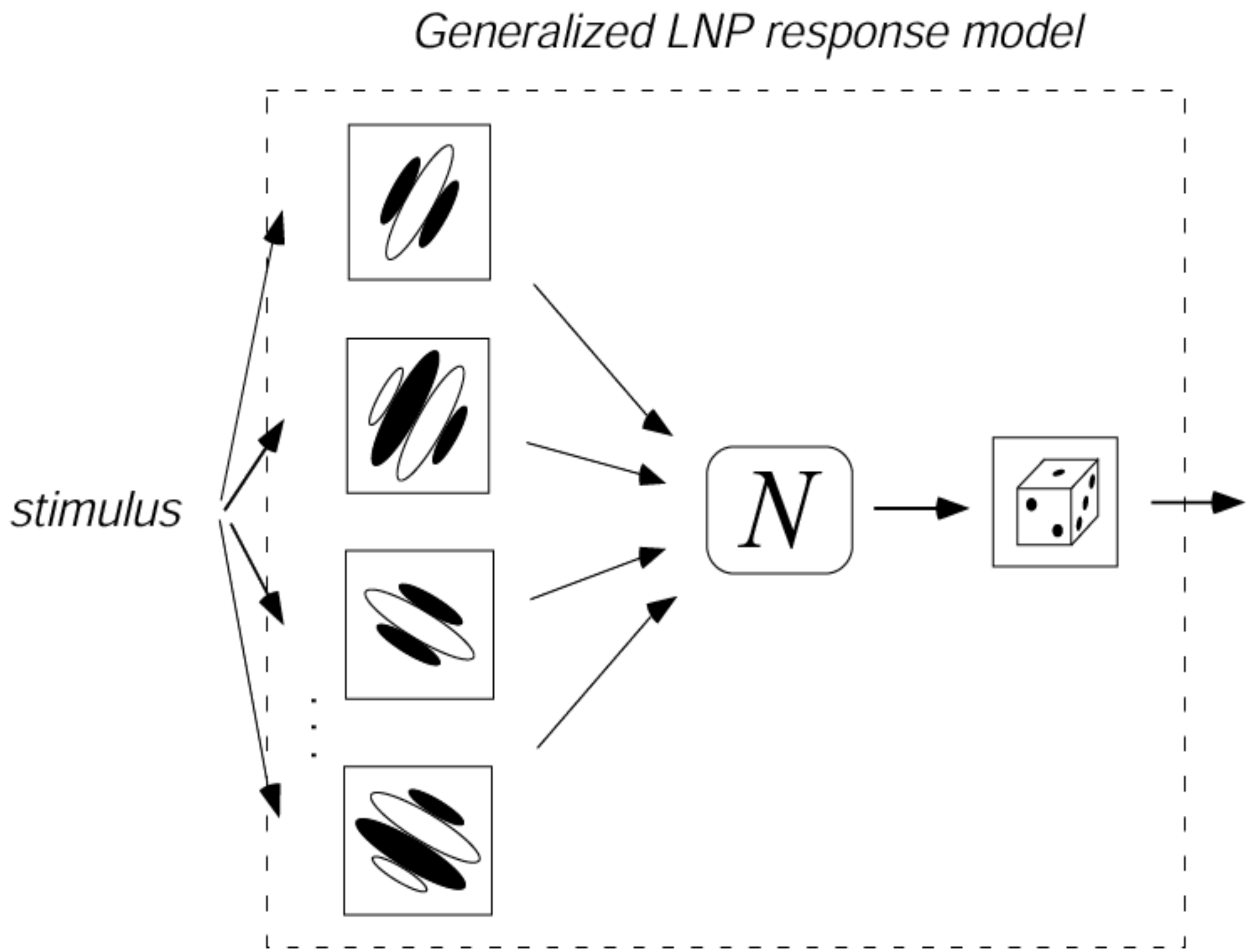
Neural coding: Part 3

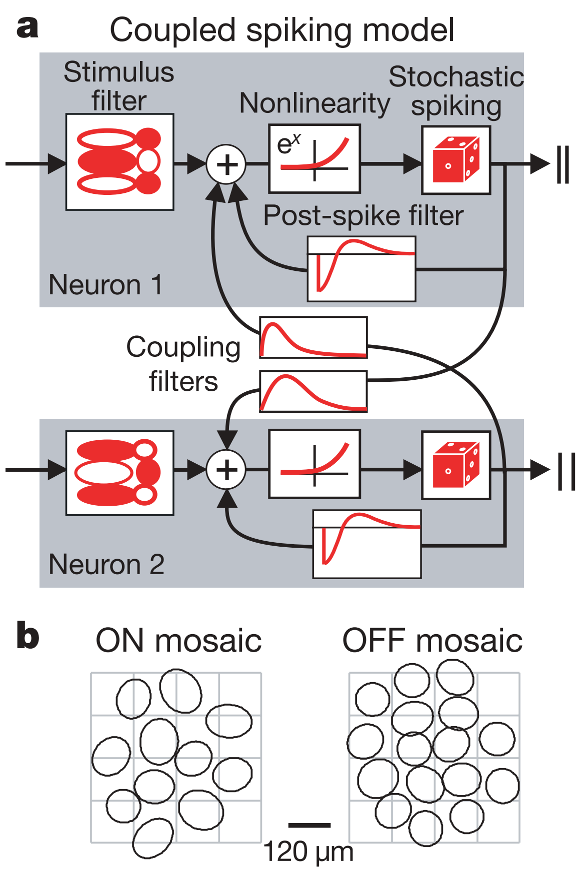
Odelia Schwartz

Last time



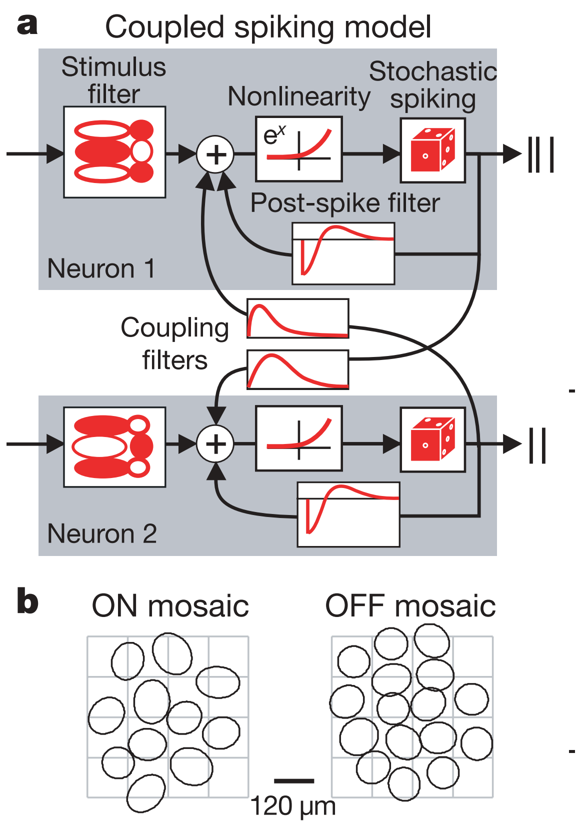
Methods paper on solving with Spike-triggered approaches:
Schwartz, Pillow, Rust, Simoncelli 2006

More complete visual system (e.g., retina)



Pillow et al., Nature, 2008

More complete visual system (e.g., retina)



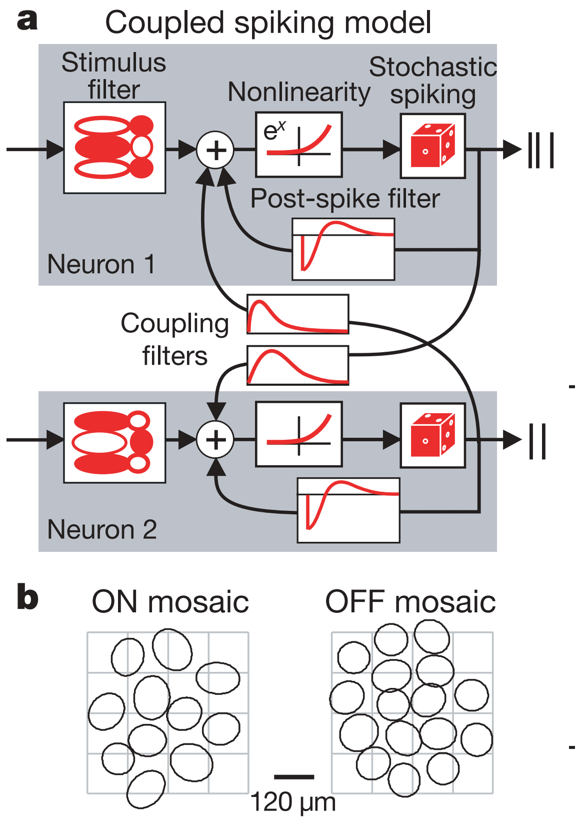
Stimulus filter: spatio-temporal light integration

Post-spiking filter: Voltage-activated currents (time course of recovery after a spike)

