Computational Neuroscience: Neuronal Dynamics of Cognition

EPFL
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

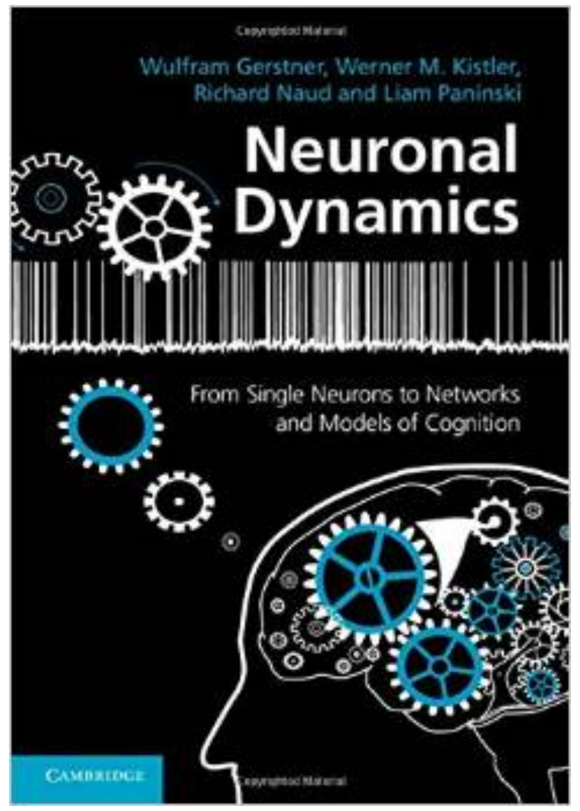
Attractor Networks and Generalizations of the Hopfield model

Wulfram Gerstner

EPFL, Lausanne, Switzerland

Reading for week 6:
NEURONAL DYNAMICS
- Ch. 17.2.5 – 17.4

Cambridge Univ. Press



1. **Attractor networks**
2. **Stochastic Hopfield model**
3. **Energy landscape**
4. **Towards biology (1)**
   - low-activity patterns
5. **Towards biology (2)**
   - spiking neurons

