

# Lecture 7: Stable Hebbian plasticity

Jochen Braun

Otto-von-Guericke-Universität Magdeburg,  
Cognitive Biology Group

Theoretical Neuroscience II, SS 2020

## Lecture 7: Stable Hebbian plasticity

We seek stable Hebbian plasticity, without exponential growth of synaptic weights. We compare and analyze three schemes: **Weight saturation**, **synaptic normalization** (rigidly keeping constant sum of weights), and **Oja rule** (dynamically approaching steady-state of total squared weights). We test performance for biological example: development of 'ocular dominance' (monocular cortical neurons arise by 'growing' and 'pruning' synaptic weights from both eyes). First, analyze different input statistics (matrices of correlation  $\mathbf{Q}$  and covariance  $\mathbf{C}$ ), correlated, anti-correlated, partially correlated, with eigenvectors in 'same' and 'different' directions. Next compare three schemes. Desired outcome (monocular neurons) obtained only from combination of **Oja and covariance rules**: development of synaptic weights is governed by 'different' eigenvector and guarantees monocular steady-state.

# Organization of lecture

- ▶ 1 Recap
- ▶ 2 Stable Hebbian learning rules
  - ▶ Weight saturation
  - ▶ Subtractive normalization
  - ▶ Oja rule
- ▶ 3 Biological example: left and right eye statistics
  - ▶ General statistics
  - ▶ Correlation – Anti-correlation – Partial correlation
- ▶ 4 Biological example: development of ocular dominance
  - ▶ Weight saturation
  - ▶ Subtractive normalization
  - ▶ Oja rule
- ▶ 5 Summary

# 1 Recap

- ▶ In a recurrently connected network, synaptic strengths determine the propagation of activity.
- ▶ When synapses exhibit Hebbian plasticity, the pre- and post-synaptic activities in turn determine the development of synaptic strengths.
- ▶ Controlling this coupled dynamics of activity (faster) and synaptic strengths (slower) has been likened to “taming the beast”.

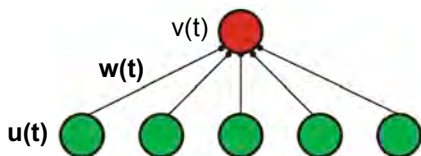


# Correlation-driven plasticity

Hebbian plasticity *without* a threshold is driven by *correlations* (average products of absolute inputs):

$$\tau_w \frac{d\mathbf{w}}{dt} = \langle \mathbf{v} \mathbf{u} \rangle$$

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w}, \quad \mathbf{Q} \equiv \langle \mathbf{u} \mathbf{u} \rangle, \quad v = \mathbf{w} \cdot \mathbf{u}$$

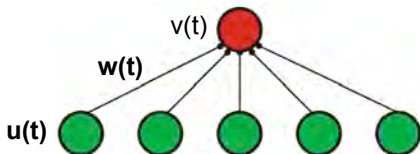


# Covariance-driven plasticity

Hebbian plasticity *with* a threshold is driven by *covariances* (average products of relative inputs, i.e., deviations from mean):

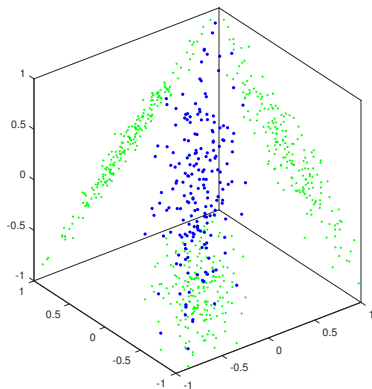
$$\tau_w \frac{d\mathbf{w}}{dt} = \langle v(\mathbf{u} - \langle \mathbf{u} \rangle) \rangle$$

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{C} \mathbf{w}, \quad \mathbf{C} \equiv \langle \mathbf{u}\mathbf{u} \rangle - \langle \mathbf{u} \rangle^2, \quad v = \mathbf{w} \cdot \mathbf{u}$$



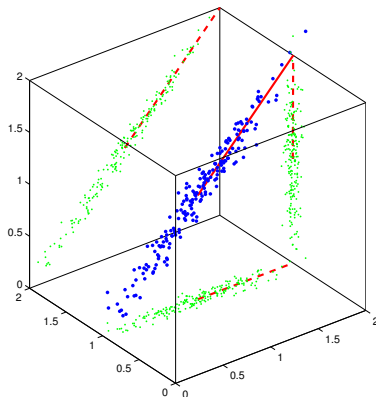
# Correlations and covariances

Correlations reveal which inputs *vary together* around zero ('form a team'). Covariances reveal which inputs *vary together* around their mean. Correlation and covariance are identical for zero-mean input.



