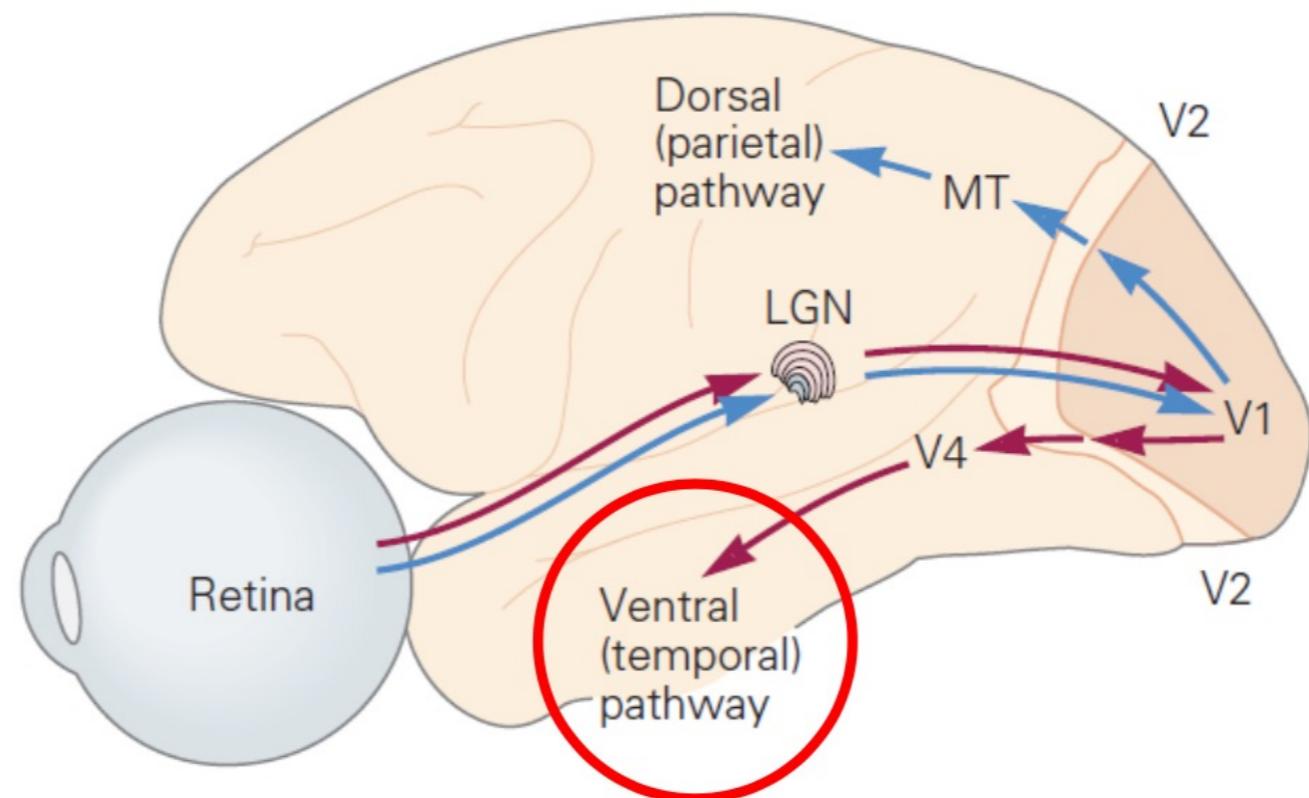
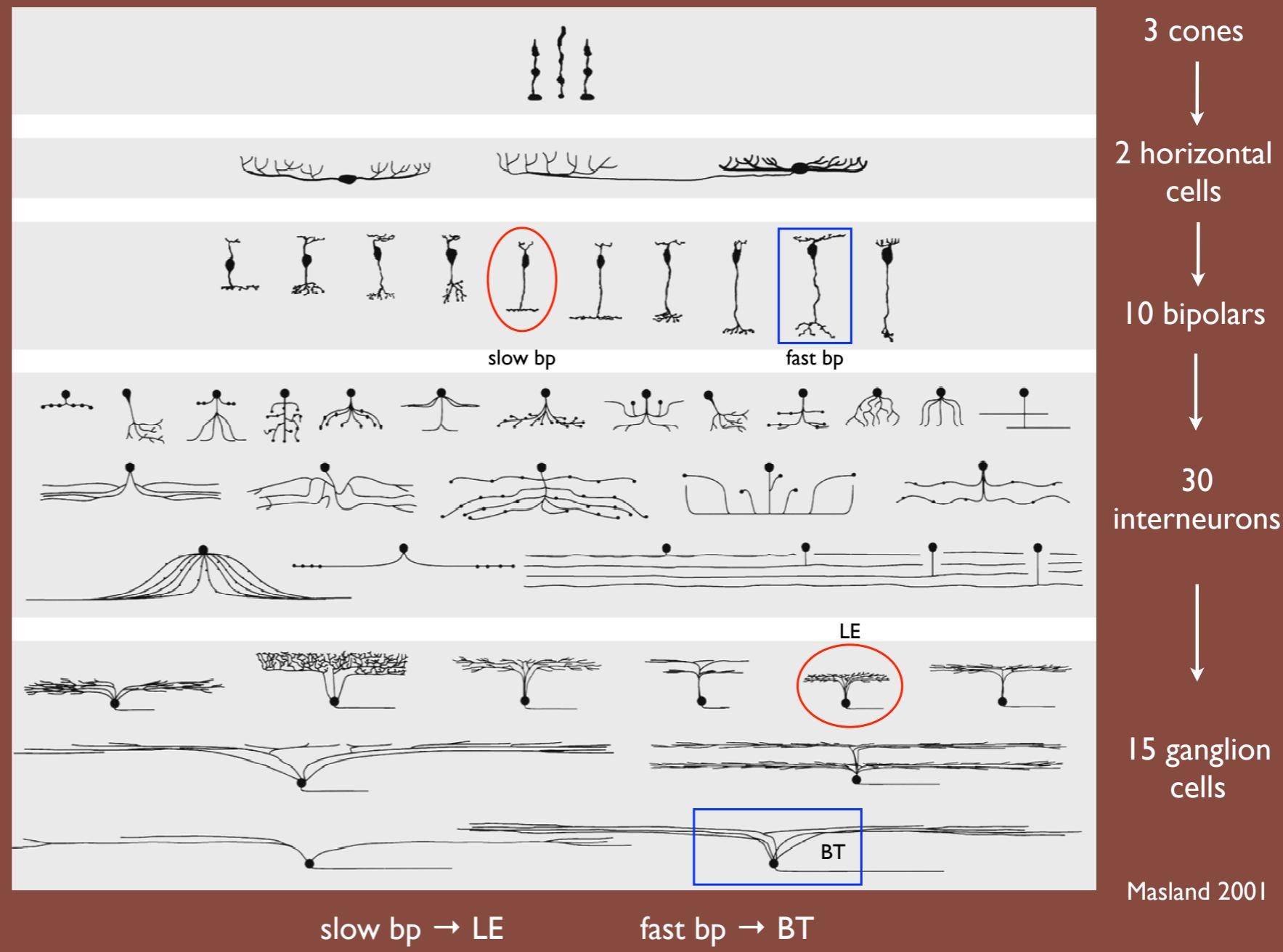
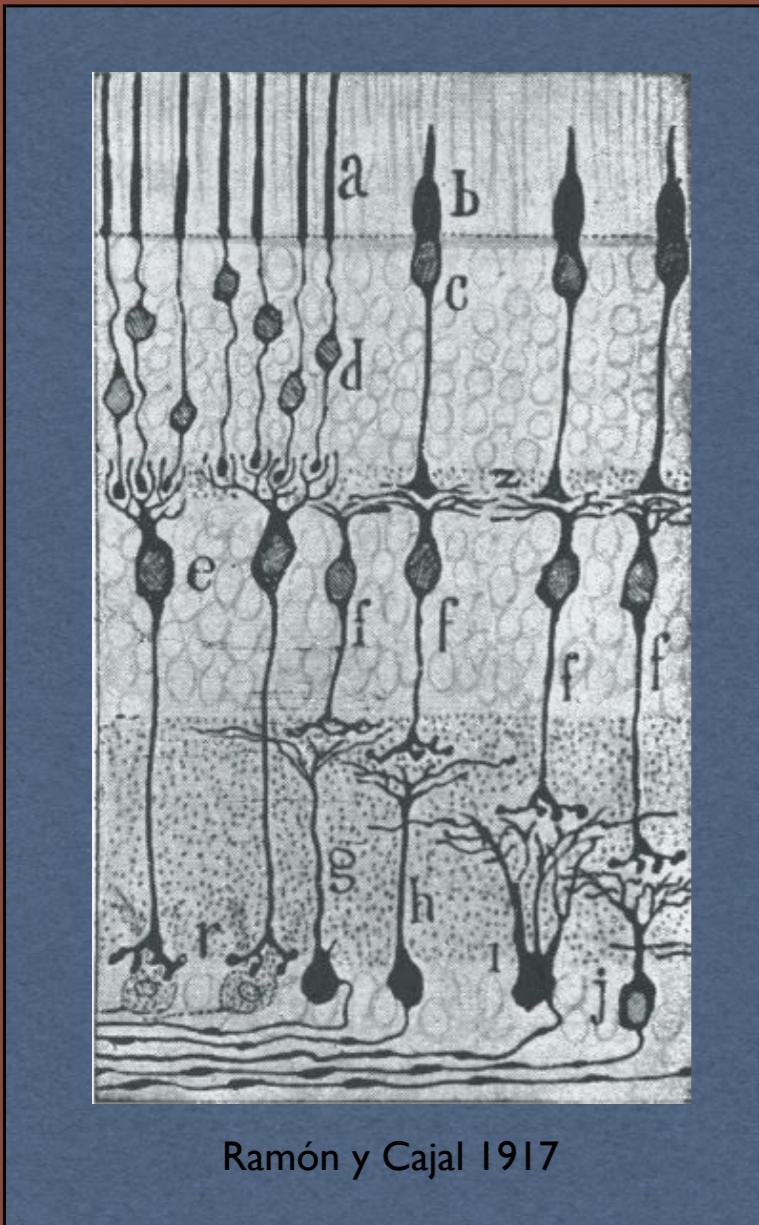


# **Sensory coding**

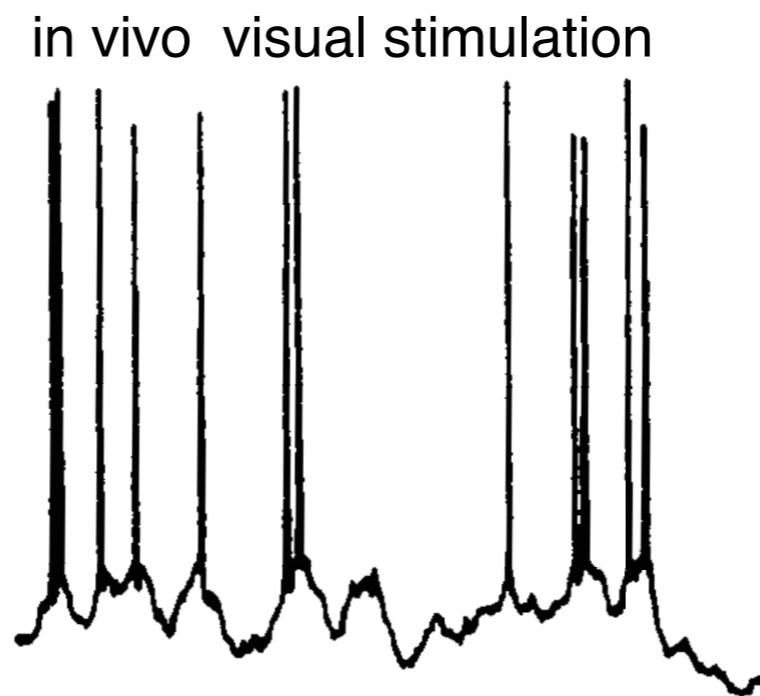
# Ventral visual pathway



# It really is mind-bogglingly diverse!

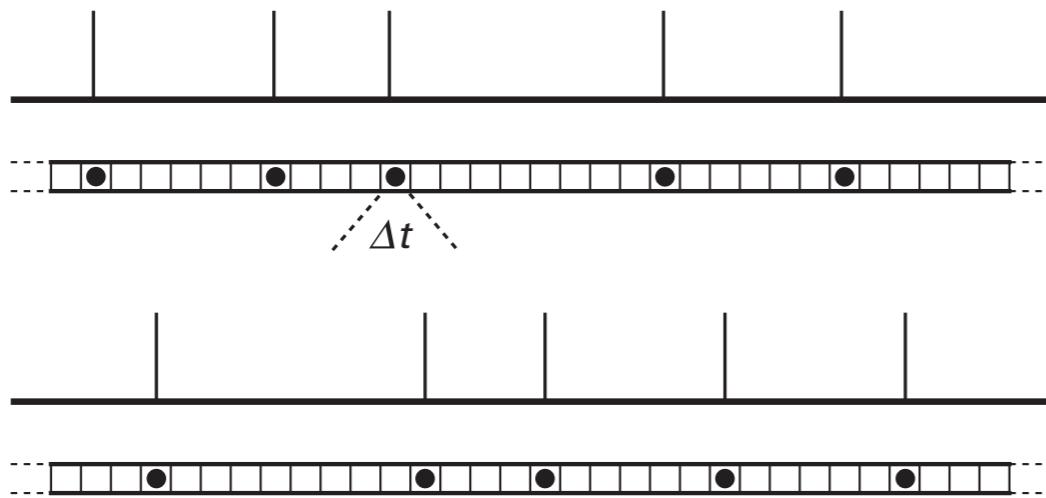


# Spike trains convey information



$$\rho(t) = \sum_i \delta(t - t_i)$$

# Information is entropy

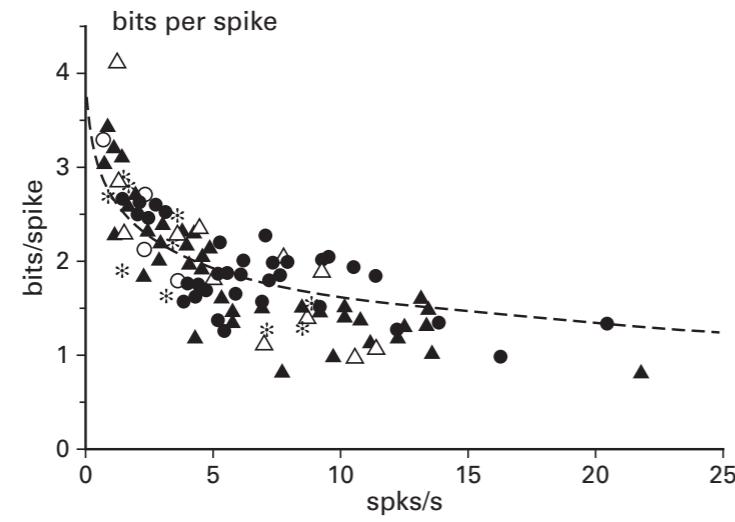
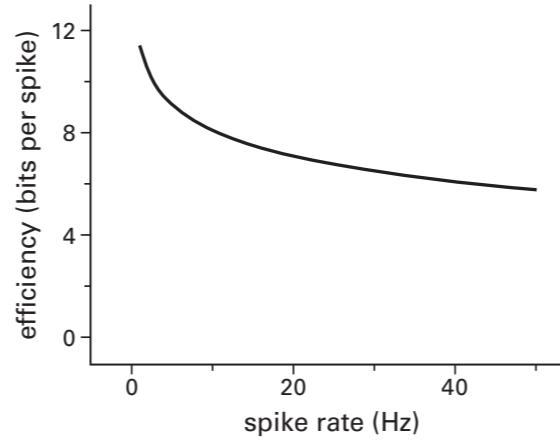
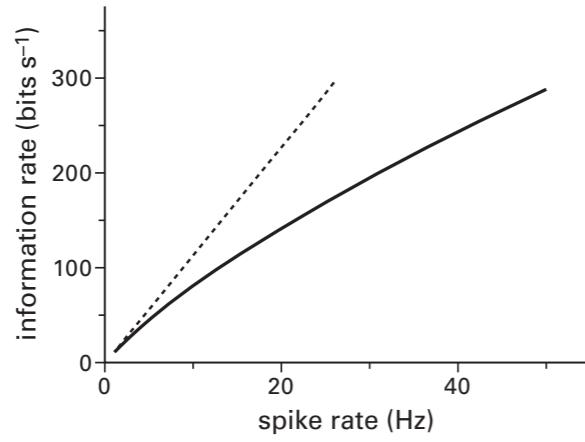


$$T = \frac{1 \text{ sec}}{\Delta t}$$

$$M = \frac{T!}{R!(T-R)!}$$

$$I = \log M$$

# Does every spike count?



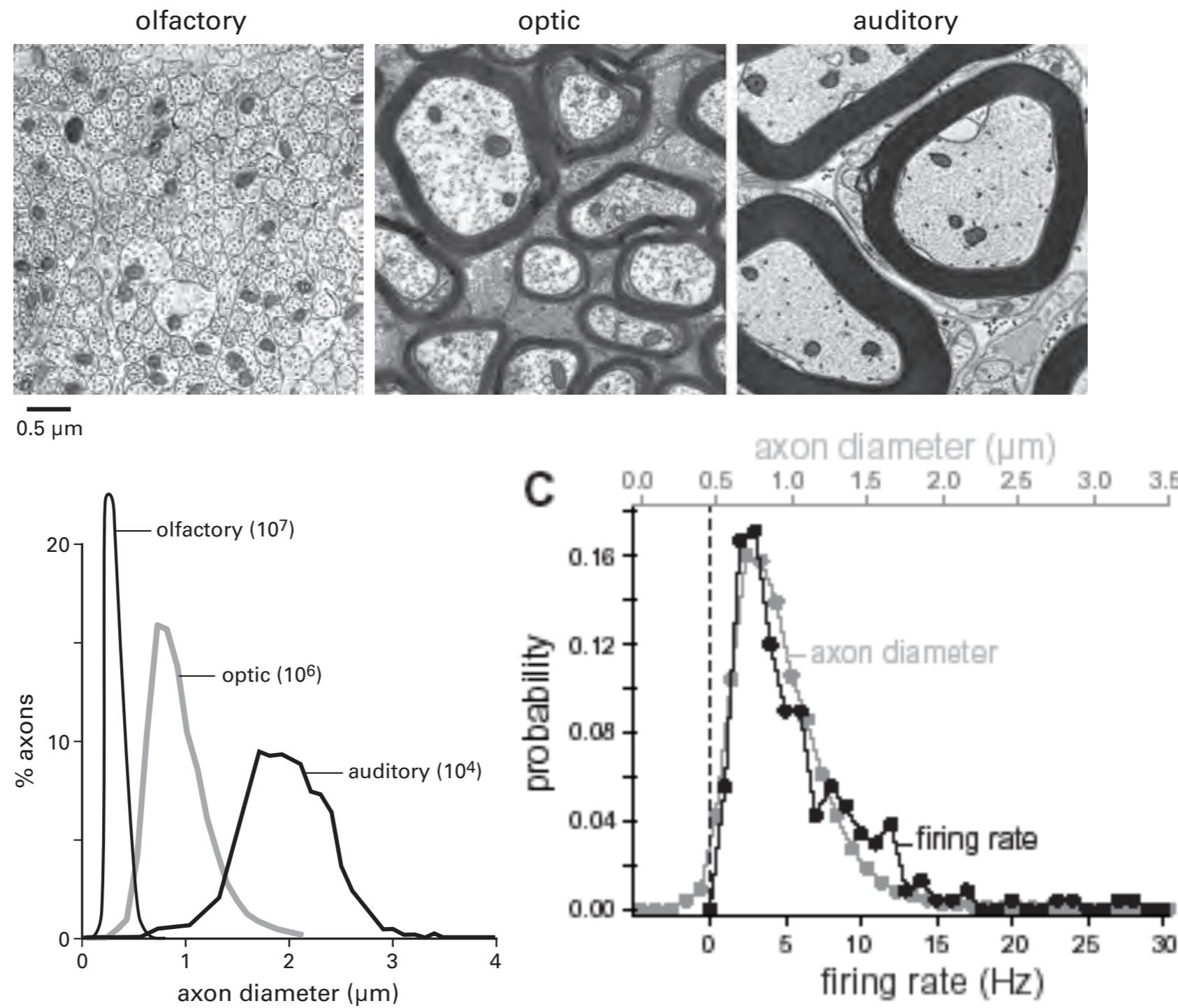
$$T = \frac{1 \text{ sec}}{\Delta t}$$

$$M = \frac{T!}{R!(T-R)!}$$

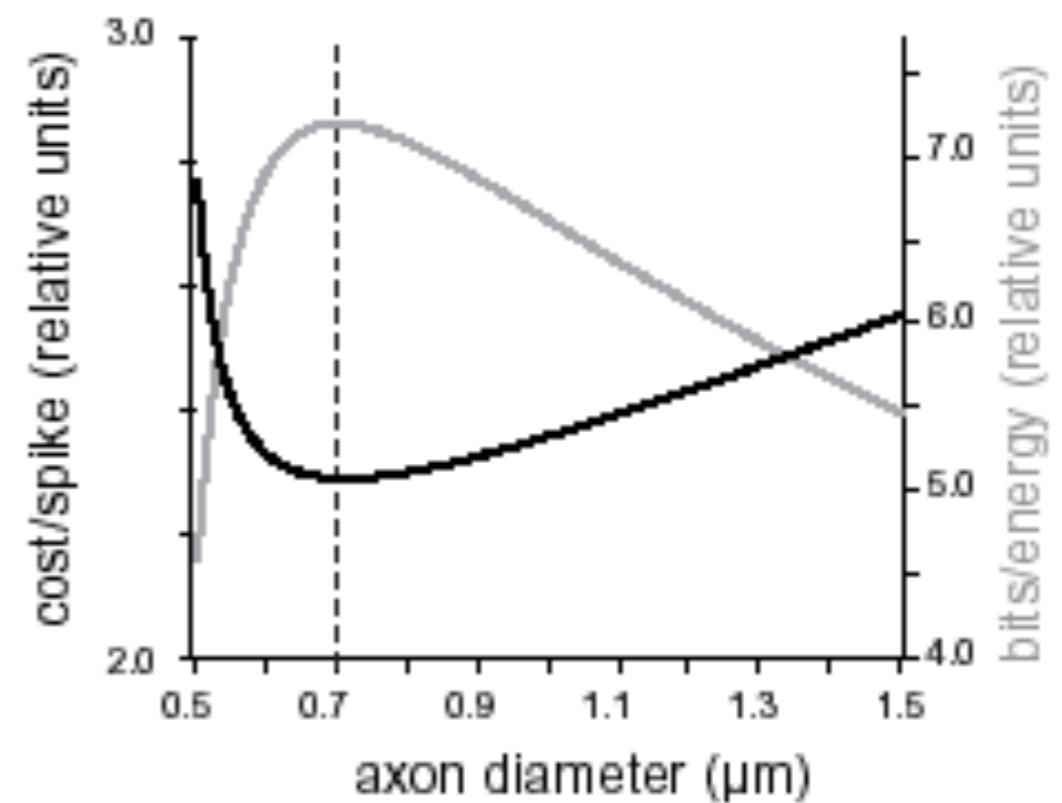
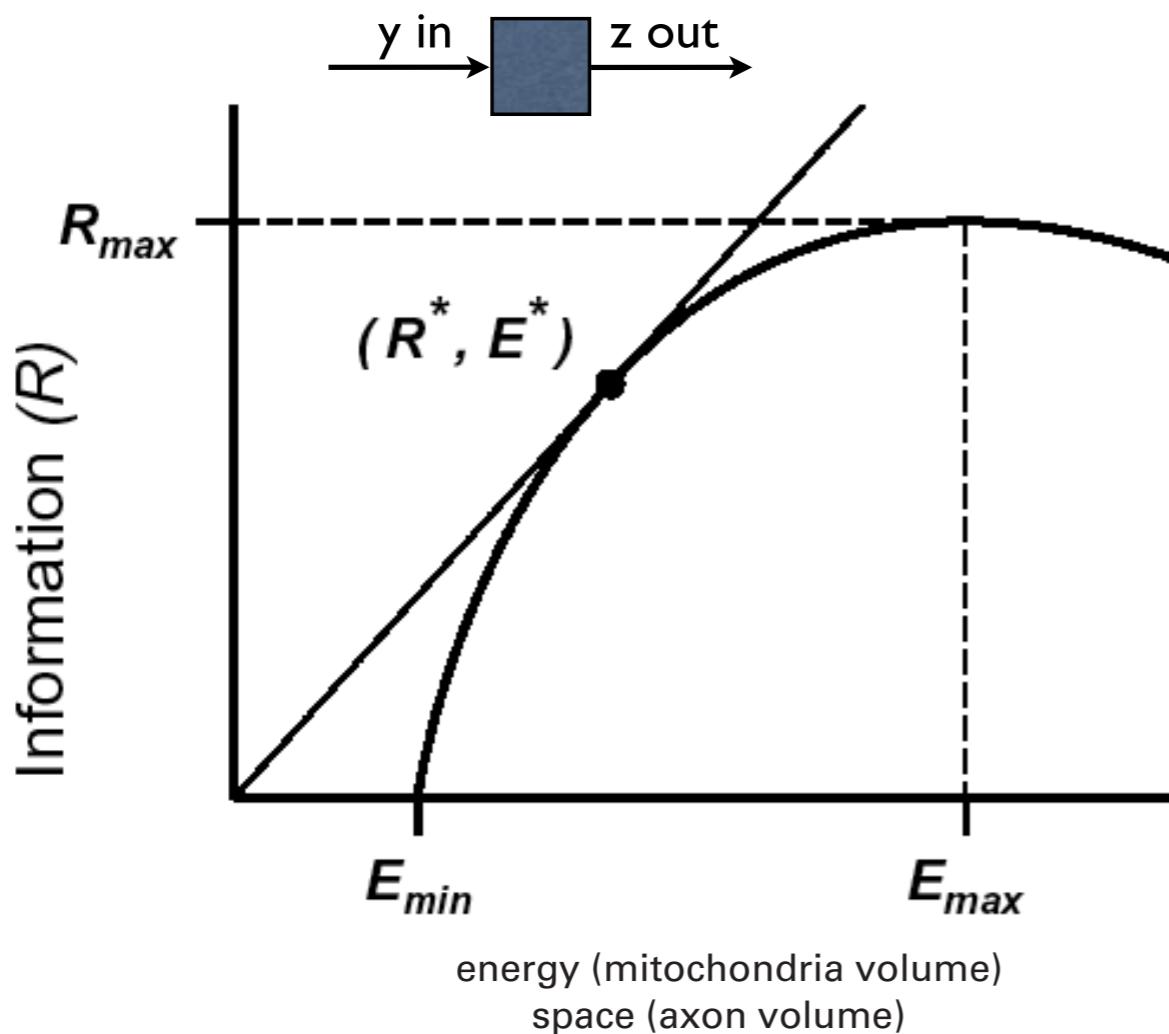
$$I = \log M$$

