Harvard Bauer Center, 02/17/05

Thinking about information processing in biological systems

Ilya Nemenman

Joint Centers for Systems Biology
Columbia

Goals & Philosophy

Bauer Center for Genomics Research

The 20th century put biology on a molecular footing. This was achieved by reducing problems to defined questions that could be isolated as much as possible from the complexities of living organisms. Biochemists purified proteins and studied them in vitro, crystallographers solved their structure and geneticists used mutations to focus on the role of individual genes. Although these approaches have produced enormous advances, they have not solved the ultimate challenge of biology: how can we explain the behavior, function, structure, and evolution of cells?

Major unanswered questions include the following:

1) How is the behavior of the thousands of different types of molecules (proteins, nucleic acids, carbohydrates, and small organic molecules) coordinated and integrated to allow cells to survive and reproduce?

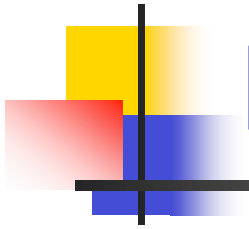
2) What enables key properties of cells to be robust to a wide range of perturbations in their genotype and environment?

3) How does evolution change biological molecules and the interactions between them so that organisms adapt to and exploit long term environmental changes?

4) How can short-term robustness co-exist with long-term adaptability?

We believe that there are design principles to be discovered in biology, which will help us address the major questions of how living organisms work

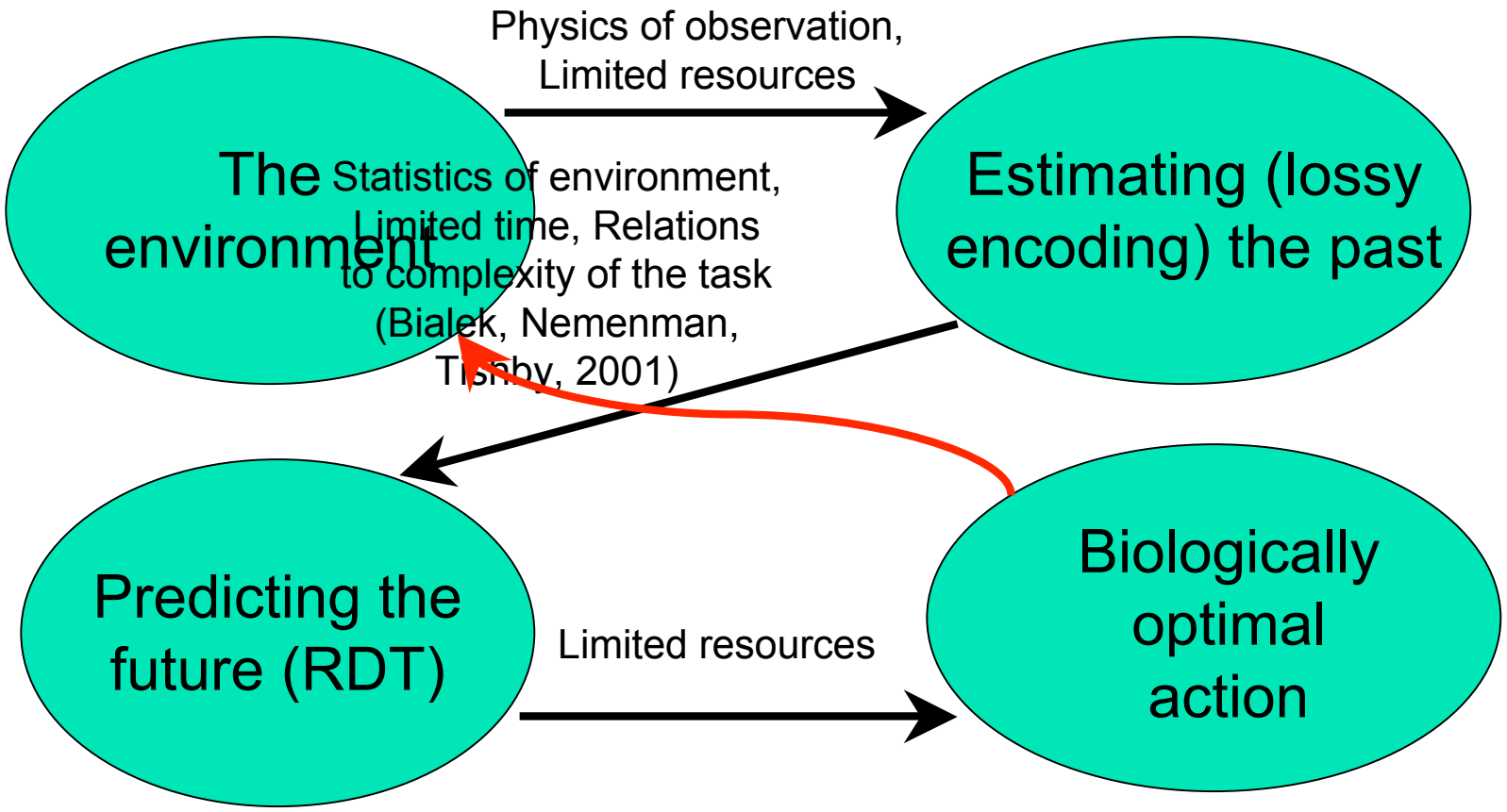
Next page



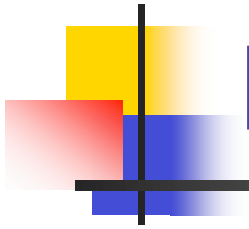
Design principles

- Not *how*, but *what* and *why.* Understanding.
- More than one data set per principle.
- Prediction not postdiction.
- Falsifiable.
- Possibly spanning many scales, systems - universality.
- Mathematizable.
- Constraints.
- Not everything is optimal - look where optimality matters… Behavior?

