Coding and computation by neural ensembles in the retina

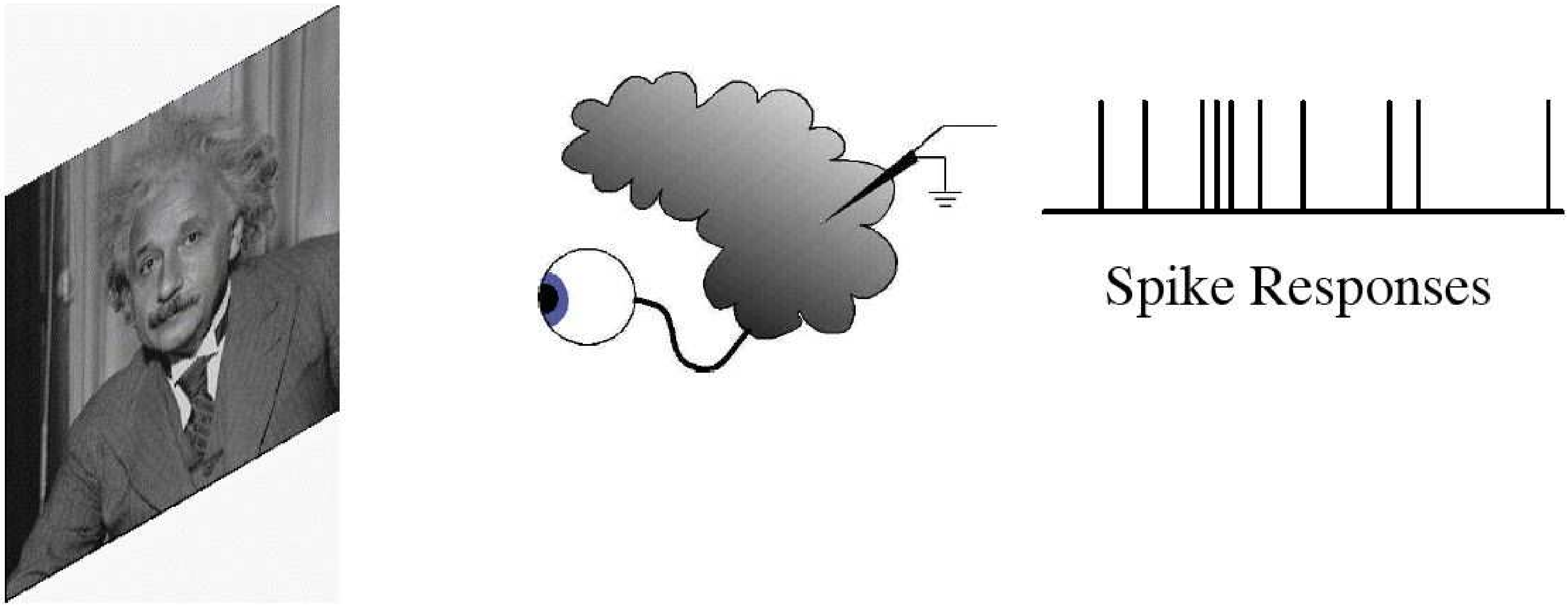
Liam Paninski

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

May 29, 2008

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

The neural code



Input-output relationship between

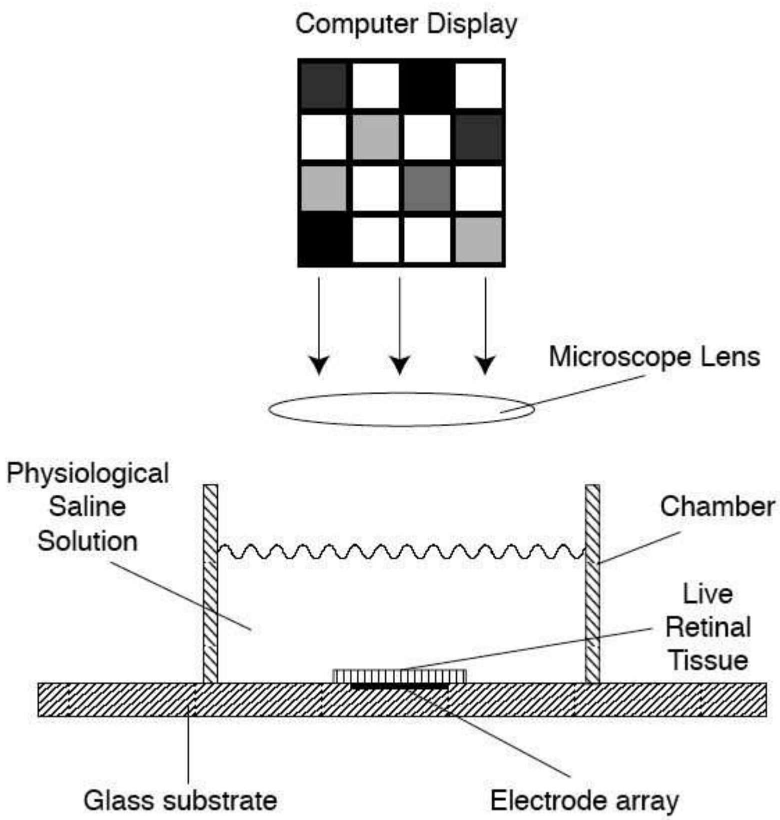
- External observables x (sensory stimuli, motor responses...)
- Neural variables y (spike trains, population activity...)

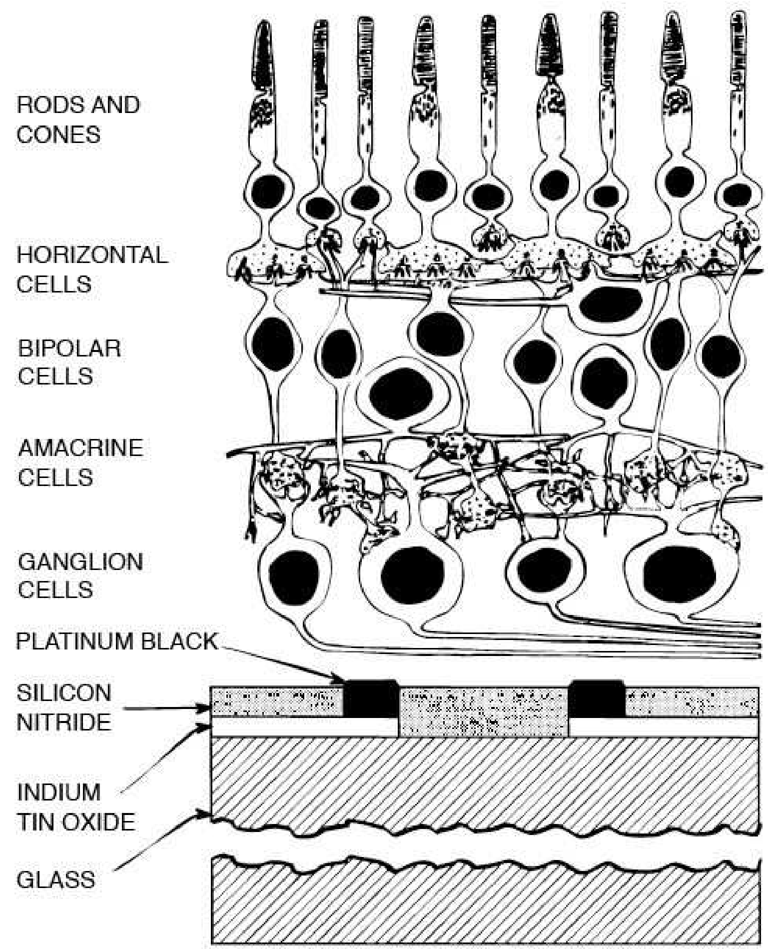
Encoding problem: $p(y|x)$; decoding problem: $p(x|y)$

