# HOMEOSTASIS AND LEARNING THROUGH SPIKE-TIMING DEPENDENT PLASTICITY

L.F. Abbott[1] and Wulfram Gerstner[2]

*(1) Volen Center for Complex Systems and Department of Biology*
*Brandeis University, Waltham MA 02454-9110*
*(2) Laboratory of Computational Neuroscience, I&C and Brain-Mind institute*
*EPFL, CH-1015 Lausanne EPFL*

# Contents

**Abstract** Synaptic plasticity is thought to be the neuronal correlate of learning. Moreover, modification of synapses contributes to the activity-dependent homeostatic maintenance of neurons and neural networks. In this chapter, we review theories of synaptic plasticity and show that both homeostatic control of activity and detection of correlations in the presynaptic input can arise from spike-timing dependent plasticity (STDP). Relations to classical rate-based Hebbian learning are discussed.

## 1. Introduction

Neurons interact with each other through synapses, which are not simply static connections between neurons, but dynamic elements that display plasticity over a wide range of time scales involving numerous biophysical processes. These notes cover the mathematical description of a subset of this rich dynamics that may have particular relevance to the functioning of neural circuits. Spike-timing dependent plasticity (STDP) is the main theme here but, in addition, rate-based Hebbian learning and short-term synaptic depression are covered, both because they are of intrinsic interest, and because they appear to have a close connection to and interaction with STDP.

The essential steps of signal transmission across a chemical synapse can be summarized as follows. If an action potential that has been sent off by a presynaptic neuron arrives at time $t^{\mathrm{pre}}$ at a presynaptic terminal, voltage-gated calcium channels are opened. Due to calcium influx into the cell, vesicles docked onto the presynaptic membrane release the transmitter they contain into the synaptic cleft. Transmitter molecules diffuse to the postsynaptic side of the synapse where they bind to receptors in the membrane of the postsynaptic neuron (Fig. 1). Binding of neurotransmitter at the receptors opens ion channels in the membrane, so that, shortly after $t^{\mathrm{pre}}$ an ion current starts to flow into the postsynaptic neuron. Once opened, ion channels stay open for some time before they close stochastically. If the average open time is $\tau_{\mathrm{syn}}$, the total ion current will decay exponentially for $t > t^{\mathrm{pre}}$ and can be expressed as

$$I_{\mathrm{syn}}(t) = G \exp(-\frac{t - t^{\mathrm{pre}}}{\tau_{\mathrm{syn}}}) \left[u(t) - E_{\mathrm{syn}}\right] \tag{1.1}$$

where $u(t)$ is the postsynaptic membrane potential at the location of the synapse and $E_{\mathrm{syn}}$ the reversal potential of the synapse. The factor $G$ can be thought of as the 'strength' or 'weight' of the synapses and will be the focus of the following discussions. Changes of $G$ induced by the activity of pre- and postsynaptic neurons are referred to as 'synaptic plasticity' and are the topic of this chapter. The synaptic strength represented by $G$ depends on several factors including the number of postsynaptic receptors and the maximal conductance of the associated ion channels, the number of presynaptic vesicles available for release, and the probability that these vesicles actually are released when
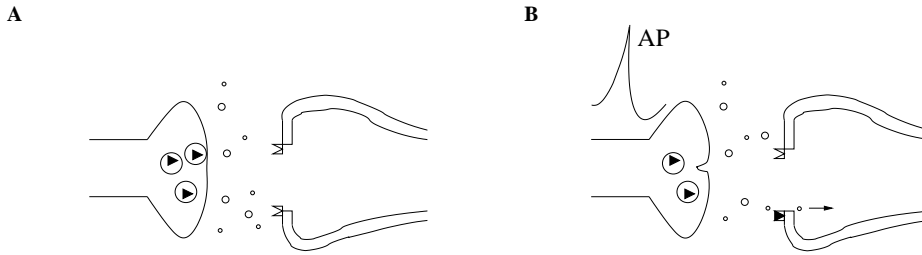
**A**                                                     **B**



Fig. 1. Chemical Synapse. **A**. Vesicles in the presynaptic terminal contain neurotransmitter molecules (filled triangles). **B**. If an action potential (AP) arrives at the presynaptic terminal, the vesicle merges with the cell membrane, transmitter diffuses into the synaptic cleft, binds to receptors on the postsynaptic membrane and triggers channel opening.

a presynaptic action potential occurs. The total postsynaptic receptor conductance is often what people mean when they say synaptic strength, but presynaptic factors are important as well and have interesting dynamic implications.

In order not to clutter the discussion with references, citations will be given here and not repeated. The basic binomial framework for modeling synaptic transmission is classic, and the reader can consult any text or review on synaptic physiology for an account of it. Standard rate-based learning rules related to Hebb's proposal in 1949 are found in Sejnowski (1977), Bienenstock et al. (1982), Oja (1982), Kohonen (1984), Sejnowski and Tesauro (1989) and many others. The problem of long-term stability has been emphasized by Grossberg (1987) and Fusi et al. (2000) among others. The analysis of correlation-based Hebbian plasticity for linear neurons is covered in any textbook on neural networks. The role of normalization and constraints is discussed in Miller and MacKay (1994).

A review of the relevant experimental results and references to the original work on STDP are contained in Abbott and Nelson (2000) or Bi and Poo (2001). More recent references include Froemke and Dan (2002) and Söjström et al. (2003). The theoretical framework for modeling short-term depression of synapses presented here is developed in Abbott et al. (1997) and Tsodyks and Markram (1997), and the relationship of STDP to short-term depression is analyzed in Senn et al. (2001). The approach described for treating STDP appears in Kempter et al. (1999 and 2001), Kistler and van Hemmen (2000), Rubin et al. (2001), Van Rossum et al. (2000), Gütig et al. (2002), and Câteau et al. (2003), and results arising from this and other analyses include Minai and Levy (1993), Blum and Abbott (1996), Gerstner et al. (1996), Abbott and Blum (1996), Gerstner and Abbott (1997), Roberts (1999), Song et al. (2000), Roberts and Bell (2000), Mehta et al. (2000), Song and Abbott (2001), and Rao and Sejnowski (2001). Models of STDP are reviewed in Gerstner and Kistler (2002).

Observations relevant to synaptic equalization and anti-STDP appear in Magee and Cook (2000), London and Segev (2001), Goldberg et al. (2002), and Rumsey and Abbott (2003).

.

## 2. The Binomial Synapse

We begin by reviewing a simple, standard model of presynaptic vesicle release that is relevant for any discussion of synaptic transmission. Imagine that there are $N_{\max}$ vesicle docking sites at a presynaptic terminal, and at any given time $N$ of them are occupied. Thus, $N$ is in the range $0 \leq N \leq N_{\max}$. Each vesicle has a probability $p$ of being released when there is a presynaptic action potential, and we assume each vesicle acts independently. In this case, the probability that $n$ vesicles are released in response to the presynaptic spike is given by a binomial distribution

$$P[n] = \frac{N!\, p^n\, (1-p)^{N-1}}{(N-n)!\, n!} \,.$$  (2.1)

The average number of vesicles released is then

$$\langle n \rangle = N\, p$$  (2.2)

and the trial-to-trial variance of the number of vesicles released is

$$\sigma_n^2 = N\, p\, (1-p)\,.$$  (2.3)

Unfortunately, $n$ cannot be measured experimentally. Instead, what can be determined is the change in postsynaptic conductance caused by transmitter released from the vesicles. This is a factor $g$, which is the conductance activated in the postsynaptic neuron per vesicle released, times the number of vesicles released. Denoting the total postsynaptic conductance by $G$, we have $G = gn$, or

$$\langle G \rangle = g\,N\,p \tag{2.4}$$

and, for the variance of $G$,

$$\sigma_G^2 = g^2\,N\,p\,(1-p)\,. \tag{2.5}$$

Note that the postsynaptic conductance activated by a presynaptic action potential depends on three parameters $p$, $N$, and $g$. The first two of these refer to properties of the presynaptic side of the synapse, whereas the last refers to the postsynaptic side. To isolate presynaptic dependences, experimentalists often quote the quantity

$$\frac{1}{\mathrm{CV}^2} = \frac{\langle G \rangle^2}{\sigma_G^2} = \frac{Np}{1-p} \tag{2.6}$$

where CV is the coefficient of variation, which is the ratio of the mean of $G$ to its standard deviation. Mechanisms of synaptic plasticity can target $g$, in which case they are called postsynaptic, or either $p$ or $N_{\mathrm{max}}$, in which case they are called presynaptic.

## 3.  Short-Term Depression through Vesicle Depletion

In the previous section, we showed how the postsynaptic conductance activated by a presynaptic action potential depends on $g$, $p$, and $N$, but we did not discuss the relationship between $N$, the actual number of vesicles available to release, and $N_{\mathrm{max}}$, the maximum number that can be available. This is the the subject we now address.

Each time a presynaptic action potential occurs and $n$ vesicles are released, the number of available vesicles is reduced, $N \to N - n$. Because $\langle n \rangle = Np$, we have, on average for each presynaptic action potential,

$$N \to N - Np = N(1-p)\,. \tag{3.1}$$

Vesicles are continually replenished at the synapse, so $N$ increases steadily back to $N_{\mathrm{max}}$ at a certain rate to oppose the depletion of vesicles due to presynaptic action potentials.

It is convenient to parameterize the number of available vesicles in terms of the percentage of available vesicles,

$$D = \frac{N}{N_{\mathrm{max}}} \tag{3.2}$$

which varies from 0 to 1 and acts as a synaptic depression factor. On average, when a presynaptic action potential occurs,

$$D \to D\,(1-p) \quad \text{or} \quad \Delta D = -p\,D \tag{3.3}$$

In addition, $D \to 1$, as vesicles are replenished. If presynaptic action potentials occur at a rate $\nu$, we can summarize these processes by the following differential equation:

$$\tau_D \frac{dD}{dt} = 1 - D - \tau_D p\,D\,\nu\,. \tag{3.4}$$

The last term on the right side of this equation represents the loss of vesicles through release, and the remaining terms represent replenishment. The constant $\tau_D$ determines the rate at which vesicles are replenished.

