The BCM theory of synaptic plasticity.

<u>The BCM theory of cortical plasticity</u>

BCM stands for Bienestock Cooper and Munro, it dates back to 1982. It was designed in order to account for experiments which demonstrated that the development of orientation selective cells depends on rearing in a patterned environment.

BCM Theory

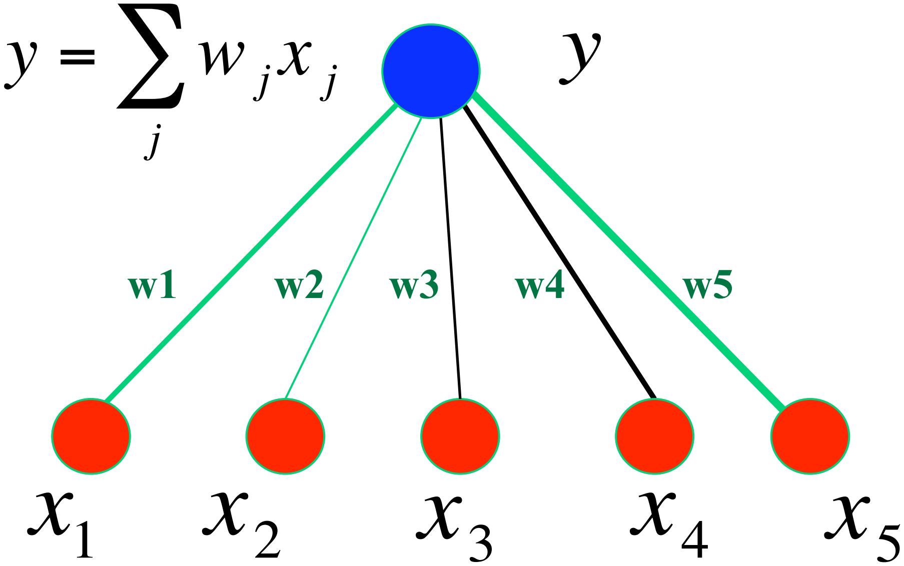
(Bienenstock, Cooper, Munro 1982; Intrator, Cooper 1992)

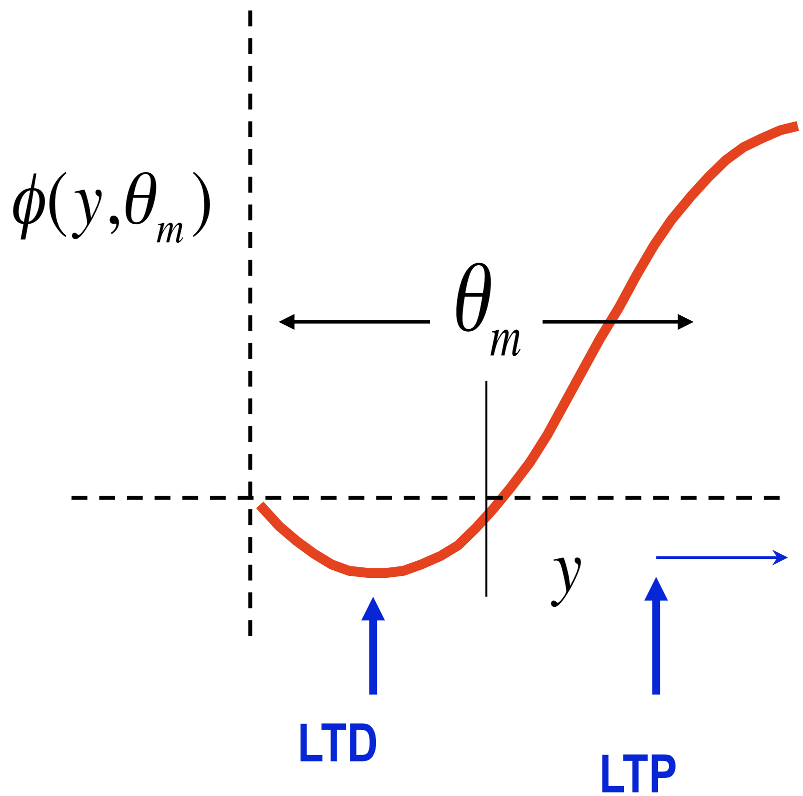
1) The learning rule:

$$\frac{dw_j}{dt} = \eta x_j \phi(y,\theta_m)$$

For simplicity – linear neuron

$$y = \sum_j w_j x_j$$



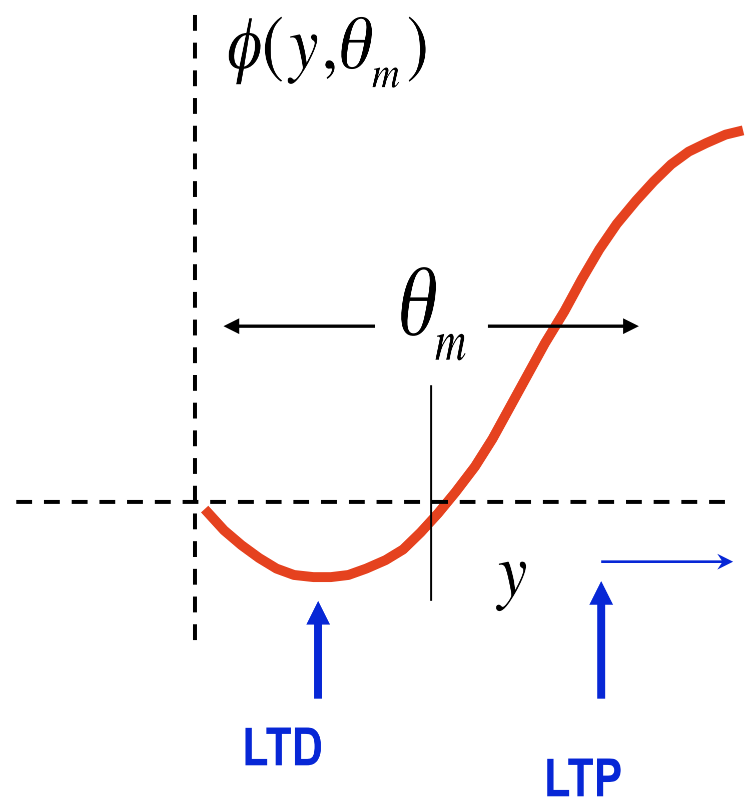


BCM Theory

1) $$\frac{dw_j}{dt} = \eta x_j \phi(y,\theta_m)$$

2) The sliding threshold

$$\theta_m \propto E\left[y^2\right] = \frac{1}{\tau}\int_{-\infty}^{t} y^2(t')e^{-(t-t')/\tau}dt'$$



Requires

- Bidirectional synaptic modification LTP/LTD
- Sliding modification threshold
- The fixed points depend on the environment, and in a patterned environment only selective fixed points are stable.

The integral form of the average:

Is equivalent to this differential form:

$$\theta_m \propto E\left[y^2\right] = \frac{1}{\tau}\int_{-\infty}^{t} y^2(t')e^{-(t-t')/\tau}dt'$$

$$\frac{d\theta_m}{dt} = \frac{1}{\tau}(y^2 - \theta_m)$$

Note, it is essential that θ_m is a superlinear function of the history of C, that is:

$$\frac{d\theta_m}{dt} = \frac{1}{\tau}(y^p - \theta_m) \qquad \text{with } p>1$$

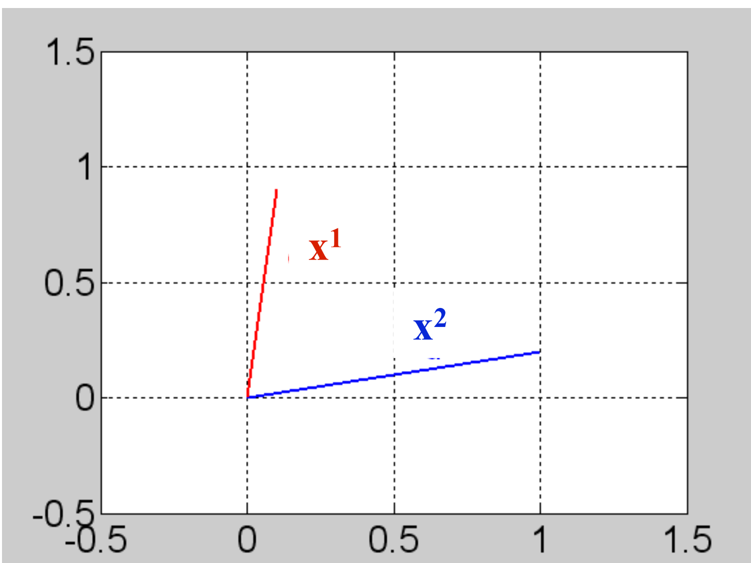
Note also that in the original BCM formulation (1982) $\theta_m \propto E\left[y\right]^2$ rather then $\theta_m \propto E\left[y^2\right]$

What is the outcome of the BCM theory?

