

# Spike-timing-dependent plasticity: common themes and divergent vistas

Adam Kepecs<sup>2</sup>, Mark C.W. van Rossum<sup>1</sup>, Sen Song<sup>2</sup>, Jesper Tegner<sup>3,\*</sup>

Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454, USA

<sup>1</sup>ANC, University of Edinburgh, 5 Forest Hill, Edinburgh, EH1 2QL, UK

<sup>2</sup>Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

<sup>3</sup>Department of Physics, IFM, Computational Biology, Linköpings Institute of Technology, S-581 83 Linköping, Sweden

Received: 4 June 2002 / Accepted: 12 July 2002

**Abstract.** Recent experimental observations of spike-timing-dependent synaptic plasticity (STDP) have revitalized the study of synaptic learning rules. The most surprising aspect of these experiments lies in the observation that synapses activated shortly after the occurrence of a postsynaptic spike are weakened. Thus, synaptic plasticity is sensitive to the temporal ordering of pre- and postsynaptic activation. This temporal asymmetry has been suggested to underlie a range of learning tasks. In the first part of this review we highlight some of the common themes from a range of findings in the framework of predictive coding. As an example of how this principle can be used in a learning task, we discuss a recent model of cortical map formation. In the second part of the review, we point out some of the differences in STDP models and their functional consequences. We discuss how differences in the weight-dependence, the time-constants and the non-linear properties of learning rules give rise to distinct computational functions. In light of these computational issues raised, we review current experimental findings and suggest further experiments to resolve some controversies.

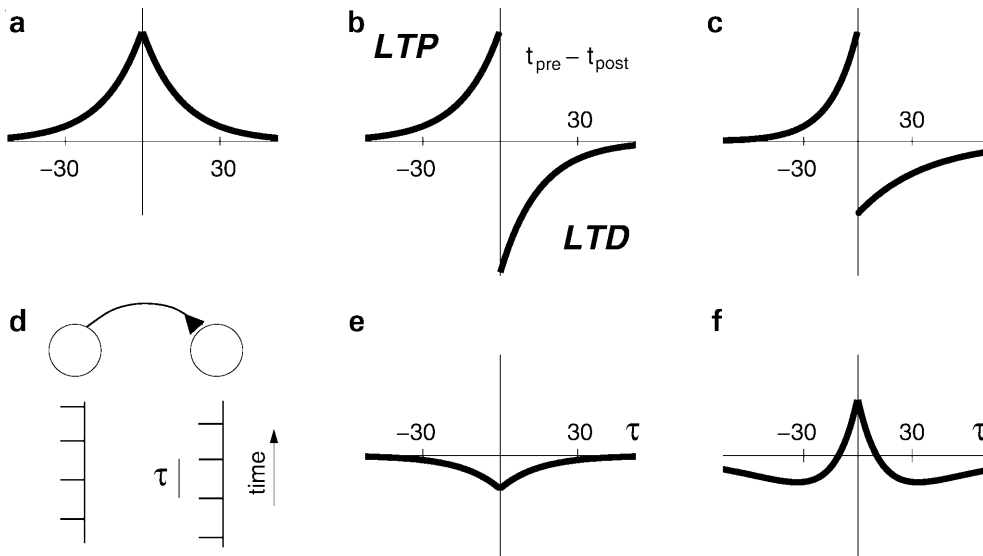
## 1 Temporally asymmetric plasticity

Early models of synaptic learning used the correlations between the firing rates of pre- and postsynaptic neurons as the signal for plasticity (von der Malsburg 1979; Bienenstock et al. 1982; Miller et al. 1989). While there were some indications that temporal order might be important for plasticity (Levy and Steward 1983; Gustafsson et al. 1987; Debanne et al. 1994), initial phenomenological characterizations of long-term synaptic plasticity largely focused on the requirements of coincident activity between pre- and postsynaptic

neurons (Barrionuevo and Brown 1983; Malinow and Miller 1986) leading to rate based learning rules. Markram et al. (1997), however, showed that synapses can be robustly weakened if the presynaptic spike arrived shortly after the postsynaptic spike and that the transition between potentiation and depression is very sharp (Fig. 1B). Later studies confirmed this in several brain regions and refined the temporal properties of plasticity (Magee and Johnston 1997; Zhang et al. 1998; Bi and Poo 1998; Feldman 2000; Sjöström et al. 2001). Fueled by these experiments demonstrating the importance of spike-timing in determining both the direction and magnitude of synaptic modification, there has been increasing interest in learning rules that take the timing of spikes into account (Abbott and Nelson 2000). Here we review recent developments in STDP but first we discuss a property common to many computational models of STDP and discuss how it relates to predictive coding. We illustrate how this principle can be used in a learning task by reviewing a recent model of cortical map formation. In the second part of the review, we consider some of the differences in STDP models and their functional consequences.

These recent studies on STDP come in the backdrop of a large body of work based on Hebb's postulate of learning (Sejnowski 1999). Hebb's original postulate is consistent with the observed temporal asymmetry, stressing the importance of *causality* in contrast to coincidence as a condition for synaptic strengthening. For a spike in neuron *A* to contribute to neuron *B*'s firing, it would need to arrive before a spike was initiated in neuron *B*. Thus causality provides a temporally asymmetric condition for synaptic strengthening. Hebb did not specify what should happen if the spike in neuron *A* lags behind the spike in neuron *B*. In fact, Hebb's postulate did not address the issue of synaptic weakening at all. However, a reasonable extension in the spirit of causality suggests that synaptic weakening should occur when inputs systematically "miss" the opportunity to fire their postsynaptic neuron. Such a temporally asymmetric Hebbian rule can be thought of as a natural extension of the causality criterion of the original postulate,

\* Present address: Correspondence to: Sen Song  
(e-mail: songs@cshl.edu)



**Fig. 1.** Different time windows for spike-timing-dependent plasticity. Plasticity time windows show the amount of change in synaptic weight as a function of the time difference (ms) between pre- and postsynaptic spikes. **A** Hypothetical symmetric time window as illustrates the “coincidence” interpretation of classical activity-dependent Hebbian learning. **B** Observed temporally asymmetric rule with an anti-symmetric time window. The time constants for potentiation and depression are the same. **C** Temporally asymmetric rule with an

asymmetric time window. The depression time window is longer as has been observed at some synapses. **D** Two connected neurons are clamped with periodic spike trains. The trains have a period  $\tau$  and are  $180^\circ$  out of phase. **E** The “effective time window” for the network in **D** for the plasticity window shown in **B**. Note that depression dominates. **F** The “effective time window” for the plasticity window shown in **C**. Note that either strengthening or weakening can occur depending on the input frequency

namely, that inputs which did not participate in spike generation, but fired briefly after, are weakened. This interpretation of the rule also implies that there is a natural time scale for the inputs received by a neuron, which is set by the time window of plasticity.

These principles define a new class of learning rules with two characteristic properties. First, they are temporally asymmetric (Fig. 1B), which should be contrasted with classic, associative rules based on covariance (Fig. 1A). Second, plasticity depends on the timing of spikes within a short time window. While it is not clear how such temporal precision (a complete reversal in the sign of plasticity in a few milliseconds!) is achieved biophysically, the computational consequences of this are substantial.

## 2 Temporal asymmetry and predictive coding

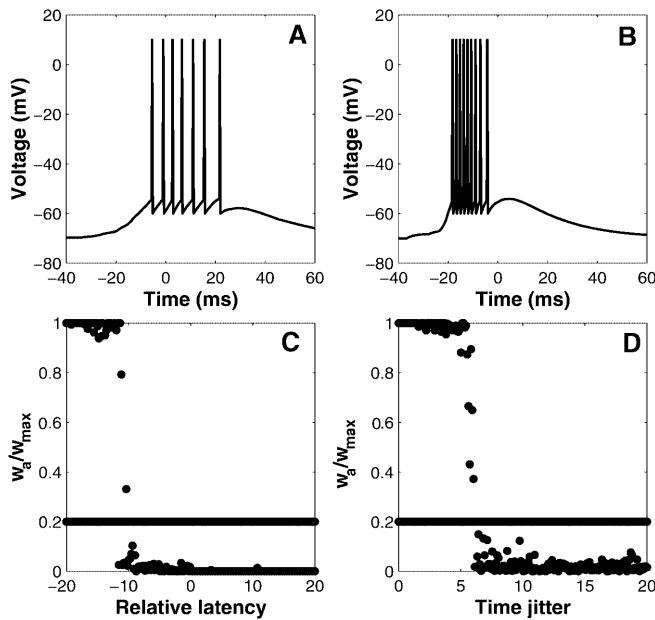
The temporal asymmetry of learning allows for a wide range of learning rules not covered by classical theories of Hebbian learning. For instance, as demonstrated by several authors, learning rules based on temporally asymmetric STDP can yield a differential Hebbian learning rule, where synaptic strength is changed according to the correlation between the derivatives of the rates rather than the correlation between rates (Roberts 1999; Seung and Xie 2000; Rao and Sejnowski 2000).

