Lecture 4 – Optimal sensory coding

Functional advantages of response properties and changes in those properties

Why do cells have a particular nonlinear response function?

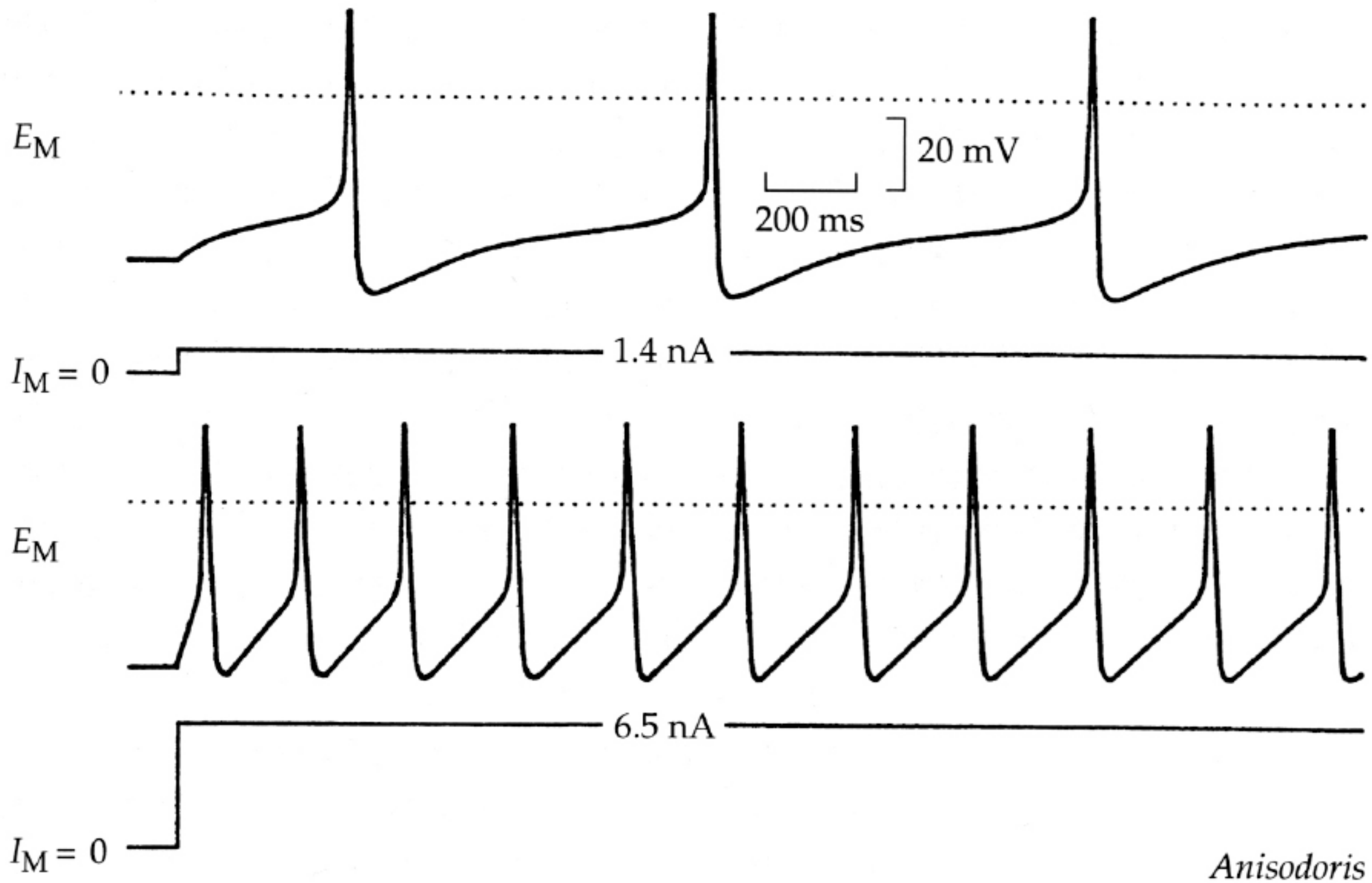
Why does the nonlinearity change?

Why do cells have a certain duration filter?

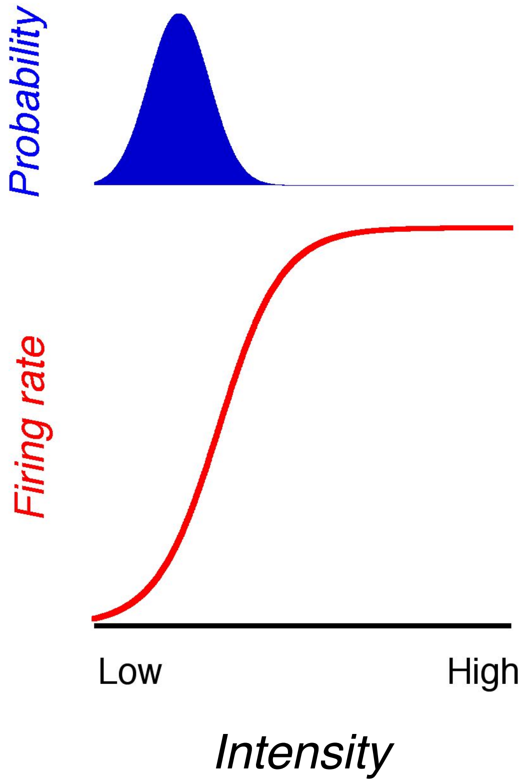
Why do they have a certain shape filter?

Why does the filter change?

Neurons have a limited dynamic range
set by maximum and minimum output levels, and by noise