Assume a neuron with N inputs (N synapses), and an environment composed of N different input vectors.

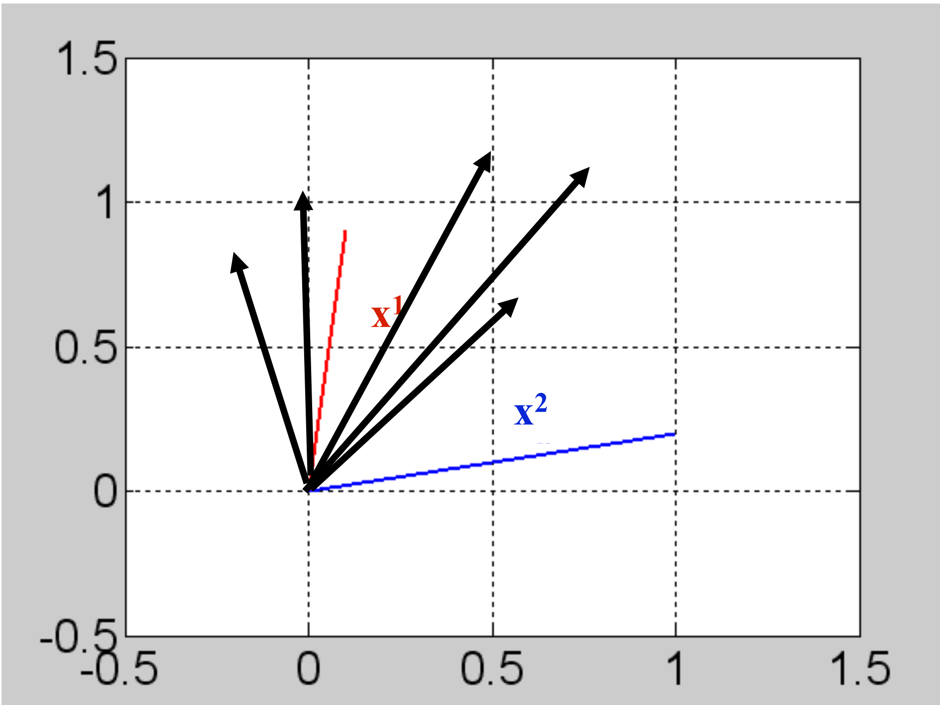
An N=2 example:

$$\mathbf{x}^1 = \begin{pmatrix} 1.0 \\ 0.2 \end{pmatrix} \quad \mathbf{x}^2 = \begin{pmatrix} 0.1 \\ 0.9 \end{pmatrix}$$



What are the stable fixed points of **W** in this case?

(Notation: $\mathbf{y}^{\mathbf{i}} = \mathbf{w}^T \cdot \mathbf{x}^{\mathbf{i}}$)



Note:
Every time a new input is presented, **W** changes, and so does θ_m

What are the fixed points? What are the stable fixed points?

(Show matlab)

Two examples with N= 5

Note: The stable FP is such that for one pattern $y^i = w^T x^i = \theta_m$ while for the others $y^{(i \neq j)} = 0$.

Show movie

(note: here c=y)

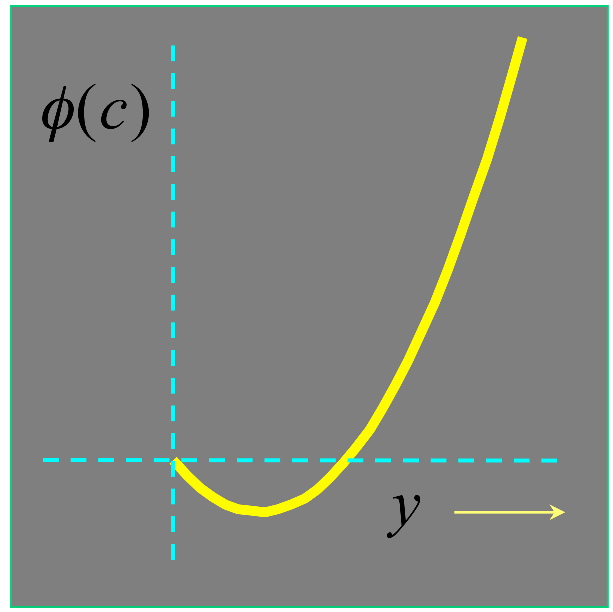
BCM Theory
Stability

- One dimension
- Quadratic form
- **Instantaneous limit**

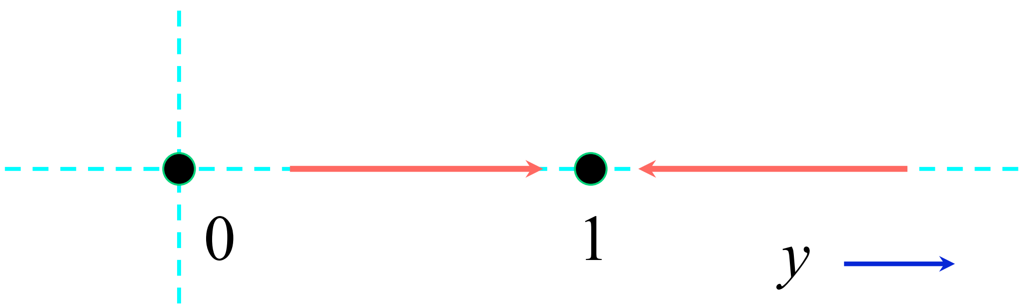
$$y = \mathbf{w} \cdot \mathbf{x}^T$$

$$\frac{dw}{dt} = \eta y \left(y - \theta_M \right) x$$

$$\theta_M = y^2$$

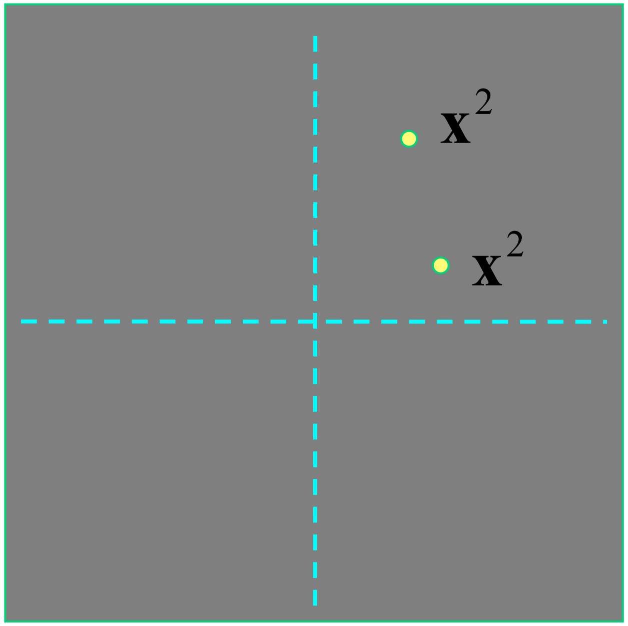


$$\begin{aligned} \frac{dw}{dt} &= \eta y \left(y - y^2 \right) x \\ &= \eta y^2 (1 - y) x \end{aligned}$$



BCM Theory
Selectivity

- Two dimensions $$y = w_1 x_1 + w_2 x_2 = \mathbf{w} \cdot \mathbf{x}^T$$
- Two patterns $$y^1 = \mathbf{w} \cdot \mathbf{x}^1 \ , \ \ y^2 = \mathbf{w} \cdot \mathbf{x}^2$$
- Quadratic form $$\frac{dw}{dt} = \eta \ y^k \left(y^k - \theta_M \right) \mathbf{x}^k$$
- Averaged threshold $$\theta_M = E\left[y^2 \right]_{\text{patterns}} = \sum_{k=1}^{2} p_k (y^k)^2$$