Differential Hebbian learning is closely related to temporal difference (TD) learning in reinforcement learning (Rao and Sejnowski 2001), which can be used to model classical conditioning. It allows the proper

arrangement of the synaptic weights to pick out the most predictive features in the synaptic inputs. Closely related to this, STDP can be used to learn temporal delays with a high precision, which has been used to model auditory processing in the barn-owl (Gerstner et al. 1996). Sequence learning, an important task for an organism to establish directional association between events, can also be accomplished using the temporal asymmetry inherent in STDP (Blum and Abbott 1996). Interestingly, these last two theoretical studies explored the computational importance of temporally asymmetric Hebbian learning before STDP was clearly established experimentally.

Temporally asymmetric STDP connects Hebbian learning with predictive coding. If a feature in the synaptic input pattern can reliably predict the occurrence of a postsynaptic spike and seldom comes after a postsynaptic spike, the synapses related to that feature are strengthened, giving that feature more control over the firing of the postsynaptic cell. The most predictive feature can be different in different situations. Two scenarios are considered in Fig. 2.

First, we consider a situation where different inputs to a neuron have different latencies, perhaps corresponding to inputs received from different pathways. Here the most predictive feature is the one with the shortest latency. To show how STDP selects these inputs we simulated a neuron driven by 200 excitatory synapses, each having a typical delay with a small jitter. Synapses were modified according to the STDP model introduced by Song et al. (2000). Figure 2C shows that learning selectively strengthened the low latency inputs. In this state, the neuron fired earlier (Fig. 2B) than before learning



**Fig. 2.** Reinforcing predictive features with STDP. An integrate and fire neuron was driven by 200 excitatory inputs. Each afferent fired once at a typical latency and jitter every 500 ms. In (A-C) each afferent had a different latency from  $-20$  to  $20$  ms and had the same time jitter of  $2$  ms. **A** Firing before learning. **B** After learning, firing occurs earlier and for a shorter period. **C** Synaptic weights before (line) and after learning (dots). Low latency inputs are strengthened. **D** STDP also selects precise inputs. Each afferent arrived at the same mean latency but with a different time jitter ranging from  $0$  to  $20$  ms. The time jitter was exponentially distributed and was either positive or negative. The synaptic weights are plotted before (line) and after learning (dots). Inputs with low jitter are strengthened while others are weakened (Song et al. 2000)

(Fig. 2A), which is a direct indication of a predictive shift in firing.

If instead, the different inputs have different temporal precision, then the most predictive inputs are the ones with the smallest jitter. We simulated this scenario similarly to the previous one. Now, however, all input spikes arrived at the same time on average, but with different time jitter. The jitter was drawn from an exponential distribution with each afferent having a unique value (Fig. 2D). STDP selectively strengthened the inputs with the smallest jitter (Fig. 2D) and thereby produced more reliable firing of the postsynaptic neuron. Similar results can be obtained if the reliability of an afferent corresponds to the probability of release (not shown). These results show that STDP preferentially strengthens the most reliable and the shortest latency inputs.

### 2.1 Consequences for network function

In networks of neurons, the predictive property of temporally asymmetric STDP can be used to perform predictive sequence learning. STDP can strengthen the appropriate pattern of synapses so that activity in the network at one instant activates the predicted pattern of neurons at the next instant. Temporally asymmetric

STDP has been suggested as a possible mechanism for sequence learning in the hippocampus (Abbott and Blum 1996; Minai and Levy 1993). A backward shift in hippocampal place cells was predicted based on these studies (Blum and Abbott 1996), which was later observed experimentally during route learning in rats (Mehta et al. 1997; Mehta and Wilson 2000). The same principles have been suggested to underlie the emergence of directionally sensitive visual neurons (Rao and Sejnowski 2000; Mehta and Wilson 2000; Senn and Buechs 2002).

When the time constants of potentiation and depression in the window function are identical (Fig. 1B, E), STDP discourages the formation of short, mutually excitatory loops. Namely, if neuron *A* is predictive of the firing of neuron *B*, then neuron *B* cannot be predictive of the firing of neuron *A* and must lag behind. Therefore, if the synapses from neuron *A* to neuron *B* are strengthened, the synapses from neuron *B* to neuron *A* will be weakened, making a mutually excitatory configuration unstable. This shows why the formation of feed-forward structures (Song and Abbott 2001) and synfire chains (Horn et al. 2002) is favored. Horn et al. (2002) found that, in a sparsely connected network, temporally asymmetric STDP can spontaneously break up the assembly of cells into synchronous groups of cells that fire in cyclic manner.

Various refinements of STDP models can explain the existence of bi-directional connections, so ubiquitous in cortex. First, when the time scale of LTD is longer than the time scale of LTP, bi-directional connections are made possible (Song and Abbott 2001). This can be understood by considering two neurons driven by periodic inputs (Fig. 1D). The net effect of STDP will be to strengthen recurrent synapses that connect them. This happens due to the differences in the LTP and LTD time windows (Fig. 1C), resulting in an effective time window for plasticity that has both a strengthening and weakening lobe (Fig. 1F). Synapses in a reciprocally connected pair of neurons can strengthen if the correlation between them is tight enough.

Furthermore, STDP has been observed to be dependent on the pairing frequency (Markram et al. 1997; Sjöström et al. 2001). For high stimulus frequencies LTP becomes stronger; a high frequency spike train pre- and postsynaptically, will strengthen the connections in a manner largely independent of the precise pre-post timing.

