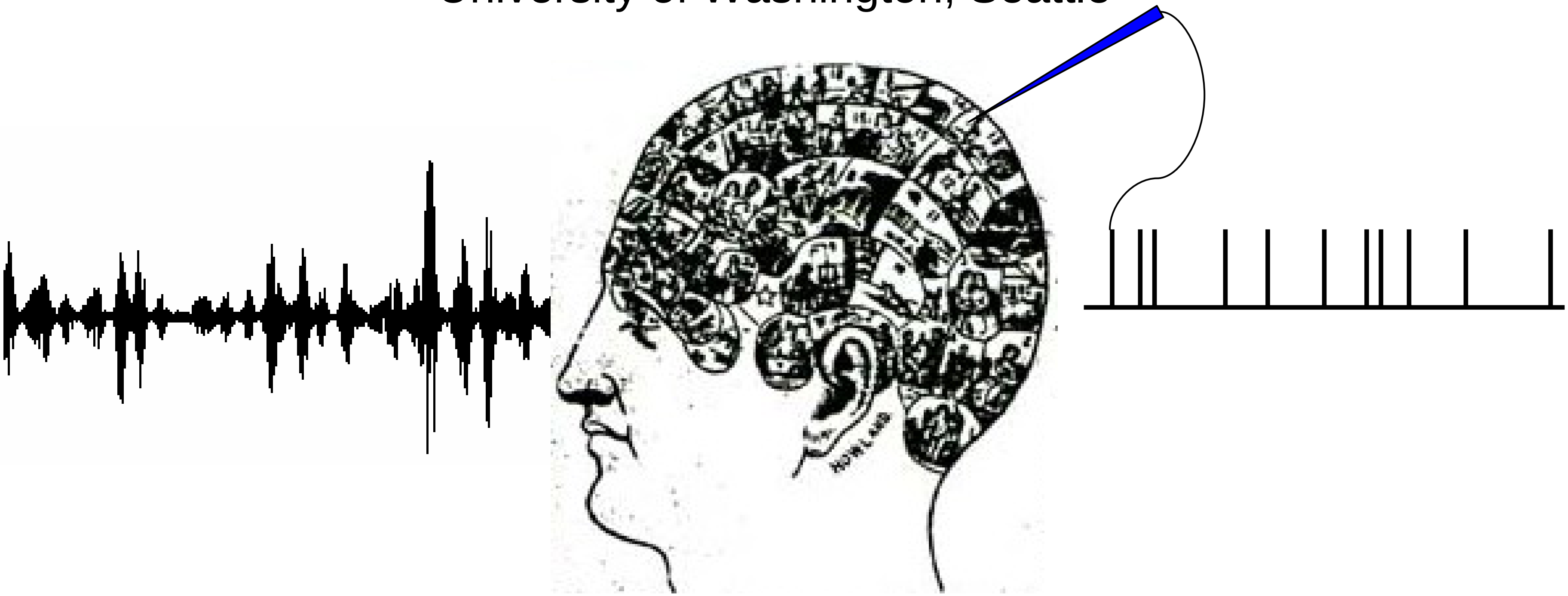
Adapting to the statistics of the environment

Adrienne Fairhall
Physiology and Biophysics
University of Washington, Seattle



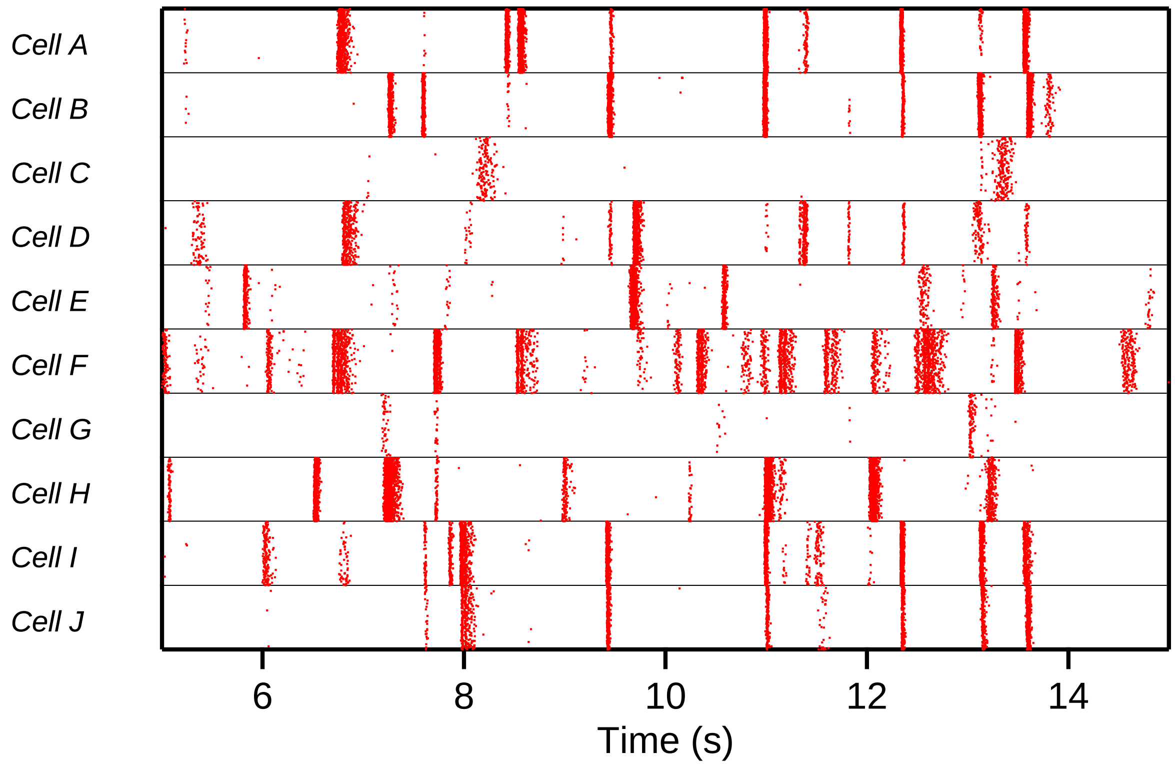
Sensory encoding

Multielectrode array recording from retinal ganglion cells during repeated viewing of a natural movie

Different neurons select and encode different features in the stimulus

What are the rules governing the encoding and representation of stimuli?

Are they fixed or can they be modulated by context?



Courtesy M. Berry and J. Puchalla

Natural stimuli show complex structure in space and time

1. Huge dynamic range: variations over many orders of magnitude


SIEGFRI


The New York Times
SIEGFRI


Thanks to Fred Soo!

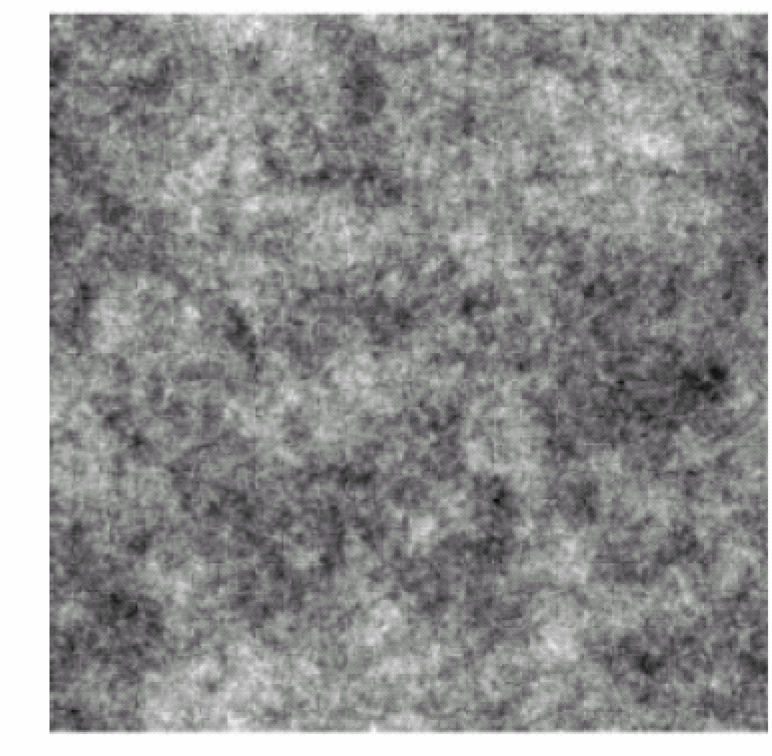
Natural stimuli show complex structure in space and time

1. Huge dynamic range: variations over many orders of magnitude

Natural stimuli show complex structure in space and time

1. Huge dynamic range: variations over many orders of magnitude

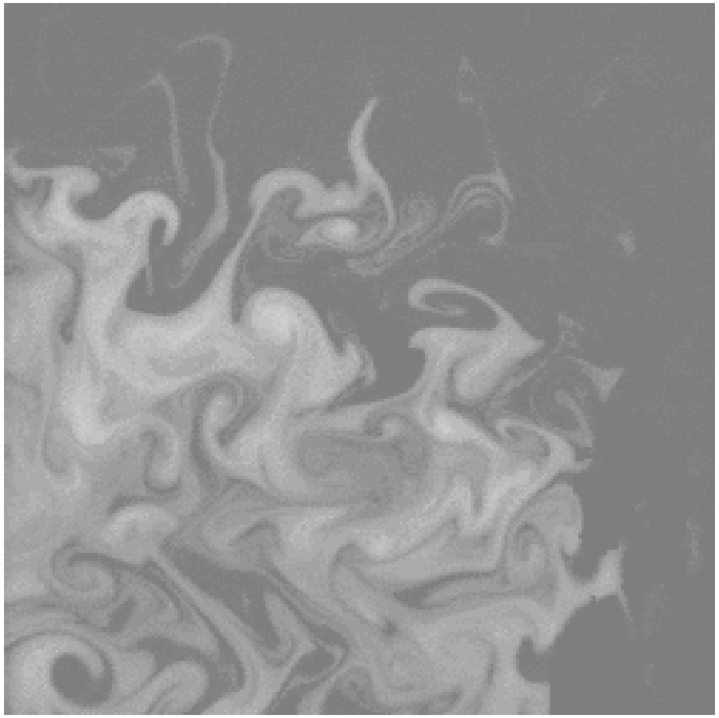
2. Power law scaling: highly nonGaussian



Natural stimuli show complex structure in space and time

1. Huge dynamic range: variations over many orders of magnitude

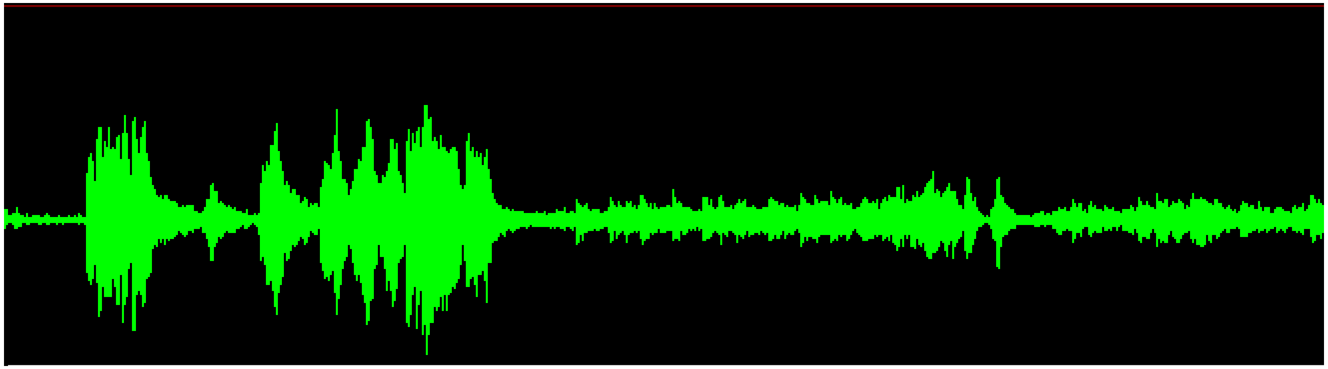
2. Power law scaling: highly nonGaussian

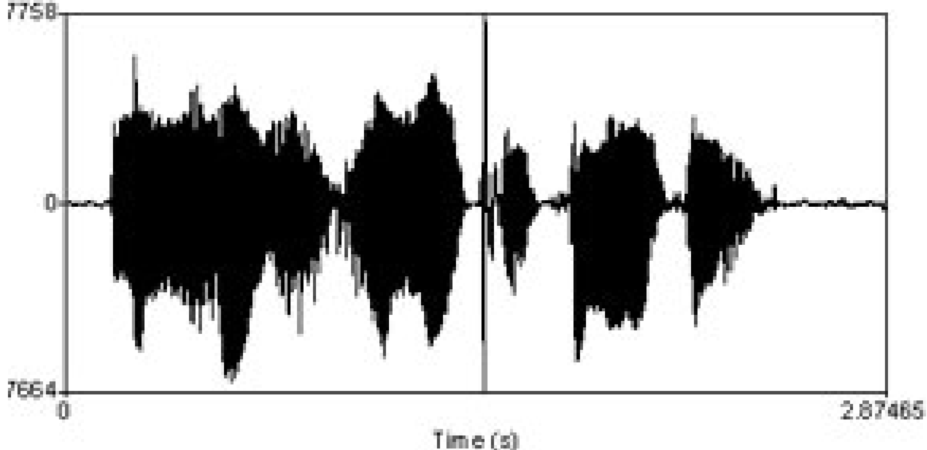


Natural stimuli show complex structure in space and time

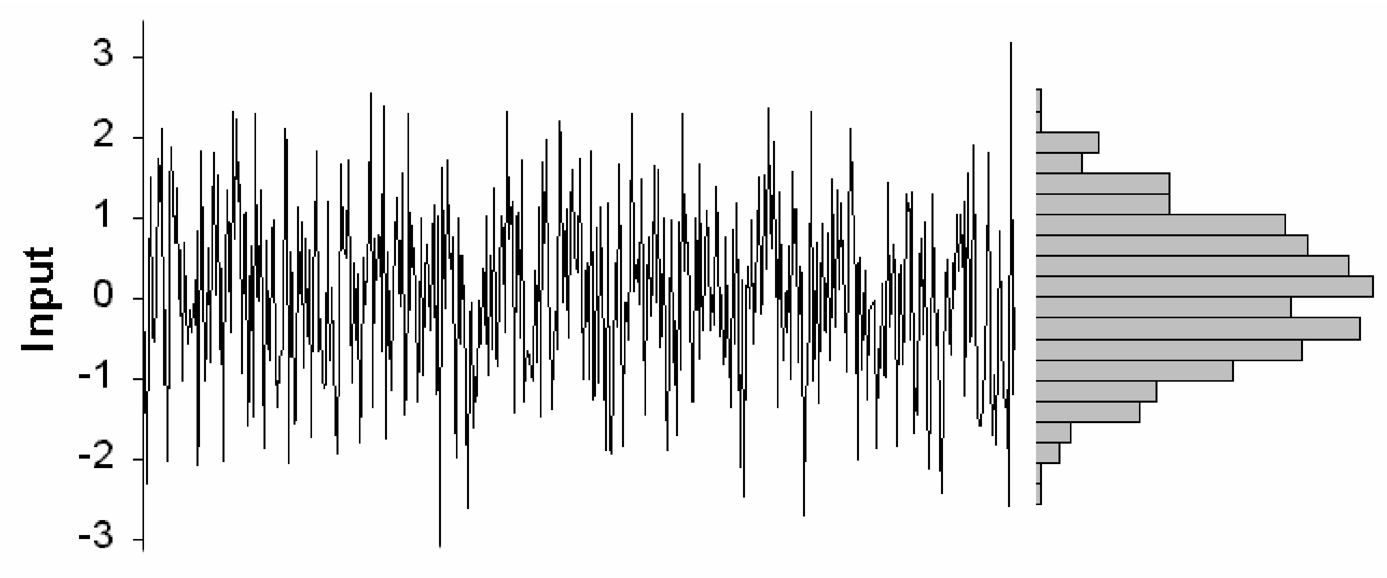
1. Huge dynamic range: variations over many orders of magnitude

2. Power law scaling: highly nonGaussian

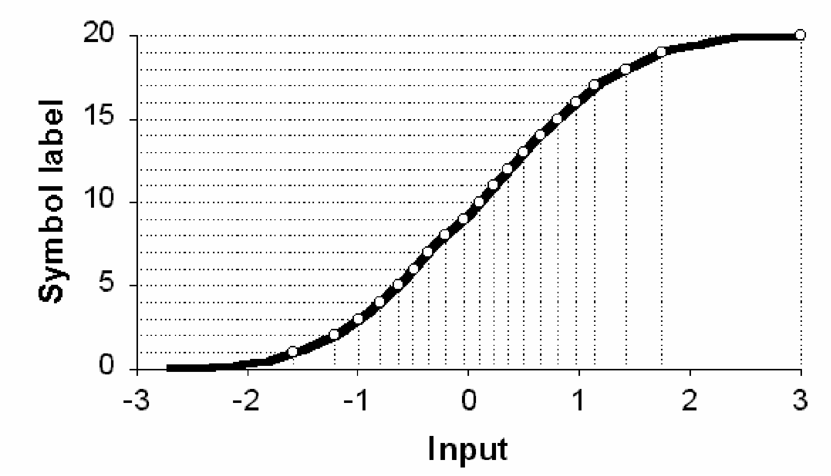




Efficient coding



Input/output function



Shape of the I/O function should be determined by the distribution of natural inputs