For constant firing rate $\nu$, Eq. (3.4) has the steady-state solution

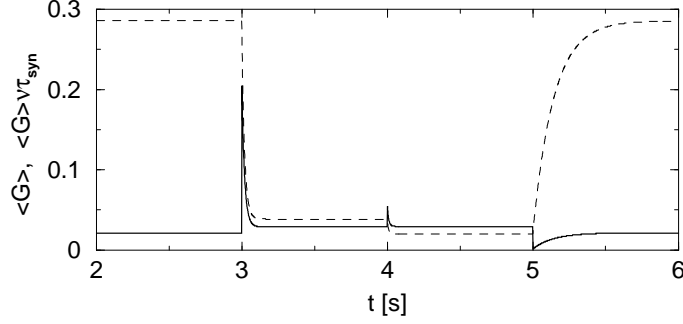$$D = \frac{1}{1 + \tau_D p\,\nu}\,, \tag{3.5}$$

Fig. 2. Short-Term Depression. The factor $\langle G \rangle = gN_{\max}D$ (dotted line) with $D$ given by Eq. (3.4) as well as the average conductance $\langle G \rangle \nu \tau_{\mathrm{syn}}$ (solid line) as a function of time. The neuron is stimulated by stationary rates that are switched at time $t = 3$s from $\nu = 5$Hz to 50Hz, one second later to 100Hz, and then one second later back to 5Hz. Note that, after a transient, the average postsynaptic conductance at 100Hz is the same as during 50Hz stimulation. Parameters: $g = N_{\max} = p = 1$, $\tau_D = 0.5$s, $\tau_{\mathrm{syn}} = 15$ms.

which has some surprising properties. First, recall that the total postsynaptic conductance is given by $\langle G \rangle = gNp = gN_{\max}Dp$. $D$ is called a depression factor because it measures how far $\langle G \rangle$ is below its maximum value $gN_{\max}p$. Using Eq. (3.5), we find

$$\langle G \rangle = gN_{\max}Dp = \frac{g\,N_{\max}\,p}{1 + \tau_D\,p\,\nu} \,. \tag{3.6}$$

For large presynaptic rates (rates for which $\nu\tau_D \gg 1/p$, this approaches

$$\langle G \rangle = \to \frac{g\,N_{\max}}{\tau_D\,\nu} \tag{3.7}$$

which is independent of $p$. This is an important point, *the strength of a synapses (which is what $G$ characterizes) at high presynaptic firing rates is independent of the single-vesicle release probability $p$.* The second point concerns the total conductance arising from a large number of synapses similar to the one we have been discussing. The average postsynaptic conductance per synapse due to such a population of synapses is $\langle G \rangle \nu \tau_{\mathrm{syn}}$, where $\tau_{\mathrm{syn}}$ is the synaptic time constant. At high presynaptic firing rates, we have

$$\langle G \rangle\,\nu\,\tau_{\mathrm{syn}} \to \frac{gN_{\max}\nu\,\tau_{\mathrm{syn}}}{\tau_D\,\nu} = \frac{gN_{\max}\tau_{\mathrm{syn}}}{\tau_D} \,. \tag{3.8}$$

Thus, *at high presynaptic rates the average postsynaptic conductance is independent of both $p$ and the presynaptic firing rate $\nu$.* This is illustrated in Fig. 2.

## 4. Hebbian Long-Term Plasticity

While synaptic changes induced by short-term depression decay rapidly, so that synapses return to their normal value on a time scale of tens or hundreds of milliseconds, there are other types of synaptic modification that remain for hours, days, and potentially even years. These modifications are called long-term potentiation (LTP) and long-term depression (LTD), and they are thought to be the neuronal correlate of learning. Over the last 50 years, a large body of experimental and theoretical work on synaptic plasticity and learning has been inspired by Hebb's postulate :

When an axon of cell $A$ is near enough to excite cell $B$ or 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.

An illustration of this principle is given in Fig. 3. Today Hebb's postulate is often rephrased in the sense that modifications in the synaptic transmission efficacy are driven by the correlations in the firing activity of pre- and postsynaptic neurons. Even though the idea of learning through correlations dates further back in the past, correlation-based learning rules are now generally called Hebbian learning.
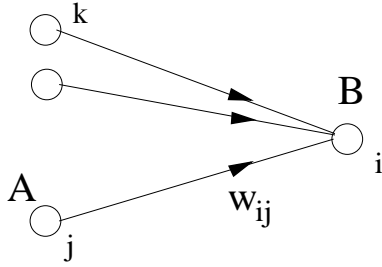
Fig. 3. The change at synapse $w_{ij}$ depends on the state of the presynaptic neuron $j$ and the postsynaptic neuron $i$ and the present efficacy $w_{ij}$, but not on the state of other neurons $k$.

Hebb formulated his principle on purely theoretical grounds. He realized that such a principle would help to stabilize specific neuronal activity patterns in the brain. If neuronal activity patterns correspond to behavior, then stabilization of specific patterns implies learning of specific behaviors. Even though Hebb's statement was essentially a theoretical one, he did not formulate it himself in mathematical terms. We discuss first classical formulations of Hebbian learning based on firing rates, before we turn to the more interesting case of spike-timing dependent plasticity (STDP).

## 5. Rate-Based Hebbian Learning

In rate-based neuron models, the activity of a given neuron $i$ is described by its firing rate $\nu_i$ which is related to the 'membrane potential' $u_i$ by a nonlinear monotonously increasing function $g$, i.e.,

$$\nu_i = g(u_i) \, . \tag{5.1}$$

Although we refer to $u_i$ as the membrane potential, it is not the actual membrane potential of the neuron (because this is oscillating rapidly as the neuron fires action potentials), but rather a time-average of the actual membrane potential or, equivalently, some measure of the total synaptic current. This time-average membrane potential can be estimated from the presynaptic firing rates $\nu_j$ and the synaptic weights $w_{ij}$ as $u_i = \sum_j w_{ij} \nu_j$. In the following we assume that the firing rates $\nu_i, \nu_j$ of pre- and postsynaptic neurons are constant during one trial of an experiment. For several classical experiments on long-term potentiation (LTP), this is a reasonable assumption. LTP can, for example, be introduced by high-frequency trains of presynaptic pulse at several synapses during a time $\mathcal{T}$. In such a situation the temporal resolution is rather coarse and a description of pre- and postsynaptic activity by fixed rates is appropriate. Time-dependent stimulation paradigms will be treated in the context of the spike-based formulation outlined in Section 6.

We consider the weight change $\Delta w_{ij}$ during one learning trial. Since the total weight change during a trial depends on the duration of the trial, we focus on the rate of change $\mathrm{d}w_{ij}/\mathrm{d}t = \Delta w_{ij}/\mathcal{T}$.

According to Hebb's postulate, the weight should increase if, during an experimental trial, both neurons are active together. Can we give a more precise mathematical description of Hebb's ideas? Our aim is to formulate the weight change during Hebbian learning as a function of the activity of pre- and postsynaptic neurons. From our point of view, at least six aspects are important for the formulation of a useful plasticity model.

(i) *Locality*. The learning rule for the synapse $w_{ij}$ connecting neuron $j$ to neuron $i$ should depend only on the activity of $j$ and $i$ and not on the state of other neurons $k \neq i, j$. In a rate model the state of neuron $i$ is fully characterized by its firing rate $\nu_i$. Given $\nu_i$, the membrane potential $u_i$ follows from Eq. (5.1), i.e., $u_i = g^{-1}(\nu_i)$ where $g^{-1}$ denotes the inverse of the transfer function $g$. The only other variable in a rate model that is locally available at the synapse is the present value $w_{ij}$ of the synaptic efficacy itself. Mathematically, we may therefore write

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = F(w_{ij}; \nu_i, \nu_j) \tag{5.2}$$

where $F$ is a yet unspecified function. Hebb's formulation is clearly consistent with the locality requirement.

(ii) *Cooperativity*. Hebb's formulation 'takes part in firing it' implies that both presynaptic *and* postsynaptic neuron must be active to induce a weight increase. It furthermore suggests a causal relationship between the firings. We will

| post $i$ | pre $j$ | $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \propto$ $\nu_i \, \nu_j$ | $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \propto$ $\nu_i \, \nu_j - c_0$ | $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \propto$ $(\nu_i - \nu_\theta) \, \nu_j$ | $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \propto$ $\nu_i \, (\nu_j - \nu_\theta)$ | $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \propto$ $(\nu_i - \langle \nu_i \rangle)(\nu_j - \langle \nu_j \rangle)$ |
|---|---|---|---|---|---|---|
| ON | ON | + | + | + | + | + |
| ON | OFF | 0 | - | 0 | - | - |
| OFF | ON | 0 | - | - | 0 | - |
| OFF | OFF | 0 | - | 0 | 0 | + |

Table 1

The change $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}$ of a synapse from $j$ to $i$ for various Hebb rules as a function of pre- and postsynaptic activity. 'ON' indicates a neuron firing at maximal rate ($\nu = \nu^{\mathrm{max}}$), whereas 'OFF' means an inactive neuron ($\nu = 0$). From left to right: standard Hebb rule, Hebb with decay, pre- and postsynaptic gating, covariance rule. The parameters satisfy $0 < \nu_\theta < \nu^{\mathrm{max}}$ and $0 < c_0 < (\nu^{\mathrm{max}})^2$.

discuss an implementation of causality in Section 6. For the moment, we restrict ourselves to the requirement of *simultaneous activity* of pre- and postsynaptic neuron. How can we implement this information in the function $F$ in Eq. (5.2)? $F$, which is a function of the rates $\nu_i$ and $\nu_j$, can be expanded in a Taylor series about $\nu_i = \nu_j = 0$. An expansion to second order in the rates yields

$$
\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \approx c_2^{\mathrm{corr}}(w_{ij}) \, \nu_i \, \nu_j + c_2^{\mathrm{post}}(w_{ij}) \, \nu_i^2 + c_2^{\mathrm{pre}}(w_{ij}) \, \nu_j^2
$$
$$
+ c_1^{\mathrm{pre}}(w_{ij}) \, \nu_j + c_1^{\mathrm{post}}(w_{ij}) \nu_i + c_0(w_{ij}) + \mathcal{O}(\nu^3). \tag{5.3}
$$

The first term on the right side of (5.3), picks up the correlations between pre- and postsynaptic activity. In fact, it is this bilinear term $\nu_i \, \nu_j$, viz., a term that is sensitive to the *correlations* between presynaptic and postsynaptic activity, that makes Hebbian learning a useful concept. The simplest implementation of Hebbian plasticity would be to require $c_2^{\mathrm{corr}} > 0$ and set all other parameters in the expansion (5.3) to zero

$$
\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_2^{\mathrm{corr}}(w_{ij}) \, \nu_i \, \nu_j \, . \tag{5.4}
$$

Eq. (5.4) with fixed parameter $c_2^{\mathrm{corr}} > 0$ is the prototype of Hebbian learning. A learning rule with $c_2^{\mathrm{corr}} < 0$ is usually called anti-Hebbian. We note that, if we continue the expansion on the right side of Eq. (5.3), more and more complex learning rules can be constructed. The next terms would be of order $\nu^3$, e.g., terms of the form $\nu_i \, \nu_j^2$, $\nu_j \, \nu_i^2$, etc.

*(iii) Synaptic Depression*. Hebb's original proposal gives no rule for a decrease of synaptic weights, but only refers to the conditions under which a strengthening of synapses should occur. It is clear, however, that a system where synapses can only increase and never decrease, is bound to be useless. An option for decreasing the weights (synaptic depression) is therefore a necessary requirement for any useful learning rule. This can, for example, be achieved by a weight decay, i.e., we take the parameter $c_0$ in Eq. (5.3) as

$$
c_0(w_{ij}) = -\gamma_0 \, w_{ij} \tag{5.5}
$$

with $\gamma_0 > 0$. Synaptic depression can also be implemented by several other combinations of the factors $c_1^{\mathrm{post}}$, $c_1^{\mathrm{pre}}$, and $c_0$. For example, the rule $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = (\nu_i - \nu_\theta) \, \nu_j$ could be implemented by the choice $c_2^{\mathrm{corr}} = 1, c_1^{\mathrm{pre}} = -\nu_\theta < 0$, and all other parameters equal to zero. Such a rule is called presynaptically gated, because presynaptic activity is a necessary requirement for any change; the activity level of the postsynaptic neuron determines the direction of the change. An overview of various possibilities of implementing synaptic depression in the framework of Eq. (5.3) is given in Table 1.