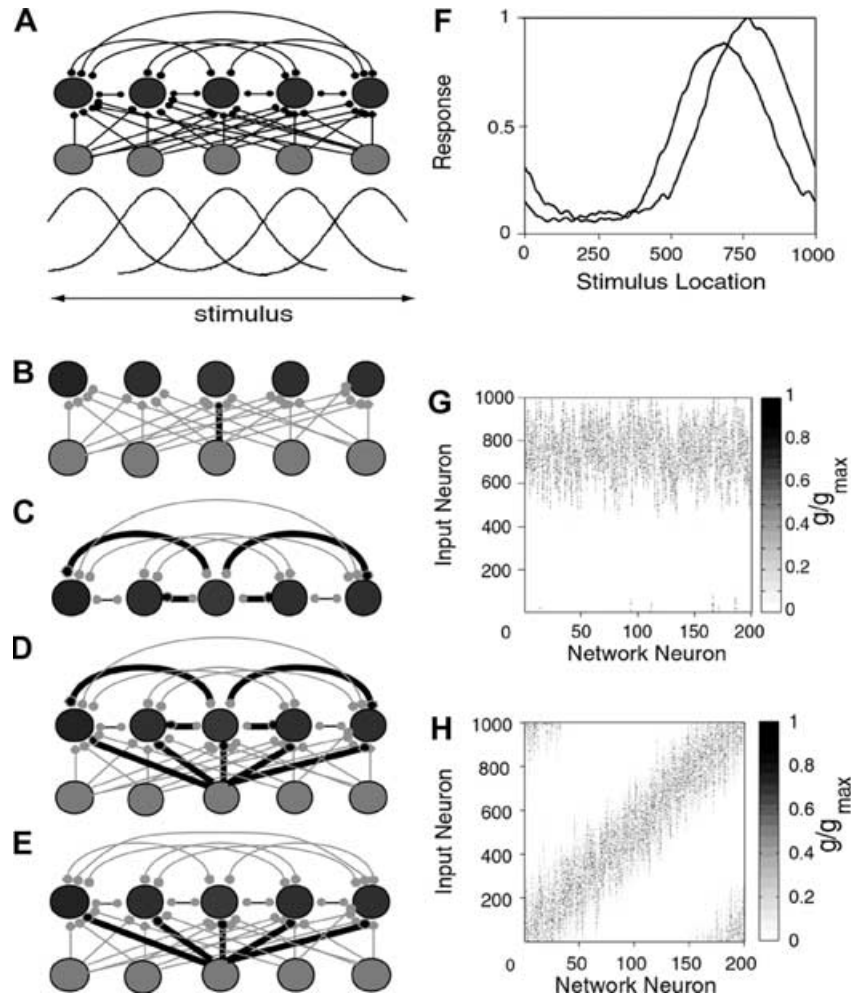
## 3 Formation of maps

The temporal asymmetry of STDP plays an important role in a recent model of column and map formation (Song and Abbott 2001). This study shows that STDP can enable information contained in the connectivity of one part of the network to be transferred to other parts of the network (Song and Abbott 2001). This model contains a feed-forward input layer and a recurrent network layer. Using STDP with minimal constraints, columns of neurons tuned to a similar location, and

hypercolumns consisting of collections columns in orderly progression, can be developed in the network layer. The model differs from previous models in that both the feed-forward connections and the recurrent connections are plastic at all times.

Considering a network of these neurons, without the recurrent connections, neurons in the network layer develop receptive fields at random locations. However, when recurrent connections are present, a single column of neurons all tuned to the same location is formed (Fig. 3G). The sequence of events leading to the for-

mation of a column is presented in Figs. 3B–E. Each circle in the graph represents a group of neurons. For the ease of presentation, a seed has been added to the original network consisting of increased strength for neurons from the center input group to the center network group (Fig. 3B). This arrangement introduces a bias into the network by making the central network group correlated with each other. This bias breaks the symmetry in the network group and makes the process of column formation faster and easier to visualize. It is not required for the development of a single column. In



**Fig. 3.** Formation of columns and maps under STDP. **A** Network diagram. Two layers are simulated. The input layer contained neurons that fire Poisson trains at different rates and the network layer contained integrate-and-fire neurons. Every neuron in the input layer is randomly connected to one fifth of the neurons in the network layer and all the neurons in the network layer are recurrently connected. *Circles* in the figure represent groups of neurons. All synaptic connections are governed by STDP. Input correlations are introduced by moving a Gaussian hill of activity along the input layer. The Gaussian hill is centered on a random input neuron for a period of time and then shifted to another random location at the end of the period. The length of the period is chosen from an exponential distribution and the time constant is similar to the time window of STDP. **B** First stage of column formation. Seed placed in the feed-forward connections creates a correlated group of network neurons.

**C** Second stage of column formation. Correlated group of network neurons send out connections to other neurons in the network. **D** Third stage of column formation. Transfer of information from recurrent layer to feed-forward layer. **E** Last stage of column formation. Recurrent connections weaken as feed-forward connections become well formed. **F** Receptive fields of two network neurons. **G** Feed-forward synaptic strengths define a column. *Dark dots* represent strong synapses. The *horizontal stripe* indicates that the network neurons have similar input connections, i.e., receptive fields. **H** When short range excitatory and global inhibitory recurrent connections are introduced in the network layer, a map forms in the feed-forward connection. The diagonal bar reflects the progression of receptive field centers as we move across the sheet of network neurons (Song and Abbott 2001)

the absence of this arrangement, a group of neurons in the network layer will automatically become more correlated than others to break the symmetry. However, the process of symmetry breaking can be time consuming and the correlated group of neurons are not always adjacent to each other on the network neuron making visualization difficult.

The group of correlated network neurons will strengthen connections to other neurons in the network. This is easily understood from the perspective of a neuron that is not part of the correlated group (Fig. 3C). From this perspective, STDP strengthens only the synapses of the most correlated inputs. At this stage of the development of a column, activity originates in the input layer, passes unto the correlated group of neurons, and then unto other neurons in the network layer. However, the firing of the input layer precedes that of the other neurons in the network layer. Therefore, the connections from the input layer unto the other neurons in the network layer would be strengthened (Fig. 3D). Once this pathway reaches sufficient strength, the recurrent connections weaken because the direct pathway has a shorter latency and is therefore the preferred pathway. Eventually, a column defined wholly by the feed-forward connections forms (Fig. 3E), completing the process of the transfer of connectivity information from the recurrent connections in the network to the feed-forward connections. The transfer of information from the recurrent connections to the feed-forward connections is possible due to the predictive properties of STDP. If the recurrent connections are restricted to a neighborhood and uniform lateral inhibition is imposed on the network layer, a map consisting of a collection of columns in orderly progression of receptive field center is formed (Fig. 3H).

The model predicts that the development and plasticity of column and map structure on the network layer precedes the development and plasticity of feed-forward inputs as is supported by recent experimental evidence (Trachtenberg and Stryker 2001).

#### 4 Dependence of STDP on synaptic strength

So far we have discussed consequences of STDP that appear to be common to different formulations of the learning rule. Importantly, however, models proposed for STDP differ in subtle ways that turn out to have significant computational consequences. Both the shape of the plasticity time window (anti-symmetric, asymmetric, etc.; see Fig. 1) and the rule used to update the weights after a plasticity “event” can be different. In this and the following section we address two such differences in models: the dependence of the weight update rule on the present weight and the non-linearity of the interaction between two plasticity events.

One of the crucial issues – how the specific form of STDP determines the functional consequences of learning – has to do with whether the magnitude of synaptic plasticity is dependent on the present strength of the synapse. In the learning rules studied by

Song et al. (2000) and Kempter et al. (2001), the weight increases and decreases are independent of the present weight of a synapse. In this case an upper and a lower bound has to be placed on synaptic strength for stable learning. The resulting weight distribution will cluster near the extreme weights and can be bimodal.

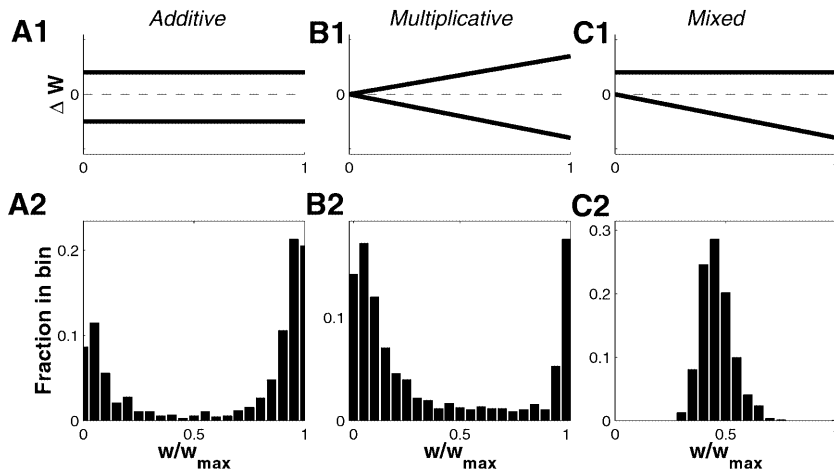
If the changes in synaptic weights are dependent on the present weights then the resulting learning rule will have quite different properties. Specifically, if synaptic weakening is inversely proportional to the strength of a synapse, while synaptic strengthening is independent of the present weight of a synapse, the resulting distribution of synaptic strengths is unimodal rather than bimodal (Kistler and van Hemmen 2000; van Rossum et al. 2000). However, there are many other possible ways in which the update rule can incorporate weight-dependence (Rubin et al. 2001).

Here we examine three different update rules to emphasize how they determine the computational properties of a synaptic plasticity. First, it is possible that there is no weight-dependency and the weight changes are fully determined by their timing. This is referred to as an *additive* update rule (Rubin et al. 2001). On the other hand, some data suggest that synaptic depression, but not potentiation, depends on the present weight of the synapse (Bi and Poo 1998). Other dependencies are also possible, for example, potentiation may depend on the difference between the present weight and the maximal weight (Rubin et al. 2001). We will refer to these as *mixed* update rules.<sup>1</sup> Finally, we consider what happens if both potentiation and depression are multiplicatively dependent on the present synaptic weight. We call this scenario *multiplicative* update rule.

