Adaptation on multiple timescales

Adrienne Fairhall
Department of Physiology and Biophysics
University of Washington

The plan

- Adaptive neural coding
- Example: the fly visual system
- Long timescales: a phenomenological model
- Short timescales: mechanisms
  - functional description of the neural computation
  - Adaptation through intrinsic properties

Collaborators:

Bill Bialek
Experimental work done in the lab of Rob de Ruyter,
with a lot of assistance from Geoff Lewen.
Blaise Aguera y Arcas

Michael Berry
Naama Brenner

Adaptation

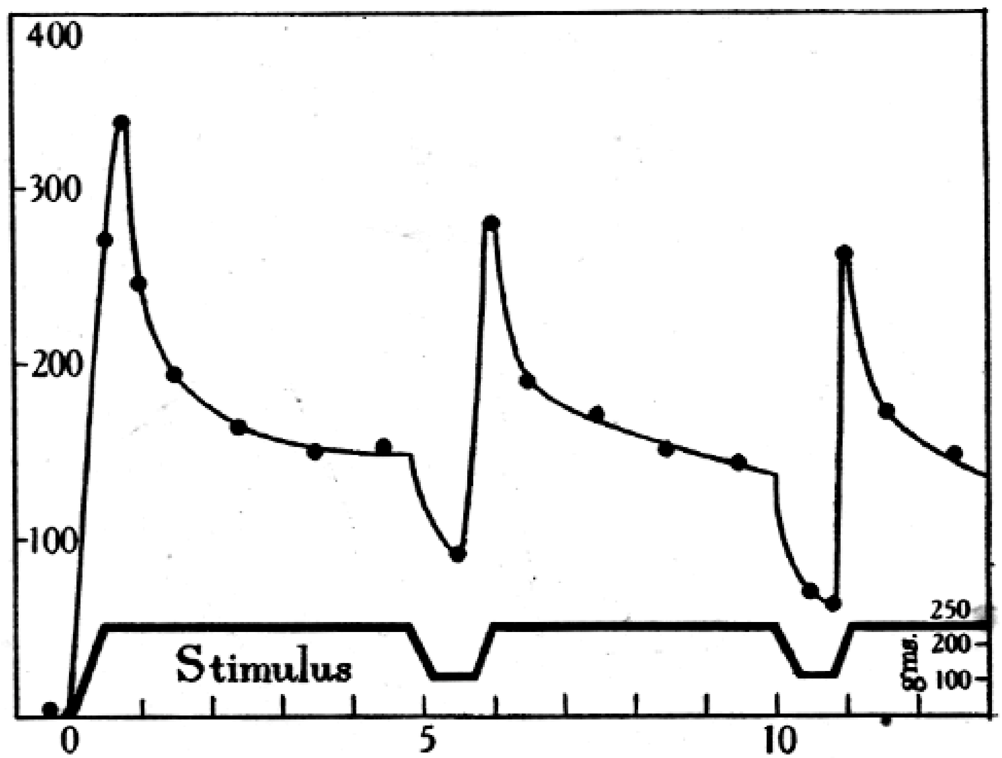


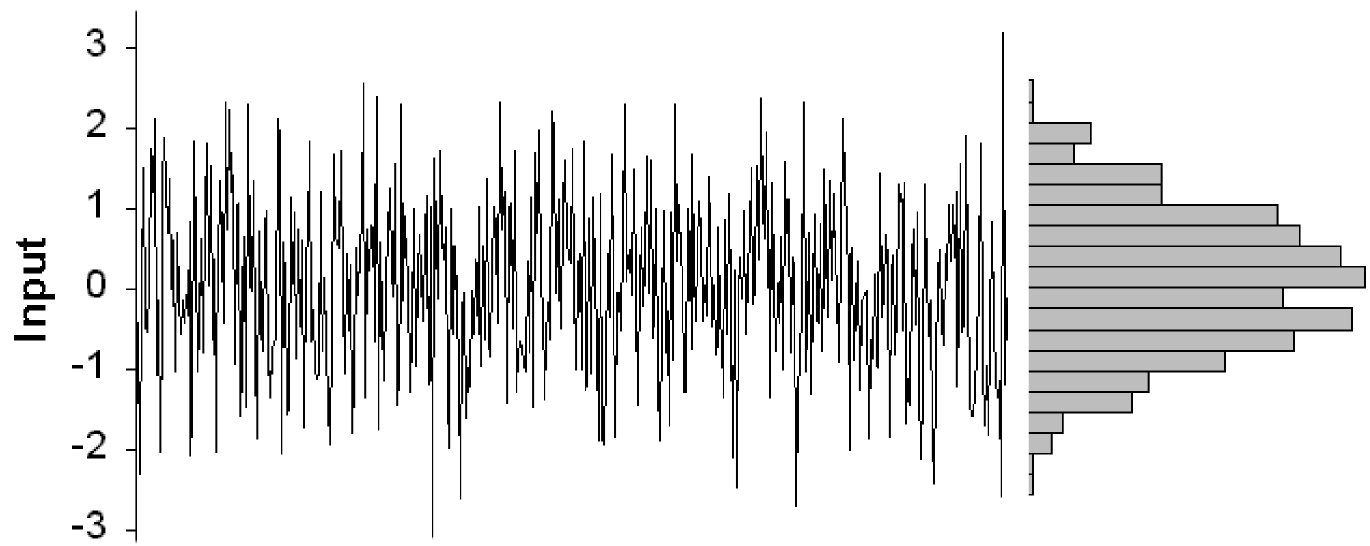
FIG. 29. IMPULSES PRODUCED BY PRESSURE ON CAT'S TOE-PAD. STIMULUS REMOVED PARTIALLY AND RENEWED.

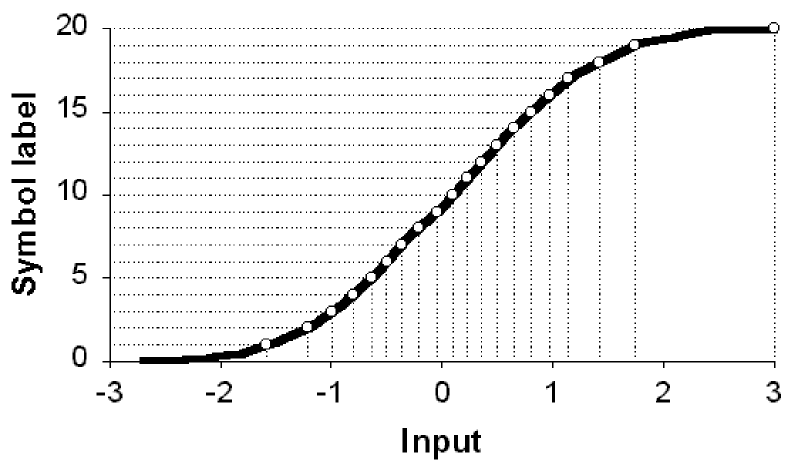
The Basis of Sensation, Adrian (1929)

Functional role of adaptation for information processing

- prevents a neuron from continuing to respond to repetitive stimuli (Adrian)
- redundancy reduction (Barlow and Attneave)
- increases neuron’s dynamic range
- improves information transmission
- should consider adaptation not to a sustained stimulus but to a changing ensemble of stimuli

Coding: adapting to a distribution



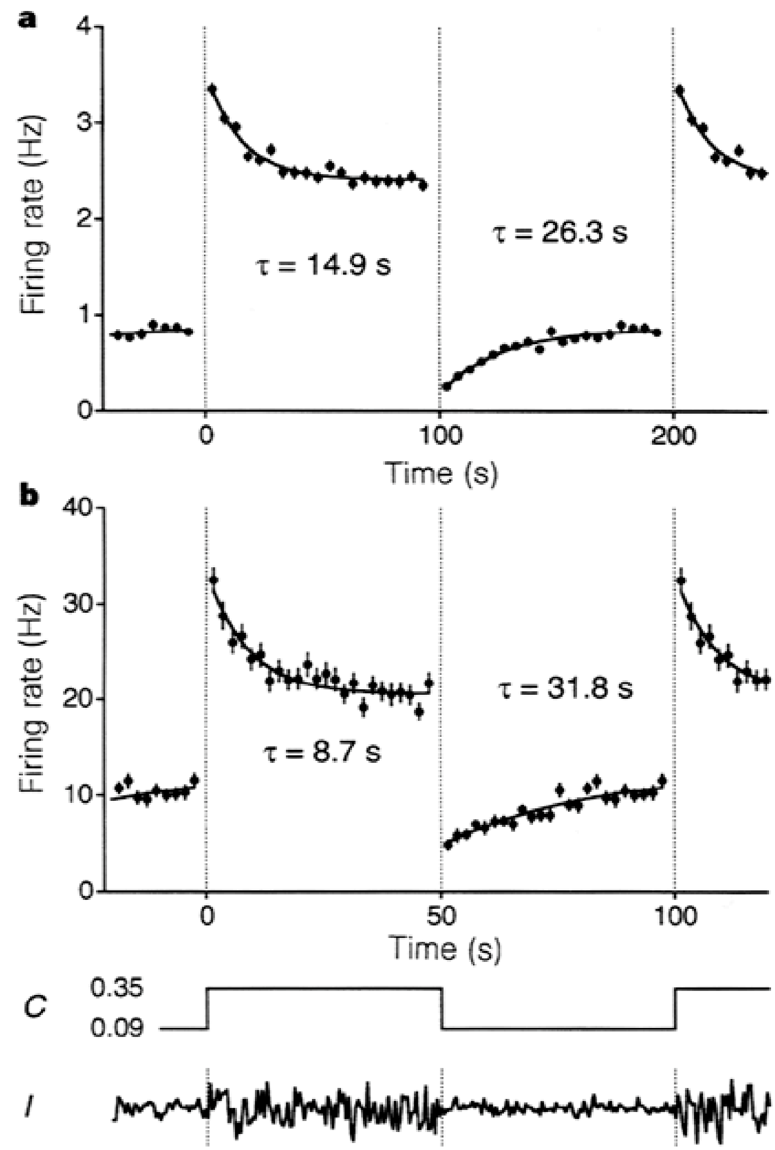


→ Input/output curve
which depends on
the stimulus distribution

Different aspects of adaptation are relevant for information processing

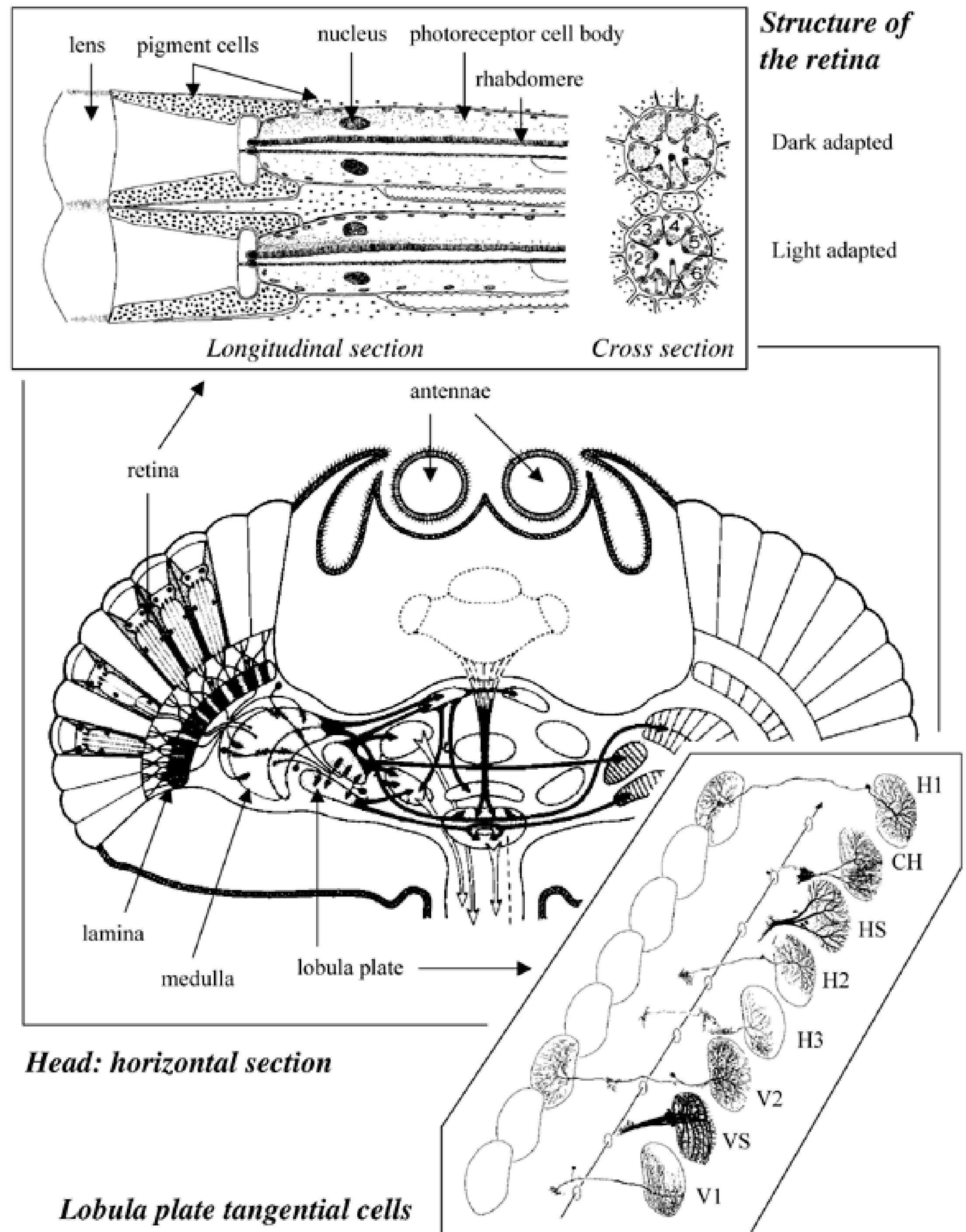
- rate accommodation
- changes in the neural input/output relation
- changes in the neuron’s feature selectivity

Example: contrast adaptation in the retina



Smirnakis et al., Nature (1997)

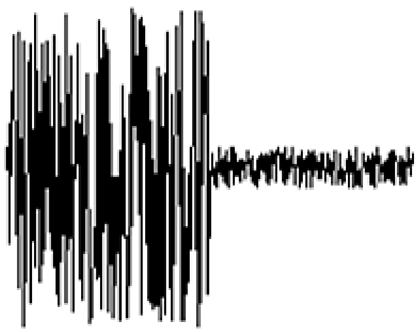
H1: a large identified motion-sensitive neuron in the fly lobula plate

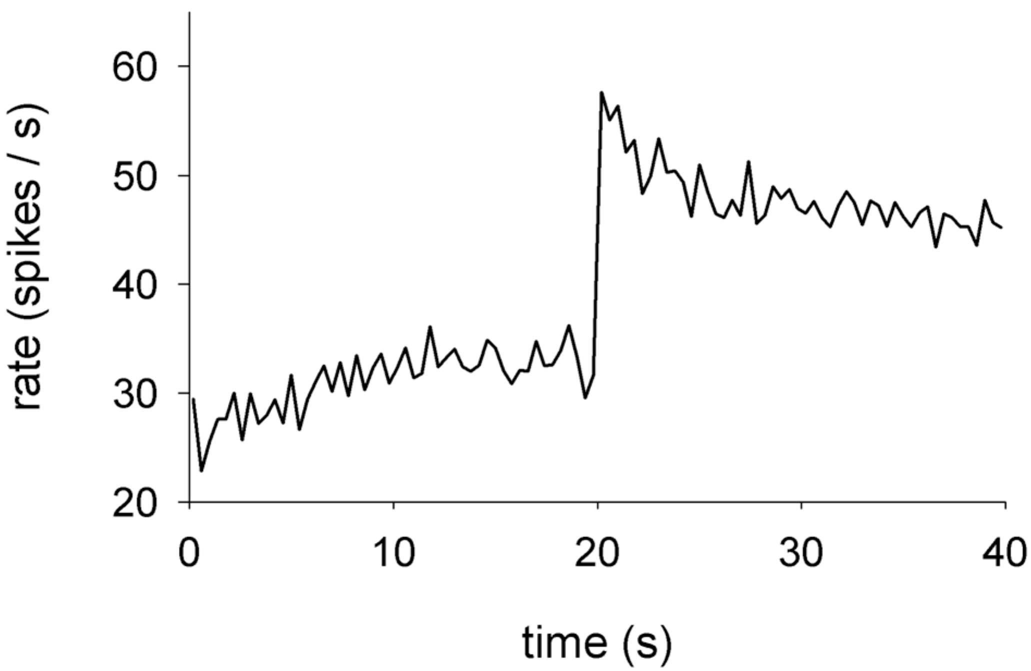


de Ruyter van Steveninck and Bialek, FIG 3.