## Principal / dominant eigenvector

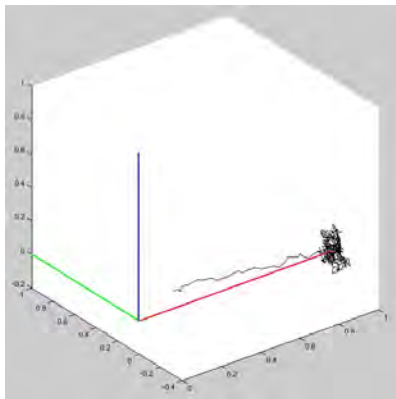
The principal or dominant eigenvector of a correlation/covariance matrix reveals the the direction in input space around which the inputs are most correlated (vary together most closely). It identifies the 'strongest team' among the inputs.





## Synaptic weights align with principal eigenvector

Synaptic weights grow exponentially along all eigenvectors of  $\mathbf{C}$ , but most rapidly along the principal eigenvector. Thus, synaptic weights align with the principal eigenvector.



## 2 Stable Hebbian Learning rules

We now consider ways of restraining the exponential growth of synaptic weights and take a first step towards “taming the beast”.



## 2.1 Weight saturation

If synaptic weights are bounded, limitless growth is (trivially) prevented:

$$\tau_w \frac{d\mathbf{w}}{dt} = \langle v(\mathbf{u} - \langle u \rangle) \rangle, \quad 0 \leq w_i \leq 1$$

However, as we will see later, final weights depend largely on initial weights (and the choice of maximal weight). Input statistics plays only a limited role. In short, this rule does not guarantee the desired outcome (ocular dominance).

## 2.2 Synaptic normalization

If the total synaptic weight is fixed, one synapse can be strengthened only at the expense of weakening others. Such a constraint can be enforced **rigidly** at all times:

$$\tau_w \frac{d\mathbf{w}}{dt} = \langle v(\mathbf{u} - \bar{u}) \rangle, \quad \bar{u} = \frac{1}{N} \sum_{i=1}^N u_i$$

Averaging over all  $w_i$ , we see that this rule obviously imposes stability

$$\tau_w \frac{d\bar{w}}{dt} = \langle v(\bar{u} - \bar{u}) \rangle = 0, \quad \bar{w} = \frac{1}{N} \sum_{i=1}^N w_i$$

Plus point: development is activity-driven. Minus points: summed weights don't grow, requires 'cross-talk' between synapses, does not guarantee desired outcome (as we will see later).

## ctd (on board)

For the 2D case, we reformulate this in terms of the correlation matrix  $\mathbf{Q}$ , substituting  $v = \mathbf{w} \cdot \mathbf{u}$ :

$$\begin{aligned}\tau_w \frac{d\mathbf{w}}{dt} &= \langle \mathbf{w} \cdot \mathbf{u} \left[ \mathbf{u} - \frac{1}{2} (u_1 + u_2) \right] \rangle = \\ &= \left\langle \left( \begin{array}{cc} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{array} \right) \right\rangle \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} - \left\langle (w_1 u_1 + w_2 u_2) \frac{1}{2} (u_1 + u_2) \right\rangle = \\ &= \left\langle \left( \begin{array}{cc} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{array} \right) \right\rangle \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} - \frac{1}{2} \langle w_1 u_1^2 + w_2 u_1 u_2 + w_1 u_1 u_2 + w_2 u_2^2 \rangle = \\ &= \mathbf{Q} \cdot \mathbf{w} - \frac{1}{2} \mathbf{n}^T \cdot \mathbf{Q} \cdot \mathbf{w}\end{aligned}$$

This expresses *synaptic normalization* in terms of the correlation matrix  $\mathbf{Q}$ .

## 2.3 Oja rule

Alternatively, normalization can be imposed **dynamically** in the limit of large times ('in the long run'), with the 'Oja rule':

$$\tau_w \frac{d\mathbf{w}}{dt} = \langle v \mathbf{u} - \alpha v^2 \mathbf{w} \rangle$$

Multiplying (dot-product) this equation with  $\mathbf{w}$ , we find a steady-state for the total squared weights:

$$\tau_w \frac{d|\mathbf{w}|^2}{dt} = \langle v^2 \rangle - \alpha \langle v^2 \rangle |\mathbf{w}|^2$$

$$\tau_w \frac{d|\mathbf{w}|^2}{dt} = 0 \quad \Rightarrow \quad |\mathbf{w}_{ss}|^2 = 1/\alpha$$

The Oja rule is highly competitive and small weights are reduced disproportionately ('poor get poorer').

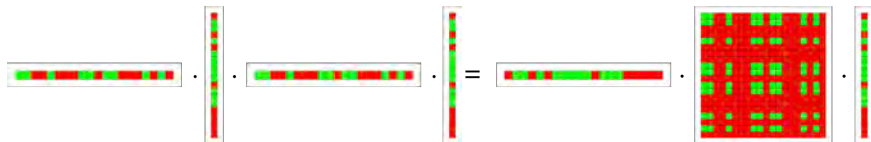
The Oja rule can also be reformulated in terms of *correlations*:

$$\tau_w \frac{d\mathbf{w}}{dt} = \langle v \mathbf{u} - \alpha v^2 \mathbf{w} \rangle, \quad v = \mathbf{u} \cdot \mathbf{w}$$

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \alpha \left( \mathbf{w}^T \mathbf{Q} \mathbf{w} \right) \mathbf{w}$$

where we have used

$$\langle (\mathbf{u} \cdot \mathbf{w})^2 \rangle = \mathbf{w}^T \mathbf{Q} \mathbf{w}$$



Graphical illustration of linear algebra equivalence.

