Sensory encoding models

Eero P. Simoncelli
CNS / Courant

Computational Modeling of Neuronal Systems
Spring, 2010
Role of Models

- Dayan & Abbott
Role of Models

- Descriptive (what?)

- Dayan & Abbott
Role of Models

- Descriptive (what?)
  - eg: tuning curves, receptive field, LNP

- Dayan & Abbott
Role of Models

• Descriptive (what?)
  • eg: tuning curves, receptive field, LNP

• Mechanistic (how?)

- Dayan & Abbott
Role of Models

• Descriptive (what?)
  • eg: tuning curves, receptive field, LNP

• Mechanistic (how?)
  • eg: compartmental models, Hodgkin-Huxley

- Dayan & Abbott
Role of Models

- Descriptive (what?)
  - eg: tuning curves, receptive field, LNP
- Mechanistic (how?)
  - eg: compartmental models, Hodgkin-Huxley
- Interpretive/Explanatory (why?)

- Dayan & Abbott
Role of Models

• Descriptive (what?)
  • eg: tuning curves, receptive field, LNP

• Mechanistic (how?)
  • eg: compartmental models, Hodgkin-Huxley

• Interpretive/Explanatory (why?)
  • eg: efficient coding, optimal estimation/decision, wiring length, metabolic cost, etc

- Dayan & Abbott
Interaction with Experiments
Interaction with Experiments

• Fit existing data
Interaction with Experiments

• Fit existing data
• Make predictions...
Interaction with Experiments

• Fit existing data
• Make predictions...
  - for other neurons, under other conditions
Interaction with Experiments

• Fit existing data
• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
Interaction with Experiments

• Fit existing data
• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
Interaction with Experiments

• Fit existing data

• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
Interaction with Experiments

• Fit existing data

• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
  - that can be tested with new experiments....
Interaction with Experiments

• Fit existing data

• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
  - that can be tested with new experiments....

• Develop new experiments...
Interaction with Experiments

• Fit existing data
• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
  - that can be tested with new experiments....
• Develop new experiments...
  - to refine model
Interaction with Experiments

• Fit existing data

• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
  - that can be tested with new experiments....

• Develop new experiments...
  - to refine model
  - to differentiate models
Interaction with Experiments

• Fit existing data

• Make predictions...
  - for other neurons, under other conditions
  - of mechanisms not yet understood (e.g., HH)
  - of behavior
  - in other animals/species
  - that can be tested with new experiments....

• Develop new experiments...
  - to refine model
  - to differentiate models
  - with optimized stimuli, to characterize cells
Descriptive Response Models (outline)

- Receptive fields and tuning curves
- Linear models
- Rate models
- Wiener/Volterra (polynomial) models
- LN models
- Poisson spiking
- Fitting/validating LNP models
Rate coding

Some of Adrian's first recordings from very small numbers of individual nerve fibres. Each spiky deflection is a single nerve impulse. These records were taken from the sensory nerves of a cat's toe. The toe was flexed slowly, more quickly and very rapidly to produce these three traces. The frequency of firing depends on the strength of the stimulus – Adrian's law.
Receptive Fields

• Classical: A region of the retina (visual field) that must be stimulated directly in order to obtain a response in a neuron
  – Sherrington (1906), Hartline (1938), Kuffler (1953)
Receptive Fields

• Classical: A region of the retina (visual field) that must be stimulated directly in order to obtain a response in a neuron
  – Sherrington (1906), Hartline (1938), Kuffler (1953)

• Modern generalization: Kernel that captures those attributes of the stimulus that generate/modulate responses. Often assumed linear.
- Dayan & Abbott, after Hubel & Wiesel ‘62
Estimating firing rates

- Dayan & Abbott
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]

# pars:
- const: 1
- vector: \(n\)
  \(n\) \(=\) (20)
- matrix: \(n^2\)
  \(n^2\) \(=\) (400)
- 3-tensor: \(n^3\)
  \(n^3\) \(=\) (8000)
Polynomial Model  
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]

- estimate kernels using moments of spike-triggered stimuli
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]

- estimate kernels using moments of spike-triggered stimuli
- in practice, insufficient data to go beyond 2\textsuperscript{nd} order
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + k_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]

# pars:
- const: 1
- vector: \( n \) (20)
- matrix: \( n^2 \) (400)
- 3-tensor: \( n^3 \) (8000)
Polynomial Model
(Volterra/Weiner Kernels)

\[ r(\vec{x}) = k_0 + \vec{k}_1 \cdot \vec{x} + \vec{x}^T K_2 \vec{x} + \ldots \]

- \# pars: 1
  - const: 1 (20)
  - vector: \(n\)
  - matrix: \(n^2\)
  - 3-tensor: \(n^3\)

- in practice, insufficient data to go beyond 2\textsuperscript{nd} order
Low-order polynomials do a poor job of representing the nonlinearities found in neurons.
Low-order polynomials do a poor job of representing the nonlinearities found in neurons.
Low-order polynomials do a poor job of representing the nonlinearities found in neurons.
• Threshold-like nonlinearity => linear classifier
• Classic model for Artificial Neural Networks
  - McCullough & Pitts (1943), Rosenblatt (1957), etc
• No spikes (output is firing rate)
LNP cascade model

- Simplest successful descriptive spiking model
- Easily fit to (extracellular) data
- Descriptive, and interpretable (although not mechanistic)
Geometric view of Poisson models

1D stimulus over time
(e.g., flickering bars)

- 8 x 6 stimulus block
  = 48-dimensional vector
Geometric picture

Stimulus

Response

time →

• raw stimuli
• spiking stimuli
Neural response is captured by relationship between the distribution of red points (spiking stim) and blue points (raw stim)

Expressed in terms of Bayes’ rule:

$$P(\text{spike}|\text{stim}) = \frac{P(\text{spike, stim})}{P(\text{stim})}$$

Cannot be estimated directly
ML estimation of LNP

[on board]
ML estimation of LNP

If $f_\theta(\vec{k} \cdot \vec{x})$ is convex (in $x$ and theta), and $\log f_\theta(\vec{k} \cdot \vec{x})$ is concave, then the likelihood of the LNP model is convex (for all observed data, $\{n(t), \vec{x}(t)\}$).

