

Hebbian Learning and Plasticity

Wulfram Gerstner

*School of Computer and Communication Sciences
and Brain-Mind Institute
Ecole Polytechnique Fédérale de Lausanne
1015 Lausanne EPFL
wulfram.gerstner@epfl.ch*

..

To appear in:
From Neuron to Cognition via Computational Neuroscience,
edited by Michael Arbib and Jimmy Bonaiuto,
MIT Press Cambridge
Chapter 9

9. Hebbian Learning and Plasticity

The elementary processing units in the brain are neurons (see Chapter 2) which are connected to each other via cable-like extensions, called axons and dendrites (see Chapter 3). The contact points between an axon terminal of a neuron A and the dendrite of another neuron B are called synapses. The most important synapses are chemical synapses, but there exist also electrical synapses which we will not consider further. When an action potential of neuron A arrives at the axon terminating at a chemical synapse, a chemical signal (neurotransmitter) is ejected into the synaptic cleft and taken up by receptors sitting on the membrane of the dendrite of the receiving neuron. Upon transmitter binding, an ion channel opens, ions flow into the cell and cause a change in the properties of neuron B. This change can be measured as an Excitatory or Inhibitory Postsynaptic Potential (EPSP or IPSP) at the soma of the receiving neuron (see Chapter 3). The strength of an excitatory synaptic connection can be quantified by the amplitude of the EPSP. The most important insight for all the rest of chapter is that the strength of a synapse is not fixed, but can change.

9.1 Introduction to Hebbian Plasticity

Changes of synapses in the network of neurons in the brain are called synaptic plasticity. These changes are thought to be the basis of learning.

9.1.1 What is Learning?

A small child learns to unscrew the lid of a bottle. An older child learns to ride a bicycle, to ski, or to skateboard. Learning activities such as in these examples typically uses an internal reward system of the brain: it hurts if we fall from the bicycle and we are happy if we achieve for the first time a slalom slope with the skies.

Learning is also the basis of factual and episodic memories: we know the name of the current president of the United States because we have heard it often enough; we know the date of the French Revolution because we have learned it in school etc. These are examples of factual memories. We can remember the first day we went to school; we can recall a beautiful scene encountered during our last vacation. These are examples of episodic memories, that have been generated most often without explicit learning, but are still acquired (as opposed to inborn) and have therefore been 'learned' in the loose sense of the word.

Finally, it is the result of learning if a musician is able to distinguish between tones that sound absolutely identical to the ear of a normal untrained human. In experiments with monkeys measurable differences were found between the auditory areas of animals exposed to specific tones and others living in a normal environment (Recanzone et al., 1993). More generally, it is believed that the cortex adapts itself such that more neurons are devoted to stimuli that appear more frequently or are more important and less neurons to less relevant ones (Buonomano and Merzenich, 1998). This adaptation of cortex (see Chapter 14, Neural Maps) is also subsumed under the term of 'learning' in the wider sense.

Most likely, all the different forms of learning that we have mentioned (action

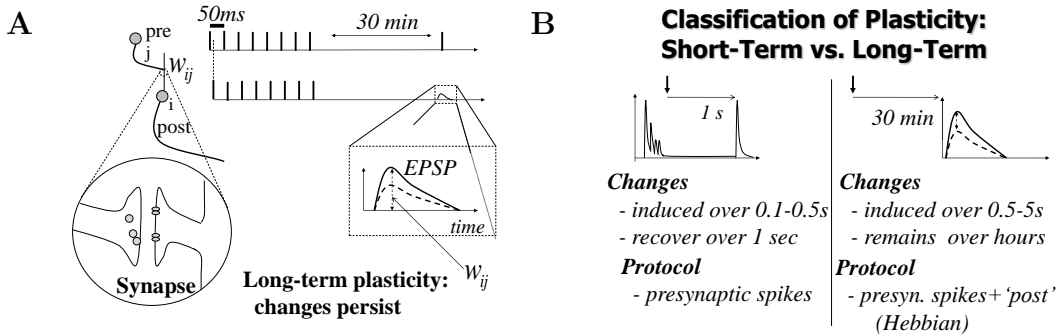


Fig. 9.1: **A.** Two neurons, a presynaptic neuron j and a postsynaptic neuron i are connected by a synapse with weight w_{ij} . The weight is determined by the amplitude of the excitatory postsynaptic potential (EPSP) that is measured as the response of the postsynaptic neuron to an isolated input spike (inset lower right). The synapse itself functions via a chemical signaling chain (inset lower left). The weight of the synapse is changed if the two neurons are activated in the sequence pre- before postsynaptic neuron. **B.** Short-Term plasticity recovers rapidly whereas Long-Term plasticity persists for a long time.

learning, formation of factual or episodic memories, adaptation of cortical organization) are in one way or other related to changes in the strength of synapses.

9.1.2 Classification of Synaptic Plasticity

Changes in synaptic strength can be induced in a controlled way in preparations of neuronal brain slices. First, the strength of the connection is measured by generating a single test spike in the presynaptic (=signaling) neuron while recording the postsynaptic potential (or postsynaptic current) in the postsynaptic (=receiving) neuron. Then an appropriate stimulus is given to induce a change of the synapse. Finally, a second test spike is evoked in the presynaptic neuron and the change in the amplitude of the postsynaptic potential is noted; see Fig. 9.1.

An appropriate stimulus, for example, a sequence of 10 spikes at 100 Hz in the presynaptic neuron, can increase or decrease the amplitude of the measured EPSP by a factor of two or more. However, if another test pulse is given 5 seconds later, this change has disappeared (Markram et al., 1998; Abbott et al., 1997). Since this type of plasticity lasts only for one or a few seconds, it is called short-term plasticity (Fig. 9.1B).

We now consider a different stimulus. The presynaptic neuron is stimulated so as to produce 60 spikes at 20Hz. In parallel the postsynaptic neuron is also stimulated to produce 60 spikes at 20Hz, but the two stimulation protocols are slightly shifted so that the postsynaptic neuron fires always 10ms after the presynaptic one (Fig. 9.1A). Note that the stimulus only lasts three seconds in total. Nevertheless, it introduces an increase in the EPSP that persists for minutes or hours. Hence it is an example of persistent plasticity, also called Long-Term Potentiation (LTP). If the relative timing is reversed so that the presynaptic neuron fires always *after* the postsynaptic one, the protocol with 60 spikes induces Long-Term Depression (LTD).

The specific protocol discussed here is called Spike Timing Dependent Plasticity (STDP) (Markram et al., 1997; Abbott and Nelson, 2000), but it is only one example of a much broader class of stimulation protocols that are all suitable to induce LTP and LTD. Instead of driving the postsynaptic neuron to firing, it

can also be held at weakly or strongly depolarized level. If the depolarization is combined with presynaptic spike arrival, this causes LTD or LTP (Artola et al., 1990; Ngezahayo et al., 2000). Instead of controlling both presynaptic and postsynaptic neurons precisely, one can also stimulate many presynaptic fibers by a high-frequency sequence of extracellular current pulses while recording (intracellularly or extracellularly) from postsynaptic neurons. If enough presynaptic fibers are stimulated, the postsynaptic neuron is depolarized or even firing, and LTP can be induced. Such an extracellular stimulation protocol is particularly useful to confirm that changes induced by LTP or LTD indeed last for many hours (Frey and Morris, 1998).

The first and most important classification is that between short-term plasticity and long-term plasticity; see Fig. 9.1B. In the following we will focus on persistent plasticity (LTP and LTD) and neglect short-term plasticity. But within the realm of long-term plasticity further classifications are possible. We mentioned already a distinction between spike-timing based protocols on one side and traditional protocols on the other side.

Another important distinction is the presence or absence of neuromodulators during the plasticity inducing protocol. We will come back to this distinction in Section 4 of this chapter.

9.1.3 Long-Term Potentiation as Hebbian Learning

LTP and LTD are thought to be the synaptic basis of learning in the brain. Many experiments on LTP and LTD, and nearly all synaptic theories of learning, have been inspired by a formulation of Hebb (Hebb, 1949) which has roots that can in fact be traced back much further in the past (Makram et al., 2011). It states that it would be useful to have a rule that synapses are modified whenever the two neurons that are connected to each other are active together. It is sometimes summarized in the slogan 'fire together - wire together', but the exact wording is worth a read:

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.

In contrast to the compressed slogan it contains a 'causal' tone, because 'taking part in firing' the postsynaptic neuron implies that the presynaptic neuron is, at least partly, causing the spike of the postsynaptic one. We will come back to this aspect when we speak more about STDP.