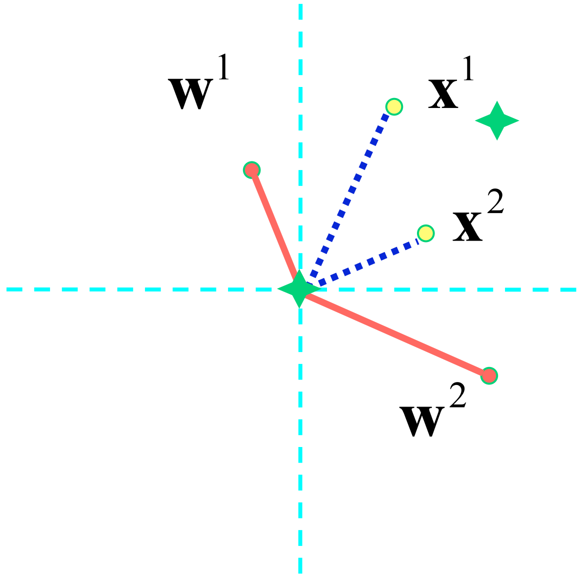


- Fixed points $$\left\langle \frac{d\mathbf{w}}{dt} \right\rangle = 0$$

BCM Theory: Selectivity

- Learning Equation $\frac{d\mathbf{w}}{dt} = \eta y^k \left(y^k - \theta_M \right) \mathbf{x}^k$
- Four possible fixed points

(unselective)	$y^1 = 0$,	$y^2 = 0$
(Selective)*	$y^1 = \theta_M$,	$y^2 = 0$
(Selective)	$y^1 = 0$,	$y^2 = \theta_M$
(unselective)	$y^1 = \theta_M$,	$y^2 = \theta_M$



- Threshold* $\theta_M = p_1 (y^1)^2 + p_2 (y^2)^2 = p_1 (y^1)^2$

$$= y^1 = 1/p_1$$

Stability of Fixed point

Assume w* is a fixed point: $F(w^*) = 0$

Assume that Δw is the deviation from this fixed point.
w=w*+Δw.

Then: $$\frac{dw}{dt} = \frac{dw^*}{dt} + \frac{d\Delta w}{dt} = \frac{d\Delta w}{dt}$$

If $\frac{dw}{dt} = F(w)$ **then**

$$\frac{d\Delta w}{dt} = F(w^* + \Delta w) \approx F\left(w^*\right) + \Delta w \frac{\partial F}{\partial W}\bigg|_{w^*}$$

Consider a selective F.P (w^1) where: $$w^1 \cdot x^1 = \theta_m^1$$ $$w^1 \cdot x^2 = 0$$

and $$\theta_m^1 = E[y^2] = \frac{1}{2}\left[(w^1 \cdot x^1)^2 + (w^1 \cdot x^2)^2\right] = \frac{1}{2}\left[\theta_m^1\right]^2$$

So that $\theta_m^1 = 2$

for a small pertubation from the F.P such that $w = w^* + \Delta w$

The two inputs result in: $$w \cdot x^1 = \theta_m^1 + \Delta w \cdot x^1$$ $$w \cdot x^2 = \Delta w \cdot x^2$$

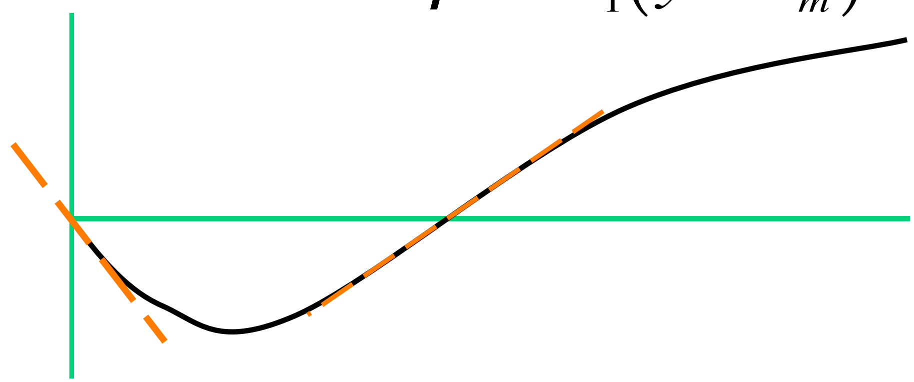
So that

$$\theta_m = \frac{1}{2}\left\{(w^1 \cdot x^1 + \Delta w \cdot x^1)^2 + (w^1 \cdot x^2 + \Delta w \cdot x^2)^2\right\}$$

$$\approx \theta_m^* (1 + \Delta w \cdot x^1) + O(\Delta w^2) = \theta_m^* + 2\Delta w \cdot x^1$$

At y≈0 and at y≈θ_m we make a linear approximation

$$\eta\Phi \approx -\varepsilon_2 y \qquad\qquad \eta\Phi \approx \varepsilon_1 (y - \theta_m)$$



In order to examine whether a fixed point is stable we examine if the average norm of the perturbation $||\Delta w||$ increases or decreases.

<u>Decrease ≡ Stable   Increase ≡ Unstable</u>

Remember: $\theta_m = \theta_m^* + 2\Delta w \cdot x^1$

For the preferred input x^1:

$$\frac{d\Delta w}{dt} = \varepsilon_1 (y - \theta_m) x^1 = \varepsilon_1 (\theta_m^* + \Delta w \cdot x^1 - \theta_m) x^1$$

$$\varepsilon_1 (\theta_m^* + \Delta w \cdot x^1 - \theta_m^* - 2\Delta w \cdot x^1) x = -\varepsilon_1 (\Delta w \cdot x_1) x_1$$

Similarly, for the non preferred input x^2:

$$\dot{\Delta w} = -\varepsilon_2 y \cdot x^2 = -\varepsilon_2 (\Delta w \cdot x^2) x^2$$

Use trick: $\frac{d}{dt}[\Delta w]^2 = \frac{d}{dt}[\Delta w \cdot \Delta w] = 2\Delta w \cdot \Delta \dot{w}$

And average over two input patterns

$$E\left[\frac{d}{dt}[\Delta w]^2\right] = \frac{1}{2}\left[\Delta w \cdot \Delta \dot{w}\right]_{x^1} + \frac{1}{2}\left[\Delta w \cdot \Delta \dot{w}\right]_{x^2}$$

Insert previous result to show that:

$$E\left[\frac{d}{dt}[\Delta w]^2\right] = -\left[\varepsilon_1(\Delta w \cdot x^1)^2 + \varepsilon_2(\Delta w \cdot x^2)^2\right] < 0$$

For the non selective F.P we get:

$$E\left[\frac{d}{dt}[\Delta w]^2\right] \geq 0$$

Phase plane analysis of BCM in 1D

Previous analysis assumed that $\theta_m = E[y^2]$ exactly.
If we use instead the dynamical equation

Will the stability be altered?

$$\frac{d\theta_m}{dt} = \frac{1}{\tau}(y^2 - \theta_m)$$

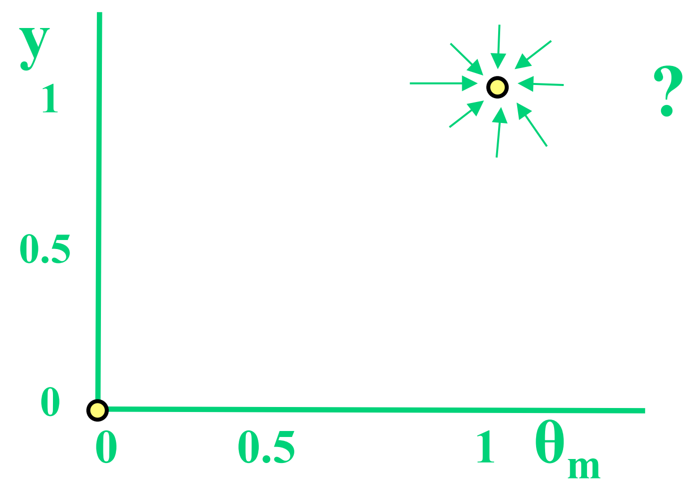
Look at 1D example

Phase plane analysis of BCM in 1D

Assume x=1 and therefore y=w. Get the two BCM equations:

$$\frac{dy}{dt} = \eta y(y - \theta_m)$$

$$\frac{d\theta_m}{dt} = \frac{1}{\tau}(y^2 - \theta_m)$$



There are two fixed points y=0, θ_m=0, and y=1, θ_m=1. The previous analysis shows that the second one is stable, what would be the case here?

How can we do this?

(supplementary homework problem)

Linear stability analysis:

Summary

• The BCM rule is based on two differential equations, what are they?