Optimizes information between output and input

Adaptation of the input/output relation

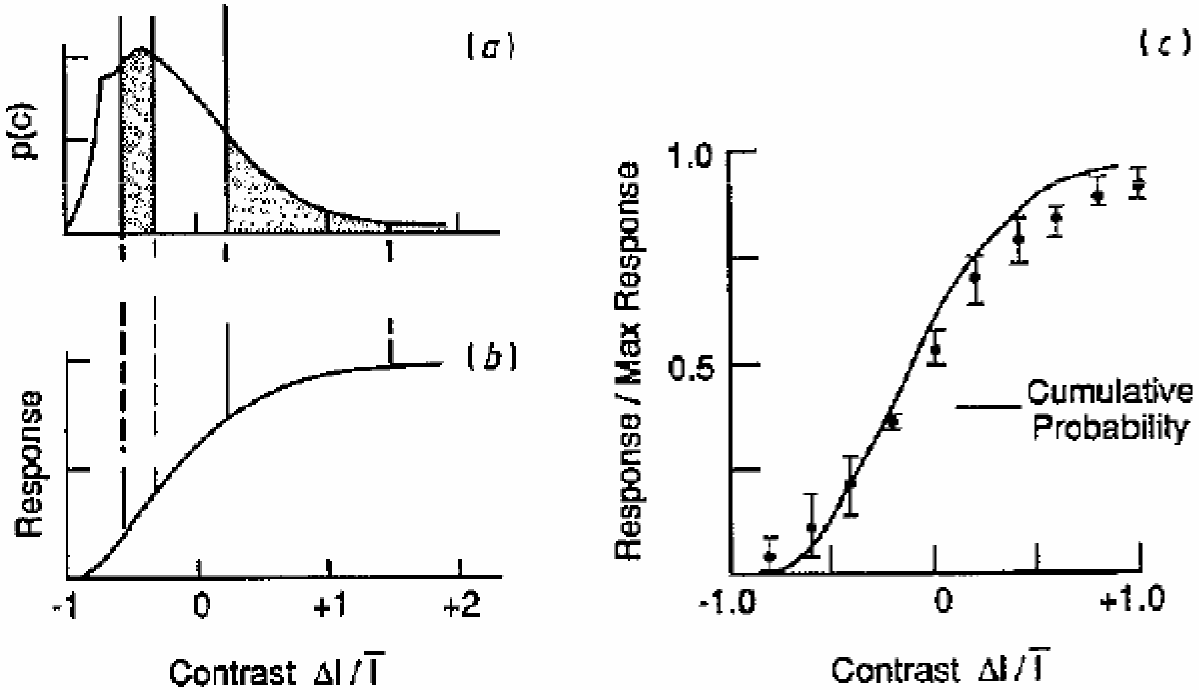
If we constrain the *maximum output*, the solution for the distribution of outp symbols is P(r) = constant = a.
Take the output to be a nonlinear transformation on the input: r = g(s).

$$P(r)dr = P(s)ds$$

$$\rightarrow \quad r = g(s) = \frac{1}{\alpha}\int_{-1}^{s} ds'\, P(s').$$

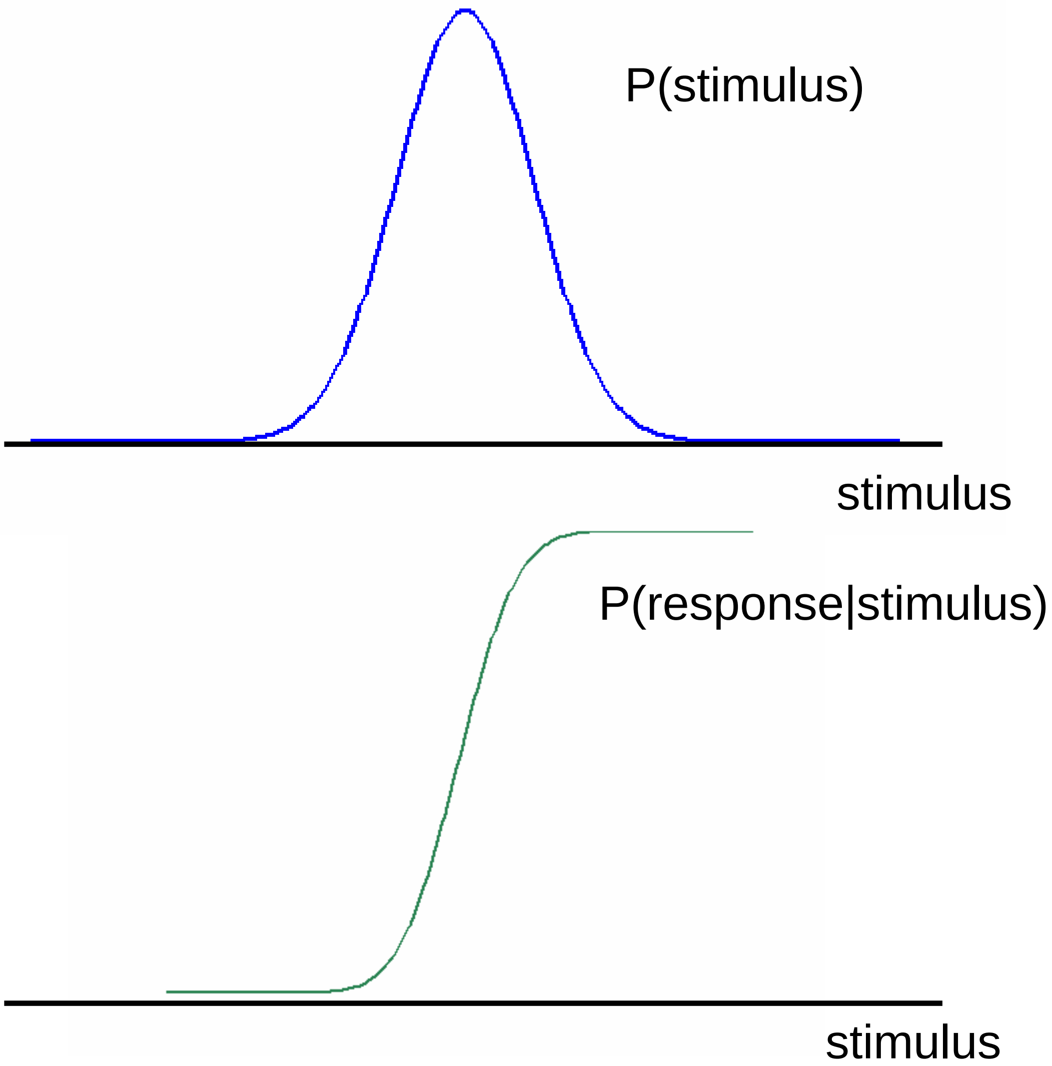
Fly LMC cells.

Measured contrast in natural scenes.



Laughlin '81.

Adaptation of the input/output relation



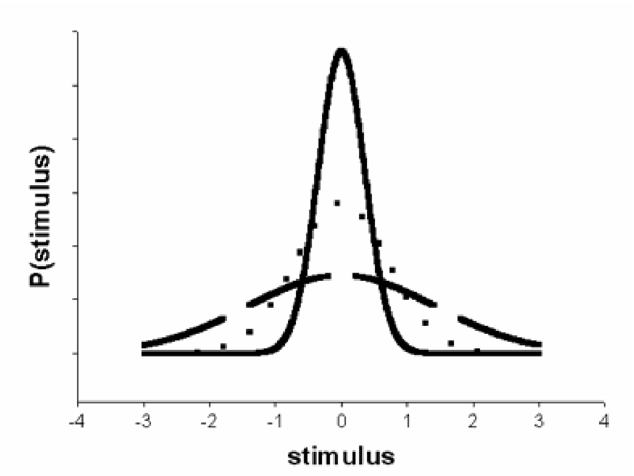
Changing the variance of the distribution changes the slope of the best input/output curve

One can show that an input/output curve which matches the variance of the stimulus maximizes mutual information between spikes and stimulus
(Brenner et al.2000)

What is the appropriate timescale for adaptation?

Contrast varies hugely in time.

Should a neural system optimize over evolutionary time or locally?



Natural image statistics: Bialek and Ruderman '94

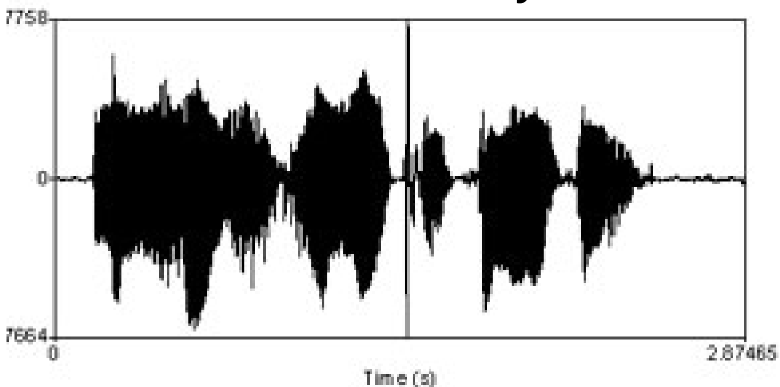
Distribution of contrast values in an image is highly nonGaussian, with long tails

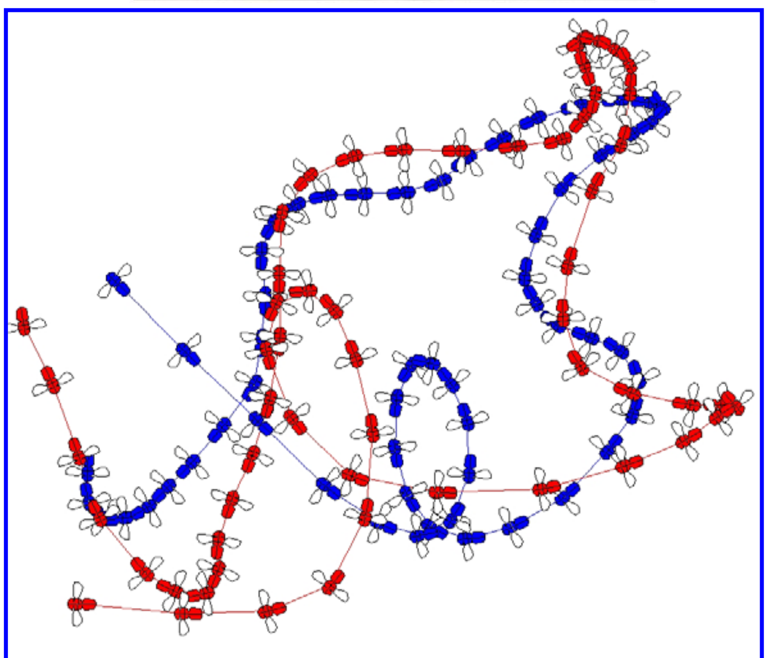
Separate local fluctuations and the local varia

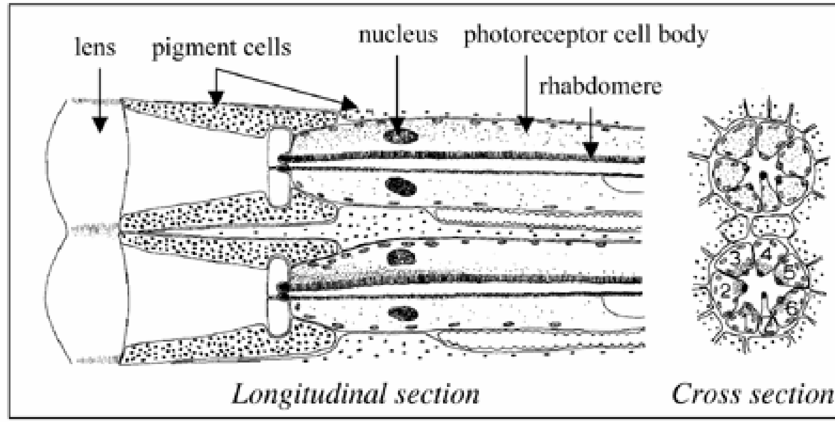
Local variance has a long-tailed distribution
Normalized contrast is nearly Gaussian

→Approximate structure:
  locally Gaussian, modulated by larger
  lengthscale variance envelope

Gaussian distributions are maximum entropy, efficiently encoded
If a neural system can normalize out local variance it can efficiently encode local fluctuations



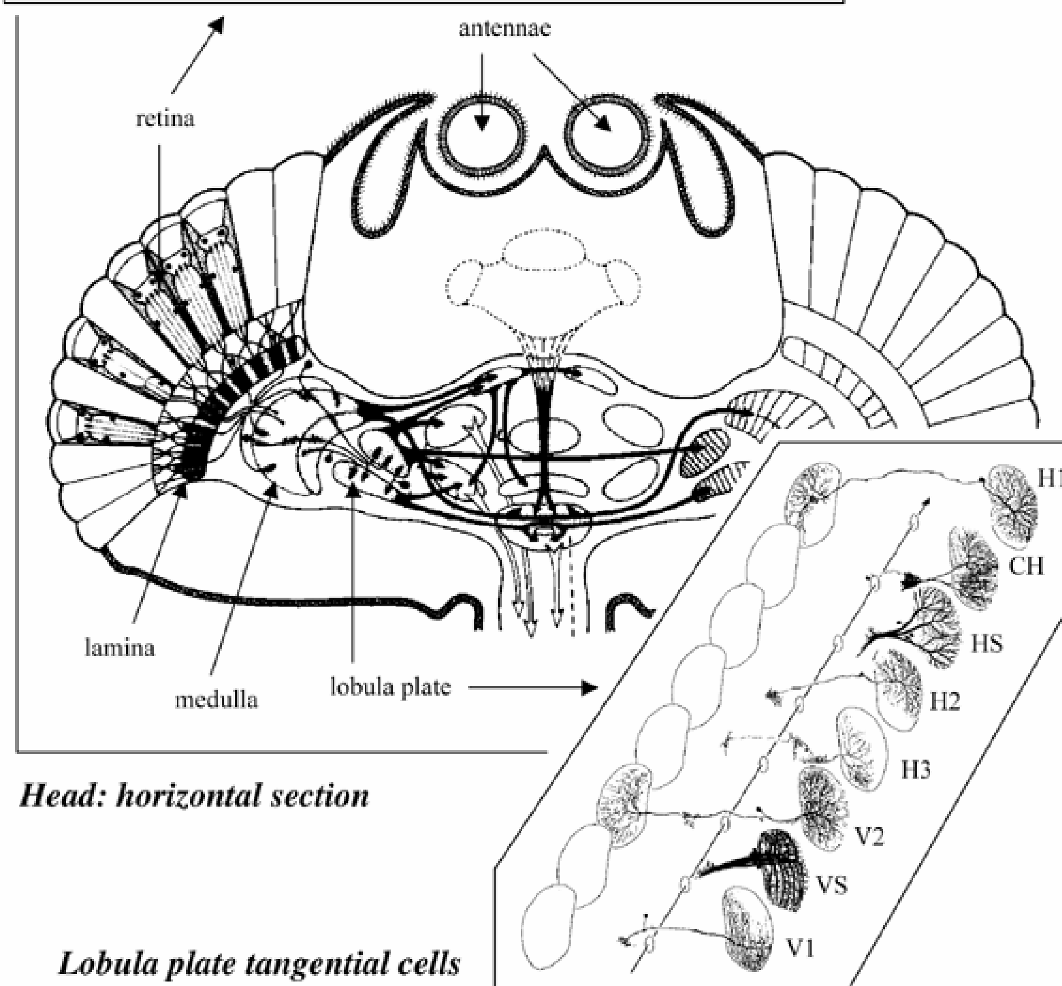




Structure of the retina

Dark adapted

Light adapted

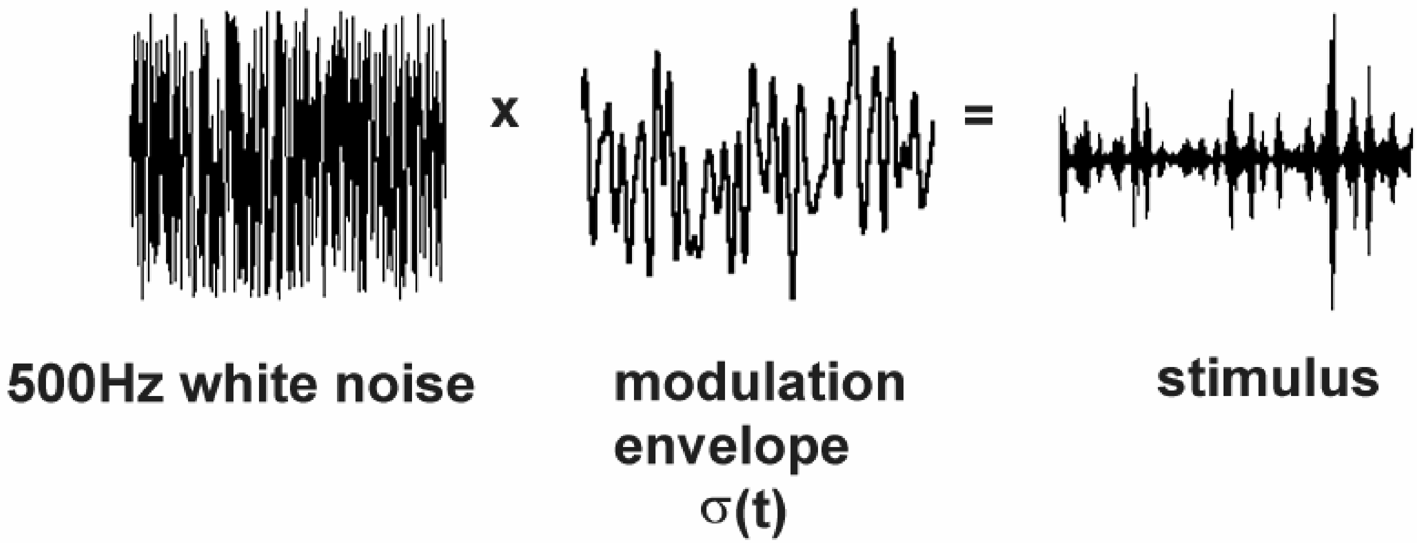


Head: horizontal section

Lobula plate tangential cells

de Ruyter van Steveninck and Bialek, FIG 3.