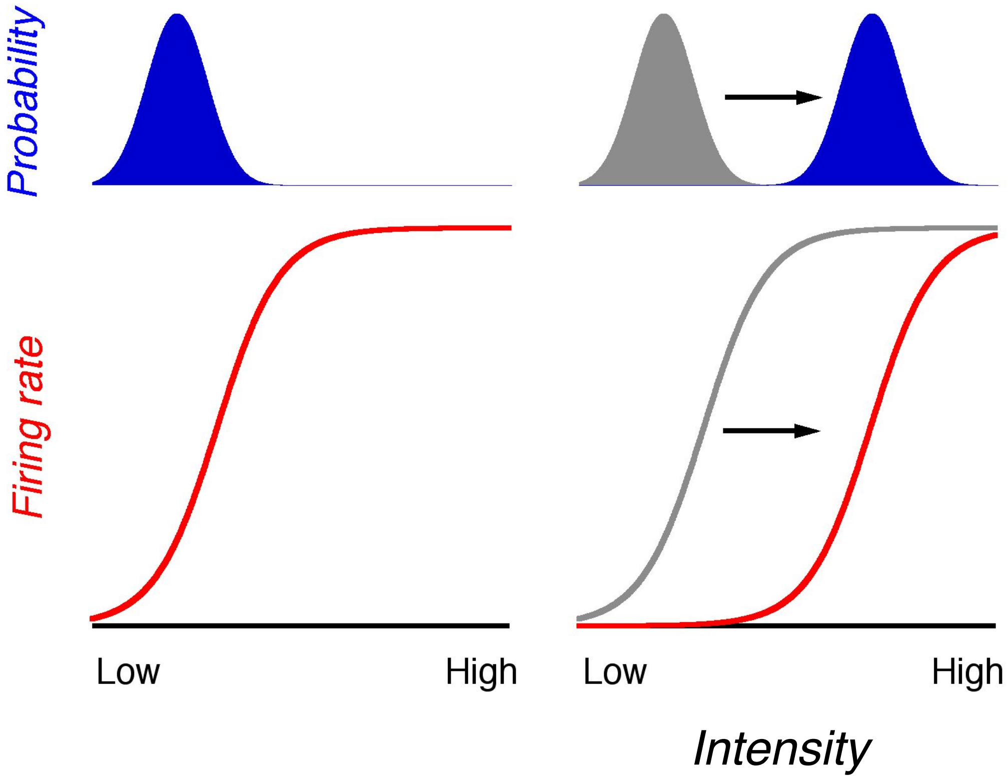


Adaptation to the average input

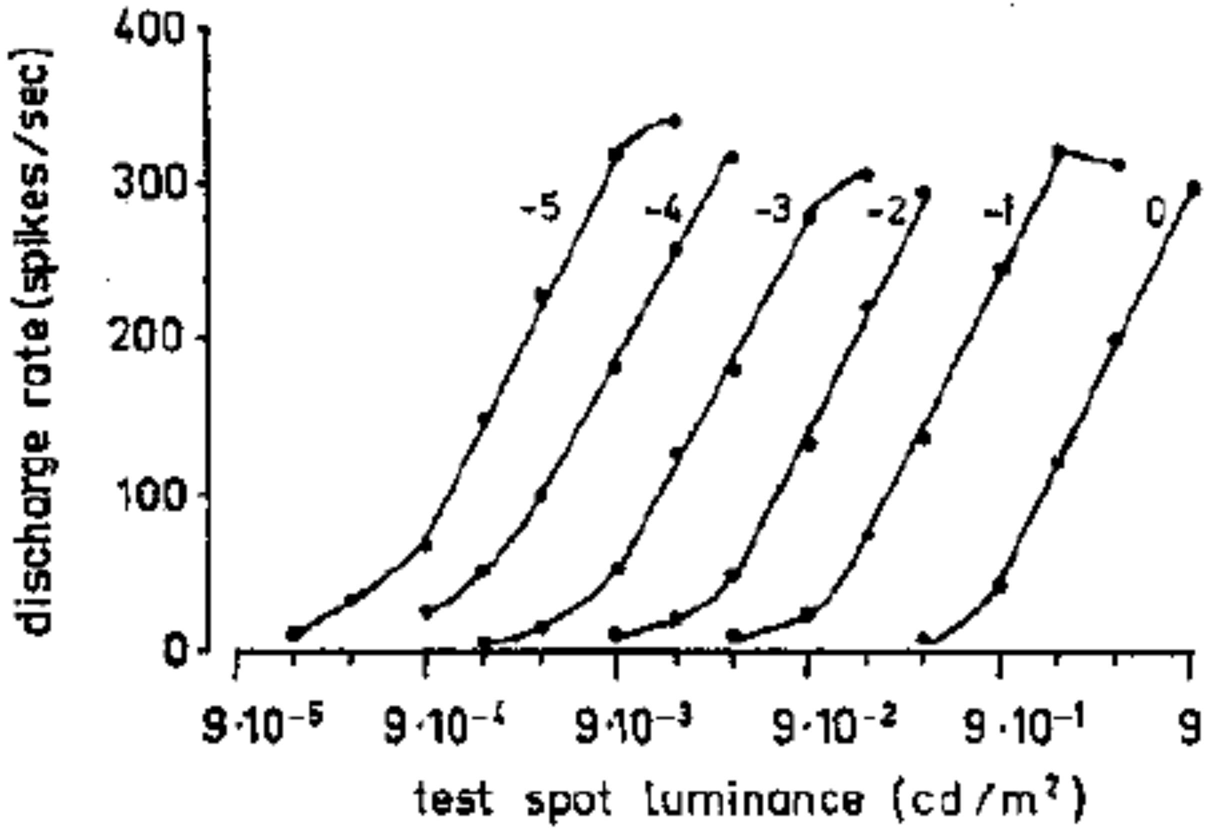


Adaptation to the average input

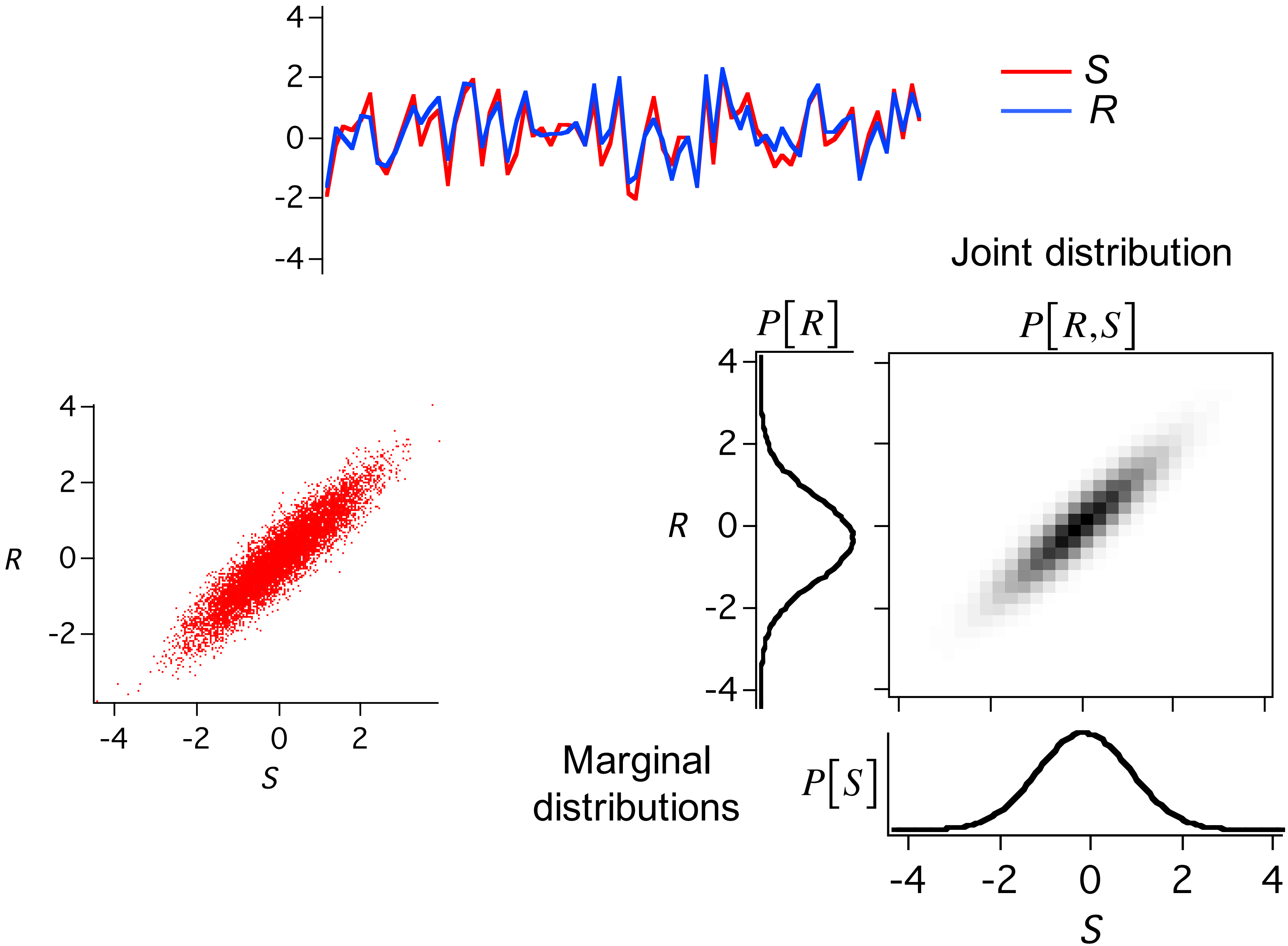
Light adaptation



Ganglion cell response curves shift to the mean light intensity



Sakmann and Creuzfeldt, Scotopic and mesopic light adaptation in the cat's retina (1969)



S
R
Joint distribution
P[R]
P[R,S]
Marginal distributions
P[S]
4
2
0
-2
-4

A Mathematical Theory of Communication

Claude Shannon (1948)

What is information?

Entropy*

A measure of uncertainty of a random variable in bits. The maximum possible amount of information there is to be learned from a variable.

$$H(X) = -\sum_i P[x_i] \log P[x_i]$$

Entropy of a fair coin =
- 1/2 log(1/2) - 1/2 log(1/2) = 1 bit

of an unfair coin =
- 3/4 log(3/4) - 1/4 log(1/4) = ~0.8 bits

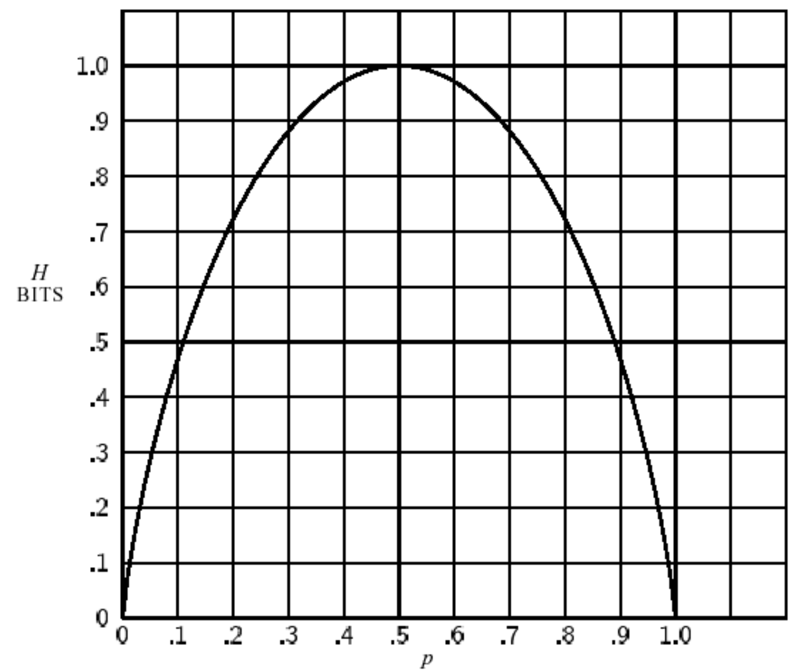


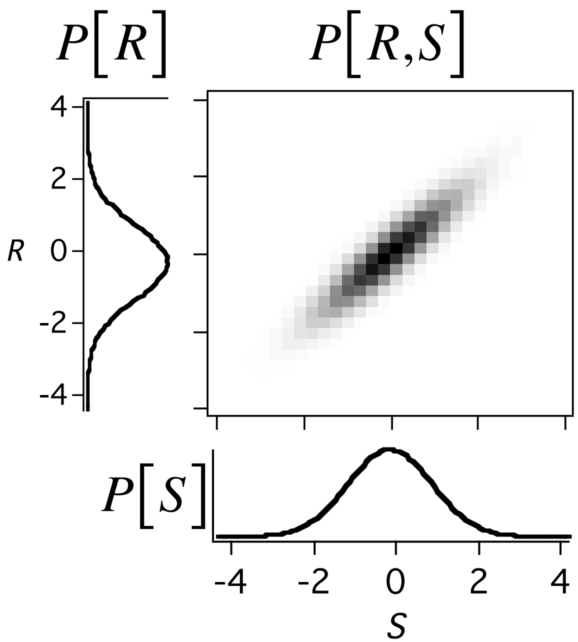
Fig. 7—Entropy in the case of two possibilities with probabilities p and $(1-p)$.