Life is…

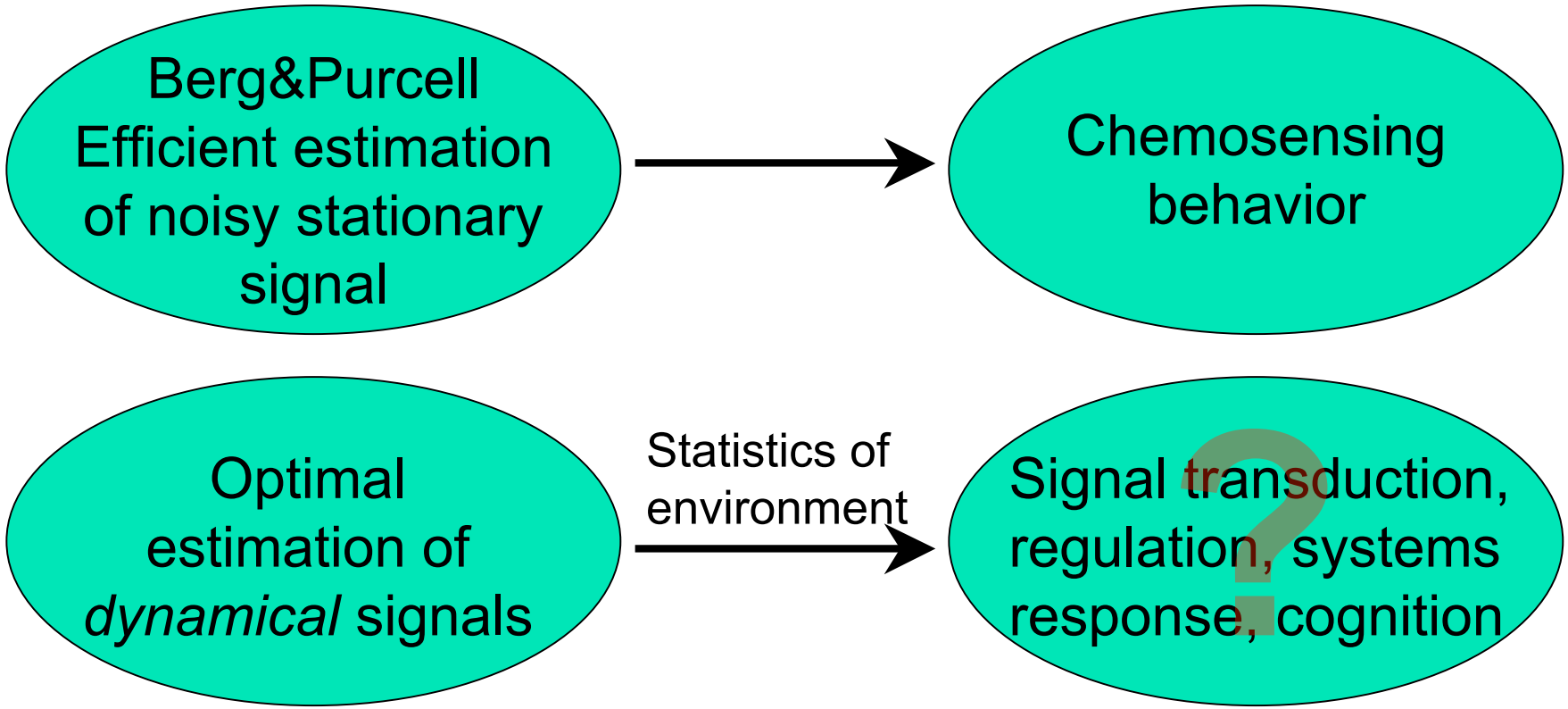

Physics of observation,
Limited resources
The environment
Statistics of environment,
Limited time, Relations
to complexity of the task
(Bialek, Nemenman,
Tishby, 2001)
Estimating (lossy encoding) the past
Predicting the future (RDT)
Limited resources
Biologically optimal action

Efficient estimation as a biological design principle

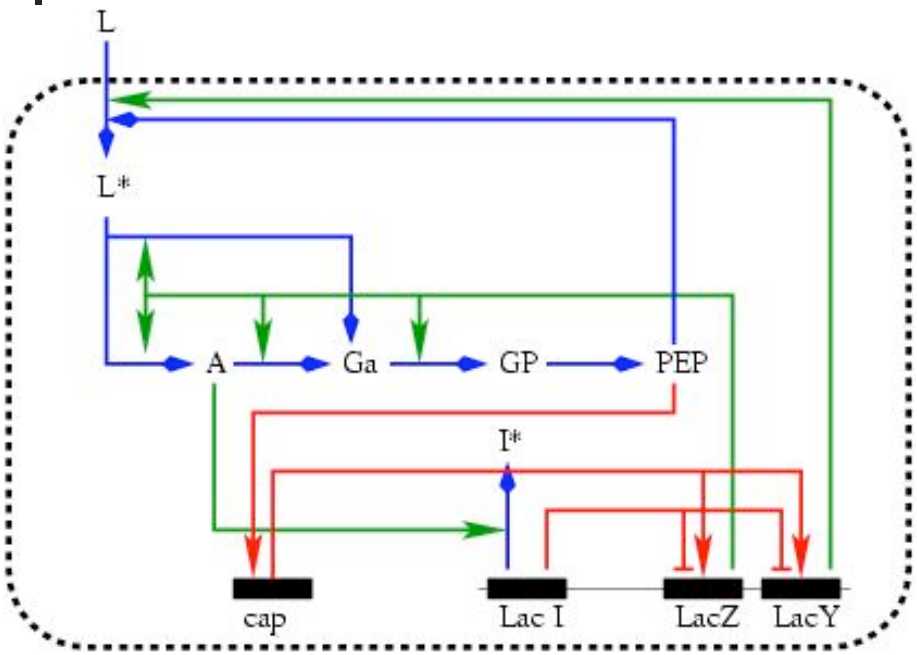


- Berg and Purcell (1977). Chemosensing precision and reliability is limited by physical noise sources.
- Since then: single photon responses, transcription, chemotaxis run length, motion estimation,… - all are at physical limits to sensing.
- The second arrow? (estimation of and reaction to a *dynamical* environment).

In time learning/prediction: *necessary* for active response



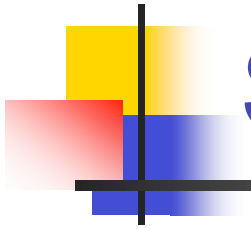
Lac and PTS: Do we understand?



- Very slow positive feedback (cap), ~1hr
- Slow positive feedback (lac I), ~10min
- Medium-fast positive feedback (PEP), ~10s
- Fast negative PEP feedback, ~100ms
- Very fast low pass filter (receptor), ~10ms

PEP - phosphoenolpyruvate

Why?



Statistics of environment?

- Long scale statistics of lactose food appearance - cap averaging over hours sets the mean operating point.
- Lactose appears with time scales of minutes and disappear in tens of minutes - PEP activation and lac shut-off.
- Chemotaxis(?)/bad mixing leads to higher concentrations on scales of seconds - PEP feedback.

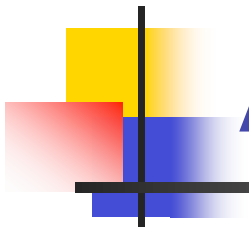
Statistics of environment?

- Negative PEP feedback at PEP saturation stabilizes energy production.
- Low path filtering at receptors removes statistical noise.

Maybe (near) optimal for this environment?

(with Wall, Bettencourt, Hlavacek)

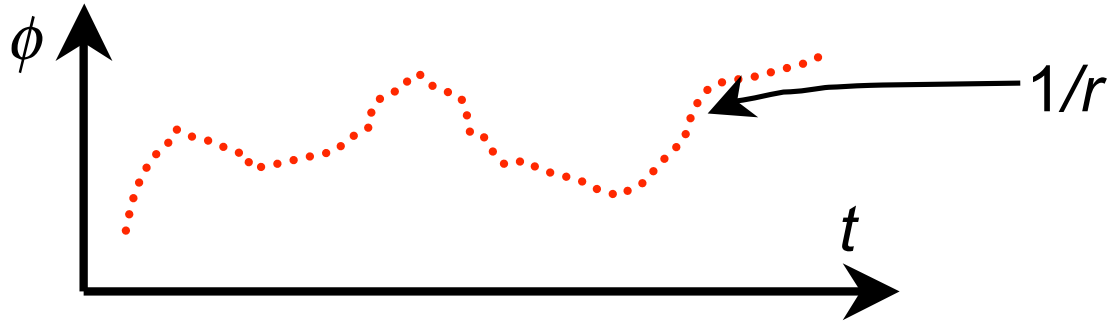
Mathematics of prediction: A limited form of prediction



- Estimation of dynamical signal “right now” (t=0) from observations of its past (t<0).
- Need to know time statistics of the signal.