• When there are two linearly independent inputs, what will be the BCM stable fixed points? What will θ be?

•When there are K independent inputs, what are the stable fixed points? What will θ be?

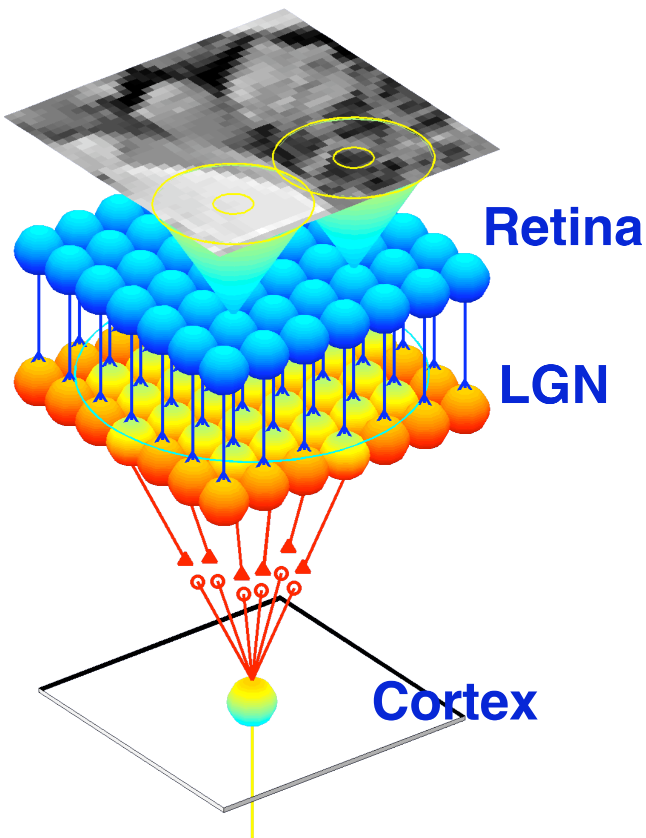
Homework 2: due on Feb 1

1. **Code a single BCM neuron, apply to case with 2 linearly independent inputs with equal probability**
2. **Apply to 2 inputs with different probabilities, what is different?**
3. **Apply to 4 linearly indep. Inputs with same prob.**

Extra credit 25 pt

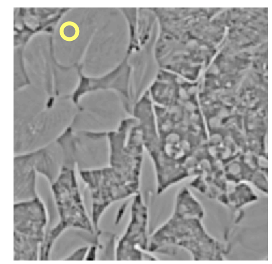
4. **a. Analyze the f.p in 1D case, what are the stable f.p as a function of the systems parameters. b. Use simulations to plot dynamics of $y(t)$, $\theta(t)$ and their trajectories in the m θ plane for different parameters. Compare stability to analytical results (Key parameters, η τ)**