*By analogy to entropy in statistical mechanics,

$$S = k \log W$$

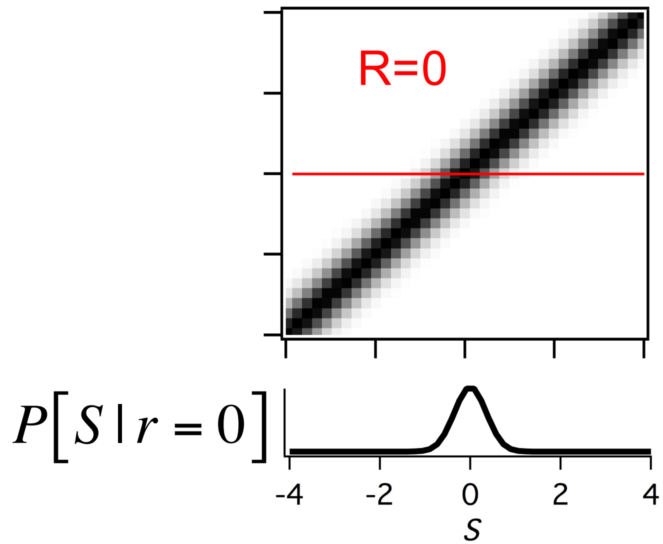
k: Boltzmann constant W: Number of possible microscopic states

Information is a reduction in entropy



Conditional distribution

$$P[S \mid R] = P[R,S] / P[R]$$



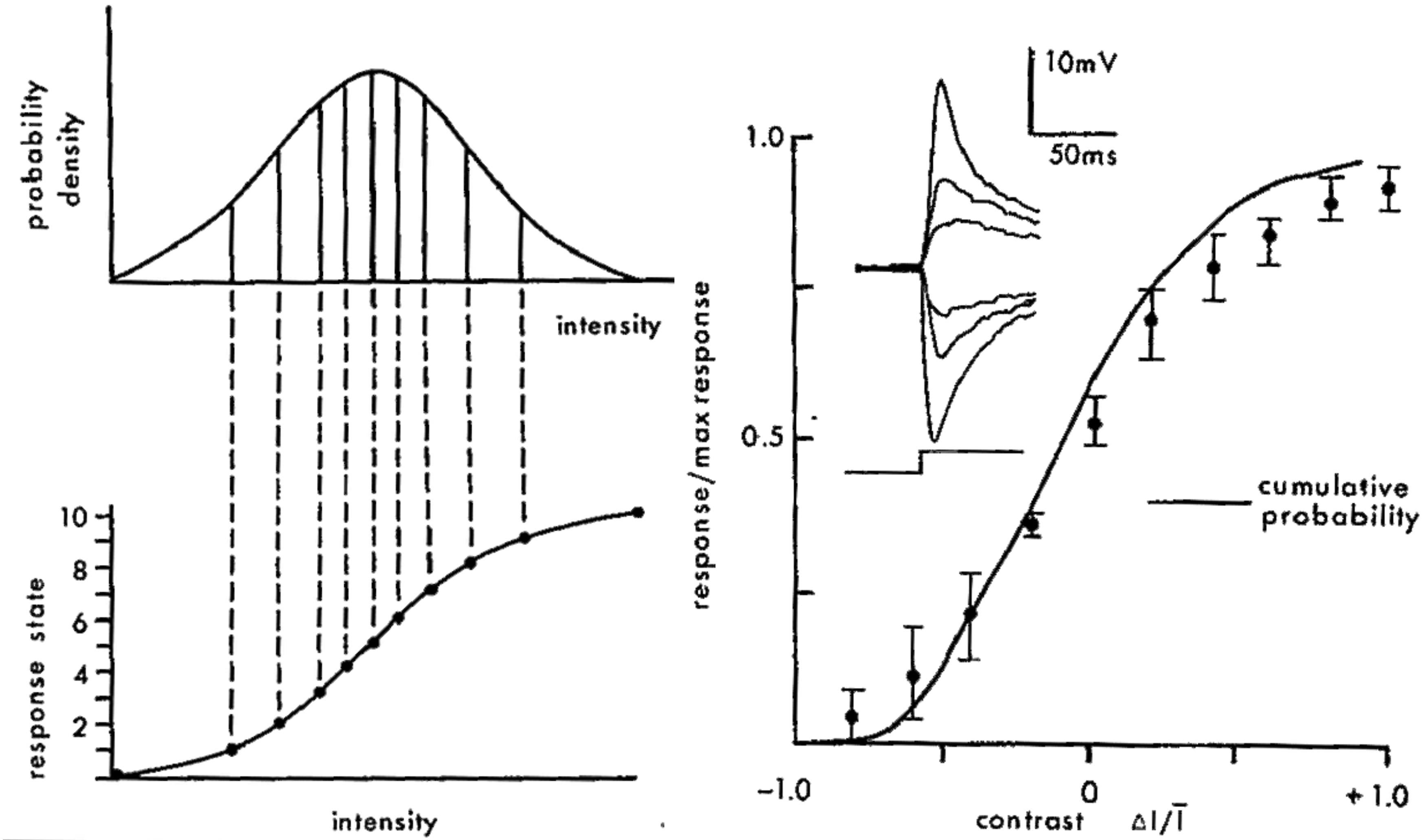
Conditional entropy

$$H(S \mid R) = -\sum_s \sum_r P(r,s) \log\left(P(s \mid r)\right)$$

Mutual information

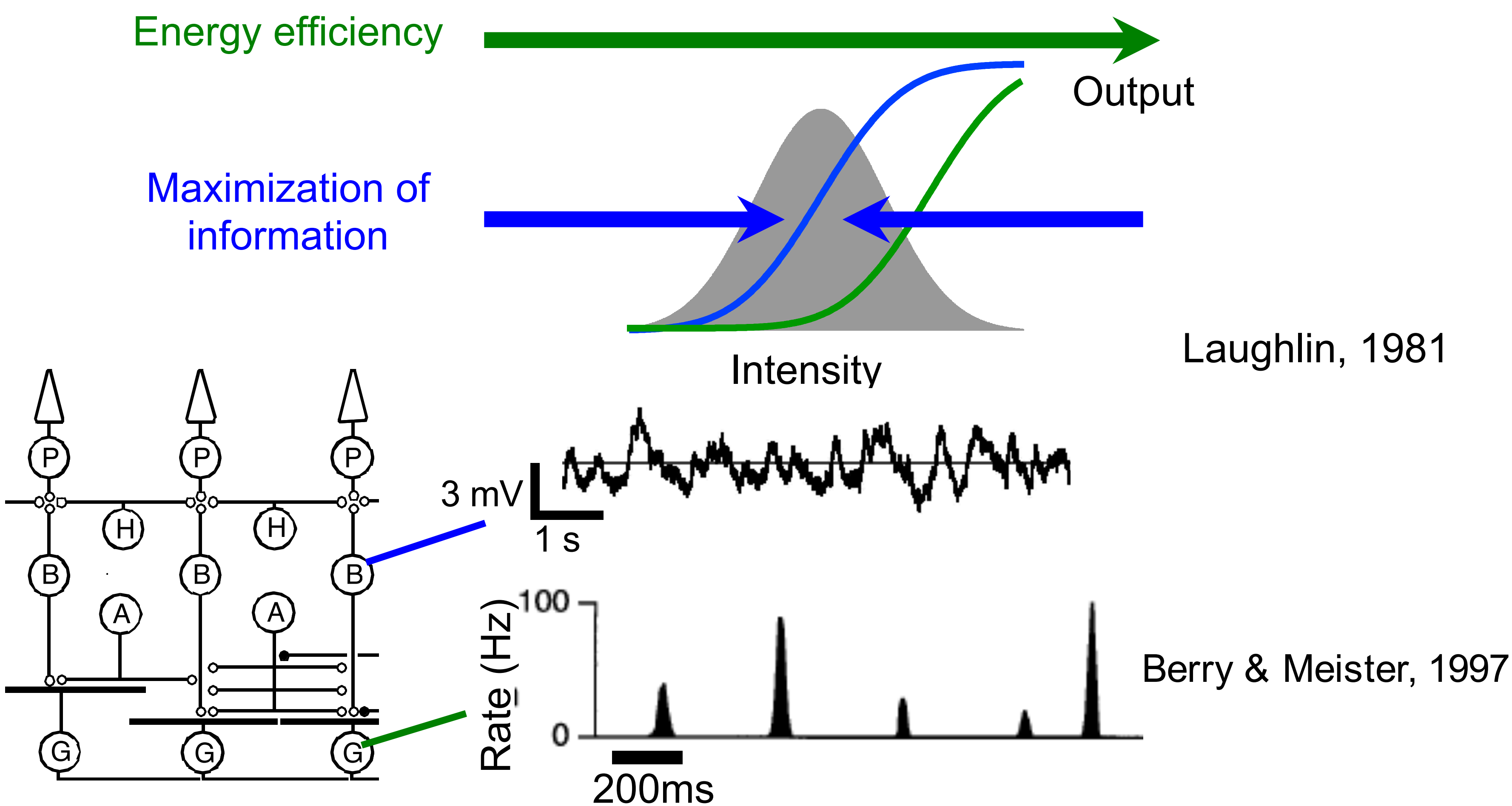
A measure, in bits, of how much information is conveyed by one random variable about another random variable. It is equal to the entropy minus the conditional entropy.

