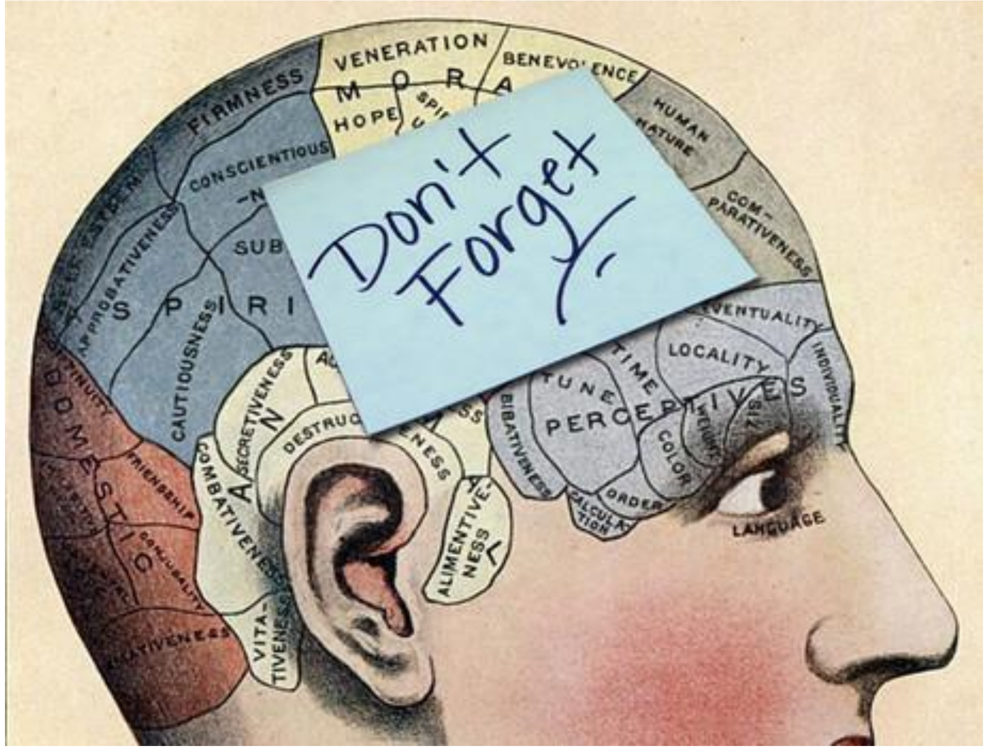
Lecture 6 – Spiking Neural Networks Part 1

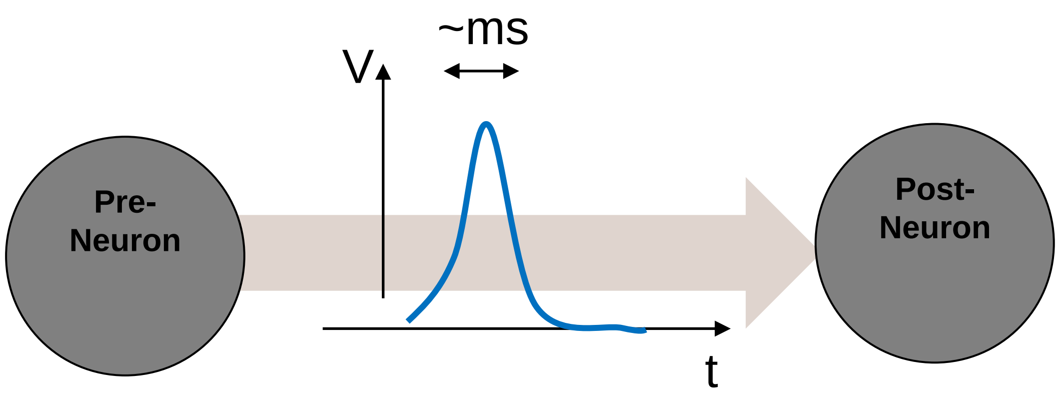
Outline – Lecture 6

- Structure of SNN's
- Spike transmission between neurons
- Learning rules
- Teaching SNN hardware



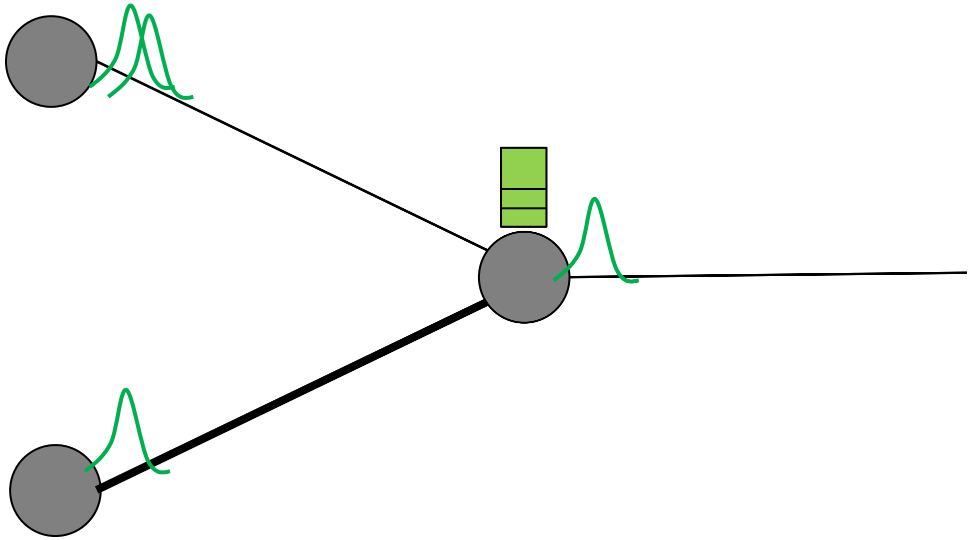
Spiking neural networks (SNNs)

- Based on biological process in brain (neuromorphic)
- Information transmitted between neurons via voltage spikes
- **Pre-neuron:** neuron that sends spikes into a synapse
- **Post-neuron:** neuron that receives spikes from a synapse
- **Activity:** The average rate of spikes firing from a neuron



Important features of SNNs
- Analog nature of synapse weights
- Nonlinearity of synaptic response
- Temporal correlations important

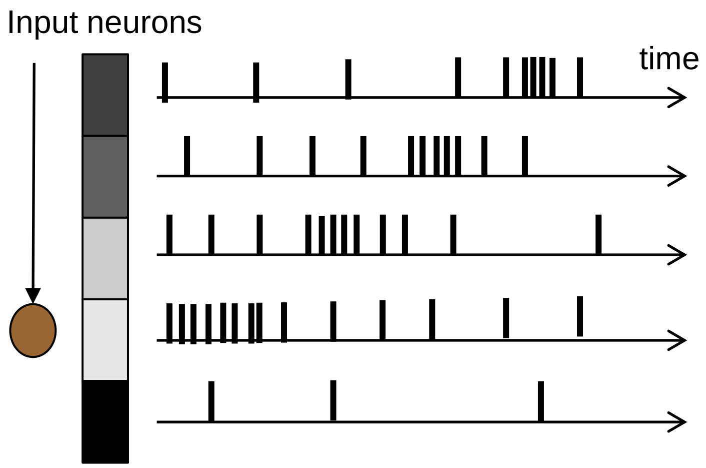
Example of SNN functionality



Encoding schemes

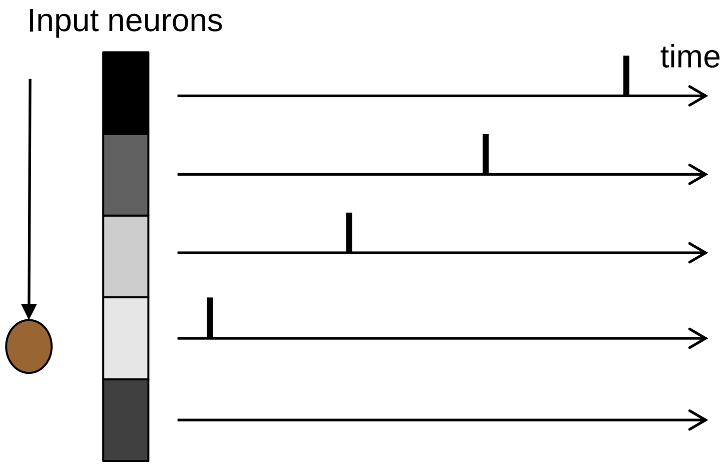
Example: object moving in front of input sensors