$$\frac{H}{R} \text{ bits per spike}$$

# Information transmission is costly

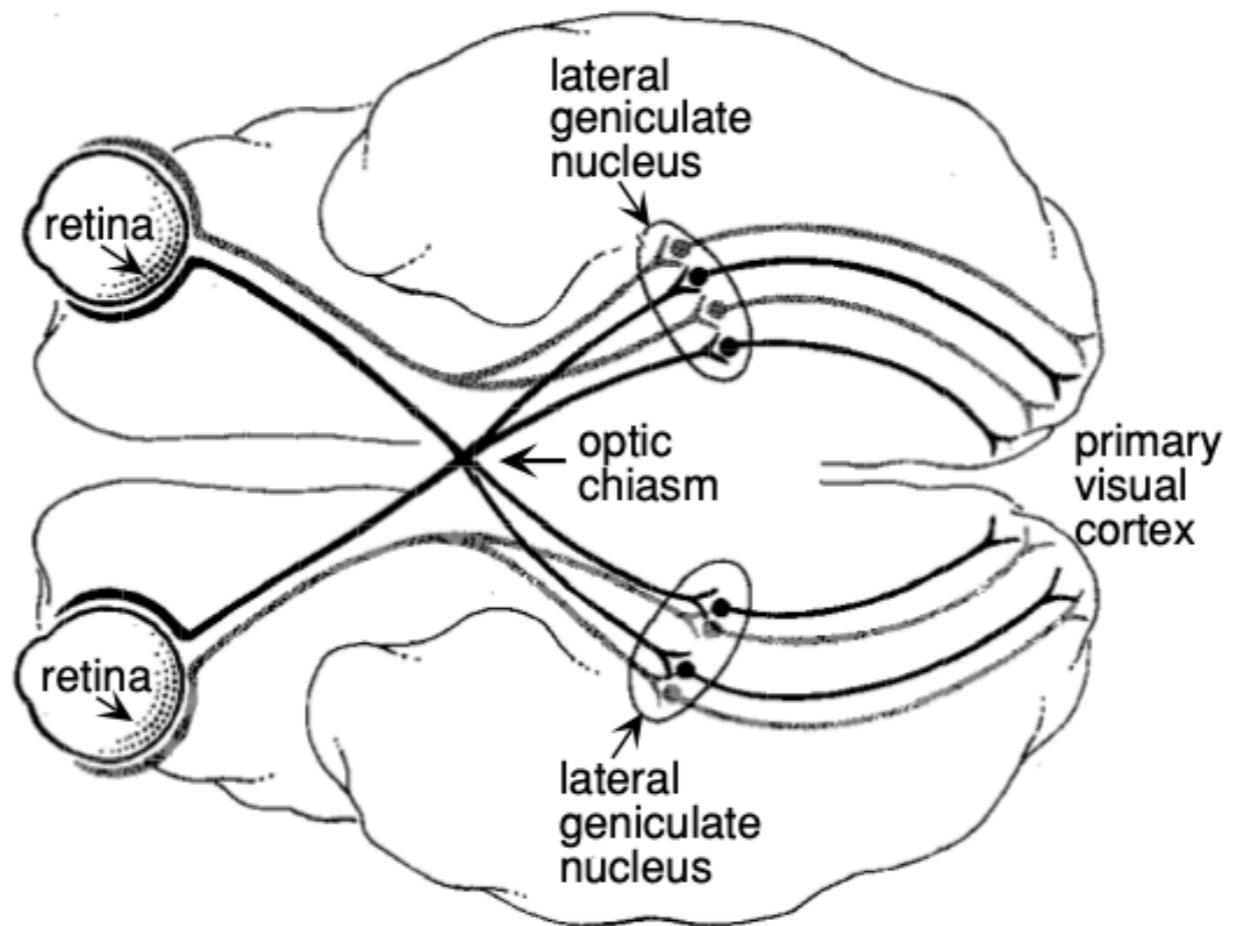


# Law of diminished returns

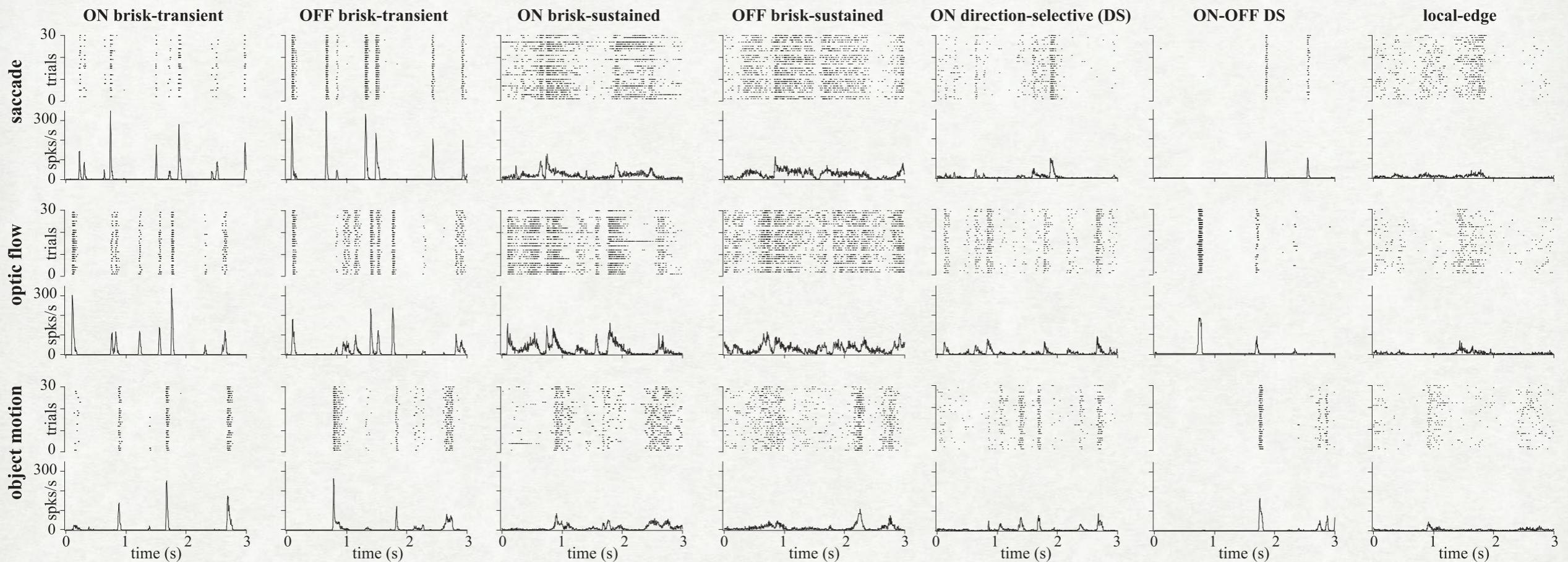


Heterogeneity & efficiency in the brain Vijay Balasubramanian 2015

Perge et al, 2018 How the optic nerve allocates space, energy capacity, and information

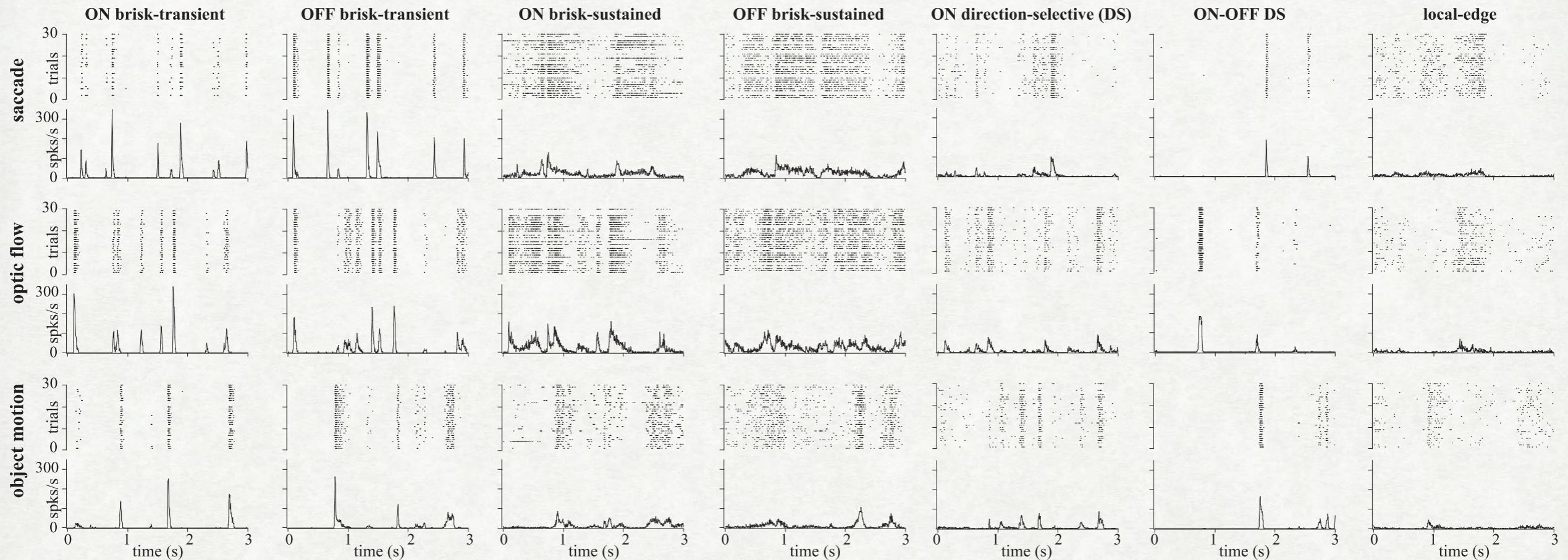


# Application: the distribution of information traffic in the optic nerve



Guinea Pig (Koch et al., 2005; Koch et al., 2006 “How much the eye tells the brain”)

# Application: the distribution of information traffic in the optic nerve

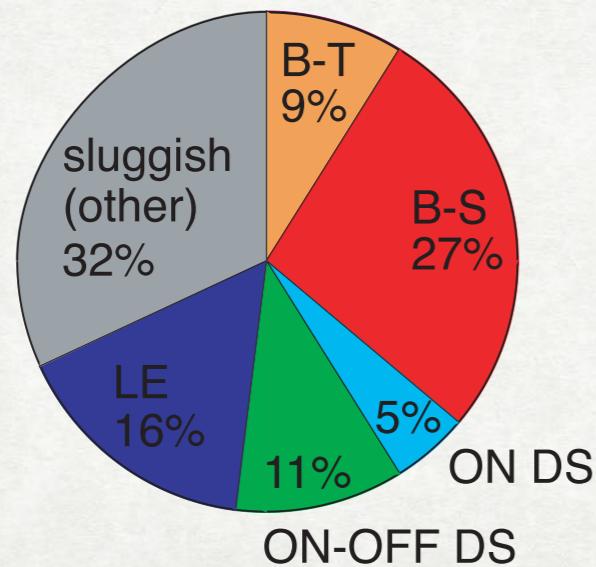


Guinea Pig (Koch et al., 2005; Koch et al., 2006 “How much the eye tells the brain”)

How and why is the total information traffic divided up between so many different types of channels?

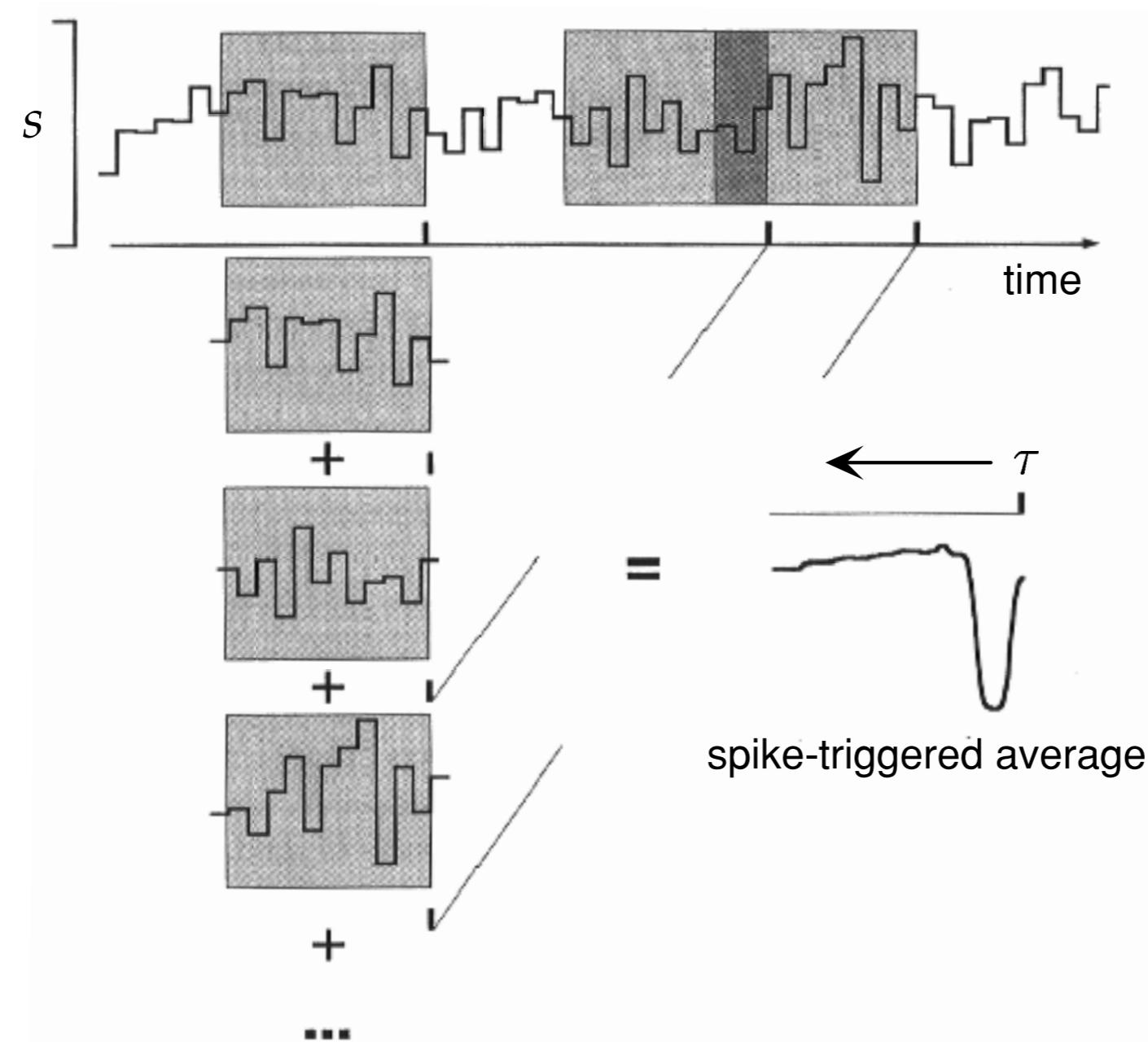
# Information traffic along the optic nerve

cell type	bits/s	bits/spike	cells	bits	spikes
Brisk-transient	13	1.9	6,000	78,000	41,000
Brisk-sustained	10	1.8	24,000	240,000	133,000
ON DS	6	2.2	7,000	42,000	19,000
ON-OFF DS	8	2.2	12,000	96,000	44,000
Local-edge	7	2.1	20,000	140,000	67,000
Sluggish (other)	9	2.2	31,000	279,000	127,000
<b>Total</b>			<b>100,000</b>	<b>875,000</b>	<b>431,000</b>



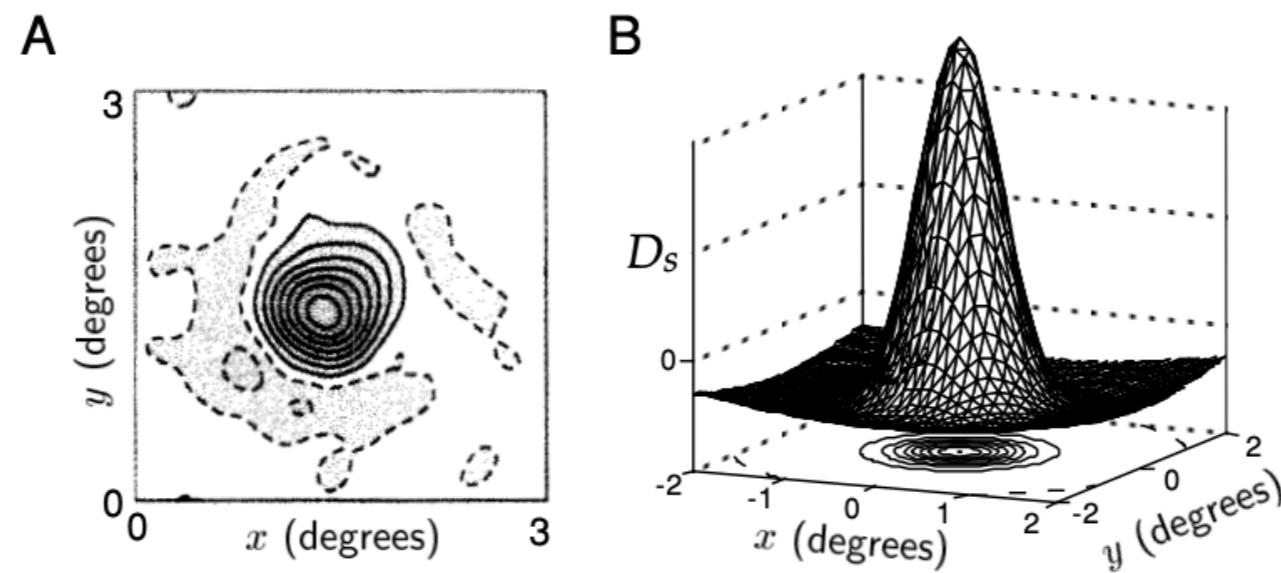
- Messages are transmitted asymmetrically - less studied sluggish cells account for most of the traffic
- Scaling up to human optic nerve (million axons) gives a traffic of order an Ethernet cable. This is the same order of magnitude as the amount of information in natural scenes according to Ruderman, 1994.

# What kind of information a spike train represent?



# **Linear response theory and optimal kernel**

# Receptive field of retinal ganglion cells



Differences between Gaussians

$$D_s(x, y) = \pm \left( \frac{1}{2\pi\sigma_{\text{cen}}^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_{\text{cen}}^2}\right) - \frac{B}{2\pi\sigma_{\text{sur}}^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_{\text{sur}}^2}\right) \right).$$

# Receptive field of simple cells in the visual cortex

$$D_s(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right) \cos(kx - \phi)$$

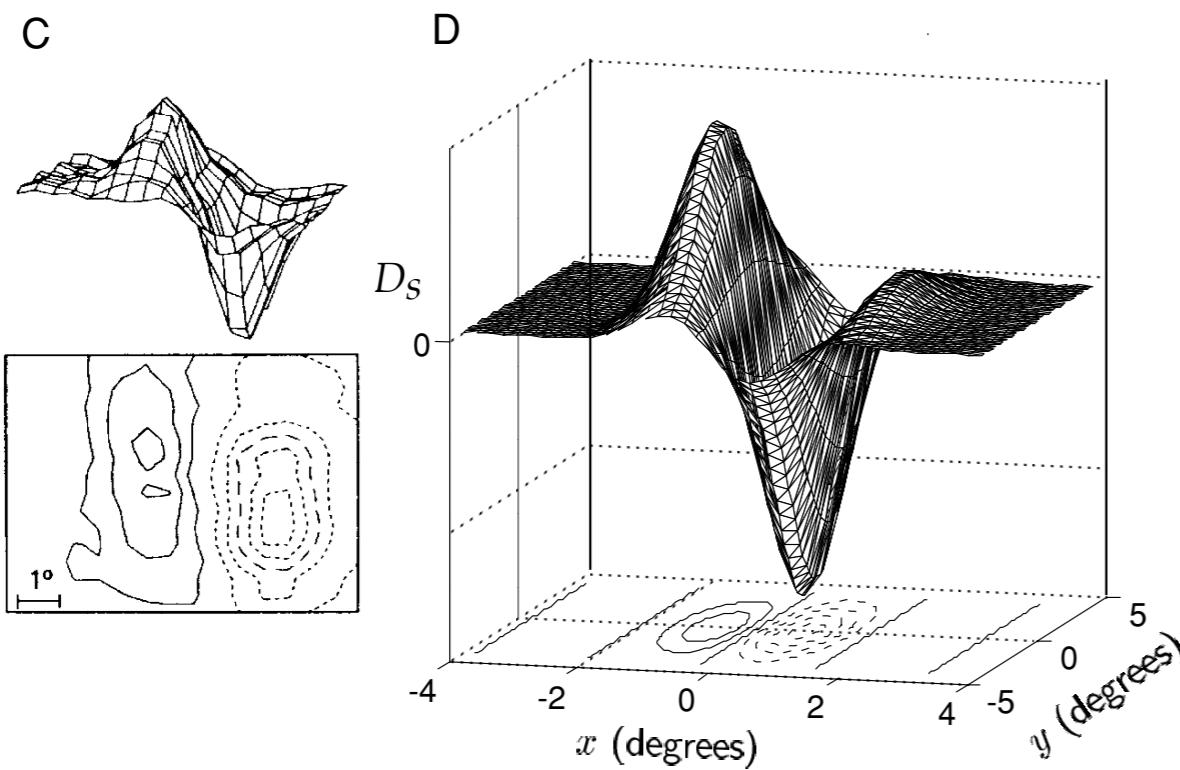


Figure 2.10 Spatial receptive field structure of simple cells. (A) and (C) Spatial structure of the receptive fields of two neurons in cat primary visual cortex determined by averaging stimuli between 50 ms and 100 ms prior to an action potential. The upper plots are three-dimensional representations, with the horizontal dimensions acting as the  $x$ - $y$  plane and the vertical dimension indicating the magnitude and sign of  $D_s(x, y)$ . The lower contour plots represent the  $x$ - $y$  plane. Regions with solid contour curves are ON areas where  $D_s(x, y) > 0$ , and regions with dashed contours are OFF areas where  $D_s(x, y) < 0$ . (B) and (D) Gabor functions (equation 2.27) with  $\sigma_x = 1^\circ$ ,  $\sigma_y = 2^\circ$ ,  $1/k = 0.56^\circ$ , and  $\phi = 1 - \pi/2$  (B) or  $\phi = 1 - \pi$  (D), chosen to match the receptive fields in A and C. (A and C adapted from Jones and Palmer, 1987a.)