1. Review and next steps

6.1. Attractor networks

6.2. Stochastic Hopfield model

6.3. Energy landscape

6.4. Towards biology (1)

- low-activity patterns

6.5 Towards biology (2)

- spiking neurons

1. Review and next steps

Hopfield model

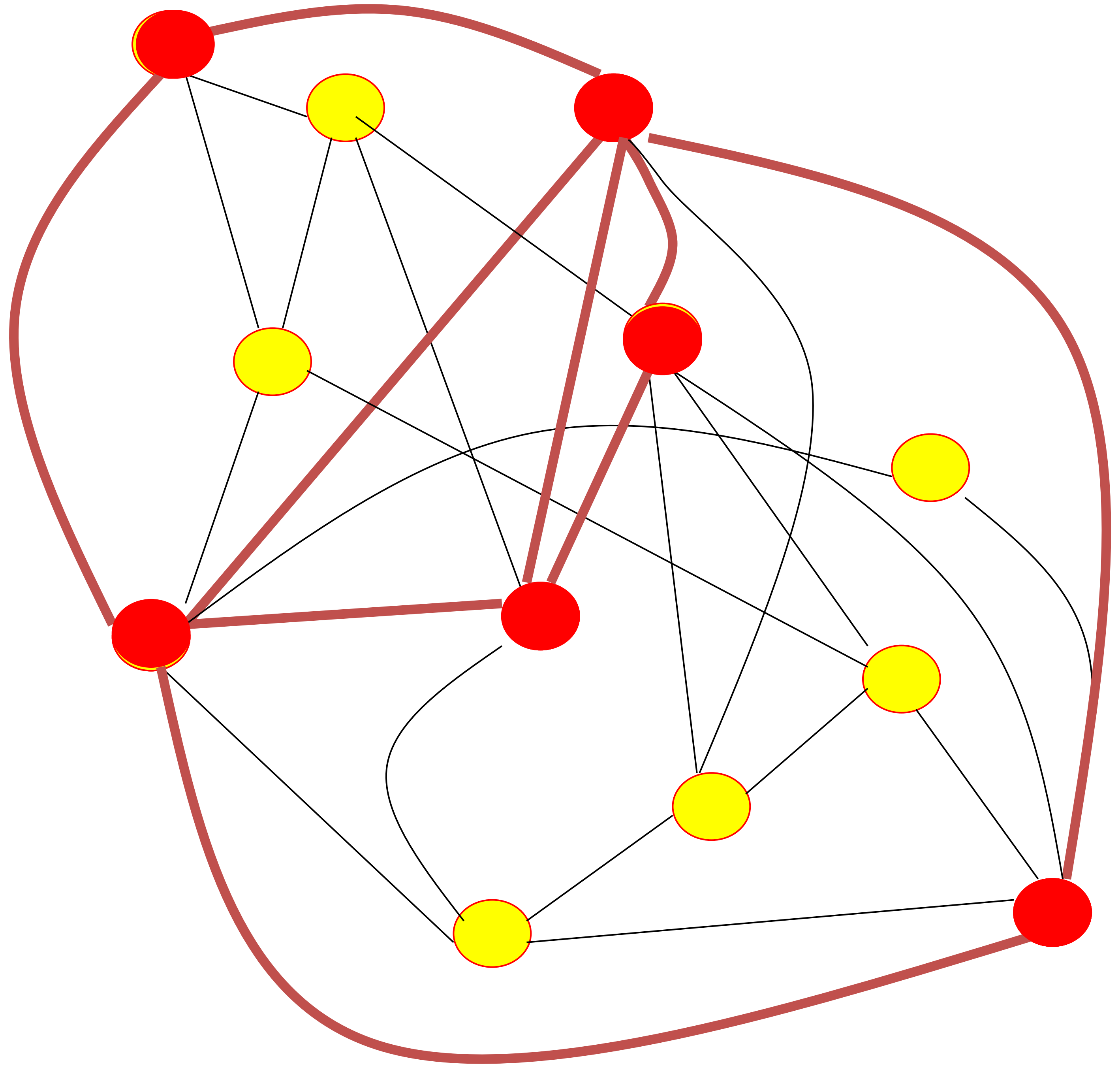
special case

attractor

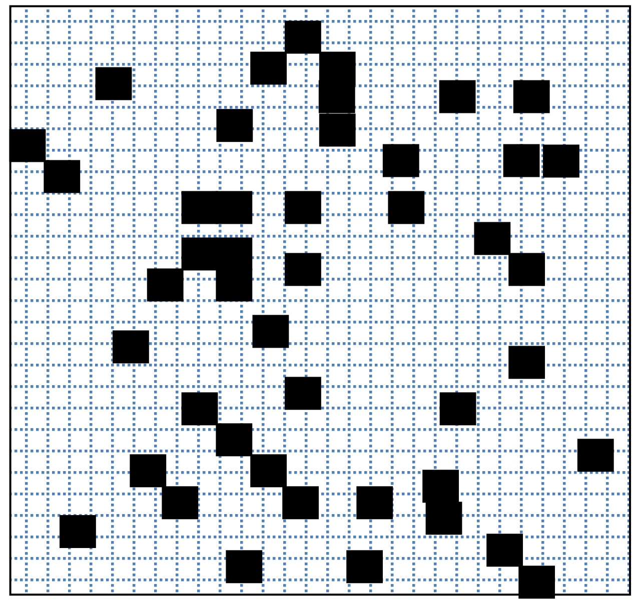
energy

biology

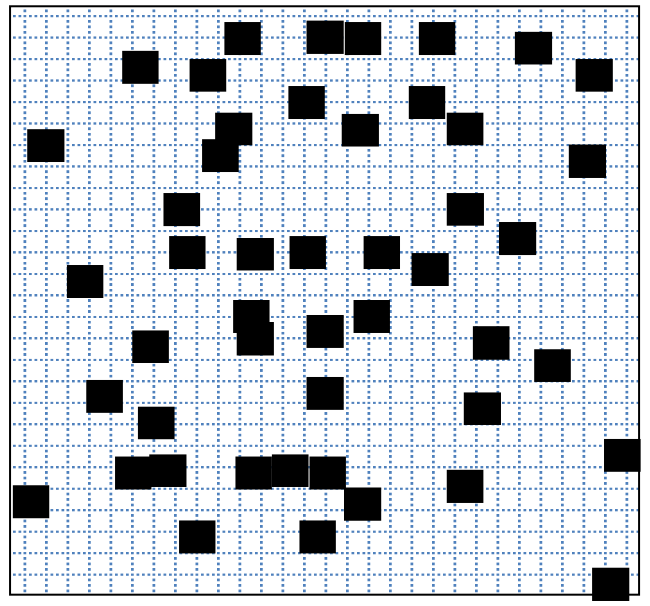
1. Review of last week 5



1. Review of last week: Deterministic Hopfield model



Prototype $\vec{p}^1$



Prototype $\vec{p}^2$

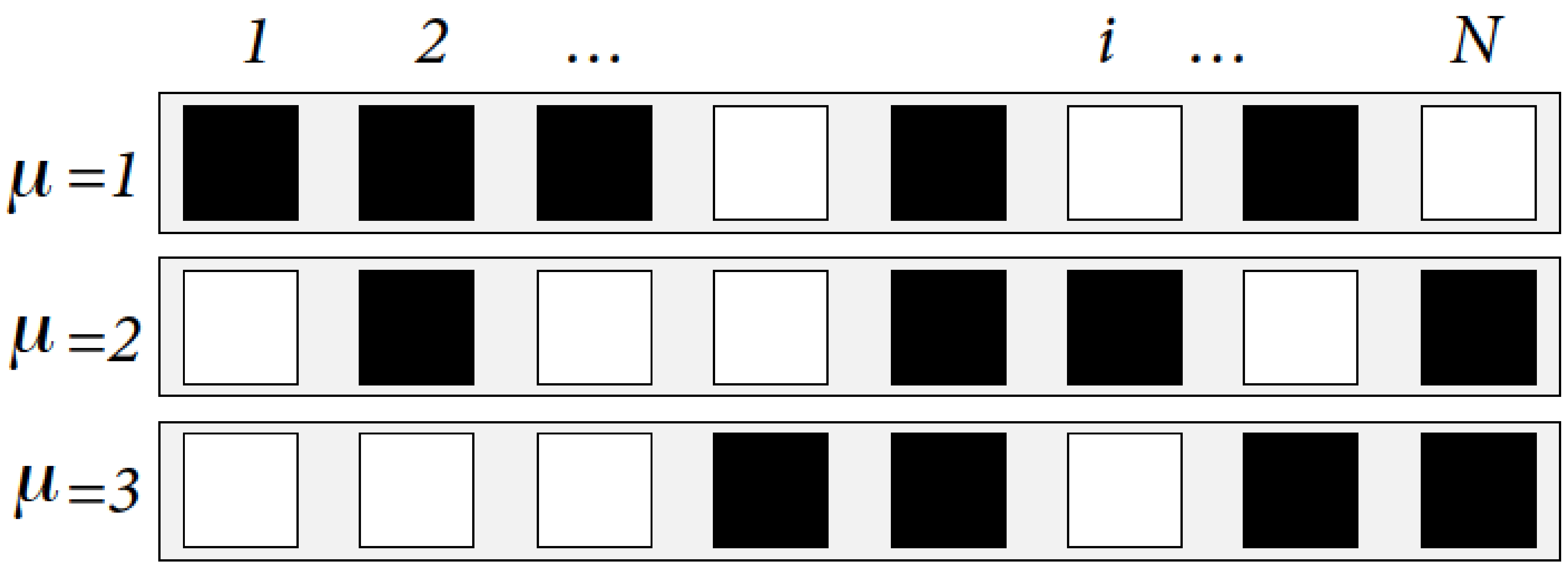
interactions $$w_{ij} = \frac{1}{N}\sum_{\mu} p_i^{\mu} p_j^{\mu}$$

Sum over all prototypes

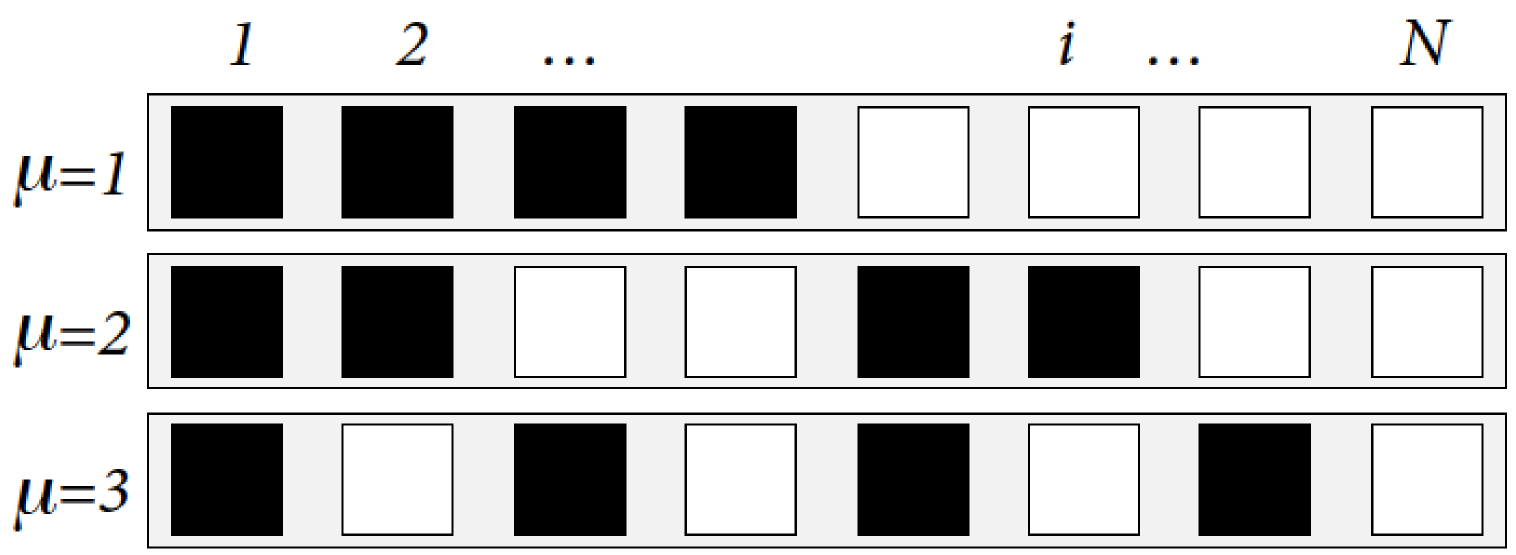
- each prototype has black pixels with probability 0.5
- prototypes are random patterns, chosen once at the beginning

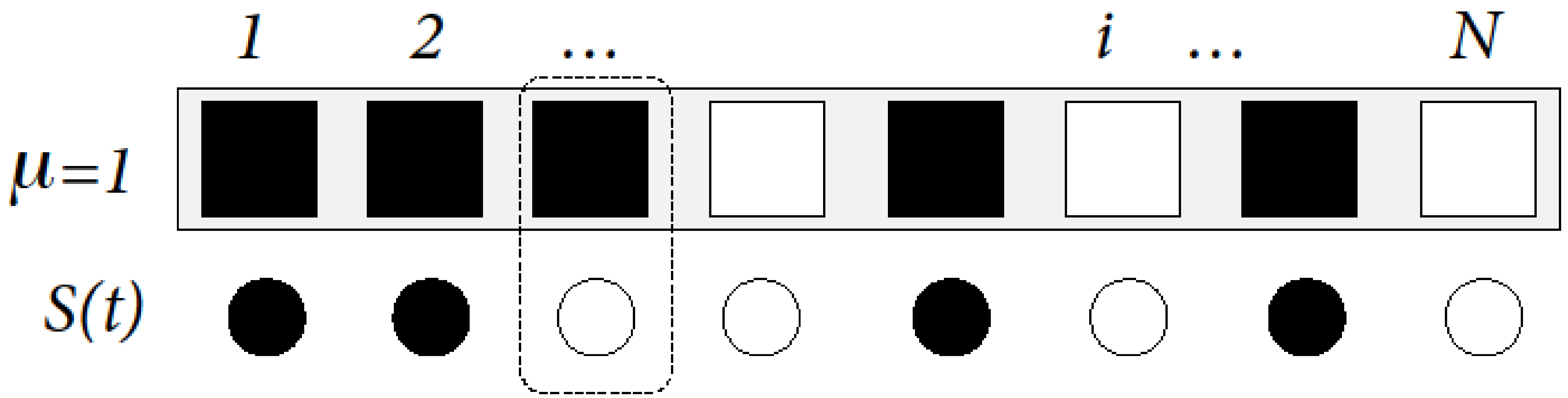
1. Review of last week: overlap / correlation

