19

Synaptic plasticity and learning

In the network models discussed in Parts III and IV, each synapse has so far been characterized by a single *constant* parameter w_{ij}, called the synaptic weight, synaptic strength, or synaptic efficacy. If w_{ij} is constant, the amplitude of the response of a postsynaptic neuron i to the arrival of action potentials from a presynaptic neuron j should always be the same. Electrophysiological experiments, however, show that the response amplitude is not fixed but can change over time. In experimental neuroscience, changes of the synaptic strength are called synaptic plasticity.

Appropriate stimulation paradigms can induce changes of the postsynaptic response that last for hours or days. If the stimulation paradigm leads to a persistent *increase* of the synaptic efficacy, the effect is called long-term potentiation of synapses, or LTP for short. If the result is a *decrease* of the synaptic efficacy, it is called long-term depression (LTD). These persistent changes are thought to be the neuronal correlate of learning and memory. LTP and LTD are different from short-term synaptic plasticity such as synaptic facilitation or depression that we have encountered in Section 3.1. Facilitated or depressed synapses decay back to their normal strength within less than a few seconds, whereas, after an LTP or LTD protocol, synapses keep their new values for hours. The long-term storage of the new values is thought to be the basis of long-lasting memories.

In the formal theory of neural networks, the weight w_{ij} of a connection from neuron j to i is considered a parameter that can be adjusted so as to optimize the performance of a network for a given task. The process of parameter adaptation is called *learning* and the procedure for adjusting the weights is referred to as a *learning rule*. Here learning is meant in its widest sense. It may refer to synaptic changes during development just as much as to the specific changes necessary to memorize a visual pattern or to learn a motor task. There are many different learning rules, all of which we cannot cover in this chapter. In particular, we leave aside the large class of "supervised" learning rules which are an important topic in the fields of artificial neural networks and machine learning. Here we focus on two other classes of learning rules that are of biological relevance.

In Section 19.1 we introduce the Hebb rule and discuss its relation to experimental protocols for long-term potentiation (LTP) and spike-timing-dependent plasticity (STDP). In Section 19.2 we formulate mathematical models of Hebbian plasticity. We shall see in Section 19.3 that Hebbian plasticity causes synaptic connections to tune to the statistics of the

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

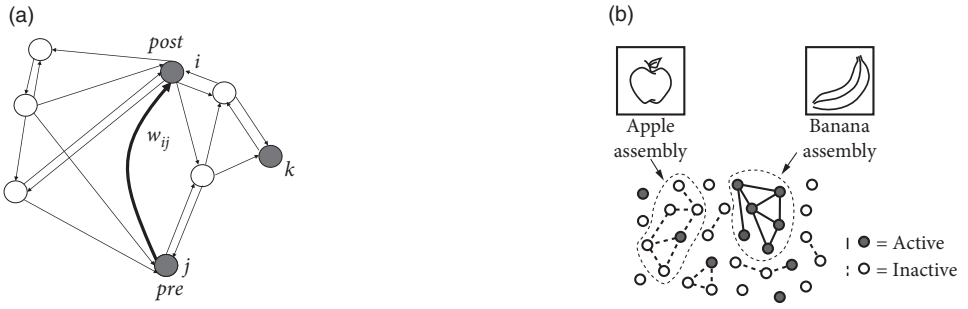

Fig. 19.1 Hebbian learning. (a) The change of a synaptic weight w_{ij} depends on the state of the presynaptic neuron j and the postsynaptic neuron i and the present efficacy w_{ij}, but not on the state of other neurons k. (b) Hebbian learning strengthens the connectivity within assemblies of neurons that fire together, for example during the perception of a banana. Schematic figure.

input. Such a self-tuning of network properties is an example of unsupervised learning. While unsupervised learning is thought to be a major drive for developmental plasticity in the brain, it is not sufficient to learn specific behaviors such as pressing a button in order to receive a reward. In Section 19.4 we discuss reward-based learning rules in the form of STDP modulated by reward. Reward-modulated synaptic plasticity is thought to be the basis of behavioral learning observed in animal conditioning experiments.

19.1 Hebb rule and experiments

Since the 1970s, a large body of experimental results on synaptic plasticity has been accumulated. Many of these experiments are inspired by Hebb's postulate (Hebb, 1949), which describes how the connection from a presynaptic neuron A to a postsynaptic neuron B should be modified:

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

Today this famous postulate is often rephrased in the sense that modifications of the synaptic transmission efficacy are driven by correlations in the firing activity of pre- and postsynaptic neurons; see Fig. 19.1a. The shortest summary is: neurons that "fire together, wire together" (Shatz, 1992). Note that the term "fire together" is less precise than Hebb's original formulation which contains an asymmetry since a neuron that "contributes to firing" another one has to be active slightly *before* the latter. Even though the idea of learning through correlations dates further back in the past (James, 1890), correlation-based learning is now generally called *Hebbian learning*.

Hebb formulated his principle on purely theoretical grounds. He realized that such a mechanism would help to stabilize specific neuronal activity patterns in the brain; see

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

Fig. 19.1b. If neuronal activity patterns correspond to behavior, then stabilization of specific patterns implies learning of specific types of behaviors (Hebb, 1949). We emphasize that Hebbian learning is unsupervised, because there is no notion of "good" or "bad" changes of a synapse. Synaptic changes happen whenever there is joint activity of pre- and postsynaptic neurons, i.e., they are driven by the neuronal firing patterns. These patterns may reflect sensory stimulation as well as ongoing brain activity, but there is no feedback signal from a "supervisor" or from the environment.

In this section we review experimental protocols that induce lasting synaptic changes and discuss their relation to Hebbian learning.

19.1.1 Long-term potentiation

The classic paradigm of LTP induction is, very schematically, the following (Brown *et al.*, 1989; Bliss and Collingridge, 1993). Neuronal activity is monitored by an extracellular or intracellular electrode, while presynaptic fibers are stimulated by means of a second (extracellular) electrode. Small pulses are applied to the presynaptic fibers in order to measure the strength of the postsynaptic response (Fig. 19.2a). The amplitude of the test pulse is chosen such that the stimulation evokes a postsynaptic potential, but no action potentials.

In a second step, the input fibers are strongly stimulated by a sequence of high-frequency pulses so as to evoke postsynaptic firing (Fig. 19.2b). After that, the strength of the postsynaptic response to small pulses is tested again and a significantly increased amplitude of postsynaptic potentials is found (Fig. 19.2c). This change in the synaptic strength persists over many hours and is thus called *long-term potentiation* or LTP (Fig. 19.2d).

The increase of the synaptic weights can be interpreted as "Hebbian," because it occurred after an episode of joint activity of pre- and postsynaptic neurons. Early LTP experiments were done with two extracellular electrodes (one for the stimulation of presynaptic fibers, the other for the measurement of the neuronal response), but in later experiments LTP was also studied with intracellular recordings.

Example: Voltage dependence of LTP

With an intracellular electrode, an experimenter can not only record the response to an incoming spike, but also manipulate the membrane potential of the postsynaptic neuron; see Fig. 19.3. If presynaptic spikes arrive during a period where the neuron is strongly depolarized, LTP is induced at the activated synapses. On the other hand, spike arrival combined with weak depolarization causes LTD (Artola *et al.*, 1990; Artola and Singer, 1993; Ngezahayo *et al.*, 2000). These and similar experiments reveal the importance of the postsynaptic voltage during the induction of synaptic plasticity. If we interpret strong depolarization of the postsynaptic neuron as a substitute of neuronal activity, the above protocol for LTP induction can be called "Hebbian."

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

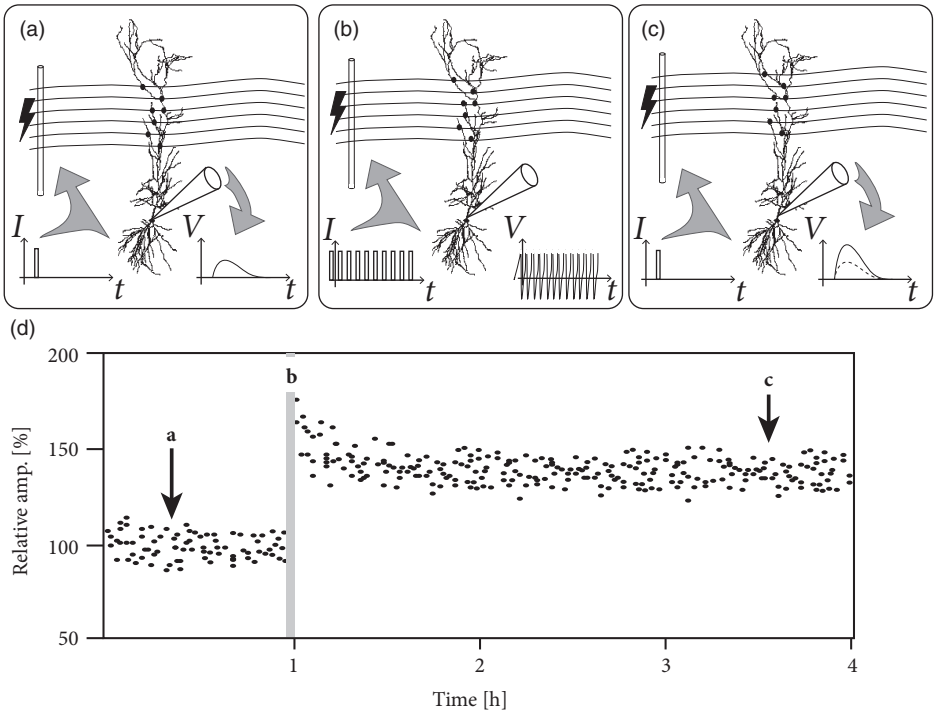

Fig. 19.2 Schematic drawing of a paradigm of LTP induction. (a) A weak test pulse (left) evokes the postsynaptic response sketched on the right-hand side of the figure. (b) A strong stimulation sequence (left) triggers postsynaptic firing (right, the peak of the action potential is out of bounds). (c) A test pulse applied some time later evokes a larger postsynaptic response (right; solid line) than the initial response. The dashed line is a copy of the initial response in (a). (d) The relative amplitude as measured with the test pulses illustrated in (a) and (c) is increased after the strong stimulation at $t = 1$ h. Schematic figure.

19.1.2 Spike-timing-dependent plasticity

Pairing experiments with multiple intracellular electrodes in synaptically coupled neurons have opened the possibility of studying synaptic plasticity at an excellent spatial and temporal resolution (Markram *et al.*, 1997; Zhang *et al.*, 1998; Debanne *et al.*, 1998; Bi and Poo, 1998, 1999; Sjöström *et al.*, 2001); see Bi and Poo (2001) and Sjöström and Gerstner (2010) for reviews.

Figure 19.4 illustrates a pairing experiment with cultured hippocampal neurons where the presynaptic neuron (j) and the postsynaptic neuron (i) are forced to fire spikes at time t_j^f and t_i^f, respectively (Bi and Poo, 1998). The resulting change in the synaptic efficacy Δw_{ij} after several repetitions of the experiment turns out to be a function of the difference $t_j^f - t_i^f$ between the firing times of the pre- and postsynaptic neuron. This observation

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

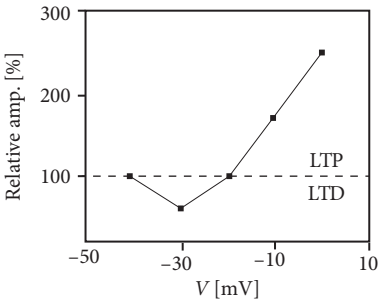

Fig. 19.3 Voltage dependence of LTP. Repeating the experiment shown in Fig. 19.2 while holding the membrane potential of the postsynaptic neuron shows a decrease in EPSP amplitude when holding at −30 mV and an increase when holding a higher voltage. Vertical axis: Ratio of EPSP amplitude $w_{ij}(T)/w_{ij}(0)$ before $(t = 0)$ and after $(t = T)$ the plasticity inducing protocol. Adapted from Ngezahayo *et al*. (2000).

has given rise to the term "spike-timing-dependent plasticity" (STDP). Most notably, the direction of the change depends critically on the relative timing of pre- and postsynaptic spikes on a millisecond time scale (Markram *et al*., 1997). The synapse is strengthened if the presynaptic spike occurs shortly before the postsynaptic neuron fires, but the synapse is weakened if the sequence of spikes is reversed; see Fig. 19.4b. This observation is indeed in agreement with Hebb's postulate because presynaptic neurons that are active slightly *before* the postsynaptic neuron are those which "take part in firing it" whereas those that fire later obviously did not contribute to the postsynaptic action potential. An asymmetric learning window such as the one in Fig. 19.4f is thus an implementation of the causality requirement that is implicit in Hebb's principle.

