Statistical models for neural encoding, decoding, and optimal stimulus design

Liam Paninski

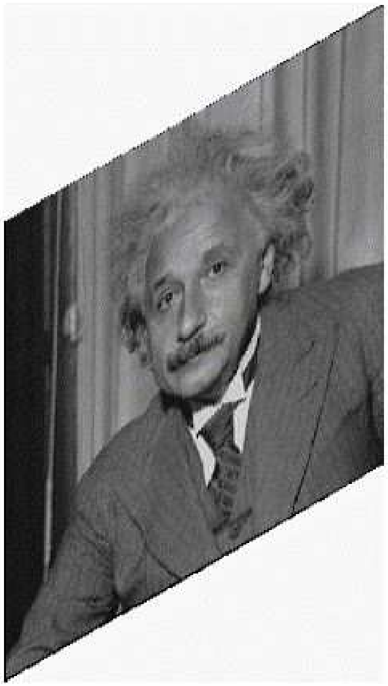
Department of Statistics and Center for Theoretical Neuroscience
Columbia University
http://www.stat.columbia.edu/∼liam
liam@stat.columbia.edu

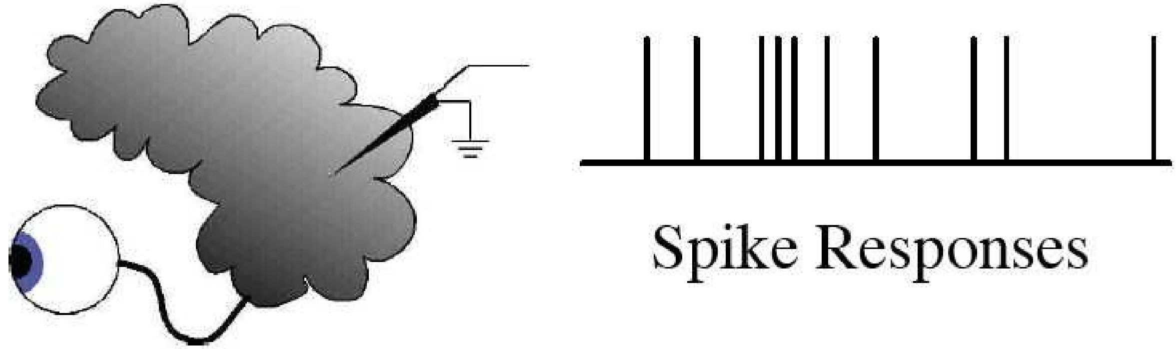
June 1, 2009

— with J. Pillow (UT Austin), E. Simoncelli (NYU), E.J. Chichilnisky (Salk), J. Lewi (Georgia Tech), Y. Ahmadian, S. Woolley (Columbia).

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

The neural code



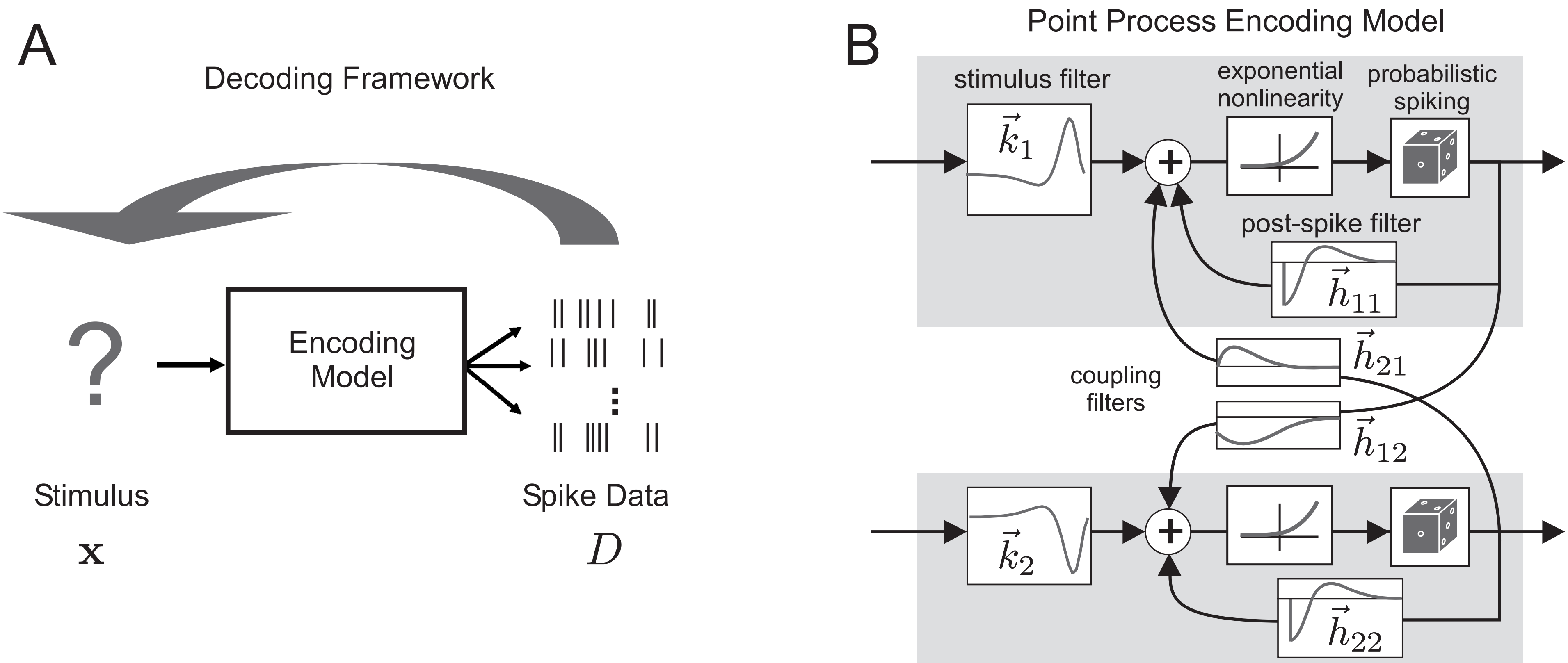


Input-output relationship between

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

Probabilistic formulation: $p(y|x)$

Multineuronal point-process GLM



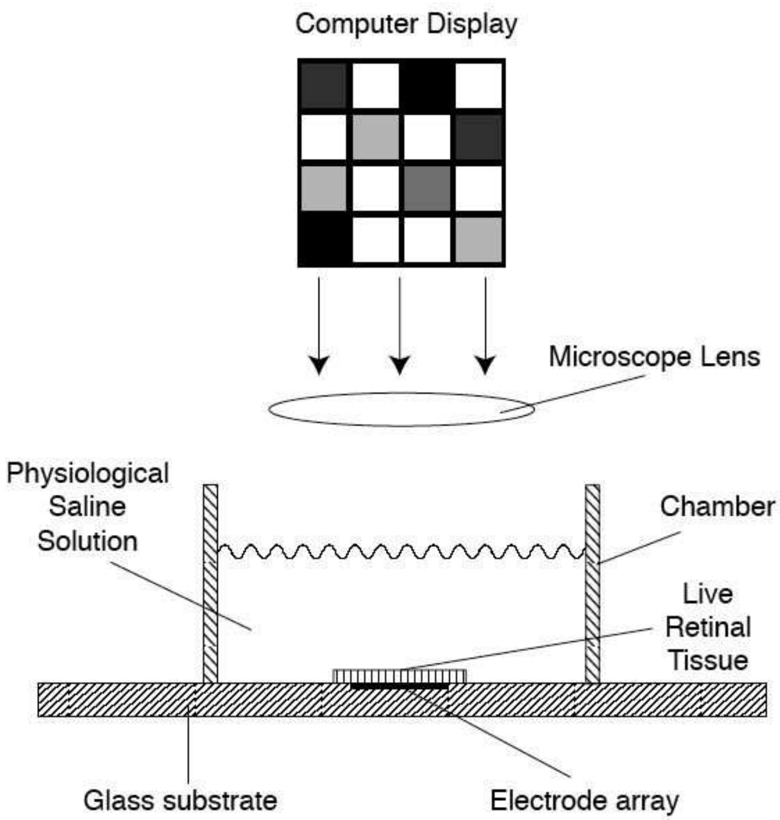
$$\lambda_i(t) = f\left(b + \vec{k}_i \cdot \vec{x}(t) + \sum_{i',j} h_{i',j} n_{i'}(t-j)\right),$$

