CN510: Principles and Methods of Cognitive and Neural Modeling

Spike-Timing Dependent Plasticity I

Lecture 13

Instructor: Anatoli Gorchetchnikov <anatoli@bu.edu>

General Modeling Concerns

Plasticity shall only depend on the local variables that are accessible at the synapse:

- presynaptic components:
  - membrane potential
  - Ca concentration
  - transmitter concentration, production rate, depletion rate
- postsynaptic components:
  - membrane potential
  - Ca concentration
  - second messenger concentration
  - receptor density/total amount
  - channel density/total amount

Possibly also time-averaged versions of these variables

General Modeling Concerns

Phenomenological models are more tractable and less computationally expensive

Depending on the level of detail, variables can be lumped together

Extreme case:

- one variable representing all presynaptic effects
- one variable representing all postsynaptic effects
- one variable representing efficacy

Other variables such as transmitter depletion can be added if necessary

Rules can be determined from theoretical considerations, or macroscopic phenomena like receptive field formation, changes in response curves, etc

Theoretical Concepts

- Feature extraction based on correlation between input and output
  - If multiple inputs all activate the same output then these inputs are related
  - A basis for unsupervised categorization, generalization, and receptive field formation
- Stability of resulting representations
  - Encoding of a new memory shall not destroy old memories
  - Encoded memories shall not erode over time
- Take into account the quality of task performance
  - Positive and negative reinforcement

Practicalities

These concepts can have different value for different subsystems

No single rule implements all three

Sometimes even rules that are designed to implement one of the concepts do not align well with the data

- BCM was designed to explain visual receptive field formation, but it shows bimodal weight distribution, while experiments show unimodal distribution in visual cortex of young animals
- On the other hand unimodal rules often show instability of resulting representations over long term, while BCM forms stable receptive fields

Locally Computable Measures

Given three variables: w, x_{pre}, x_{post}

Zero order terms e.g. fixed growth/decay

First order terms

- Weight-based decay
- Effects of postsynaptic firing
- Effects of presynaptic firing

Second order terms

- Gated weight-based decay
- Conjunction of presynaptic and postsynaptic signal

Possibly a third order term

None of the variables have to have an immediate effect

Hebbian Postulate

When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased

Hebb emphasized causality and, therefore, a temporal order of neuronal firing

Spike-timing dependent plasticity is a generalization of Hebbian learning

Spike-timing Dependent Plasticity (STDP)

Discovered as early as 1983 by *Levy and Steward*, although formulated not as precise as in later work and in terms of rate-based Hebbian learning

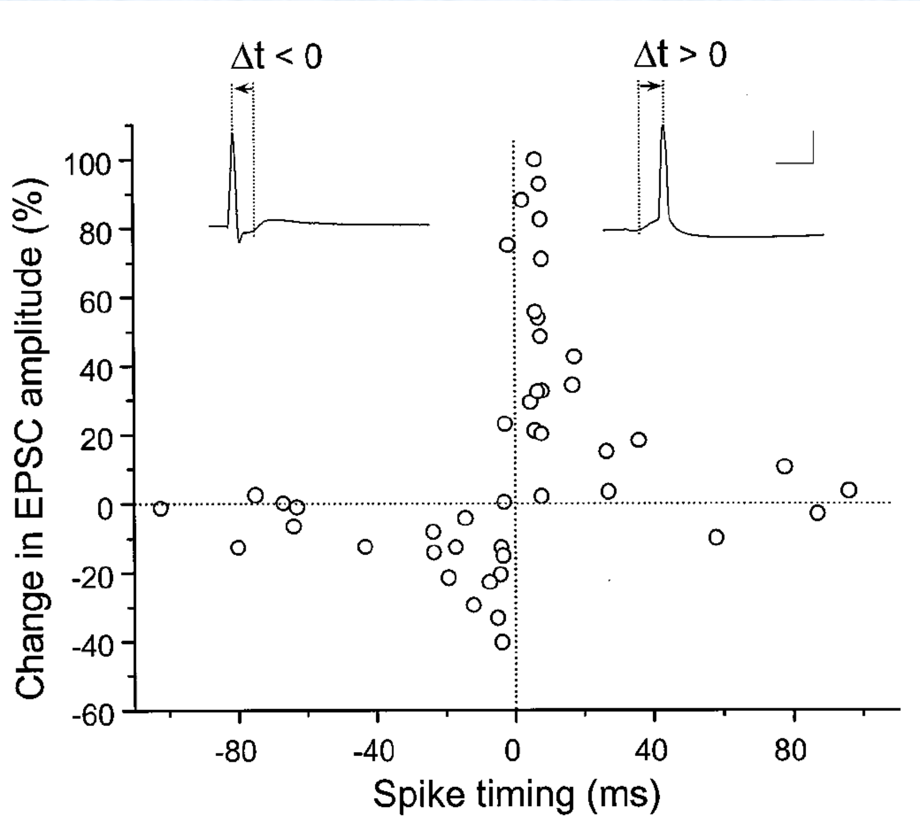
More precisely described by *Markram et al, (1997)* and by *Bi and Poo (1998)*

Long-term plasticity depends on the exact timings of presynaptic and postsynaptic spikes on the time scale of milliseconds

Note that single pairing rarely has a noticeable effect, the results are usually based on 50 or more pairings

Spike-timing Dependent Plasticity (STDP)

Bi and Poo (1998)



Potentiation of the synapse if the presynaptic spike precedes the postsynaptic spike

Depression if the presynaptic spike follows the postsynaptic spike

Variations

Repetition frequency can be important: same 60 pairings at low frequency have no effect, while at 20Hz frequency show strong effect

Postsynaptic potential matters: if the neuron is clamped slightly above rest – synapses are depressed, if it is clamped at higher depolarization – the same stimulation leads to potentiation

There is a wide variety of STDP curves depending on cell types

- In the neocortex negative window width depends on layer
- Interneurons tend to have symmetric curves
- Inverted curves have been reported in some cultures

Models of STDP

Timing-based models for rate neurons

- *Levy et al (1983+)*

Abstract models:

- *Gerstner et al. (1999)*
- *Song et al. (2000)*
- *Porr et al. (2004)*

Biophysical models based on Ca^{++} dynamics:

- *Rubin et al. (2005)*
- *Hartley et al. (2005)*

Levy et al, (1995)

A network of “local context” neurons

Goal is to disambiguate sequences, including goal-directed navigation with shortcuts

CA3 model uses

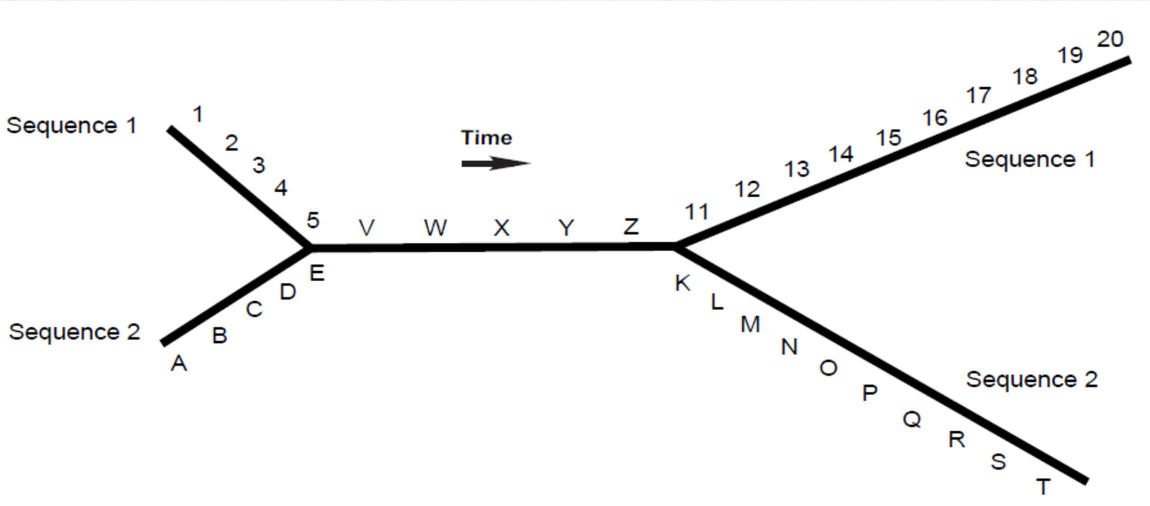
- High degree of random recurrent connectivity
- One step delay in all projections
- Postsynaptically gated decay rule