Figure 4A illustrates the case when the update rule is additive and the final weight distribution is bimodal (Song et al. 2000; Kempter et al. 2001). If both potentiation and depression depend multiplicatively on the weight, but depression dominates (Fig. 4B), then the final weight distribution is again bimodal. But if only depression depends on the weight then the final weight distribution is unimodal (Fig. 4C; van Rossum et al. 2000; Rubin et al. 2001). Aharonov et al. (unpublished) proposed a way to interpolate between a purely additive and a mixed update rule. They show that one can optimally adjust the functional properties of this rule with an interpolation parameter.

How update rules determine the final weight distribution can be determined using a Fokker-Planck analysis explained in the Appendix (van Rossum et al. 2000; Rubin et al. 2001). Such analysis has only been done for Poisson driven neurons in which the final weight distribution is determined by the balance between potentiation and depression. In the additive model, potentiation and depression almost cancel each other for all weights. The change in strength of a particular synapse will strongly depend on the correlation between this input and the output spikes. Once a synapse becomes strong

<sup>1</sup>Note that this type of rule has also been called multiplicative in previous papers (Rubin et al. 2001).



**Fig. 4.** The effects of differential weight-dependence of potentiation and depression on the final synaptic weight distributions. *Top panels* show the dependence of LTP (*upper curve*) and LTD (*lower curve*) on the present synaptic weight. *Bottom panels* show the final synaptic weight distributions of 1000 afferents driven by Poisson inputs. The parameters of the models are the same as in Song et al. (2000) except for the update rule in *B* and *C*. **A** Additive rule. The final synaptic weight distribution is bimodal. **B** Multiplicative rule. Simulations using a multiplicative rule for both potentiation and depression. **C** Mixed rule. Simulations using the (noisy) update rule after van Rossum et al. (2000)

enough, it will be strengthened further. When potentiation and depression do not balance out, such as in Fig. 4C the effect of correlations is much weaker.

#### 4.1 Synaptic competition and rate stabilization

These differences in update rules and final weight distributions have important computational consequences. Specifically, they are related to whether STDP results in synaptic competition and the way learning is stabilized.

Models that result in a bimodal weight distribution show strong competition and rate stabilization. However, competition only occurs for a limited range of input rates and requires the fine tuning of parameters (Sect. 5). In contrast, there is no rate stabilization in models that lead to unimodal weight distributions. Synaptic competition is also lacking because higher postsynaptic rates are not compensated by a concomitant decrease in synaptic weights. Instead the postsynaptic neuron roughly extracts the principal component of the inputs (van Rossum et al. 2000).

Synaptic competition is of great importance as it is thought to underlie the formation of cortical maps (Blakemore et al. 1976; Garraghty et al. 1986) and places an important constraint on synaptic learning rules (Lara and di Prisco 1983; Miller 1996). Since other competitive mechanisms are also possible, such as those based on the conservation of resources (von der Malsburg 1973) or on neurotrophic factors (Harris et al. 1997), an integrated view is warranted (Van Ooyen 2002). For instance, homeostatic synaptic scaling (Turrigiano et al. 1998) could account for competition in models with mixed update rules that show no intrinsic competition (van Rossum et al. 2000).

Importantly, the mixed update rule STDP models introduced by Rubin et al. 2001 keeps the total synaptic weight roughly constant and does not control firing rates. On the other hand, the additive rule can stabilize the rate (Kempler et al. 2001) and keep a neuron in a balanced input regime where it is sensitive to the timing of its input (Abbott and Song 1999). Even when extra mechanisms for synaptic competition are added to

models with mixed update rules (van Rossum et al. 2000), their computational consequences can be quite different. For instance, as noted by Blais et al. (1999), traditional correlational rules using heterosynaptic competition cannot account for the activity-dependence of cortical remapping (Rittenhouse et al. 1999) unlike BCM which explicitly stabilizes the output activity (Bienenstock et al. 1982). In this regard it is important to note that the competition in additive STDP is also homosynaptic, arising via activity, which has important consequences for the dynamics of cortical remapping (Tegnér and Kepecs 2002b).

## 5 Beyond linear models of STDP

To fully understand the functional implications of STDP, it is essential to anchor computational models in available physiological data. However, as discussed in the above sections, reduced models have been instrumental in guiding our ideas toward the functional consequences of STDP. Simplified models have, for example, revealed the conditions for predictive coding (Sect. 2), map formation (Sect. 3), and synaptic competition occurrences (Sect. 4). A key assumption in these “first generation” STDP models is that the contribution from every pre/post spike interaction to synaptic plasticity is independent from one another. This implies that plasticity changes induced by closely spaced pairs of spikes sum linearly. Not surprisingly, recent experimental results have revealed a number of features not taken into account in the first generation of STDP models. For instance, Markram et al. (1997) showed that there is a threshold for plasticity at 5 Hz and saturation at 40 Hz. From a modeler’s perspective, it is natural to move beyond the first generation of STDP models and try to incorporate plausible biophysical mechanisms which could account for the range of experimental findings using different stimulus protocols. Clearly, a long-term objective is to have a complete model for synaptic plasticity which can account for results from both classical LTP/LTD experiments and STDP protocols.

Senn et al. (2001) took a first step toward this objective when they constructed a second generation

plasticity model using biophysically inspired kinetic schemes for the different states of the channels and second messengers that are thought to underlie the plasticity. Their model incorporates both a threshold and synaptic saturation and can therefore reproduce the results of Markram et al. (1997) and some more recent observations of non-linear interaction in STDP (Senn et al. 2001). This approach, however, leaves unclear what the functional implications of non-linear interactions might be.

Instead of designing detailed kinetic models to account for new experimental results, Tegnér and Kepecs (2002a, c) took another route and extended the competitive version of the STDP rule with an adaptive control loop to make it more robust. Song et al. (2000) observed that the exact value of depression to potentiation ( $\alpha = LTD/LTP$ ) in the update rule is critical for stable learning. For instance, to reach a balanced firing regime requires a learning ratio larger than unity (Abbott and Song 1999). Moreover, the synaptic weight distribution depends on the fine-tuning of the balance between parameters in the STDP model. Consequently, since the input statistics can easily change, it would be advantageous to put the highly sensitive learning parameter,  $\alpha$ , under activity-dependent control. Interestingly, when the learning ratio is controlled by the post-synaptic activity, pairs of spikes sum non-linearly (Tegnér and Kepecs 2002a), similar to what has been recently observed by Froemke and Dan (2002).

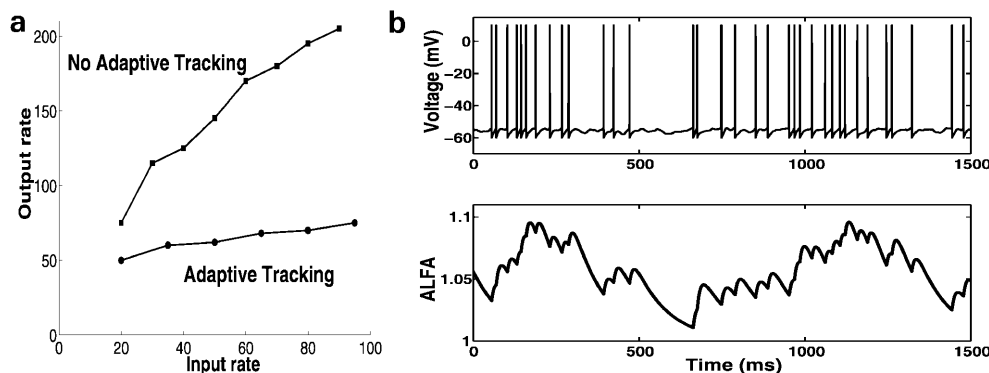
The basic idea of the adaptive extension to STDP is to allow the ratio of LTD to LTP ( $\alpha$ ) to dynamically change. In this model, postsynaptic calcium is used to measure the postsynaptic firing rate (Tegnér and Kepecs 2002c). This is useful because calcium signals are thought to underlie synaptic plasticity (Lisman 1989). Therefore it is plausible that the asymmetry between LTP and LTD will depend on the level of postsynaptic calcium. For example, it is known that increased resting calcium levels inhibit NMDA channels (Rosenmund et al. 1995; Umekiya et al. 2001) and thus calcium influx due to synaptic input. Additionally, the calcium levels required for depression (Lisman 1989) are easier to

reach. Both of these effects in turn increase the probability of LTD induction. These effects can be incorporated into a simple kinetic scheme, where  $\alpha$  changes as a function of calcium, which in turn is due to action potentials (Tegnér and Kepecs 2002a).

