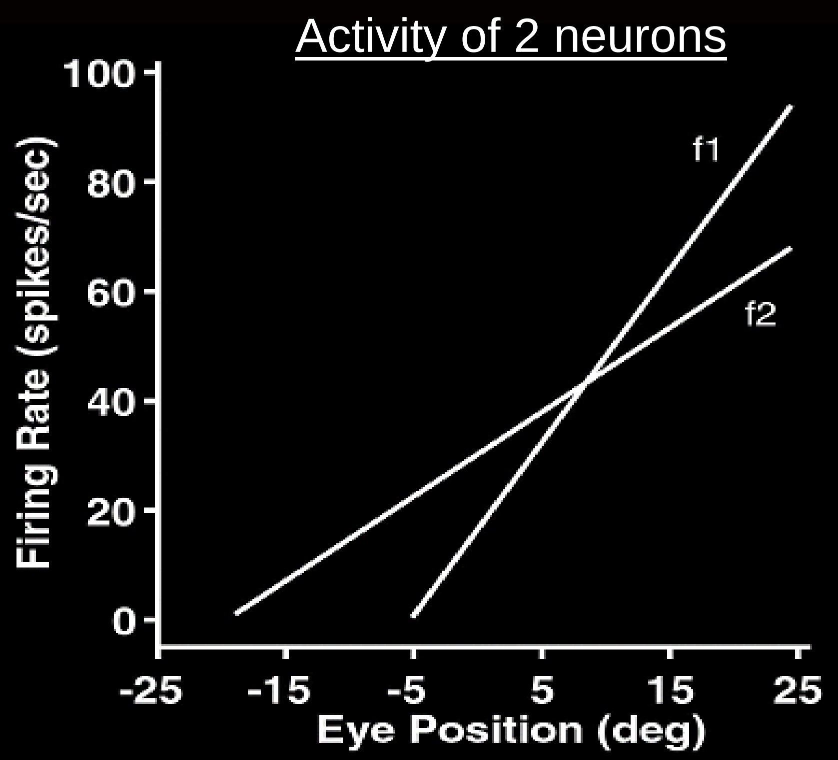
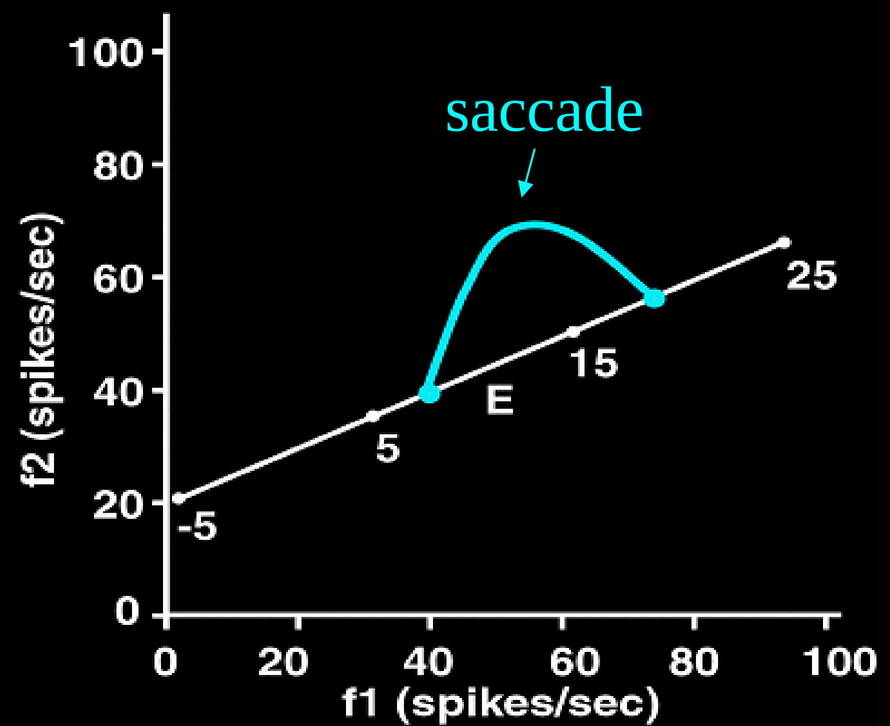
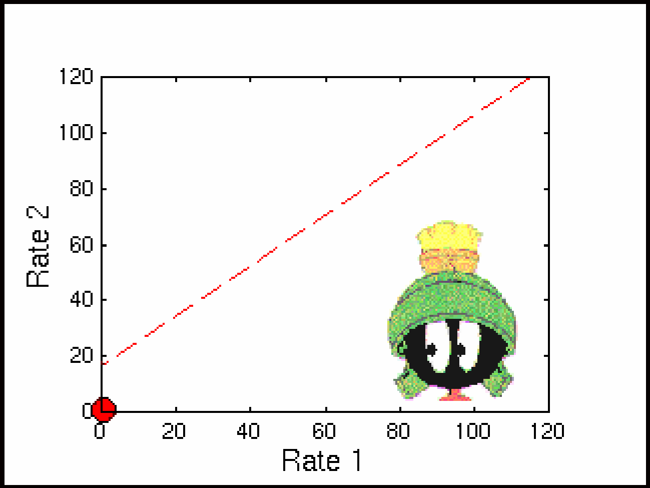
Robustness in Neurons & Networks

Mark Goldman

Many-neuron Patterns of Activity Represent Eye Position





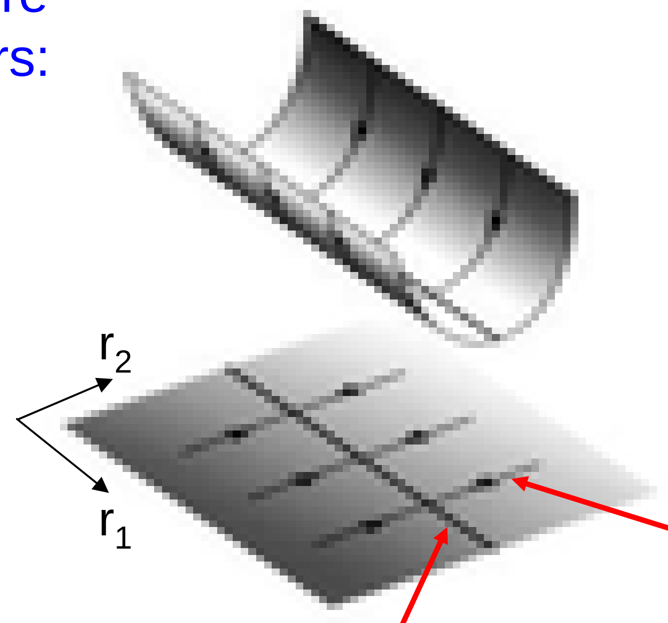


eye position represented by location along a low dimensional manifold ("line attractor")

(H.S. Seung, D. Lee)

Line Attractor Picture of the Neural Integrator

Geometrical picture
of eigenvectors:



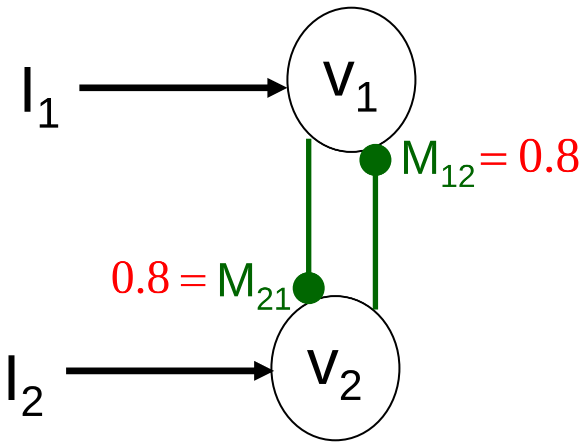
Axes = firing rates
of 2 neurons

Decay along direction
of decaying eigenvectors

No decay or growth along direction
of eigenvector with eigenvalue = 1

“Line Attractor” or “Line of Fixed Points”

Examples



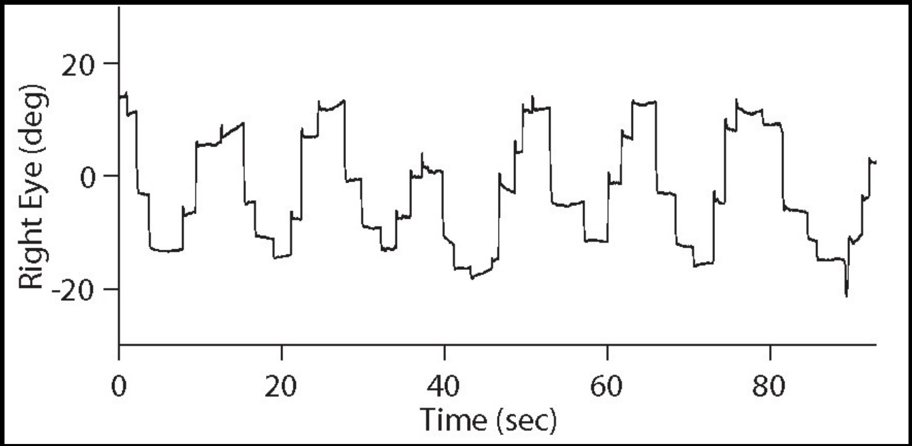
Q: Can you guess what input pattern **I** will be amplified most?
(i.e. eigenvector with largest λ)

Which will be compressed most?
(i.e. eigenvector with smallest λ)

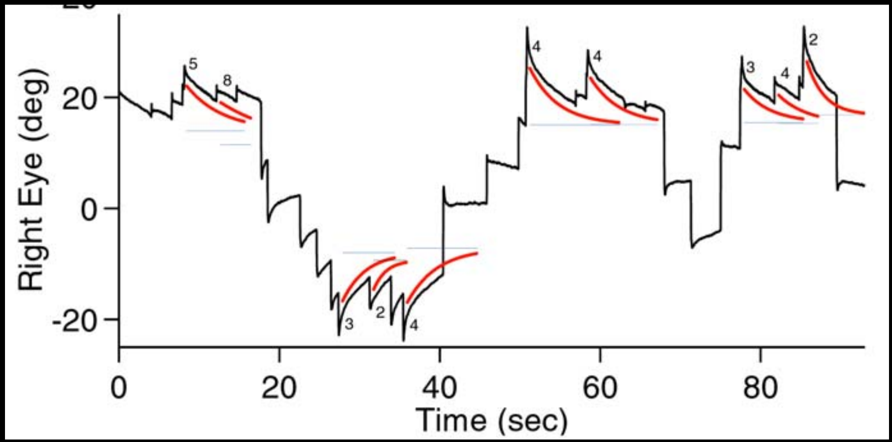
A: [1 1] is amplified most → amplifies common input
[1 -1] is compressed most → attenuates differences

Effect of Bilateral Network Lesion

Control



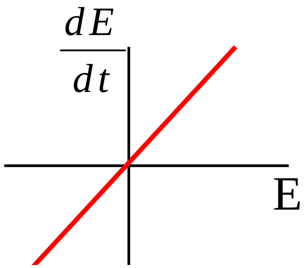
Bilateral Lidocaine:
Remove Positive Feedback

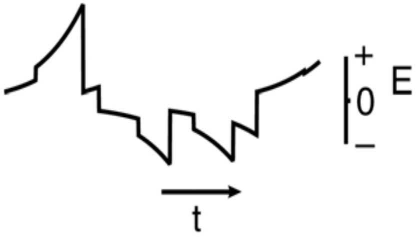


Unstable Integrator

Human with unstable integrator:

$$\frac{dE}{dt} = (w-1)E$$

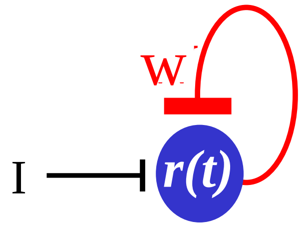




Issue: Robustness of Integrator

Integrator equation:

$$\tau_{bio}\frac{dr}{dt} = -1r + wr + I$$



$$\tau_{network} = \frac{\tau_{bio}}{|1-w|}$$

Experimental values:

Single isolated neuron: $\tau_{bio} \sim 100$ ms

Integrator circuit: $\tau_{network} \sim 10$ sec

➡ Synaptic feedback w must be tuned to accuracy of:

$$|1-w| = \frac{\tau_{bio}}{\tau_{network}} \sim 1\%$$

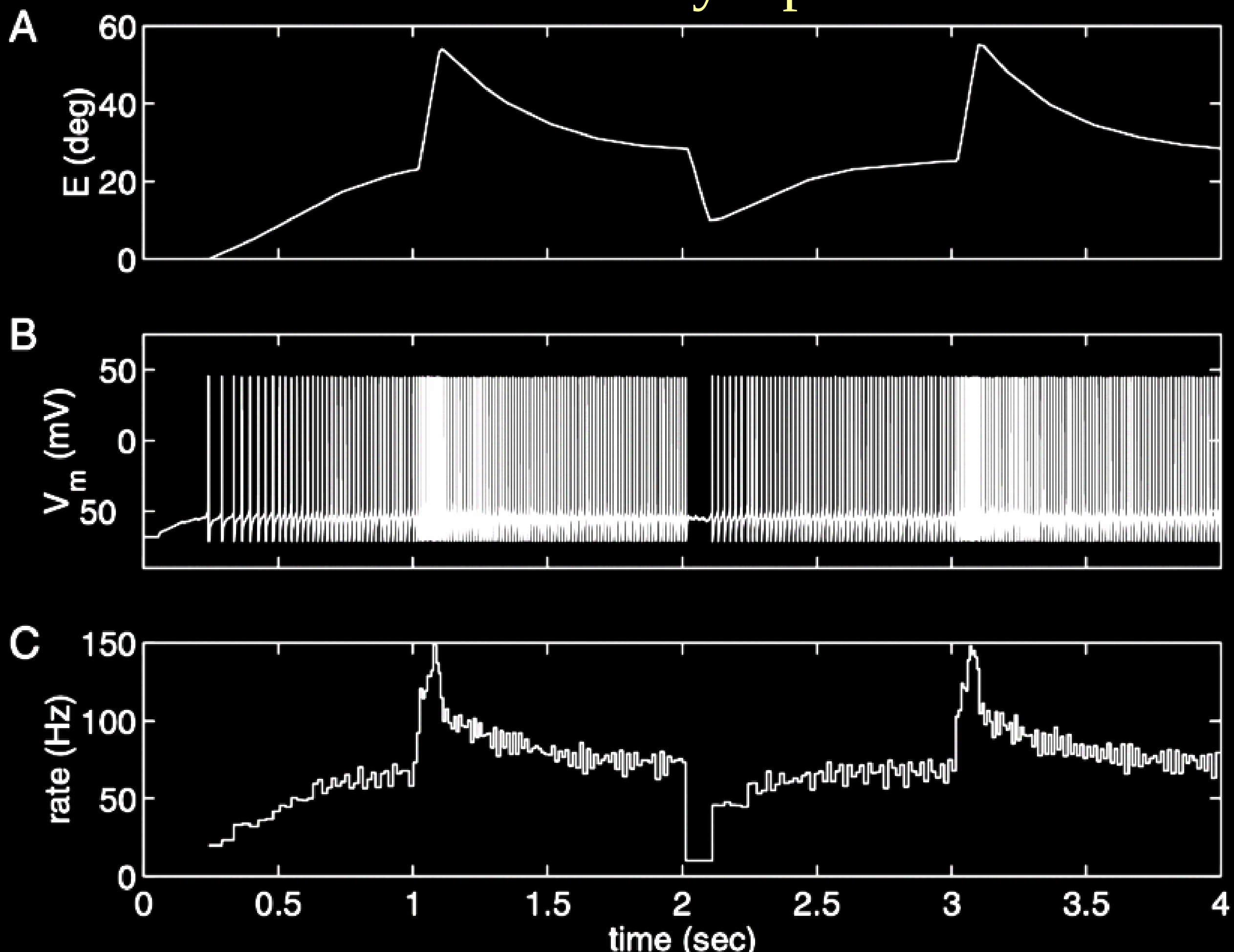
Weakness: Robustness to Perturbations

Imprecision in accuracy of feedback connections severely compromises performance (memory drifts: leaky or unstable)

Model:

(Seung et al., 2000)

10% decrease in synaptic feedback



Robustness in Dynamical Systems

Robustness refers to:

A. Low sensitivity of a system to perturbations

B. Ability to recover, over time, from a perturbation (e.g. plasticity, drug tolerance)

Issues to consider:

1) Time scale for robust behavior

2) What perturbations is a system robust against?
   - Design systems to resist the most common perturbations

3) What features of a system's output are robust to a particular perturbation?

4) What are the signatures of a system exhibiting various robustness mechanisms?

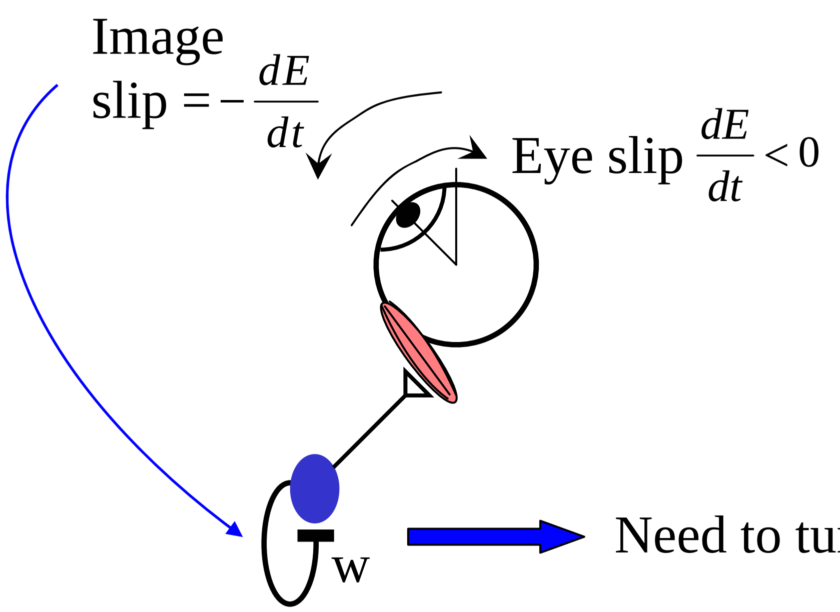
Learning to Integrate

How accomplish fine tuning of synaptic weights?

→ IDEA: Synaptic weights w learned from “image slip”

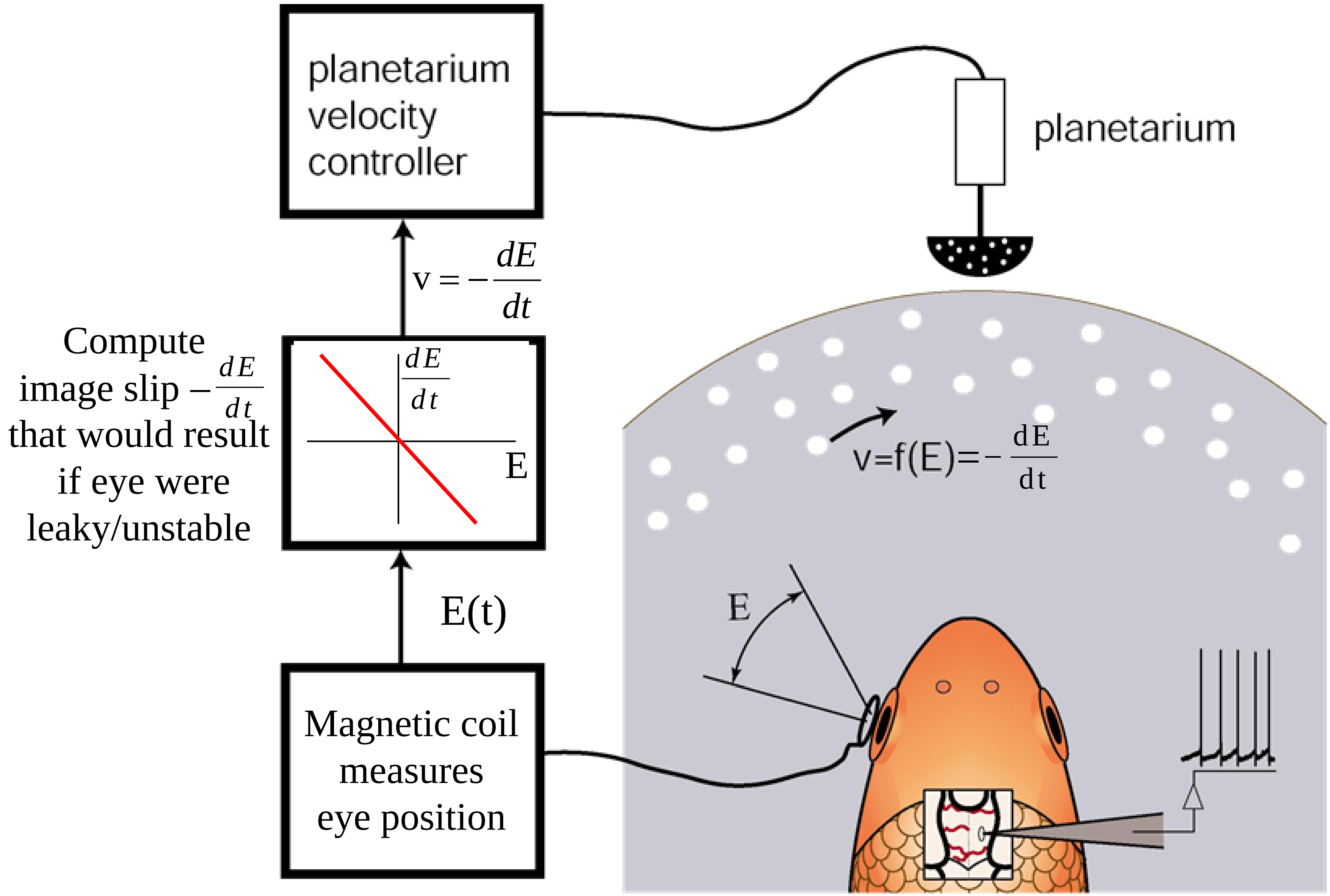
(Arnold & Robinson, 1992)

E.g. leaky integrator:



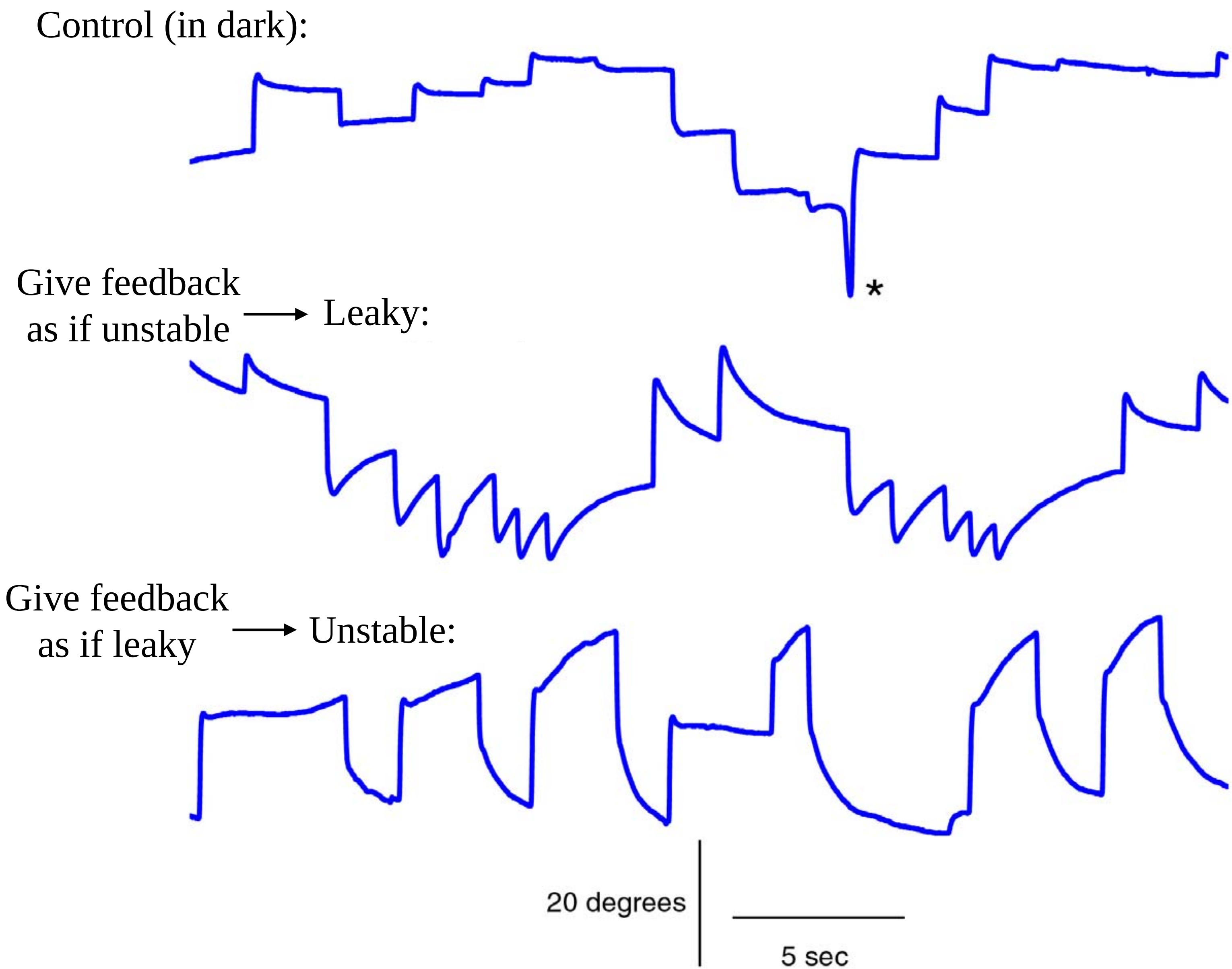
→ Need to turn up feedback → $\Delta w \propto -\frac{dE}{dt}$

Experiment: Give Feedback as if Eye is Leaky or Unstable




LOOK HERE
Planetarium
Search Coils
z-axis rate table

Integrator Learns to Compensate for Leak/Instability!



Previous Example:

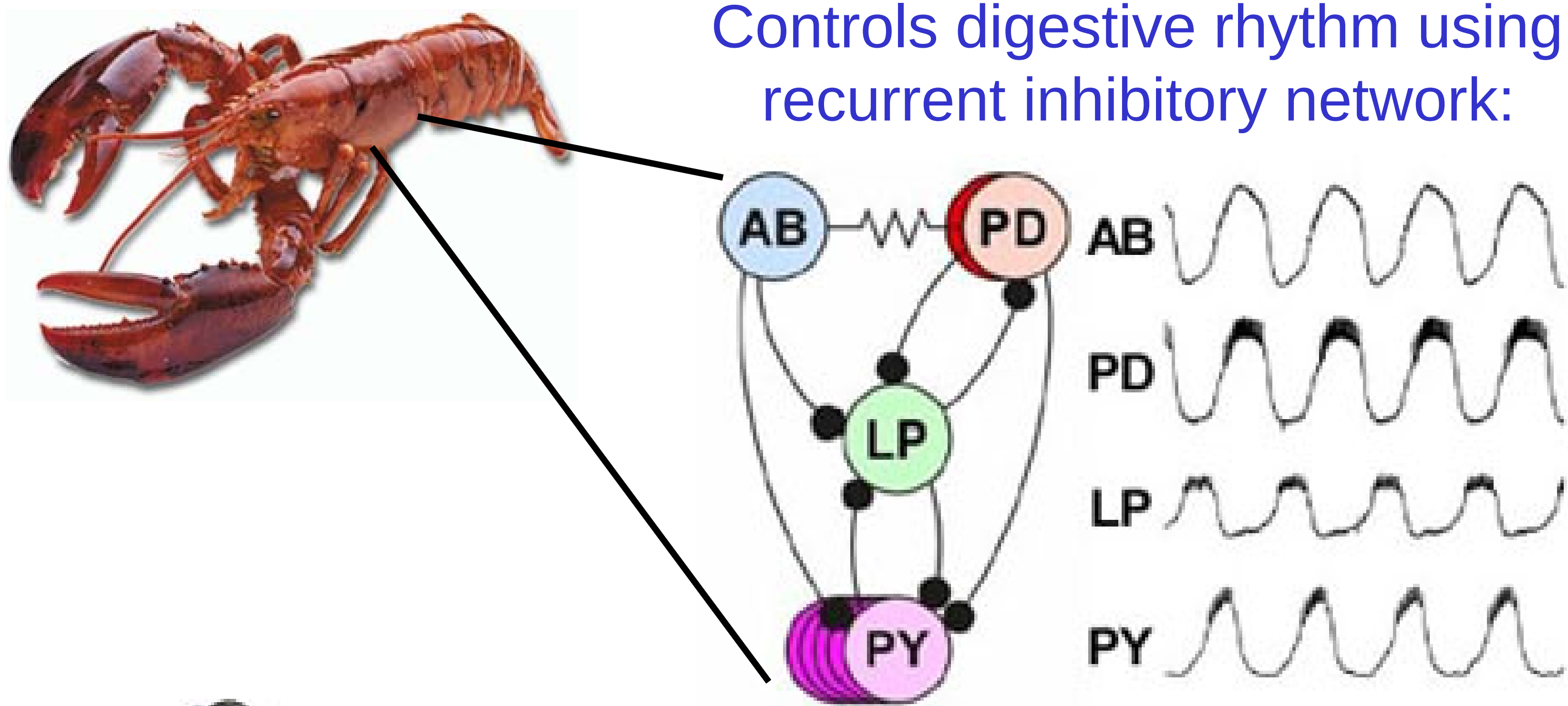
Error signal to tune network is due to *sensory error* (image slip on retina)

Question:

- Might systems have *intrinsic* monitors of activity to accomplish tuning?
- What might be the signatures of a system that utilizes such a mechanism?

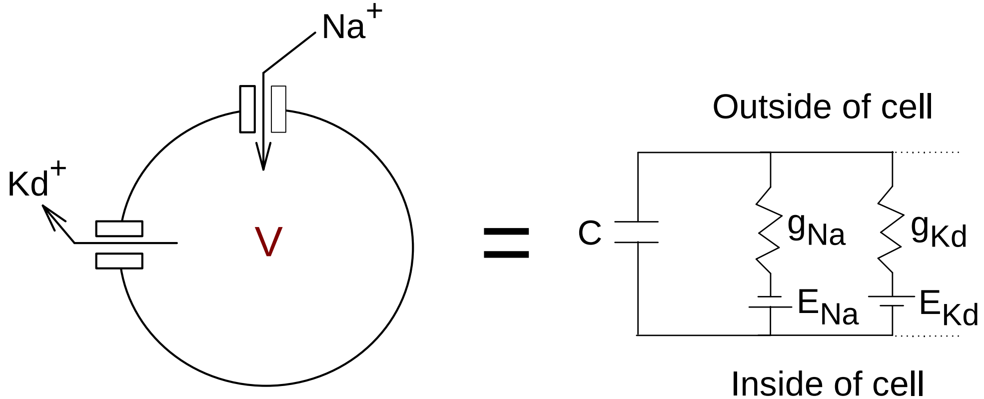
Pattern Generating Network:
Stomatogastric ganglion (STG) of crab/lobster stomach

Controls digestive rhythm using recurrent inhibitory network:



Conductance-based neuron models

Electrical circuit model of neuron:



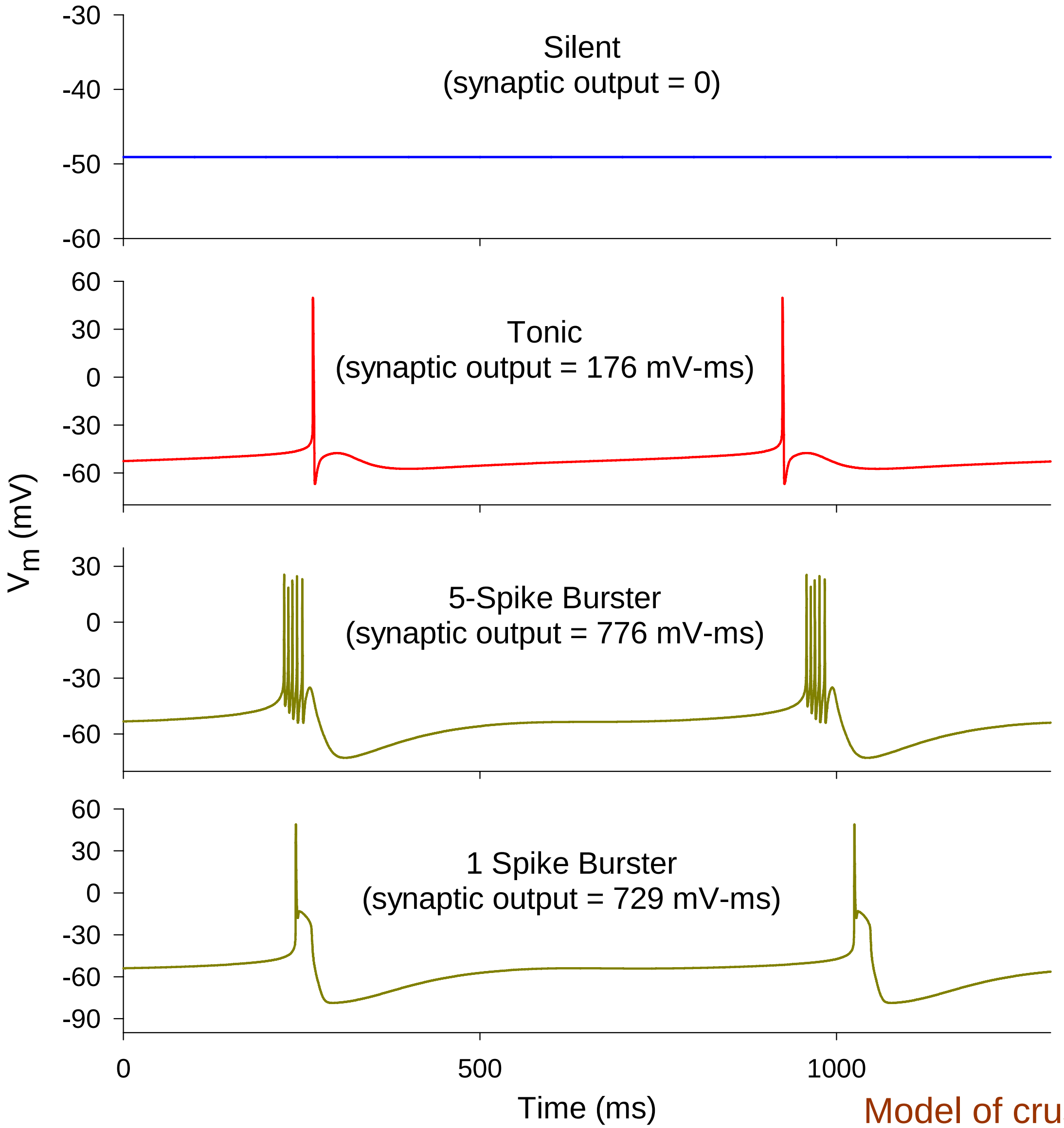
$$C \frac{dV}{dt} = \sum_i g_{max,i}\, p_{open,i}(V)\ (E_i - V)$$

i = conductance type = Na, Ca, A, KCa, Kd

$p_{open}(V)$ = probability channel is open

Inward currents (increase V)	Outward currents (decrease V)
Na (fast)	Kd (fast)
Ca (slower)	A (slower)
	KCa (slowest)