Approximating a “natural stimulus”

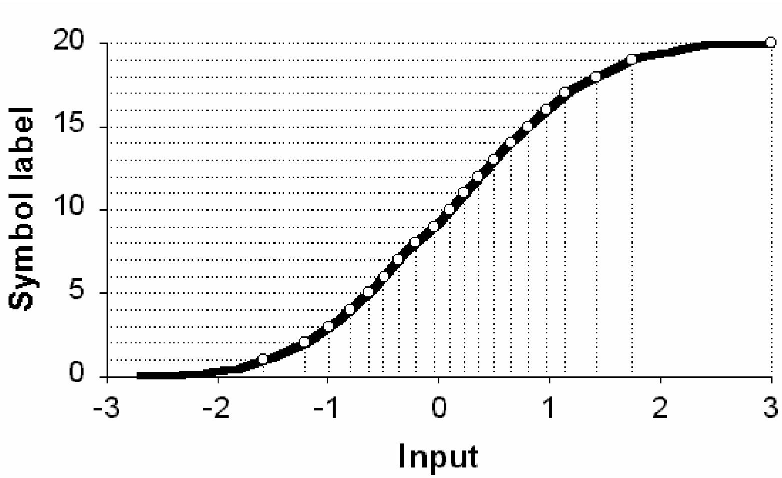


Random stimulus whose variance varies in time

Technical digression I: measuring input/output relations

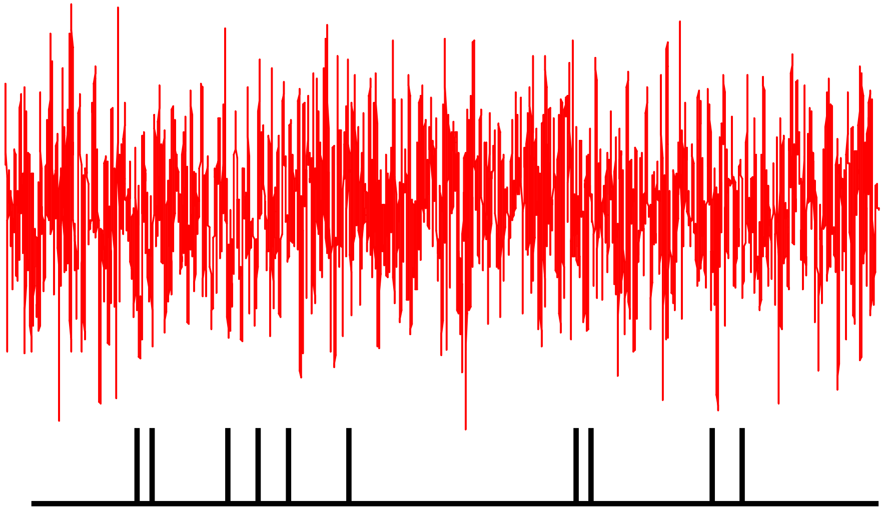
What we want:

Input/output function



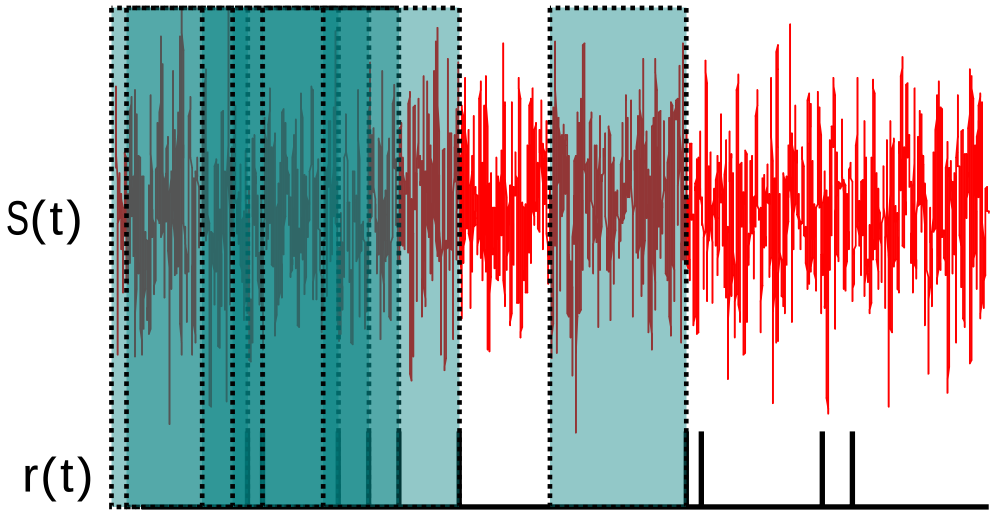
What we have:

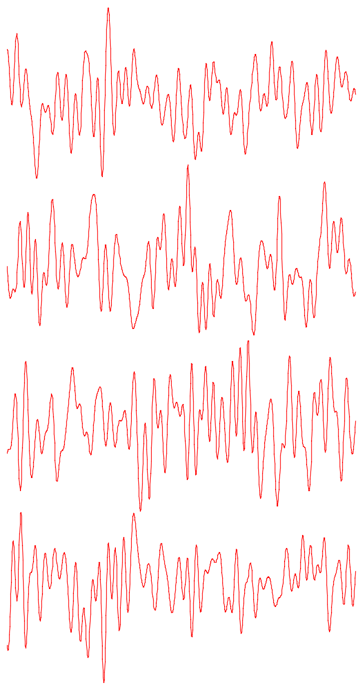
stimulus s(t)



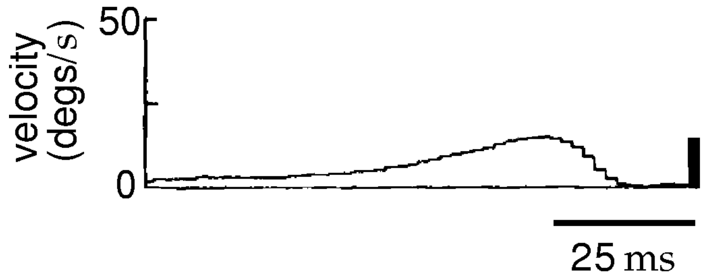
responser(t)

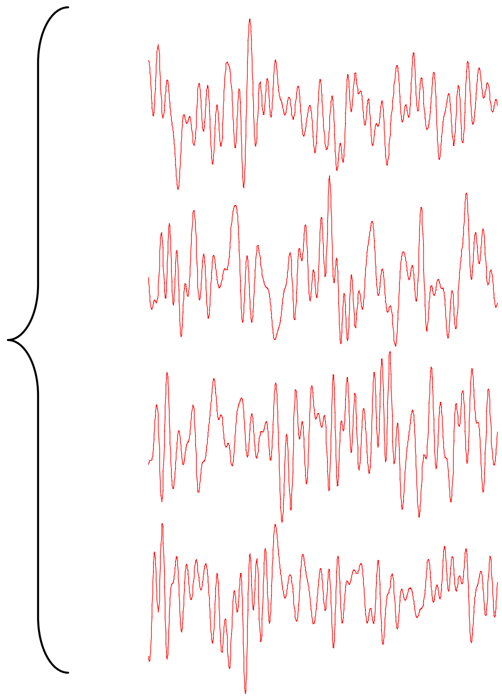
Measuring input/output relations from data





Measuring input/output relations from data





We are only interested in this “relevant” component of each stimulus sample

For simplicity we are using only one component, the average, but one could find multiple relevant stimulus components..

Determining the input/output function

The input/output curve is the function

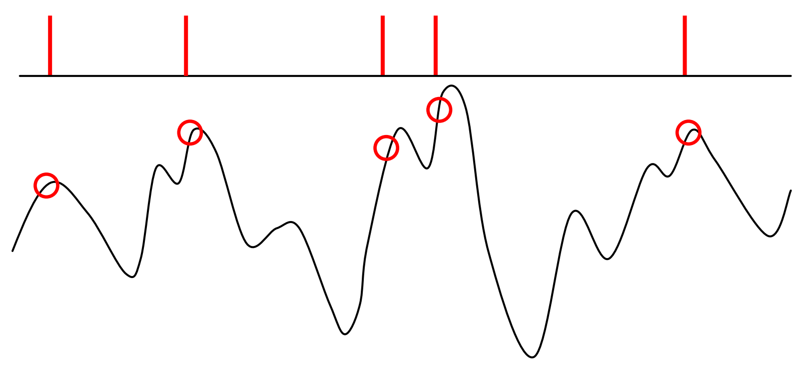
$$P(\text{spike}|\text{stimulus})$$

which we can derive from data using Bayes' rule:

$$P(\text{spike}|s_1) = \frac{P(s_1|\text{spike})P(\text{spike})}{P(s_1)}$$

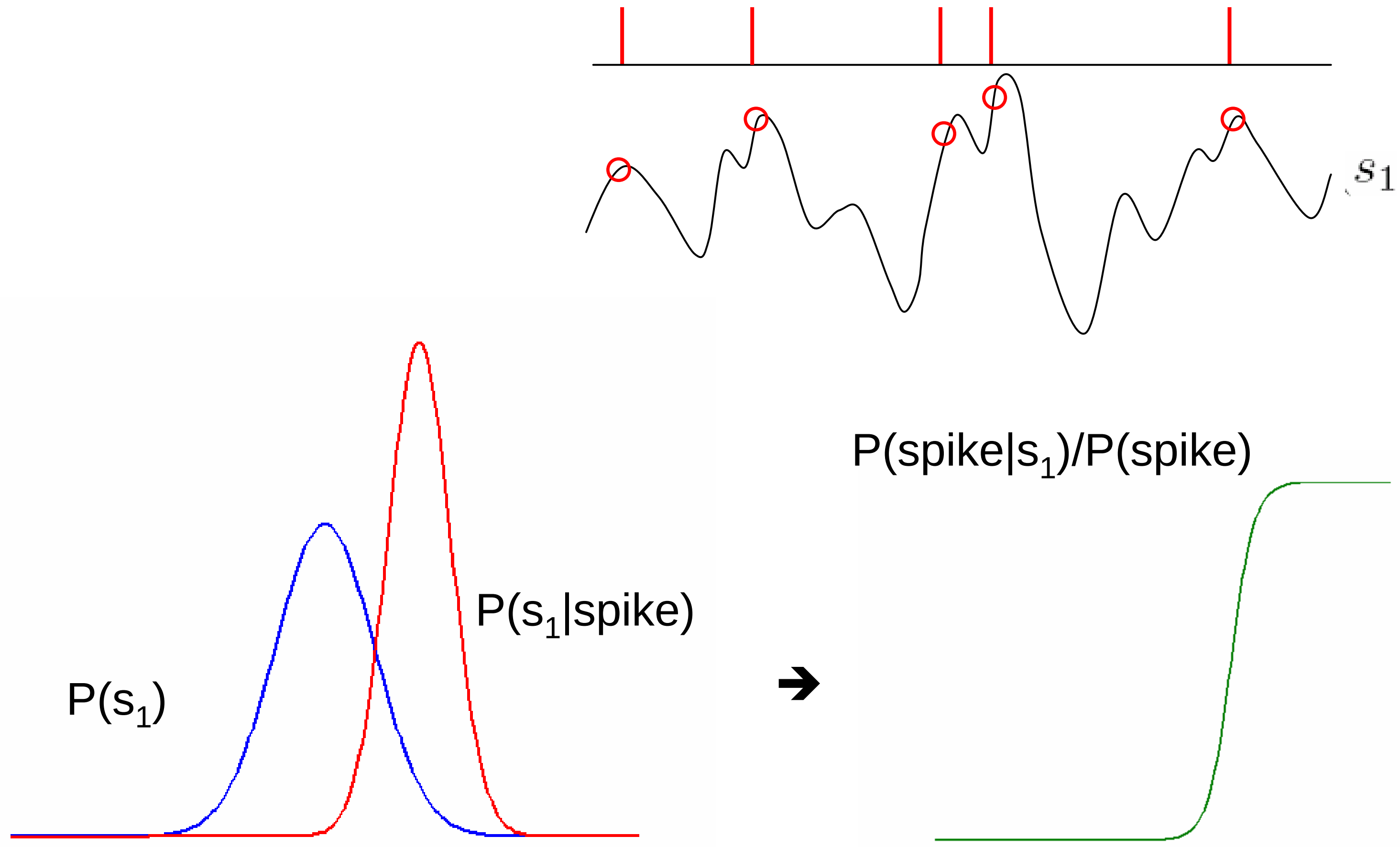
$P(s_1)$

$P(s_1|\text{spike})$



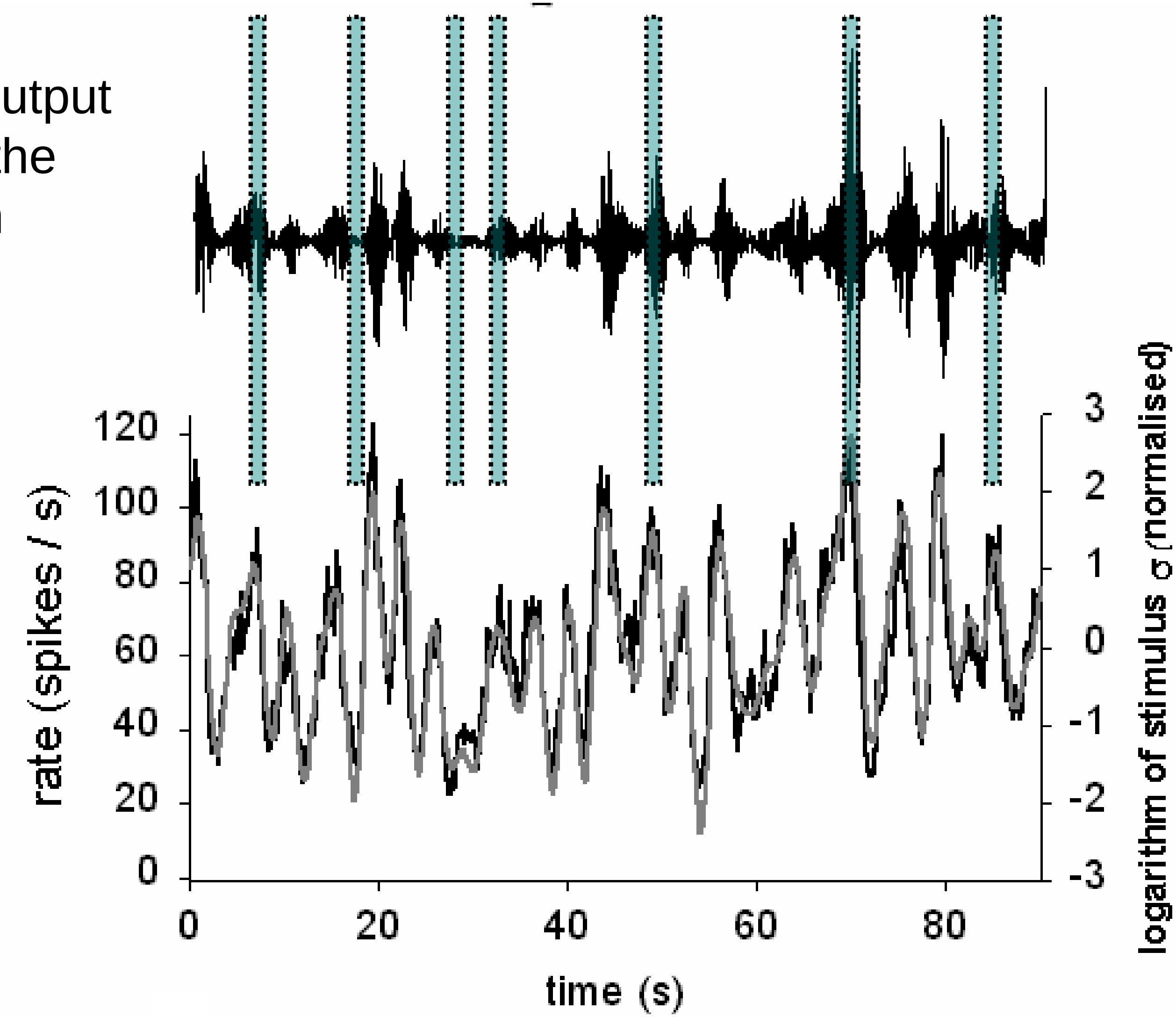
s_1

Determining the input/output function



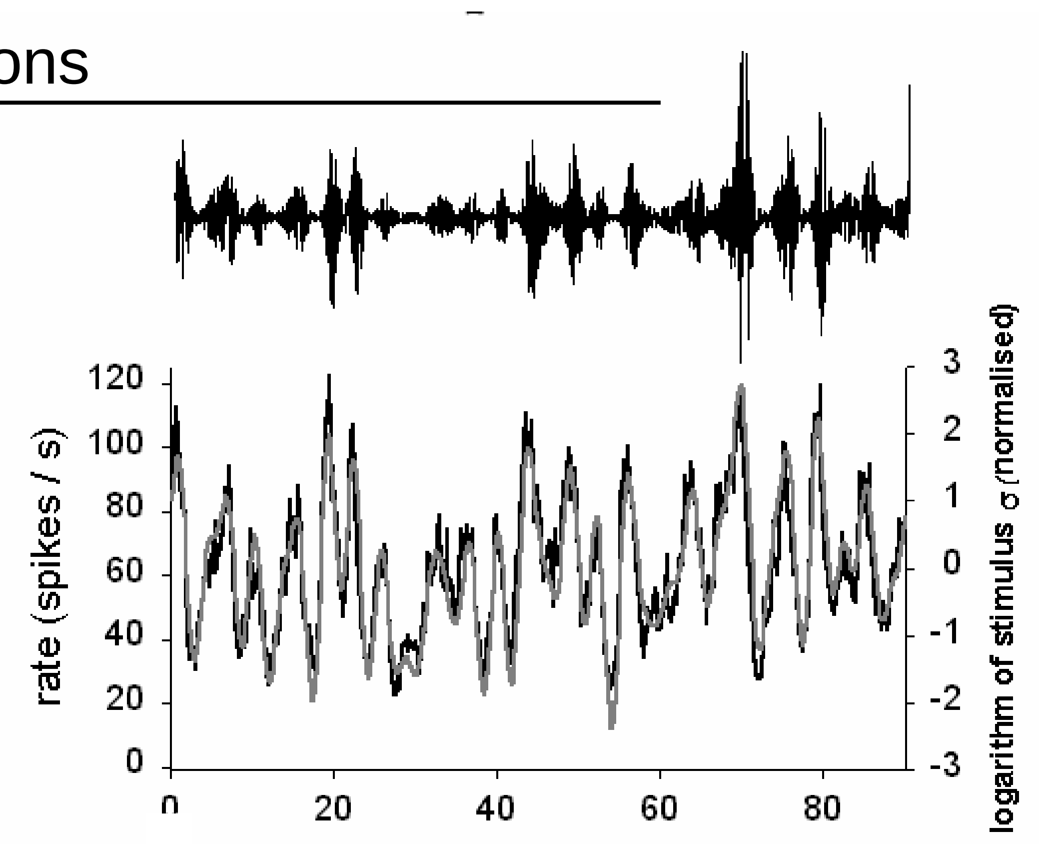
Dynamic input/output relations

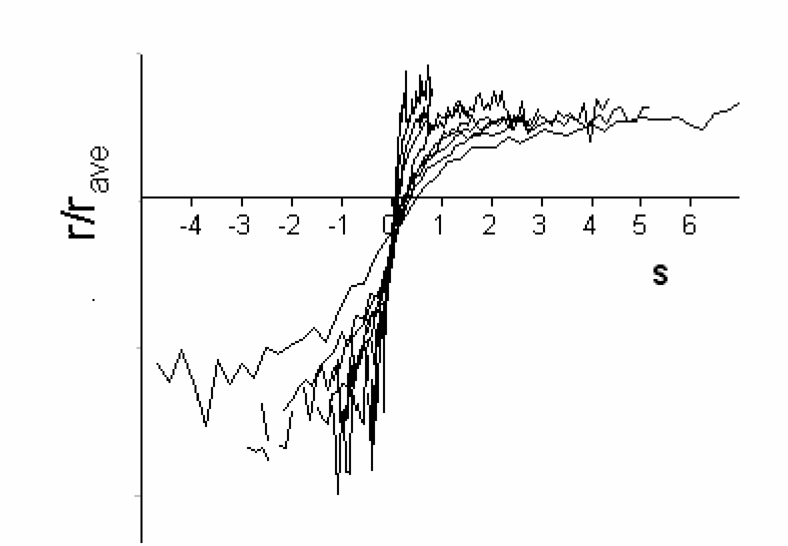
For fly neuron H1, determine the input/output relations throughout the stimulus presentation



Dynamic input/output relations

For fly neuron H1, determine the input/output relations throughout the stimulus presentation

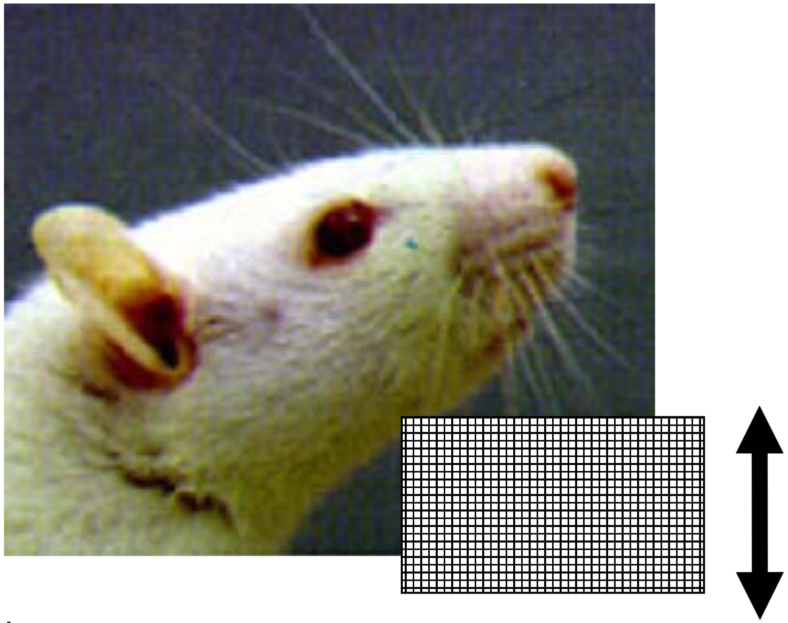




Several neural systems use variance-normalized encoding

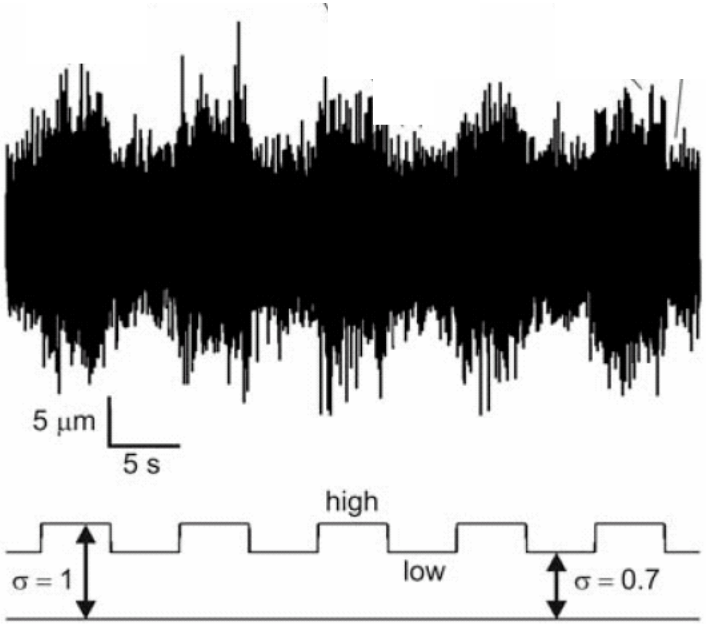
The fly is encoding motion with stimulus values that are locally normalized by variance

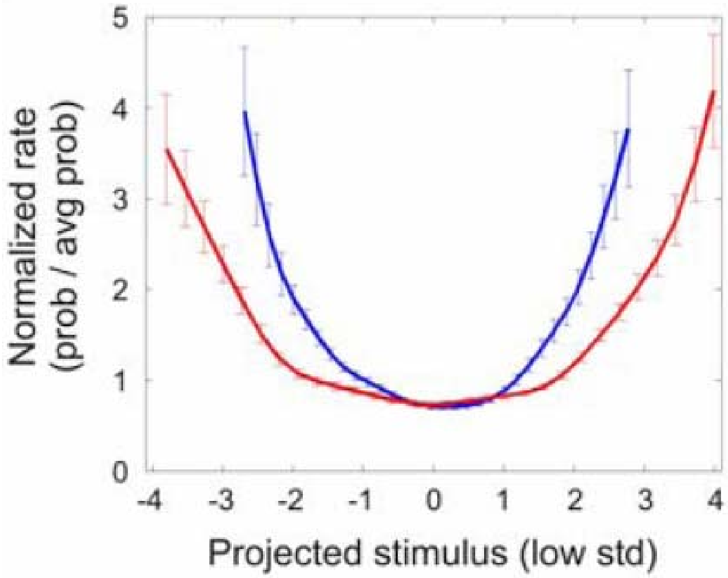
Similar results have been seen in other systems, e.g. retina (*Kim and Rieke '01*), and rat barrel cortex:

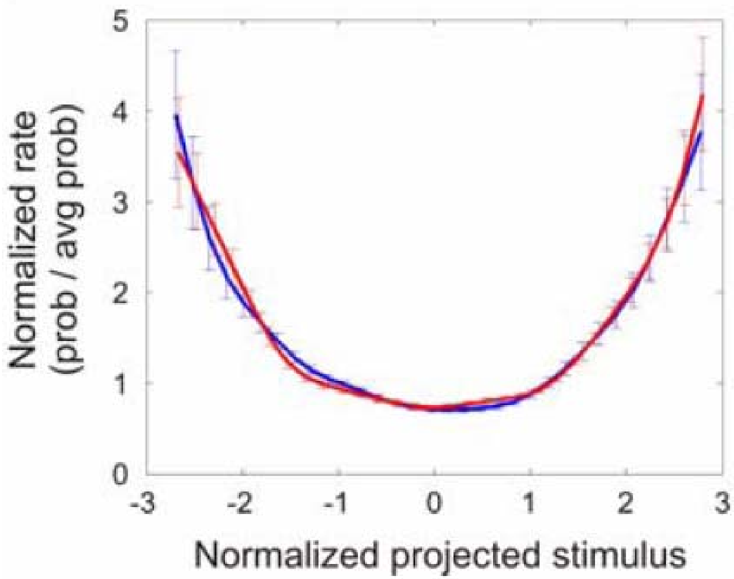


Picture courtesy Ras Petersen

Neurons of rat barrel cortex encoding white noise whisker vibratio







Maravall, Petersen, Fairhall, Arabzadeh and Diamond, PLoS (2006)

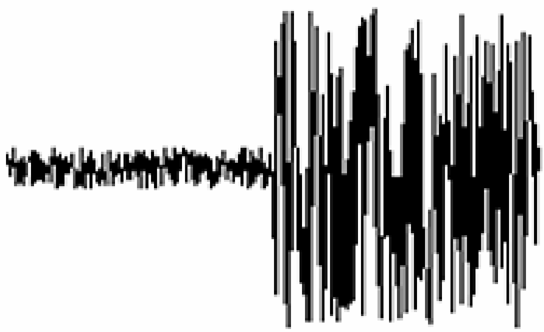
Temporal contrast normalization is a kind of learning

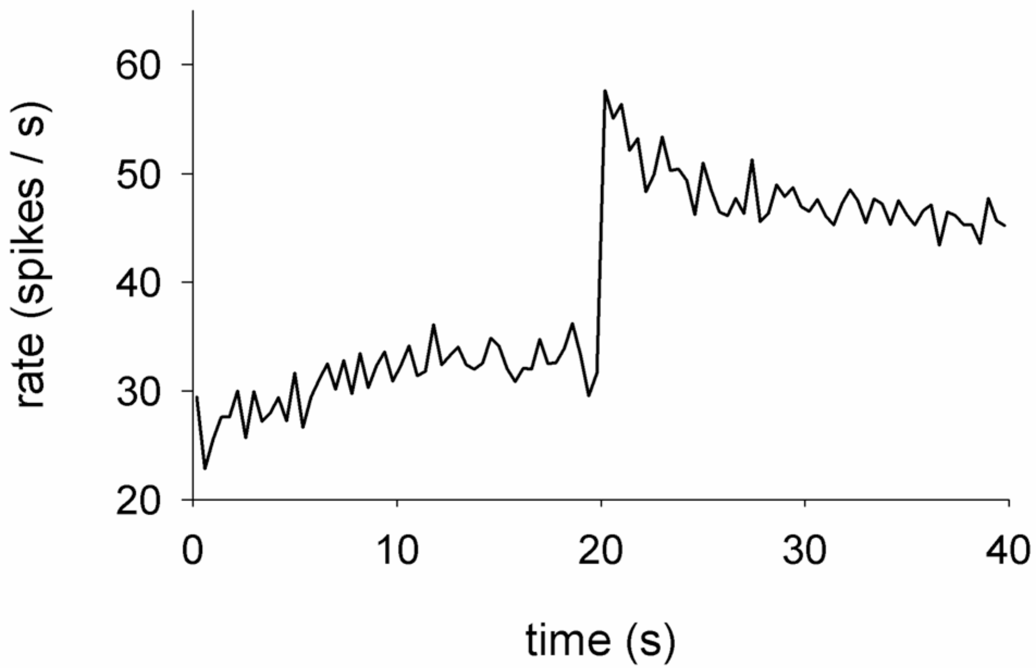
→ the system is “learning” about the evolving characteristics of the stimulus statistics and adjusting its coding strategy appropriately

→ How long does this take?

Back to the fly..