Retinal ganglion neuronal data

Preparation: dissociated macaque retina

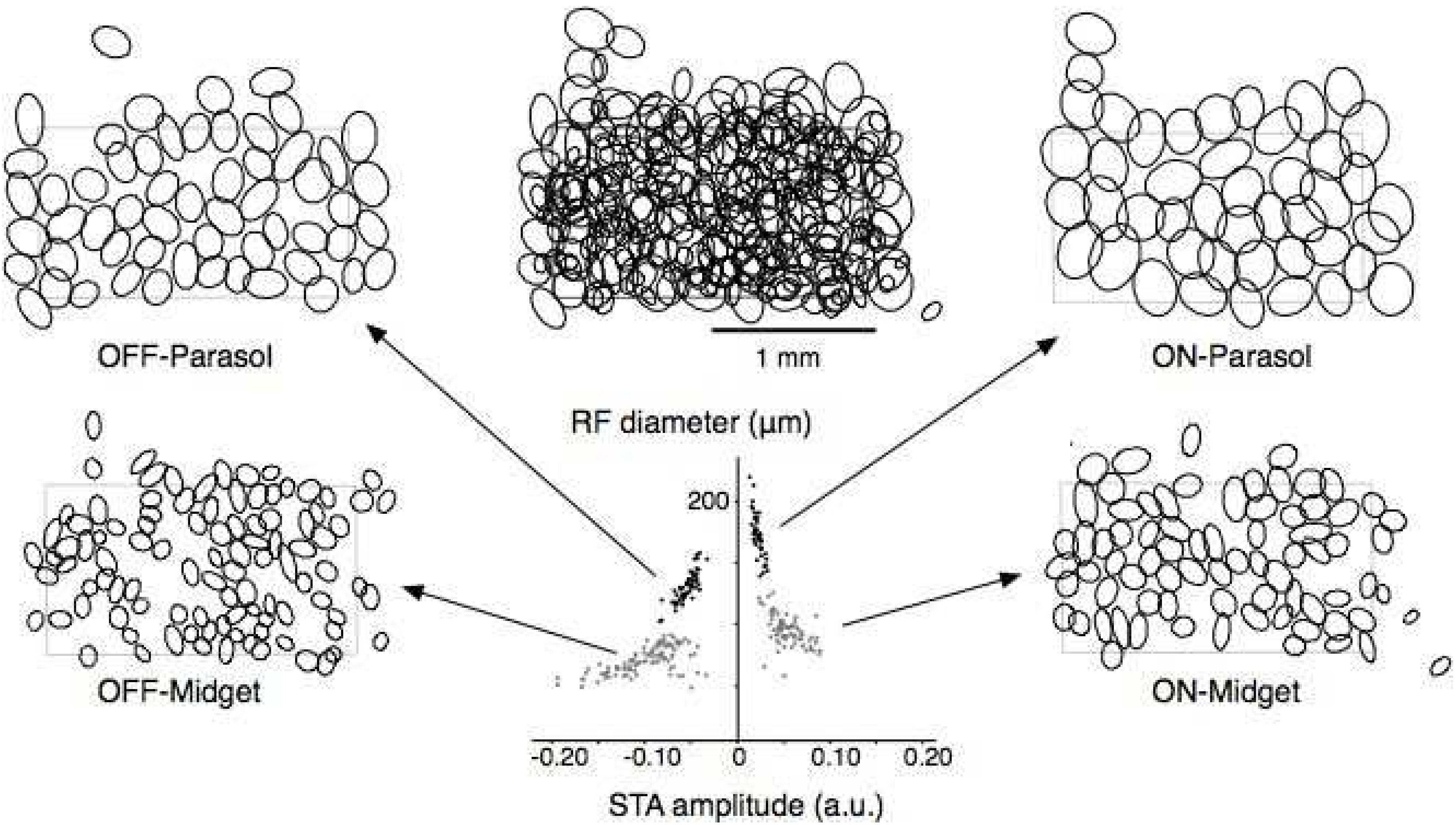
— extracellularly-recorded responses of populations of RGCs



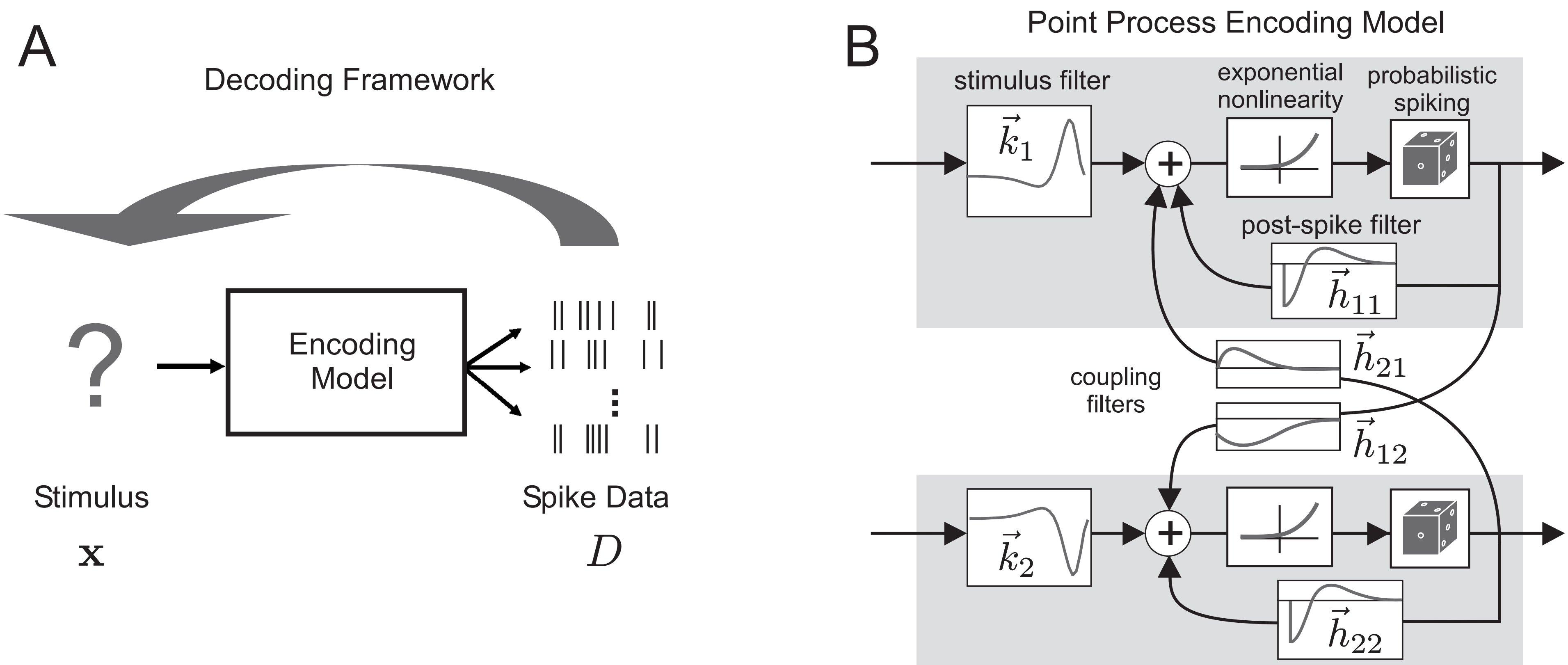


Stimulus: random spatiotemporal visual stimuli (Pillow et al., 2008)