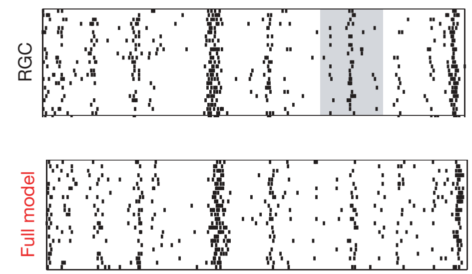
Coupling filters: e.g., synaptic interactions between cells

Pillow et al., Nature, 2008

More complete visual system (e.g., retina)

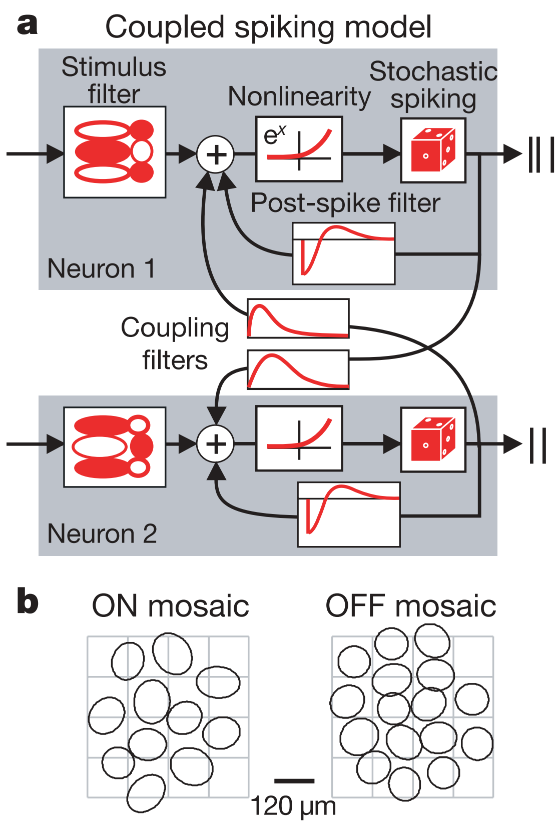


- Predict spike trains with coupled and uncoupled model

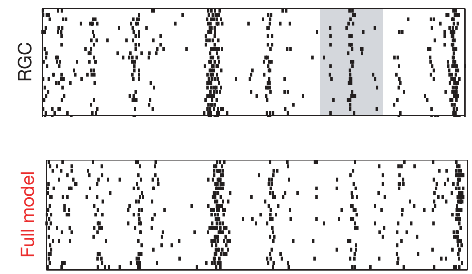


Pillow et al., Nature, 2008

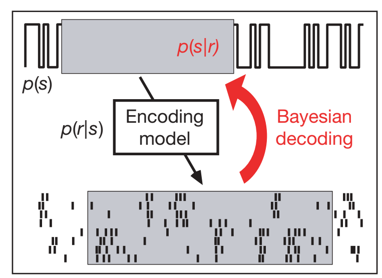
More complete visual system (e.g., retina)



- Predict spike trains with coupled and uncoupled model



- Opposite direction: from spike train model, predict input stimuli

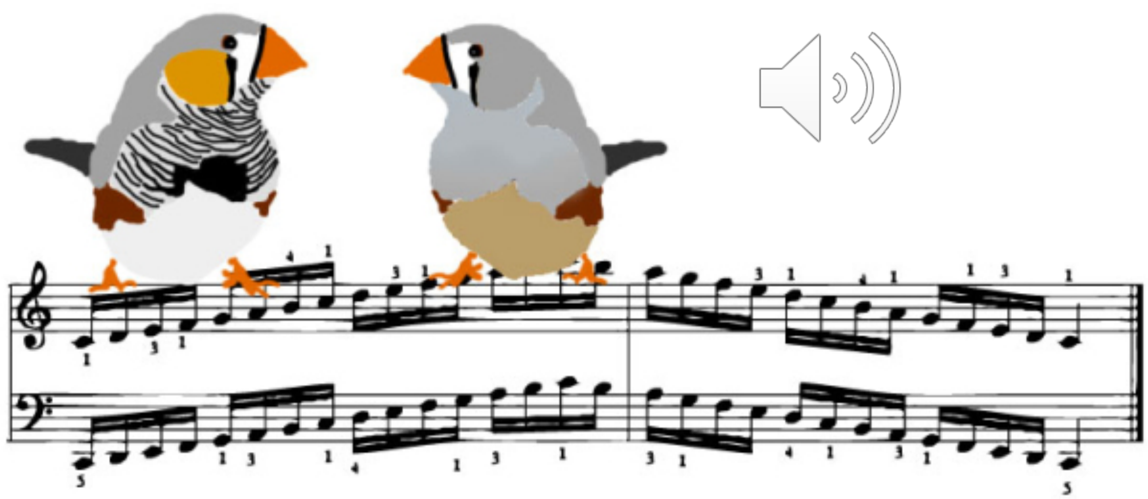


Pillow et al., Nature, 2008
Also used beyond vision, for instance:
Pillow et al., Nature Neuroscience 2014: Encoding and decoding in parietal cortex during sensorimotor decision-making

