activity in coarsely-defined brain areas (fMRI, EEG, MEG...)

Three challenges

1. Reconstructing the full spatiotemporal voltage on a dendritic tree given noisy, intermittently-sampled subcellular measurements
2. Decoding behaviorally-relevant information from multiple neuronal responses
3. Inferring circuit connectivity from large populations of noisily-observed responses

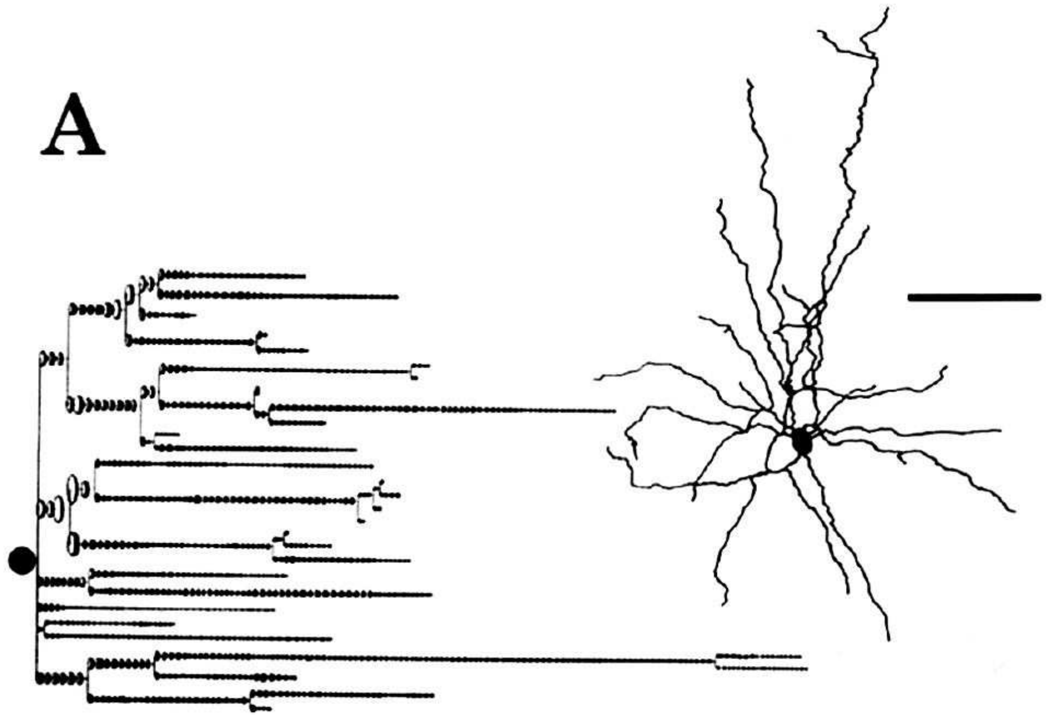
The filtering problem

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



(Djurisic et al., 2004; Knopfel et al., 2006)

Basic paradigm: compartmental models



- write neuronal dynamics in terms of equivalent nonlinear, time-varying RC circuits (Koch, 1999)
- leads to a coupled system of stochastic differential equations

Basic paradigm: the Kalman filter

Variable of interest, q_t, evolves according to a noisy differential equation (Markov process):

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

Make noisy observations:

$$y_t = g(q_t) + \eta_t.$$

We want to infer $E(q_t|Y)$: optimal estimate given observations.

If $f(.)$ and $g(.)$ are linear, and ϵ_t and η_t are Gaussian, then solution is classical: Kalman filter. More general problems: particle filter (Huys and Paninski, 2009).

Basic Kalman filter requires $O(\dim(q)^3 T)$ time. Reduction to $O(\dim(q)T)$ by exploiting tree structure of dendrite (Paninski, 2009).

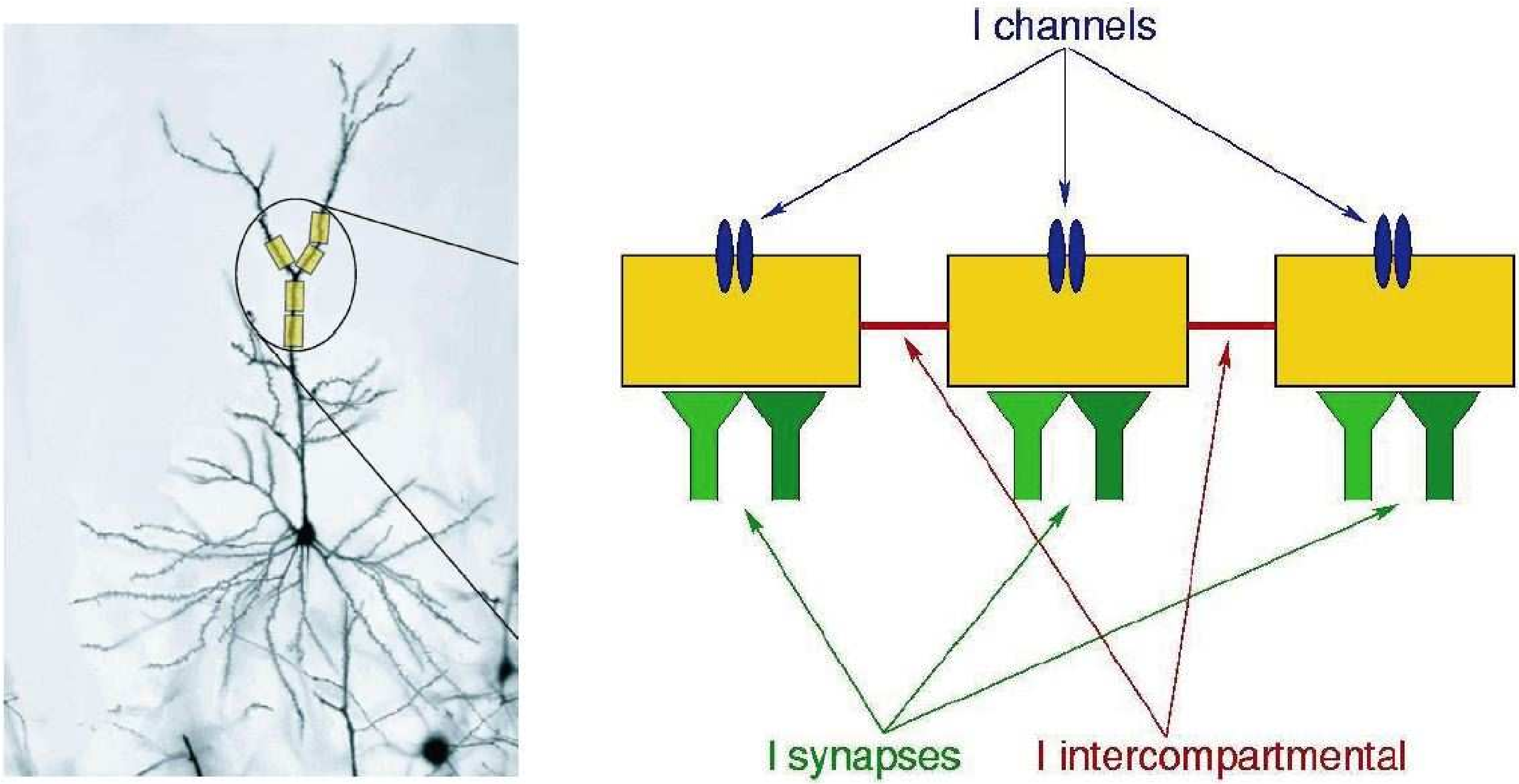
Example: inferring voltage from subsampled observations