Receptive fields tile visual space

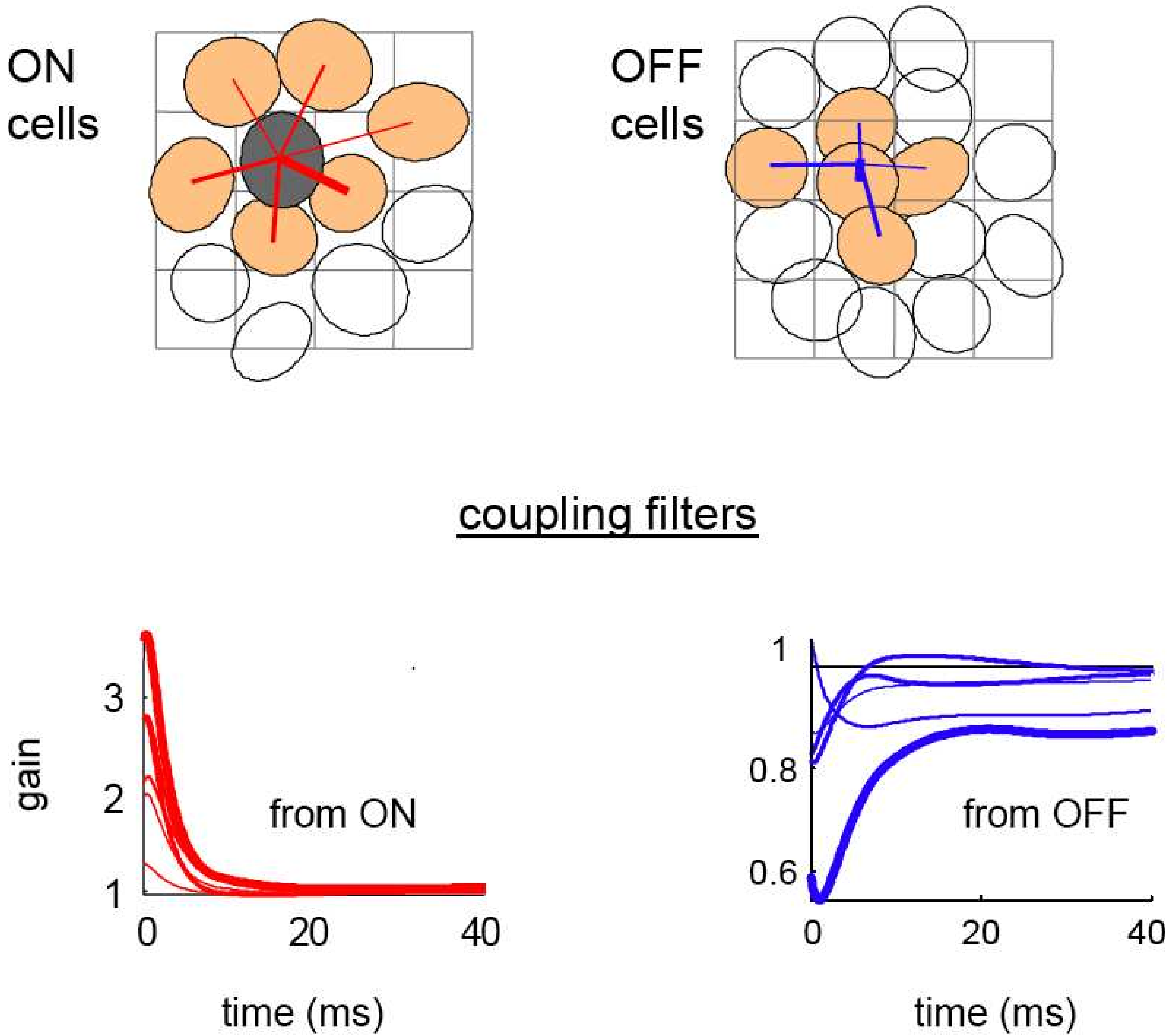


Multineuronal point-process model

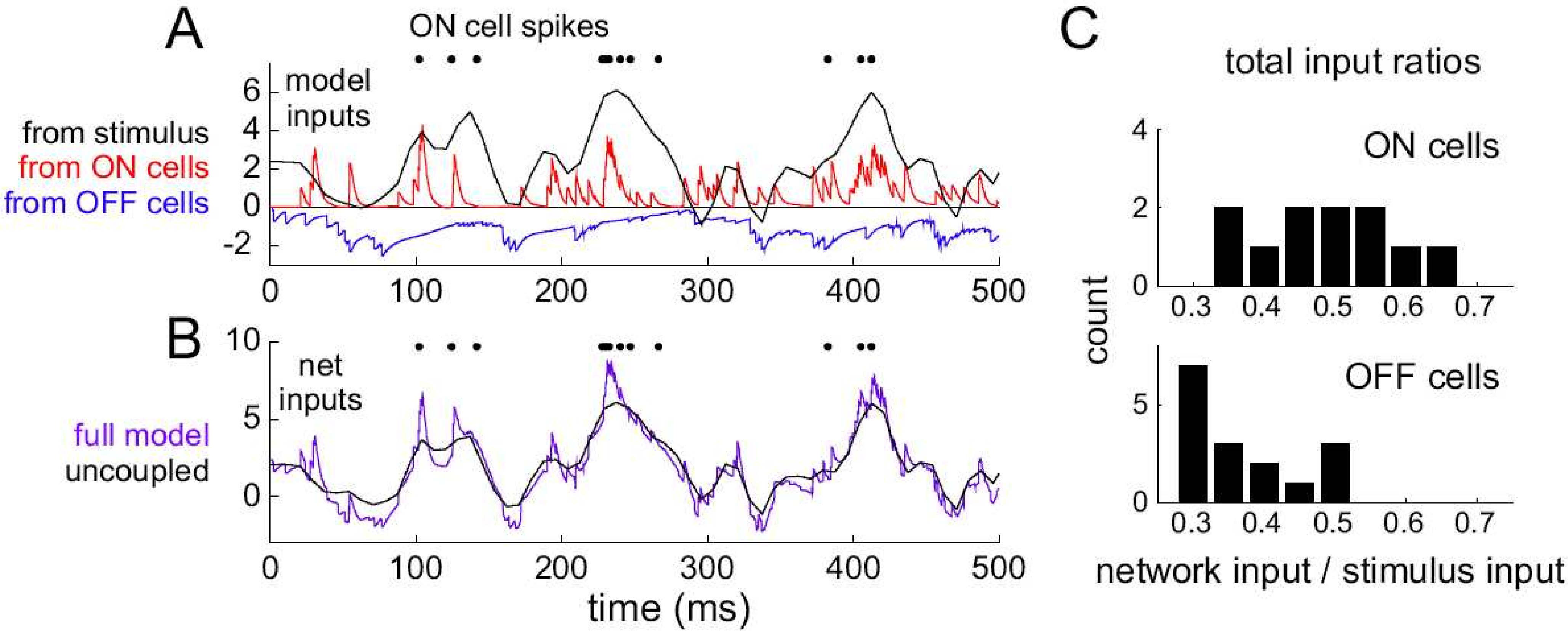


$$\lambda_i(t) = f\left(b + \vec{k}_i \cdot \vec{x}(t) + \sum_{i',j} h_{i',j} n_{i'}(t-j)\right),$$