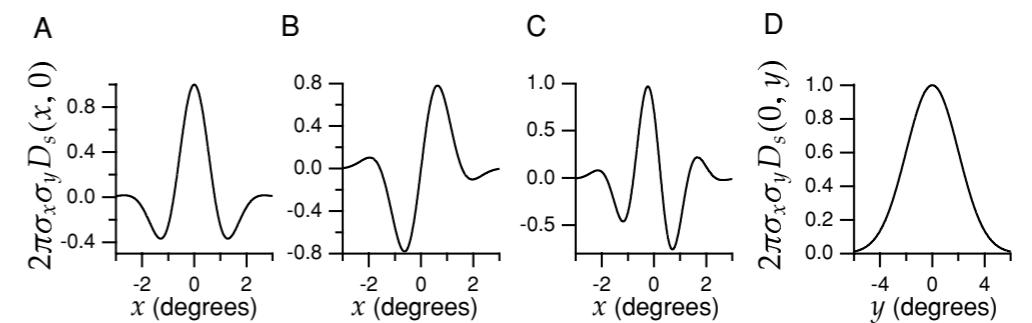
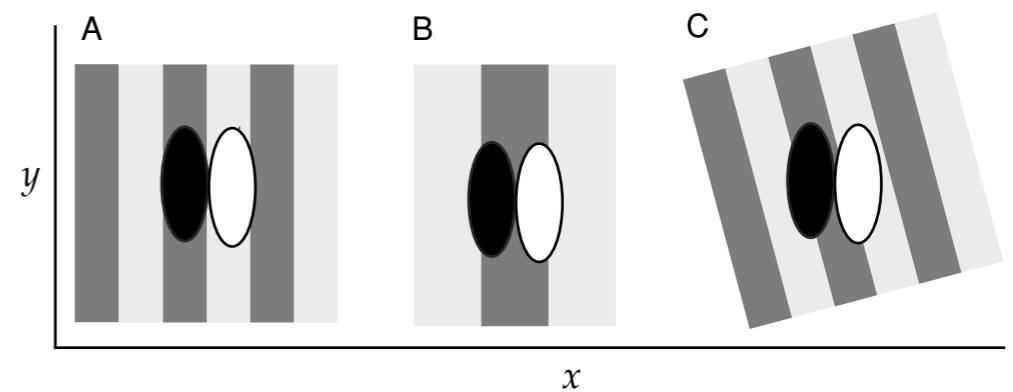


Figure 2.12 Gabor functions of the form given by equation 2.27. For convenience we plot the dimensionless function  $2\pi\sigma_x\sigma_y D_s$ . (A) A Gabor function with  $\sigma_x = 1^\circ$ ,  $1/k = 0.5^\circ$ , and  $\phi = 0$  plotted as a function of  $x$  for  $y = 0$ . This function is symmetric about  $x = 0$ . (B) A Gabor function with  $\sigma_x = 1^\circ$ ,  $1/k = 0.5^\circ$ , and  $\phi = \pi/2$  plotted as a function of  $x$  for  $y = 0$ . This function is antisymmetric about  $x = 0$  and corresponds to using a sine instead of a cosine function in equation 2.27. (C) A Gabor function with  $\sigma_x = 1^\circ$ ,  $1/k = 0.33^\circ$ , and  $\phi = \pi/4$  plotted as a function of  $x$  for  $y = 0$ . This function has no particular symmetry properties with respect to  $x = 0$ . (D) The Gabor function of equation 2.27 with  $\sigma_y = 2^\circ$  plotted as a function of  $y$  for  $x = 0$ . This function is simply a Gaussian.

# Temporal receptive field

$$D_t(\tau) = \alpha \exp(-\alpha\tau) \left[ \frac{(\alpha\tau)^5}{5!} - \frac{(\alpha\tau)^7}{7!} \right]$$

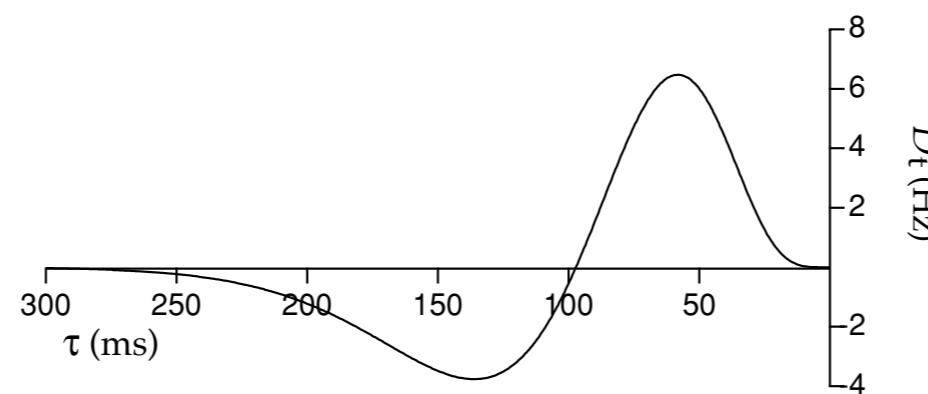
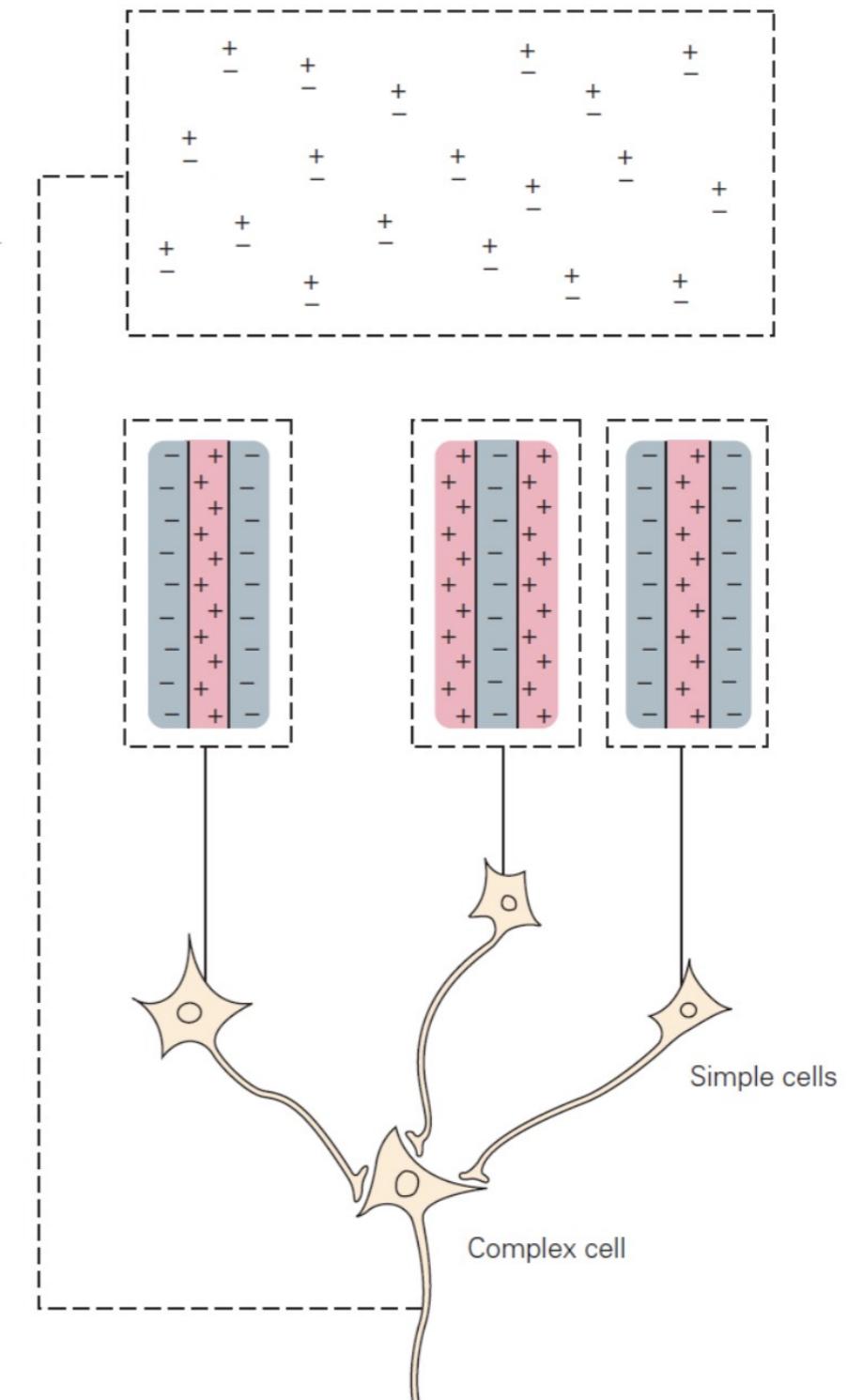


Figure 2.14 Temporal structure of a receptive field. The function  $D_t(\tau)$  of equation 2.29 with  $\alpha = 1/(15 \text{ ms})$ .

# 复杂细胞



Torsten N. Wiesel   David H. Hubel  
1981 Nobel prize



# Hubel-Wiesel model of orientation selectivity

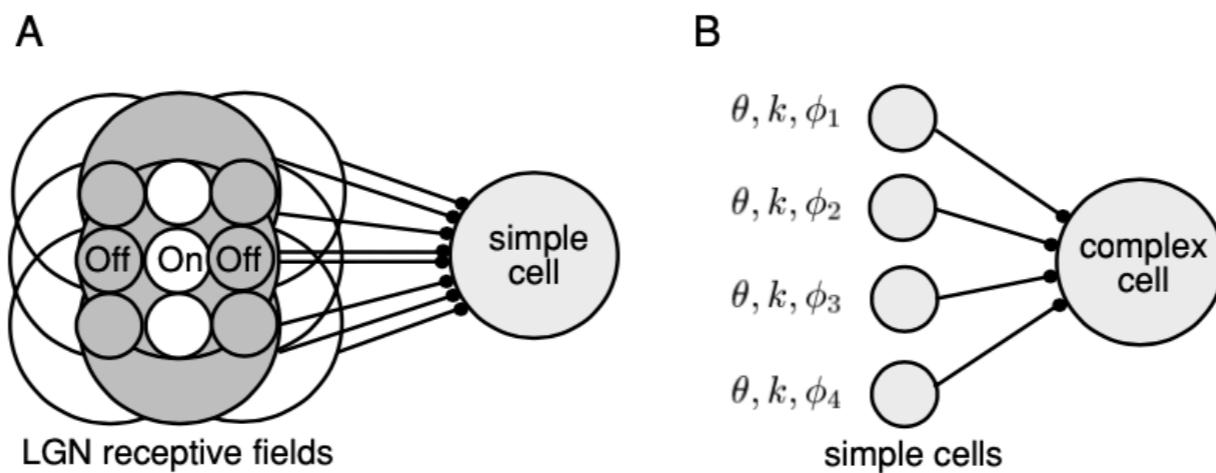
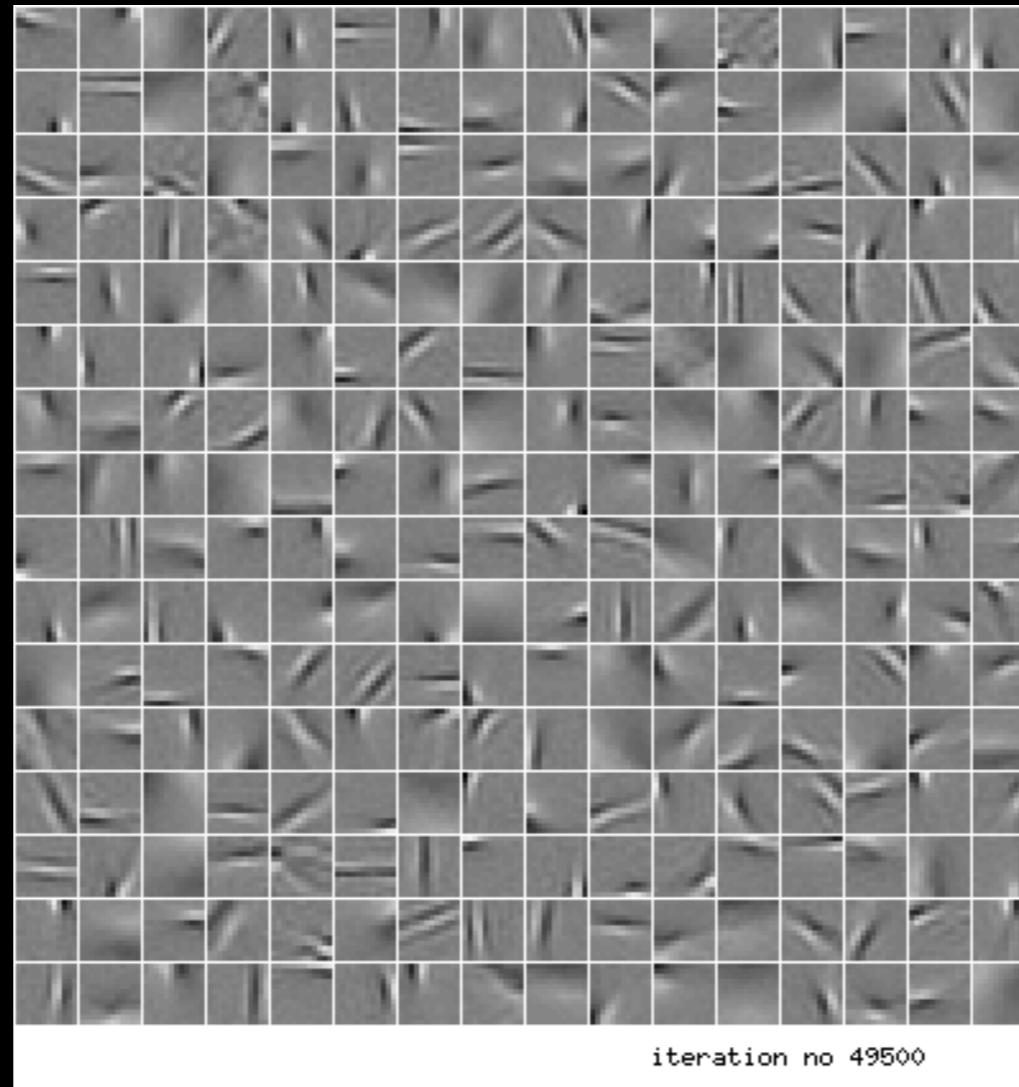


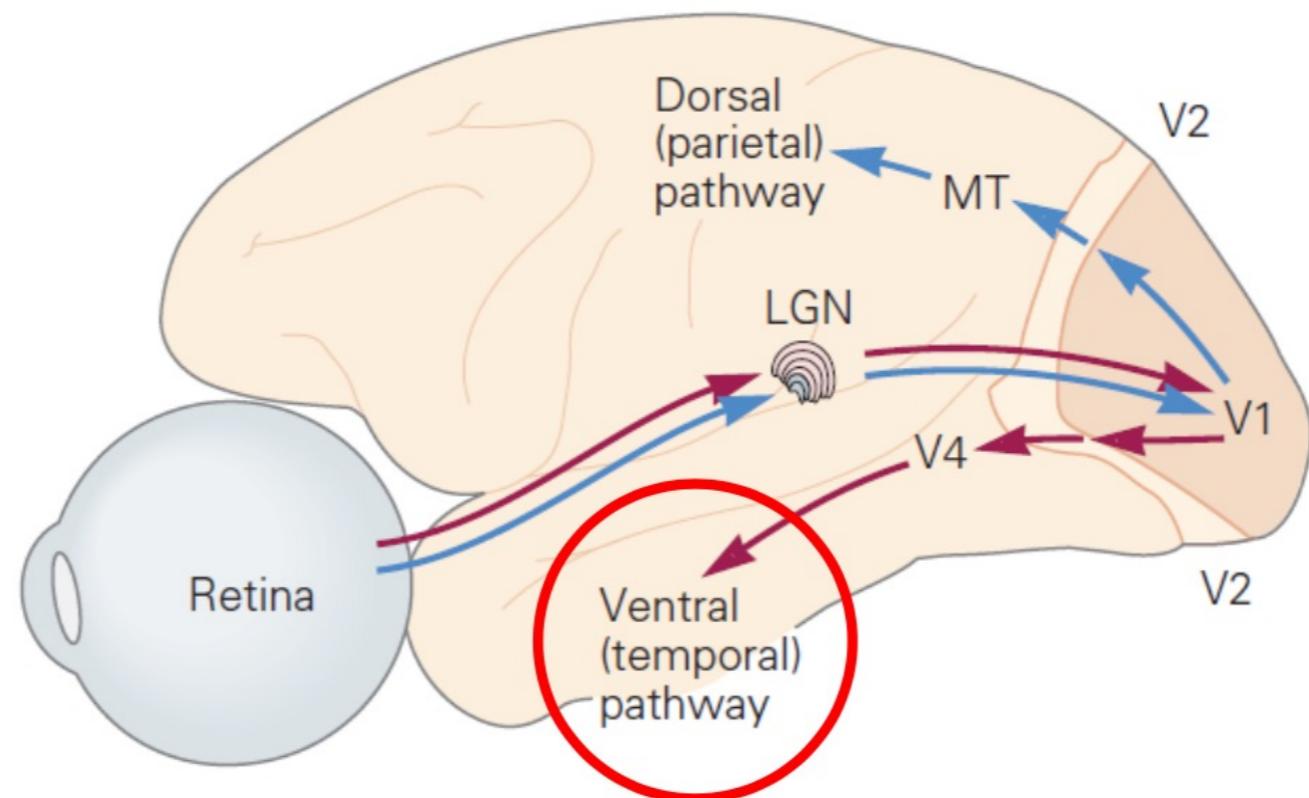
Figure 2.27 (A) The Hubel-Wiesel model of orientation selectivity. The spatial arrangement of the receptive fields of nine LGN neurons are shown, with a row of three ON-center fields flanked on either side by rows of three OFF-center fields. White areas denote ON fields and gray areas, OFF fields. In the model, the converging LGN inputs are summed by the simple cell. This arrangement produces a receptive field oriented in the vertical direction. (B) The Hubel-Wiesel model of a complex cell. Inputs from a number of simple cells with similar orientation and spatial frequency preferences ( $\theta$  and  $k$ ), but different spatial phase preferences ( $\phi_1, \phi_2, \phi_3$ , and  $\phi_4$ ), converge on a complex cell and are summed. This produces a complex cell output that is selective for orientation and spatial frequency, but not for spatial phase. The figure shows four simple cells converging on a complex cell, but additional simple cells can be included to give a more complete coverage of spatial phase.

Convolutional neural net (CNN) was inspired by information processing in human visual pathway

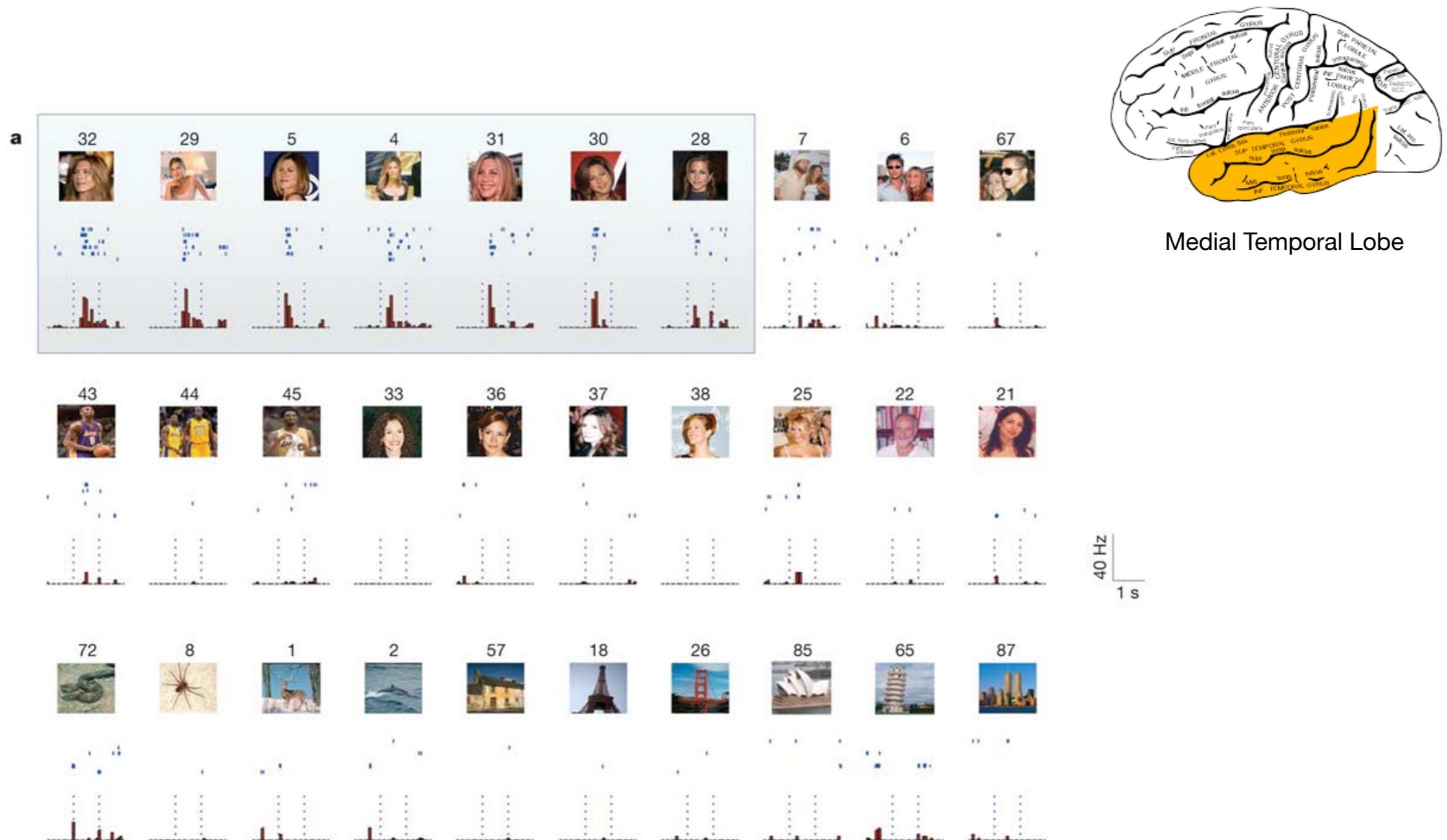


In the first layer of CNN, neurons can generate receptive field just like simple cells in visual cortex