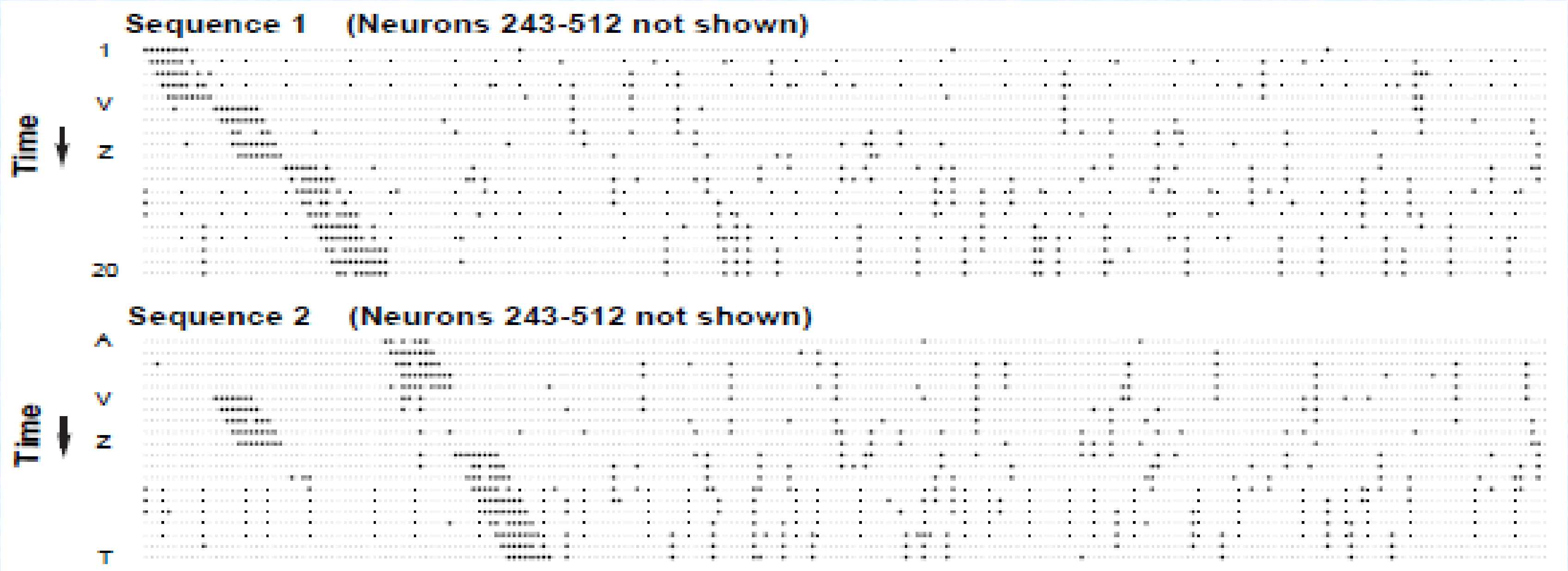
External input was provided to 10% of the cells, feedforward inhibition normalized the input

Delay in projections leads to inclusion of the order

- If presynaptic was active on the previous step and postsynaptic is active on this step…

Levy et al, (1995)





Gerstner et al. (1999) Model

The weight change is defined as

$$\Delta w_{ij} = \int_0^T \int_0^T f\left(t - t'\right) x_i\left(t'\right) y_j\left(t\right) dt dt'$$

where

$$f\left(s\right) = \begin{cases} \left[A_+ - A_-\right] e^{-\frac{s - s^*}{\tau}} & if \quad s > s^* \\ A_+ e^{-\frac{s - s^*}{\tau_+}} - A_- e^{-\frac{s - s^*}{\tau_-}} & if \quad s < s^* \end{cases}$$

Spatially local: does not require any information that is not available at the synapse the rule is applied to

Flexible: can fit any experimental data

Reduces to Hebbian for the continuous firing rate coding

Gerstner et al. (1999) **Model**

The weight change is defined as

$$\Delta w_{ij} = \int_0^T \int_0^T f\left(t - t'\right) x_i\left(t'\right) y_j\left(t\right) dt dt'$$

where

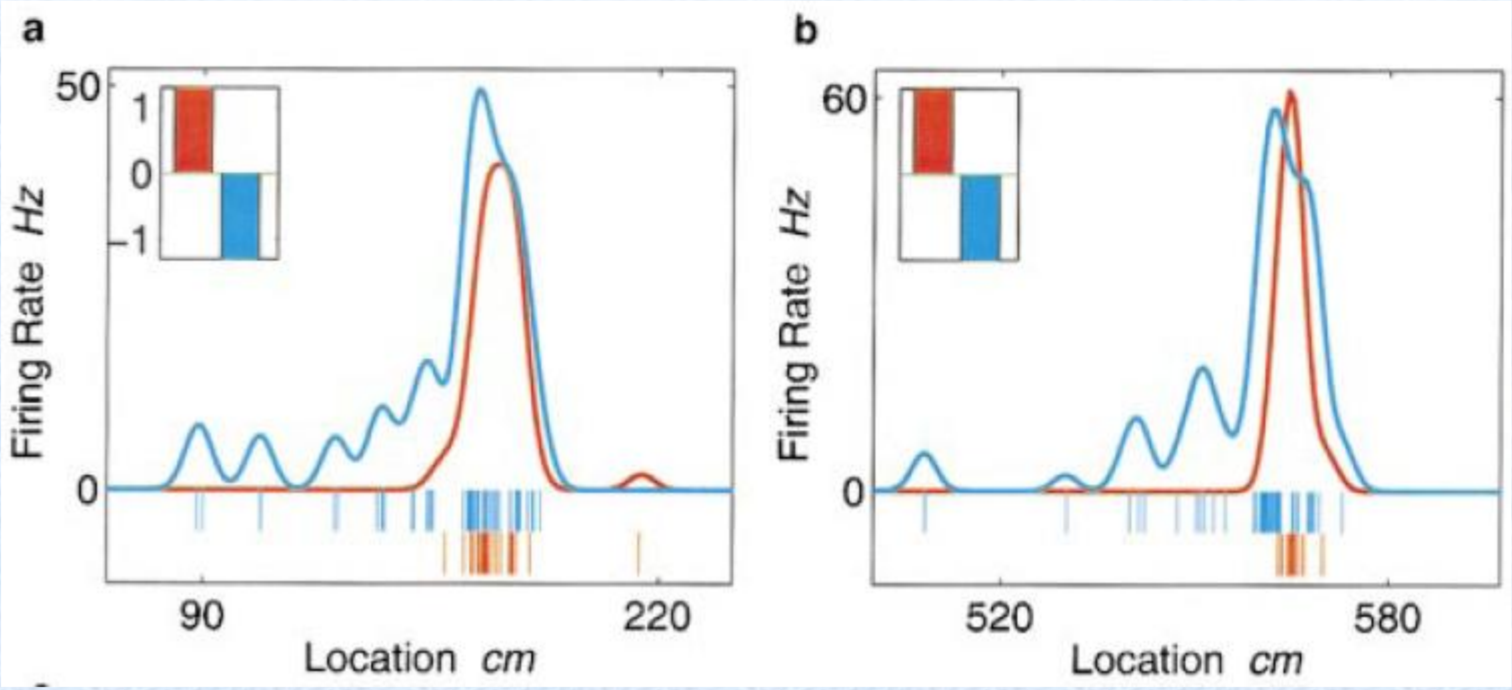
$$f\left(s\right) = \begin{cases} \left[A_+ - A_-\right] e^{-\frac{s - s^*}{\tau}} & if \quad s > s^* \\ A_+ e^{-\frac{s - s^*}{\tau_+}} - A_- e^{-\frac{s - s^*}{\tau_-}} & if \quad s < s^* \end{cases}$$