$$I(S;R) = H(S) - H(S \mid R) \qquad I(R;S) = I(S;R)$$

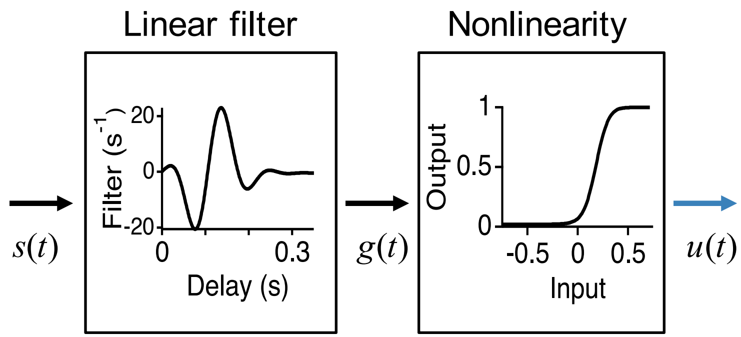


Simon Laughlin, A simple coding procedure enhances a neuron's information capacity Z. Naturforsch, 36c: 910-912 (1981)

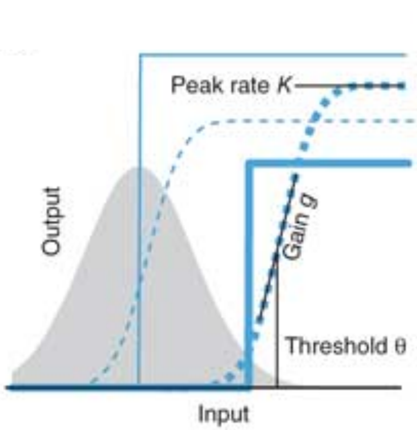
Tradeoff of information and energy efficiency

