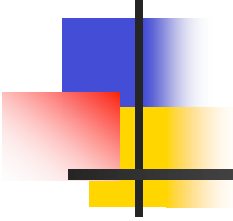# Statistical physics of information processing in biological systems

Ilya Nemenman

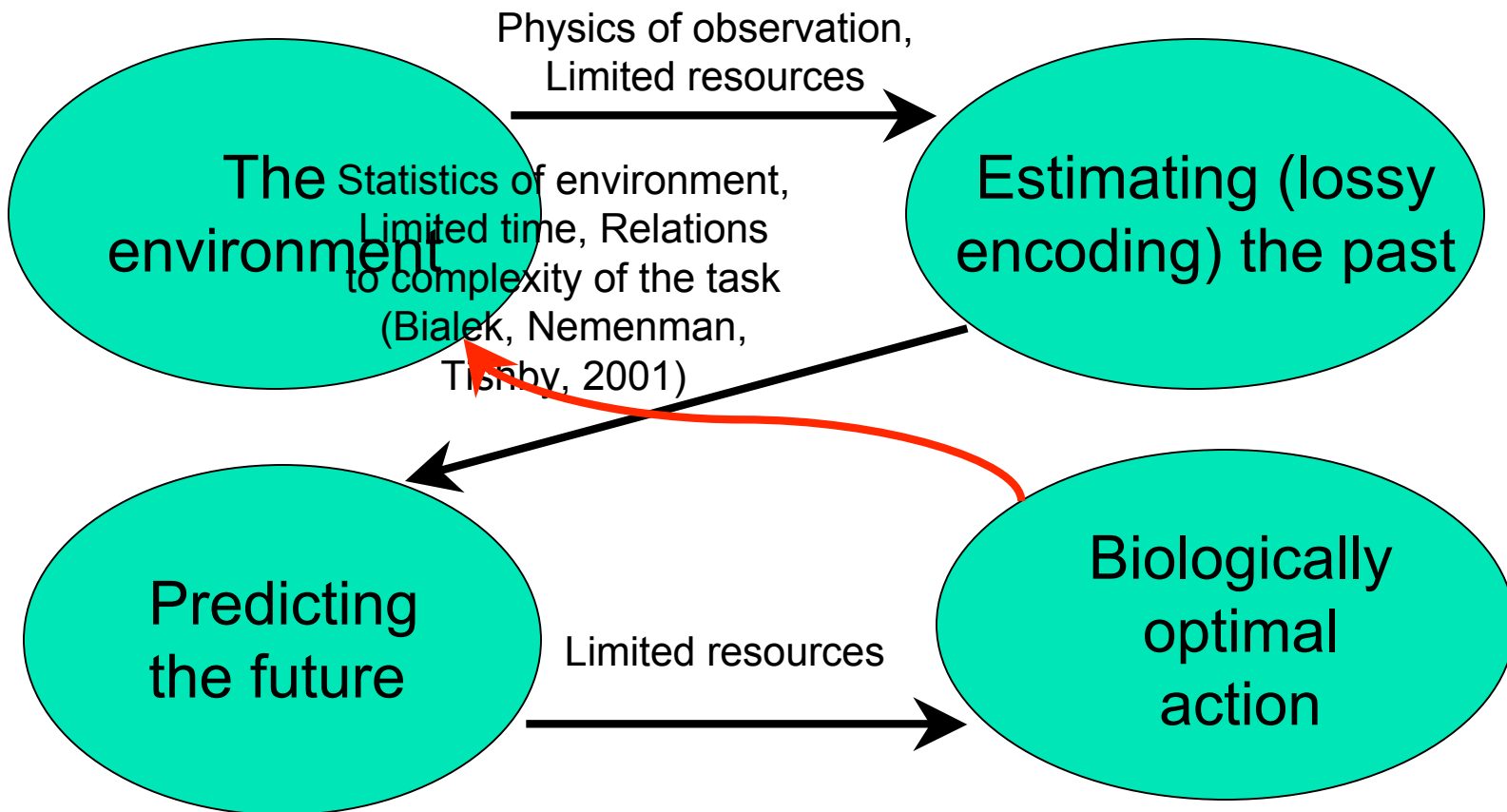Joint Centers for Systems Biology

Columbia

What is theoretical biophysics?

- Risk of answering just (soft, hard condensed matter) physics questions about biological materials/systems.

- Will then know how things happen, not why they happen.

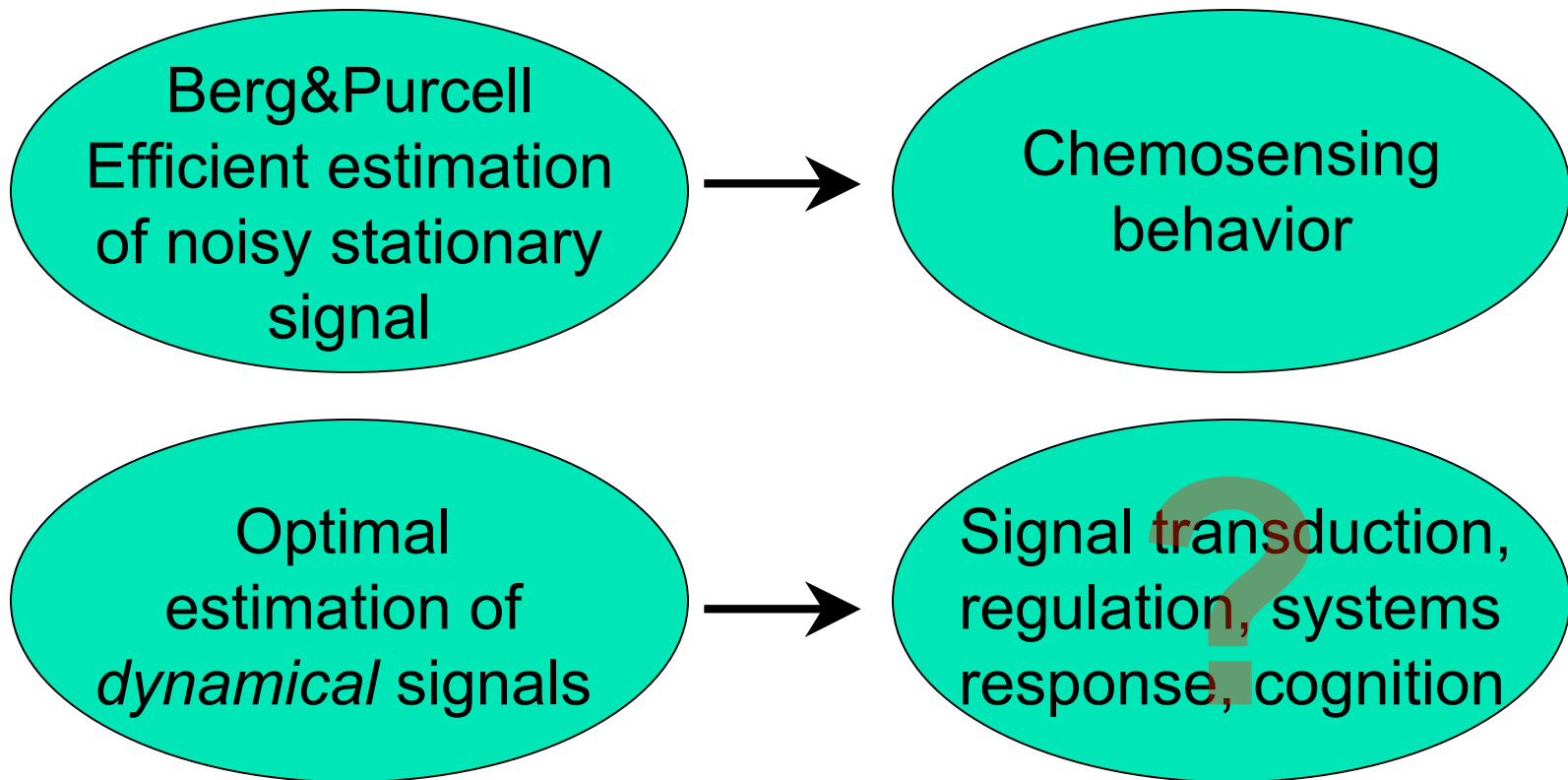- To be sure: ask questions relevant to behavior.