Uses both presynaptic and postsynaptic signals for both potentiation and depression

Implementation of this rule requires integration over *[0, T]* as well as memory to store *t'*

Example of Use: *Mehta et al. (2000)*

Attempt to explain the experimentally observed change in the shape of place fields



During about 20 laps the place fields of CA1 cells transform from symmetric to asymmetric and shift to earlier position

Example of Use: *Mehta et al. (2000)*

Model: point rate neurons with Ca++ concentration and spike frequency adaptation

Learning: STDP rule adapted for a firing rate model

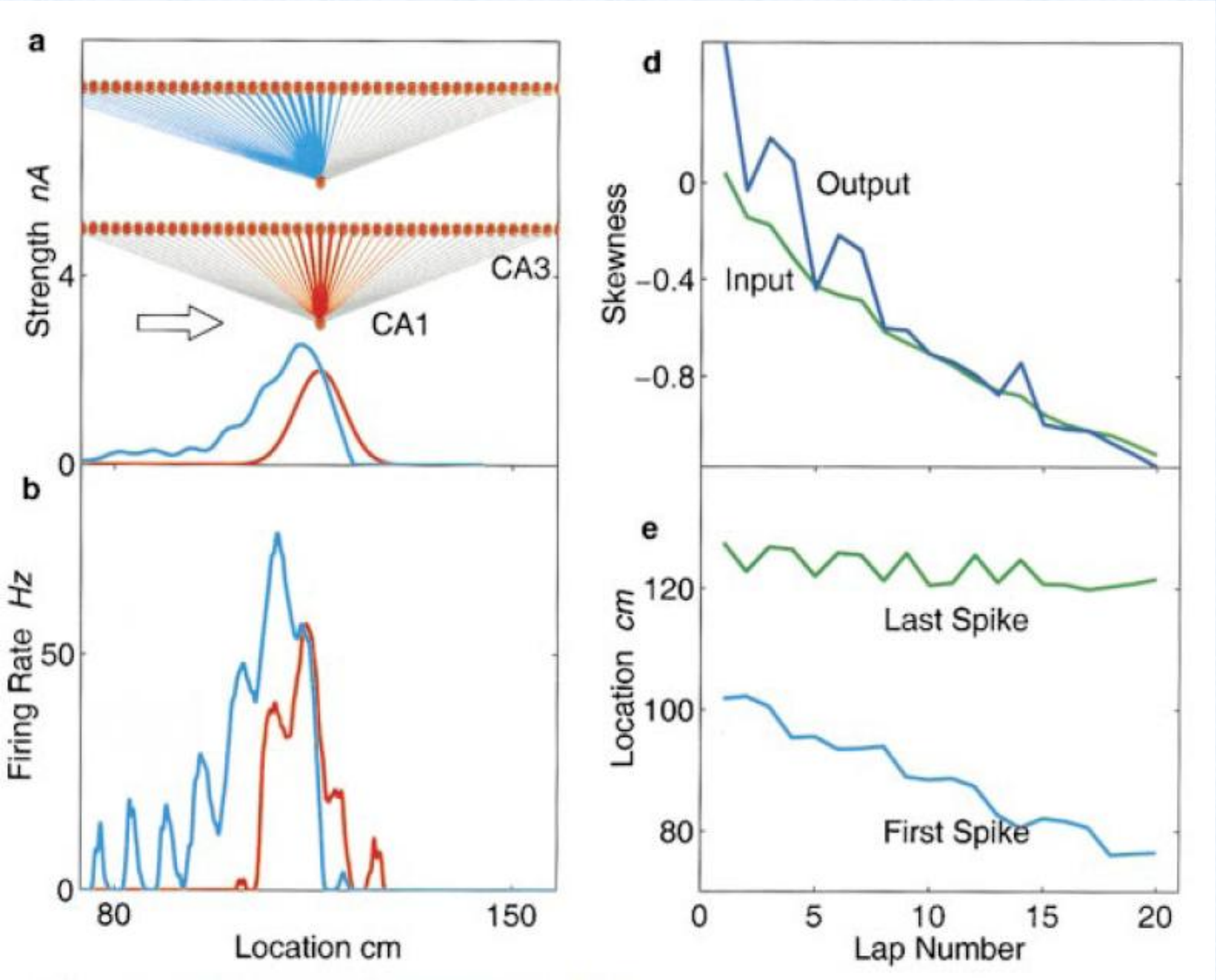
$$\Delta w_{ij} = \int_0^T dt \int_1^{50} \left[A_{LTP} e^{-\frac{\tau}{\tau_{LTP}}} x_i\left(t-\tau\right) y_j\left(t\right) - A_{LTD} e^{-\frac{\tau}{\tau_{LTD}}} x_i\left(t+\tau\right) y_j\left(t\right) \right] d\tau$$