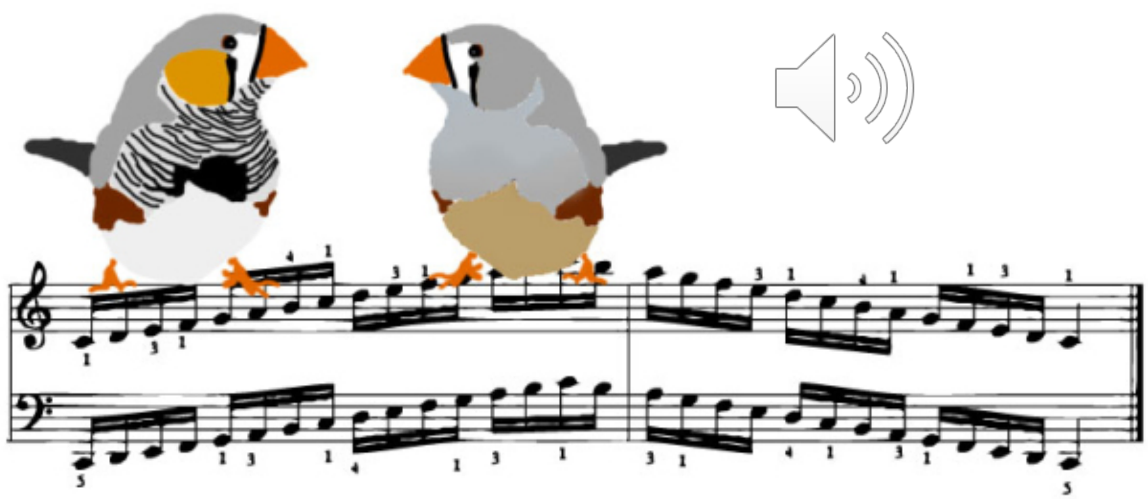
Another example system and coding

Ultra Sparse Song Bird System

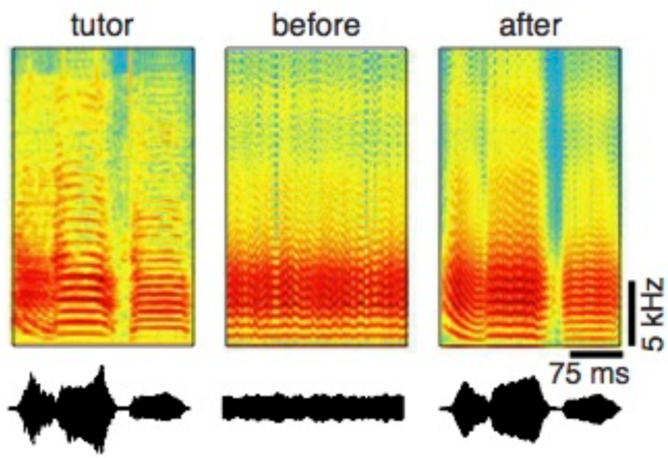
Song before learning



Song after learning

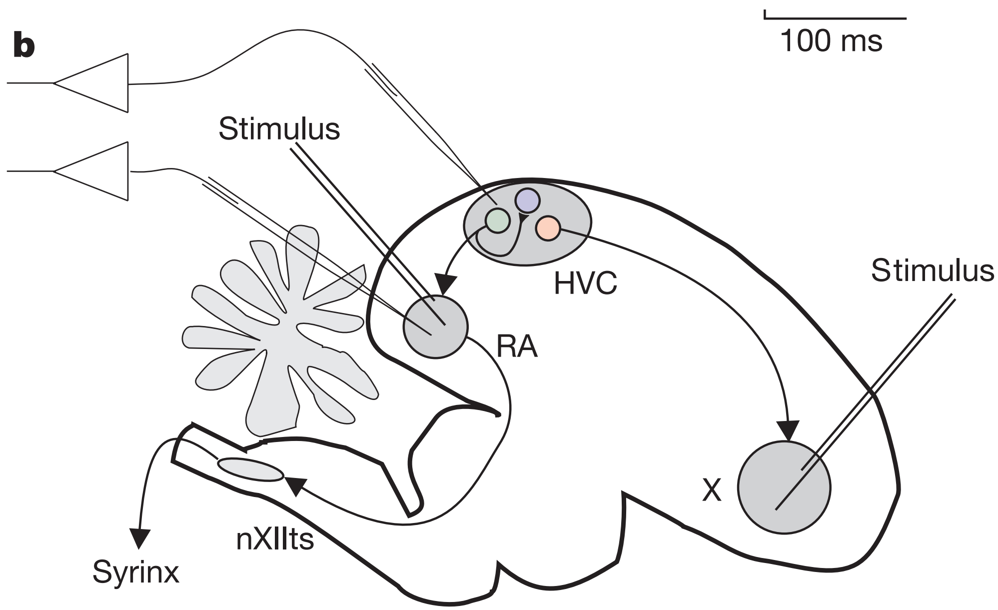


Songbird



Fiete et al. 2009 review paper

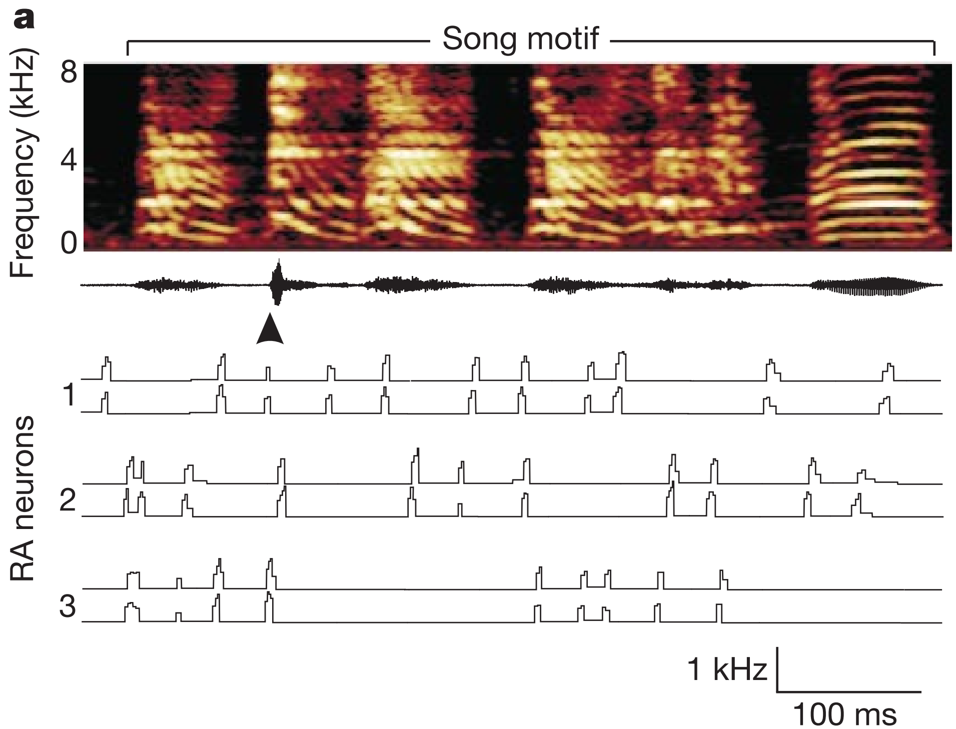
Songbird



Hahnloser et al. 2002, Nature

HVC neurons connect to RA neurons, which control muscles

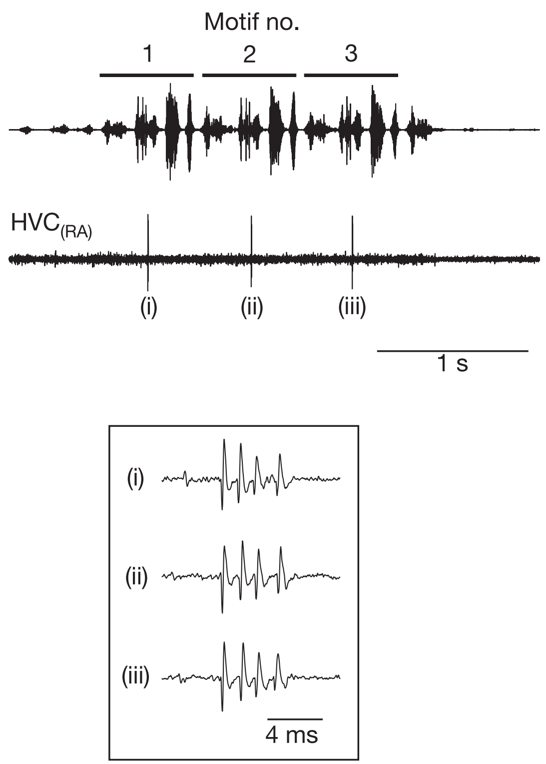
Songbird



Hahnloser et al. 2002, Nature

RA neurons fire at multiple times during a song

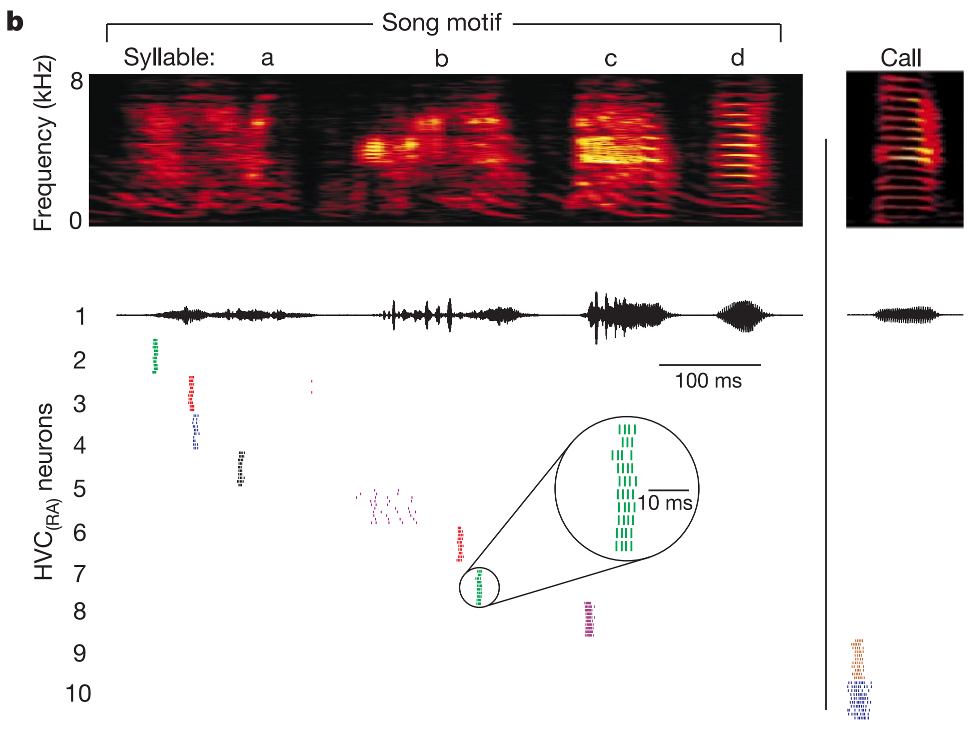
Songbird



Hahnloser et al. 2002, Nature

HVC neurons burst reliably at a single precise time in the song or call!

Songbird



Hahnloser et al. 2002, Nature

HVC neurons burst reliably at a single precise time in the song or call

Songbird model

Why ultra sparse responses in the songbird??

Songbird model

Why ultra sparse responses in the songbird??

“Intuitively … minimizing interference between different synapses during learning … In this paper we make the intuitive argument more concrete.”

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong
Also: Doya and Sejnowski 1995 (**considered sparseness in a model before known**)

Songbird model

Why ultra sparse responses in the songbird??