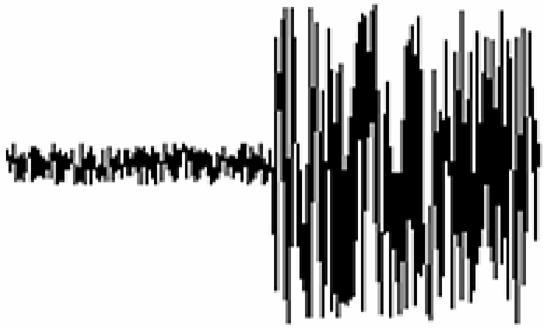
Dynamical adaptive coding

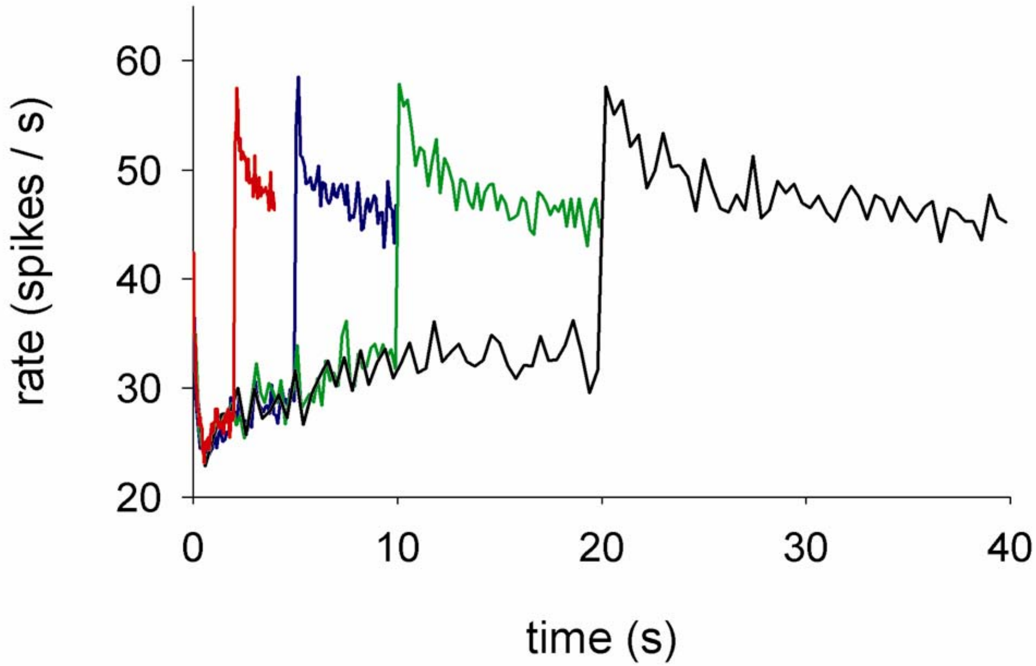




The timescale of the rate adaptation reflects the timescale for learning of the variance… **?**

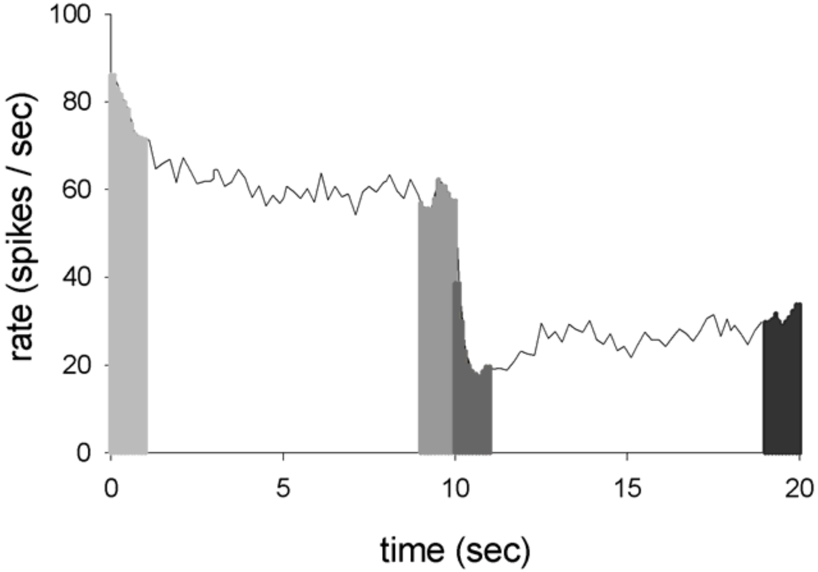
Dynamical adaptive coding

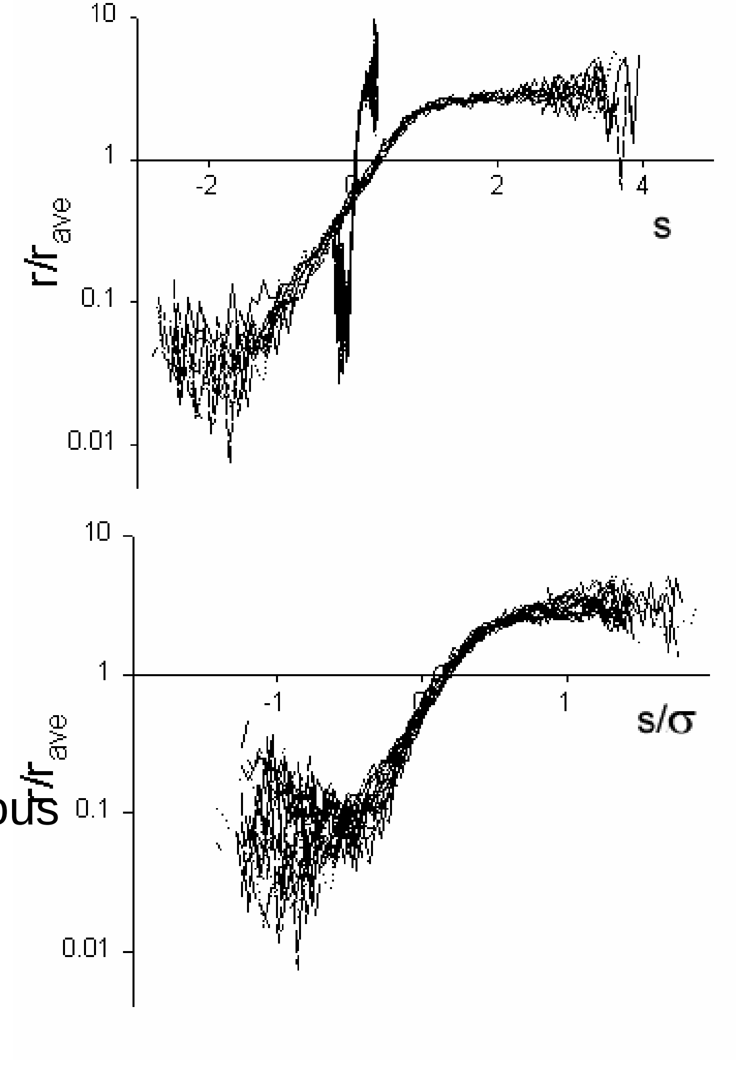




The timescale of the rate adaptation reflects the timescale for learning of the variance… **Not so!**

Dynamical adaptive coding



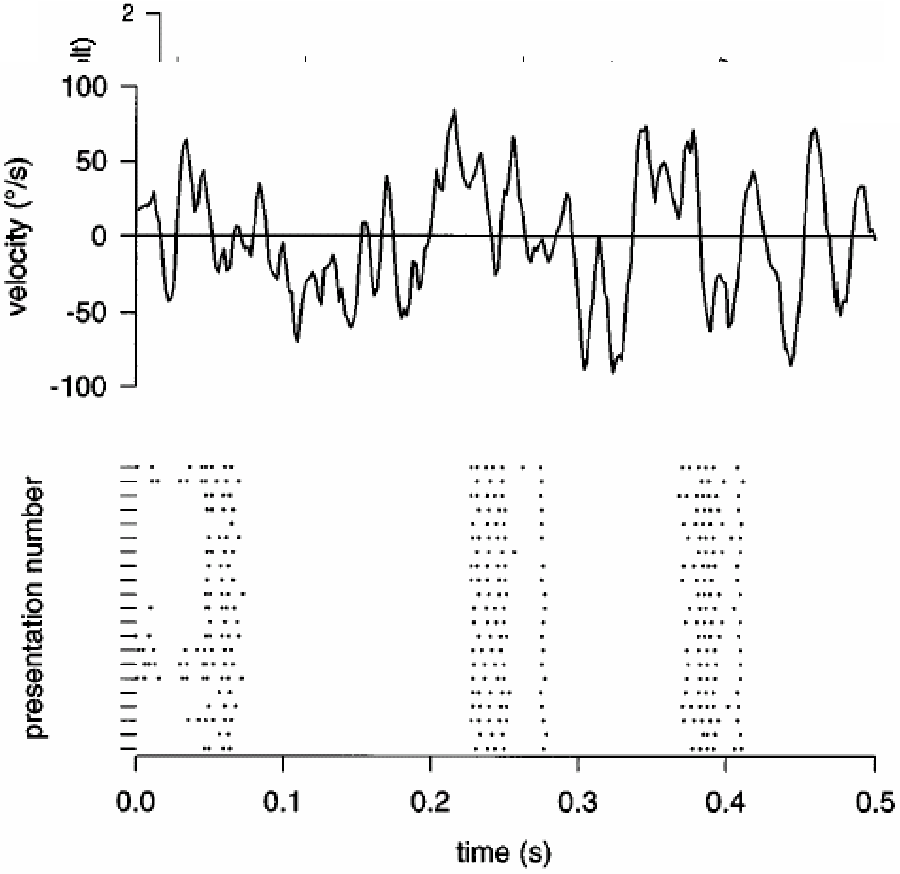


Find variance normalization is almost instantaneous

Determining a precise timescale for adaptation using information

Recall that this type of adaptation should maximize information transmission.

Technical digression II: computing information in spike trains



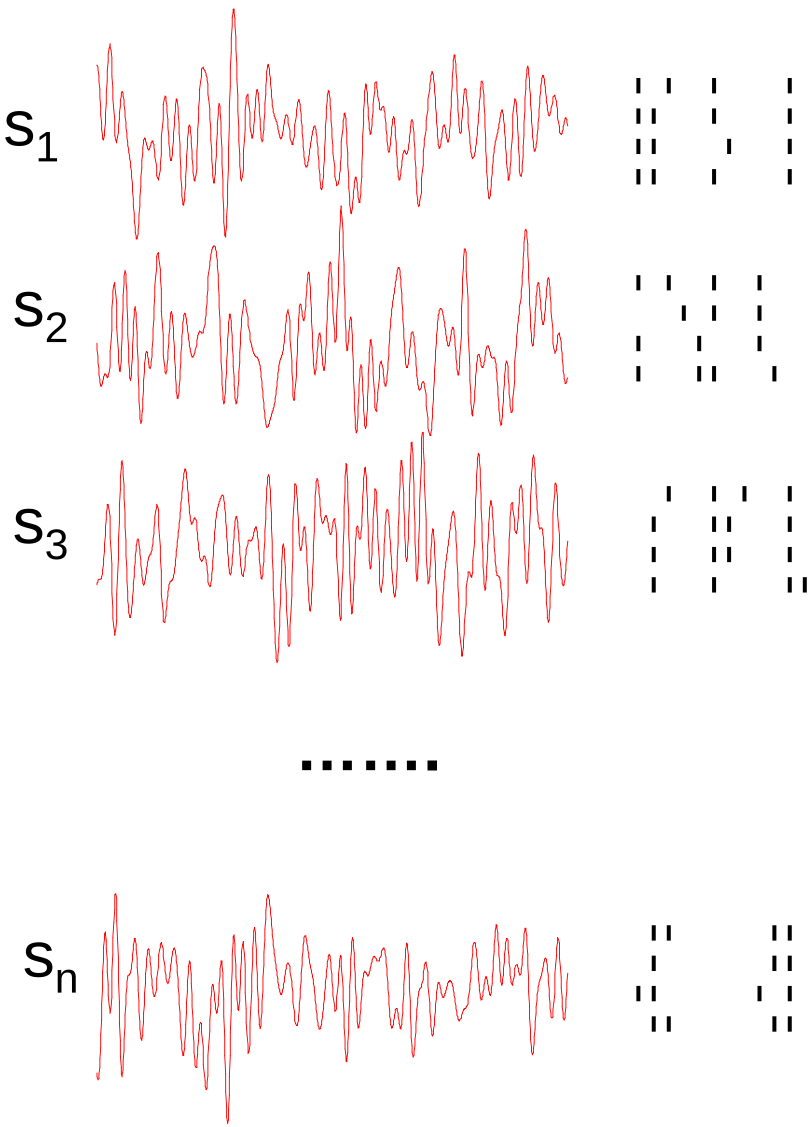
The “direct method”, Strong et al., 199

Represent spike train as binary words

At every time t throughout the stimulus presentation, there will be a distribution of word ouputs

P(w,t)

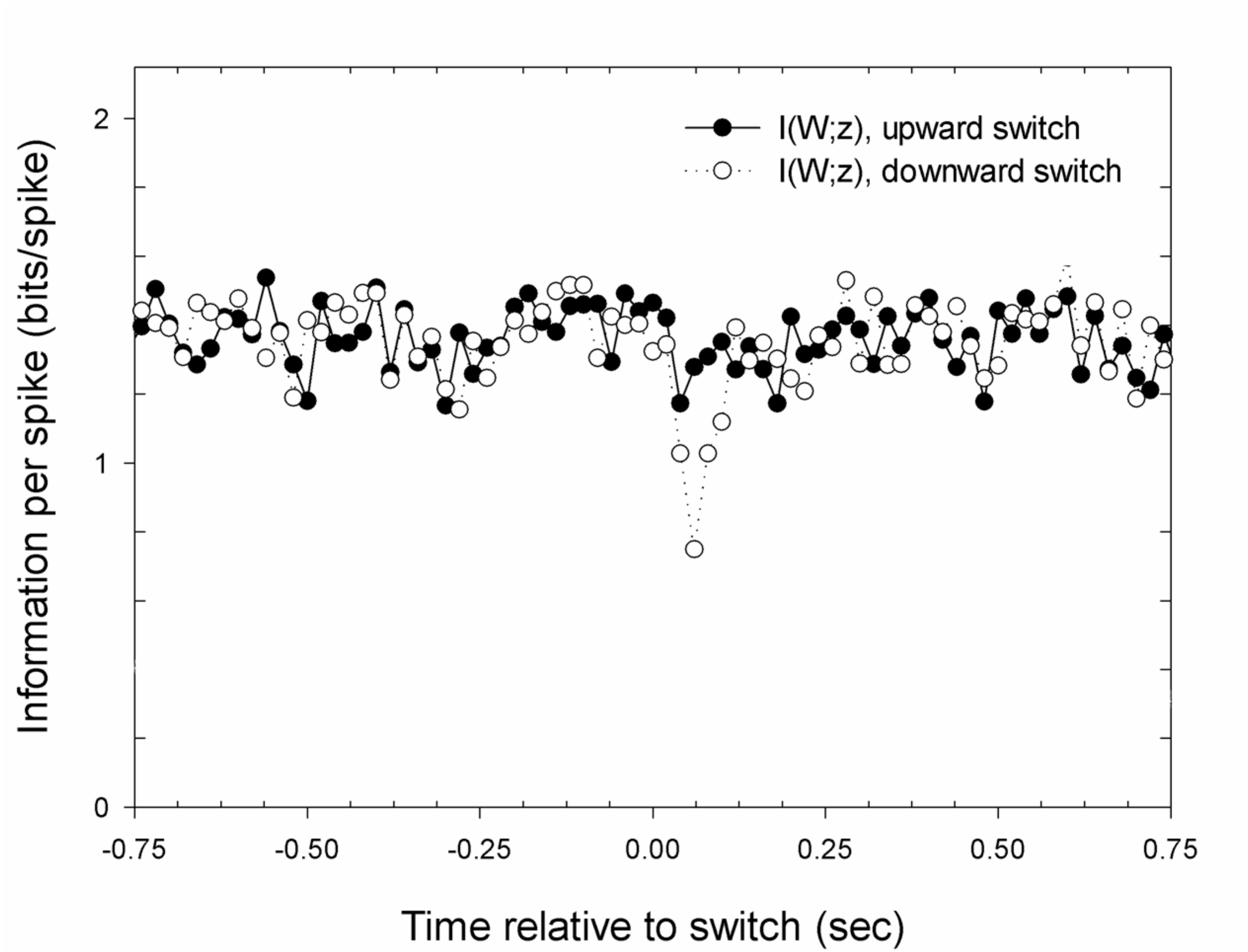
Determining a precise timescale for adaptation using information