*(iv) Boundedness*. In any reasonable rules, weights at excitatory synapses should remain bounded in a range $0 \leq w_{ij} \leq w^{\mathrm{max}}$ where $w^{\mathrm{max}}$ is the maximal weight value that is sustainable by the bio-chemical machinery implementing the synaptic connection. To achieve boundedness, we make use of the dependence of the parameters in Eq. (5.3) on $w_{ij}$. Because $F$ on the right side of Eq. (5.2) can be a function of $w_{ij}$, the expansion coefficients $c_2^{\mathrm{corr}}, c_2^{\mathrm{post}}, c_2^{\mathrm{pre}}, c_1^{\mathrm{post}}, c_1^{\mathrm{pre}}, c_0$ can also depend on the weight $w_{ij}$. A suitable choice of the $w_{ij}$ dependence of the positive parameters guarantees that the weight cannot get larger than an upper bound $w^{\mathrm{max}}$; similarly, a suitable choice of the negative parameters assures that the $w_{ij}$ cannot decrease below zero.

In a simple formulation of a 'saturating' weight dependence, we take the parameter $c_2^{\mathrm{corr}}$ in Eq. (5.3) as

$$c_2^{\mathrm{corr}}(w_{ij}) = \eta_0 \left( w^{\mathrm{max}} - w_{ij} \right) \tag{5.6}$$

with a constant $\eta_0 > 0$. The factor $(w^{\mathrm{max}} - w_{ij})$ can be seen as an implementation of 'soft' bounds. The closer a weight is to its maximum the smaller the effect of an experimental trial with an LTP-induction protocol. As a modeling alternative to the 'soft' bounds, we can also use 'hard' bounds: growth of the synaptic weights has a constant factor $c_2^{\mathrm{corr}}$ as long as $w_{ij} < w^{\mathrm{max}}$ and stops if $w_{ij} = w^{\mathrm{max}}$. Hence, hard bounds correspond to the replacement $(w^{\mathrm{max}} - w_{ij}) \longrightarrow \Theta(w^{\mathrm{max}} - w_{ij})$. Here $\Theta(\cdot)$ denotes the Heaviside step function. Similarly, in a hard-bound formulation weight decrease would simply stop at $w_{ij} = 0$ whereas in the soft-bound formulation all negative factors would be taken as proportional to $w_{ij}$.

*(v) Competition.* A further useful feature of learning rules is *competitiveness*. If some weights grow, they do so at the expense of others that must decrease. Ideally, competitiveness should be a *consequence* of the learning rule (5.3) and should not require any additional assumptions. One specific implementation of competitiveness relies on the normalization of the set of weights $w_{ij}$ of all synapses converging onto the same postsynaptic neuron. While, at a first glance, such a normalization step would appear to violate the requirement of locality, it can in fact be realized using a purely local rule. An example is Oja's rule

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_2^{\mathrm{corr}} \nu_i \nu_j - \gamma w_{ij} \nu_i^2 \tag{5.7}$$

which is found from Eq. (5.3) if we take $c_2^{\mathrm{post}} = -\gamma w_{ij}$, keep $c_2^{\mathrm{corr}} > 0$, and set all other parameters to zero, $c_0 = c_1^{\mathrm{pre}} = c_2^{\mathrm{pre}} = c_1^{\mathrm{post}} = 0$.

*(vi) Long-term stability.* Most of the learning theories concentrate on the *induction* of weight changes. Once the 'learning session' is over, weights are taken as fixed parameters. Most neural systems, however, are subject to constantly changing input. If systems continue to remain adaptive, care must be taken that previously learned information is not lost. Grossberg has coined the term 'stability-plasticity dilemma' for this problem. A simple overwriting of previously stored information, sometimes called a 'palimpsest' property, must be avoided.

To approach this problem, we need to consider the *consolidation* of weights. Consolidation of previously learned items is conceivable with a weight dynamics that converges to binary weight values $w_{ij} = 0, 1$. In our framework, such a dynamics can be implemented by setting

$$c_0(w_{ij}) = -\gamma w_{ij} \left( 1 - w_{ij} \right) \left( w_\theta - w_{ij} \right) \tag{5.8}$$

where $0 < w_\theta < 1$ and $\gamma > 0$. Small weights $w_{ij} < w_\theta$ decay to zero. Large weights $w_{ij} > w_\theta$ increase towards one.

If Eq. (5.8) is inserted into Eq. (5.3), the effects of learning persist (or are even increased) after the end of a learning trial. A combination of Eqs. (5.8) with other linear and second-order terms in the plasticity equation (5.3) can therefore be considered as a model of consolidation of synaptic plasticity. In many formulations of synaptic plasticity, the problem of weight consolidation and overwriting are ignored or treated insufficiently.

**BCM rule and Covariance rule**. While the requirements (i) - (vi) pose a number of constraints for the formulation of learning rules, the framework sketched in Eq. (5.3) is general enough to classify various well-known learning rules. The Bienenstock-Cooper-Monroe (BCM) rule, for example,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = \eta \, \phi(\nu_i - \nu_\theta) \, \nu_j - \gamma w_{ij} \tag{5.9}$$

is obtained if we expand the function $F$ in Eq. (5.2) to linear order in the presynaptic rate $\nu_j$ while keeping all higher-order terms in the postsynaptic variable $\nu_i$. The function $\phi$ can in fact be identified with $\mathrm{d}F/\mathrm{d}\nu_j$ evaluated at $\nu_j = 0$. In the BCM-theory, $\phi$ and $\nu_\theta$ are chosen so that $\phi(0) = 0$ and $\phi'(0) > 0$, that is, $\phi(x)$ has a zero-crossing at $x = 0$ with positive slope; e.g., $\phi(x) = x - x^3 + \dots$. It is easy to demonstrate that an output rate $\nu_i = \nu_\theta$ is an unstable fixed point under the dynamics (5.9). To see this, we simply note that for fixed input rates $\nu_j > 0$ and monotone gain function $g$, all weights increase without bounds if $\nu_i > \nu_\theta$. In order to avoid unlimited growth of weights, $\nu_\theta$ is therefore taken in the BCM theory as an adaptive parameter $\nu_\theta = f(\langle \nu_i \rangle)$, where $\langle \nu_i \rangle$ is a short-term average of the output rate. A suitable functional dependence is $\nu_\theta = \langle \nu_i \rangle^2 / \nu_0$. The mean firing $\nu_i$ will then be attracted towards a value in the neighborhood of $\nu_0$. The notion of a running short-term average goes beyond the current framework, since it requires some memory, i.e., it is non-local in time. We will see, however, that it can be incorporated in the more general framework that is developed below for STDP.

The short-term average plays also a role in Sejnowski's covariance rule

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = (\nu_i - \langle \nu_i \rangle) \, (\nu_j - \langle \nu_j \rangle) \, . \tag{5.10}$$

It is based on an expansion of $F$ about a running average of $\nu_i$ and $\nu_j$, rather than on an expansion about zero. For fixed $\langle \nu_j \rangle$ and $\langle \nu_i \rangle$, an identification with Eq. (5.3) is possible and yields $c_0 = \langle \nu_i \rangle \langle \nu_j \rangle$, $c_1^{\mathrm{pre}} = -\langle \nu_i \rangle$ $c_1^{\mathrm{post}} = -\langle \nu_j \rangle$, and $c_2^{\mathrm{corr}} = 1$.

## 6. Spike-based Hebbian Plasticity

### 6.1. A theoretical framework

The approach taken in this section can be seen as a generalization of the Taylor expansion in the rate model of Section 5 to the case of spiking neurons. We recall that we started our formulation of rate-based Hebbian learning from a general formula

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = F(w_{ij}; \nu_i, \nu_j) \tag{6.1}$$

where weight changes are given as a function of the weight $w_{ij}$ as well as of the pre- and postsynaptic rates $\nu_j, \nu_i$. The essential assumption was that neuronal activity is characterized by firing rates that change slowly enough to be considered as stationary. Hebbian rules followed then from a Taylor expansion of Eq. (6.1). In the following, we keep the idea of an expansion, but drop the assumption that the output of a neuron can described solely by its firing rate.

For spiking neurons, the weight change observed after activity of duration $\mathcal{T}$, depends on the relative timing of pre- and postsynaptic spikes as well as of the total number of spikes involved. Let us denote the presynaptic spike train by $S_j(t) = \sum_f \delta(t - t_j^{(f)})$ and the postsynaptic one by $S_i(t) = \sum_f \delta(t - t_i^{(f)})$. In general the weight change $\Delta w_{ij}$ can be described by

$$\Delta w_{ij} = F_{\mathcal{T}}[w_{ij}; S_i(t'), S_j(t'')] \tag{6.2}$$

where $F_{\mathcal{T}}$ is a *functional* of the pre- and postsynaptic spike trains in the interval $\mathcal{T}$. Our notation with $t'$ and $t''$ is intended to indicate that the weight changes do not only depend on the momentary situation at time $t$, but also on the spiking history $t' < t$ and $t'' < t$. The weight value $w_{ij}$ and the local value of pre- and postsynaptic membrane potential are the essential variables that are available at the site of the synapse to control the up- and down-regulation of synaptic weights. In detailed neuron models, $F_{\mathcal{T}}$ would depend not only on the weight $w_{ij}$ and membrane potentials, but also on all other variables that are locally available at the site of the synapse. In particular, there could be a dependence upon the local calcium concentration and the time course of the membrane potential. In the following we adopt the point of view that the calcium concentration and membrane potential are largely determined by the previous firing history, so that there is no need to introduce additional explicit variables for them.

In analogy to the approach taken for rate based plasticity, we now expand the right side of Eq. (6.2) in terms of spikes. We start with a drift term $c_0$ in the absence of spikes the effect of which accumulates over the interval $\mathcal{T}$, then we add the effect of isolated spikes, pairs of spikes, triplets of spikes, and so forth. The first few terms are

$$\Delta w_{ij} = c_0(w_{ij}) \, \mathcal{T} + \sum_{t_j^{(f)} \in [t, t+\mathcal{T})} c_1^{\mathrm{pre}} + \sum_{t_i^{(f)} \in [t, t+\mathcal{T})} c_1^{\mathrm{post}} + \sum_{t_i^{(f)}, t_j^{(f)} \in [t, t+\mathcal{T})} W(t_j^{(f)} - t_i^{(f)}) + \dots \, . \tag{6.3}$$