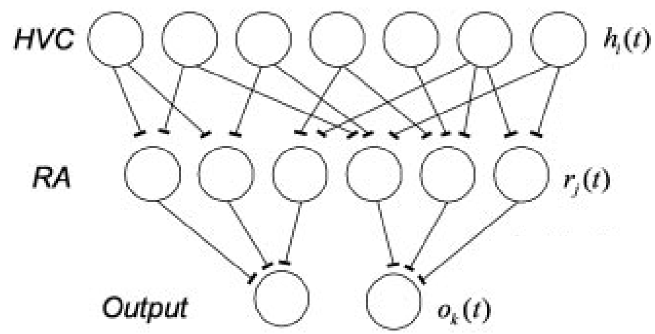
We'll look at modeling work, and also introduce network modeling approaches…

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled;
7 in real system)



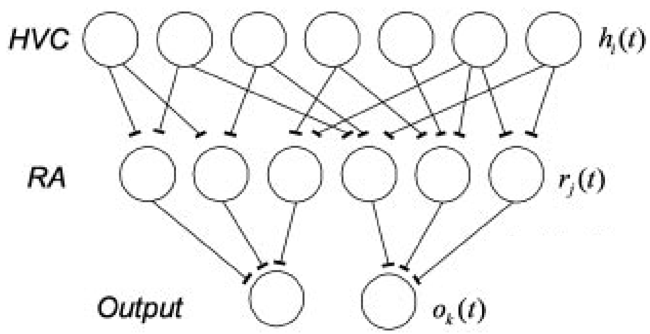
Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons (number of vocal muscles controlled; 7 in real system)



Input known (binary burst pulses chosen randomly of either 1, 2, 4 or 8 bursts per motif)

Hidden

Desired output known

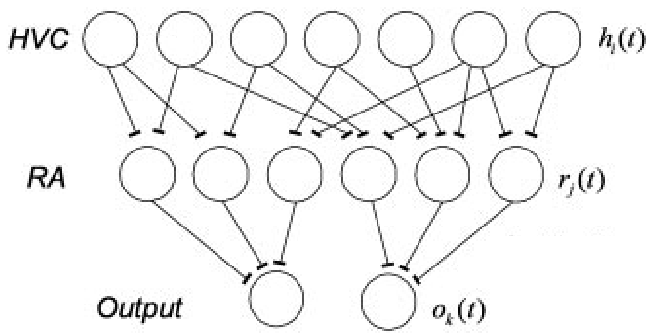
Fiete et al. 2004: Temporal Sparseness of the Premotor Drive Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled;
7 in real system)



Input known (binary burst
pulses chosen randomly of either
1, 2, 4 or 8 bursts per motif)

Hidden

Desired output known

Goal: minimize error between network output and desired output

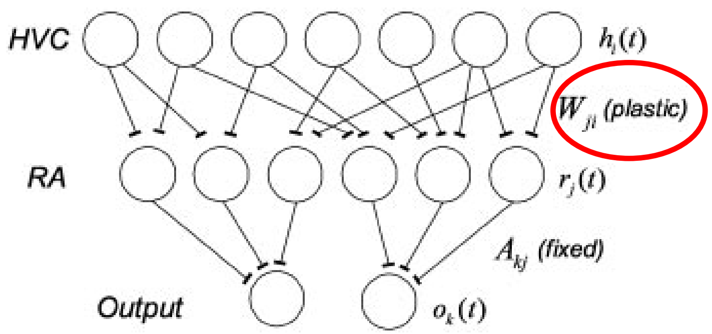
Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons (number of vocal muscles controlled; 7 in real system)



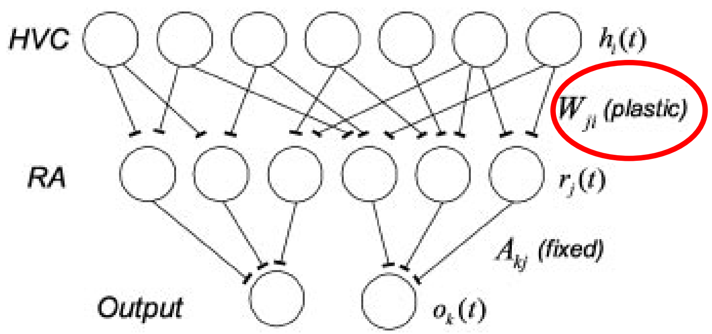
Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)

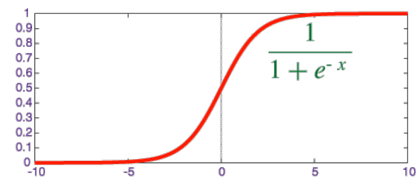


$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

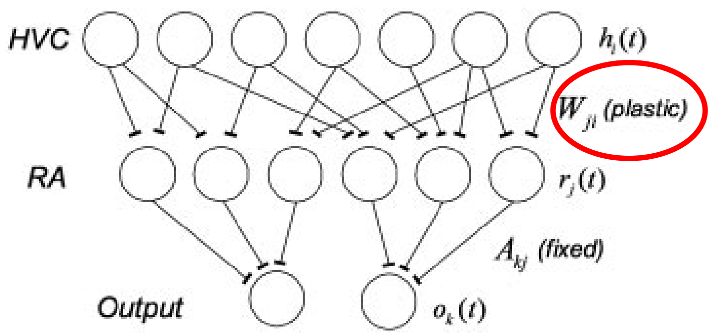


sigmoid

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)



threshold

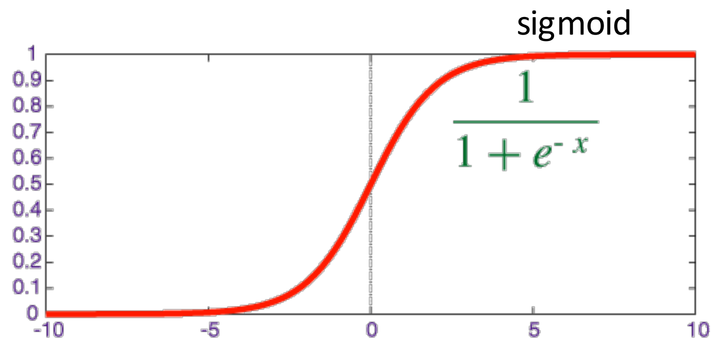
$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Sigmoid curve

$$r_j(t) = f(x)$$



$$x = \sum_{i=1}^{N_h} w_{ji} h_i(t) - \Phi_j$$

$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

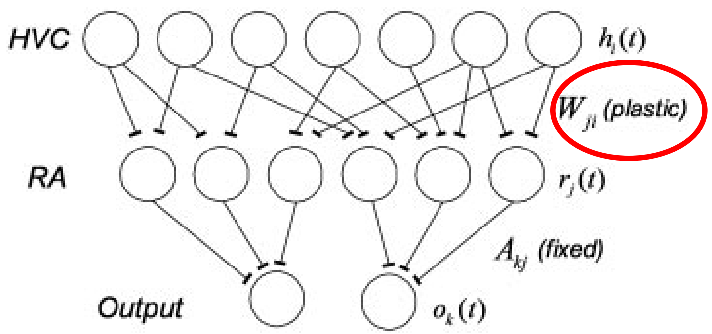
Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)



$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

We know inputs and desired outputs

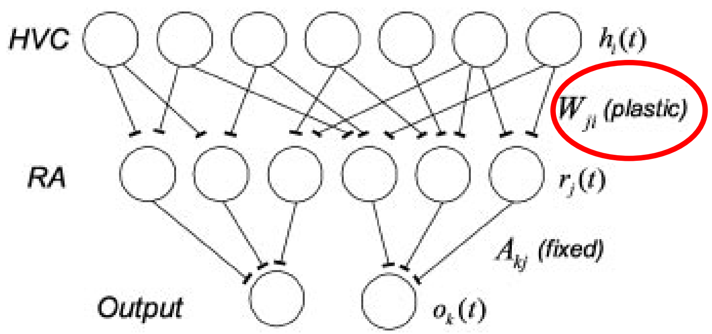
Back Propagation: 1970s; Rumelhart, Williams, Hinton,
Nature, 1986; and prominent again today in deep networks

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)



$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:
Compare outputs with correct answer to get error

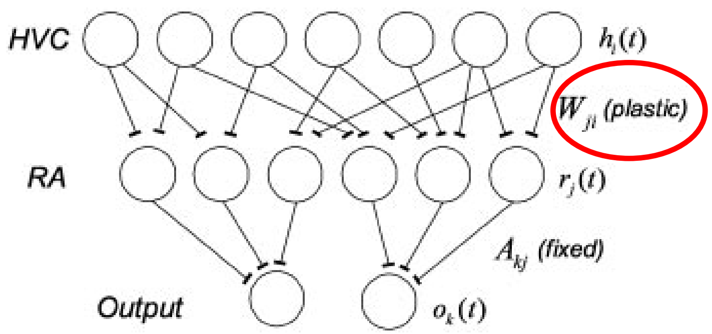
Back Propagation: 1970s; Rumelhart, Williams, Hinton,
Nature, 1986; and prominent again today in deep networks

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)



$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:
Compare outputs with correct answer to get error
What kind of machine learning approach is this??

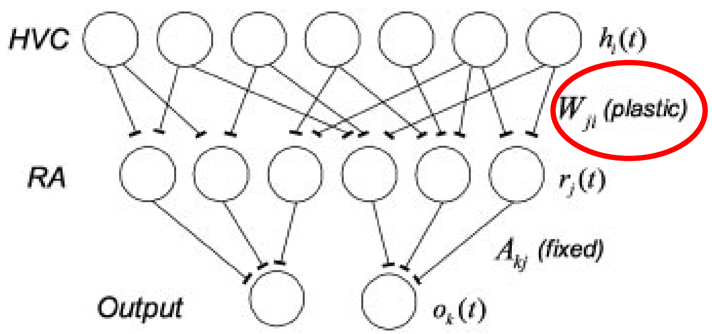
Back Propagation: 1970s; Rumelhart, Williams, Hinton, Nature, 1986; and prominent again today in deep networks

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons (number of vocal muscles controlled)



$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:

Compare outputs with correct answer to get error

What kind of machine learning approach is this?? Supervised learning

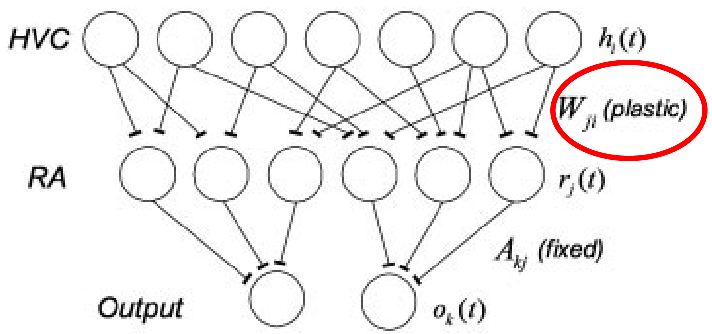
Back Propagation: 1970s; Rumelhart, Williams, Hinton, Nature, 1986; and prominent again today in deep networks

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)



$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:
Compare outputs with correct answer to get error
(What other approach could be relevant here??)