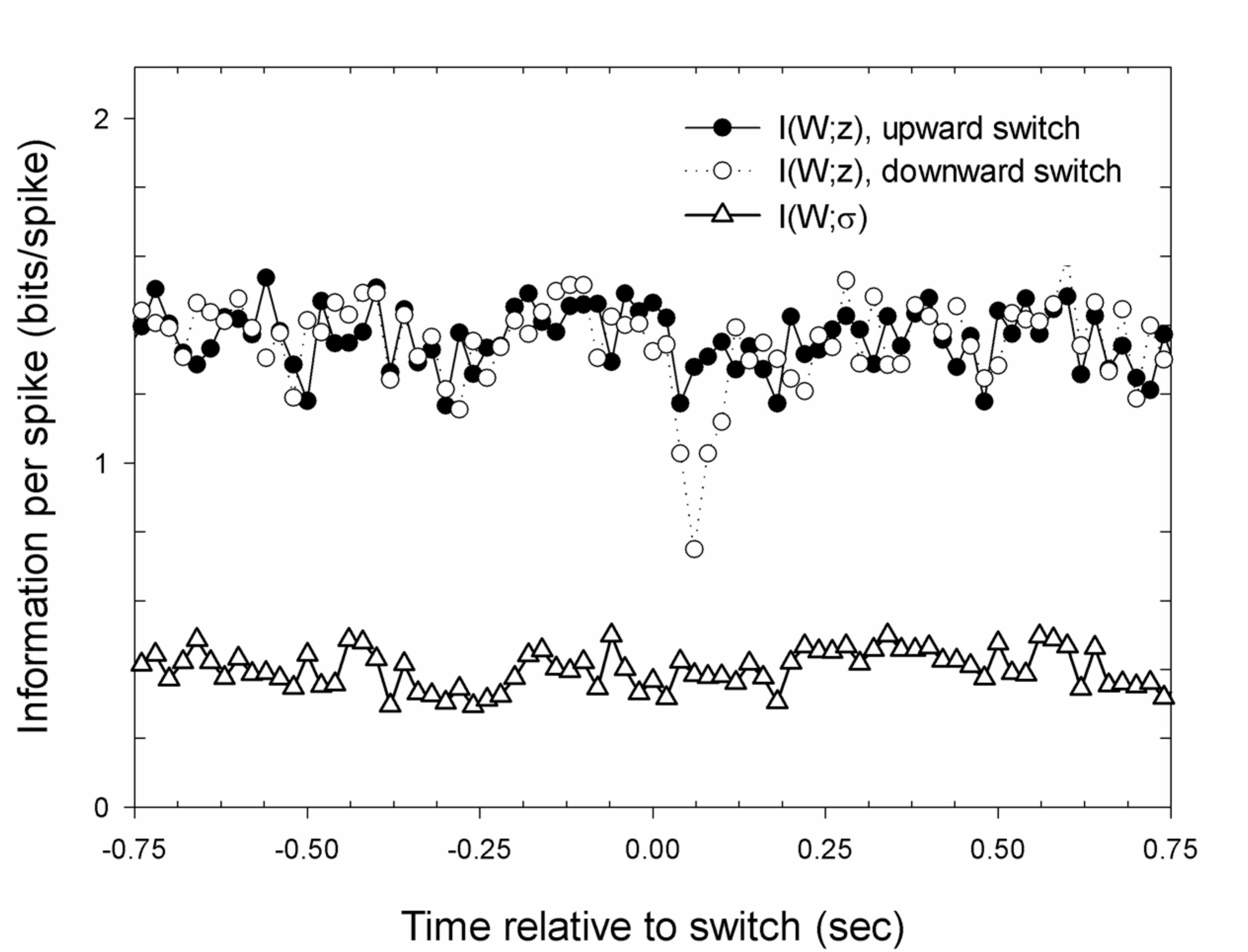
$P(w)$: prior distribution

$P(w|s_i)$: stimulus-conditional distribution

Determining a precise timescale for adaptation using information



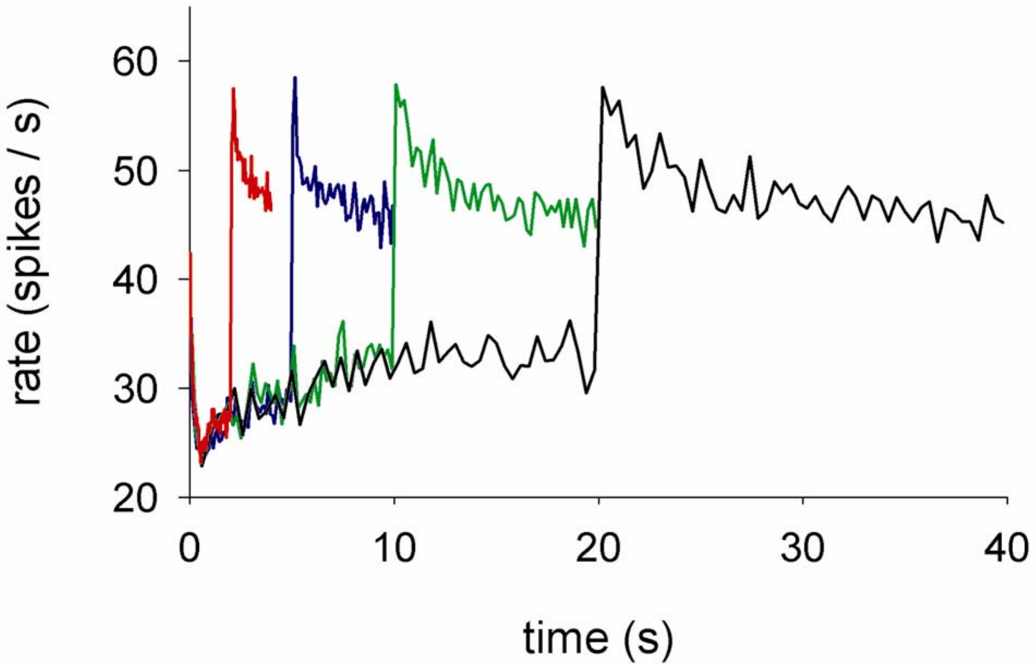
Determining a precise timescale for adaptation using information



Multiple timescales of adaptation

Variance normalization occurs as fast as is statistically possible

The timescale for the adaptation of the firing rate is independent of the timescale of the variance normalization



$$P(\text{spike}|\text{stimulus}) = g(s/s(t))\, R[s(t)]$$

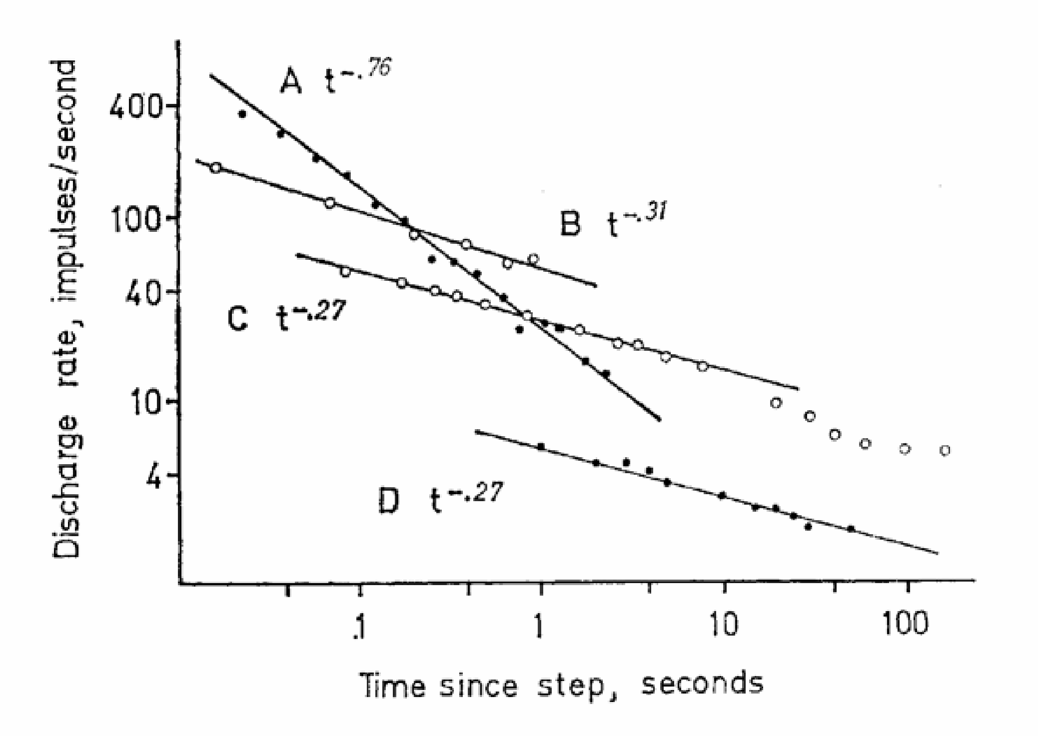
Normalized input/output function

Rate modulation envelope

The rate envelope

• *no fixed timescale*
• Consistent with power-law adaptation

Suggests that rate behaves like
fractional differentiation
of the log-variance envelope

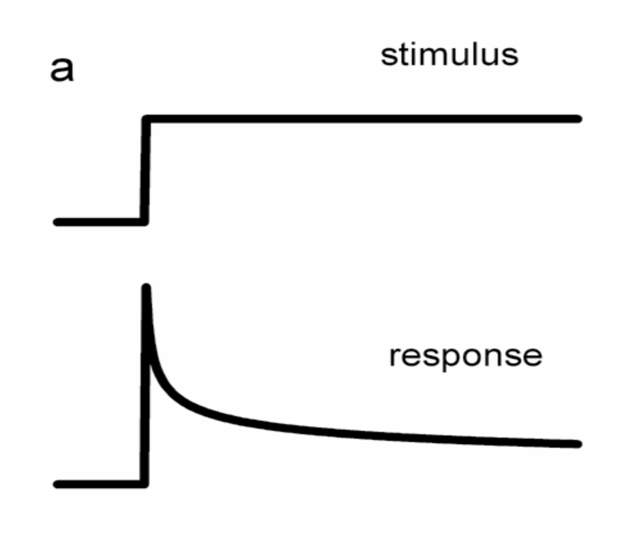


A. Cockroach leg mechanoreceptor, to spinal distortion
B. Spider slit sensillum, to 1200 Hz sound
C. Stretch receptor of the crayfish
D. Limulus eccentric-cell, to increase in light intensity

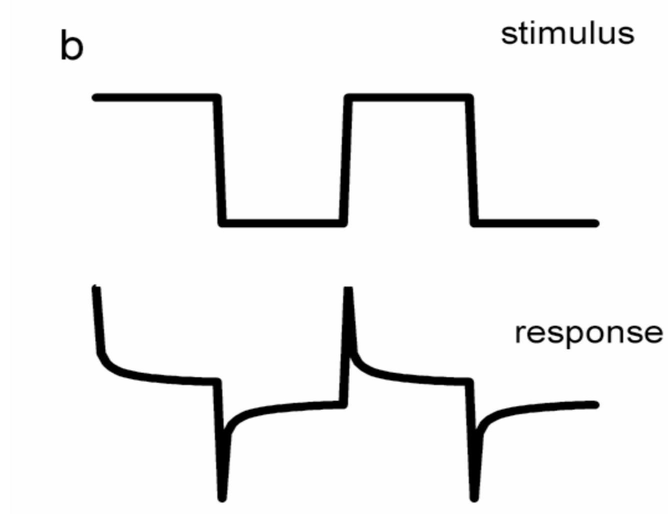
Thorson and Biederman-Thorson, Science (1974)

Fractional differentiation

power-law response to a step:



scaling "adaptive" response to a square wave:



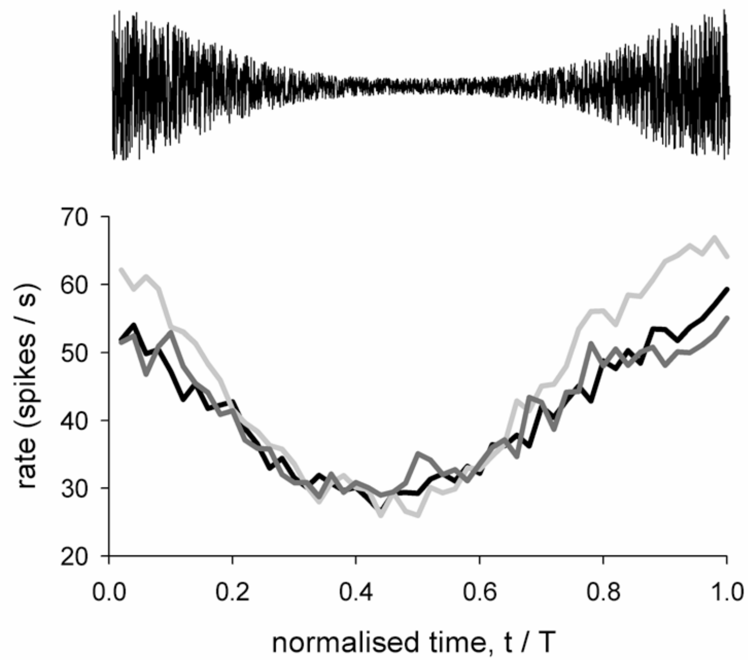
Fourier representation $(i\omega)^{\alpha}$:

each frequency component scaled by ω^{α}

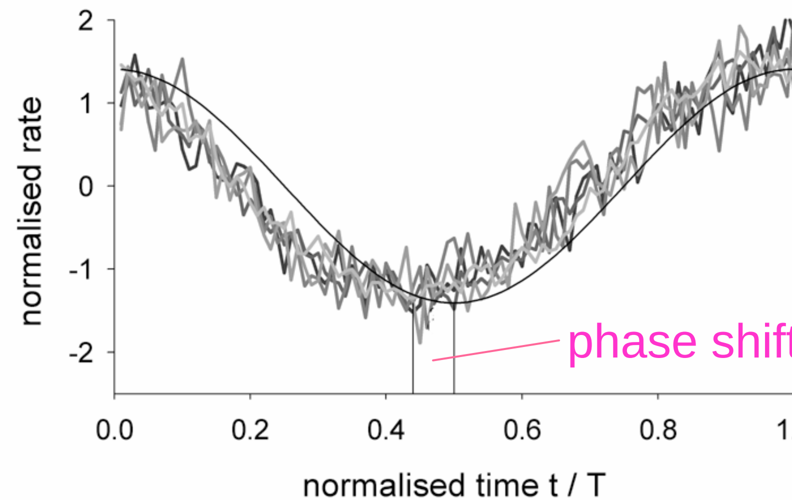
and with phase shifted by a constant phase $i^{\alpha} \rightarrow \alpha\pi/2$

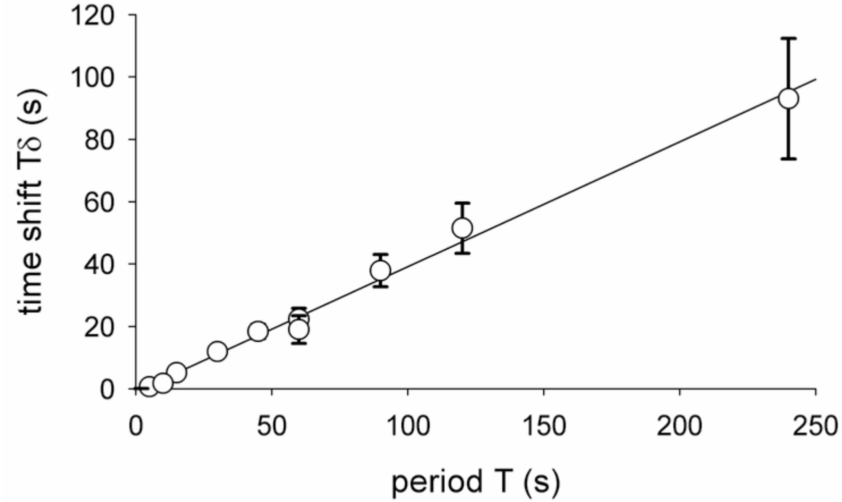
Linear analysis agrees

- Stimulate with a set of sine waves at different frequencies
- Variance envelope ~ exp[sin t/T] for a range of frequencies 1/T



T = 30s, 60s, 90s

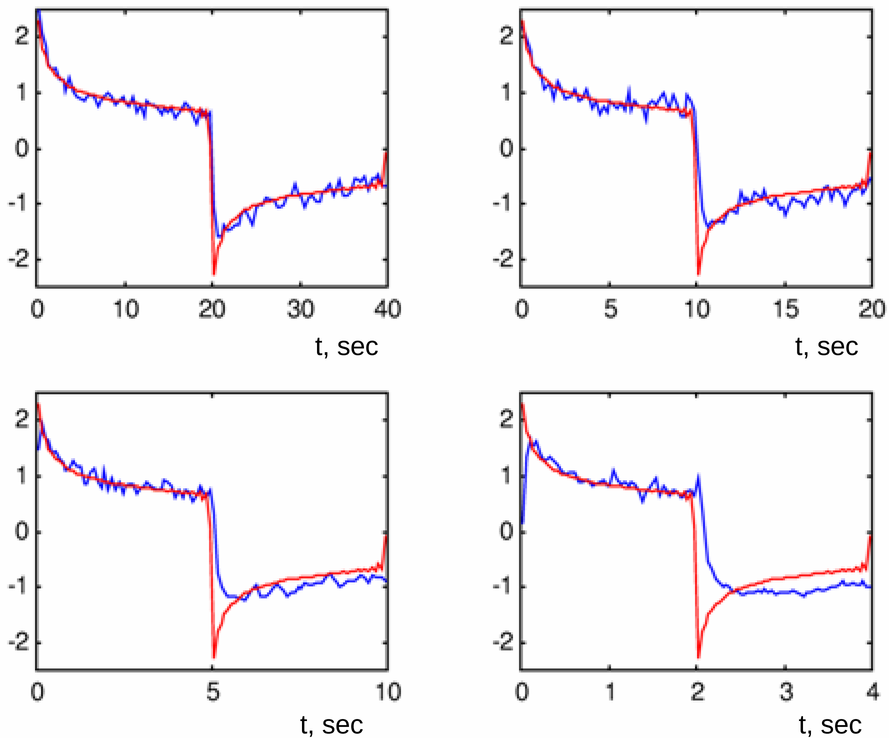




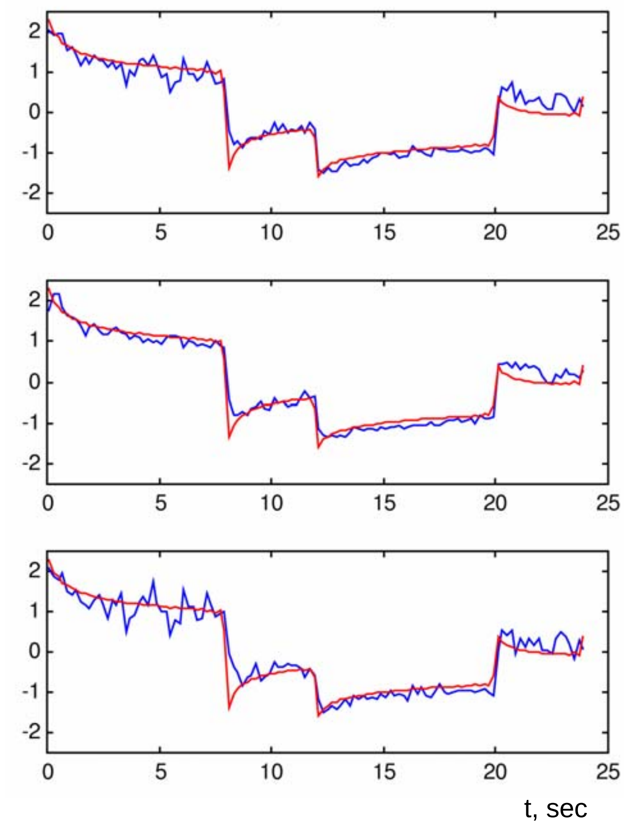
A single parameter fits multiple experiments

From sinusoid experiments, find exponent $\alpha \sim 0.2$

Two-state switching



Three-state switching



Why a fractional differentiator?