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


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

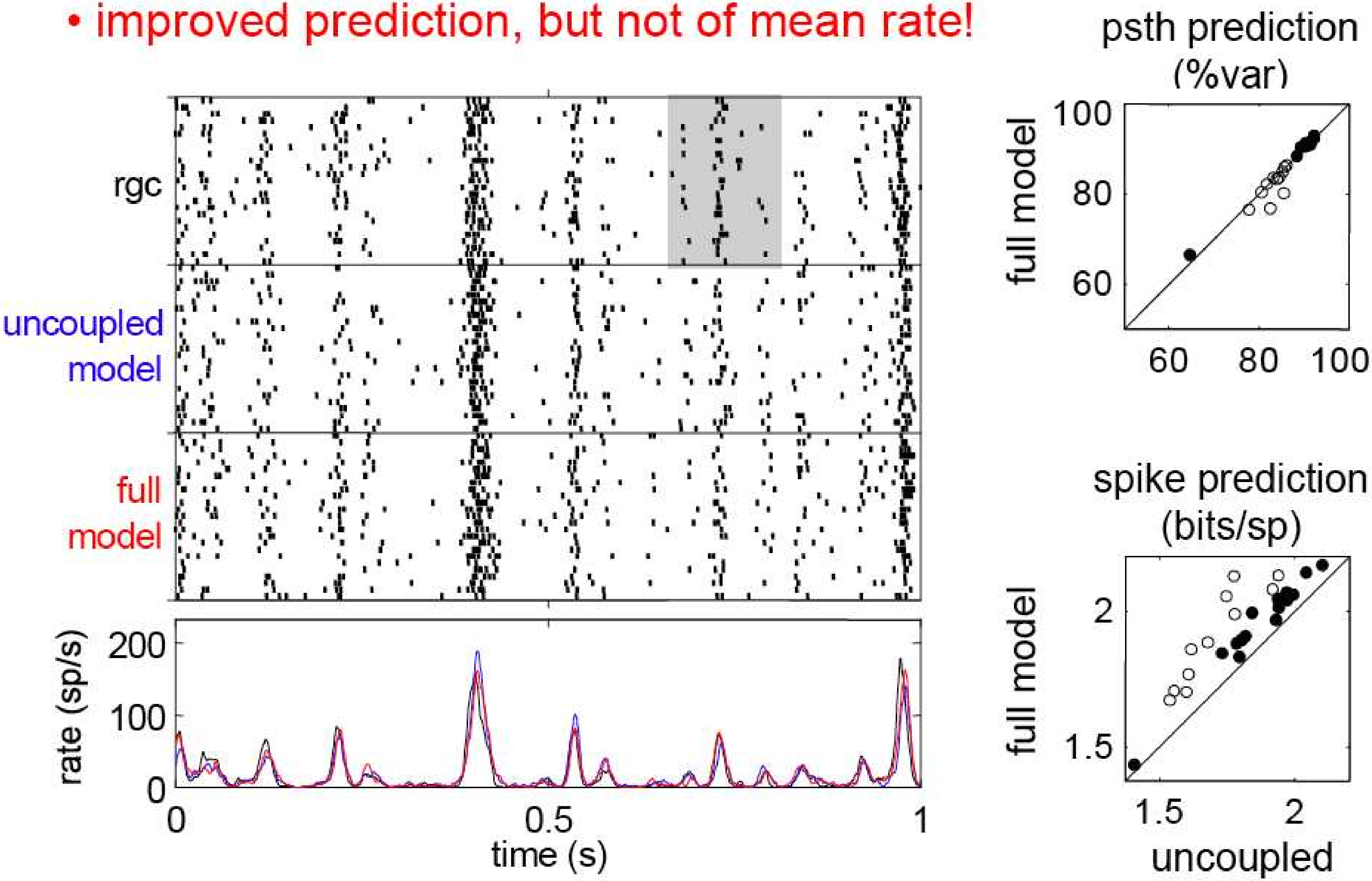
Network vs. stimulus drive



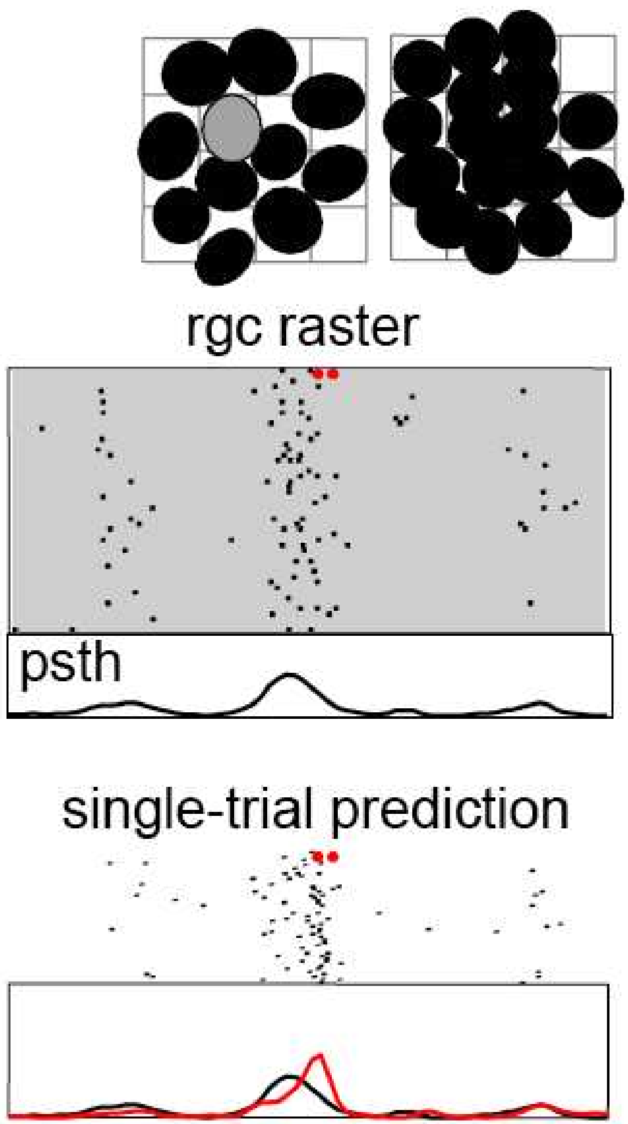
— Network effects are ≈ 50% as strong as stimulus effects

Spike Train Prediction

- improved prediction, but not of mean rate!