Similar results on spike-time-dependent synaptic plasticity have been found in various neuronal systems (Abbott and Nelson, 2000; Bi, 2002; Caporale and Dan, 2008), but there are also characteristic differences. Synapses between parallel fibers and "Purkinje-cells" in the cerebellar-like structure of electric fish, for example, show the opposite dependence on the relative timing of presynaptic input and the (so-called "broad") postsynaptic spike (Bell *et al*., 1997). In this case the synapse is weakened if the presynaptic input arrives shortly before the postsynaptic spike (anti-Hebbian plasticity).

19.2 Models of Hebbian learning

Before we turn to spike-based learning rules, we first review the basic concepts of correlation-based learning in a firing rate formalism. Firing rate models (see Chapter 15) have been used extensively in the field of artificial neural networks; see Hertz *et al*. (1991); Haykin (1994) for reviews.

19.2.1 A mathematical formulation of Hebb's rule

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 time being we content ourselves with a description in

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

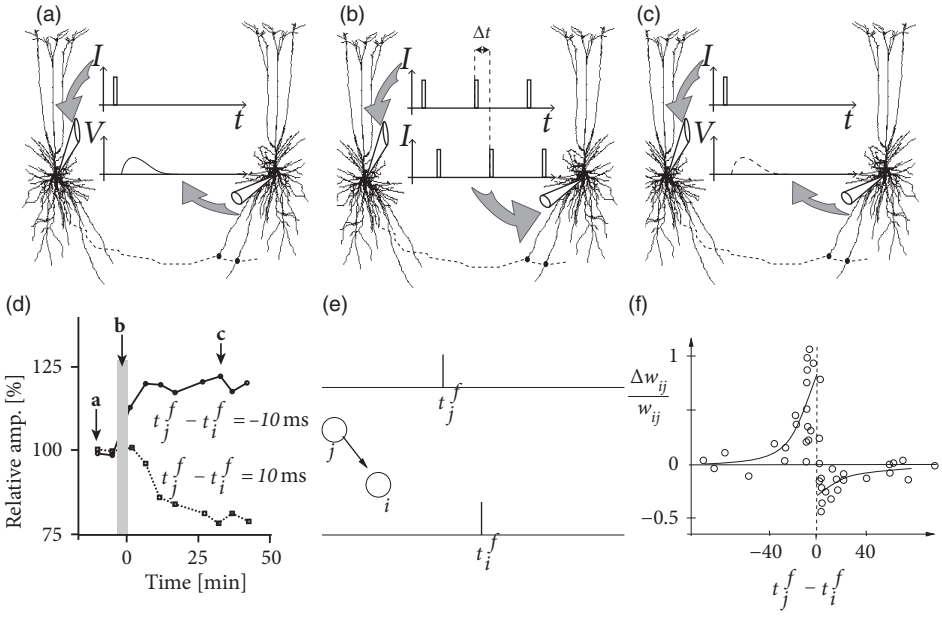

Fig. 19.4 Spike-Timing Dependent Plasticity. (a) Intracellular electrodes are used to manipulate two synaptically coupled neurons (axons are shown as dashed lines). A test pulse (I) injected into the presynaptic neuron causes an EPSP in the postsynaptic neuron (V). (b) During the plasticity induction protocol of a few seconds ("pairing"), both neurons are stimulated with current pulses forcing spikes at precise moments in time. (c) After the pairing protocol, the presynaptic neuron is stimulated by another current pulse, testing the level of potentiation of the synapse (before pairing protocol, dashed line; after, full line). (d) Amplitude of EPSP relative to initial amplitude as a function of time after the pairing protocol. If the presynaptic spike is 10 ms before the postsynaptic one, potentiation occurs (full line). If the order of the spikes is inverted, depression occurs (data points redrawn after Markram *et al*. (1997)). (e) Synaptic changes Δw_{ij} occur only if presynaptic firing at t_j^f and postsynaptic activity at t_i^f occur sufficiently close to each other. (f) The STDP window summarizes the timing requirements between pre- and postsynaptic spikes. Experimentally measured weight changes (circles) as a function of $t_j^f - t_i^f$ in milliseconds overlaid on a schematic two-phase learning window (solid line). A positive change (LTP) occurs if the presynaptic spike *precedes* the postsynaptic one; for a reversed timing, synaptic weights are decreased (data points redrawn after the experiments of Bi and Poo (1998)).

terms of mean firing rates. In what follows, the activity of the presynaptic neuron is denoted by ν_j and that of the postsynaptic neuron by ν_i.

There are two aspects of Hebb's postulate that are particularly important: *locality* and *joint activity*. Locality means that the change of the synaptic efficacy can depend only 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 write down a rather general formula for the change of the synaptic efficacy,

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = F(w_{ij}; \nu_i, \nu_j) \,. \tag{19.1}$$

Here, $\mathrm{d}w_{ij}/\mathrm{d}t$ is the rate of change of the synaptic coupling strength and F is a so-far-undetermined function (Sejnowski and Tesauro, 1989). We may wonder whether there are other local variables (e.g., the input potential h_i; see Chapter 15) 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 input potential h_i is uniquely determined by the postsynaptic firing rate, $\nu_i = g(h_i)$, with a monotone gain function g.

The second important aspect of Hebb's postulate is the notion of "joint activity" which implies that pre- and postsynaptic neurons 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 $\nu_i = \nu_j = 0$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_0(w_{ij}) + c_1^{\mathrm{pre}}(w_{ij})\, \nu_j + c_1^{\mathrm{post}}(w_{ij}) \nu_i + c_2^{\mathrm{pre}}(w_{ij})\, \nu_j^2 \\ + c_2^{\mathrm{post}}(w_{ij})\, \nu_i^2 + c_{11}^{\mathrm{corr}}(w_{ij})\, \nu_i \nu_j + \mathscr{O}(\nu^3) \,. \end{aligned} \tag{19.2}$$

The term containing c_{11}^{corr} on the right-hand side of (19.2) is bilinear in pre- and postsynaptic activity. This term implements the AND condition for joint activity. If the Taylor expansion had been stopped before the bilinear term, the learning rule would be called "non-Hebbian," because pre- or postsynaptic activity alone induces a change of the synaptic efficacy, and joint activity is irrelevant. Thus a Hebbian learning rule needs either the bilinear term $c_{11}^{\mathrm{corr}}(w_{ij})\, \nu_i \nu_j$ with $c_{11}^{\mathrm{corr}} > 0$ or a higher-order term (such as $c_{21}(w_{ij})\, \nu_i^2\, \nu_j$) that involves the activity of both pre- and postsynaptic neurons.

Example: Hebb rules, saturation, and LTD

The simplest choice for a Hebbian learning rule within the Taylor expansion of Eq. (19.2) is to fix c_{11}^{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,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c_{11}^{\mathrm{corr}}\, \nu_i \nu_j \,. \tag{19.3}$$

We note in passing that a learning rule with $c_{11}^{\mathrm{corr}} < 0$ is usually called anti-Hebbian because it weakens the synapse if pre- and postsynaptic neuron are active simultaneously, a behavior that is just contrary to that postulated by Hebb.

Note that, in general, the coefficient c_{11}^{corr} may depend on the current value of the weight w_{ij}. This dependence can be used to limit the growth of weights at a maximum value $w^{\max}$. The two standard choices of weight-dependence are called "hard bound" and "soft bound," respectively. Hard bound means that $c_{11}^{\mathrm{corr}} = \gamma_2$ is constant in the range $0 < w_{ij} < w^{\max}$ and zero otherwise. Thus, weight growth stops abruptly if w_{ij} reaches the upper bound $w^{\max}$.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

Post ν_i	Pre ν_j	$dw_{ij}/dt \propto \nu_i\,\nu_j$	$dw_{ij}/dt \propto \nu_i\,\nu_j - c_0$	$dw_{ij}/dt \propto (\nu_i - \nu_\theta)\,\nu_j$	$dw_{ij}/dt \propto \nu_i\,(\nu_j - \nu_\theta)$	$dw_{ij}/dt \propto (\nu_i - \langle\nu_i\rangle)(\nu_j - \langle\nu_j\rangle)$
ON	ON	+	+	+	+	+
ON	OFF	0	−	0	−	−
OFF	ON	0	−	−	0	−
OFF	OFF	0	−	0	0	+

Table 19.1 *The change* $\frac{d}{dt}w_{ij}$ *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 high rate* ($\nu > 0$)*, whereas "OFF" means an inactive neuron* ($\nu = 0$)*. From left to right: Standard Hebb rule, Hebb with decay, Hebb with postsynaptic or presynaptic LTP/LTD threshold, covariance rule. The parameters are* $0 < \nu_\theta < \nu^{\max}$ *and* $0 < c_0 < (\nu^{\max})^2$.

A soft bound for the growth of synaptic weights can be achieved if the parameter $c_{11}^{\rm corr}$ in Eq. (19.3) tends to zero as w_{ij} approaches its maximum value $w^{\max}$,

$$c_{11}^{\rm corr}(w_{ij}) = \gamma_2\,(w^{\max} - w_{ij})^\beta\,, \tag{19.4}$$

with positive constants γ_2 and β. The typical value of the exponent is $\beta = 1$, but other choices are equally possible (Gütig *et al.*, 2003). For $\beta \to 0$, the soft-bound rule (19.4) converges to the hard-bound one.

Note that neither Hebb's original proposal nor the simple rule (19.3) contains a possibility for a decrease of synaptic weights. However, in a system where synapses can only be strengthened, all efficacies will eventually saturate at their upper maximum value. Our formulation (19.2) is sufficiently general to allow for a combination of synaptic potentiation and depression. For example, if we set $w^{\max} = \beta = 1$ in (19.4) and combine it with a choice $c_0(w_{ij}) = -\gamma_0\,w_{ij}$, we obtain a learning rule

$$\frac{d}{dt}w_{ij} = \gamma_2\,(1 - w_{ij})\,\nu_i\,\nu_j - \gamma_0\,w_{ij}\,, \tag{19.5}$$

where, in the absence of stimulation, synapses spontaneously decay back to zero. Many other combinations of the parameters $c_0, \ldots, c_{11}^{\rm corr}$ in Eq. (19.2) exist. They all give rise to valid Hebbian learning rules that exhibit both potentiation and depression; see Table 19.1.

Example: Covariance rule

Sejnowski (1977) has suggested a learning rule of the form

$$\frac{d}{dt}w_{ij} = \gamma\,(\nu_i - \langle\nu_i\rangle)\,(\nu_j - \langle\nu_j\rangle)\,, \tag{19.6}$$

called the covariance rule. This rule is based on the idea that the rates $\nu_i(t)$ and $\nu_j(t)$ fluctuate around mean values $\langle\nu_i\rangle, \langle\nu_j\rangle$ that are taken as running averages over the recent

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

firing history. To allow a mapping of the covariance rule to the general framework of Eq. (19.2), the mean firing rates $\langle \nu_i \rangle$ and $\langle \nu_j \rangle$ have to be constant in time.

Example: Oja's rule

All of the above learning rules had $c_2^{\rm pre} = c_2^{\rm post} = 0$. Let us now consider a nonzero quadratic term $c_2^{\rm post} = -\gamma w_{ij}$. We take $c_{11}^{\rm corr} = \gamma > 0$ and set all other parameters to zero. The learning rule

$$\frac{\rm d}{{\rm d}t} w_{ij} = \gamma [\nu_i \nu_j - w_{ij} \nu_i^2] \tag{19.7}$$

is called Oja's rule (Oja, 1982). Under some general conditions 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. (19.3); see Exercises. 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.