Image: Neuronal Dynamics, Gerstner et al., Cambridge Univ. Press (2014),



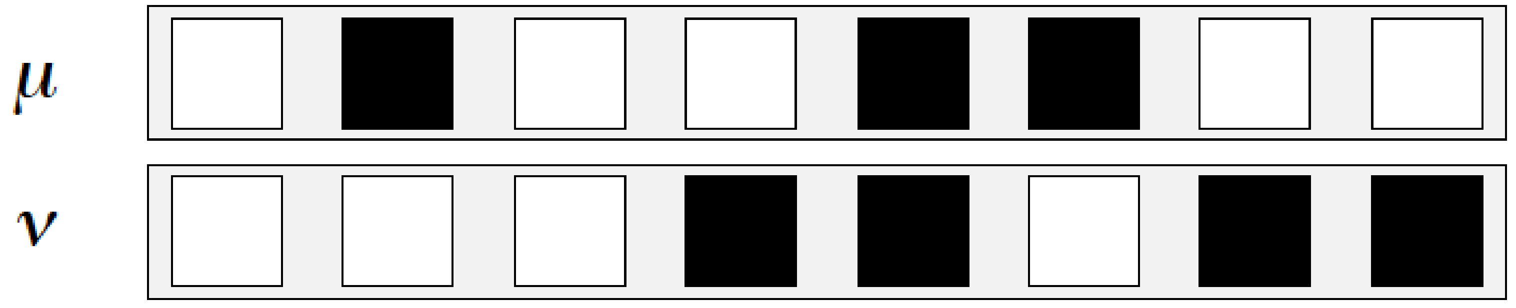
Correlation: overlap between *one pattern* and *another*





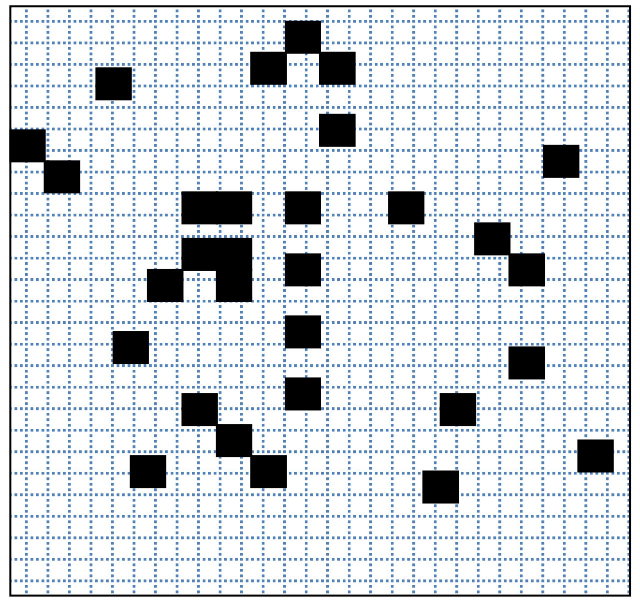
Overlap: similarity between state *S(t)* and pattern

$$m^{\mu} = \frac{1}{N} \sum_{j} p_j^{\mu} S_j$$

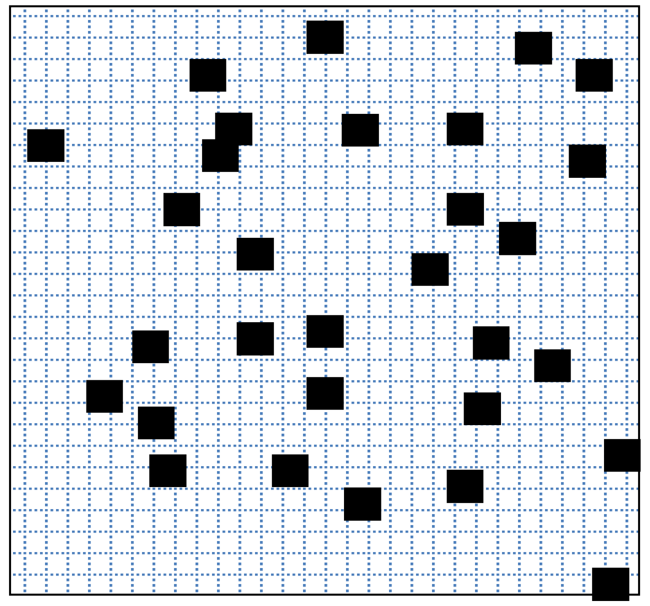


Orthogonal patterns

1. Review of last week: Deterministic Hopfield model



Prototype $\vec{p}^1$



Prototype $\vec{p}^2$

Deterministic dynamics

interactions $$w_{ij} = \frac{1}{N} \sum_{\mu} p_i^{\mu} p_j^{\mu}$$

Sum over all prototypes

Input potential $$h_i = \sum_j w_{ij} S_j$$

Sum over all inputs to neuron *i* prototypes

dynamics $$S_i(t+1) = \text{sgn}[h_i(t)] = \text{sgn}[\sum_j w_{ij} S_j(t)]$$

Similarity measure: Overlap w. pattern 17: $$m^{17}(t+1) = \sum_j p_j^{17} S_j$$

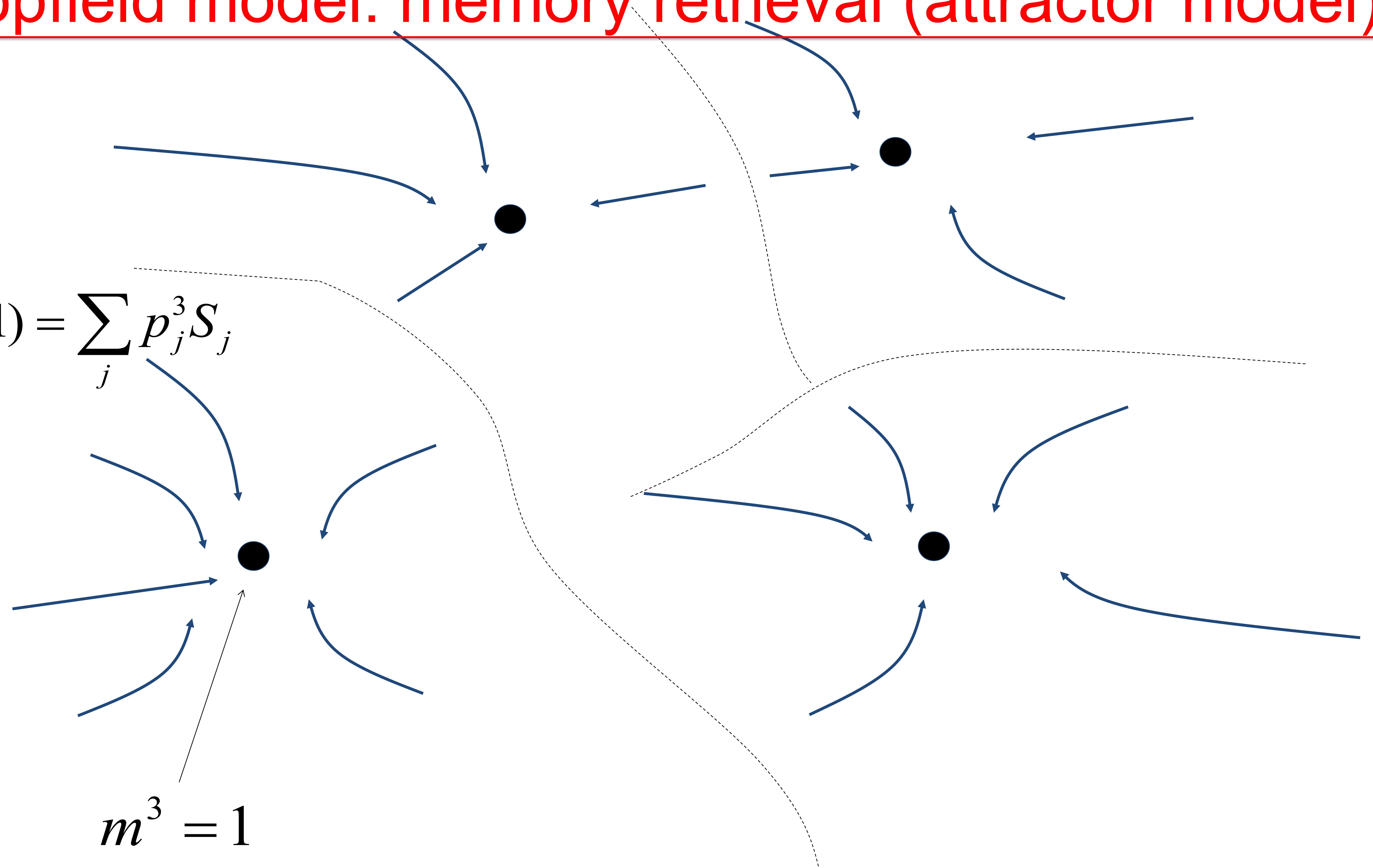
1. Hopfield model: memory retrieval (with overlaps)

$$S_i(t+1) = \mathrm{sgn}[h_i(t)] = \mathrm{sgn}[\sum_j w_{ij} S_j(t)]$$

$$S_i(t+1) = \mathrm{sgn}[\sum_\mu p_i^\mu m_j^\mu(t)]$$

$$m_j^\mu(t+1) \leftarrow \quad m_j^\mu(t)$$

1. Hopfield model: memory retrieval (attractor model)



$$m^3(t+1) = \sum_j p_j^3 S_j$$

$$m^3 = 1$$

1. Hopfield model: memory retrieval (attractor model)

Attractor networks:

dynamics moves network state to a fixed point

Hopfield model:

for a small number of patterns, states with overlap 1 are fixed points

Aim for today:

generalize!