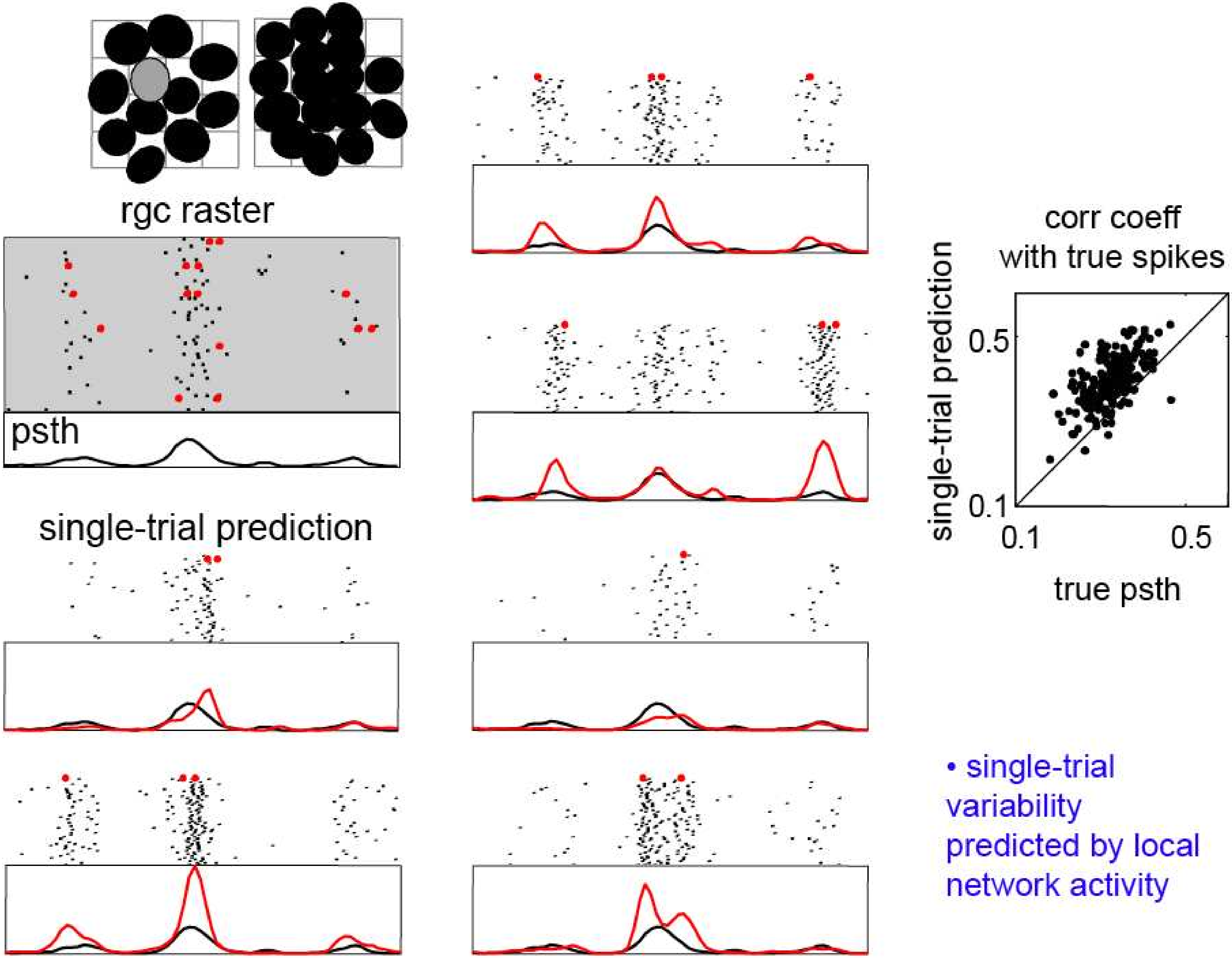


Network predictability analysis

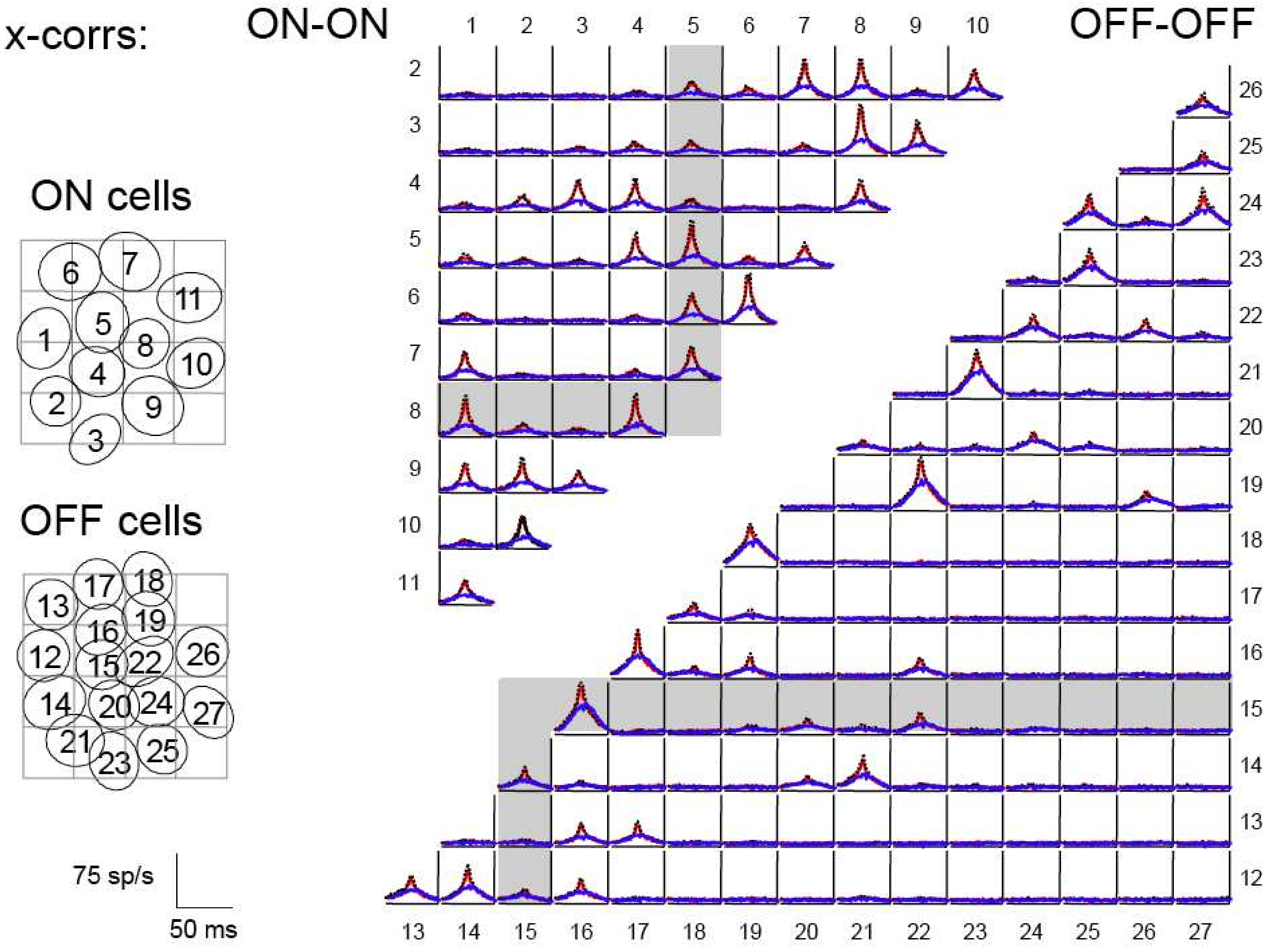


• fix all other neurons for a single trial

• draw single-trial predictions of this cell's spike train


rgc raster
psth
single-trial prediction
corr coeff
with true spikes
single-trial prediction
0.5
0.1
0.1
0.5
true psth
• single-trial
variability
predicted by local
network activity

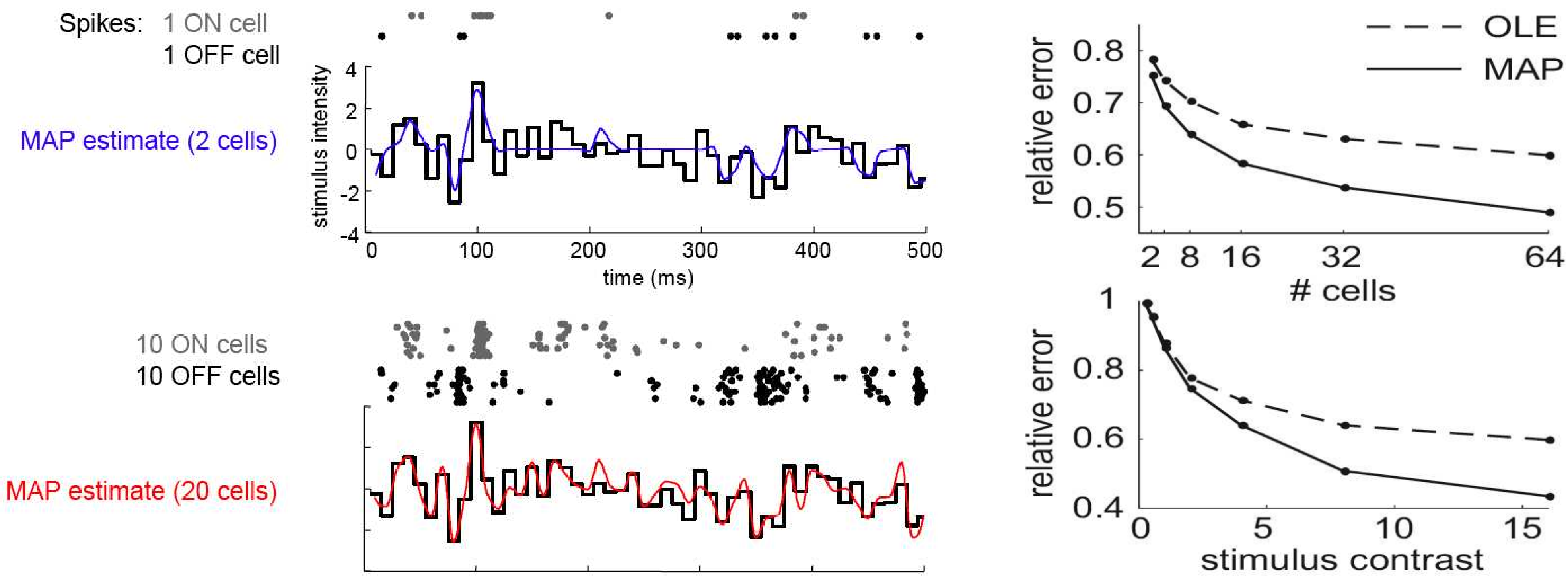
Model captures spatiotemporal cross-corrs



Maximum a posteriori decoding

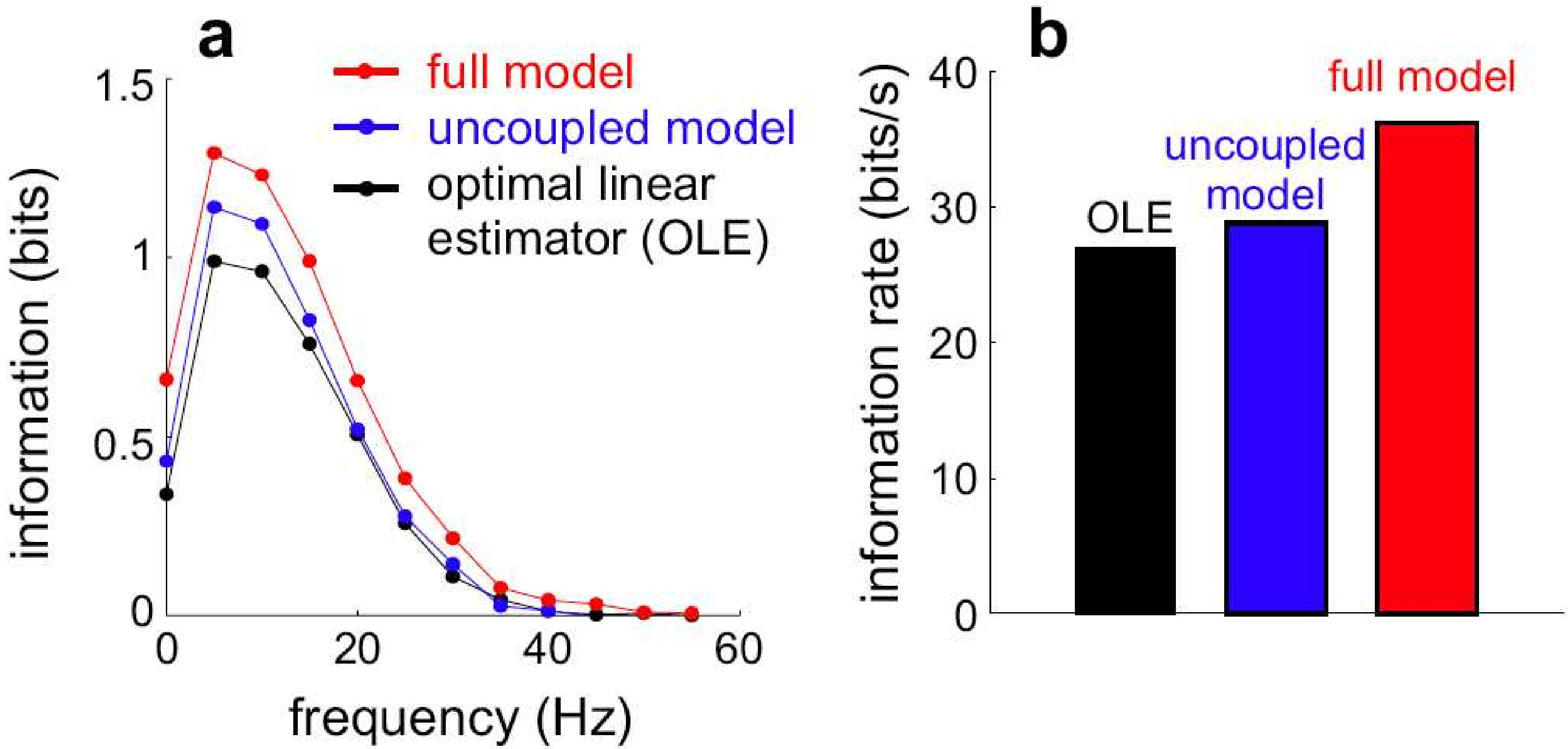
$$\arg\max_{\vec{x}} \log P(\vec{x}|spikes) = \arg\max_{\vec{x}} \log P(spikes|\vec{x}) + \log P(\vec{x})$$