The next terms would contain at least two presynaptic or postsynaptic spikes and have been neglected. Eq. (6.3) is the central result of this section. It provides a framework for the formulation of spike-based learning rules and may be seen as the generalization of the general rate-based model that we have derived in Section 5. The considerations summarized in points (iii) - (vi) of the preceding section apply also to the case of spike-based learning rules and will not be repeated here.
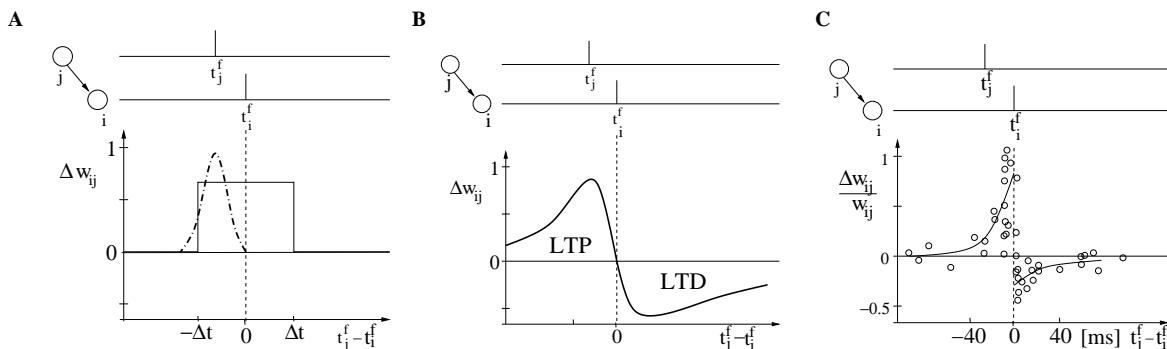
Fig. 4. Learning window. The change $\Delta w_{ij}$ of the synaptic efficacy depends on the timing of pre- and postsynaptic spikes. **A**. The solid line indicates a rectangular time window as is often used in standard Hebbian learning. The synapse is increased if pre- and postsynaptic neuron fire sequentially with an interspike interval smaller than $\Delta t$. The dot-dashed line shows an asymmetric learning window useful for sequence learning. The synapse is strengthened only if the presynaptic spike arrives slightly before the postsynaptic one and is therefore partially 'causal' in firing it. **B**. An asymmetric bi-phasic learning window similar to the one used in many modeling studies. A synapse is strengthened (long-term potentiation, LTP), if the presynaptic spike arrives slightly before the postsynaptic one, but is decreased (long-term depression LTD), if the timing is reversed. **C**. Experimental results have confirmed the existence of bi-phasic learning windows. Data points redrawn after the experiments of Bi and Poo (1998).

## 6.2. STDP

Spike-timing dependent plasticity, or STDP, is a form of long-term modification of synaptic strength that depends on the timing between pre- and postsynaptic action potentials. STDP is described by the function $W(t_{\mathrm{pre}} - t_{\mathrm{post}})$ in Eq. (6.3) that determines how the strength of a synapses is modified by a pair of action potentials, a presynaptic action potential occurring at time $t_{\mathrm{pre}}$ and a postsynaptic action potential occurring at time $t_{\mathrm{post}}$. This function, which is called the STDP window function, determines the fractional change in synaptic strength induced by such a pair. In principle, it could be symmetric, or asymmetric; positive only or biphasic with positive and negative parts (Fig. 4). The typical observed dependence can be approximated by

$$W(t_{\mathrm{pre}} - t_{\mathrm{post}}) = \begin{cases} A_+ \exp(-(t_{\mathrm{post}} - t_{\mathrm{pre}})/\tau_+) & \text{if } t_{\mathrm{pre}} < t_{\mathrm{post}} \\ A_- \exp(-(t_{\mathrm{pre}} - t_{\mathrm{post}})/\tau_-) & \text{if } t_{\mathrm{pre}} > t_{\mathrm{post}} . \end{cases} \qquad (6.4)$$

The parameters $\tau_-$ and $\tau_+$ determine the temporal ranges of the two sides of the window function, while $A_-$ and $A_+$ determine the nature and size of the changes induced by the appropriate spike pairings.

Standard STDP refers to the case $A_+ > 0$ and $A_- < 0$. In this case, the upper line on the right side of Eq. (6.4), corresponding to pairings in which the presynaptic spike occurs before the postsynaptic spike, produces a long-lasting strengthening of the synapses called long-term potentiation or LTP. The lower line in Eq. (6.4), corresponding to pairing in which the presynaptic spike occurs after the postsynaptic spike, produces a long-lasting weakening of the synapses called long-term depression or LTD.

SDTP is interesting, in part, because its timing dependence reflects a sensitivity to causality. Presynaptic action potentials that occur before the postsynaptic response are predictive of that response and therefore useful. It makes sense to strengthen such inputs, and this is what Fig. (4C) tells us happens. On the other hand, a presynaptic action potential that occurs after the postsynaptic response is clearly of no predictive value, and it makes sense to weaken such an input, which is also what happens.

In addition to STDP, we will also consider in Section 9 a reversed form of timing-dependent long-term plasticity called anti-STDP. For anti-STDP, $A_+ < 0$ and $A_- \geq 0$ so, in particular, pre-before-post ordering results in a weakening rather than a strengthening of the synapse.

## 7. Mean Field Approach for uncorrelated input

### 7.1. Weight evolution equation

Before we deal with the most general form of the learning rule, we will discuss the specific effects of the window function. The STDP window function indicates the amount that a synapse changes in strength as a function of a

single pair of pre- and postsynaptic spikes. In active neural circuits, such spike pairs will occur in large numbers and complex temporal patterns. Statistical methods are needed to compute the change in strength that will arise in such a situation. In particular, we make the approximation that the effects of different spike pairs add linearly. In this case, we can compute the rate at which a particular synapse from neuron $j$ to the postsynaptic neuron $i$ with strength $w_{ij}$, will change strength as the product of the STDP window function and a function that describes how often pairs of a particular temporal spacing occur. This function is the correlation function of the pre- and postsynaptic action potential sequences, which we denote by $\Gamma_{ij}$. Specifically, $\Gamma_{ij}(T)$ is the probability density for spike pairs separated by an interval $T = t_{\text{pre}} - t_{\text{post}}$. The rate of change of synaptic strength for synapse $j$ at time $t$ is given in terms of the correlation function $\Gamma_{ij}$ and the window function $F$ by

$$\frac{dw_{ij}}{dt} = \int_{-\infty}^{t} dt' W(t' - t)\Gamma_{ij}(t' - t) + \int_{-\infty}^{t} dt' W(t - t')\Gamma_{ij}(t - t'). \tag{7.1}$$

To proceed, we need a specific form for the correlation function so that we can calculate the integrals in Eq. (7.1). We will assume a simple form in which the correlation between pre- and postsynaptic spikes is described by two terms. The first term arises from chance pairings of pre- and postsynaptic spikes, which would occur even for a synapse of zero strength. This term is simply the product of the pre- and postsynaptic firing rates, $\nu_j \nu_i$. The second term arises from postsynaptic spikes evoked by the presynaptic spike and is proportional to the efficacy of the synapse. The efficacy, denoted by $w_{ij}^{\text{eff}}$, is a measure of the ability of the synapse to evoke a postsynaptic response. The efficacy $w_{ij}^{\text{eff}}$ is closely related to, and often assumed proportional to, the strength $w_{ij}$, but the two should be distinguished. Synaptic efficacy ($w_{ij}^{\text{eff}}$) is affected by global properties of a neuron such as the location of the synapses on the dendritic tree relative to the soma and the morphology and conductance profile of the neuron. Synaptic strength ($w_{ij}$) only depends on local factors, such as $g$, $N_{\text{max}}$, and $p$, as outlined in Section 2.

Putting the two terms discussed in the previous paragraph together we write

$$\Gamma_{ij}(T) = \nu_i \nu_j + w_{ij}^{\text{eff}} \nu_j C^{\text{spike}}(-T) \tag{7.2}$$

where, for example,

$$C^{\text{spike}}(T) = \frac{\Theta(T)}{\tau_C} \exp\left(-\frac{T}{\tau_C}\right). \tag{7.3}$$

The correlation time $\tau_C$ determines the time scale over which a presynaptic spike affects the postsynaptic response. We assume an exponential form just for simplicity.

Inserting Eqs. (7.2) and (7.3) into Eq.7.1, we find

$$\frac{dw_{ij}}{dt} = A_- \int_{-\infty}^{t} dt' \exp((t' - t)/\tau_-)\Gamma_{ij}(t' - t) + A_+ \int_{-\infty}^{t} dt' \exp((t' - t)/\tau_+)\Gamma_{ij}(t - t') \tag{7.4}$$

or

$$\frac{dw_{ij}}{dt} = (A_-\tau_- + A_+\tau_+)\nu_j\nu_i + \frac{A_+ w_{ij}^{\text{eff}} \nu_j}{\tau_C} \int_{-\infty}^{t} dt' \exp\left(-\frac{t - t'}{\tau_+}\right) \exp\left(-\frac{t - t'}{\tau_C}\right), \tag{7.5}$$

which gives

$$\frac{dw_{ij}}{dt} = \left((A_-\tau_- + A_+\tau_+)\nu_i + \frac{A_+\tau_+ w_{ij}^{\text{eff}}}{\tau_+ + \tau_C}\right)\nu_j. \tag{7.6}$$

To simplify the notation, we define

$$\alpha = -(A_-\tau_- + A_+\tau_+) \tag{7.7}$$

and

$$\beta = \frac{A_+\tau_+}{\tau_+ + \tau_C}, \tag{7.8}$$

**A**

**B**
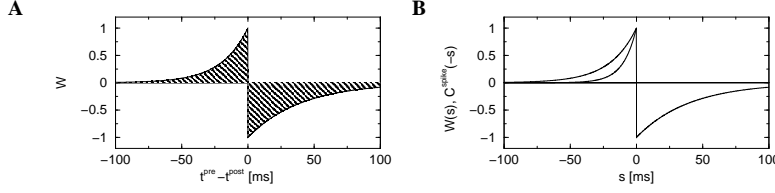


Fig. 5. STDP learning window with $A_+ = -A_- = 1$ and time constants $\tau_+ = 20$ms and $\tau_- = 40$ms. **A.** The integral over the learning window $\overline{W} = \int W(s)\,ds$ is negative so that $\alpha = -\overline{W} > 0$. **B.** The parameter $\beta$ is defined by the overlap between the window function (thin line) and the spike correlations, $\beta = \int W(s)\,C^{\mathrm{spike}}(-s)\,ds$ (thick solid line). Here we assumed a time constant $\tau_C = 10$ms for $C^{\mathrm{spike}}$.

so that Eq. (7.6) becomes simply

$$\frac{dw_{ij}}{dt} = (\beta w_{ij}^{\mathrm{eff}} - \alpha\nu_i)\nu_j \,. \tag{7.9}$$

This is the basic result that we will analyze. Note that $\alpha = -\int W(s)\,ds$ so that $\alpha$ is positive if the integral over the STDP window is negative as in Fig. 5. The integral over the window function will occasionally also be denoted by $\overline{W}$.

### 7.2. Constrained Hebbian Plasticity from STDP

We first consider the case of STDP for which $\beta > 0$, but we also impose the condition that $\alpha > 0$, which does not follow automatically, but agrees with the data. Given that $w_{ij}^{\mathrm{eff}}$ is a monotonically increasing function of $w_{ij}$, it is immediately obvious that Eq. (7.9) is unstable for fixed postsynaptic rate $\nu_i$, making $w_{ij}$ either increase or decrease without bound as a function of time.

The instability of Eq. (7.9) is a typical feature of Hebbian plasticity schemes; it is what makes them capable of learning, which is essentially the amplification of small changes in input features into large state changes. To control this instability, we require the synaptic strength to be bounded in the range $0 \leq w_{ij} \leq w_{\max}$, for some fixed value of $w_{\max}$. The bounds are imposed by setting any strength that attempts to move beyond a bound equal to that bound.