Life is…

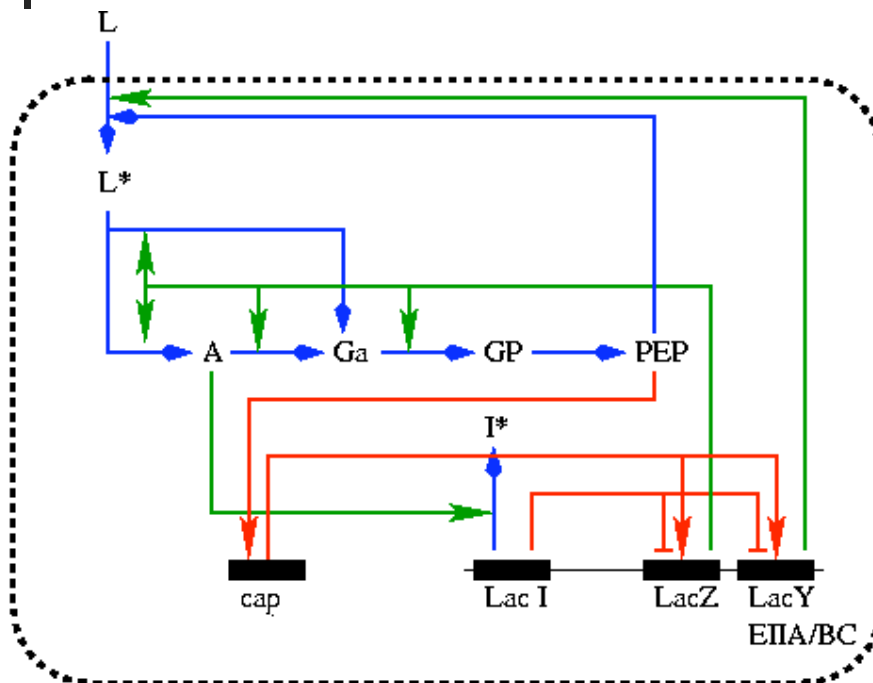# Efficient estimation as a biological design principle

- 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 in insects, chemotaxis network design - all are at physical limits to sensing.

- Many behaviors and designs need consideration of the remaining arrows (estimation of and reaction to a dynamical stimulus).

Fast learning and active response?

Berg&Purcell Efficient estimation of noisy stationary signal → Chemosensing behavior

Optimal estimation of *dynamical* signals → Signal transduction, regulation, systems response, cognition ?

Lac operon and phoshotransferase system

PEP - phosphoenolpyruvate

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

Why?

Statistics of environment

- In a newborn baby (because of the milk-enriched diet) and a non-European adult (because of a certain mutation), lactose enters lower small intestines every few hours - cap averaging sets the mean operating point.

- Bursts appear with time scales of minutes and disappear in tens of minutes - PEP activation, and lac shut-off.

- Chemotaxis leads to higher concentrations on scales of seconds - PEP feedback.
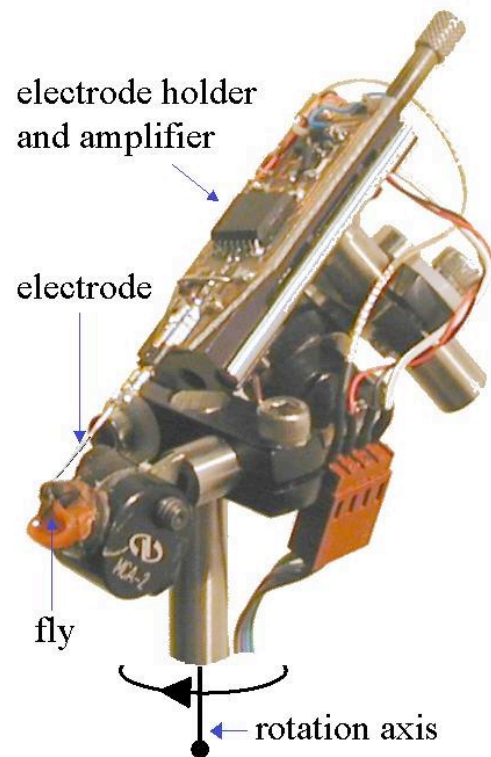
Statistics of environment

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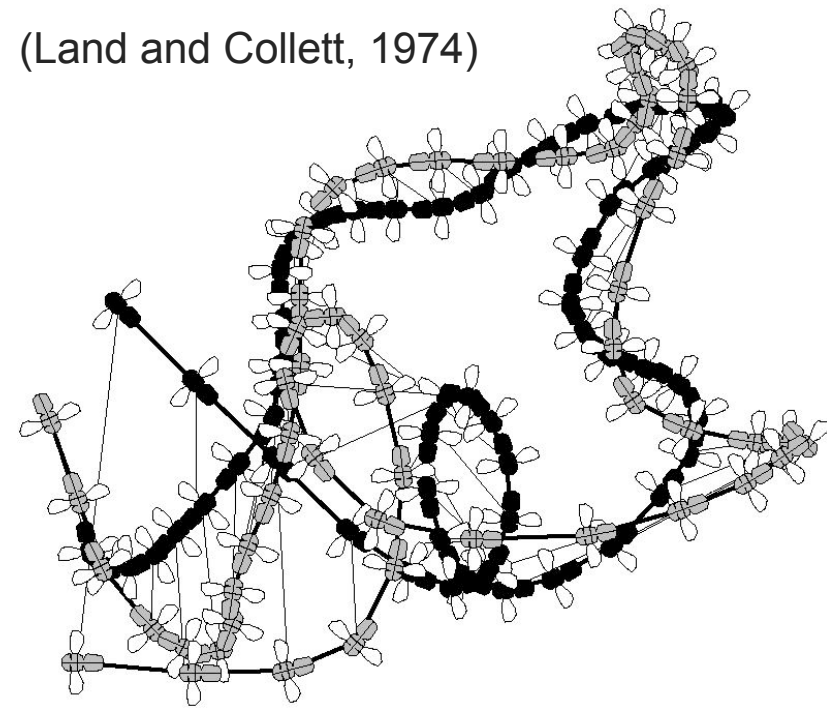- Low path filtering at receptors removes statistical noise.

Maybe (near) optimal for this environment?