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



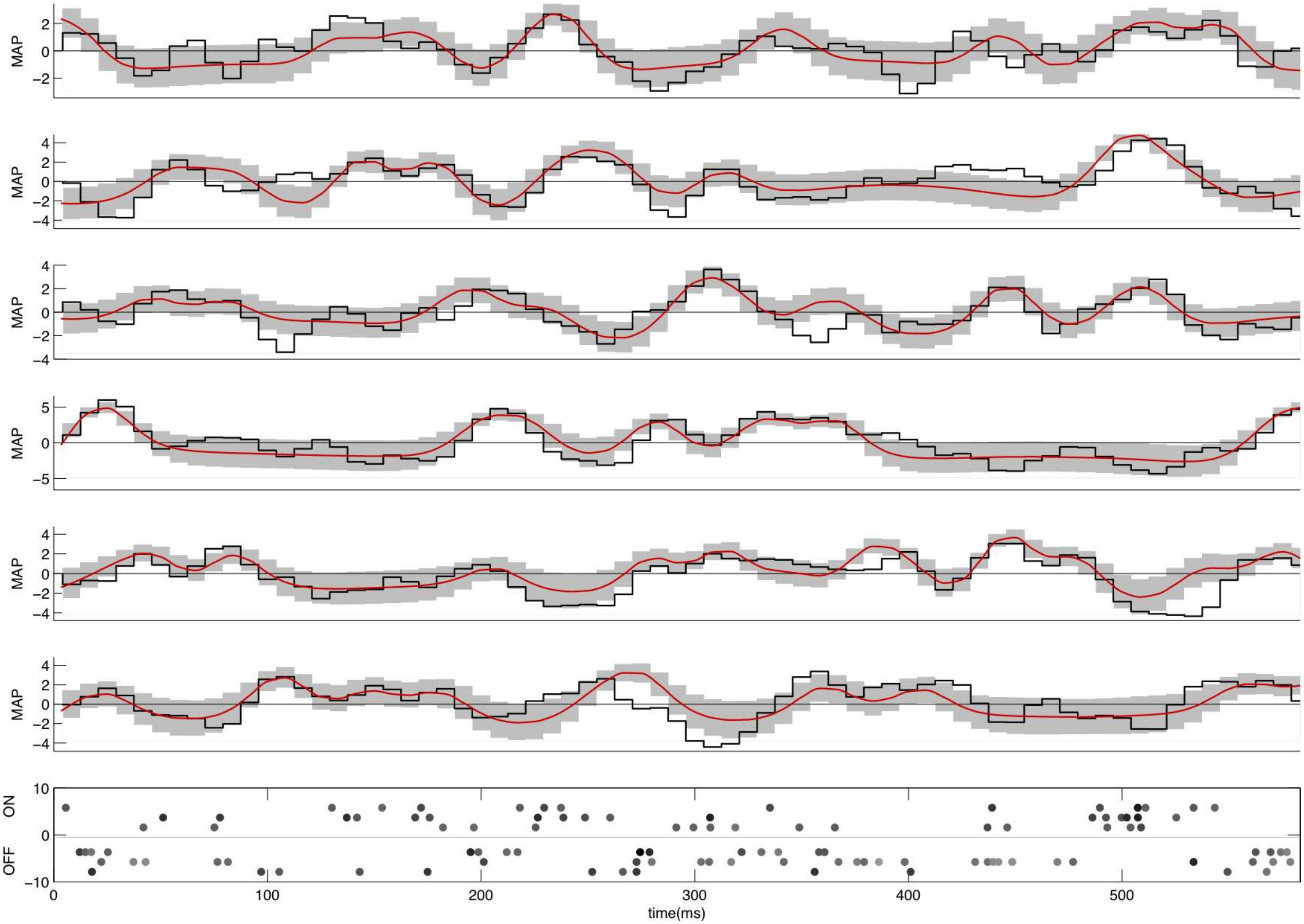
— Decoding can be done in linear time via standard Newton-Raphson methods, since Hessian of $\log P(\vec{x}|spikes)$ w.r.t. $\vec{x}$ is banded (Pillow and Paninski, 2007).

Does including correlations improve decoding?



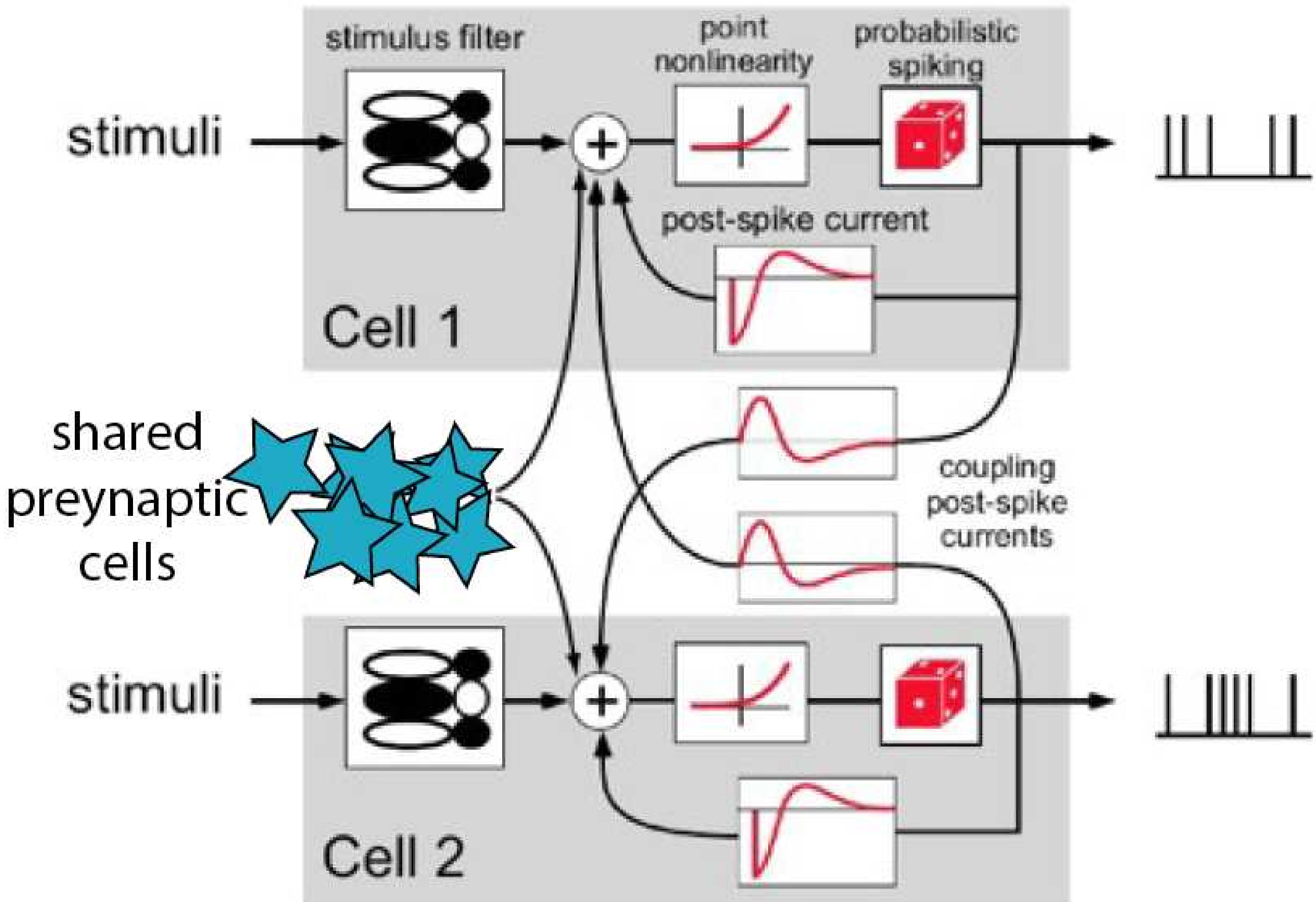
— Including correlations improves decoding accuracy.