How does this rule differ from the original with a static  $\alpha$ ? In the static case, when a neuron receives correlated input, the normalization property (Kempster et al. 2001) of competitive STDP is fragile. Driving a neuron with increasing input rates increases the output rate significantly. Adding the adaptive control loop normalizes the output rates (Fig. 5A). This simulation shows that the average postsynaptic firing rate is regulated by the adaptive tracking scheme. Depending on the parameters of the rule, different regimes are possible. For instance, the final firing rate of a neuron after learning can become entirely independent of input rate and dependent only upon input correlations (Tegnér and Kepecs 2002a).

Figure 5B shows the evolution of the membrane potential (top) and the learning ratio  $\alpha$  (bottom). Note that because the adaptive rule tracks fast changes in firing by adjusting the learning ratio for every spike, the strength plasticity is different for each spike. Interestingly, the learning ratio  $\alpha$  fluctuates around 1.05, the value used in previous studies (Song et al. 2000). The adaptive STDP rule adjusts the learning ratio on a millisecond time-scale based on the rapid feedback regulation of NMDA channels (Umekiya et al. 2001). This is in contrast to other slow homeostatic controllers considered previously (LeMasson et al. 1993; Turrigiano et al. 1998; Turrigiano and Nelson 2000; van Rossum et al. 2000). Due to the rapid changes, the adaptive rule is highly sensitive to the spike-to-spike dynamics. Consequently, the preceding pattern of post-synaptic spikes influence the size of synaptic changes. Therefore, rapid self-regulation implies non-linear addition of plasticity events.

These results are directly relevant to the experiments of Froemke and Dan (2002), designed to reevaluate the assumption of linear summation of spike interactions in STDP. To this end, they delivered more complex stimulus protocols using either: two presynaptic spikes



**Fig. 5.** Adaptive STDP **A** Steady-state response with (*squares*) or without (*circles*) the adaptive tracking scheme. When the STDP rule is extended with an adaptive control loop, the output rates are normalized in the presence of correlated input. **B** Fast adaptive tracking. Since the adaptive rule tracks changes in intracellular

calcium on a rapid time-scale, every spike experiences a different learning ratio,  $\alpha$ . Note that the adaptive scheme approximates the learning ratio ( $\alpha = 1.05$ ) used in Song et al. (2000) and Tegnér and Kepecs (2002a)

with a single postsynaptic spike in between; or, the reverse, a single presynaptic spike with two postsynaptic spikes, one occurring prior and the other following the presynaptic spike. Their results suggest a surprisingly simple conclusion: each spike (pre or post) suppresses the plasticity of the following spike.

While the mechanisms of the suppressive effects in STDP are not known, two models suggest that calcium inhibition of NMDA receptors could play an important role during the suppression by postsynaptic spikes (Senn et al. 2001; Tegnér and Kepecs 2002a). Short-term synaptic depression could possibly contribute in the presynaptic suppression effect. It will be an interesting challenge to uncover the mechanisms underlying suppression in the STDP scheme proposed by Froemke and Dan (2002) and relate these to computational questions. It is tempting to suggest that the observed suppressive non-linearity is a signature of an adaptive STDP rule.

An interesting application of an adaptive, or self-regulating STDP rule, may be found in a model for column formation (Sect. 3). Here, the average amount of input correlations changes quite dramatically because the origin of the synapses to the neurons are initially randomly assigned. As a consequence, each neuron normally receives a large number of synapses and the inputs are relatively uncorrelated. After learning, neurons receive input from their neighbors only. As a result, the inputs which the neuron receives are highly correlated with each other. Therefore, a static  $\alpha$  cannot control the firing rate or the irregularity of the firing. Simulations show that an adaptive STDP rule can produce neurons with irregular firing and reasonable rates without disrupting the process of column formation (Song 2002).

## 6 Summary

Experimental observations of the timing-dependence of synaptic plasticity are changing computational theories of learning. Here we have reviewed some of the recent computational studies that examined the consequences of an emerging class of learning rules based on the spike-timing-dependence of synaptic plasticity. While we have tried to include the many different directions in research, we have tried not to be exhaustive. For instance, “anti-Hebbian” learning rules based on timing have been observed in weakly electric fish (Bell et al. 1997; Han et al. 2000) and at cortical inhibitory synapses (Holmgren and Zilberter 2001), both of which have distinct computational consequences (Roberts 1999; Seung and Xie 2000).

The temporal asymmetry of the learning window implies that the ordering of spikes at the time scale of tens of milliseconds determines both the direction (potentiation or depression) and the degree of change in synaptic strength. If the inputs are changing together at the time scale of the learning rule, then the low latency and high reliability inputs will be strengthened (Fig. 2). Such effects might underlie temporal difference learning at the synaptic level (Rao and Sejnowski 2001) and the

formation of direction selective responses (Rao and Sejnowski 2000; Mehta and Wilson 2000; Senn and Buchs 2002).

At the network level the predictive properties of STDP allow for map formation (Song and Abbott 2001). This can be thought of as a form of self-supervised learning, where, if taught by the recurrent connections, STDP ends up selecting the fastest and the most reliable feed-forward inputs. Obviously, such latency reduction is highly advantageous in sensory systems.

### 6.1 Consequences of different timing dependent rules

While these common aspects of STDP are important, we would like to emphasize that some of the observed differences in plasticity at different synapses and the different interpretations of these findings have significant computational implications. The dependence of plasticity on the present synaptic strength is an important open experimental question. As we have seen, the different versions of STDP can show very different properties. Synaptic competition and output rate stabilization are crucially dependent on the nature of weight updating. If the weight-dependence of LTP and LTD is similar then it appears that STDP will result in synaptic competition (Fig. 4A,B). If LTP and LTD have mixed update rules then no synaptic competition is observed (Fig. 4C). Further, experiments are needed to uncover the different forms of plasticity at different classes of synapses (Sect. 6.2.1).

While most models have only considered a linear summation of different plasticity events, the vast literature of synaptic plasticity and recent reports regarding STDP (Markram et al. 1997; Sjöström et al. 2001; Froemke and Dan 2002) suggest that this assumption needs to be reconsidered. It is likely that a certain threshold is needed for plasticity and that plasticity saturates (Petersen et al. 1998). A similar model by Senn et al. (2001) demonstrated that timing and frequency-dependence need not be exclusive properties of synaptic learning rules. The work of Froemke and Dan (2002) suggests that there might be important simplifications possible at the descriptive level of plasticity which can be summarized as preceding plasticity events to suppress future plasticity. Tegnér and Kepecs (2002a) suggested that similar non-linearities in learning are required for a self-regulating, adaptive plasticity rule. However, the frequency-dependent enhancement of LTP (Markram et al. 1997; Sjöström et al. 2001) and the suppressive effects of previous spikes (Froemke and Dan 2002) do not seem consistent with one another. Further work will be required to clarify the circumstances under which some of the same biochemical and biophysical processes result in different plasticity properties.

### 6.2 Experimental issues

These theoretical studies raise questions that lead us to propose experiments in order to clarify the properties



and computational implications of spike-timing-dependent plasticity.

**6.2.1 Weight-dependence.** We have seen that the dependence of plasticity on synaptic weight is important for competition and the steady-state weight distribution. Experimental observations about the weight-dependence of plasticity do not provide a clear cut picture. Some data from cultured cells support the differential dependency of LTP and LTD on synaptic weight (Bi and Poo 1998). However, synapses are scarce in cell cultures and, as a result, often unusually large. In slice experiments the inverse relationship is not particularly strong and depends mostly on a few large synapses (Sjöström et al. 2001).

