

# **CN510: Principles and Methods of Cognitive and Neural Modeling**

## **Derivation of Shunting Network to Solve Noise-Saturation Dilemma**

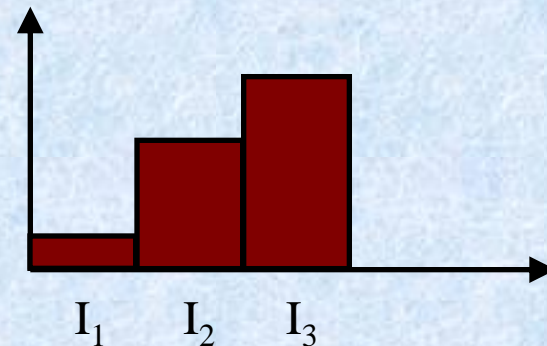
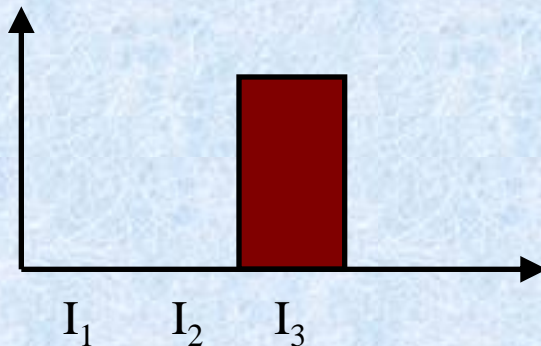
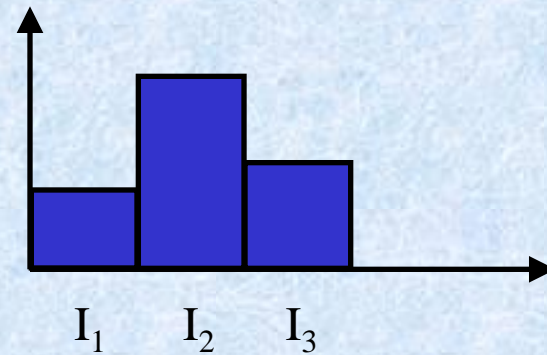
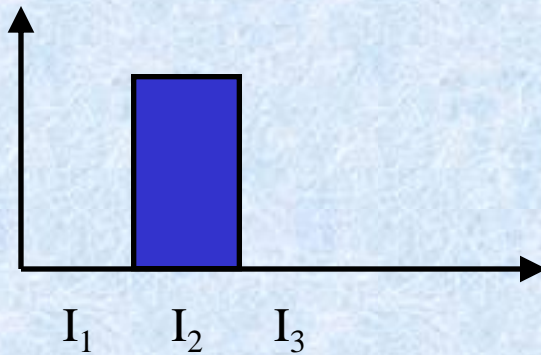
### **Lecture 7**

Instructor: Anatoli Gorchetchnikov <[anatoli@bu.edu](mailto:anatoli@bu.edu)>

# What are the Inputs?

In a simple setup one input channel  $I_i$  signals a stimulus

In real setting usually a combination of input channels (a pattern) signals every stimulus