(Loading low-rank-speckle.mp4)

Example: summed observations

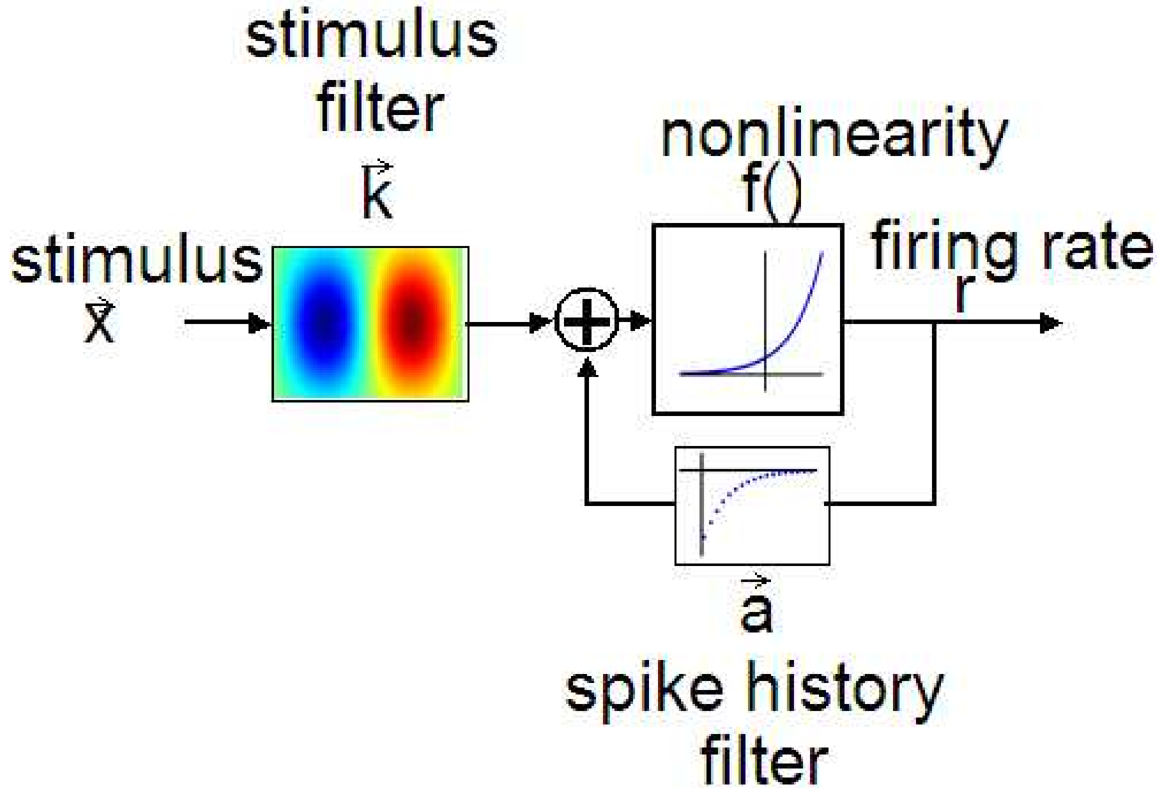
(Loading low-rank-horiz.mp4)

Application: inferring biophysical parameters



Given the spatiotemporal voltage $V(x, t)$, it turns out that we can estimate these biophysical parameters via standard convex nonnegative regression methods (Huys and Paninski, 2009).