Mathematics of prediction: A limited form of prediction

For a signal ϕ sampled at rate r and with $C(t) \xrightarrow[t\to 0]{} \sigma^2\left(1-\left(\frac{t}{\tau}\right)^{\nu}\right)$



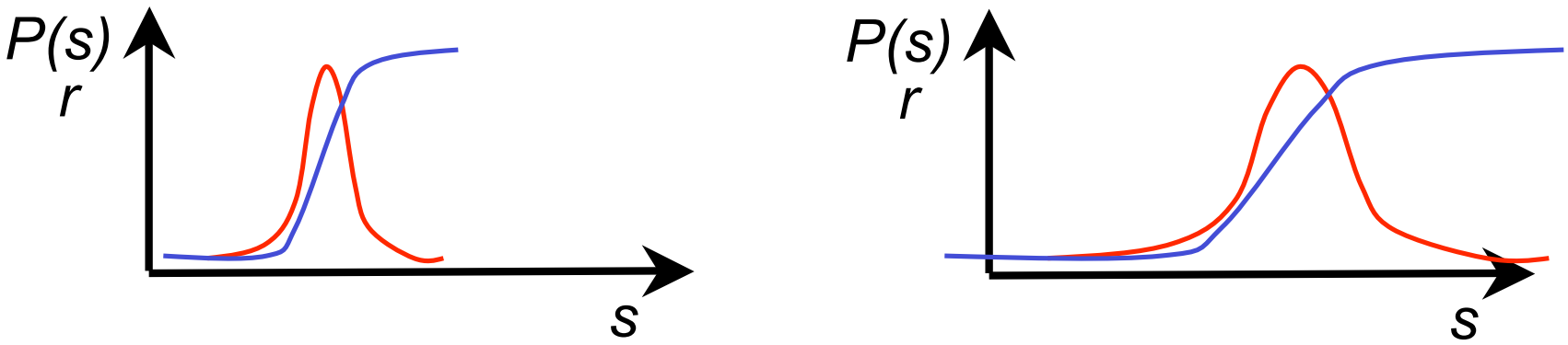
$$\left\langle \delta^2\phi \right\rangle \sim \frac{\text{noise}}{rt} + \left(\frac{t}{\tau}\right)^{\nu} = \min \quad \rightarrow \quad t \sim \left(\frac{\tau^{\nu}}{r}\right)^{1/(\nu+1)}$$

Var of the mean

Var of decorrelation

A note on optimal information transmission

$$P(s,r) \rightarrow I[P] = \left\langle \log \frac{P(s,r)}{P(s)P(r)} \right\rangle$$

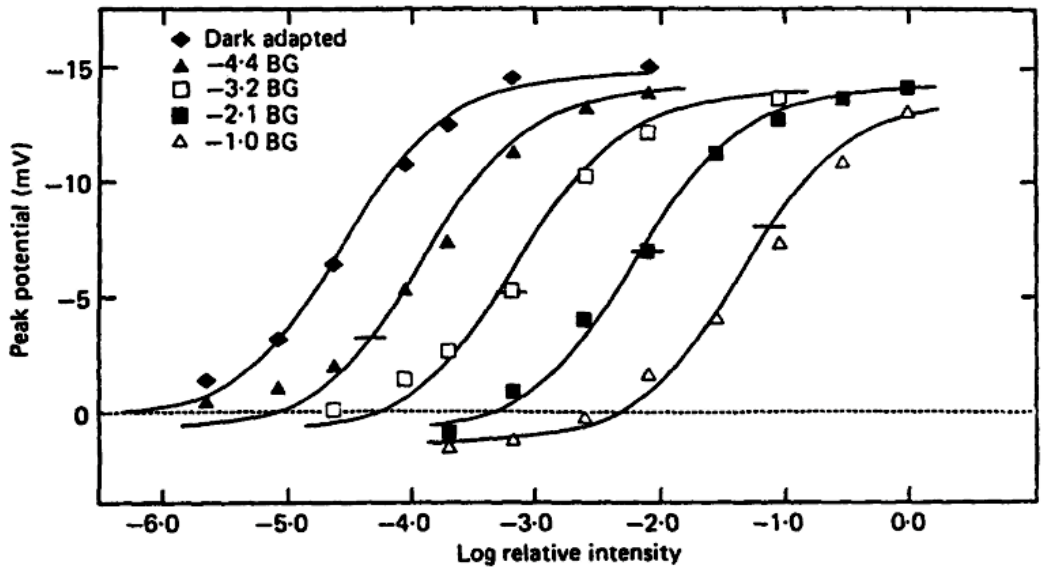


Matching mean and variance maximizes information transmission.

(Laughlin, 1981)

Turtle cone background light intensity adaptation

$$P(I) \propto \exp\left[-\frac{1}{2\sigma^2}\left(\log\frac{I}{I_0}\right)^2\right]$$



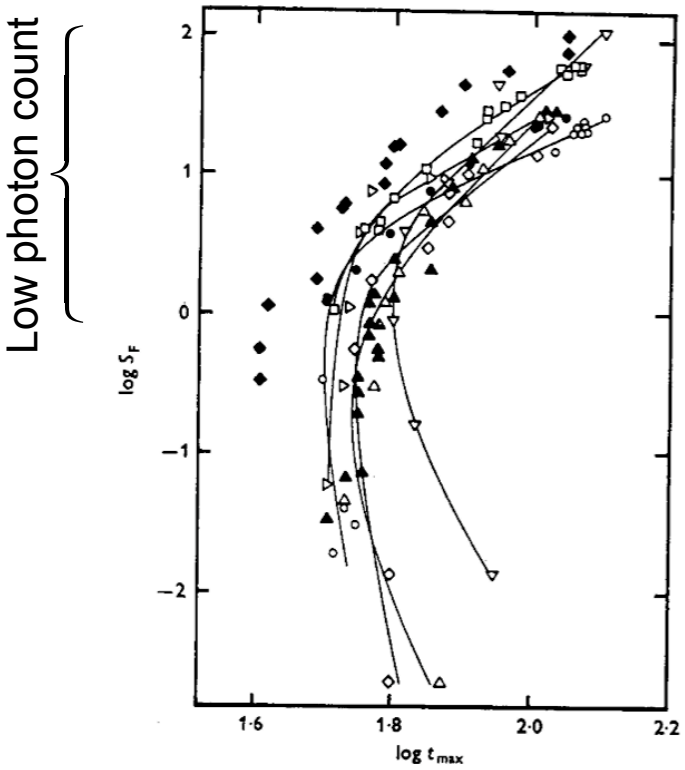
(Normann & Perlman, 1979)

Bckgr, log I	Adap, log I
dark	-4.4
-4.4	-3.8
-3.2	-3
-2.1	-2.3
-1	-1.3

$$I_a \propto I_0^{0.73}$$

Bad!