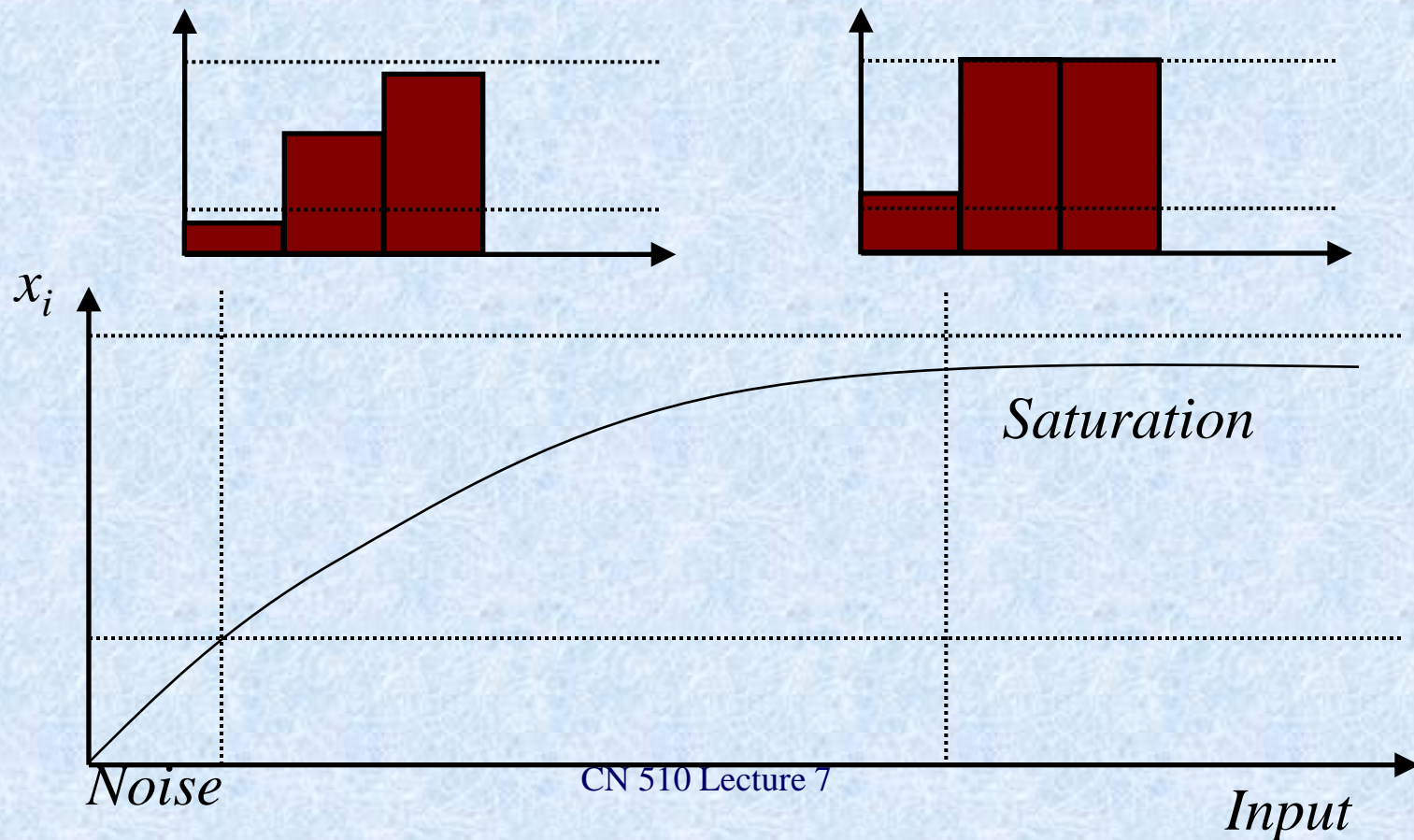


We would like to take into account every component of an input pattern

# Noise-Saturation Dilemma

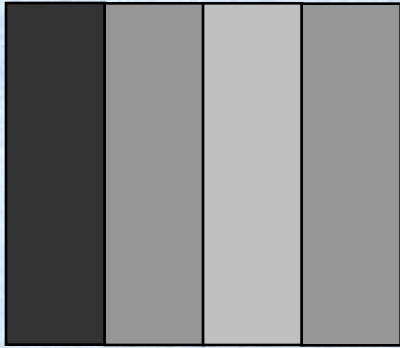
If the component of an input pattern is too small it can be lost in random fluctuations of noise

But if we simply amplify all input, then the stronger components will saturate and become indistinguishable

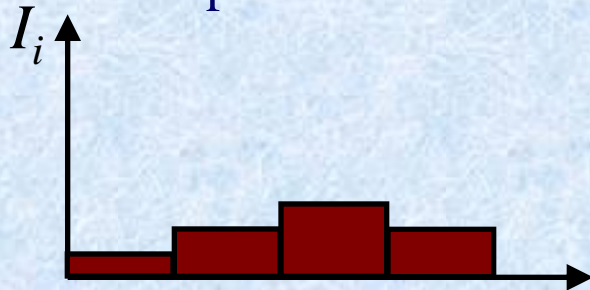


# Example: Vision

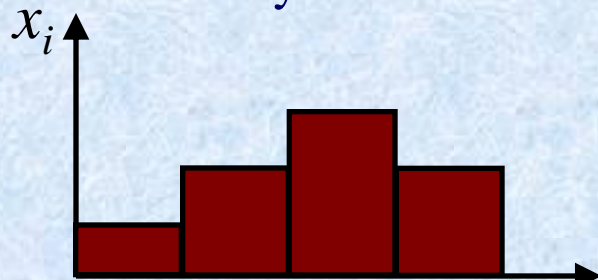
Dim light



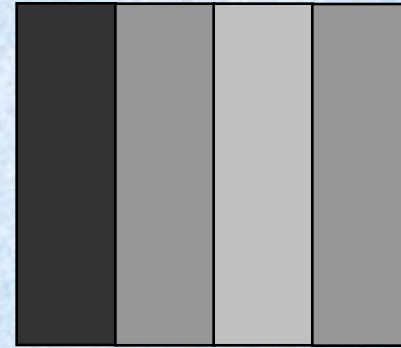
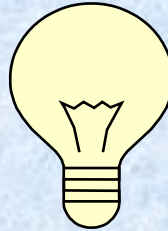
Luminance profile:



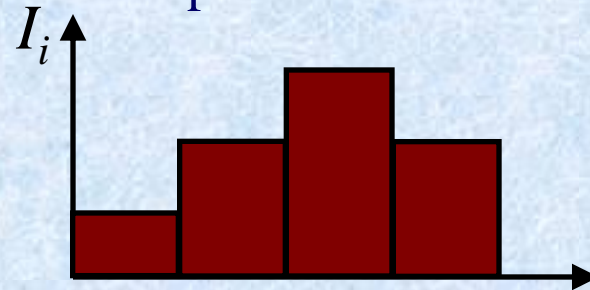
Desired activity:



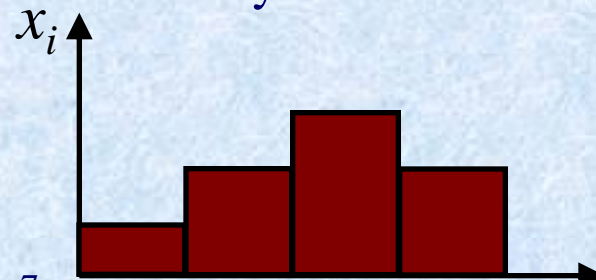
Bright light



Luminance profile:



Desired activity:



# Example: Vision

We want to discount the intensity of the light source (energy) and register the relative luminances of different patches in the visual field (pattern)

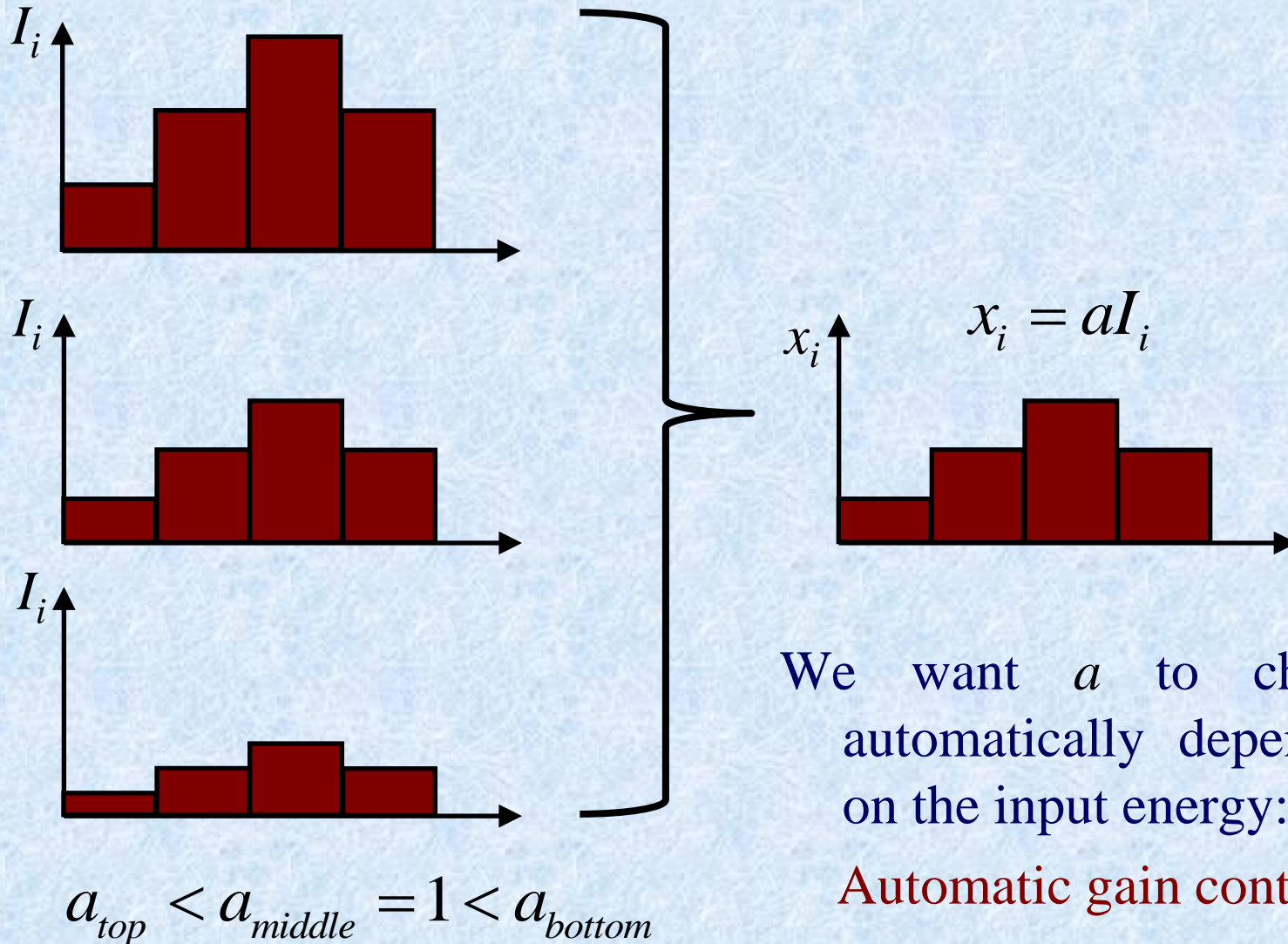
This will allow us to recognize the same object independently of lighting conditions