Slow rate adaptation

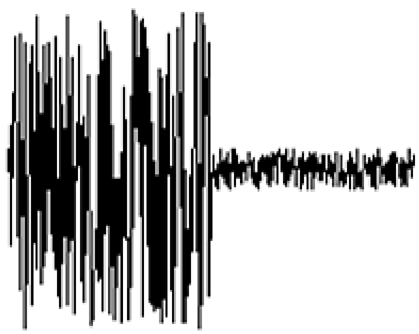
Periodically switch stimulus variance

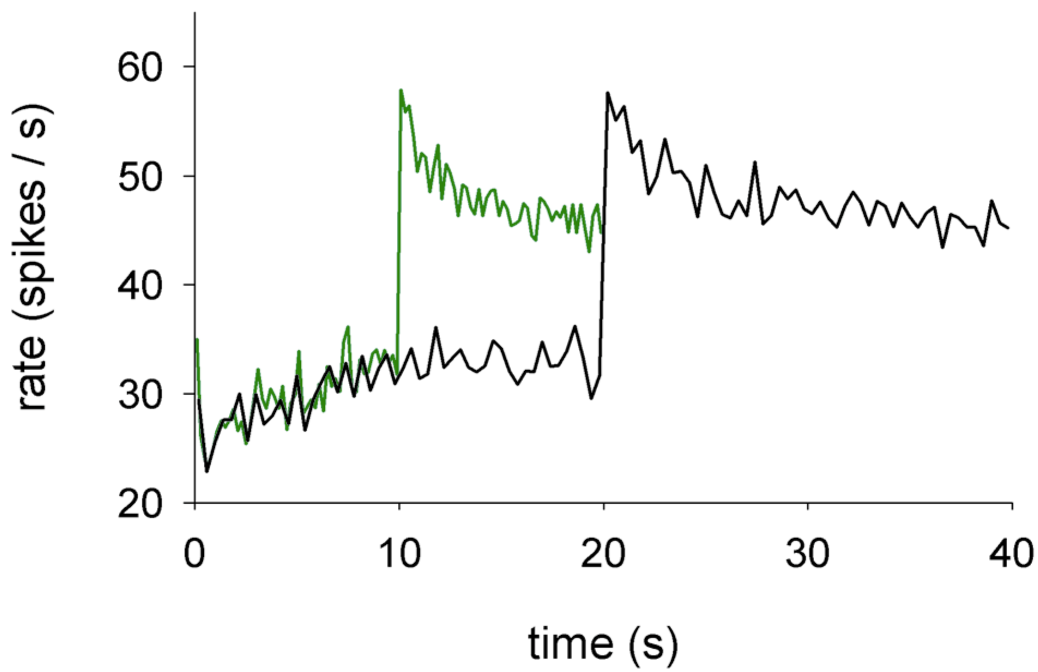




Slow rate adaptation

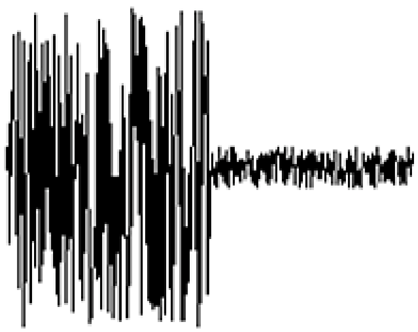
Periodically switch stimulus variance

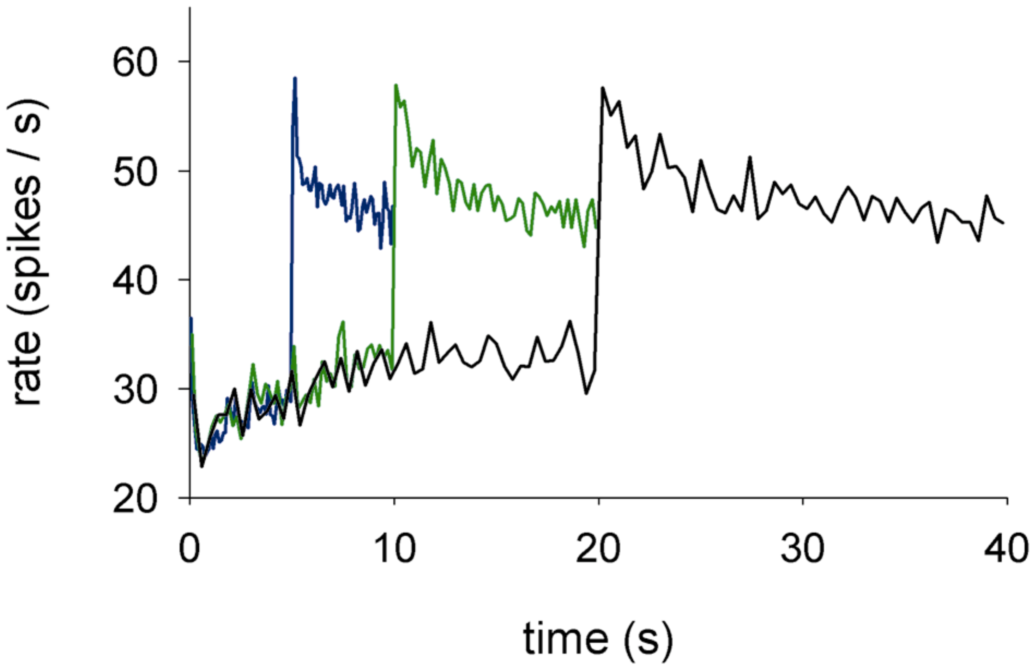




Slow rate adaptation

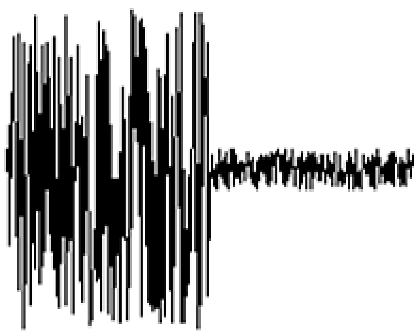
Periodically switch stimulus variance

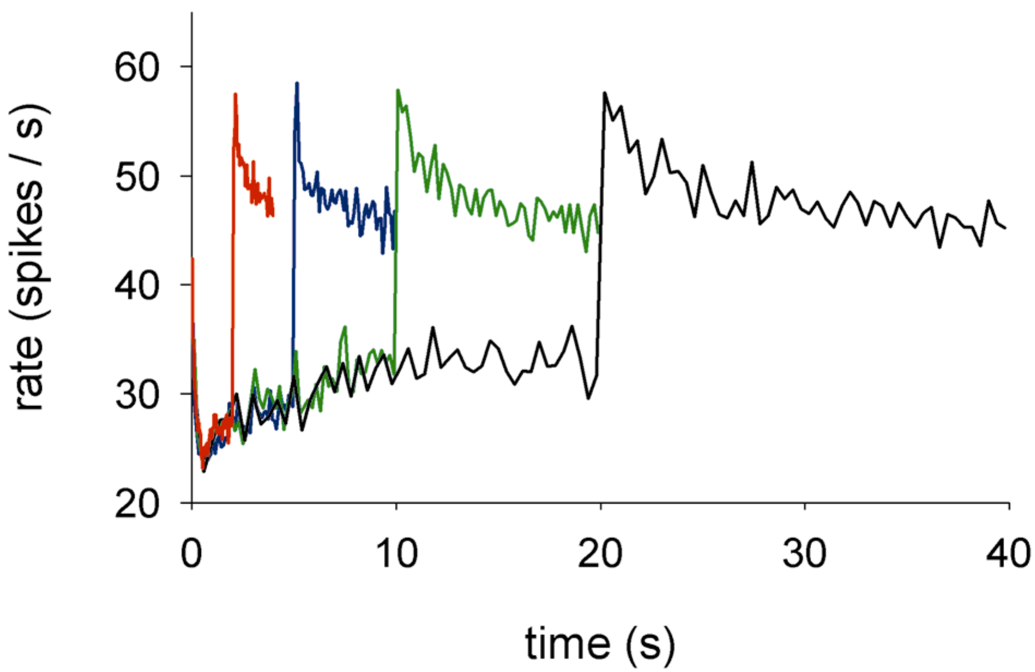




Slow rate adaptation

Periodically switch stimulus variance



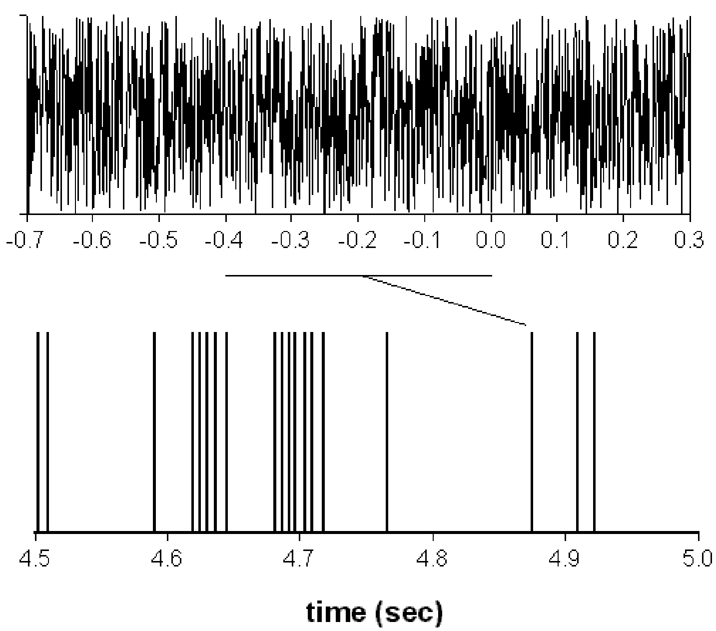


Fairhall, Bialek, Lewen and de Ruyter, “Efficiency and ambiguity in an adaptive neural code”, Nature (2001)

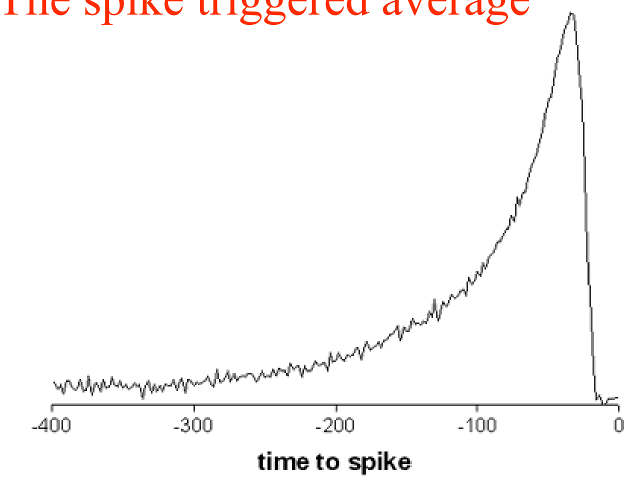
Slow rate adaptation over tens of seconds
Timescale scales with timescale of the experiment

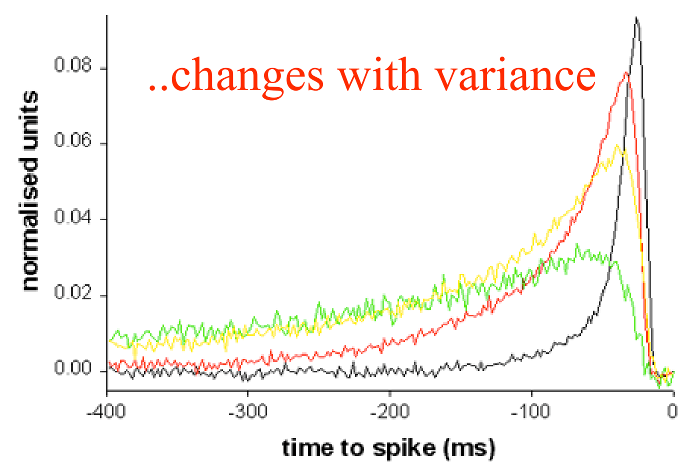
Constructing the neuron’s input/output relation

• stimulus is reduced to its projection onto the spike triggered average

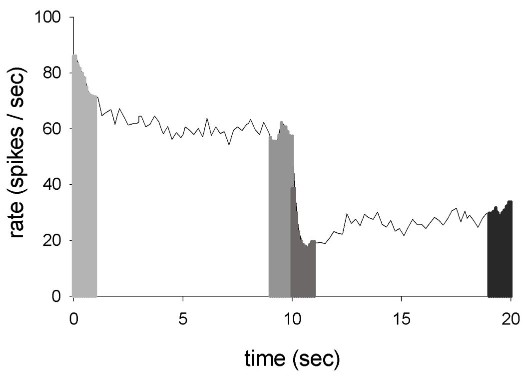


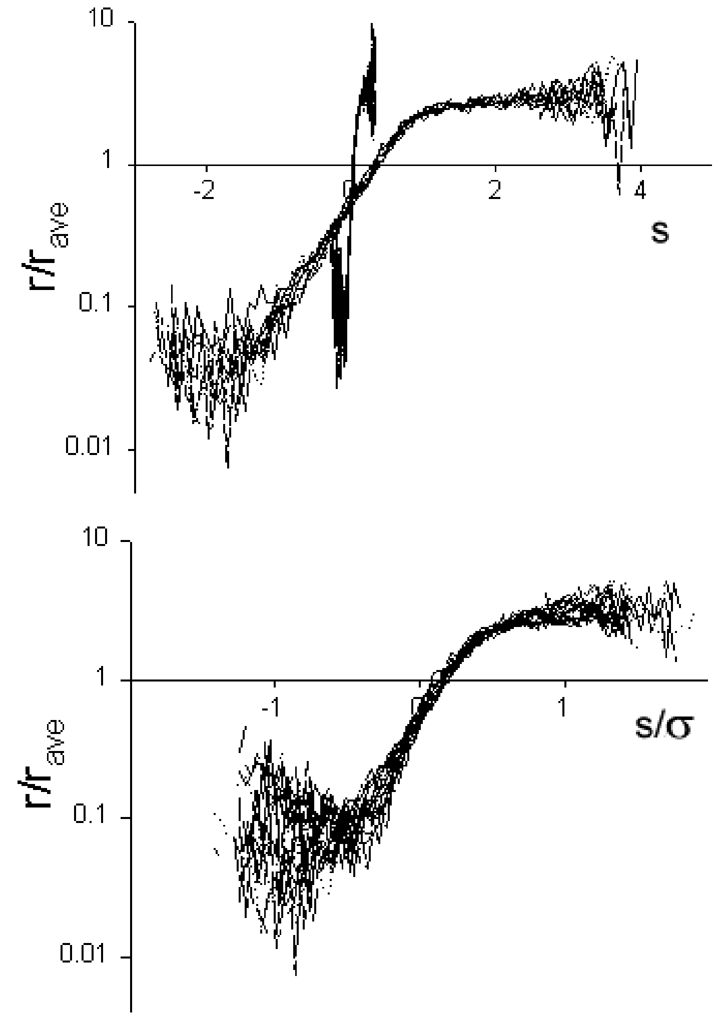
The spike triggered average





Fast adaptation of input/output relations



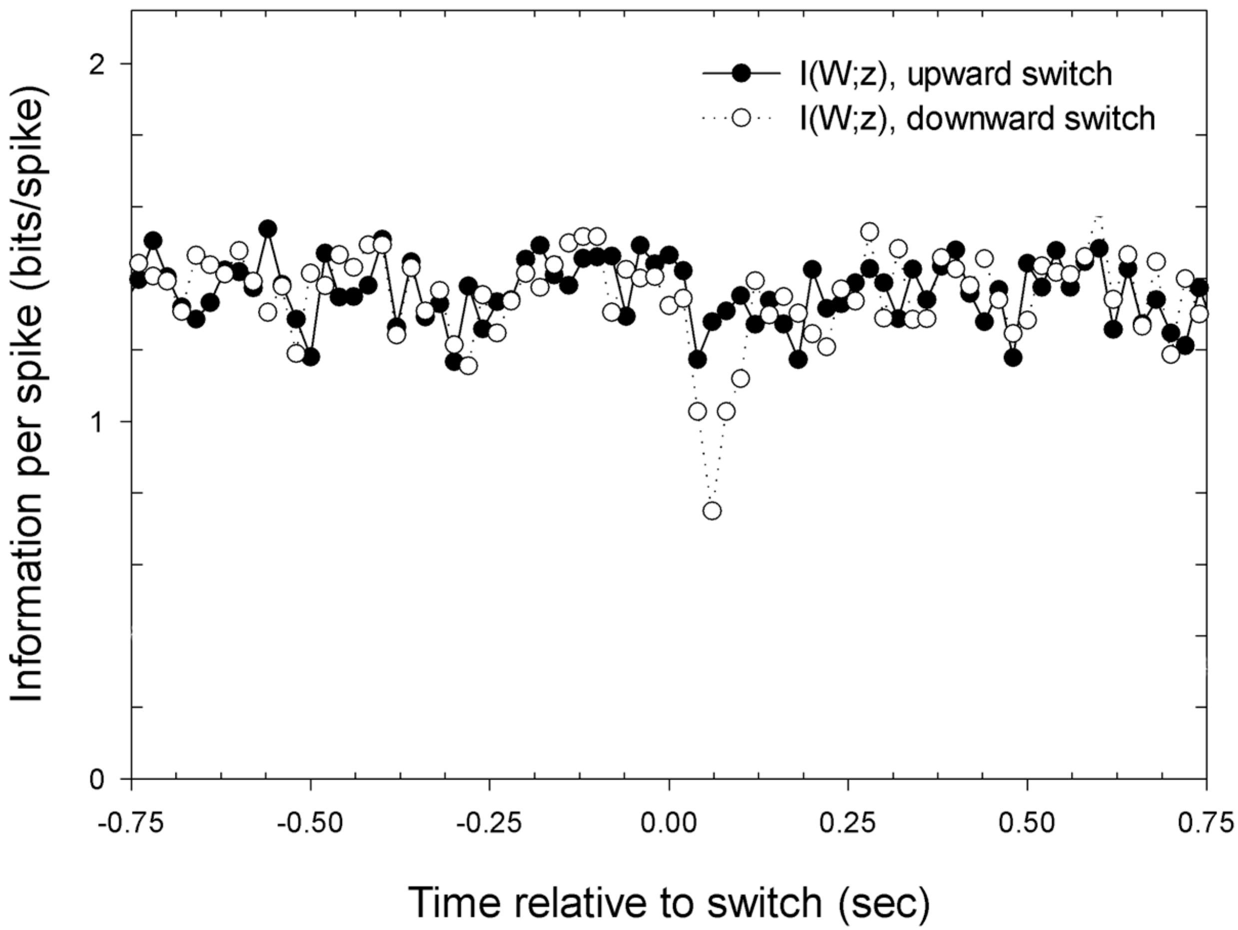


• dynamical rescaling to variance of distribution

• adaptation of filters and I/O relations on timescales of ~100ms: at statistical limits

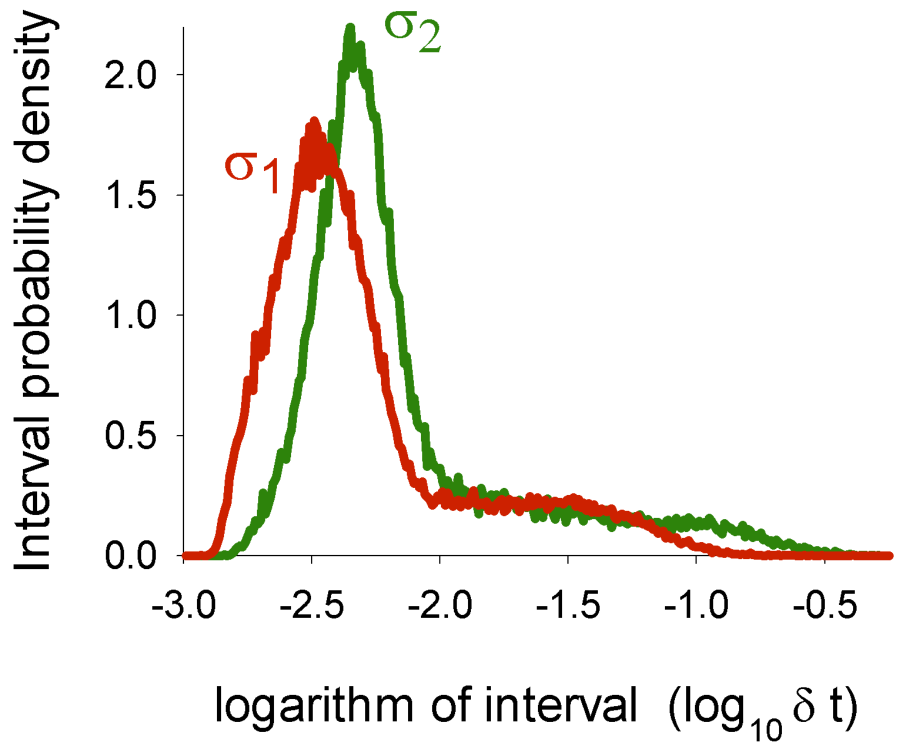
• Similar results in the retina
(Baccus and Meister, 2002)

Adaptation preserves information transmission rate.