Part 2: modeling spike trains

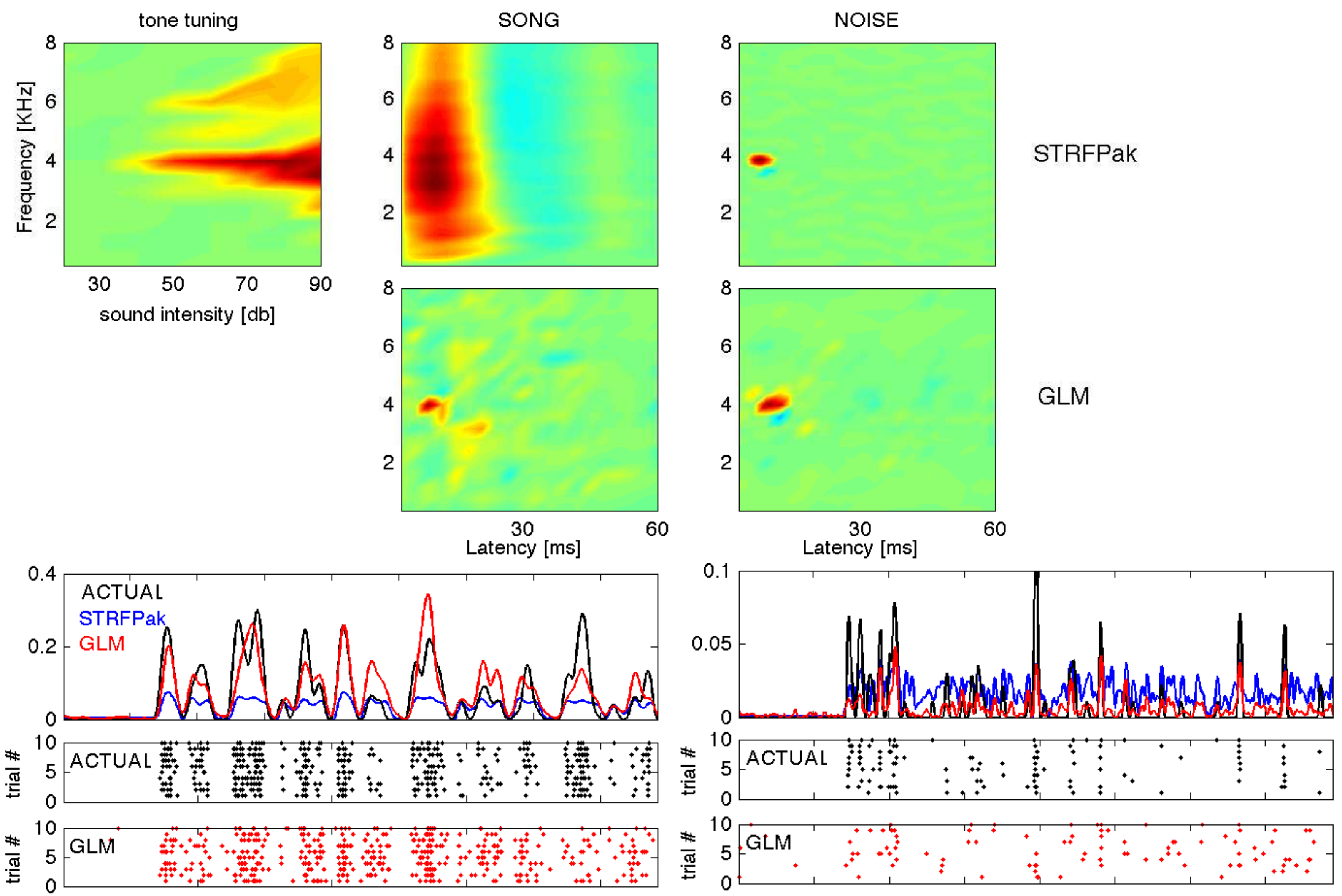


$$p(r_t = 1) = \lambda_t dt$$

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

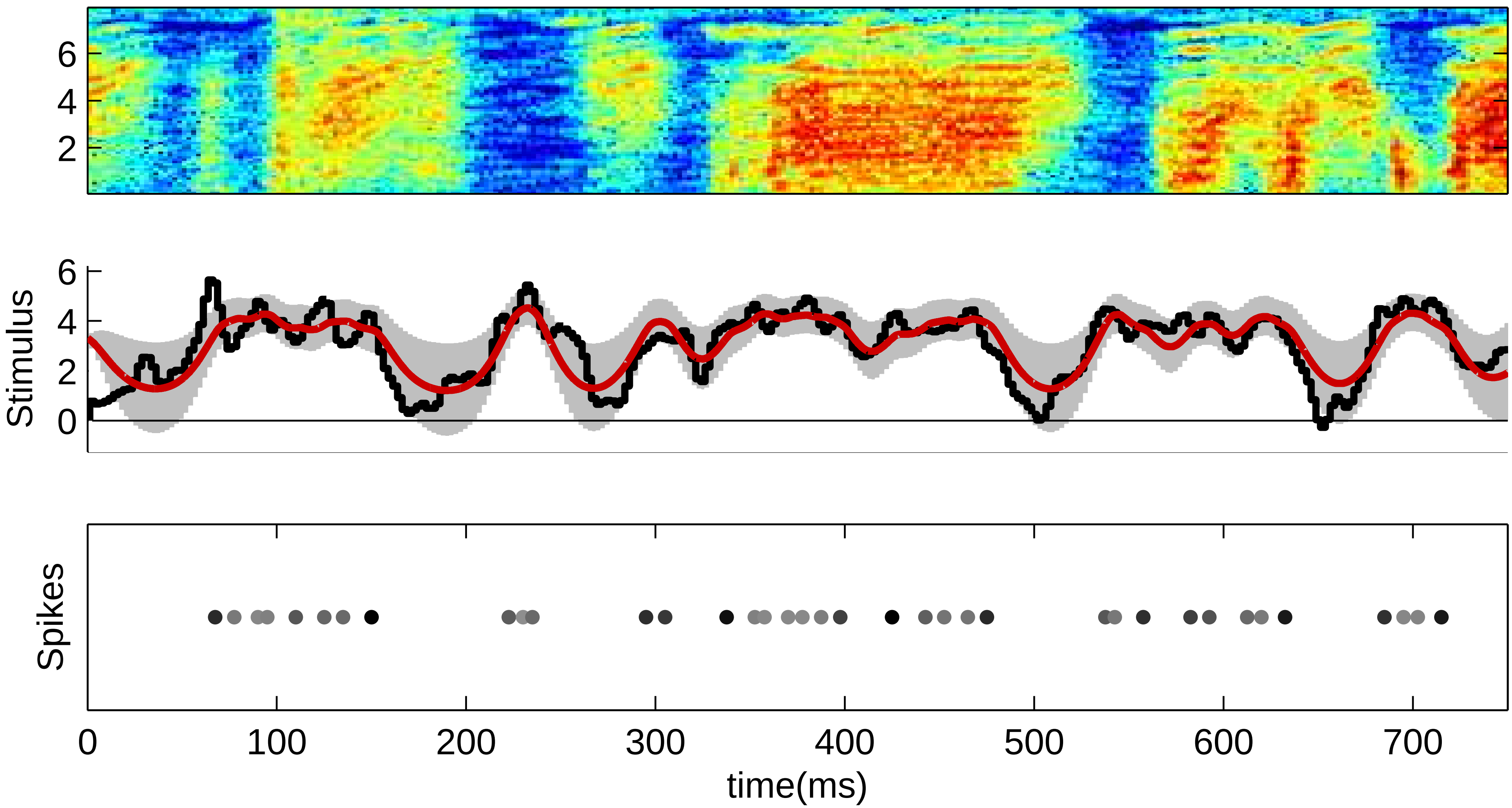
Generalized linear model: log-likelihood is concave $\Longrightarrow$ easy to estimate parameters via maximum likelihood.