In general:

Often we will want the activity pattern across a set of cells to represent the pattern of an input rather than its total energy

**Factorization of pattern and energy**

# Factorization of Pattern and Energy



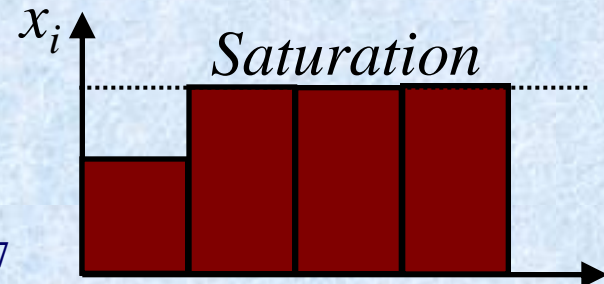
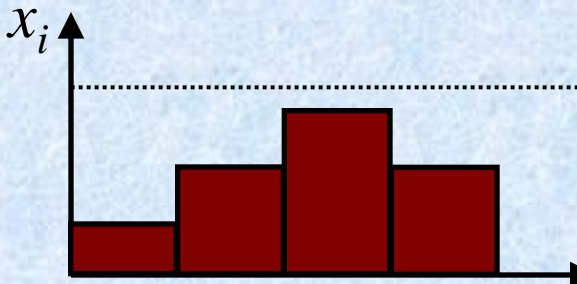
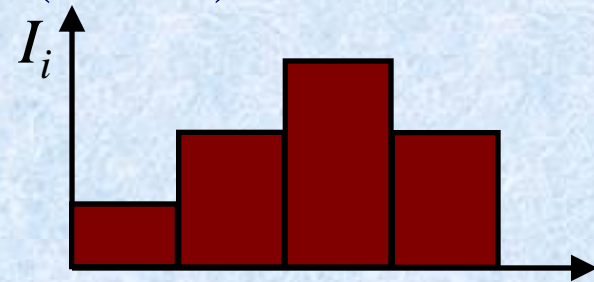
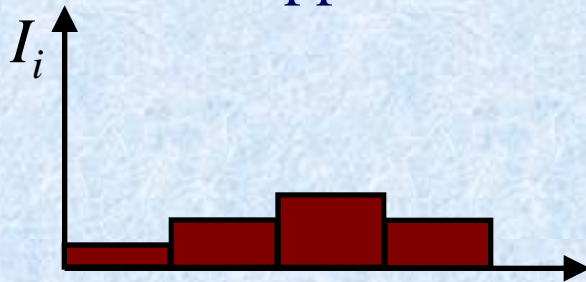
# Let's Try Leaky Integrator

$$\frac{dx_i}{dt} = -Ax_i + I_i$$

With constant inputs at equilibrium we have  $x_i = \frac{I_i}{A}$

This is perfect since  $x_i \sim I_i$ , but unrealistic since if  $I_i \rightarrow \infty$  then  $x_i \rightarrow \infty$

If we add an upper bound for  $x_i$ , then ( $A=0.5$ )



# Pattern Variables

How can we have  $x_i \sim I_i$  but at the same time not be dependent on input's energy?

We can normalize the input pattern

Let's define a pattern variable  $J_i = \frac{I_i}{\sum_j I_j}$

Note that by definition  $0 \leq J_i \leq 1$ , but  $I_i/I_j = J_i/J_j$

Thus we have a limited range input dependent variable that preserves the input pattern

To achieve factorization of pattern and energy we need to ensure  $x_i \sim J_i$

To solve noise-saturation dilemma we must ensure that the maximal value of  $x$  (let's call it  $B$ ) is no less than  $\max(J_i)$  which is always true if  $1 \leq B$



# Reverse Engineering

Let's adapt leaky integrator so that instead of  $x_i = \frac{I_i}{A}$  at

equilibrium it will have something like  $x_i \sim \frac{I_i}{\sum_j I_j}$

Attempt 1:  $x_i = \frac{I_i}{A \sum_j I_j} \rightarrow Ax_i \sum_j I_j = I_i$

$$\frac{dx_i}{dt} = -Ax_i \sum_j I_j + I_i$$

Issues:

- For a really small input  $I_i$  activation  $x_i$  will be even smaller
- No passive decay

# Reverse Engineering

$$\text{Attempt 2: } x_i = \frac{I_i}{A + \sum_j I_j} \rightarrow Ax_i + x_i \sum_j I_j = I_i$$

$$\frac{dx_i}{dt} = -Ax_i + I_i - x_i \sum_j I_j$$

Has passive decay

For small inputs ( $\sum_j I_j \ll A$ )  $x_i \sim I_i$

For large inputs ( $\sum_j I_j \gg A$ )  $x_i \sim J_i$

Issue:

- $I_i$  is excitatory in the second term and inhibitory in the third term

# Reverse Engineered Shunting Equation

$$\frac{dx_i}{dt} = -Ax_i + I_i - x_i \sum_j I_j$$

Lets pull out  $I_i$  from the sum

$$\frac{dx_i}{dt} = -Ax_i + I_i - x_i I_i - x_i \sum_{j \neq i} I_j$$

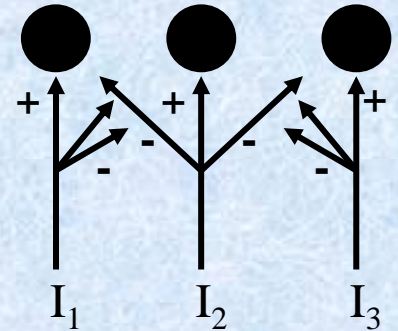
And recombine the terms

$$\frac{dx_i}{dt} = -Ax_i + (1 - x_i)I_i - x_i \sum_{j \neq i} I_j$$

Here we got the simplest version of a shunting equation

# Reverse Engineered Shunting Network

$$\frac{dx_i}{dt} = -Ax_i + (1 - x_i)I_i - x_i \sum_{j \neq i} I_j$$



Passive decay that brings neurons to rest when there is no input

Activation  $x_i$  is bound from above by  $1$ , the rate of its growth is controlled by  $I_i$

All other inputs  $I_j$  provide inhibition, their effect is controlled by activation  $x_i$

Activation  $x_i$  is bound from below by  $0$  thus the neuron cannot be inhibited below the resting potential

The maximal value of  $x$  is no less than  $\max(J_i)$  (both are  $1$ )

# Original Derivation

Let's assume a neuron has fixed number  $B$  of binary “sites” that can be turned on or off

Let's further assume that neural activity  $x_i$  is the number of sites that are **on**

Then cell  $i$  has  $x_i$  excited (**on**) sites and  $(B-x_i)$  not excited (**off**) sites

The probability that at any given moment an excited site will turn off is constant, let's call it  $A$

Then the rate of turning these sites off is  $Ax_i$

The probability of unexcited site turning on is proportional to input  $I_i$

Then the rate of turning these sites on is  $(B-x_i)I_i$

# Original Derivation

Combining both terms into differential equation gives

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i$$

Solving at equilibrium:

$$Ax_i + x_i I_i = BI_i$$

$$x_i = \frac{BI_i}{A + I_i}$$

Not good, since if  $I_i \gg A$  then  $x_i \rightarrow B$  – still saturates

How can we avoid saturation?

Let's make  $A$  very large

# Original Derivation

$$x_i = \frac{BI_i}{A + I_i}$$

If  $A$  is large, then for small inputs  $I_i$  the activity  $x_i \rightarrow 0$  or lost in noise

Hence the term noise-saturation dilemma

Another solution to saturation is to rescale  $A$  depending on input, for example on total input  $I = \sum_j I_j$

$$\frac{dx_i}{dt} = -(A + f(I))x_i + (B - x_i)I_i$$

# Original Derivation

$$\frac{dx_i}{dt} = -(A + f(I))x_i + (B - x_i)I_i$$

Solving at equilibrium for simplest  $f(I)=I$ :

$$Ax_i + Ix_i + x_i I_i = BI_i$$

$$x_i = \frac{BI_i}{A + I + I_i}$$

In this case  $x_i$  does not saturate

Is this solution good enough?

Let's do some case studies



## Case Studies

Let's assume  $I_i$  is large ( $I_i \gg A$ ) and all other  $I_j$  are small, so that  $I \approx I_i$

Then

$$x_i = \frac{BI_i}{A + I + I_i} = \frac{B}{2}$$

Thus we are wasting half of our dynamic range

Now let's assume  $I_i=80$ ,  $I_j=20$ ,  $A=1$

$$x_i = \frac{BI_i}{A + I + I_i} = \frac{80B}{1 + 80 + 100} = 0.442B$$
$$x_j = \frac{BI_j}{A + I + I_j} = \frac{20B}{1 + 20 + 100} = 0.165B$$

The relative magnitude of inputs is not preserved

## What's the Problem?

Remember in the beginning we looked at  $x_i = \frac{I_i}{A + \sum_j I_j}$  and it seemed reasonable end goal

Now we are looking at  $x_i = \frac{BI_i}{A + I + I_i}$  where  $I = \sum_j I_j$  and it does not work

The main difference is extra  $I_i$  in denominator

How we can get rid of this?