The dynamics of the rate envelope is independent of the rescaling of the input/output relations.

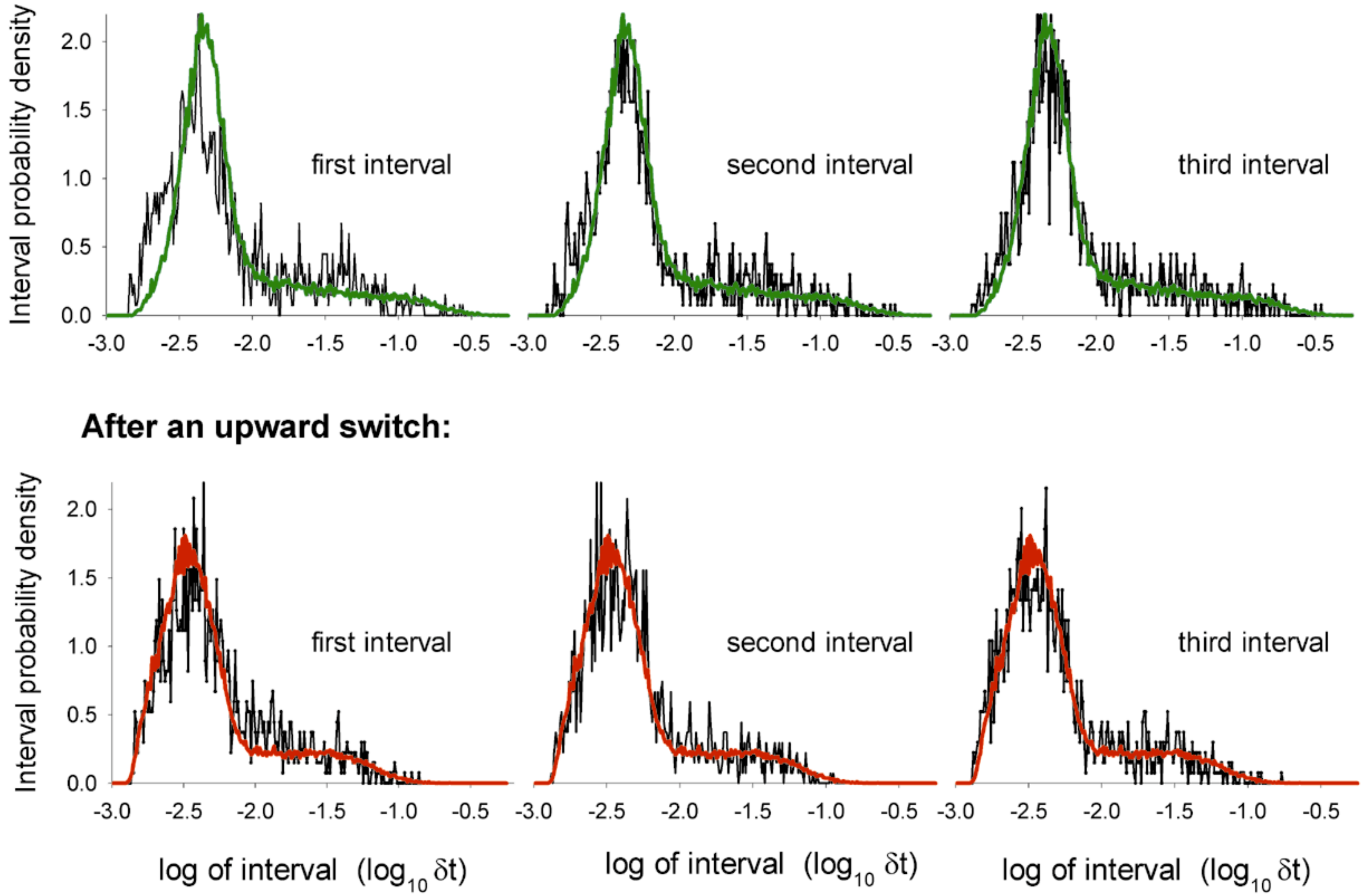
Disambiguating the variance: the interspike interval distribution



Fairhall, Bialek, Lewen and de Ruyter, "Efficiency and ambiguity in an adaptive neural code", Nature (2001)

After a downward switch:

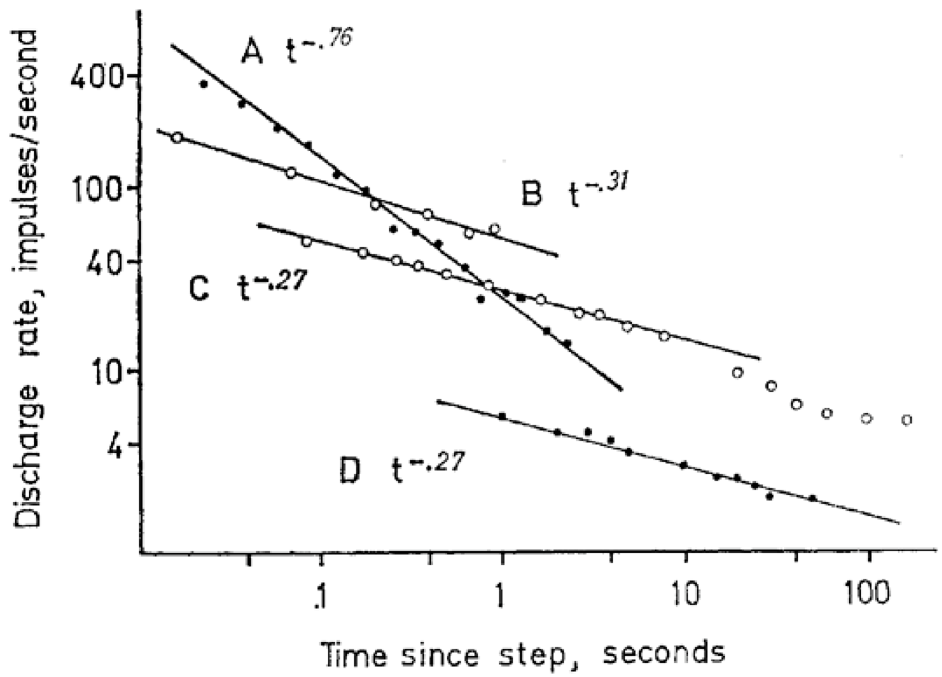
After an upward switch:



The rate dynamics: what's going on

- Recall: *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)

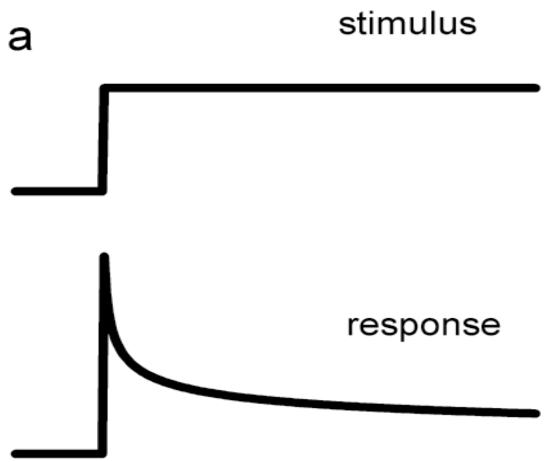
Fractional differentiation

Fourier representation $(\mathrm{i}\square)^{\square}$:

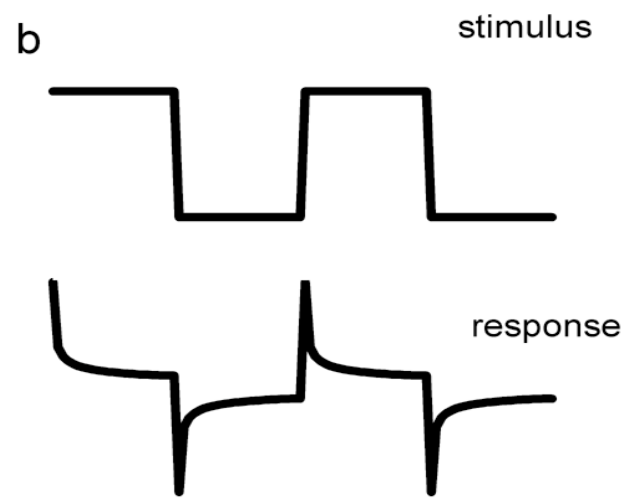
each frequency component scaled by $\square^{\square}$

and with phase shifted by a constant phase $\mathrm{i}^{\square} \rightarrow \square\square$

power-law response to a step:

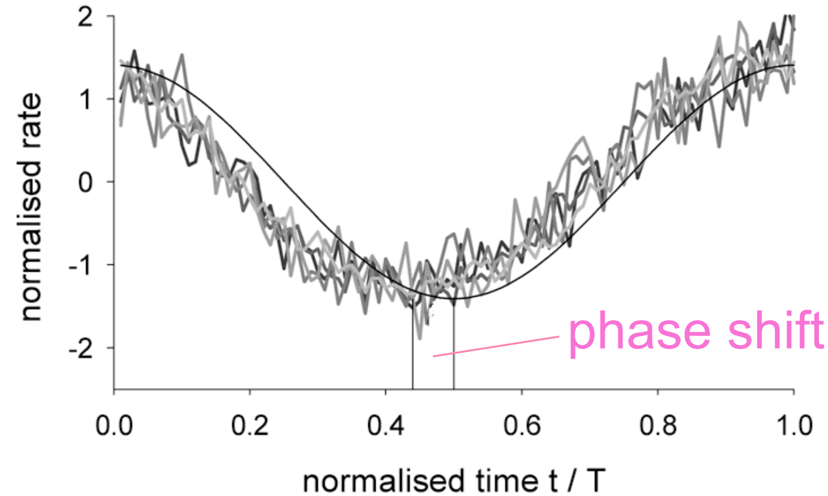


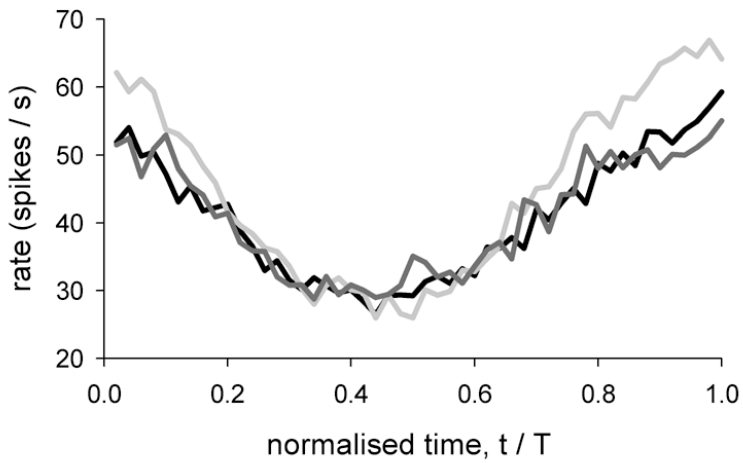
scaling “adaptive” response to a square wave:



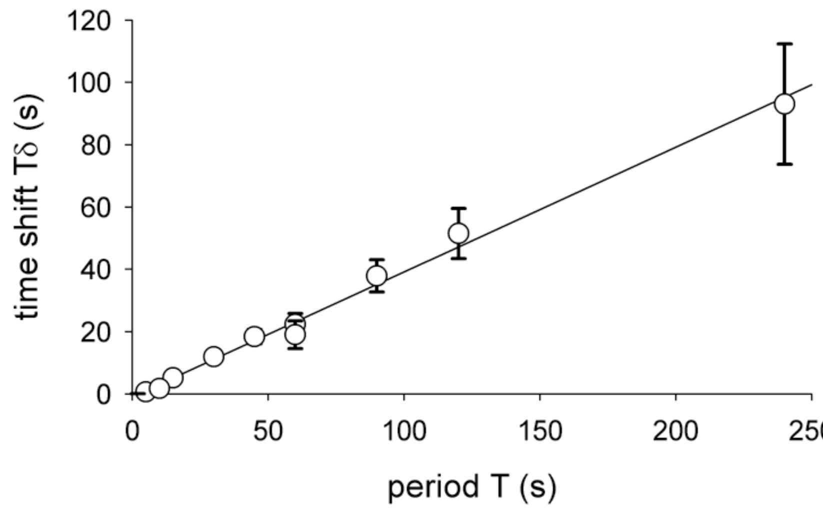
Linear analysis agrees

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





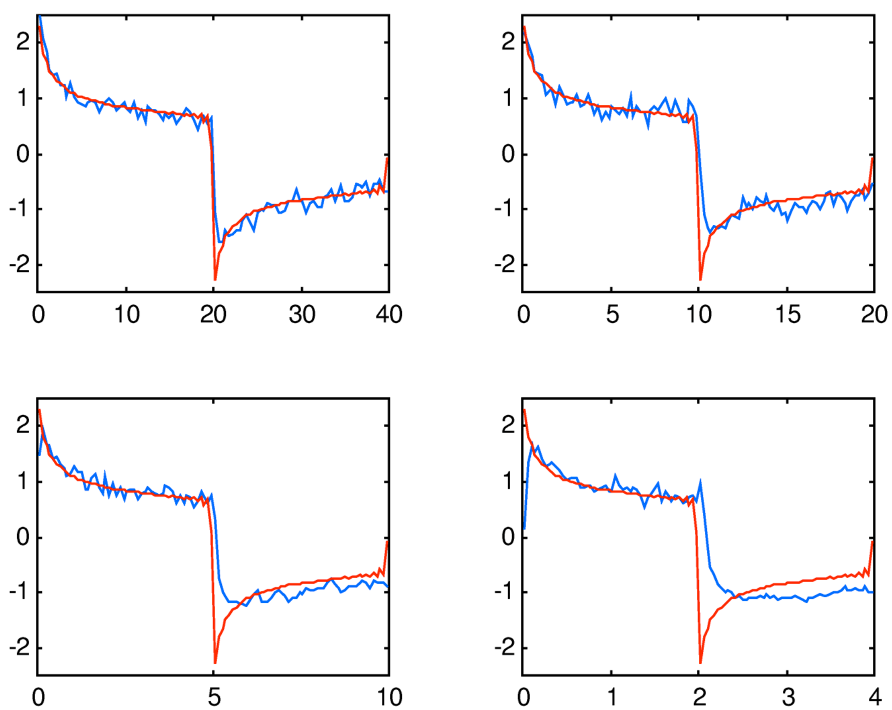
T = 30s, 60s, 90s



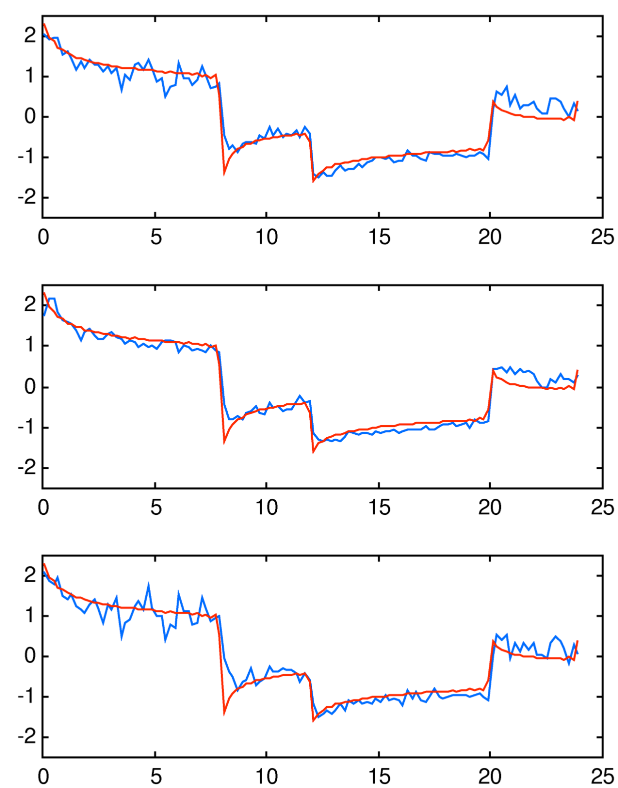
Fits very well

From sinusoid experiments, find exponent □ ~ 0.2

Two-state switching



Three-state switching



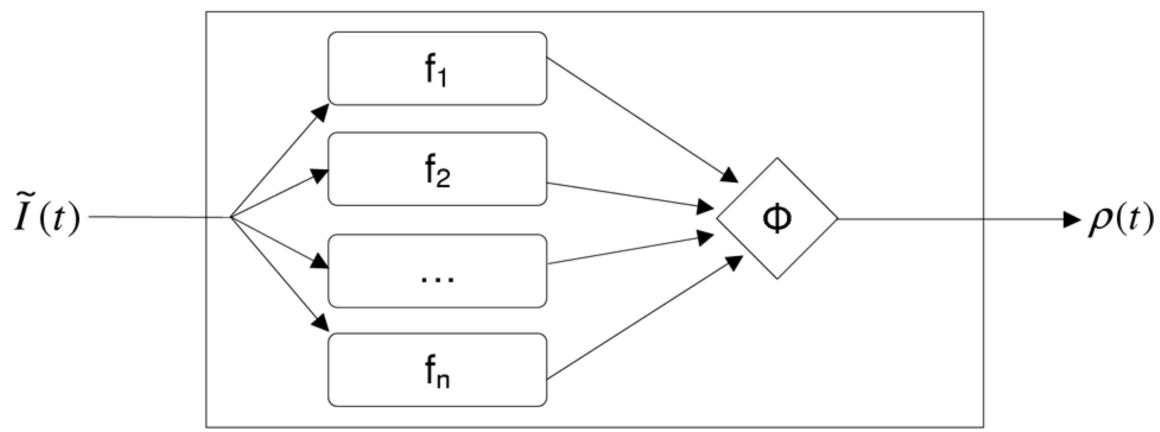
So it’s a fractional differentiator…

- connects with ‘universal’ power-law behaviour of receptors
- unusual to see this in a “higher computation”
- functional interpretation: whitening stimulus spectrum (van Hateren)
- introduces long history dependence but *linear:* invertible/decodable
- emphasizes rapid changes and extends dynamic range, but does not throw out information in the steady state
- what’s the mechanism? Some ideas..