How important is timing?



(Ahmadian et al., 2008)

Extension: common input effects



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

Direct state-space optimization methods

$$
\begin{aligned}
\lambda_i(t) &= f\left[b + \vec{k}_i \cdot \vec{x}(t) + \sum_{i',j} h_{i',j} n_{i'}(t-j) + q_i(t)\right] \\
&= f\left[X_t \theta + q_i(t)\right] \\
\vec{q}_{t+dt} &= \vec{q}_t + A\vec{q}_t dt + \sigma\sqrt{dt}\vec{\epsilon}_t
\end{aligned}
$$

— Parameter θ is high-d; standard point-process filter EM is very slow. Instead, optimize Laplace-approximated marginal likelihood directly:

$$
\begin{aligned}
\log p(spikes|\theta) &= \log \int p(Q|\theta) p(spikes|\theta, Q) dQ \\
&\approx \log p(\hat{Q}_\theta|\theta) + \log p(spikes|\hat{Q}_\theta) - \frac{1}{2}\log|J_{\hat{Q}_\theta}| \\
\hat{Q}_\theta &= \arg\max_Q \{\log p(Q|\theta) + \log p(spikes|Q)\}
\end{aligned}
$$

— all terms can be computed in linear time via block-tridiagonal matrix methods (Koyama et al., 2008). Number of applications (Vogelstein et al., 2008).

Optimal velocity decoding

How to decode behaviorally-relevant signals, e.g. image velocity?

If image I is known, use Bayesian estimate (Weiss et al., 2002):

$$p(v|spikes, I) \propto p(v)p(spikes|v, I)$$

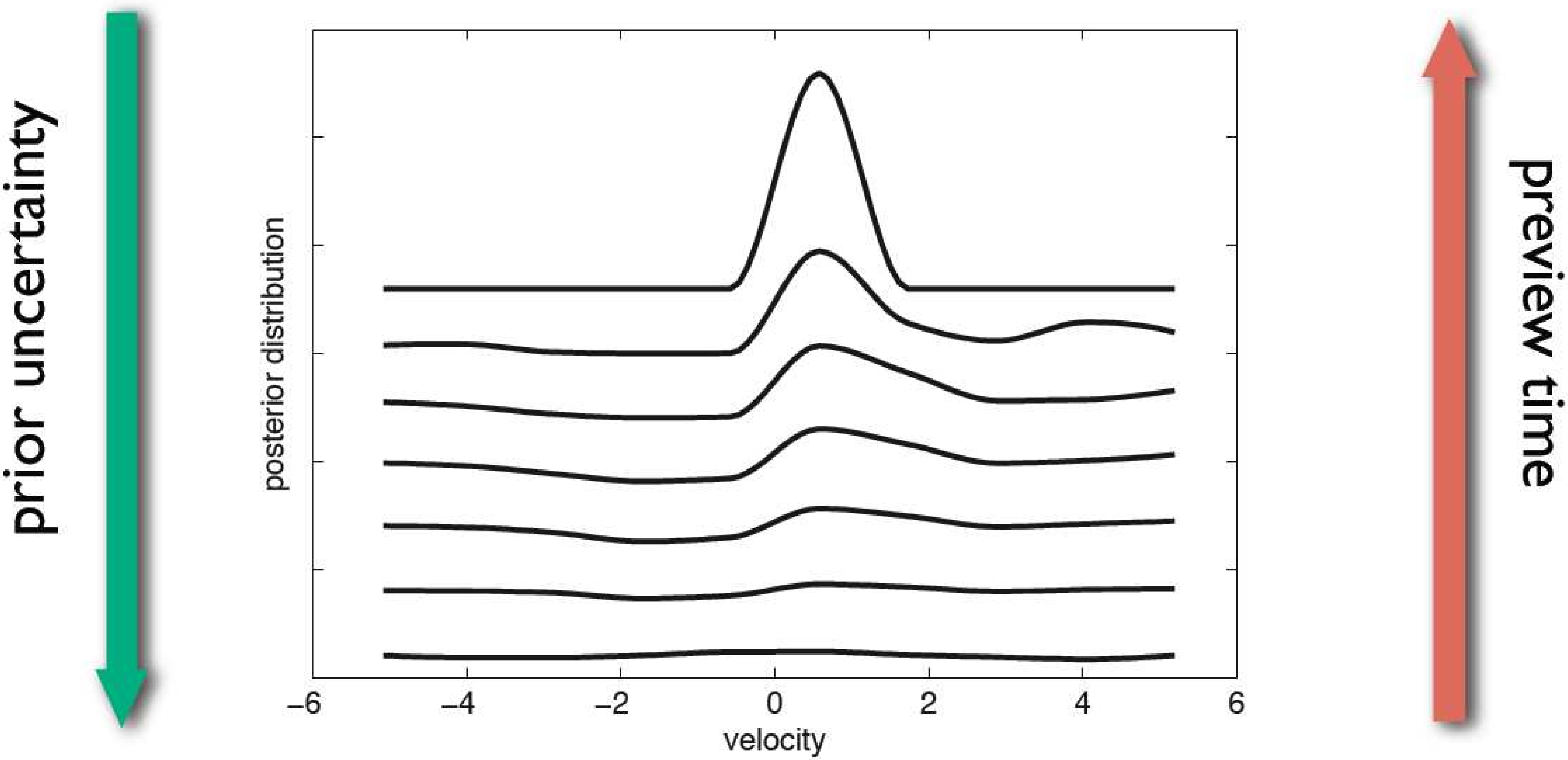
If image is unknown, we have to integrate out:

$$p(v|spikes) \propto p(v)p(spikes|v) = p(v) \int p(I)p(spikes|v, I)dI;$$

$p(I)$ denotes *a priori* image distribution.