Accompanying the local instability discussed in the previous paragraph, Hebbian plasticity rules typically have a global instability. This is due to the fact that all synapses are subject to the same plasticity rule and are modified independently. If, as in the case we consider, they receive input with the same statistical structure (Poisson spike trains at a given rate), they should all undergo similar modifications. Thus, in a typical Hebbian scheme, we would expect all the synapses to attain either values near 0 or values near $w_{\max}$. This causes the postsynaptic firing rate $\nu_i$ to go either to 0 or to some large value. An interesting feature of STDP is that this does not happen, even though this form of plasticity is locally unstable and acts independently at each synapse. Instead, if there are $N$ synapses for a particular postsynaptic neuron, these divide into two groups: a group of $xN$ synapses that take values near $w_{\max}$ and a group of $(1-x)N$ synapses that take values near 0. For a given partitioning, the postsynaptic neuron attains a firing rate that depends on $x$, $\nu_i(x)$, and this partitioning is stable around a particular value of $x$. For appropriate choice of parameters, this value of $x$ always produces a reasonable postsynaptic firing rate $\nu_i(x)$, no matter what values the presynaptic rates take. This is an unusual stabilizing or homeostatic role taken by a plasticity mechanism that is otherwise Hebbian. Furthermore, this mechanism introduces competitiveness between synapses.

To see how the stabilization discussed in the previous paragraph takes place, we note that Eq. (7.9) has an unstable fixed point at the value of synaptic strength, $w_{ij}$, that sets the efficacy of synapse $j$, $w_{ij}^{\mathrm{eff}}$ equal to the value

$$w_{\mathrm{sep}}^{\mathrm{eff}} = \frac{\alpha\nu_i}{\beta} \,. \tag{7.10}$$

This unstable fixed point acts as a separatrix in that synapses that have efficacies larger than this value get pushed to the maximum allowed strength, and synapses with efficacies lower than this value get pushed to zero. Mathematically, if $w_{ij}^{\mathrm{eff}} > w_{\mathrm{sep}}^{\mathrm{eff}}$, then $w_{ij} \to w_{\max}$; and if $w_{ij}^{\mathrm{eff}} < w_{\mathrm{sep}}^{\mathrm{eff}}$, then $w_{ij} \to 0$. So far, this is just the typical behavior of an unstable system, but it means that the partitioning variable $x$ discussed above is determined by $w_{\mathrm{sep}}^{\mathrm{eff}}$. Recalling that the postsynaptic rate $\nu_i$ depends on $x$, we find that equation (7.10) is actually a self-consistency condition for $w_{\mathrm{sep}}^{\mathrm{eff}}$,

$$w_{\mathrm{sep}}^{\mathrm{eff}} = \frac{\alpha\nu_i(x(w_{\mathrm{sep}}^{\mathrm{eff}}))}{\beta} \,. \tag{7.11}$$

The left side of this equation is obviously an increasing function of $w_{\mathrm{sep}}^{\mathrm{eff}}$, whereas the right side is a decreasing function of $w_{\mathrm{sep}}^{\mathrm{eff}}$ because high values of $w_{\mathrm{sep}}^{\mathrm{eff}}$ push synapses to low values and decrease the postsynaptic firing rate. Furthermore, provided that the neuron can fire at all, the left side is zero and the right side is positive for $w_{\mathrm{sep}}^{\mathrm{eff}} = 0$, the left side is positive, and the right side is zero for sufficiently large $w_{\mathrm{sep}}^{\mathrm{eff}}$. Therefore, this equation always has a solution that determines $w_{\mathrm{sep}}^{\mathrm{eff}}$, $x$, and $\nu_i$, and this solutions displays the features discussed above.

STDP involves a curious combination of an unstable form of plasticity that we would associate with learning, and a stable form of plasticity that we would call homeostatic. Individual synapses are forced to the extreme values of 0 or $w_{\mathrm{max}}$ by a Hebbian process that is highly sensitive to the correlation and temporal structure of the total synaptic input (see below). This corresponds to a Hebbian learning process with an added temporal wrinkle. The fraction $x$ of synapses that take the maximal value of strength is, on the other hand, modified in a homeostatic manner that regulates the firing rate of the postsynaptic neuron, keeping it in a reasonable operating range.

## 8. Detecting Correlations in the Input

Hebbian learning rules are sensitive to correlations, which we have not treated up to this point but will address in this section. In order to focus on the essential phenomenon, we start with the simplest possible plasticity rule, i.e., a standard rate model. We will then see that, by turning to STDP rules, several problems of the basic model can be solved while retaining the essential aspect of sensitivity to correlations.

### 8.1. Rate-Based Model

For the analysis of rate-based Hebbian learning, we suppose that presynaptic neurons $j$, $k$, have firing rates $\nu_j$ and $\nu_k$ with mean $\langle \nu_j \rangle = \langle \nu_k \rangle = \nu^{\mathrm{pre}}$ and synapse-to-synapse or 'spatial' correlation

$$C_{jk} = \langle \nu_k(t)\, \nu_j(t) \rangle. \tag{8.1}$$

As previously, angular brackets denote an expectation value. It is straightforward to see that rate-based learning rules are sensitive to the spatial correlations in the input. If the postsynaptic neuron is modeled as a linear unit

$$\nu_i(t) = \sum_k w_{ik} \nu_k(t), \tag{8.2}$$

the standard Hebbian learning rule $\mathrm{d}w_{ij}/\mathrm{d}t = c_2^{\mathrm{corr}}\, \nu_i\, \nu_j$ yields an expected weight evolution

$$\langle \frac{\mathrm{d}}{\mathrm{d}t} w_{ij} \rangle = c_2^{\mathrm{corr}} \sum_k w_{ik} \langle \nu_k(t)\, \nu_j(t) \rangle \tag{8.3}$$

which is in direction of the principal eigenvector of the spatial correlation matrix. If learning is slow[1], the actual weight vector stays close to the expected one, and the angular brackets on the left side of Eq. (8.3) can be dropped.

We introduce a vector $\vec{w}_i$ to describe the set of synapses that converge on the postsynaptic neuron $i$. The correlation between presynaptic neurons can be denoted by a matrix $C$ with components $C_{jk}$ defined in (8.1). In matrix notation, Eq. (8.3) is

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{w}_i = c_2^{\mathrm{corr}}\, C\, \vec{w}_i. \tag{8.4}$$

Because the correlation matrix is positive definite the weight vector grows exponentially in standard Hebbian learning ($c_2^{\mathrm{corr}} > 0$) (Fig. 6A).

To avoid unlimited growth of weights and to introduce competitiveness between the synapses converging onto the same postsynaptic neuron, Eq. (8.3) is usually combined with a suitable normalization procedure. As a first example, we consider Oja's rule, Eq. (5.7). Repeating the same arguments as before we get the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{w}_i = c_2^{\mathrm{corr}}\, C\, \vec{w}_i - \gamma\, [\vec{w}_i\, C\, \vec{w}_i]\, \vec{w}_i \tag{8.5}$$

---

[1]The weight vector should change only by small amount during the time needed to get a representative sample of the input statistics. This can always be achieved by taking $c_2^{\mathrm{corr}}$ to be small enough so that the time scale of learning and that of the input are well separated.
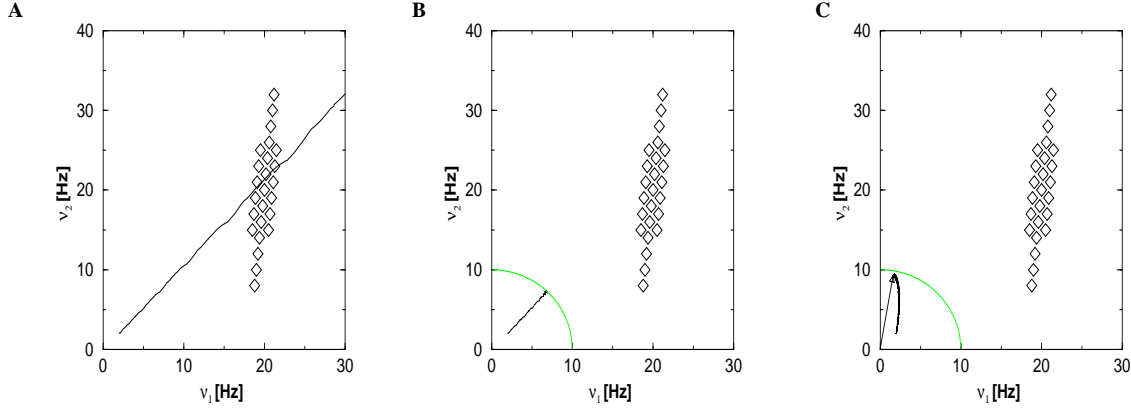
Fig. 6. Rate-based learning. The two presynaptic input rates $\nu_1$ and $\nu_2$ are chosen stochastically from a set of 25 firing patterns (diamonds). The mean rate averaged over all patterns is $\langle \nu_1 \rangle = \langle \nu_2 \rangle = 20$Hz. The initial value of the weights is $w_{i1} = w_{i2} = 2$. **A**. Standard correlation based learning (5.4) leads to a movement of the weight vector (solid line) in direction of the pattern cluster. The length of the weight vector grows without bound. **B**. Using Oja's rule (5.7), the weight vector moves in the same direction as before, but is limited to a maximum length of 10, indicated by the circle. **C**. The covariance rule (5.10) combined with Oja's normalization method yields a weight vector that points in direction of the principal axis of the data cloud.

with fixed points $\vec{w}_i = a\, e_k$ where $e_k$ is a normalized eigenvector of $C$ and $a = \sqrt{c_2^{\mathrm{corr}}/\gamma}$. It is possible to show that only the solution parallel to the eigenvector $e_1$ with maximal eigenvalue $\lambda_1 \geq \lambda_k$ for all $k$ is stable. Hence, during learning the weight vector $\vec{w}_i$ remains bounded and turns in direction of the dominant eigenvector of the input correlation matrix $C$. This is illustrated in Fig. 6B.

As a second example, we consider subtractive normalization. To see how this works, we assume that during a single trial of duration $\mathcal{T}$, the standard Hebbian rule $dw_{ij}/dt = c_2^{\mathrm{corr}} \nu_i \nu_j$ would yield a 'raw' weight change $\Delta \tilde{w}_{ij}$. In order to guarantee that the sum of the weights $\sum_k w_{ik}$ does not change, all weights are reduced *a posteriori* by an amount $N^{-1} \sum_k \Delta \tilde{w}_{ik}$ where $N$ is the number of synapses converging onto the same postsynaptic neuron $i$. Overall the two steps (i.e., 'raw' change and subsequent reduction) amount to a new learning rule

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_2^{\mathrm{corr}} \nu_i \nu_j - N^{-1} c_2^{\mathrm{corr}} \nu_i \sum_{k=1}^{N} \nu_k \,. \tag{8.6}$$

The first term on the right side of Eq. (8.6) is the standard learning rule and the second term the subtractive normalization term (Fig. 7A). If we write the postsynaptic rate as $\nu_i = \sum_k w_{ik} \nu_k$ and take, as before, the ensemble average, we find

$$\frac{\mathrm{d}}{\mathrm{d}t} \vec{w} = c_2^{\mathrm{corr}} [C - \overline{C}] \vec{w} \tag{8.7}$$

where $\overline{C}$ is a matrix with components $\overline{C}_{kj} = \sum_n C_{nj}$. Eq. (8.7) is the analog of Eq. (8.4) for the case of subtractive weight normalization. It is usually combined with hard bounds $0 \leq w_{ij} < w^{\mathrm{max}}$.