Possible mechanisms for fast statistical adaptation:
intrinsic neural properties

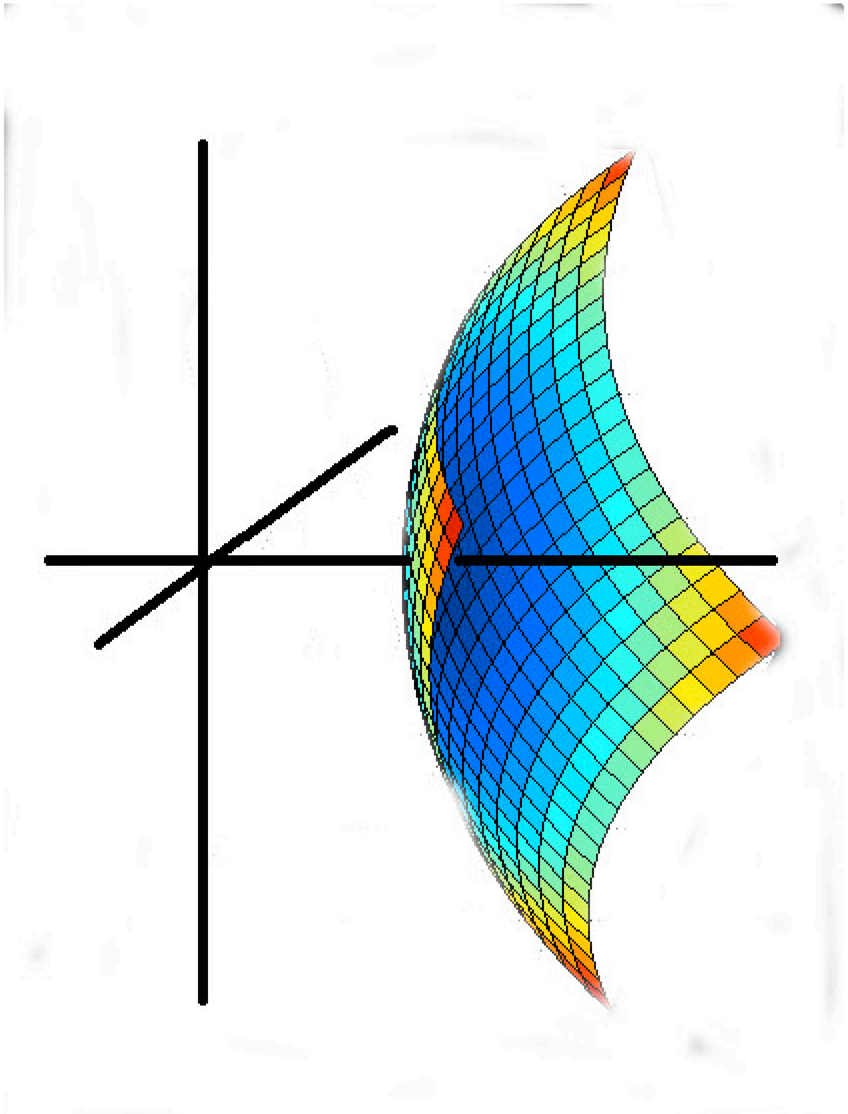
Functional neural computation

Start from biophysical/dynamical systems description
or from experimental data;
want a *functional* characterisation of the neural computation.
Basic idea:
feature detection followed by a nonlinearity

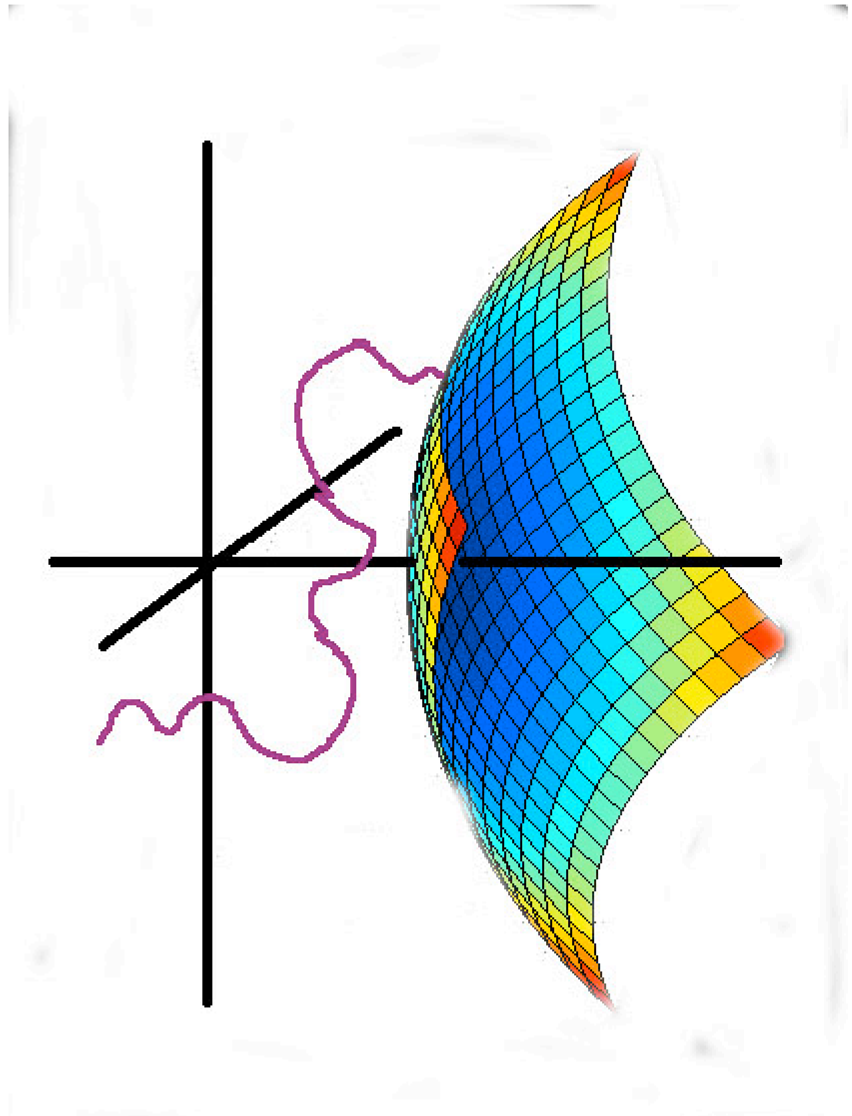


Perceptron (Rosenblatt); LN model; dimensionality reduction (Bialek et al.)

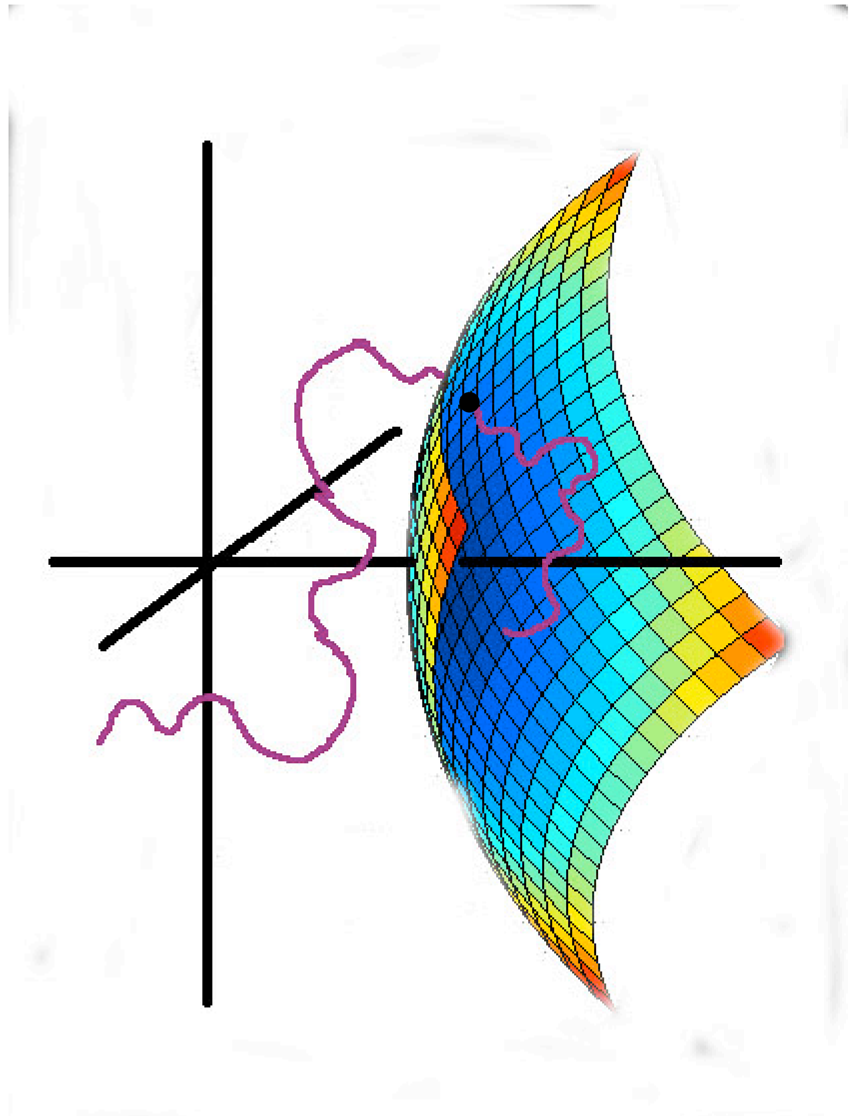
Spiking surface in stimulus space



A given time-dependent stimulus is a trajectory in this space

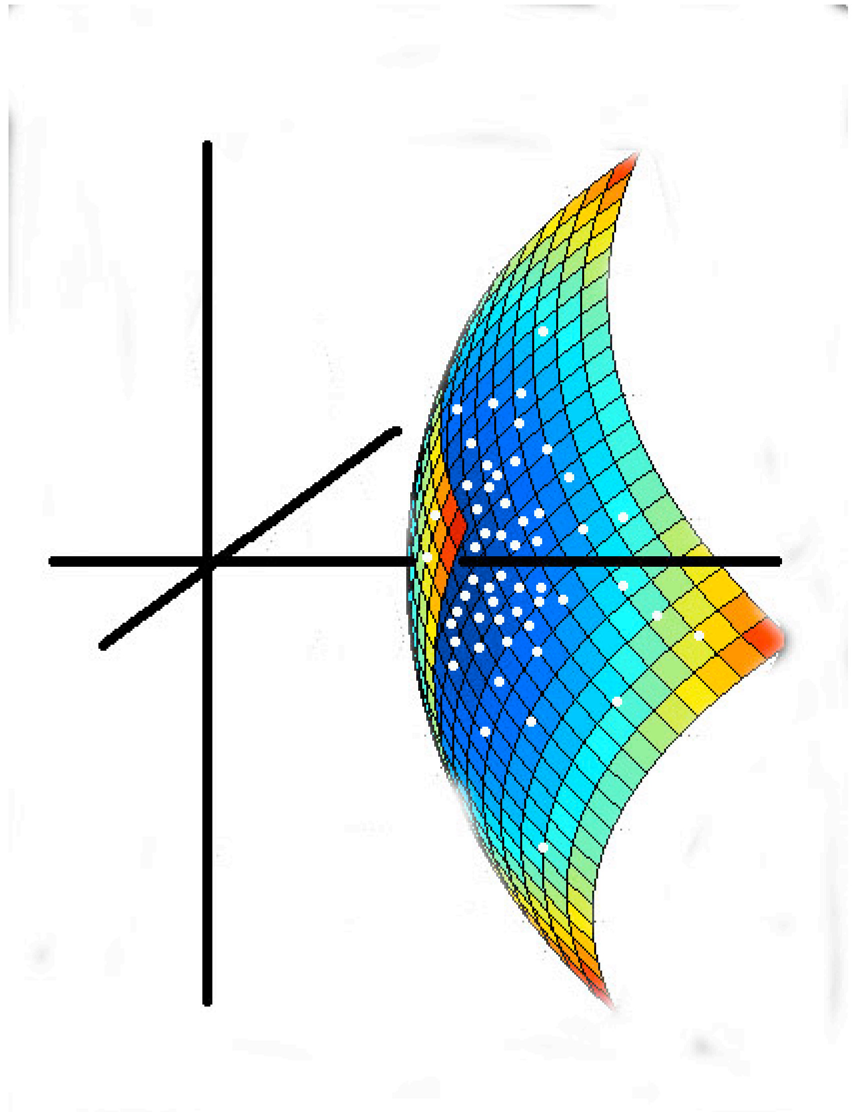


A spike..

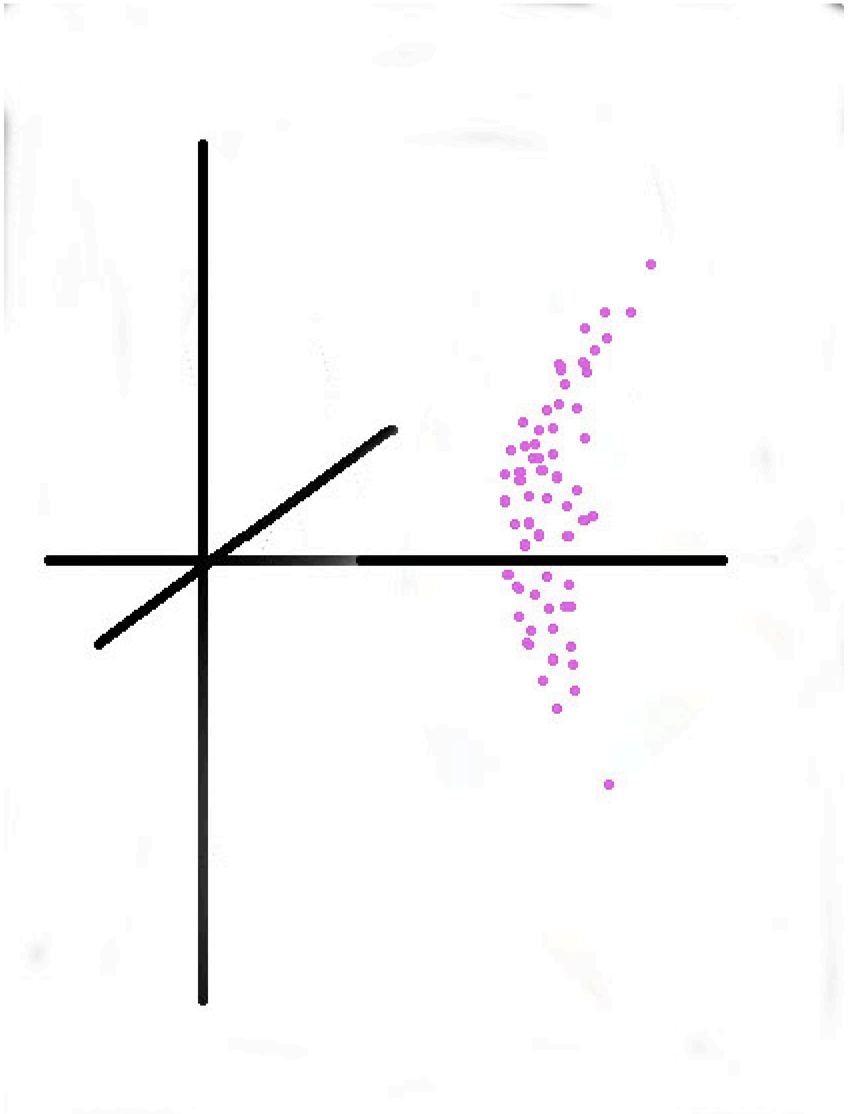


What happens after a spike? (Aguera y Arcas, Fairhall and Bialek 2000)