[Paninski, ’04]
ML estimation of LNP

If $f_\theta(\vec{k} \cdot \vec{x})$ is convex (in x and theta), and $\log f_\theta(\vec{k} \cdot \vec{x})$ is concave, then the likelihood of the LNP model is convex (for all observed data, $\{n(t), \vec{x}(t)\}$ )

Examples: $e^{(\vec{k} \cdot \vec{x}(t))}$

$(\vec{k} \cdot \vec{x}(t))^\alpha$, $1 < \alpha < 2$

[Paninski, ’04]
Simple LNP fitting

• Assuming:
  – stochastic stimuli, spherically distributed
  – spike counts in small time bins (0,1)
  – neural response is such that mean of spike-triggered ensemble is shifted

• Reverse correlation gives an unbiased estimate of $k$ [on board]

• For exponential $f$, this is same as ML

- Bussgang 52; de Boer & Kuyper 68
Computing the STA

Stimulus

Response

time

STA

+ 

raw stimuli

spiking stimuli
STA corresponds to a “direction” in stimulus space
Projecting onto the STA

\[ P \left( \text{spike}(t) \mid \vec{k} \cdot \vec{s}(t) \right) = \frac{P \left( \text{spike}(t) \& \vec{k} \cdot \vec{s}(t) \right)}{P \left( \vec{s}(t) \right)} \]
Projecting onto the STA
Projecting onto the STA
Projecting onto the STA
Projecting onto an axis orthogonal to the STA
Projecting onto an axis orthogonal to the STA
Projecting onto an axis orthogonal to the STA
The spike-triggered average L-cone contrast during random adaptation operates over tens of seconds (Fig. 2) (Smirnakis et al., 1997; Kim and Rieke, 2001), responses from the adapting low and high contrast. Thus, the stimulus that controlled the state of the ON cell was simultaneously used to probe light responses in the second stage of the LN model captured a significant feature of the input strength filtering of recent visual inputs.

To examine temporal contrast adaptation, RGC light responses were characterized as above using random filtering of contrast stimuli. Were RGC light responses linear in the present conditions, the functions in Figure 3, could be meaningfully compared if scaling the amplitude of the high contrast linear analysis to the data. Stimulus: 33 Hz flicker stimuli of low intensity.

Figure 4 shows low and high contrast STAs from a representative OFF cell. The corresponding nonlinearities are shown in Figure 4A–E simultaneously recorded in salamander retina. A–E, Fractional change in peak divided by the peak of the low contrast STA, and the high contrast STA for altering low and high contrast stimulation. The high nonlinear STA, this was scaled so that its peak divided by the peak of the low contrast STA equals the ratio of the peak sensitivity (H11006) and time to zero crossing (H17040).

Characterization of light response in one ON cell (A, B) and one OFF cell (C, D) simultaneously recorded in salamander retina. A, C, B, D.

- Chander & Chichilnisky 01
- Pillow et al., 2004
74% of var

- Pillow et al., 2004
V1 simple cell

- Ozhawa, et al.
LNP summary

- LNP is the de facto standard descriptive model, and is implicit in much of the experimental literature
- Accounts for basic RF properties
- Accounts for basic spiking properties (rate code)
- Easily fit to data
- Easily interpreted
- BUT, non-mechanistic, and exhibits striking failures (esp. beyond early sensory/motor) ...
LNP limitations

- Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells)
LNP limitations

- Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells)
  ➞ Subspace LNP
an estimate of the network's functional connectivity [7] (see Methods).

Using a likelihood-based "pruning" procedure, we were able to reduce the number of coupling filters to a minimal set, which reduces the model's computational complexity and provides an efficient set, which reduces the model's computational complexity and provides...
V1 simple cell

[STA variance time tf STC]

[Rust, Schwartz, Movshon, Simoncelli, ‘05]
LNP limitations

• Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells)
  ➡ Subspace LNP

• Responses depend on spike history, other cells
LNP limitations

- Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells)
  ➔ Subspace LNP

- Responses depend on spike history, other cells
  ➔ Recursive models (GLM)
Linear-Nonlinear-Poisson (LNP)
Recursive LNP

[Truccolo et al ‘05; Pillow et al ‘05]
LNP limitations

• Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells) ➔ Subspace LNP

• Responses depend on spike history, other cells ➔ Recursive models (GLM)

• White noise doesn’t drive mid- to late-stage neurons well
LNP limitations

- Symmetric nonlinearities and/or multi-dimensional front-end (e.g., V1 complex cells)
  ➡ Subspace LNP

- Responses depend on spike history, other cells
  ➡ Recursive models (GLM)

- White noise doesn’t drive mid- to late-stage neurons well
  ➡ Specialized “afferent” stimuli
Credits

• Spike-triggered covariance: Odelia Schwartz, Jonathan Pillow, Liam Paninski, Nicole Rust

• Stochastic integrate-and-fire & rLNP models: Jonathan Pillow, Liam Paninski

• V1/MT physiology/modeling: Nicole Rust, Tony Movshon (NYU)

• Retinal physiology/modeling: Jonathon Shlens, Valerie Uzzell, Divya Chander, EJ Chichilnisky (Salk Institute)