Let's set our  $f(I) = \sum_{j \neq i} I_j$

# Original Derivation

$$\frac{dx_i}{dt} = -(A + f(I))x_i + (B - x_i)I_i$$

Solving at equilibrium for  $f(I) = \sum_{j \neq i} I_j$  :

$$Ax_i + x_i \sum_{j \neq i} I_j + x_i I_i = BI_i$$

$$Ax_i + x_i (\sum_{j \neq i} I_j + I_i) = BI_i$$

$$x_i = \frac{BI_i}{A + I}$$

In this case  $x_i$  does not saturate, if  $I \gg A$  then  $x_i = BI_i/I = BJ_i$

Thus we achieved factorization of pattern and energy

## Back to Case Studies

Let's assume  $I_i$  is large ( $I_i \gg A$ ) and all other  $I_j$  are small, so that  $I \approx I_i$

Then

$$x_i = \frac{BI_i}{A+I} = B$$

Thus we are using our full dynamic range

Now let's assume  $I_i=80$ ,  $I_j=20$ ,  $A=1$

$$x_i = \frac{BI_i}{A+I} = \frac{80B}{1+100} = 0.792B$$

$$x_j = \frac{BI_j}{A+I} = \frac{20B}{1+100} = 0.198B$$

The relative magnitude of inputs is preserved

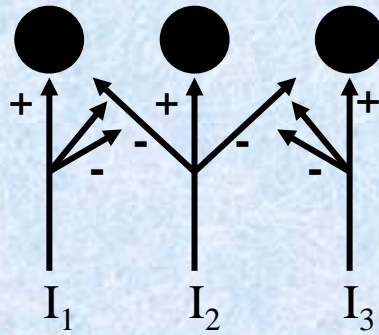
Equation 
$$\frac{dx_i}{dt} = -(A + f(I))x_i + (B - x_i)I_i$$

where  $f(I) = \sum_{j \neq i} I_j$  we can rewrite as

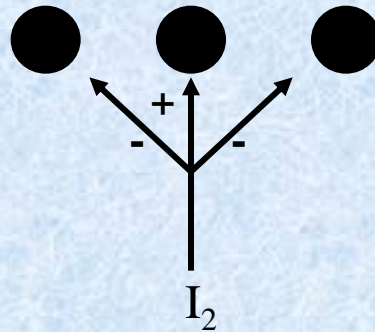
$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - x_i \sum_{j \neq i} I_j$$

and interpret from neural perspective the last term as shunting inhibition from competing input channels

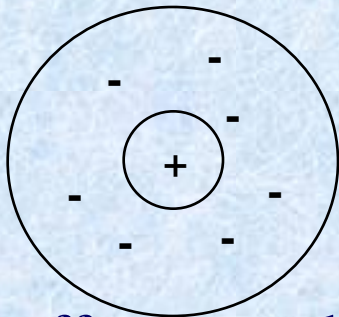
Thus we can claim that if we want to achieve factorization of pattern and energy we have to have a feedforward competitive shunting network



Consider it from the perspective of single input:



Or in 2D:



On-center off-surround competitive network

# Automatic Gain Control

Given 
$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - x_i \sum_{j \neq i} I_j$$

we can ask how much activity cell  $i$  gains with increase of  $I_i$

Our equation ensures that this gain is **automatically controlled** by what is happening everywhere else in the network

$$x_i = \frac{BI_i}{A + I}$$

Gain in  $x_i$  with increase of  $I_i$  is automatically scaled by equal increase of the total input  $I$

# Conservation of Total Activation

Let's sum the activities

$$x_i = \frac{BI_i}{A + I}$$

$$\sum_i x_i = \frac{B \sum_i I_i}{A + I} = \frac{BI}{A + I}$$

So when  $I \gg A$

$$\sum_i x_i = B$$

Thus with large inputs total activity in the network approaches the maximal activation for a single cell

In other words the total activity of the network is conserved or normalized



Later in the course we will see that recurrent competitive field (RCF) network also normalizes the activity, but it does it even for small inputs

This might be a problem if small inputs are just noise

$$x_i = \frac{BI_i}{A + I}$$

Here for small  $I_i$

We will get  $x_i \rightarrow 0$  if  $I$  is large and  $x_i \rightarrow BI_i/A$  if  $I$  is also small

Which means that if we have lots of good inputs we can disregard small ones as noise, otherwise we can keep them (if decay rate  $A$  is low)

# Morphing Conductance-based Model into Shunting

$$C_m \frac{dV}{dt} = -g_l V_m + \overset{\text{excitatory input}}{g_e (E_e - V_m)} + \overset{\text{inhibitory input}}{g_i (E_i - V_m)}$$

Divide both sides by capacitance

$$\frac{dV}{dt} = -\frac{g_l}{C_m} V_m + \frac{g_e}{C_m} (E_e - V_m) + \frac{g_i}{C_m} (E_i - V_m)$$