Natural Images, Noise, and Learning



image

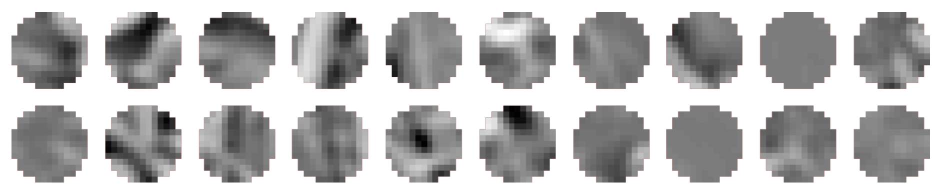
retinal activity



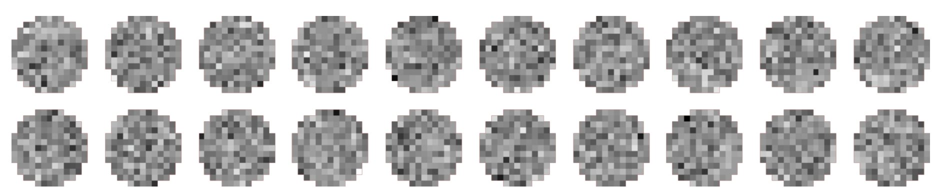
•present patches

•update weights

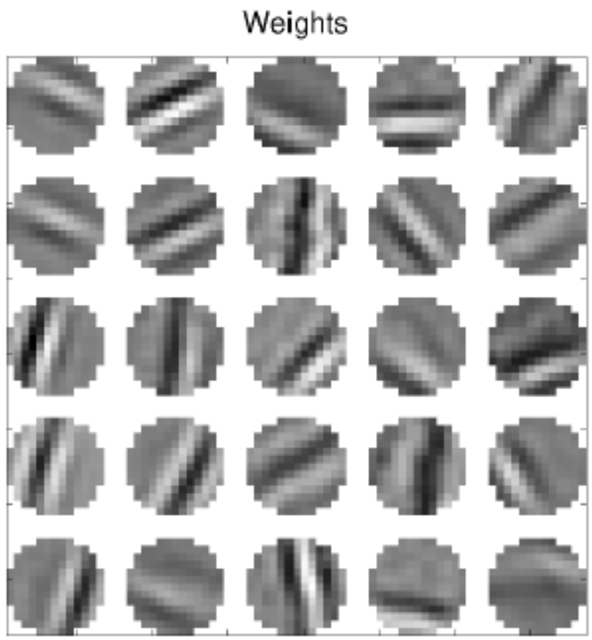
•Patches from retinal activity image



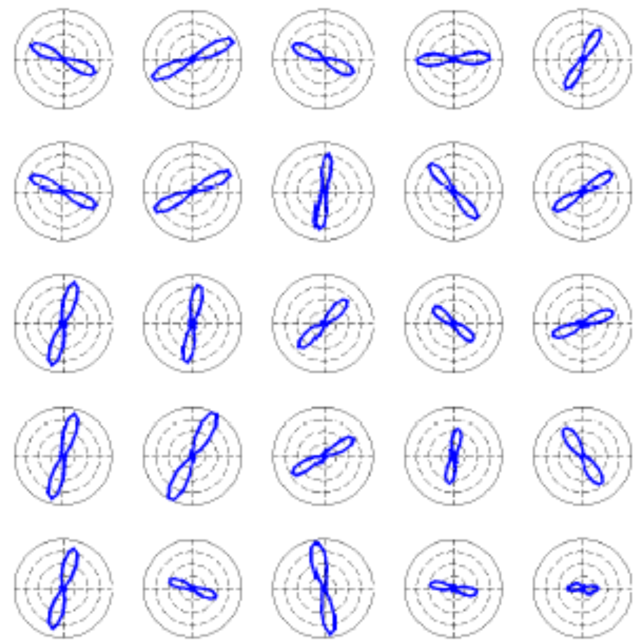
•Patches from noise



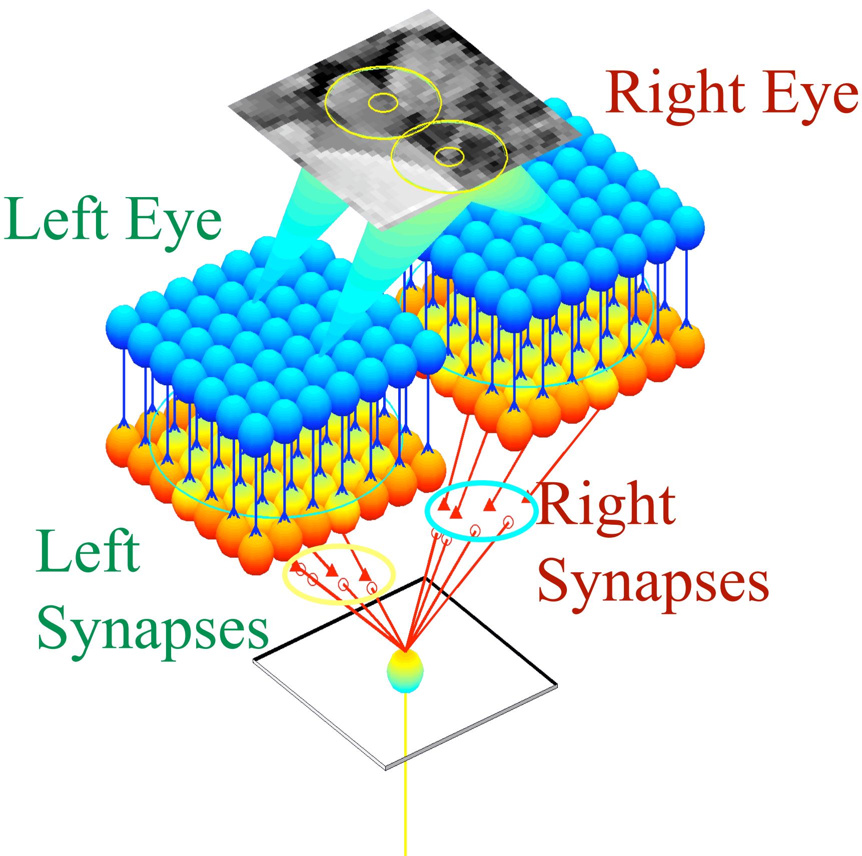
Resulting receptive fields

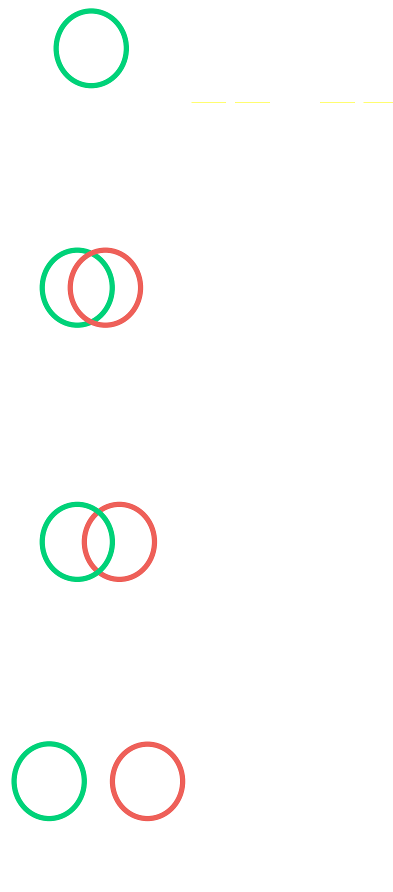


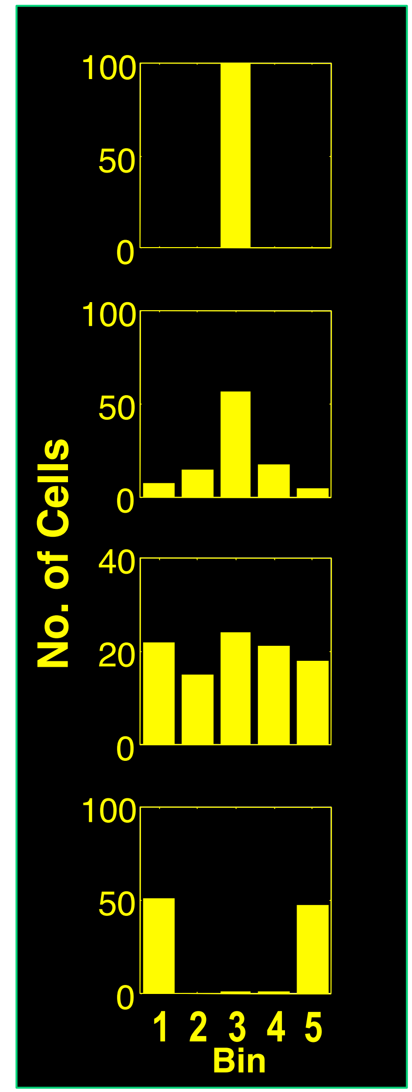
Corresponding tuning curves (polar representation)



BCM neurons can develop both orientation selectivity and varying degrees of Ocular Dominance







Shouval et. al., *Neural Computation*, 1996

The distinction between homosynaptic and heterosynaptic models

Examples:

BCM – homosynaptic

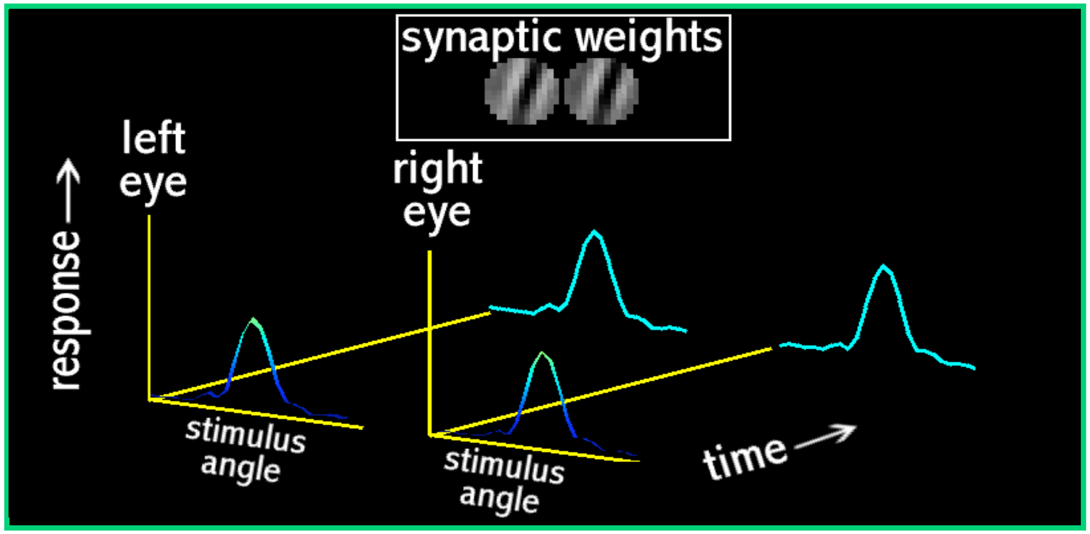
$$\frac{dw_j}{dt} = \eta x_j \phi(y, \theta_m)$$

PCA (Oja) - heterosynaptic

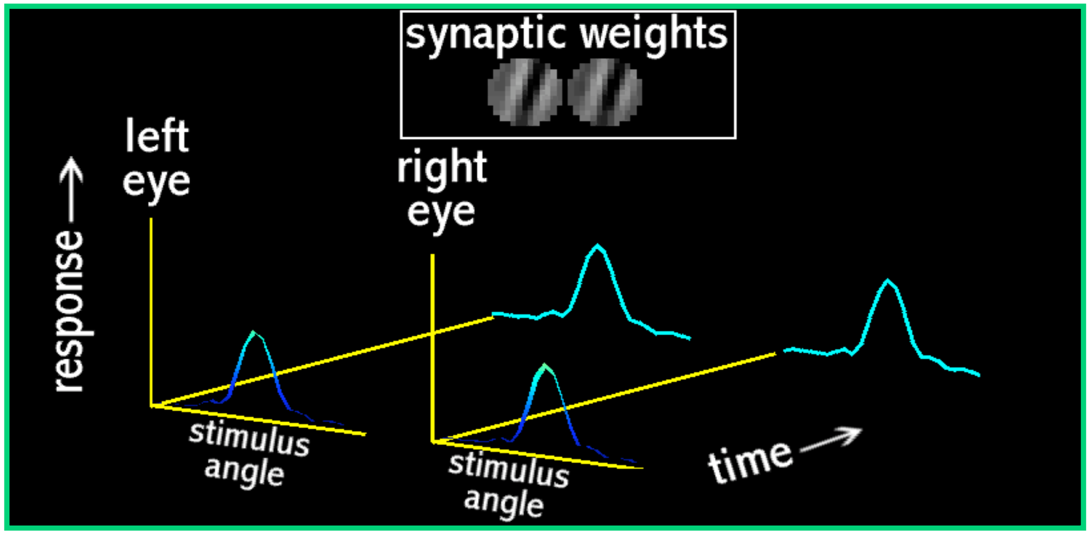
$$\frac{dw_i}{dt} = \eta(x_i y - w_i y^2)$$

Monocular Deprivation
Homosynaptic model (BCM)