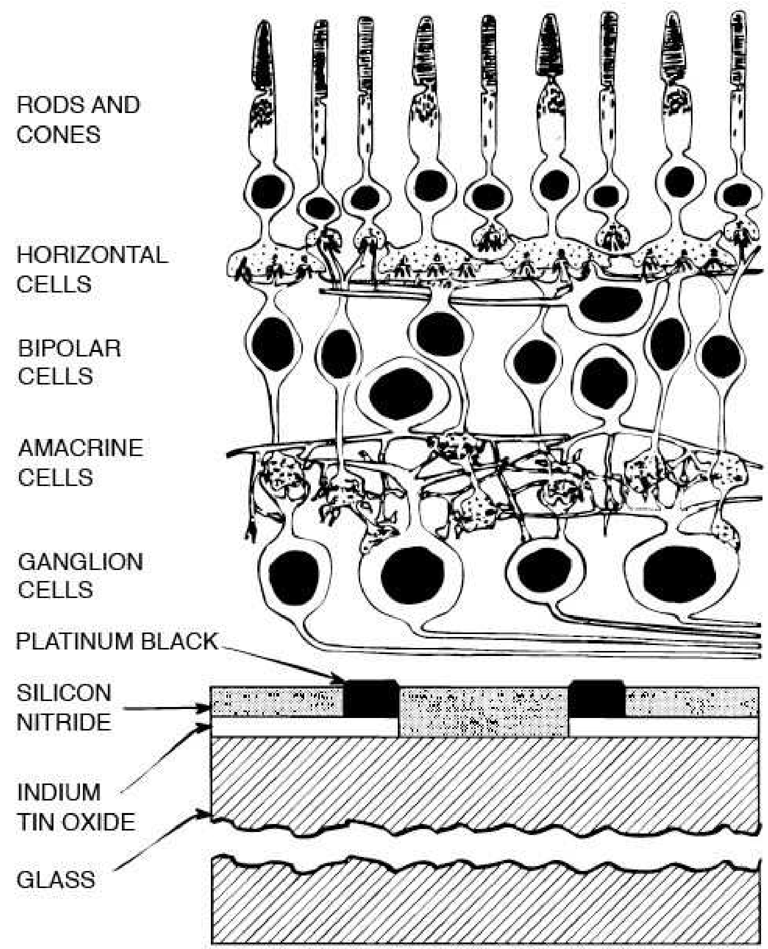
— Fit by L_1-penalized maximum likelihood (concave optimization)
(Brillinger, 1988; Paninski, 2004; Truccolo et al., 2005)

Retinal ganglion neuronal data

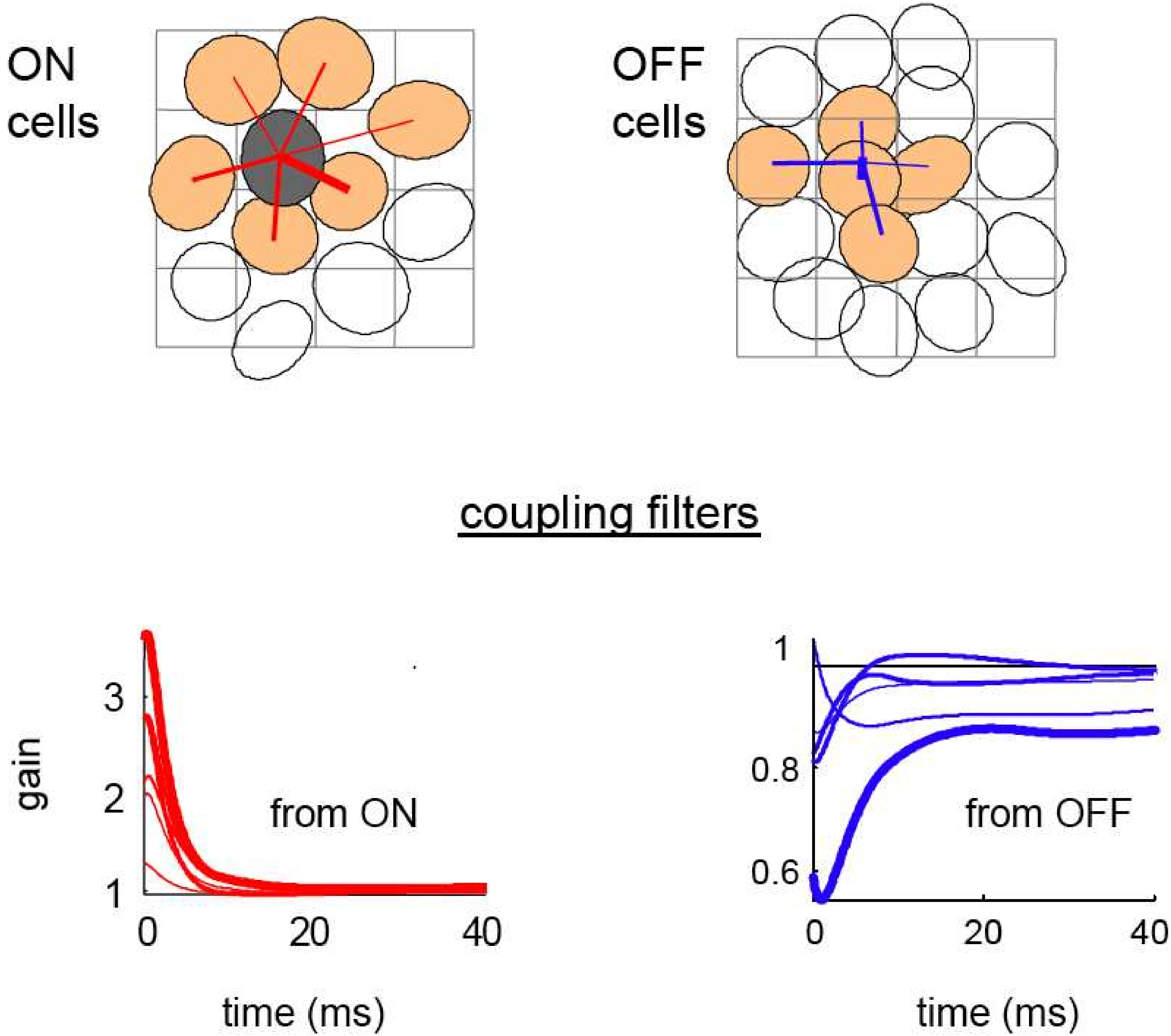
Preparation: dissociated salamander and macaque retina