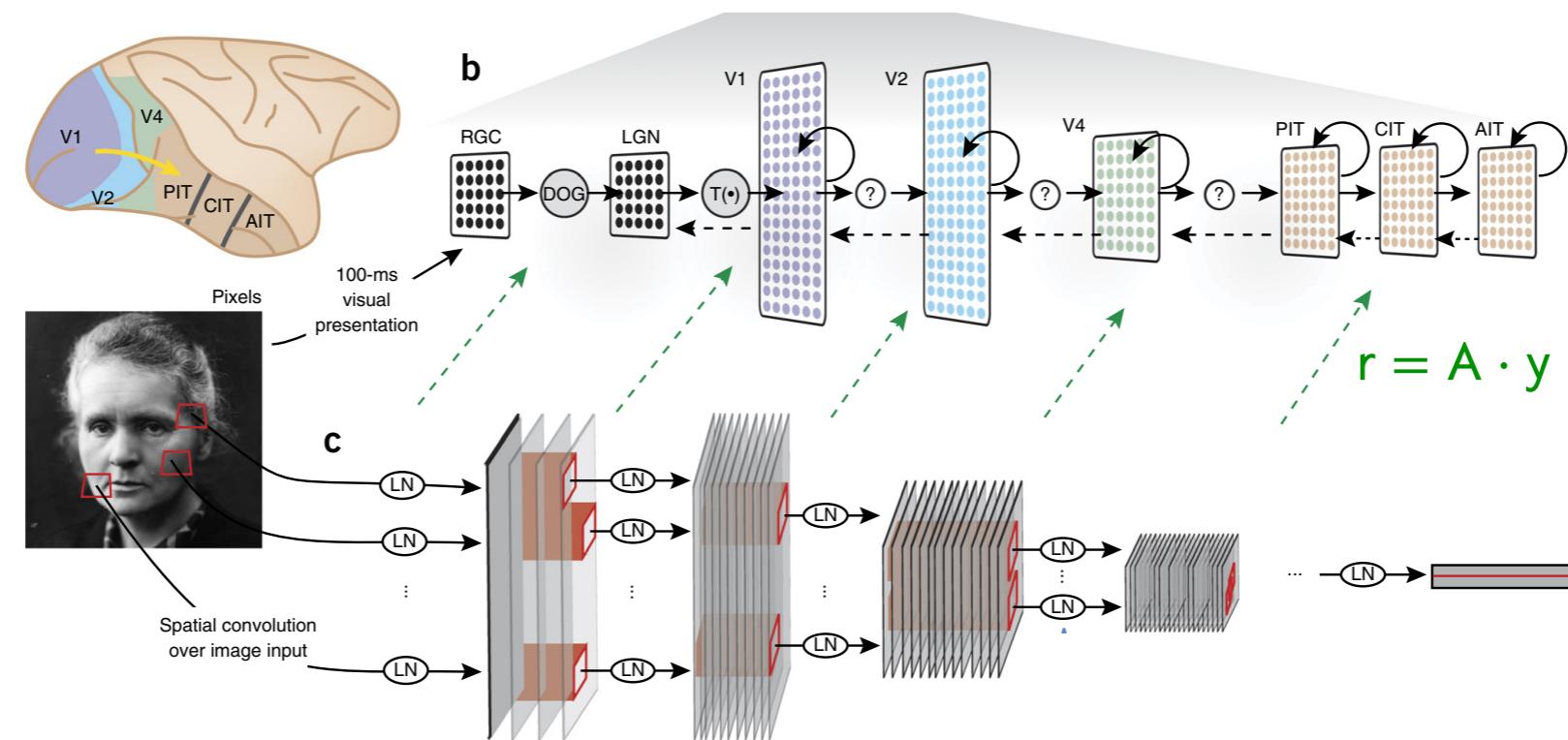
# Ventral visual pathway



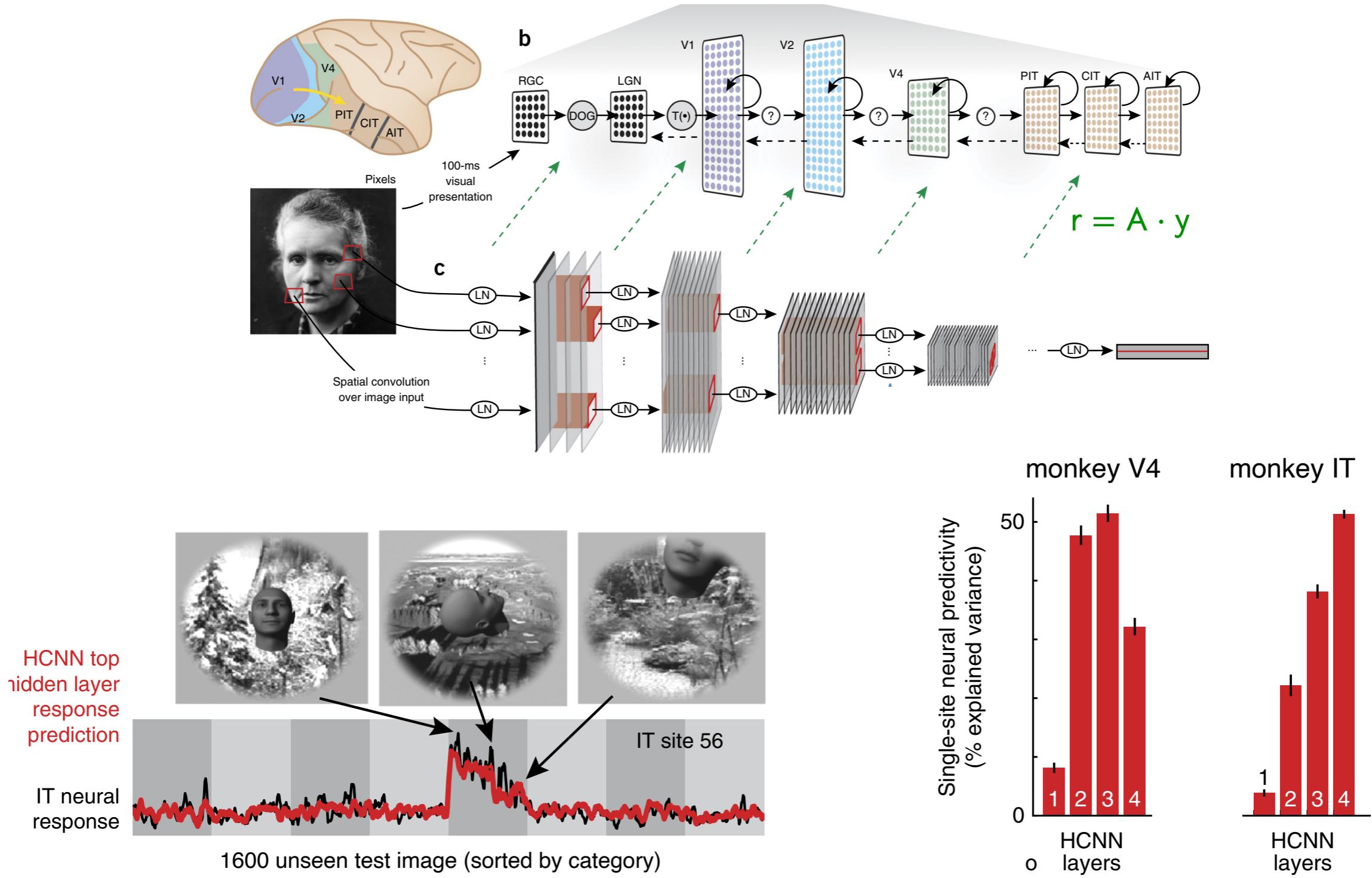
# The Jennifer Aniston cell



# Deep neural network and biological neurons

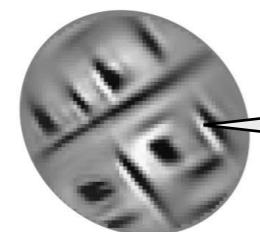


# Deep neural network and biological neurons

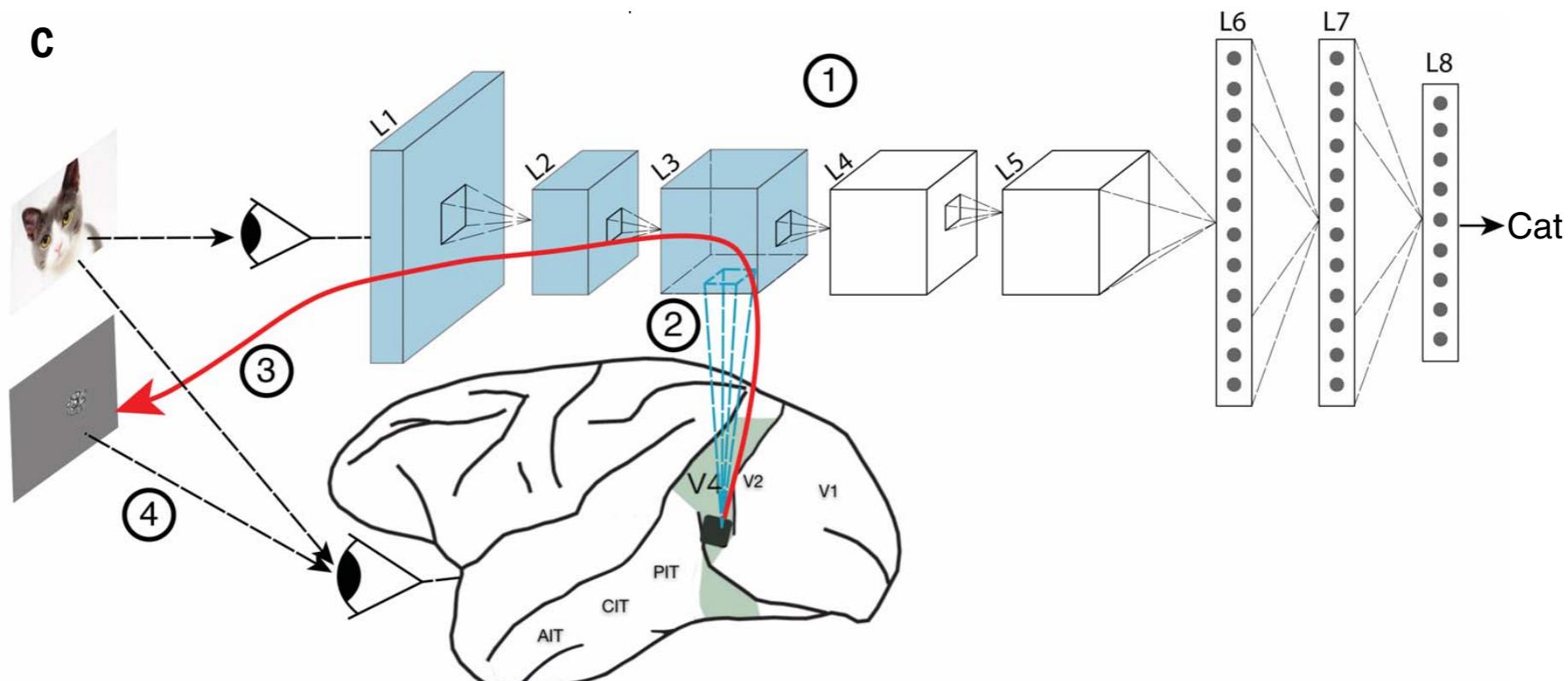


# Test deep neural networks on V4 neurons

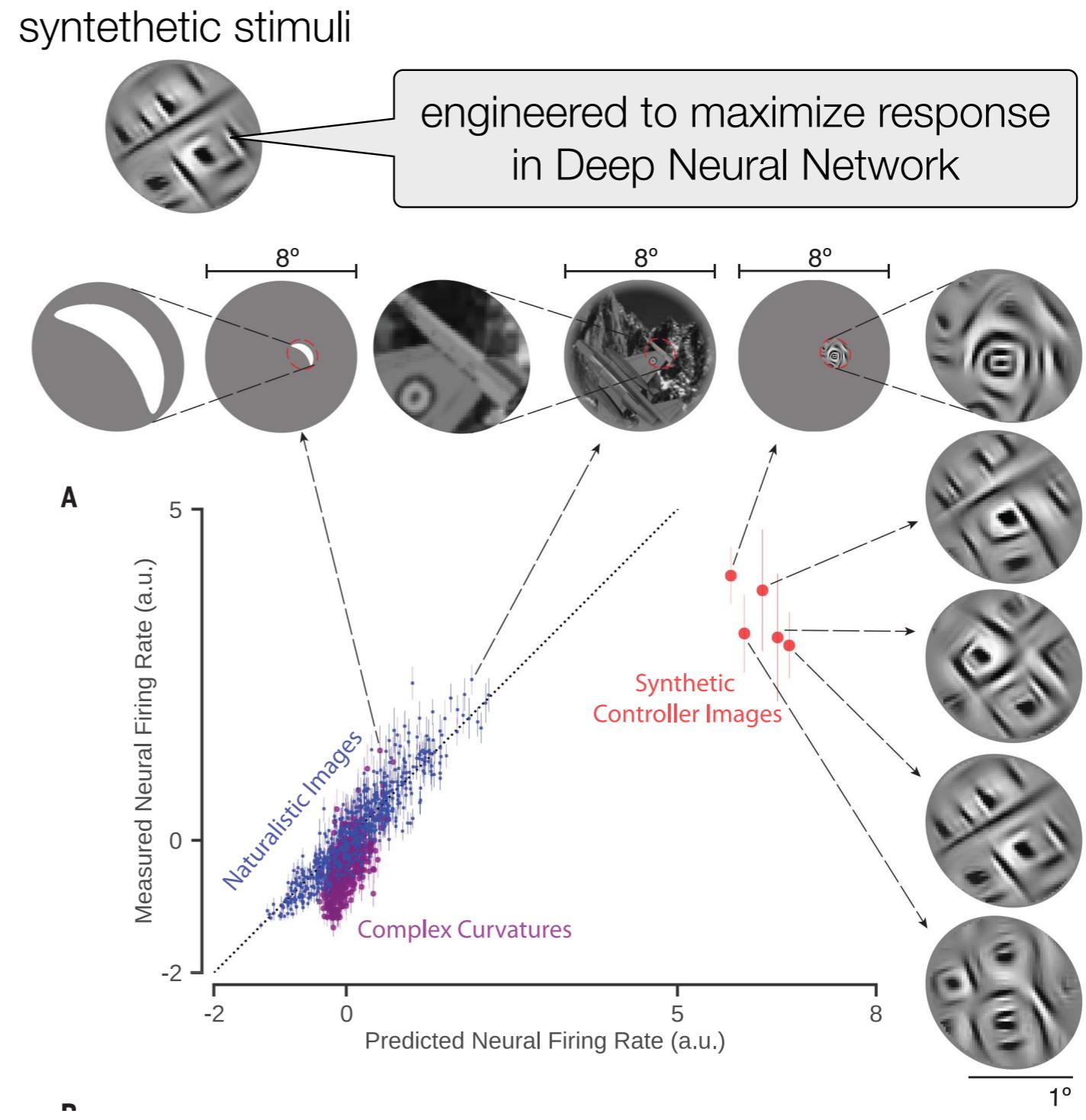
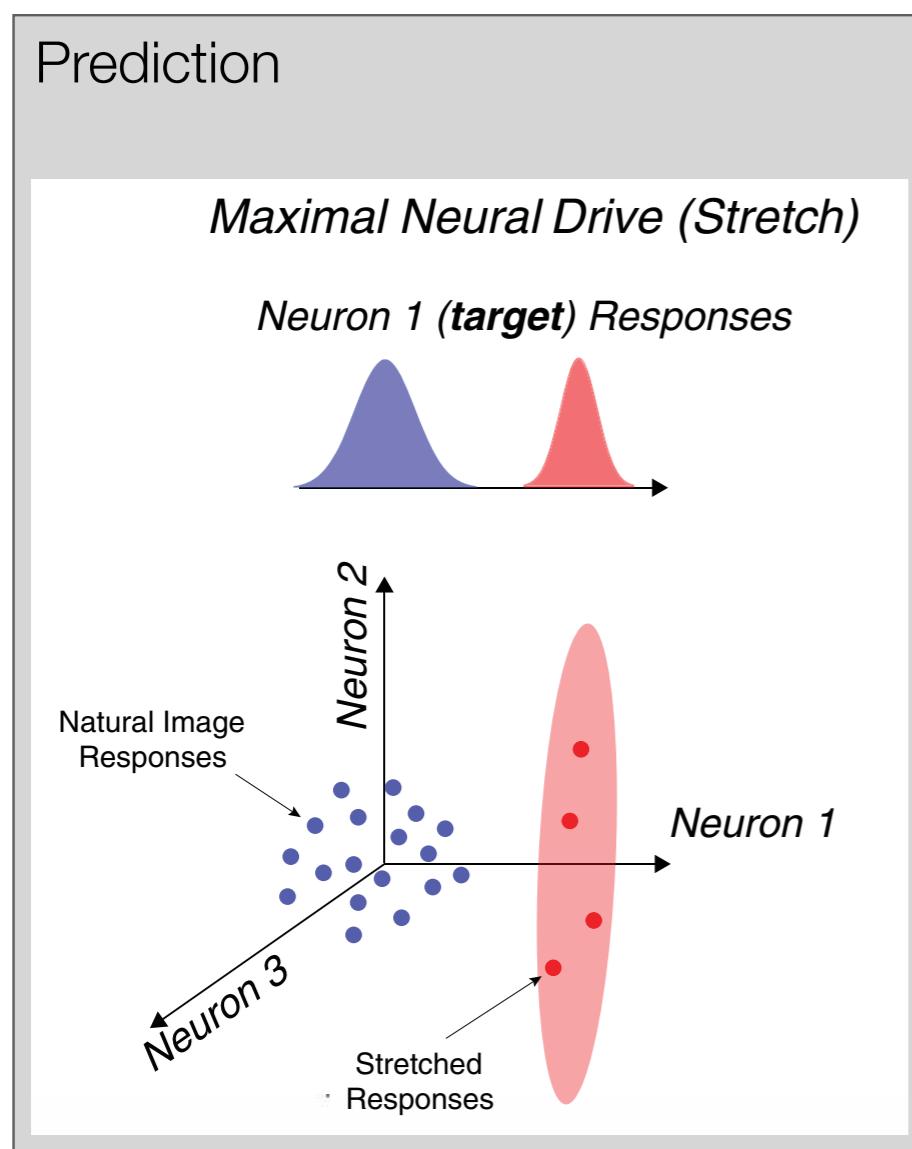
synthetic stimuli



engineered to maximize response  
in Deep Neural Network

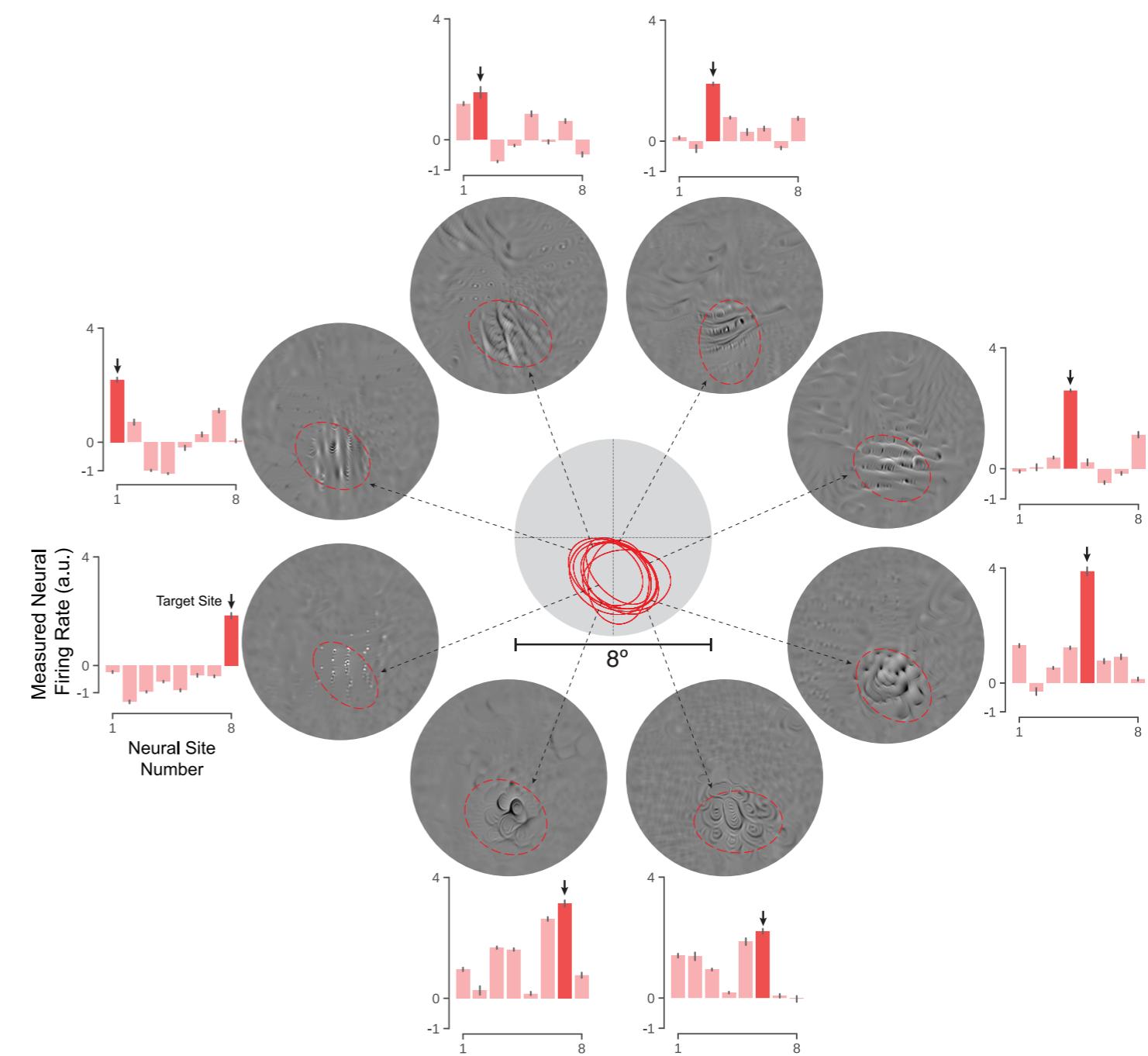
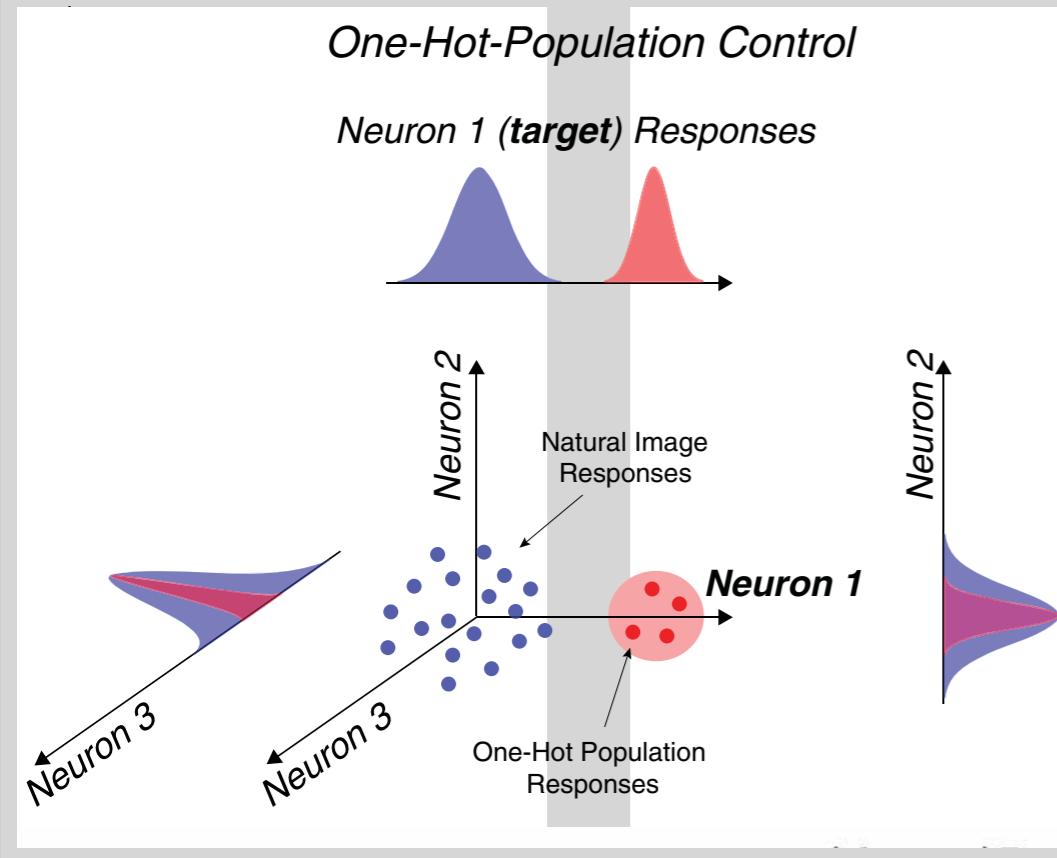