Example: Bienenstock–Cooper–Munro rule

Higher-order terms in the expansion on the right-hand side of Eq. (19.2) lead to more intricate plasticity schemes. Let us consider

$$\frac{\rm d}{{\rm d}t} w_{ij} = \phi\,(\nu_i - \nu_\theta)\,\nu_j \tag{19.8}$$

with a nonlinear function ϕ and a reference rate ν_θ. If we take ν_θ to be a function $f(\langle \nu_i \rangle)$ of the average output rate $\langle \nu_i \rangle$, then we obtain the so-called Bienenstock–Cooper–Munro (BCM) rule (Bienenstock *et al.*, 1982).

The basic structure of the function ϕ is sketched in Fig. 19.5. If presynaptic activity is combined with moderate levels of postsynaptic excitation, the efficacy of synapses activated by presynaptic input is *decreased*. Weights are *increased* only if the level of postsynaptic activity exceeds a threshold, ν_θ. The change of weights is restricted to those synapses which are activated by presynaptic input. A common choice for the function ϕ is

$$\frac{\rm d}{{\rm d}t} w_{ij} = \eta\, \nu_i\,(\nu_i - \nu_\theta)\,\nu_j = c_{21} \nu_i^2 \nu_j - c_{11}^{\rm corr} \nu_i \nu_j, \tag{19.9}$$

which can be mapped to the Taylor expansion of Eq. (19.2) with $c_{21} = \eta$ and $c_{11}^{\rm corr} = -\eta\,\nu_\theta$.

For stationary input, it can be shown that the postsynaptic rate ν_i under the BCM-rule (19.9) has a fixed point at ν_θ which is unstable (see Exercises). To avoid the postsynaptic firing rate blowing up or decaying to zero, it is therefore necessary to turn ν_θ into an adaptive variable which depends on the average rate $\langle \nu_i \rangle$. The BCM rule leads to input

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

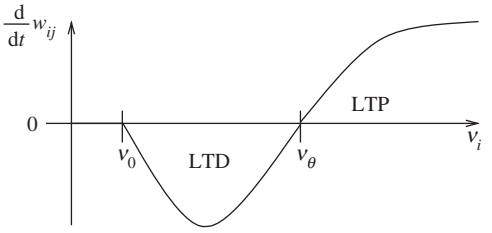

Fig. 19.5 BCM rule. Synaptic plasticity is characterized by two thresholds for the postsynaptic activity (Bienenstock *et al.*, 1982). Below ν_0 no synaptic modification occurs, between ν_0 and ν_θ synapses are depressed, and for postsynaptic firing rates beyond ν_θ synaptic potentiation can be observed. Often ν_0 is set to zero.

selectivity (see Exercises) and has been successfully used to describe the development of receptive fields (Bienenstock *et al.*, 1982).

19.2.2 Pair-based models of STDP

We now switch from rate-based models of synaptic plasticity to a description with spikes. Suppose a presynaptic spike occurs at time $t_{\rm pre}$ and a postsynaptic one at time $t_{\rm post}$. Most models of STDP interpret the biological evidence in terms of a pair-based update rule, i.e., the change in weight of a synapse depends on the temporal difference $|\Delta t| = |t_{\rm post} - t_{\rm pre}|$; see Fig. 19.4f. In the simplest model, the updates are

$$\begin{aligned} \Delta w_+ &= A_+(w) \cdot \exp(-|\Delta t|/\tau_+) \text{ at } t_{\rm post} \quad \text{for } t_{\rm pre} < t_{\rm post}, \\ \Delta w_- &= A_-(w) \cdot \exp(-|\Delta t|/\tau_-) \text{ at } t_{\rm pre} \quad \text{for } t_{\rm pre} < t_{\rm post}, \end{aligned} \tag{19.10}$$

where $A_\pm(w)$ describes the dependence of the update on the current weight of the synapse. The update of synaptic weights happens immediately after each presynaptic spike (at time $t_{\rm pre}$) and each postsynaptic spike (at time $t_{\rm post}$). A pair-based model is fully specified by defining: (i) the weight-dependence of the amplitude parameter $A_\pm(w)$; (ii) which pairs are taken into consideration to perform an update. A simple choice is to take all pairs into account. An alternative is to consider for each postsynaptic spike only the nearest presynaptic spike or vice versa. Note that spikes that are far apart hardly contribute because of the exponentially fast decay of the update amplitude with the interval $|\Delta t|$. Instead of an exponential decay (Song *et al.*, 2000), some other arbitrary time dependence, described by a learning window $W_+(s)$ for LTP and $W_-(s)$ for LTD is also possible (Gerstner *et al.*, 1996a; Kempter *et al.*, 1999a).

If we introduce $S_j = \sum_f \delta(t - t_j^f)$ and $S_i = \sum_f \delta(t - t_i^f)$ for the spike trains of pre- and postsynaptic neurons, respectively, then we can write the update rule in the form (Kistler and van Hemmen, 2000)

$$\begin{aligned} \frac{\rm d}{{\rm d}t} w_{ij}(t) = S_j(t) \left[a_1^{\rm pre} + \int_0^\infty A_-(w_{ij}) W_-(s)\, S_i(t-s)\, {\rm d}s \right] \\ + S_i(t) \left[a_1^{\rm post} + \int_0^\infty A_+(w_{ij}) W_+(s)\, S_j(t-s)\, {\rm d}s \right], \end{aligned} \tag{19.11}$$

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

where $W_\pm$ denotes the time course of the learning window while $a_1^{\rm pre}$ and $a_1^{\rm post}$ are non-Hebbian contributions, analogous to the parameters $c_1^{\rm pre}$ and $c_1^{\rm post}$ in the rate-based model of Eq. (19.2). In the standard pair-based STDP rule, we have $W_\pm(s) = \exp(-s/\tau_\pm)$ and $a_1^{\rm pre} = a_1^{\rm post} = 0$; see 19.10.

Example: Implementation by local variables

The pair-based STDP rule of 19.10 can be implemented with two local variables, i.e., one for a low-pass filtered version of the presynaptic spike train and one for the postsynaptic spikes. Suppose that each presynaptic spike at synapse j leaves a trace x_j, i.e., its update rule is

$$\frac{{\rm d}x_j}{{\rm d}t} = -\frac{x_j}{\tau_+} + \sum_f \delta(t - t_j^f), \tag{19.12}$$

where t_j^f is the firing time of the presynaptic neuron. In other words, the variable is increased by an amount of one at the moment of a presynaptic spike and decreases exponentially with time constant τ_+ afterward. Similarly, each postsynaptic spike leaves a trace y_i

$$\frac{{\rm d}y_i}{{\rm d}t} = -\frac{y_i}{\tau_-} + \sum_f \delta(t - t_i^f)\,. \tag{19.13}$$

The traces x_j and y_i play an important role during the weight update. At the moment of a presynaptic spike, a decrease of the weight is induced proportional to the value of the postsynaptic trace y_i. Analogously, potentiation of the weight occurs at the moment of a postsynaptic spike proportional to the trace x_j left by a previous presynaptic spike,

$${\rm d}w_{ij}/{\rm d}t = -A_-(w_{ij}) y_i(t) \sum_f \delta(t - t_j^f) + A_+(w_{ij}) x_j(t) \sum_f \delta(t - t_i^f)\,. \tag{19.14}$$

The traces x_j and y_i correspond here to the factors $\exp(-|\Delta t|/\tau_\pm)$ in 19.10. For the weight dependence of the factors A_- and A_+, one can use either hard bounds or soft bounds; see Eq. (19.4).

19.2.3 Generalized STDP models

There is considerable evidence that the pair-based STDP rule discussed above cannot give a full account of experimental results with STDP protocols. Specifically, they reproduce neither the dependence of plasticity on the repetition frequency of pairs of spikes in an experimental protocol, nor the results of triplet and quadruplet experiments.

STDP experiments are usually carried out with about 50−60 pairs of spikes. The temporal distance of the spikes in the pair is of the order of a few to tens of milliseconds, whereas

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

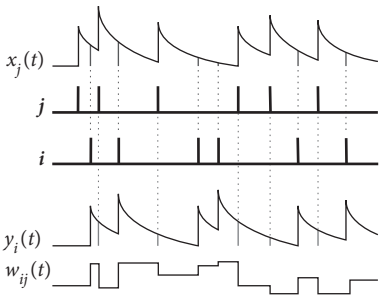

Fig. 19.6 Implementation of pair-based plasticity by local variables: The presynaptic spikes leave a trace $x_j(t)$, postsynaptic spikes a trace $y_i(t)$. The weight increases at the moment of a postsynaptic spike proportional to the momentary value of the trace $x_j(t)$ left by previous presynaptic spike arrivals. Analogously we get depression for post-before-pre pairings at the moment of a presynaptic spike (vertical dashed lines highlight moments of spike firing). From Morrison *et al*. (2008).

the temporal distance between the pairs is of the order of hundreds of milliseconds to seconds. In the case of a potentiation protocol (i.e., pre-before-post), standard pair-based STDP models predict that if the repetition frequency ρ is increased, the strength of the depressing interaction (i.e., post-before-pre) becomes greater, leading to less net potentiation. However, experiments show that increasing the repetition frequency leads to an increase in potentiation (Sjöström *et al*., 2001; Senn *et al*., 2001). Other experimenters have employed multiple-spike protocols, such as repeated presentations of symmetric triplets of the form pre-post-pre and post-pre-post (Bi, 2002; Froemke and Dan, 2002; Wang *et al*., 2005; Froemke *et al*., 2006). Standard pair-based models predict that the two sequences should give the same results, as they each contain one pre-post pair and one post-pre pair. Experimentally, this is not the case.

Here we review two examples of simple models which account for these experimental findings (Pfister and Gerstner, 2006; Clopath *et al*., 2010), but there are other models which also reproduce frequency dependence, (e.g., Senn, 2002).

Triplet model

One simple approach to modeling STDP which addresses the issues of frequency dependence is the triplet rule developed by Pfister and Gerstner (2006). In this model, LTP is based on sets of three spikes (one presynaptic and two postsynaptic). The triplet rule can be implemented with local variables as follows. Similarly to pair-based rules, each spike from the presynaptic neuron j contributes to a trace x_j at the synapse:

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = -\frac{x_j}{\tau_+} + \sum_{t_j^f} \delta\left(t - t_j^f\right),$$

where t_j^f denotes the firing times of the presynaptic neuron. Unlike pair-based rules, each spike from postsynaptic neuron i contributes to a fast trace $y_{i,1}$ and a slow trace $y_{i,2}$ at the

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

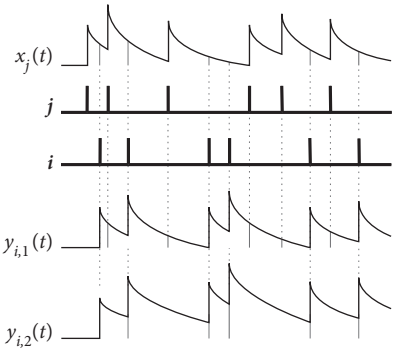

Fig. 19.7 Implementation of the triplet rule by local variables. The spikes of a presynaptic neuron j contribute to a trace $x_j(t)$, the spikes of postsynaptic neuron i contribute to a fast trace $y_{i,1}(t)$ and a slow trace $y_{i,2}(t)$. The update of the weight w_{ij} at the moment of a presynaptic spike is proportional to the momentary value of the fast trace $y_{i,1}(t)$, as in the pair-based model of Fig. 19.6. The update of the weight w_{ij} at the moment of a postsynaptic spike is proportional to the momentary value of the trace $x_j(t)$ and the value of the slow trace $y_{i,2}(t)$ just before the spike. Moments of weight update are indicated by vertical dashed lines. From Morrison *et al*. (2008).

synapse:

$$\frac{\mathrm{d}y_{i,1}}{\mathrm{d}t} = -\frac{y_{i,1}}{\tau_1} + \sum_f \delta(t - t_i^f),$$
$$\frac{\mathrm{d}y_{i,2}}{\mathrm{d}t} = -\frac{y_{i,2}}{\tau_2} + \sum_f \delta(t - t_i^f),$$

where $\tau_1 < \tau_2$; see Fig. 19.7. The new feature of the rule is that LTP is induced by a triplet effect: the weight change is proportional to the value of the presynaptic trace x_j evaluated at the moment of a postsynaptic spike and also to the slow postsynaptic trace $y_{i,2}$ remaining from previous postsynaptic spikes:

$$\Delta w_{ij}^+ \left(t_i^f\right) = A_+ \left(w_{ij}\right) x_j \left(t_i^f\right) y_{i,2} \left(t_i^{f-}\right) \tag{19.15}$$

where t_i^{f-} indicates that the function $y_{i,2}$ is to be evaluated before it is incremented due to the postsynaptic spike at t_i^f. LTD is analogous to the pair-based rule, given in (19.14), i.e., the weight change is proportional to the value of the fast postsynaptic trace $y_{i,1}$ evaluated at the moment of a presynaptic spike.