Low noise

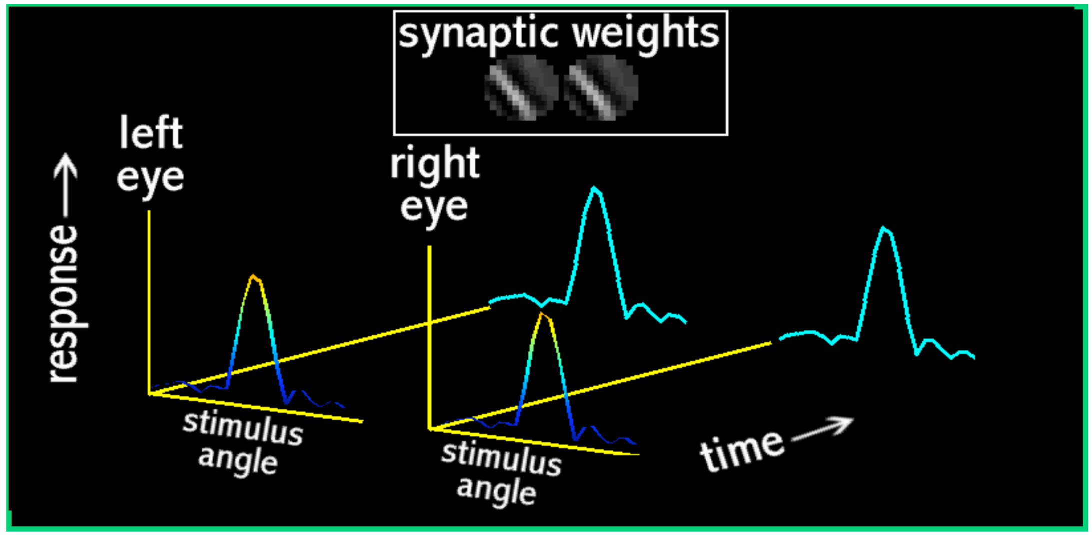


High noise

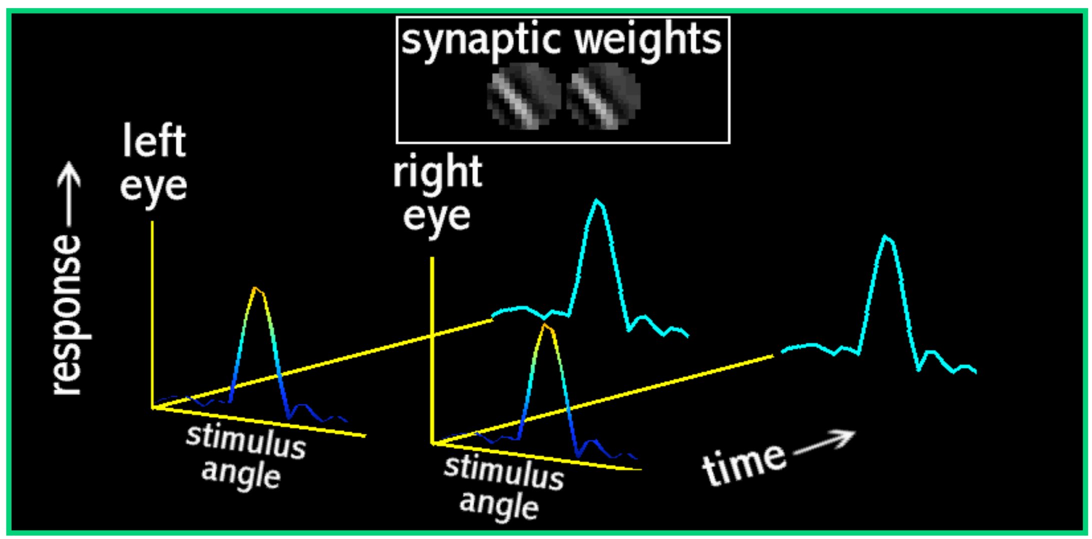


Monocular Deprivation
Heterosynaptic model (K2)