Results: qualitatively similar to experimental data

Issues: very simple model, only accounts for the shift during running,

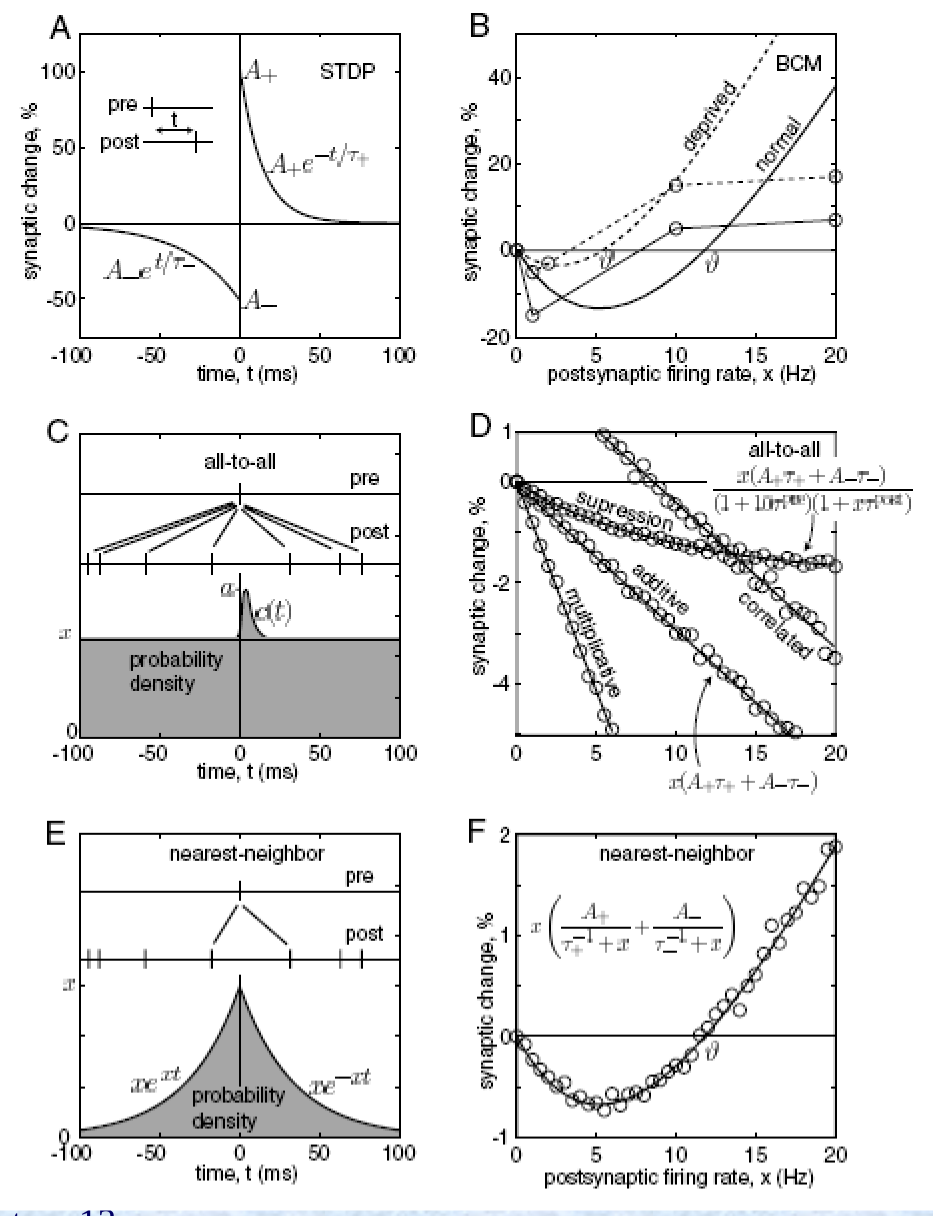
- does not account for the saturation of this shift
- does not account for the reset between training sessions

Example of Use: *Mehta et al. (2000)*



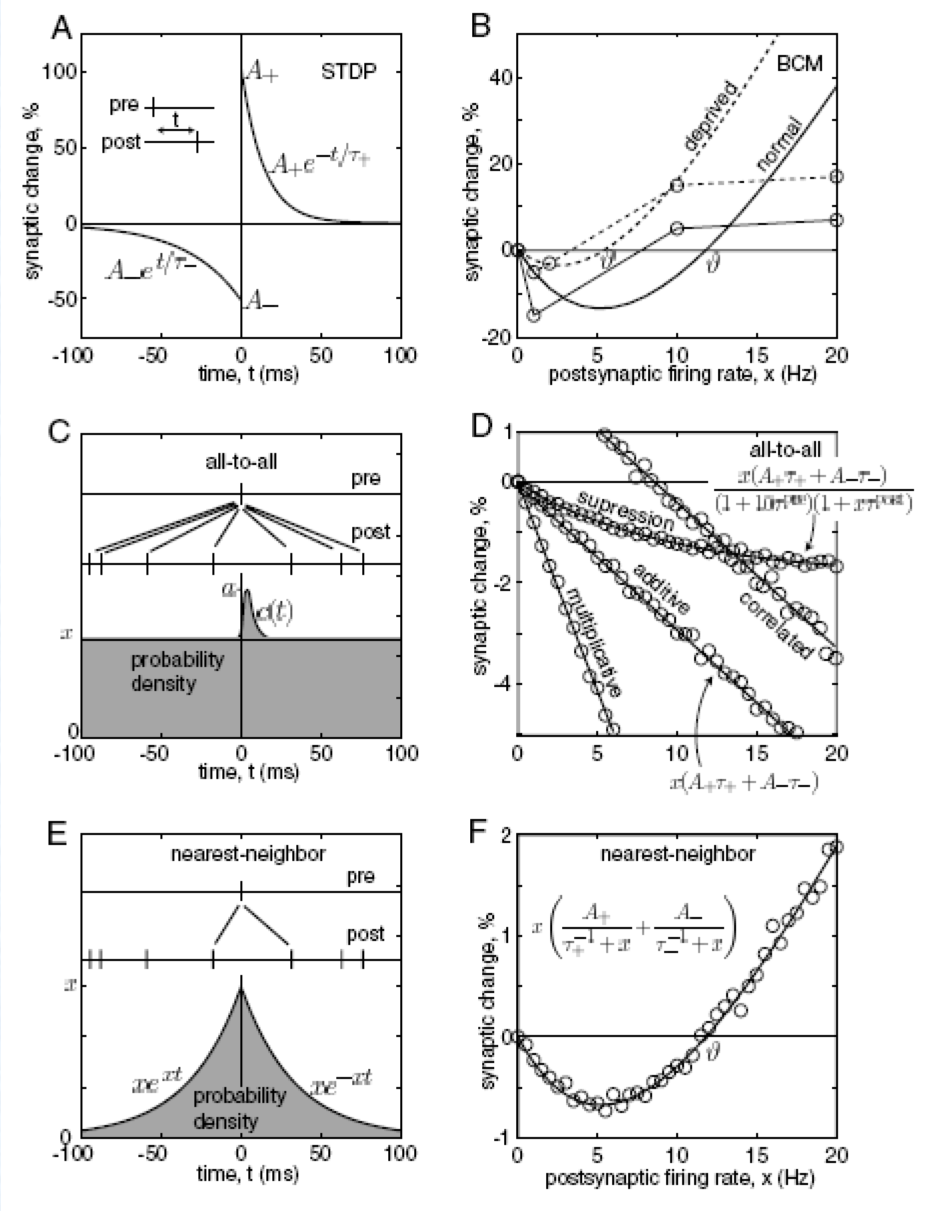
STDP and BCM

- High peak of potentiation might result in net potentiation for high firing rate
- Long tail of depression might result in net depression for low firing rate
- In fact with random firing it is the ratio of areas under the curves that determines the net effect
- Correlation between pre and post does not make it much better



STDP and BCM

- If only nearest neighbor spikes are considered though, the BCM-like curve results
- Is the threshold sliding?
- If the potentiation time constant depends on NMDA receptor time constant, which has been shown to increase with low postsynaptic firing, then yes



Pair-based STDP Rule

General form:

$$\Delta w = \begin{cases} F_+(w)e^{-\frac{|\Delta t|}{\tau_+}} & if \quad \Delta t > 0 \\ -F_-(w)e^{-\frac{|\Delta t|}{\tau_-}} & if \quad \Delta t \le 0 \end{cases}$$