Rate-encoding scheme



”Stronger input, more spiking”

Time-encoding scheme



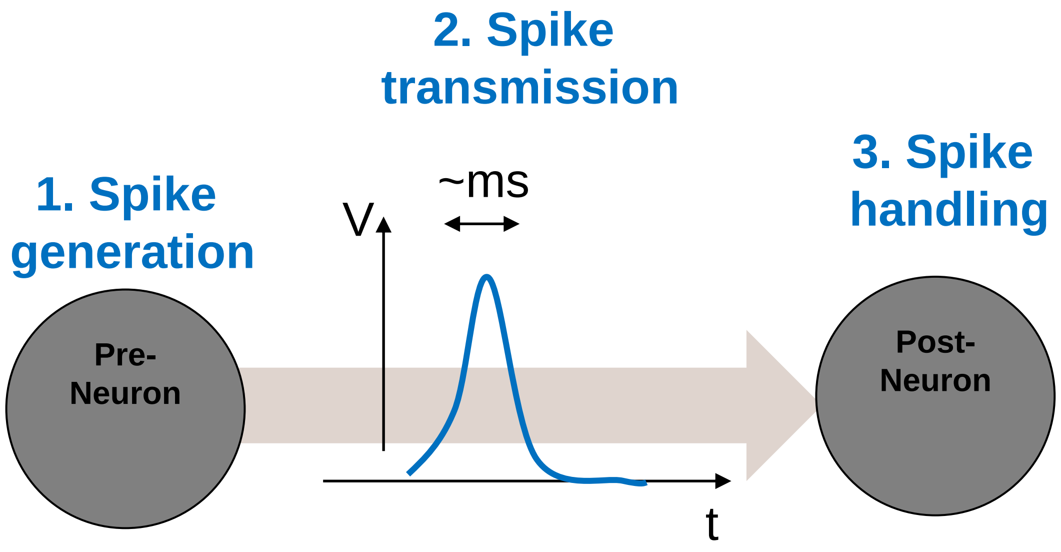
”Only generate spikes on strong input”

Information contained in spiking frequency
Many spikes if strong neuron activation

Information contained in timing of spikes
Sparser spikes possible

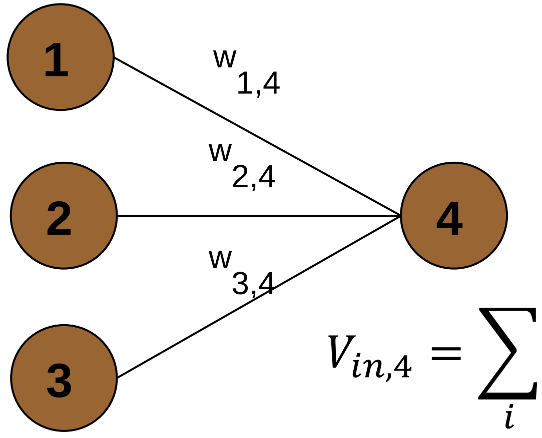
Information processing in SNNs

- 1. Spike generation at neuron (**NEXT TIME**)
- 2. Spike transmission between neurons (**TODAY**)
- 3. Spike handling among neurons (**NEXT TIME**)

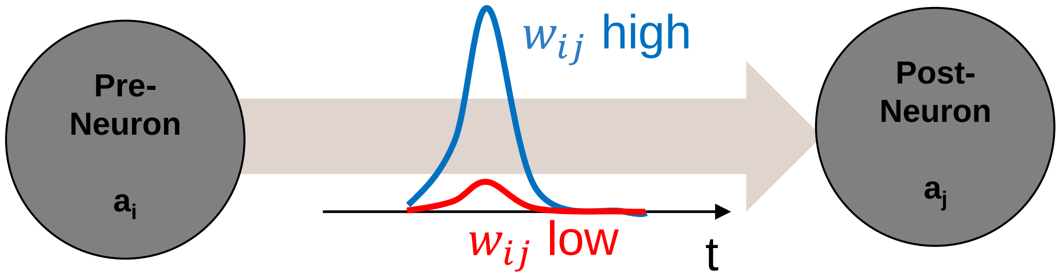


Spike transmission

- Signal attenuation will depend on the ”weight” of the synapse
- The weight can be changed → ***plasticity***
- Strengthening of weight → ***potentiation***
- Weakening of weight → ***depression***



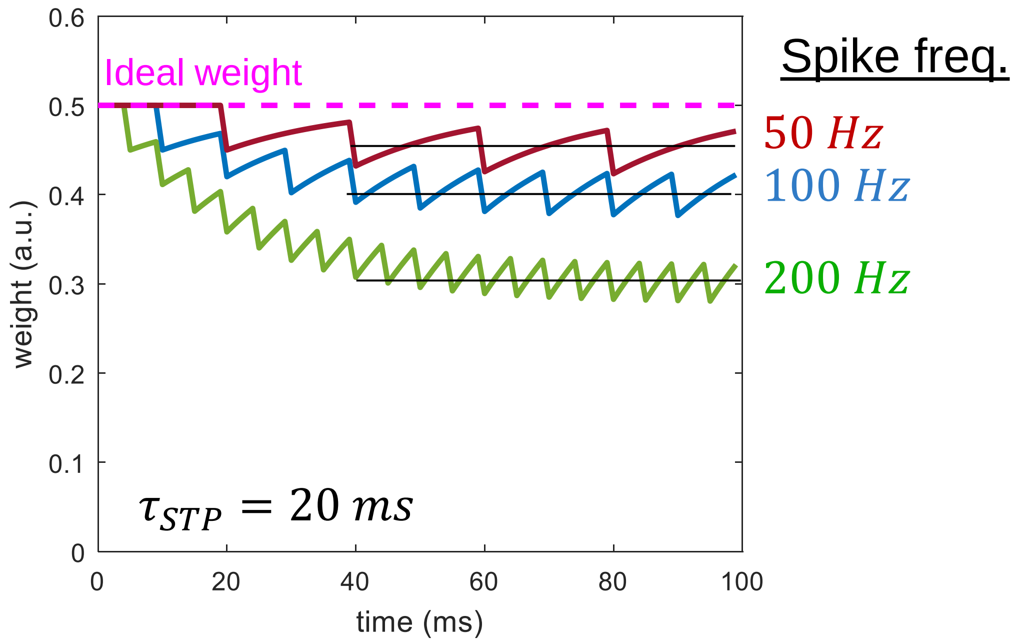
$$V_{in,4} = \sum_i w_{i,4} V_{out,i}$$



Short-Term Plasticity

- **Short-Term Potentiation** (STP)
- **Short-Term Depression** (STD)

- Temporary increase (STP) or decrease (STD) of synapse weight away from equilibrium weight
- Time scale of recovery ~ 10-100's of ms
- Results in dynamic synapses
  → Counteracts temporary changes in transmission activity
- Ex: STD: Strong activity decreases effective weight



Long-Term Plasticity (LTP)

- **Long Term Potentiation** (LTP)
- **Long Term Depression** (LTD)

- ”Permanent weight change”

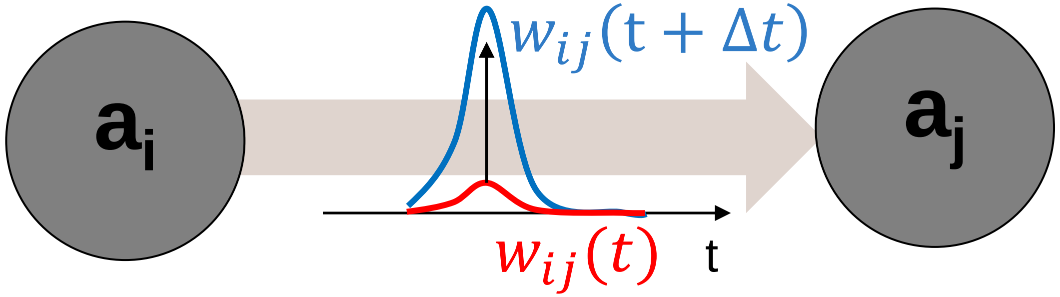
→ time scale is ~ years

- Facilitates learning/training like in ANNs

Hebbian learning

- Long lasting form of synaptic modifications that are:
  - synapse-specific
  - Depend on correlation between pre- and postsynaptic activity
- $\frac{dw_{ij}}{dt} = \eta(a_i * a_j)$, *"who fires together, wire together"*