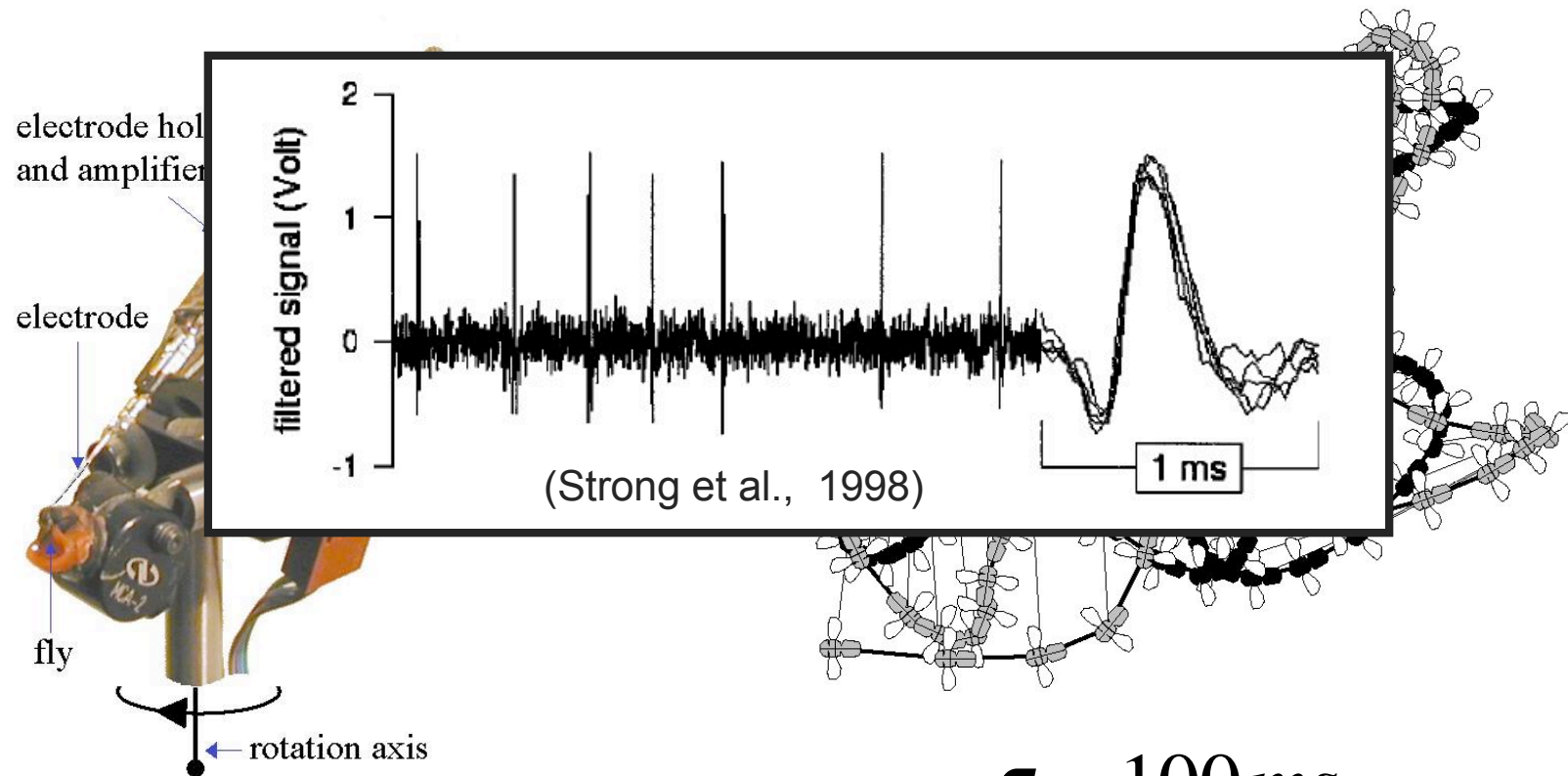(with Wall, Bettencourt, Hlavacek)
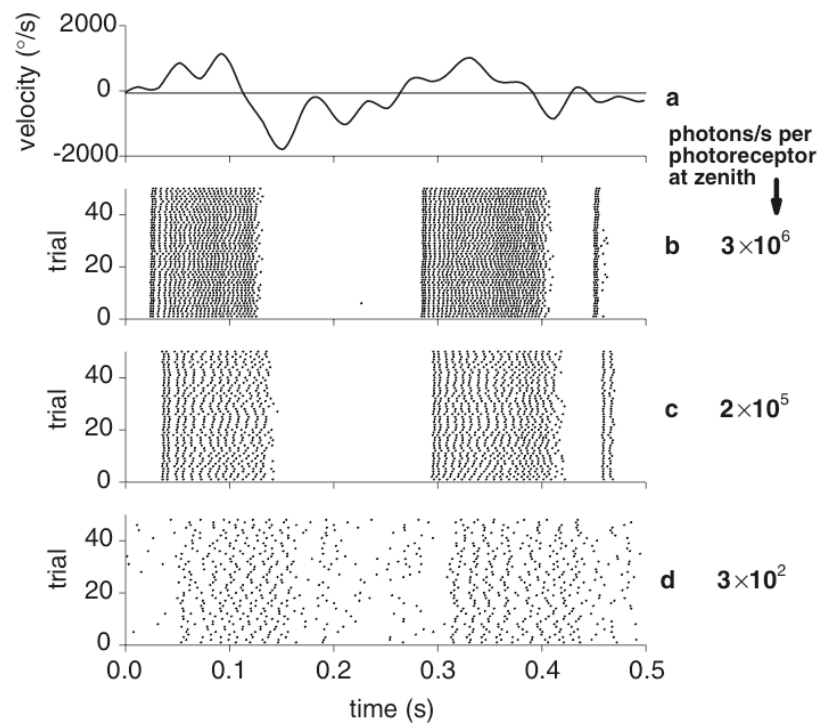
Motion estimation in fly H1

(Lewen et al, 2001)

(Land and Collett, 1974)

$$\tau = 100\,ms$$

Diffusion in a quadratic potential

Motion estimation in fly H1

(Strong et al., 1998)

$$\tau = 100\,ms$$

Precision is limited by physical noise sources

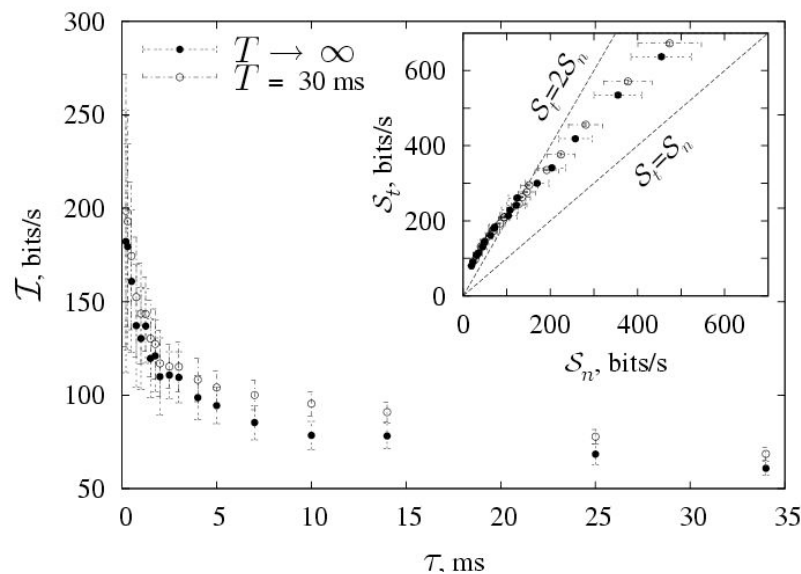(Lewen, et al 2001)

In an upcoming paper we show that

$$I(s,r) = S(r) - \langle S(r \mid s) \rangle_s$$

$$= \lim_{T \to \infty} \frac{1}{2} \sum_{\omega} \log[1 + SNR(\omega)]$$

also grows with the light intensity.

Not just optimal estimation: just in time estimation!

(IN, Bialek, de Ruyter van Steveninckl, 2005)

- An effort to make information available "now"
- Information grows with resolution

Why does the fly use such high resolution and short observation times to code this relatively slow signal?

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.