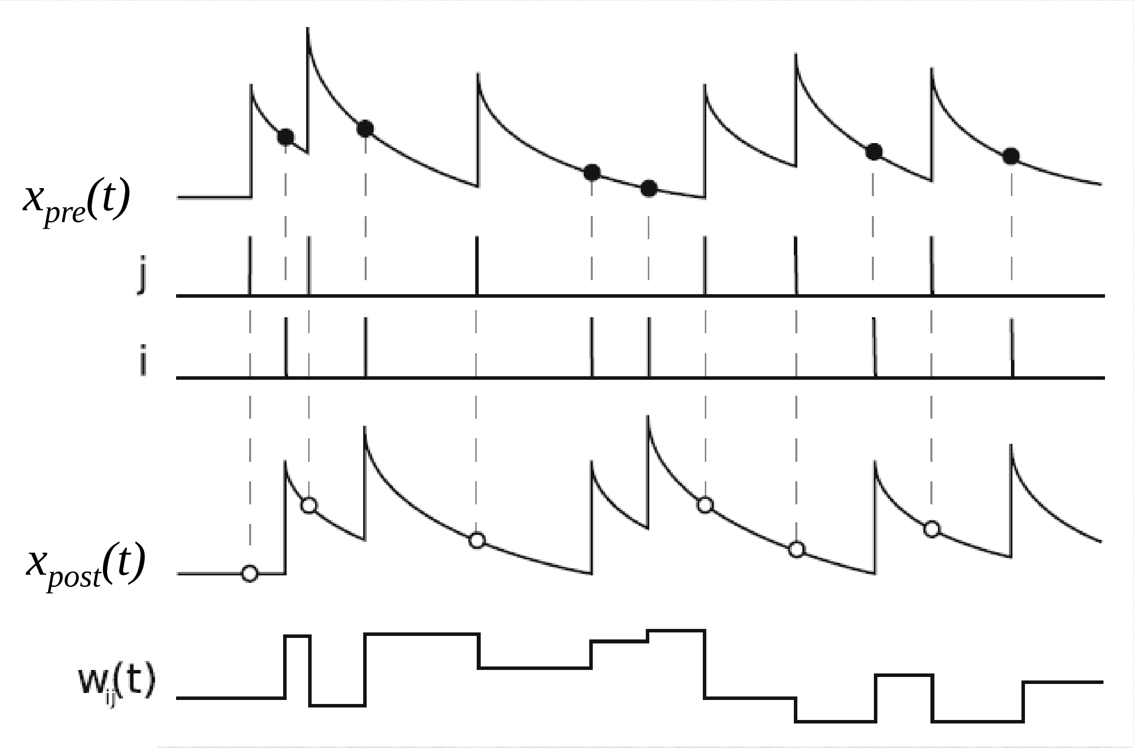
To specify:

- Define *F(w)*
- Define which spikes to pair
- Define delay contributions

Implicit Pair-based STDP Rule

Making exponents and time intervals implicit

$$\Delta w = \begin{cases} F_+(w)x_{pre}(t) & \text{if} \quad \delta(t - t_{post}) \neq 0 \\ -F_-(w)x_{post}(t) & \text{if} \quad \delta(t - t_{pre}) \neq 0 \end{cases}$$

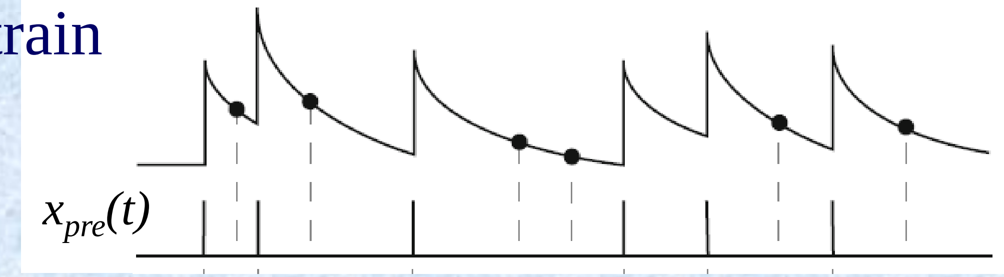


Implicit Pair-based STDP Rule

Differential form:

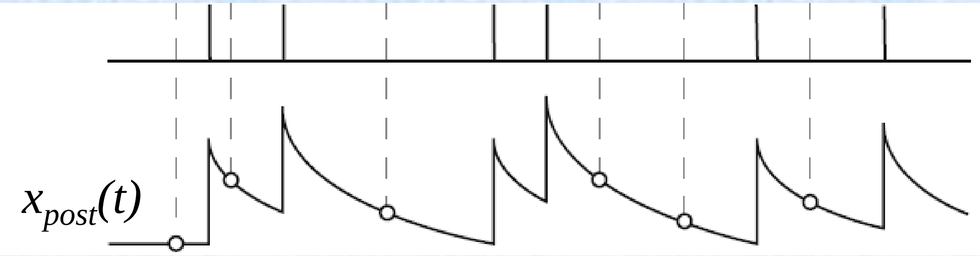
$$\dot{w} = F_+(w)x_{pre}(t)\delta(t - t_{post}) - F_-(w)x_{post}(t)\delta(t - t_{pre})$$

Note that a trace of presynaptic activity is similar to the synaptic potential/current/conductance caused by a presynaptic spike train



So we do not need to maintain a special variable for it!

For the postsynaptic train we can use a K voltage gated conductance or current as the trace variable



Implicit Pair-based STDP Rule

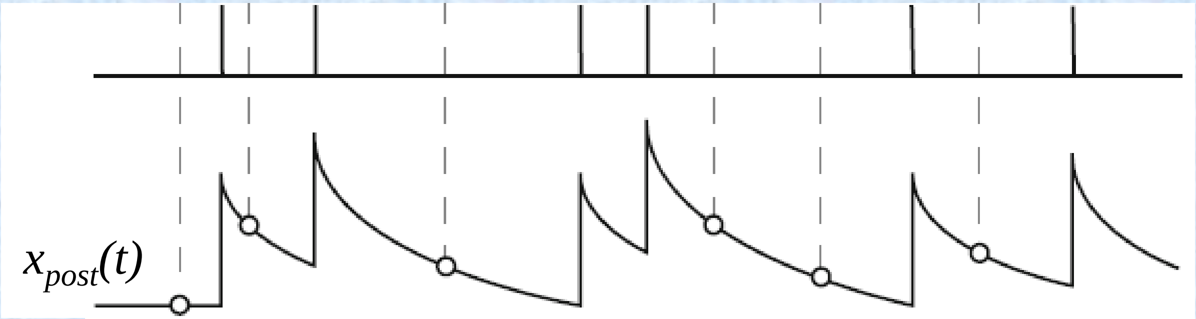
$$\dot{w} = F_+(w)x_{pre}(t)\delta(t - t_{post}) - F_-(w)x_{post}(t)\delta(t - t_{pre})$$

If we use Izhikevich simple neuron

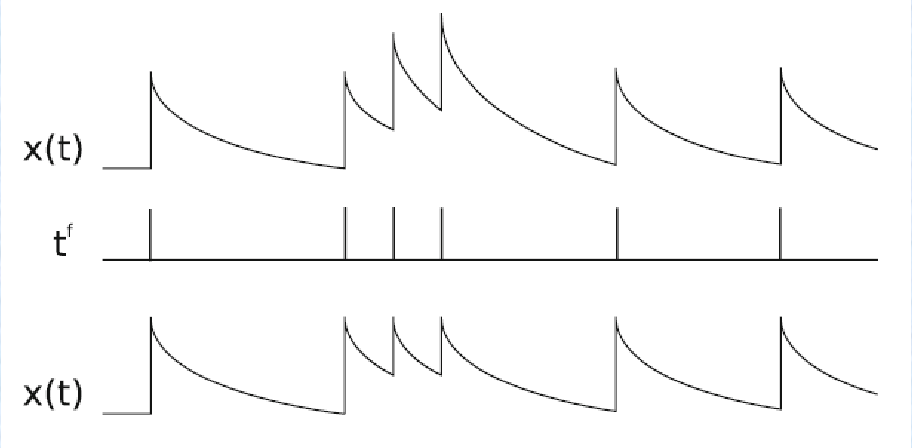
$$\frac{dV}{dt} = 0.04V^2 + 5V + 140 - u + I$$

$$\frac{du}{dt} = a(bV - u)$$

then for the low values of b the variable u also behaves similar to this trace due to $u=u+d$ at the moment of the spike