Learning rate



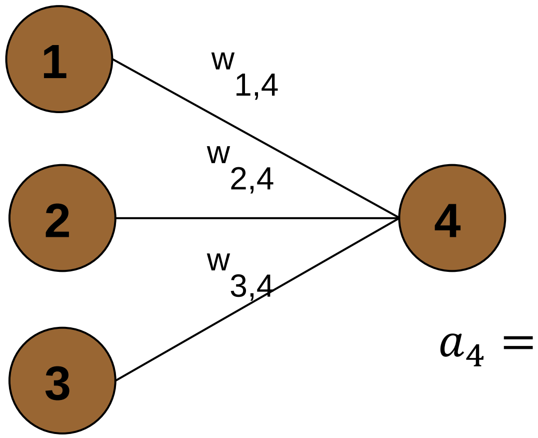
Abbott & Nelson Nat. Neuro 2000

2 min exercise - Hebb

$$\frac{dw_{ij}}{dt} = \eta\left(a_i * a_j\right)$$ Hebb's rule applies

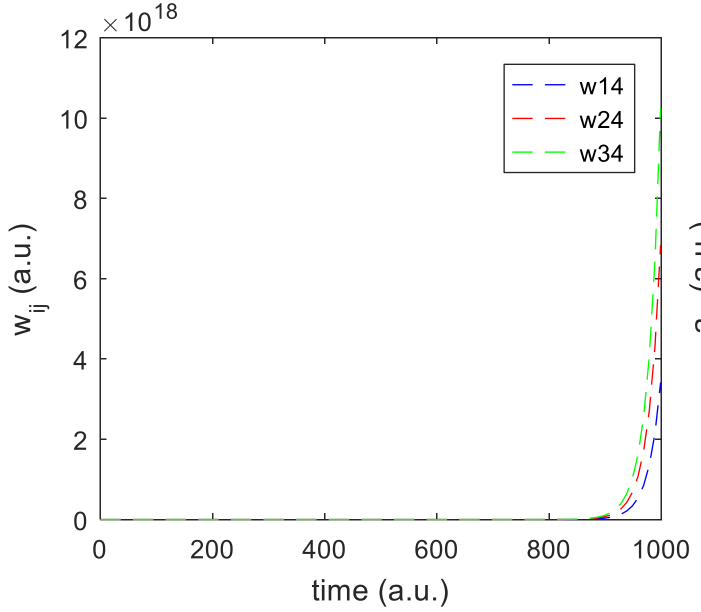
1. Which *weight* will grow fastest?
2. As $t \to \infty$, what will happen to a_4?

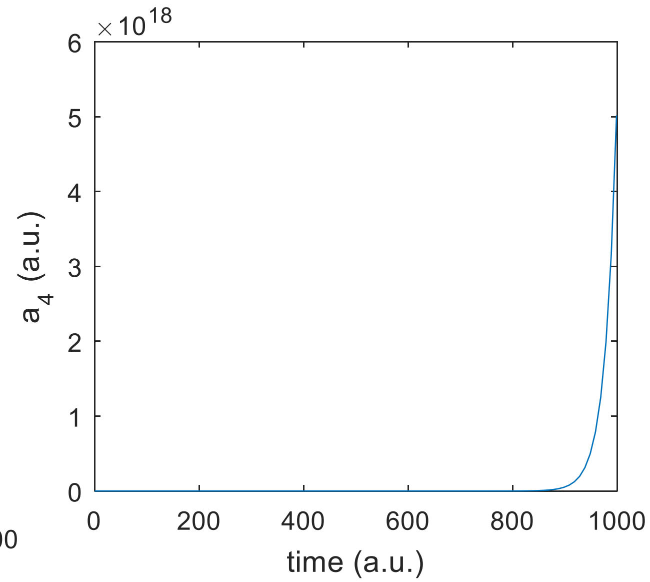
w(t=0) = 0.1



$$a_4 = \frac{\sum_{i=1}^{3} w_{i4} a_i}{\sum_{i=1}^{3} a_i}$$

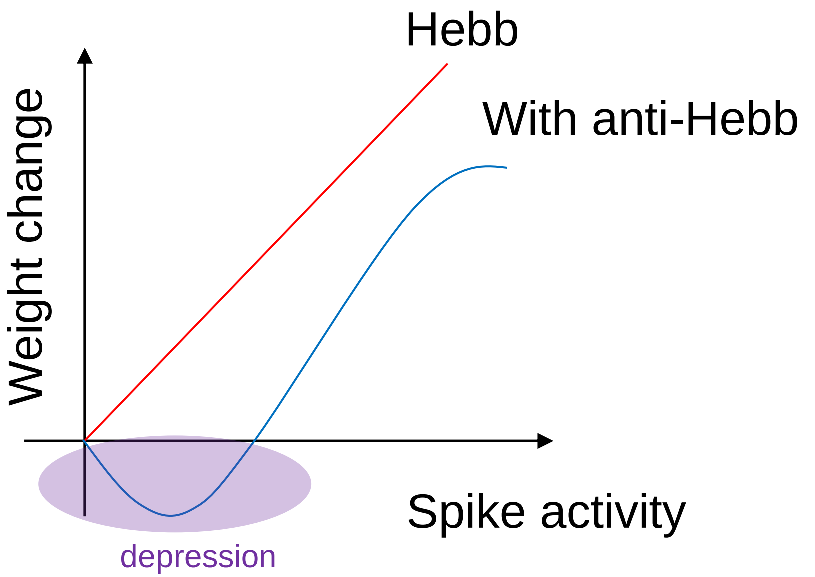
a_1 = 0.05
a_2 = 0.10
a_3 = 0.15





Amendments to Hebbian learning

- **Hebb's rule: only Potentiation (LTP) →** Risk of exploding activity level in network
- Anti-Hebb rule adds depression (LTD)
- Can stabilize weights…



$$\frac{dw_{ij}}{dt} = \underbrace{\eta_1(a_i a_j)}_{\text{Hebb}} - \underbrace{f(a_i, a_j)}_{\text{Anti-Hebb}}$$

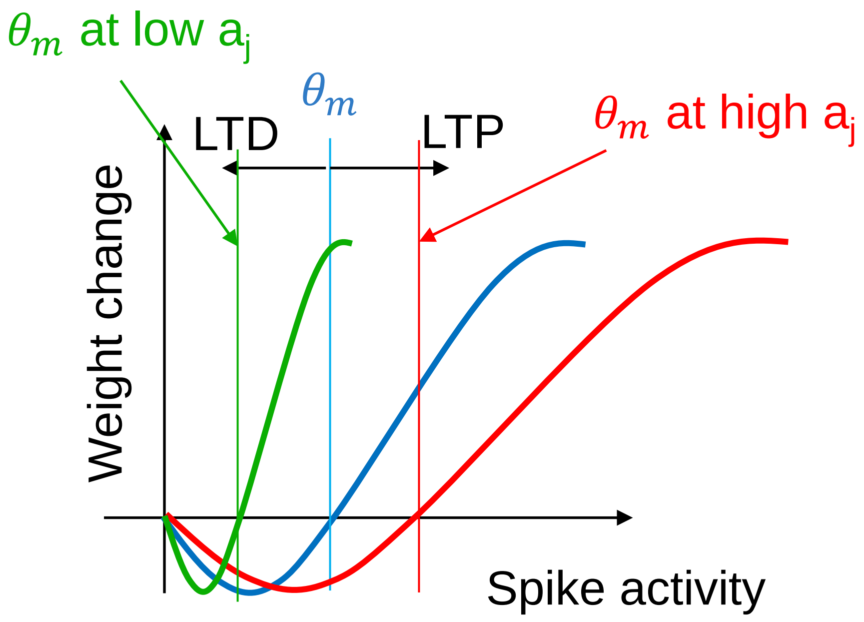
Bienenstock-Cooper-Munro (BCM rule)

- Sliding threshold between Hebb and Anti-Hebb.
- Threshold increased by strong post-neuron activity
- Threshold decreased by weak post-neuron activity

$$\frac{dw_{ij}}{dt} = \varphi(a_j(t))a_i$$

$$\varphi(a_j) < 0, \qquad \forall\, a_j < \theta_m$$

$$\varphi(a_j) > 0, \qquad \forall\, a_j > \theta_m$$



Journal of neuroscience 2, p32, (1982)