Quiz 1: overlap and attractor dynamics

[ ] The overlap is maximal
if the network state matches one of the patterns.
[ ] The overlap increases during memory retrieval.
[ ] The mutual overlap of orthogonal patterns is one.
[ ] In an attractor memory, the dynamics converges to a stable
fixed point.
[ ] In a perfect attractor memory network, the network state
moves towards one of the patterns.
[ ] In a Hopfield model with N random patterns stored in a
network N neurons, the patterns are attractors.
[ ] In a Hopfield model with 200 random patterns stored in a
network 1000 neurons, all fixed points have overlap one.

Computational Neuroscience: Neuronal Dynamics of Cognition

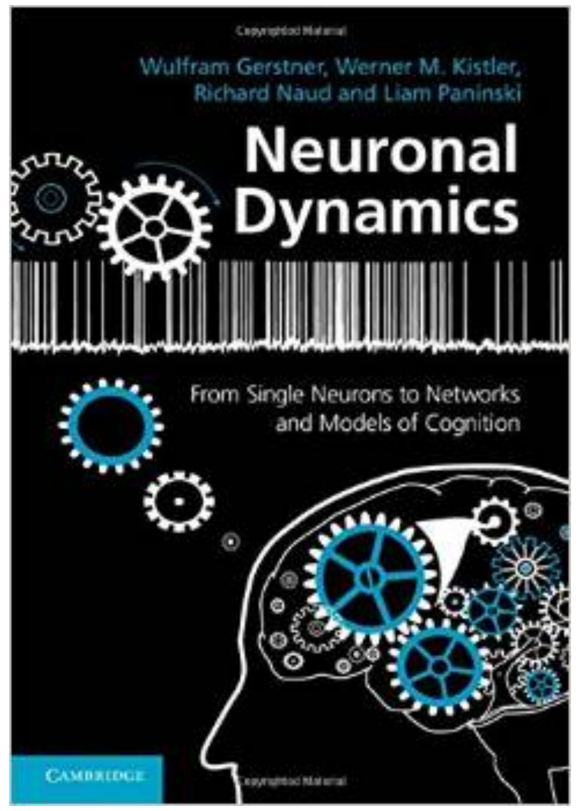
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Attractor Networks and Generalizations of the Hopfield model

Wulfram Gerstner

EPFL, Lausanne, Switzerland

Reading for week 6:
NEURONAL DYNAMICS
- Ch. 17.2.5 – 17.4



Cambridge Univ. Press

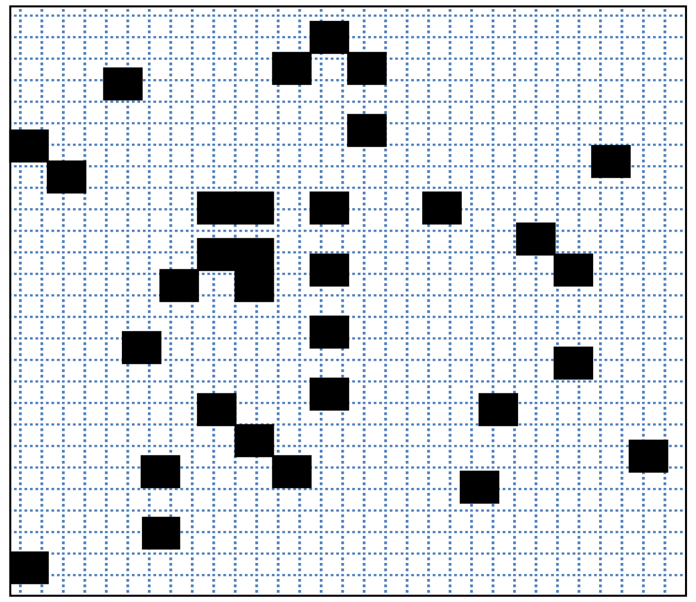
1. Attractor networks
2. **Stochastic Hopfield model**
3. Energy landscape
4. Towards biology (1)
   - low-activity patterns
5. Towards biology (2)
   - spiking neurons

2. Stochastic Hopfield model

Neurons may be noisy:

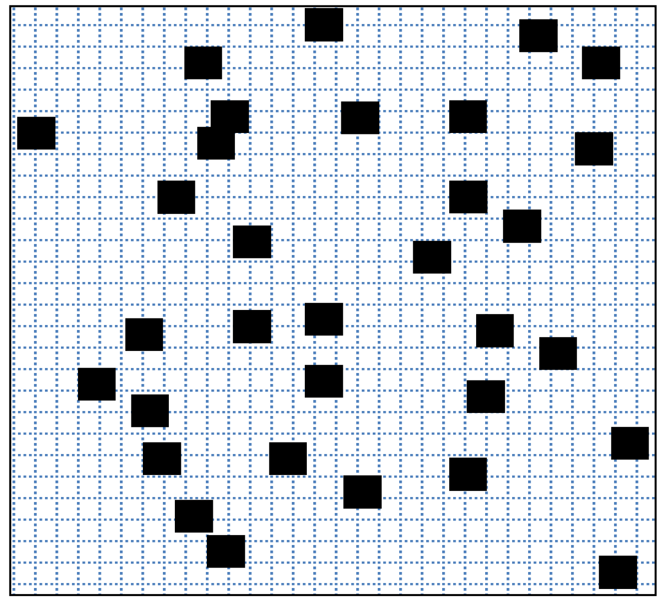
What does this mean for attractor dynamics?

2. Stochastic Hopfield model



Prototype

$\vec{p}^1$



Prototype

$\vec{p}^2$

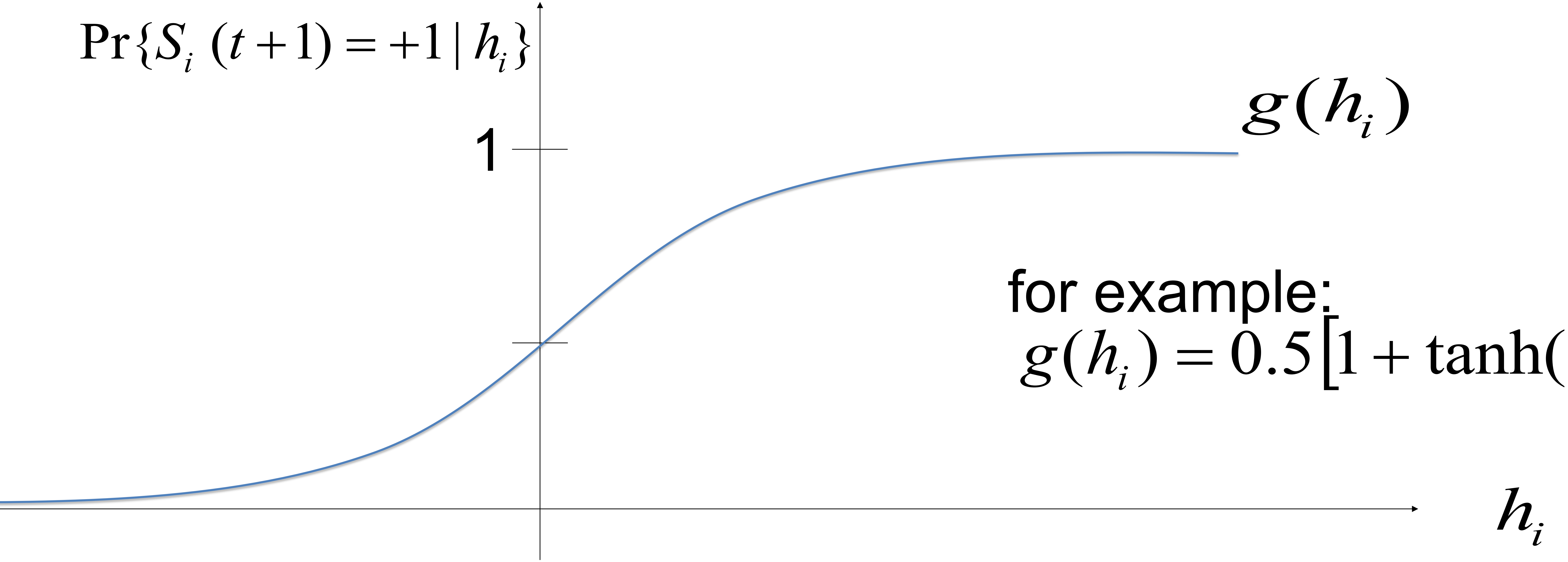
Random patterns

Interactions (1) $w_{ij} = \frac{1}{N} \sum_{\mu} p_i^{\mu} p_j^{\mu}$

Dynamics (2)

$$\Pr\{S_i(t+1) = +1 \mid h_i\} = g[h_i] = g\left[\sum_j w_{ij} S_j(t)\right]$$

2. Stochastic Hopfield model: firing probability



for example:
$$g(h_i) = 0.5\left[1 + \tanh(2h)\right]$$

$$\Pr\{S_i\,(t+1)=+1\,|\,h_i\} = g[h_i] = g\left[\sum_j w_{ij} S_j(t)\right] = g\left[\sum_\mu p_i^\mu m^\mu(t)\right]$$