Response time adaptation



(Baylor & Hodgkin, 1974)

$$\tau \propto I_0^{1/3\cdots1/5}$$

Other animals range from 1/2 to 1/7.

$$I_a \propto I_0 \tau$$

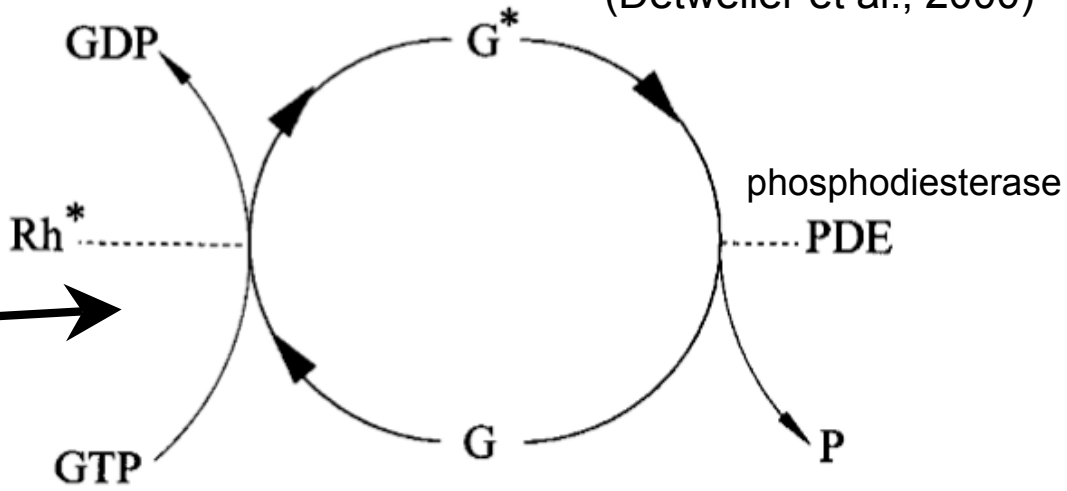
Probably not a coincidence:
Adapting to integrated flux.

What should τ be?

(Detweiler et al., 2000)

Cone: 3 low pass filters (at least):

- $\gamma + Rh \rightarrow Rh^*$
- $Rh^* \rightarrow PDE^*$
- $PDE^* \rightarrow GC$



$$\tau_R \frac{d\delta R}{dt} = -\delta R + g_R\left(\phi(t) + \eta_\phi(t)\right), \quad \phi = \log\frac{I}{I_0}, \left\langle \eta_\phi(t)\eta_\phi(0)\right\rangle = 1/I_0\delta(t)$$

$$\tau_P \frac{d\delta P}{dt} = -\delta P + g_P\left(\delta R + \eta_R(t)\right), \qquad \ldots$$

…

Linear due to Ca feedback!

Solution (for signal-limited precision)

$$I_0 \equiv I[\phi(t=0); v(t=0)] = \log \frac{\langle \phi^2 \rangle}{\langle \phi^2 \rangle - \frac{\prod g_i^2 \langle \phi_f^2 \rangle^2}{\langle v^2 \rangle}}$$

Note that this is not same as

$$I[\phi(t); v(t)] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T/2}^{T/2} \frac{d\omega}{2\pi} \log\left(1 + SNR(\omega)\right)$$

which is the channel capacity.

Solution (for signal-limited precision)

$$I_0 \equiv I[\phi(t=0); v(t=0)] = \log \frac{\left\langle \phi^2 \right\rangle}{\left\langle \phi^2 \right\rangle - \frac{\prod g_i^2 \left\langle \phi_f^2 \right\rangle^2}{\left\langle v^2 \right\rangle}}$$

Can also maximize total predictive information:

$$I\left[\phi(t>0); v(t\le 0)\right]$$

Finding τ

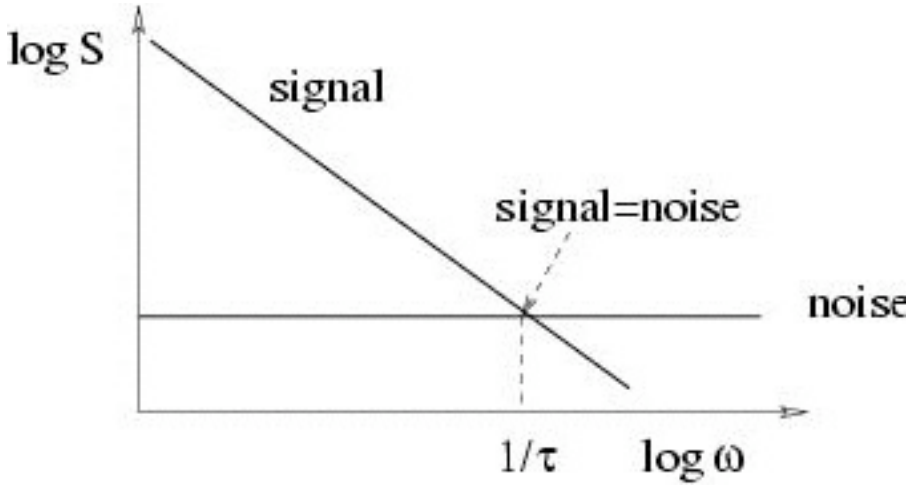
Maximize I_0 *w.r.t* τ

For:

$$S_\phi(\omega) \xrightarrow[\omega\to\infty]{} \omega^{-\alpha}$$

get:

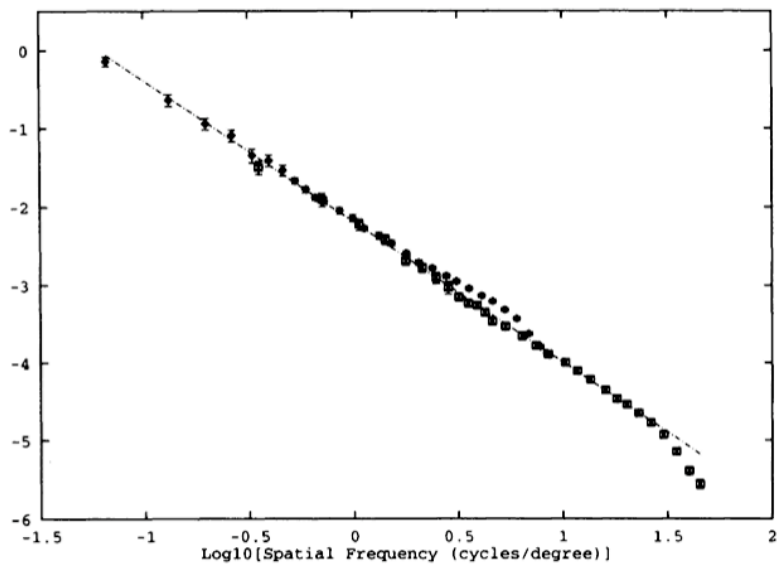
$$\tau \sim I_0^{-1/\alpha}$$



Best possible matched filter
(limited by biochemical mechanisms)

Also predicted by variance balance argument.

A problem



(Ruderman & Bialek, 1994)

- $1/k^{2-\varepsilon}$ spatial spectrum
- ~10 phoreceptors/fixation drift
- $1/\omega^{2-\varepsilon}$ temporal spectrum
- Should have $\tau \sim I_0^{-1/2}$