Rename the variable and parameters

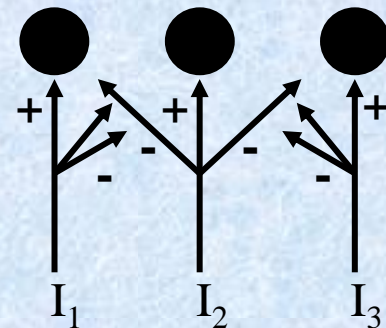
$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_E + (C_N - x_i)I_I$$

Make all parameters non-negative and adjust signs

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_E - (C + x_i)I_I$$

Set the inputs accordingly to network architecture

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$



# Similarities

$$C_m \frac{dV}{dt} = -g_l V_m + g_{Na} (E_{Na} - V_m) + g_K (E_K - V_m)$$

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$

Parameters  $B$  and  $C$  are similar to reverse potentials for excitatory and inhibitory channels, respectively

If  $C=0$  we can assume that inhibition is provided by Cl channels with Nernst potential equal to resting potential

Cell activity  $x$  is similar to membrane potential

Leakage parameter  $A$  is directly related to membrane time constant  $1/RC$

# Differences

$$C_m \frac{dV}{dt} = -g_l V_m + g_{Na} (E_{Na} - V_m) + g_K (E_K - V_m)$$

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$

Voltage-gated properties are not included

As a result  $x$  can achieve steady state anywhere between  $-C$  and  $B$ , while in real neuron potentials above spiking threshold will cause firing of an action potential, not a steady state

But there is a relationship between firing rate of a neuron and the strength of excitatory input

## Discussion Question

How can we manipulate  $x$  to achieve better correspondence between output of Grossbergian neuron and Hodgkin-Huxley neuron?

Threshold output function  $f(x)$  will take care of converting  $x$  into firing rate

Problem is the relationship between net input to neuron and its spiking frequency is complex and depends on many factors

Thus output functions might (and do) differ in different models

# Comparison Between Two Networks

From HH equation we derived

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$

From noise-saturation dilemma we derived

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - x_i \sum_{j \neq i} I_j$$

Does the equation from HH have the same properties as the one derived from noise-saturation dilemma?

$$Ax_i + x_i \sum_{j \neq i} I_j + x_i I_i = BI_i - C \sum_{j \neq i} I_j$$

$$Ax_i + x_i I = BI_i + CI_i - CI$$

$$x_i = \frac{(B+C)I_i - CI}{A+I} = \frac{(B+C)I_i}{A+I} - \frac{CI}{A+I}$$

If  $I \ll A$  then  $x_i = \frac{(B+C)I_i}{A}$  so it has similar properties as before

If  $I \gg A$  then  $x_i = \frac{(B+C)I_i}{I} - C$ . Here

- If  $I_i$  is small  $x_i \rightarrow -C$ ,
- If  $I_i$  is large ( $I \approx I_i$ ) then  $x_i \rightarrow B$

So when the total input to the system is large we use full dynamic range between  $-C$  and  $B$  for activation

- Smallest inputs are pushed down to  $-C$
- Largest inputs are closer to  $B$

When the total input is small even the smallest inputs are around zero and thus have more effect than smallest inputs in the first case

Does this network have any other additional properties?



# Adaptation Level of the Network

Rearranging 
$$x_i = \frac{(B + C)I_i}{A + I} - \frac{CI}{A + I} = \frac{(B + C)I}{A + I} \left( \frac{I_i}{I} - \frac{C}{B + C} \right)$$

This implies that  $x_i$  will be less than 0 if  $\frac{I_i}{I} < \frac{C}{B + C}$

Thus all inputs that are between 0 and  $\frac{C}{B + C}$  can be treated as noise and suppressed

The constant  $\frac{C}{B + C}$  is called adaptation level of the network

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$

For biological  $B$  and  $C$  adaptation level is about  $20mV/140mV$   
or  $1/7$

# Noise Suppression of Uniform Inputs

If we let  $C$  be number of cells excited by  $I_i$  ( $I$ ) and  $B$  be the number of cells inhibited by  $I_i$  ( $N-I$ ) where  $N$  is the number of cells in the network then