## Linear analysis of Oja rule

The linear analysis is carried out most easily in eigenvector coordinates. Each weight component  $w_\mu$  represents an eigenvector  $\mathbf{e}_\mu$ :

$$\mathbf{w}(t) = \sum_{\mu} w_{\mu}(t) \mathbf{e}_{\mu} \qquad \mathbf{Q} \cdot \mathbf{e}_{\mu} = \lambda_{\mu} \mathbf{e}_{\mu}$$

Using

$$\mathbf{Q} \cdot \mathbf{w} = \sum_{\mu} w_{\mu}(t) \mathbf{Q} \cdot \mathbf{e}_{\mu} = \sum_{\mu} w_{\mu}(t) \lambda_{\mu} \mathbf{e}_{\mu}$$

we can rewrite

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \cdot \mathbf{w} - \alpha \left( \mathbf{w}^T \cdot \mathbf{Q} \cdot \mathbf{w} \right) \mathbf{w}$$

as

$$\tau_w \frac{dw_{\mu}}{dt} = \left( \lambda_{\mu} - \alpha \sum_{\nu} \lambda_{\nu} w_{\nu}^2 \right) w_{\mu}$$



$$\tau_w \frac{dw_\mu}{dt} = \left( \lambda_\mu - \alpha \sum_\nu \lambda_\nu w_\nu^2 \right) w_\mu$$

The bracket term is the learning rate for components aligned with eigenvector  $\mu$ . As this rate is largest for the dominant eigenvector 1, the sum will come to be dominated by its largest term

$$\alpha \sum_\nu \lambda_\nu w_\nu^2 \longrightarrow \alpha \lambda_1 w_1^2$$

so that

$$\tau_w \frac{dw_1}{dt} = (1 - \alpha w_1^2) \lambda_1 w_1$$

Accordingly, the weights will align with the largest eigenvector.

## 2 Points to note

- ▶ Several schemes have been proposed to stabilize Hebbian learning.
- ▶ Weight saturation stabilizes development, but loses input dependence.
- ▶ Synaptic normalization subtracts the average synaptic weight, *rigidly* keeping constant summed weights:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \frac{1}{2} \mathbf{n}^T \mathbf{Q} \mathbf{w}$$

- ▶ The Oja rule subtracts a more complex term, *dynamically* approaching a steady-state for total squared weights:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \alpha \left( \mathbf{w}^T \mathbf{Q} \mathbf{w} \right) \mathbf{w}$$

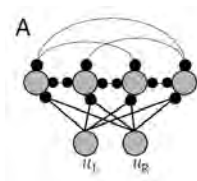
- ▶ Next we compare these schemes for a biological example.

### 3 Biological example: left and right eye statistics

We compare stabilization schemes for a biological example.

In young animals, the development of sensory pathways is shaped by an interplay of sensory experience and developmental mechanisms. The development of **ocular dominance** in the early visual pathway is a well-studied example

Initially, visual neurons receive projections from both left and right eyes. During certain 'critical periods' of development, these connections are pruned such that most neurons come to be 'dominated' by input from one or the other eye.



**A:** Two input units (left and right eye) and four output units (visual cortex).

**B:** Ocular dominance columns in area 17 of kitten (top) and schematic illustration of alternating dominance (bottom).

## 3.1 General statistics of left and right eye inputs

For simplicity, we assume that the left and right eye provide binary inputs with balanced probabilities:

$$u_L, u_R \in \{0, 1\}, \quad p_0 = p_1 = 1/2$$

In other words, both  $u_L$  and  $u_R$  take values of 0 or 1 with equal probability.

We leave as a variable parameter the probability  $p_{11}$  that both eye inputs are 1. The probabilities for other combinations follow:

$$p_{01} = p_1 - p_{11} = 1/2 - p_{11}, \quad p_{10} = p_1 - p_{11} = 1/2 - p_{11},$$

$$p_{00} = 1 - p_{11} - p_{01} - p_{10} = p_{11}$$

# Probability table

	$u_L = 1$	$u_L = 0$	
$u_R = 1$	$p_{11}$	$p_1 - p_{11}$	$p_1$
$u_R = 0$	$p_1 - p_{11}$	?	$1 - p_1$
	$p_1$	$1 - p_1$	1

$$p_{00} = 1 - p_{01} - p_{10} - p_{11} = 1 - 2p_1 + p_{11}$$

	$u_L = 1$	$u_L = 0$	
$u_R = 1$	$p_{11}$	$p_1 - p_{11}$	$p_1$
$u_R = 0$	$p_1 - p_{11}$	$1 - 2p_1 + p_{11}$	$1 - p_1$
	$p_1$	$1 - p_1$	1