The triplet rule reproduces experimental data from visual cortical slices (Sjöström *et al*., 2001) that increasing the repetition frequency in the STDP pairing protocol increases net potentiation (19.8). It also gives a good fit to experiments based on triplet protocols in hippocampal culture (H.-X.Wang *et al*., 2005).

The main functional advantage of such a triplet learning rule is that it can be mapped to the BCM rule of Eqs. (19.8) and (19.9): if we assume that the pre- and postsynaptic spike trains are governed by Poisson statistics, the triplet rule exhibits depression for low

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

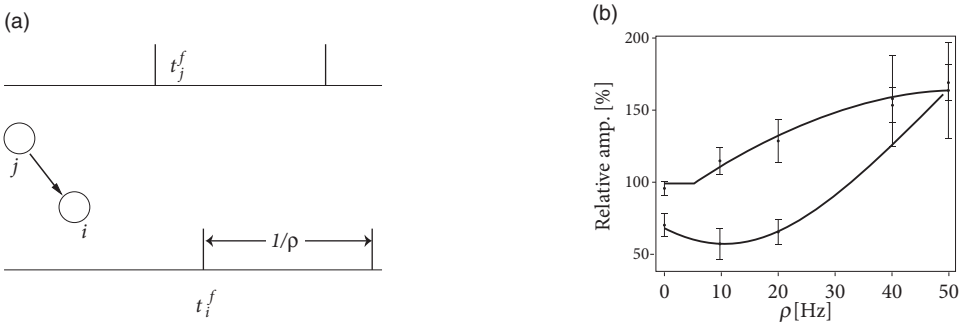

Fig. 19.8 Frequency dependence of STDP. (a) The experimental protocol depicted in Fig. 19.4 was repeated for different frequency, ρ, of pre-post pairs. (b) The triplet rule reproduces the finding that increased frequency of pair repetition leads to increased potentiation in visual cortex pyramidal neurons. Top curve $t_j^f - t_i^f = 10$ ms, bottom curve –10 ms. Data from Sjöström *et al*. (2001), figure adapted from Pfister *et al*. (2006).

postsynaptic firing rates and potentiation for high postsynaptic firing rates (Pfister and Gerstner, 2006); see Exercises. If we further assume that the triplet term in the learning rule depends on the mean postsynaptic frequency, a sliding threshold between potentiation and depression can be defined. In this way, the learning rule matches the requirements of the BCM theory and inherits the properties of the BCM learning rule such as the input selectivity (see Exercises). From the BCM properties, we can immediately conclude that the triplet model should be useful for receptive field development (Bienenstock *et al*., 1982).

Example: Plasticity model with voltage dependence

Spike timing dependence is only one of several manifestations of synaptic plasticity. Apart from spike timing, synaptic plasticity also depends on several other variables, in particular on postsynaptic voltage (Fig. 19.3). In this example, we present the voltage-dependent model of Clopath *et al*. (2010).

The Clopath model exhibits separate additive contributions to the plasticity rule, one LTD and another for LTP. For the LTD part, presynaptic spike arrival at a synapse from a presynaptic neuron j to a postsynaptic neuron i induces depression of the synaptic weight w_{ij} by an amount $-A_{\rm LTD}\,[\overline{u}_{i,-}(t) - \theta_-]_+$ that is proportional to the average postsynaptic depolarization $\overline{u}_{i,-}$. The brackets $[]_+$ indicate rectification, i.e., any value $\overline{u}_{i,-} < \theta_-$ does not lead to a change; see Artola *et al*. (1990) and Fig. 19.3. The quantity $\overline{u}_{i,-}(t)$ is an low-pass filtered version of the postsynaptic membrane potential $u(t)$ with a time constant τ_-:

$$\tau_- \frac{\mathrm{d}}{\mathrm{d}t}\overline{u}_{i,-}(t) = -\overline{u}_{i,-}(t) + u_i(t).$$

Introducing the presynaptic spike train $S_j(t) = \sum_f \delta(t - t_j^f)$, the update rule for

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

depression is (Fig. 19.9)

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij}^{\mathrm{LTD}} = -A_{\mathrm{LTD}}(\bar{u}_i)\, S_j(t)\, [\overline{u}_{i,-}(t) - \theta_-]_+ \qquad \text{if } w_{ij} > w_{\min}, \tag{19.16}$$

where $A_{\mathrm{LTD}}(\bar{u}_i)$ is an amplitude parameter that depends on the mean depolarization $\bar{u}$ of the postsynaptic neuron, averaged over a time scale of 1 second. A choice $A_{\mathrm{LTD}}(\bar{u}_i) = \alpha \frac{\bar{u}_i^2}{u_{\mathrm{ref}}^2}$ where u_{ref}^2 is a reference value, is a simple method to avoid a run-away of the rate of the postsynaptic neuron, analogous to the sliding threshold in the BCM rule of Eq. (19.9). A comparison with the triplet rule above shows that the role of the trace y_i (which represents a low-pass filter of the postsynaptic spike train, see Eq. (19.13)) is taken over by the low-pass filter $\overline{u}_{i,-}$ of the postsynaptic voltage.

For the LTP part, we assume that each presynaptic spike at the synapse w_{ij} increases the trace $\bar{x}_j(t)$ of some biophysical quantity, which decays exponentially with a time constant τ_+ in the absence of presynaptic spikes; see Eq. (19.12). The potentiation of w_{ij} depends on the trace $\bar{x}_j(t)$ and the postsynaptic voltage via (see also Fig. 19.9)

$$\frac{\mathrm{d}}{\mathrm{d}t} w_i^{\mathrm{LTP}} = +A_{\mathrm{LTP}}\, \bar{x}_j(t)\, [u_i(t) - \theta_+]_+\, [\overline{u}_{i,+}(t) - \theta_-]_+ \quad \text{if } w_{ij} < w_{\max}. \tag{19.17}$$

Here, $A_{\mathrm{LTP}} > 0$ is a constant parameter and $\overline{u}_{i,+}(t)$ is another low-pass filtered version of $u_i(t)$ similar to $\overline{u}_-(t)$ but with a shorter time constant τ_+ around 10 ms. Thus positive weight changes can occur if the momentary voltage $u_i(t)$ surpasses a threshold θ_+ and, at the same time the average value $\overline{u}_{i,+}(t)$, is above θ_-. Note again the similarity to the triplet STDP rule. If the postsynaptic voltage is dominated by spikes, so that $u_i(t) = \sum_f \delta(t - t_i^f)$, the Clopath model and the triple STDP rule are in fact equivalent.

The Clopath rule is summarized by the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = -A_{\mathrm{LTD}}(\bar{u})\, S_j(t)\, [\overline{u}_{i,-}(t) - \theta_-]_+ + A_{\mathrm{LTP}}\, \bar{x}_i(t)\, [u_i(t) - \theta_+]_+\, [\overline{u}_{i,+}(t) - \theta_-]_+ \tag{19.18}$$

combined with hard bounds $0 \le w_{ij} \le w_{\max}$.

The plasticity rule can be fitted to experimental data and can reproduce several experimental paradigms (Sjöström *et al.*, 2001) that cannot be explained by pair-based STDP or other phenomenological STDP rules without voltage dependence.

19.3 Unsupervised learning

In artificial neural networks some, or even all, neurons receive input from external sources as well as from other neurons in the network. Inputs from external sources are typically described as a statistical ensemble of potential stimuli. Unsupervised learning in the field of artificial neural networks refers to changes of synaptic connections which are driven by the statistics of the input stimuli – in contrast to supervised learning or reward-based learning where the network parameters are optimized to achieve, for each stimulus, an optimal behavior. Hebbian learning rules, as introduced in the previous section, are the prime example of unsupervised learning in artificial neural networks.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

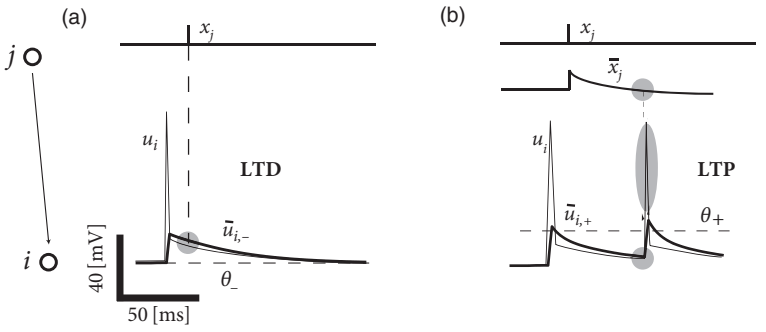

Fig. 19.9 Clopath model of voltage-dependent plasticity. Synaptic weights react to presynaptic events (top) and postsynaptic membrane potential (bottom) (a) The synaptic weight is decreased if a presynaptic spike x_j (dashed vertical line) arrives when the low-pass filtered value $\overline{u}_{i,-}$ (thick black line) of the membrane potential is above θ_- (dashed horizontal line). (b) The synaptic weight is increased if three criteria are met (shaded ellipses): (i) the membrane potential u_i (thin black line) is above a threshold θ^+ (horizontal dashed line); (ii) the low-pass filtered value of the membrane potential $\overline{u}_{i,+}$ (thick black line) is higher than a threshold θ^-; and (iii) the presynaptic low-pass filter $\bar{x}$ is above zero; adapted from Clopath *et al.* (2010).

In the following we always assume that there are N input neurons $1 \leq j \leq N$. Their firing rates ν_j are chosen from a set of P firing rate patterns with index $1 \leq \mu \leq P$. While one of the patterns, say pattern μ with $\xi^\mu = (\xi_1^\mu, \ldots, \xi_N^\mu)$, is presented to the network, the firing rates of the input neurons are $\nu_j = \xi_j^\mu$. In other words, the input rates form a vector $\nu = \xi^\mu$ where $\nu = (\nu_1, \ldots, \nu_N)$. After a time Δt a new input pattern is randomly chosen from the set of available patterns. We call this the static pattern scenario.

19.3.1 Competitive learning

In the framework of Eq. (19.2), we can define a Hebbian learning rule of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = \gamma \nu_i [\nu_j - \nu_\theta(w_{ij})], \tag{19.19}$$

where γ is a positive constant and ν_θ is some reference value that may depend on the current value of w_{ij}. A weight change occurs only if the postsynaptic neuron is active, $\nu_i > 0$. The direction of the weight change depends on the sign of the expression in the rectangular brackets.

Let us suppose that the postsynaptic neuron i is driven by a subgroup of highly active presynaptic neurons ($\nu_i > 0$ and $\nu_j > \nu_\theta$). Synapses from one of the highly active presynaptic neurons onto neuron i are strengthened while the efficacy of other synapses that have not been activated is decreased. Firing of the postsynaptic neuron thus leads to LTP at the active pathway ("homosynaptic LTP") and at the same time to LTD at the inactive synapses ("heterosynaptic LTD"); for reviews see Brown *et al.* (1991) and Bi and Poo (2001).

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

A particularly interesting case from a theoretical point of view is the choice $\nu_\theta(w_{ij}) = w_{ij}$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = \nu_i [\nu_j - w_{ij}] . \tag{19.20}$$

The synaptic weights thus move toward the fixed point $w_{ij} = \nu_j$ whenever the postsynaptic neuron is active. In the stationary state, the set of weight values w_{ij} reflects the presynaptic firing pattern $\nu_j, 1 \le j \le N$. In other words, the presynaptic firing pattern is *stored* in the weights.