At the moment, two aspects are important. First, for the connection from A to B only two neurons are important, namely A and B, but not any other neuron C that might make a connection onto A or B. We summarize this insight by saying that the learning rule is 'local': only information that is available at the location of the synapse can be used to change the weight of that synapse. Second, the wording 'cell A ... takes part in firing' cell B implies that both cells have to be active. We summarize this insight by saying that the learning rule must be sensitive to the correlations between the action potentials of the two neurons.

With these two aspects in mind, let us now return to the plasticity protocols for LTP that we discussed in the previous subsection. Indeed, the STDP protocol

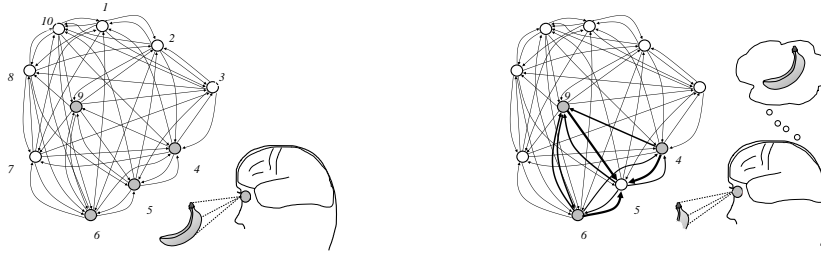


Fig. 9.2: A novel percept causes simultaneous activity of a subset of neurons, e.g. those numbered as 4,5,6, and 9 (left). Under the assumption of Hebbian learning the connections between these neurons will be strengthened. A partial cue later (right) is sufficient to recall the complete memory item, since neurons that are not activated by the sensory cue will be excited by the strong lateral connections.

makes both neurons, the presynaptic neuron A and the postsynaptic neuron B, fire together (and therefore induces correlations), while other neurons C do not play a role (and therefore we have locality). Hence, it is a Hebbian protocol. The traditional extracellular stimulation protocol excites many presynaptic fibers to generate spikes at high frequencies and these presynaptic spikes are likely to fire the postsynaptic neuron, too. Therefore, again a Hebbian protocol. One may argue that postsynaptic firing can be replaced by strong postsynaptic depolarization, so that the other protocols mentioned above fall also into the category of Hebbian protocols.

Why should Hebbian Learning be Useful?

Suppose that, somewhere in the brain, we have a network of neurons that can mutually excite each other. The network could consist of 10 000 neurons or more, but only a few of these are depicted in Fig. 9.2. We assume now that, while a human or animal sees for the first time a banana, a subset of these neurons are active. The neurons may represent in an abstract form the different sensations associated with the percept, such as form, color or the smell of the banana. Maybe somebody standing nearby says in addition the word banana, or breaks it open and starts to eat it.

If the brain has a Hebbian learning rule, the result of the co-activation of different neurons in our network, all the connections between the simultaneously active neurons are strengthened. The claim is that this means that the item 'banana' has now been memorized. The memory concept 'banana' has been formed.

How can we check that the memory concept 'banana' works? Let us suppose that, the next day, the subject sees a banana, partly hidden behind another object. The banana as a whole is not visible, but there are some yellow cues, organized into a slightly bent structure. Let us now *suppose* that seeing part of the banana stimulates again a subset of the neurons that have been active the day before when the banana was seen for the first time. Because of the previously established strong connections, the neurons that are part of the 'banana concept' but are currently inactive will now be activated by neighboring neurons that send spikes across the strong connections. As a result, after a few iterations, the memory concept 'banana' is fully activated and the memory item is retrieved.

The basic idea of memory retrieval discussed here is at the heart of models of working memory and long-term memory (see Chapter 13).

A Family of Hebbian Rules

Hebb was a theoretically inclined psychologist, who developed the essential insights of why the principle, which we now call the Hebb-rule, would be useful; but he was not a mathematician and never wrote down a mathematical formulation of his rule. Finding an appropriate mathematical description is our task now. In this subsection we follow closely the treatment in Chapter 10.2 of Gerstner and Kistler (2002).

In order to find a mathematically formulated learning rule based on Hebb's postulate we focus on a single synapse with efficacy w_{ij} that transmits signals from a presynaptic neuron j to a postsynaptic neuron i . For the moment we focus on a description in terms of mean firing rates. In the following, the activity of the presynaptic neuron is denoted by ν_j and that of the postsynaptic neuron by ν_i .

As mentioned before, there are two aspects in Hebb's postulate that are particularly important, viz. *locality* and *cooperativity*. Locality means that the change of the synaptic efficacy can only depend on local variables, i.e., on information that is available at the site of the synapse, such as pre- and postsynaptic firing rate, and the actual value of the synaptic efficacy, but not on the activity of other neurons. Based on the locality of Hebbian plasticity we can make a rather general ansatz for the change of the synaptic efficacy,

$$\frac{d}{dt}w_{ij} = F(w_{ij}; \nu_i, \nu_j). \quad (9.1)$$

Here, dw_{ij}/dt is the rate of change of the synaptic coupling strength and F is a so far undetermined function.

We may wonder whether there are other local variables (e.g., the membrane potential u_i) that should be included as additional arguments of the function F . It turns out that in standard rate models this is not necessary, since the membrane potential u_i is uniquely determined by the postsynaptic firing rate, $\nu_i = g(u_i)$, with a monotone gain function g .

The second important aspect of Hebb's postulate, cooperativity, implies that pre- and postsynaptic neuron have to be active *simultaneously* for a synaptic weight change to occur. We can use this property to learn something about the function F . If F is sufficiently well-behaved, we can expand F in a Taylor series about $\nu_i = \nu_j = 0$,

$$\begin{aligned} \frac{d}{dt}w_{ij} &= c_0(w_{ij}) + c_1^{\text{post}}(w_{ij})\nu_i + c_1^{\text{pre}}(w_{ij})\nu_j \\ &+ c_2^{\text{pre}}(w_{ij})\nu_j^2 + c_2^{\text{post}}(w_{ij})\nu_i^2 + c_2^{\text{corr}}(w_{ij})\nu_i\nu_j + \mathcal{O}(\nu^3). \end{aligned} \quad (9.2)$$

The term containing c_2^{corr} on the right-hand side of (9.2) is bilinear in pre- and postsynaptic activity. This term implements the AND condition for cooperativity which makes Hebbian learning a useful concept.

The simplest choice for our function F is to fix c_2^{corr} at a positive constant and to set all other terms in the Taylor expansion to zero. The result is the prototype of Hebbian learning,

$$\frac{d}{dt}w_{ij} = c_2^{\text{corr}} \nu_i \nu_j. \quad (9.3)$$

The dependence of F on the synaptic efficacy w_{ij} is a natural consequence of the fact that w_{ij} has to be bounded. If F were independent of w_{ij} , then the synaptic efficacy would grow without limit if the same potentiating stimulus is applied over and over again. Explosion of weights can be avoided by 'hard bounds': $c_2^{\text{corr}}(w_{ij}) = \gamma_2 > 0$ for $0 < w_{ij} < 1$ and zero otherwise. A more subtle saturation of synaptic weights can be achieved, if the parameter c_2^{corr} in Eq. (9.2) tends to zero as w_{ij} approaches its maximum value, say $w^{\text{max}} = 1$, e.g.,

$$c_2^{\text{corr}}(w_{ij}) = \gamma_2 (1 - w_{ij}) \quad (9.4)$$

with a positive constant γ_2 . An interpolation between the soft bounds and hard bounds can be implemented by writing $c_2^{\text{corr}}(w_{ij}) = \gamma_2 (1 - w_{ij})^\eta$ with a parameter $0 \leq \eta \leq 1$. For $\eta = 0$ we retrieve the hard bounds while for $\eta = 1$ we are back to the linear soft bounds.

Obviously, setting all parameters except c_2^{corr} to zero is a very special case of the general framework developed in Eq. (9.2). Are there other 'Hebbian' learning rules in this framework?

First we note that a learning rule with $c_2^{\text{corr}} = 0$ and only first-order terms (such as $c_1^{\text{post}} \neq 0$ or $c_1^{\text{pre}} \neq 0$) would be called non-Hebbian plasticity, because pre- or postsynaptic activity alone induces a change of the synaptic efficacy. Hence these learning rules miss the correlation aspect of Hebb's principle. Thus a learning rule in the family of Eq. (9.2) needs a term $c_2^{\text{corr}} > 0$ so as to qualify as Hebbian. But more complicated learning rules can be constructed if in addition to the linear terms, and the terms nu_i^2 or nu_j^2 other terms in the expansion of Eq. (9.2), such as $\nu_i \nu_j^2$, $\nu_i^2 \nu_j$, $\nu_i^2 \nu_j^2$, etc., are included as well. A learning rule with a positive coefficient in front of $\nu_i^2 \nu_j$ would also qualify as Hebbian, even if c_2^{corr} vanishes or is negative, because at high postsynaptic firing rates the positive correlations dominate the dynamics, as we will see further below in the context of the Bienenstock-Cooper-Munro rule.