Wrong! But…

Structural constraint

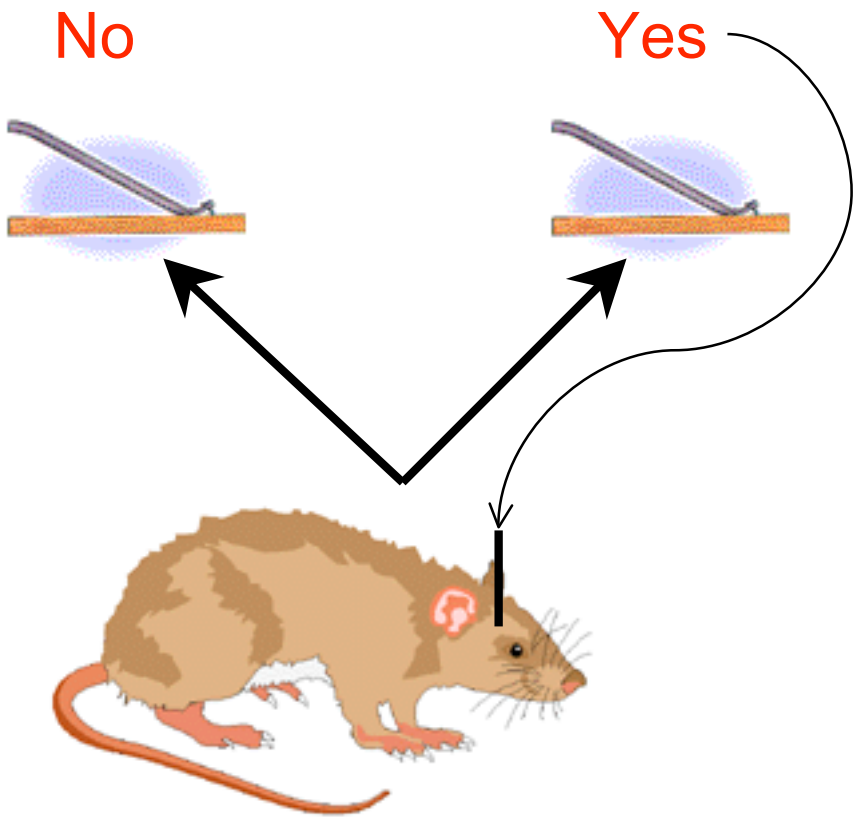
Rh* is the signal (for the adapting rest of the circuit), its temporal response is uncontrollable
(and badly known - Rieke & Baylor, 1998)

$$S_{Rh^*} \sim \frac{1}{\omega^2} S_{\phi} \sim \frac{1}{\omega^4}$$

Given this signal, the rest of the biochemistry should adapt in agreement with experiment

$$\tau \sim I_0^{-1/4}$$

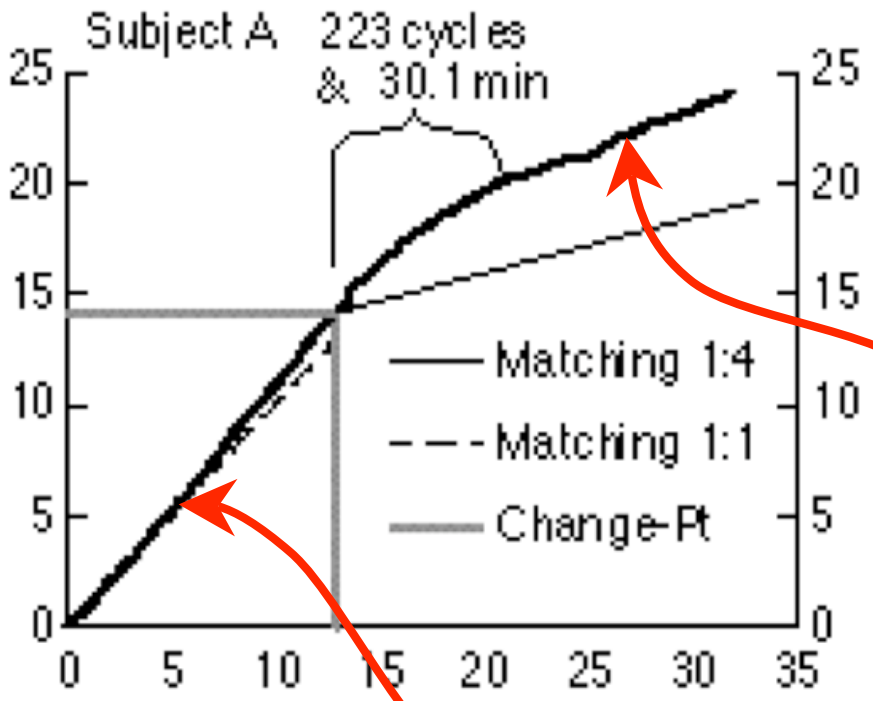
Rat matching experiments



(with Gallistel)

- Poisson deposition of rewards
- Rewards do not accumulate
- Possibly variable rate
- Changeover delay
- Rat matches

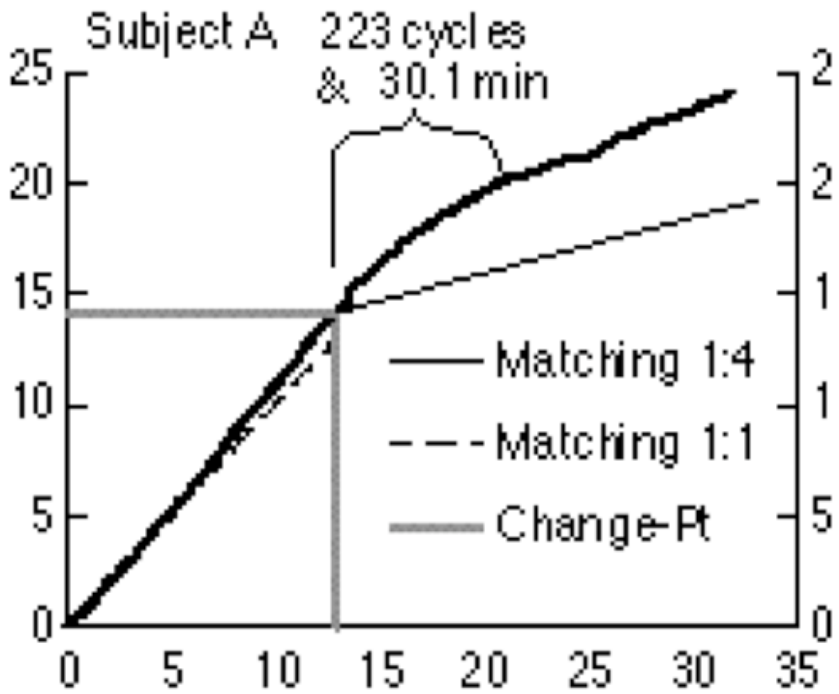
Rat matching experiments



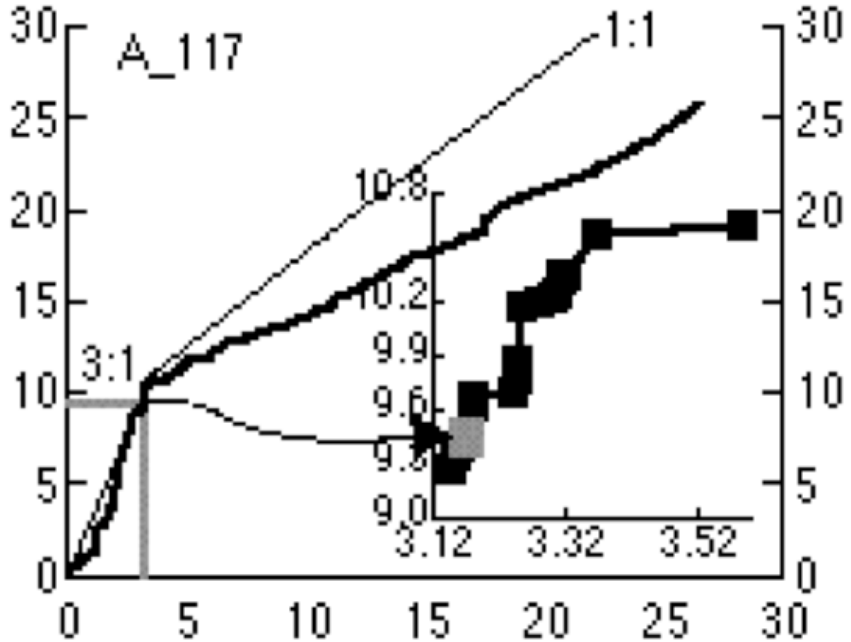
(Gallistel et al 2001)

- Poisson deposition of rewards
- Rewards do not accumulate
- Possibly variable rate
- Changeover delay
- Rat matches

But: Time scales are history dependent. Can we explain?