The above learning rule is an important ingredient of competitive unsupervised learning (Kohonen, 1984; Grossberg, 1976). To implement competitive learning, an array of K (postsynaptic) neurons receive input from the same set of N presynaptic neurons which serve as the input layer. The postsynaptic neurons inhibit each other via strong lateral connections, so that, whenever a stimulus is applied at the input layer, the K postsynaptic neurons compete with each other and only a single postsynaptic neuron responds. The dynamics in such competitive networks where only a single neuron "wins" the competition have already been discussed in Chapter 16.

In a learning paradigm with the static pattern scenario, all postsynaptic neurons use the same learning rule (19.20), but only the active neuron i' (i.e., the one which "wins" the competition) will effectively update its weights (all others have zero update because $\nu_i = 0$ for $i \neq i'$). The net result is that the weight vector $w_{i'} = (w_{i'1} \dots w_{i'N})$ of the winning neuron i' moves closer to the current vector of inputs $\nu = \xi^\mu$. For a different input pattern μ' the same or another postsynaptic neuron may win the competition. Therefore, different neurons specialize for different subgroups ("clusters") of patterns and each neuron develops a weight vector which represents the center of mass of "its" cluster.

Example: Developmental learning with STDP

The results of simulations of the Clopath model shown in Fig. 19.10 can be interpreted as a realization of a soft form of competitive learning. Neurons form subnetworks that specialize on the same features of the input. Because of inhibition, different subnetworks specialize on different segments of the input space.

Ten excitatory neurons (with all-to-all connectivity) are linked to three inhibitory neurons. Each inhibitory neuron receives input from eight randomly selected excitatory neurons and randomly projects back to six excitatory neurons (Clopath *et al.*, 2010). In addition to the recurrent input, each excitatory and inhibitory neuron receives feedforward spike input from 500 presynaptic neurons j that generate stochastic Poisson input at a rate ν_j. The input neurons can be interpreted as a sensory array. The rates of neighboring input neurons are correlated, mimicking the presence of a spatially extended object stimulating the sensory layer. Spiking rates at the sensory layer change every 100 ms.

Feedforward connections and lateral connections between model pyramidal neurons are plastic whereas connections to and from inhibitory neurons are fixed. In a simulation

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

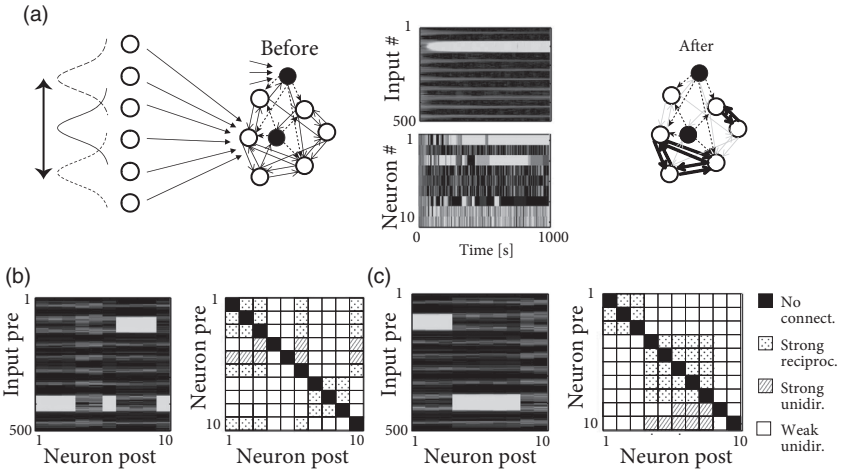

Fig. 19.10 Receptive field development and lateral connectivity. (a) A network of ten excitatory neurons (empty circles, not all neurons are shown) is connected to three inhibitory neurons (solid circles) and receives feedforward inputs from 500 Poisson spike trains with a Gaussian profile of firing rates. The center of the Gaussian is shifted randomly every 100 ms. Between the two schematic figures representing the network before (left) and after the plasticity experiment (right), we depict the evolution of input weights and recurrent excitatory weights onto one selected excitatory neuron. (b) Mean feedforward weights (left) and recurrent excitatory weights (right) averaged over 100 s. The gray level graph for the feedforward weights (left) indicates that neurons develop receptive fields that are localized in the input space. The diagonal in the matrix of recurrent connectivity is black, since self-connections do not exist in the model. (c) Same as (b) but for the sake of visual clarity the index of neurons is reordered so that neurons with similar receptive fields have adjacent numbers, highlighting that neurons with similar receptive fields (e.g., neurons 1 to 3) have strong bilateral connections. Adapted from Clopath *et al*. (2010).

of the model network, the excitatory neurons developed localized receptive fields, i.e., weights from *neighboring* inputs to the same postsynaptic neuron become either strong or weak *together* (Fig. 19.10a). Similarly, lateral connections onto the same postsynaptic neuron develop strong or weak synapses, that remain, apart from fluctuations, stable thereafter (Fig. 19.10a) leading to a structured pattern of synaptic connections (Fig. 19.10b). While the labeling of the excitatory neurons at the beginning of the experiment was randomly assigned, we can relabel the neurons after the formation of lateral connectivity patterns so that neurons with similar receptive fields have similar indices. After reordering we can clearly distinguish that two groups of neurons have been formed, characterized by similar receptive fields and strong bidirectional connectivity within the group, and different receptive fields and no lateral connectivity between groups (Fig. 19.10c).

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

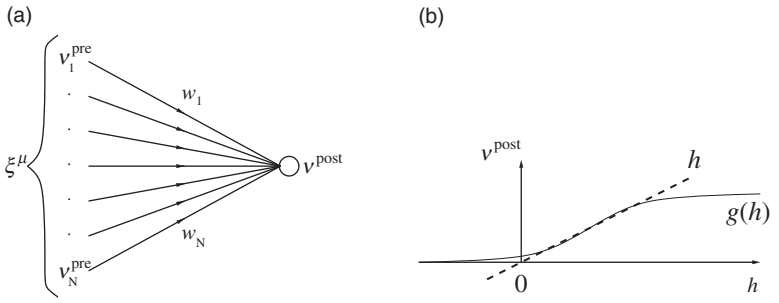

Fig. 19.11 Elementary model. (a) Patterns ξ^μ are applied as a set of presynaptic firing rates ν_j, i.e., $\xi_j^\mu = \nu_j^{\text{pre}}$ for $1 \le j \le N$. (b) The gain function of the postsynaptic neuron is taken as linear, i.e., $\nu^{\text{post}} = h$. It can be seen as a linearization of the sigmoidal gain function $g(h)$.

19.3.2 Learning equations for rate models

We focus on a *single* analog neuron that receives input from N presynaptic neurons with firing rates ν_j^{pre} via synapses with weights w_j; see Fig. 19.11a. Note that we have dropped the index i of the postsynaptic neuron since we focus in this section on a single output neuron. We think of the presynaptic neurons as "input neurons," which, do not, however, 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. As before, the firing rate of the input neurons is modeled by the static pattern scenario. We will show that the statistical properties of the input control the evolution of synaptic weights. In particular, we identify the conditions under which unsupervised Hebbian learning is related to principal component analysis (PCA).

In the following we analyze the evolution of synaptic weights using the simple Hebbian learning rule of Eq. (19.3). The presynaptic activity drives the postsynaptic neuron and the joint activity of pre- and postsynaptic neurons triggers changes of the synaptic weights:

$$\Delta w_i = \gamma \nu^{\text{post}} \nu_i^{\text{pre}} . \tag{19.21}$$

Here, $0 < \gamma \ll 1$ is a small constant called the "learning rate." The learning rate in the static pattern scenario is closely linked to the correlation coefficient c_{11}^{corr} in the continuous-time Hebb rule introduced in Eq. (19.3). In order to highlight the relation, let us assume that each pattern ξ^μ is applied during an interval Δt. For Δt sufficiently small, we have $\gamma = c_{11}^{\text{corr}} \Delta t$.

In a general rate model, the firing rate ν^{post} of the postsynaptic neuron is given by a nonlinear function of the total input

$$\nu^{\text{post}} = g\left(\sum_j w_j \nu_j^{\text{pre}} \right) ; \tag{19.22}$$

see Fig. 19.11b and Chapter 15. For the sake of simplicity, we restrict our discussion in the

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

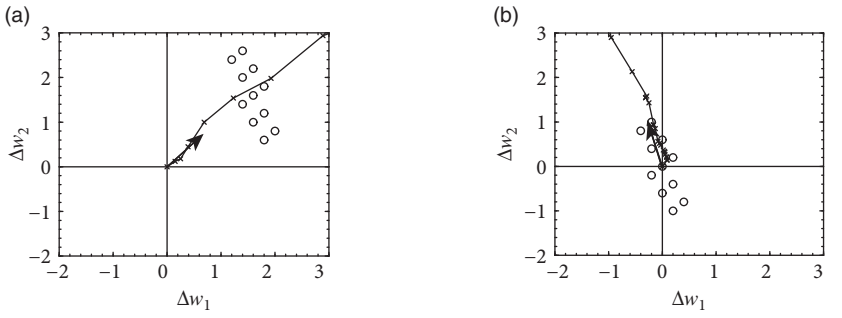

Fig. 19.12 Weight changes induced by the standard Hebb rule. Input patterns $\xi^\mu \in \mathbb{R}^2$ are marked as circles. The sequence of weight vectors $w(1)$, $w(2)$, ... is indicated by crosses connected by a solid line. (a) The weight vector evolves in the direction of the center of mass of the cloud of data points, because this is the dominant eigenvector of the (non-normalized) correlation matrix of the data. (b) If the input patterns are normalized so that their center of mass is at the origin, then the weight vector becomes parallel to the first principal component e_1 of the dataset.

following to a *linear* rate model with

$$\nu^{\text{post}} = \sum_j w_j \nu_j^{\text{pre}} = w \cdot \nu^{\text{pre}} , \tag{19.23}$$

where we have introduced vector notation for the weights $w = (w_1, \ldots, w_N)$, the presynaptic rates $\nu^{\text{pre}} = (\nu_1, \ldots, \nu_N)$ and the dot denotes a scalar product. Hence we can interpret the output rate ν^{post} as a projection of the input vector onto the weight vector.

If we combine the learning rule (19.21) with the linear rate model of Eq. (19.23) we find after the presentation of pattern ξ^μ

$$\Delta w_i = \gamma \sum_j w_j \nu_j^{\text{pre}} \nu_i^{\text{pre}} = \gamma \sum_j w_j \xi_j^\mu \xi_i^\mu . \tag{19.24}$$

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

$$w_i(n+1) = w_i(n) + \gamma \sum_j w_j \xi_j^{\mu_n} \xi_i^{\mu_n} , \tag{19.25}$$

where μ_n denotes the pattern that is presented during the nth time step. Eq. (19.25) is called an "online" rule, because the weight update happens immediately after the presentation of each pattern. The evolution of the weight vector $w = (w_1, \ldots, w_N)$ during the presentation of several patterns is shown in Fig. 19.12.

If the learning rate γ is small, a large number of patterns has to be presented in order to induce a substantial weight change. In this case, there are two equivalent routes to proceed with the analysis. The first one is to study a version of learning where all P patterns are

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

presented before an update occurs. Thus, Eq. (19.25) is replaced by

$$w_i(n+1) = w_i(n) + \tilde{\gamma} \sum_j w_j \sum_{\mu=1}^{P} \xi_j^\mu \, \xi_i^\mu \tag{19.26}$$

with a new learning rate $\tilde{\gamma} = \gamma/P$. This is called a "batch update." With the batch update rule, the right-hand side can be rewritten as

$$w_i(n+1) = w_i(n) + \gamma \sum_j C_{ij} \, w_j(n) \,, \tag{19.27}$$

where we have introduced the correlation matrix

$$C_{ij} = \frac{1}{P} \sum_{\mu=1}^{P} \xi_i^\mu \, \xi_j^\mu = \langle \xi_i^\mu \, \xi_j^\mu \rangle_\mu \,. \tag{19.28}$$

Thus, the evolution of the weights is driven by the correlations in the input.