Hebb's original proposal does not contain a rule for a decrease of synaptic weights. In a system where synapses can only be strengthened, all efficacies will finally saturate at their upper maximum value. An option of decreasing the weights (synaptic depression) is therefore a necessary requirement for any useful learning rule. This can, for example, be achieved by weight decay, which can be implemented in Eq. (9.2) by setting

$$c_0(w_{ij}) = -\gamma_0 w_{ij}. \quad (9.5)$$

Here, γ_0 is (small) positive constant that describes the rate by which w_{ij} decays back to zero in the absence of stimulation. Our formulation (9.2) is hence sufficiently general to allow for a combination of synaptic potentiation and depression. If we combine (9.4) and (9.5) we obtain the learning rule

$$\frac{d}{dt}w_{ij} = c_2^{\text{corr}}(w_{ij}) \nu_i \nu_j - \gamma_0 w_{ij}. \quad (9.6)$$

The last term leads to an exponential decay to $w_{ij} = 0$ in the absence of stimulation, only one of the two neurons is active.

Another interesting aspect of learning rules is *competition*. The idea is that synaptic weights can only grow at the expense of others so that if a certain subgroup of synapses is strengthened, other synapses to the same postsynaptic neuron have to be weakened. Competition is essential for any form of self-organization and pattern formation. Practically, competition can be implemented in simulations by normalizing by an explicit algebraic step the sum of all weights converging onto the same postsynaptic neuron (Miller and MacKay, 1994). Though this can be motivated by a limitation of common synaptic resources such a learning rule violates locality of synaptic plasticity. Much more elegant, however, is a formulation that remains local, but makes use of the second-order term ν_i^2 in Eq. (9.2). Specifically, we take $c_2^{\text{corr}} = \gamma > 0$ and $c_2^{\text{post}} = -\gamma w_{ij}$ and set all other parameters to zero. The learning rule

$$\frac{d}{dt}w_{ij} = \gamma [\nu_i \nu_j - w_{ij} \nu_i^2] \quad (9.7)$$

is called Oja's rule (Oja, 1982). If combined with a linear model neuron, Oja's rule converges asymptotically to synaptic weights that are normalized to $\sum_j w_{ij}^2 = 1$ while keeping the essential Hebbian properties of the standard rule of Eq. (9.3). We note that normalization of $\sum_j w_{ij}^2$ implies competition between the synapses that make connections to the same postsynaptic neuron, i.e., if some weights grow others must decrease.

9.1.4 The Bienenstock-Cooper-Munro Rule as an Example of Hebbian Learning

Higher terms in the expansion on the right-hand side of Eq. (9.2) lead to more intricate plasticity schemes. As an example, let us consider the Bienenstock-Cooper-Munro rule

$$\frac{d}{dt}w_{ij} = \eta \phi(\nu_i) \nu_j - \gamma w_{ij} \quad (9.8)$$

with a nonlinear function $\phi(\nu_i)$ which we take as $\phi(\nu_i) = \nu_i (\nu_i - \theta)$ and a parameter θ as a reference rate (Bienenstock et al., 1982). A simple calculation shows that Eq. (9.8) can be classified in the framework of Eq. (9.2) with a term $c_2^{\text{corr}} = -\eta \nu_\theta$ and a higher-order term proportional to $\nu_i^2 \nu_j$ that comes with a positive coefficient $\eta > 0$. What does this learning rule do? If the presynaptic neuron is inactive ($\nu_j = 0$), the synaptic weight does not change. Let us now suppose that the presynaptic neuron is active ($\nu_j > 0$) while at the same time the postsynaptic neuron is firing at a high rate $\nu_i > \theta$ (see Fig. 9.3). The synaptic weight increases since both neurons are jointly active. Hence it is a Hebbian rule. If the presynaptic neuron is firing while the postsynaptic neuron is only weakly active ($\nu_i < \theta$), then the synaptic weight decreases. Thus the parameter θ marks the transition point between the induction of LTD and LTP.

In order to get an understanding of how Hebbian plasticity works, the reader is asked to turn now to the first exercise.

Two important insights can be derived from an analysis of Eq. (9.8) [see also Exercise 1 for an intuitive approach].

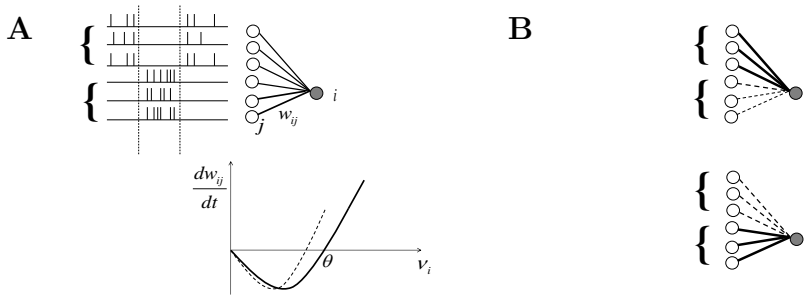


Fig. 9.3: Learning in the Bienenstock-Cooper-Munro model of synaptic plasticity. **A.** A postsynaptic neuron i receives input from several presynaptic neurons, one of them is marked with the index j . The connection w_{ij} from j to i increases or decreases as a function of the postsynaptic firing rate ν_i (bottom), provided that the presynaptic neuron j is active. The activity of presynaptic neurons (top left) switches from time to time (vertical lines) between two groups of correlated inputs. **B.** During learning one of the input groups develops strong synapses, while the synapses of the others are weakened.

First, we can convince ourselves that the postsynaptic rate has a fixed point at θ , but that this fixed point is unstable. In order to avoid that the postsynaptic firing rate blows up or decays to zero, it is therefore necessary to turn θ into an adaptive variable (Bienenstock et al., 1982). More precisely, θ must have a super-linear dependence upon the averaged firing rate $\bar{\nu}_i$. A good choice is to set $\theta = (\bar{\nu}_i)^2 / \nu_{\text{target}}$ where ν_{target} is a parameter that plays the role of a target rate.

Second, if the input is structured into two groups and if the parameter θ is in an appropriate regime, then the learning rule separates the weights into two groups: Some weights increase toward the upper bound at the expense of other weights that decrease to zero (Fig. 9.3B). Thus the learning rule exhibits competition.

Competition is the key ingredient to understand the development of receptive fields: because of synaptic plasticity the postsynaptic neuron becomes specialized to a subset of the inputs. Functional consequences of Hebbian learning, in particular the development of receptive fields, is the topic of the next section.

9.2 Functional Consequences of Hebbian Learning

In the previous section, a family of Hebbian Learning rules has been introduced. We will now study functional consequences of Hebbian plasticity. One particularly intriguing result is that Hebbian learning can lead to the development of receptive fields.

9.2.1 What are Receptive Fields?

A given neuron in cortex does not respond to all stimuli, but only to a small subset of potential stimuli. For example, a neuron in primary visual cortex responds to visual stimuli - but not all of them. Suppose a monkey fixates on a red cross in the center of a gray computer screen while an electrophysiologist records the activity of one single neuron in visual cortex of the monkey (Fig. 9.4A). On the screen localized light blobs are switched on at various locations, in some random

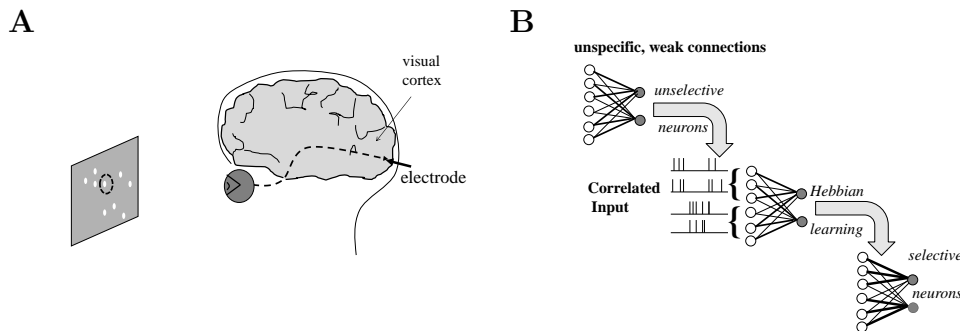


Fig. 9.4: **A** The receptive field of a visual neuron is the zone on the gray screen (left) in which a small light blob evokes a response. The response is measured with an electrode in proximity of the neuron (right). **B** If model neurons are stimulated with spatially correlated input so that neighboring presynaptic neurons change their activity together, the postsynaptic neurons develop localized receptive field as a result of Hebbian synaptic plasticity.

sequence. For most possible locations the neuron remains quiet. However, if the light blob occurs in one specific small region of the screen, the neuron becomes active. The zone in the visual field where the neuron is responsive is called the (visual) receptive field. Receptive fields occur not only in vision, but also in other modalities such as audition or touch.

