Simultaneous activity measurements on intact mammalian retina

Phil Nelson
University of Pennsylvania

For these slides see:
www.physics.upenn.edu/~pcn
Part IV: Parallel recordings from dozens of individual neurons

- Sometimes our model is not obviously connected with what we can actually measure experimentally, and we need to make a connection.

- Sometimes the model that interests us involves the behavior of actors that we can only see indirectly in our data; theory may be needed to separate them out from each other, and from noise.
Sources of energy

Experiments done in the lab of Vijay Balasubramanian (Penn).

Kristy Simmons, Penn Neuroscience

Jason Prentice, Penn Physics (now at Princeton)

(plus Gasper Tkacik.)
(Many thanks to Michael Berry and Olivier Marre, Princeton; Bart Borghuis, Janelia Farms; Michael Freed and others at Penn Retina Lab; Joerg Sander, U Alberta; Ronen Segev, BGU.)

Jan Homann (now Princeton)
A piece of tissue **computes** something important for survival. It’s amazing. When we get over being amazed, we ask:

**What does that tissue compute, and how?**

A time-honored approach is to subject that tissue to a known input, find its output, and determine which input motifs elicited each of the repertoire of outputs.

★ In neural systems, the outputs are **spikes**.
★ Retina is popular, in part because we have total control over the inputs.
Retina is also an approachable, yet still complex, part of the brain. It’s a 2D carpet consisting of “only” three layers of neurons.
Future medicine

(Work of Sheila Nirenberg, Weill Cornell)
1. Experiment
2. Clustering
3. Fitting
4. Performance

Get data ➔ Cluster ➔ Fit ➔ Interpret
Model organism

*Cavia porcellus.*
OK, mammals are harder than amphibians. But not that much harder.
Cf Meister, Pine, and Baylor 1994. Incredibly, one can keep a mammalian retina alive in a dish for over 6 hours while presenting it stimuli and recording its activity.
What's in the dish

Simple events

67 ms of data, viewed as a movie.
[data taken at 10kHz have been smoothed. Biggest spikes about 400µV.]

Some spikes move across the array:

Mostly we are hearing retinal ganglion cells, as desired, because they’re the ones that spike.

The spike-sorting problem is: Given raw data like these, convert to a list of discrete events (which cells fired at what times).

Big Data means Big Headaches. Here’s a tiny fragment of what comes out of a multielectrode array every few milliseconds.

Unfortunately many events are complex, with multiple overlapping spikes in many locations. And of course these may be the most interesting ones!

Our algorithm assessed “which neuron fired when” by using a Bayesian inference approach.

Many authors say bursts are a big problem, but here is a nice fit that we obtained with no special effort.

We even handle overlapping spikes, which some algorithms do not attempt.

*JS Prentice, J Homann, KD Simmons, G Tkacik, V Balasubramanian, PCN, PLoS ONE 6(7): e19884 (2011).*
Typical cluster

Superposing 50 traces chosen from 284 in this cluster shows that they really do all resemble each other.

Occasional events in which this event collides with another don’t affect the “archetype waveform” (template) (next slide).

Although the shape of each instance of the template is quite constant, still its amplitude has significant variation.

*JS Prentice, J Homann, KD Simmons, G Tkacik, V Balasubramanian, PCN, PLoS ONE 6(7): e19884 (2011).*
We scaled each instance of each template to get best agreement with the others, then took the median at each time point to find our best estimate of the consensus waveform (blue). As a check, the pointwise mean waveform looks the same (red).
Suppose we measure some experimental data, and wish to make an inference about some situation that we cannot directly observe. That is, we imagine a variety of worlds with different values of $X$, and ask which is most probable given the observed data.

If we know the probability that those data would have arisen in a world with a particular value of $X$, then the Bayes formula gives us what we actually want:

$$P(X|\text{observed data}) = P(\text{data}|X) \frac{P(X)}{P(\text{data})}$$

We can ignore the denominator, if all we want is to compare two hypotheses (e.g. maximize over $X$).

For our application, we’d like $P(\text{spikes} \mid \text{data})$, where ”data” is an observed waveform and ”spikes” refers to a collection of spike templates $\mu_1, \ldots$ occurring at times $t_1, \ldots$ with amplitudes $A_1, \ldots$ relative to the amplitude of the corresponding template (neuron). Bayes’s formula gives what we want as

$$K \times (\text{likelihood}) \times (\text{prior}) = KP(\text{data} \mid \text{spikes})P(\text{spikes})$$
Here “spikes” refers to a collection of spike templates $\mu_1, \ldots$ occurring at times $t_1, \ldots$ with amplitudes $A_1, \ldots$ relative to the amplitude of the corresponding template.