Maintaining Traces



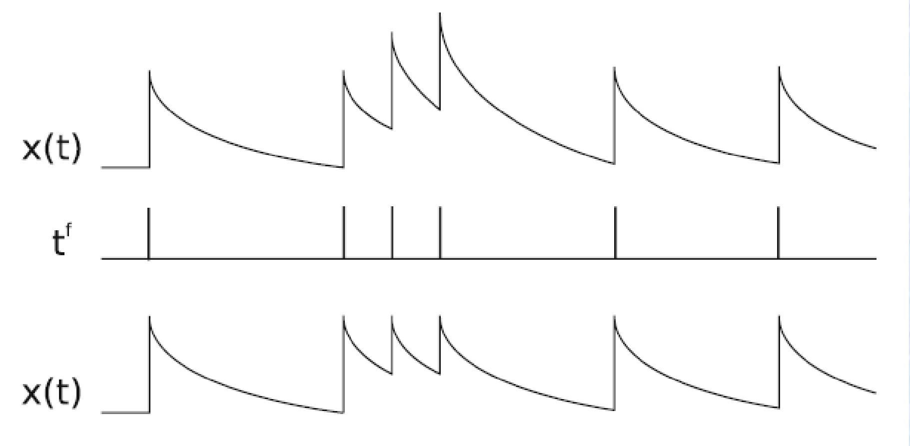
- A trace variable can be written as

$$\dot{x} = -Ax + \sum_{t_f} (B - Cx\_)\delta(t - t_f)$$

- Depending on the parameters it can be between two extreme cases on the right
- E.g. if *C=0.5* and *B=1* then on each spike *x* will jump by *(2-x)/2*

- Top trace: *C=0*, *x(t)* is increased by *B* on each spike – traces all previous spikes
- Bottom trace: *C=1*, *x(t)* is bumped to *B* on each spike – traces only the last spike

Maintaining Traces



Note that if we use conductance or current as a trace variable it has to saturate: we only have a fixed number of channels

If we use potential as a trace then it does not have to saturate, we will integrate all the inputs

In Izhikevich neuron we use the top case

Top trace: *C=0*, *x(t)* is increased by *B* on each spike – traces all previous spikes

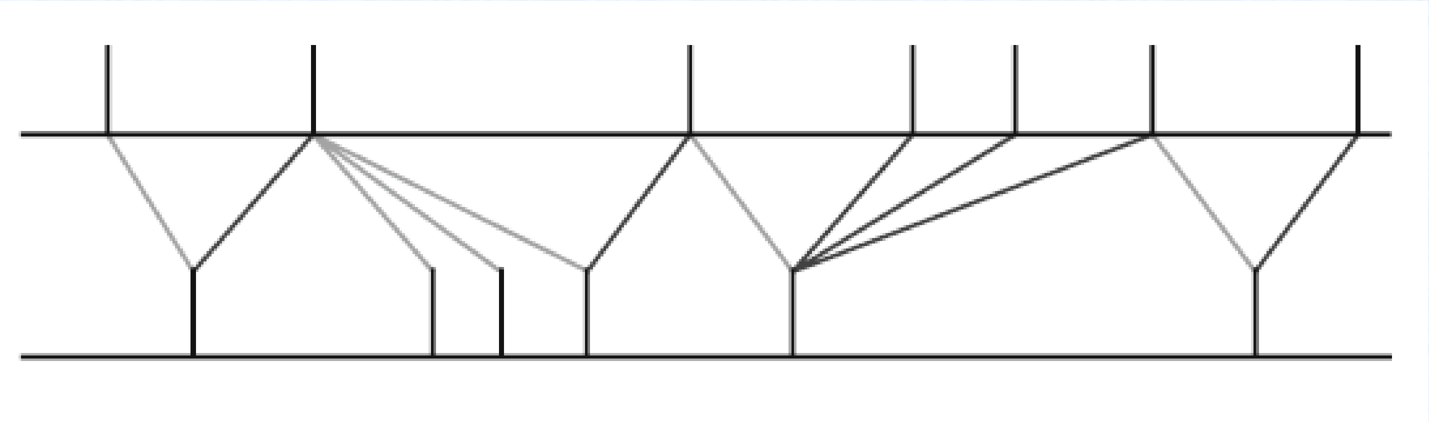
Bottom trace: *C=1*, *x(t)* is bumped to *B* on each spike – traces only the last spike

Which Spikes Do We Pair?

With non-saturating traces we pair all-to-all

With instantly saturating traces we pair only nearest neighbors where

- each presynaptic spike is paired with the last postsynaptic spike for depression component
- each postsynaptic spike is paired with the last presynaptic spike for potentiation component

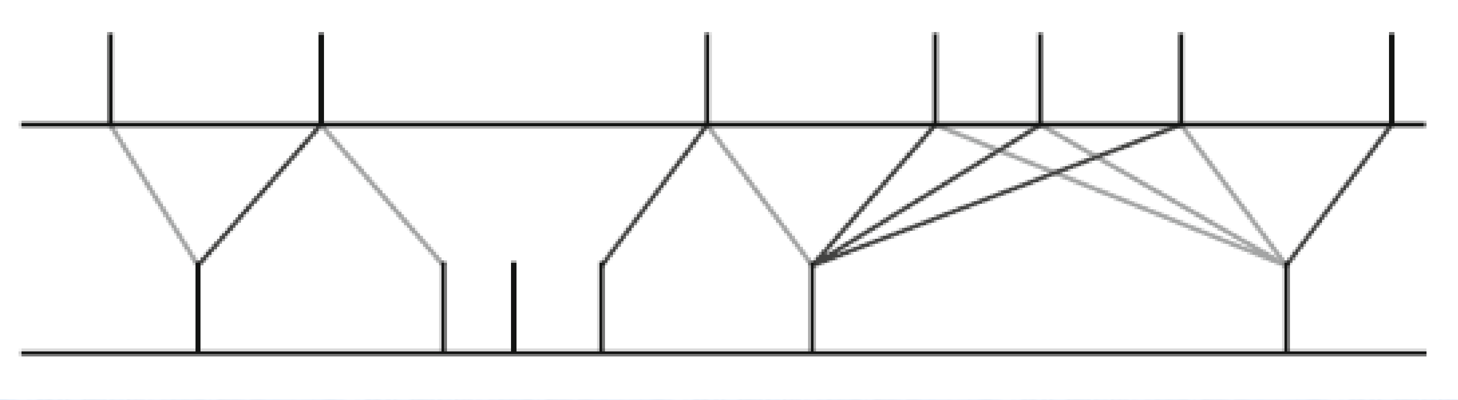


With partial saturation we weight the contribution of spikes

Which Spikes Do We Pair?

More complicated schemes can be achieved with more elaborate traces:

- Reset postsynaptic trace to *B* on every postsynaptic spike (saturating)
- Reset presynaptic trace to *B* on every presynaptic spike and to 0 on every postsynaptic spike



Feasible if we consider a postsynaptic spike as a shock/reset of a neuronal/synaptic state

Which Spikes Do We Pair?

Data suggests that for the potentiation component multiple spike pairings provide contribution, but not in the linear fashion as all-to-all scheme suggests

Possibly a partially saturated trace is the best… Conductance or current that is not fully activated by a single presynaptic spike?

For the depression component the best match seem to come from both traces having double reset

Weight Dependence

$$\dot{w} = F_+(w)x_{pre}(t)\delta(t - t_{post}) - F_-(w)x_{post}(t)\delta(t - t_{pre})$$

What are the shapes of $F_+(w)$ and $F_-(w)$?