## Correlations and covariances

Correlations are 'mean products'

$$\langle u_L u_L \rangle = p_0 \cdot 0 \cdot 0 + p_1 \cdot 1 \cdot 1 = p_1 = 1/2$$

$$\langle u_R u_R \rangle = p_0 \cdot 0 \cdot 0 + p_1 \cdot 1 \cdot 1 = p_1 = 1/2$$

$$\langle u_R u_L \rangle = p_{00} \cdot 0 \cdot 0 + p_{01} \cdot 0 \cdot 1 + p_{10} \cdot 1 \cdot 0 + p_{11} \cdot 1 \cdot 1 = p_{11}$$

Covariances are 'mean products' minus 'product of means'

$$\langle u_L u_L \rangle - \langle u_L \rangle^2 = 1/2 - p_1^2 = 1/4$$

$$\langle u_R u_R \rangle - \langle u_R \rangle^2 = 1/2 - p_1^2 = 1/4$$

$$\langle u_R u_L \rangle - \langle u_L \rangle \langle u_R \rangle = p_{11} - p_1^2 = p_{11} - 1/4$$

$$\mathbf{Q} = \begin{pmatrix} \langle u_L u_L \rangle & \langle u_L u_R \rangle \\ \langle u_R u_L \rangle & \langle u_R u_R \rangle \end{pmatrix} = \begin{pmatrix} 1/2 & p_{11} \\ p_{11} & 1/2 \end{pmatrix} \equiv \begin{pmatrix} q_S & q_D \\ q_D & q_S \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 1/4 & p_{11} - 1/4 \\ p_{11} - 1/4 & 1/4 \end{pmatrix} = \begin{pmatrix} c_S & c_D \\ c_D & c_S \end{pmatrix}$$

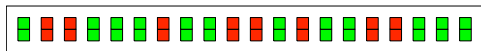
Development will be governed by eigenvectors and eigenvalues! For a symmetric 2D matrix, we expect eigenvectors in 'same' and in the 'different' direction.

$$\mathbf{e}_{same} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \mathbf{e}_{diff} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

Which direction dominates and has largest eigenvalue?

## 3.2 Eyes with correlated activity

At one extreme, activities  $u_R$  and  $u_L$  are *perfectly correlated* ( $p_{11} = 1/2$ ) (i.e., half on, half off, but always the same):



$p(u_L, u_R)$		$u_L$		$p(u_R)$
		0	1	
$u_R$	0	0.5	0	0.5
	1	0	0.5	0.5
$p(u_L)$		0.5	0.5	

$$q_S = p_1 \cdot 1 \cdot 1 = 1/2$$

$$q_D = p_{11} \cdot 1 \cdot 1 = 1/2$$

$$c_S = p_1 \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = 1/4$$

$$c_D = p_{11} \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = 1/4$$

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$$



The resulting eigenvalues and eigenvectors of  $\mathbf{Q}$  and  $\mathbf{C}$ , respectively, are

$$\mathbf{e}_{1,2} = \left( \begin{array}{c} 1/\sqrt{2} \\ 1/\sqrt{2} \end{array} \right), \left( \begin{array}{c} 1/\sqrt{2} \\ -1/\sqrt{2} \end{array} \right) \quad \lambda_{1,2} = 1, 0$$

$$\mathbf{e}_{1,2} = \left( \begin{array}{c} 1/\sqrt{2} \\ 1/\sqrt{2} \end{array} \right), \left( \begin{array}{c} 1/\sqrt{2} \\ -1/\sqrt{2} \end{array} \right) \quad \lambda_{1,2} = 1/2, 0$$

In both cases,  $\mathbf{e}_{same}$  dominates!

For *perfectly correlated* inputs, synaptic weights can grow only in 'same' direction, never in 'different' direction!

### 3.3 Eyes with anti-correlated activity

At another extreme, activities  $u_R$  and  $u_L$  are *perfectly anti-correlated* ( $p_{11} = 0$ ) (i.e., half on, half off, always different):



$p(u_L, u_R)$		$u_L$		$p(u_R)$
		0	1	
$u_R$	0	0	0.5	0.5
	1	0.5	0	0.5
$p(u_L)$		0.5	0.5	1

$$q_S = p_1 \cdot 1 \cdot 1 = 1/2$$

$$q_D = p_{11} \cdot 1 \cdot 1 = 0$$

$$Q = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

$$c_S = p_1 \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = 1/4$$

$$c_D = p_{11} \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = -1/4$$

$$C = \begin{pmatrix} 1/4 & -1/4 \\ -1/4 & 1/4 \end{pmatrix}$$

The resulting eigenvalues and eigenvectors of  $\mathbf{Q}$  and  $\mathbf{C}$ , respectively, are

$$\mathbf{e}_{1,2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \lambda_{1,2} = 1/2, 1/2$$

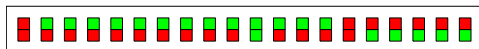
$$\sqrt{2} \mathbf{e}_{1,2} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \lambda_{1,2} = 1/2, 0$$

In both cases,  $\mathbf{e}_{different}$  dominates!