$\tau = 30 \ldots 100 + \text{min}$



$\tau = 1 \ldots 2 \,\text{min}$

(also note imperfect matching)

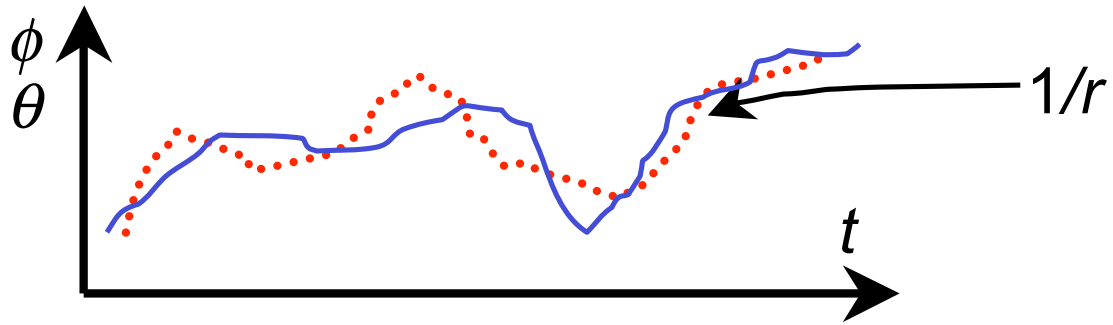
Optimal estimation: Bayes theorem

$$P[\phi(t) \mid \text{data}] = \frac{P\left(\text{data} \mid \phi(t)\right) P[\phi(t)]}{Z}$$

$$P[\phi(t)] \propto \exp\left[-\frac{l}{2}\int (\partial_t \phi)^2 dt\right], \quad \phi(t) = \log r(t)$$

Sampling

Sampling from a target $\theta(t) \rightarrow \phi_i \quad \text{at} \quad t_i$



Log-likelihood

$$P[\phi(t) \mid \{x_i\}, \{t_i\}] \propto \exp\left[-\frac{l}{2} \int dt \, (\partial_t \phi)^2 - \sum_i V(\phi_i - \phi(t_i)) \right]$$

Evolution in a random (time and space) potential

A better solution (WKB): Learning a Poisson variable

Bialek, Callan, & Strong, 1996, Nemenman and Bialek, 2002

$$\tau = \sqrt{l / e^{\theta}}, \qquad \left\langle \delta^2 \phi_0 \right\rangle \approx \frac{1}{4} \frac{1}{\sqrt{e^{\theta} l}}$$

Time scales

Correlation time: $\tau \propto \sqrt{l / r}$

For stable period $(\tau_0 \sim 1\,\mathrm{hr})$:

$$l \approx 3 \cdot 10^6\, s, \quad r \approx 1/10 s^{-1}, \quad \tau \sim 5 \cdot 10^3\, s \approx 1.5\,\mathrm{hrs}$$

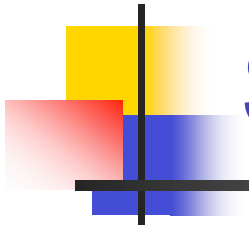
For variable schedule $(\tau_0 \sim 1\,\mathrm{min})$:

$$l \approx 1800\, s, \quad r \approx 1/10 s^{-1}, \quad \tau \sim 130\, s \approx 2\,\mathrm{min}$$

For monkeys (Sugrue et al, 2004) $(\tau_0 \sim 17\,\mathrm{samples})$:

$$l \approx 300 / r, \quad r, \quad \tau \sim 15\,\mathrm{samples}$$

Importantly, estimate starts to change immediately in both cases

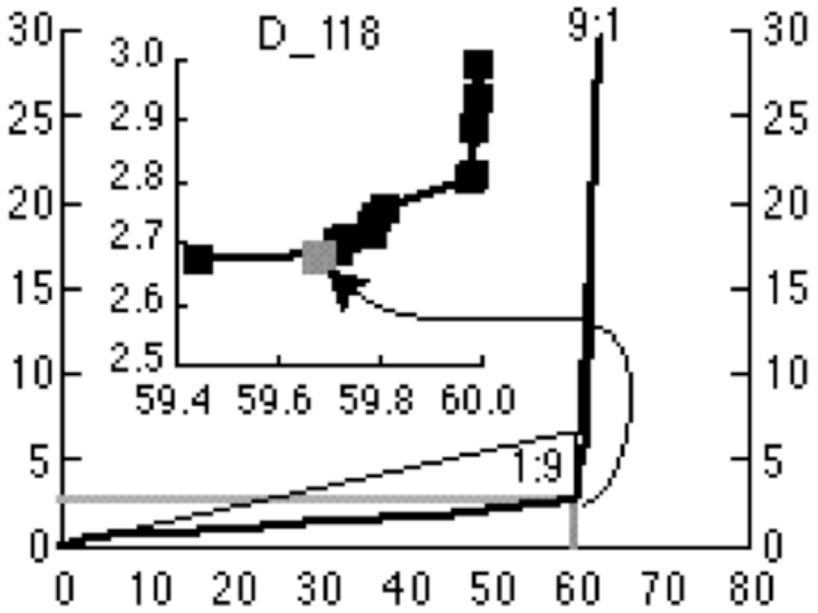


Self-consistent estimation of *l*

Averaging over $P[l]$ leads to correct estimation of the smoothness scale for fixed l (Nemenman and Bialek, 2002).

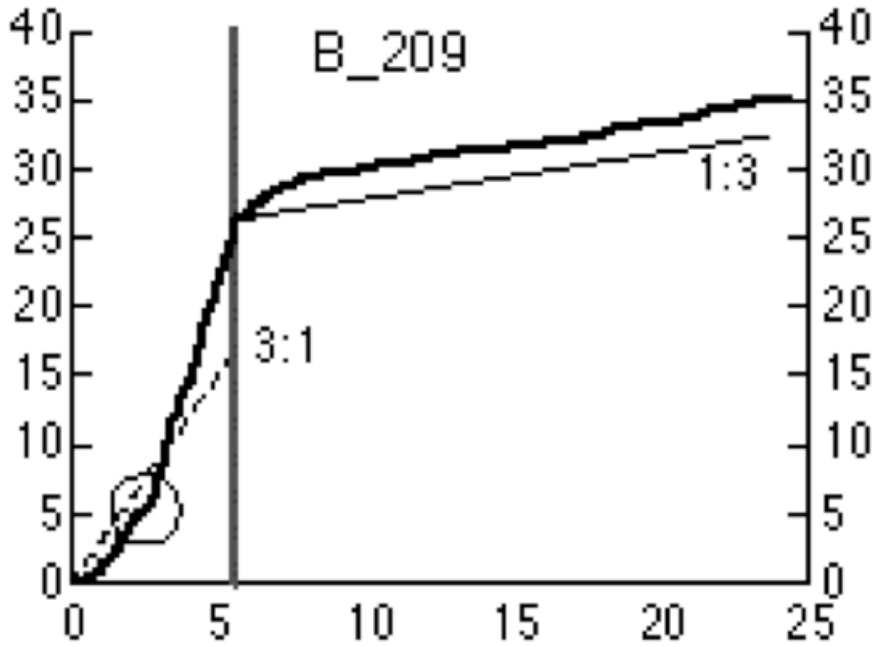
Can do the same for dynamic l.