2. Stochastic Hopfield model

Dynamics (2)

$$\Pr\{S_i(t+1)=+1 \mid h_i\} = g[h_i] = g\left[\sum_j w_{ij} S_j(t)\right]$$

$$\Pr\{S_i(t+1)=+1 \mid h_i\} = g\left[\sum_\mu p_i^\mu m^\mu(t)\right]$$

Assume that there is **only** overlap with pattern 17:

two groups of neurons: those that should be ‘on’ and ‘off’

2. Stochastic Hopfield model

Dynamics (2)

$$\Pr\{S_i(t+1) = +1 \mid h_i\} = g[h_i] = g\left[\sum_j w_{ij} S_j(t)\right]$$

$$\Pr\{S_i(t+1) = +1 \mid h_i\} = g\left[\sum_\mu p_i^\mu m^\mu(t)\right]$$

Assume that there is only overlap with pattern 17:

two groups of neurons: those that should be 'on' and 'off'

$$\Pr\{S_i(t+1) = +1 \mid h_i = h^+\} = g\left[m^{17}(t)\right]$$

$$\Pr\{S_i(t+1) = +1 \mid h_i = h^-\} = g\left[-m^{17}(t)\right]$$

Overlap (definition) $\quad m^{17}(t+1) = \sum_j p_j^{17} S_j$

2. Stochastic Hopfield model

Overlap (definition) $$m^{17}(t+1) = \frac{1}{N}\sum_{i=1}^{N} p_j^{17} S_j(t+1)$$

Suppose initial overlap with pattern 17 is 0.4;
Find equation for overlap at time *(t+1)*,
given overlap at time *(t)*.
Assume overlap with other patterns stays zero.

Hint: Use result from previous slide and consider 4 groups of neurons

- Those that should be ON and are ON
- Those that should be ON and are OFF
- Those that should be OFF and are ON
- Those that should be OFF and are OFF

2. Stochastic Hopfield model

Overlap $$m^{17}(t+1) = \frac{1}{N}\sum_{i=1}^{N} p_j^{17} S_j(t+1)$$

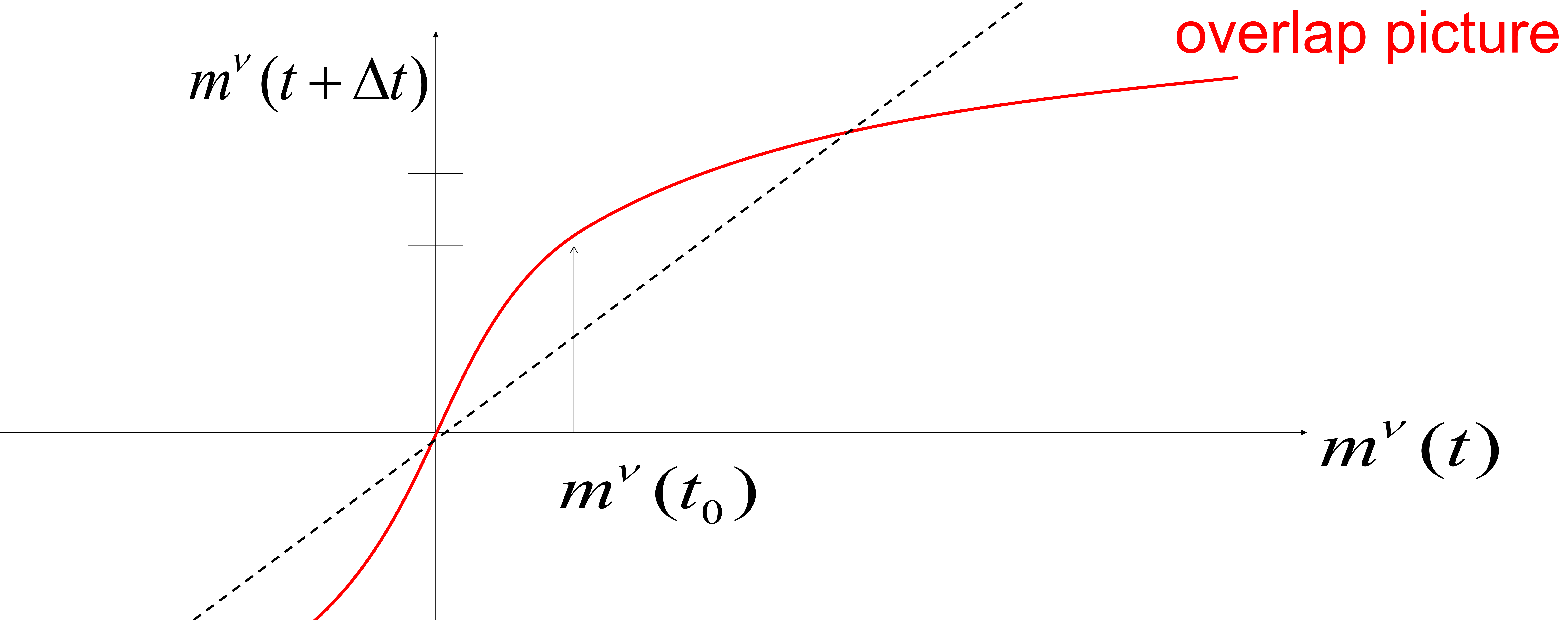
2. Stochastic Hopfield model: memory retrieval

Overlap:

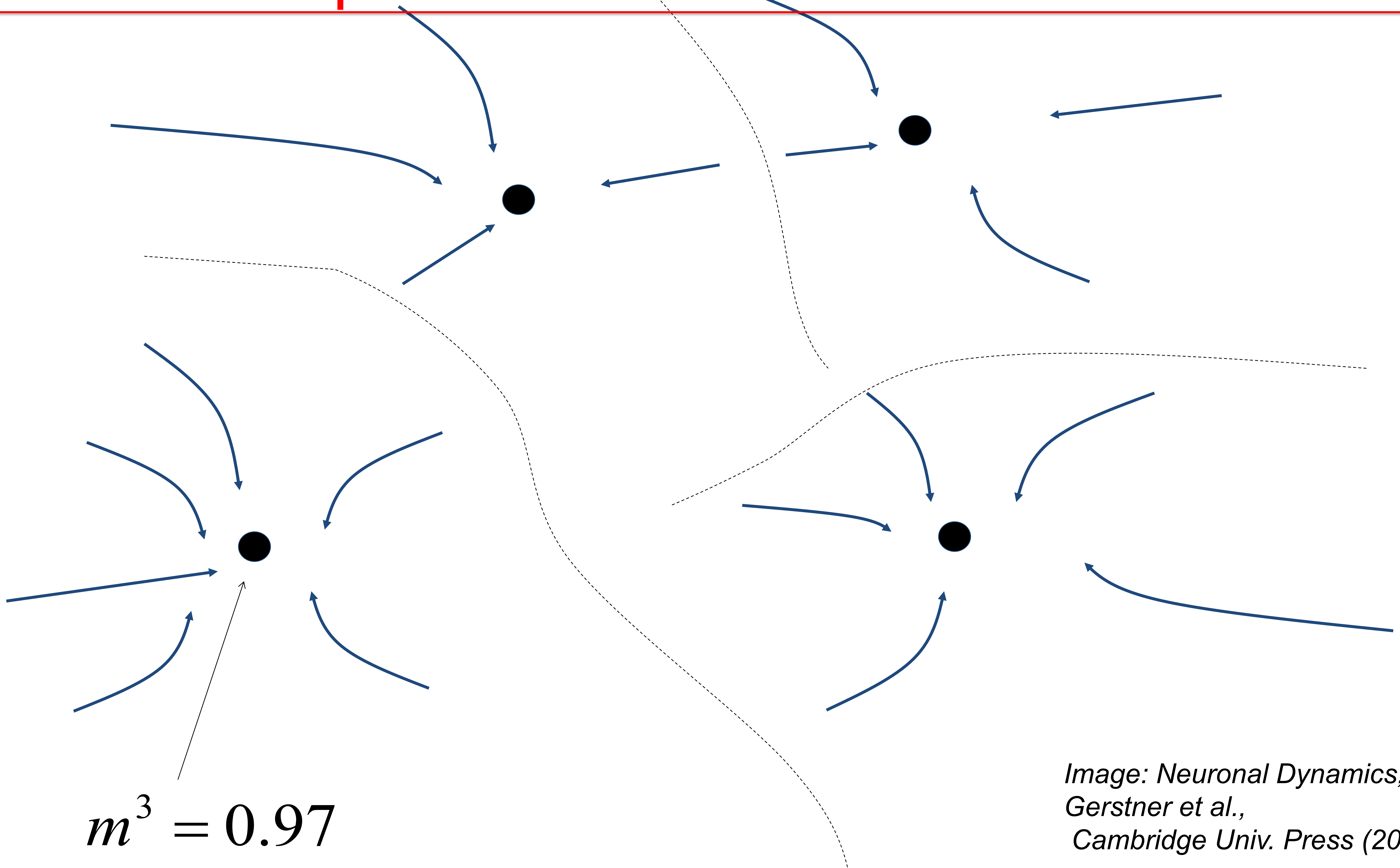
Neurons that should be 'on' Neurons that should be 'off'