Apart from the fact that the analysis has so far been restricted to rate-based learning, there are two potential problems. First, the rule (8.6) is *non-local* because the weight change at the synapse from $j$ to $i$ depends on the firing rate of other presynaptic neurons $k$. We see in Section 8.3 that implicit subtractive normalization can also be achieved by *local* rules via a stabilization of postsynaptic firing rates. Second, plasticity is driven by correlations whereas it would seem more sensible to first subtract the mean input rates averaged over the statistical ensemble (which are typically less informative) and focus on the *covariance*. This could be achieved by replacing the standard Hebb rule by covariance rule (5.10) and postulate that the mean input and output rates are controlled by some homeostatic or adaptive mechanism (Fig. 6C). Again, we will see below that, for a suitable choice of parameters, STDP rules automatically subtract the mean rates so that the spike based learning rules are driven by the covariance of the input.
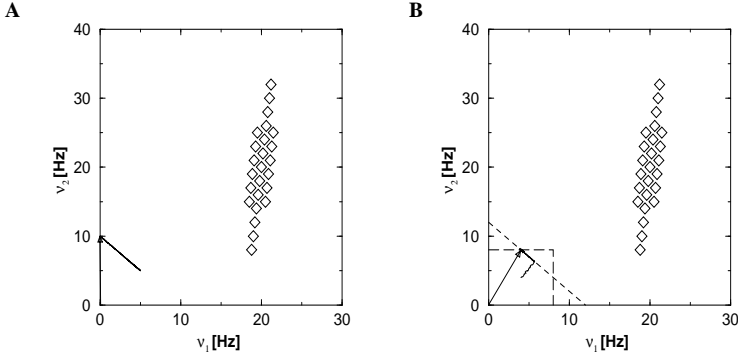
**A**



**B**

Fig. 7. Subtractive normalization in rate-based learning. **A**. Two presynaptic inputs with rate patterns indicated by the diamonds. The initial value of the weights is $w_{i1} = w_{i2} = 5$. During learning combined with subtractive normalization as in Eq. (8.6) the rate vector moves along a line defined by the normalization condition $w_{i1} + w_{i2} = const$. The final position of the weight vector is indicated by an arrow. **B**. Three presynaptic inputs. The first two components $\nu_1, \nu_2$ are the same as above while the third component has a fixed rate $\nu_3 = 20$Hz which is the same for all patterns. The initial value of the weights is $w_{i1} = w_{i2} = w_{i3} = 4$. During the first phase of learning, $w_{i3}$ decreases to zero (not shown) while the first two components $(w_{i1}, w_{i2})$ of the weight vector move towards the cloud of data points until they hit the normalization bound $w_{i1} + w_{i2} = 12$ (dashed diagonal line). Weight evolution then continues along the diagonal line until one of the weights has reached its maximum value $w_{\max} = 8$. The final position of the weight vector $(w_{i1}, w_{i2})$ is shown by an arrow. The hard bounds $w_{ij} \leq w_{\max}$ are indicated by the long-dashed line.
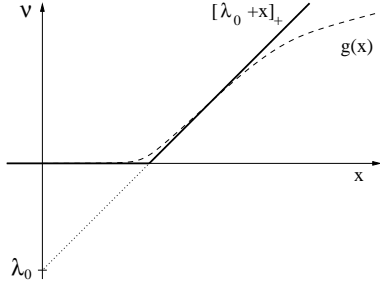


Fig. 8. Spike generation by the postsynaptic neuron is given by an inhomogeneous Poisson process with rate $\nu_i(t) = g[x(t)]$ which can be approximated by a piecewise linear function $\nu_i(t) = [\lambda_0 + x(t)]_+$. Note that in the example shown in the figure $\lambda_0$ is negative.

### 8.2. Spike-based Model

In order to extend the above results to spike-based learning, the linear rate model in Eq. (8.2) is replaced by a stochastically spiking neuron model with instantaneous firing rate

$$\nu_i(t) = \left[\lambda_0 + \rho_0 \sum_k w_{ik} \sum_f \epsilon(t - t_k^{(f)})\right]_+ = \left[\lambda_0 + \rho_0 \sum_k w_{ik} \int_0^\infty \epsilon(s) S_k(t - s) \, ds\right]_+ \tag{8.8}$$

where $t_k^{(f)}$ denotes the time of presynaptic spike arrival at the synapse from neuron $k$ to neuron $i$. $S_k(t) = \sum_f \delta(t - t_k^{(f)})$ is the presynaptic spike train. The time course of the postsynaptic potential is described by the function $\epsilon$ which is normalized to $\int_0^\infty \epsilon(s) ds = 1$. The bracket $[.]_+$ denote rectification, i.e., $[x]_+ = x$ for $x > 0$ and zero otherwise. The left side can be taken as the instantaneous rate (or stochastic intensity) of an inhomogeneous Poisson process. We call Eq. (8.8) a piecewise linear Poisson neuron. Note that the piecewise linear neuron model can be considered as an approximation to a general nonlinear Poisson model $\nu_i(t) = g(\sum_k w_{ik} \sum_f \epsilon(t - t_k^{(f)}))$(Fig. 8).

We assume that input spike trains have stationary statistics with known mean and correlations. Specifically, input spikes at a given synapse $j$ are generated by a doubly stochastic point process. Spikes are generated by an inhomogeneous Poisson process with instantaneous rate $\nu_j(t)$. The rate itself is drawn from a distribution with constant expectation value $\langle \nu_j(t) \rangle = \nu^{\text{pre}}$ and correlations by $\langle \nu_j(t) \nu_k(t') \rangle = C_{jk}(t - t')$. We suppose that all presynaptic spike trains have identical properties. In particular, we require that $\langle \nu_j(t) \rangle = \nu^{\text{pre}}$ independent of $j$ and

$N^{-1} \sum_{k=1}^{N} C_{jk}(s) = \overline{C}(s)$ independent of $j$. Some of these restrictions are not necessary but simplify the discussion. Since the expected input rates are constant, the expected output rate $\nu^{\text{post}} \equiv \langle \nu_i(t) \rangle = \nu^{\text{pre}} \sum_j w_{ij}(t)$ is a slowly varying function of time which changes only as the synaptic efficacies adapt.

In order to discuss the weight dynamics we start from the general spike-based learning equation that has been developed in Section 6. Throughout the following we assume that weight changes $\Delta w_{ij}/|w_{ij}|$ are small during the time that is necessary to approximately sample the input statistics. In this case, we can separate the time scale of learning from that of the neuronal dynamics. The right side of Eq. (6.3) is then 'self-averaging', so that the evolution of the weight vector (6.3) is given by

$$
\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}(t) = c_0 + c_1^{\text{pre}} \langle\langle S_j(t) \rangle\rangle + c_1^{\text{post}} \langle\langle S_i(t) \rangle\rangle
$$
$$
+ \int_{-\infty}^{\infty} W(s) \langle\langle S_j(t) \, S_i(t-s) \rangle\rangle \, \mathrm{d}s \; ; \tag{8.9}
$$

Double angular brackets in Eq. (8.9) denote the average over the input statistics defined by the above process. We first have to calculated the expected number of spikes, given the rates; and then we have to average over the rates. The double angular brackets on the left side have been dropped, because of the self-averaging property.

For an interpretation of Eq. (8.9) we start with the terms that are linear in the spike trains. We first calculate the expected number of spikes given the rates and take then the expectation of the rates

$$
\langle\langle S_j(t) \rangle\rangle = \langle \nu_j(t) \rangle = \nu^{\text{pre}} \tag{8.10}
$$
$$
\langle\langle S_i(t) \rangle\rangle = \lambda_0 + \rho_0 \sum_j w_{ij} \int_0^{\infty} \epsilon(s) \langle\langle S_j(t-s) \rangle\rangle \, \mathrm{d}s = \lambda_0 + \rho_0 \sum_j w_{ij}(t)\nu^{\text{pre}} = \nu^{\text{post}}(t) \tag{8.11}
$$

For the second line we have assumed that the fluctuations of the postsynaptic rate are small so that the neuron always stays in the linear regime. Finally, the term $\langle\langle S_j S_i \rangle\rangle$ on the right side of Eq. (8.9) describes the correlation between input and output *on the level of spikes*. We may summarize Eq. (8.9) by saying that the evolution of the weight vector is driven by the expected firing rates and by *correlations on the time scale of the learning window*.

The correlation term

$$
\langle\langle S_j(t) \, S_i(t-s) \rangle\rangle = \Gamma_{ij}(-s) \tag{8.12}
$$

depends not only on the input statistics, but also on the dynamics of the neuron model under consideration. Since we have assumed that learning is a slow process the correlation term can then be evaluated for constant weights $w_{ij}$, $1 \le i \le N$. As we have seen before, the correlations $\Gamma$ between pre- and postsynaptic spike trains can be written as

$$
\langle\langle S_j(t) \, S_i(t-s) \rangle\rangle = \langle \nu_i(t-s) \, \nu_j(t) \rangle + \langle \nu_j(t) \rangle \, w_{ij}^{\text{eff}}(t) \, C^{\text{spike}}(-s) \, , \tag{8.13}
$$

with a postsynaptic firing rate $\nu_i(t)$. If the linear Poisson neuron model is working in the linear regime, the efficacy is $w_{ij}^{\text{eff}}(t) = w_{ij}(t)$ and the spike-spike correlation term is $C^{\text{spike}}(s) = \rho_0 \, \epsilon(s)$, i.e., the temporal correlations between pre- and postsynaptic neuron reflect the time course of the postsynaptic potential. The same would hold for integrate-and-fire neurons in the high-noise limit, whereas the correlations of integrate-and-fire neurons are more complicated in the limit of low noise.

Substituting Eqs. (8.10) – (8.13) into Eq. (8.9) we find

$$
\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}(t) = c_0 + c_1^{\text{pre}} \nu^{\text{pre}} + c_1^{\text{post}} \nu^{\text{post}}(t) + w_{ij}^{\text{eff}} \nu^{\text{pre}} \int_{-\infty}^{0} W(s) \, C^{\text{spike}}(-s) \, \mathrm{d}s
$$
$$
+ \; \nu^{\text{post}}(t-s) \, \nu^{\text{pre}} \int_{-\infty}^{\infty} W(s) \, \mathrm{d}s
$$
$$
+ \; \int_{-\infty}^{\infty} W(s) \, \langle (\nu_i(t-s) - \nu^{\text{post}}(t-s)) \, (\nu_j(t) - \nu^{\text{pre}}) \rangle \mathrm{d}s \tag{8.14}
$$

Here $\nu^{\mathrm{post}}(t) = \langle \nu_i(t) \rangle$ is the mean postsynaptic rate. The last term in (8.14) is the covariance between pre- and postsynaptic rates. We now use the linear Poisson neuron Eq. (8.8) to express this term by the input covariance. This yields

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}(t) = c_0 + c_1^{\mathrm{pre}} \nu^{\mathrm{pre}} + c_1^{\mathrm{post}} \nu^{\mathrm{post}}(t) + \beta\, w_{ij}\, \nu^{\mathrm{pre}} + \overline{W}\, \nu^{\mathrm{post}}(t)\, \nu^{\mathrm{pre}}$$

$$+ \sum_{k=1}^{N} w_{ik}(t) Q_{jk} \tag{8.15}$$

where $\overline{W} = \int_{-\infty}^{\infty} W(s)\, \mathrm{d}s$ and $\beta = \int_{-\infty}^{0} W(s)\, C^{\mathrm{spike}}(-s)\, \mathrm{d}s$, as before. The matrix $Q$ with elements

$$Q_{jk} = \int_{-\infty}^{\infty} \mathrm{d}s\, W(s) \int_{0}^{\infty} \mathrm{d}s'\, \epsilon(s')\, C_{jk}^0(s + s') \,. \tag{8.16}$$

contains the spatio-temporal input covariance function

$$C_{jk}^0(s) = \langle (\nu_j(t) - \nu^{\mathrm{pre}})\, (\nu_k(t-s) - \nu^{\mathrm{pre}}) \rangle \tag{8.17}$$

which is convolved with the learning window $W$ and the postsynaptic potential $\epsilon$. Thus, the correlations between pre- and postsynaptic neuron in Eq. (8.9) have been transformed into spatio-temporal covariance in the input.