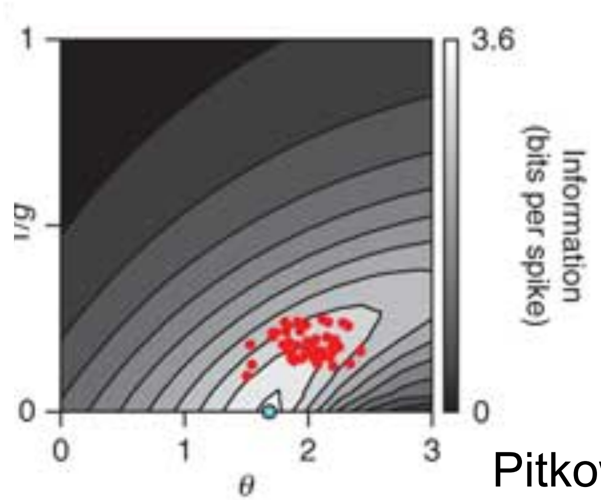


Given a rate constraint, the retina maximizes information



Poisson-like noise & rate constraint





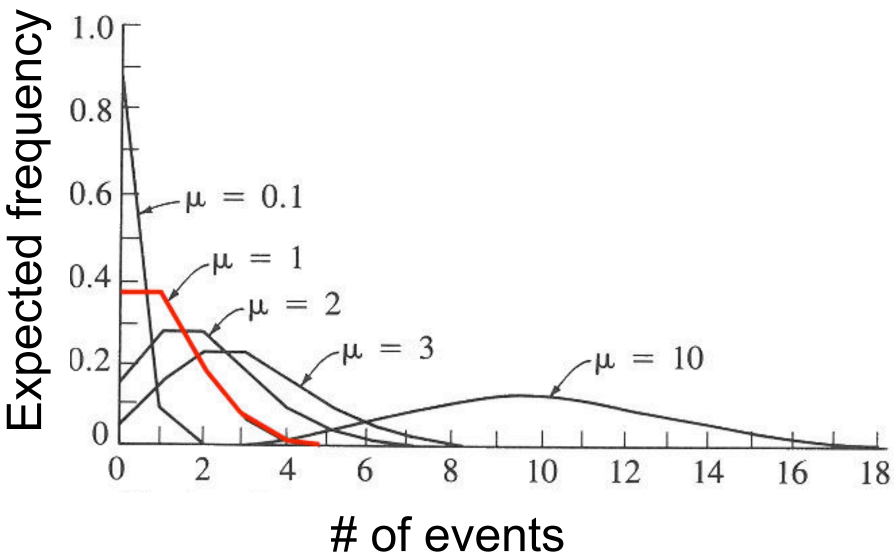
Pitkow & Meister, 2012

Events with Poisson statistics $\mathrm{P}[n,\mu]$

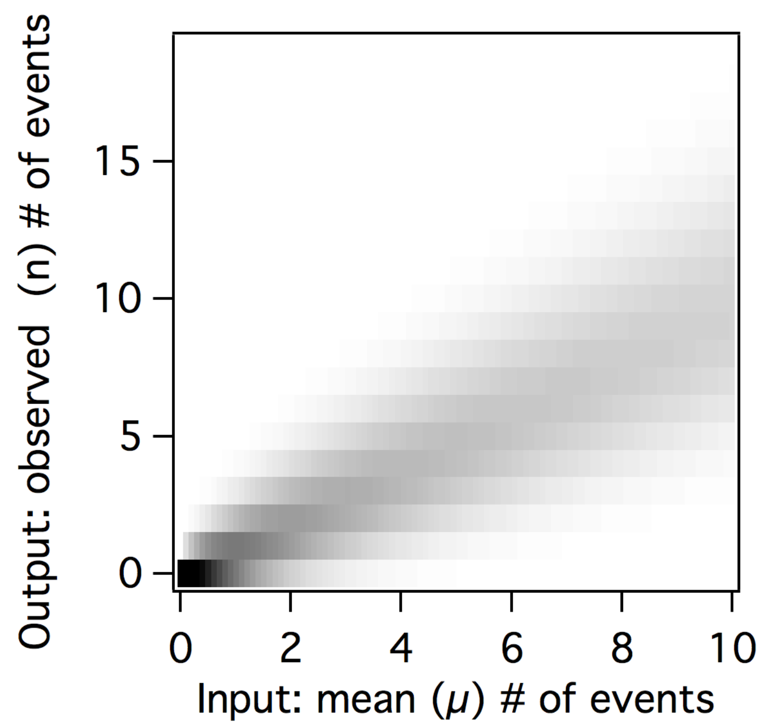
$$\frac{e^{-\mu}\mu^{n}}{n!}$$

μ = mean # of events in a time interval

n = events in a time interval

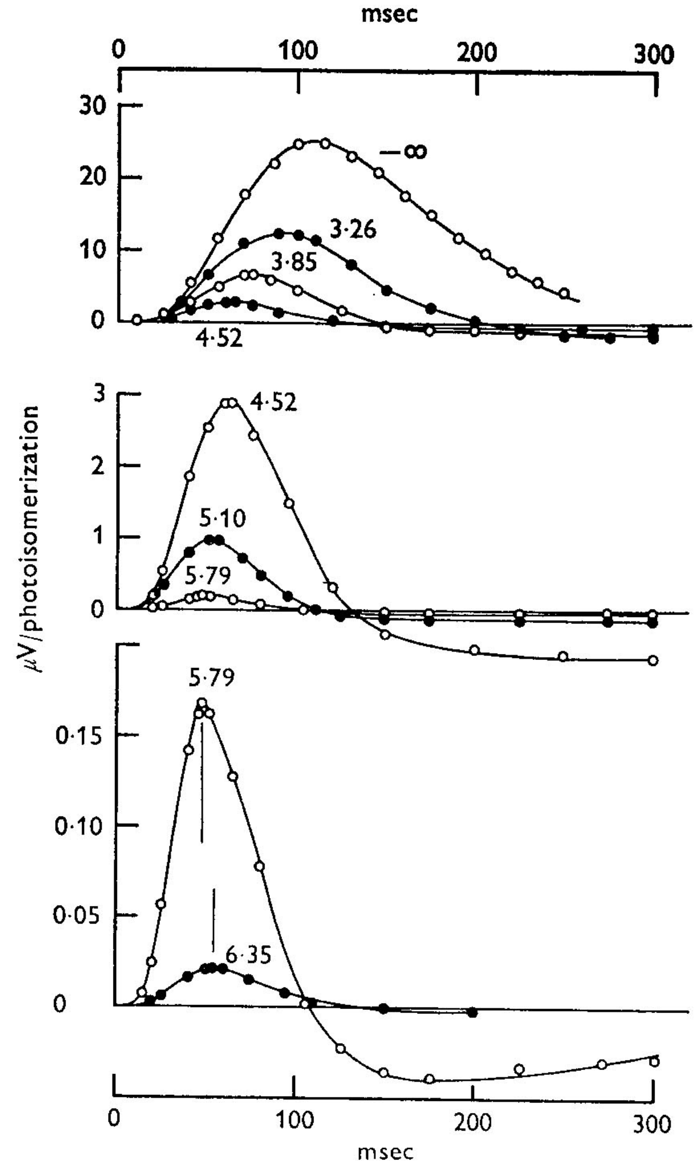


Joint probability distribution $\mathrm{P}[n,\mu]$

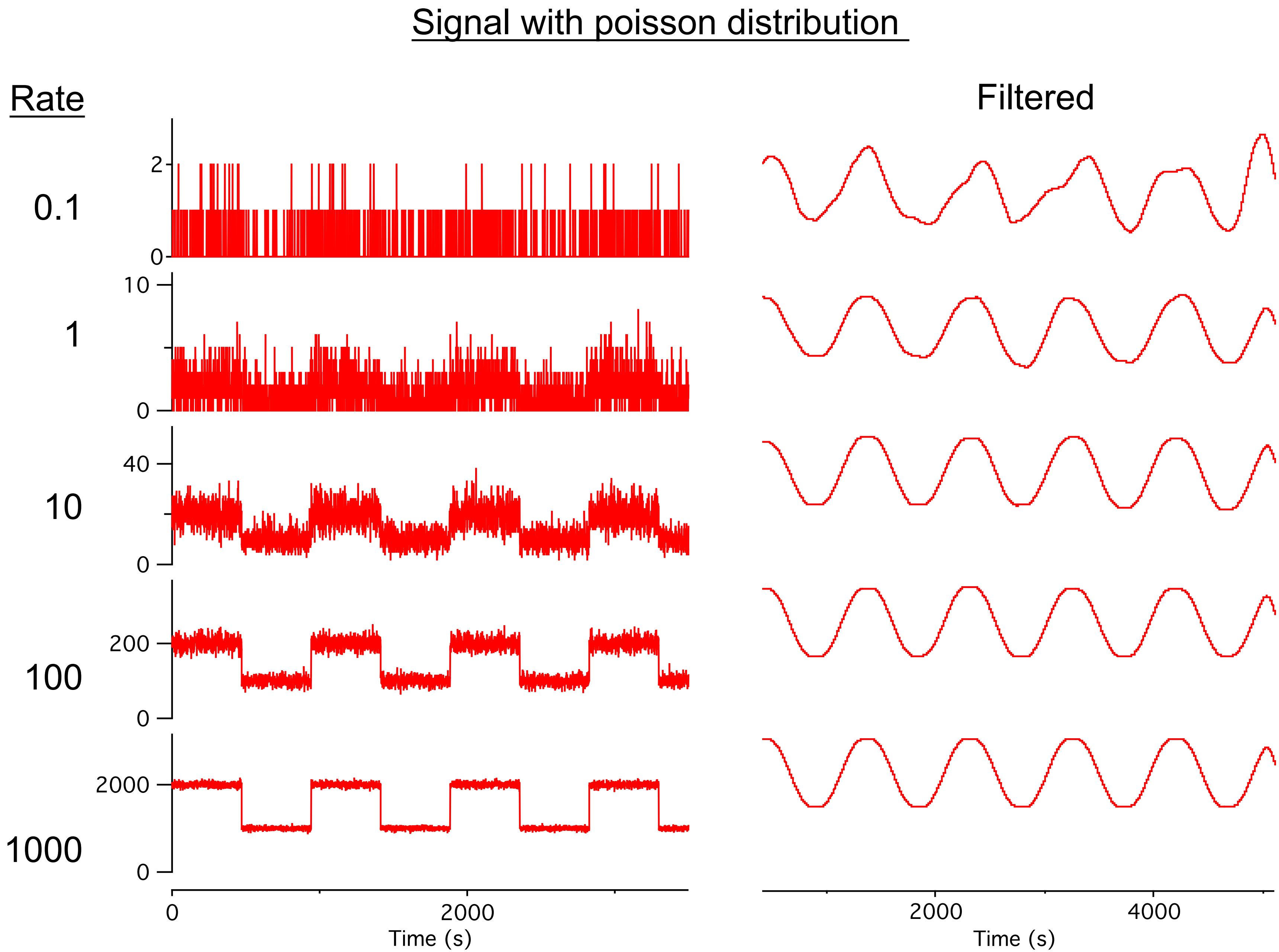


variance=mean=μ

Turtle Cones: Sensitivity *and Kinetics* change with mean luminance



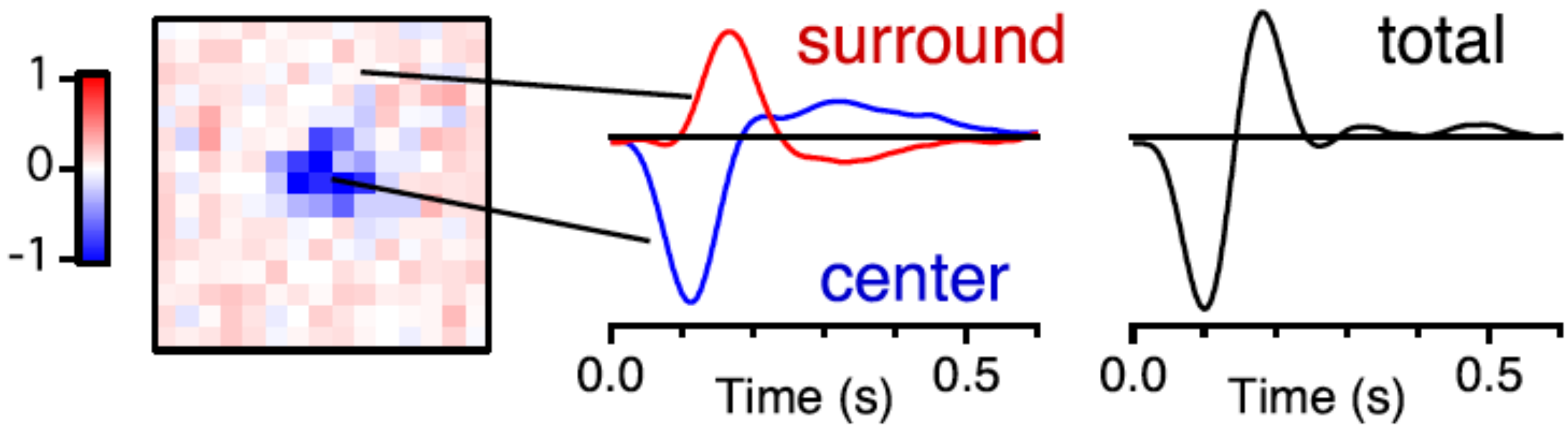
Baylor & Hodgkin 1974



Signal with poisson distribution
Rate
Filtered
0.1
1
10
100
1000
2
0
10
0
40
0
200
0
2000
0
0
2000
Time (s)
2000
4000
Time (s)

What receptive field maximizes information transmission?

Retinal bipolar cell receptive field

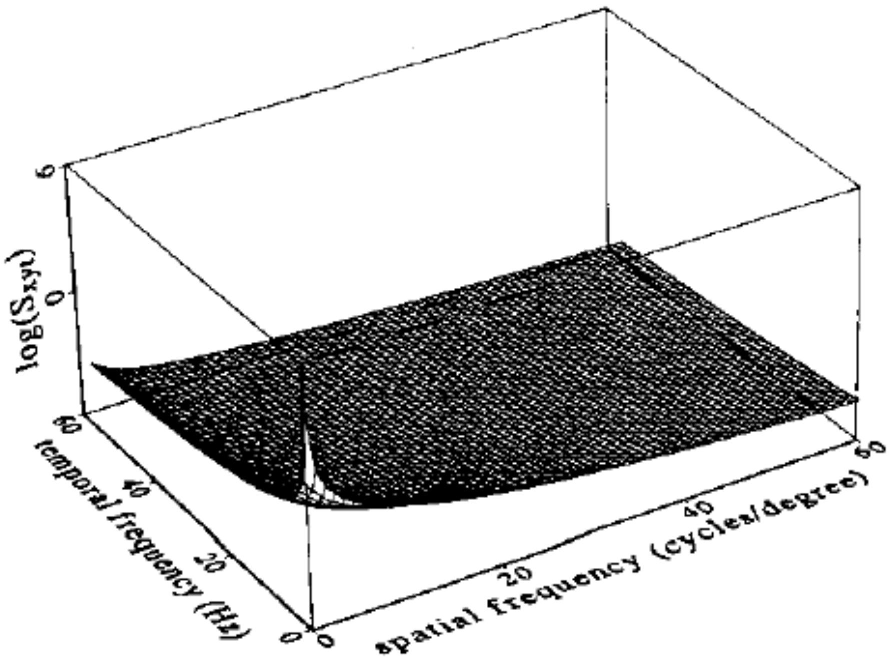


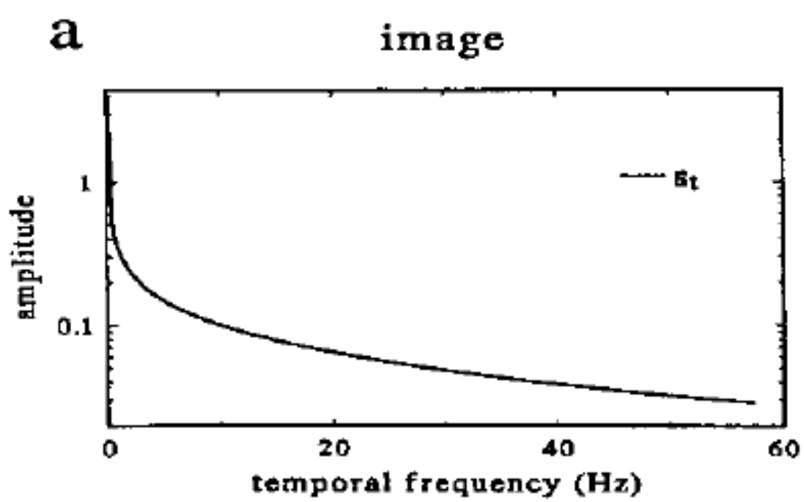
Baccus, Olveczky, Manu & Meister, 2008

Theory of maximizing information in a noisy neural system

‘Efficient Coding’ - Horace Barlow

Natural visual scenes are dominated by low spatial and temporal frequencies





J.H. van Hateren. Real and optimal neural images in early vision. *Nature* 360:68-70 (1992)