Predicting songbird auditory responses



(Calabrese, Schneider, Woolley et al. 2009)

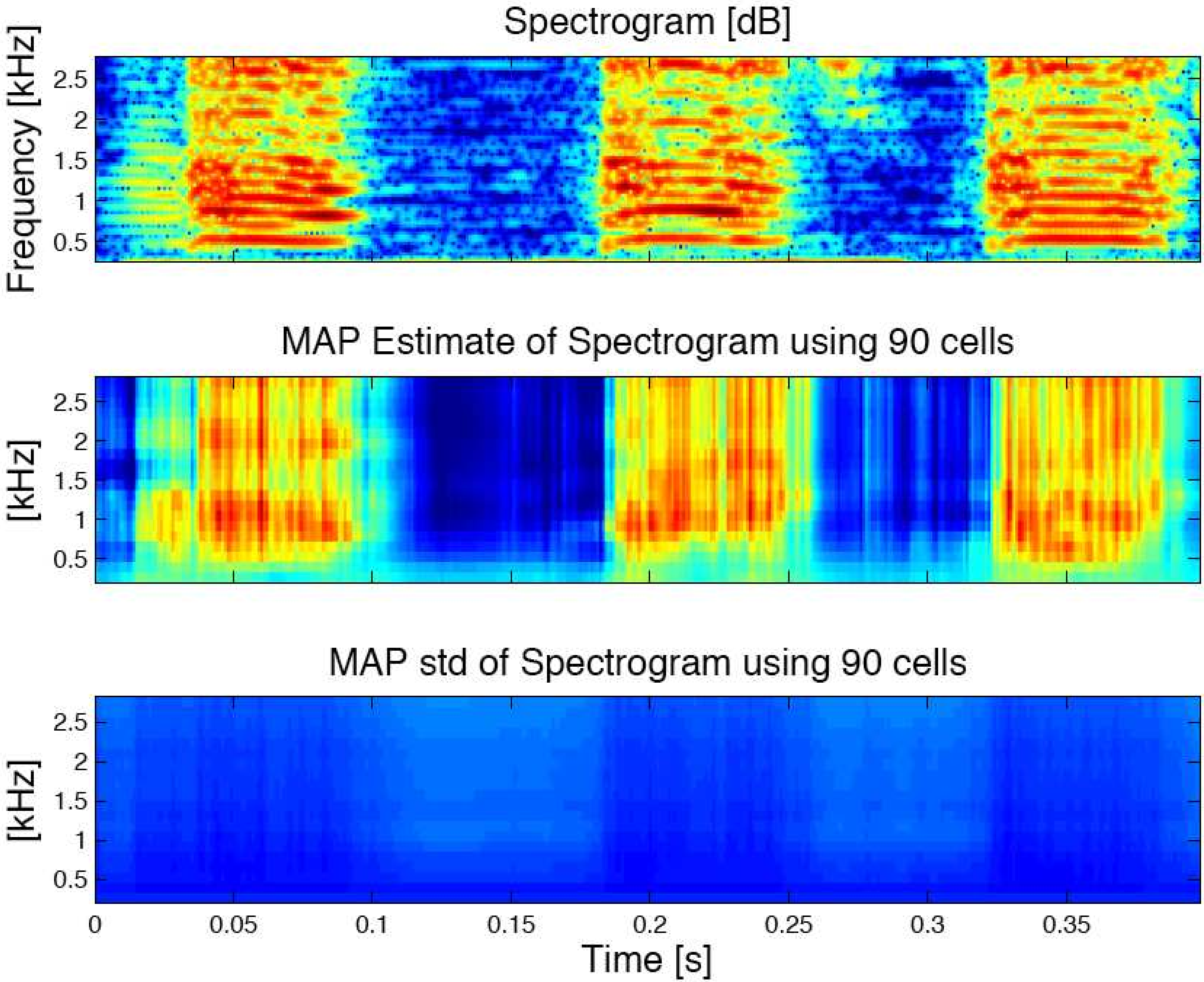
Application: fast optimal decoding



Concave optimization: $\log p(\vec{x}|r, \vec{\theta}) = \log p(r|\vec{x}, \vec{\theta}) + \log p(\vec{x})$ w.r.t. $\vec{x}$.

Banded Hessian $\implies$ fast computation: $O(T)$ time (Pillow et al., 2009).

Decoding a full song



Application: optimal stimulus design

Idea: we have full control over the stimuli we present. Can we choose stimuli $\vec{x}_t$ to maximize the informativeness of each trial?

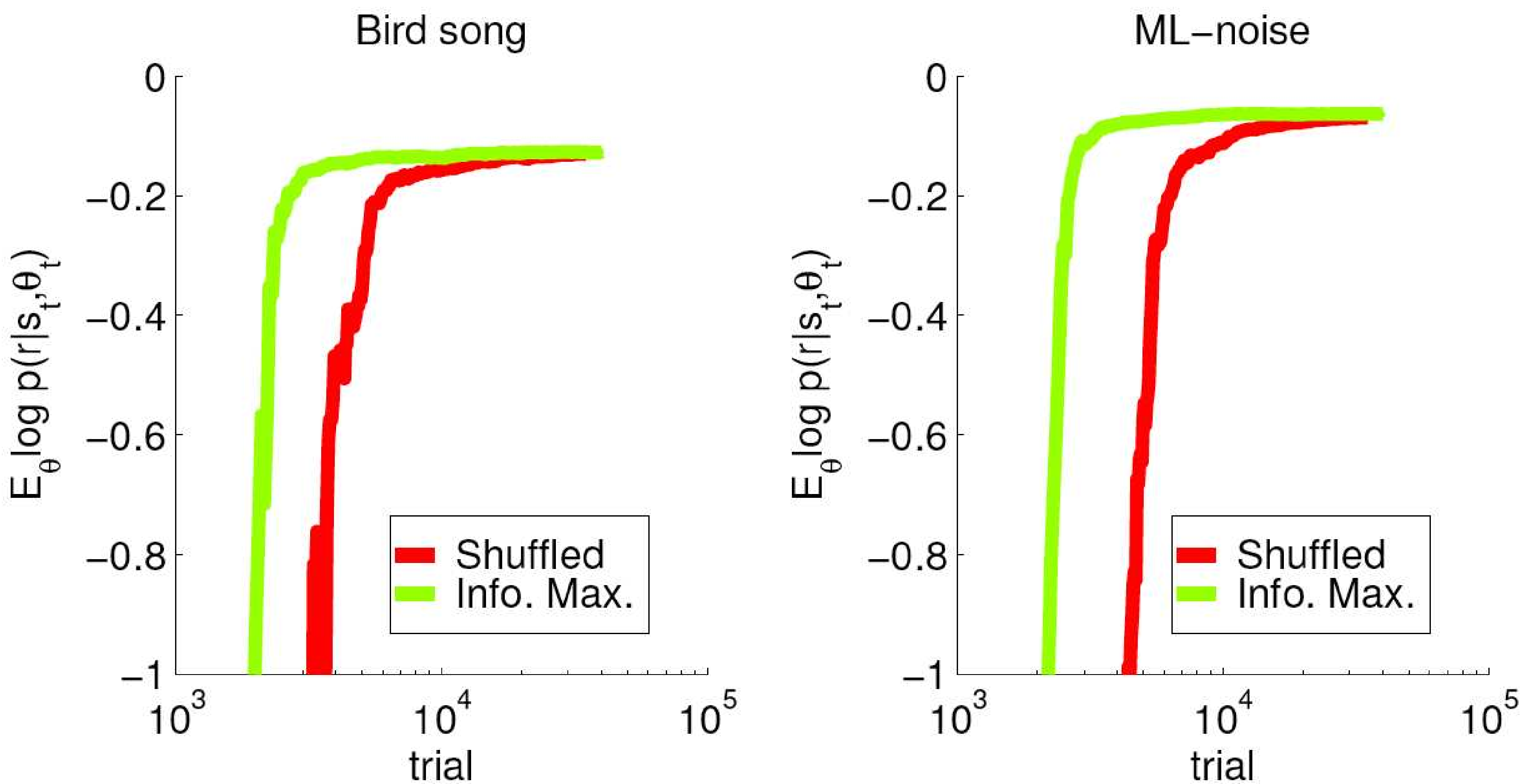
— More quantitatively, optimize $I(r_t; \theta | \vec{x}_t)$ with respect to $\vec{x}_t$.