For *perfectly anti-correlated* inputs, synaptic weights can grow only in 'different' direction, never in 'same' direction!

## 3.4 Eyes with partially correlated activity

A third possibility is that activities are *partially correlated*



$p(u_L, u_R)$		$u_L$		$p(u_R)$
		0	1	
$u_R$	0	$1/8$	$3/8$	$1/2$
	1	$3/8$	$1/8$	$1/2$
$p(u_L)$		$1/2$	$1/2$	1

$$p_0 = p_1 = 1/2, \quad p_{00} = p_{11} = 1/8, \quad p_{01} = p_{10} = 3/8$$

$$q_S = p_1 \cdot 1 \cdot 1 = 1/2$$

$$q_D = p_{11} \cdot 1 \cdot 1 = 1/8$$

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 1/8 \\ 1/8 & 1/2 \end{pmatrix}$$

$$c_S = p_1 \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = 1/4$$

$$c_D = p_{11} \cdot 1 \cdot 1 - (p_1 \cdot 1)^2 = -1/8$$

$$\mathbf{C} = \begin{pmatrix} 1/4 & -1/8 \\ -1/8 & 1/4 \end{pmatrix}$$

## Partial correlation

The resulting eigenvectors and eigenvalues are:

$$\mathbf{e}_{1,2} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \quad \lambda_{1,2} = 5/8, 3/8$$

*Correlations* are dominated by the “same” eigenvector.

$$\mathbf{e}_{1,2} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \lambda_{1,2} = 3/8, 1/8$$

*Covariances* are dominated by the “different” eigenvector.

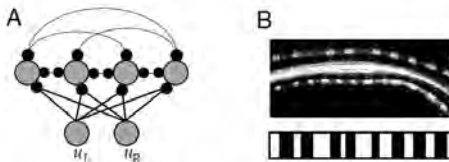
### 3 Points to note

- ▶ Binary inputs from two eyes may exhibit different correlational statistics.
- ▶ The correlation/covariance matrices and their principal eigenvectors capture these differences.
- ▶ Correlation entails a large “same” eigenvector  $(1, 1)$ .
- ▶ Anti-correlation entails a large “different” eigenvector  $(1, -1)$ .
  
- ▶ In intermediate cases, both eigenvectors have non-zero eigenvalues.
- ▶ Specifically, the “same” eigenvector dominates *correlations*  $\mathbf{Q}$ , the “different” eigenvector dominates *covariances*  $\mathbf{C}$ .

**We now apply different stable learning schemes to the development of ocular dominance.**

## 4 Biological example: development of ocular dominance

What plasticity rule could 'prune' the initial, unspecific connections such that neurons come to be **completely dominated** by one eye or the other? On other words, such that neurons become **completely monocular**?

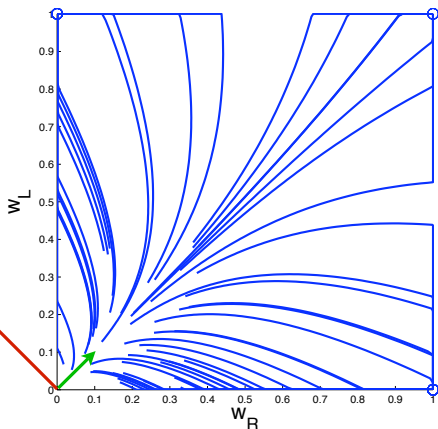
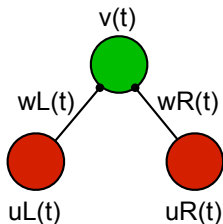


To obtain complete 'eye dominance', input statistics must be dominated by a 'different eigenvector'  $(1, -1)$ . Thus, input must be at least partially anti-correlated and development must follow a *covariance* rule!

## 4.1 Weight saturation

As a first attempt to stabilize the development of ocular dominance, we add weight saturation to a *covariance* rule:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{C} \mathbf{w}, \quad 0 \leq w_{L,R} \leq 1$$



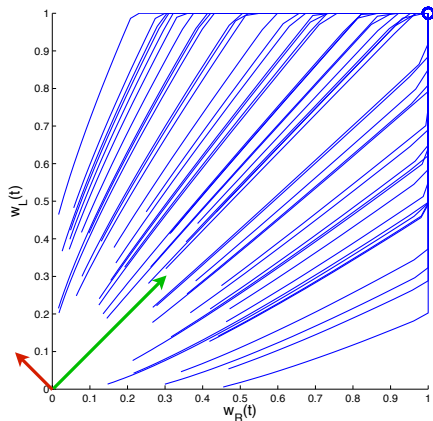
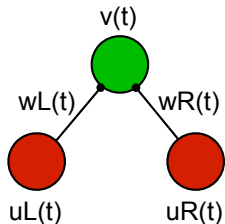


# Weight saturation

- ▶ 'Different' eigenvector dominates, due to *partial anti-correlation* and *covariance* rule.
- ▶ Development results sometimes in different weights ( $w_L \neq w_R$ ), sometimes in identical weights ( $w_L = w_R$ ).
- ▶ Outcome depends on initial weights and saturation level (see Exercise 3).
- ▶ Thought experiment: use *correlation* rule instead instead of *covariance* rule.
- ▶ 'Same' eigenvector dominates.
- ▶ Development always results in identical weights ( $w_L = w_R$ ).