Two factors complicate the situation. First, the maximal strength is likely to be different for each synapse, for instance, as constrained by the size of the postsynaptic density (Lisman and Harris 1993; Shepherd and Harris 1998). Therefore, the inverse relationship observed could be a dependence of STDP on the *maximal* synaptic strength of each synapse and not the actual strength of the synapse. Another possibility is that some strong synapses close to their maximal value show saturation. This can act as a soft-bound and the rule could be competitive for some choices of parameters.

Few experiments have applied a weakening protocol to previously potentiated synapses. Such experiments would be crucial to distinguish the different forms of the weight update rules. From a theoretical point of view these experiments could factor out the dependence of plasticity on the maximal weight of a synapse. This is an interesting issue in light of previous experimental observations which state that LTD and de-potentialization might have very different properties (Dudek and Bear 1992; Montgomery and Madison 2002).

**6.2.2 Timing, rate and location-dependence of plasticity.** Most models reviewed here make the simplifying assumption that spike-spike interactions during plasticity add linearly. While it is clear that plasticity requires a certain threshold as well as shows saturation, the frequency-dependence of plasticity provides fertile ground for further experiments. For instance, in some spike-timing based plasticity protocols potentiation is enhanced by high frequency pairing of spikes and no depression is observed beyond 40 Hz (Markram et al. 1997; Sjöström et al. 2001). On the other hand, preceding spikes seem to suppress plasticity in some induction protocols (Froemke and Dan 2002). The circumstances under which either enhancement or suppression dominate need to be further explored. For instance, the firing rates on the pre- and postsynaptic sides could be manipulated independently. Such protocols would enable the dissection of pre- and postsynaptic factors controlling plasticity and would directly test the adaptive STDP rule (Sect. 5).

Clearly timing is not the only determinant of plasticity. Uncovering the biophysical mechanisms of how synaptic plasticity is triggered will help in elucidating the balance between frequency, depolarization and spike-

timing based factors in controlling plasticity. For instance, further work should characterize NMDA channel activation during multi-spike interactions. One important question in this regard is whether calcium inhibition of NMDA channels can account for timing-dependent LTD and the suppressive effects of nearby spikes.

In most STDP experiments the postsynaptic spike was injected into the soma. Under physiological conditions many synapses participate in firing a postsynaptic spike and excitatory inputs arrive onto the dendrites. Experiments are required to determine whether the rules of plasticity differ when postsynaptic spikes are triggered due to synaptic activation as opposed to somatic current injection. Such experiments could refine the associativity requirement *inherent* in STDP and shed light onto the roles of dendritic voltage-gated conductances. For instance, A-type potassium currents were found to contribute to the spike dependence of plasticity (Watanabe et al. 2002). It would be interesting to examine whether A-type channels also enable spikes to selectively back-propagate and/or get boosted along the particular dendritic branches where EPSPs successfully forward propagated to cause the action potential. Such a mechanism would enable STDP only at the synapses that actually contributed to spike generation. Furthermore, A-type and other voltage-gated currents result in location-dependent changes in action potential shape, which could lead to dendritic input location specific differences in synaptic plasticity.

**6.2.3 Regional specialization of STDP.** Different spike-timing-dependent plasticity rules have different computational consequences, which may be related to regional specializations of neuronal function. For instance, if the time-window of LTP and LTD is the same, feed-forward processing is favored and bi-directional connections cannot be made. When the LTD time window is longer, then under some conditions recurrent networks can be formed. The time window of learning also determines the range of correlations to which plasticity is sensitive. It will be interesting to compare the time windows of STDP at different classes of synapses to their functional specialization at the network level. For instance, the plasticity of visual cortical orientation maps is sensitive to the order of stimulus presentations at the time scale of tens of milliseconds (Yao and Dan 2001; Schuett et al. 2001). Interestingly, plasticity is minimal at pinwheel centers which are regions of visual cortex where all orientations are represented near one another (Schuett et al. 2001). Similar stimulus timing-dependent plasticity has also been observed in auditory cortex (Ahissar et al. 1998). More studies are necessary to connect the functional specialization of different cortical areas with the timing-dependence of their plasticity. It would also be useful to investigate whether STDP is itself subject to state-dependent and/or developmental regulation.

In conclusion, our knowledge of synaptic plasticity is still limited. For instance, Egger et al. (1999) have reported spike-timing-dependent but symmetric plasticity at synapses of layer four stellate cells. Here, coincident

activity decreases synaptic strength regardless of the ordering of pre- and postsynaptic spikes. However, no mechanism for strengthening was found. Such observations illustrate our lack of knowledge about the spectrum of processes which together control synaptic strength. Since such observations could be due to the limitations of the preparation, it also highlights the need for experimental conditions which match real life as closely as possible. Undoubtedly, such studies will lead to further insights, bringing a closer match between experimental and theoretical work on spike-timing-dependent plasticity.

### Appendix: Fokker-Planck analysis of STDP

It is advantageous to analytically calculate the final synaptic weight distribution resulting from an STDP rule. Such calculations highlight the roles of the different learning parameters. Furthermore, since the rate of convergence can be slow for these models there is always an issue whether the observed synaptic distribution reflects a transient phenomenon rather than the true steady-state synaptic weight distribution. Several authors have developed similar methods to derive analytical expressions for final synaptic weight distributions after STDP (van Rossum et al. 2000; Rubin et al. 2001; Cateau et al. 2002). In the two following sections we use this framework to analyze the different weight-dependent formulations of STDP (Sect. 4) and the adaptive version of STDP (Sect. 5).

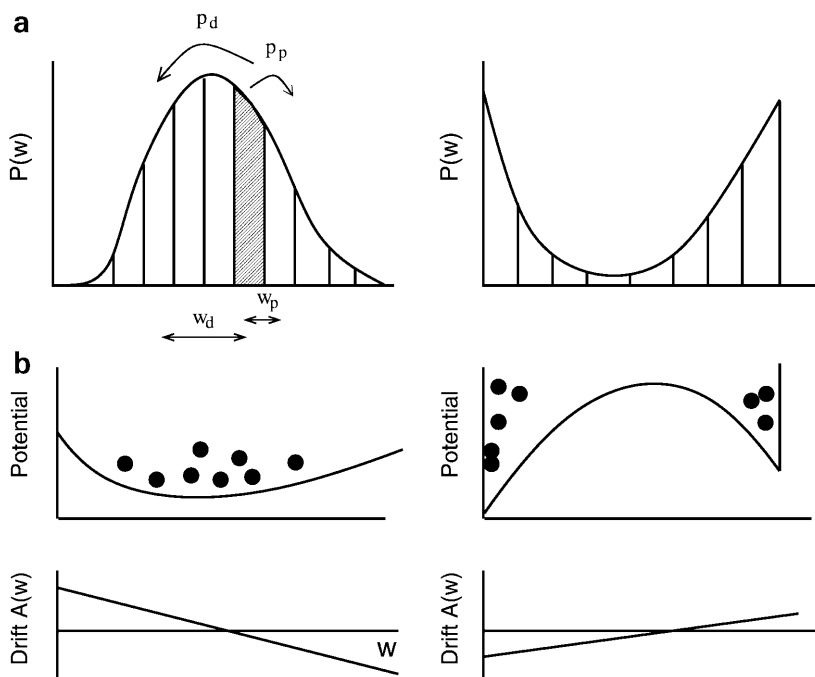
Here, we analyze a neuron receiving uncorrelated Poisson inputs via synapses that are modified according to an STDP learning rule. The basic idea is to count the influx and outflux of weights for a given bin in a synaptic weight distribution  $P(w, t)$  (Fig. 6A left). Then we examine how  $P(w, t)$  changes over time when synapses

evolve according to the STDP rule. Hence, collecting all the loss and gain terms for a bin in the synaptic weight distribution gives a Fokker-Planck problem describing how the synaptic weight distribution changes over time. Taylor expansion of  $P(w)$  gives

$$\frac{1}{r_{\text{pre}}} \frac{\partial P(w, t)}{\partial t} = - \frac{\partial}{\partial w} [A(w)P(w, t)] + \frac{\partial^2}{\partial w^2} [B(w)P(w, t)] \quad (1)$$