Phenomenology: Abrupt changes



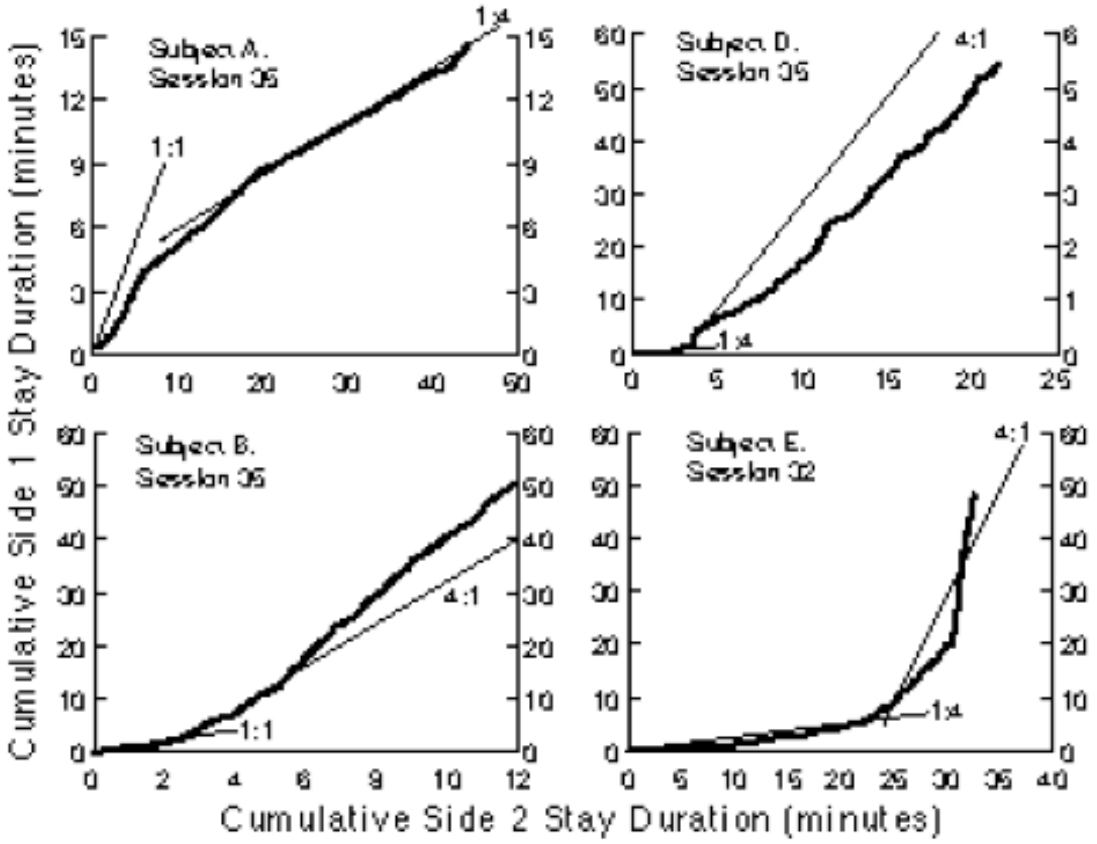
- Only after a few changes have been experienced
- Common during fast changes epochs

Phenomenology: Abrupt changes

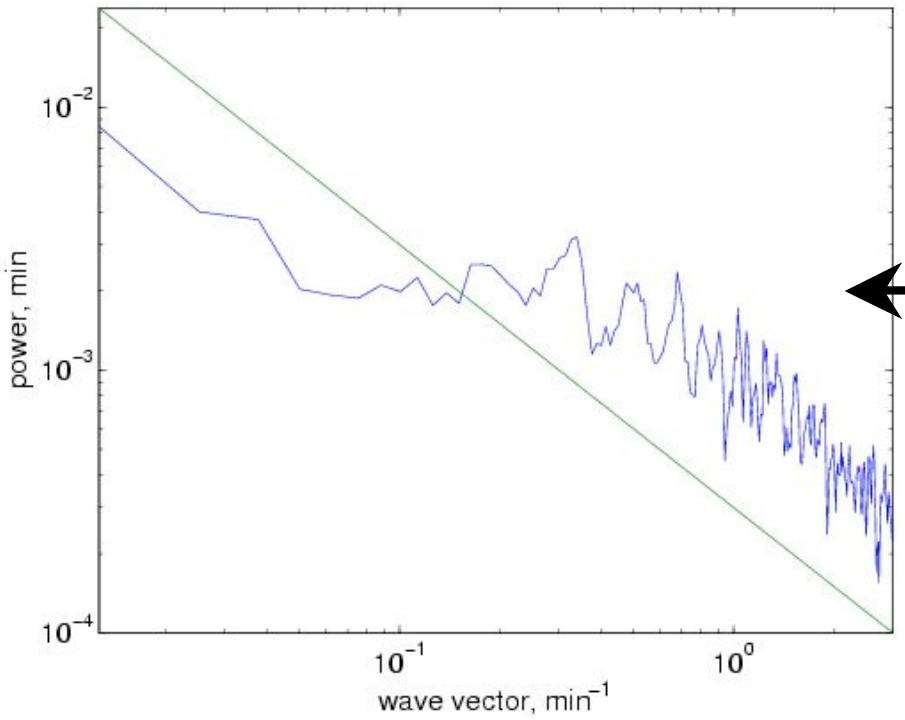


- Only after a few changes have been experienced
- Common during fast changes epochs
- (Metastable states)?