$$\frac{C}{B+C} = \frac{1}{N-1+1} = \frac{1}{N}$$

$$x_i = \frac{NI}{A+I} \left( \frac{I_i}{I} - \frac{1}{N} \right)$$

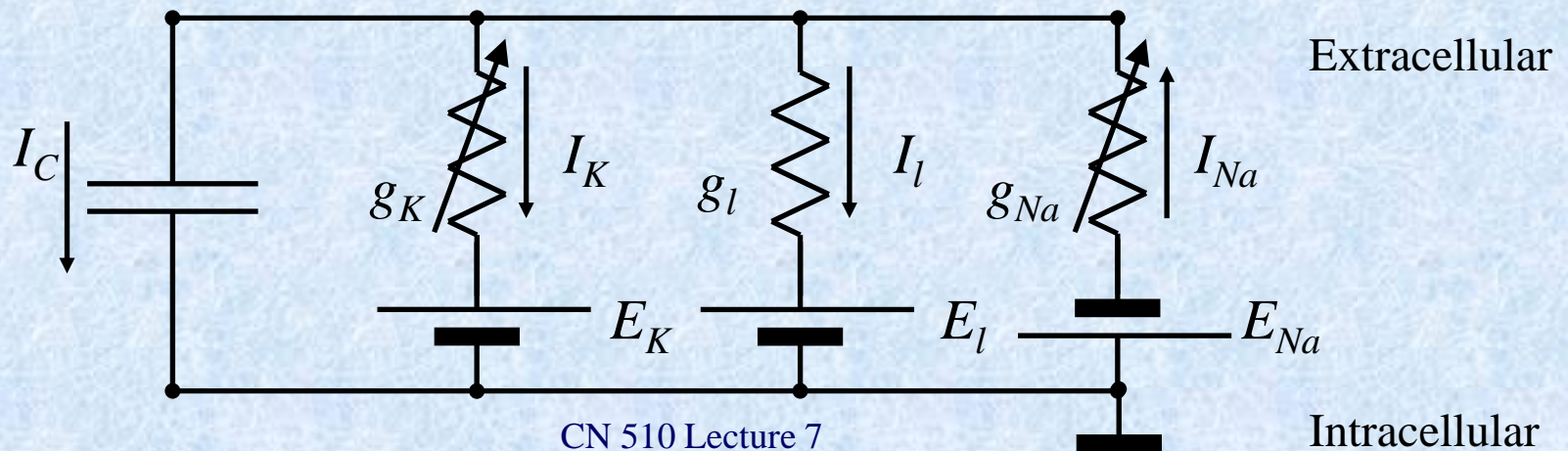
And if all inputs are uniform ( $I_i=I/N$ ) then all  $x_i=0$

# Finally, What Is Shunting?

To shunt is to divert, but in electronics a shunt is

- A conductor joining two points in an electrical circuit so as to form a parallel or alternative path through which a portion of the current may pass (as for regulating the amount passing in the main circuit)

In HH circuit we can consider any of the currents as a shunt, but the term shunting network usually refers to the inhibitory projections implemented as a shunt



# Another Way to Look at It

Inputs do not affect the currents directly

Current depends upon the gradient of concentration and the difference of potentials

Input only gates the current, or opens up shunts whose effect depends on the current state of the neuron

For example, additive inhibition would always have inhibitory effect, while inhibition based on shunt will drive the voltage towards reverse potential of respective ion and can be excitatory if membrane potential is below the Nernst potential

# Additive Equation

Solving

$$\frac{dx_i}{dt} = -Ax_i + BI_i - \sum_{j \neq i} I_j$$

At equilibrium yields

$$Ax_i = BI_i + I_i - I$$
$$x_i = \frac{(B+1)I_i - I}{A}$$

All inputs smaller than  $I/(B+1)$  will be suppressed

Let's assume  $I_i$  is large and all other  $I_j$  are small, so that  $I \approx I_i$

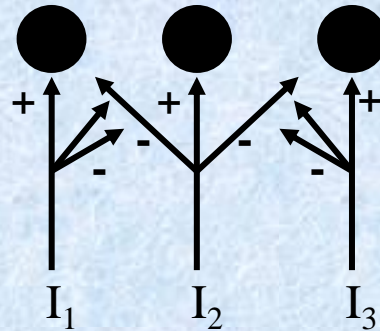
Then

$$x_i = \frac{BI}{A}$$

which obviously saturates if we put upper limit on  $x_i$

# Summary to This Point

We are investigating feedforward competitive networks



We derived a simple feedforward additive network as

$$\frac{dx_i}{dt} = -Ax_i + BI_i - \sum_{j \neq i} I_j$$

This is simple, but shown to have problems with correct pattern representation and noise-saturation dilemma

## Summary to This Point

Attempting to solve noise-saturation dilemma we derived a simple feedforward shunting network as

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - x_i \sum_{j \neq i} I_j$$

This network is non-linear, and was shown to solve noise-saturation dilemma as well as problem of factorization of pattern and energy

We then extended our derivation to feedforward shunting network with hyperpolarization term

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i) \sum_{j \neq i} I_j$$

# Summary to This Point

Network

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)I_i - (C + x_i)\sum_{j \neq i} I_j$$

was also derived from conductance-based model

This network has an additional property of an adaptation level, which suppresses uniform inputs



# Next Time

We will look at some additional properties of shunting networks

- Shift property as seen in retinal cells
- Weber's law of difference detection thresholds

Readings:

- Werblin, F.S. (1971). Adaptation in a vertebrate retina: Intracellular recording in *Necturus*. *Journal of Neurophysiology*, **34**, pp. 228-241.