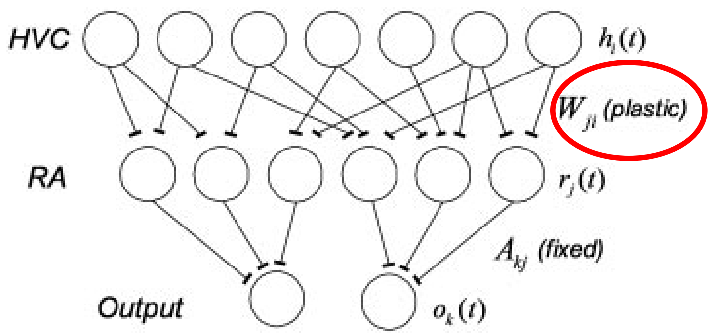
Back Propagation: 1970s; Rumelhart, Williams, Hinton,
Nature, 1986; and prominent again today in deep networks

Songbird model

500 HVC neurons

800 RA neurons

2 Output Neurons (number of vocal muscles controlled)



$$r_j(t) = f\Big[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\Big]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:

Compare outputs with correct answer to get error

(What other approach could be relevant here?? Reinforcement learning)

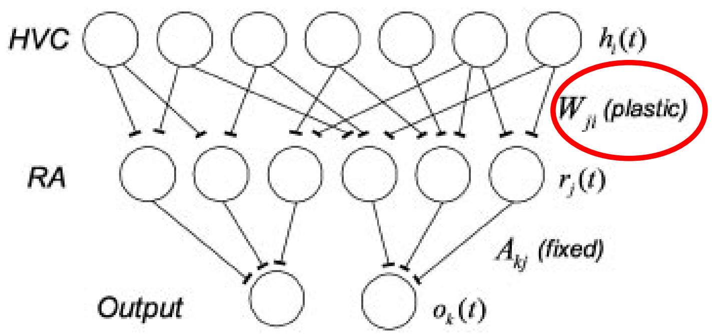
Back Propagation: 1970s; Rumelhart, Williams, Hinton, Nature, 1986; and prominent again today in deep networks

Songbird model

Nh = 500 HVC neurons

Nr = 800 RA neurons

No = 2 Output Neurons (number of vocal muscles controlled)



$$r_j(t) = f\Big[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\Big]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Back propagation:

- Compare current outputs with correct desired answer to get error
- Update weights by small step down gradient

Back Propagation: 1970s; Rumelhart, Williams, Hinton, Nature, 1986; and prominent again today in deep networks

Back propagation

Error:

$$C = \int_0^T dt \sum_{k=1}^{N_o} [\mathrm{d}_k(t) - o_k(t)]^2$$

Desired minus actual outputs

$$r_j(t) = f[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j]$$

Back Propagation Gradient descent:

Derivative of RA neuron rj with respect to weights

$$\Delta W_{ji} = -\eta \frac{\partial C}{\partial W_{ji}} = \eta \int_0^T \mathrm{d}t \sum_{k=1}^{N_o} 2[\mathrm{d}_k(t) - o_k(t)] A_{kj} f_j' h_i$$

Learning rate

Back Propagation: 1970s; Rumelhart, Williams, Hinton, Nature, 1986; and prominent again today in deep networks

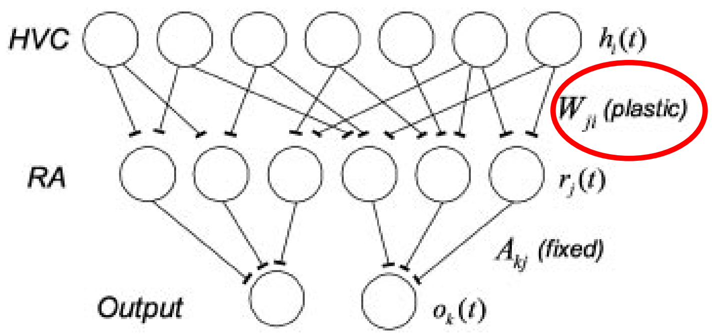
Songbird model

Do sparse HVC responses help learning??

500 HVC neurons

800 RA neurons

2 Output Neurons
(number of vocal
muscles controlled)

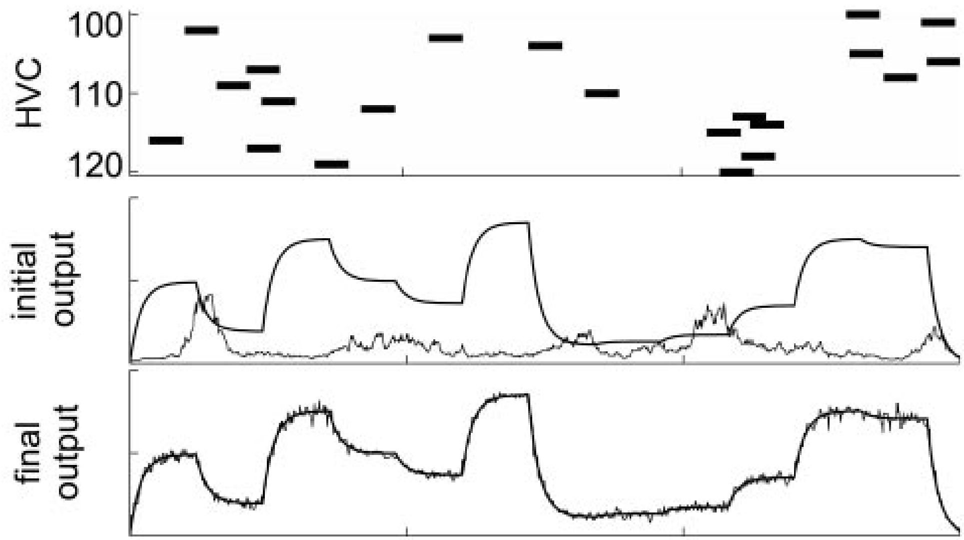


$$r_j(t) = f\left[\sum_{i=1}^{N_h} W_{ji} h_i(t) - \theta_j\right]$$

$$O_k(t) = \sum_{j=1}^{N_r} A_{kj} r_j(t)$$

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model



20 HVC neurons each bursting once

Initial versus desired

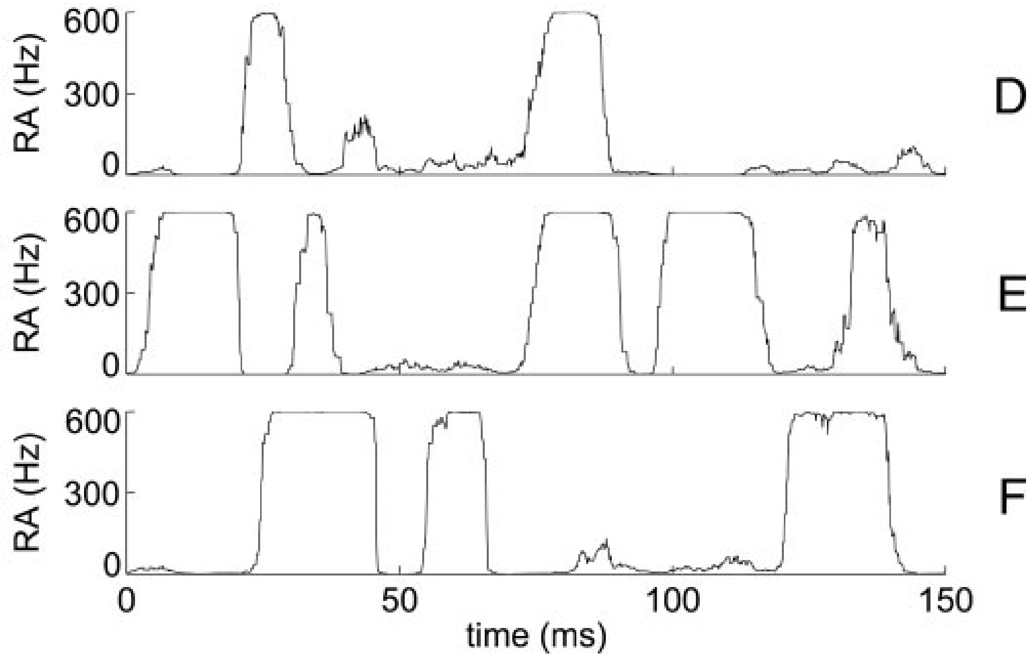
final versus desired

Top: HVC units; middle: initial network output; and bottom: final network output matching desired output for one of the two output units

output units

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive Is Important for Rapid Learning in a Neural Network Model of Birdsong