$$2m^{17}(t+1) = g\left[m^{17}(t)\right] - \{1 - g\left[m^{17}(t)\right]\} - g\left[-m^{17}(t)\right] + \{1 - g\left[-m^{17}(t)\right]\}$$

$$m^{17}(t+1) = \tilde{F}\left[m^{17}(t)\right]$$



2. Stochastic Hopfield model = attractor model



$$m^3 = 0.97$$

Image: Neuronal Dynamics, Gerstner et al., Cambridge Univ. Press (2014),

2. Stochastic Hopfield model: memory retrieval

- Memory retrieval possible with stochastic dynamics

- Fixed point at value with large overlap (e.g., 0.95)

- Need to check that overlap of other patterns remains small

- Random patterns: nearly orthogonal but ‘noise’ term

Quiz 2: Stochastic networks and overlap equations

[ ] The update of the overlap leads always to a fixed point with overlap m=1

[ ] The update equation as derived here implicitly assumed **orthogonal** patterns because otherwise we would have to analyze overlaps with several patterns in **parallel**

[ ] The update equation as derived here requires a function

$$g(h_i) = 0.5\left[1 + \tanh(2h)\right]$$

Computational Neuroscience: Neuronal Dynamics of Cognition

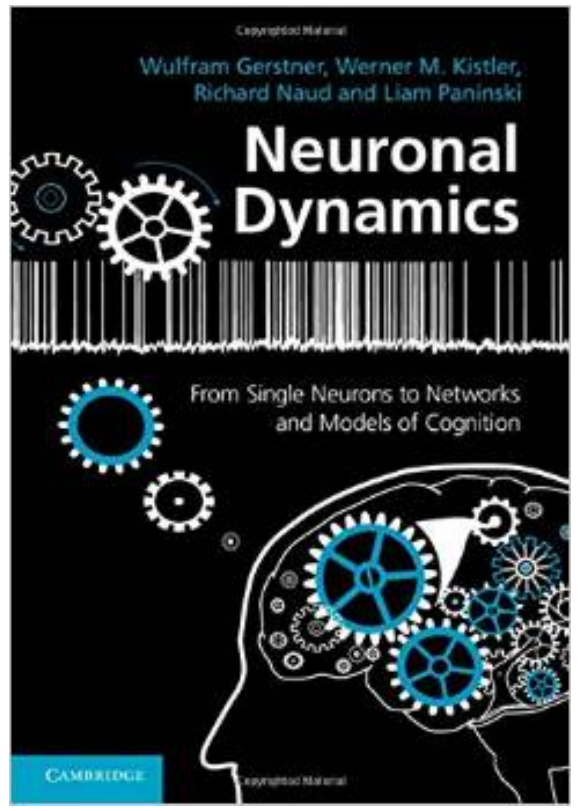
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Attractor Networks and Generalizations of the Hopfield model

Wulfram Gerstner

EPFL, Lausanne, Switzerland

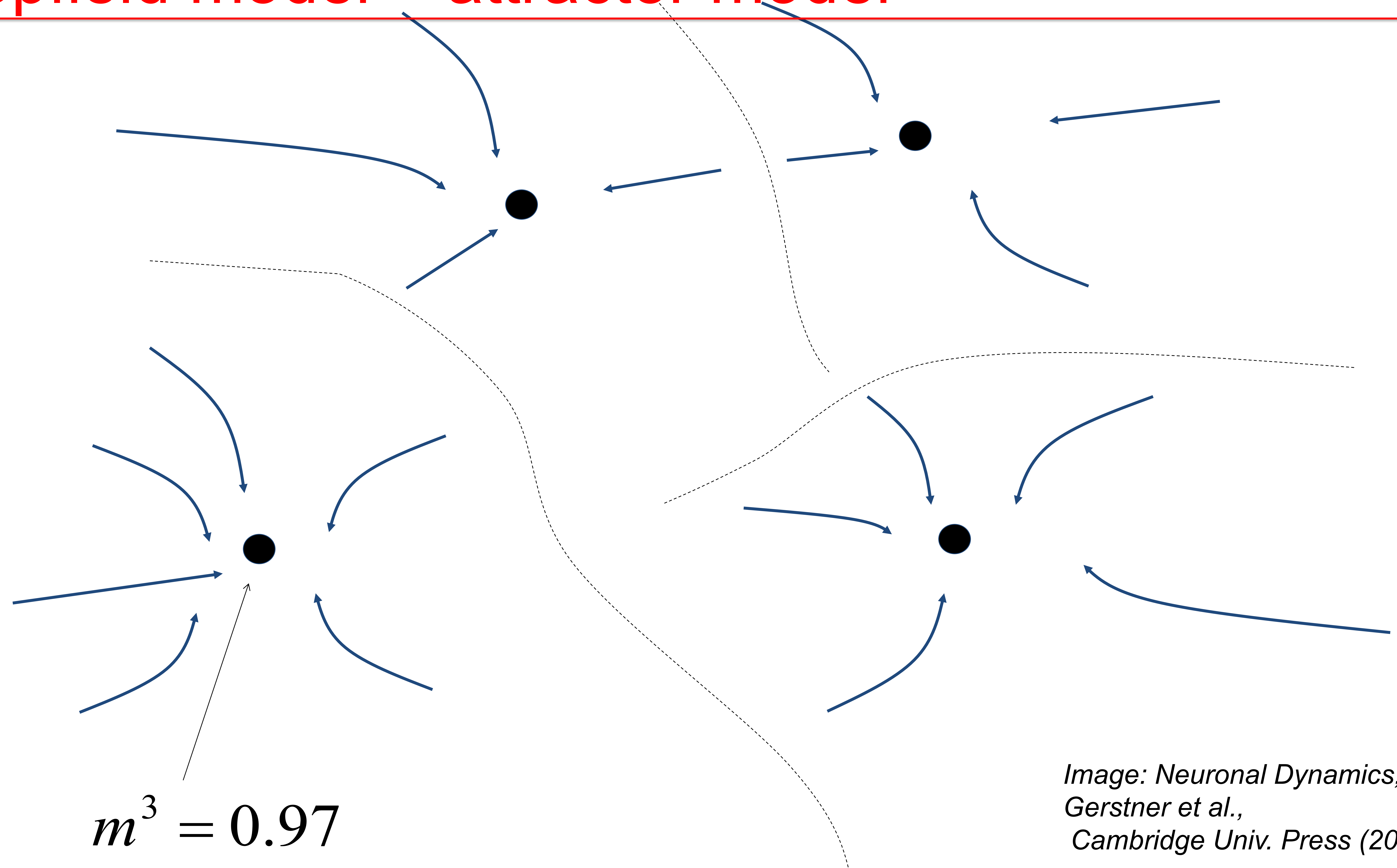
Reading for week 6:
NEURONAL DYNAMICS
- Ch. 17.2.5 – 17.4



Cambridge Univ. Press

1. Attractor networks
2. Stochastic Hopfield model
3. **Energy landscape**
4. Towards biology (1)
   - low-activity patterns
5. Towards biology (2)
   - spiking neurons