To understand the concept of receptive fields from the point of view of modeling, let us consider a group of 10 postsynaptic neurons, each being connected to the same set of 100 presynaptic neurons. The presynaptic neurons play the role of potential locations of a stimulus.

Suppose that you test the response properties of one specific postsynaptic neuron by giving an input to a single presynaptic neuron. Before plasticity started all presynaptic neurons were connected with the same weak weight to this postsynaptic neuron. Therefore the postsynaptic neuron responds unspecifically to all possible input locations. After learning, however, the postsynaptic neuron responds only to a small subset of the presynaptic neurons. Hence, the neuron has developed a receptive field.

Exercises 1c has highlighted competition and receptive field development for the case where there are two input groups which differ in their firing rates. Input pattern A consisted of a group of neurons firing at 10 Hz while input pattern B consisted of another group of neurons firing at 30 Hz. However, this looks somewhat artificial. Why should there be such a difference? One would expect that input in nature is balanced, so that averaged over all potential input patterns, all inputs should be more or less equally often active.

In order to construct a first example with balanced inputs, let us again consider a sequence of inputs that alternates between two patterns. In pattern A, a first group of 50 presynaptic neurons fires at 20Hz and a second group of 50 presynaptic neurons fires at 15Hz. In pattern B, the first group of neurons fires again at 20Hz while the second group now fires at 25 Hz. Averaged across both patterns, all neurons have the same mean rate $\bar{\nu}_j = 20Hz$. Nevertheless Hebbian learning rules will still prefer one input group over the other, so that the postsynaptic neuron will specialize and develop a receptive field. The reason is, that Hebbian learning is sensitive to correlations in the input.

9.2.2 Correlations and Principal Component Analysis

We now turn to a formal derivation of the functional properties of Hebbian learning rules. We show mathematically that simple models of Hebbian learning are sensitive to correlations in the input. More precisely, a standard Hebb rule combined with a linear rate model for the postsynaptic neuron performs Principal Component Analysis, also called PCA. This section follows closely the text in Chapter 11.1 of Gerstner and Kistler (2002).

We analyze the evolution of synaptic weights using the simple Hebbian learning rule Eq. (9.3) and then generalize to the Oja's rule (9.7). For the analysis, we consider a highly simplified scenario consisting of an analog neuron that receives input from N presynaptic neurons with firing rates ν_i^{pre} via synapses with weights w_i ; cf. Fig. 9.5. We think of the presynaptic neurons as 'input neurons', which, however, do not have to be sensory neurons. The input layer could, for example, consist of neurons in the lateral geniculate nucleus (LGN) that project to neurons in the visual cortex. We will see that the statistical properties of the input control the evolution of synaptic weights.

For the sake of simplicity, we model the presynaptic input as a set of static patterns. Let us suppose that we have a total of p patterns $\{\vec{\xi}^\mu; 1 \leq \mu \leq p\}$. At each time step one of the patterns $\vec{\xi}^\mu$ is selected at random and presented to the network by fixing the presynaptic rates at $\nu_i^{\text{pre}} = \xi_i^\mu$. We call this the *static-pattern scenario*. The presynaptic activity drives the postsynaptic neuron and the joint activity of pre- and postsynaptic neurons triggers changes of the synaptic weights. The synaptic weights are modified according to a Hebbian learning rule, i.e., according to the correlation of pre- and postsynaptic activity; cf. Eq. (9.3). Before the next input pattern is chosen, the weights are changed by an amount

$$\Delta w_j = \gamma \nu^{\text{post}} \nu_j^{\text{pre}} \quad (9.9)$$

Here, $0 < \gamma \ll 1$ is a small constant called 'learning rate'. Since there is only one postsynaptic neuron, we have suppressed the index i of the postsynaptic cell. The learning rate in the static-pattern scenario is closely linked to the correlation coefficient c_2^{corr} in the continuous-time Hebb rule introduced in Eq. (9.3). In order to highlight the relation, let us assume that each pattern $\vec{\xi}^\mu$ is applied during an interval Δt . For Δt sufficiently small, we have $\gamma = c_2^{\text{corr}} \Delta t$.

In a general rate model, the firing rate ν^{post} of the postsynaptic neuron is given by a nonlinear function of the total input $\nu^{\text{post}} = g(\sum_i w_i \nu_i^{\text{pre}})$ but for the sake of simplicity, we restrict our discussion in the following to a *linear* rate model with

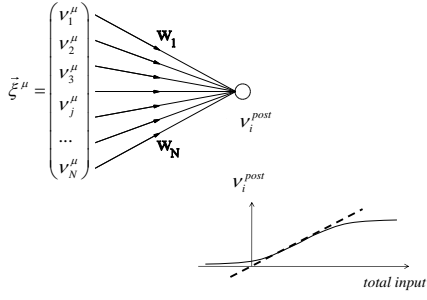
$$\nu^{\text{post}} = \sum_j w_j \nu_j^{\text{pre}}. \quad (9.10)$$

Obviously, this is a highly simplified neuron model, but it will serve our purpose of gaining some insights in the evolution of synaptic weights.

If we combine the learning rule (9.3) with the linear rate model of Eq. (9.10) we find after the presentation of pattern $\vec{\xi}^\mu$ the synaptic weight w_i connecting the presynaptic neuron i to the postsynaptic cell is changed by an amount

$$\Delta w_i = \gamma \sum_j w_j \nu_j^{\text{pre}} \nu_i^{\text{pre}} = \gamma \sum_j w_j \xi_j^\mu \xi_i^\mu. \quad (9.11)$$

A



B

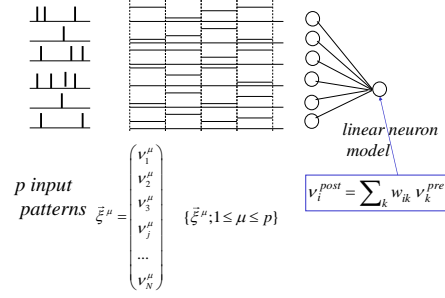


Fig. 9.5: **A.** Patterns $\vec{\xi}^\mu$ are applied as a set of presynaptic firing rates ν_j , i.e., $\xi_j^\mu = \nu_j^\mu$ for $1 \leq j \leq N$. The output rate of the postsynaptic neuron is taken as a linear function of the total input, an approximation to a sigmoidal gain function. Adapted from Gerstner and Kistler (2002). **B** Patterns are applied one after the other in a sequence.

The evolution of the weight vector $\vec{w} = (w_1, \dots, w_N)$ is thus determined by the iteration

$$w_i(n+1) = w_i(n) + \gamma \sum_j w_j \xi_j^{\mu_n} \xi_i^{\mu_n}, \quad (9.12)$$

where μ_n denotes the pattern that is presented during the n th time step.

We are interested in the long-term behavior of the synaptic weights. To this end we assume that the weight vector evolves along a more or less deterministic trajectory with only small stochastic deviations that result from the randomness at which new input patterns are chosen. This is, for example, the case if the learning rate is small so that a large number of patterns has to be presented in order to induce a substantial weight change. In such a situation it is sensible to consider the expectation value of the weight vector, i.e., the weight vector $\langle \vec{w}(n) \rangle$ averaged over the sequence $(\vec{\xi}^{\mu_1}, \vec{\xi}^{\mu_2}, \dots, \vec{\xi}^{\mu_n})$ of all patterns that so far have been presented to the network. From (9.12) we find

$$\begin{aligned} \langle w_i(n+1) \rangle &= \langle w_i(n) \rangle + \gamma \sum_j \langle w_j(n) \xi_j^{\mu_{n+1}} \xi_i^{\mu_{n+1}} \rangle \\ &= \langle w_i(n) \rangle + \gamma \sum_j \langle w_j(n) \rangle \langle \xi_j^{\mu_{n+1}} \xi_i^{\mu_{n+1}} \rangle \\ &= \langle w_i(n) \rangle + \gamma \sum_j C_{ij} \langle w_j(n) \rangle. \end{aligned} \quad (9.13)$$

The angular brackets denote an ensemble average over the whole sequence of input patterns $(\vec{\xi}^{\mu_1}, \vec{\xi}^{\mu_2}, \dots)$. The second equality is due to the fact that input patterns are chosen *independently* in each time step, so that the average over $w_j(n)$ and $(\xi_j^{\mu_{n+1}} \xi_i^{\mu_{n+1}})$ can be factorized. In the final expression we have introduced the correlation matrix C_{ij} ,

$$C_{ij} = \frac{1}{p} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu = \langle \xi_i^\mu \xi_j^\mu \rangle_\mu. \quad (9.14)$$

Expression (9.13) can be written in a more compact form using matrix notation,

$$\langle \vec{w}(n+1) \rangle = (\mathbf{I} + \gamma C) \langle \vec{w}(n) \rangle = (\mathbf{I} + \gamma C)^{n+1} \langle \vec{w}(0) \rangle, \quad (9.15)$$

where $\vec{w}(n) = (w_1(n), \dots, w_N(n))$ is the weight vector and \mathbf{I} is the identity matrix.