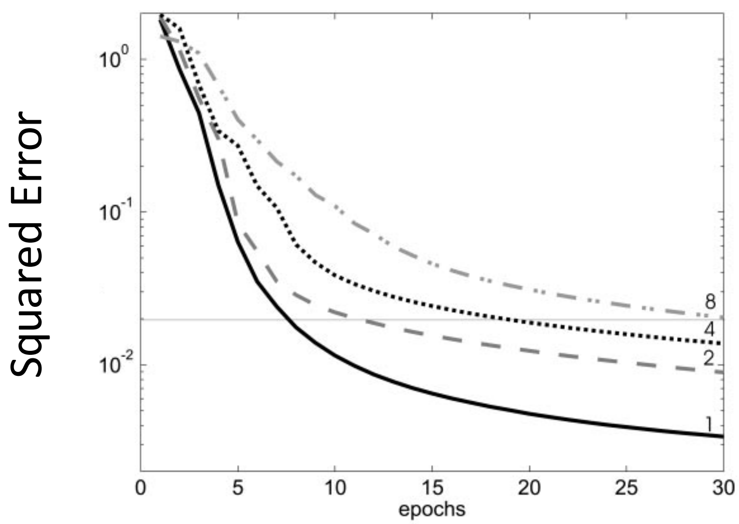
Songbird model



3 RA units after learning

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

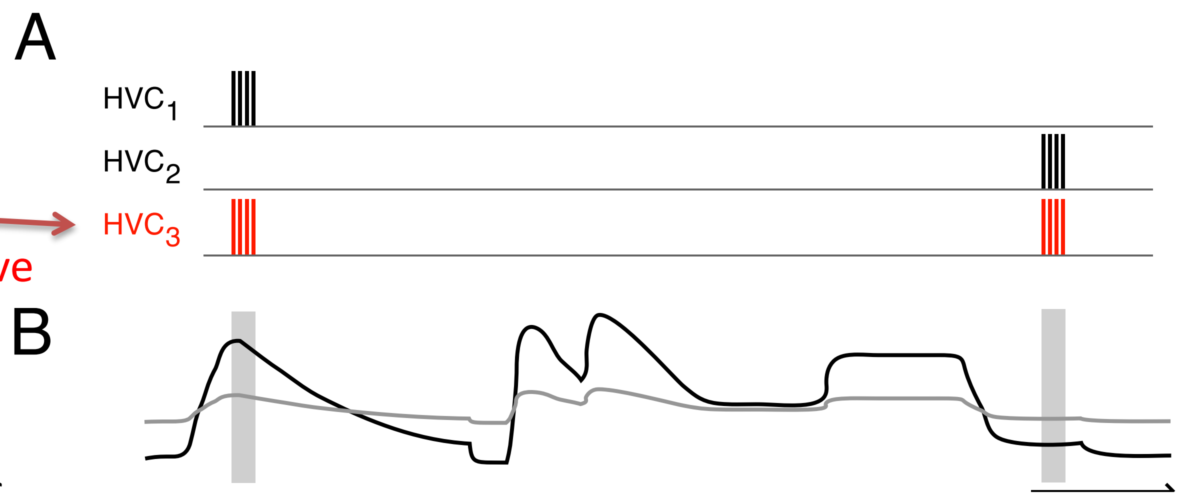


Most sparse

Each line plot: Varying number of bursts per motif of simulated HVC neurons.
Lowest error for 1 burst per motif.

Fiete et al. 2004: Temporal Sparseness of the Premotor Drive
Is Important for Rapid Learning in a Neural Network Model of Birdsong

Songbird model

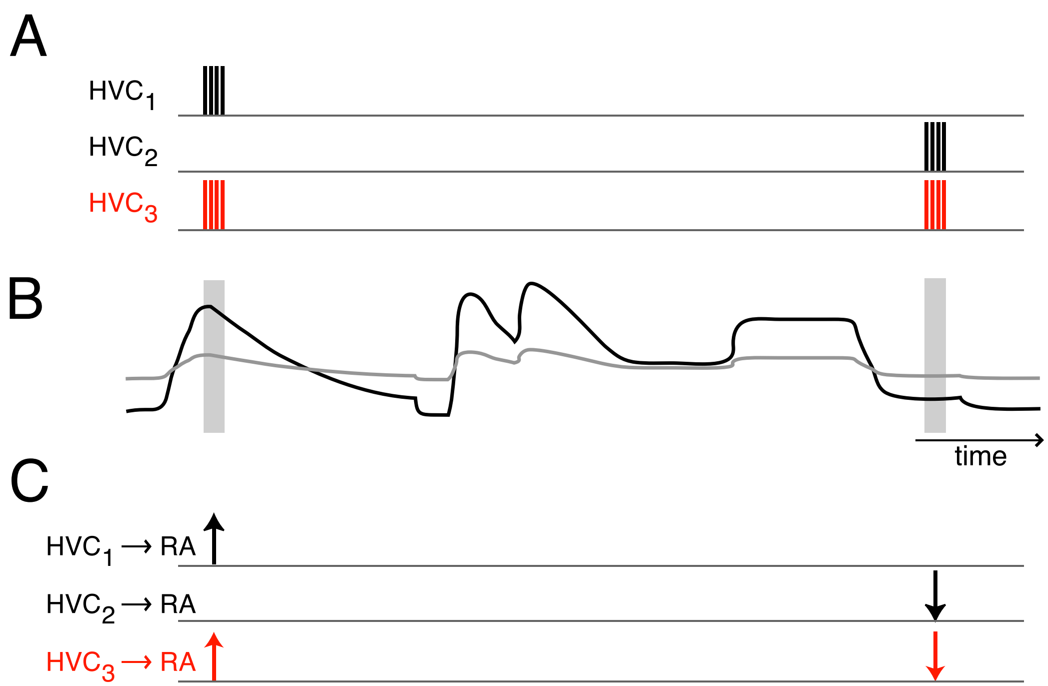


Mixed (neuron active twice)

Tutor versus pupil outputs

Fiete et al. 2009, review

Songbird model

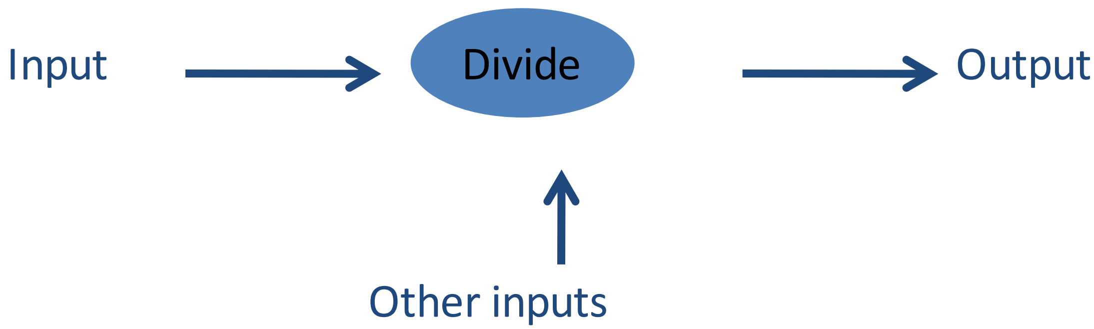


Fiete et al. 2009, review

Synapse should be strengthened and weakened -> Conflicting demands causes slowdown of learning

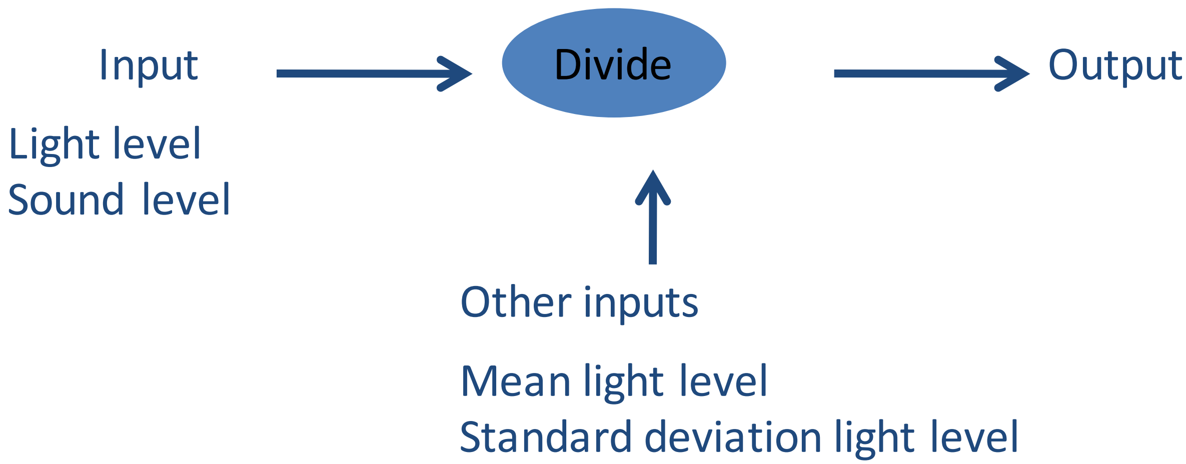
Canonical computations in the brain??