Sample of Firing States Observed



Model of crustacean STG neurons
based on data of Turrigiano et al., 1995

Identified neurons, yet different conductances

Identified neurons:

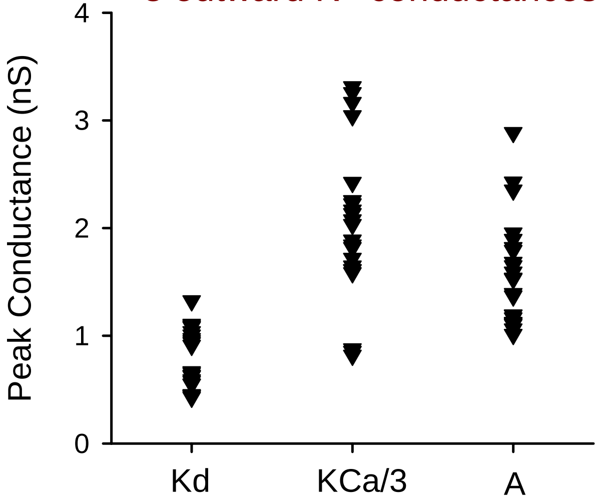
>Same location, morphology, function

>Traditional view:

- Same conductances
- Each conductance has unique role

Data:

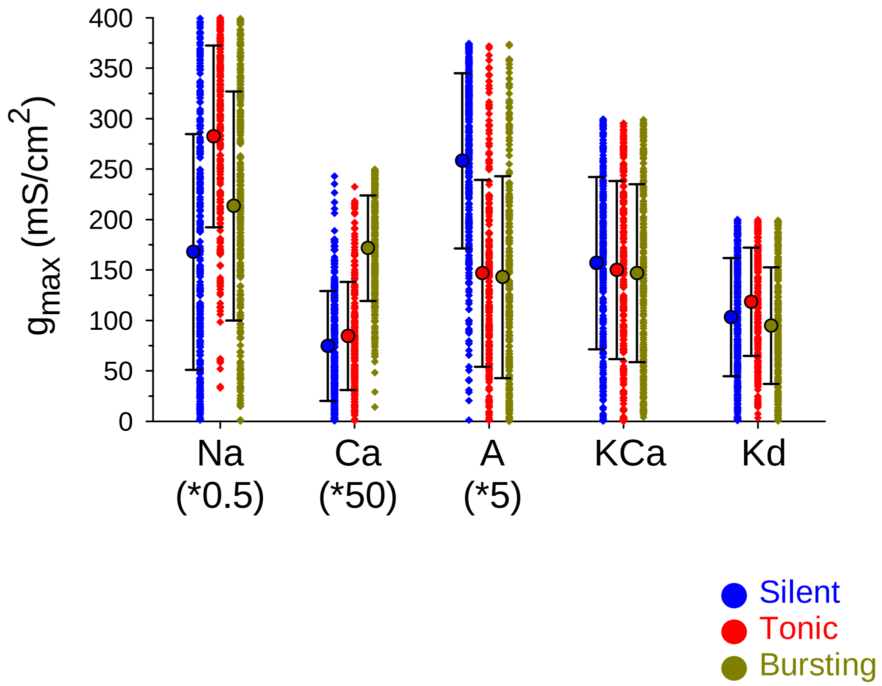
3 outward K^+ conductances



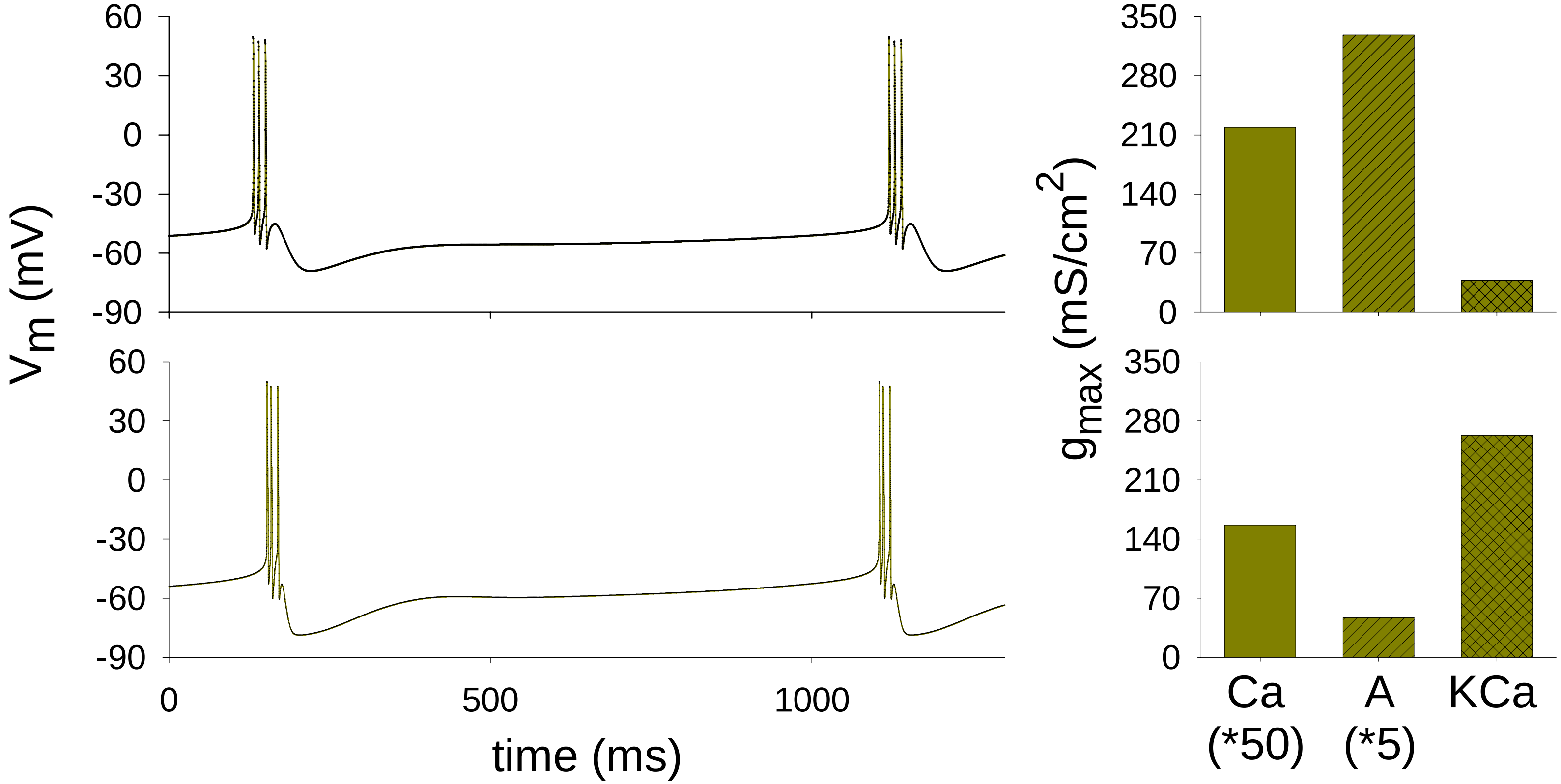
(Crab IC neuron; Golowasch et al., 1999)

Can different conductances give similar firing?

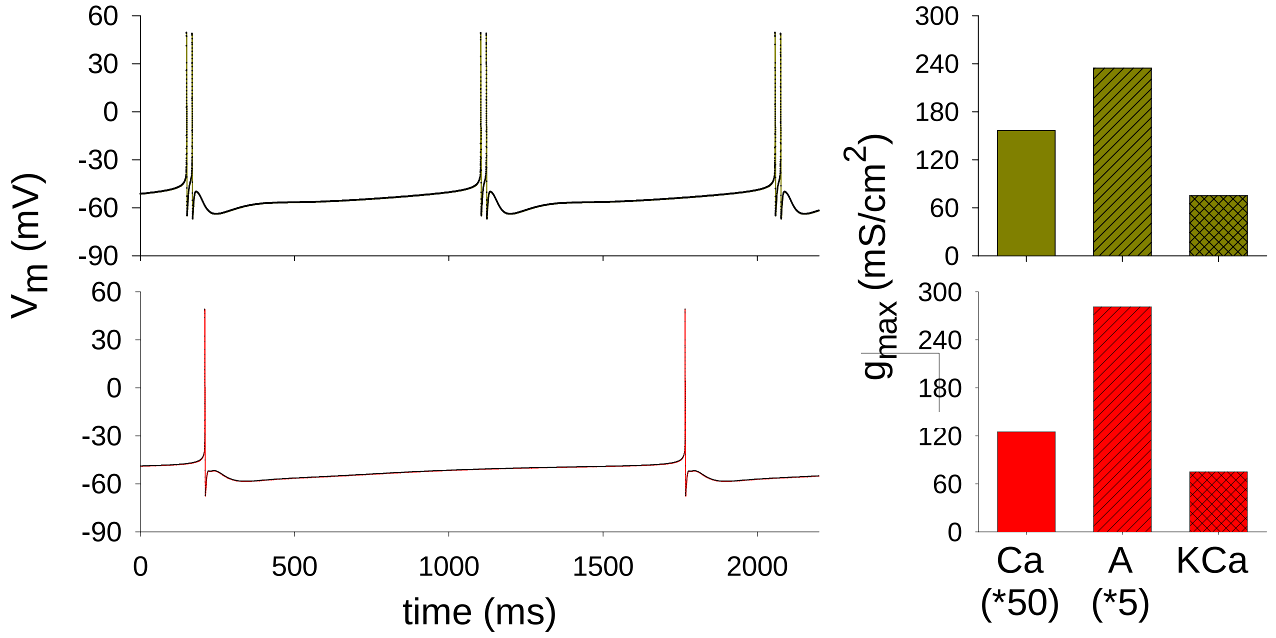
Single Conductances Do Not Determine Firing State



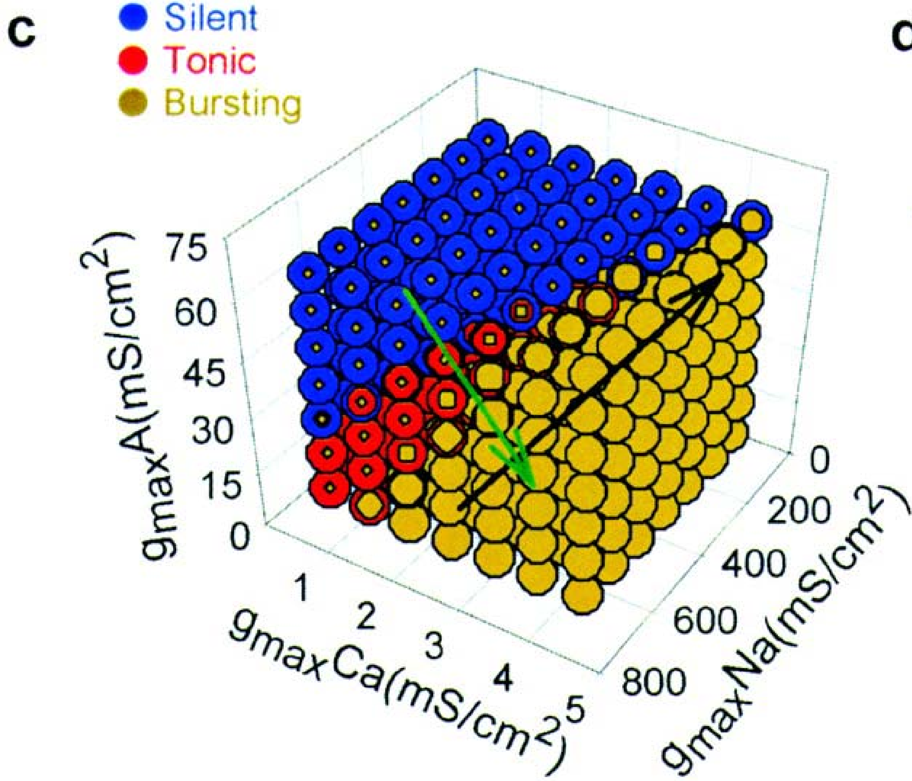
Similar firing, different conductances:



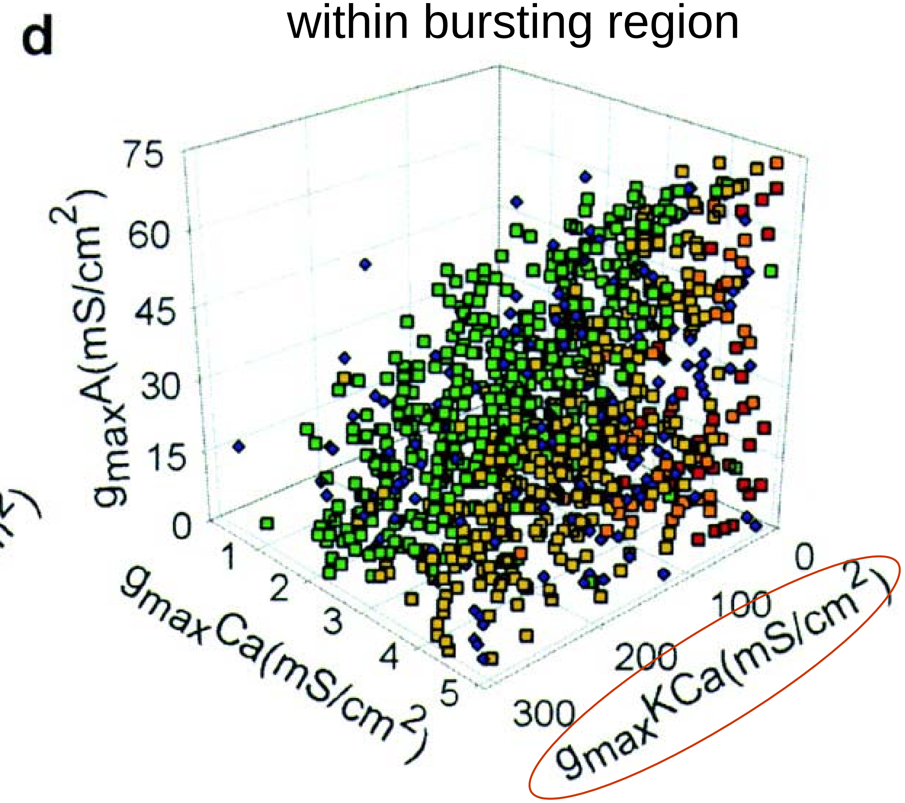
Different firing, similar conductances:



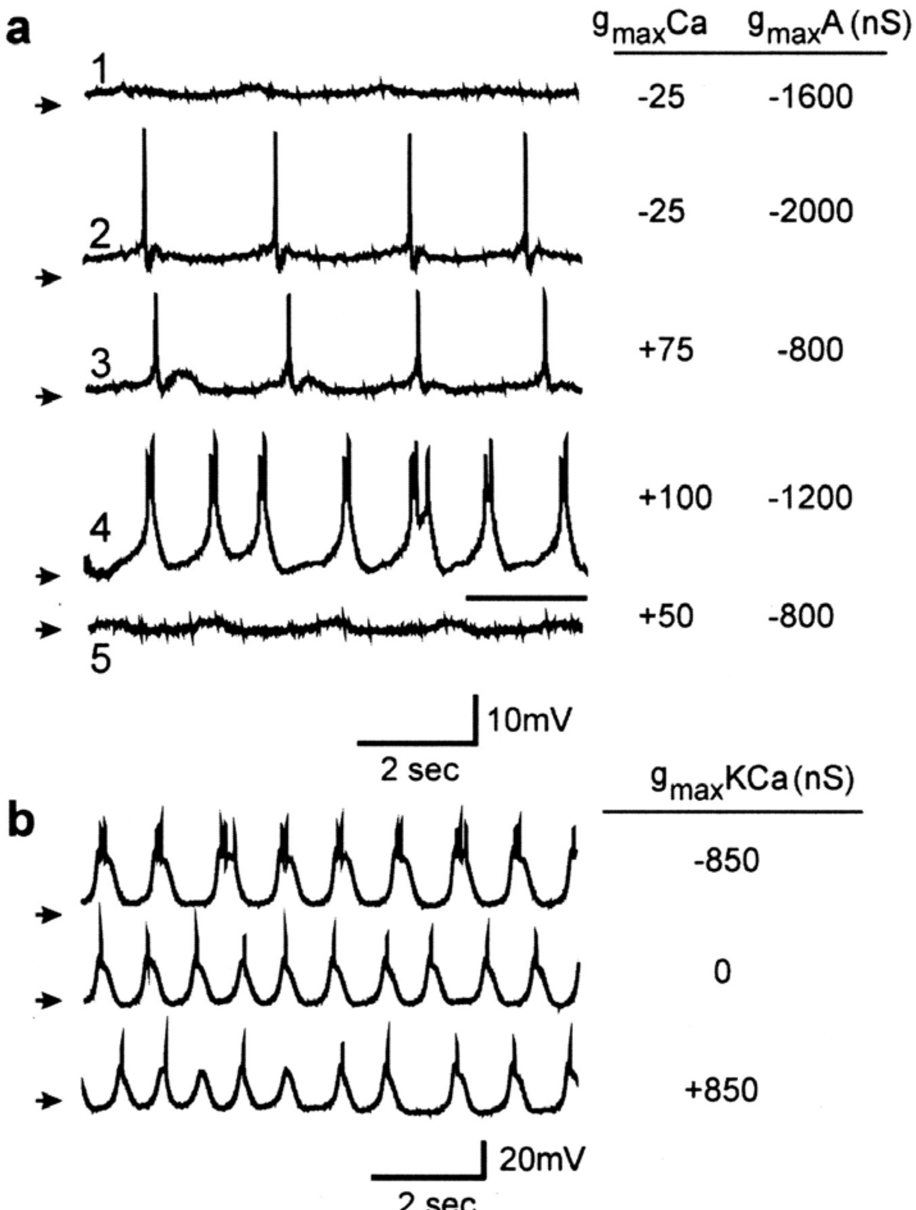
Firing State Diagram:
Combinations of Conductances Better Determine Firing State

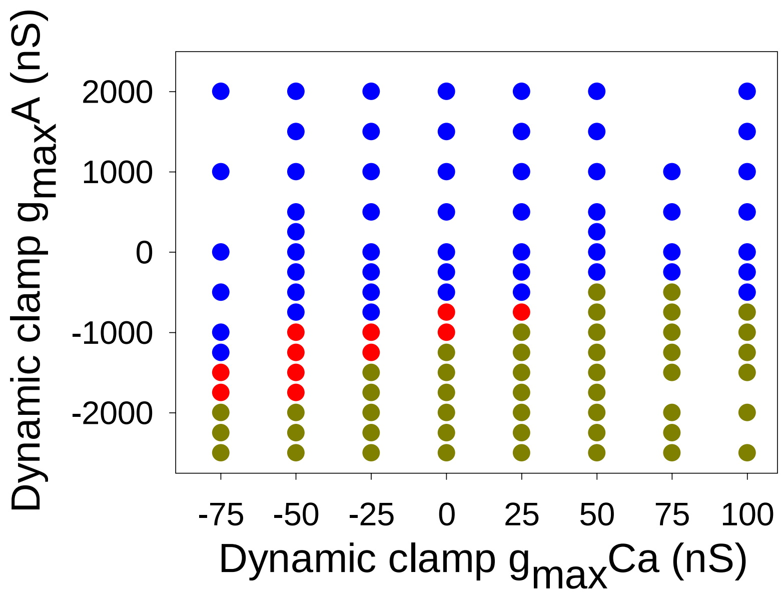


# of spikes/burst
within bursting region



Real Data

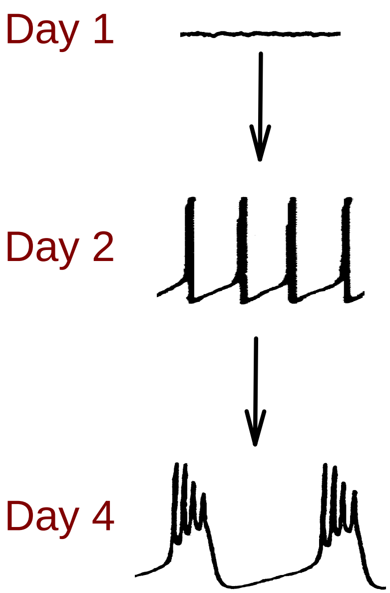

a
gmaxCa gmaxA (nS)
1
-25 -1600
2
-25 -2000
3
+75 -800
4
+100 -1200
5
+50 -800
10mV
2 sec
b
gmaxKCa (nS)
-850
0
+850
20mV
2 sec

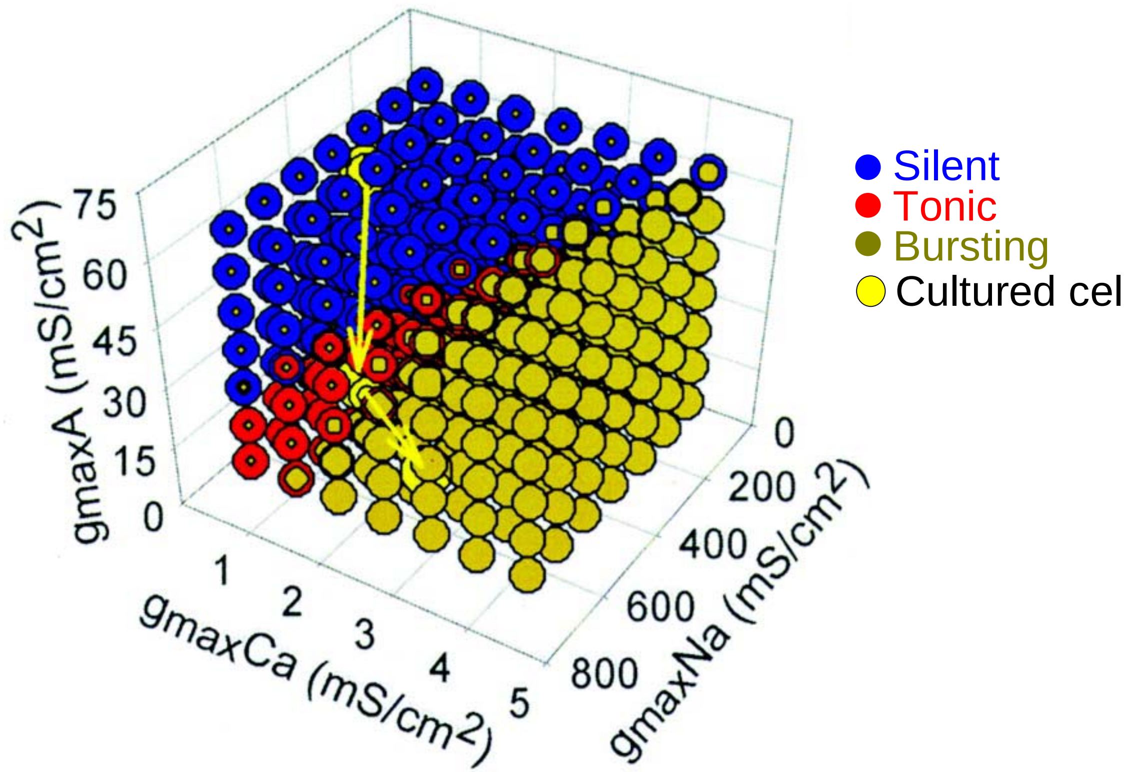

Dynamic clamp gmaxA (nS)
2000
1000
0
-1000
-2000
-75
-50
-25
0
25
50
75
100
Dynamic clamp gmaxCa (nS)