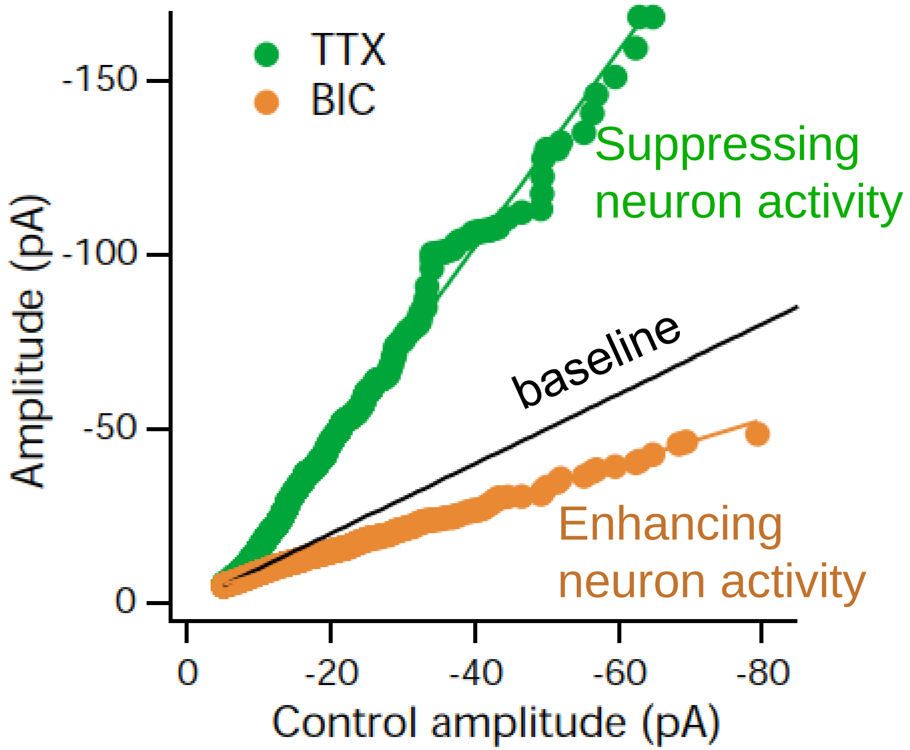
Synaptic scaling

- Biologically present (neucortical, hippocampal and spinal-cord)
- Blocking post-neuron activity increases synaptic weights!
- Enhancing post-neuron activity suppresses synaptic weights

→ post-neuron activity-based negative term

$$\frac{dw_{ij}}{dt} = \eta_1(a_i a_j) - w_{ij} a_j^2$$

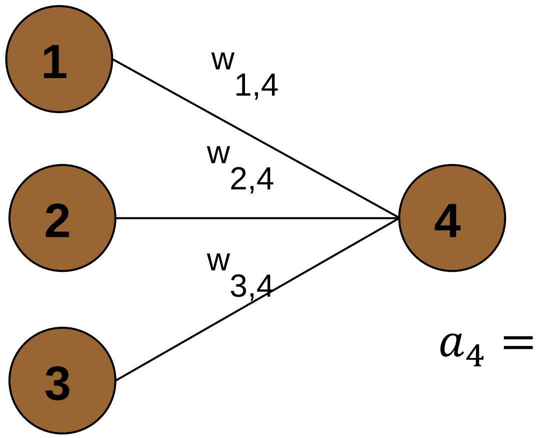
Oja's rule



Abbott & Nelson Nat. Neuro 2000

2 min exercise - Oja

w(t=0) = 0.1

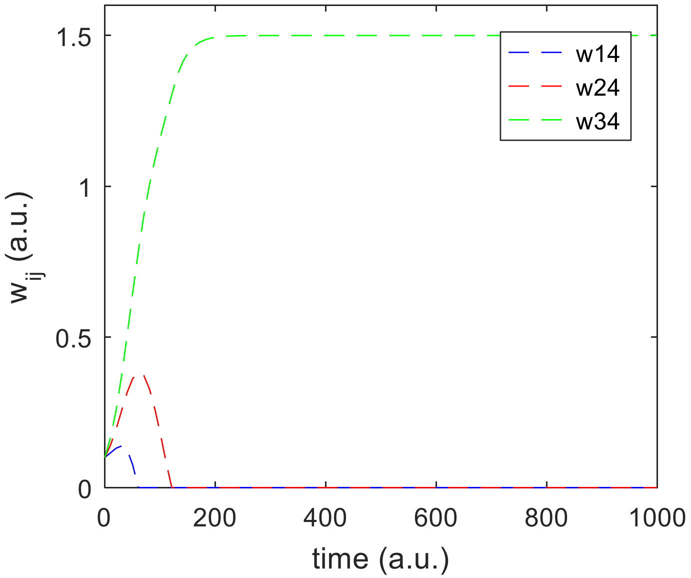


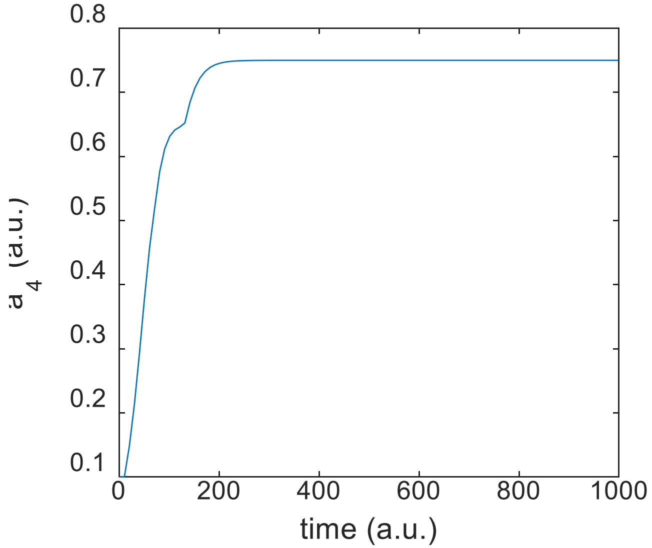
$$a_4 = \frac{\sum_{i=1}^{3} w_{i4} a_i}{\sum_{i=1}^{3} a_i}$$

$a_1 = 0.5$
$a_2 = 0.10$
$a_3 = 0.15$

$$\frac{d\boldsymbol{w_{ij}}}{d\boldsymbol{t}} = \boldsymbol{\eta_1(a_i a_j) - w_{ij} a_j^2}$$ Oja's rule applies

1. Describe what will happen to the weights as time passes.
2. As $t \rightarrow \infty$, what will happen to a_4?

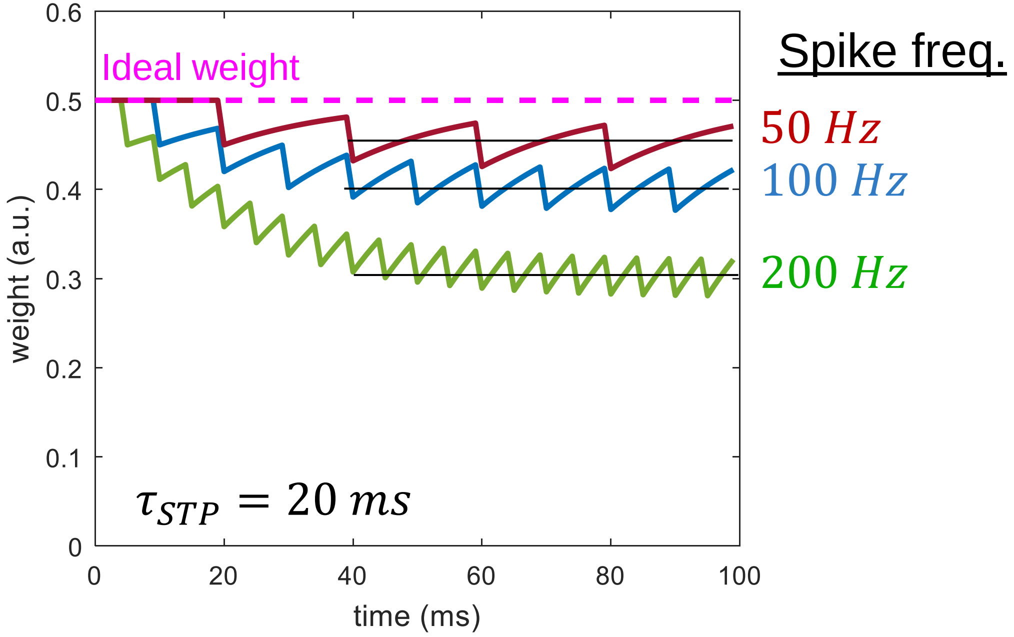




Synaptic realignment

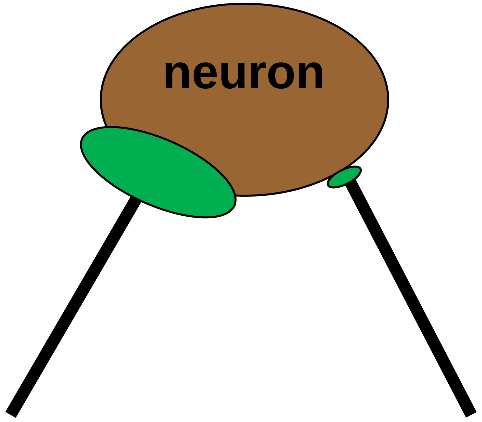
- STD (short-term depression) decreases steady-state weight