# Test deep neural networks on V4 neurons

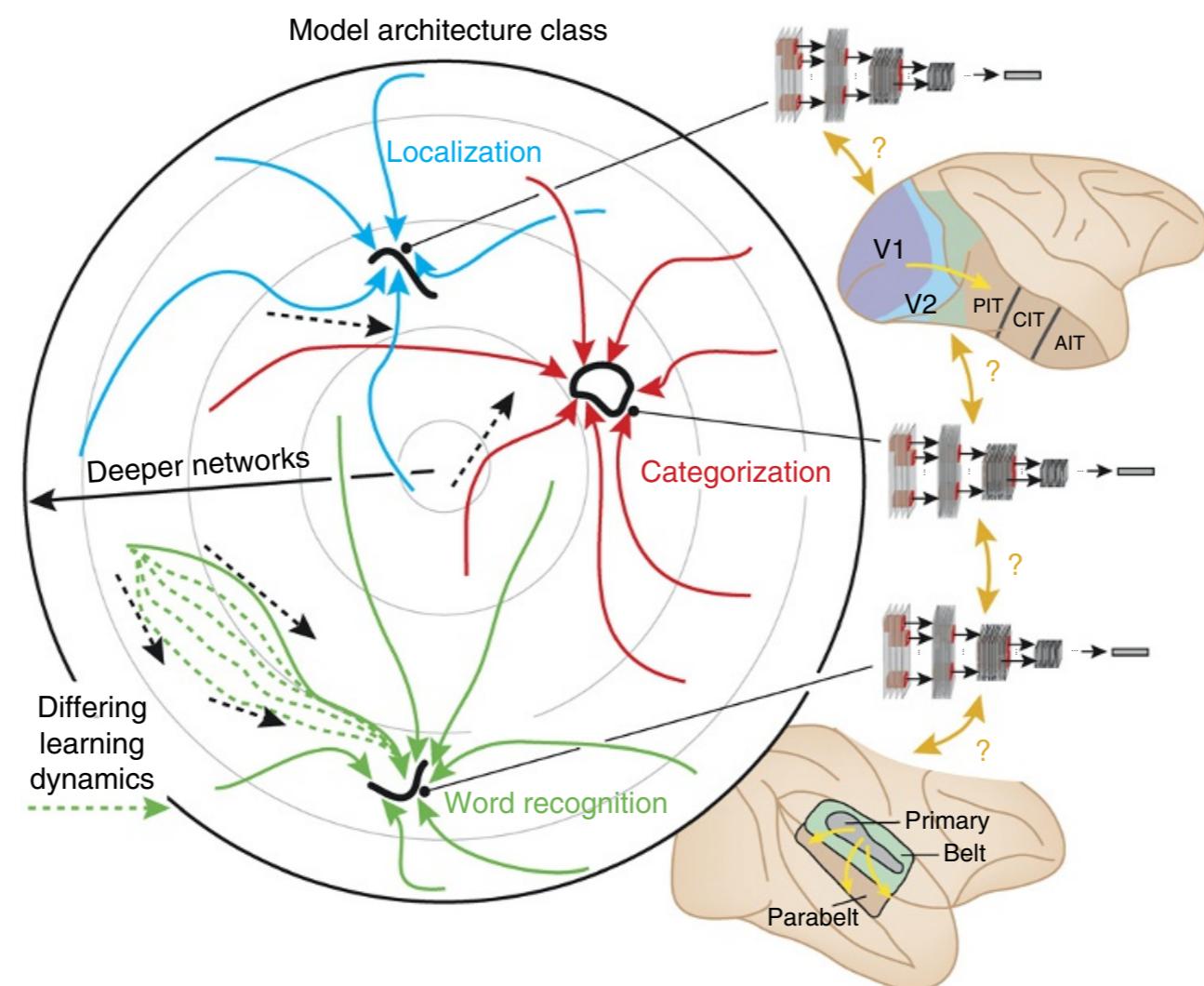


# Test deep neural networks on V4 neurons

Prediction

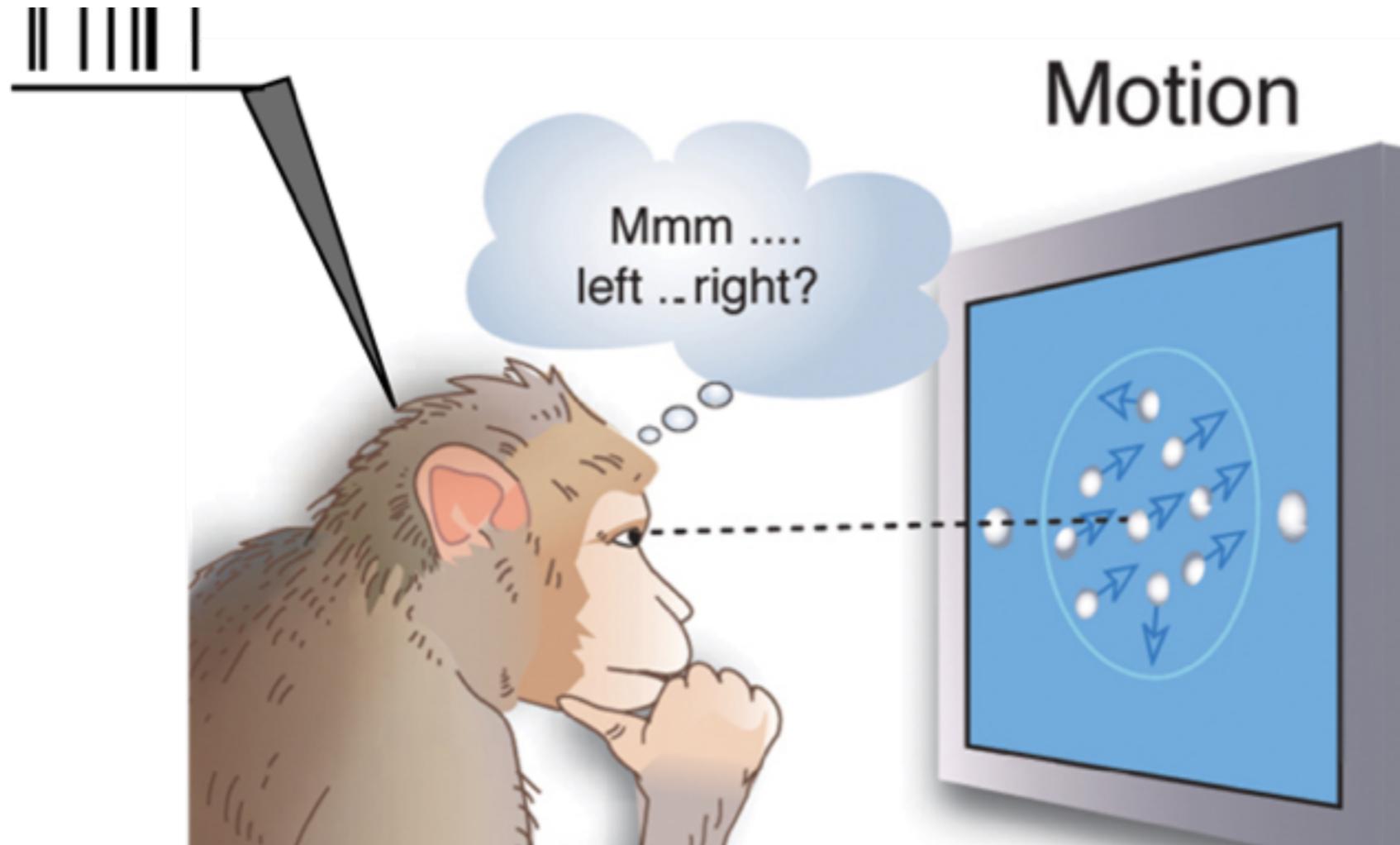


# Task specific deep neural networks

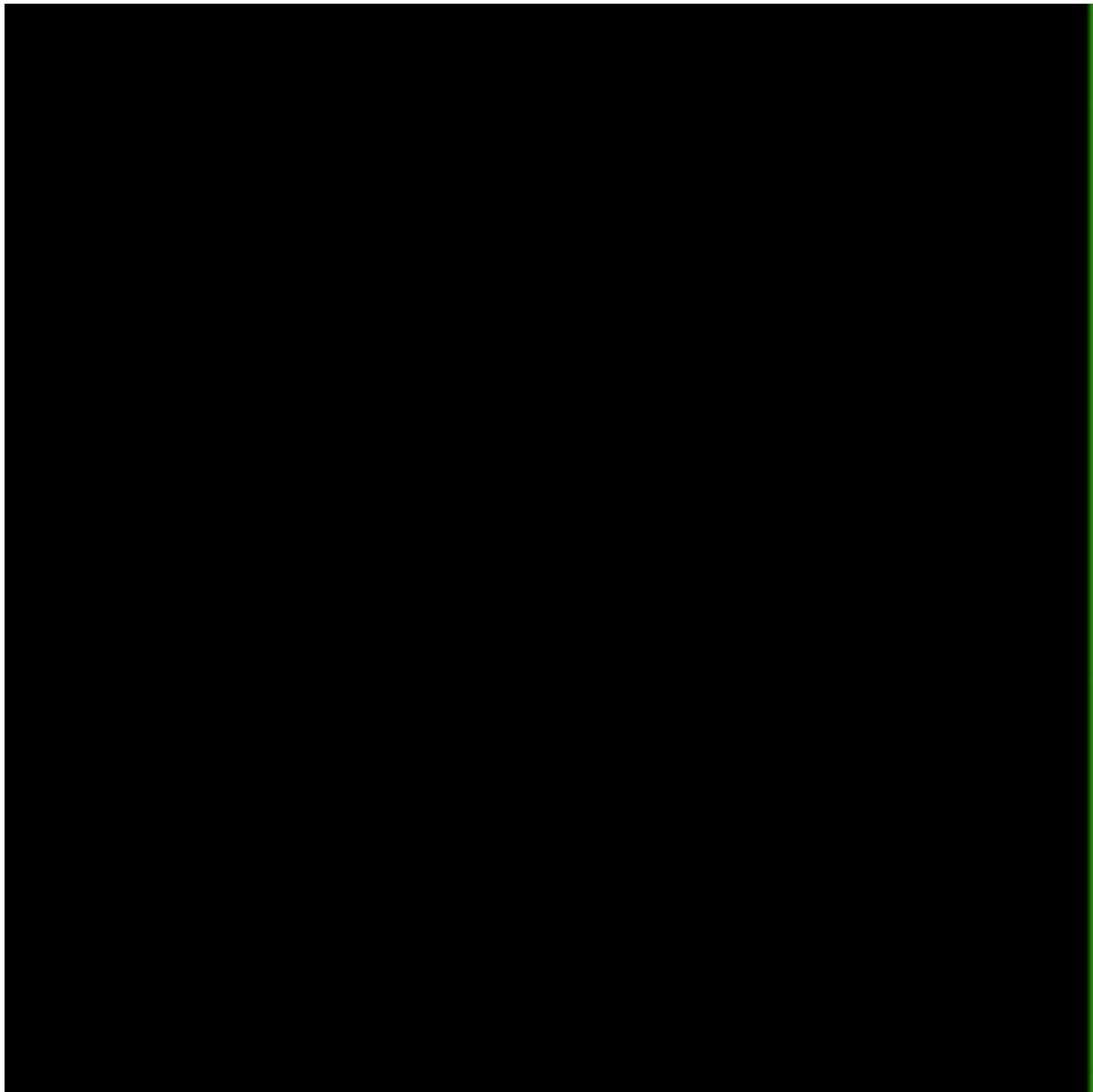


# Neuron Decoding

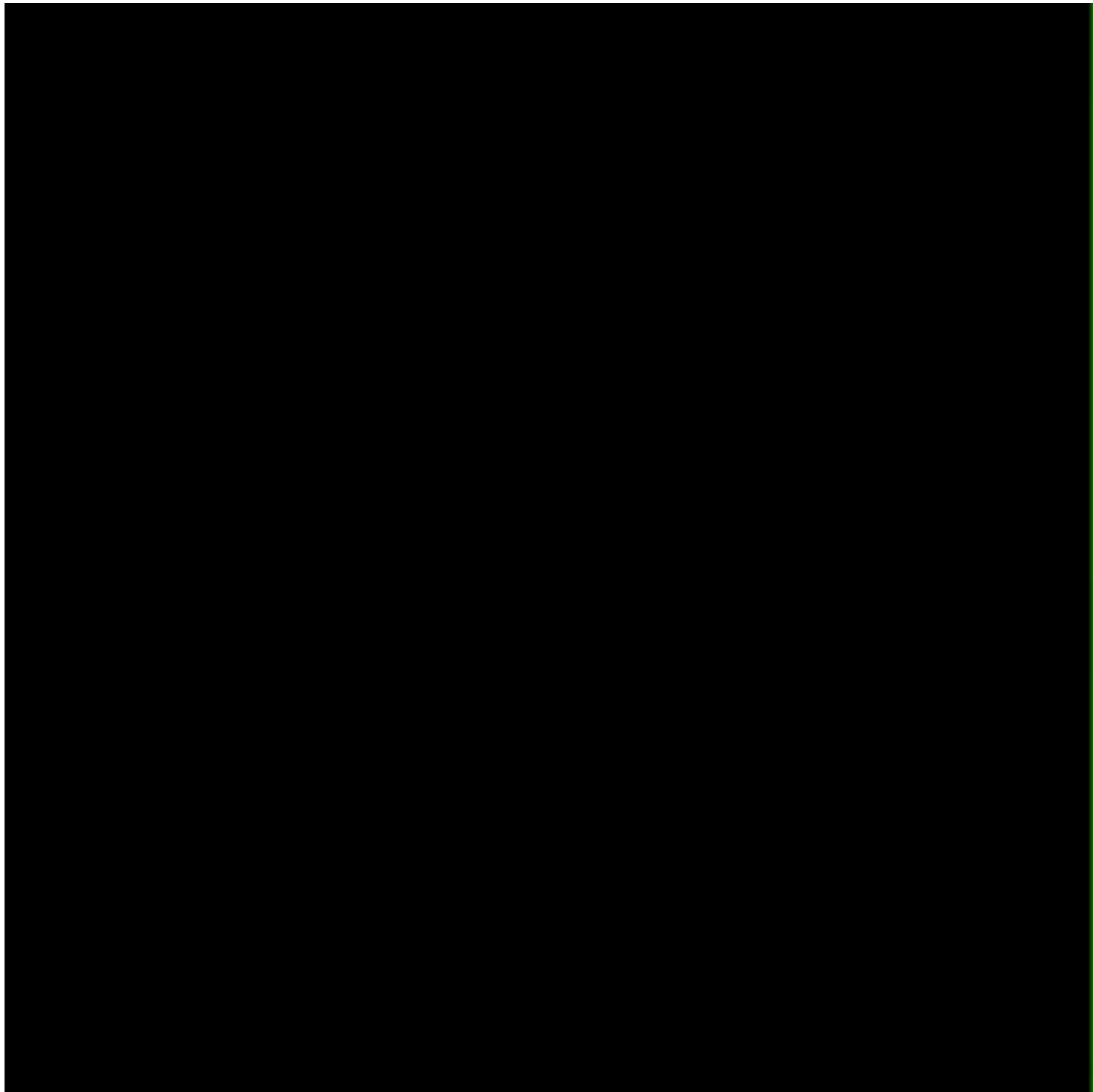
# Random dots motion discrimination task



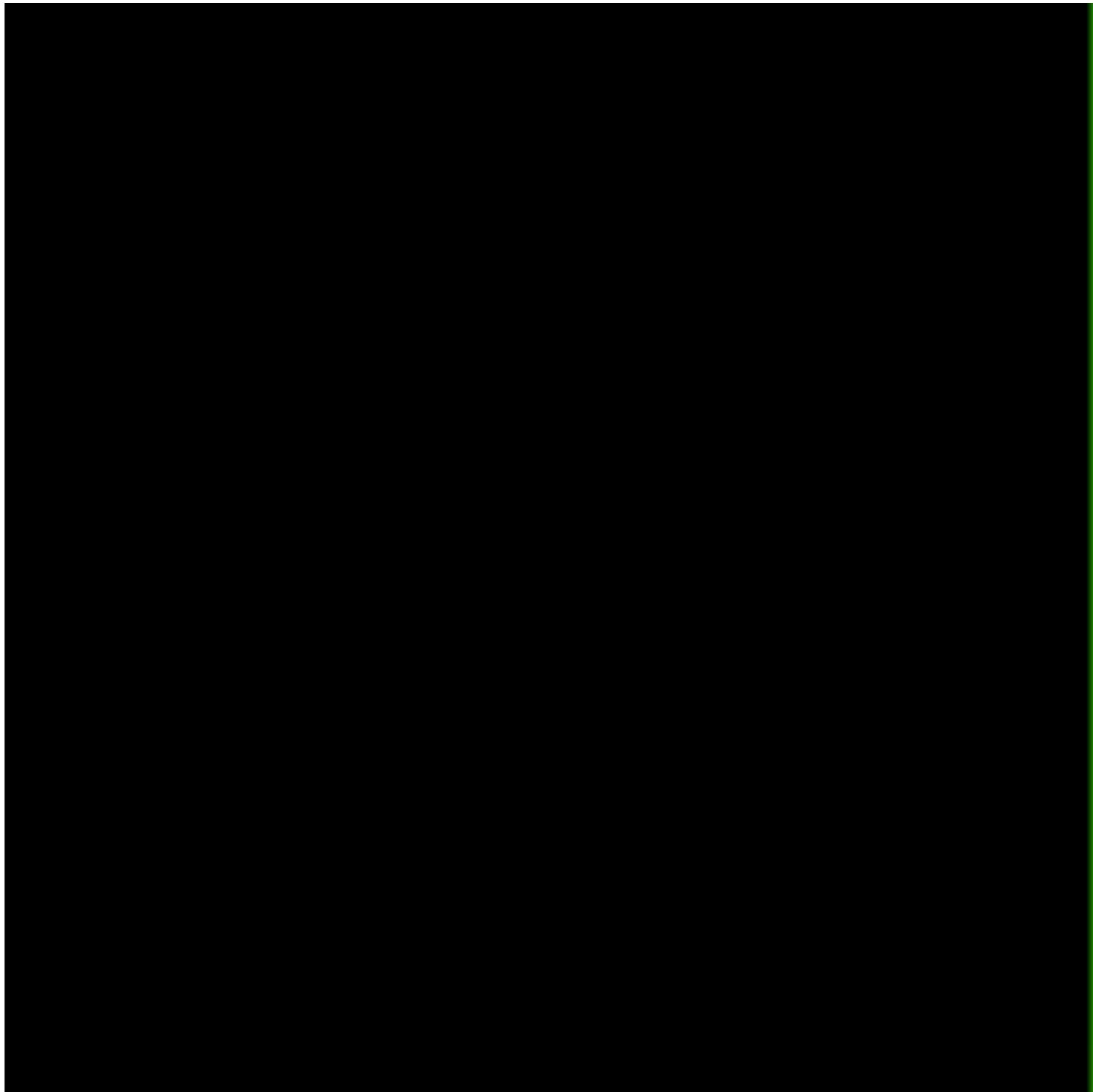
This slide and the following movies are adapted from Professor Tianming Yang's lectures