If we express the weight vector in terms of the eigenvectors \vec{e}_k of C ,

$$\langle \vec{w}(n) \rangle = \sum_k a_k(n) \vec{e}_k, \quad (9.16)$$

we obtain an explicit expression for $\langle \vec{w}(n) \rangle$ for any given initial condition $a_k(0)$, viz.,

$$\langle \vec{w}(n) \rangle = \sum_k (1 + \lambda_k)^n a_k(0) \vec{e}_k. \quad (9.17)$$

Since the correlation matrix is positive semi-definite all eigenvalues λ_k are real and positive. Therefore, the weight vector is growing exponentially, but the growth will soon be dominated by the eigenvector with the largest eigenvalue, i.e., the *first principal component*,

$$\langle \vec{w}(n) \rangle \xrightarrow{n \rightarrow \infty} (1 + \lambda_1)^n a_1(0) \vec{e}_1; \quad (9.18)$$

Recall that the output of the linear neuron model (9.10) is proportional to the projection of the current input pattern $\vec{\xi}^\mu$ on the direction \vec{w} . For $\vec{w} \propto \vec{e}_1$, the output is therefore proportional to the projection on the first principal component of the input distribution. A Hebbian learning rule such as (Eq. (9.11)) is thus able to extract the first principal component of the input data; see Fig. 9.6A.

From a data-processing point of view, the extraction of the first principle component of the input data set by a biologically inspired learning rule seems to be very compelling. There are, however, a few drawbacks and pitfalls. First, the above statement about the Hebbian learning rule is limited to the *expectation value* of the weight vector. We will see below that, if the learning rate is sufficiently low, then the actual weight vector is in fact very close to the expected one.

Second, principal components are only meaningful if the input data is normalized, i.e., distributed around the origin. This requirement is *not* consistent with a rate interpretation because rates are usually positive. This problem, however, can be overcome by learning rules with appropriately chosen linear terms c^{pre} and c^{post} in Eq. (9.2).

Third, while the direction of the weight vector moves in the direction of the principal component, the *norm* of the weight vector grows without bounds. However, we can use additional terms in the expansion indicated in Eq. (9.2) so as to guarantee a normalization of the weight vector. Let us recall Oja's rule from Eq. (9.7). After the presentation of n patterns the weight vector has a weight $w_j(n)$ and the next update step is

$$\Delta w_j = \gamma \nu^{\text{post}} \nu_j - \gamma w_j(n) (\nu^{\text{post}})^2. \quad (9.19)$$

In order to see that Oja's learning rule selects the first principal component we show that the eigenvectors $\{\vec{e}_1, \dots, \vec{e}_N\}$ of C are fixed points of the dynamics. For any fixed weight vector \vec{w} we can calculate the expectation of the weight change in the next time step by averaging over the whole ensemble of input patterns $\{\vec{\xi}^1, \vec{\xi}^2, \dots\}$. With $\langle \Delta \vec{w}(n) \rangle = \gamma C \vec{w}$ we find from Eq. (9.19)

$$\langle \Delta \vec{w} \rangle = \gamma C \vec{w} - \gamma \vec{w} [\vec{w} \cdot C \vec{w}], \quad (9.20)$$

We claim that any eigenvector \vec{e}_i of the correlation matrix C is a fixed point of Eq. (9.20). Indeed, if we substitute $\vec{w} = \vec{e}_i$ in the above equation we find that $\langle \Delta \vec{w} \rangle = 0$.

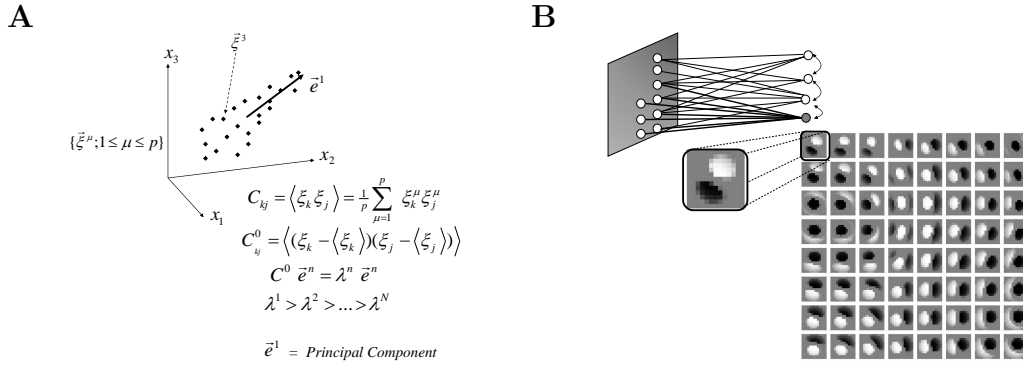


Fig. 9.6: **A.** A cloud of data points ξ^μ , the third data point is labeled. The principal component, i.e., the eigenvector of the correlation matrix with the largest eigenvalue, points in the direction of maximal variance of the data distribution. **B.** Receptive fields (small squares) of 64 cortical neurons (large grid). Each small square shows the distribution of weights for one of the cortical neurons. The receptive field of the neuron in the upper left corner of the grid is enlarged and the interpretation of the receptive field schematically indicated in the upper left part of the figure [adapted from Wimbauer et al. (1998)].

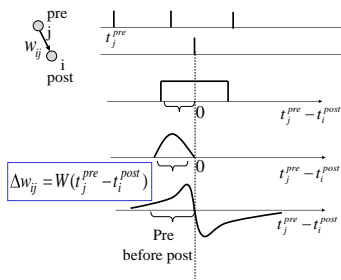
9.2.3 Application to Receptive Field Development

During Hebbian learning, neurons become sensitive to a subset of the inputs, viz. those inputs that show the strongest correlations. So far the analysis has been restricted to a single postsynaptic neuron. However, it is possible to repeat the essential mathematical steps for a model with a large number of output neurons located on a grid. We assume that model neurons are linear and have locally restricted lateral connectivity. Moreover, we assume that because of genetically encoded chemical gradients, a rough first arrangement is pre-wired at the beginning of cortical development, so that neurons in the upper left corner of the grid have receptive fields located in the the upper left region of the input space. However, even the rather rich genetic code does not contain enough information to tell all the billions of potential synapses between pre- and postsynaptic neurons where to form. Rather synaptic plasticity rules are in place during development that control the formation, strengthening, and weakening of synapses. Therefore we assume in the model that the synaptic weights from the input layer to the grid of postsynaptic neurons are subject to a Hebbian learning rule with a normalization term as in Oja's rule.

An analysis of such a model is possible (MacKay and Miller, 1990; Wimbauer et al., 1998). Here, however, we will just summarize some results from a computer simulation consisting of an array of 8×8 cortical neurons and two times 20×20 LGN neurons. Figure 9.6 shows a typical outcome of such a simulation after stimulation of local clusters of presynaptic neurons at random locations. Each of the small rectangles shows the receptive field of the corresponding cortical neuron. A bright color means that the neuron responds with an increased firing rate to a bright spot at that particular position within its receptive field; dark colors indicate inhibition.