## Weight saturation with *correlation* rule

$$\tau_w \frac{dw}{dt} = \mathbf{Q} \mathbf{w}, \quad 0 \leq w_{L,R} \leq 1$$



## 4.2 Subtractive normalization

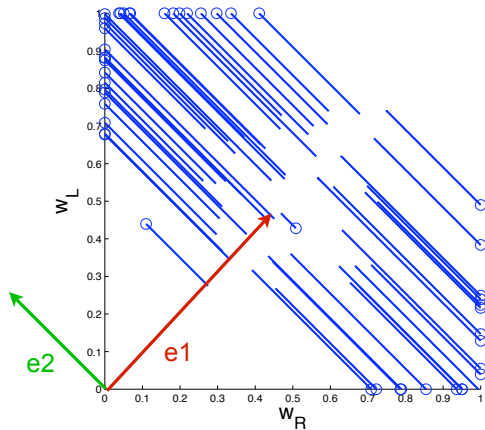
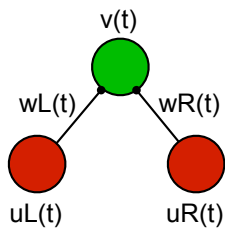
Let's try to stabilize a *correlation rule* by adding subtractive normalization:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \frac{1}{2} \left( \mathbf{n}^T \mathbf{Q} \mathbf{w} \right) \mathbf{n}, \quad 0 \leq w_{L,R} \leq 1$$

Recall that the “same” eigenvector dominates, so Hebbian development should favor ‘binocular’ inputs.

We will see that synaptic normalization prevents this outcome.

ctd



To understand this unexpected outcome, we arrange the dynamic equation in eigenvector coordinates

$$\mathbf{w} = \mathbf{w}_+ + \mathbf{w}_-, \quad \mathbf{w}_+ \equiv w_+ \mathbf{e}_1, \quad \mathbf{w}_- \equiv w_- \mathbf{e}_2,$$

We find that

$$\tau_w \frac{d\mathbf{w}_+}{dt} = \mathbf{Q} \mathbf{w}_+ - \frac{1}{2} \left( \mathbf{n}^T \mathbf{Q} \mathbf{w}_+ \right) \mathbf{n} = 0$$

$$\tau_w \frac{d\mathbf{w}_-}{dt} = \mathbf{Q} \mathbf{w}_- - \frac{1}{2} \left( \mathbf{n}^T \mathbf{Q} \mathbf{w}_- \right) \mathbf{n} = (q_S - q_D) \mathbf{w}_-$$

Normalization prevents *any* growth along  $\mathbf{n}$ , which aligns with to  $\mathbf{e}_1 = \mathbf{e}_{same}$ . All growth is along  $\mathbf{e}_2 = \mathbf{e}_{diff}$ .

# Proof

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \frac{1}{2} \left( \mathbf{n}^T \mathbf{Q} \mathbf{w} \right) \mathbf{n}$$

$$\tau_w \frac{d}{dt} \begin{pmatrix} w_L \\ w_R \end{pmatrix} = \begin{pmatrix} q_S & q_D \\ q_D & q_S \end{pmatrix} \begin{pmatrix} w_L \\ w_R \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} q_S & q_D \\ q_D & q_S \end{pmatrix} \begin{pmatrix} w_L \\ w_R \end{pmatrix}$$

$$\tau_w \frac{d}{dt} \begin{pmatrix} w_L \\ w_R \end{pmatrix} = \begin{pmatrix} q_S w_L + q_D w_R \\ q_D w_L + q_S w_R \end{pmatrix} - \frac{1}{2} (q_S w_L + q_D w_R + q_D w_L + q_S w_R)$$

(on board)

$$w_+ = w_L + w_R$$

$$\begin{aligned} \tau_w \frac{d}{dt} w_+ &= (q_S w_L + q_D w_R + q_D w_L + q_S w_R) - \\ &- (q_S w_L + q_D w_R + q_D w_L + q_S w_R) = 0 \end{aligned}$$

$$w_- = w_L - w_R$$

$$\tau_w \frac{d}{dt} w_- = (q_S w_L + q_D w_R) - (q_D w_L + q_S w_R) = (q_S - q_D) w_-$$

# Subtractive normalization

- ▶ Normalization prevents growth along the dominant (same) eigenvector.
- ▶ Any growth is along the second (different) eigenvector.
- ▶ Development always results in different weights ( $w_L \neq w_R$ ).
- ▶ The typical outcome, partial dominance, is *not* the desired outcome.
- ▶ Development rigidly maintains the initial sum of weights.
- ▶ Requires interactions ('cross talk') between synapses.
- ▶ Is considered biologically unrealistic.