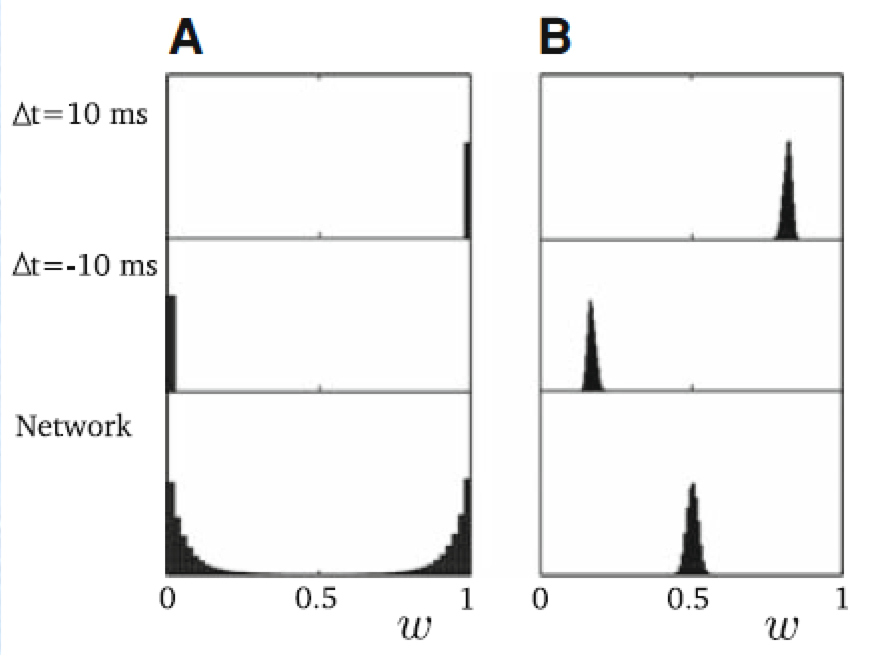
Additive: $F_-(w) = \eta_-$ $F_+(w) = \eta_+$

Multiplicative: $F_-(w) = \eta_- w$ $F_+(w) = \eta_+(w_{\max} - w)$

Power: $F_-(w) = \eta_- w^{\mu_+}$ $F_+(w) = \eta_+(w_{\max} - w)^{\mu_-}$

Weight Dependence

For 1000 uncorrelated Poisson spike trains



Additive rule leads to

- each synapse being either at 0 or at maximal weight
- network distribution being bimodal

Multiplicative rule leads to

- individual weights being not so extreme
- distribution being unimodal

A – additive rule

$$F_{-}(w) = \eta_{-} \qquad F_{+}(w) = \eta_{+}$$

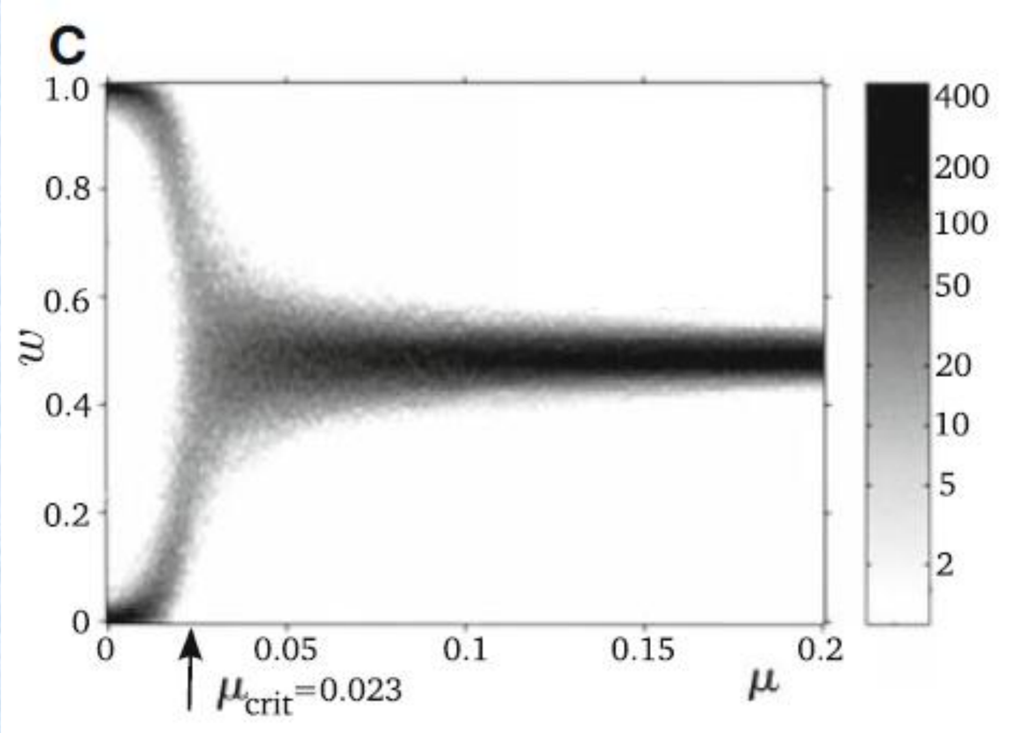
B – multiplicative rule

$$F_{-}(w) = \eta_{-} w$$

$$F_{+}(w) = \eta_{+}(w_{\max} - w)$$

Weight Dependence

- For uncorrelated trains unimodal distribution seem to be the rule
- Critical value of μ decreases as the effective number of the incoming synapses increases
- But! For correlated or partially correlated inputs – bimodal distribution for all μ



Power rule

$$F_-(w) = \eta_- w^{\mu}$$

$$F_+(w) = \eta_+ (w_{max} - w)^{\mu}$$

Are Synapses Binary?

Some data suggests that individual synapses are all or nothing *(Sjolstrom et al, 2001)*

Other data shows unimodal distribution of synaptic weights across multiple pairs of neurons/synapses *(Petersen et al, 1998)*

Might be due to data collection protocols, statistical properties, or different cell/synapse types, e.g.

- Upper bounds might be different for different synapses, resulting in broad distributions when the data is pooled
- Untrained neurons might have unimodal distribution while trained neurons might show bimodal
- …

Weight Dependence

$$\dot{w} = F_+(w)x_{pre}(t)\delta(t - t_{post}) - F_-(w)x_{post}(t)\delta(t - t_{pre})$$

It seems like power law STDP is set up to pick up correlated input traces and suppress uncorrelated ones

What about the data?

For a depression component the best fit to the *Bi and Poo* data comes from $F_-(w) = \eta_- w$

For a potentiation component the best fit is $F_+(w) = \eta_+(w_{\max} - w)^{\mu}$

with $\mu = 0.4$

The Role of Delays

Experiments usually measure

- presynaptic timing as the start of EPSP (so axonal delay of the presynaptic cell is subtracted automatically)
- postsynaptic timing as a spike timing (so delay to backpropagate the influence to the synapse is not subtracted

Playing with the delays can introduce a systematic bias in the learning and in some cases produces different spiking regimes in the resulting networks

Pair-based STDP Rule

$$\dot{w} = F_{+}(w)x_{pre}(t)\delta(t - t_{post}) - F_{-}(w)x_{post}(t)\delta(t - t_{pre})$$

Analytical and simulation studies suggest to use

$$F_{-}(w) = \eta_{-}w$$

$$F_{+}(w) = \eta_{+}(w_{\max} - w)^{\mu}$$

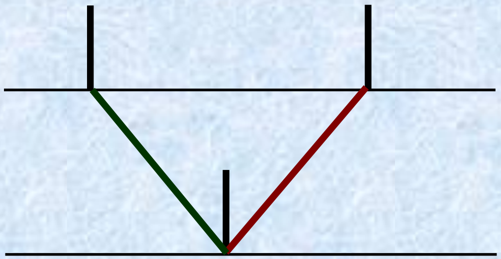
synaptic current or conductance as a presynaptic trace variable (ideally so that single input spike does not open all the channels at once)