Neurons regulate firing to achieve rhythmic pattern

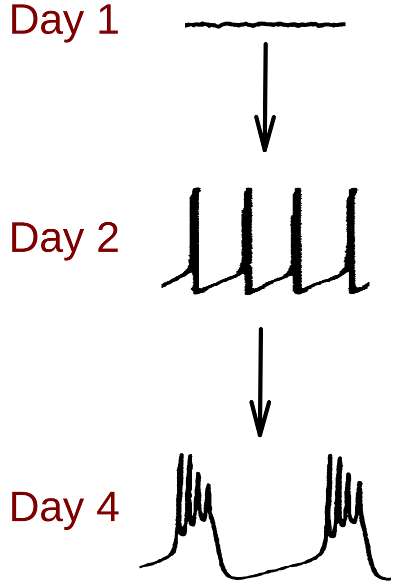
Change in firing during development:

(Turrigiano et al., 1995)





Homeostatic learning rule to recover activity



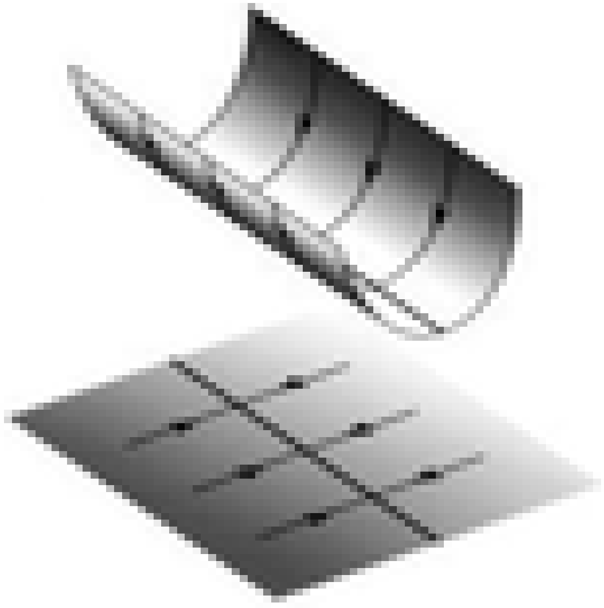
Idea: Cell may have target levels of activity on different time scales

1. Monitor Ca++ entry on:
   a) fast (action potential) time scale
   b) medium (burst) time scale
   c) slow (average voltage) time scale

2. Feed back the error from target onto:
   a) fast (action potential-generating) channels
   b) Medium time-scale generating channels
   c) Burst rate-controlling channels

Extra slides: Models for Robustness of the Integrator

Geometry of Robustness
& Hypotheses for Robustness on Faster Time Scales

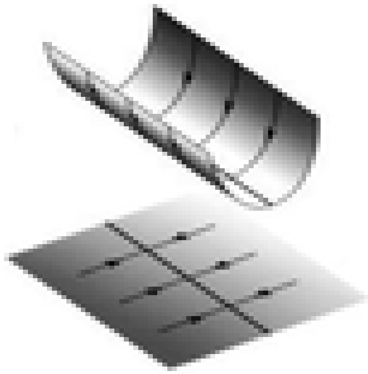


Plasticity on slow time scales:
Reshapes the trough to make it flat

1) Extrinsic Perturbations (in external inputs to system):
- -Only input component along integrating mode persists
- -Strategy: make integrating mode orthogonal to perturbations

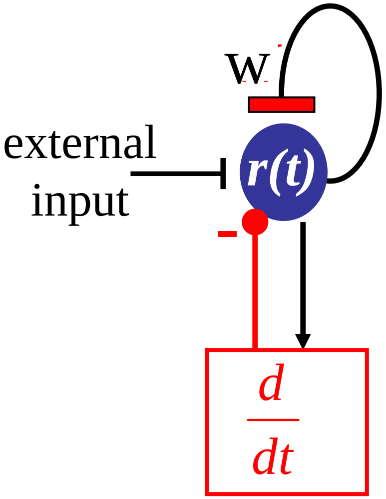
2) Intrinsic Perturbations (in weights):
- -Need eigenvalue of integrating mode = 1:
  - Many different network structures can obey this condition.
  - Goal: find structures that resist common perturbations in weights

Geometry of Robustness



3a) Add “friction” by effectively “putting system in viscous fluid”

-Sliding friction: if a separate circuit monitors integrator slip, it can feed back an opposing input:



$$\tau_{bio}\frac{dr}{dt} = -r + \mathrm{w}r + I - A\frac{dr_{estimate}}{dt}$$

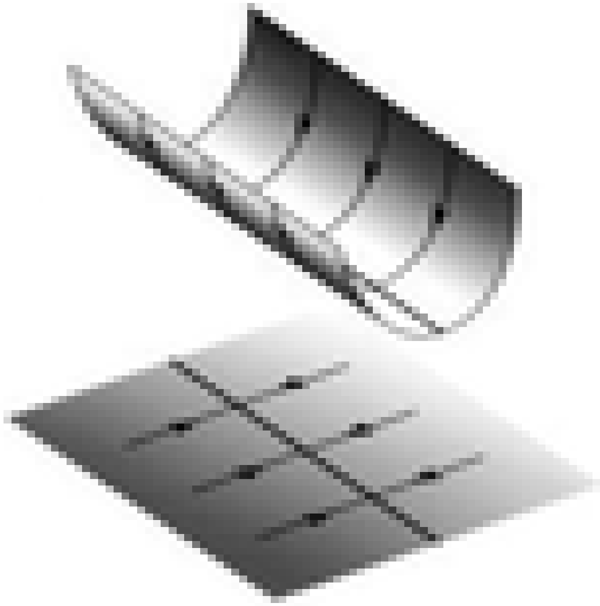
$$(\tau_{bio} + A)\frac{dr}{dt} = -r + \mathrm{w}r + I$$

Control theory: Integral feedback can make perfect adaptation/derivative

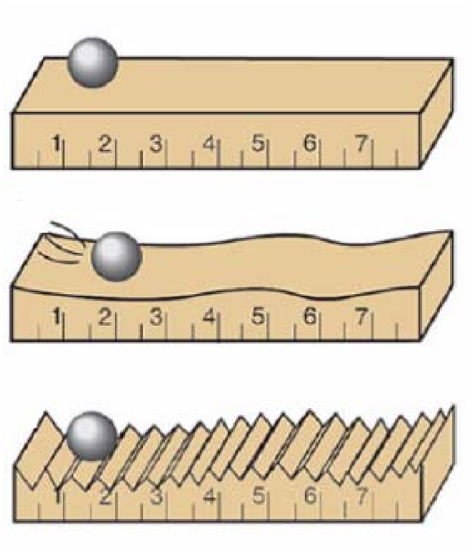
Derivative feedback can make perfect integrator

But….how does one make a perfect derivative???

Geometry of Robustness



Trough of energy function:



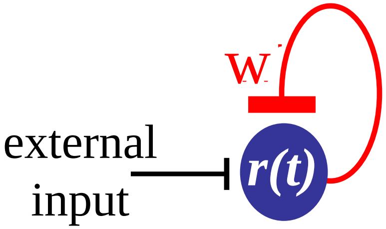
3b) Add “friction” by “roughening” energy surface

Need for fine-tuning in linear feedback models

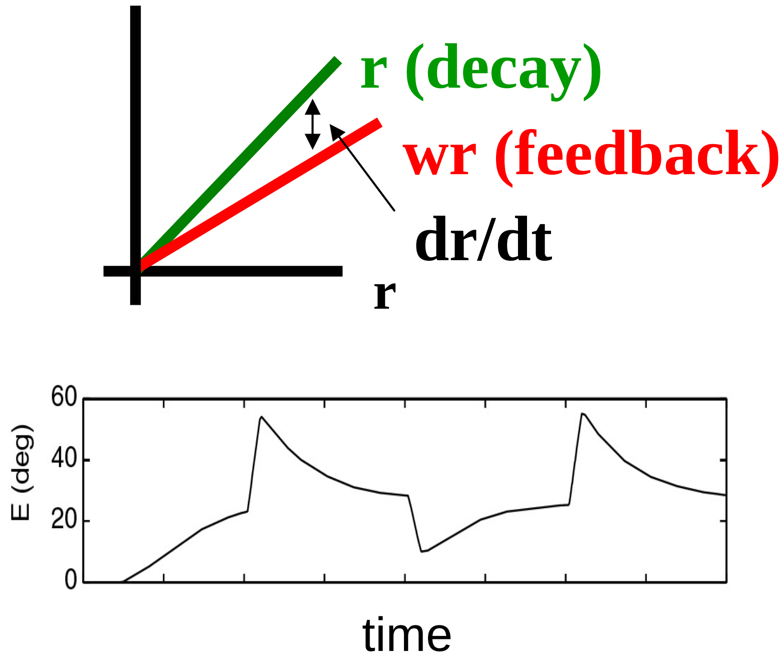
Fine-tuned model:

$$\tau_{neuron}\frac{dr}{dt} = -r + wr + external\ input$$

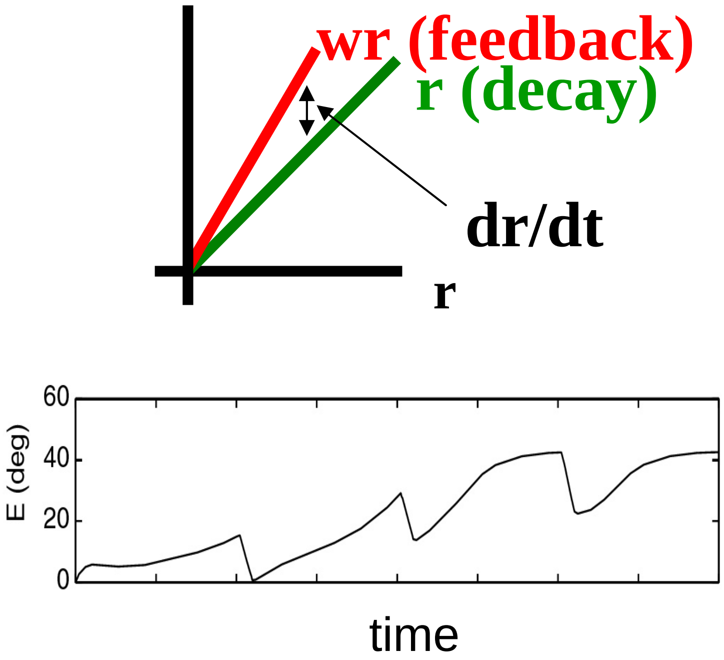
decay feedback



<u>Leaky behavior</u>



<u>Unstable behavior</u>

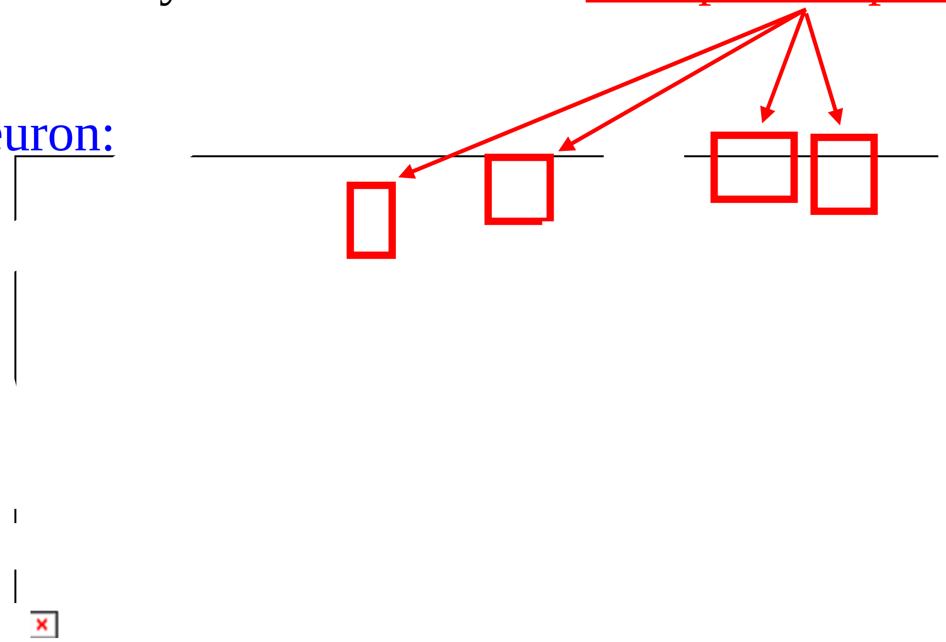


IDEA: Can bistability add robustness to persistent neural activity?