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

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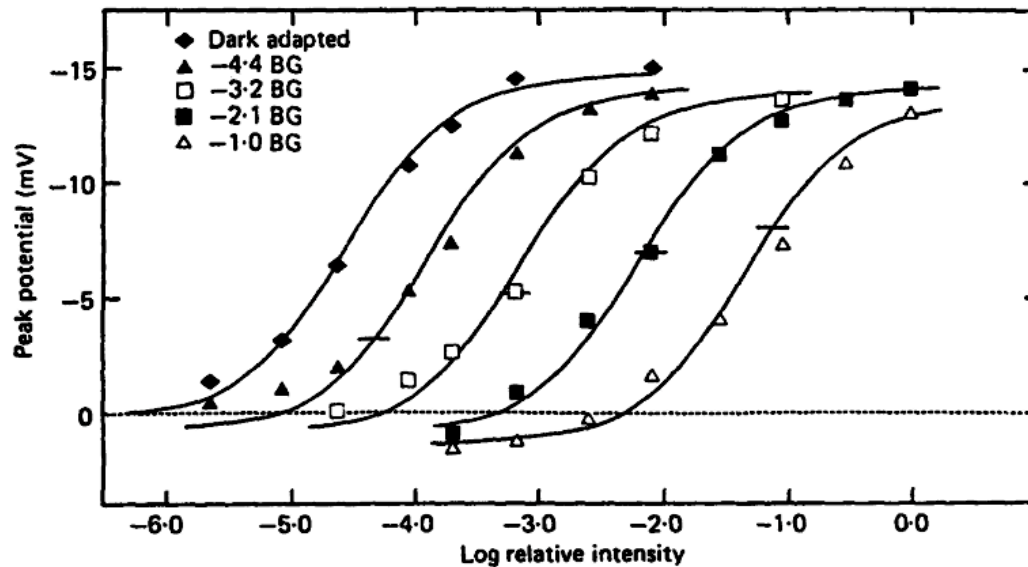
Var of decorrelation

Var of the mean

Motion prediction by fly

- Receptor delay (sampling) ~10ms
- Correlation time 100ms
- Efficient estimation ~30ms (also behavioral response time of the fly)
- Thus ability to send most info in 30ms windows is not surprising
- For 30ms windows, coding at <1ms may be needed.
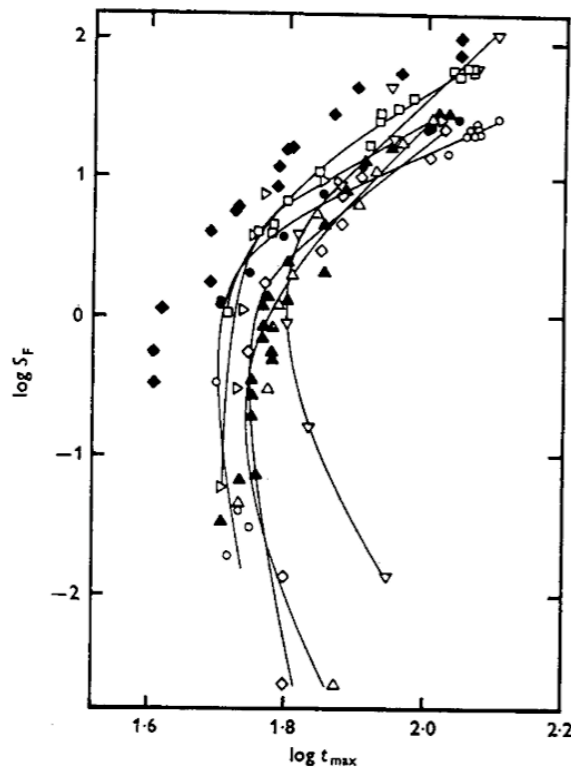
Turtle cone background light intensity adaptation

(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

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

Response time adaptation

(Baylor & Hodgkin, 1974)

$$\tau \propto I_0^{1/3\cdots1/5}$$
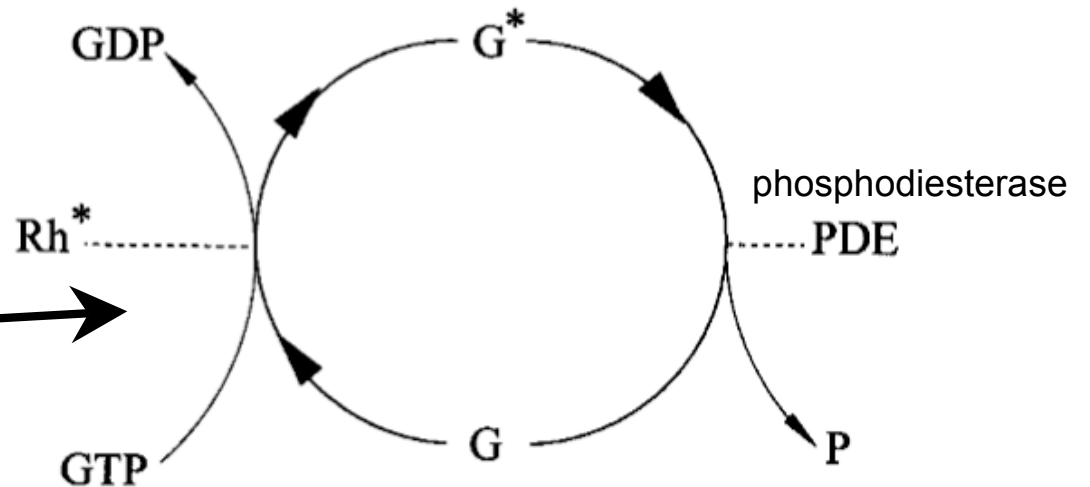
Other animals range from 1/2 to 1/7.

$$I_a \propto I_0\tau$$

Probably not a coincidence:
Adapting to integrated flux.

What should τ be?

Cone: 3 low pass
filters (at least):

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

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

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

$\ldots$

Solution
(for signal-limited precision)

Represent SDE's as SFT, integrate out all fields except fields at moment $t=0$. Get

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

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)

Represent SDE's as SFT, integrate out all fields except fields at moment $t=0$. Get

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

Can also maximize total predictive information:

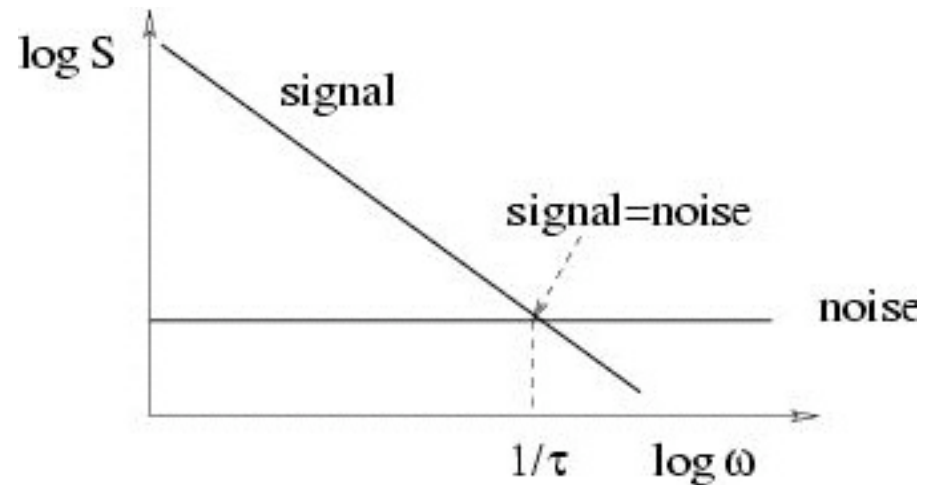$$I\left[\phi(t>0); v(t \leq 0)\right]$$

Finding τ

Maximize I_0 w.r.t τ

For:

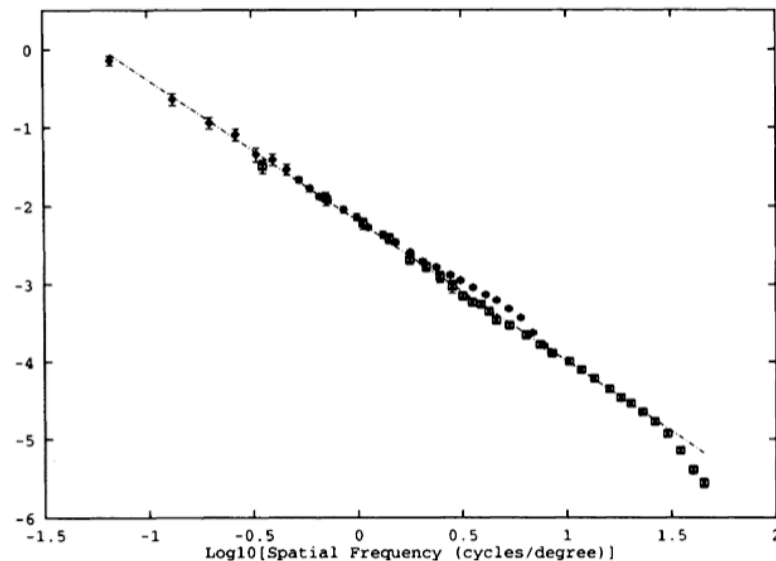$$S_\phi(\omega) \xrightarrow[\omega \to \infty]{} \omega^{-\alpha}$$

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 $\tau \sim I_0^{-1/2}$

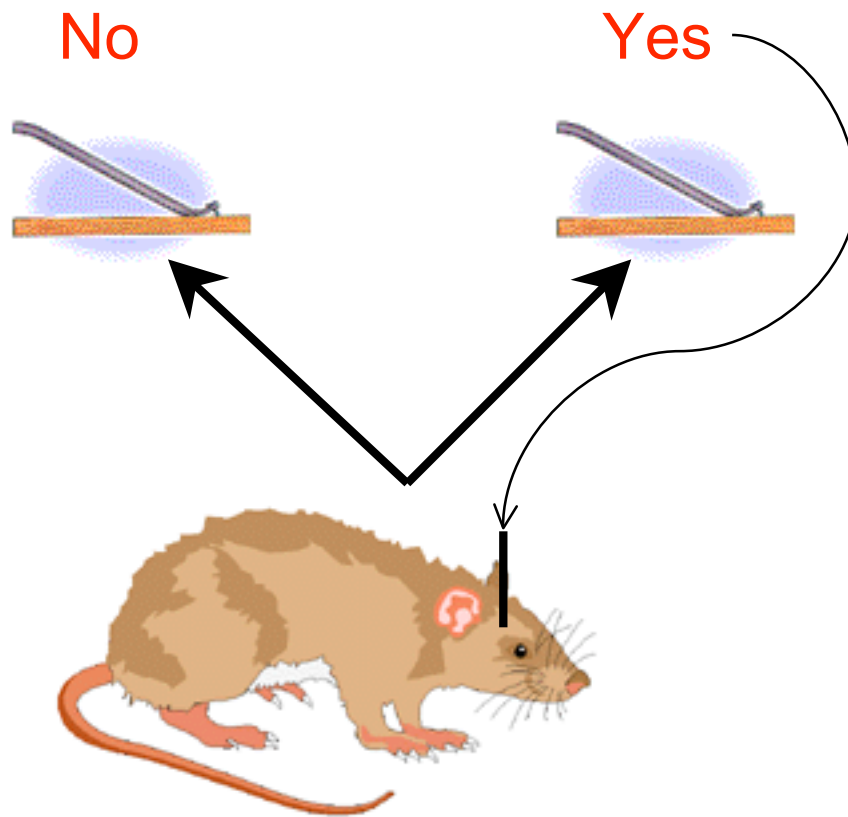Wrong! But…

Biochemical constraint

Rh* is the signal, 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
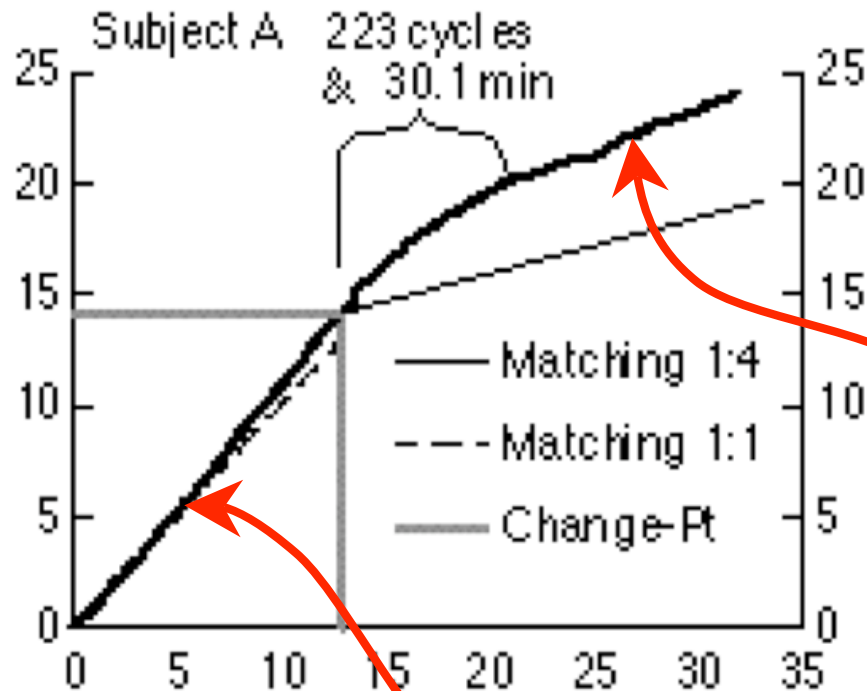should adapt in agreement with experiment

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

Rat matching experiments

No   Yes

(with Gallistel)

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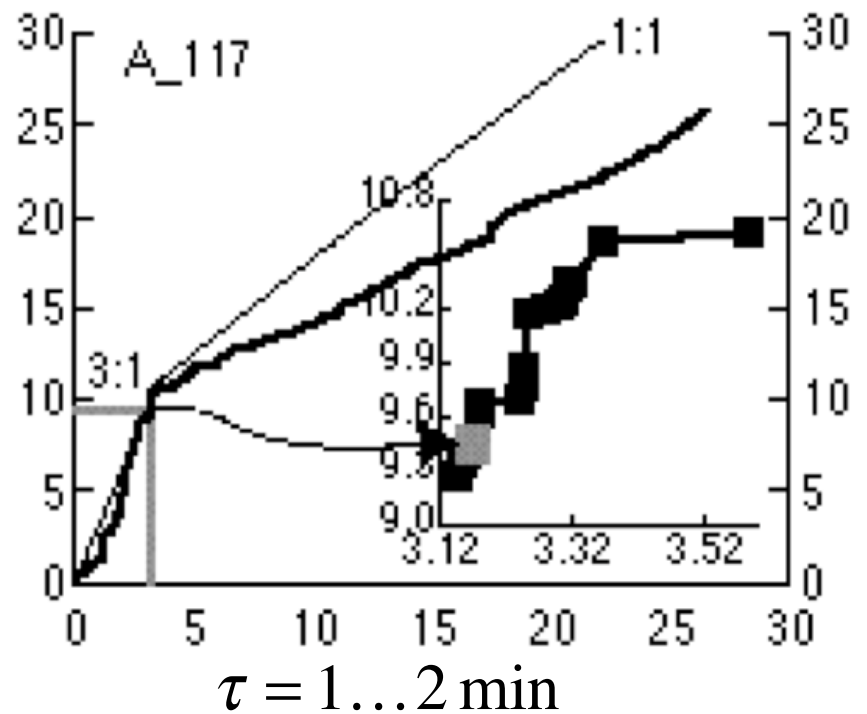- Possibly variable rate
- Changeover delay
- Rat matches
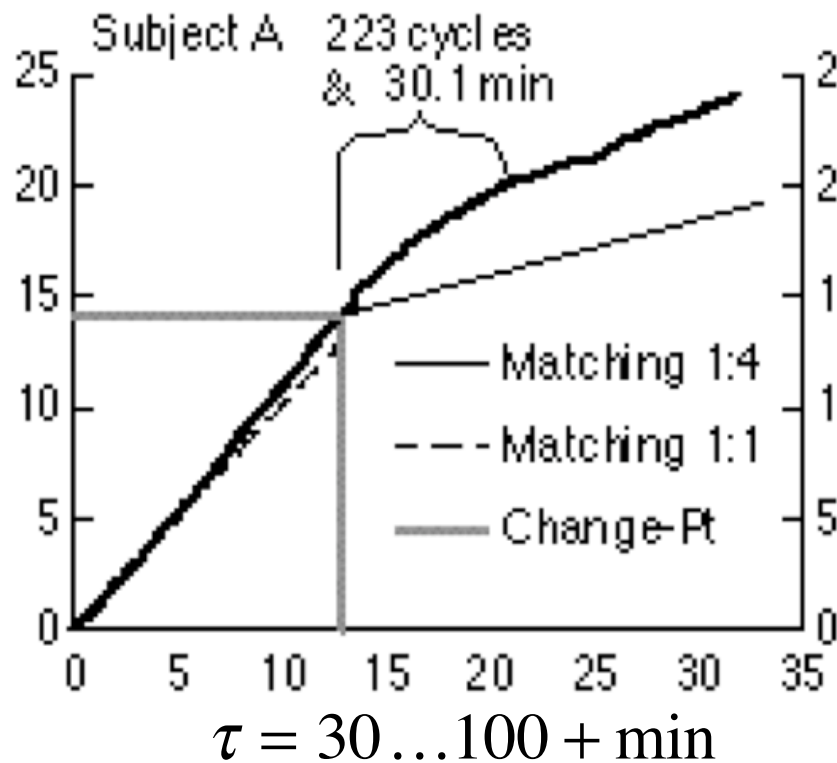
Rat matching experiments