- STD does not affect transient spiking
- Can help to enhance response to transients
  - Spontaneous bursts of activity allowed
  - Prolonged high activity suppressed
- Recurrent networks possible without loss of stability



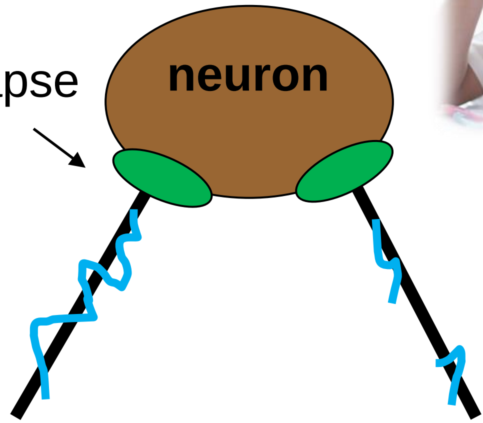
Synaptic shrinking

- Wake: Synaptic patterns are learned (strenghtened)
- During sleep synapses are weakened in the brain
  - Consolidates learning → decreases overall brain activity
- **Weak synapses are ”shrinking” more than strong.**



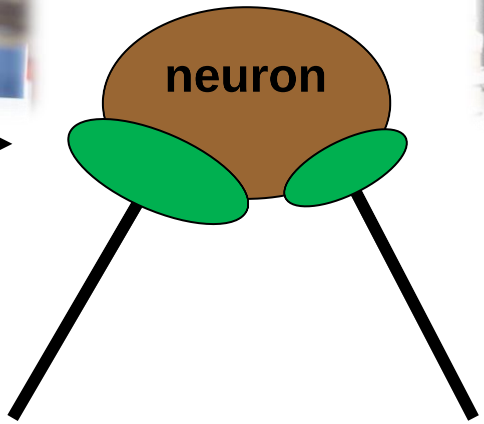
several days

synapse



High activity

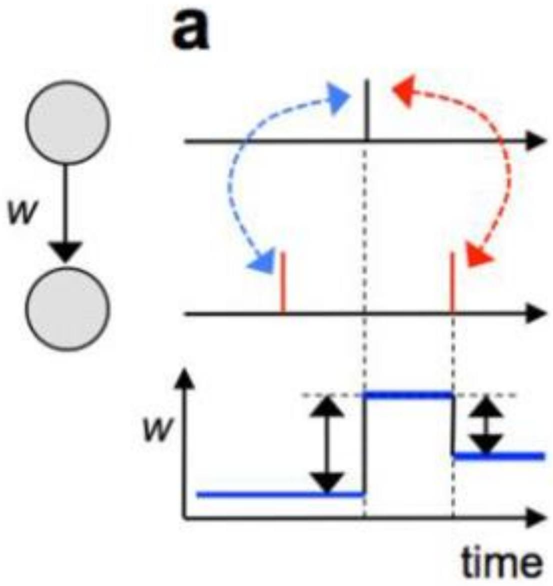
Low activity

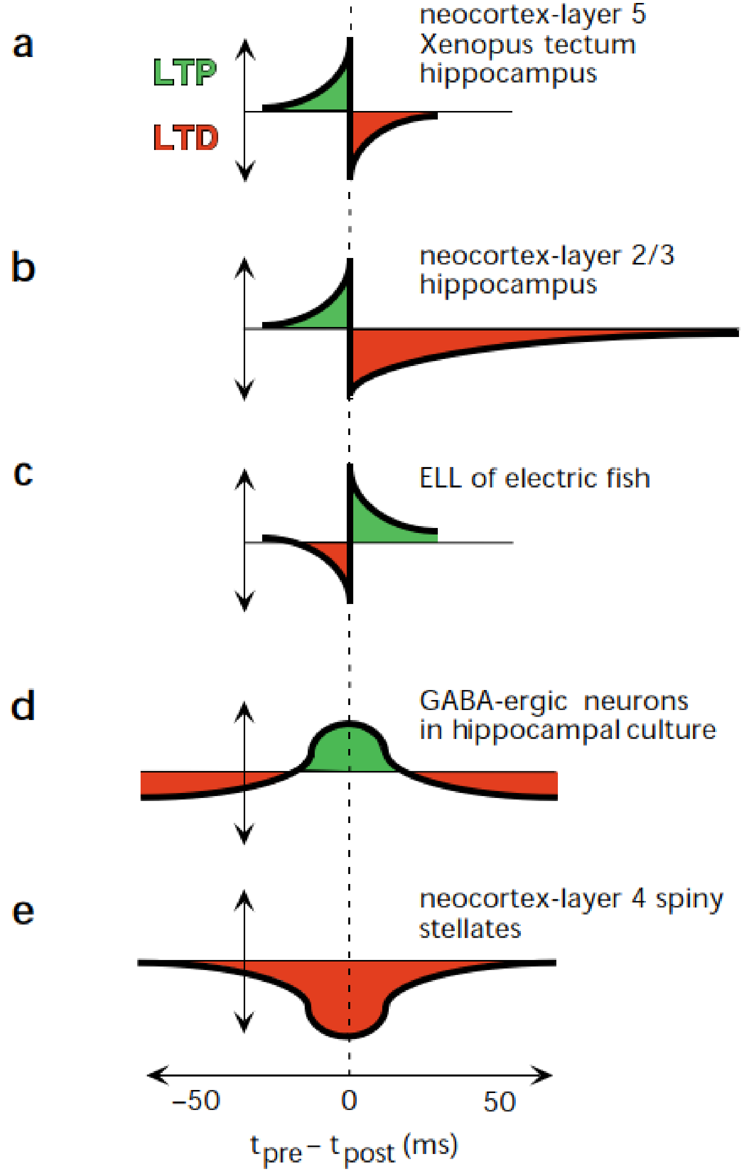


In the evening

Spike Time Dependent Plasticity

- Weight change determined by relative timing of pre- and post-neuron spikes.
- A purely Hebbian form of plasticity
  (direct correlation between post AND pre-neuron activity)
- Commonly observed in biology: Exact STDP-function differs