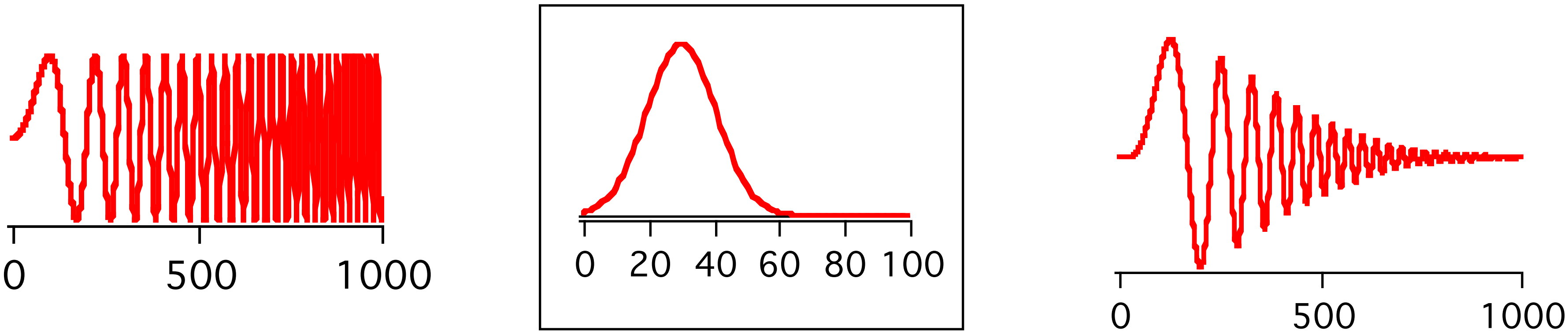
J.H. van Hateren, Spatiotemporal contrast sensitivity of early vision. *Vision Res.*, 33:257-67 (1993)

Linear filter and frequency response

Stimulus

Filter

Response



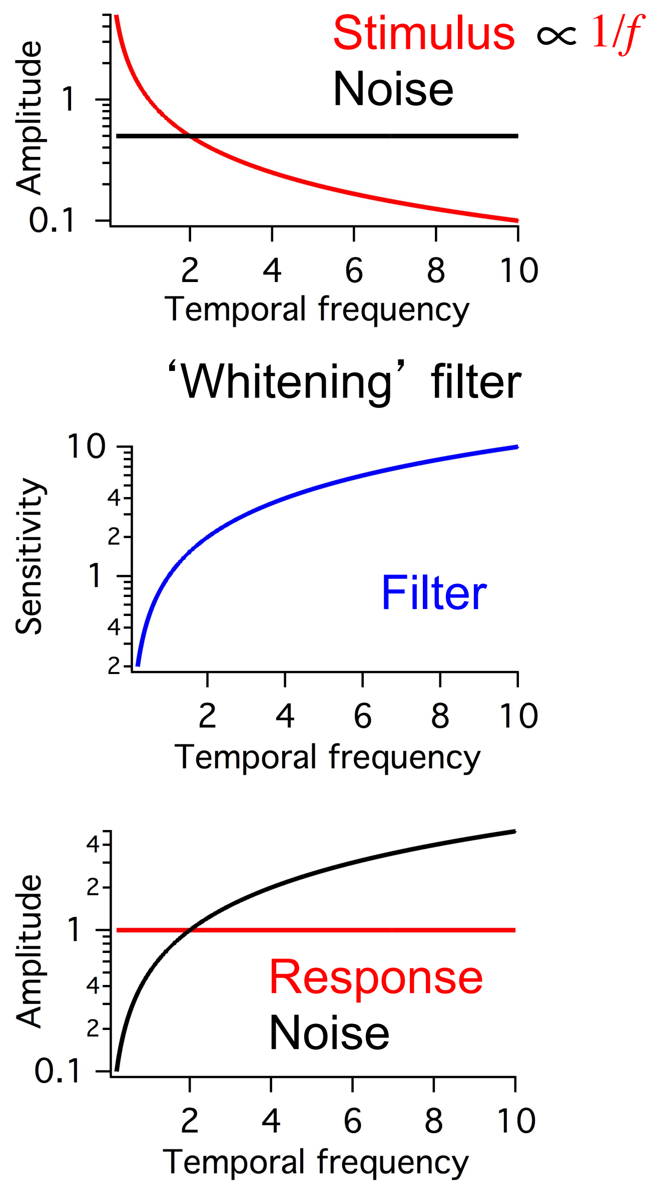
Convolution theorem

$$h(t) = f(t) * g(t) \quad \Leftrightarrow \quad \tilde{h}(\omega) = \tilde{f}(\omega)\tilde{g}(\omega)$$

a convolution in the time domain

is a simple product in the frequency domain

Optimal filter whitens but also cuts out noise



Filter to whiten in the presence of noise

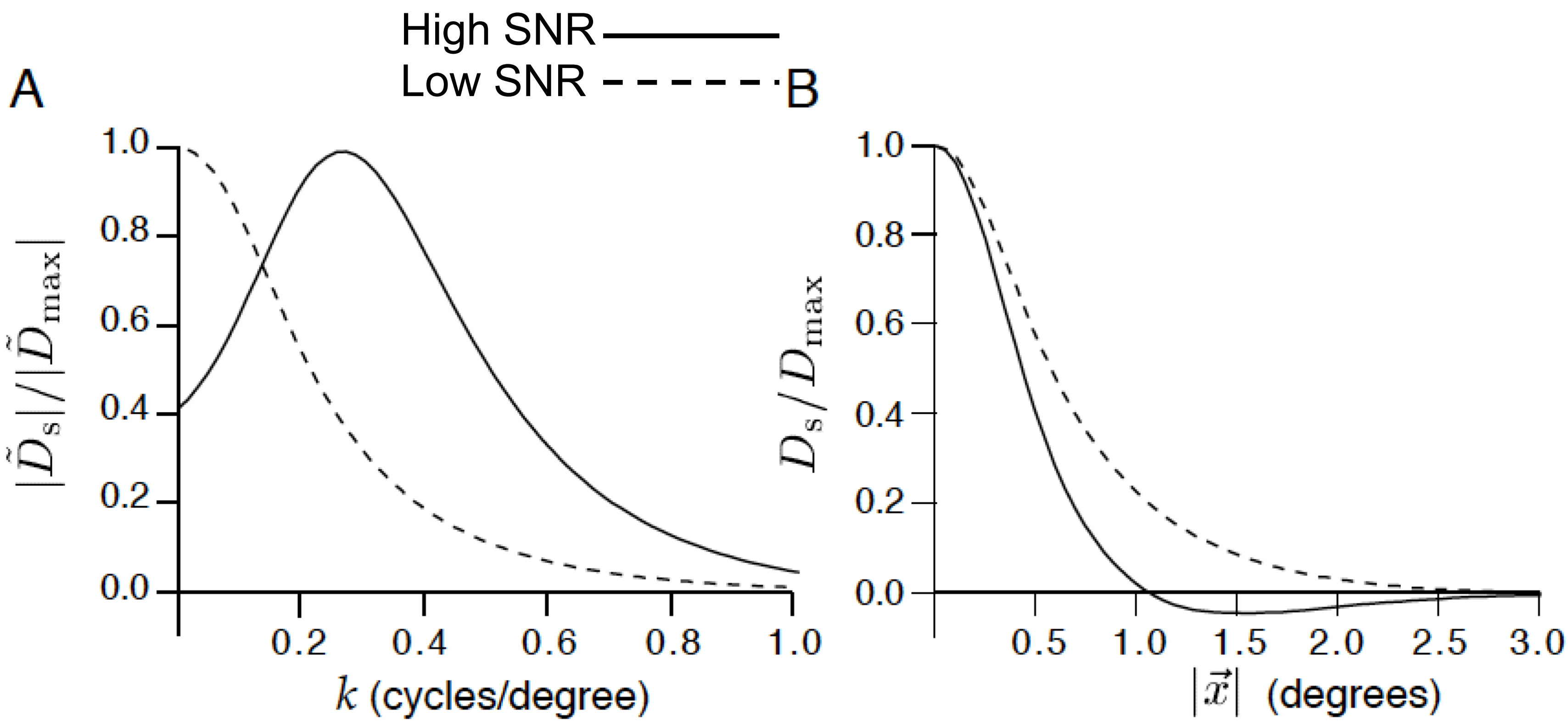
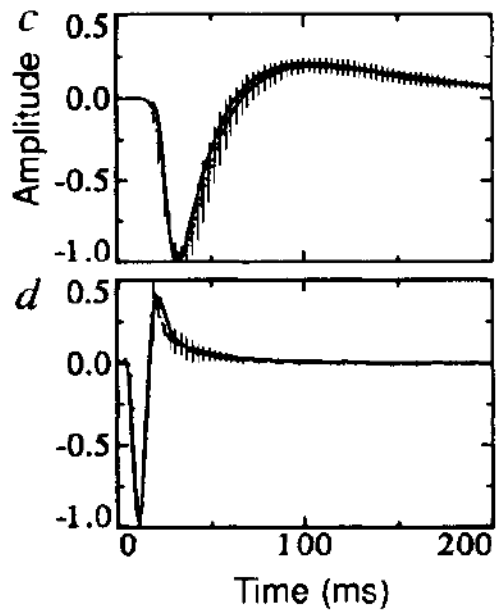


Figure 4.3: Receptive field properties predicted by entropy maximization and

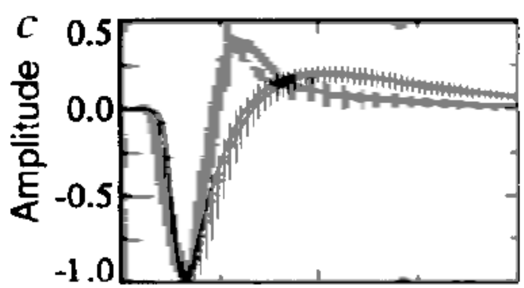
Theory of maximizing information in a noisy neural system