Subject A 223 cycles & 30.1 min
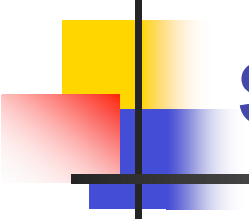
Matching 1:4
Matching 1:1
Change-Pt

(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?

$$\tau = 30\ldots100 + \min$$

$$\tau = 1\ldots2\min$$

(also note imperfect matching)

Can we explain the time scales

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

Sampling from a target $\theta(t) \to \phi_i$ at t_i

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

Evolution in a random (time and space) potential

Examples of potentials

Poisson process:

$$\sum_i V_i = \int dt\, e^{\phi(t)} - \sum_i \phi(t_i) \quad \rightarrow \quad \int dt\, e^{\phi(t)} - \int dt\, e^{\theta(t)} \phi(t)$$

Learning a fluctuating mean of a Gaussian:

$$\sum_i V_i = \sum_i \frac{1}{2\sigma^2}\left(\phi(t_i) - \phi_i\right)^2 \quad \rightarrow \quad \frac{1}{2\sigma^2} \int dt\, r(t)(\phi - \theta)^2$$

In general, for fast sampling rate:

$$\sum_i V_i \quad \rightarrow \quad \int dt\, r(t)\, V(\phi - \theta)^2 \quad \rightarrow \quad r \int dt\, V(\phi - \theta)^2$$

Learning a Poisson variable

$$l \frac{d^2 \phi_0}{dt^2} = e^{\phi_0} - \sum_i \delta(t - t_i)$$

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

$$\langle \delta^2 \phi_0 \rangle \approx \frac{1}{4} \frac{1}{\sqrt{e^\theta l}}$$

(WKB, for large $e^\theta l$)

Compare to the crude estimate above:

$$\tau \sim \sqrt{\frac{l}{e^\theta}}, \quad \langle \delta^2 \phi_0 \rangle \sim \frac{1}{\sqrt{e^\theta l}}$$

Time scales

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

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

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

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

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

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

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

Importantly, estimate starts to
change immediately in both cases

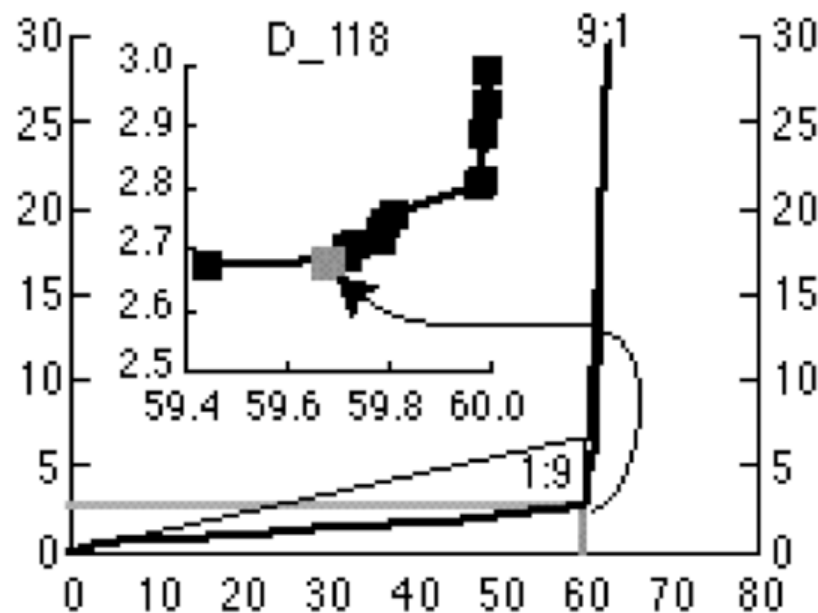# Self-consistent estimation of *l*

Averaging over $P[l]$ leads to correct estimation of the smoothness scale for fixed l (Nemenman and Bialek, 2002) if one takes into the account first loop corrections around ϕ_0.

For time-dependent l:

$$P[l] \propto \exp\left[-\int dt\, l\,(\partial_t \log l)^2\right]$$
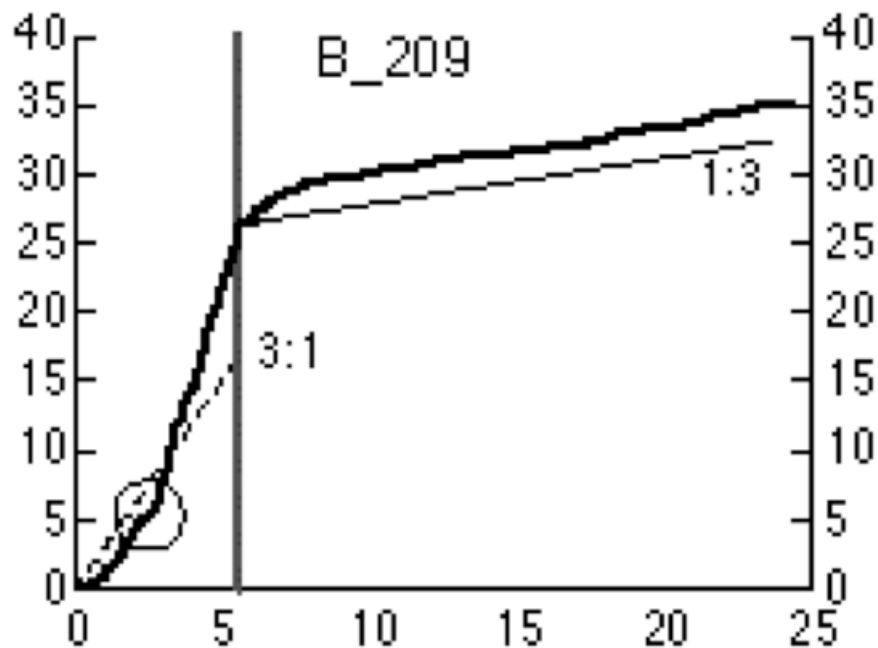
leads to self-consistent selection of l. Do not investigate particular form due to lack of data.
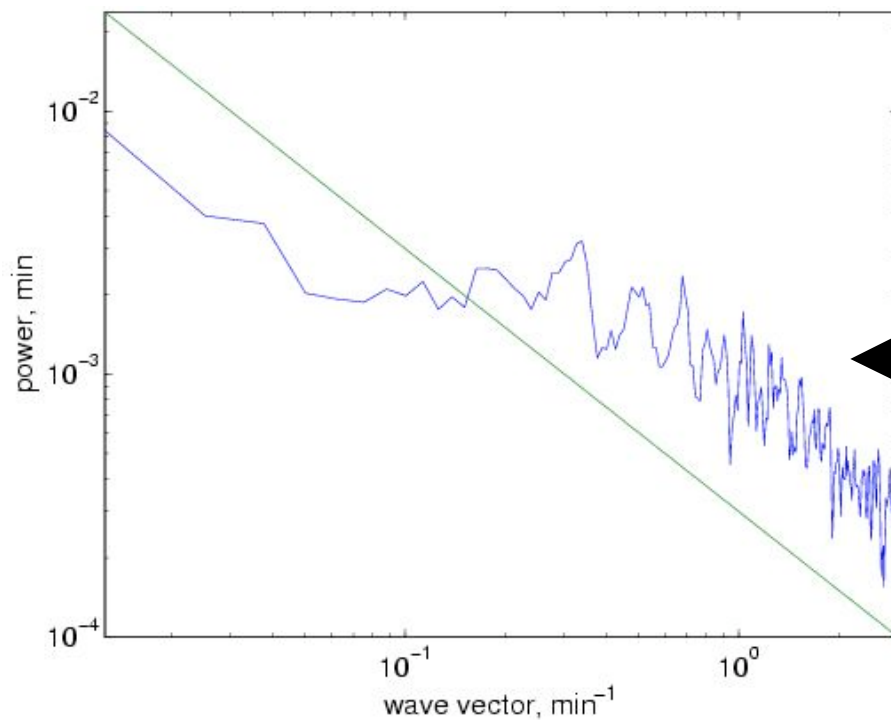
Abrupt changes

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

Abrupt changes

- Only after a few changes have been experienced
- Common during fast changes epochs
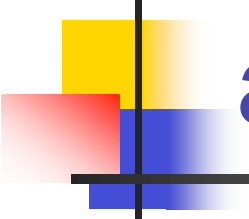- Sometimes "unwarranted"
- Metastable states?

Caused by memory (non-local QFT)

- Overestimation of rate immediately leads to higher rate and persists
- Nonlocality, friction (see also Atwal, Bialek, 2004)
- Power spectrum of reward histories
- Two regimes clearly seen
- Peak at 0 - long range correlations

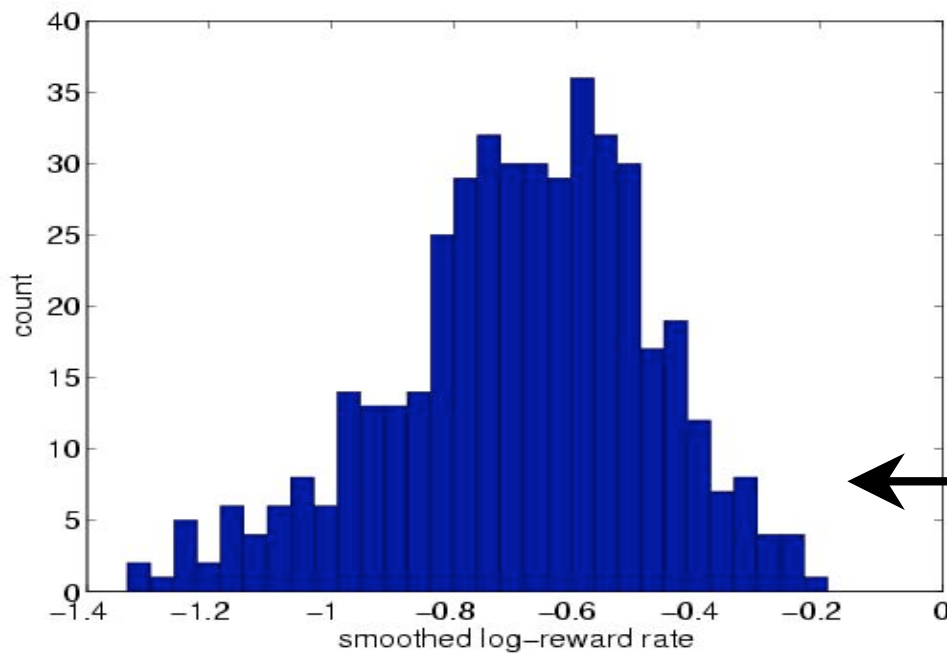# Memory, metastability, and non-Gaussianity
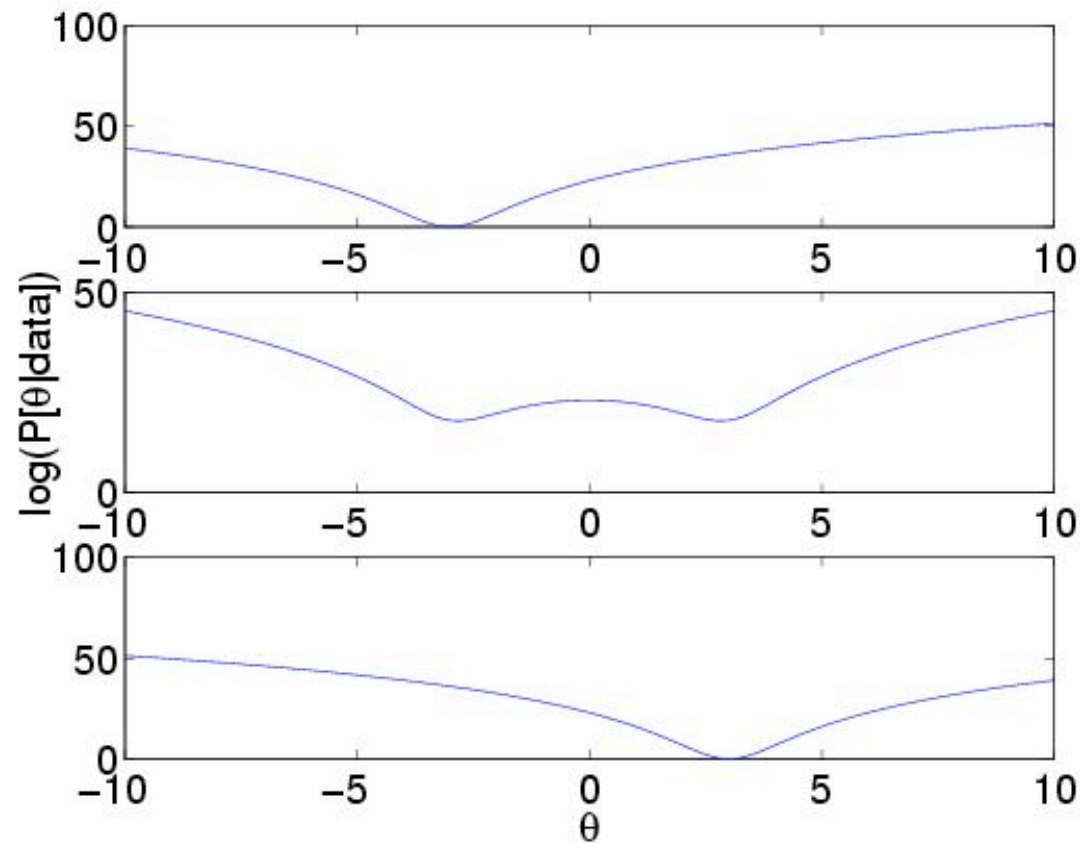
Modeling memory: two scales and rapid readjustments