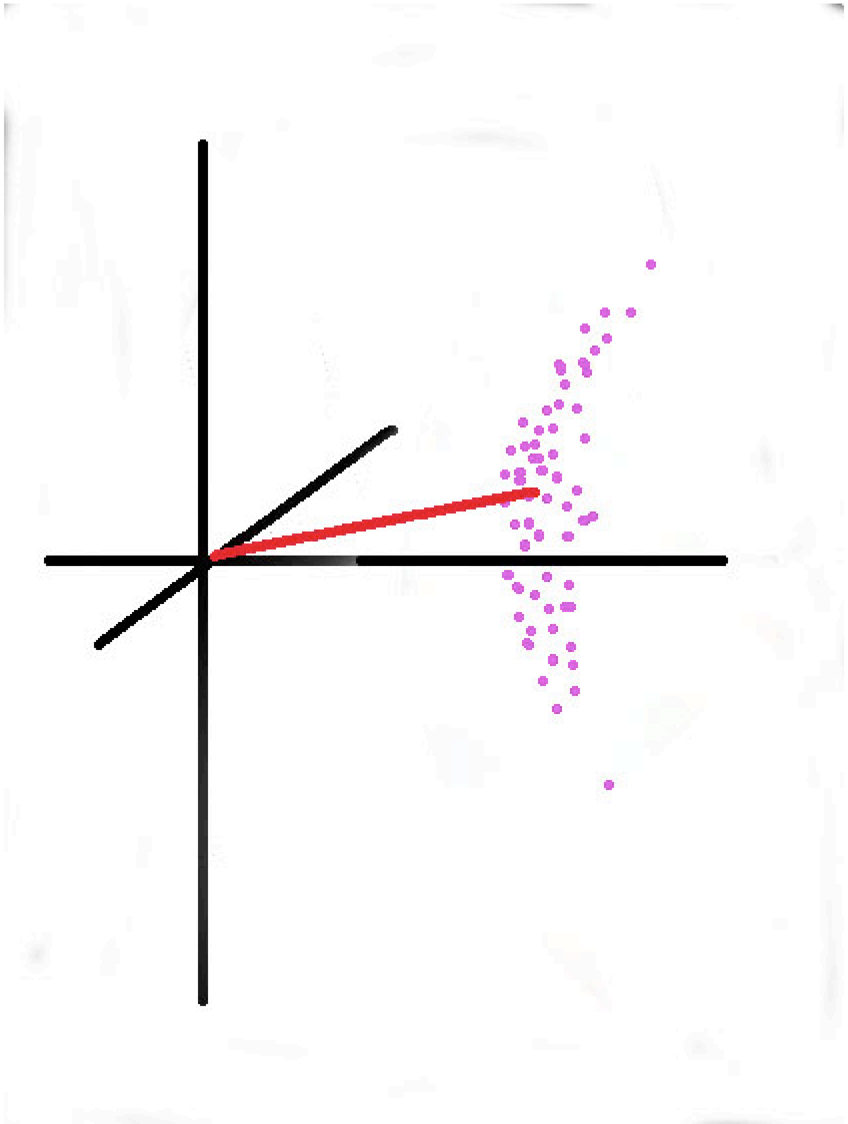
The spike-conditional distribution



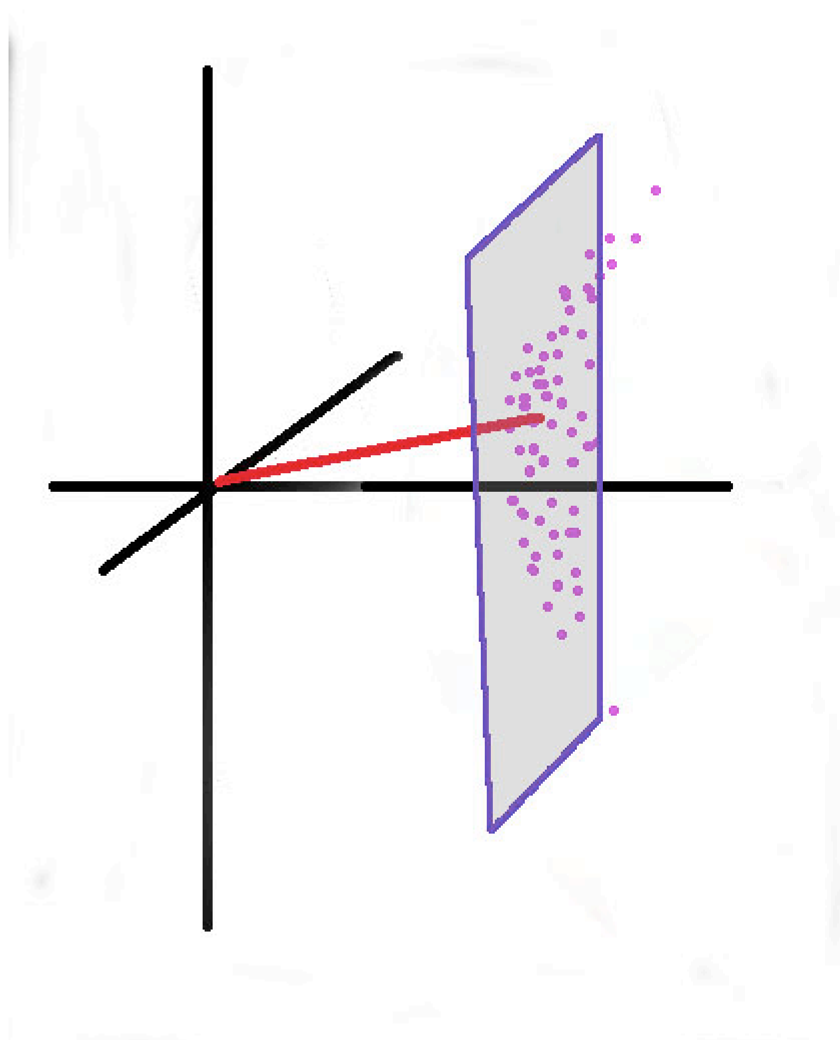
The spike-conditional distribution



The STA is the centroid

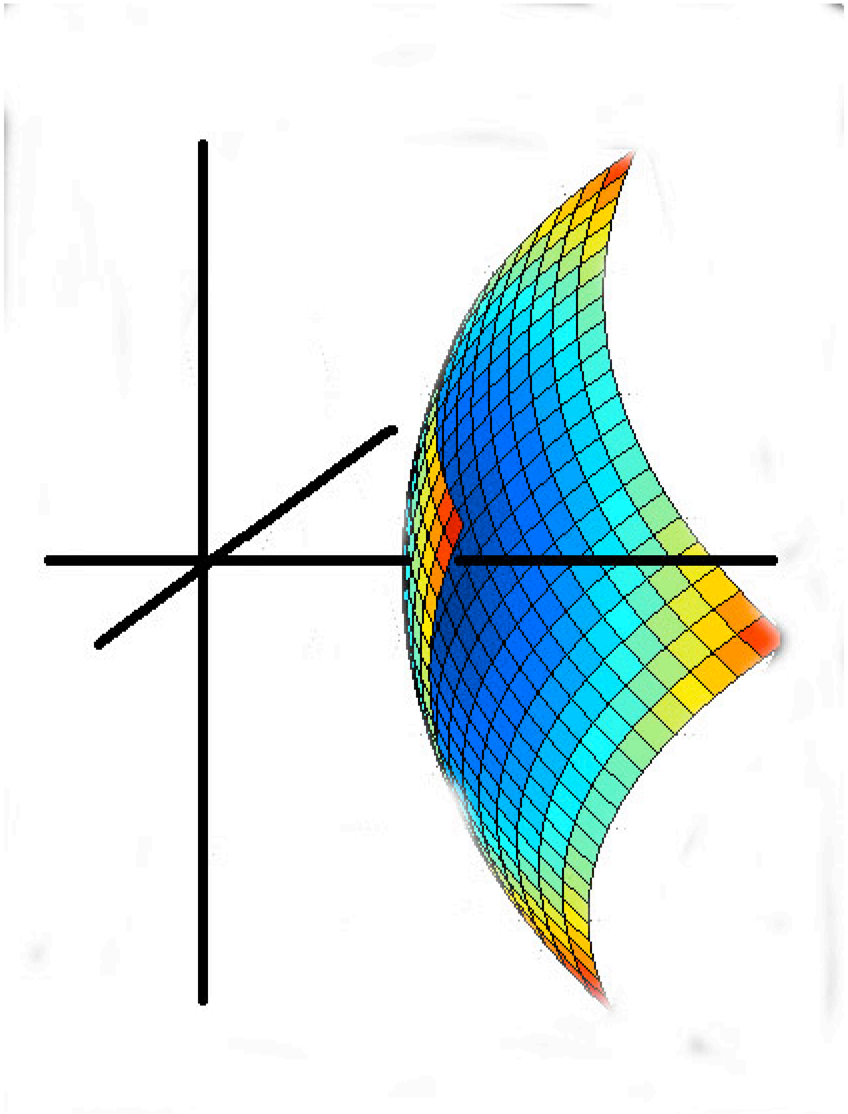


Covariance methods define a linear subspace which fits a hyperplane to the distribution



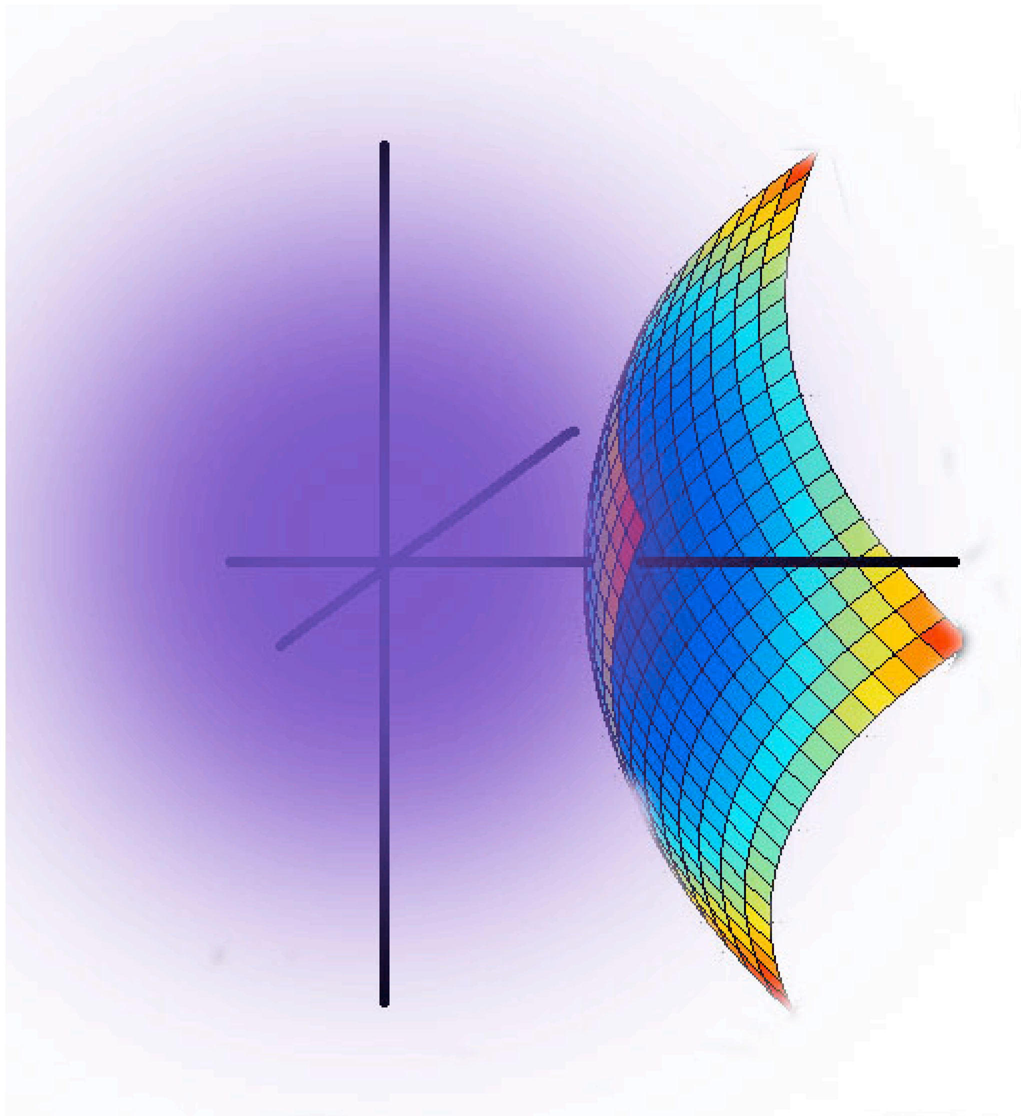
Brenner et al., Aguera y Arcas et al., Bialek et al., Schwartz et al., Rust et al., Petersen et al.

Spiking surface generally is curved



Stimulus dimensions with a nonempty normal intersection with the spiking surface are the relevant feature space

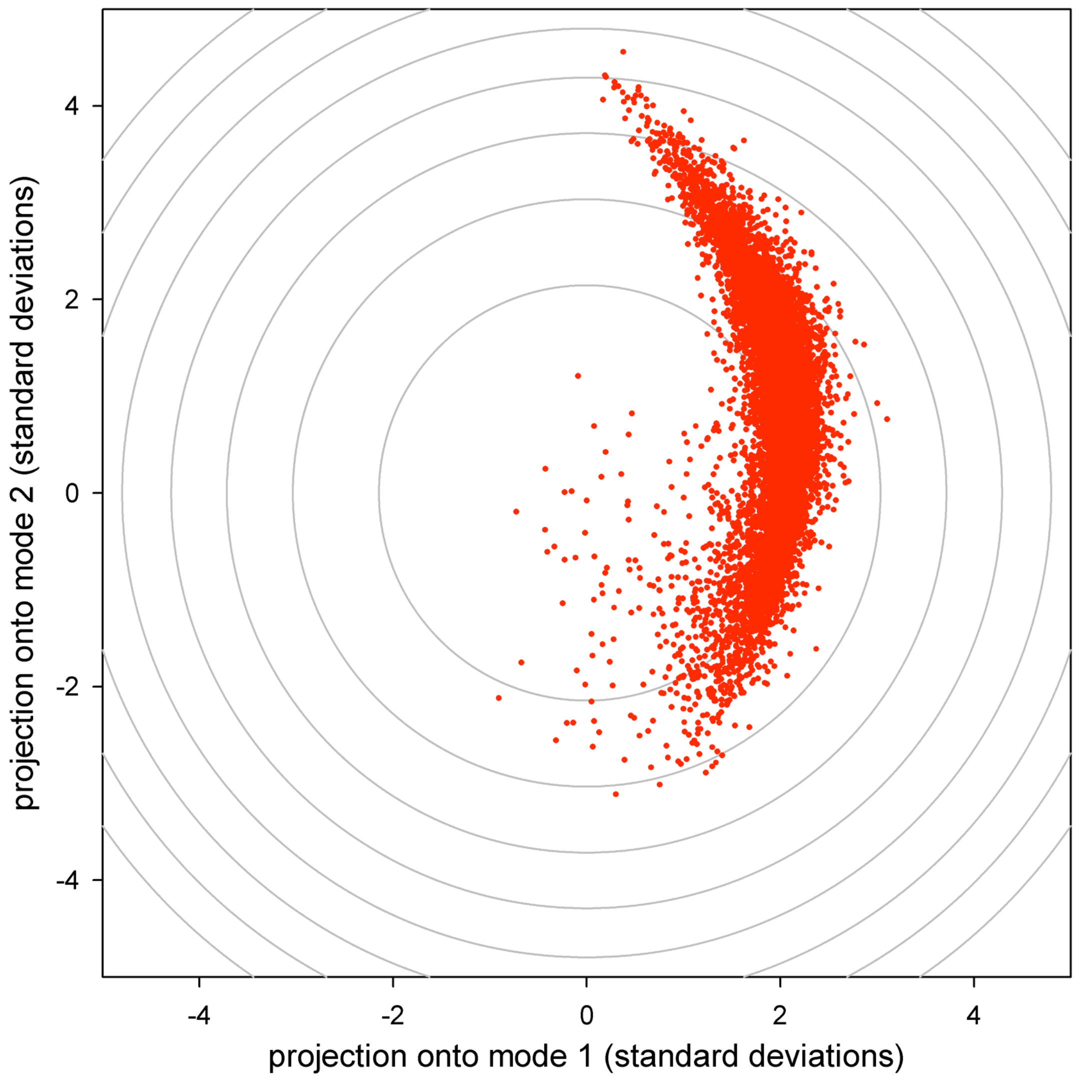
Different statistical ensembles explore the surface differently



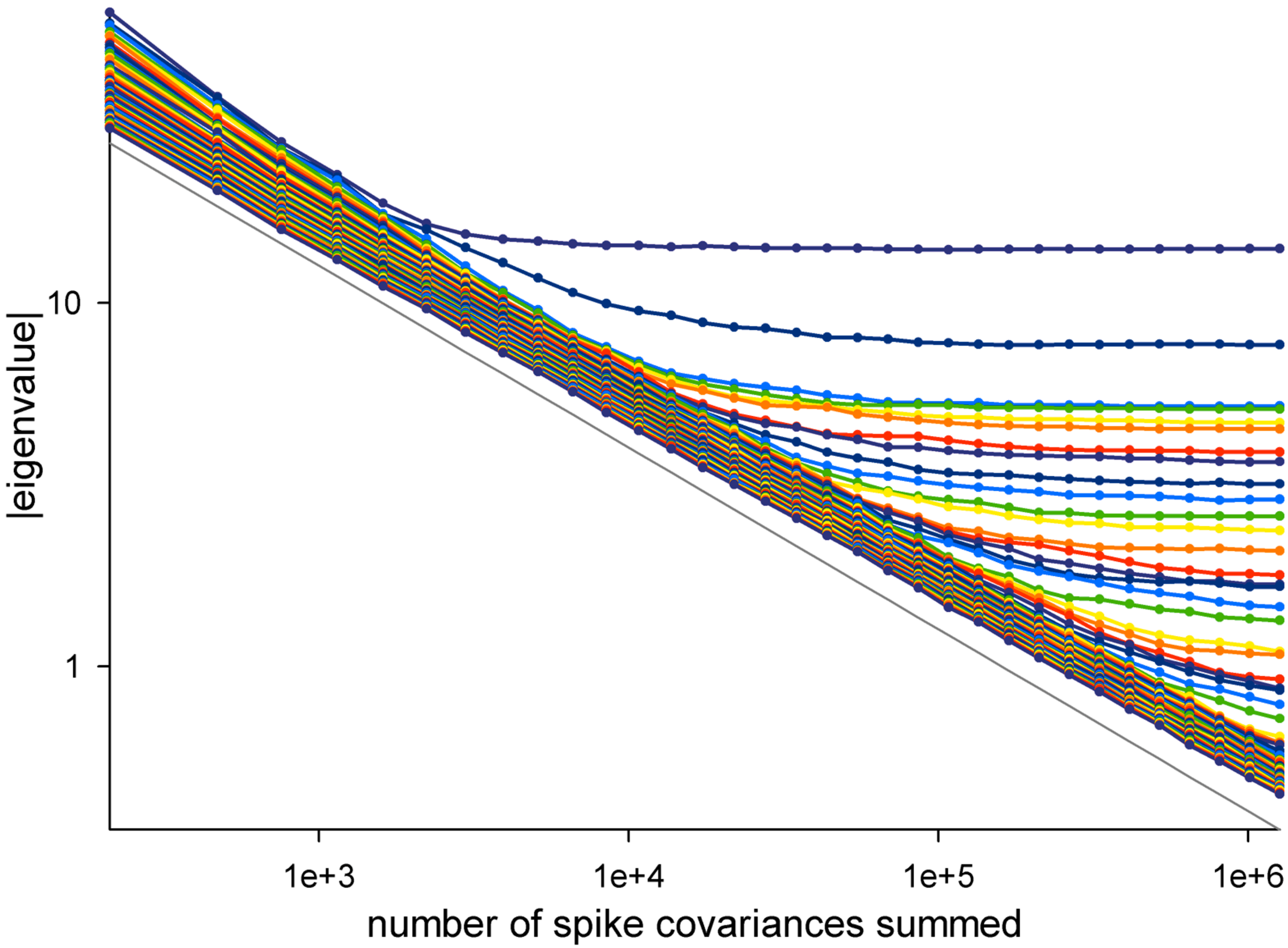
.. changes the STA, eigenmodes, input/output relation..