- Bistability in firing rate relationships
  - → Jumps in firing rate **not observed experimentally**

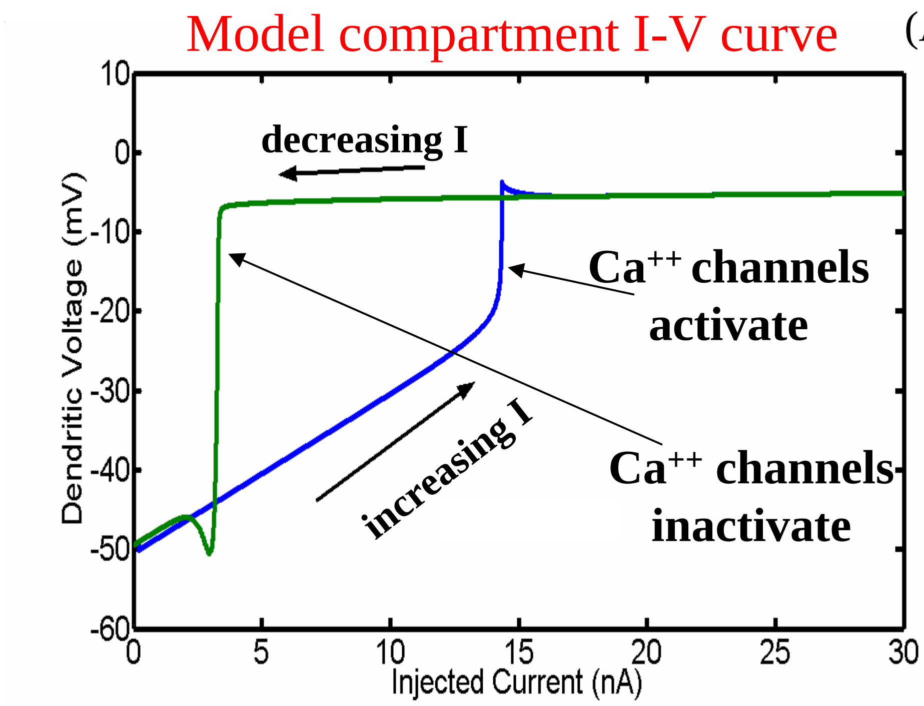
- Dendritic bistability distributed across multiple independent dendrites

Integrator neuron:



Evidence for dendritic bistability & independence

1) Dendritic bistability has been observed experimentally
   - due to the *self-sustaining* properties of NMDA, NaP, or Ca^{++} channels:



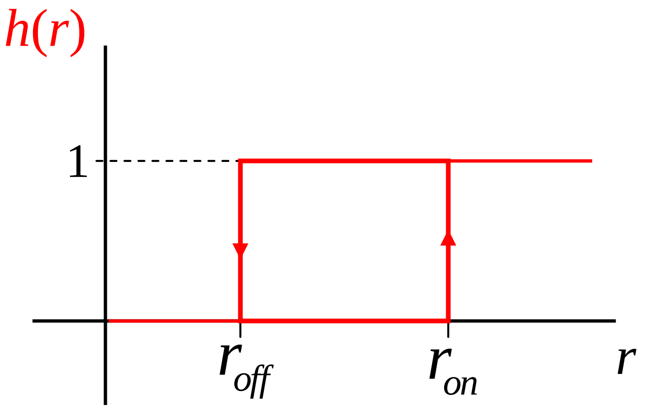
(Adapted from spinal motoneuron model of Booth & Rinzel, 1995)

2) Anatomically realistic models suggest that different dendritic branches may behave approximately independently (Koch et al., 1983; Poirazi et al., 2003)

Simplified analytic model

Dendrite dynamics:

$$\tau_{rec}\frac{dD(r,t)}{dt} = -D(r,t) + h(r)$$



where

r = presynaptic input firing rate

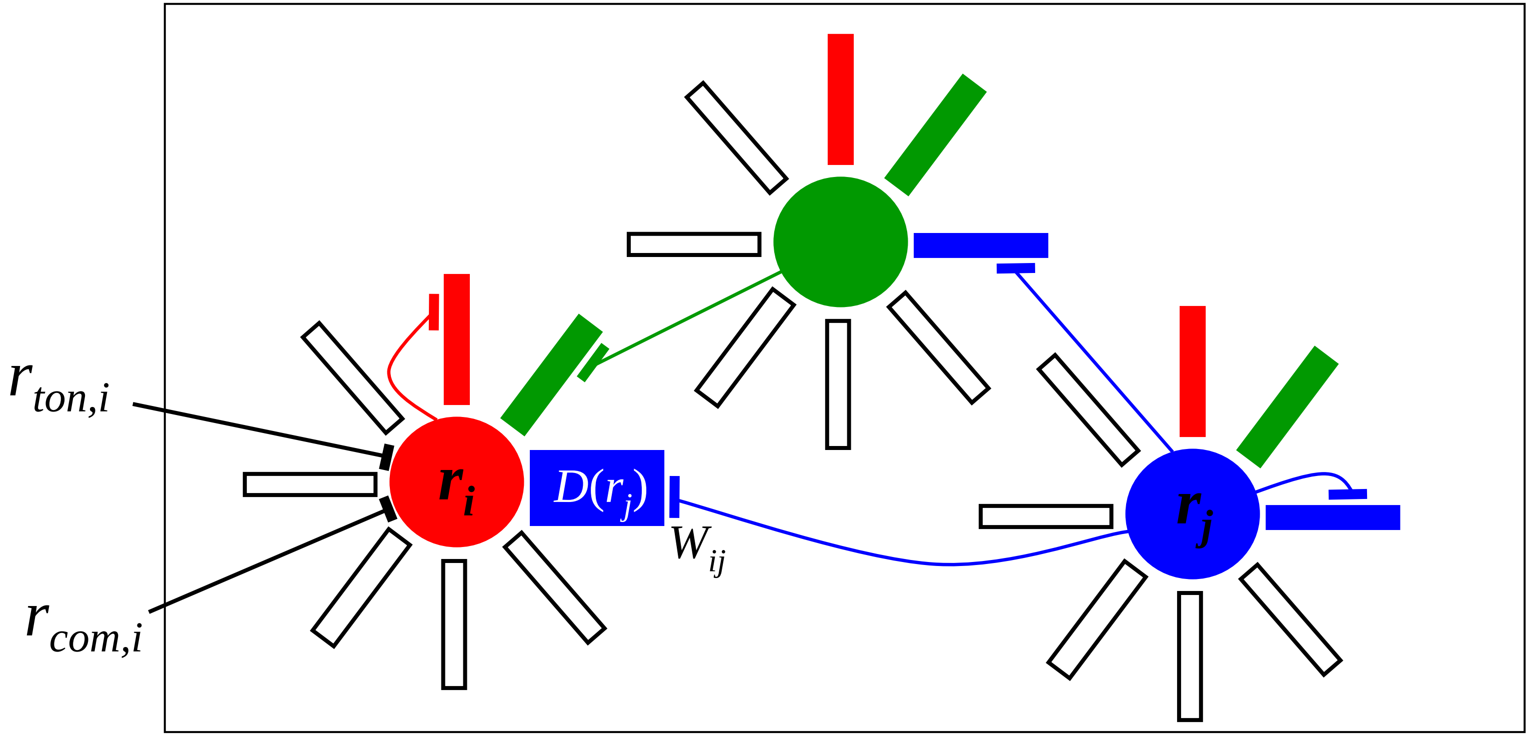
$D(r, t)$ = dendritic compartment activation

$h(r)$ = steady-state dendritic compartment activation

τ_{rec} = time scale for dendrite to reach steady state activation

Network with bistable dendrites

Network of N neurons, each with N identical dendrites:

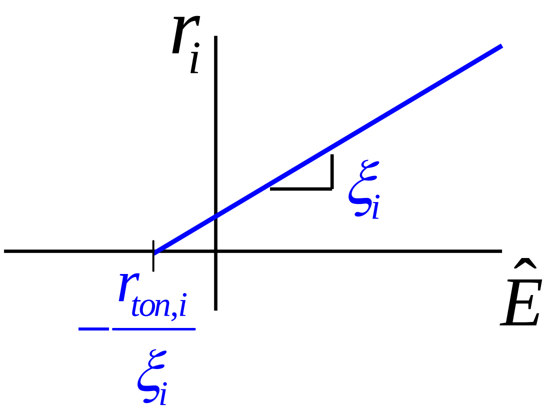


$$r_i(t) = [\underbrace{\sum_{j=1}^{N} W_{ij} D(r_j, t)}_{\text{recurrent input}} + \underbrace{r_{ton,i}}_{\text{background input}} + \underbrace{r_{com,i}}_{\text{eye movement commands}}]_+$$

Result 1: Firing rate is linear in eye position

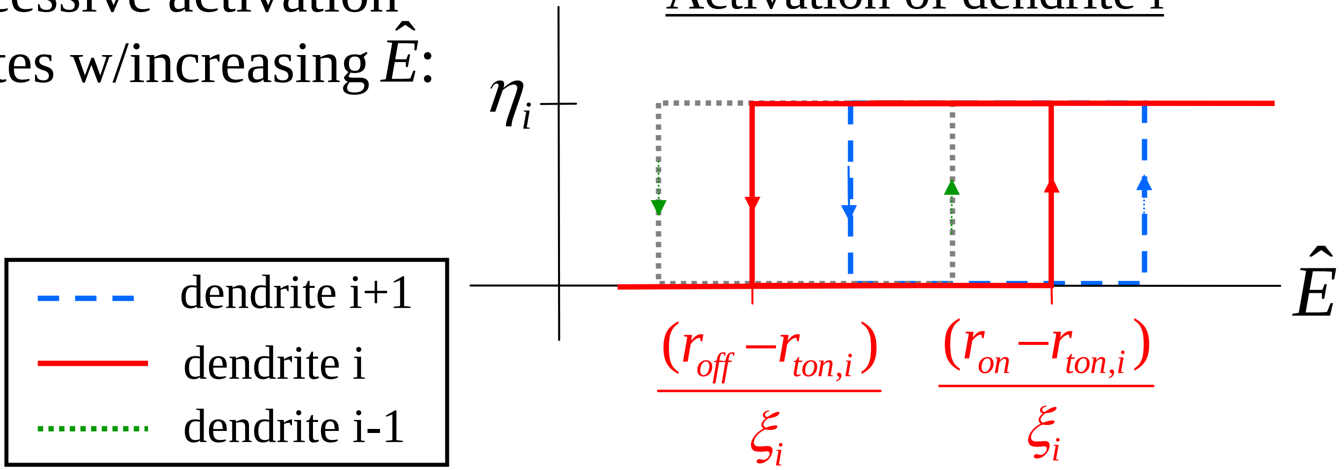
Simplify weight matrix to 1-D (rank 1) form: $W_{ij} = \xi_i \eta_j$

$$r_i(t) = [\xi_i \underbrace{\sum_{j=1}^{N} \eta_j D(r_j, t)}_{\equiv \hat{E}} + r_{ton,i} + r_{com,i}]_+$$