The drift term  $A(w)$  is the average net “force” experienced by an individual synapse. Whether a synapse increases or decreases depends on whether  $A(w)$  is positive or negative for the given weight  $w$ . In analogy with statistical physics we can introduce a potential  $V(w) = \int_0^w dw' A(w')$ . The synaptic weights jump around on this potential surface shown in Fig. 6B.  $A(w)$  is given by  $p_d w_d + p_p w_p$  where  $p_p$  and  $p_d$  denote the probability for potentiation and depression respectively, and  $w_p$  and  $w_d$  describe how much a weight changes as a result of potentiation and depression respectively. With a few simplifying assumptions, closed expressions for  $p_p$  and  $p_d$  can be found. The time window  $t_w$  can be defined through the probability for depression ( $p_d = t_w/t_{\text{isi}}$ ) that a synaptic event occurs with the time window ( $t_w < t_{\text{isi}}$ ). Treating the synaptic input as a brief square pulse causing a jump in the potential gives  $p_p = p_d(1 + w/W_{\text{tot}})$  for the potentiation. Here  $W_{\text{tot}} = t_w r_{\text{pre}} N(w)$ , is a competition parameter for  $N$  excitatory afferents.

We will consider four cases: (a) multiplicative update for the LTD only, (b) multiplicative update in both LTD and LTP, (c) multiplicative in only LTP, and (d) the additive rule. In the first case (a), considered by van Rossum et al. (2000) and Rubin et al. (2001), the amount the synaptic weight changes from potentiation and depression respectively is given by  $w_p = c_p$  and  $w_d = -c_d w$



**Fig. 6.** Evolution of the synaptic weight distribution. **A** A mixed update rule is shown on the left, while the additive rule is shown on the right. **A** The synaptic weights continually change due to potentiation and depression (arrows). When the synaptic weight distribution is in equilibrium, the influx into each bin of the histogram matches the outflux. **B** In analogy, the synaptic weights hop around on a potential surface. The shape of the potential given by the drift, determines the final distribution. **C** The drift term  $A(w)$  determines the net weight change. Here small weights increase as  $A(w)$  is positive whereas a negative  $A(w)$  reduces large weights thus leading to a unimodal weight distribution

where  $c_p$  and  $c_d$  play the role of  $A_{LTP}$  and  $w_d = A_{LTD}$  respectively. Here a larger weight is depressed more than a smaller weight. This creates a negative slope for  $A(w)$

$$A(w) \propto -c_d w + (1 + w/W_{tot})c_p \quad (2)$$

Thus  $A(0)$  is positive and  $A(w)$  decreases with increasing  $w$ . The steady-state synaptic weight distribution becomes unimodal (Fig. 6 left).

However, if the update rule for the potentiation is also multiplicative (case b) then the resulting synaptic weight distribution becomes bimodal. Introducing  $w_p = c_p w$  gives

$$A(w) \propto -c_d w + w(1 + w/W_{tot})c_p \quad (3)$$

and  $A(0) = 0$  and  $A(w) > 0$  for large  $w$ . Inspection of  $\partial A(w)/\partial w$  reveals that  $A(w)$  becomes negative for moderate  $w$  provided that depression is somewhat stronger than potentiation. Under these conditions the steady-state distribution is bimodal. This explains the simulations in Fig. 4B since the multiplicative update in LTP rewards strong synapses to become even larger. When the update rule is multiplicative only for LTP but not for LTD (case c) we find that

$$A(w) \propto -c_d + w(1 + w/W_{tot})c_p \quad (4)$$

while  $A(0) = -c_d p_d$  and  $A(w) > 0$  for large  $w$ . Here  $A(w) = 0$  when  $w \approx c_d/c_p$ . Note that when depression is sufficiently strong compared to potentiation, the synaptic weight distribution becomes bimodal.

Song et al. (2000) simulated the situation when the synaptic weights were updated with an *additive* amount as  $w_p = A_{LTP}$  and  $w_d = -A_{LTD}$ . The existence of steady-state distribution requires a maximal synaptic weight  $w_{max}$  when the update rules are additive. Now, letting  $\alpha = A_{LTD}/A_{LTP}$ , we can then write the drift term  $A(w)$  as

$$A(w) \propto w/W_{tot} - 1 + 1/\alpha \quad (5)$$

When  $w_{max} > (1 - 1/\alpha)W_{tot}$ , the synaptic weight distribution becomes bimodal. The positive slope of  $A(w)$  means that the  $w$  value at which  $A(w) = 0$  becomes unstable will increase or decrease respectively depending on whether a given synapse is larger or smaller than that particular  $w$  value. The synaptic steady-state weight distribution will therefore be bimodal. The analogy with the potential explains why the system takes so long to equilibrate in this case: some synapses have to climb over a barrier to reach their steady-state (Fig. 6 right).

In the additive version of the STDP rule both the synaptic competition and rate normalization are sensitive to parameter changes and a delicate balance between depression and potentiation is required. Small changes in input rates can easily create a unimodal synaptic weight distribution near zero or maximal weight. One approach to fix this problem is to place a particularly sensitive parameter,  $\alpha$ , under dynamic control. Self-consistent Fokker-Planck calculations can be used to determine how the asymmetry in the learning ratio,  $\alpha$ , should depend on presynaptic firing rate in order to produce a given neuronal input-output

relationship. This approach has been used to derive constraints for an adaptive STDP rule where the parameter  $\alpha$  is under activity-dependent control (Tegnér and Kepecs 2002a).

Whether a bimodal distribution and/or rate normalization occurs is also highly dependent on the particular shape of the learning function. For example, in weakly electric fish an anti-Hebbian inverted STDP has been found (Bell et al. 1997). The stability of this learning rule was first described by Roberts (1999). Cateau et al. (2002) have developed an extension to the Fokker-Planck method presented here where the drift term  $A(w)$  can be calculated for arbitrary window functions. Note, however, that the Fokker-Planck analysis presented here was developed for feed-forward models and does not seem feasible when recurrent network effects are included.

*Acknowledgements.* We are very grateful to Larry Abbott for discussions and comments on the manuscript. We thank the W.M. Keck Foundation for support. A.K. and M.v.R. were supported by the Sloan-Swartz Foundation. A.K. was supported by the NINDS grant RO1-27337, S.S. by a Howard Hughes Predoctoral Fellowship and J.T. by the Wennergren Foundation, the Foundation for Strategic Research, the Ake Wibergs Foundation, and The Swedish Royal Academy for Science.

## References

- Abbott LF, Blum KI (1996) Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex* 6: 406–416
- Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. *Nature Neurosci* 3: 1178–1183
- Abbott LF, Song S (1999) Temporally asymmetric Hebbian learning, spike-timing and neuronal response variability. In: Solla SA; Leen TK; Muller KR (eds) *Adv Neural Inf Process Syst*, vol 11. MIT Press, Cambridge, MA pp 69–75
- Ahissar E, Abeles M, Ahissar M, Haidarliu S, Vaadia E (1998) Hebbian-like functional plasticity in the auditory cortex of the behaving monkey. *Neuropharmacology* 37: 633–655
- Barrionuevo G, Brown TH (1983) Associative long-term potentiation in hippocampal slices. *Proc Natl Acad Sci USA* 80: 7347–7351
- Bell CC, Han VZ, Sugawara Y, Grant K (1997) Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* 387: 278–281
- Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike-timing, synaptic strength, and postsynaptic cell type. *J Neurosci* 18: 10464–10472
- Bienenstock EL, Cooper LN, Munro PW (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci* 2: 32–48
- Blais BS, Shouval HZ, Cooper LN (1999) The role of presynaptic activity in monocular deprivation: comparison of homosynaptic and heterosynaptic mechanisms. *PNAS* 96: 1083–1087
- Blakemore C, Van Sluyters CV, Movshon JA (1976) Synaptic competition in the kitten's visual cortex. *Cold Spring Harb Symp Quant Biol* 40: 601–609
- Blum KI, Abbott LF (1996) A model of spatial map formation in the hippocampus of the rat. *Neural Comput* 8: 85–93
- Cateau H, Kitano K, Fukait T (2002) An accurate and widely applicable method to determine the distribution of synaptic strengths formed by spike-timing-dependent learning. *Neurocomput* 44–46