The second, alternative, route is to stick to the online update rule, but study the *expectation* value of the weight vector, i.e., the weight vector $\langle w(n) \rangle$ averaged over the sequence $(\xi^{\mu_1}, \xi^{\mu_2}, \ldots, \xi^{\mu_n})$ of all patterns that so far have been presented to the network. From Eq. (19.25) we find

$$\begin{aligned} \langle w_i(n+1) \rangle &= \langle w_i(n) \rangle + \gamma \sum_j \left\langle w_j(n) \, \xi_j^{\mu_{n+1}} \, \xi_i^{\mu_{n+1}} \right\rangle \\ &= \langle w_i(n) \rangle + \gamma \sum_j \langle w_j(n) \rangle \left\langle \xi_j^{\mu_{n+1}} \, \xi_i^{\mu_{n+1}} \right\rangle \\ &= \langle w_i(n) \rangle + \gamma \sum_j C_{ij} \, \langle w_j(n) \rangle \,. \end{aligned} \tag{19.29}$$

The angle brackets denote an ensemble average over the whole sequence of input patterns $(\xi^{\mu_1}, \xi^{\mu_2}, \ldots)$. 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. Note that Eq. (19.29) for the *expected* weights in the online rule is equivalent to Eq. (19.27) for the weights in the *batch* rule.

Expression (19.29), or equivalently (19.27), can be written in a more compact form using matrix notation (we drop the angle brackets in the following)

$$w(n+1) = (\mathbb{1} + \gamma C) \, w(n) = (\mathbb{1} + \gamma C)^{n+1} \, w(0) \,, \tag{19.30}$$

where $w(n) = (w_1(n), \ldots, w_N(n))$ is the weight vector and $\mathbb{1}$ is the identity matrix.

If we express the weight vector in terms of the eigenvectors e_k of C,

$$w(n) = \sum_k a_k(n) \, e_k \,, \tag{19.31}$$

we obtain an explicit expression for $w(n)$ for any given initial condition $a_k(0)$, namely,

$$w(n) = \sum_k (1 + \lambda_k)^n \, a_k(0) \, e_k \,. \tag{19.32}$$

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

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*,

$$w(n) \xrightarrow{n\to\infty} (1+\lambda_1)^n a_1(0)\, e_1 \,. \tag{19.33}$$

Recall that the output of the linear neuron model (19.23) is proportional to the projection of the current input pattern ξ^μ on the direction w. For $w \propto e_1$, the output is therefore proportional to the projection on the first principal component of the input distribution. A Hebbian learning rule such as (19.21) is thus able to extract the first principal component of the input data.

From a data-processing point of view, the extraction of the first principal component of the input dataset by a biologically inspired learning rule seems to be very compelling. There are, however, a few drawbacks and pitfalls when using the above simple Hebbian learning scheme. Interestingly, all three can be overcome by slight modifications in the Hebb rule.

First, the above statement about the Hebbian learning rule is limited to the *expectation value* of the weight vector. However, it can be shown that if the learning rate is sufficiently low, then the actual weight vector is very close to the expected one so that this is not a major limitation.

Second, 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, variants of Hebbian learning such as the Oja learning rule (19.7) allow us to normalize the length of the weight vector without changing its direction; see Exercises.

Third, 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 can, however, be overcome by learning rules such as the covariance rule of Eq. (19.6) that are based on the deviation of the rates from a certain mean firing rate. Similarly, STDP rules can be designed in such a way that the output rate remains normalized so that learning is sensitive only to deviations from the mean firing rate and can thus find the first principal component even if the input is not properly normalized (Kempter *et al*., 1999a; Song *et al*., 2000; Kempter *et al*., 2001).

Example: Correlation matrix and principal component analysis

For readers not familiar with principal component analysis (PCA) we review here the basic ideas and main results. PCA is a standard technique to describe statistical properties of a set of high-dimensional data points and is performed to find the direction in which the data shows the largest variance. If we think of the input dataset as a cloud of points in a high-dimensional vector space centered around the origin, then the first principal component is the direction of the longest axis of the ellipsoid that encompasses the

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

cloud; see Fig. 19.13a. In what follows, we will explain the basic idea and show that the first principal component gives the direction where the variance of the data is maximal.

Let us consider an ensemble of data points $\{\xi^1, \ldots, \xi^P\}$ drawn from a (high-dimensional) vector space, for example $\xi^\mu \in \mathbb{R}^N$. For this set of data points we define the *correlation matrix* C_{ij} as

$$C_{ij} = \frac{1}{P} \sum_{\mu=1}^{P} \xi_i^\mu \, \xi_j^\mu = \left\langle \xi_i^\mu \, \xi_j^\mu \right\rangle_\mu \,. \tag{19.34}$$

Angle brackets $\langle \cdot \rangle_\mu$ denote an average over the whole set of data points. In a similar way to the variance of a single random variable we can also define the *covariance matrix* V_{ij} of our dataset,

$$V_{ij} = \left\langle (\xi_i^\mu - \langle \xi_i^\mu \rangle_\mu)(\xi_j^\mu - \langle \xi_j^\mu \rangle_\mu) \right\rangle_\mu \,. \tag{19.35}$$

Here we will assume that the coordinate system is chosen so that the center of mass of the set of data points is located at the origin, i.e., $\langle \xi_i \rangle_\mu = \langle \xi_j \rangle_\mu = 0$. In this case, correlation matrix and covariance matrix are identical.

The *principal components* of the set $\{\xi^1, \ldots, \xi^P\}$ are defined as the eigenvectors of the covariance matrix V. Note that V is symmetric, i.e., $V_{ij} = V_{ji}$. The eigenvalues of V are thus real-valued and different eigenvectors are orthogonal (Horn and Johnson, 1985). Furthermore, V is positive semi-definite since

$$y^{\mathrm{T}} V \, y = \sum_{ij} y_i \left\langle \xi_i^\mu \, \xi_j^\mu \right\rangle_\mu y_j = \left\langle \left[\sum_i y_i \, \xi_i^\mu \right]^2 \right\rangle_\mu \geq 0 \tag{19.36}$$

for any vector $y \in \mathbb{R}^N$. Therefore, all eigenvalues of V are non-negative.

We can sort the eigenvectors e_i according to the size of the corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$. The eigenvector with the largest eigenvalue is called the first principal component.

The first principal component points in the direction where the variance of the data is maximal. To see this we calculate the variance of the projection of ξ^μ onto an arbitrary direction y (Fig. 19.13b) that we write as $y = \sum_i a_i \, e_i$ with $\sum_i a_i^2 = 1$ so that $\|y\| = 1$. The variance σ_y^2 along y is

$$\sigma_y^2 = \left\langle \left[\xi^\mu \cdot y \right]^2 \right\rangle_\mu = y^{\mathrm{T}} V \, y = \sum_i \lambda_i \, a_i^2 \,. \tag{19.37}$$

The right-hand side is maximal under the constraint $\sum_i a_i^2 = 1$ if $a_1 = 1$ and $a_i = 0$ for $i = 2, 3, \ldots, N$, i.e., if $y = e_1$.

19.3.3 Learning equations for STDP models (*)

The evolution of synaptic weights in the pair-based STDP model of 19.10 can be assessed by assuming that pre- and postsynaptic spike trains can be described by Poisson processes.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

Fig. 19.13 Principal component analysis. (a) Ellipsoid approximating the shape of a cloud of data points. The first principal component e_1 corresponds to the principal axis of the ellipsoid. (b) Sample distribution of data points in two dimensions. The first principal component e_1 points in the direction where the variance of the data is maximal. Projection (dashed line) of data points onto an arbitrary other axis y gives a distribution of smaller variance.

For the postsynaptic neuron, we take the linear Poisson model in which the output spike train is generated by an inhomogeneous Poisson process with rate

$$\nu_i(u_i) = [\alpha\, u_i - \nu_0]_+ \tag{19.38}$$

with scaling factor α, threshold ν_0 and membrane potential $u_i(t) = \sum_j w_{ij} \varepsilon \left(t - t_j^f\right)$, where $\varepsilon(t)$ denotes the time course of an excitatory postsynaptic potential generated by a presynaptic spike arrival. The notation $[x]_+$ denotes a piecewise linear function: $[x]_+ = x$ for $x > 0$ and zero otherwise. In the following we assume that the argument of our piecewise linear function is positive so that we can suppress the square brackets.

For the sake of simplicity, we assume that all input spike trains are Poisson processes with a *constant* firing rate ν_j. In this case the *expected* firing rate of the postsynaptic neuron is simply:

$$\langle \nu_i \rangle = -\nu_0 + \alpha \bar{\varepsilon} \sum_j w_{ij}\, \nu_j\,, \tag{19.39}$$

where $\bar{\varepsilon} = \int \varepsilon(s) \mathrm{d}s$ is the total area under an excitatory postsynaptic potential. The *conditional* rate of firing of the postsynaptic neuron, given an input spike at time t_j^f, is given by

$$\nu_i(t) = -\nu_0 + \alpha \bar{\varepsilon} \sum_j w_{ij} \nu_j + \alpha w_{ij} \varepsilon \left(t - t_j^f\right)\,. \tag{19.40}$$

Since the conditional rate is different from the expected rate, the postsynaptic spike train $S_i(t)$ is correlated with the presynaptic spike trains $S_j(t')$. The correlations can be calculated to be

$$\Gamma_{ji}(s) = \langle S_i(t+s) S_j(t) \rangle = \langle \nu_i \rangle \nu_j + \alpha\, w_{ij} \nu_j\, \varepsilon(s)\,. \tag{19.41}$$

In a similar way to the expected weight evolution for rate models in Section 19.3.2, we now

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

study the expected weight evolution in the spiking model with the pair-based plasticity rule of Eq. (19.10). The result is (Kempter *et al.*, 1999a)

$$\langle \dot{w}_{ij} \rangle = \nu_j \langle \nu_i \rangle [-A_-(w_{ij})\tau_- + A_+(w_{ij})\tau_+] + \alpha w_{ij} \nu_j A_+(w_{ij}) \int W_+(s)\varepsilon(s)\mathrm{d}s\,. \quad (19.42)$$

The first term is reminiscent of rate-based Hebbian learning, Eq. (19.3), with a coefficient c_{11}^{corr} proportional to the integral under the learning window. The last term is due to pre-before-post spike timings that are absent in a pure rate model. Hence, despite the fact that we started off with a pair-based STDP rule, the synaptic dynamics contains a term of the form $\alpha \nu_j w A_+(w) \int W_+(s)\varepsilon(s)\mathrm{d}s$ that is *linear* in the presynaptic firing rate (Kempter *et al.*, 1999a, 2001).

Example: Stabilization of postsynaptic firing rate

If spike arrival rates $\nu_j = \nu$ at all synapses are identical, we expect a solution of the learning equation to exist where all weights are identical, $w_{ij} = w$. For simplicity we drop the averaging signs. Eq. (19.42) then becomes

$$\frac{\dot{w}}{\nu} = \nu_i [-A_-(w)\tau_- + A_+(w)\tau_+] + \alpha w A_+(w) \int W_+(s)\varepsilon(s)\mathrm{d}s\,. \quad (19.43)$$

Moreover, we can use Eq. (19.39) to express the postsynaptic firing rate in terms of the input rate ν:

$$\nu_i = -\nu_0 + \alpha \nu \bar{\varepsilon} N w\,. \quad (19.44)$$

If the weight w increases, the postsynaptic firing rate also increases. We now ask whether the postsynaptic firing has a fixed point ν_{FP}.