Spiking on curved subspaces

Not so esoteric:
the Hodgkin-Huxley
model

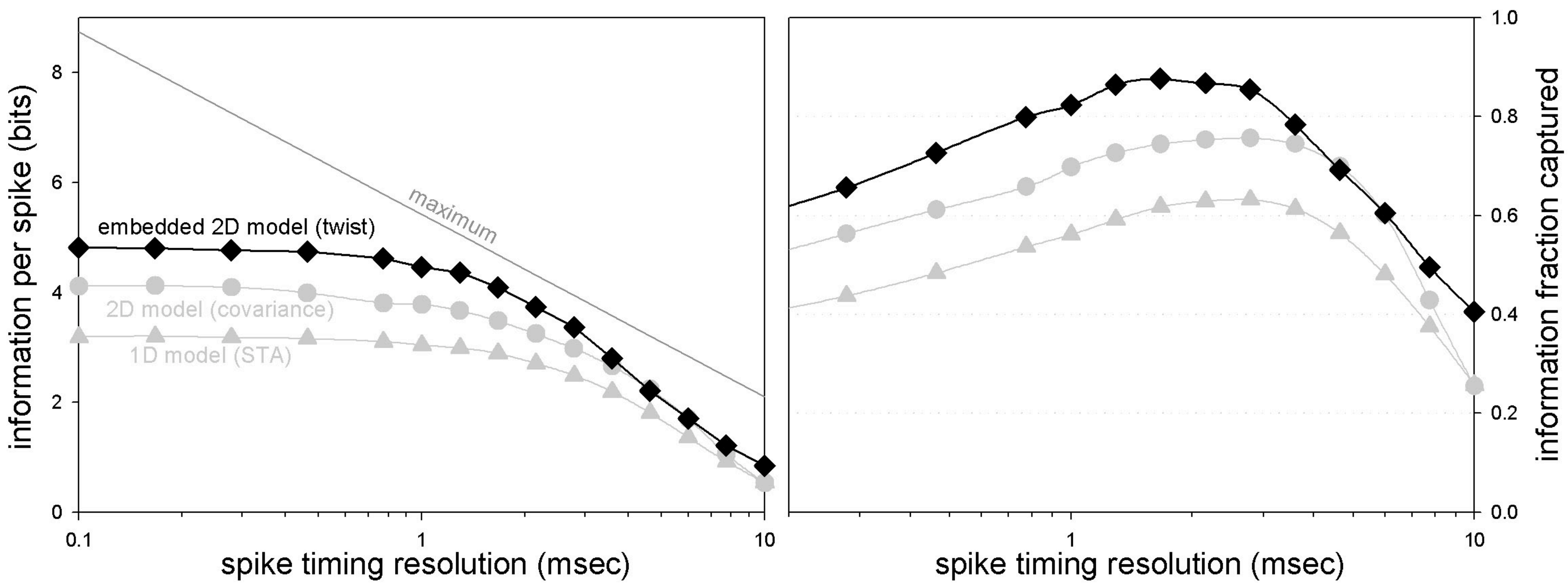


Spectrum is still high dimensional, although most of these modes are not spike related. (more later)



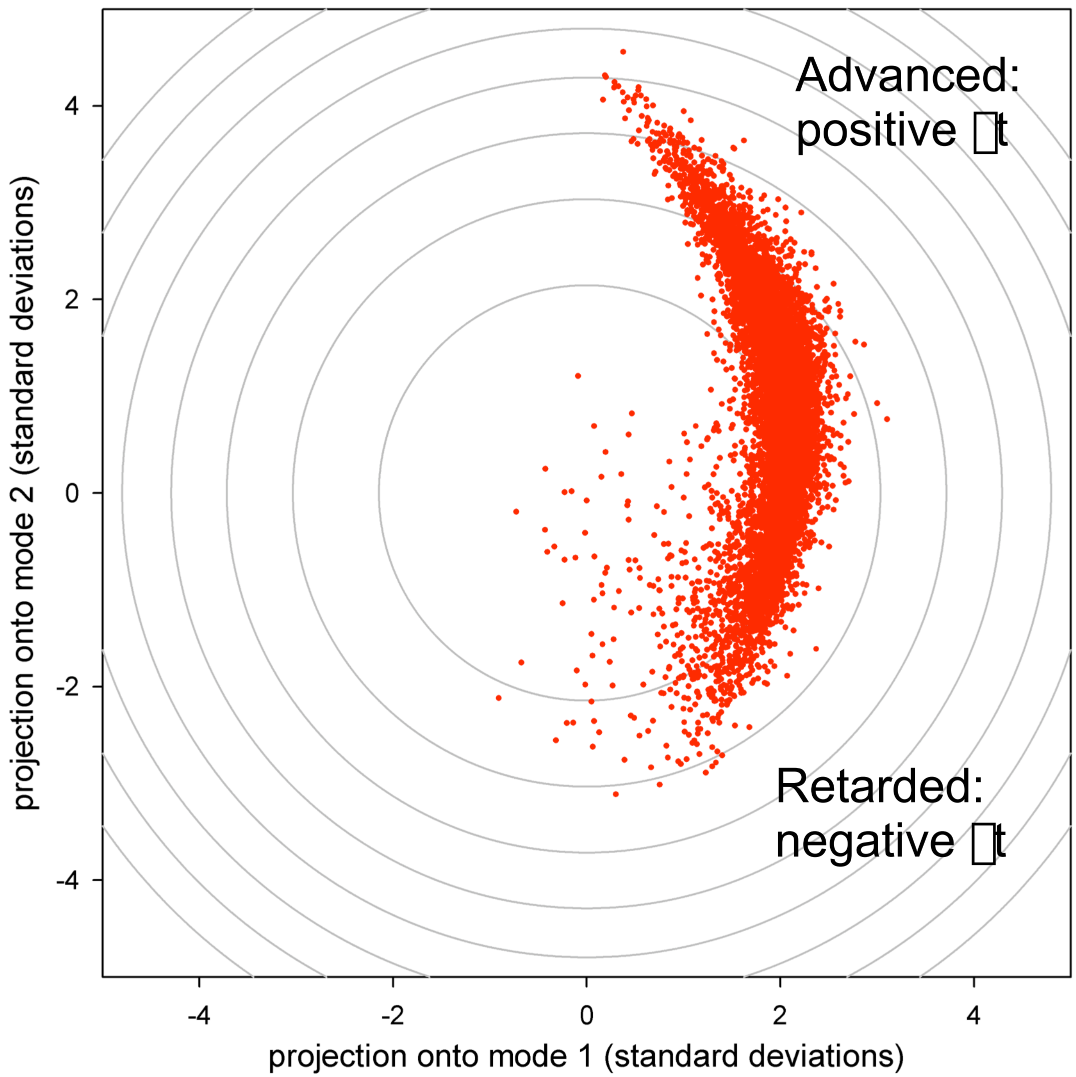
Beyond covariance

a simple attempt to capture the curvature
recovers 90% of the information in isolated spikes



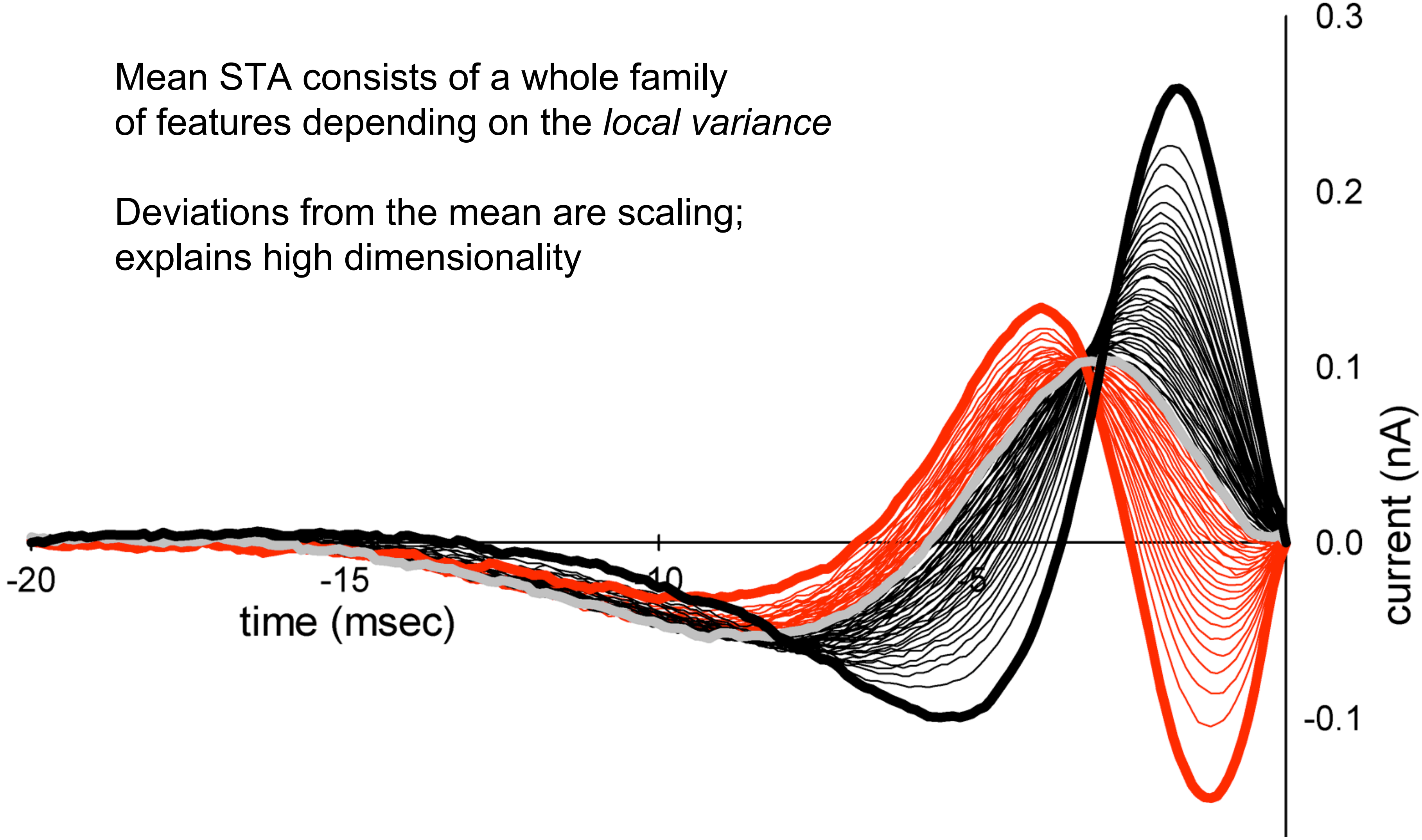
Aguera y Arcas et al., Neural Computation 2003

Helps to understand where the banana is coming from: slice it up

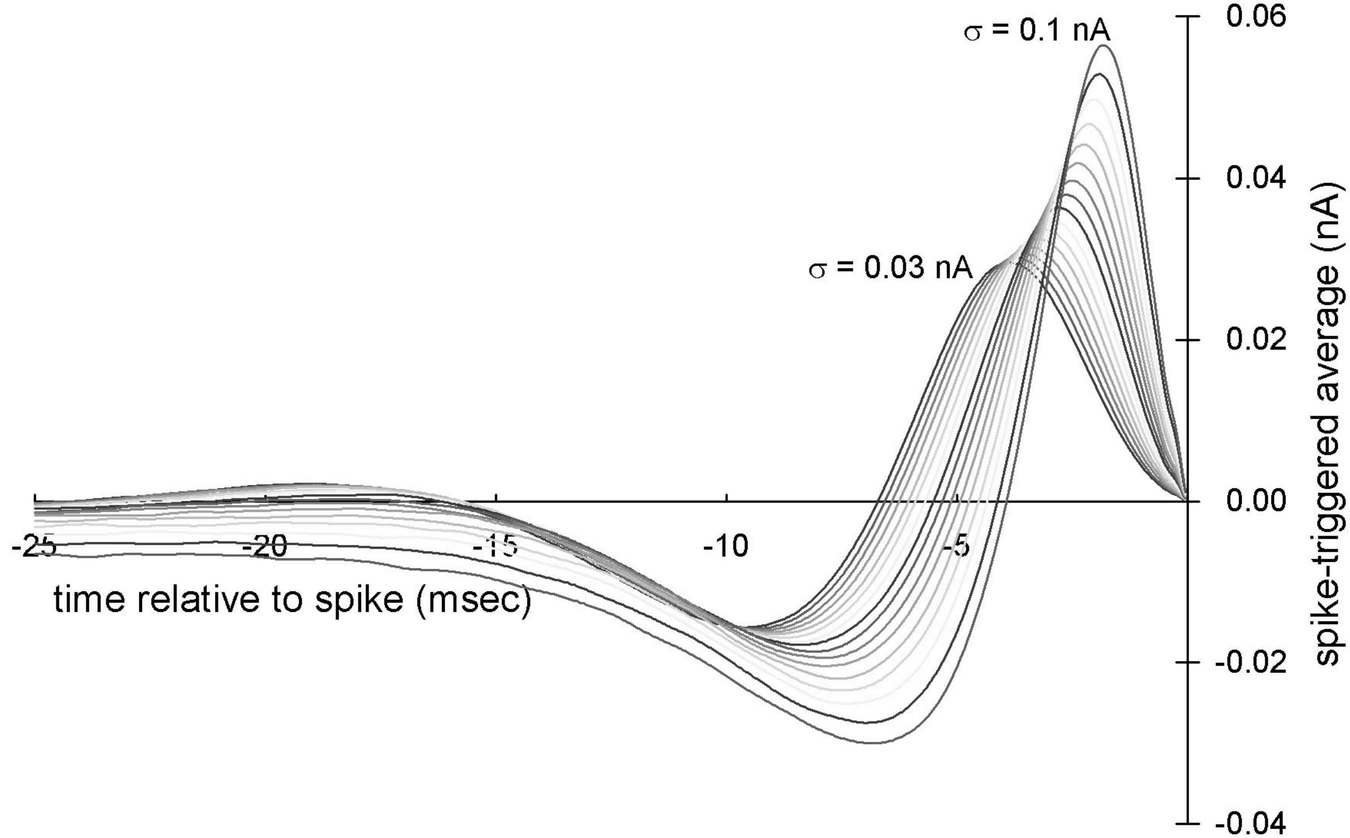


Mean STA consists of a whole family of features depending on the *local variance*

Deviations from the mean are scaling; explains high dimensionality



(not surprisingly) HH has instantaneous “feature adaptation”



But: no learning and no long timescales

Intrinsic nonlinearities may be responsible for some forms of adaptation to the stimulus distribution

There are experimental examples where the form of the adapting filter seems to maximise information transmission (van Hateren).

“Designability” of spiking surfaces?

White noise analysis can introduce confounds that look like adaptation.

Testbed: leaky integrate and fire.

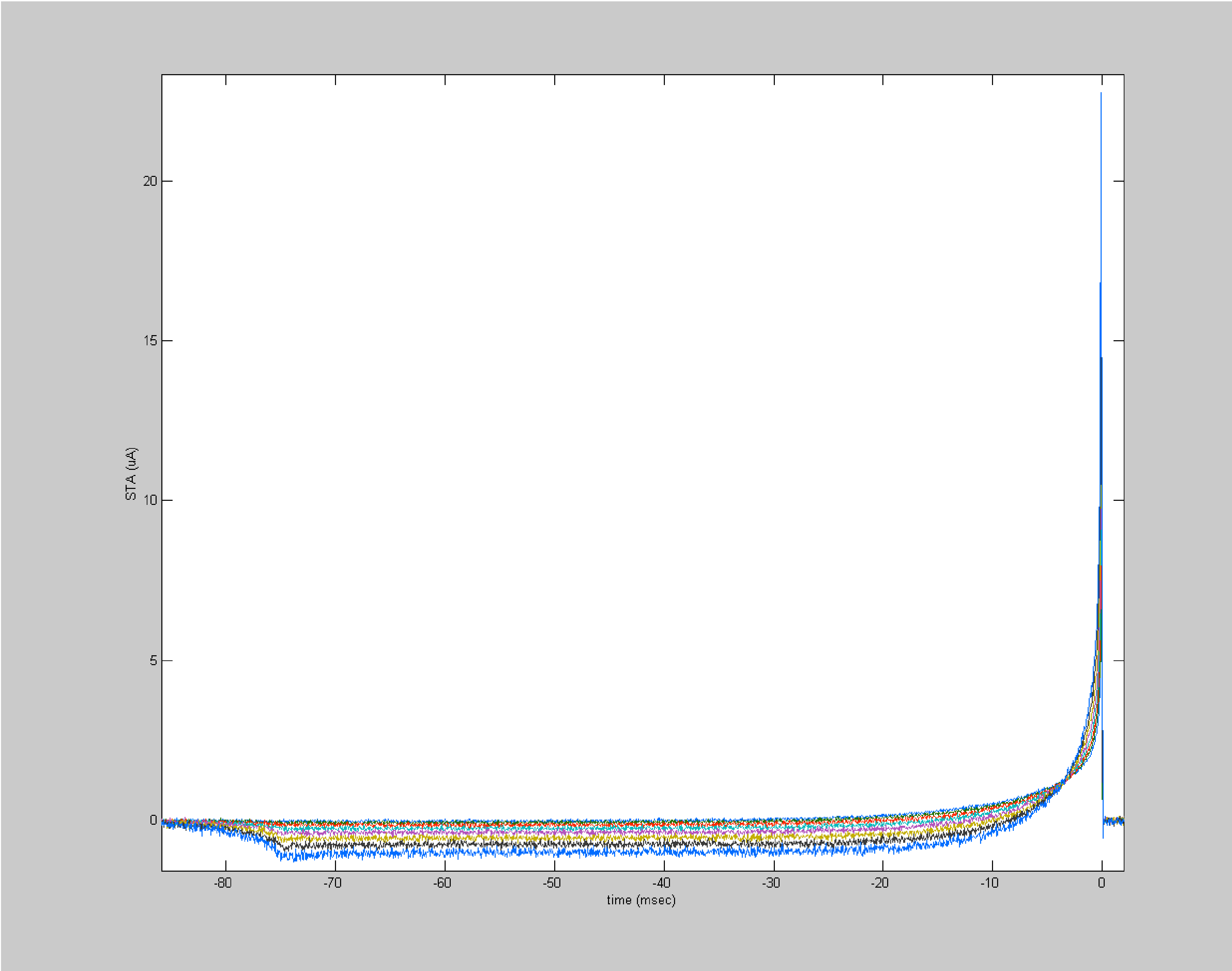
By definition, ONE dimension/linear filter controls the spiking decision, the exponential.