To summarize this section, we have solved the dynamics of spike-time dependent plasticity under the assumption that learning is slow compared to the variations in the input. For the piecewise linear Poisson neuron, i.e., a stochastically spiking neuron model, the spike-spike correlations between pre- and postsynaptic firing can be evaluated to give $C^{\mathrm{spike}}(s) = \rho_0 \epsilon(s)$ in the linear regime. The final result is a learning equation where weight changes are driven by the expected input rates as well as the spatial and temporal correlations of the input.

### 8.3. Homeostasis and correlation learning

We now extend our mean-field analysis of STDP and show that the interaction between stabilizing homeostatic processes and learning of input correlations gives rise to an interesting learning dynamics with intrinsic subtractive weight normalization. The analysis is facilitated if we add on the right side of Eq. (8.15) a couple of terms that sum to zero, so as to arrive at the following form:

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}(t) = \left\{ c_0 + c_1^{\mathrm{pre}} \nu^{\mathrm{pre}} + c_1^{\mathrm{post}} \nu^{\mathrm{post}}(t) \right.$$

$$\left. + \overline{W} \nu^{\mathrm{pre}} \nu^{\mathrm{post}}(t) + \beta\, (\rho_0 N)^{-1} \left[ \nu^{\mathrm{post}}(t) - \lambda_0 \right] + \overline{Q} \sum_k w_{ik} \right\}$$

$$+ \beta \left[ w_{ij}(t)\, \nu^{\mathrm{pre}} - (\rho_0 N)^{-1} (\nu^{\mathrm{post}}(t) - \lambda_0) \right] + \sum_k w_{ik} \left[ Q_{kj} - \overline{Q} \right] \,. \tag{8.18}$$

$\overline{Q} = N^{-1} \sum_k Q_{jk}$ with $Q_{jk}$ given by Eq. (8.16) is independent of $j$ for the stimulus ensemble under consideration. For a discussion of Eq. (8.18) let us first consider the expected postsynaptic rate $\nu^{\mathrm{post}}(t) = \lambda_0 + \rho_0 \sum_j w_{ij}(t) \nu^{\mathrm{pre}}$, under the assumption that the neuron operates in the linear regime. Even for constant input rates, the postsynaptic rate changes on the slow time scale of learning due to synaptic plasticity. The rate of change can be found by taking the sum over $j$ on the right side of Eq. (8.18) and multiplying by $\nu^{\mathrm{pre}}$. The terms outside the curly braces in Eq. (8.18) cancel each other after summation, and the terms inside the braces yield a linear equation

$$\frac{\mathrm{d}\nu^{\mathrm{post}}(t)}{\mathrm{d}t} = \gamma [\nu^{\mathrm{post}} - \nu_{\mathrm{FP}}] \tag{8.19}$$

with a fixed point

$$\nu_{\mathrm{FP}} = -\frac{c_0 + c_1^{\mathrm{pre}} \nu^{\mathrm{pre}} - \lambda_0\, \beta\, (\rho_0 N)^{-1} - \lambda_0\, \overline{Q}\, (\rho_0 \nu^{\mathrm{pre}})^{-1}}{c_1^{\mathrm{post}} + \nu^{\mathrm{pre}} \overline{W} + \beta\, (\rho_0 N)^{-1} + \overline{Q}\, (\rho_0 \nu^{\mathrm{pre}})^{-1}} \,. \tag{8.20}$$
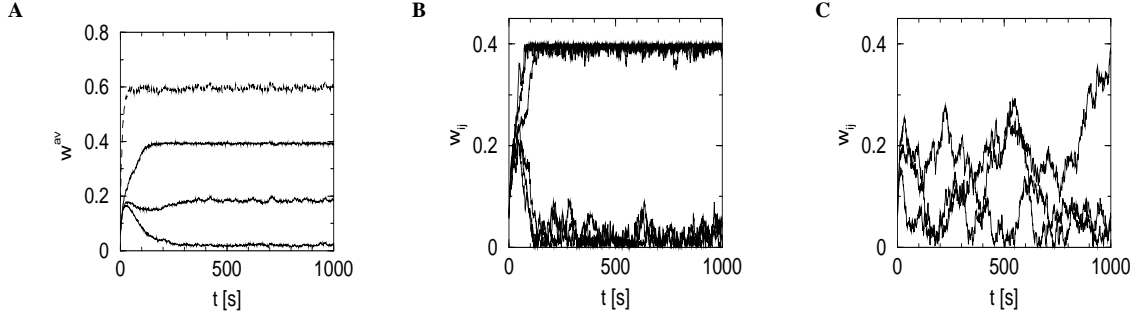
Fig. 9. . A postsynaptic neuron is driven by 300 presynaptic inputs. A first group of 100 inputs are generated by a common inhomogeneous Poisson process with a rate that switches every 20ms between zero and 40 Hz (group 1 - strong correlations). The second group of 100 inputs has a rate that switches simultaneously between 15 and 25 Hz (group 2 - weak correlations). The third group has a fixed rate of 20 Hz (group 3 - no correlations). **A**. The *average weights* $w^{av}$ of synapses in group 3 (bottom trace), in group 2 (second from below), and group 1. The sum of the average weights (top trace) stabilizes rapidly after 50 seconds. During the next 100 seconds synapses in group 1 grow at the expense of those in group 3. **B**. Weights of individual synapses $w_{ij}$ in group 1 (top 3 traces) and group 3 (bottom 3 traces). Weights are bounded between 0 and a maximum value of $w^{\max}$=0.4. **C**. Weights of three synapses in group 2.

The fixed point of the mean postsynaptic rate is found to be stable if

$$\gamma = c_1^{\text{post}} + \nu^{\text{pre}}\,\overline{W} + (\rho_0 N)^{-1}\,\beta + \overline{Q}\,(\rho_0\nu^{\text{pre}})^{-1} < 0\,. \tag{8.21}$$

To proceed with the analysis let us suppose that $\gamma \ll 0$. In our theoretical framework, this can always be achieved if either $c_1^{\text{post}}$ or $\overline{W}$ is sufficiently negative. In this case, the effective time constant $\tau_{\text{eff}} = -1/\gamma$ is short so that the fixed point is attained rapidly provided that it is in the accessible range $\nu_{\text{FP}} \geq 0$. After convergence to the fixed point, the summed weights $\sum_j w_{ij}$ remain constant and the terms in the curly braces in Eq. (8.18) cancel each other. The remaining terms on the right side of Eq. (8.18) determine the evolution of the weight vector. If we switch to matrix notation we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{w} = \left[Q - \overline{Q}\right]\vec{w} + \beta\,\nu^{\text{pre}}\left[\vec{w} - |\vec{w}|_1\,\vec{n}\right] \tag{8.22}$$

where $\vec{n} = (1,1,\ldots,1)^T$, $|\vec{w}|_1 = \sum_k w_{ik}/N$, and $\overline{Q}$ denotes the matrix where all elements have the same value $\overline{Q}$. The result is similar to subtractive weight normalization (cf. Eq. 8.7) with a few specific differences; the spatial correlation matrix $C$ that appears in rate-formulations of Hebbian learning, has been replaced by the matrix $Q$ that describes the spatio-temporal *covariance* on the time scale of the learning window and postsynaptic potential. Thus, the learning rule behaves similar to Sejnowski's covariance rule (cf. Eq. 5.10). The learning dynamics is illustrated in Fig. 9.

The additional spike-spike correlations that appear in spiking neuron models give rise to the second term on the right side of Eq. (8.22), which tends to stabilize synapses that have large weights and decrease synapses with small weights. If we rewrite the vector equation (8.22) in terms of components, the last term gives $\propto (w_{ij} - |\vec{w}|_1)\,\nu^{\text{pre}}$ which shows that the term has exactly the structure of the STDP term in Eq. (7.9) that we have discussed earlier. Hence, if we neglect correlations in the input all arguments developed earlier apply. In particular, synapses that have a value larger than $|\vec{w}|_1$ will be strengthened and others will be decreased. Synapses that are close to the bounds $w^{\max}$ or zero will therefore remain stable over long times, as desired.

Let us now focus on the first term and neglect the spike-spike correlations contained in the second term. The dynamics of the weight vector is then dominated by the eigenvector of the matrix $Q - \overline{Q}$ with the largest eigenvalue. Eq. (8.22) can be seen as the generalization of the simple Hebbian learning rule (8.4) to the case of spike-based learning. *Thus STDP rules with appropriate choice of parameters will automatically implement subtractive weight normalization.*

We may wonder whether the above results are restricted to the specific scenario introduced above or whether they also hold in a generic setting. Generalizations could occur in several directions. First, we recall that the coefficients $c_0$, $c_1^{\text{pre}}$, $c_1^{\text{post}}$, and the learning window $W$ depend, in general, on the current weight value $w_{ij}$. In the above derivation we have assumed that these values are constant. It is, however, possible to set upper and lower bounds for the synaptic

efficacies $w_{ij}$, i.e., weight changes are zero, if $w_{ij} > w^{\max}$ or $w_{ij} < 0$. It is straightforward to extend the above arguments to this case. More realistically, we could also assume an explicit weight dependence where all positive terms have a saturating factor $(w_{ij} - w^{\max})$ and all negative terms a decay factor $w_{ij}$. With these dependencies, it is again possible to analyze the weight dynamics.