There are two interesting aspects. First, the evolution of the synaptic weights has lead to asymmetric receptive fields, which give rise to orientation selectivity. Second, the structure of the receptive fields of neighboring cortical neurons are

A



B

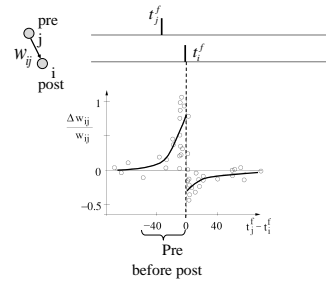


Fig. 9.7: **A** Coincidence of a presynaptic spike with a postsynaptic one needs to be defined can be defined by a learning window which could be symmetric and square (top) asymmetric (middle) or even contain two phases (bottom). **B** Experimentally found STDP window; figure adapted from Bi and Poo (1998).

similar; neuronal response properties thus vary continuously across the cortex. The neurons are said to form a *map* for, e.g., orientation. The formation of cortical maps is the topic of Chapter 14.

9.3 Spike Timing Dependent Plasticity

Traditional models of Hebbian learning have been formulated at the level of firing rates or abstract variables describing the activity of pre- and postsynaptic neuron. However, neuronal activity consists of spikes, short electrical pulses (action potentials). We show that the formulation of Hebbian learning rules on the level of spikes necessarily leads to the notion of Spike Timing Dependent Plasticity (STDP) with a window of simultaneity that describes the co-occurrence of spikes.

Models of STDP can be formulated on several levels. We do not discuss the description of detailed biophysics based models, mostly based on calcium, but focus directly on phenomenological or 'minimal' models. Spike timing dependent plasticity can be understood as a causal interaction between traces left at the synapses by spikes of presynaptic neurons (e.g. neurotransmitter) and backpropagating action potentials in the dendrite of the postsynaptic neuron. However, the voltage plays also an important role, when there is no backpropagating action potential.

9.3.1 Spikes versus Rates: The learning window

At the beginning of the chapter we emphasized two aspect of Hebbian learning: locality and correlation. When we formulated Hebbian learning on the level of rates, this led us a family of Hebbian rules. Locality implied that the rule can depend only on pre- and postsynaptic activity; correlation implied that post- and presynaptic firing rates are important *together*.

What does change if we turn to spikes? In a spiking neuron model, there are many more local variables than just the rate (see Chapter 2). There is the postsynaptic voltage, there are postsynaptic spikes, there may also be explicit variables for postsynaptic calcium. The local voltage at the synapse might be

different from that at the soma and the same is true for calcium, and all of these aspects may play a role. But let us keep things simple and focus on just postsynaptic spikes. Spikes are events that can be defined, e.g., as the onset of an action potential in real neurons or at the moment of threshold crossing in an integrate-and-fire neuron. As far as locality is concerned, a Hebbian learning rule could then depend on the spike timings t_j^f and t_i^f of a presynaptic neuron j and a postsynaptic neuron i .

Let us now turn to the second aspect, the notion of 'correlation that we substituted for Hebb's wording 'takes part in firing it'. A first, naive, approach would be to say that a synaptic change happens if pre- and postsynaptic firing occurs 'at the same time'. However, since firing times are threshold crossing events, they are infinitely short so that it never happens that two firings occur at the same time. To solve this issue, we need to set, somewhat arbitrarily, a time window or temporal resolution for our definition of simultaneity: If pre- and postsynaptic spikes occur within less than y ms, then we call the event simultaneous and a synaptic change occurs.

Such a definition can be visualized as a rectangular time window of width $2y$ centered symmetrically around the postsynaptic spike. However, why should the window be rectangular? And why should it be symmetrical? In fact, the Hebbian formulation that the presynaptic neuron 'takes part in firing' the postsynaptic one, suggest a causal relation and corresponds to a temporal order 'pre-before-post'. Hence, simultaneity in the Hebbian sense should be defined as a temporal window that is shifted slightly to the left as depicted in Fig. 9.7.

In the context of rate models of Hebbian learning, we have already seen that Hebb did not specify any conditions for depression of synapses. We are therefore free to complement the asymmetric learning window by a negative part for post-before-pre timing (Fig. 9.7), and this is indeed what has been postulated by theoreticians (Gerstner et al., 1996) and found in experiments (Markram et al., 1997; Bi and Poo, 1998; Caporale and Dan, 2008). Hebbian plasticity at the level of spikes as been termed Spike-Timing Dependent Plasticity (STDP).

STDP with an asymmetric learning window can be described mathematically as a window function $W(s)$ that is positive for $s < 0$ and negative for $s > 0$ and decays exponentially with time constants τ_+ and τ_- , respectively

$$W(s) = \begin{cases} A_+ \exp(s/\tau_+), & \text{if } s < 0 \\ A_- \exp(-s/\tau_-), & \text{if } s > 0 \end{cases} \quad (9.21)$$

with $A_+ > 0$ and $A_- < 0$.

Apart from the novel learning window, synaptic plasticity on the level of spikes is implemented analogous to rate-based plasticity. Let us describe a presynaptic spike train $S_j(t) = \sum_f \delta(t - t_j^f)$ as the sequence of presynaptic firing times and analogously the postsynaptic spike train as $S_i(t) = \sum_f \delta(t - t_i^f)$. Synaptic weights change whenever presynaptic spikes arrive or when postsynaptic action potentials are triggered,

$$\begin{aligned} \frac{d}{dt} w_{ij}(t) = & a_0 + a_1^{\text{pre}} S_j(t) + a_1^{\text{post}} S_i(t) \\ & + S_j(t) \int_0^\infty W(s) S_i(t-s) ds + S_i(t) \int_0^\infty W(-s) S_j(t-s) ds; \end{aligned} \quad (9.22)$$

In analogy to the rate picture, presynaptic spikes alone could cause a change proportional to a_1^{pre} and postsynaptic spikes along proportional to a_1^{post} and there could be a spontaneous decay or increase proportional to a_0 . The essential feature of STDP is implemented by the two terms containing the learning window W . The first one accounts for post-before-pre timings, the second one for pre-before-post.

9.3.2 A minimal model of STDP

The correlation condition in Hebb's postulate suggests that at least two biochemical components are involved in the induction of LTP. We do not wish to speculate on the nature of these components, but simply call them a and b . We assume that the first component is generated by a chemical reaction chain triggered by presynaptic spike arrival. In the absence of further input, the concentration $[a]$ decays with a time constant τ_a back to its resting level $[a] = 0$. A simple way to describe this process is

$$\frac{d}{dt}[a] = -\frac{[a]}{\tau_a} + \alpha_+ \sum_f \delta(t - t_j^f), \quad (9.23)$$

where the sum runs over all *presynaptic* firing times t_j^f . Equation (9.23) states that $[a]$ is increased at each arrival of a presynaptic spike by an amount α_+ . Since it is a linear equation, it can be integrated

$$[a](t) = \sum_{t_j^f} \alpha_+ \exp\left(-\frac{t - t_j^f}{\tau_a}\right) = \int_0^\infty \alpha_+ \exp\left(-\frac{s}{\tau_a}\right) S_j(t - s) ds \quad (9.24)$$

A high level of $[a]$ sets the synapse in a state where it is susceptible to changes in its weight. The variable $[a]$ by itself, however, does not yet trigger a weight change.

To generate the synaptic change, another substance b is needed. The production of b is controlled by a second process triggered by *postsynaptic* spikes,

$$\frac{d}{dt}[b] = -\frac{[b]}{\tau_b} + \frac{1}{\tau_b} \sum_f \delta(t - t_i^f), \quad (9.25)$$

where τ_b is another time constant. The sum runs over all postsynaptic spikes t_i^f . Note that the second variable $[b]$ does not need to be a biochemical quantity; it could, for example, be the electrical potential caused by the postsynaptic spike itself. In the following, we assume that the time constant $\tau_b \ll \tau_a$ is so short, that the process $[b]$ can be considered as instantaneous. In the limit $\tau_b \rightarrow 0$, the process $[b]$ can be approximated by a sequence of short pulses at the moment of the postsynaptic spikes, $[b](t) = S_i(t) = \sum_f \delta(t - t_i^f)$.

Hebbian learning needs both 'substances' to be present at the same time, thus

$$\frac{d}{dt}w_{ij}^{\text{corr}} = \gamma [a(t)] [b(t)], \quad (9.26)$$

with some rate constant γ . The upper index *corr* is intended to remind us that we are dealing only with the correlation term on the right-hand side of Eq. (9.22). If the process b is fast, then the process $[b]$ simply 'reads' out the value of $[a(t)]$ at the moment of each postsynaptic spike:

$$\frac{d}{dt}w_{ij}^{\text{corr}} = \gamma [a(t)] S_i(t) \quad (9.27)$$

Using our previous result for $[a(t)]$ we have

$$\frac{d}{dt}w_{ij}^{\text{corr}} = S_i(t) \int_0^\infty \gamma \alpha_+ \exp(-\frac{s}{\tau_a}) S_j(t-s) ds \quad (9.28)$$

which corresponds to the pre-before-post term in Eq. (9.22). It does not need a lot of work to convince ourselves that we need another set of substances $[c]$ and $[d]$ so as to construct a process with post-before-pre timing (see Exercise 3).