3. Hopfield model = attractor model

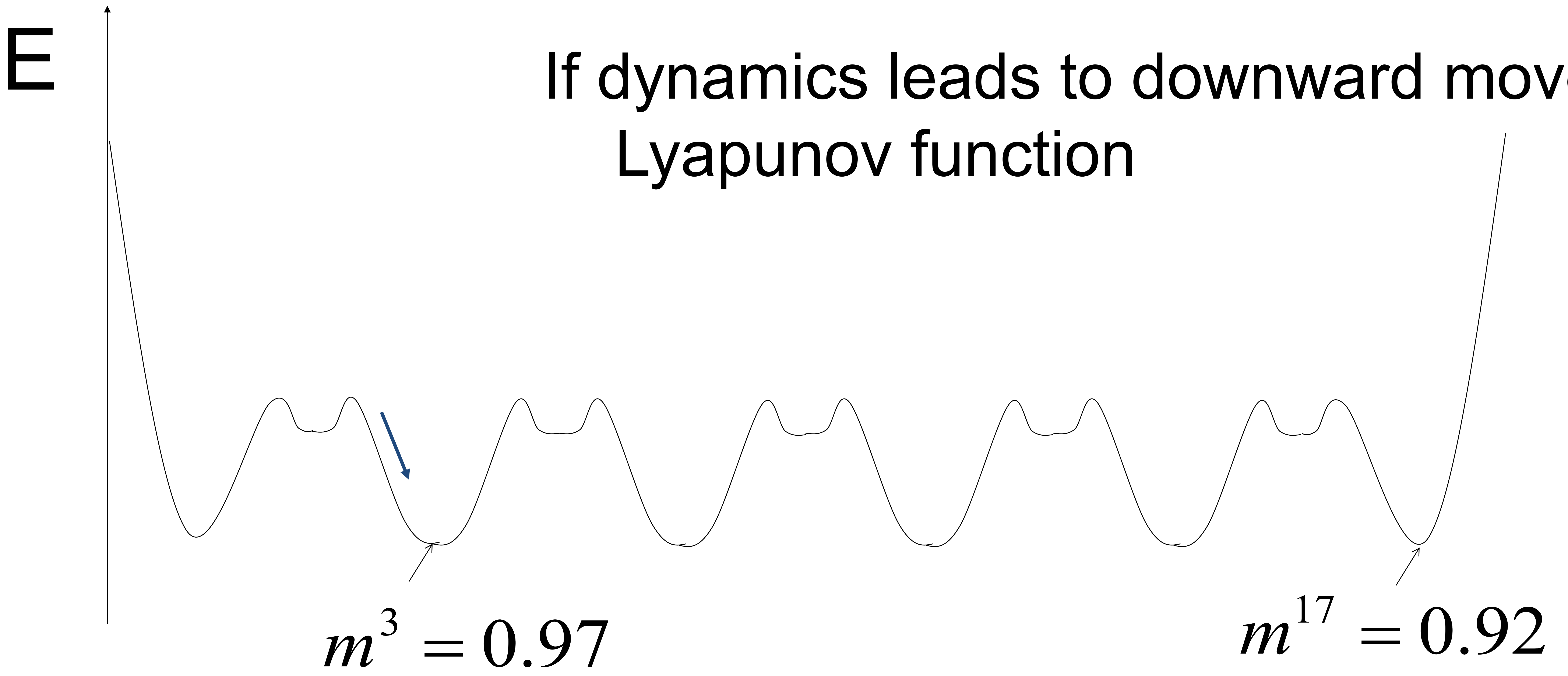


$$m^3 = 0.97$$

Image: Neuronal Dynamics, Gerstner et al., Cambridge Univ. Press (2014),

3. Symmetric interactions: Energy picture

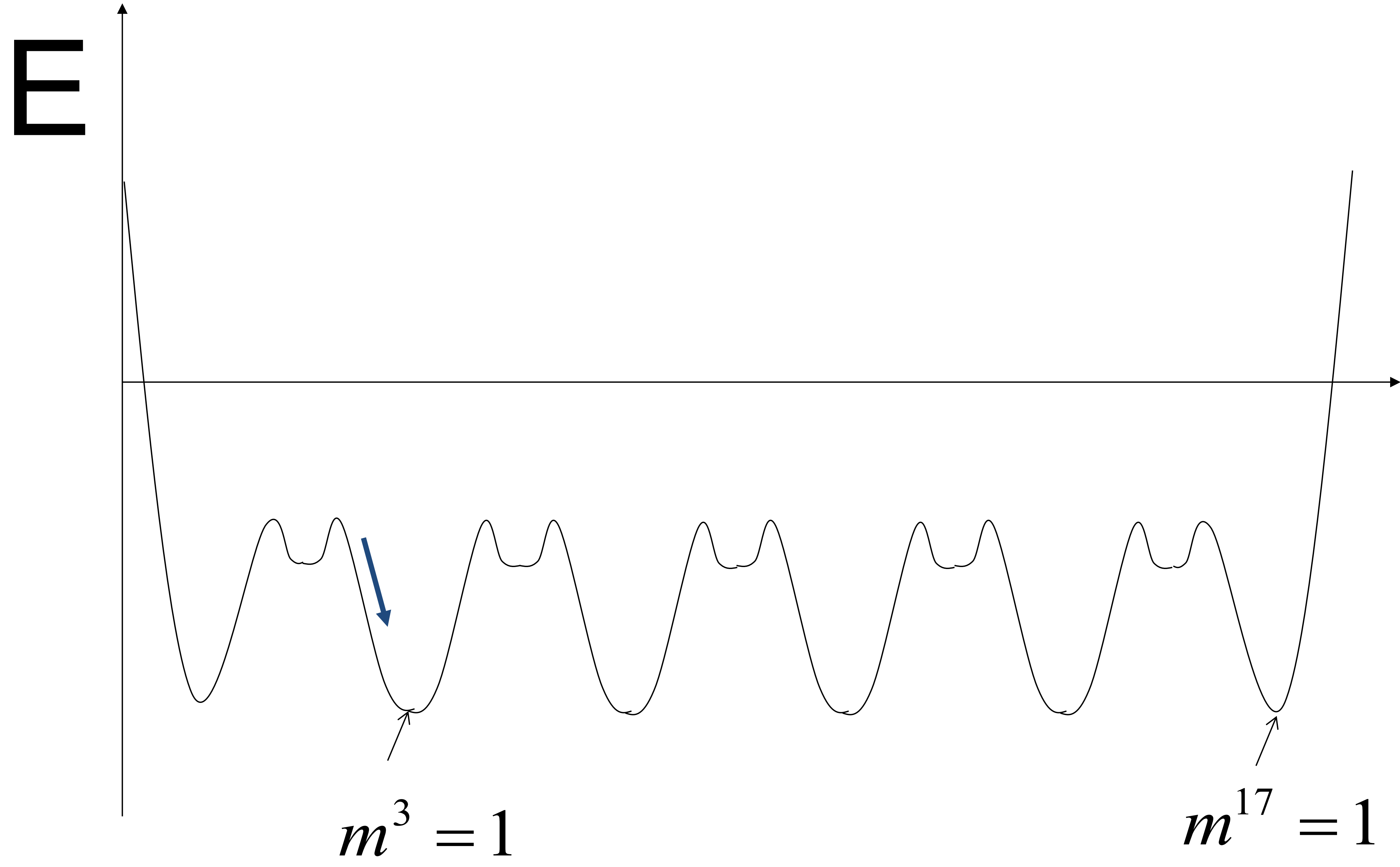
If dynamics leads to downward movement:
Lyapunov function



3. Symmetric interactions: Energy picture

$$E = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j$$

- Rewrite in terms of overlaps
- Random patterns vs. orthogonal patterns
- Random state vs. overlap state