## 4.3 Oja with covariance rule

We stabilize a *covariance* rule by subtracting an Oja term:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{C} \mathbf{w} - \alpha \left( \mathbf{w}^T \mathbf{C} \mathbf{w} \right) \mathbf{w} =$$

$$= \begin{pmatrix} c_S & c_D \\ c_D & c_S \end{pmatrix} \begin{pmatrix} w_L \\ w_R \end{pmatrix} - \alpha (c_S w_L^2 + 2c_D w_L w_R + c_S w_R^2) \begin{pmatrix} w_L \\ w_R \end{pmatrix},$$

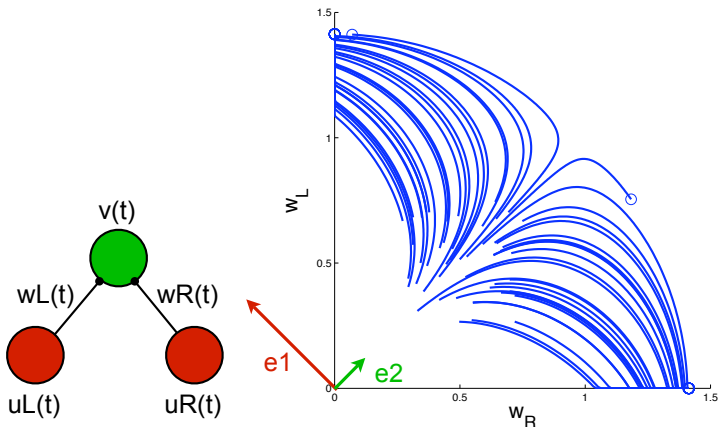
with

$$w_{L,R} \geq 0$$

Recall that the Oja term limits the synaptic weights to

$$|\mathbf{w}^2| = w_L^2 + w_R^2 \leq 1/\alpha$$

For  $\alpha = 1/2$  the weights are limited to  $|\mathbf{w}^2| \leq 2$



The weights consistently grow in the direction of the principal eigenvector ('different' direction), leading to **complete dominance** ('monocular neurons'), the desired outcome.

## Oja with *correlation* rule

Thought experiment: use *correlation* rule instead instead of *covariance* rule. Can an Oja term stabilize a *correlation* rule?

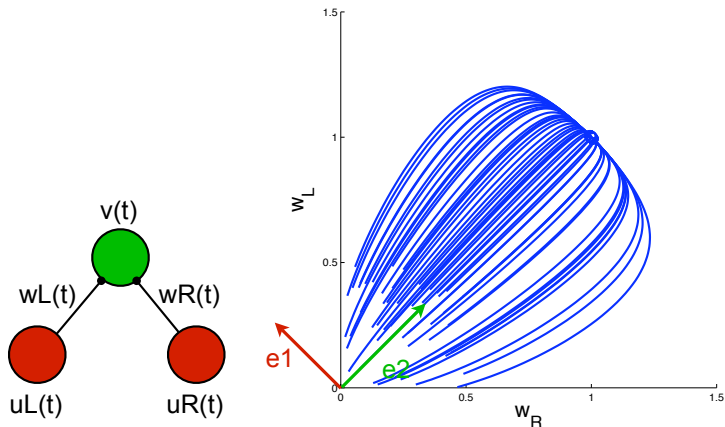
$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \alpha \left( \mathbf{w}^T \mathbf{Q} \mathbf{w} \right) \mathbf{w} =$$

$$= \begin{pmatrix} q_S & q_D \\ q_D & q_S \end{pmatrix} \begin{pmatrix} w_L \\ w_R \end{pmatrix} - \alpha (q_S w_L^2 + 2q_D w_L w_R + q_S w_R^2) \begin{pmatrix} w_L \\ w_R \end{pmatrix},$$

with

$$w_{L,R} \geq 0$$

Development produces stable outcome. 'Same' eigenvector dominates and outcome is identical weights ( $w_L = w_R$ ), or **binocular neurons**.



The weights consistently grow in the direction of the principal eigenvector ('same' direction), leading to **binocular neurons**.

