Boulder Theoretical Biophysics summer school:
Introduction to neuroscience and information theory

Lecturer:
Stephanie E. Palmer
Associate Professor
Dep’t of OBA
Dep’t of Physics
University of Chicago
sepalmer@uchicago.edu
stimuli $\rightarrow$ neurons

neurons $\rightarrow$ neurons

neurons $\rightarrow$ ... $\rightarrow$ neurons $\rightarrow$ behavior
Anatomy of a neuron:
Anatomy of a spike:

- **Absolute Refractory Period**
- **Relative Refractory Period**

- **Threshold**
- **Subthreshold**

- **Pump maintains Na⁺ outside, K⁺ inside**
- **Na⁺ in = Na⁺ channels open, K⁺ channels close**
- **K⁺ out = K⁺ channels open, Na⁺ channels close**
- **Pump = 3 Na⁺ pumped out, 2 K⁺ pumped in**
Properties of spiking:

- Resting neuron

- Single stimulus

- Sustained stimulus

How can we account for this behavior quantitatively?
We seek a quantitative description of this behavior:
Some basics:
Some basics:

chalkboard interlude
Squid giant axon:
The 1939 letter to Nature:
The final model:

The final model:


\[
C_m \frac{dV}{dt} = -g_L(V - V_L) - \bar{g}_{Na} m^3 h (V - V_{Na}) - \bar{g}_K n^4 (V - V_K)
\]

\[
\frac{dm}{dt} = \alpha_m(V)(1 - m) - \beta_m(V)m
\]

\[
\frac{dh}{dt} = \alpha_h(V)(1 - h) - \beta_h(V)h
\]

\[
\frac{dn}{dt} = \alpha_n(V)(1 - n) - \beta_n(V)n
\]
The final model:

\[ C_m \frac{dV}{dt} = -g_L(V-V_L) - \overline{g}_{Na} m^3 h(V-V_{Na}) - \overline{g}_K n^4 (V-V_K) \]

\[ \frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m \quad \alpha_m = 0.1(V_m + 35.0)/(1 - e^{-(V_m+35.0)/10.0}) \]
\[ \beta_m = 4.0 e^{-(V_m+60.0)/18.0} \]

\[ \frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h \quad \alpha_h = 0.07 e^{-(V_m+60.0)/20.0} \]
\[ \beta_h = 1.0/(1 + e^{-(V_m+30.0)/10.0}) \]

\[ \frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n \quad \alpha_n = 0.01(V_m + 50.0)/(1 - e^{-(V_m+50.0)/10.0}) \]
\[ \beta_n = 0.125 e^{-(V_m+60.0)/80.0} \]
Box 3A The Voltage Clamp Technique

- Measure $V_m$
- Command voltage
- Voltage clamp amplifier
- Measure current
- Reference electrode
- Saline solution
- Squid axon
- Recording electrode
- Current-passing electrode
Box 3A The Voltage Clamp Technique

chalkboard interlude
How can I \( I_{Na} \) and \( I_K \) be separated?

Make recordings, separate currents:
...and get conductances:

\[ g_{Na} = \frac{I_{Na}}{V - E_{Na}} \]

\[ g_{K} = \frac{I_{K}}{V - E_{K}} \]
...and get conductances:

g_{Na} = \frac{I_{Na}}{(V - E_{Na})}

g_{K} = \frac{I_{K}}{(V - E_{K})}

chalkboard interlude
Potassium channels are tetramers!
Structure of the potassium ion channel (tetramer)

View down the pore

A single subunit
The final model:

The final model:


\[
C_m \frac{dV}{dt} = -g_L(V-V_L) - g_{Na}m^3h(V-V_{Na}) - g_Kn^4(V-V_K)
\]

\[
\frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m
\]

\[
\frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h
\]

\[
\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n
\]
The final model:

\[
C_m \frac{dV}{dt} = -g_L (V - V_L) - \bar{g}_{Na} m^3 h (V - V_{Na}) - \bar{g}_K n^4 (V - V_K)
\]

\[
\frac{dm}{dt} = \alpha_m (V)(1 - m) - \beta_m (V) m
\]

\[
\frac{dh}{dt} = \alpha_h (V)(1 - h) - \beta_h (V) h
\]

\[
\frac{dn}{dt} = \alpha_n (V)(1 - n) - \beta_n (V) n
\]

\[
\alpha_m = 0.1(V_m + 35.0)/(1. - e^{-(V_m+35.0)/10.0})
\]

\[
\beta_m = 4.0 e^{-(V_m+60.0)/18.0}
\]

\[
\alpha_h = 0.07 e^{-(V_m+60.0)/20.0}
\]

\[
\beta_h = 1./(1 + e^{-(V_m+30.0)/10.0})
\]

\[
\alpha_n = 0.01(V_m + 50.0)/(1 - e^{-(V_m+50.0)/10.0})
\]

\[
\beta_n = 0.125 e^{-(V_m+60.0)/80.0}
\]

Important note: equations and parameters were derived from voltage clamp data, action potential simulations were an independent test.

In addition to producing realistic action potentials, the model:

1) exhibits sub-threshold and supra-threshold responses
2) correctly reproduces refractoriness
3) reproduces "anode break" excitation

...works well:
...works well:
...works well:

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>
The Fitzhugh-Nagumo model:

\[
\begin{align*}
\frac{dV}{dt} &= V - V^3 - W - I \\
\frac{dW}{dt} &= 0.08(V + 0.7 - 0.8W)
\end{align*}
\]
The Fitzhugh-Nagumo model is a useful simplification:
The Fitzhugh-Nagumo model is a useful simplification:

\[ \text{return to baseline} \]
The Fitzhugh-Nagumo model is a useful simplification:

- An instantaneous increase in $V$
- A small increase in $V$
- A larger increase in $V$

return to baseline

spike!
The Fitzhugh-Nagumo model

Constant current injection (negative I) will shift V nullcline up

$I = \frac{-0.7}{V - W = V - V^3 - I - 1.5 - 0.5 - 0}$

This fixed point is now unstable!