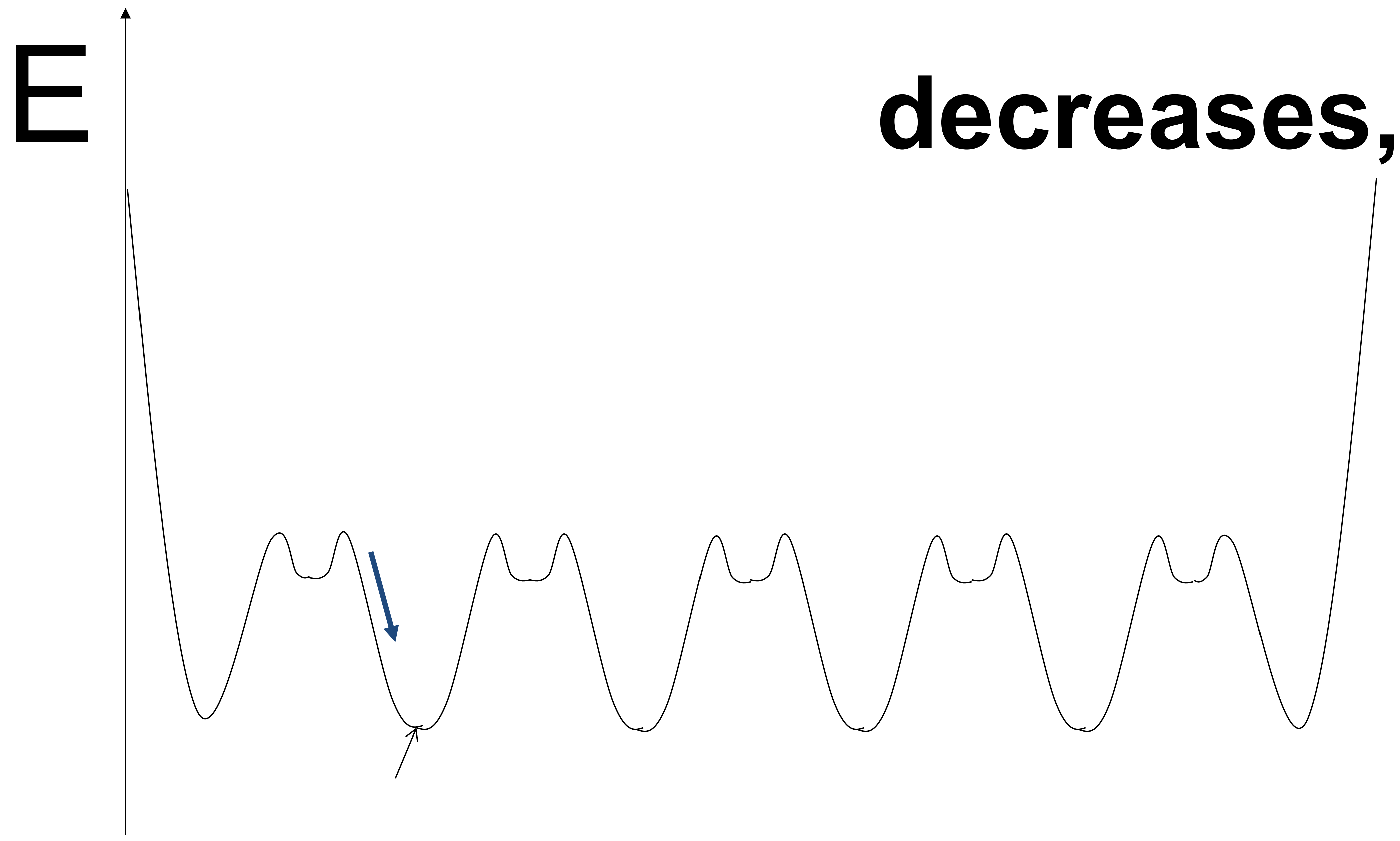
3. Symmetric interactions: Energy/Lyapunov function

Assume symmetric interaction,
Assume deterministic asynchronous update

$$S_i(t+1) = \mathrm{sgn}[h_i(t)] = \mathrm{sgn}[\sum_j w_{ij} S_j(t)]$$

Claim: the energy $E = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j$

decreases, if neuron *k* changes



J.J. Hopfield (1982) Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA 79, pp. 2554–2558

3. Symmetric interactions: Energy/Lyapunov function

$$E = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j$$

Assume symmetric interaction,

Assume deterministic asynchronous update

$$S_i(t+1) = \mathrm{sgn}[h_i(t)] = \mathrm{sgn}[\sum_j w_{ij} S_j(t)]$$

Claim:

energy decreases, if neuron *k* changes

3. Energy picture

energy picture historically important:

- capacity calculations

J.J. Hopfield (1982) Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA 79, pp. 2554–2558

D.J. Amit, H. Gutfreund and H. Sompolinsky (1987) Information storage in neural networks with low levels of activity. Phys. Rev. A 35, pp. 2293–2303.

energy picture is a side-track:

- it needs symmetric interactions

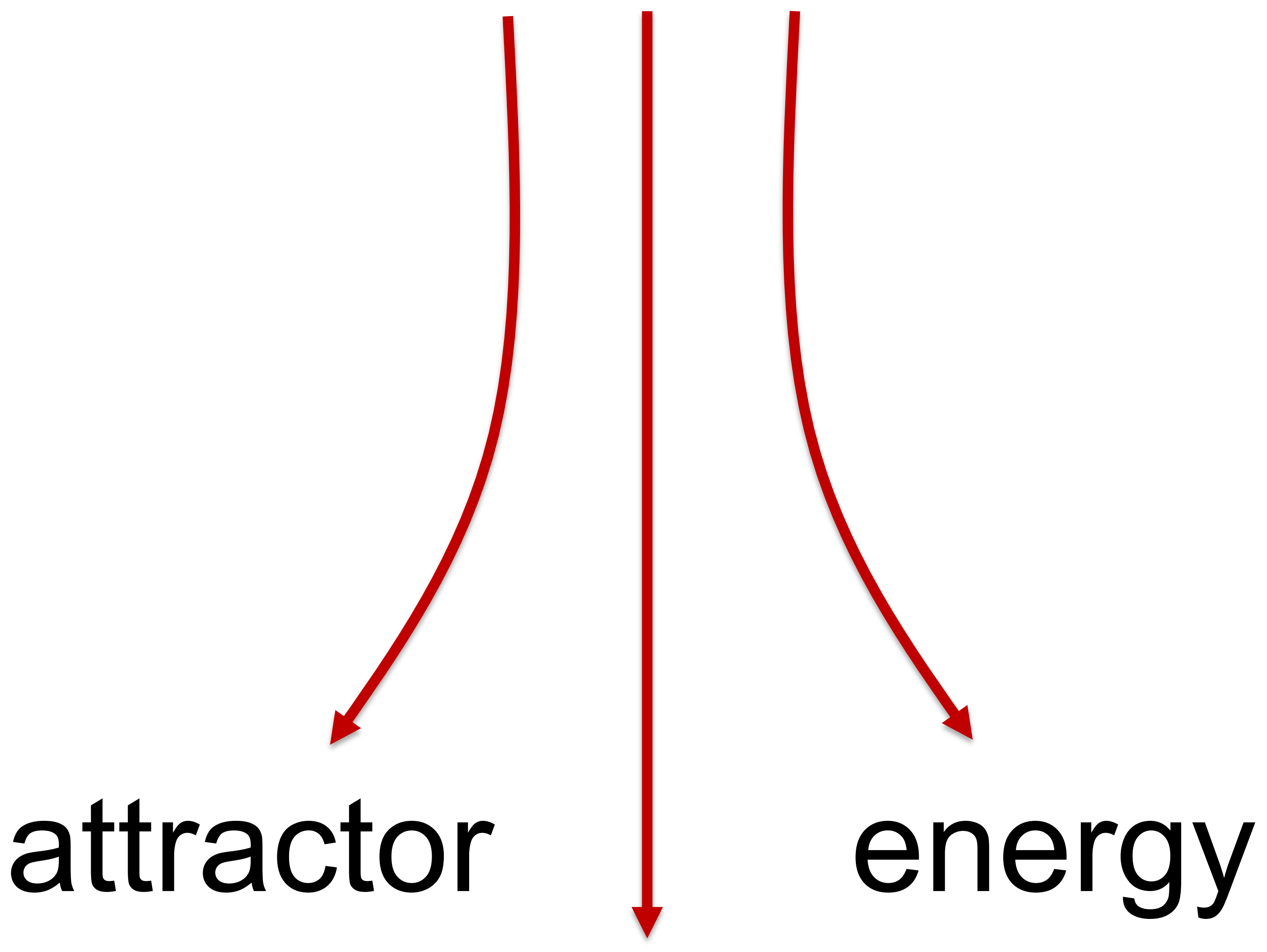
energy picture is very general:

- it shows that it should be possible to learn other patterns than mean-zero random patterns

3. Energy picture

Hopfield model

special case



biology

(asymmetric interactions)

Quiz 3: Energy picture and Lyapunov function

Let $E = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j$ be the energy of the Hopfield model

and $S_i(t+1) = \mathrm{sgn}[h_i(t)] = \mathrm{sgn}[\sum_j w_{ij} S_j(t)]$ the dynamics.

[ ] The energy picture requires random patterns with prob = 0.5
[ ] The energy picture requires symmetric weights
[ ] It follows from the energy picture of the Hopfield model that the only fixed points are those where the overlap is exactly one
[ ] In each step, the value of a Lyapunov function decreases or stays constant
[ ] Under deterministic dynamics the above energy is a Lyapunov function