Previous slide expressed $P(\text{spikes} \mid \text{data})$ as:

$$K \times (\text{likelihood}) \times (\text{prior}) = K P(\text{data} \mid \text{spikes}) P(\text{spikes})$$

To get the prior, $P(\text{spikes})$, assume that for a single spike it has the form

$$P_{\text{cell}}(\mu) P_{\text{time}}(t) P_{\text{ampl}}(A \mid \mu)$$

The three factors are respectively the popularity of this neuron, uniform in time, and a Gaussian reflecting its typical amplitude and amplitude variability. We get these priors from the data subset used in clustering.

To get the likelihood function $P(\text{data} \mid \text{spikes})$, suppose that the data consist of one template, plus noise. We measured the noise and found it to be Gaussian in character, and independent of which spikes fired.

Then the likelihood is that distribution, evaluated at the difference between the actual waveform and the idealized one. [Pouzat et. al. 2002]
Let \( V_\alpha(t) \) be measured voltage, electrode \( \alpha \) and \( F_{\mu\alpha}(t) \) be template waveform of type \( \mu \). Define the deviation \( [\delta V]_{\alpha t} = V_\alpha(t) - A F_{\mu\alpha}(t - t_1) \)

Then the probability that one spike, of type \( \mu \), is present is

\[
P(\text{spikes} \mid \text{data}) = K_\mu \exp \left[ -\frac{(A - \gamma_\mu)^2}{2\sigma_\mu^2} - \frac{1}{2}(\delta V)^t C^{-1}(\delta V) \right]
\]

which is a Gaussian in \( A \). So it’s easy to marginalize over \( A \): just complete the square! [Here \( K_\mu = P_{\text{cell}}(\mu) P_{\text{time}}(t_1) (2\pi\sigma_\mu^2)^{-1/2} \) doesn’t depend on \( A \).]

Next, we sweep over a range of \( t \) to find the best value of likelihood ratio for this spike type. [We only check \( t \) values close to the peak of the event.]

Then we choose the winner among spike types.

If the winner’s likelihood ratio is good enough (bigger than about 1), we say there’s a spike here. That’s an absolute criterion. We know we’re done when this test fails.
Vanilla least-squares fitting is not appropriate for time series, because it assumes that every sample is independent of all others—whereas actually, successive samples are correlated.

Here is the covariance of one channel with nearby channels (after doing an initial spatial filter, which we also obtained from data).

We see that the selected channel is correlated only with itself, and it has a simple covariance matrix that is easy to invert. The inverse covariance thus obtained defines our correlated Gaussian model of the noise.

[Again: The covariance is not a delta function, contrary to what is assumed in naive least-squares fitting.]
Successfully fit overlaps

Closeup of four channels, showing four fit templates found by the algorithm.

Sum of those fits (color) versus actual data (black).
Successfully fit bursts

Even though successive spikes in a burst have different amplitudes, the algorithm fit them.
Each cell has a receptive field...

... and they tile the whole visual field. MEA recording is **high throughput**: We got dozens of cells all at once. Here are cells from just one functional group, “on cells.” Each putative receptive field is a single connected region of image space, and they really do tile the region we studied.

Region of retina responded to by ganglion cell #1, etc.

Once you’ve got the spike trains, you can find receptive fields etc. Here’s a typical spike-triggered average.

Interesting--guinea pig retina has a lot of these highly anisotropic receptive fields. The “surround” doesn’t surround the “center”! Moreover, many of the receptive fields are time-dependent (motion-sensitive).
Adaptive decorrelation, (temporal)

The retina dynamically adjusts its signal processing in response to statistical properties of recently-viewed scenes, as predicted on information-theoretic grounds.

Here a particular OFF ganglion cell maintains a constant amount of temporal correlation in its output, regardless of the amount of correlation in its visual stimulus.

Also at the multi-cell level, after adaptation the degree of correlation between any two ganglion cells is nearly unchanged when we change the correlation strength in the stimulus.

I described how we identify the individual ganglion cell signals from a hash of noise and overlapping real signals:
Medical tests

Bayes formula

Poisson processes

Multi-Electrode Array

Changepoint Analysis

Theory can cut across apparently different kinds of experiment, offering useful methods to one domain from another without having to reinvent everything. Physicists are pretty good at this--especially as a part of a team involving life scientists. You do have to meet them halfway, but it’s worth it.
Wait, there’s more

A physical model -- photon theory -- helped us to extract what was going on.

A physical model -- localized spreading of potential changes in solution -- helped us to extract what was going on.

There is something weirdly -- unreasonably -- effective about approaching biological systems with a physical model. I don’t understand why. I don’t need to understand why.
Thanks

For these slides see:

www.physics.upenn.edu/~pcn