Filter of fly Large Monopolar Cells,
2nd order visual neuron



Low background intensity
Integrates over time
(real and theoretical optimum)

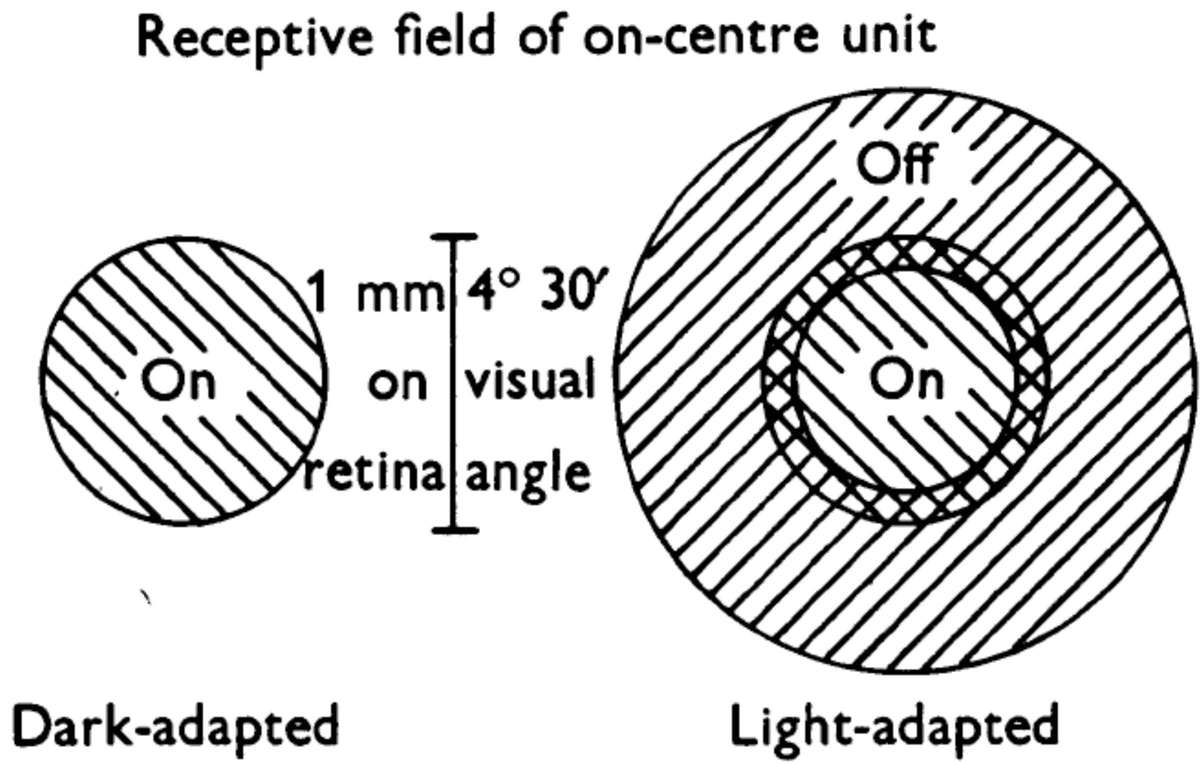
High background intensity
Emphasizes change, is more differentiating
(real and theoretical optimum)



Both, scaled in time to the first peak

J.H. van Hateren. Real and optimal neural images in early vision. *Nature* 360:68-70 (1992)

Spatial adaptation in retinal ganglion cells



Barlow, Fitzhugh & Kuffler (1957)

Theories of efficient coding:

An ideal encoder should use all output values with equal probability

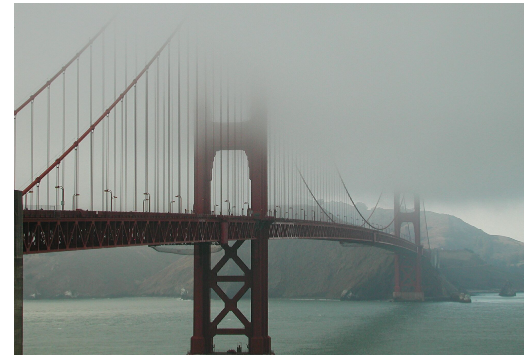
Low frequencies dominate in natural scenes

An efficient encoder should amplify higher frequencies more than low frequencies

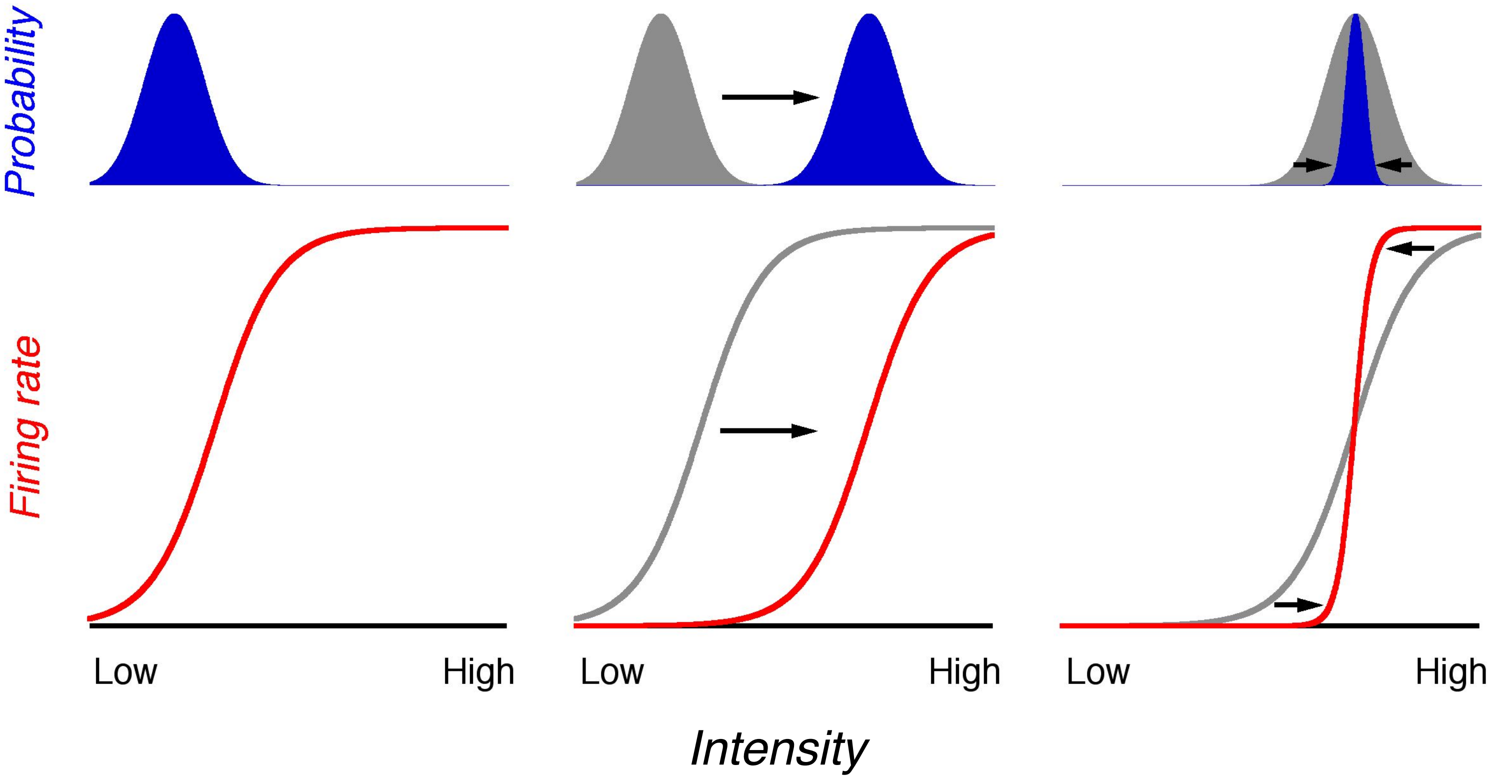
But when signals are more noisy, such as when the signal is weak, higher frequencies should be reduced, as they carry little information