Phenomenology: Reversal to status quo ante

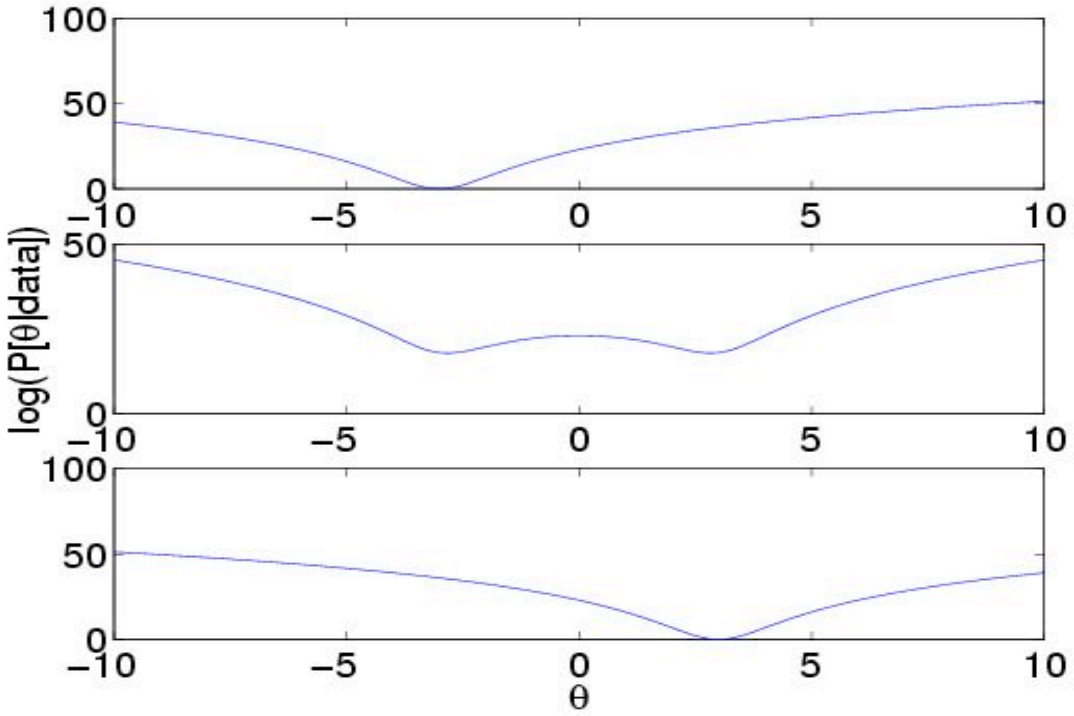


Caused by memory



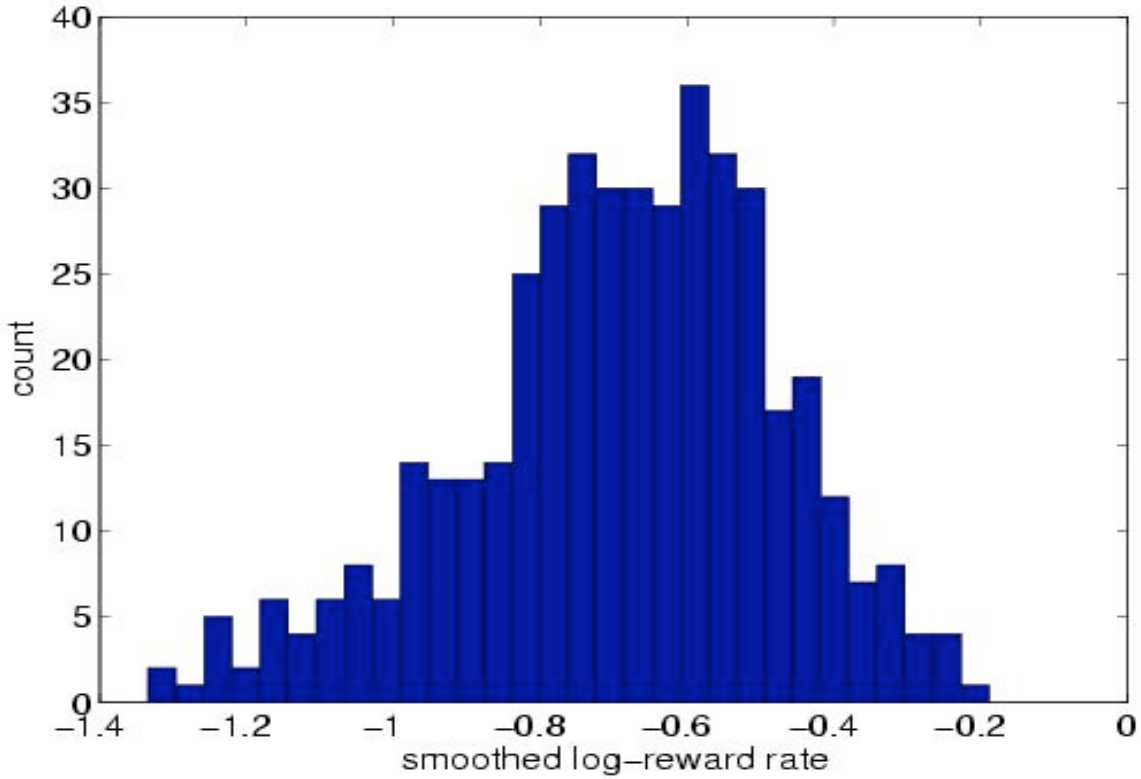
- Overestimation of rate immediately leads to higher rate and persists
- Power spectrum of reward histories
- Two regimes clearly seen
- Peak at 0 - long range correlations

Abruptness, two time scales, and non-Gaussianity



Critical periods?

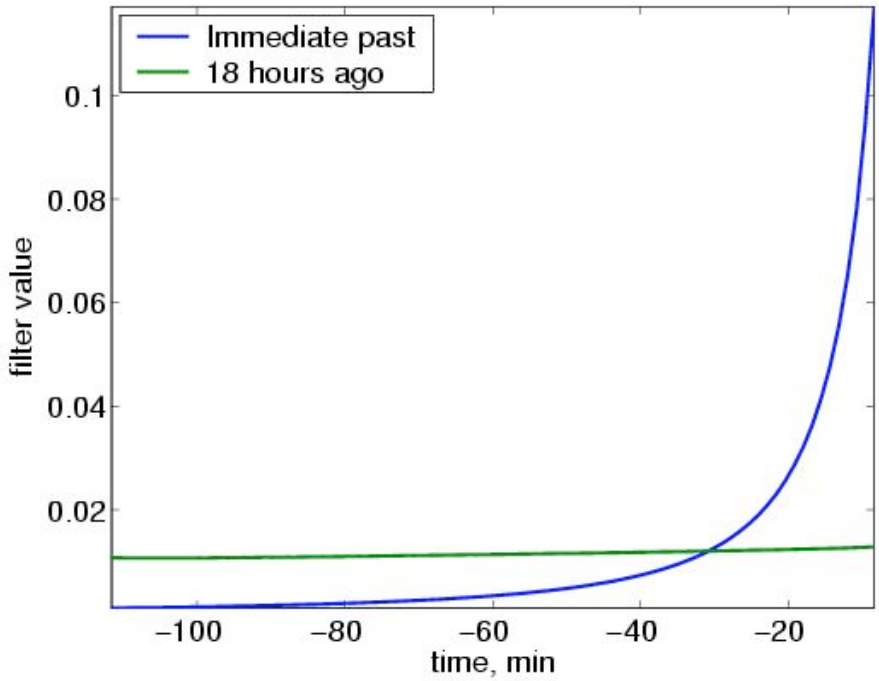
Non-Gaussianity of rate distribution



Modeling reversals: long range correlations

Bialek & Zee, 1990 - Best estimate of ϕ is approximated by

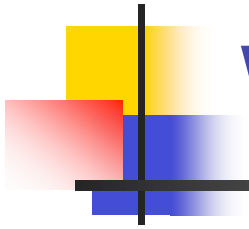
$$\phi(t) = \sum_{t_i < t} F(t_i - t)$$



Optimal *F(t)* for a Gaussian process with $C \sim t^{-2}$ for a range near *t=0* and *t=18*hrs (normalized within the window).

Long-tailed filters explain reversal

- At the end of the session, rate estimates are effected mostly by the last (post-change) observation
- After a long delay, pre-change and post-change observations are almost equally weighed, but there are much more of the former.
- Wouldn't work for exponential filters as used by Sugrue et al, 2004.
- Experiments to measure *C(t)* are now done.



Why matching?

- Matching is almost optimal for maximizing reward.
- Matching is almost optimal for tracking rate changes.
- Can it be that the bit value of a reward is higher than its food value? (Rats are curious!)
- Preliminary report: matching for accumulating rewards. Planning experiments to test matching to neutral stimuli.

Take home message:

- Optimal estimation of dynamic world seems to explain phenomenology from molecular scales, to cognitive psychology scales.
- Preliminary experimental comparisons.
- Better experiments are being done / are sought.
- For molecular networks, relation of phenomenology to structure waits to be analyzed.