— extracellularly-recorded responses of populations of RGCs

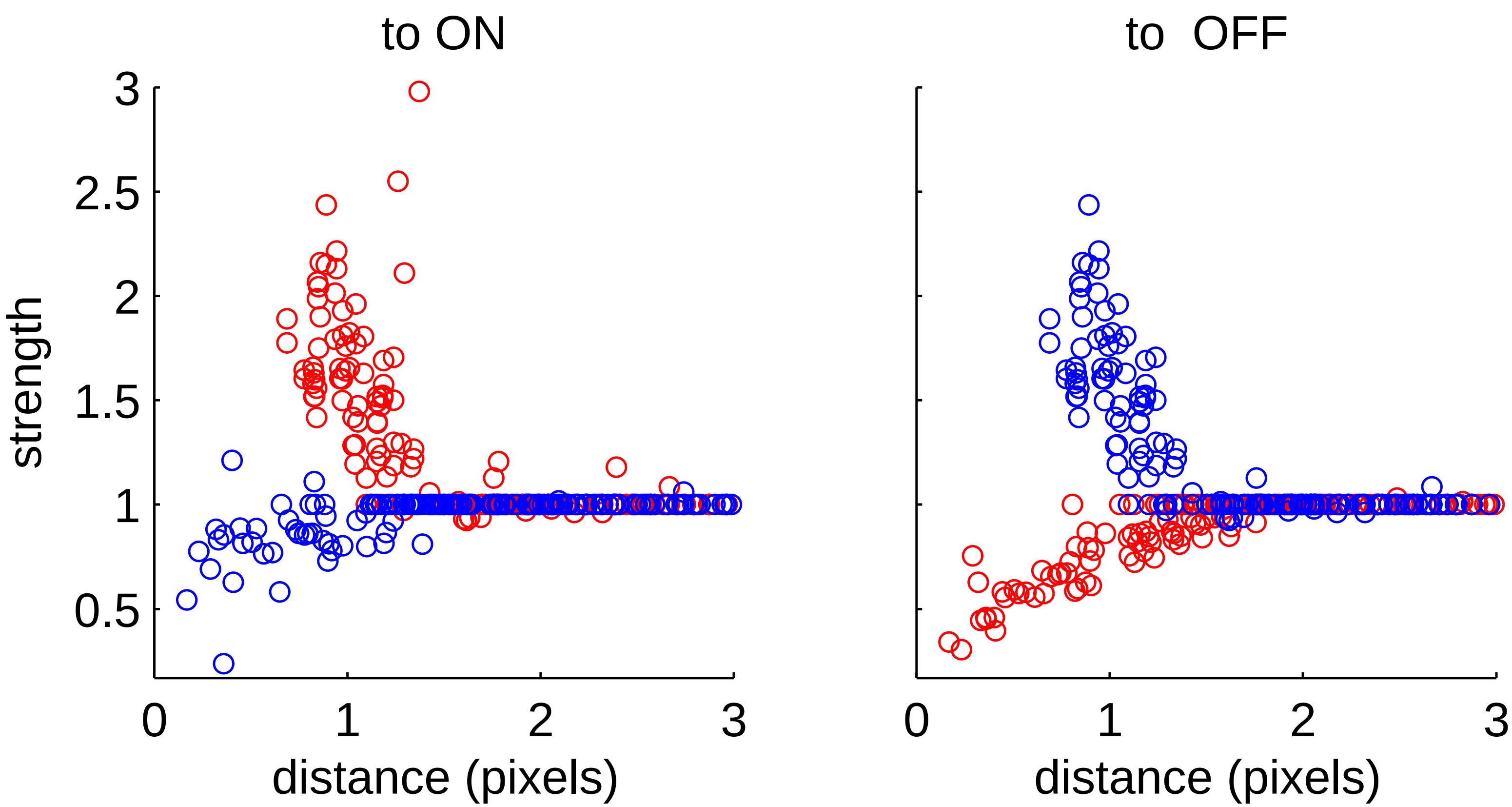




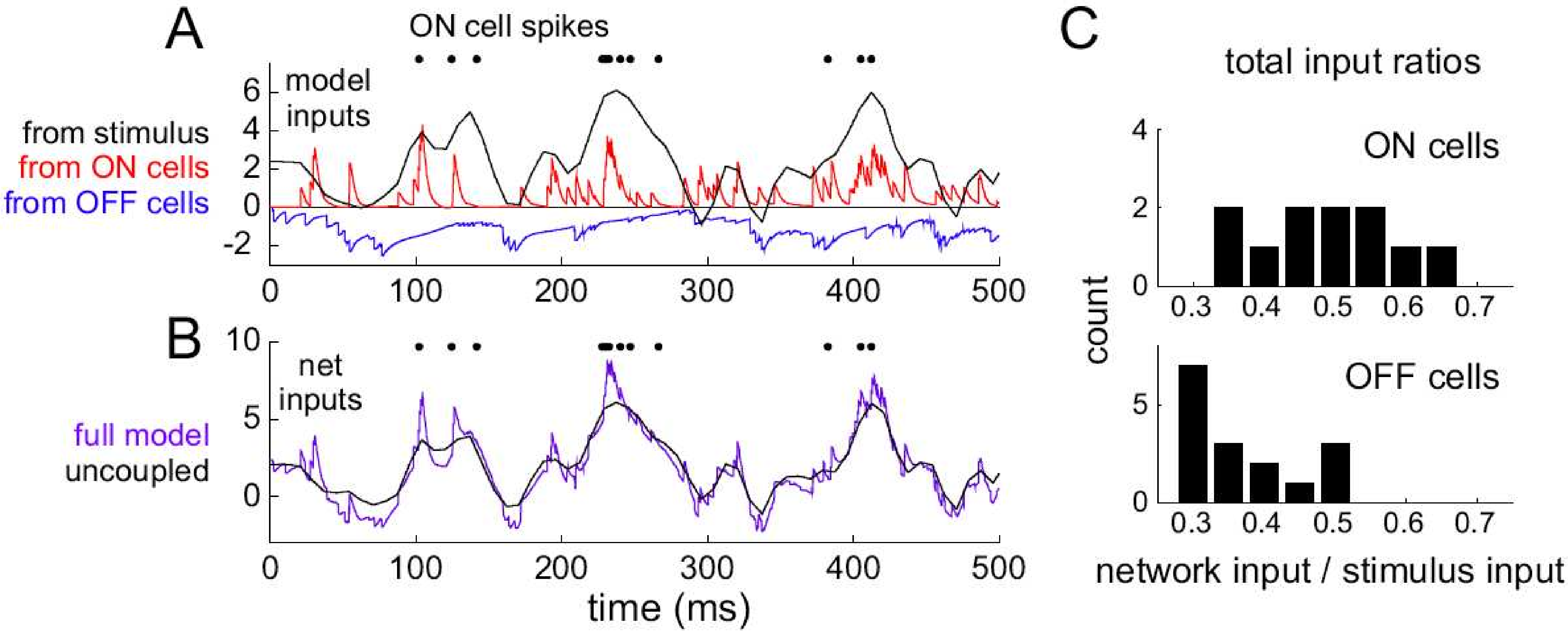
Stimulus: random spatiotemporal visual stimuli (Pillow et al., 2008)


ON
cells
OFF
cells
coupling filters
gain
3
2
1
from ON
0
20
40
time (ms)
1
0.8
0.6
from OFF
0
20
40
time (ms)