Maximizing $I(r_t; \theta; \vec{x}_t) \implies$ minimizing uncertainty about θ.

In general, very hard to do: high-d integration over θ to compute $I(r_t; \theta | \vec{x}_t)$, high-d optimization to select best $\vec{x}_t$.

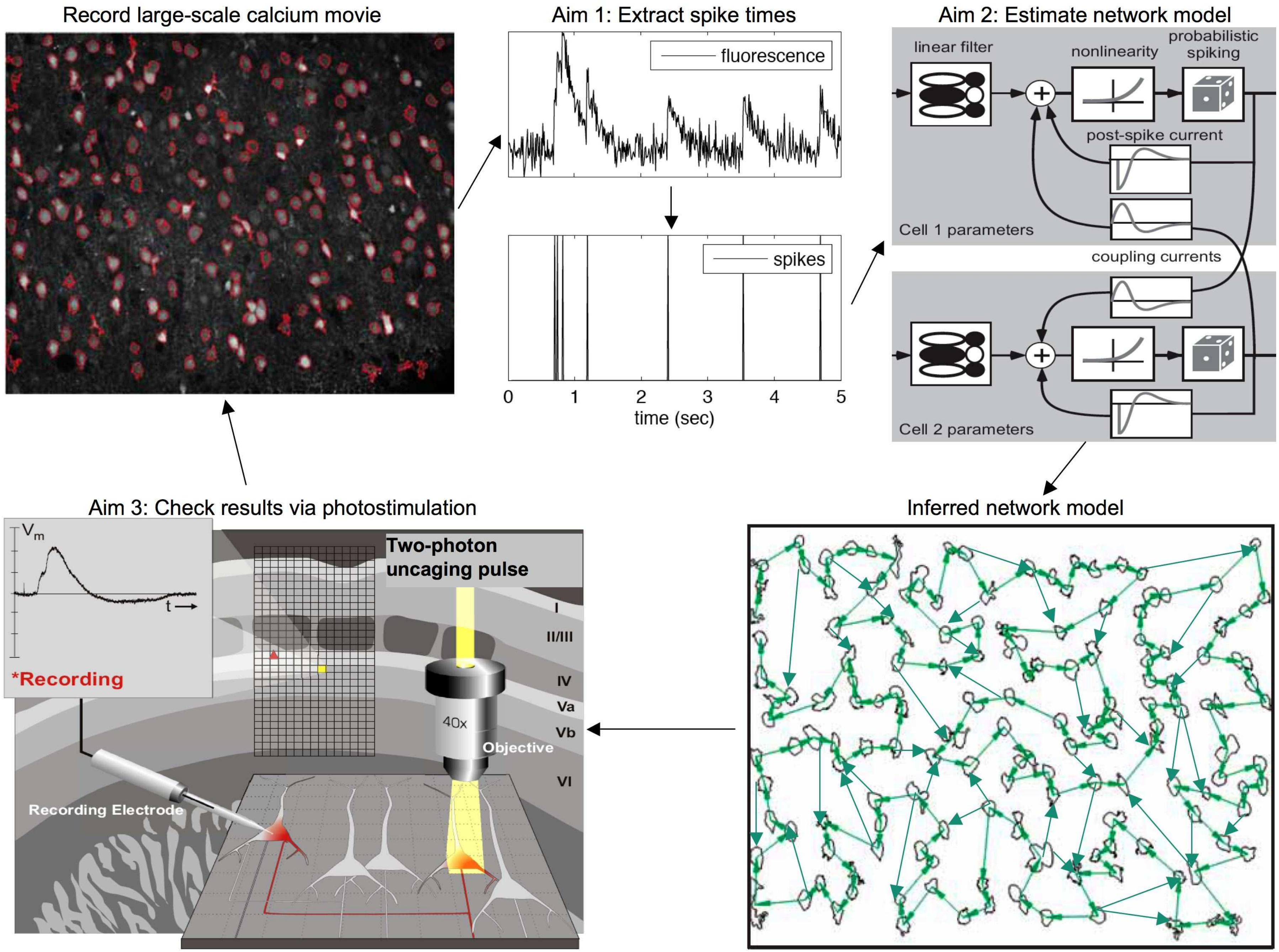
GLM setting + low-rank matrix methods make this surprisingly tractable: $O(\dim(\theta)^2)$ computation (Lewi et al., 2009).