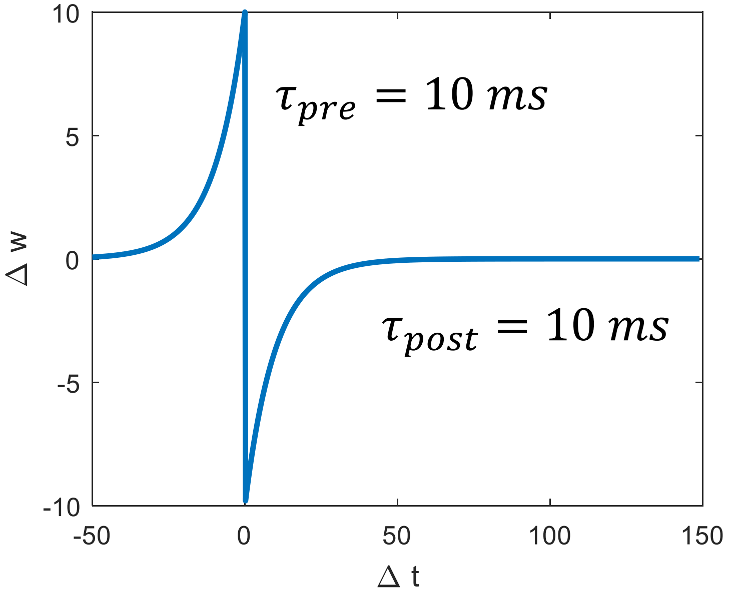


Jeong arxiv 2018; Abbot and Nelson Nat. Neuro 2000

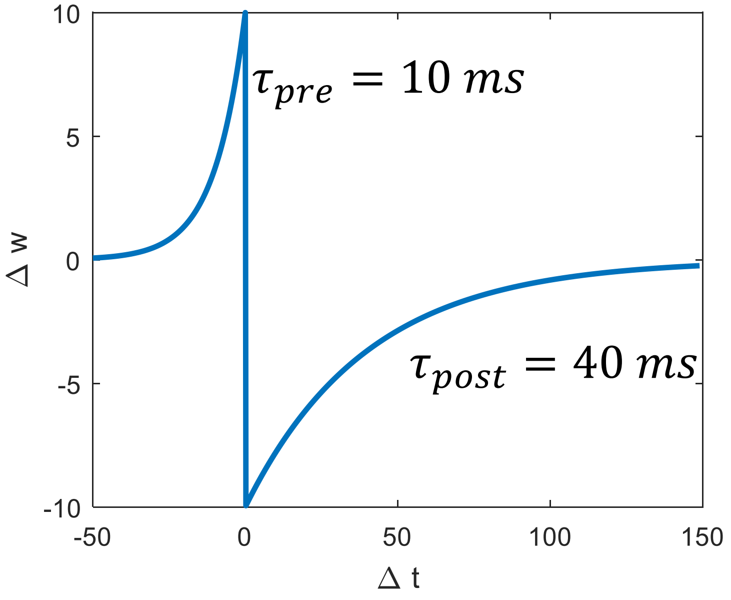
2 min exercise - Random firing in STDP

- If pre-neuron spikes randomly, which dominates (LTP or LTD)?

CASE 1



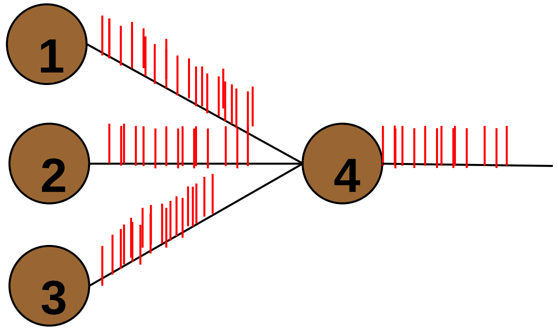
CASE 2

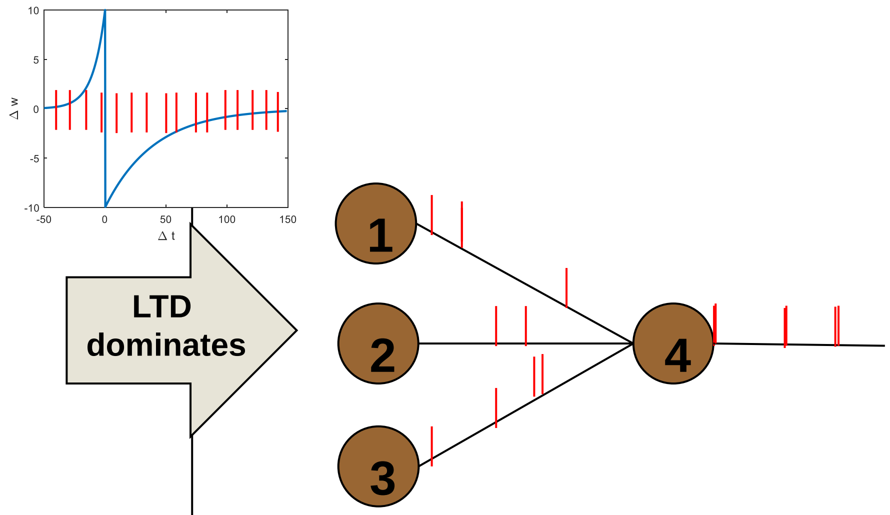


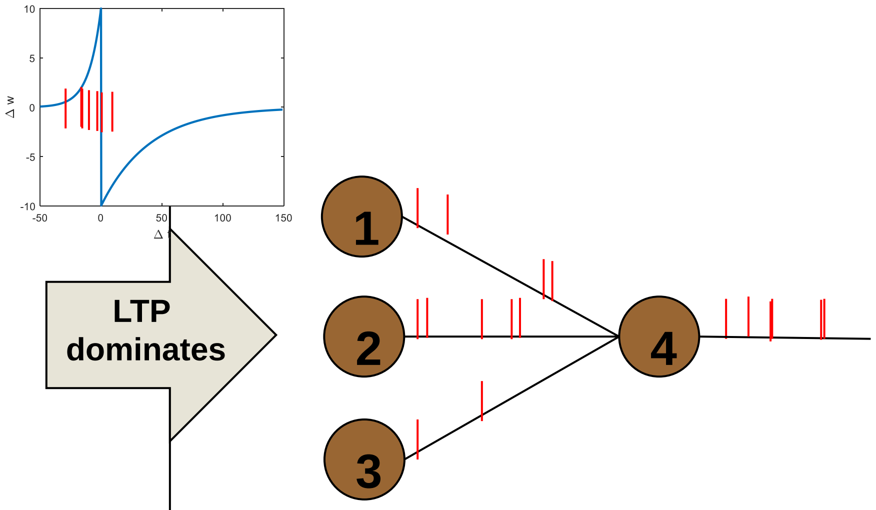
Stability in STDP

- Although truly Hebbian learning → STDP can be stable
- Requirements: Weights must be bounded between zero and maximum value

Ex:







Initial strong uncorrelated input to 4
→ Unacceptably high output activity
4 acts as an integrator

Weights are depressed until very few spikes → only clusters of spikes give output. *4 acts as coincidence detector*

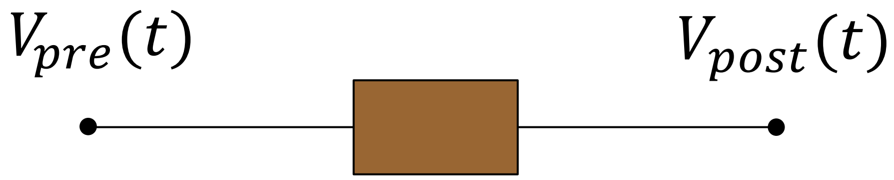
Relative timing of spikes matter
→ Nonuniform synapse weights
+ reasonable and irregular activity

Implementation of STDP in practice

- Synapse implemented with memristive device
  - ”variable resistor” R ←→ w
  - Large ΔV above threshold θ_m
    → resistance change (Δw)

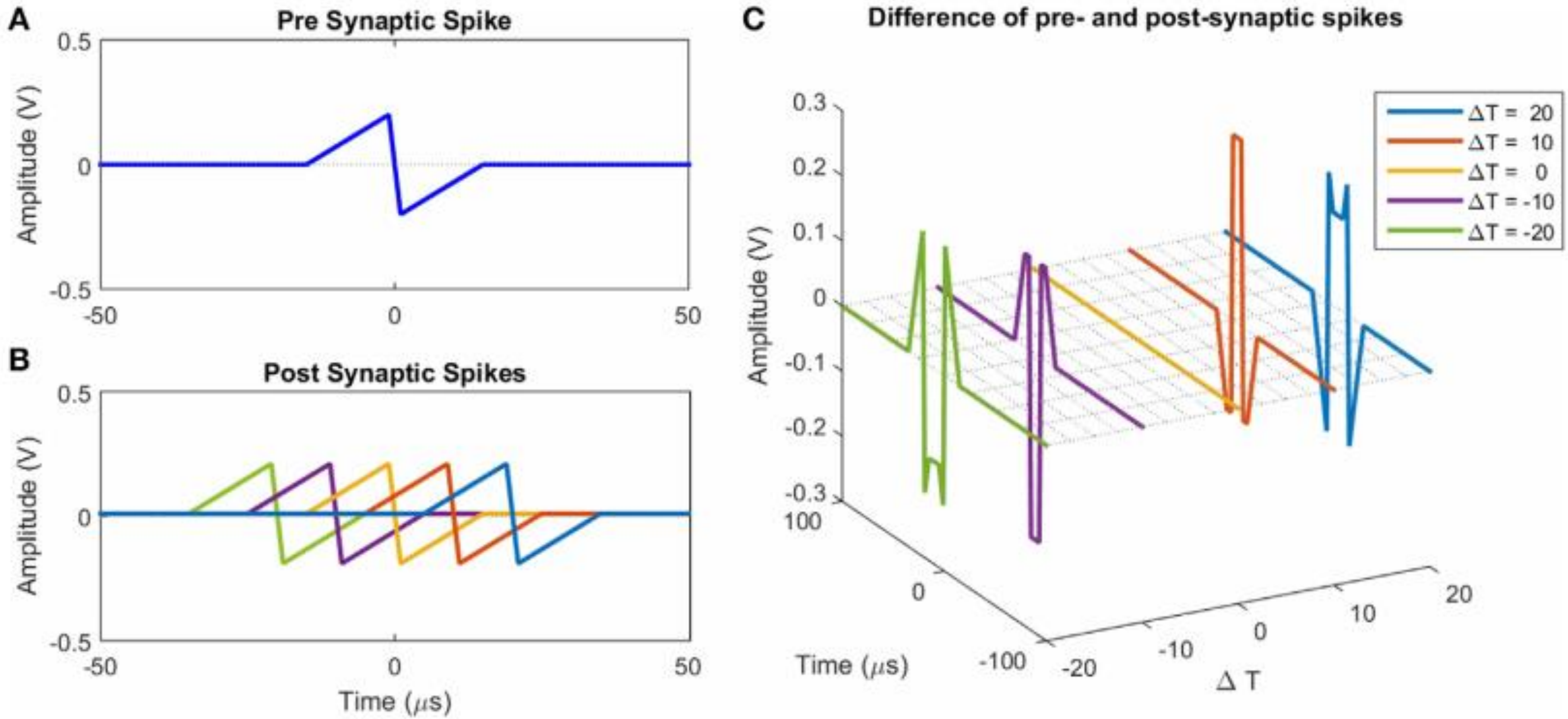
- Voltage spikes (pre and post) applied to separate electrodes, gives bias:

→ $\Delta V = V_{pre}(t) - V_{post}(t)$



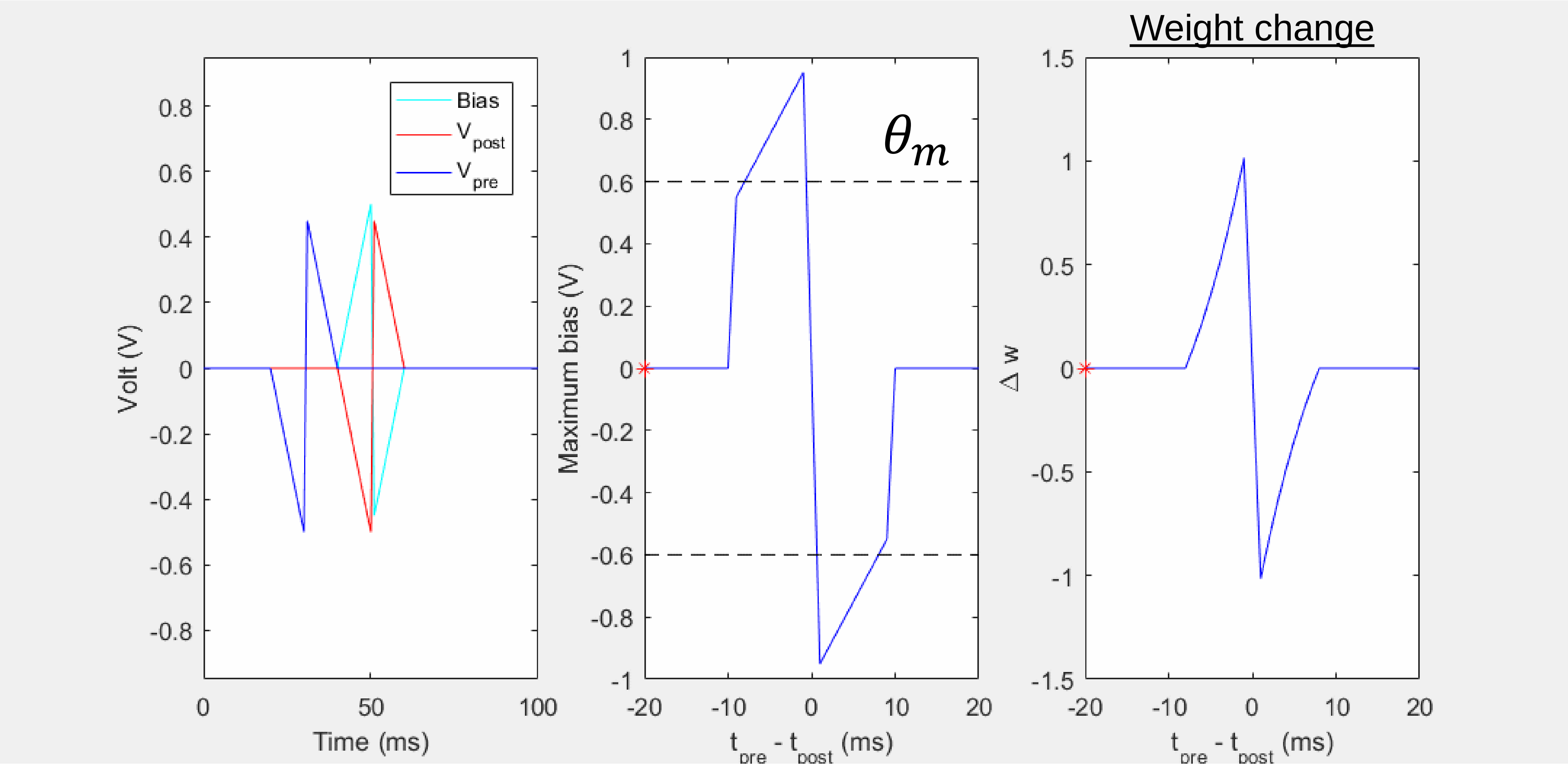
Synaptic device

Bias on memristor



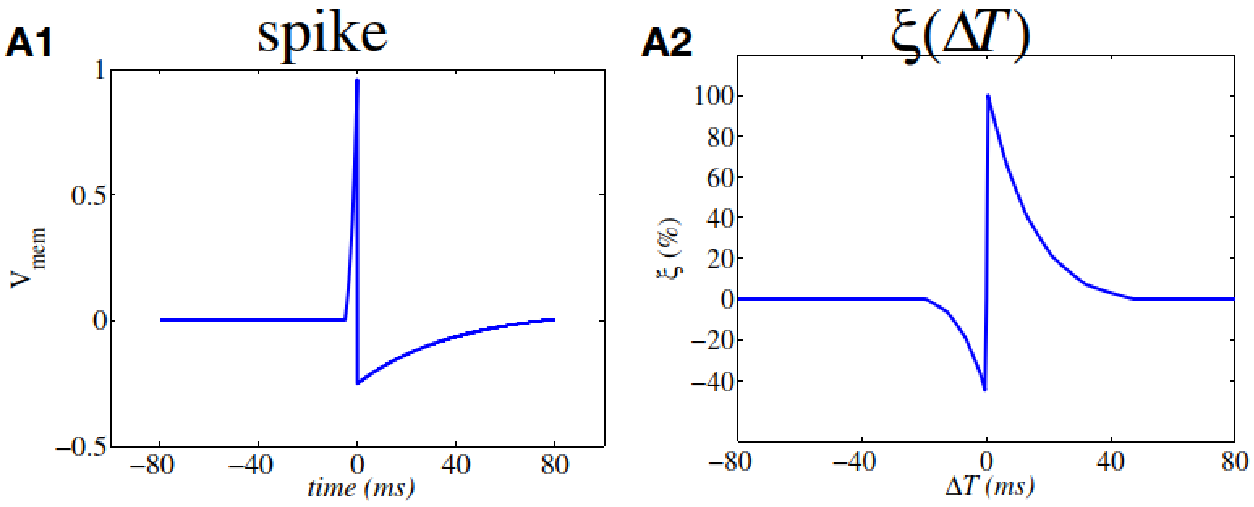
Campbell et al Frontiers in Bioeng. And Biotechn. 4, 97, 2016

