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
$
\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

from ON

from OFF

time (ms)
Network effects are $\approx 50\%$ as strong as stimulus effects
Spike Train Prediction

- improved prediction, but not of mean rate!

Examples of different models and predictions are shown in the diagram.
Network predictability analysis

- fix all other neurons for a single trial

- draw single-trial predictions of this cell’s spike train
rgc raster

psith

single-trial prediction

single-trial prediction

single-trial prediction

single-trial prediction

corr coeff with true spikes

true psith

• single-trial variability predicted by local network activity
Model captures spatiotemporal cross-corrs

x-corrs: ON-ON

ON cells

OFF cells

75 sp/s
50 ms
Maximum a posteriori decoding

\[ \arg \max_{\vec{x}} \log P(\vec{x}|\text{spikes}) = \arg \max_{\vec{x}} \log P(\text{spikes}|\vec{x}) + \log P(\vec{x}) \]

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

— Decoding can be done in linear time via standard Newton-Raphson methods, since Hessian of \( \log P(\vec{x}|\text{spikes}) \) 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)
State-space setting (Kulkarni and Paninski, 2007; Khuc-Trong and Rieke, 2008; Wu et al., 2008)

Extension: common input effects
Direct state-space optimization methods

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

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

\[
\log p(\text{spikes}|\theta) = \log \int p(Q|\theta)p(\text{spikes}|\theta, Q)dQ \\
\approx \log p(\hat{Q}_\theta|\theta) + \log p(\text{spikes}|\hat{Q}_\theta) - \frac{1}{2} \log |J_{\hat{Q}_\theta}| \\
\hat{Q}_\theta = \arg \max_Q \{\log p(Q|\theta) + \log p(\text{spikes}|Q)\}
\]

— 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):

$$p(v|\text{spikes}, I) \propto p(v)p(\text{spikes}|v, I)$$

If image is unknown, we have to integrate out:

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

$p(I)$ 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:

\[ p(I|\text{spikes}) \propto p(I)p(\text{spikes}|I) = p(I) \int p(\text{spikes}|e, I)p(e)de; \]

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

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

Theory and numerical methods

• 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