Nearest-neighbor connectivity



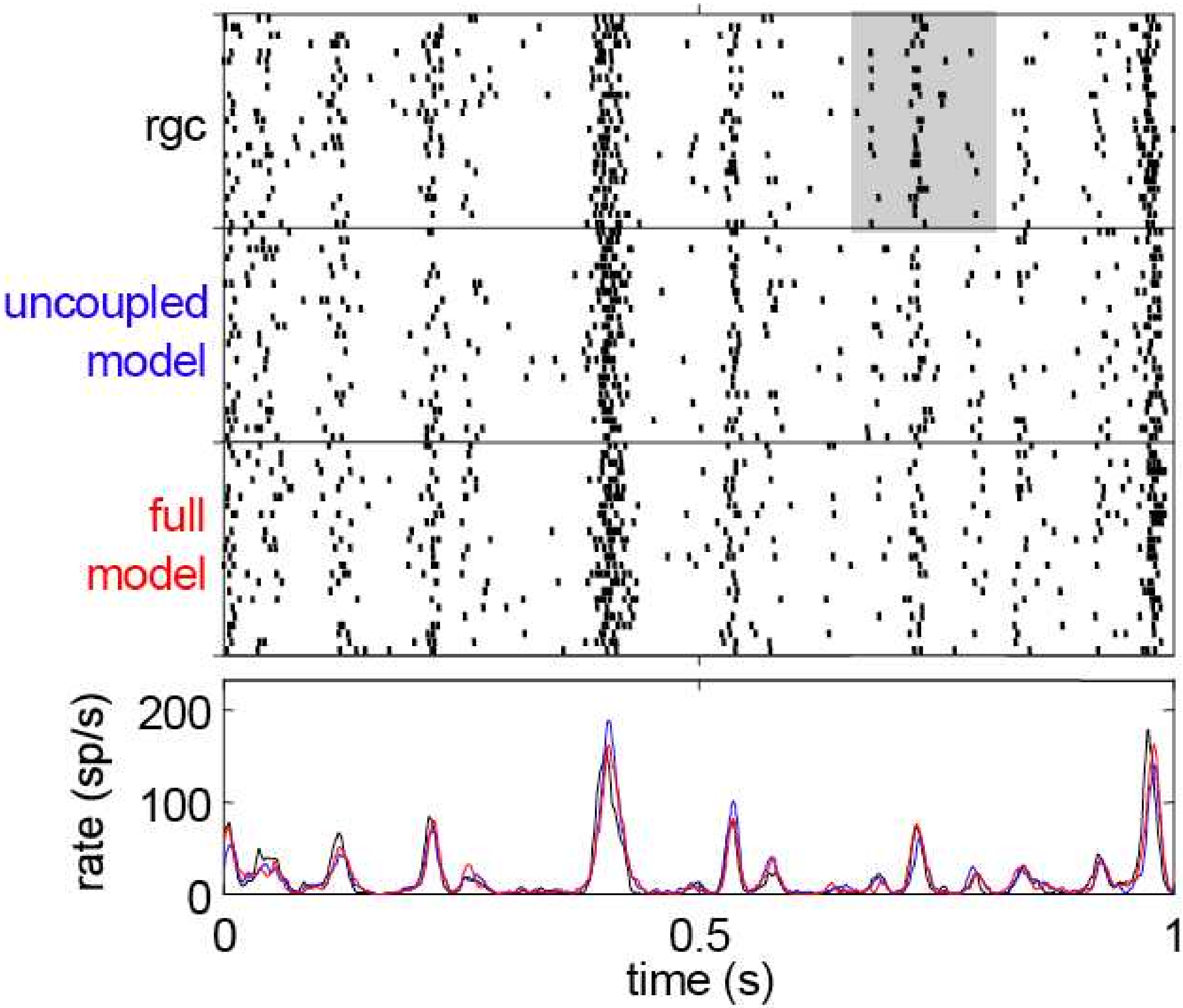
Network vs. stimulus drive

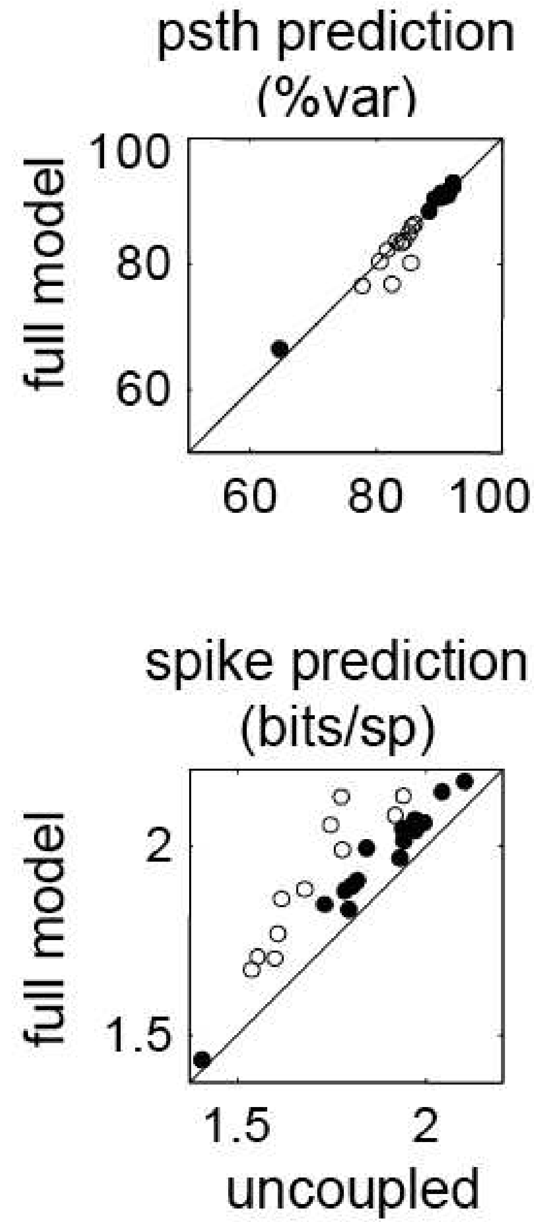


— Network effects are ≈ 50% as strong as stimulus effects

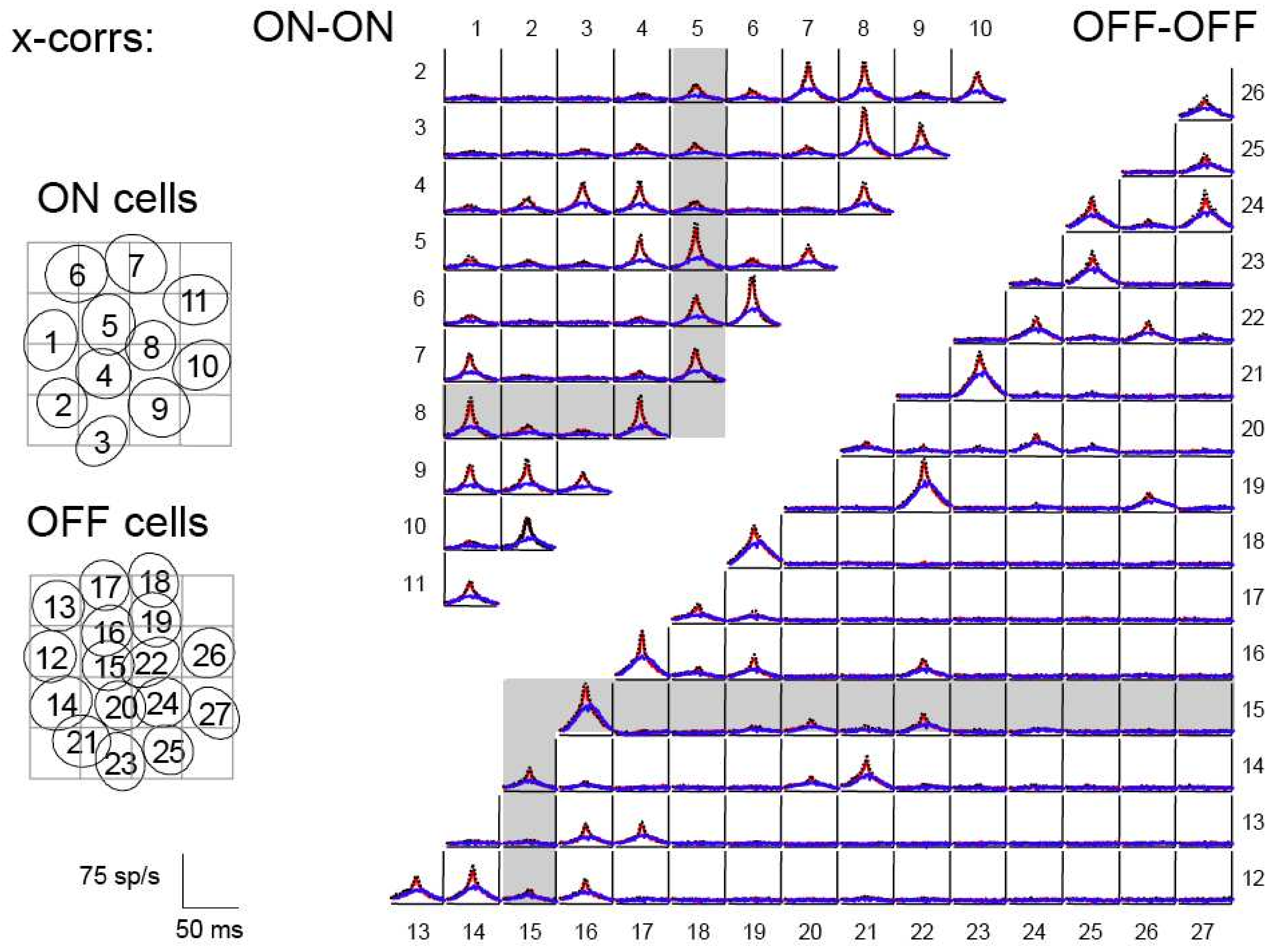
Spike Train Prediction

- improved prediction, but not of mean rate!