9.3.3 Voltage and spike timing

The learning window shows the amount of plasticity as a function of the time difference between pre- and postsynaptic spikes. The typical STDP plot, however, overemphasizes the role of pairs of spikes.

There are several important aspects that need to be highlighted. First, in experiments a single pair of one presynaptic spike followed by one postsynaptic spike has no measurable effect. Even fifty or sixty pairs of pre-before-post, if given at a repetition frequency of 1Hz or less, causes no potentiation of the synapses (Senn et al., 2001; Sjöström et al., 2001). Hence, one should consider models where the basic element is not a pair of spikes, but triplets of spikes such as post-pre-post or pre-post-post (Pfister and Gerstner, 2006).

Second, even in the absence of postsynaptic spikes potentiation of synapses is possible, if the postsynaptic neuron is sufficient depolarized (Artola et al., 1990; Ngezahayo et al., 2000). Moreover, isolated pairs of pre-post timing do induce potentiation, if they are preceded by a weak depolarization of the postsynaptic membrane (Sjöström et al., 2001). Hence, the more fundamental quantity is probably not spike timing but postsynaptic voltage. Indeed, A model where the local variable at the postsynaptic site of the synapse is not spike timing, but low-passed filtered voltage is capable of explaining a large body of experimental results (Clopath et al., 2010). Since action potentials correspond to short, but very pronounced peaks of the postsynaptic voltage, the spike-timing dependence of Hebbian plasticity follows from these models.

The exact signal processing chain that biophysically generates the change of the synapse is still a mystery. It is clear, however, that the time course of the postsynaptic calcium concentration plays a role which has led to several calcium based models of synaptic plasticity (Shouval et al., 2002; Lisman and Zhabotinsky, 2001; Rubin et al., 2005).

9.3.4 Functional Consequences of STDP

For stationary Poisson firing, STDP models can be mapped exactly to rate-based models. The main difference is that STDP models have a causal effect at the level

of spikes: if the presynaptic neuron fires, an EPSP is triggered in the postsynaptic neuron which makes the postsynaptic neuron more likely to fire. This additional effect can be included in the rate rate model. For a mathematical treatment see Chapter 11 in Gerstner and Kistler (2002).

If neurons do not fire in a stationary Poisson fashion, but in a precise temporal order, the main effect of STDP can be easily understood. Consider a network of 10 neurons with all-to-all connectivity and weight values distributed randomly between 0.5 and 1.5. Suppose an external input makes the neurons fire in the order $1 \rightarrow 2 \rightarrow 3 \rightarrow \dots 10 \rightarrow 1 \dots$. In this case lateral connections between these neurons develop a unidirectional ring structure. Thus the temporal order defined by a coding scheme with millisecond precision is reflected in the directionality of the connections (Clopath et al., 2010). In a pure rate picture of Hebbian learning, where rates are defined as the number of spikes in a time window of 100ms, the same stimulus would strengthen all the weights symmetrically.

9.4 Reward-Modulated Learning

9.4.1 Learning Depends on Feedback

At the beginning of this chapter we mentioned different types of tasks. Learning to ski, to skateboard or to ride a bicycle: all three tasks involve feedback. It hurts if you fall down from the skateboard, and you get the praise of your peers if you succeed a new trick.

Hebbian learning, as discussed in this chapter, is an unsupervised learning rule: There is no notion of good or bad, successful or unsuccessful, rewarding or painful. Hebbian learning is suitable to detect correlations in the input, and can therefore be used to explain developmental processes such as the formation of receptive fields.

Most tasks where you learn a novel activity, however, have the notion of success or reward attached to them. Therefore learning these tasks leaves the realm of Hebbian learning.

9.4.2 Unsupervised versus Reinforcement Learning

In the field of machine learning, one often distinguishes three types of learning tasks: unsupervised learning, supervised learning, and reinforcement learning. Unsupervised learning allows to detect correlations in the stream of input signals. The output of unsupervised learning can be to find the first principal component of the signal, an independent component, or the center of a cluster of data points. Hebbian learning falls into the class of unsupervised learning rules.

In supervised learning the neural network or artificial machine has to give, for every input signal, an output which is as close as possible to a target output (and it should generalize to novel inputs, too).

Reinforcement learning shares with supervised learning that there is a target output. In contrast to supervised learning, the feedback given to the learning system, however, is scarce. If we take the task of riding a bicycle, in supervised learning the supervisor would show how to turn the handle at every moment in

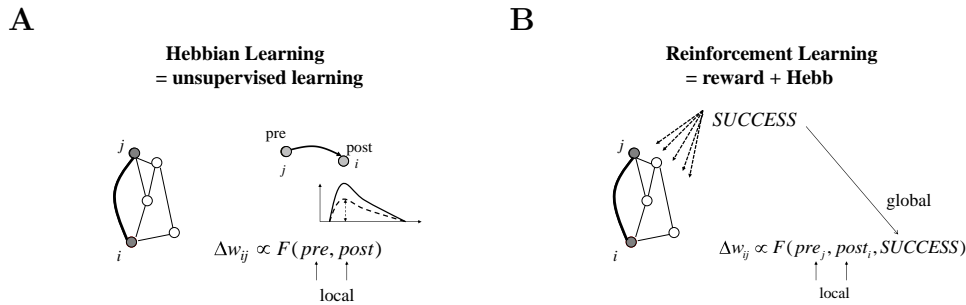


Fig. 9.8: **A** Unsupervised Hebbian learning depends only on the activity (filled circles) of the neurons i and j that are connected. **B** In addition plasticity depends on neuromodulators, that distribute information about the success (reward minus expected reward) of an action globally across the brain or a brain area.

time whereas in reinforcement learning the only feedback consists in telling the learning when it failed (i.e., it hurts when you fall). Reinforcement learning is the topic of Chapter 15.

9.4.3 Reward-based Hebbian Learning

Supervised learning methods, as studied in machine learning, suffer from the fact that it is difficult to see how they could be implemented biologically. Reinforcement learning methods, however, can be envisaged to be realized in the brain.

The basic idea is that we take a Hebbian learning rule which has two local factors (i.e. presynaptic and postsynaptic activity) and make the weight change proposed by the Hebbian learning rule subject to the presence of a neuromodulator signal S . The neuromodulator can signal 'reward' or more generally 'success S ' defined as reward minus expected reward. Indeed, the neuromodulator dopamine which is diffusively transmitted across broad areas of the brain, transmits a reward-related signal (Schultz et al., 1997). Experimental evidence indicates that Hebbian Plasticity and STDP are under the influence of neuromodulators (Reynolds and Wickens, 2002; Pawlak and Kerr, 2008; Seol et al., 2007; Pawlak et al., 2010).

It has been shown in many modeling studies that Hebbian learning can be used for reward-based learning, if the Hebbian learning rule, or STDP, is modulated by neuromodulatory signals (Sheynikhovich et al., 2009; Fremaux et al., 2010; Legenstein et al., 2008; Loewenstein and Seung, 2006; Loewenstein, 2008; Izhikevich, 2007). However, to achieve successful learning of nontrivial tasks, it is important that the success signal S defined as 'reward minus expected reward' be perfectly balanced, i.e., the brain needs an internal module capable of precisely estimating the reward expected for a given stimulus (Loewenstein, 2008; Fremaux et al., 2010).

9.4.4 Learning Novel Items

Finally, let us return to the question of how to form memories of facts or events and reconsider the example of the banana discussed at the beginning of the chapter. The Hebbian learning rules discussed in this chapter are capable of forming

an internal memory trace of the banana, but they will also form memory traces of everything else we see or feel or think: our direction of gaze changes two or three times per second so that even during a single day more than a hundred thousand perceived images would have to be stored. It can be shown that such a continued bombardment with sensory items leads to rapid overwriting of memories, with the result that what we have memorized earlier is quickly washed out and forgotten (Fusi, 2002). In order to make memory storage and retrieval stable, one needs to postulate a gating mechanism which decides which perception is novel or interesting and worth storing. Similar to the situation discussed for reward-based Hebbian learning, neuromodulators could play the role of global, brain-wide gating signal (Pawlak et al., 2010) Hebbian learning of a synapse from neuron j to i depends then not only on the activities of those two neurons, but also on the presence of a more global neuromodulator that enables the initiation or stabilization of Hebbian weight changes (Hasselmo, 1999; Clopath et al., 2008; Frey and Morris, 1998). The topics sketched here in the last few paragraphs will be discussed in greater depth in Chapter 15 (Reinforcement Learning) and Chapter 9 (Neuromodulation).