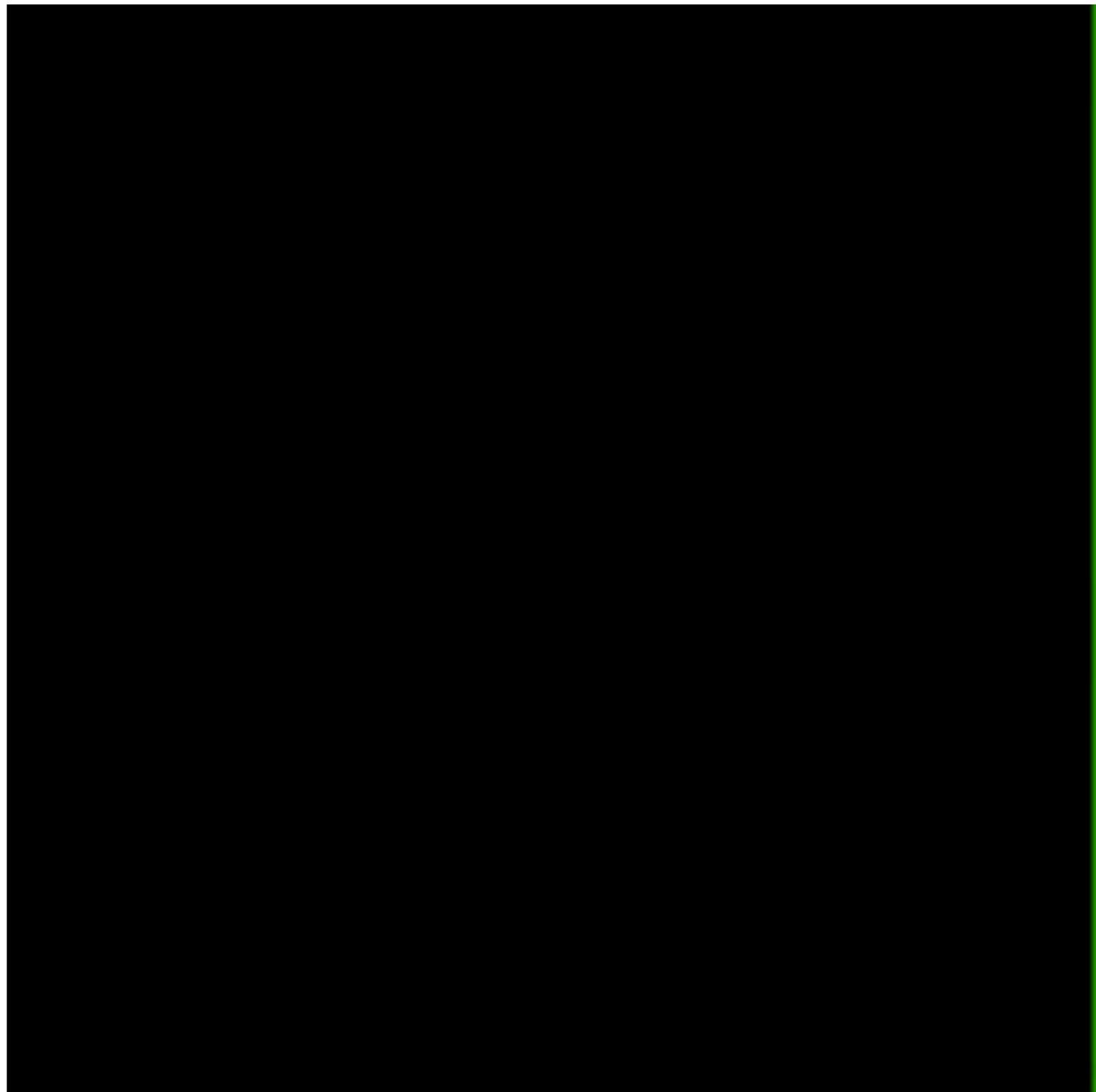
← 100%



51.2%

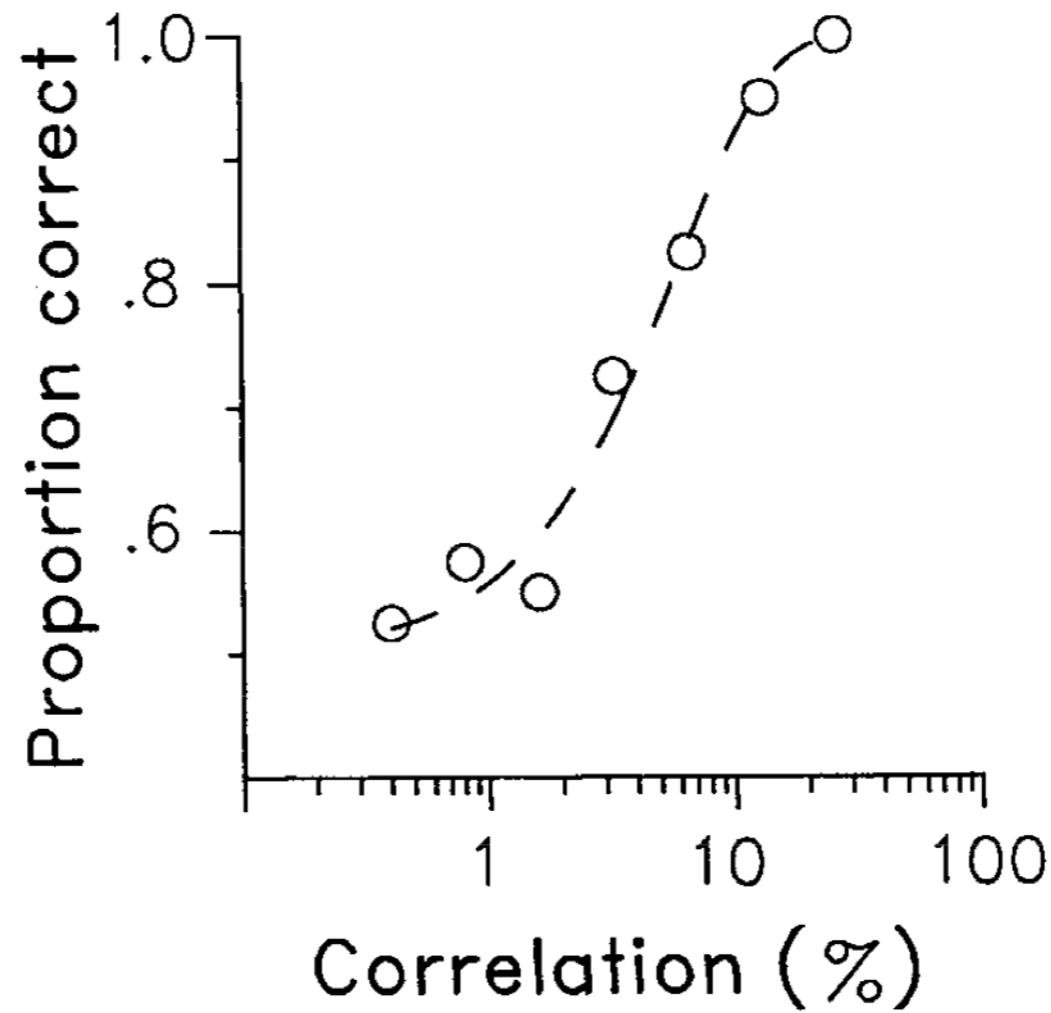


12.8%



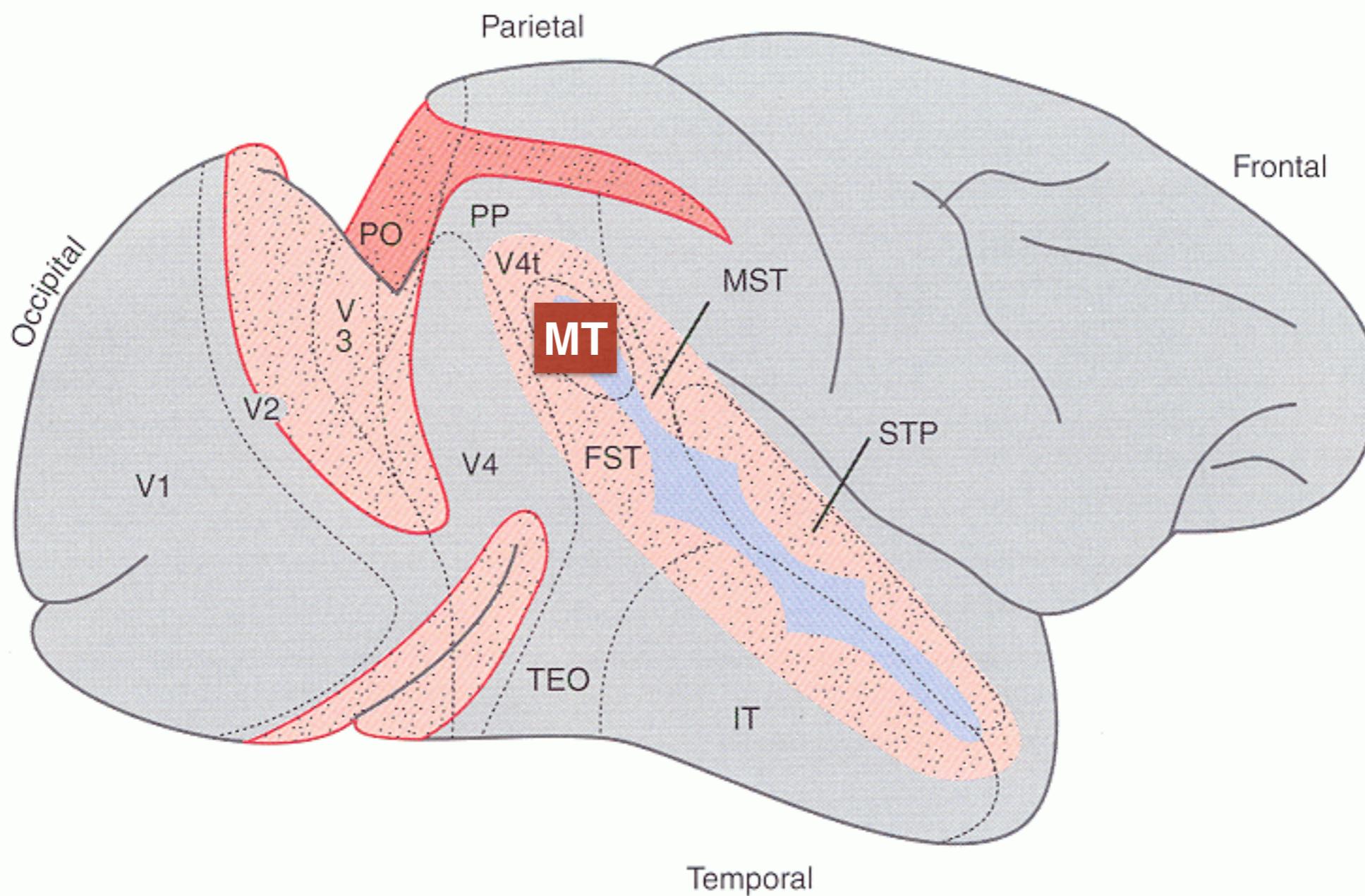
0%

# Behavior Performance

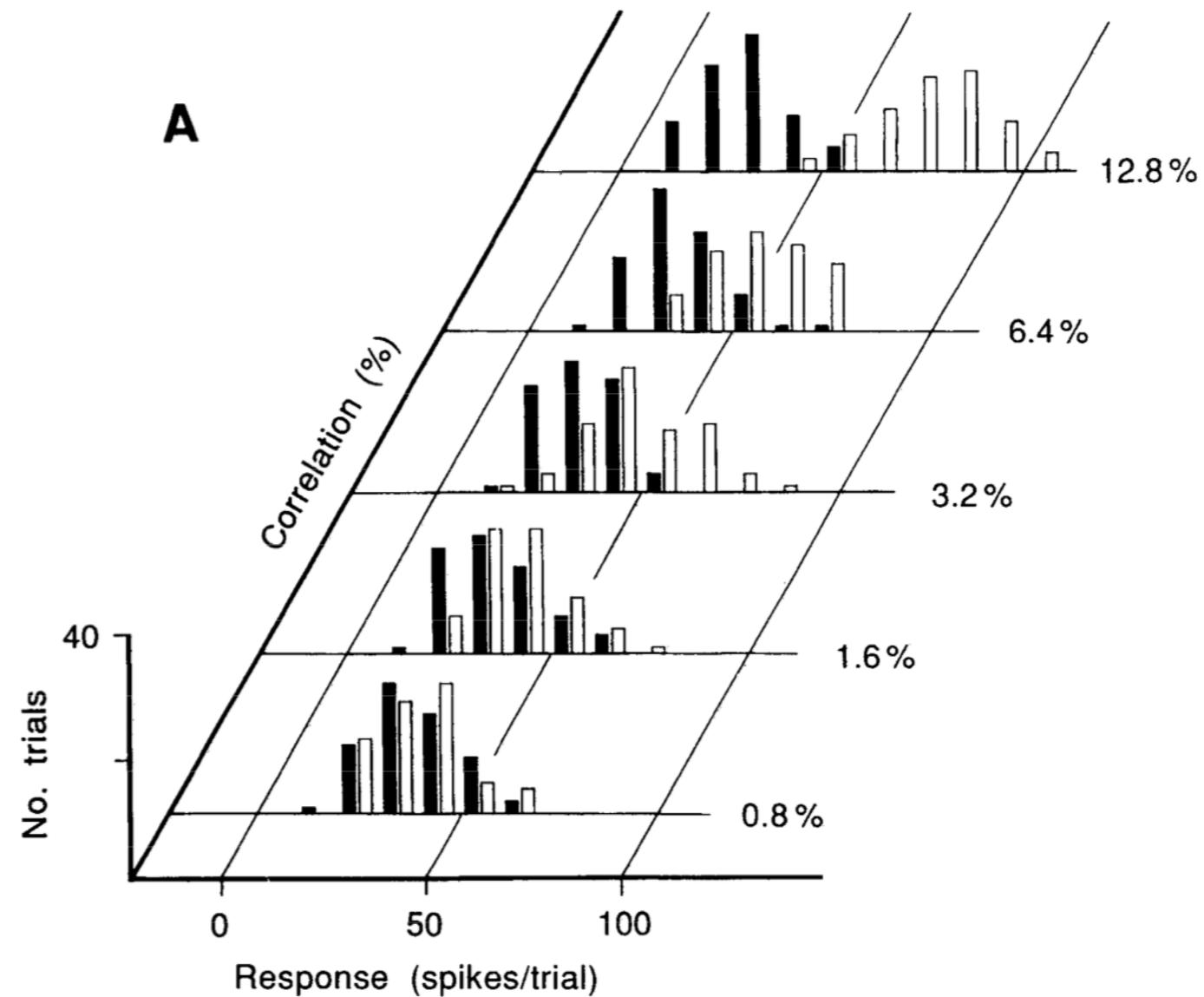


Britten, Shadlen, Newsome and Movshon, J Neurosci., 1992

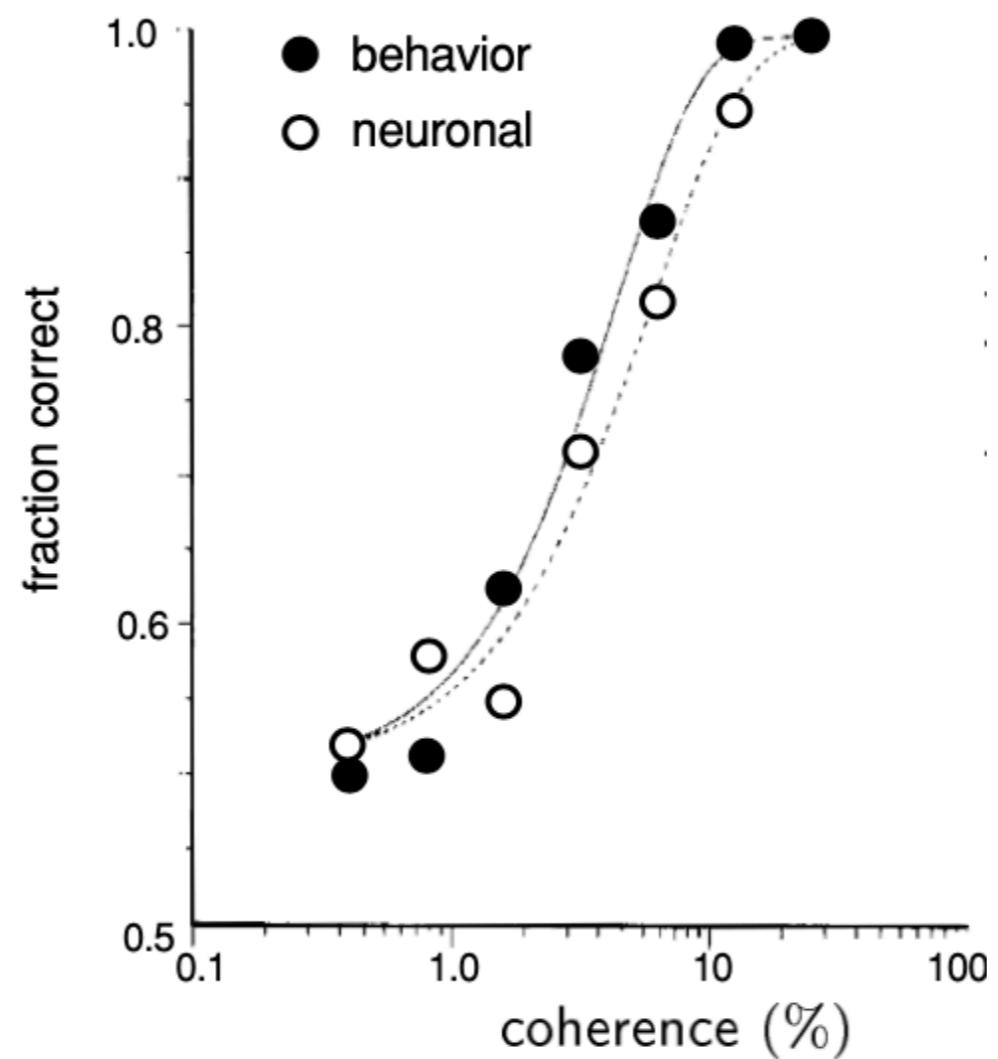
# Measuring neural activity in MT



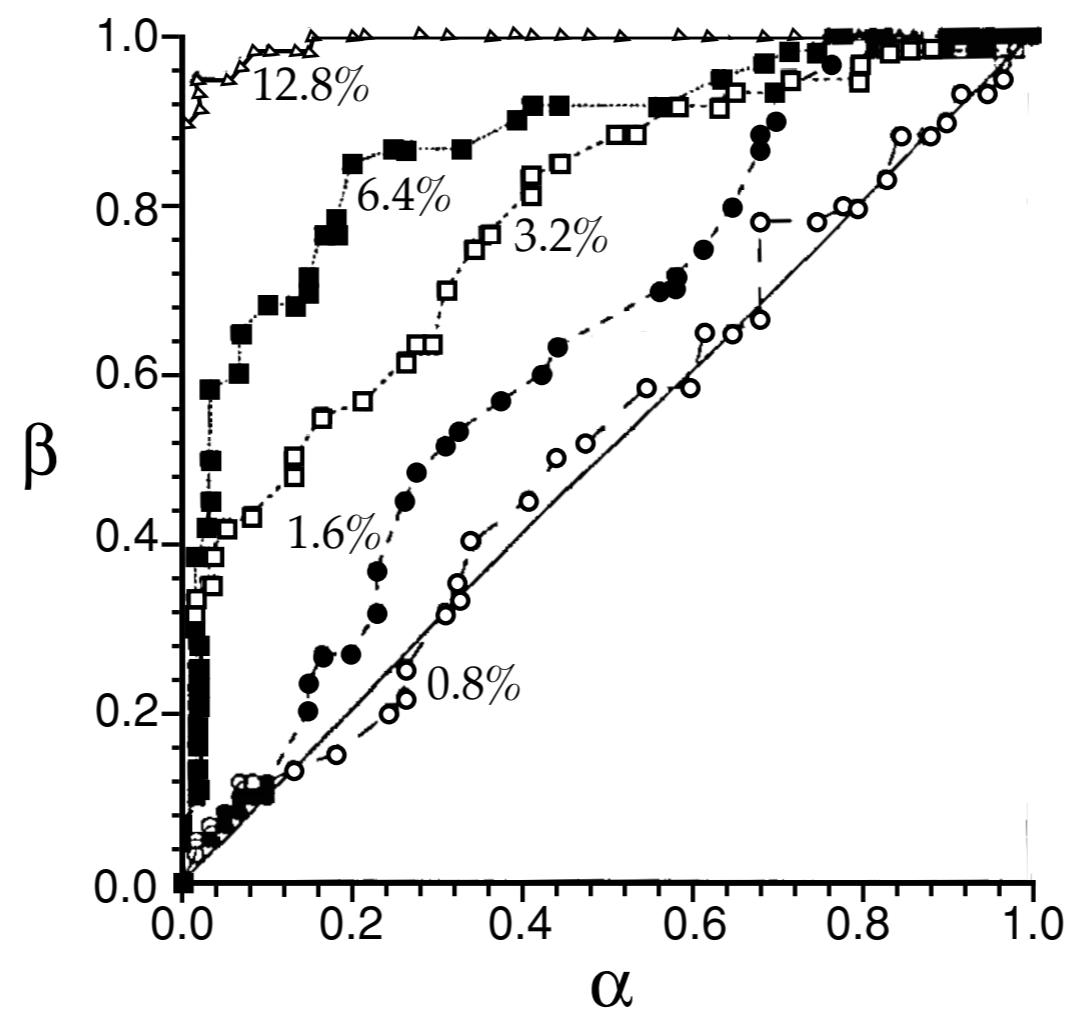
# Single Neuron recording



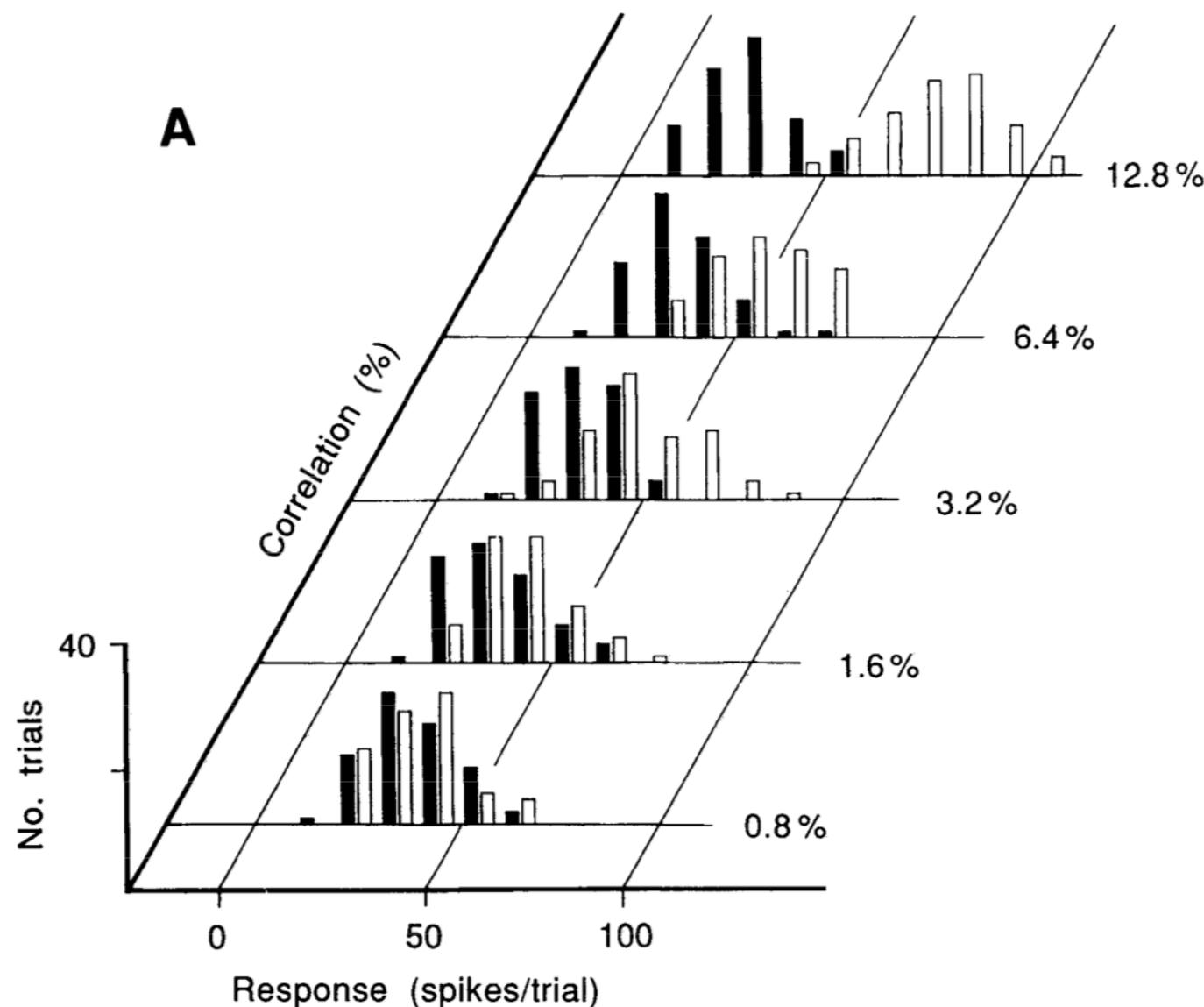
# Neural Correlates with Behavior



# ROC (receiver operating characteristic) curve



# Single Neuron recording



$$d = \frac{\langle r \rangle_+ - \langle r \rangle_-}{\sigma}$$

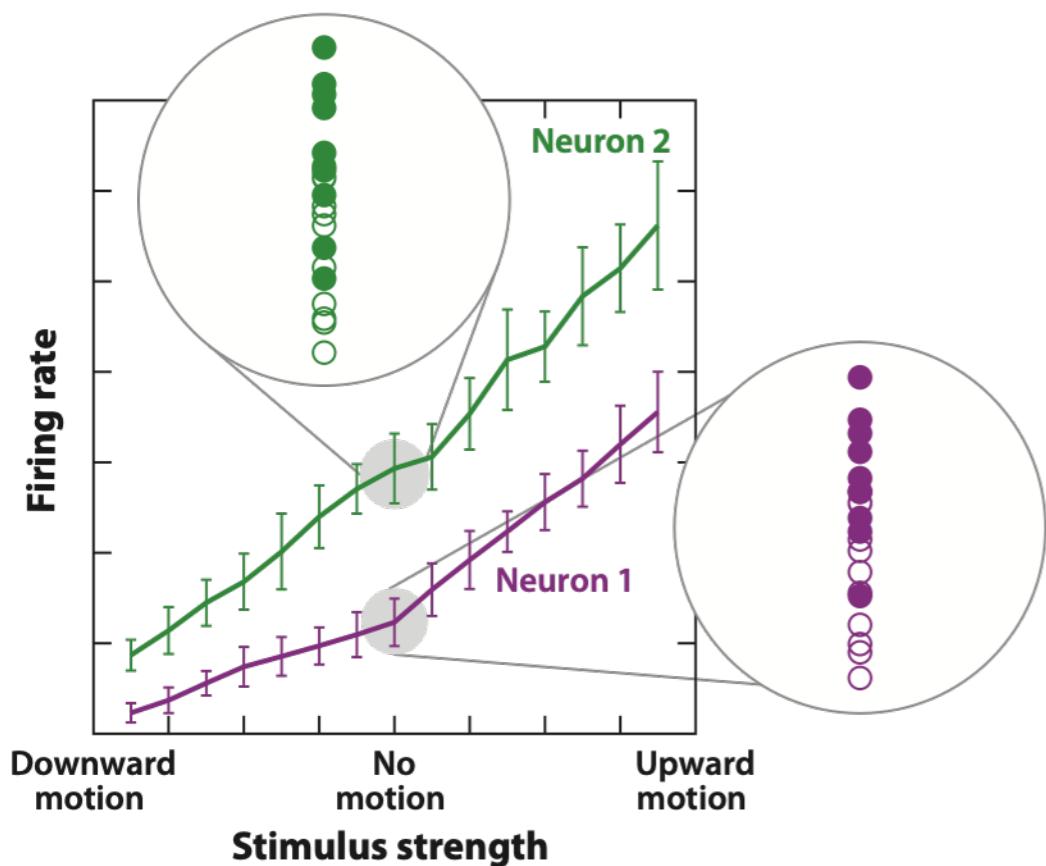
$$P(\text{correct}) = \frac{1}{\sqrt{\pi}} \int_{-d/2}^{\infty} dx \exp(-x^2)$$

# Population decoding

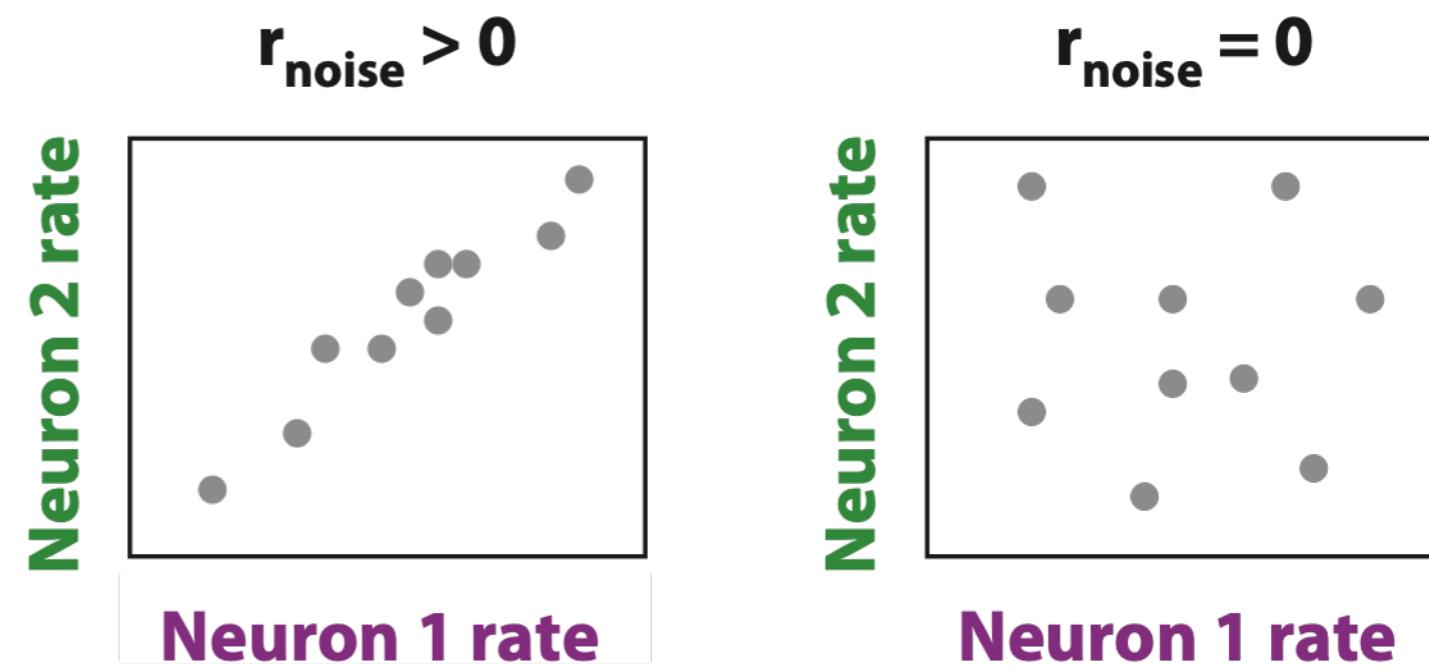
$$k_+ = \sum_{i=1}^N r_+^i \quad k_- = \sum_{i=1}^N r_-^i$$

$$d^2 = \frac{(\langle k_+ \rangle - \langle k_- \rangle)^2}{\langle (k_+ - \langle k_+ \rangle)^2 \rangle}$$

# Population decoding with correlated neurons

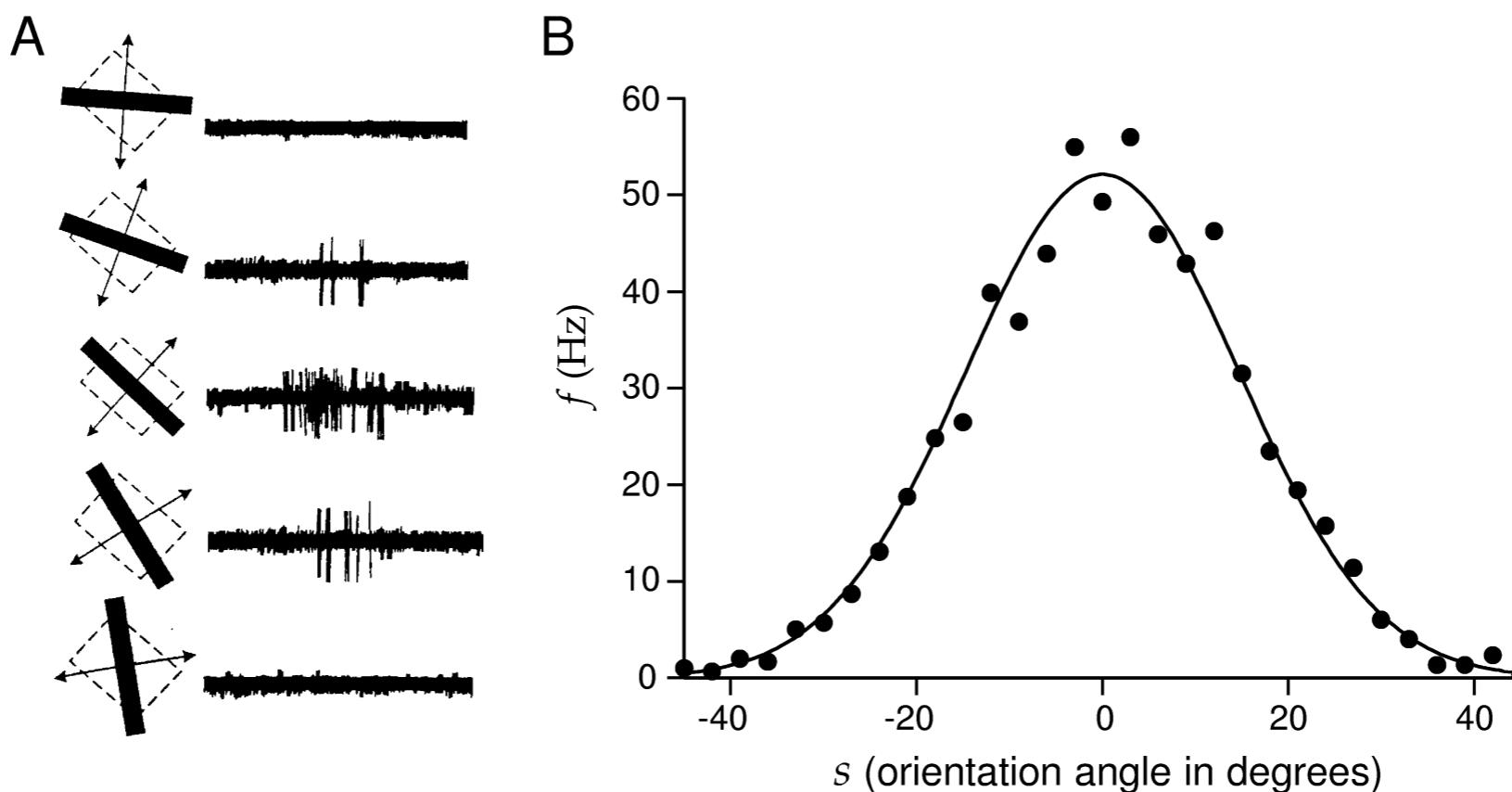


Signal Correlation



Noise Correlation

# Orientation selectivity of simple cells in the visual cortex

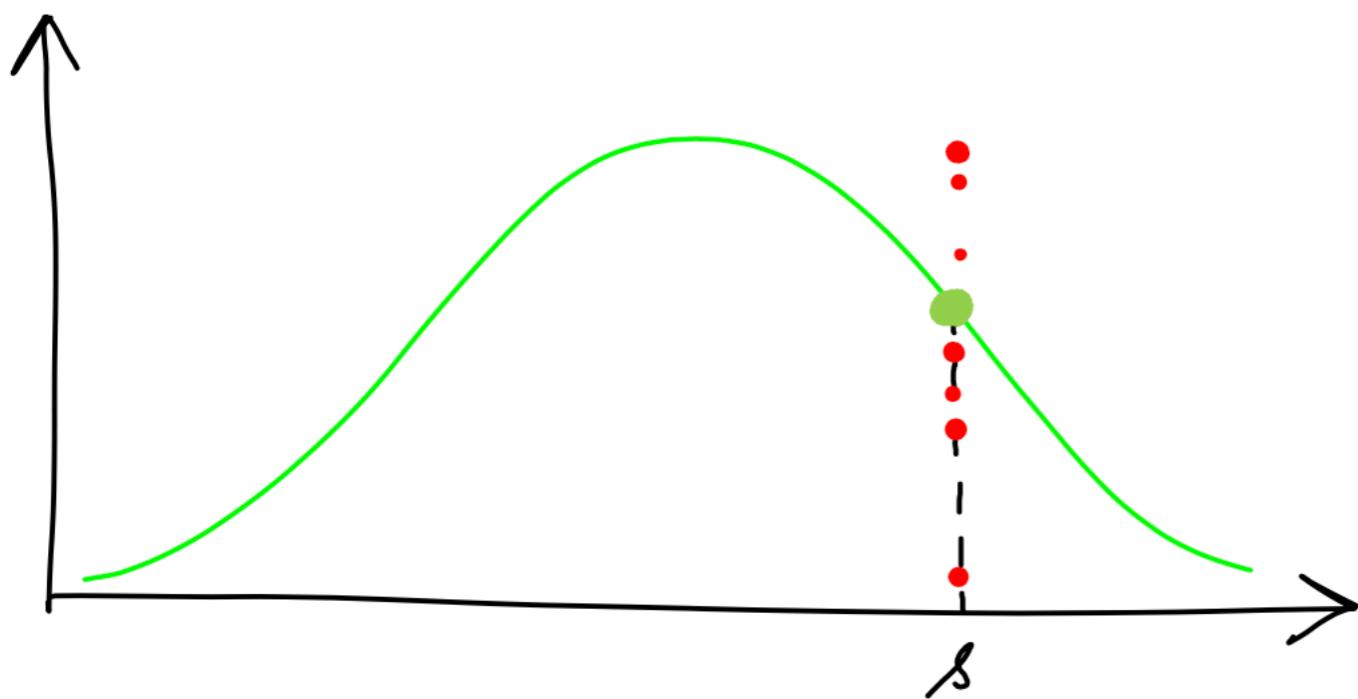


Gaussian tuning curve of a neuron in the primary visual cortex (V1)