The fixed point analysis can be performed for a broad class of STDP models. However, for the sake of simplicity we focus on the model with hard bounds in the range where $0 < w < w^{\max}$. We introduce a constant $C = A_-(w)\tau_- - A_+(w)\tau_+$. If the integral of the learning window is negative, then $C > 0$ and LTD dominates over LTP. In this case, a fixed point exists at

$$\nu_{\mathrm{FP}} = \frac{C_{\mathrm{ss}}\nu_0}{NC\nu\bar{\varepsilon} - C_{\mathrm{ss}}} \quad (19.45)$$

where $C_{\mathrm{ss}} = A_+(w)\int W_+(s)\varepsilon(s)\mathrm{d}s$ denotes the contribution of the spike–spike correlations. The mean firing rate of the neuron is, under rather general conditions, stabilized at the fixed point (Kempter *et al.*, 2001). Hence STDP can lead to a control of the postsynaptic firing rate (Kempter *et al.*, 1999a; Song *et al.*, 2000; Kempter *et al.*, 2001). We emphasize that the existence of a fixed point and its stability does not crucially depend on the presence of soft or hard bounds on the weights. Since, for constant input rates ν, we have $\nu_i = \nu_0 + \alpha\nu\bar{\varepsilon}\sum_j w_{ij}$, stabilization of the output rate automatically implies normalization of the summed weights.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

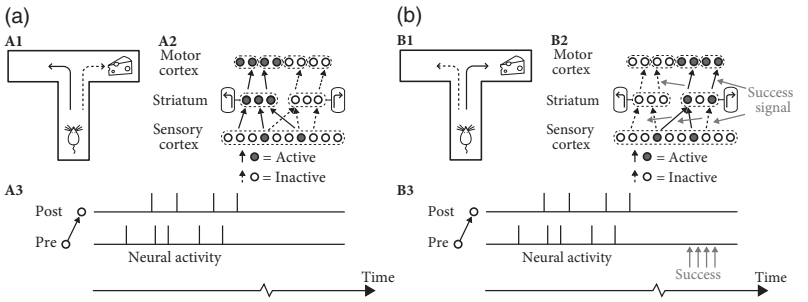

Fig. 19.14 Hebbian learning versus reward-modulated Hebbian learning in a T-maze decision task (schematic). The momentary sensory state of the animal is encoded by the layer of sensory cortex neurons, the action plan by a set of striatal neurons, and the motor output by neurons in motor cortex. (a) Hebbian learning. Joint activity of neurons in sensory, striatal, and motor areas strengthens the links between the active neurons, despite the fact that turning left does not lead to success. (b) Reward-modulated Hebbian learning. Joint activity of pre- and postsynaptic neurons strengthens connections only if, within a delay of a few seconds, a success signal is broadcast in the brain.

19.4 Reward-based learning

In conditioning experiments, animals learn complex action sequences if the desired behavior is rewarded. For example, in a simple T-maze an animal has to decide at the bifurcation point whether to turn left or right (Fig. 19.14a). In each of several trials, the same arm of the maze is baited with a piece of cheese that is hidden in a hole in the floor and therefore neither visible nor smellable. After a few trials the animal has learned to reliably turn into the baited arm of the maze.

Unsupervised Hebbian learning is of limited use for behavioral learning, because it makes no distinction between actions that do and those that do not lead to a successful outcome. The momentary sensory state at the bifurcation point is represented by activity in the sensory cortices and, possibly, in hippocampal place cells. The action plan "turn left" is represented by groups of cells in several brain areas likely to include the striatum, whereas the final control of muscle activity involves areas in the motor cortex. Therefore, during the realization of the action plan "turn left," several groups of neurons are jointly active (Fig. 19.14a). Unsupervised Hebbian learning strengthens the connections between the jointly active cells so that, at the next trial, it becomes more likely that the animal takes the same decision again. However, turning left does not lead to success if the cheese is hidden in the other branch of the maze.

In order to solve the above task, two important aspects have to be taken into account that are neglected in unsupervised Hebbian learning rules. First, rules of synaptic plasticity have to take into account the success of an action. Neuromodulators such a dopamine are ideal candidates to broadcast a success signal in the brain (Fig. 19.15), where success can loosely be defined as "reward minus expected reward" (Schultz *et al.*, 1997; Schultz, 2007,

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

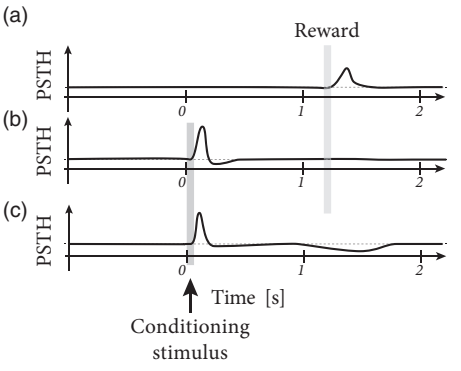

Fig. 19.15 Dopamine encodes reward minus expected reward. (a) The activity of dopaminergic neurons (PSTH) increases at the moment when a reward R occurs. (b) If a conditioning stimulus (CS) such as a tone beep or a light reliably occurs one second before the reward, the same neurons no longer respond to the reward but instead respond after the CS that predicts the reward. (c) If the reward is predicted by the CS but not given, the PSTH exhibits a dip below the baseline (dashed lines) at the moment when the reward is expected. Schematic figure summarizing data from Schultz *et al*. (1997).

2010). Second, the success often comes with a delay of a few seconds after an action has been taken; see Fig. 19.14b. Thus, the brain needs somehow to store a short-term memory of past actions. A suitable location for such a memory is the synapses themselves.

The above two points can be used to formulate a first qualitative model of reward-modulated Hebbian learning. In Hebbian learning, weight changes Δw_{ij} depend on the spikes pre_j of the presynaptic neuron j and the state $post_i$ of the postsynaptic neuron i. Correlations between the pre- and postsynaptic activity are picked up by the Hebbian function $H(pre_j, post_i)$. We assume that synapses keep track of correlations by updating a synaptic eligibility trace

$$\tau_e \frac{\mathrm{d}}{\mathrm{d}t} e_{ij} = -e_{ij} + H(pre_j, post_i)\,. \tag{19.46}$$

If the joint activity of pre- and postsynaptic neurons stops, the Hebbian term H vanishes and the eligibility trace decays back to zero with a time constant τ_e. The Hebbian term H could be modeled by one of the rate models in the framework of Eq. (19.2) or an STDP model such as the one defined in 19.10.

The update of synaptic weights requires a nonzero-eligibility trace as well as the presence of a neuromodulatory success signal M

$$\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = M \cdot e_{ij}\,. \tag{19.47}$$

While in standard Hebbian learning synaptic plasticity depends on two factors (i.e., pre- and postsynaptic activity), weight changes in Eq. (19.47) now depend on three factors, i.e., the two Hebbian factors and the neuromodulator M. The class of plasticity rules encompassed by Eq. (19.47) is therefore called three-factor learning rules. In models of reward-based learning, the modulator signal M is most often taken as "reward minus expected reward,"

$$M(t) = R(t) - \langle R \rangle\,, \tag{19.48}$$

where R denotes the reward and the expectation $\langle R \rangle$ is empirically estimated as a running

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

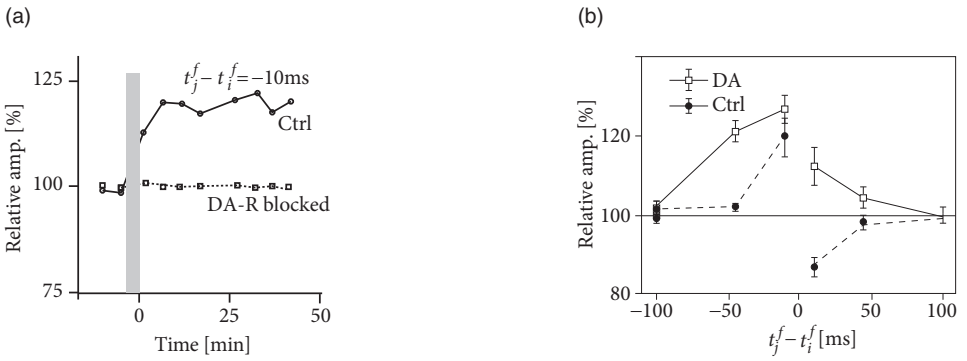

Fig. 19.16 Dopamine-modulated Hebbian learning. (a) An STDP protocol normally gives rise to long-term potentiation (pre-before-post, solid black line as in Fig. 19.4d). However, if dopamine receptors are blocked, no change occurs (schematic representation of experiments in Pawlak and Kerr (2008)). (b) The STDP window in a control situation (dashed line and solid data points) changes if additional extracellular dopamine is present (solid lines, open squares). Adapted from Zhang *et al*. (2009).

average. The time constant τ_e is typically chosen in the range of one second, so as to bridge the delay between action choice and final reward signal.

Three-factor rules have been suggested for rate-based (Reynolds and Wickens, 2002; Loewenstein, 2008) as well as spike-based Hebbian models. In spike-based Hebbian models, the Hebbian term is often taken as a standard pair-based STDP function (Izhikevich, 2007a; Legenstein *et al*., 2008; Florian, 2007) or an STDP model that also depends on postsynaptic voltage (Pfister *et al*., 2006; Florian, 2007; Baras and Meir, 2007). Experiments have shown that the shape of the STDP window is indeed modulated by dopamine as well as other neuromodulators (Fig. 19.16); for a review see Pawlak *et al*. (2010).

Example: R-STDP and learning of spike sequences

Suppose a table tennis player plays a serve or a piano player a rapid scale. In both cases the executed movements are extremely rapid, have been practiced many times, and are often performed in "open loop" mode, for example, without visual feedback during the movement. There is, however, feedback after some delay which signals the success (or failure) of the performed action, for example the ball went off the table or the scale contained a wrong note.

A rapid scale on a piano means touching about 10 different keys per second. Similarly, the complex gliding movement to give the ball its spin takes less than a second. It is likely that for such fast movements spike timing plays an important role. The motor cortex is involved in the control of limb movements. Experimental data from the arm area of the primary motor cortex indicates that populations of neurons encode the direction of hand

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

motion during reaching movements in three-dimensional space (Georgopoulos *et al.*, 1988). Each neuron i has a preferred direction of motion represented as a vector d_i. The vectors of different neurons are added up with a weighting function proportional to the cell's firing rate (Georgopoulos *et al.*, 1988). For the rapid movements of less than one second that we consider here, a single neuron is expected to emit at most a few spikes. Therefore a desired trajectory can be represented as a target spatio-temporal spike pattern; see Fig. 19.17.

Model neurons in the motor cortex receive spike input from neurons in sensory areas that represent, for example, the vertical movement of the ball that is launched at the beginning of the serve, as well as the intention of the player. During practice sessions, the aim is to associate the spatio-temporal spike pattern in the input layer with the target spike pattern in the layer of motor cortex neurons while the only feedback is the success signal available at the *end* of the movement.

Figure 19.17 shows that a two-layer network of spiking neurons can learn this task if synaptic connections use a reward-modulated STDP rule (R-STDP) where the Hebbian term H in Eq. (19.47) is the pair-based STDP rule defined in (19.10). It is important that the global neuromodulatory signal provided at the end of each trial is *not* the raw reward, but success defined as "reward – expected reward" as in Eq. (19.48). If a single task has to be learned, the expected reward can be estimated from the running average over past trials. However, if several trajectories (e.g., two different serves or two different scales) have to be learned in parallel, then the expected reward needs to be estimated separately for each trajectory (Fremaux *et al.*, 2010).

R-STDP rules have also been used for several other tasks, see, (e.g., Izhikevich, 2007a; Florian, 2007).

19.5 Summary

The Hebb rule (19.2) is an example of a *local* unsupervised learning rule. It is a local rule, because it depends only on pre- and postsynaptic firing rates and the present state w_{ij} of the synapse, i.e., information that is easily "available" at the location of the synapse. Experiments have shown that not only the firing rates, but also the membrane voltage of the postsynaptic neuron, as well as the relative timing of pre- and postsynaptic spikes, determine the amplitude and direction of change of the synaptic efficacy. To account for spike timing effects, classical pair-based models of STDP are formulated with a learning window that consists of two parts: If the presynaptic spike arrives before a postsynaptic output spike, the synaptic change is positive. If the timing is reversed, the synaptic change is negative. However, classical pair-based STDP models neglect the frequency and voltage dependence of synaptic plasticity, which are included in modern variants of STDP models.