Model captures spatiotemporal cross-corrs

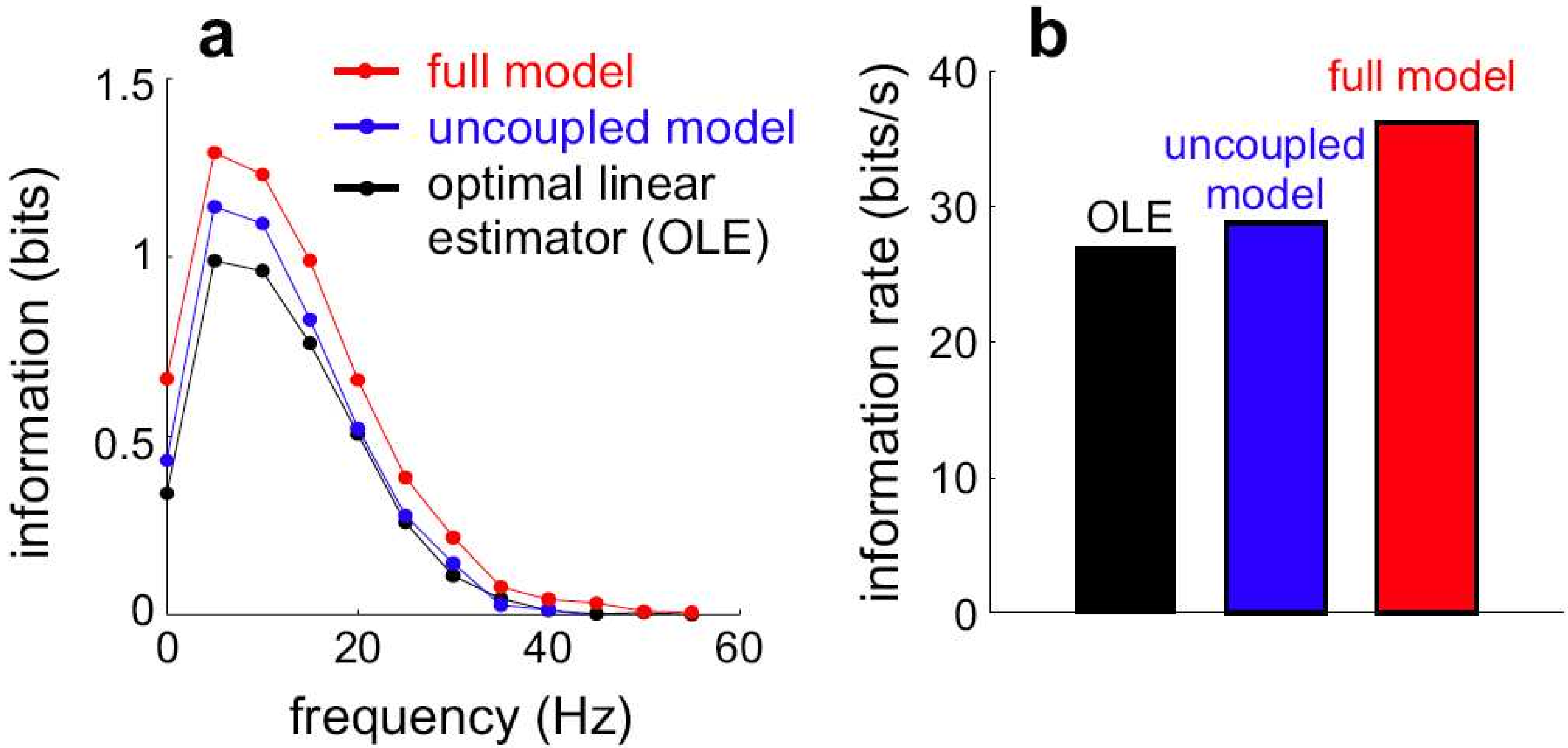


Maximum a posteriori decoding

$$\arg\max_{\vec{x}} \log P(\vec{x}|spikes) = \arg\max_{\vec{x}} \log P(spikes|\vec{x}) + \log P(\vec{x})$$

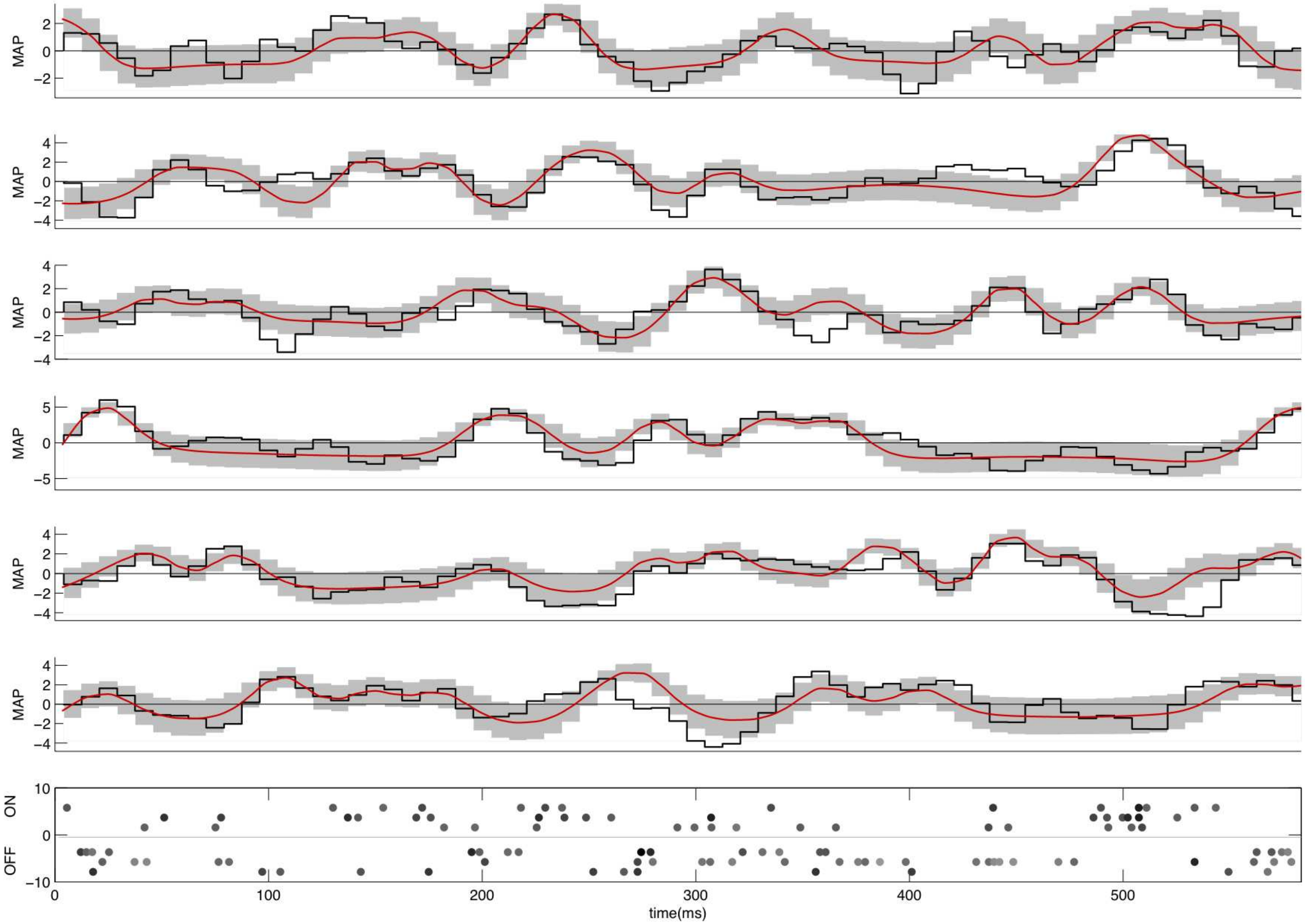
— $\log P(spikes|\vec{x})$ is concave in $\vec{x}$: concave optimization again.