9.5 Core Set of Readings:

a) Spiking Neuron Models, by W. Gerstner and W. Kistler, chapters 10-12 (Cambridge Univ. Press, 2002)

b) Theory of Cortical Plasticity by L.N. Cooper, N. Intrator, B.S. Blais and H.Z. Shouval, (World Scientific, 2004)

c) Timing is not everything: neuromodulation opens the STDP gate, V. Pawlak and J.R. Wickens and A. Kirkwood and J.N.D. Kerr, *Front. Synaptic Neurosci.* (2010)

d) J. Sjöström and W. Gerstner, Spike-Timing Dependent Plasticity, *Scholarpedia* (2010)

9.6 Exercises

Exercise 1:

Assume that plasticity is controlled by the learning rule (9.8) with a parameter $\theta = 20\text{Hz}$. The postsynaptic neuron is described as a rate neuron with firing rate $\nu_i = g(\sum_j w_{ij}\nu_j)$. For the sake of simplicity we assume that g is linear, i.e. $g(x) = x$. All synaptic weights and firing rates are positive.

a) Consider a single presynaptic neuron j that is connected with a weight w_{ij} to the postsynaptic neuron. Show that the postsynaptic rate has a fixed point at $\nu_i = \theta$.

b) Show that this fixed point is unstable.

c) Suppose that there are 20 presynaptic neurons. The input alternates between two patterns. In the first pattern, the first half of neurons $1 \leq j \leq 10$ fires at a rate of $\nu_j = 1\text{Hz}$ while the other presynaptic neurons are silent. In the second pattern, the first half of neurons is silent while the other neurons $11 \leq j \leq 20$ fire at a rate of $\nu_j = 3\text{Hz}$. The input switches several times back and forth between the two patterns. Suppose that initially all weights have the same value $w_{ij} = 1$. How do the weights evolve in the presence of the sequence of input patterns? Allow for hard bounds at $w_{ij} = 0$ and $w_{ij} = 2$.

d) As in c, but suppose that during the first pattern the first group fires at 2.5Hz . What happens?

Exercise 2.

All eigenvectors of the correlation matrix are fixed points under Oja's learning rule. Show that only the eigenvector corresponding to the largest eigenvalue is stable.

To do so, assume that the current weight vector is close to eigenvector \vec{e}_i , but has a small perturbation in the direction of \vec{e}_j . Hence, write

$\vec{w} = \vec{e}_i + c\vec{e}_j$ where $|c| \ll 1$ is the amplitude of the perturbation and analyze whether the perturbation grows or decays.

Exercise 3.

a) Derive the exponential learning window (9.21) assuming four substances with concentrations $[a]$, $[b]$, $[c]$, $[d]$ and time constants $\tau_a, \tau_b, \tau_c, \tau_d$ respectively and dynamics analogous to (9.23), where τ_c is a second presynaptic process and τ_d a further postsynaptic process. To do so, take the limit $\tau_b \rightarrow 0$ and $\tau_c \rightarrow 0$.

b) What happens if you do not take the two limits, but keep τ_b and τ_c in the range of a few milliseconds. Show that the resulting learning window is a smooth function that does not necessarily go through the origin.

Bibliography

- Abbott, L. F. and Nelson, S. B. (2000). Synaptic plasticity - taming the beast. *Nature Neuroscience*, 3:1178–1183.
- Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275:220–224.
- Artola, A., Bröcher, S., and Singer, W. (1990). Different voltage dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature*, 347:69–72.
- Bi, G. and Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.*, 18:10464–10472.
- Bienenstock, E., Cooper, L., and Munro, P. (1982). Theory of the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.*, 2:32–48.
- Buonomano, D. and Merzenich, M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience*, 21:149–186.
- Caporale, N. and Dan, Y. (2008). Spike timing-dependent plasticity: A hebbian learning rule. *Ann. Rev. Neurosci.*, 31:25–46.
- Clopath, C., Busing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: A model of voltage-based spike-timing-dependent-plasticity with homeostasis. *Nature Neuroscience*, 13:344–352.
- Clopath, C., Ziegler, L., Vasilaki, E., Busing, L., and Gerstner, W. (2008). Tag-trigger-consolidation: A model of early and late long-term-potentiation and depression. *PLOS Comput. Biol.*, 4:e1000248.
- Fremaux, N., Sprekeler, H., and Gerstner, W. (2010). Functional requirements for reward-modulated spike-timing-dependent plasticity. *J. Neurosci.*, 40:13326–13337.
- Frey, U. and Morris, R. (1998). Synaptic tagging: implications for late maintenance of hippocampal long-term potentiation. *Trends in Neurosciences*, 21:181–188.
- Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.*, 87:459–470.
- Gerstner, W., Kempter, R., van Hemmen, J., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383(6595):76–78.

- Gerstner, W. and Kistler, W. K. (2002). *Spiking Neuron Models: single neurons, populations, plasticity*. Cambridge University Press, Cambridge UK.
- Hasselmo, M. (1999). Neuromodulation: acetylcholine and memory consolidation. *Trends in Cognitive Sciences*, 3:351–359.
- Hebb, D. O. (1949). *The Organization of Behavior*. Wiley, New York.
- Izhikevich, E. (2007). Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral Cortex*, 17:2443–2452.
- Legenstein, R., Pecevski, D., and Maass, W. (2008). A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLOS Comput. Biol.*, 4:e1000180.
- Lisman, J. and Zhabotinsky, A. (2001). A model of synaptic memory: A CaMKII/PP1 switch that potentiates transmission by organizing an AMPA receptor anchoring assembly. *Neuron*, 31:191–201.
- Loewenstein, Y. (2008). Robustness of learning that is based on covariance-driven synaptic plasticity. *PLOS Comput. Biol.*, 4:e1000007.
- Loewenstein, Y. and Seung, H. (2006). Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity. *Proc. Natl. Acad. Sci. USA*, 103:15224–15229.
- MacKay, D. J. C. and Miller, K. D. (1990). Analysis of linsker’s application of hebbian rules to linear networks. *Network*, 1:257–297.
- Makram, H., Sjostrom, J., and Gerstner, W. (2011). A history of spike-timing dependent plasticity. *Frontiers in Synaptic Neuroscience*, xx:xx.
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic AP and EPSP. *Science*, 275:213–215.
- Markram, H., Wu, Y., and Tosdyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. USA*, 95:5323–5328.
- Miller, K. D. and MacKay, D. J. C. (1994). The role of constraints in hebbian learning. *Neural Computation*, 6:100–126.
- Ngezahayo, A., Schachner, M., and Artola, A. (2000). Synaptic activation modulates the induction of bidirectional synaptic changes in adult mouse hippocampus. *J. Neuroscience*, 20:2451–2458.
- Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Mathematical Biology*, 15:267–273.
- Pawlak, V. and Kerr, J. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neuroscience*, 28:2435–2446.
- Pawlak, V., Wickens, J., Kirkwood, A., and Kerr, J. (2010). Timing is not everything: neuromodulation opens the stdp gate. *Front. Synaptic Neurosci.*, 2:146.
- Pfister, J.-P. and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *J. Neuroscience*, 26:9673–9682.

- Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *The Journal of Neuroscience*, 13:87–103.
- Reynolds, J. and Wickens, J. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, 15:507–521.
- Rubin, J., Gerkin, R., Bi, G.-Q., and Chow, C. (2005). Calcium time course as a signal for spike-timing-dependent plasticity. *J. Neurophysiology*, 93:2600–2613.
- Schultz, W., Dayan, P., and Montague, R. (1997). A neural substrate for prediction and reward. *Science*, 275:1593–1599.
- 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.
- Seol, G., Ziburkus, J., Huang, S., Song, L., Kim, I., Takamiya, K., Hugarir, R., Lee, H.-K., and Kirkwood, A. (2007). Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron*, 55:919–929.
- Sheynikhovich, D., Chavarriaga, R., Strosslin, T., Arleo, A., and Gerstner, W. (2009). Is there a geometric module for spatial orientation? insights from a rodent navigation model. *Psychological Review*, 116:540–566.
- Shouval, H. Z., Bear, M. F., and Cooper, L. N. (2002). A unified model of nmda receptor dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. USA*, 99:10831–10836.
- Sjöström, P., Turrigiano, G., and Nelson, S. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32:1149–1164.
- Wimbauer, S., Gerstner, W., and van Hemmen, J. (1998). Analysis of a correlation-based model for the development of orientation-selective receptive fields in the visual cortex. *Network*, 9:449–466.