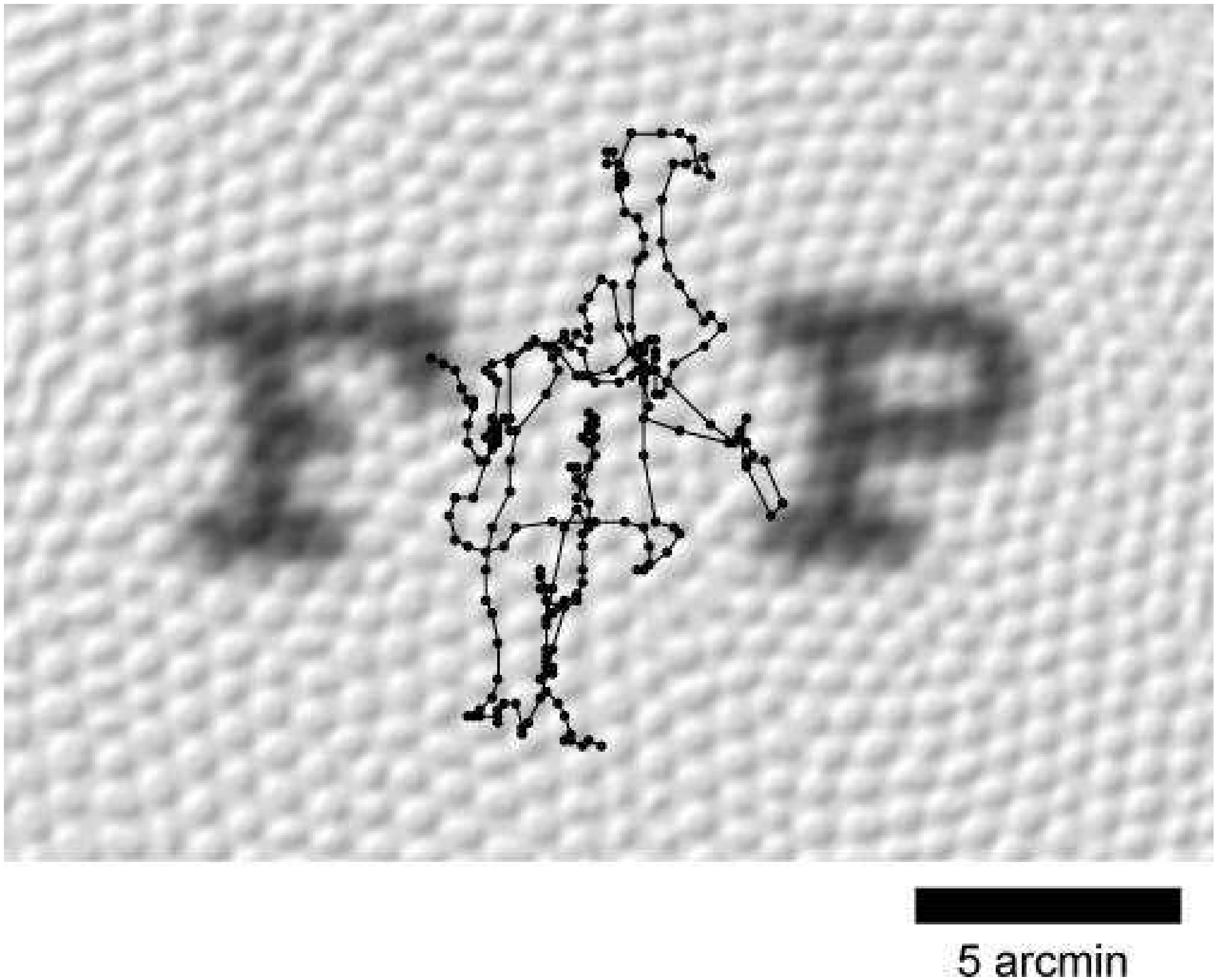
— connections to standard energy models
(Frechette et al., 2005; Lalor et al., 2008)

Optimal velocity decoding



— estimation improves with knowledge of image

Image stabilization is a significant problem



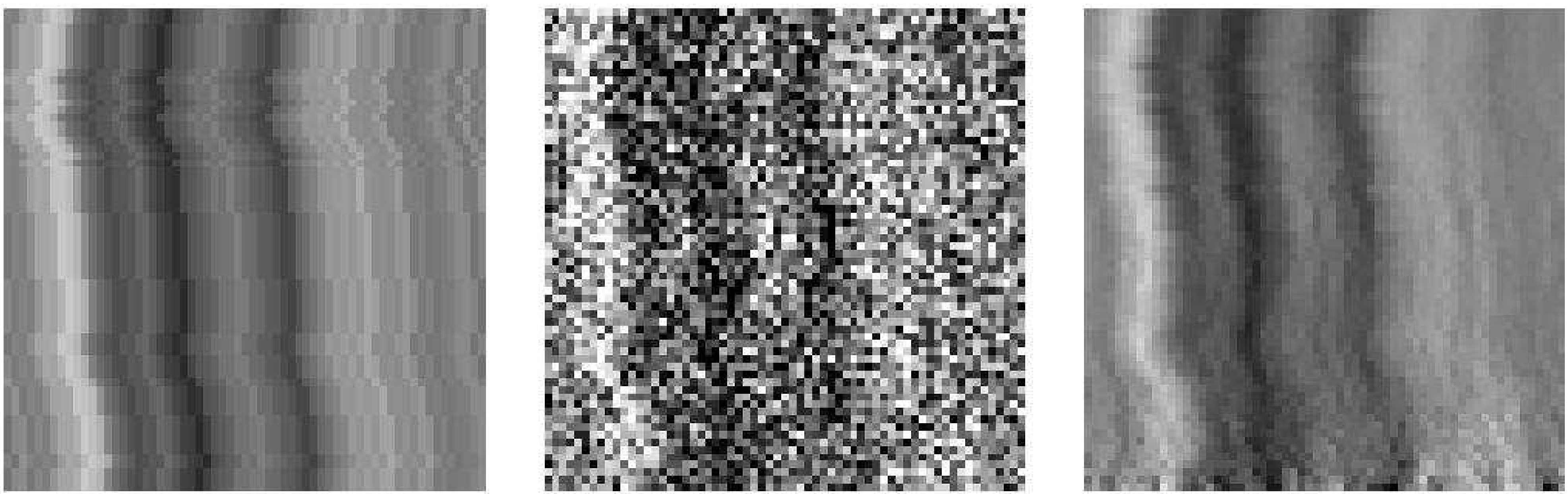
From (Pitkow et al., 2007): neighboring letters on the 20/20 line of the Snellen eye chart. Trace shows 500 ms of eye movement.

Bayesian methods for image stabilization

Similar marginalization idea as in velocity estimation:

$$p(I|spikes) \propto p(I)p(spikes|I) = p(I) \int p(spikes|e, I)p(e)de;$$

e denotes eye jitter path; integration by state-space methods.



true image w/ translations; observed noisy retinal responses; estimated image.

Collaborators

Theory and numerical methods

- Y. Ahmadian, S. Escola, G. Fudenberg, Q. Huys, J. Kulkarni, M. Nikitchenko, X. Pitkow, K. Rahnama, G. Szirtes, T. Toyoizumi, Columbia
- E. Doi, E. Simoncelli, NYU
- E. Lalor, NKI
- A. Haith, C. Williams, Edinburgh
- M. Ahrens, J. Pillow, M. Sahani, Gatsby
- S. Koyama, R. Kass, CMU
- J. Lewi, Georgia Tech
- J. Vogelstein, Johns Hopkins
- W. Wu, FSU

Retinal physiology

- E.J. Chichilnisky, J. Shlens, V. Uzzell, Salk

References

Ahmadian, Y., Pillow, J., and Paninski, L. (2008). 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.

Frechette, E., Sher, A., Grivich, M., Petrusca, D., Litke, A., and Chichilnisky, E. (2005). Fidelity of the ensemble code for visual motion in the primate retina. *J Neurophysiol*, 94(1):119–135.

Khuc-Trong, P. and Rieke, F. (2008). Origin of correlated activity between parasol retinal ganglion cells. *Submitted.*

Koyama, S., Kass, R., and Paninski, L. (2008). Efficient computation of the most likely path in integrate-andfire and more general state-space models. *COSYNE.*

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

Lalor, E., Ahmadian, Y., and Paninski, L. (2008). Optimal decoding of stimulus velocity using a probabilistic model of ganglion cell populations in primate retina. *Journal of Vision*, Under review.

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

Pillow, J. and Paninski, L. (2007). Model-based decoding, information estimation, and change-point detection in multi-neuron spike trains. *Under review, Neural Computation.*

Pillow, J., Shlens, J., Paninski, L., Simoncelli, E., and Chichilnisky, E. (2008). Visual information coding in multi-neuronal spike trains. *Nature*, In press.

Pitkow, X., Sompolinsky, H., and Meister, M. (2007). A neural computation for visual acuity in the presence of eye movements. *PLOS Biology*, 5:e331.

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.

Vogelstein, J., Babadi, B., and Paninski, L. (2008). Fast inference of spike times from noisy calcium traces via tridiagonal nonnegative deconvolution methods. *In preparation.*

Weiss, Y., Simoncelli, E., and Adelson, E. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5:598–604.

Wu, W., Kulkarni, J., Hatsopoulos, N., and Paninski, L. (2008). Neural decoding of goal-directed movements using a linear statespace model with hidden states. *Computational and Systems Neuroscience Meeting.*