Low noise

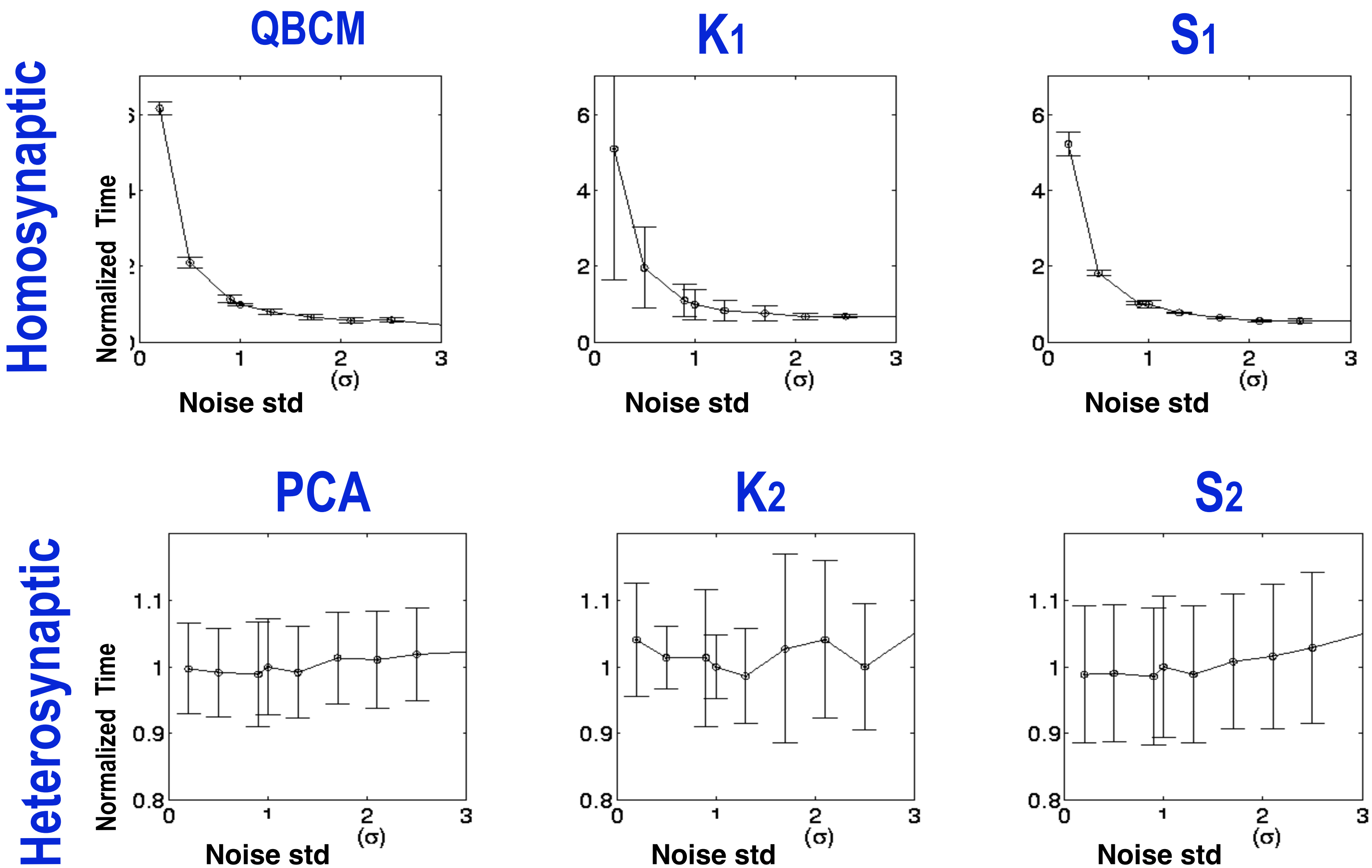


High noise



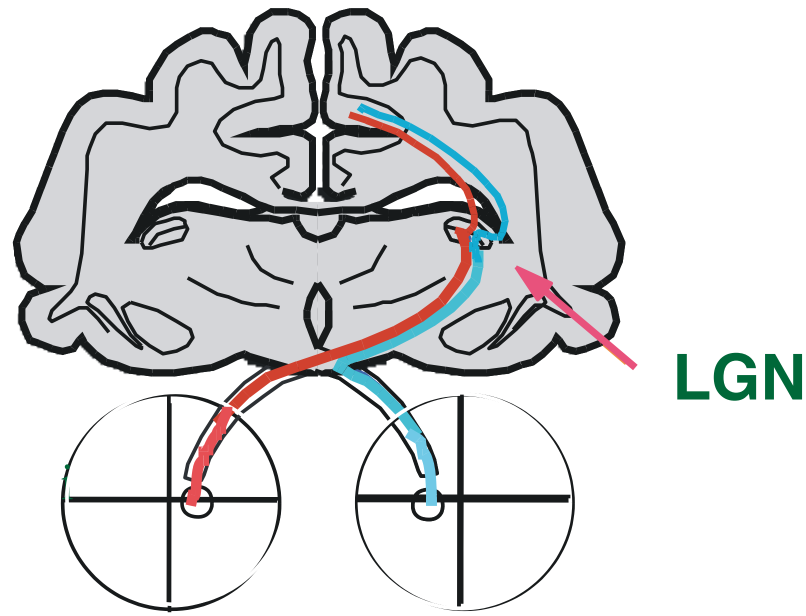
Noise Dependence of MD

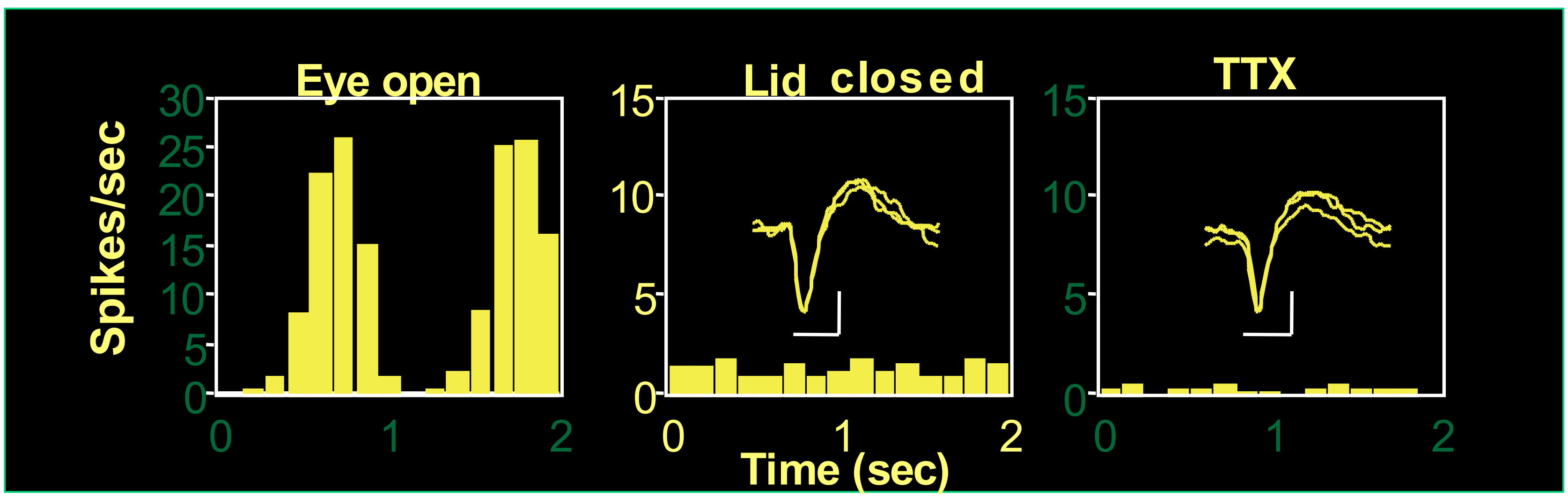
Two families of synaptic plasticity rules



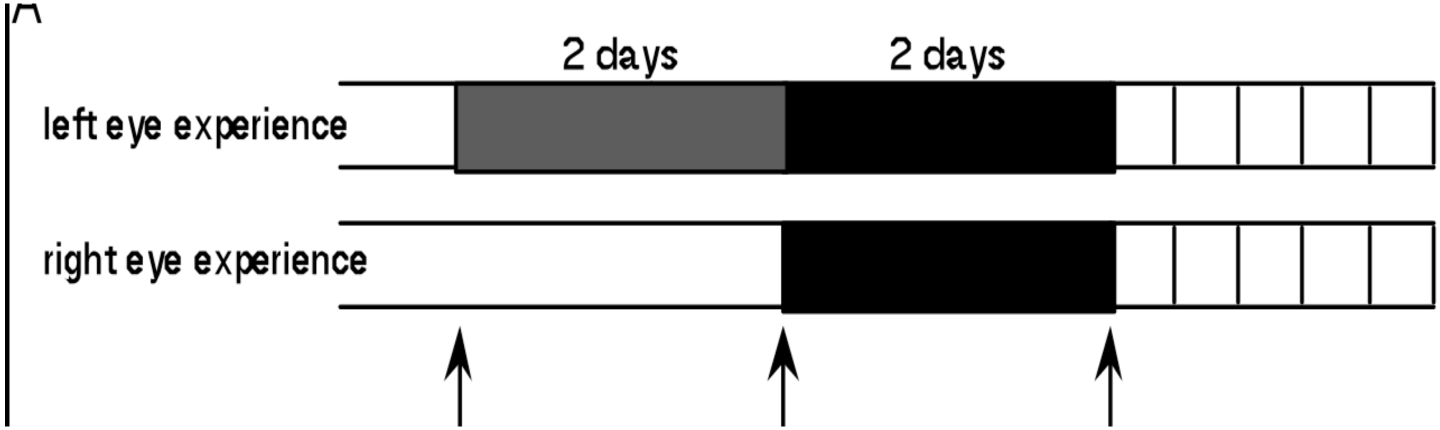
Blais, Shouval, Cooper. *PNAS,* 1999

Intraocular injection of TTX reduces activity of the "deprived-eye" LGN inputs to cortex





Experiment design

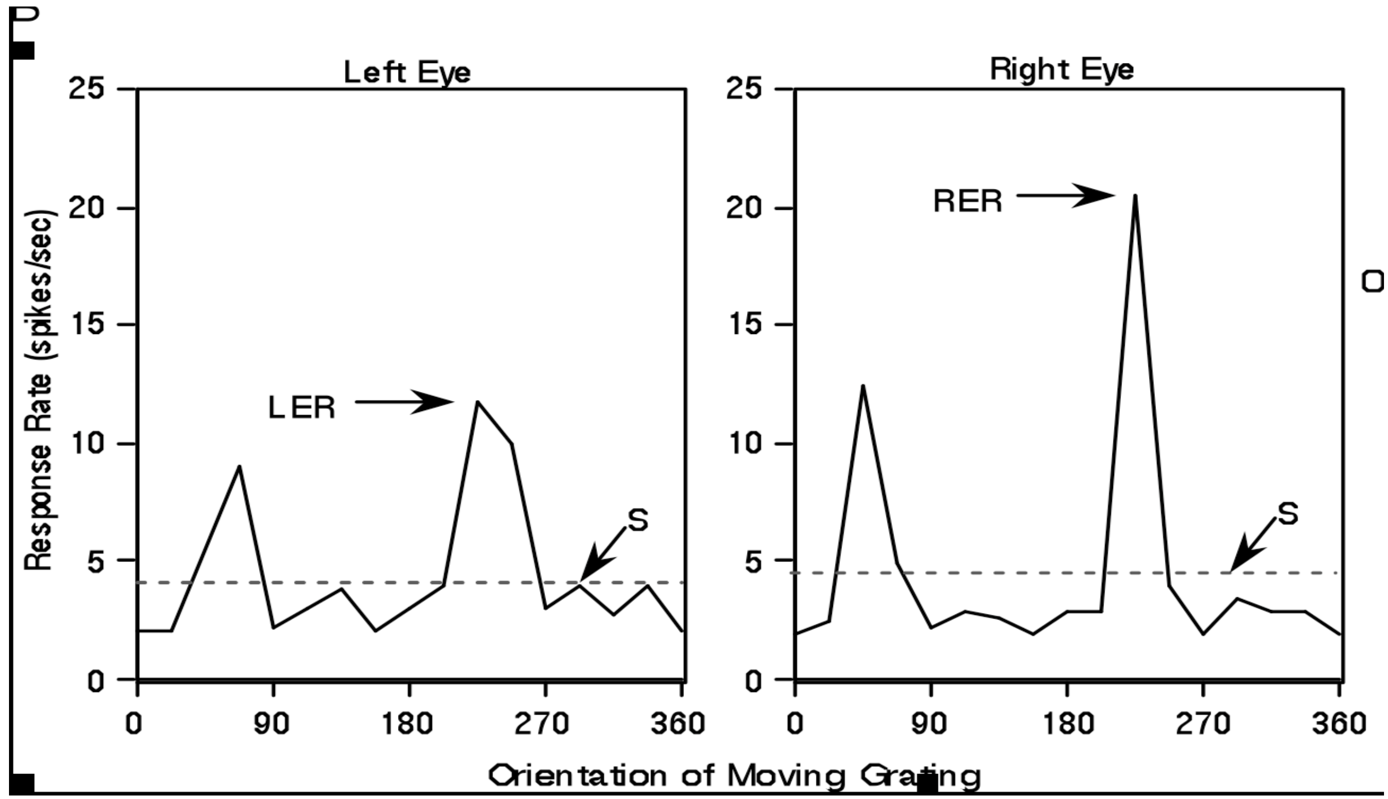


Blind injection of TTX and lid-suture (P49-61)