Divisive normalization model



- Descriptive neural model
- Canonical computation (Carandini, Heeger, Nature Reviews, 2012)
- Has mechanistic and interpretive versions
- Related to gain control in engineering

Divisive normalization model



- Descriptive neural model
- Canonical computation (Carandini, Heeger, Nature Reviews, 2012)
- Has mechanistic and interpretive versions
- Related to gain control in engineering

Divisive normalization model

Simple version of descriptive model:

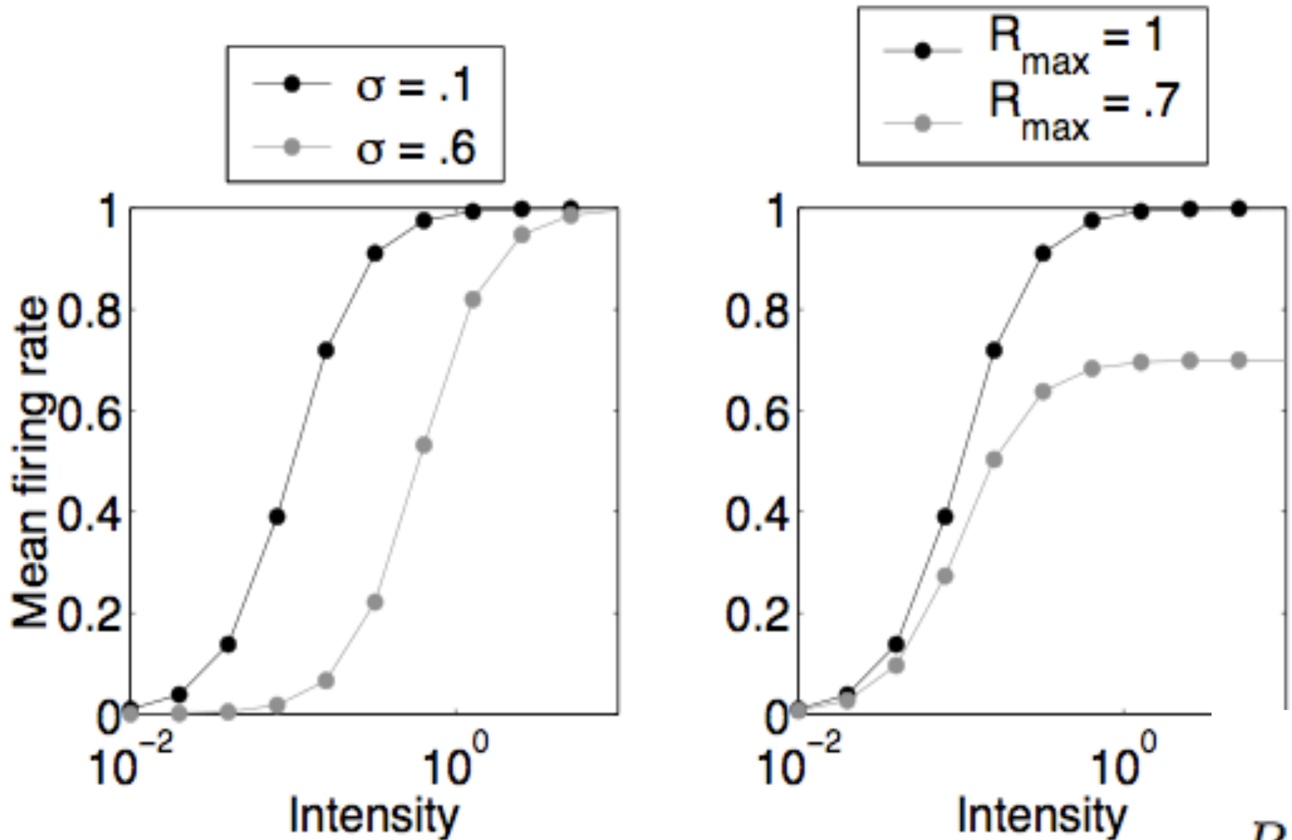
$$R = \frac{R_{max} K^2}{K^2 + \sigma^2}$$

K corresponds to illumination, contrast,

sound intensity, etc.

Divisive normalization model

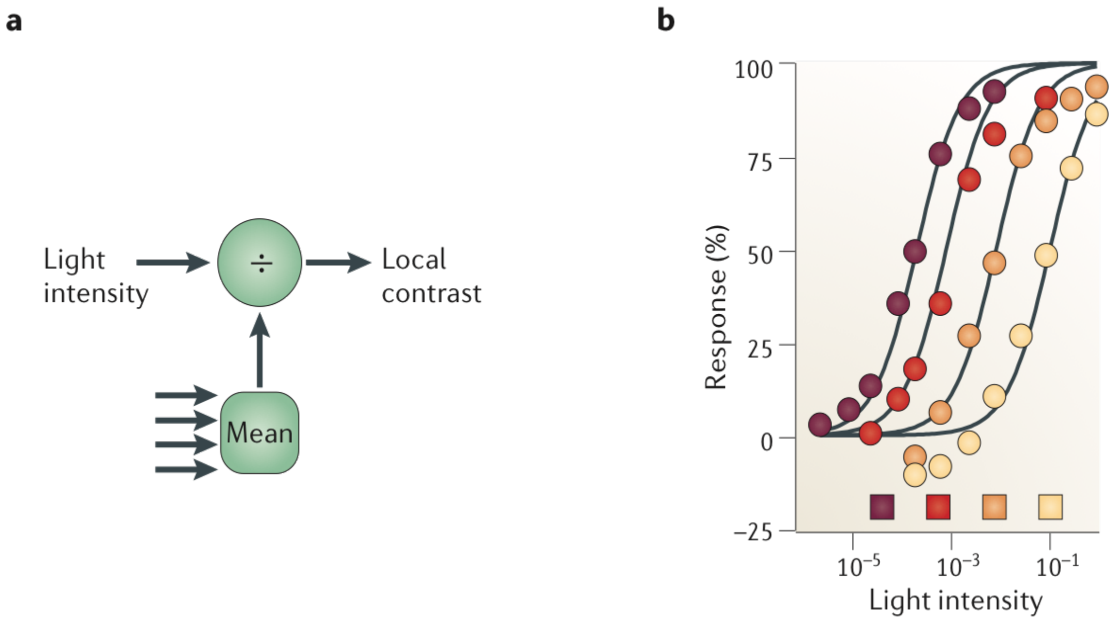
Simple version of descriptive model:



$$R = \frac{R_{max} K^2}{K^2 + \sigma^2}$$

Figure 1.3. Behavior of Naka-Rushton equation. Left, The Naka-Rushton equ tion for a constant $R_{max} = 1$ and variable σ. Note that higher values of σ shift the response curve to the right on a log axis. Right, The Naka-Rushton equation for a constant $\sigma = .1$ and variable R_{max}. Note that lower vales of R_{max} reduce the saturation level of the response curve.

Example: light adaptation



$$R = \frac{R_{max} K^2}{K^2 + \sigma^2}$$

Light adaptation to mean intensity in the retina
(in figure: turtle cone photoreceptor)
Carandini and Heeger, Nature Review Neuroscience, 2012

