

## Solution 9.1: Spike timing dependent plasticity

1. From a computational perspective, only strengthening/weakening of synaptic weights would eventually result in either all or no neurons firing in an interconnected population, which is uninteresting, as such a population would be capable of representing only a single bit of information.. From a biological perspective, both processes, forming new memories and removing old one are necessary for normal functioning.
2. The weight update rule for standard STDP is given by:

$$\dot{w} = -A^- X(t) o_1(t) + A^+ Y(t) r_1(t), \quad \text{where} \quad (1)$$

$o_1(t)$  and  $r_1(t)$  are the pre-synaptic and post-synaptic trace variables as illustrated below

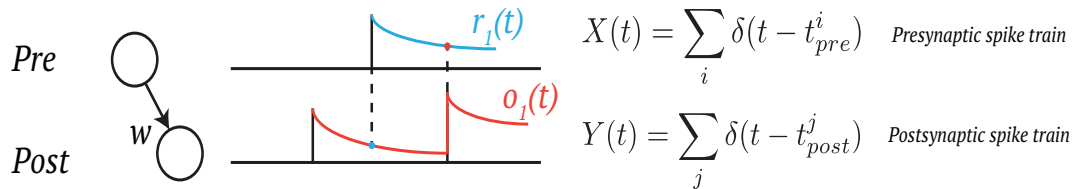


Figure 1: Schematic showing standard pair based STDP with a single set local variables  $r_1(t)$  and  $o_1(t)$  associated with the pre and post neuron respectively

The trace variables  $o_1(t)$  and  $r_1(t)$  which are low-pass filtered versions of the post-synaptic and pre-synaptic spike trains respectively, are solutions to the following differential equations (refer to lecture slides)

$$\dot{o}_1 = \frac{-o_1}{\tau_-} + Y(t) \quad (2)$$

$$\dot{r}_1 = \frac{-r_1}{\tau_+} + X(t) \quad (3)$$

Using the definitions of the dirac-delta functions (as introduced in class), the solutions to (2) and (3) can be written as:

$$o_1(t) = \int_0^\infty Y(t-s) \exp\left(-\frac{s}{\tau_-}\right) ds \quad (4)$$

$$r_1(t) = \int_0^\infty X(t-s) \exp\left(-\frac{s}{\tau_+}\right) ds \quad (5)$$

3. The expected weight change is given by:

$$\langle \dot{w} \rangle = -A^- \langle X(t) o_1(t) \rangle_{X,Y} + A^+ \langle Y(t) r_1(t) \rangle_{X,Y} \quad (6)$$

Substituting (4) and (5) in the above expression we get:

$$\langle \dot{w} \rangle = \int_0^\infty -A^- \exp\left(-\frac{s}{\tau_-}\right) \langle X(t) Y(t-s) \rangle ds + \int_0^\infty A^+ \exp\left(-\frac{s}{\tau_+}\right) \langle Y(t) X(t-s) \rangle ds \quad (7)$$

By employing a change of variable ( $s \rightarrow -s$ ) in the first integral of Eq (7), we can rewrite it as follows:

$$\langle \dot{w} \rangle = \int_{-\infty}^0 -A^- \exp\left(\frac{s}{\tau_-}\right) \langle X(t)Y(t+s) \rangle ds + \int_0^{\infty} A^+ \exp\left(-\frac{s}{\tau_+}\right) \langle Y(t)X(t-s) \rangle ds \quad (8)$$

The correlation of  $X(t)$  and  $Y(t)$  is defined as  $C(t, s) = \langle X(t)Y(t+s) \rangle$ . Because of stationarity  $C(t, s) = C(s)$ . Therefore  $\langle X(t)Y(t+s) \rangle = \langle X(t-s)Y(t) \rangle$ .

$$\begin{aligned} \langle \dot{w} \rangle &= \int_{-\infty}^0 -A^- \exp\left(\frac{s}{\tau_-}\right) C(s) ds + \int_0^{\infty} A^+ \exp\left(\frac{-s}{\tau_+}\right) C(s) ds \\ &= \int_{-\infty}^{\infty} C(s) F(s) ds \end{aligned}$$

4. If the input and output spikes are independent, then  $C(s) = \langle X(t) \rangle \langle Y(t-s) \rangle$ . Under the assumption that the rates of the neurons remain constant over time (i.e spikes are drawn from a homogenous Poisson process), we obtain the following simplification:

$$\langle X(t) \rangle = \langle X(t-s) \rangle \quad \text{and} \quad \langle Y(t) \rangle = \langle Y(t-s) \rangle$$

Defining,  $\rho_x = \langle X(t) \rangle$  and  $\rho_y = \langle Y(t) \rangle$  as the rates of the pre- and postsynaptic neuron, we can write  $C(\Delta t) = \rho_x \rho_y$  and  $\langle \dot{w} \rangle$  can be written as:

$$\langle \dot{w} \rangle = \rho_x \rho_y \int_{-\infty}^{\infty} F(s) ds \quad (9)$$

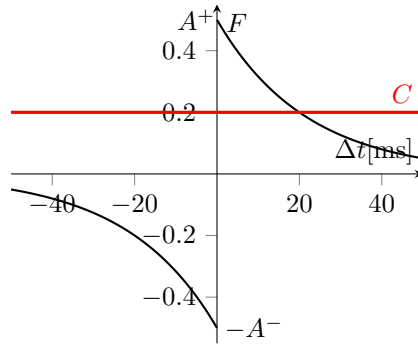


Figure 2: Correlation function  $C$  for uncorrelated pre- and postsynaptic neurons.