Application to songbird data: choosing an optimal stimulus sequence

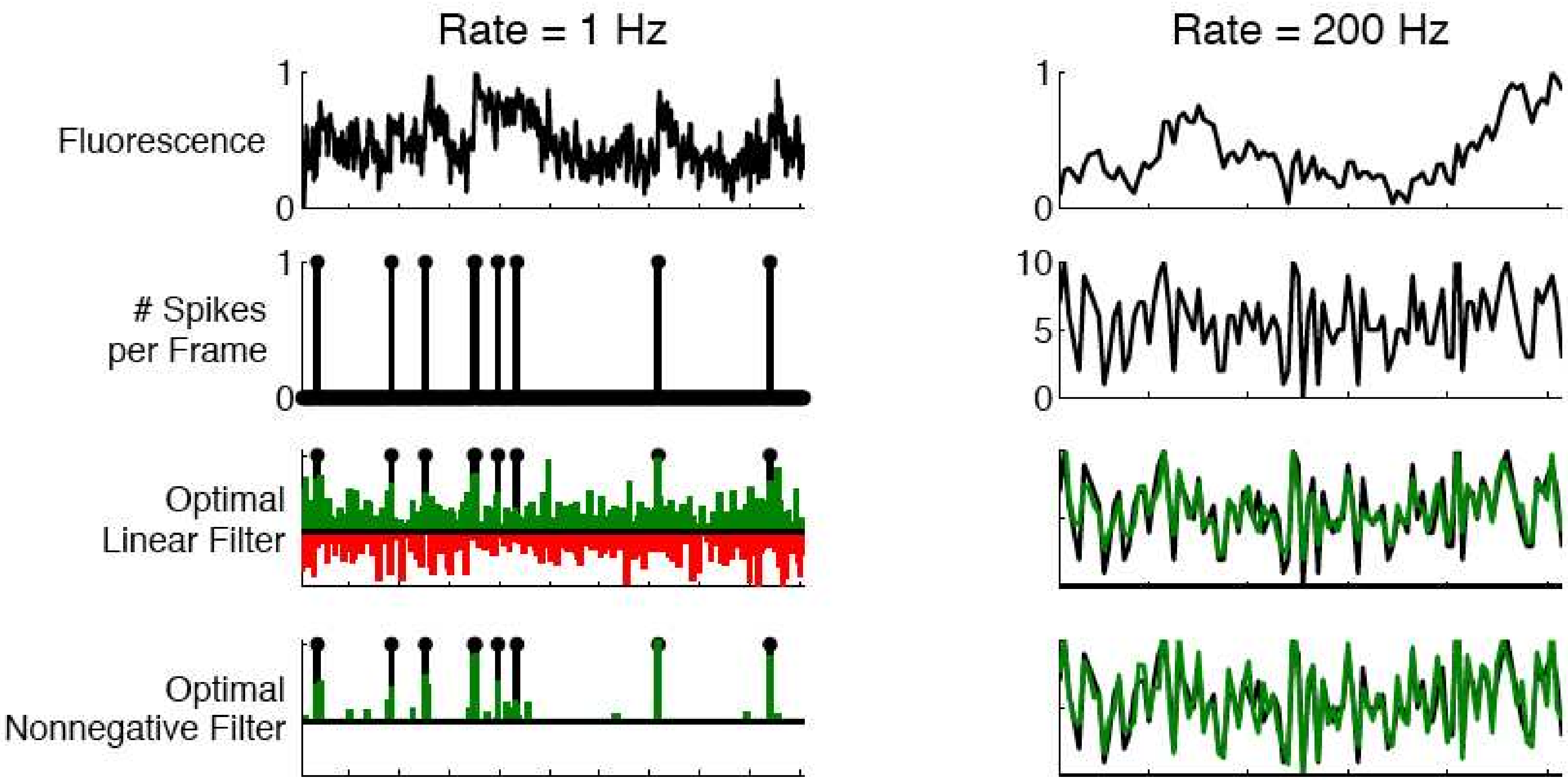


— infomax speeds convergence by a factor of three or more.

Part 3: circuit inference



Challenge: slow, noisy calcium data

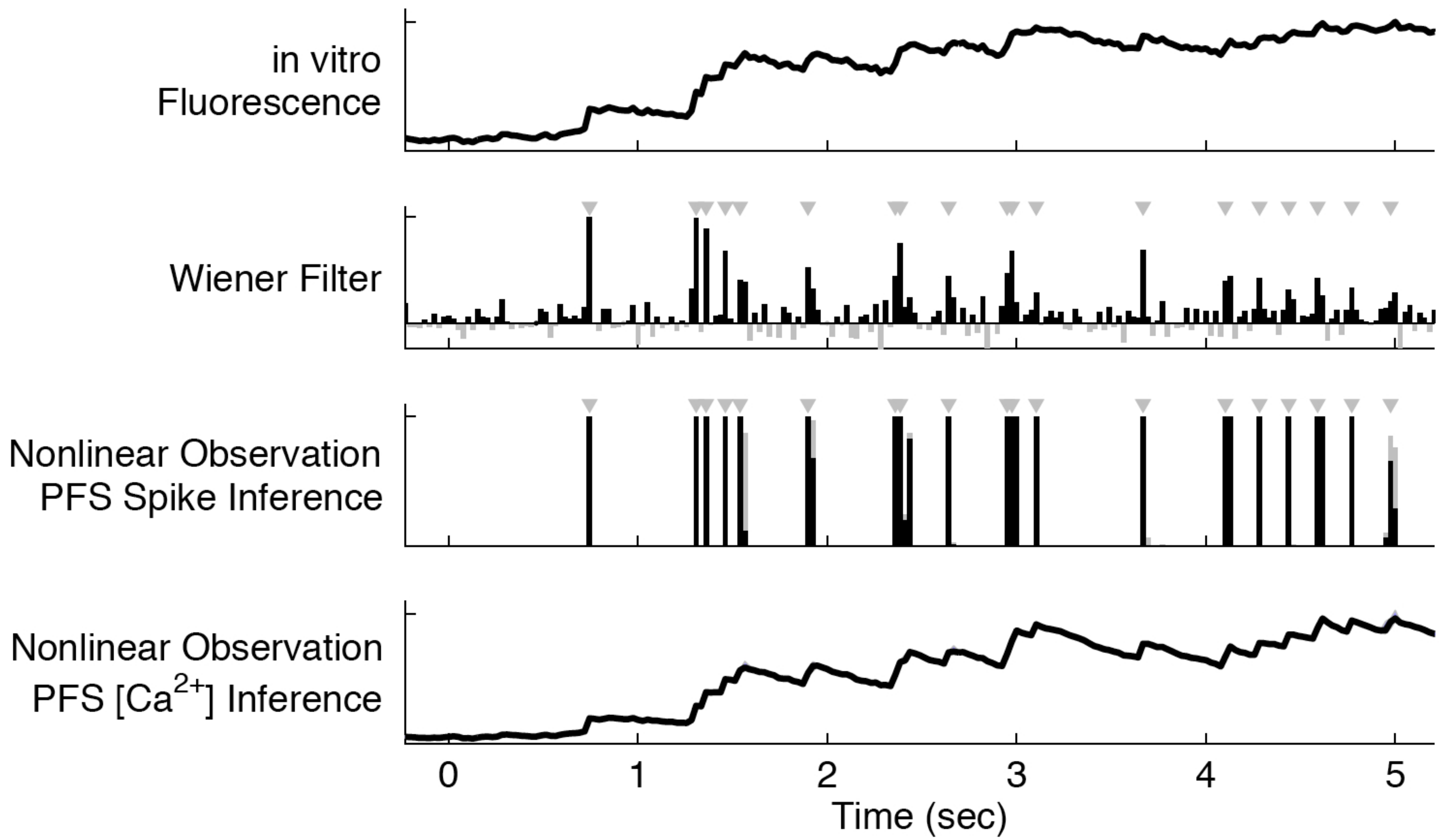


First-order model:

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

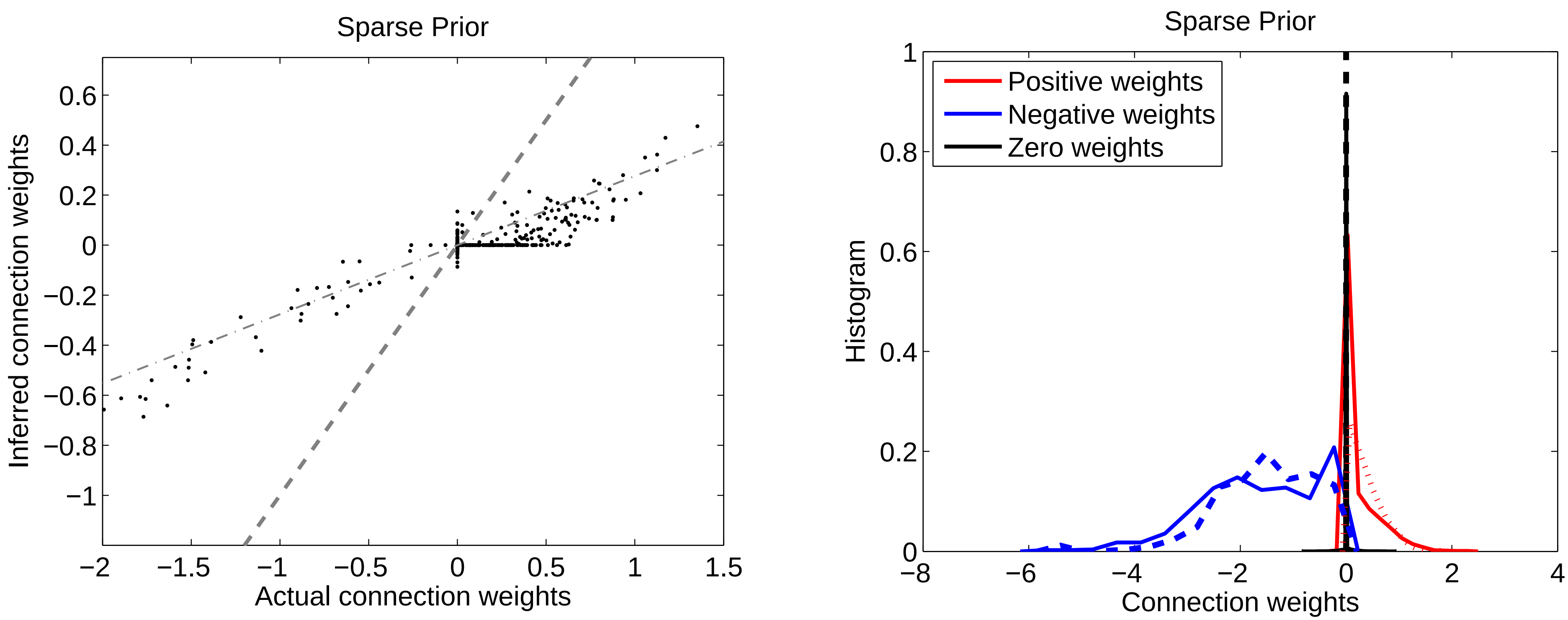
— $\tau \approx 100$ ms; nonnegative deconvolution problem. Can be solved by $O(T)$ relaxed constrained optimization methods (Vogelstein et al., 2008) or sequential Monte Carlo (Vogelstein et al., 2009).

Particle filter can extract spikes from saturated recordings



— saturation model: $y_t = g(C_t) + \epsilon_t$ (Vogelstein et al., 2009)

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).

Last example: 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?

Inputs are constrained: bounds on injected current magnitude, or laser power.

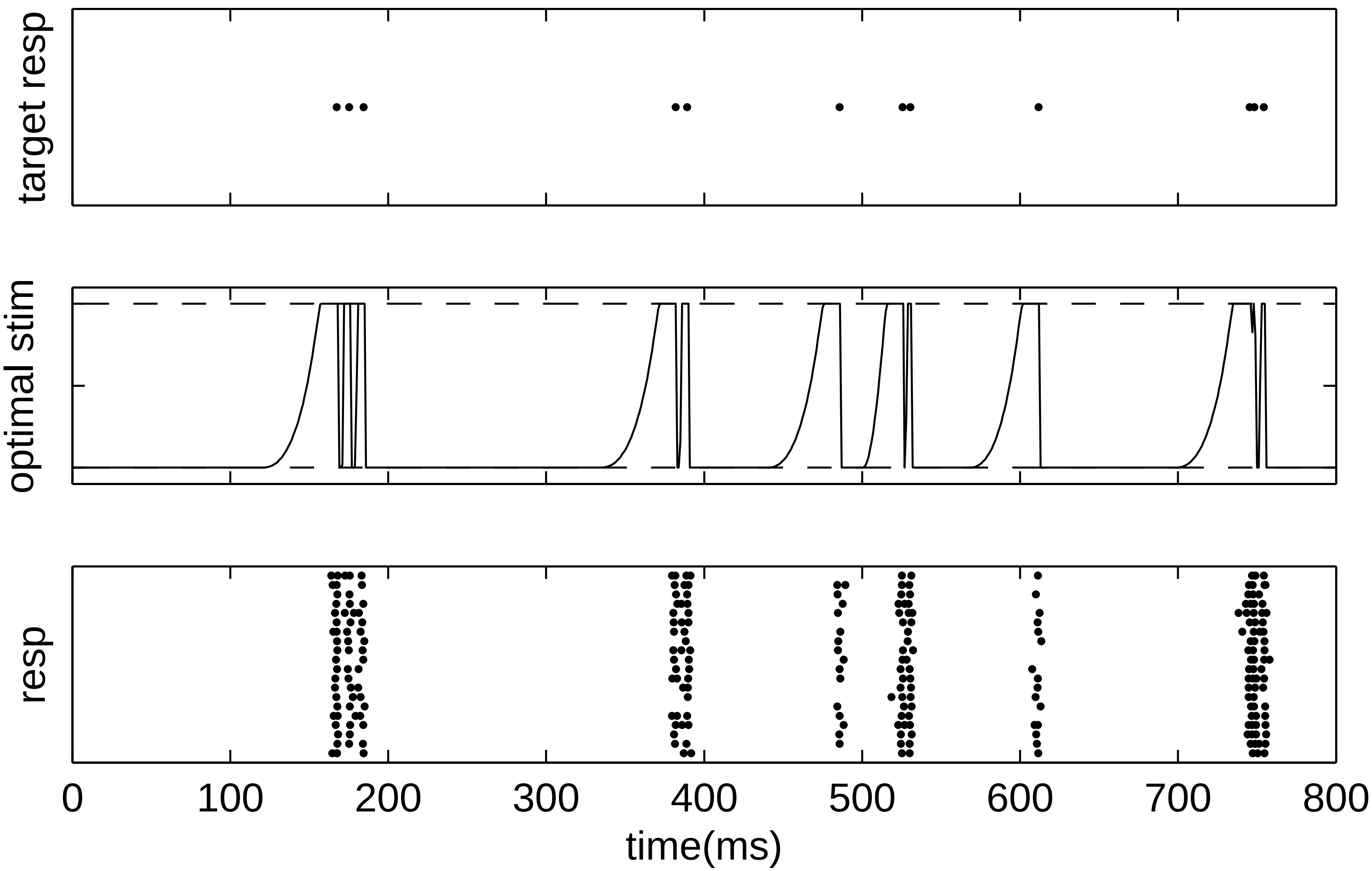
Start with a simple model:

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

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

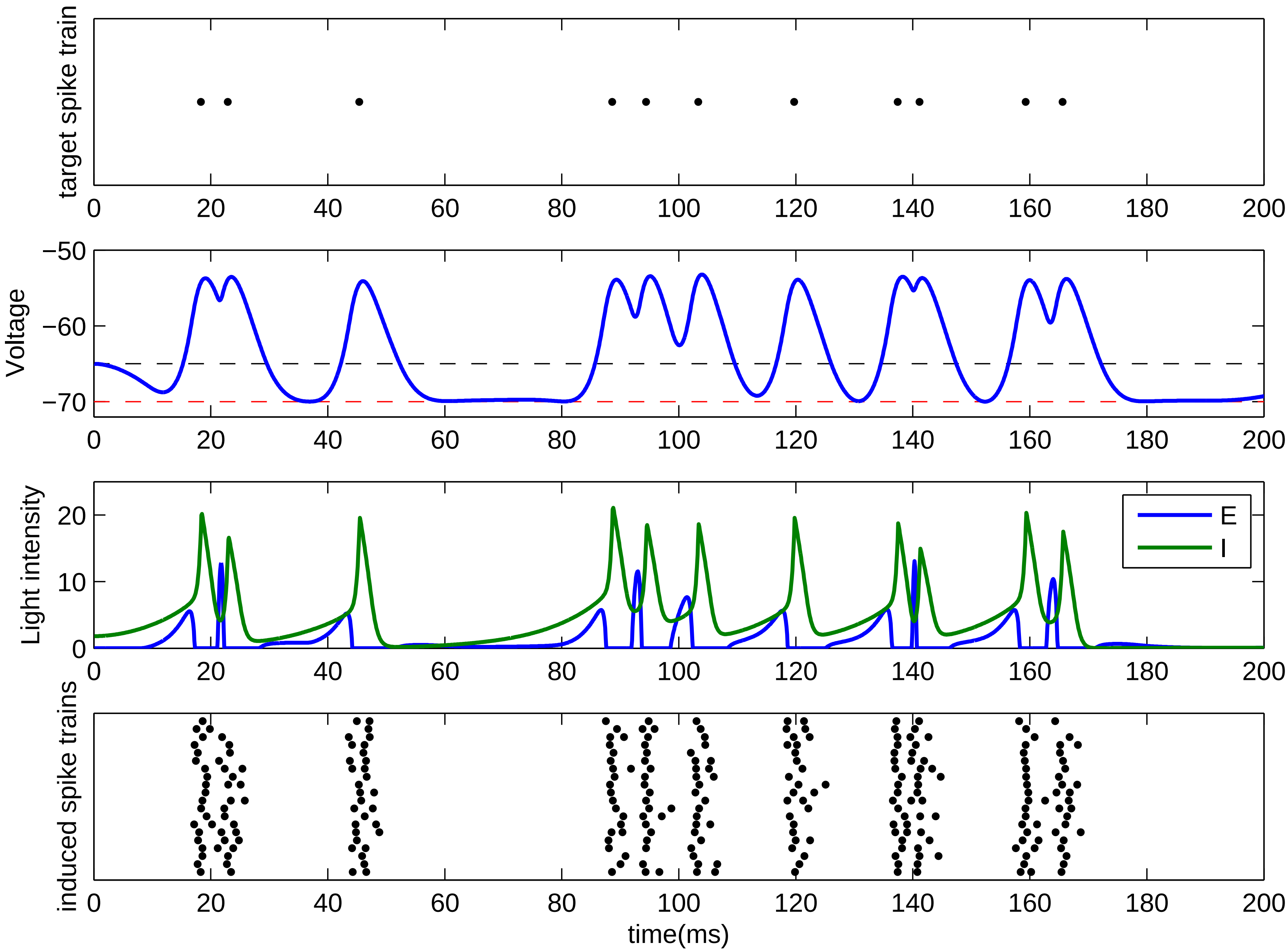
Concave objective function + convex set of inputs I_t + Hessian is banded $\Longrightarrow$ $O(T)$ optimization (Ahmadian et al., 2009).

Optimal electrical control of spike timing



Simulated data; experiments in progress...

Optical conductance-based control of spiking



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 applications of statistical ideas

References

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.

Koch, C. (1999). *Biophysics of Computation*. Oxford University Press.

Lewi, J., Butera, R., and Paninski, L. (2009). Sequential optimal design of neurophysiology experiments. *Neural Computation*, 21:619–687.

Paninski, L. (2009). Fast Kalman filtering on dendritic trees. *In progress*.

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.

Pillow, J., Ahmadian, Y., and Paninski, L. (2009). Model-based decoding, information estimation, and change-point detection in multi-neuron spike trains. *Under review, Neural Computation*.

Vogelstein, J., Babadi, B., Watson, B., Yuste, R., and Paninski, L. (2008). Fast nonnegative deconvolution via tridiagonal interior-point methods, applied to calcium fluorescence data. *Statistical analysis of neural data (SAND) conference*.

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*, In press; http://www.stat.columbia.edu/~liam/research/abstracts/vogelstein-bj08-abs.html.