- Need memory scale (long), and adaptation scale (short)

- Signal changes on long time scale, while effects of self-perpetuation of rate are on shorter scales

$$P[\phi, \psi \mid \text{data}] \propto \exp\left[-\frac{\lambda}{2} \int dt\, (\partial_t \psi)^2 - \int dt\, W(\phi - \psi) - \cdots \right.$$
$$\left. -\frac{l}{2} \int dt\, (\partial_t \phi)^2 - \sum_i V(\phi_i - \phi(t_i)) \right]$$

Non-Gaussianity

- V is not quadratic in ϕ, plateaus (self-confounding effects); exact form unclear

- W is not quadratic either; left plateau; no exact form

$$\left.\frac{\partial^2}{\partial x^2}\right|_{x=0} V(x) = \mu^2 \qquad \left.\frac{\partial^2}{\partial x^2}\right|_{x=0} W(x) = m^2$$

Dynamics of fast field

The field undergoes small fluctuations (zero point and due to irregular sampling) near the minimum of the energy

$$l \frac{d^2 \phi_0}{dt^2} = 0 = e^\theta \frac{\partial V(\phi_0 - \theta)}{\partial \phi_0} + \frac{\partial W(\phi_0 - \psi)}{\partial \phi_0}$$

Two minima

For long-tailed distributions, will have two minima

$$\phi_0^{(1)} = \psi + \varepsilon_1(\theta - \psi)$$

$$\phi_0^{(2)} = \theta + \varepsilon_2(\psi - \theta)$$

Barrier depends on the shape and vanishes
as

$$\psi \to \theta$$

Jumps possible from a metastable state near ψ to
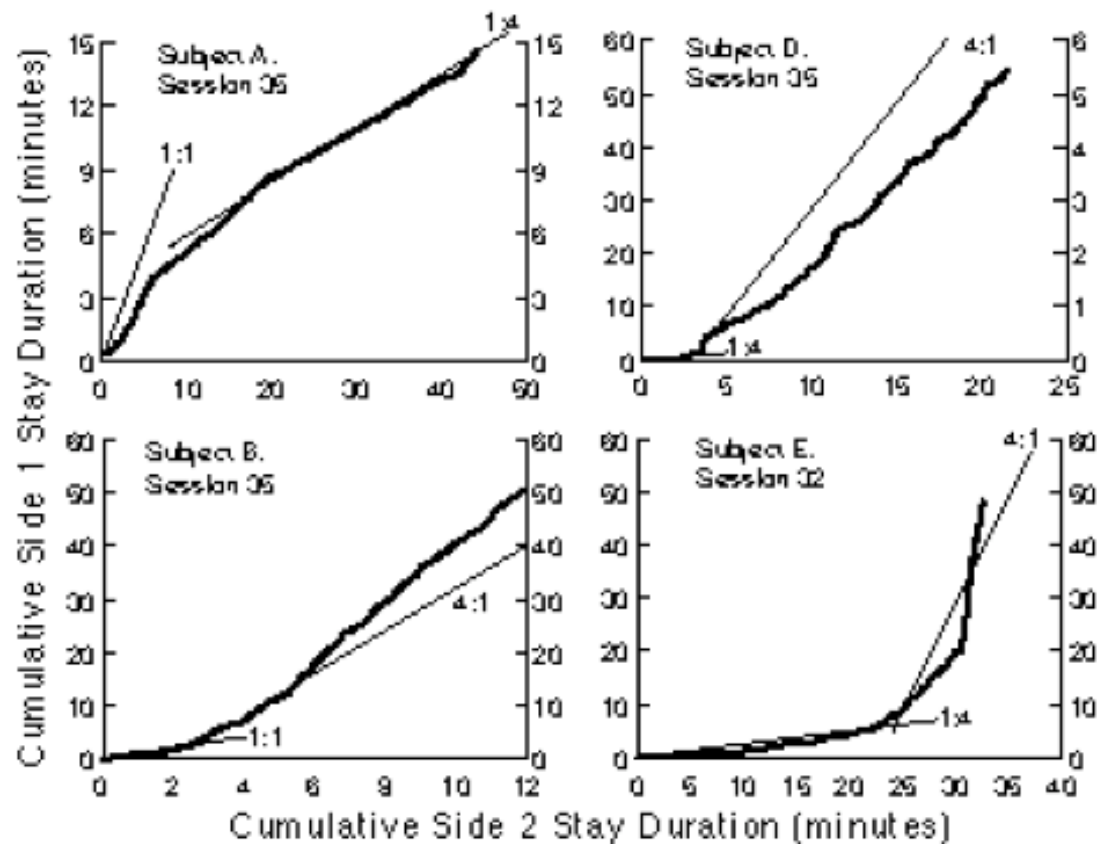new minimum near θ.

Adiabatic dynamics of the slow field

$$P[\psi \mid \text{data}] \propto \exp\left[-\frac{\lambda}{2} \int dt\, (\partial_t \psi)^2 - \int dt\, W(\phi_0^{(i)}(t) - \psi(t)) \right]$$

Either self-perpetuates, or approached the
true solution (depending on which minima the
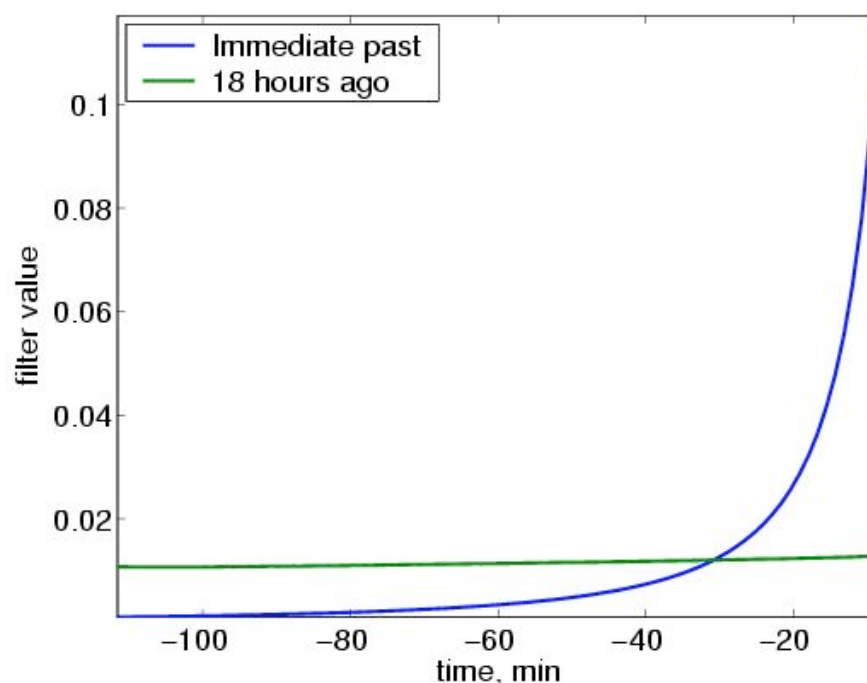fast field is in).

Critical periods?

Reversal to status quo ante

Modeling memory: nonlocality and long range correlations

Bialek & Zee, 1990 - decoding can be done linearly accurately (even if the process itself is very nonlinear). Also trivially true for Gaussian processes.

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

Optimal *F(t)* for a Gaussian process with *C~t²* 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 a bit suboptimal 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 support: matching for accumulating rewards. Planning experiments to test matching to neutral stimuli.