— Decoding can be done in linear time via standard Newton-Raphson methods, since Hessian of $\log P(\vec{x}|spikes)$ w.r.t. $\vec{x}$ is banded (Pillow et al., 2009).



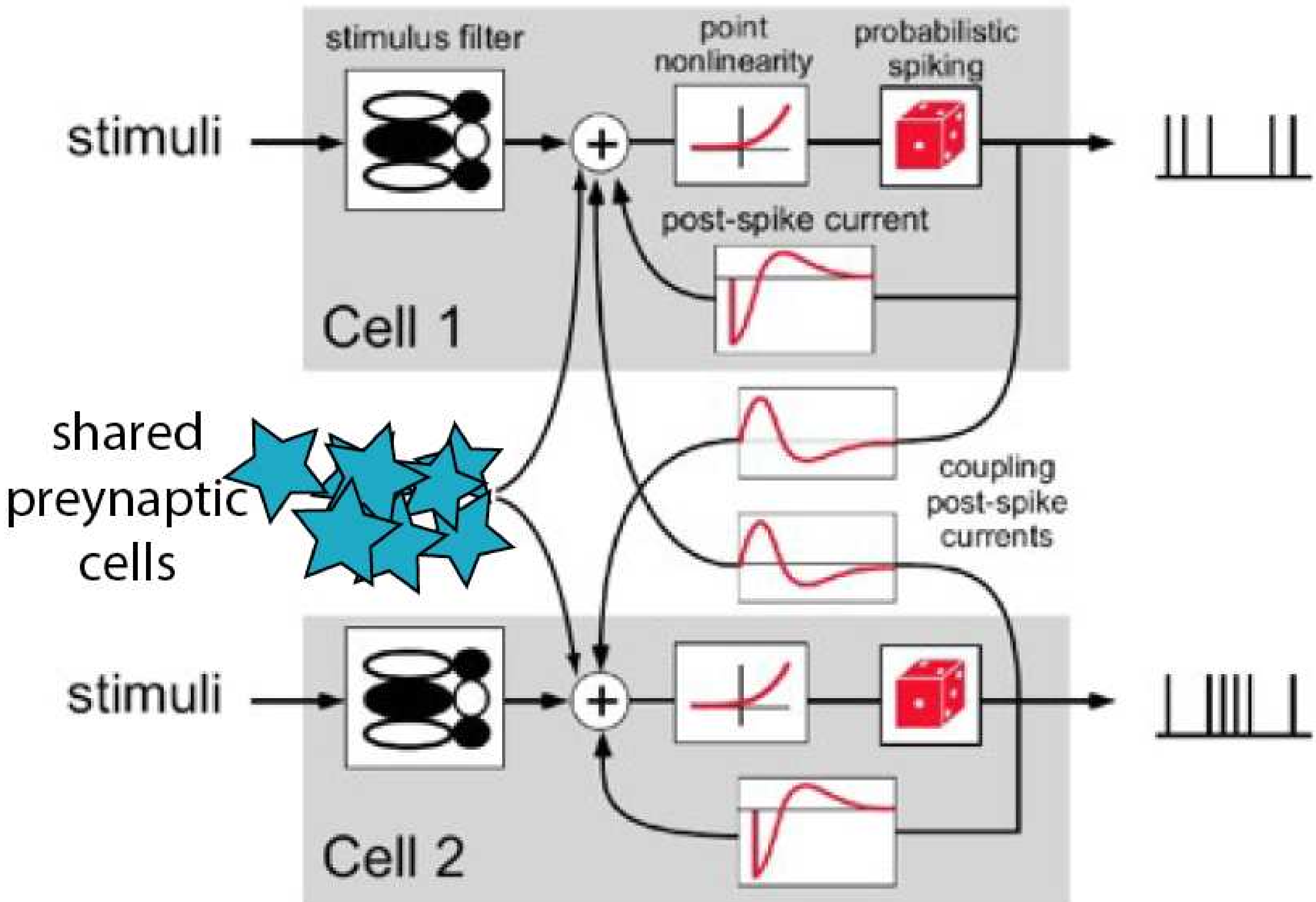
— Including network terms improves decoding accuracy.

Key question: how important is timing?



(Pillow et al., 2009; Paninski et al., 2007; Ahmadian et al., 2009)

Extension: latent “common input” effects



State-space setting; fast estimation methods (Kulkarni and Paninski, 2007; Khuc-Trong and Rieke, 2008; Wu et al., 2009; Vidne et al., 2009)

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 $I(n_t; \theta | \vec{x}_t)$ with respect to $\vec{x}_t$.

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

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

GLM setting makes this surprisingly tractable (Lewi et al., 2009).

Fast stimulus optimization

$$\lambda_i \sim Poiss(\lambda_i)$$

$$\lambda_i | \vec{x}_i, \vec{\theta} = f(\vec{k} \cdot \vec{x}_i + \sum_j a_j r_{i-j})$$

$$\log p(r_i | \vec{x}_i, \vec{\theta}) = -f(\vec{k} \cdot \vec{x}_i + \sum_j a_j r_{i-j}) + r_i \log f(\vec{k} \cdot \vec{x}_i + \sum_j a_j r_{i-j})$$

Two key points:

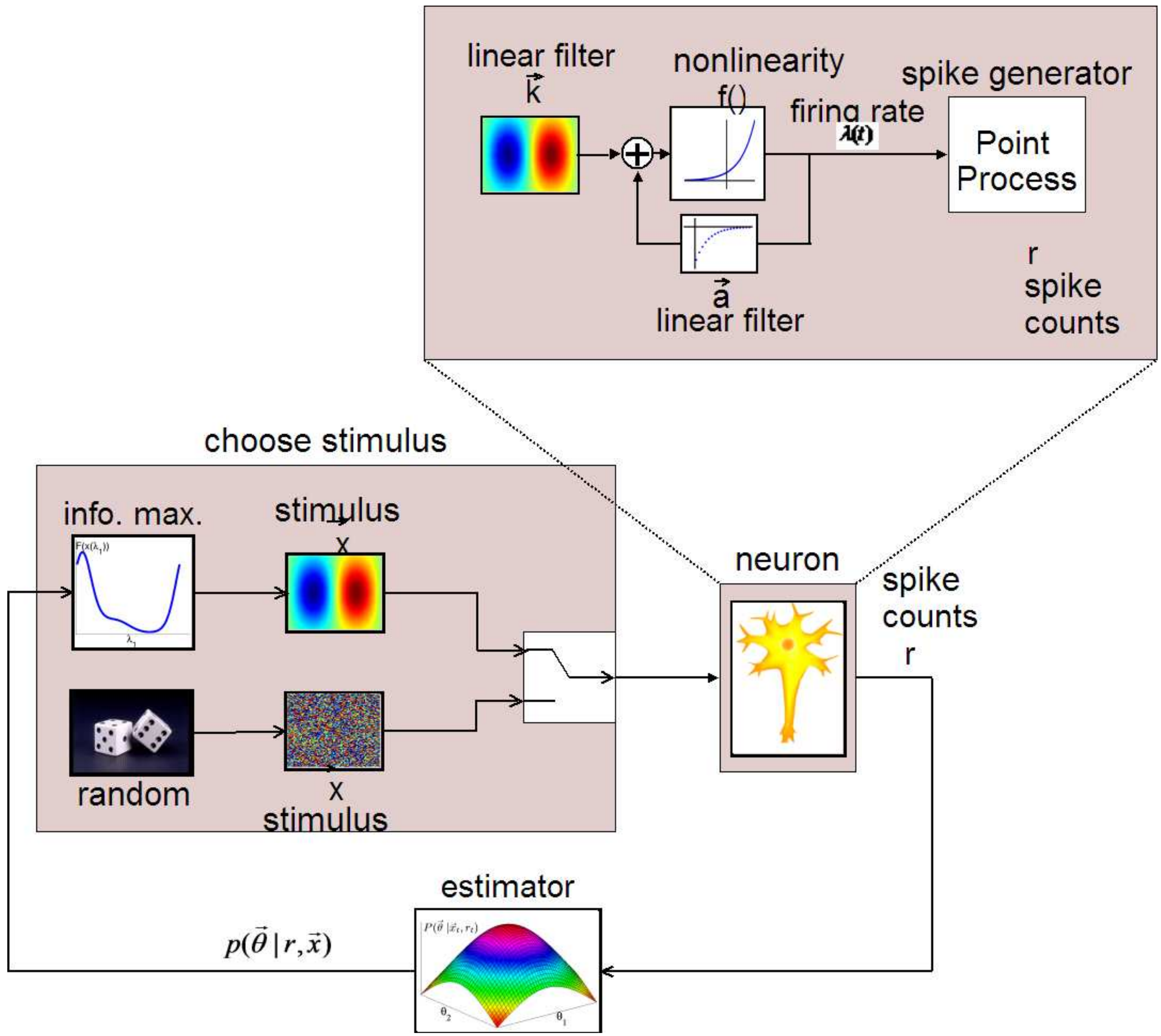
- Likelihood is "rank-1" — only depends on $\vec{\theta}$ along $\vec{z} = (\vec{x}, \vec{r})$.
- f convex and log-concave $\implies$ log-likelihood concave in $\vec{\theta}$

Idea: Laplace approximation:

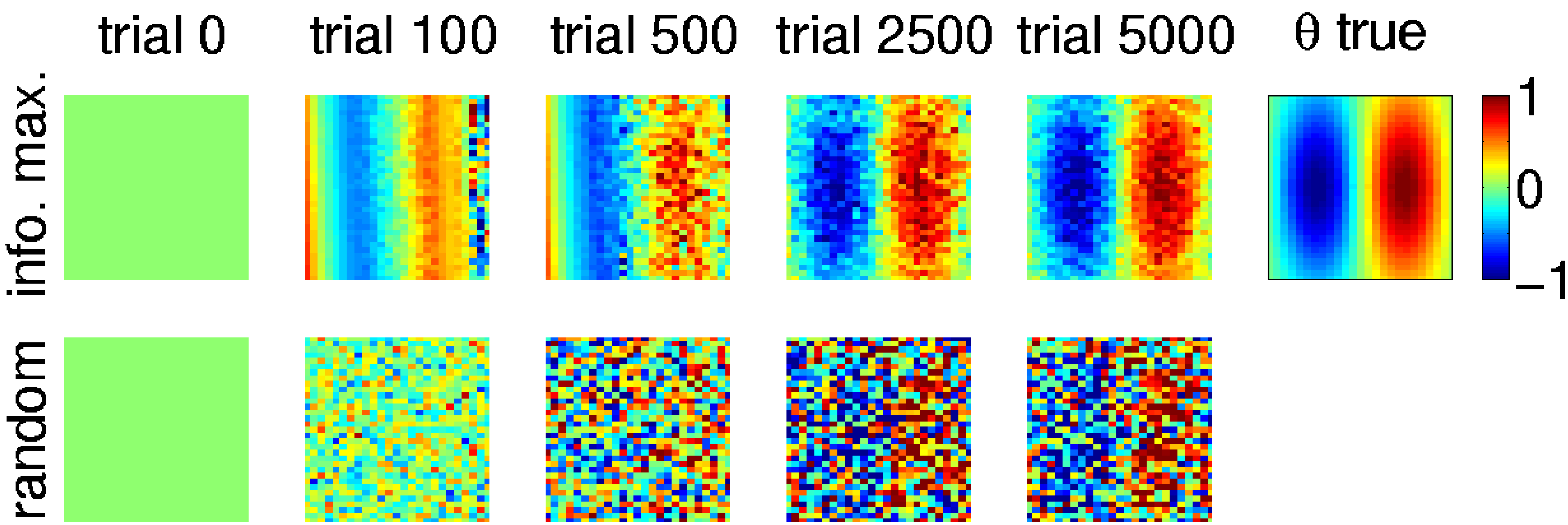
$$p(\vec{\theta} | \{\vec{x}_i, r_i\}_{i \leq N}) \approx \mathcal{N}(\mu_N, C_N)$$

— fast low-rank methods let us update μ_N, C_N and compute the optimal stimulus (maximize $I(n_t; \theta | \vec{x}_t)$) in $O(\dim(\vec{x}_t)^2)$ time.