# Bayesian decoding method

Variability

$$r = m(s) + \xi$$

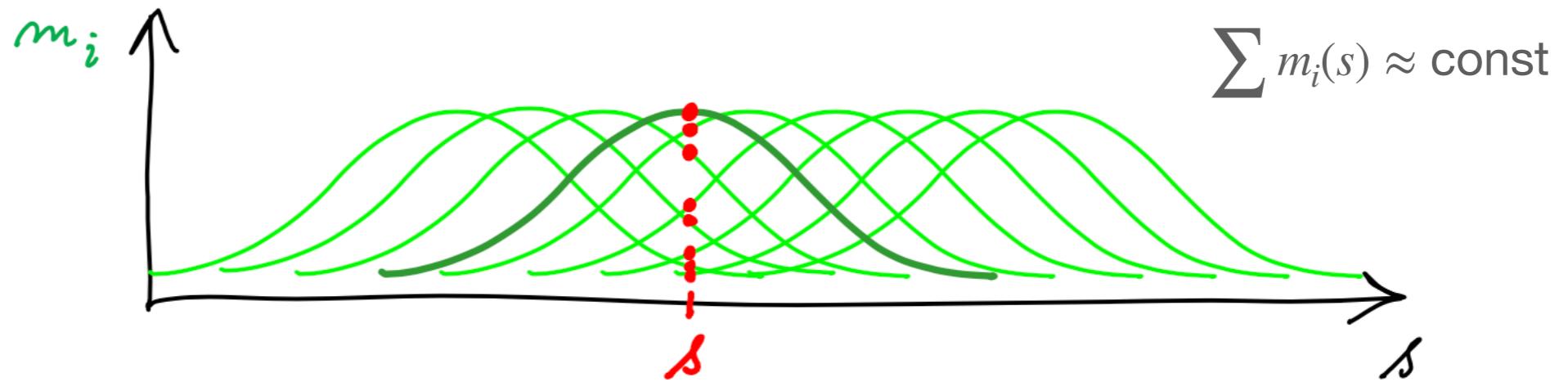


$$p(s|r) = \frac{p(r|s)p(s)}{p(r)}$$

Likelihood      Prior  
Normalization factor

$$m(s) = r_{max} \exp\left( -\frac{1}{2}(s - s_i)^2 / \sigma_i^2 \right)$$

# Population decoding



$$p(n_i | s) = \frac{[m_i(s)T]^{n_i}}{n_i!} \exp(-m_i(s)T)$$

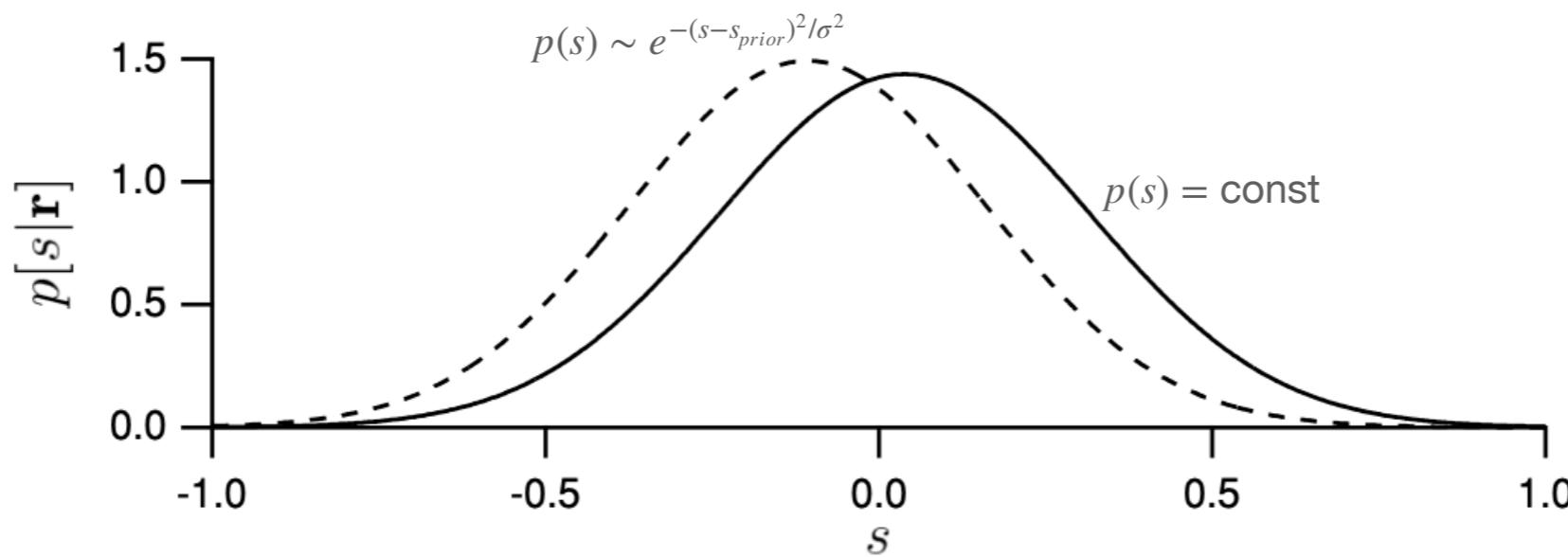
$$\underset{s}{\operatorname{argmax}} \log p(\mathbf{r} | s) = \sum_i \log p(n_i | s)$$

$$\sum_i^N r_i \frac{m'_i(s)}{m_i(s)} = 0$$

$$m'_i(s)/m_i(s) = (s_i - s)/\sigma_i^2$$

$$s = \frac{\sum r_i s_i / \sigma_i^2}{\sum r_i / \sigma_i^2}$$

# How good is our decoding method?



$$s - \langle s_{est} \rangle = b_{est} \quad \text{bias}$$

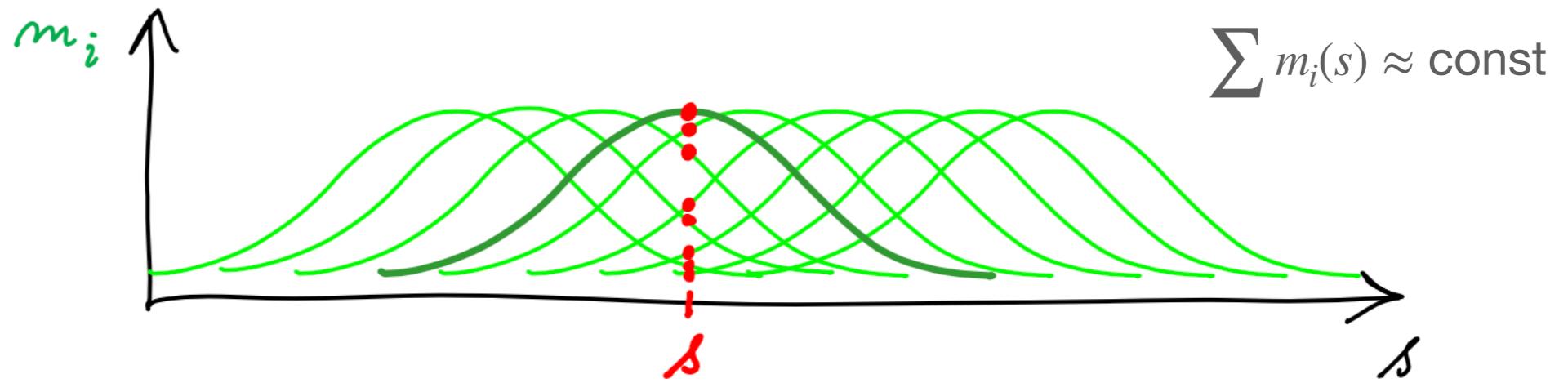
$$\langle (s - s_{est})^2 \rangle = \langle (s - \langle s_{est} \rangle + \langle s_{est} \rangle - s_{est})^2 \rangle = b_{est}^2 + \sigma_{est}^2 \quad \text{Variance}$$

# Cramer-Rao bound and Fisher information

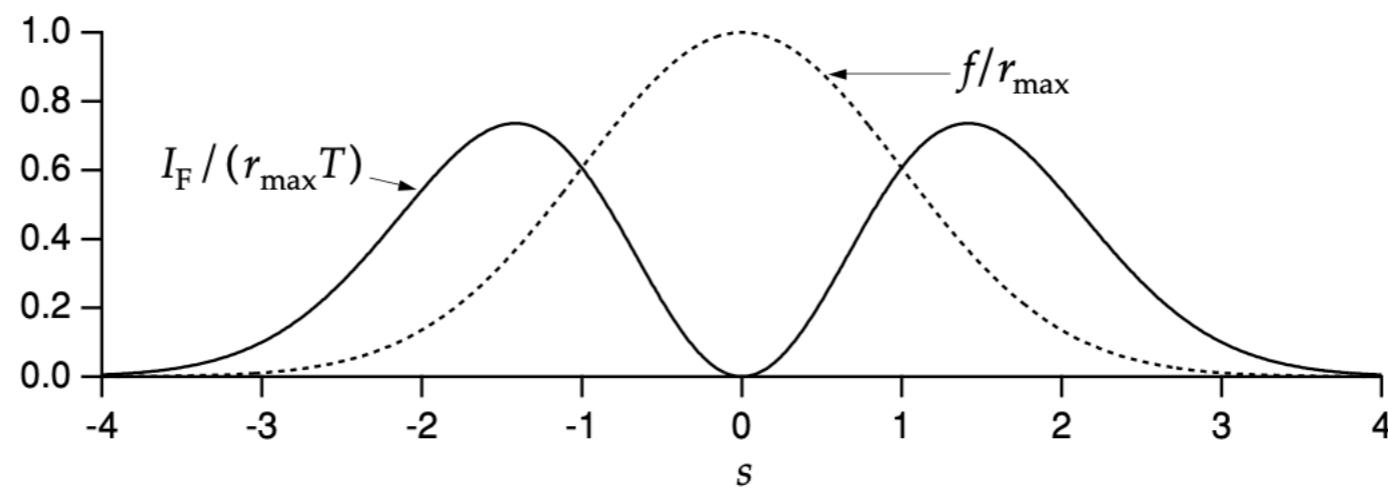
$$\sigma_{est}^2 \geq \frac{(1 + b_{est})^2}{I_F}$$

$$I_F = \left\langle -\frac{\partial^2 \log p(r|s)}{\partial s^2} \right\rangle = - \int dr p(r|s) \frac{\partial^2 \log p(r|s)}{\partial s^2}$$

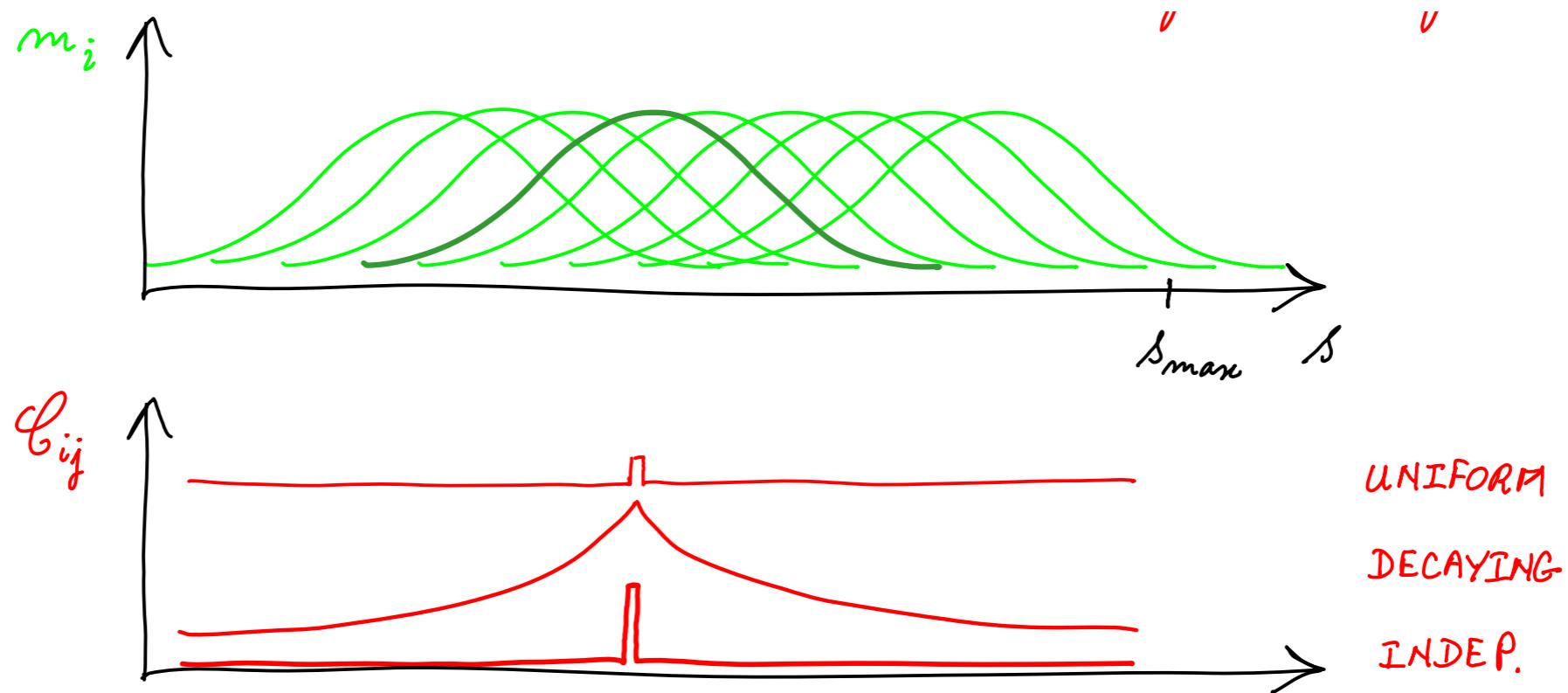
# Population decoding



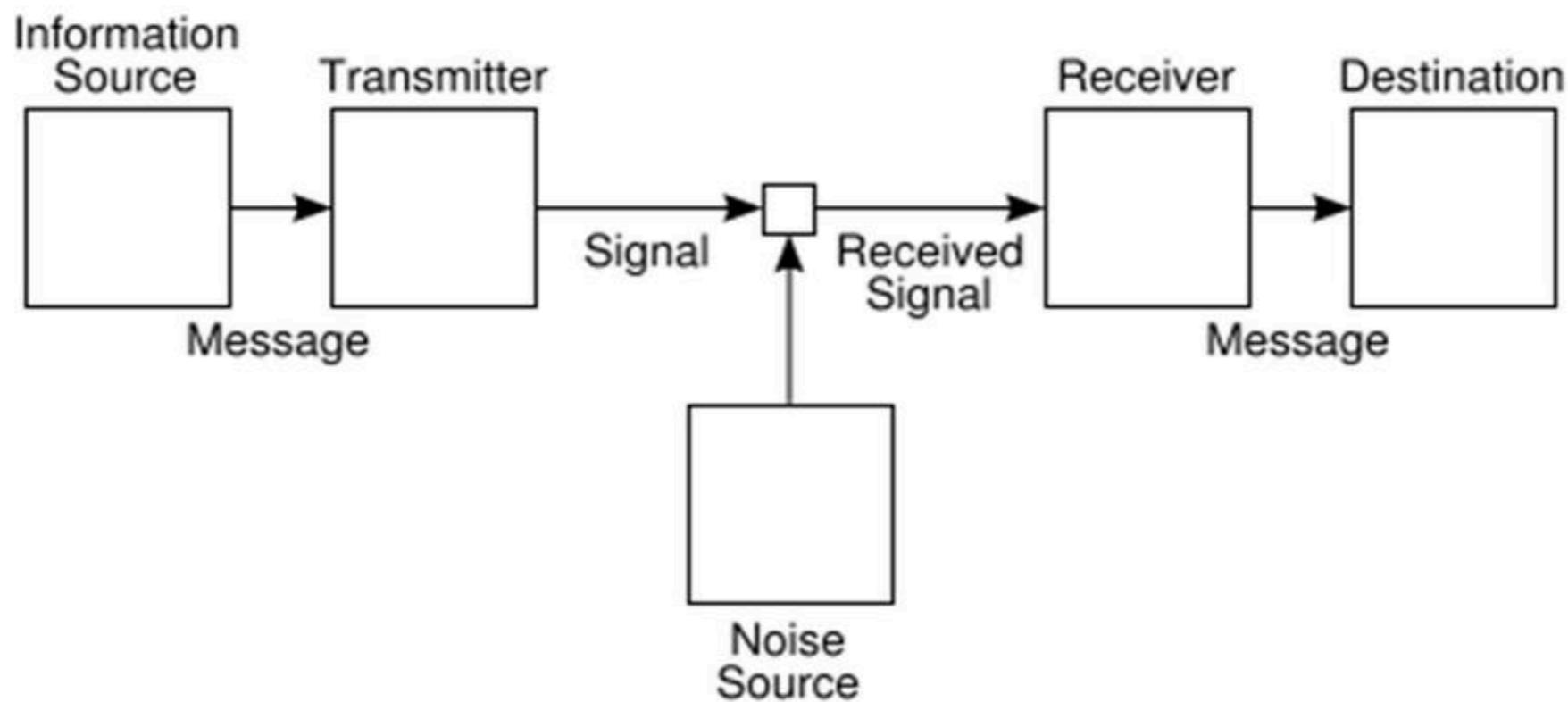
$$I_F = T \sum_i \frac{(m'_i(s))^2}{m_i(s)}$$



# Spatially correlated Neurons



# Information Theory



# Shannon Information

The game of sixty three

- $x \geq 32?$
- is  $x \bmod 32 \geq 16?$
- is  $x \bmod 16 \geq 8?$
- is  $x \bmod 8 \geq 4?$
- is  $x \bmod 4 \geq 2?$
- is  $x \bmod 2 \geq 1?$

$$I = \log_2 64 = 6 \text{ bit}$$

# Shannon Information

$$h = \log_2 N \text{ when } P(r) = \frac{1}{N}$$

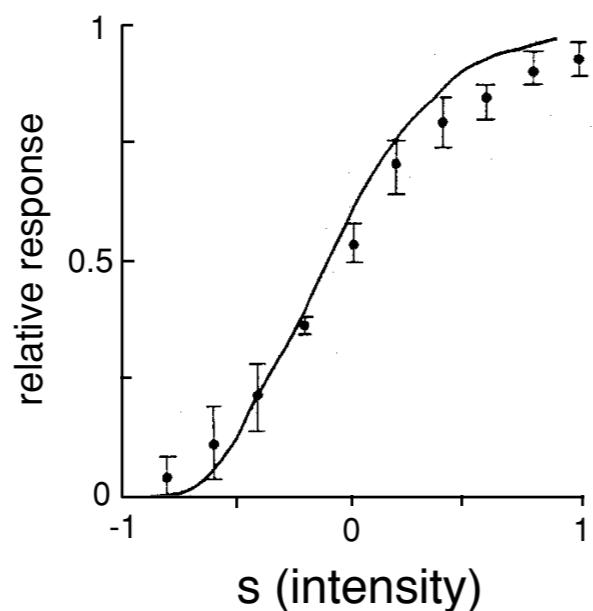
$$I = - \sum_r P(r) \log_2 P(r)$$

Continuous limit

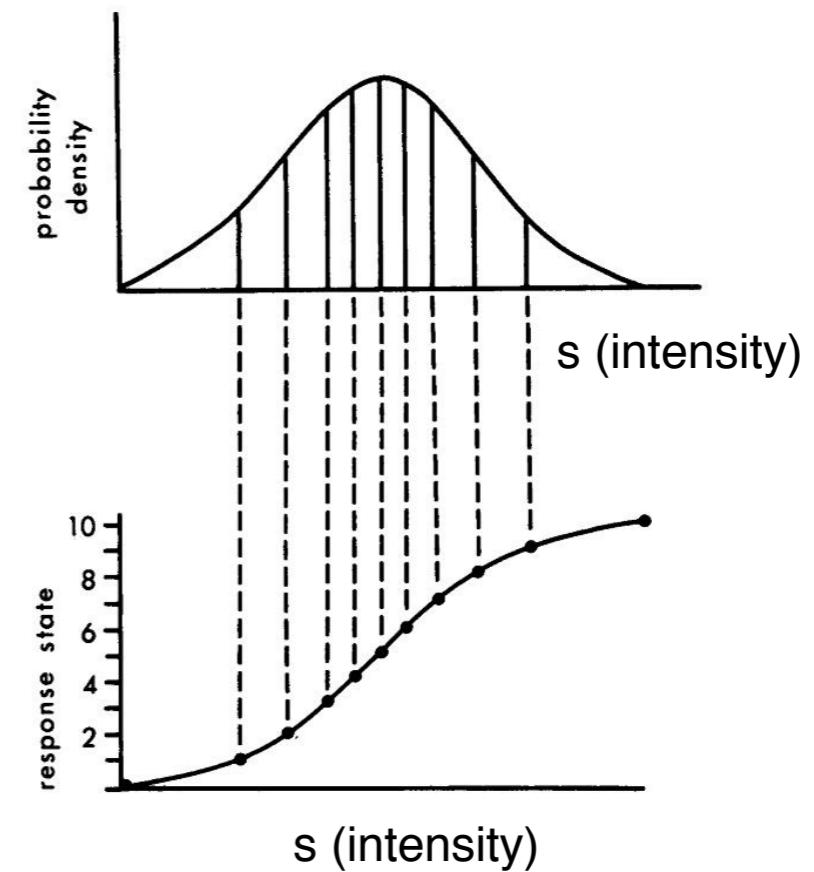
$$I = - \sum_r p(r) \Delta r \log_2 p(r) \Delta r$$

$$\lim_{\Delta r \rightarrow 0} (I + \log_2 \Delta r) = - \int dr p(r) \log_2 p(r)$$