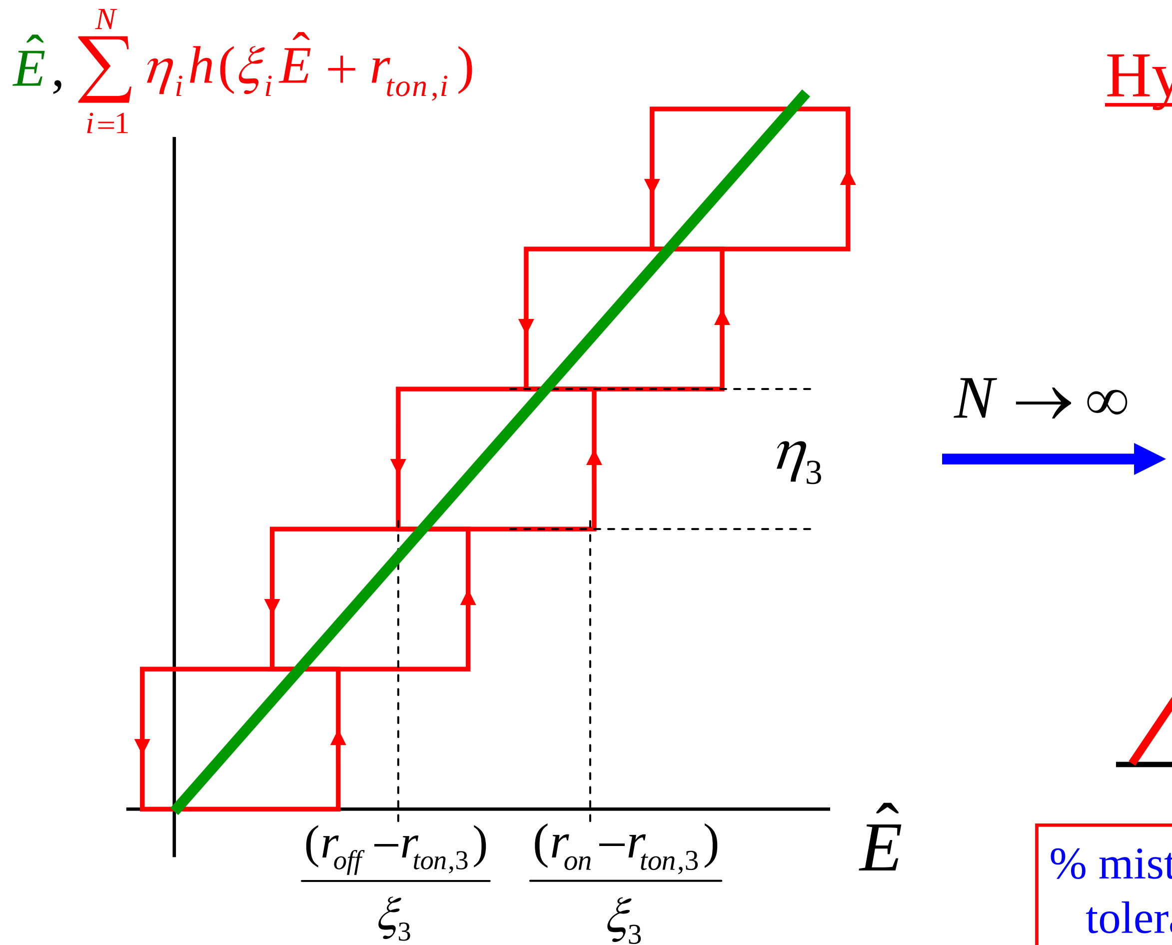
$$\Rightarrow \quad r_i = [\xi_i \hat{E} + r_{ton,i} + r_{com,i}]_+$$

Due to successive activation of dendrites w/increasing $\hat{E}$:

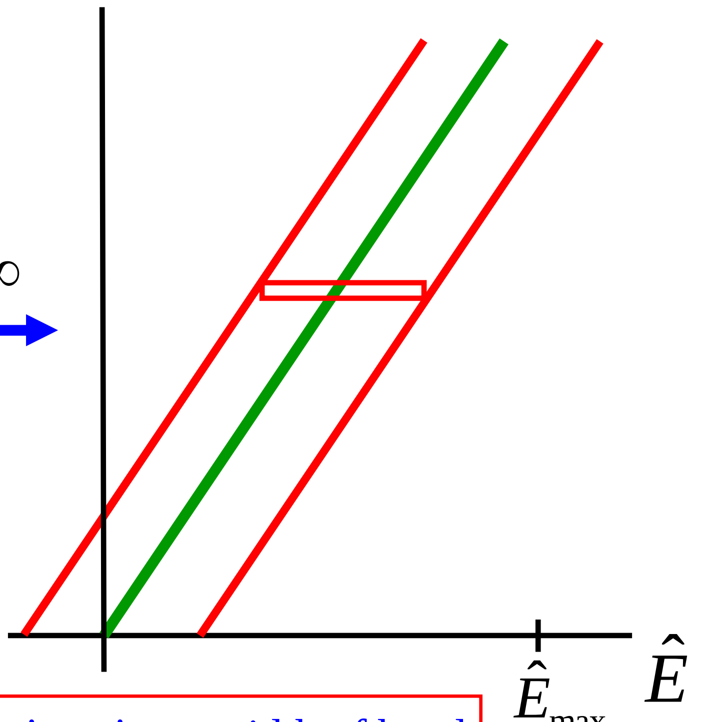


Graphical solution of balanced leak and feedback

No external input: $$\tau_{rec}\frac{d\hat{E}}{dt} = -\hat{E} + \sum_{i=1}^{N}\eta_i h(\xi_i\hat{E} + r_{ton,i})$$



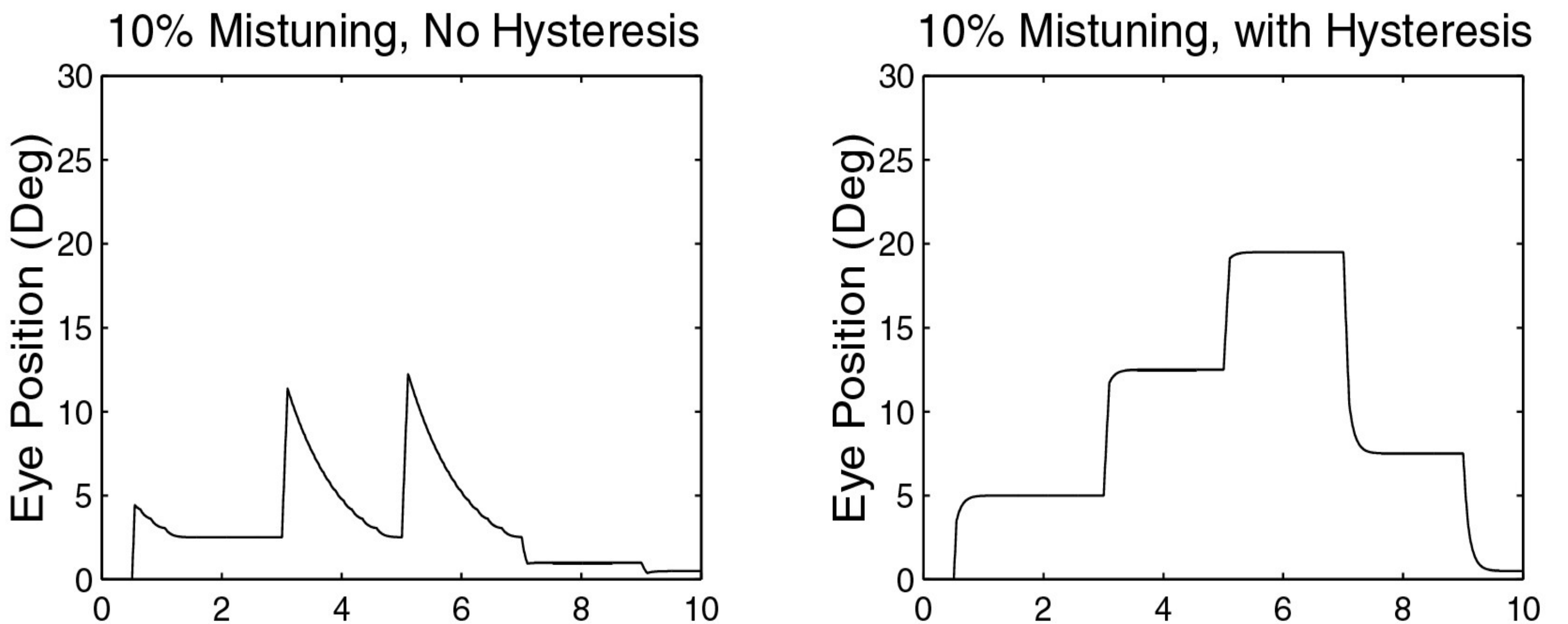
Hysteretic band of stability



% mistuning tolerated $\sim \frac{\text{width of band}}{\hat{E}_{max}}$

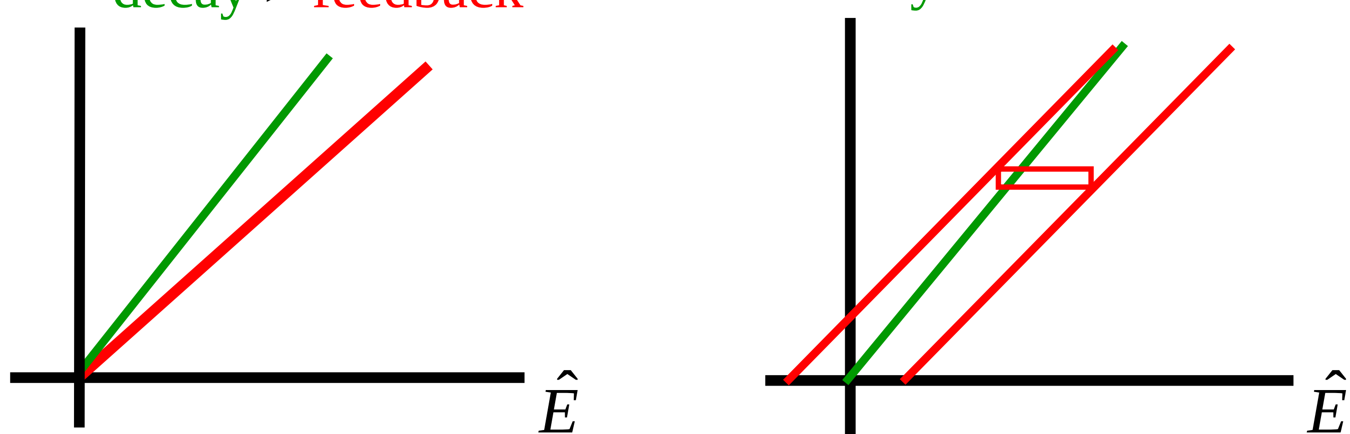
Fixations Are Robust to Mistuning of Feedback

Comparison with no-hysteresis models:



decay > feedback

decay within feedback band



Biophysical Model

Network of 20 recurrently connected neurons, each with:

(Simulations by Joseph Levine)

- Spiking model soma
- Calcium plateau mediated bistability
- Soma and dendrites ohmically coupled