Whitening power law structure of stimulus envelope ($1/k^2$ visual scenes)?

Expect that exponent is related to natural stimulus statistics

LINEAR function of stimulus history which emphasizes change without removing steady state response

.. A general neural computation?

How are these dynamics implemented?

Two distinct forms of adaptation which both have relevance for efficient processing of natural stimulus statistics

Feedback or feedforward?

Variance normalization: too fast for feedback?

Fractional differentiation: linear transformation

Adapting without learning

Variance normalization as an outcome of neural nonlinearity

Sungho Hong
Brian Lundstrom
Kate Tabor

• Nonlinearity of neurons ensures interaction between output and statistics of stimulus input

• Have reproduced this behavior predictably with a simplified generalized linear/nonlinear neural model

• Can tune parameters of a realistic neural model to find regions of greatest "adaptability"

• How much of the adaptation we see at the systems level can be carried out by single neurons? *Maravall et al.*

• Could real neural systems have evolved so as to be maximally adaptable?

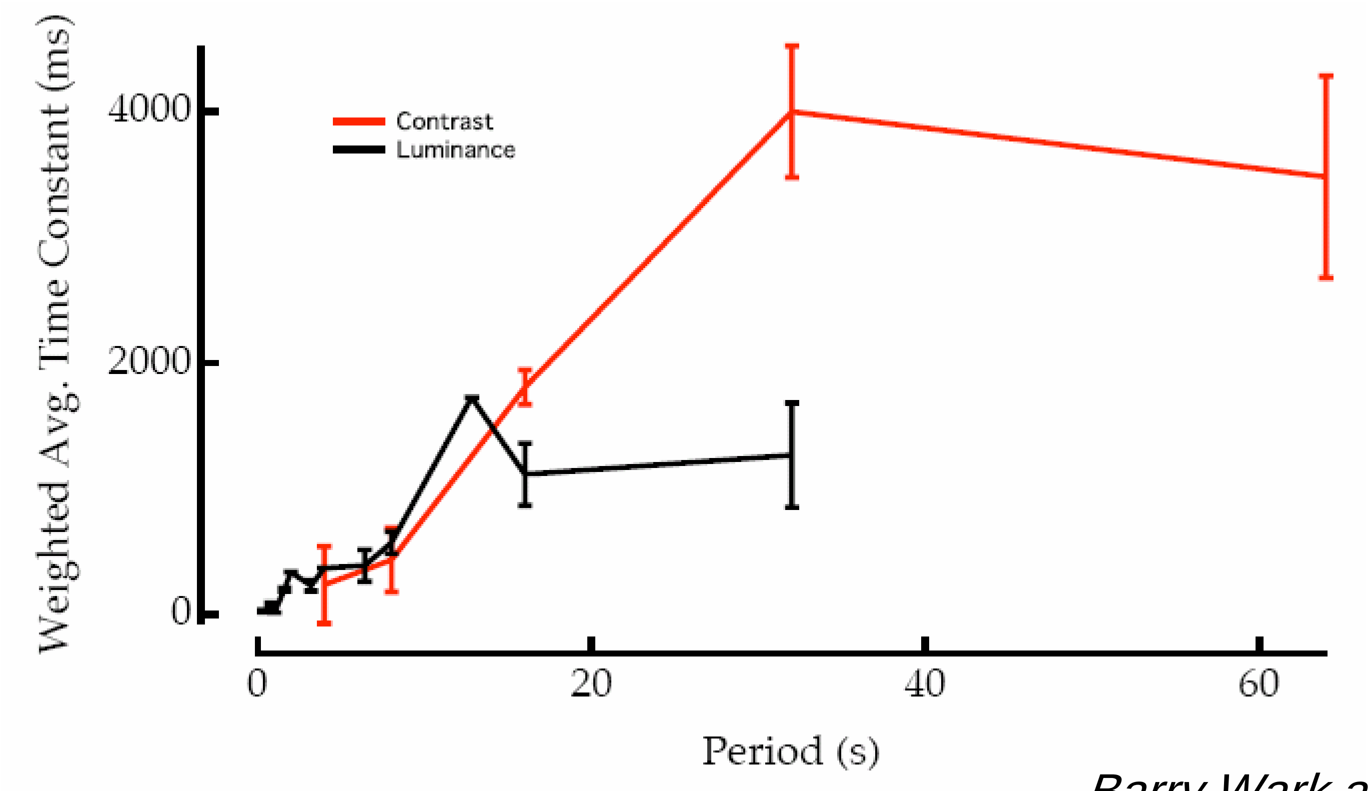
Multiplicity of timescales

Where do these come from?

- circuit based (potentially involving feedback) or single neurons?

- recent models (Brenner et al. 2005) involve multiple inactivation states of ion channels

Time-scaling adaptation in retinal ganglion cells

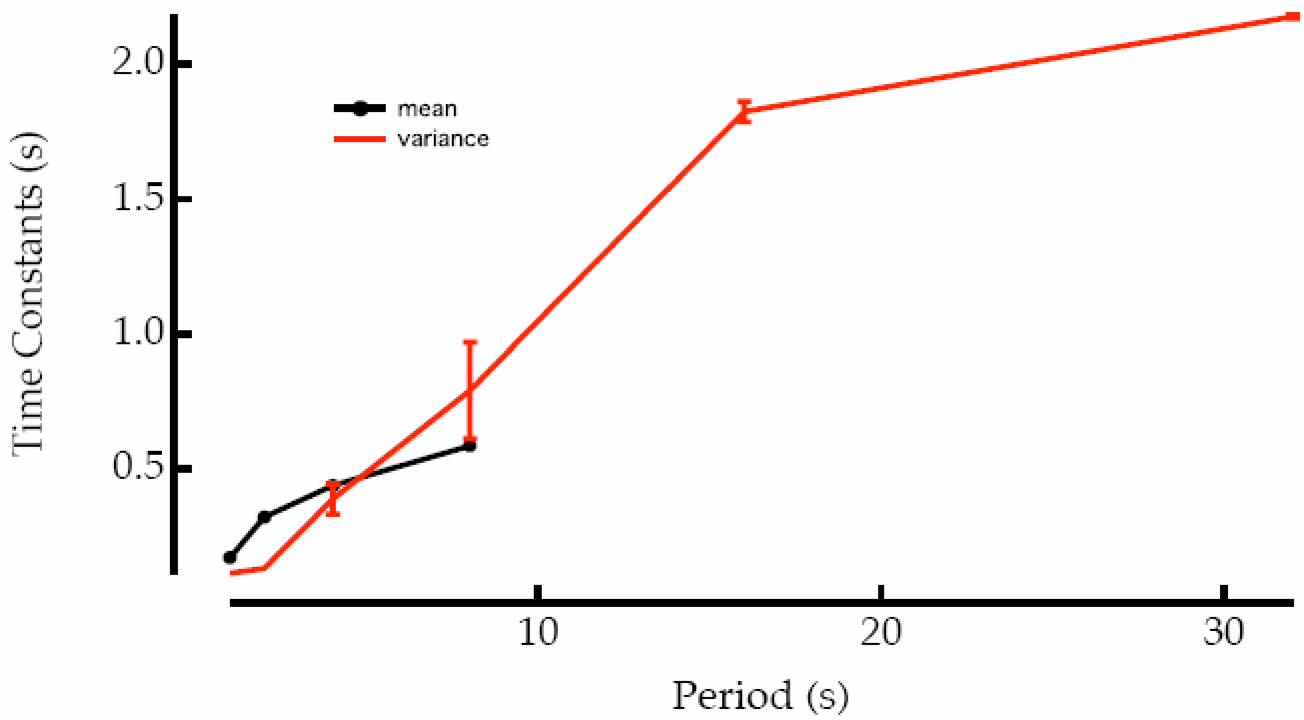


Barry Wark and Fred Rieke

→ Explore where in the circuit different adaptations occur

→ Relate to the special properties of natural images

Time-scaling adaptation in cortical pyramidal cells



Brian Lundstrom and Bill Spain

→ Very general phenomenon, not only at the sensory periphery

→ can test hypotheses of biophysical mechanism

Summary

Different types of adaptation allow neural systems to maintain high information rates in the presence of changing stimulus statistics

Separation of time/length scales for efficient encoding of different aspects of stimuli

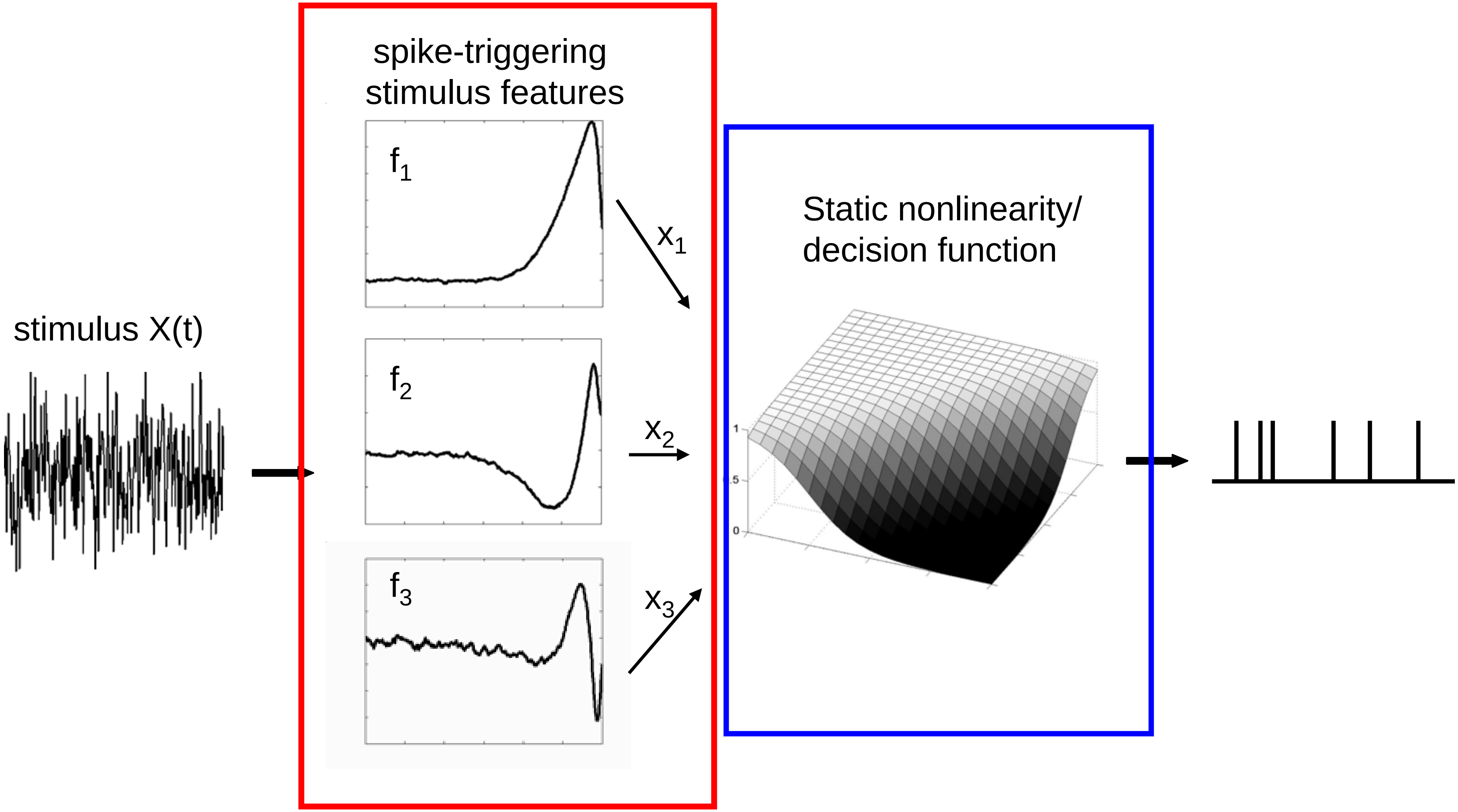
Fast adaptation allows for contrast gain control or variance normalization

Slower adaptive processes may be responsible for whitening power-law stimulus envelope statistics

These adaptive forms seem to be general across different systems

Testing models for precise biophysical mechanisms underlying these adaptive dynamics

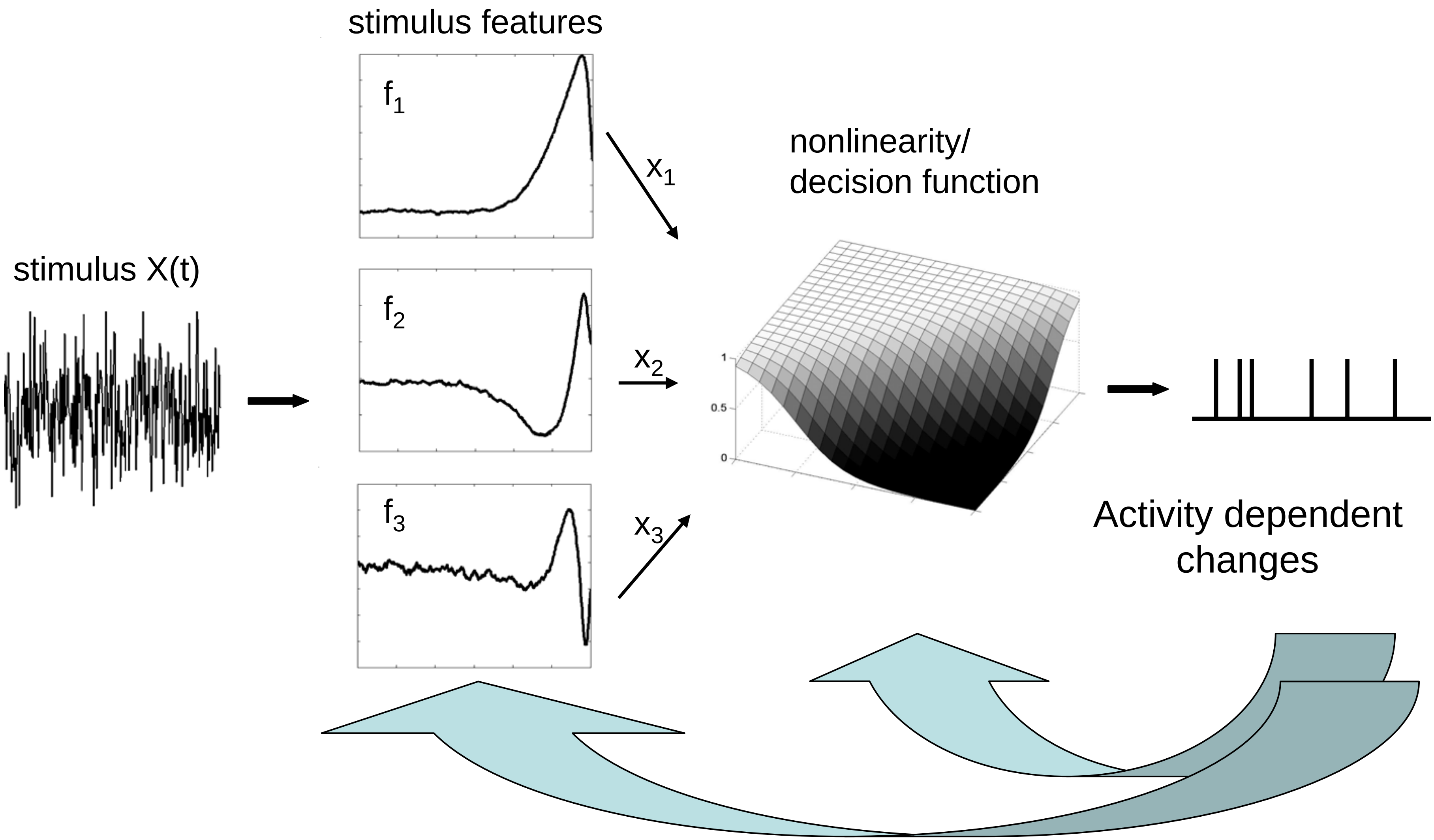
Characterizing adaptive responses: cascade models