Adaptation to mean and variance

Light adaptation



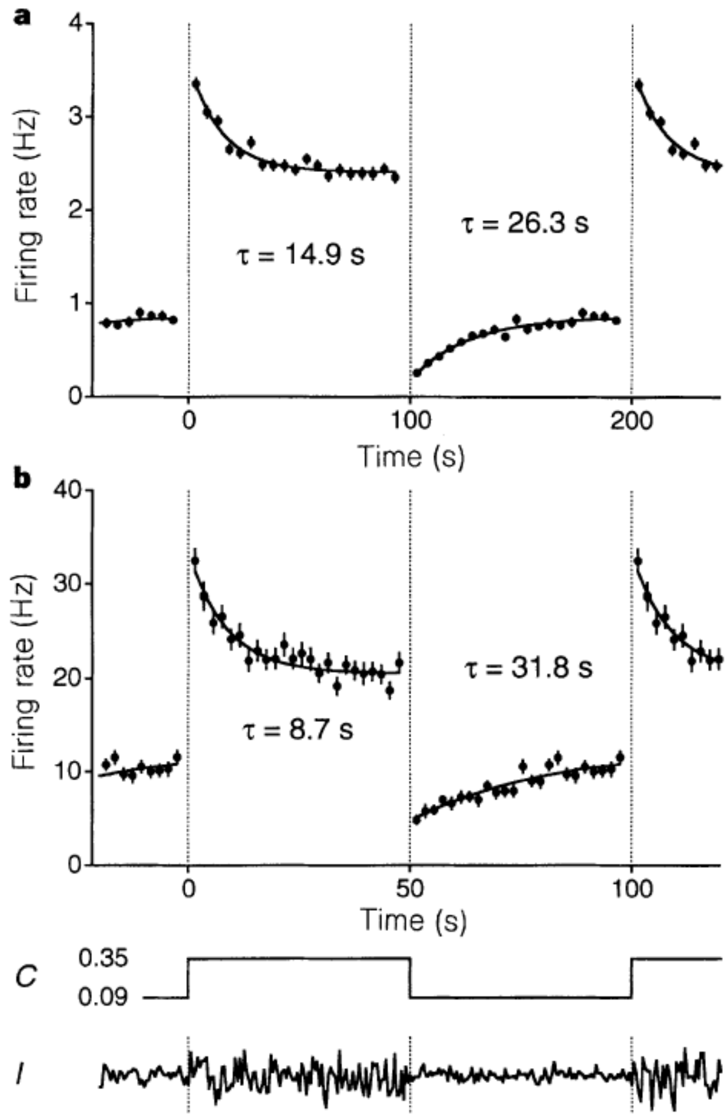
Contrast adaptation



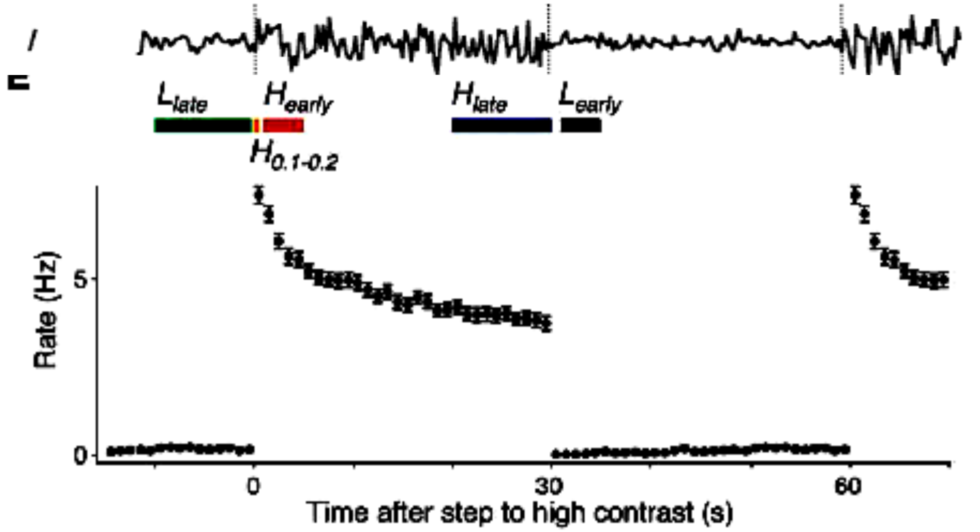
Retinal contrast adaptation

Salamander

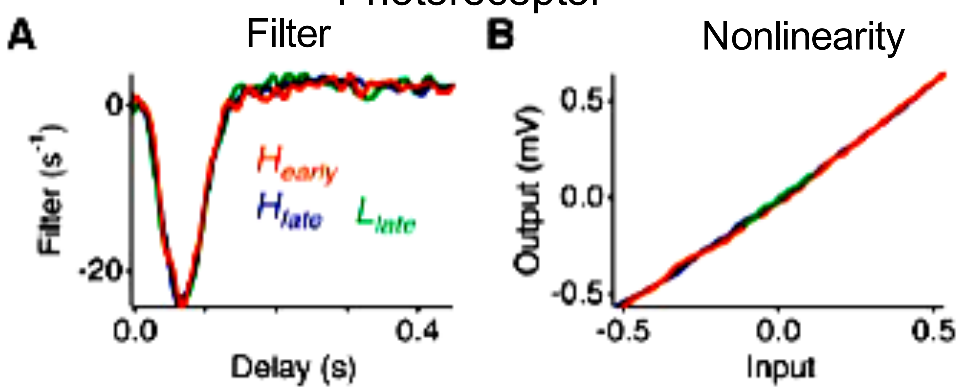
Rabbit



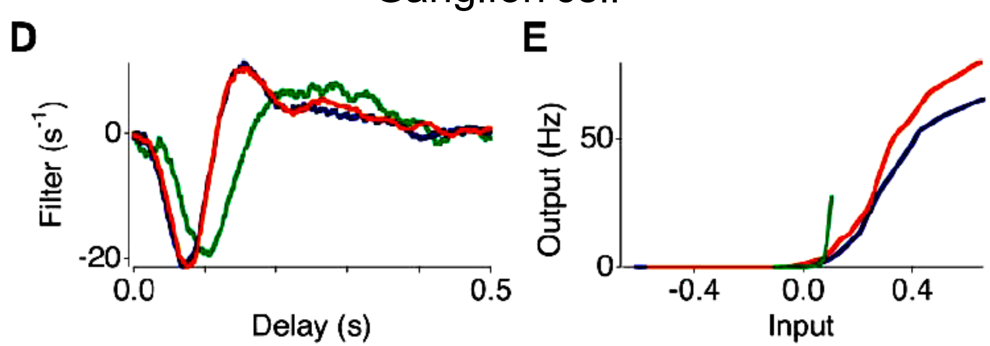
Smirnakis et al., Adaptation of retinal processing to image contrast and spatial scale. *Nature,* 386:69-73 (1997).



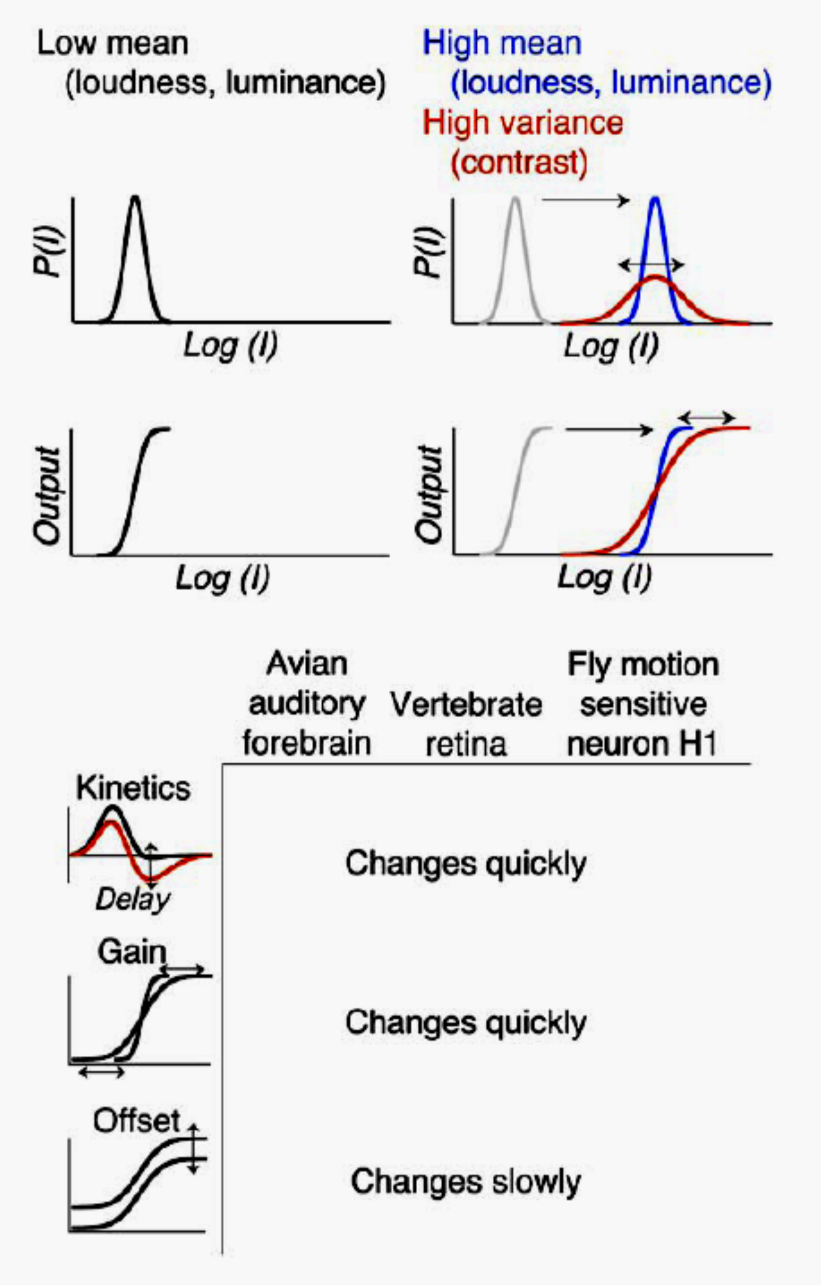
Photoreceptor



Ganglion cell



Baccus & Meister. Fast and slow contrast adaptation in retinal circuitry. (2002).



Low mean
(loudness, luminance)
High mean
(loudness, luminance)
High variance
(contrast)
P(I)
Log (I)
Output
Avian auditory forebrain
Vertebrate retina
Fly motion sensitive neuron H1
Kinetics
Delay
Changes quickly
Gain
Changes quickly
Offset
Changes slowly

Adaptation to the mean and variance of signals are similar in a number of systems:

Kinetics change as quickly as the immediate response

Gain changes as quickly as the immediate response, and over longer timescales

Offset changes more slowly, typically in a homeostatic adjustment

These adaptive properties can be interpreted as avoiding saturation and maximizing information in the presence of noise