Dark rearing to allow TTX to wear off

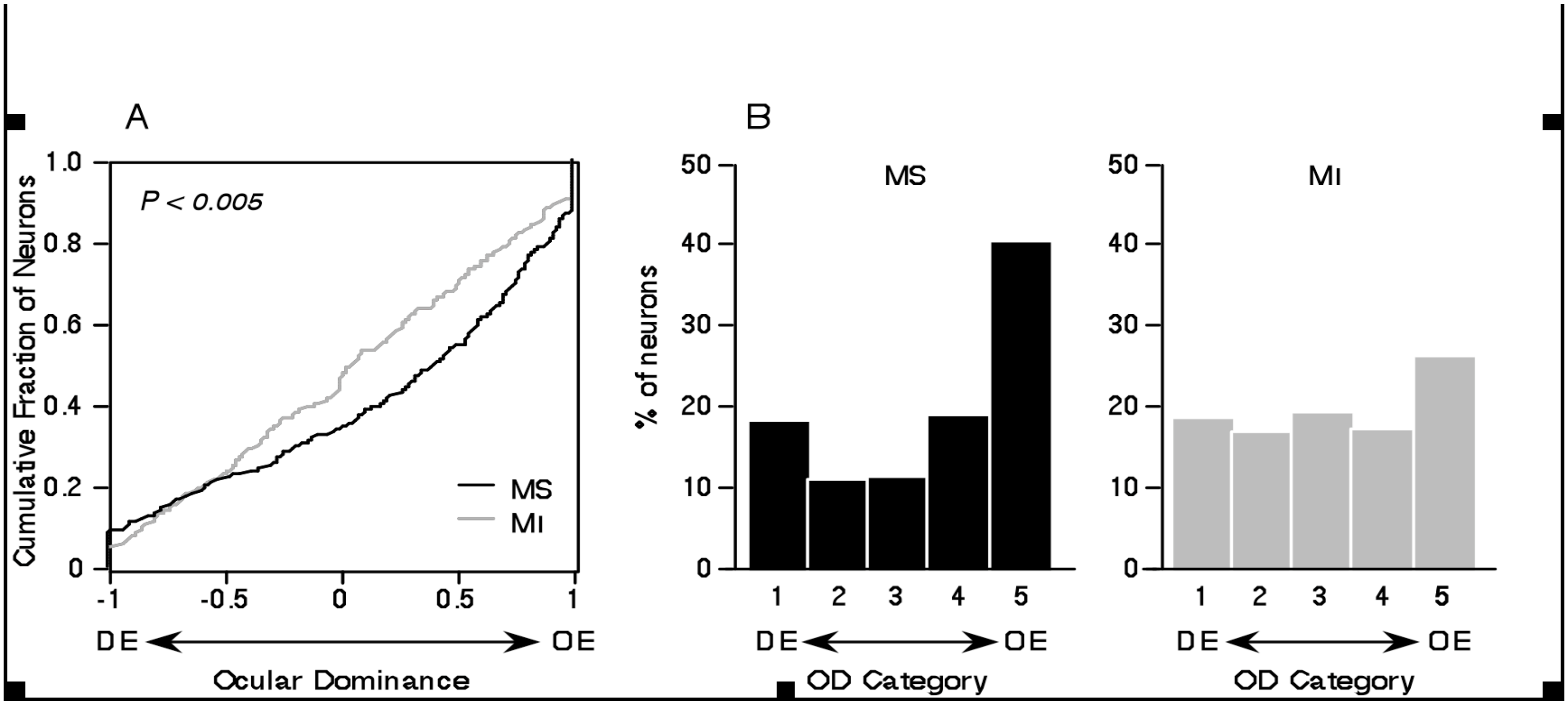
Quantitative measurements of ocular dominance

Response rate (spikes/sec)



$$\mathrm{OD} = \frac{(LER - S) - (RER - S)}{(LER - S) + (RER - S)}$$

Cumulative distribution
Of OD



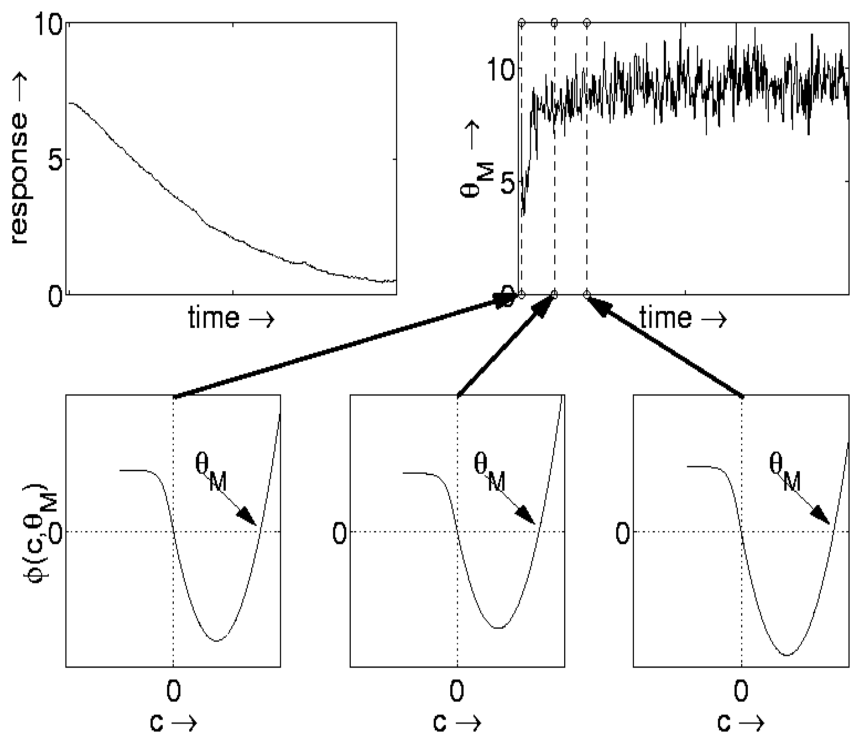
MS= Monocular lid suture

MI= Monocular inactivation (TTX)

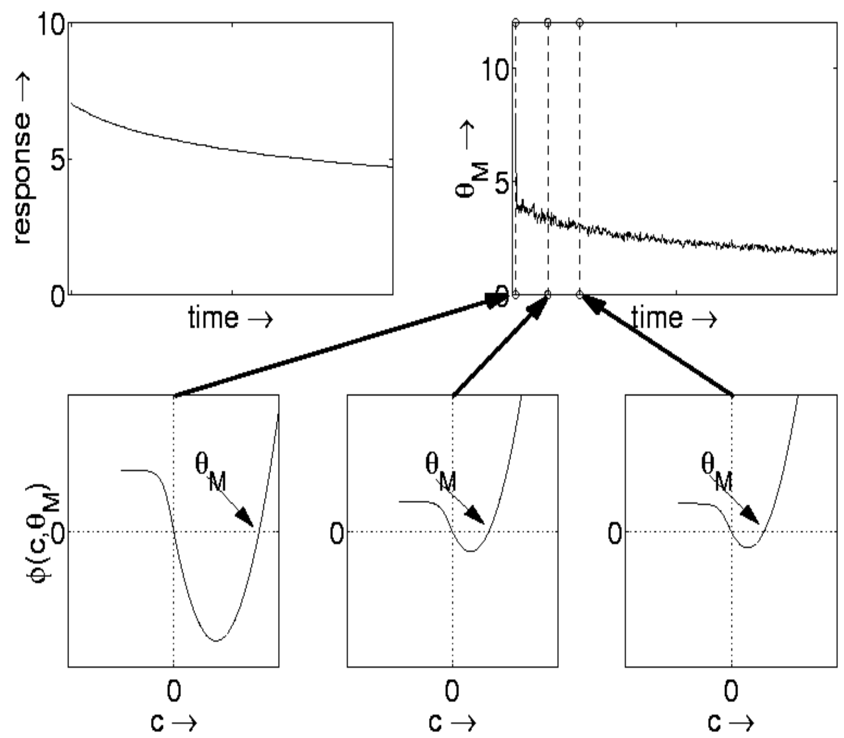
Rittenhouse, Shouval, Paradiso, Bear - *Nature* 1999

Why is Binocular Deprivation slower than Monocular Deprivation?

Monocular Deprivation



Binocular Deprivation



What did we learn up to here?