Infomax vs. randomly-chosen stimuli

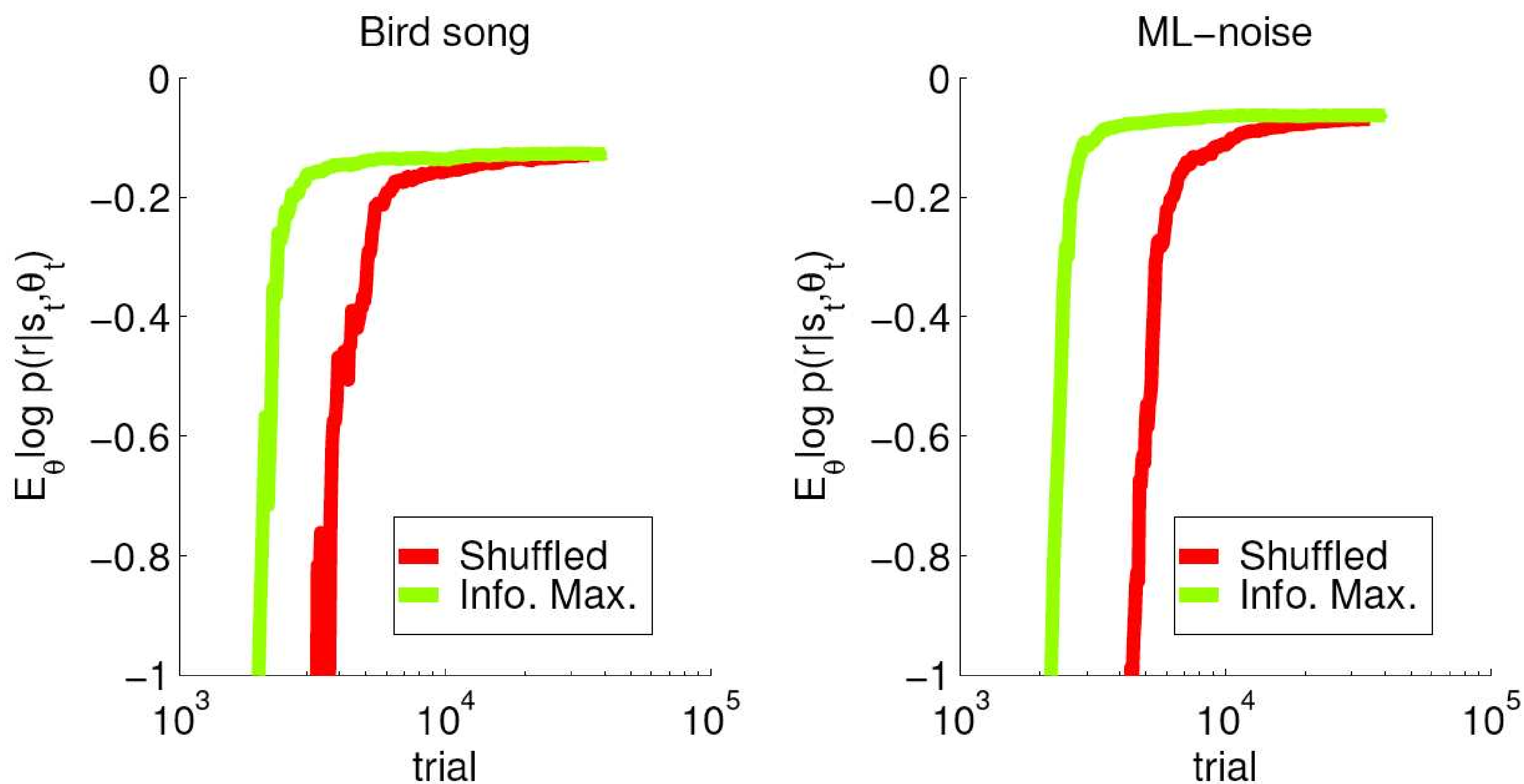


Simulated example



— infomax can be an order of magnitude more efficient.

Application to real data: choosing an optimal stimulus sequence



— stimuli chosen from a fixed pool; greater improvements expected if we can choose arbitrary stimuli on each trial.

Handling nonstationary parameters

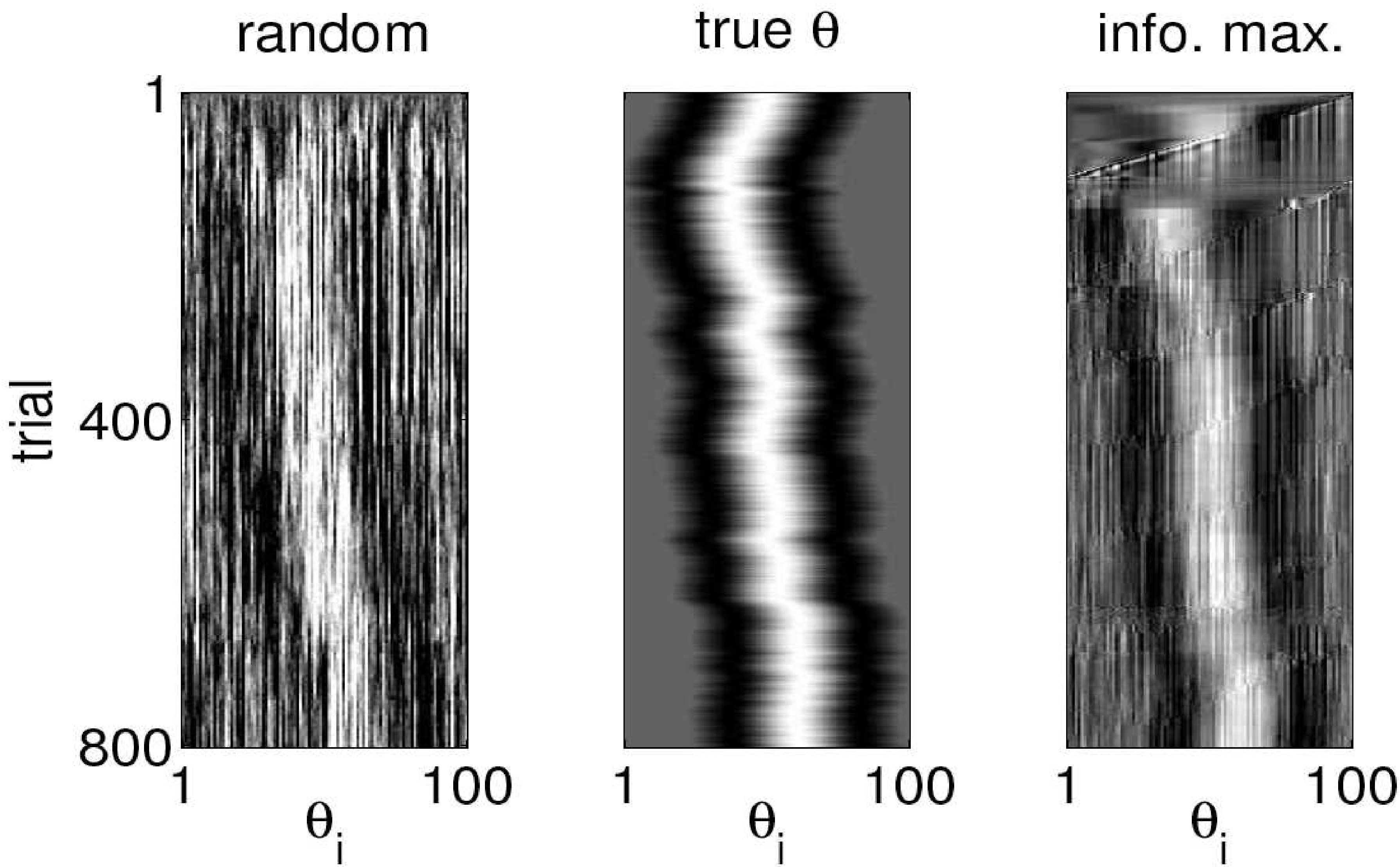
Various sources of nonsystematic nonstationarity:

- Plasticity/adaptation
- Changes in arousal / attentive state
- Changes in health / excitability of preparation

Solution: represent parameter θ in a state-space model (Czanner et al., 2008; Lewi et al., 2009):

$$\vec{\theta}_{N+1} = \vec{\theta}_N + \epsilon; \quad \epsilon \sim \mathcal{N}(0, Q)$$

Simulation: nonstationary parameters



Conclusions

- GLM and state-space approach provides 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
- Many opportunities for applications of statistical ideas in neuroscience

References

Ahmadian, Y., Pillow, J., and Paninski, L. (2009). Efficient Markov Chain Monte Carlo methods for decoding population spike trains. *Under review, Neural Computation.*

Brillinger, D. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cyberkinetics*, 59:189–200.

Czanner, G., Eden, U., Wirth, S., Yanike, M., Suzuki, W., and Brown, E. (2008). Analysis of between-trial and within-trial neural spiking dynamics. *Journal of Neurophysiology*, 99:2672–2693.

Khuc-Trong, P. and Rieke, F. (2008). Origin of correlated activity between parasol retinal ganglion cells. *Nature Neuroscience*, 11:1343–1351.

Kulkarni, J. and Paninski, L. (2007). Common-input models for multiple neural spike-train data. *Network: Computation in Neural Systems*, 18:375–407.

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

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

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.*

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.

Truccolo, W., Eden, U., Fellows, M., Donoghue, J., and Brown, E. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *Journal of Neurophysiology*, 93:1074–1089.

Vidne, M., Kulkarni, J., Ahmadian, Y., Pillow, J., Shlens, J., Chichilnisky, E., Simoncelli, E., and Paninski, L. (2009). Inferring functional connectivity in an ensemble of retinal ganglion cells sharing a common input. *COSYNE.*

Wu, W., Kulkarni, J., Hatsopoulos, N., and Paninski, L. (2009). Neural decoding of goal-directed movements using a linear statespace model with hidden states. *IEEE Trans. Biomed. Eng.*, In press.