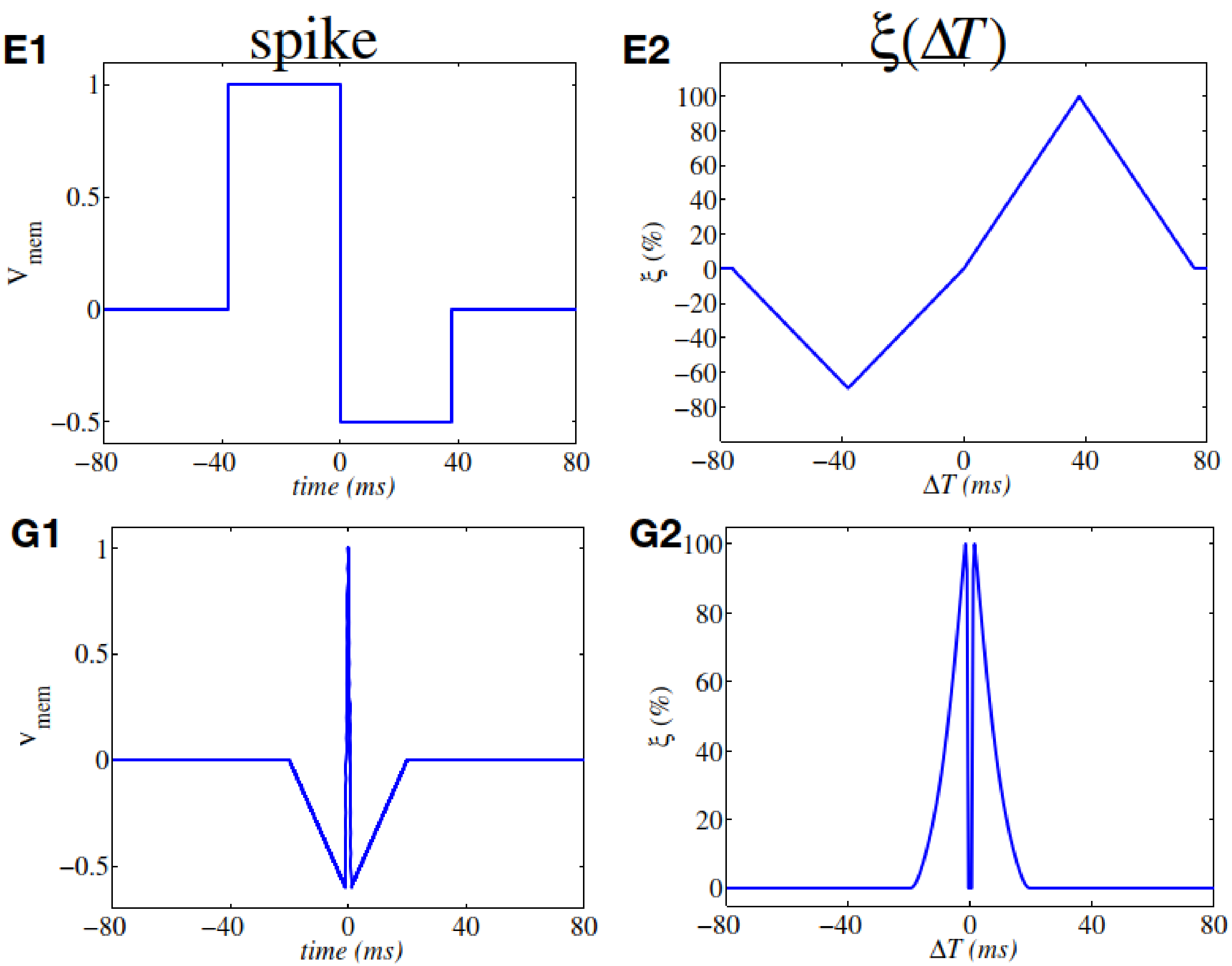
STDP on memristor

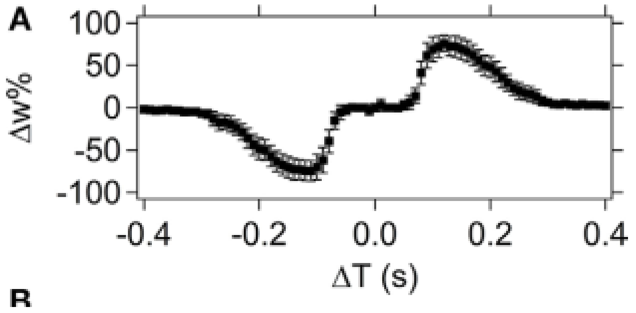


Pulse shape dependence

- STDP function very sensitive to pulse shapes



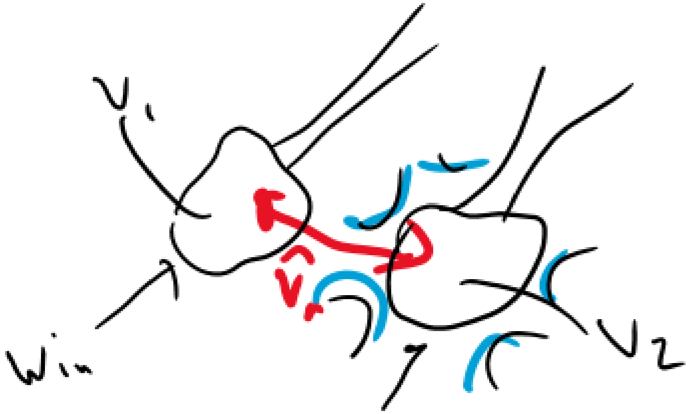




Zamarreno-Ramos et al. Frontiers NeuroSci. 2011

Effect of Noise

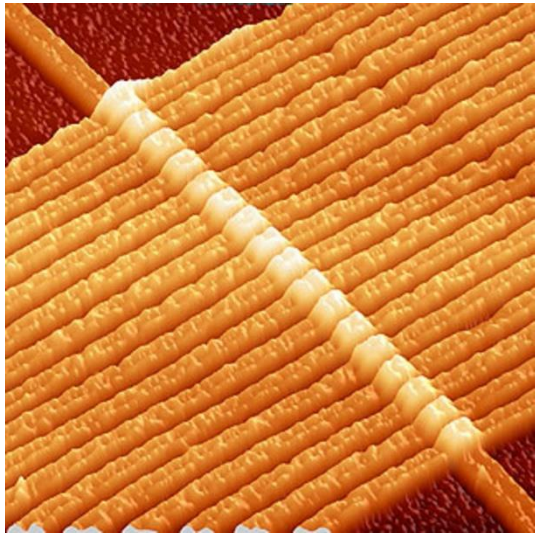
- Interactions between neighbouring neurons → potential noise
- → noise in inter-spike interval
- Which is most robust agains this?
  - Rate-coding?
  - STDP?



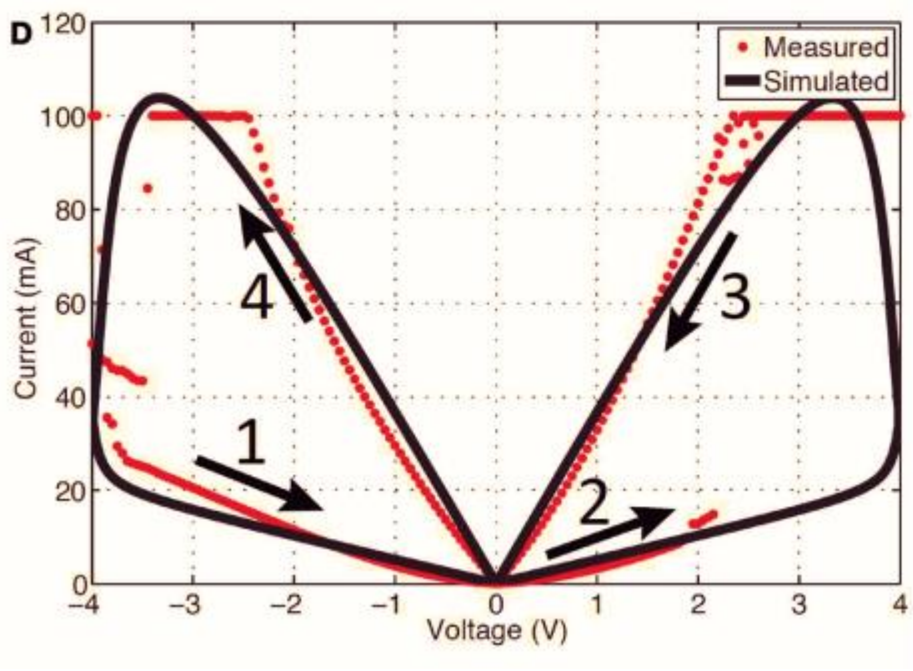
Memristor synapses in practice

- 2008: First demonstration of memristor by HP (Dmitry Strukov)
- 2009: First proposal use of memristors as synaptic devices for STDP

- Examples of memristor behavior in emerging memory technologies



This is the focus of L8-L12.



Serrano-Gotarredona et al Front. Neurosci 2013