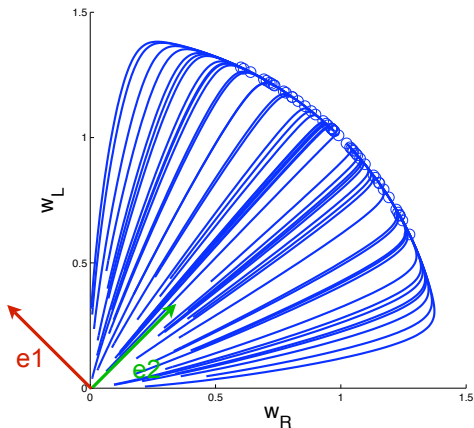
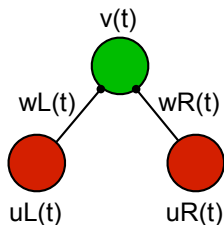
ctd

Consider highly (but not completely) anti-correlated  $\mathbf{Q}$ . Outcome unchanged:

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 1/40 \\ 1/40 & 1/2 \end{pmatrix}$$

$$\sqrt{2} \mathbf{e}_{1,2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

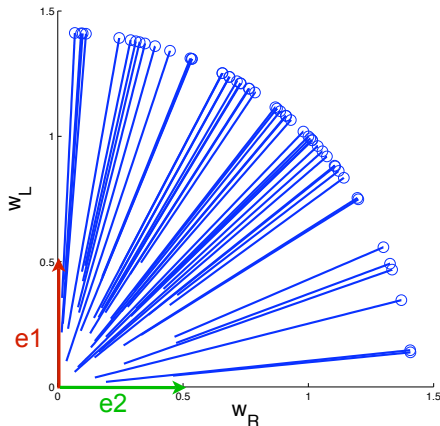
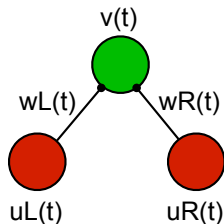
$$\lambda_{1,2} = 21/40, 19/40$$



ctd

Consider completely anti-correlated  $\mathbf{Q}$ , with degenerate eigenvectors. Development amplifies initial imbalance in  $\mathbf{w}$  until maximal sum of squared weights is reached:

$$\mathbf{Q} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \mathbf{e}_{1,2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \lambda_{1,2} = 1/2, 1/2$$



# Oja rule

- ▶ An Oja term stabilizes either *covariance* and *correlation* rules.
- ▶ The sum of squared weights approaches a fixed value  $1/\alpha$ .
- ▶ Initially, weights grow in all directions.
- ▶ Eventually, weights align with the principal eigenvector.
- ▶ The dominant eigenvector is the 'same' direction for *correlation* and 'different' direction for *covariance* rule.
- ▶ Combination of Oja and covariance rules ensure the desired outcome: **ocular dominance**.
- ▶ Next lecture: a *correlation* rule can produce **ocular dominance** with intra-cortical competition.

## 5 Summary

- ▶ We have introduced different ways of “taming” the growth of synaptic weights in Hebbian development (HD).
- ▶ Weight saturation sets an upper bound:

$$w_i \leq w_{max}$$

- ▶ Subtractive normalization *rigidly* maintains summed weights:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{Q} \mathbf{w} - \frac{1}{2} \left( \mathbf{n}^T \mathbf{Q} \mathbf{w} \right) \mathbf{n}$$

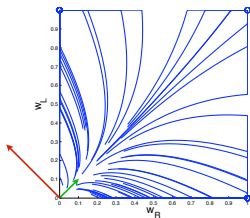
- ▶ The Oja rule *dynamically* approaches steady-state level of summed squared weights:

$$\tau_w \frac{d\mathbf{w}}{dt} = \mathbf{C} \mathbf{w} - \alpha \left( \mathbf{w}^T \mathbf{C} \mathbf{w} \right) \mathbf{w}$$



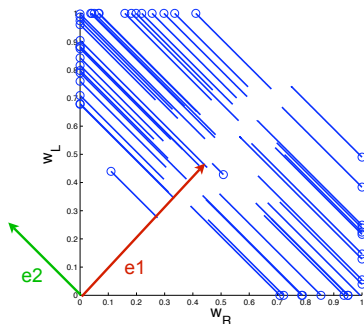
## Summary, ctd

- ▶ In the development of ocular dominance, the goal is to increase input disparity, so that the post-synaptic cells becomes “monocular” (*i.e.*, dominated by either eye).
- ▶ We tried to ensure this outcome with Hebbian rules, assuming partially anti-correlated input.
- ▶ When HD is stabilized by **weight saturation**, the post-synaptic cell often becomes “monocular” but can sometimes remain “binocular”.



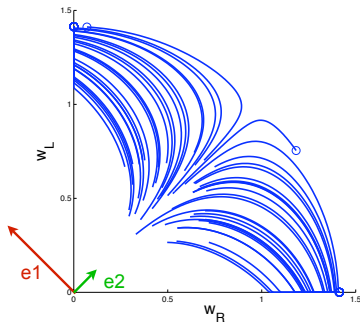
## Summary, ctd

- ▶ When HD is stabilized by **subtractive normalization**, post-synaptic cells become biased, but not completely “monocular”.
- ▶ Also, sum of weights cannot grow and synapses must interact.
- ▶ Considered biologically unrealistic.



## Summary, ctd

- ▶ When HD is stabilized by the **Oja rule**, post-synaptic cells become completely 'monocular' if we assume a covariance rule. ( $\langle u \rangle > 0$ ).



- ▶ With a correlation rule, post-synaptic cells become completely 'binocular'.

# Next: More Hebbian development