# Maximizing entropy principle



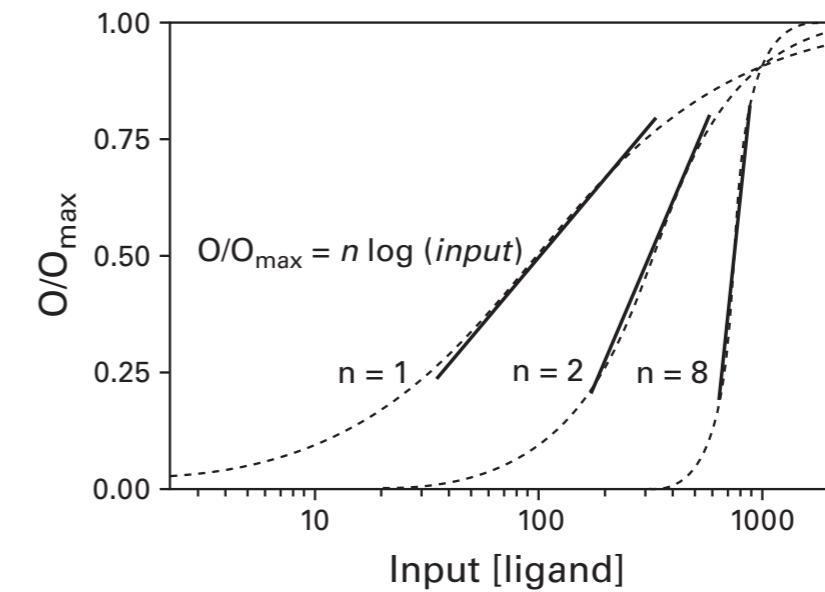
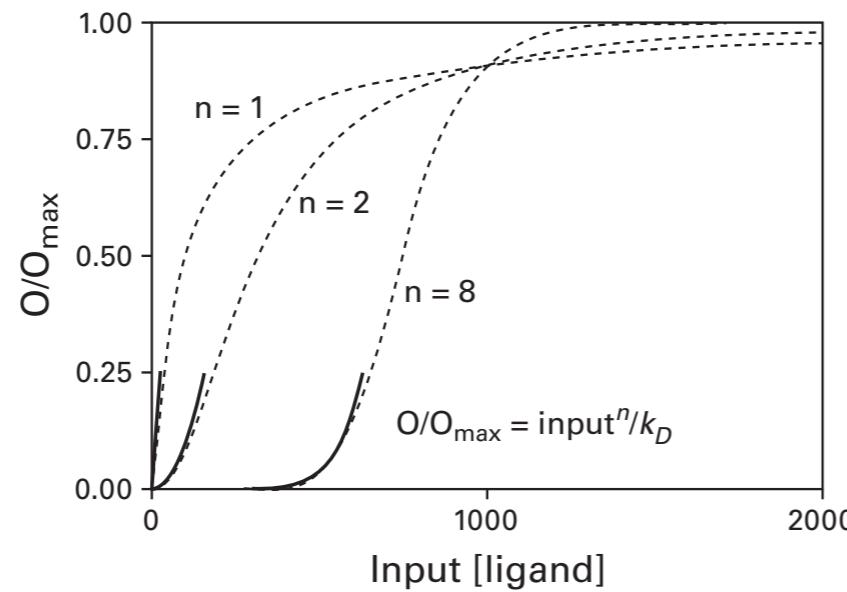
$$r = f(s)$$



$$H = - \int_0^{r_{max}} dr p(r) \ln p(r) \quad \frac{df}{ds} = r_{max} p(s)$$

$$p(r) = \frac{1}{r_{max}}$$

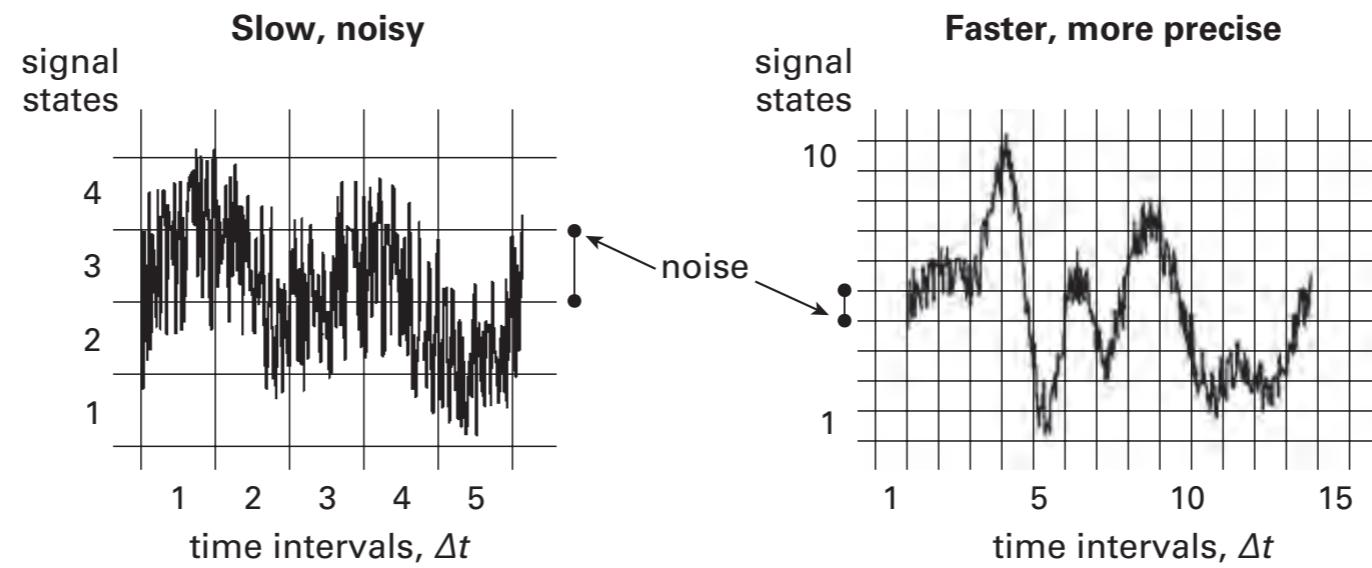
# I/O function to match input distribution



$$\frac{O}{O_{\max}} = \frac{[\text{ligand}]^n}{k_D + [\text{ligand}]^n}$$

$n$  : number of binding sites

# Information rate (analog signal)



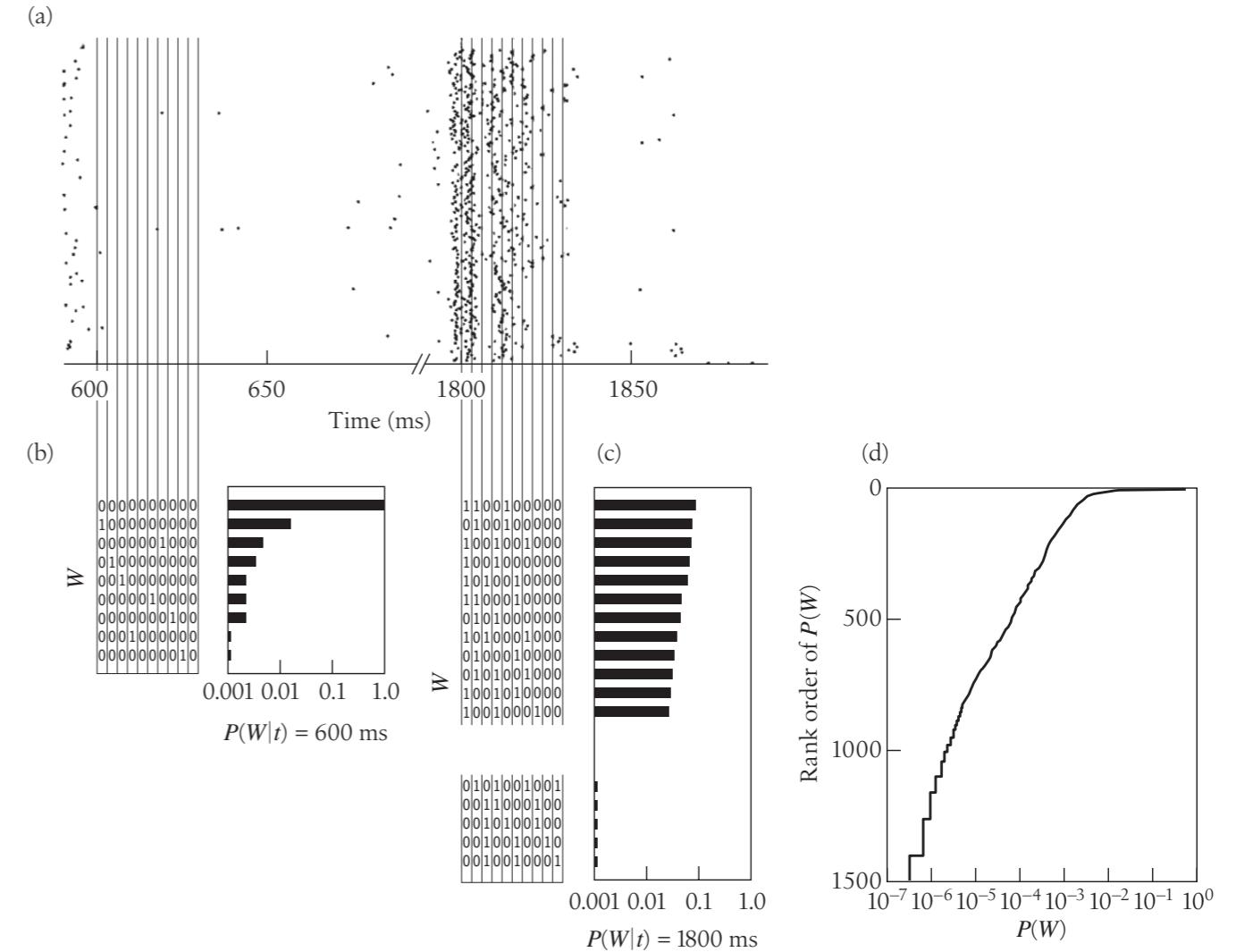
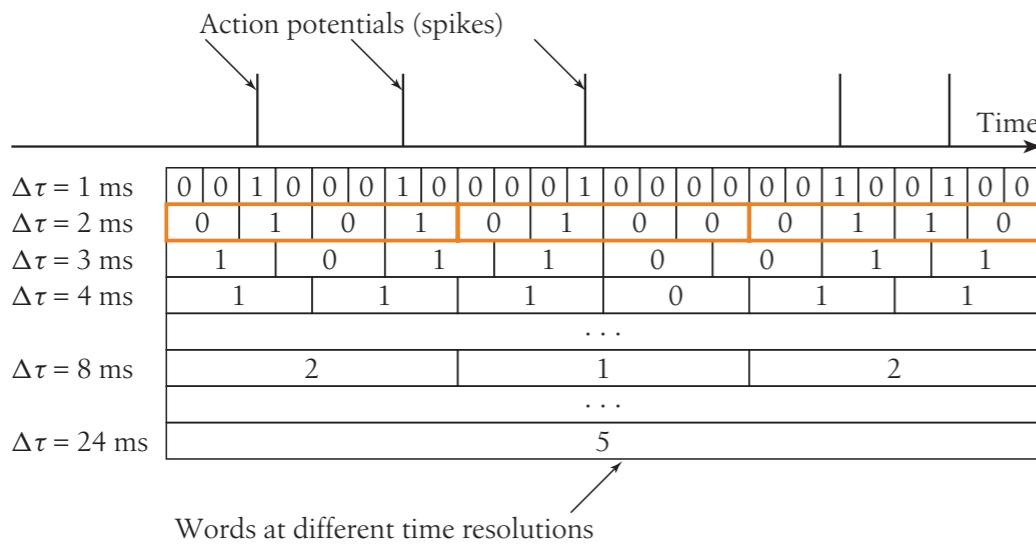
**Figure 5.3**

**Signal range, noise, and response dynamics determine the information rates of analogue signals.** Noise divides a waveform's signal range into discriminable states, and states can change at time intervals  $\Delta t$ . The faster, more reliable waveform obviously conveys more details of the signal. From equation 5.5, it also has a higher information rate because it has a higher  $S/N$  and, with shorter  $\Delta t$ , changes level at a higher rate.

$$I = \frac{1}{\Delta t} \log_2(1 + S/N)$$

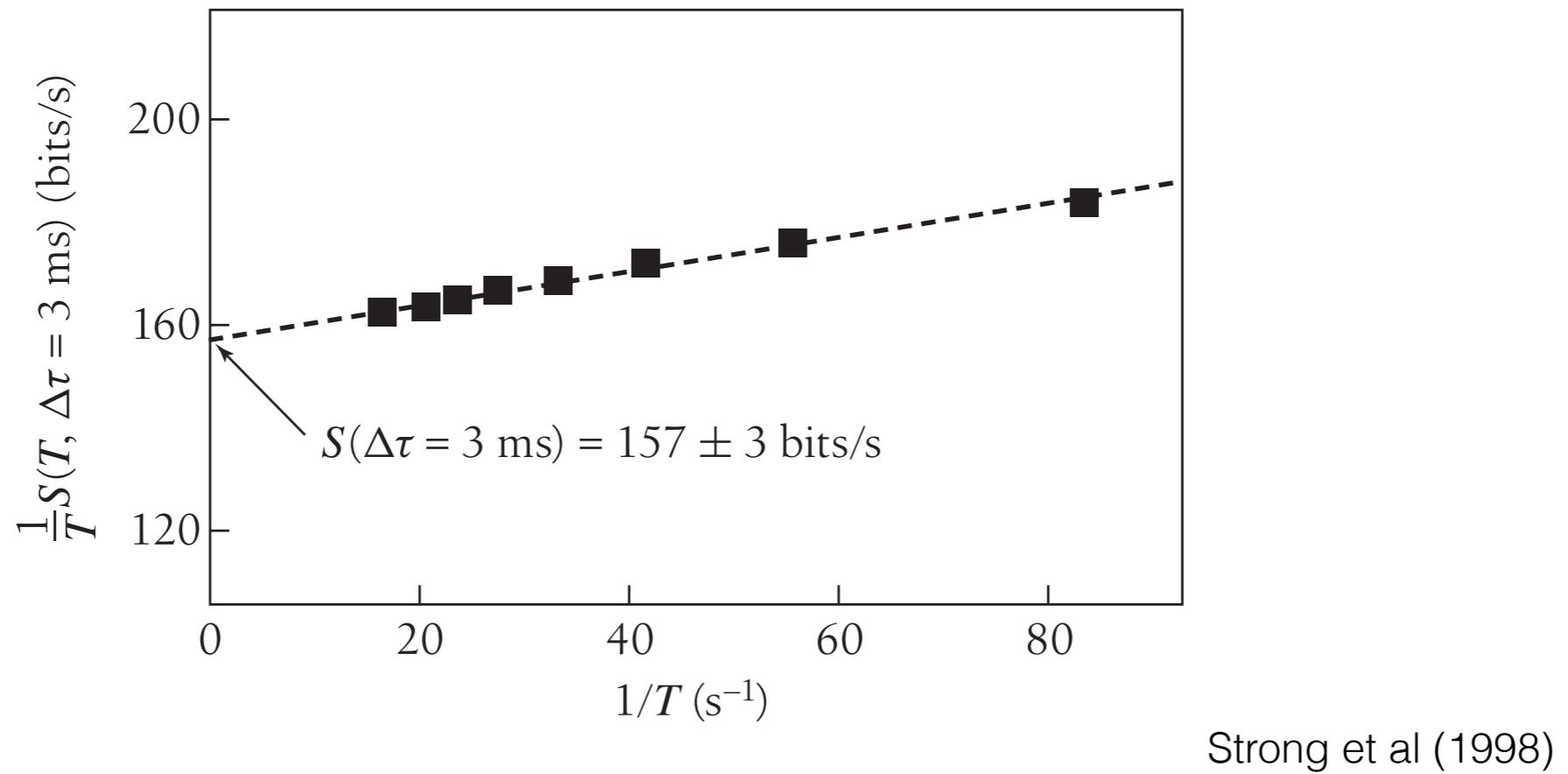
$$I = \int df \log_2[1 + S(f)/N(f)]$$

# Information rate (spikes)



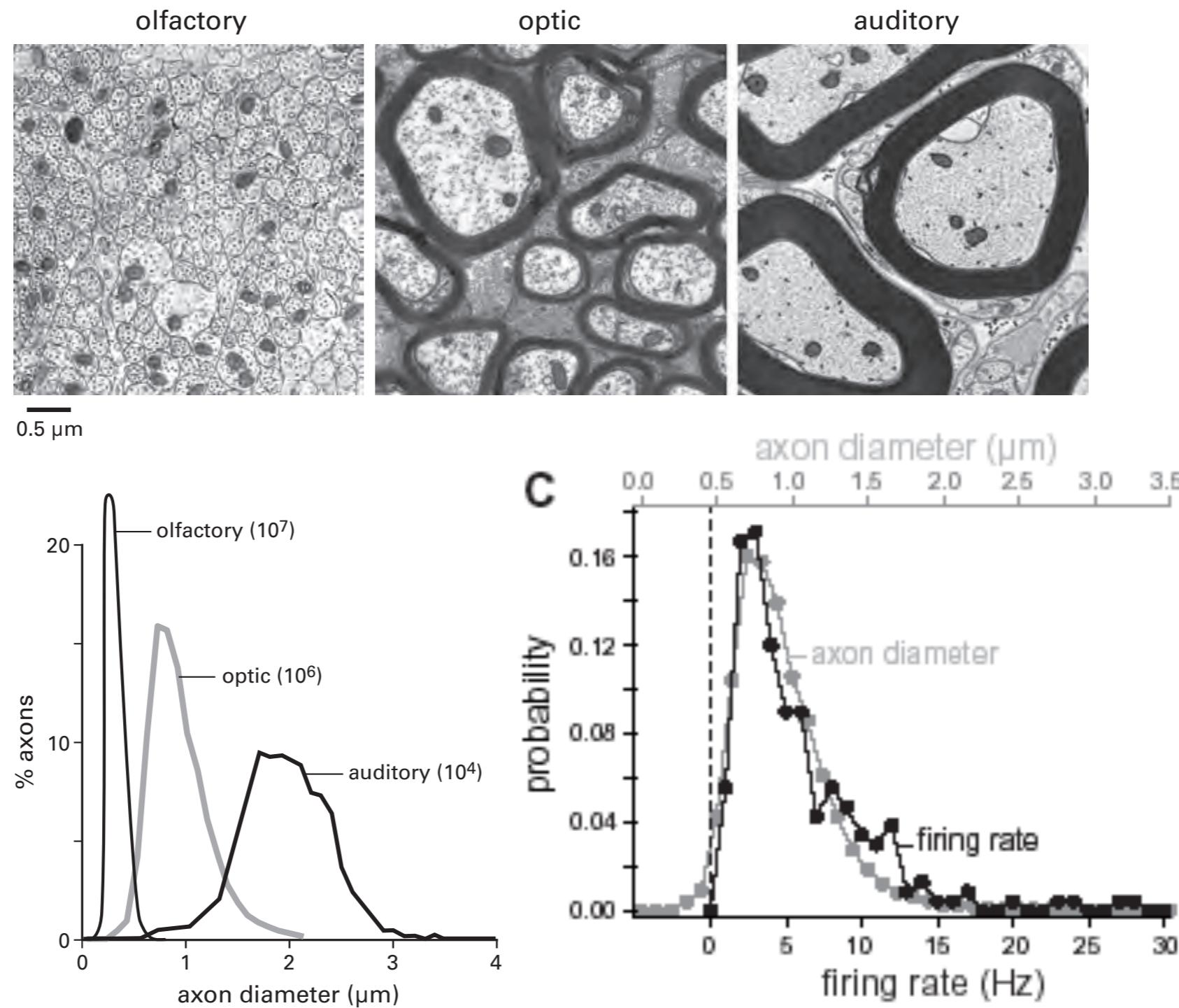
De Ruyter van Steveninck et al (1997)

# Information rate (spikes)

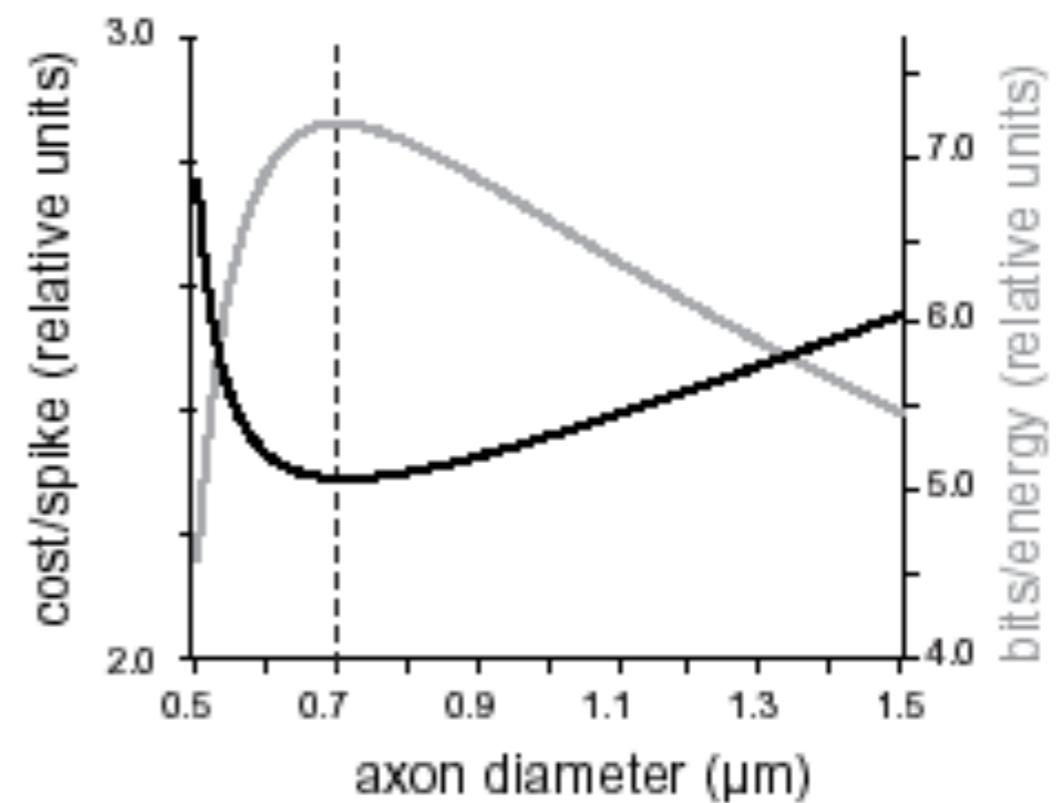
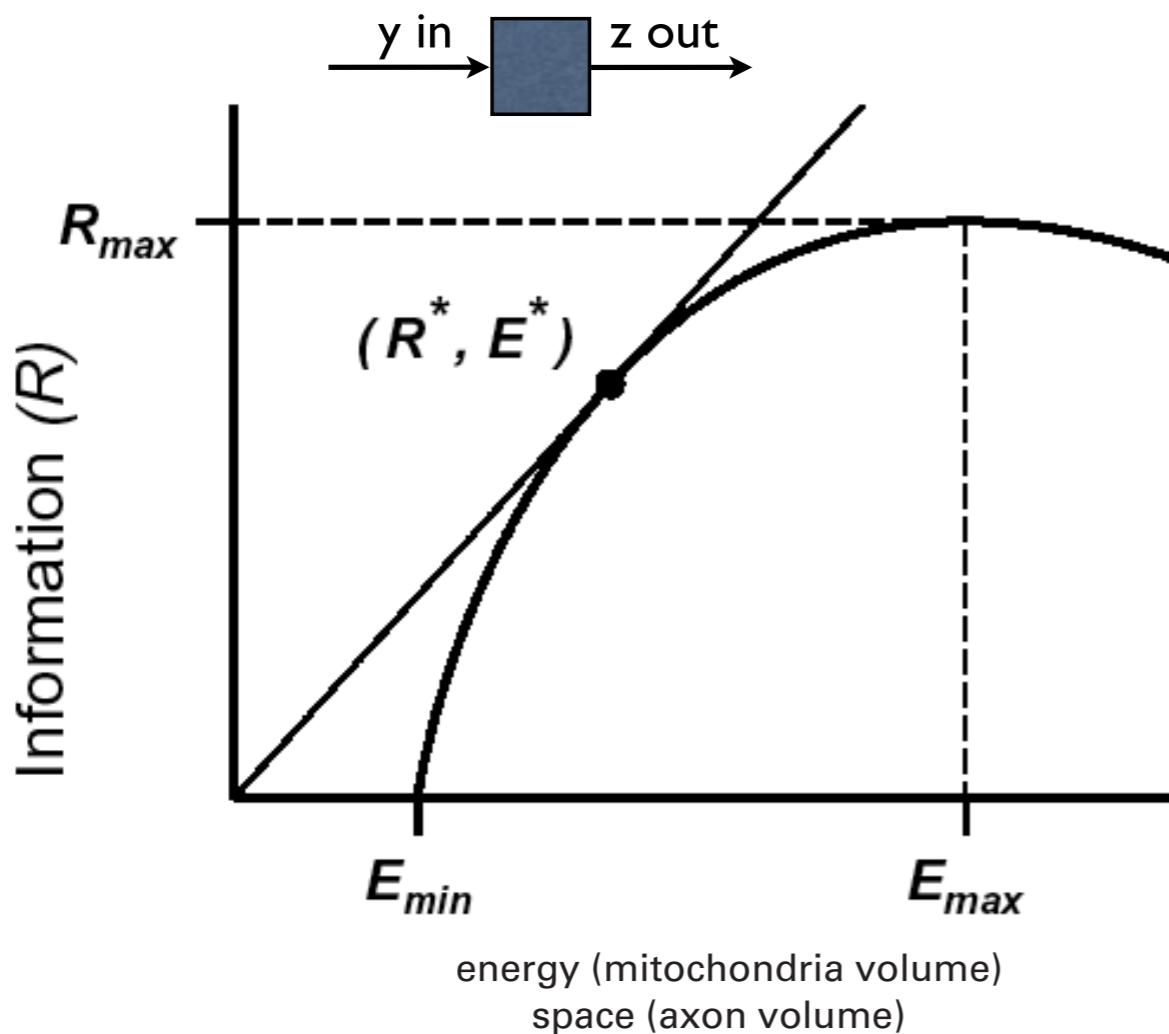


$$\frac{1}{T}S(T, \Delta\tau) = \mathcal{S}(\Delta\tau) + \frac{C(\Delta\tau)}{T} + \dots,$$

# Information transmission is costly



# Law of diminished returns



Heterogeneity & efficiency in the brain Vijay Balasubramanian 2015

Perge et al, 2018 How the optic nerve allocates space, energy capacity, and information

# Retina Circuit