Example: primary visual cortex

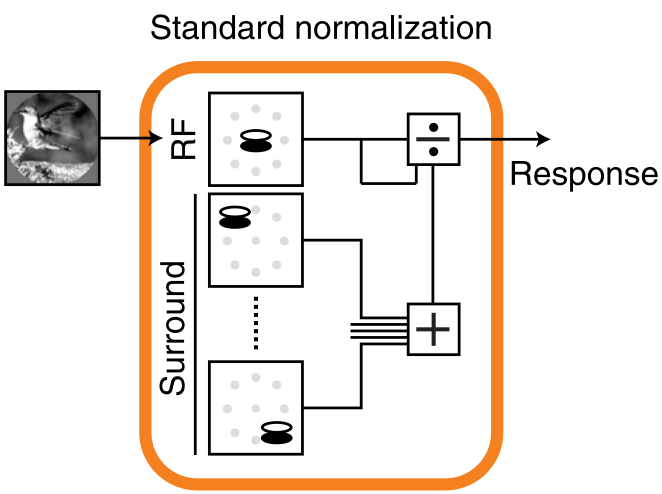
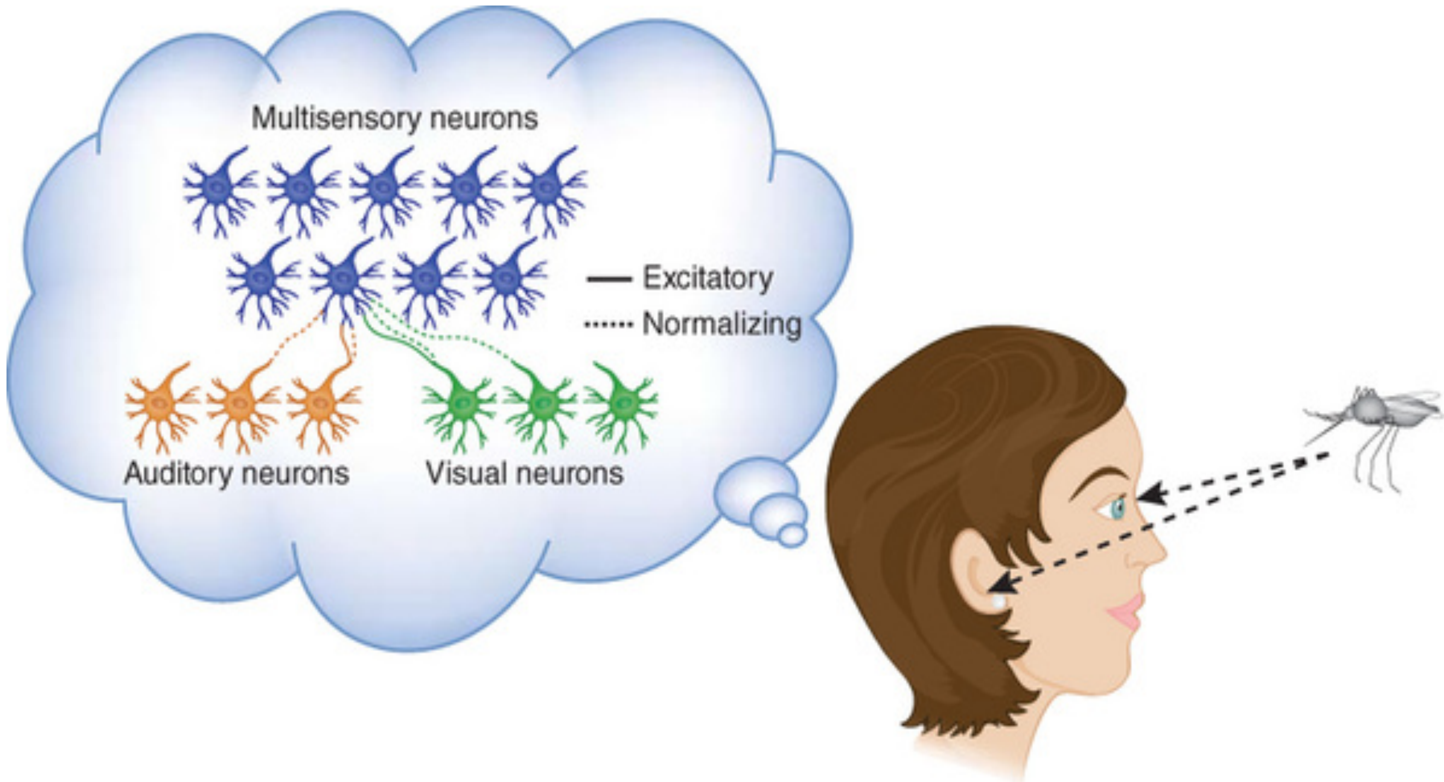


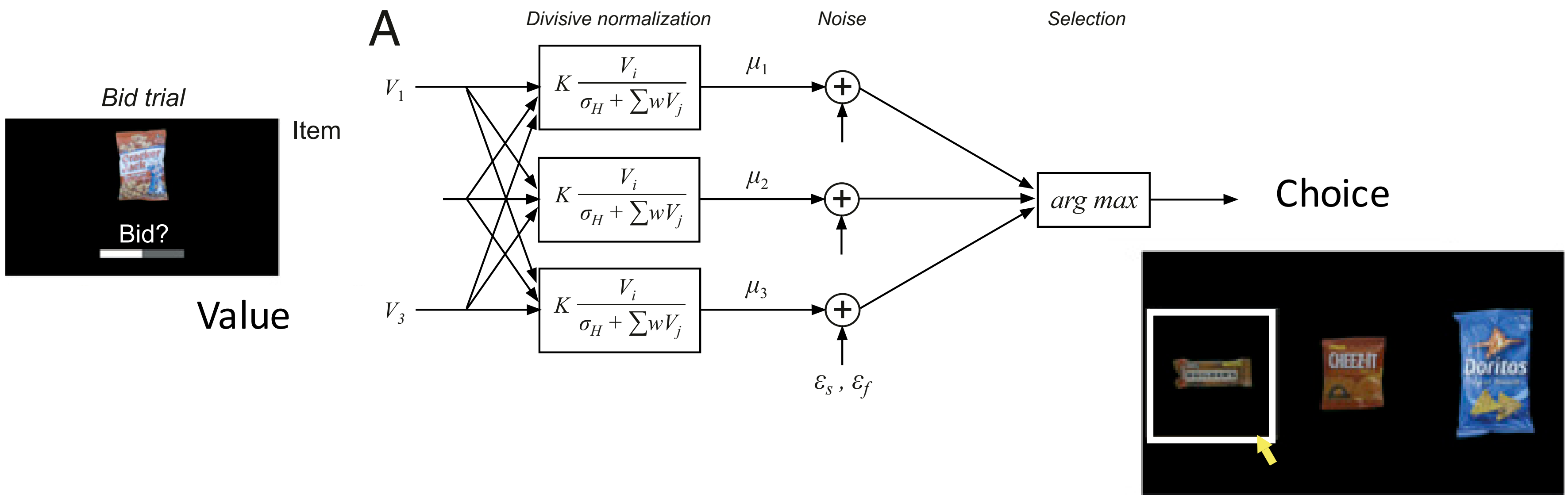
Figure from:
Cagli, Kohn, Schwartz, Nature Neuroscience 2015

Example: multisensory integration



Multisensory integration (eg, can explain change in neural responses with cue reliability)
Ohshiro, Angelaki, DeAngelis, Nature Neuroscience 2011
Figure from Churchland News and Views.

Example: decision making



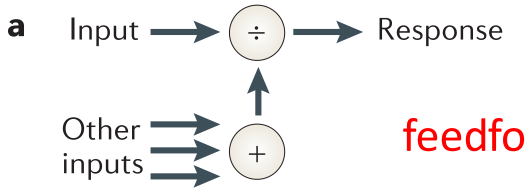
"Context-dependent choice behavior is of particular interest in economics because it violates one of the fundamental assumptions of many rational-choice theories, namely, that decisions reflect absolute valuations assigned to individual options" .. Distractors can reduce or even reverse choice"

Louie, Khaw and Glimcher, PNAS 2013: Normalization is a general neural mechanism for context-dependent decision making

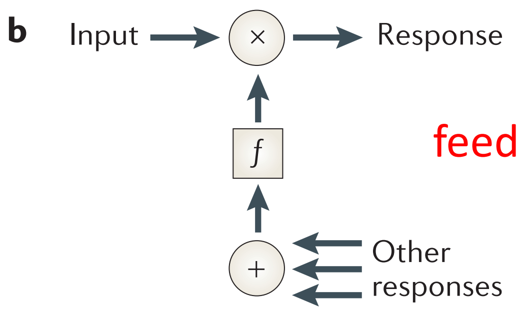
Alterations in Divisive Normalization?

- Rosenberg, Patterson, Angelaki, PNAS 2015: A computational perspective on autism
- Tibber MS, et al. (2013) Visual surround suppression in schizophrenia. Front Psychol 4:88.
- Betts LR, Taylor CP, Sekuler AB, Bennett PJ (2005) Aging reduces center-surround antagonism in visual motion processing. Neuron 45(3):361–366

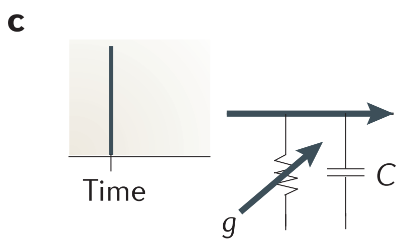
Mechanism of divisive normalization model



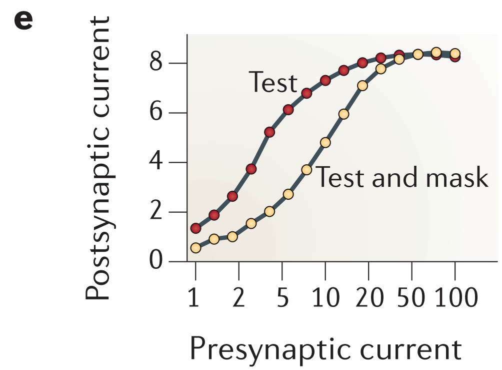
feedforward



feedback



Resistor-capacitor circuit



Synaptic depression

Carandini and Heeger, Nature Review Neuroscience, 2012