The synaptic weight dynamics of Hebbian learning can be studied analytically if weights are changing slowly compared to the time scale of the neuronal activity. Weight changes are driven by correlations between pre- and postsynaptic activity. More specifically, simple

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

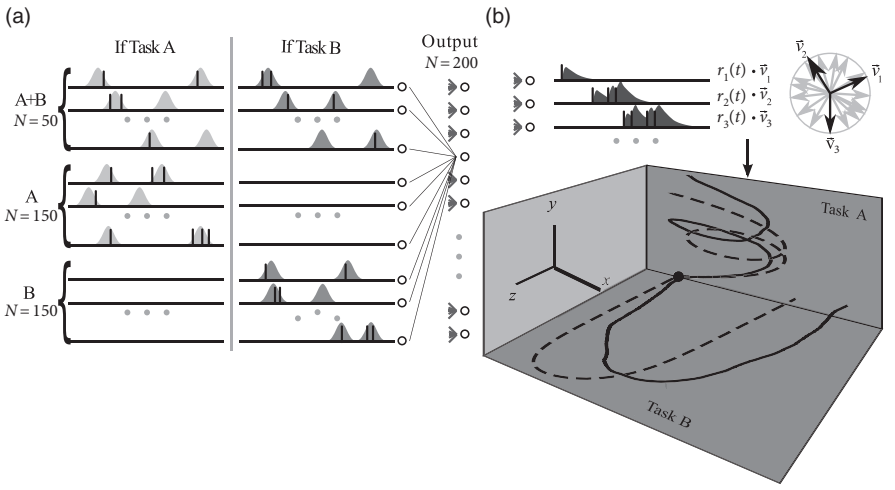

Fig. 19.17 Learning with R-STDP. (a) The input consists of 350 spike trains with a temporal precision of 20 ms. Fifty unspecific neurons fired for both tasks, whereas half of the other neurons fired only for task A or task B. The output consists of 200 spiking model neurons. (b) Output spike trains were convolved with a filter (top). The resulting continuous signal for neuron i is interpreted as the speed of movement in direction d_i, where different neurons code for different directions. The two target trajectories (bottom, dashed lines) correspond to two different target spatio-temporal spike-train patterns. After a learning period of 10 000 trials, the network output generates a trajectory (full black lines) close to the target trajectories.

Hebbian learning rules in combination with a linear neuron model find the first principal component of a normalized input dataset. Generalized Hebb rules, such as Oja's rule, keep the norm of the weight vector approximately constant during plasticity.

The interesting aspect of STDP is that it naturally accounts for *temporal* correlations by means of a learning window. Explicit expressions for temporal spike–spike correlations can be obtained for certain simple types of neuron model such as the linear Poisson model. Spike-based and rate-based rules of plasticity are equivalent as long as *temporal* spike–spike correlations are disregarded. If firing rates vary slowly, then the integral over the learning window plays the role of the Hebbian correlation term.

Hebbian learning and STDP are examples of unsupervised learning rules. Hebbian learning is considered to be a major principle of neuronal organization during development and a driving force for receptive field formation. However, Hebbian synaptic plasticity is not useful for behavioral learning, since it does not take into account the success (or failure) of an action. Three-factor learning rules combine the two Hebbian factors (i.e., pre- and postsynaptic activity) with a third factor (i.e., a neuromodulator such as dopamine) which conveys information about an action's success. Three-factor rules with an eligibility trace can be used to describe behavioral learning, in particular during conditioning experiments.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

Literature

Correlation-based learning can be traced back to Aristoteles[1] and has been discussed extensively by James (1890), who formulated a learning principle on the level of "brain processes" rather than neurons:

> When two elementary brain-processes have been active together or in immediate succession, one of them, on re-occurring, tends to propagate its excitement into the other.

A chapter of James' book is reprinted in volume 1 of Anderson and Rosenfeld's collection on neurocomputing (Anderson and Rosenfeld, 1988). The formulation of synaptic plasticity in Hebb's book (Hebb, 1949), of which two interesting sections are reprinted in the collection of Anderson and Rosenfeld (1988), has had a long-lasting impact on the neuroscience community. The historical context of Hebb's postulate is discussed in the reviews of Sejnowski (1999) and Makram *et al*. (2011).

There are several classical experimental studies on STDP (Markram *et al*., 1997; Zhang *et al*., 1998; Debanne *et al*., 1998; Bi and Poo, 1998, 1999; Sjöström *et al*., 2001), but precursors of timing-dependent plasticity can be found even earlier (Levy and Stewart, 1983). Note that for some synapses, the learning window is reversed (Bell *et al*., 1997). For reviews on STDP, see Abbott and Nelson (2000), Bi and Poo (2001), Caporale and Dan (2008), and Sjöström and Gerstner (2010).

The theory of unsupervised learning and principal component analysis is reviewed in the textbook by Hertz *et al*. (1991). Models of the development of receptive fields and cortical maps have a long tradition in the field of computational neuroscience (see, e.g., von der Malsburg, 1973; Willshaw and von der Malsburg, 1976; Sejnowski, 1977; Bienenstock *et al*., 1982; Kohonen, 1984; Linsker, 1986; Miller *et al*., 1989; MacKay and Miller, 1990; Miller, 1994; for reviews see, e.g., Erwin *et al*., 1995; Wiskott and Sejnowski, 1998). The essential aspects of the weight dynamics in linear networks are discussed in Oja (1982) and Miller and MacKay (1994). Articles by Grossberg (1976) and Bienenstock *et al*. (1982) or the book by Kohonen (1984) illustrate the early use of the rate-based learning rules in computational neuroscience.

The early theory of STDP was developed by Gerstner *et al*. (1993, 1996a); Kempter *et al*. (1999a); Roberts and Bell (2000); van Rossum *et al*. (2000); Song *et al*. (2000); Rubin *et al*. (2001) but precursors of timing-dependent plasticity can be found in earlier rate-based formulations (Herz *et al*., 1988; Sompolinsky and Kanter, 1986). Modern theories of STDP go beyond the pair-based rules (Senn *et al*., 2001; Pfister and Gerstner, 2006), consider voltage effects (Clopath *et al*., 2010), variations of boundary conditions (Gütig *et al*., 2003) or calcium-based models (Lisman *et al*., 2002; Lisman, 2003) and for reviews see Morrison *et al*. (2008) and Sjöström and Gerstner (2010).

Experimental support for three-factor learning rules is reviewed in Reynolds and Wickens (2002) and Pawlak *et al*. (2010). Model studies to reward modulated STDP are

[1] Aristoteles, "De memoria et reminiscentia": There is no need to consider how we remember what is distant, but only what is neighboring, for clearly the method is the same. For the changes follow each other by habit, one after another. And thus, whenever someone wishes to recollect he will do the following: He will seek to get a starting point for a change after which will be the change in question.

Izhikevich (2007a); Legenstein *et al*. (2008); Florian (2007); Fremaux *et al*. (2010). The consequences for behavior are discussed in Loewenstein and Seung (2006) and Loewenstein (2008). The classic reference for dopamine in relation to reward-based learning is Schultz *et al*. (1997). Modern reviews on the topic are Schultz (2007, 2010).

Exercises

1. **Normalization of firing rate.** *Consider a learning rule* $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = \gamma(\nu_i - \nu_\theta)\,\nu_j$, *i.e., a change of synaptic weights can only occur if the presynaptic neuron is active* ($\nu_j > 0$). *The direction of the change is determined by the activity of the postsynaptic neuron. The postsynaptic firing rate is given by* $\nu_i = g(\sum_{j=1}^{N} w_{ij}\nu_j)$. *We assume that presynaptic firing rates* ν_j *are constant.*
   (a) Show that ν_i *has a fixed point at* $\nu_i = \nu_\theta$.
   (b) Discuss the stability of the fixed point. Consider the cases $\gamma > 0$ *and* $\gamma < 0$.
   (c) Discuss whether the learning rule is Hebbian, anti-Hebbian, or non-Hebbian.
2. **Fixed point of BCM rule**. *Assume a single postsynaptic neuron* ν_i *which receives constant input* $\nu_j > 0$ *at all synapses* $1 \le j \le N$.
   (a) Show that the weights w_{ij} *have a fixed point under the BCM rule (19.9).*
   (b) Show that this fixed point is unstable.
3. **Receptive field development with BCM rule**. *Twenty presynaptic neurons with firing rates* ν_j *connect onto the same postsynaptic neuron which fires at a rate* $\nu_i^{\mathrm{post}} = \sum_{j=1}^{20} w_{ij}\,\nu_j$. *Synaptic weights change according to the BCM rule (19.9) with a hard lower bound* $0 \le w_{ij}$ *and* $\nu_\theta = 10\,Hz$.
   The 20 inputs are organized in two groups of 10 inputs each. There are two possible input patterns ξ^μ, *with* $\mu = 1, 2$.
   (a) The two possible input patterns are: $\mu = 1$, *a group 1 fires at 3 Hz and group 2 is quiescent; and* $\mu = 2$, *group 2 fires at 1 Hz and group 1 is quiescent. Inputs alternate between both patterns several times back and forth. Each pattern presentation lasts for* Δt. *How do weights* w_{ij} *evolve? Show that the postsynaptic neuron becomes specialized to one group of inputs.*
   (b) Similar to (a), except that that the second pattern is now $\mu = 2$*: group 2 fires at 2.5 Hz and group 1 is quiescent. How do weights* w_{ij} *evolve?*
   (c) As in (b), but you are allowed to make ν_θ *a function of the time-averaged firing rate* $\bar{\nu}_i^{\mathrm{post}}$ *of the postsynaptic neuron. Is* $\nu_\theta = \bar{\nu}_i^{\mathrm{post}}$ *a good choice? Why is* $\nu_\theta = (\bar{\nu}_i^{\mathrm{post}})^2/10\,\mathrm{Hz}$ *a better choice?*
   Hint: Consider the time it takes to update your time-averaged firing rate in comparison to the presentation time Δt *of the patterns.*
4. **Weight matrix of Hopfield model.** *Consider synaptic weights that change according to the following Hebbian learning rule:* $\frac{\mathrm{d}}{\mathrm{d}t} w_{ij} = c(\nu_i - \nu_0)\,(\nu_j - \nu_0)$.
   (a) Identify the parameters c and ν_0 *with the parameters of Eq. (19.2).*
   (b) Assume a fully connected network of N neurons. Suppose that the initial weights w_{ij} *vanish. During presentation of a pattern* μ, *activities of all neurons* $1 \le k \le N$ *are fixed to values* $\nu_k = p_k^\mu$, *where* $p_k^\mu \in \{0, 1\}$ *and synapses change according to the Hebbian learning rule. Patterns are applied one after the other, each for a time* Δt. *Choose an appropriate value for* ν_0 *so that, after application of P patterns, the final weights are* $w_{ij} = \gamma \sum_{j=1}^{P} p_i^\mu\, p_j^\mu$. *Express the parameter* γ *by* $c, \nu_0, \Delta t$.
   (c) Compare your results with the weight matrix of the Hopfield model in Chapter 17. Is the above learning procedure realistic? Can it be classified as unsupervised learning?
   Hint: Consider not only the learning phase, but also the recall phase. Consider the situation where input patterns are chosen stochastically.
5. **PCA with Oja's learning rule**. *In order to show that Oja's learning rule (19.7) selects the first principal component proceed in three steps.*
   (a) Show that the eigenvectors $\{e_1, \ldots, e_N\}$ *of C are fixed points of the dynamics.*

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press

Hint: Apply the methods of Section 19.3 to the batch version of Oja's rule and show that

$$\Delta w = \gamma C w - \gamma w [w \cdot C w] . \tag{19.49}$$

The claim then follows.
(b) Show that only the eigenvector e_1 with the largest eigenvalue is stable.
Hint: Assume that the weight vector $w = e_1 + \varepsilon e_k$ has a small perturbation $\varepsilon \ll 1$ in one of the principal direction. Derive an equation for $\mathrm{d}\varepsilon/\mathrm{d}t$ and show that the perturbation grows if $k \neq 1$.
(c) Show that the output rate represents the projection of the input onto the first principal component.

6. **Triplet STDP rule and BCM**. *Show that for Poisson spike arrival and output spike generated by an independent Poisson process of rate ν_i, the triplet STDP model gives rise to a rate-based plasticity model identical to BCM. Identify the function ϕ in Eqs. (19.8) and (19.9) with the parameters of the triplet model in (19.15).*
Hint: Use the methods of Section 19.3.3. Independent Poisson output means that you can neglect the pre-before-post spike correlations.

https://doi.org/10.1017/CBO9781107447615.023 Published online by Cambridge University Press