A good method should recover the exponential filter, independent of the stimulus variance.

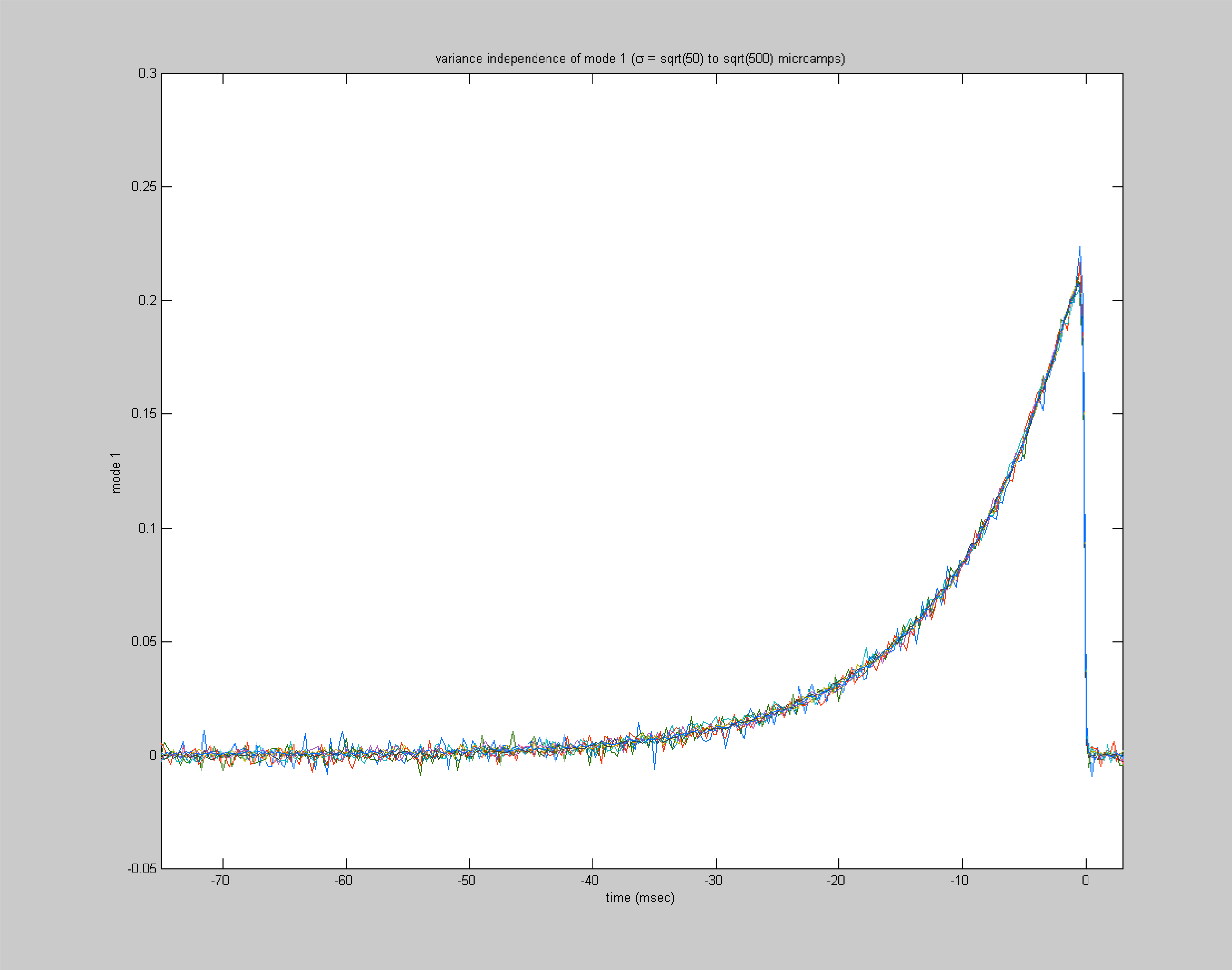
“Bare” white noise analysis is confounded by interactions between spikes. (Aguera y Arcas et al., 2000)

Solution: use only isolated spikes. We are trying to capture what in the ***stimulus*** is relevant for spiking, not the neuron’s internal state.

Using isolated spikes introduces its own complications! namely, high dimensionality. (nonGaussian prior)



STA (uA)
20
15
10
5
0
-80
-70
-60
-50
-40
-30
-20
-10
0
time (msec)


variance independence of mode 1 (σ = sqrt(50) to sqrt(500) microamps)
mode 1
time (msec)
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-70
-60
-50
-40
-30
-20
-10
0

Using covariance analysis on isolated spikes only, we are able to recover true, stimulus-invariant feature selection

Asymmetry of the question:
What is the best reconstruction filter? ("taking the organism's point of view")
vs.
What is the best predictor of a spike? (learning a functional model for spike generation from data).

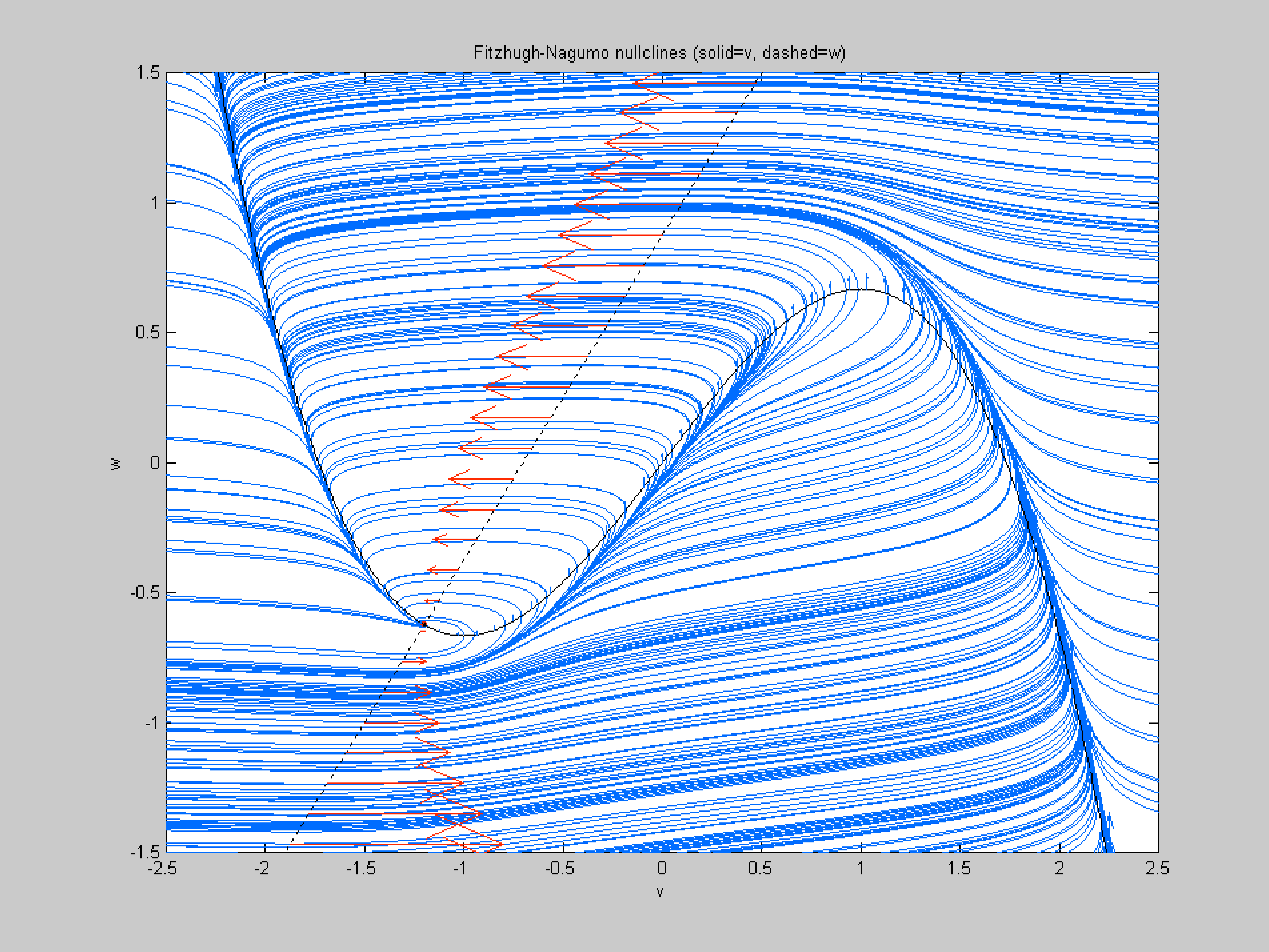
Conclusions

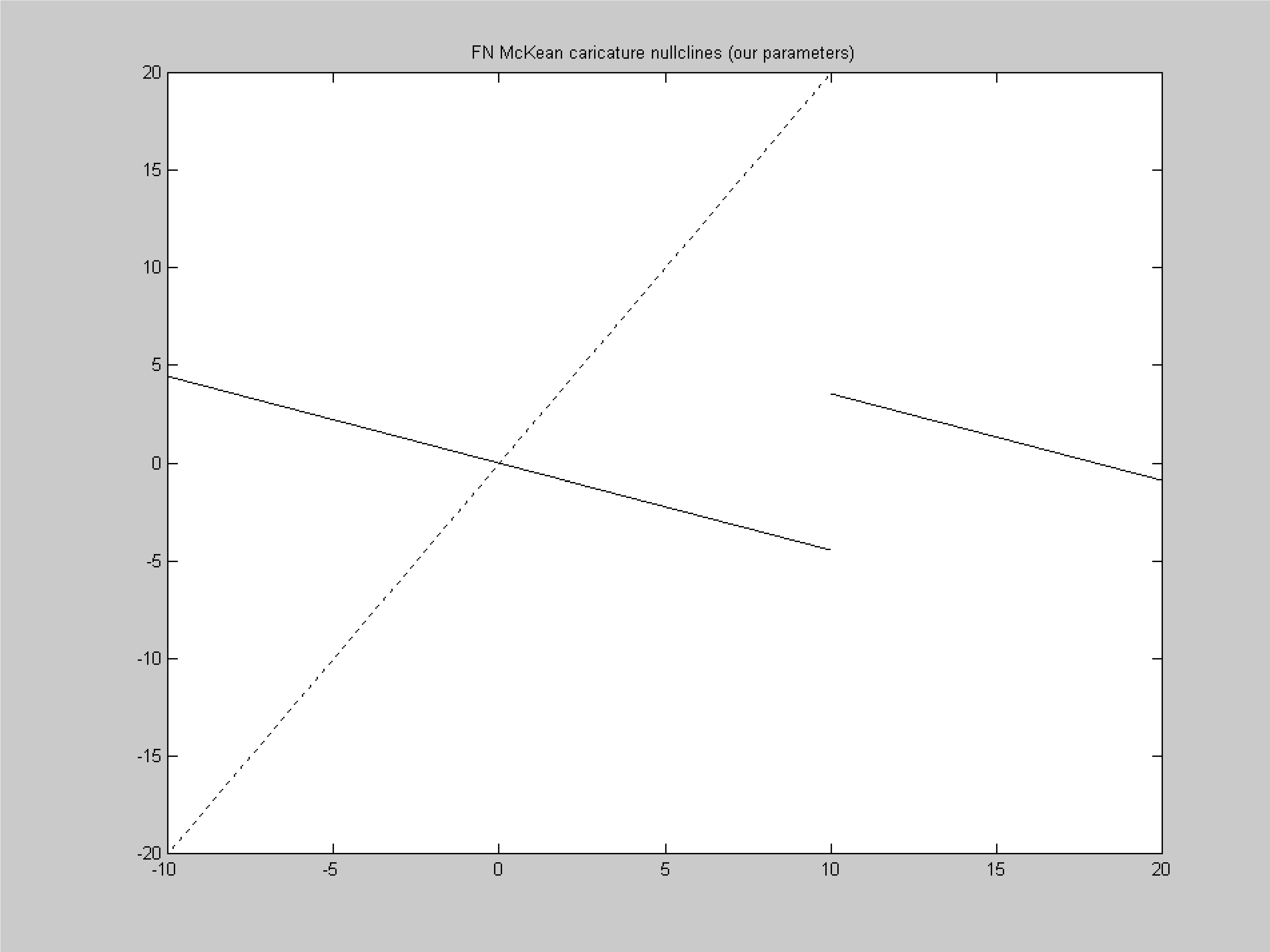
• adaptation to stimulus statistics in the fly visual system dynamically optimises information transmission

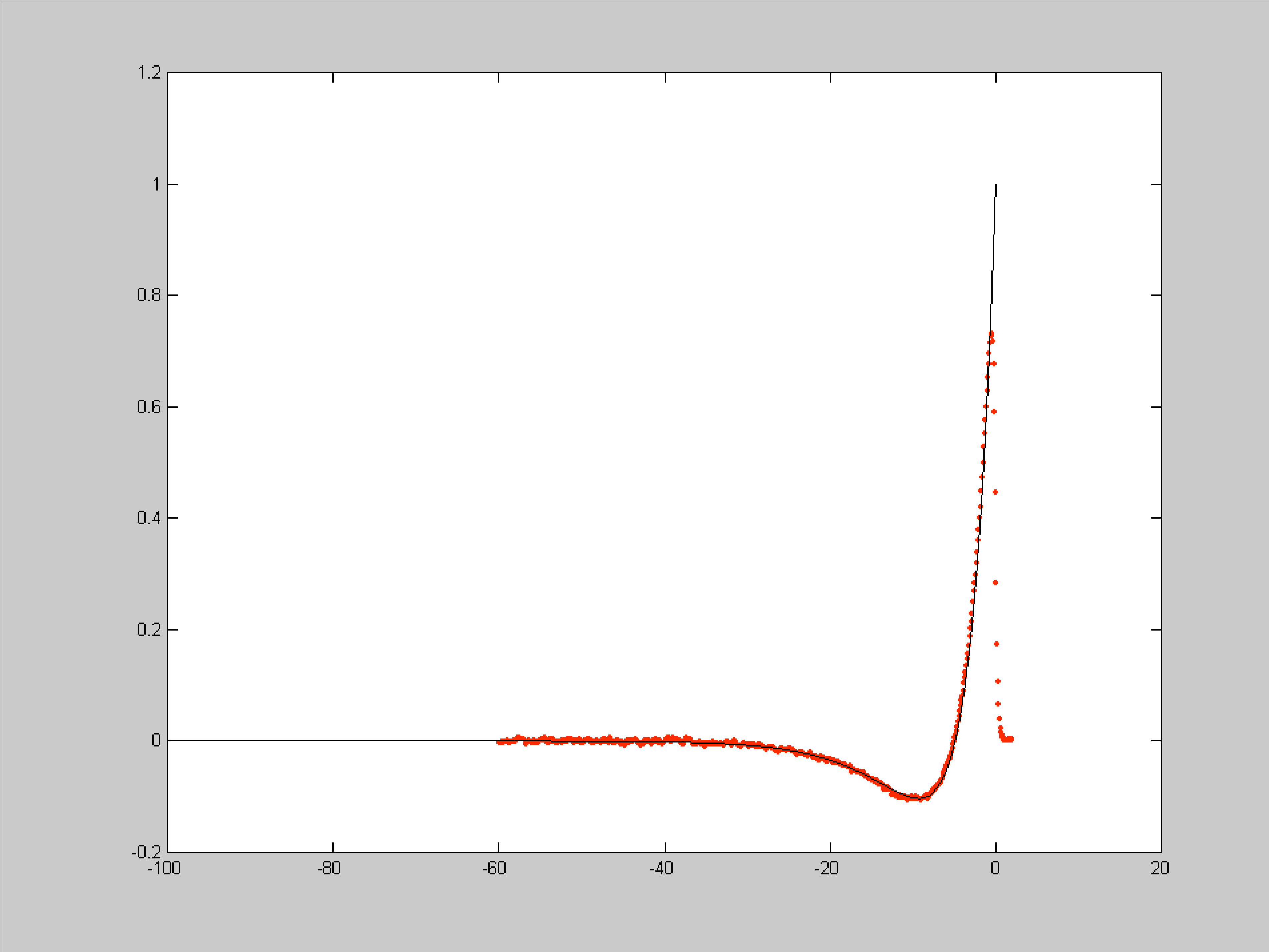
• rate dynamics are almost perfectly described by fractional differentiation.