Receptive field adaptation
Atick

Threshold function adaptation
Laughlin

Cascade models with feedback



Adaptation to stimulus statistics

Have concentrated on adaptation to variance and justified that with a connection to the properties of natural scenes

Claim: the visual system only adapts to the first two moments
Bonin et al. (2005)

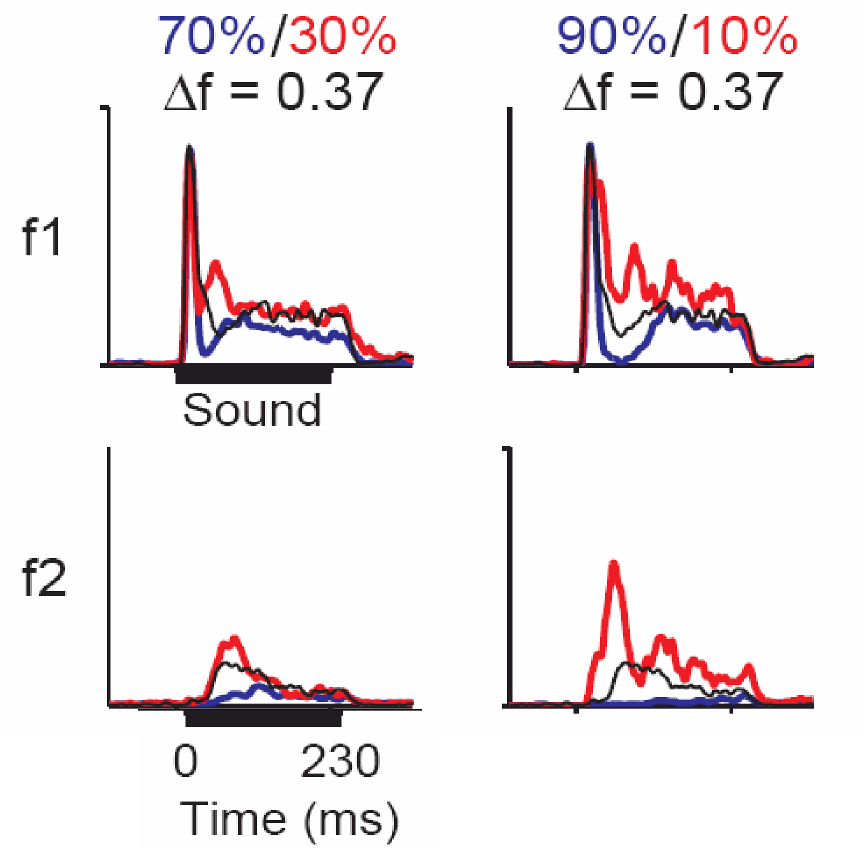
Let's look at some examples of adaptation to other stimulus statistics

Adaptation to stimulus frequency

Processing of low-probability sounds by cortical neurons
Ulanovsky, Las and Nelken, Nature Neuroscience (2003)

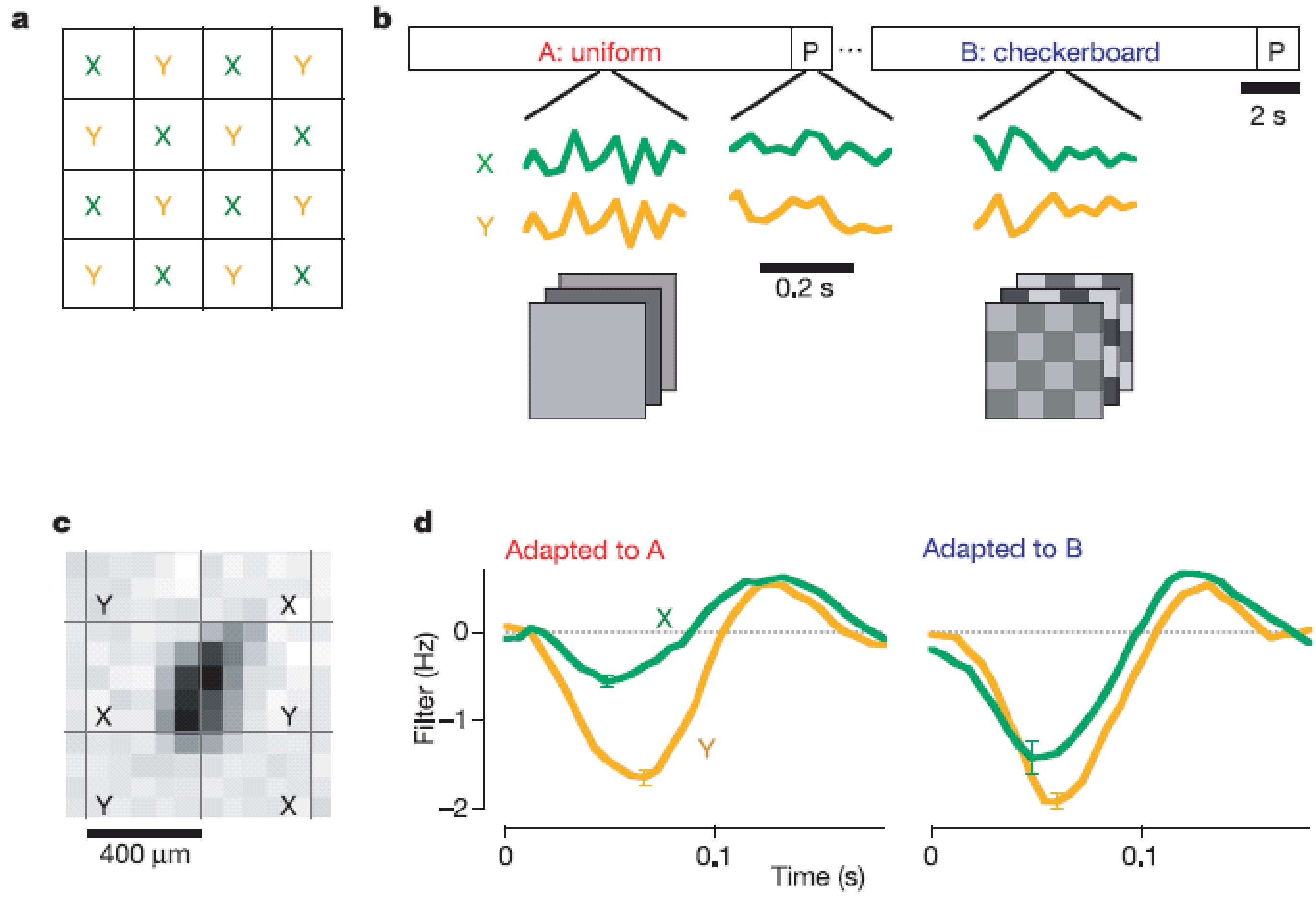
Auditory stimulus, recording in A1:

f1 f1 f1 f2 f1 f1 f1 f2 — f2 rare
f2 f2 f2 f1 f2 f2 f2 f1 — f1 rare
f1 f2 f1 f2 f1 f2 f1 f2 — neither rare



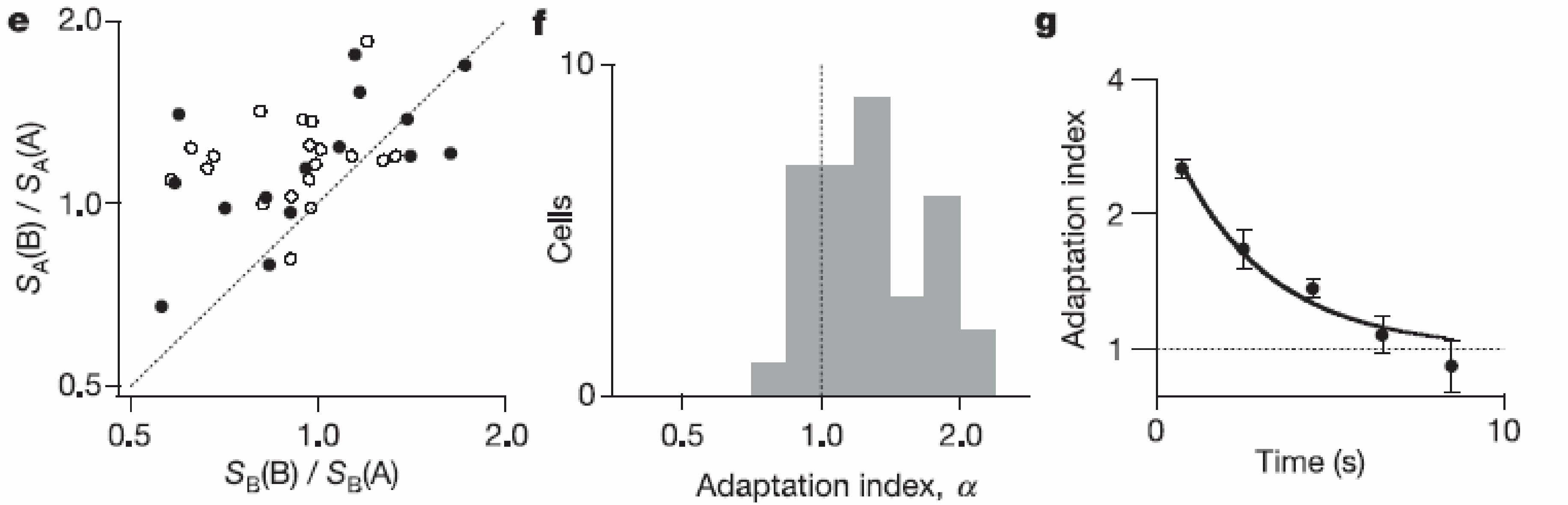
Adaptation to arbitrary spatio-temporal correlations

Dynamic predictive coding by the retina, Hosoya et al., Nature (2005)



Adaptation to arbitrary spatio-temporal correlations

Dynamic predictive coding by the retina, Hosoya et al., Nature (2005)

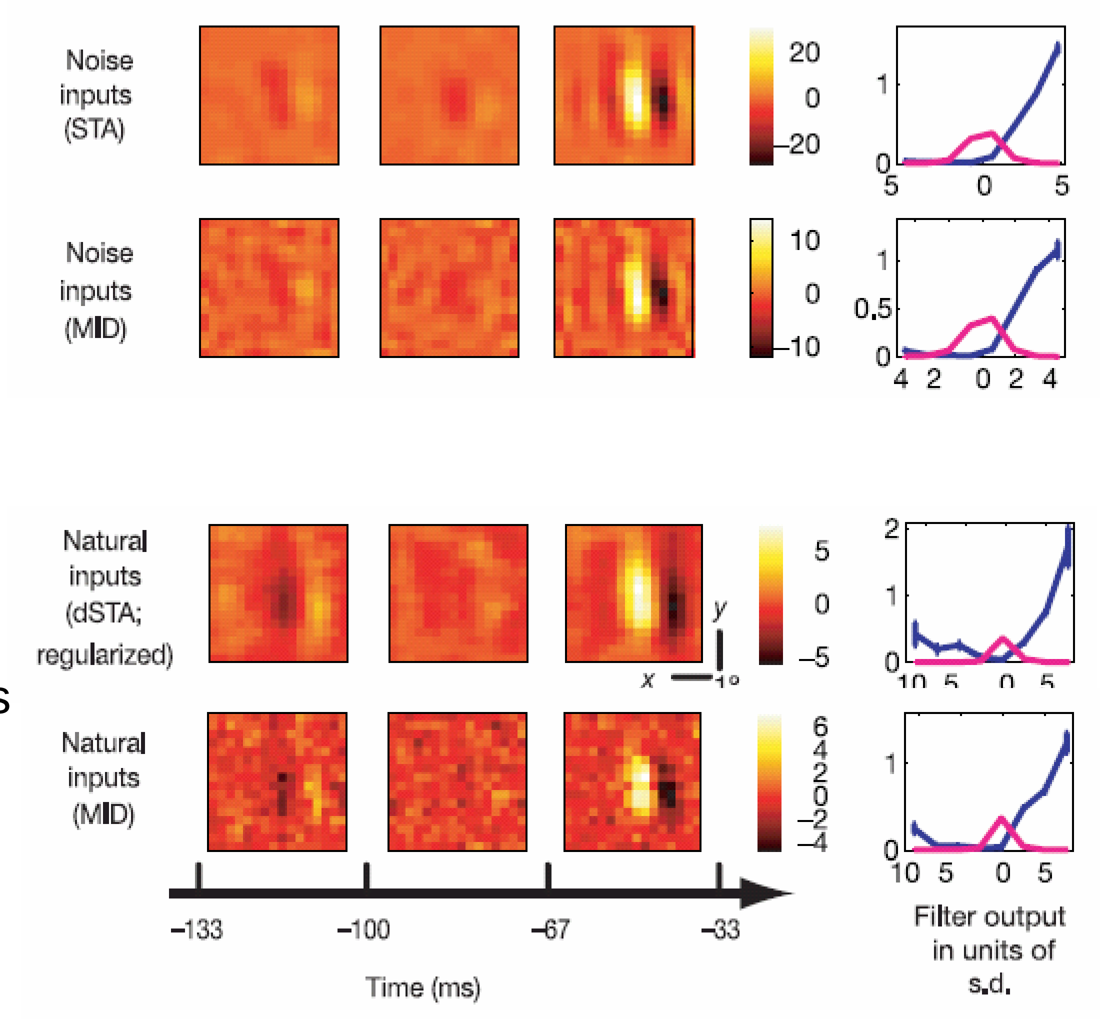


Adaptation to natural stimulus statistics

Adaptive filtering enhances information transmission in visual cortex, Sharpee, Sugihara, Kurgansky, Rebrik, Stryker and Miller, Nature (2006)

Receptive fields determined from STA and from maximally informative dimension

Compare receptive fields obtained with noise and with natural inputs



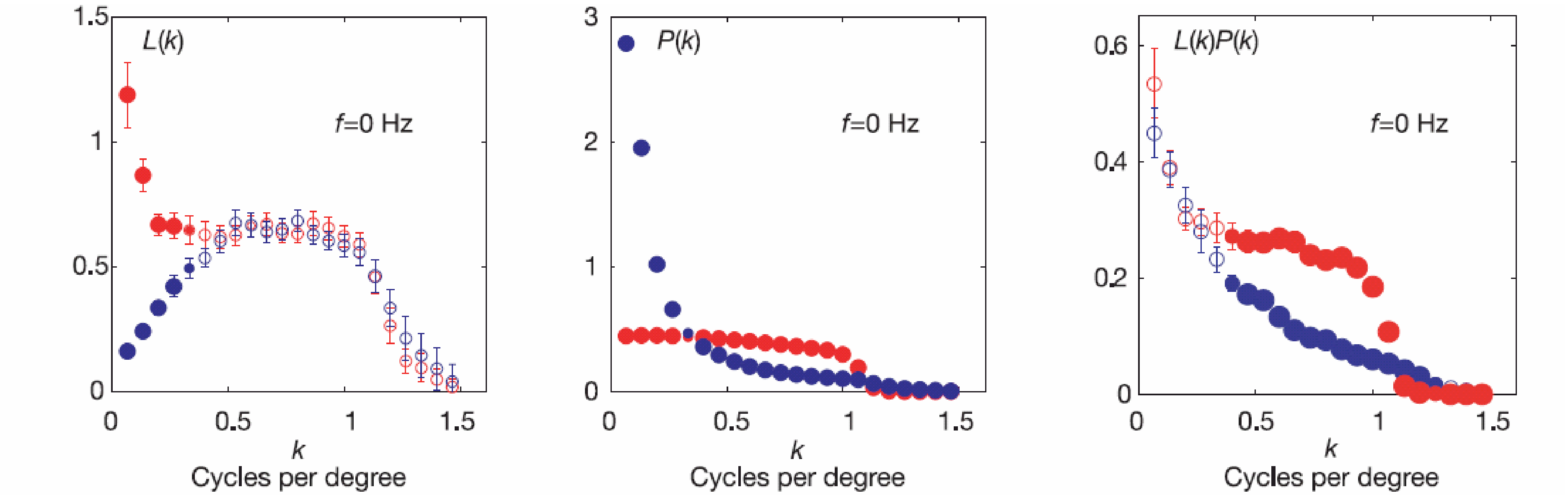
Adaptation to natural stimulus statistics

Adaptive filtering enhances information transmission in visual cortex, Sharpee, Sugihara, Kurgansky, Rebrik, Stryker and Miller, Nature (2006)

Filter spectrum

Stimulus spectrum

Filtered output



Red = white noise
Blue = natural statistics

Adaptation to natural stimulus statistics

Adaptive filtering enhances information transmission in visual cortex, Sharpee, Sugihara, Kurgansky, Rebrik, Stryker and Miller, Nature (2006)