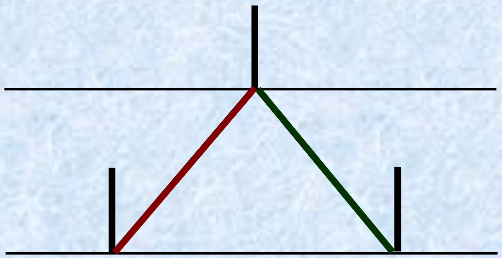
Postsynaptic trace shall be more intricate, it shall lead to only nearest neighbor dependence of the depression (reset of trace to 0 by a presynaptic spike)

Pair-based STDP Rule

Main issues:

Does not account for symmetric triplets





Does not account for frequency-dependence of the stimulation:

- Same 60 pairings give no effect at low frequency and strong effect at 20Hz frequency

Triplet Rule *(Pfister and Gerstner, 2006)*

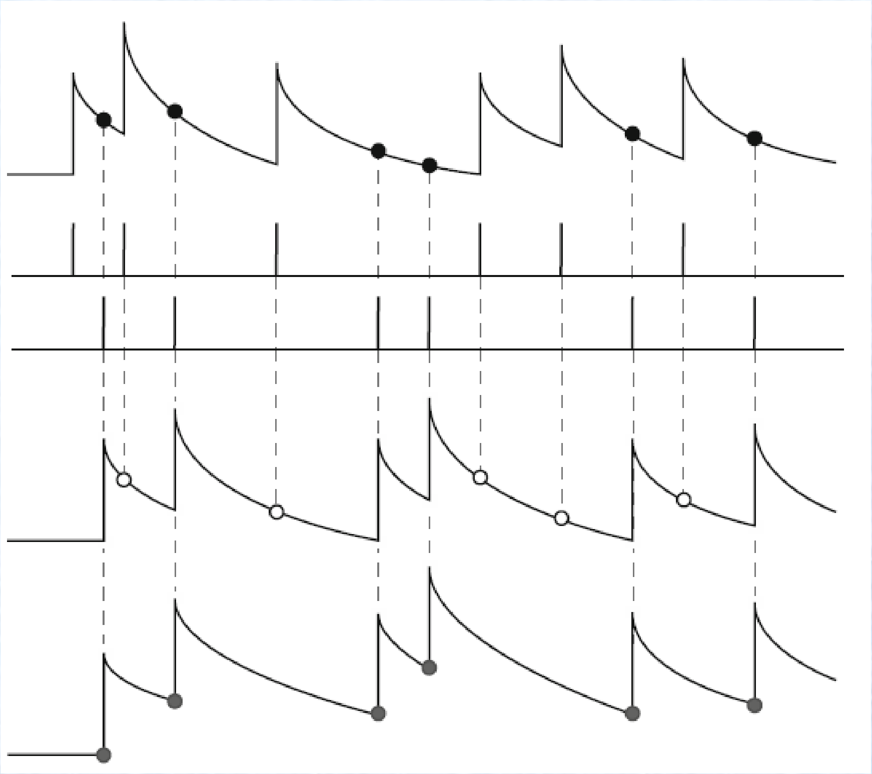
Two traces of postsynaptic activity: fast and slow

Depression is the same as in pair-based rule and uses the fast trace

$$-F_{-}(w)x_{post1}(t)\delta(t-t_{pre})$$

Potentiation uses the slow trace and the presynaptic trace

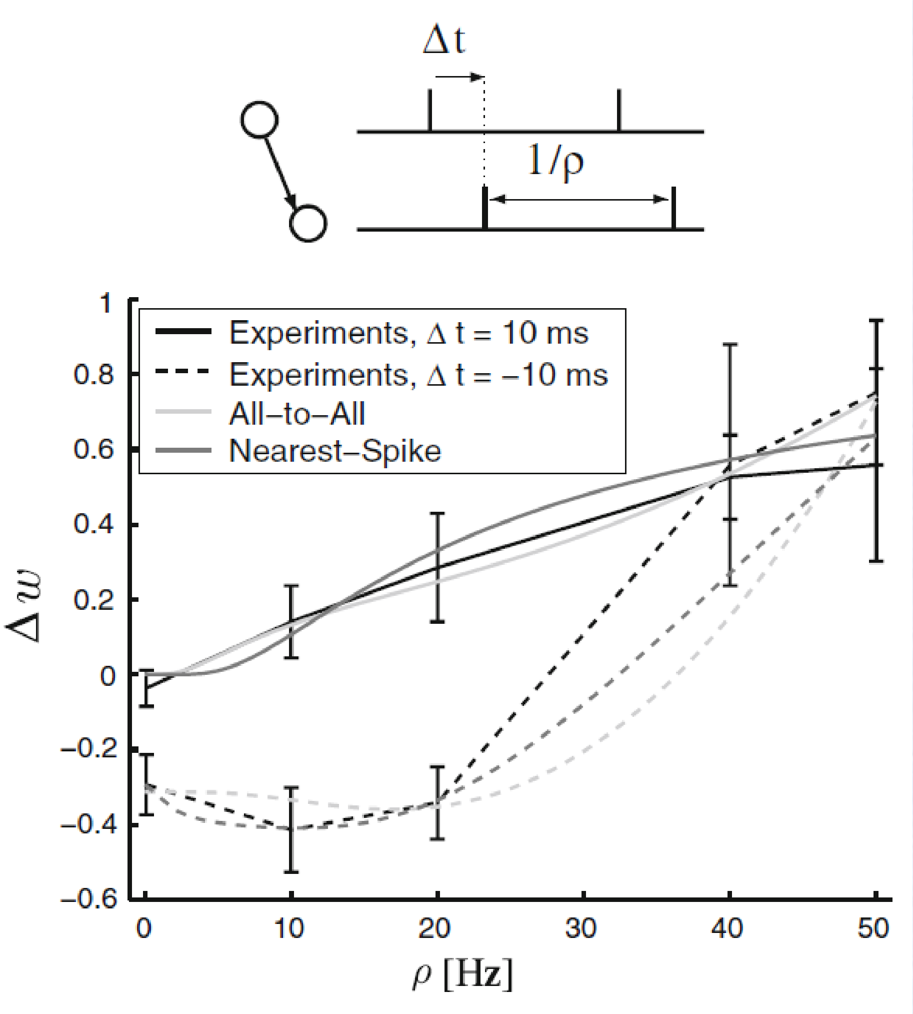
$$F_{+}(w)x_{pre}(t)x_{post2}(t)$$



Different approaches to trace saturations are still valid for this rule

Triplet Rule *(Pfister and Gerstner, 2006)*

- Matches the data on frequency dependency in visual cortex
- Good fit for triplet protocol in the hippocampus
- Also maps into BCM rule
- Furthermore, as slow trace in the triplet term depends on the postsynaptic firing rate, it naturally creates the sliding threshold



Next Week

An related implementation of STDP based on Hebbian rule and gated decays is discussed

An example of application of this rule in BEATS (Beats Encoder Algorithm for Time and Space) model is analyzed

Readings: Gorchetchnikov, A., Versace, M., and Hasselmo, M.E. (2005). A model of STDP based on spatially and temporally local information: Derivation and combination with gated decay. *Neural Networks* 18, 458-466.