• adaptation-like behaviour arises from simple models without "memory"

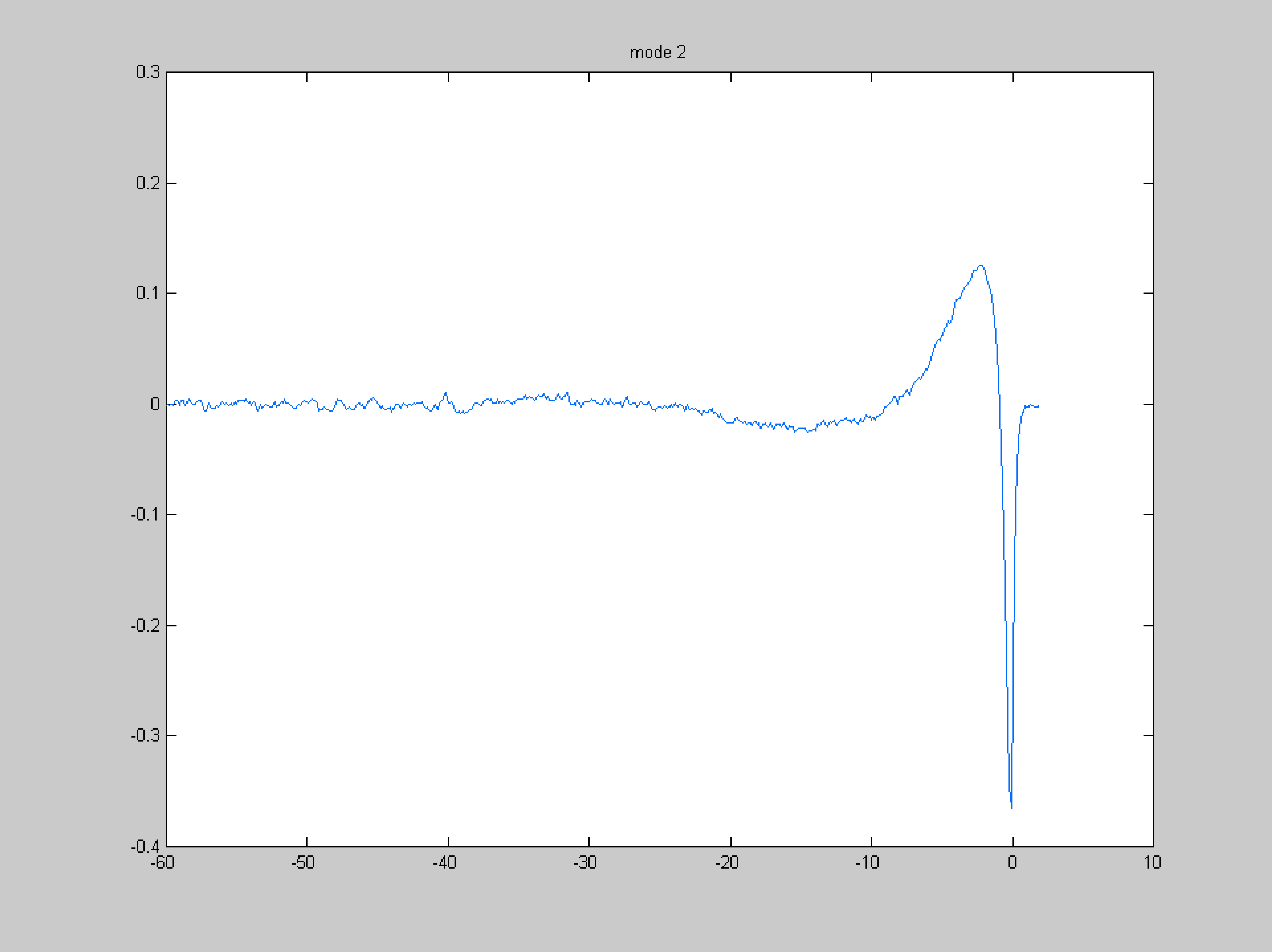
• intrinsic neural nonlinearities may be tuned to support advantageous information processing strategies


Fitzhugh-Nagumo nullclines (solid=v, dashed=w)
w
v
1.5
1
0.5
0
-0.5
-1
-1.5
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5

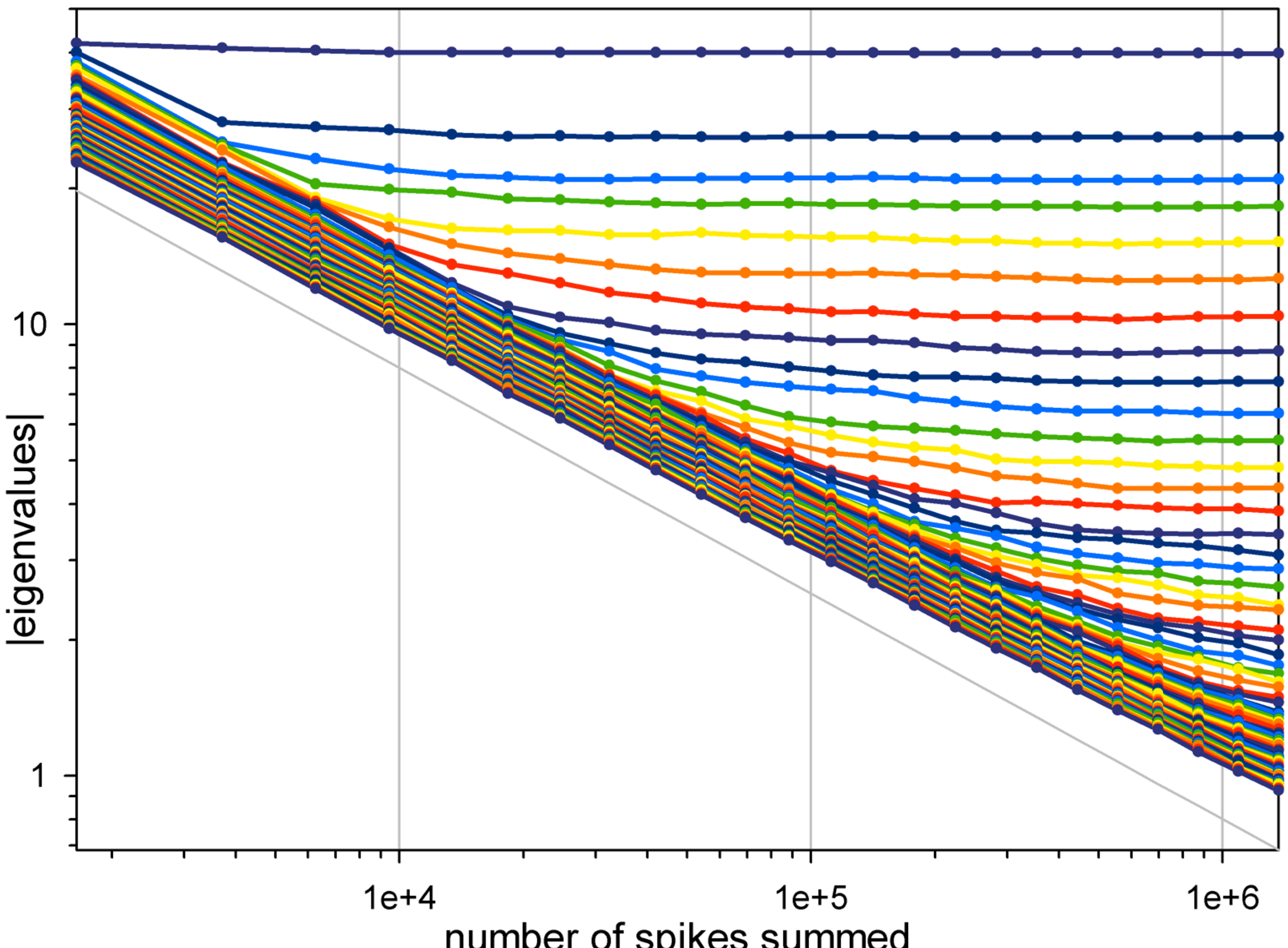

FN McKean caricature nullclines (our parameters)
20
15
10
5
0
-5
-10
-15
-20
-10
-5
0
5
10
15
20



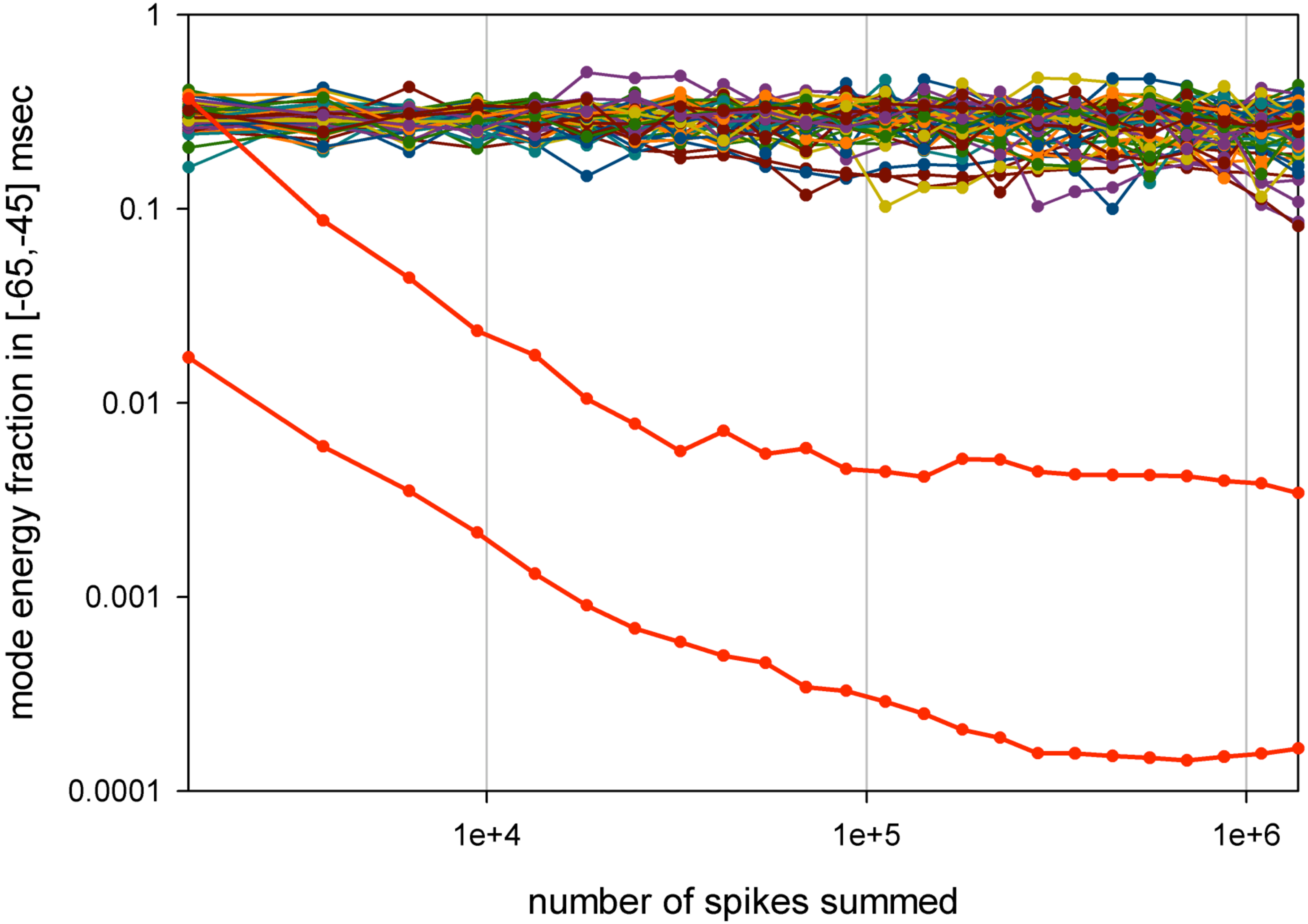
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-100
-80
-60
-40
-20
0
20


mode 2
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-60
-50
-40
-30
-20
-10
0
10

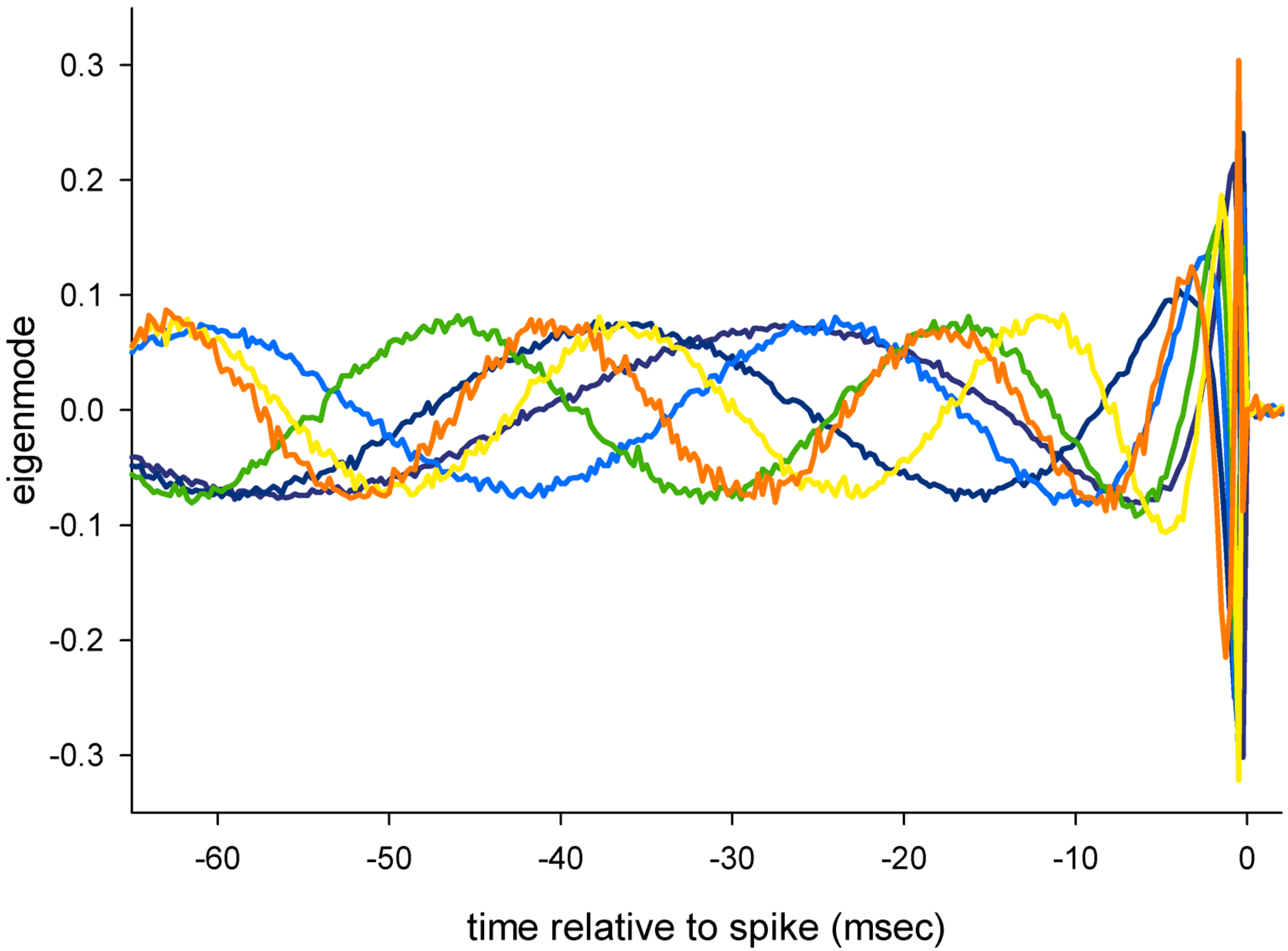
isolated spike triggered covariance spectrum



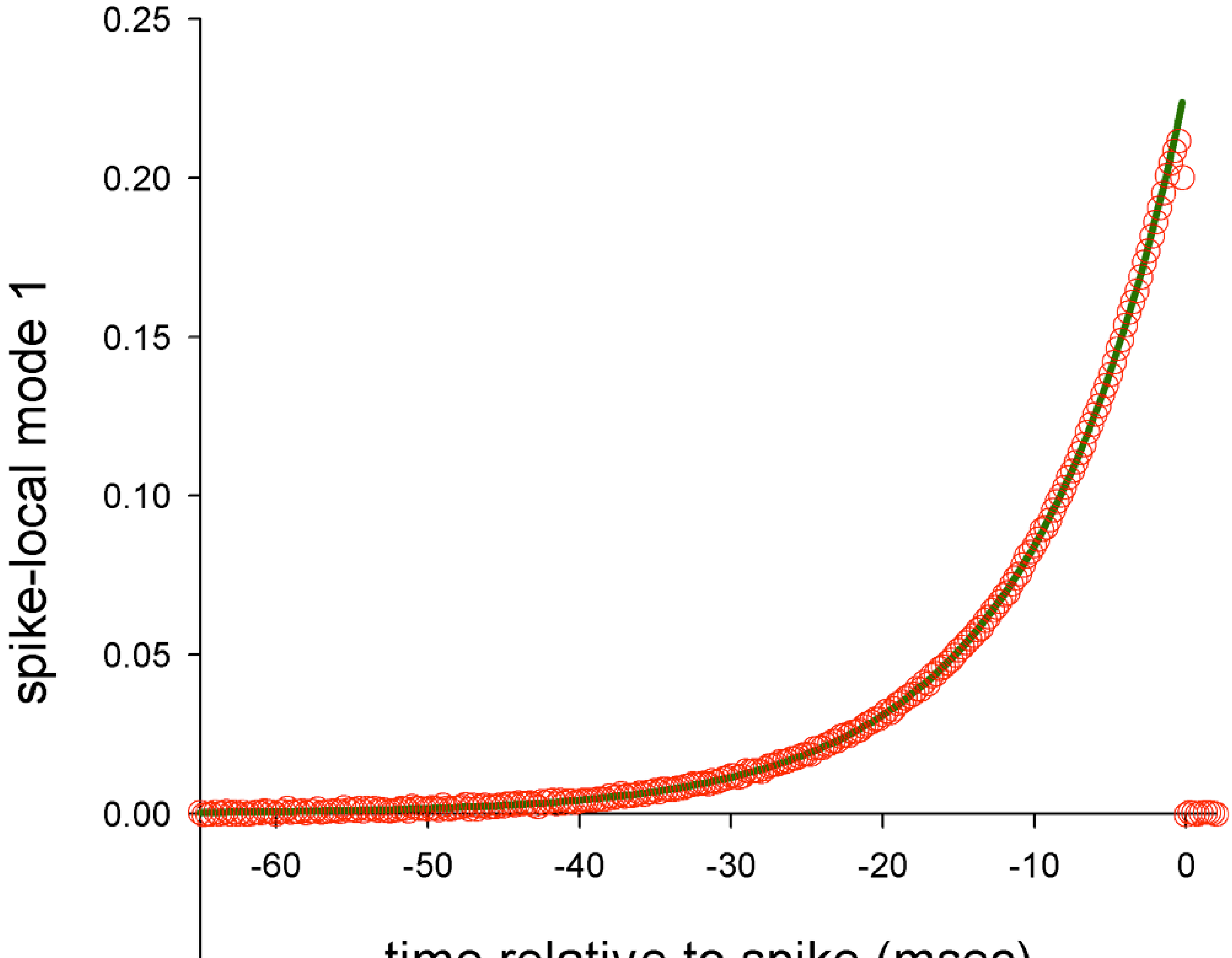
silence power diagnostic



silence modes



outstanding spike mode



threshold-crossing mode