Two main aspects in Hebbian learning are: *joint activity* and *locality*. Joint activity means that both pre- and postsynaptic neurons have to be active simultaneously to induce changes in the weight. Locality means that only neurons that are directly connected at a given synapse can change the weight of that synapse. The average learning rule fulfills both aspects through the term  $\rho_x \rho_y$ .

5. Introducing  $C(s) = C_0 + H(s)C_1 \exp(-s/\tau_C)$  in the expression of  $\langle \dot{w} \rangle$ :

$$\begin{aligned} \langle \dot{w} \rangle &= \int_{-\infty}^0 -A^- \exp\left(\frac{s}{\tau_-}\right) C(s) ds + \int_0^{\infty} A^+ \exp\left(\frac{-s}{\tau_+}\right) C(s) ds \\ &= C_0 \int_{-\infty}^{\infty} F(s) ds + C_1 A^+ \int_0^{\infty} \exp(-s/\tau_C) \exp(-s/\tau_+) ds \end{aligned}$$

Let's now consider another set of values for  $F$ :  $A^+ = 0.77$ ,  $A^- = 0.47$ ,  $\tau_- = 35\text{ms}$  and  $\tau_+ = 7\text{ms}$ , such that  $\int_{-\infty}^{\infty} F(s) ds < 0$  (because the area on the negative side of  $\Delta t$  is higher than the area on the positive side).

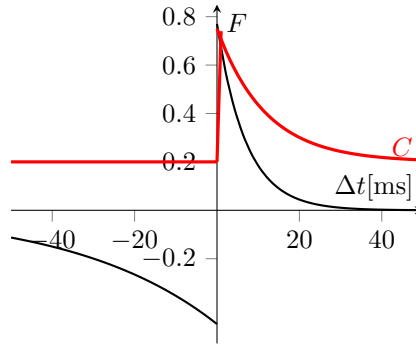


Figure 3: Correlation function  $C$  for correlated pre- and postsynaptic neurons.

When  $C_1 = 0$ :

$$\langle \dot{w} \rangle = C_0 \int_{-\infty}^{\infty} F(s) ds < 0$$

As we increase the value of  $C_1$ ,  $\langle \dot{w} \rangle$  becomes positive, because we amplify the area on the positive side of  $\Delta t$ . Therefore, for low  $C_1$  the synapse will undergo depression, and for high  $C_1$  the synapse will undergo potentiation. We call the relation between  $C_1$  and  $\langle \dot{w} \rangle$  a correlation detector, because knowing whether the synapse is depressed or potentiated, you can infer how correlated are the two connected neurons.

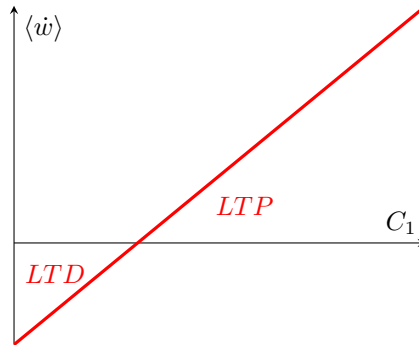


Figure 4: Correlation Detector. LTD = Long Term Depression, LTP = Long Term Potentiation

## Supplementary Info related to 9.1:

1. If two spike trains are correlated, of pre- and post-synaptic neurons, then it follows that if a pre-synaptic neuron spikes, there is a high probability that a post-synaptic neuron will spike afterwards.
2. The values for function  $F$  of  $A^+$ ,  $A^-$ ,  $\tau_+$ ,  $\tau_-$  vary from brain area to brain area.

## Solution 9.2: Triplet STDP (optional)

1. The triplet STDP rule can reproduce certain effects observed in the experimental data that the standard pair based STDP cannot. For instance, in biological experiments it has been observed that at low repetition frequency (frequency of the pairs of spikes during the STDP induction), the synapse undergoes no potentiation which is inconsistent with the standard STDP model.

Under the assumption of independence, the triplet STDP is consistent with a rate based learning rule called the **Bienenstock-Cooper-Munro (BCM)** rule [2]. The BCM rule has interesting properties such as making the post-synaptic neuron selective to only certain kinds of inputs i.e. it will fire when a specific input pattern is presented while not firing when other input patterns are presented. Hence, such a learning rule results in "input selectivity" which is important for physiological processes such as receptive field formation. A simulation of the triplet STDP (see Figure 7 in reference 1), shows that the triplet STDP inherits this property of "input selectivity" due to its resemblance to the BCM rule.

## References

1. Pfister, J.-P., and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *Journal of Neuroscience*, 26(38), 9673-9682.
2. Bienenstock, Elie L., Leon N. Cooper, and Paul W. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience* 2.1 (1982): 32-48.