Second, we had required that the mean rate $\nu^{\mathrm{pre}}$ be the same for all presynaptic neurons and that spatial correlations have the properties $\sum_k C_{jk} = \overline{C}$ independent of $j$. Both restrictions can be dropped without changing the qualitative structure of the results. The mean rate of the postsynaptic neuron has a fixed point that must be found from the two equations

$$\nu^{\mathrm{FP}} = -\frac{c_0 + c_1^{\mathrm{pre}}\psi + \beta\phi\sum_j w_{ij}(\nu_j^{\mathrm{pre}})^2 + \phi\sum_j w_{ij}(\sum_k Q_{jk}\nu_k^{\mathrm{pre}})}{c_1^{\mathrm{post}} + \psi\,\overline{W}} \tag{8.23}$$

and

$$\nu^{\mathrm{FP}} = \langle g(\sum_j w_{ij}\nu_j(t))\rangle \tag{8.24}$$

where $g$ is the gain function of the inhomogeneous Poisson process, $\nu_j^{\mathrm{pre}} = \langle\nu_j(t)\rangle$ the mean firing rate of presynaptic neuron $j$, $\phi = [\sum_k \nu_k^{\mathrm{pre}}]^{-1}$, and $\psi = [\sum_j(\nu_j^{\mathrm{pre}})^2]/\sum_j \nu_j^{\mathrm{pre}}$. One (or several) solutions always exist since (8.23) is a linear function of the weights and (8.24) as a function of the weights is bounded between zero and a maximum rate $\nu^{\max}$. If one of the fixed points has been attained, the remaining weight equation reads

$$\frac{1}{g'}\frac{\mathrm{d}}{\mathrm{d}t}w_{ij} = \sum_k w_{ik}(Q_{kj} - \overline{q}_k) + \beta\,(w_{ij}\,\nu_i^{\mathrm{pre}} - \theta) \tag{8.25}$$

where $\overline{q}_k = (\sum_n Q_{kn}\,\nu_n^{\mathrm{pre}})\phi$ and $\theta = [\sum_n w_{in}(\nu_n^{\mathrm{pre}})^2]\phi$. Equation (8.25) is the generalization of (8.22).

## 9. Synaptic Equalization from anti-STDP

Before we close the chapter, let us turn to anti-STDP. One of the main differences is that $\beta$ becomes negative,

$$\beta = \int_{-\infty}^{0} W(s)\,C^{\mathrm{spike}}(-s)\,\mathrm{d}s < 0\,. \tag{9.1}$$

For the sake of simplicity, we also take $\alpha = 0$, which corresponds to an antisymmetric window function. This second condition is not necessary, but it simplifies the analysis considerably. Under these conditions, Eq. (7.9) reduces to

$$\frac{dw_{ij}}{dt} = -|\beta|w_{ij}^{\mathrm{eff}}\nu_j\,, \tag{9.2}$$

which results simply in a reduction of all synaptic strengths to zero. To avoid this uninteresting consequence, we add to the anti-STDP an addition form of plasticity seen, for example, in the electric fish experiments (Roberts and Bell, 2000) where anti-STDP is accompanied by a non-associative form of LTP. The non-associative LTP increases synaptic strength at a rate proportional to the presynaptic firing rate (and independent of the postsynaptic firing rate), so Eq. (9.2) becomes

$$\frac{dw_{ij}}{dt} = -|\beta|w_{ij}^{\mathrm{eff}}\nu_j + c_1^{\mathrm{pre}}\nu_j\,. \tag{9.3}$$

This has an (obviously) stable equilibrium point at

$$w_{ij}^{\mathrm{eff}} = \frac{c_1^{\mathrm{pre}}}{\beta}\,. \tag{9.4}$$

The interesting thing about the result (9.4) is that it implies a regulation not of synaptic strength, but of synaptic efficacy. In a number of different neuron types, it has been observed that synapses increase in strength the further they are from the soma on the dendrites. This increase appears to compensate for the attenuation that the potentials

generated by these synapses experience, so that it appears as if synaptic efficacy may be staying constant. Eq. (9.4) provides a way of explaining this result from a simple process that occurs locally and independently at each synapse but that nevertheless compensates for the global structure of the neuron.

## 10. Conclusion - Presynaptic versus Postsynaptic STDP

These notes began with a discussion of short-term synaptic depression through vesicle depletion, and proceeded to a discussion of different forms of STDP without drawing a connection between these two types of synaptic plasticity. That is the subject of the present section. There is experimental evidence that STDP may be primarily, or at least partially, expressed presynaptically. Specifically, STDP modifies the probability of vesicle release $p$. It is undoubtedly a simplification to say that this is all that STDP does, especially in the case of LTP, but it is worth entertaining the implications of a solely presynaptic form of STDP.

With presynaptic STDP, synapses do not get weaker and stronger in the usual sense because, as noted at the beginning of this chapter, the average postsynaptic conductance induced by a high-frequency train of presynaptic action potentials does not depend on $p$. Rather, synapses become more or less willing to transmit the first spikes in a train when they are "strengthened" or "weakened" by STDP. But what about postsynaptic modifications? One speculation is that these are subject to the effects of anti-STDP. In such a scheme, all synapses have the same efficacy in that all presynaptic spikes have the same probability of evoking a postsynaptic response. Selectivity arises in this case, because some synapses transmit a higher percentage of the spikes that reach them, not because these spikes induce a larger effect in the postsynaptic neuron.

## References

[Abbott, 2000] Abbott, L. (2000) and Nelson, S.B.. Synaptic plasticity - taming the beast. *Nature Neuroscience*, 3:1178–1183.

[Abbott and Blum, 1996] Abbott, L. F. and Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, 6:406–416.

[Abbott et al., 1997] Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275:220–224.

[Bi and Poo, 2001] Bi, G. and Poo, M. (2001). Synaptic modification of correlated activity: Hebb's postulate revisited. *Ann. Rev. Neurosci.*, 24:139–166.

[Bienenstock et al., 1982] Bienenstock, E., Cooper, L., and Munroe, P. (1982). Theory of the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2:32–48. reprinted in Anderson and Rosenfeld, 1990.

[Blum and Abbott, 1996] Blum, K. and Abbott, L. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural Comput.*, 8:85–93.

[Chateau and Fukai, 2003] Chateau, H. and Fukai, T. (2003). A stochastic model to predict the consequences of arbitrary forms of spike-timing dependent plasticity. *Neural Computation*, 15:597–620.

[Froemke and Dan, 2002] Froemke, R. and Dan, Y. (2002). Spike-timing dependent plasticity induced by natural spike trains. *Nature*, 416:433–438.

[Fusi et al., 2000] Fusi, S., Annunziato, M., Badoni, D., Salamon, A., and D.J.Amit (2000). Spike-driven synaptic plasticity: theory, simulation, vlsi implementation. *Neural Computation*, 12:2227–2258.

[Gerstner and Abbott, 1996] Gerstner, W. and Abbott, L.F. (1997) Learning Navigational Maps Through Potentiation and Modulation of Hippocampal Place Cells. *J. Computational Neurosci.*, 4:79–94.

[Gerstner et al., 1996] Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383:76–78.

[Gerstner and Kistler, 2002] Gerstner, W. and Kistler, W. K. (2002). *Spiking Neuron Models*. Cambridge University Press.

[Goldberg et al., 2002] Goldberg, J., Holthoff, K., and Yuste, R. (2002). A problem with hebb and local spikes. *Trends in Neurosciences*, 25:433–435.

[Grossberg, 1987] Grossberg, S. (1987). *The Adaptive Brain I*. Elsevier.

[Gütig et al., 2003] Gütig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through non-linear temporally asymmetry hebbian plasticity. *J. Neuroscience*, 23:3697–3714.

[Hebb, 1949] Hebb, D. O. (1949). *The Organization of Behavior*. Wiley, New York.

[James, 1890] James, W. (1890). *Psychology (Briefer Course), ch. 16*. Holt, New York.

[Kempter et al., 1999] Kempter, R., Gerstner, W., and van Hemmen, J. L. (1999). Hebbian learning and spiking neurons. *Phys. Rev. E*, 59:4498–4514.

[Kempter et al., 2001] Kempter, R., Gerstner, W., and van Hemmen, J. L. (2001). Intrinsic stabilization of output rates by spike-based hebbian learning. *Neural Computation*, 13:2709–2741.

[Kistler, 2002] Kistler, W. M. (2002). Spike-timing dependent plasiticity: a phenomenological framework. *Biological Cybernetics*, xx:xx.

[Kistler and van Hemmen, 2000] Kistler, W. M. and van Hemmen, J. L. (2000). Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic potentials. *Neural Comput.*, 12:385–405.

[Kohonen, 1984] Kohonen, T. (1984). *Self-Organization and Associative Memory*. Springer-Verlag, Berlin Heidelberg New York.

[London and Segev, 2001] London, M. and Segev, I. (2001). Synaptic scaling in vitro and in vivo. *Nature Neuroscience*, 4:853–854.

[Magee and Cook, 2000] Magee, J. C. and Cook, E. P. (2000). Somatic epsp amplitude is independent of synapse location in hippocampal pyramidal neurons. *Nature Neuroscience*, 3:895–903.

[Mehta et al., 2000] Mehta, M. R., Quirk, M., and Wilson, M. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron*, 25:707–715.

[Miller and MacKay, 1994] Miller, K. D. and MacKay, D. J. C. (1994). The role of constraints in hebbian learning. *Neural Computation*, 6:100–126.

[Minai and Levy, 1993] Minai, A. and Levy, W. B. (1993). Sequence learning in a single trial. In *Proc. of the INNS World Congress on Neural Networks, Portland, Oregon, Vol. II*, pages 505–508. International Neural Network Society,.

[Oja, 1982] Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Mathematical Biology*, 15:267–273.

[Rao and Sejnowski, 2001] Rao, R. P. N. and Sejnowski, T. J. (2001). Spike-timing dependent Hebbian plasticity as temporal difference learning. *Neural Computation*, 13:2221–2237.

[Roberts, 1999] Roberts, P. (1999). Computational consequences of temporally asymmetric learning rules: I. Differential Hebbian learning. *J. Computational Neuroscience*, 7:235–246.

[Roberts and Bell, 2000] Roberts, P. and Bell, C. (2000). Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. *Computational Neuroscience*, 9:67–83.

[Rubin et al., 2001] Rubin, J., Lee, D. D., and Sompolinsky, H. (2001). Equilibrium properties of temporally asymmetric Hebbian plasticity. *Physical Review Letters*, 86:364–367.

[Rumsey and Abbott, 2003] Rumsey, C. and Abbott, L. (2003). Equalization of synaptic efficacy by activity- and timing-dependent synaptic plasticity. *J. Neurophysiology*, xx:xx. in press.

[Sejnowski, 1977] Sejnowski, T. (1977). Storing covariance with nonlinearly interacting neurons. *J. Mathematical Biology*, 4:303–321.

[Sejnowski and Tesauro, 1989] Sejnowski, T. J. and Tesauro, G. (1989). The Hebb rule for synaptic plasticity: algorithms and implementations. In Byrne, J. H. and Berry, W. O., editors, *Neural Models of Plasticity*, chapter 6, pages 94–103. Academic Press.

[Senn et al., 2001] Senn, W., Tsodyks, M., and Markram, H. (2001). An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. *Neural Computation*, 13:35–67.

[Sjöström et al., 2003] Sjöström, P., Turrigiano, G., and Nelson, S. (2003). Neocortical ltd via coincident activation of presynaptic nmda and cannabinoid receptors. *Neuron*, 39:641–654.

[Song and Abbott, 2001] Song, S. and Abbott, L. (2001). Column and map development and cortical re-mapping through spike-timing dependent plasticity. *Neuron*, 32(339-350).

[Song et al., 2000] Song, S., Miller, K., and Abbott, L. (2000). Competitive Hebbian learning through spike-time-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926.

[Tsodyks and Markram, 1997] Tsodyks, M. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Academy of Sci., USA*, 94:719–723.

[van Rossum et al., 2000] van Rossum, M. C. W., Bi, G. Q., and Turrigiano, G. G. (2000). Stable Hebbian learning from spike timing-dependent plasticity. *J. Neuroscience*, 20:8812–8821.