Repetitive action potentials with constant current = conversion from stable fixed point to stable limit cycle

Even explains sustained firing!
Extracellular recording:
Neuronal response is ‘noisy’:
Population recordings:
The basic learning problem: what about the stimulus do neurons respond to?
The basic learning problem: what about the stimulus do neurons respond to?
What about the stimulus do neurons respond to?

LGN Receptive Fields
- On-Centered
- Off-Centered

Red -> ON subregions
Green -> OFF subregions
Both types are roughly circularly symmetrical

Cortical "Simple" Cell Receptive Field
- Oriented along a particular axis
- Segregated On and Off subfields
What about the stimulus do neurons respond to?
Tools we use:

- statistics
- learning / inference
- optimization
- dynamical systems
- information theory
A brief introduction to information theory:

*When is information theory useful? When you...*

- want to go beyond linear correlation
- have enough data to sample $P(x,y)$
- are not sure what your ‘code’ is
A brief introduction to information theory:

When is information theory useful? When you...

- want to go beyond linear correlation
- have enough data to sample $P(x,y)$
- are not sure what your ‘code’ is
Zero correlation, no information:
Zero correlation, but obvious information:
Entropy as a measure of uncertainty:

uncertainty = \log (n)
Entropy as a measure of uncertainty:

\[
\text{uncertainty} = \log (n) \\
= \log (1/p) \\
= - \log (p)
\]
Entropy as a measure of uncertainty:

\[
\text{uncertainty} = \log(n) \\
= \log(1/p) \\
= -\log(p)
\]

\[
u_i = -\log(p_i)
\]

\[
\langle u_i \rangle = - \sum_i p_i \log(p_i)
\]
Entropy as a measure of uncertainty:

\[
\text{uncertainty} = \log (n) \\
= \log (1/p) \\
= - \log (p)
\]

\[
u_i = - \log (p_i)
\]

\[
\langle u_i \rangle = - \sum_i p_i \log (p_i)
\]

\[
S(X) = - \sum_x p(x) \log_2 (p(x))
\]
Information == reduction in uncertainty

Mutual information:

\[ I(A; B) = S(A) - S(A|B) \]
\[ = S(B) - S(B|A) \]
\[ = \sum_{a,b} P(a, b) \log_2 \left( \frac{P(a, b)}{P(a)P(b)} \right) \]
\[ = \sum_{a,b} P(a)P(b|a) \log_2 \left( \frac{P(b|a)}{P(b)} \right) \]
Information == reduction in uncertainty

Mutual information:

\[
I(A; B) = S(A) - S(A|B) = S(B) - S(B|A) = \sum_{a,b} P(a,b) \log_2 \left( \frac{P(a,b)}{P(a)P(b)} \right) = \sum_{a,b} P(a)P(b|a) \log_2 \left( \frac{P(b|a)}{P(b)} \right)
\]
useful formulae

**Product rule:**

\[ P(a, b) = P(a|b)P(b) \]
useful formulae

**Product rule:**

\[ P(a, b) = P(a|b)P(b) \]

**Sum rule:**

\[ P(a) = \sum_b P(a, b) \]

\[ = \sum_b P(a|b)P(b) \]
useful formulae

**Product rule:**

\[ P(a, b) = P(a|b)P(b) \]

**Sum rule:**

\[ P(a) = \sum_b P(a, b) \]
\[ = \sum_b P(a|b)P(b) \]

**Bayes’ rule:**

\[ P(a|b) = \frac{P(b|a)P(a)}{P(b)} \]
\[ = \frac{P(b|a)P(a)}{\sum_{a'} P(b|a')P(a')} \]
more useful formulae
more useful formulae

Additivity:

\[ S(A, B) = S(A) + S(B) \iff P(a, b) = P(a)P(b) \]
more useful formulae

Additivity:

\[ S(A, B) = S(A) + S(B) \iff P(a, b) = P(a)P(b) \]

Chain rule:

\[ S(A, B) = S(A) + S(B|A) = S(B) + S(A|B) \]
more useful formulae

Additivity:

\[ S(A, B) = S(A) + S(B) \iff P(a, b) = P(a)P(b) \]

Chain rule:

\[ S(A, B) = S(A) + S(B|A) = S(B) + S(A|B) \]

Kullback-Liebler divergence (\(D_{KL}\)):

\[ D_{KL}(P, Q) = \sum_a P(a) \log_2 \frac{P(a)}{Q(a)} \]
more useful formulae

**Additivity:**

\[ S(A, B) = S(A) + S(B) \iff P(a, b) = P(a)P(b) \]

**Chain rule:**

\[ S(A, B) = S(A) + S(B|A) = S(B) + S(A|B) \]

**Kullback-Liebler divergence \( D_{KL} \):**

\[ D_{KL}(P, Q) = \sum_a P(a) \log_2 \frac{P(a)}{Q(a)} \]

**Mutual information \( \geq 0 \):**

\[ I(A; B) = S(A) - S(A|B) \]
Efficient coding in single neurons
The efficient coding hypothesis, brief history:

- Claude Shannon (1948) *A Mathematical Theory of Communication*
- Fred Attneave (1954) *Some informational aspects of visual perception*
- Horace Barlow (1961) *Possible principles underlying the transformation of sensory messages*

Are sensory systems optimized for information transmission?
Information theory example, the weighing problem:
Information theory example, the weighing problem:

chalkboard interlude