- Debanne D, Gahwiler BH, Thompson SM (1994) Asynchronous pre- and postsynaptic activity induces associative long-term depression in area c1 of the rat hippocampus in vitro. *Proc Natl Acad Sci, USA* 91: 1148–1152
- Dudek SM, Bear MF (1992) Homosynaptic long-term depression in area ca1 of hippocampus and effects of n-methyl-d-aspartate receptor blockade. *Proc Natl Acad Sci, USA* 89: 4363–4367
- Egger V, Feldmeyer D, Sakmann B (1999) Coincidence detection and changes of synaptic efficacy in spiny stellate neurons in rat barrel cortex. *Nat Neurosci* 2: 1098–1105
- Feldman DE (2000) Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex. *Neuron* 27: 45–56
- Froemke RC, Dan Y (2002) Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* 416: 433–438
- Gütig, Robert and Aharonov, Ranit and Rotter, Stefan and Aertsen, Ad and Sompolinsky, Haim (2001) Generalized Synaptic Updating in Temporally Asymmetric Hebbian Learning. *Proceedings of the 28th Göttingen Neurobiology Conference*. p. 254. Stuttgart: Thieme
- Gütig R, Aharonov R, Rotter S, Sompolinsky H (2002) Learning Input Correlations Through Non-Linear Temporally Asymmetric Hebbian Plasticity. Submitted
- Garraghty PE, Sur M, Sherman SM (1986) Role of competitive interactions in the postnatal development of x and y retinogeniculate axons. *J Comp Neurol* 251: 216–239
- Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383: 76–81
- Gustafsson B, Wigstrom H, Abraham WC, Huang YY (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *J Neurosci* 7: 774–780
- Han VZ, Grant K, Bell CC (2000) Reversible associative depression and non-associative potentiation at a parallel fiber synapse. *Neuron* 26: 611–622
- Harris AE, Ermentrout GB, Small SL (1997) A model of ocular dominance column development by competition for trophic factor. *Proc Natl Acad Sci, USA* 94: 9944–9949
- Holmgren CD, Zilberter Y (2001) Coincident spiking activity induces long-term changes in inhibition of neocortical pyramidal cells. *J Neurosci* 21: 8270–8277
- Horn D, Levy N, Meilijson I, Ruppin E (2002) Distributed synchrony of spiking neurons in a Hebbian cell assembly. In: Leen TK; Solla SA; Muller KR (eds) *Adv Neural Inf Process Syst* 12, MIT Press Cambridge, MA, pp 129–135
- Kempter R, Gerstner W, van Hemmen JL (2001) Intrinsic stabilization of output rates by spike-based Hebbian learning. *Neural Comput* 13: 2709–2742
- Kistler WM, van Hemmen JL (2000) Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic action potentials. *Neural Comput* 12: 385–405
- Lara R, di Prisco GV (1983) Synaptic competition in developmental binocular plasticity. *J Theor Biol* 101: 503–527
- LeMasson G, Marder E, Abbott LF (1993) Activity-dependent regulation of conductances in model neurons. *Science* 259: 1915–1917
- Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *J Neurosci* 4: 791–797
- Lisman JE (1989) A mechanism for the hebb and the anti-hebb processes underlying learning and memory. *PNAS* 86: 9574–9578
- Lisman JE, Harris KM (1993) Quantal analysis and synaptic anatomy—integrating two views of hippocampal plasticity. *Trends Neurosci* 16: 141–147
- Magee JC, Johnston D (1997) A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* 275: 209–213
- Malinow R, Miller JP (1986) Postsynaptic hyperpolarization during conditioning reversibly blocks induction of long-term potentiation. *Nature* 320: 529–530
- Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275: 213–215
- Mehta MR, Barnes CA, McNaughton BL (1997) Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc Natl Acad Sci, USA* 94: 8918–8921
- Mehta MR, Wilson MA (2000) From hippocampus to v1: effect of LTP on spatio-temporal dynamics of receptive fields. *Neurocomput* 32: 905–911
- Mehta MR, Quirk MC, Wilson MA (2000) Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron* 25: 791–715
- Miller KD (1996) Synaptic economics. *Neuron* 7: 371–374
- Miller KD, Keller JB, Stryker MP (1989) Ocular dominance column development: analysis and simulation. *Science* 245: 605–615
- Minai AA, Levy WB (1993) Sequence learning in a single trial. In: *INNS World Congress of Neural Netw II* pp 505–508
- Montgomery JM, Madison DV (2002) State-dependent heterogeneity in synaptic depression between pyramidal cell pairs. *Neuron* 33: 765–777
- Petersen CC, Malenka RC, Nicoll RA, Hopfield JJ (1998) All-or-none potentiation at CA3-CA1 synapses. *Proc Natl Acad Sci, USA* 95: 4732–4737
- Rao RPN, Sejnowski TJ (2000) Predictive sequence learning in recurrent neocortical circuits. In: Solla SA; Leen TK; Muller KR (eds) *Adv Neural Inf Process Syst*, vol 12. MIT Press, Cambridge, MA, pp 164–170
- Rao RPN, Sejnowski TJ (2001) Spike-timing-dependent Hebbian plasticity as temporal difference learning. *Neural Comput* 13: 2221–2237
- Rittenhouse CD, Shouval HZ, Paradiso MA, Bear MF (1999) Monocular deprivation induces homosynaptic long-term depression in visual cortex. *Nature* 397: 347–350
- Roberts PD (1999) Computational consequences of temporally asymmetric learning rules: I. differential Hebbian learning. *J Comp Neurosci* 7: 235–246
- Rosenmund C, Feltz A, Westbrook GL (1995) Calcium-dependent inactivation of synaptic NMDA receptors in hippocampal neurons. *J Neurophysiol* 73: 427–430
- Rubin J, Lee DL, Sompolinsky H (2001) Equilibrium properties of temporally asymmetric Hebbian plasticity. *Phys Rev Lett* 86: 364–367
- Schuett S, Bonhoeffer T, Hubener M (2001) Pairing-induced changes of orientation maps in cat visual cortex. *Neuron* 32: 315–323
- Sejnowski TJ (1999) The book of Hebb. *Neuron* 24: 773–776
- Senn W, Buchs NJ (2002) Spike-based synaptic plasticity and the emergence of direction selective simple cells. *J Comput Neurosci*
- Senn W, Markram H, Tsodyks M (2001) An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike-timing. *Neural Comput* 13: 35–67
- Seung S, Xie X (2000) Spike-based learning rules and stabilization of persistent neural activity in recurrent neocortical circuits. In: Solla SA; Leen TK; Muller KR (eds) *Adv Neural Inf Process Syst*, vol 12., MIT Press, Cambridge, MA, pp 2–3
- Shepherd GM, Harris KM (1998) Three-dimensional structure and composition of CA3→CA1 axons in rat hippocampal slices: implications for presynaptic connectivity and compartmentalization. *J Neurosci* 18: 8300–8310
- Sjöström PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32: 1149–1164
- Song S (2002) Competitive Hebbian learning through spike-timing dependent plasticity. PhD thesis, Brandeis University, Waltham, MA

- Song S, Abbott LF (2001) Column and map development and cortical re-mapping through spike-timing dependent plasticity. *Neuron* 32: 339–350
- Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neurosci* 3: 919–926
- Tegnér J, Kepecs A (2002) Why neuronal dynamics should control synaptic learning rules. In: Dietterich TG; Becker S; Ghahramani Z (eds) *Adv Neural Inf Process Syst 14*, MIT Press, Cambridge, MA
- Tegnér J, Kepecs A (2002) Sliding-threshold correlational learning follows from spike-timing-dependent plasticity. *Soc Neurosci Abstracts* 28
- Tegnér J, Kepecs A (2002) An adaptive spike-timing-dependent plasticity rule. *Neurocomput* 44–46
- Trachtenberg JT, Stryker MP (2001) Rapid anatomical plasticity of horizontal connections in the developing visual cortex. *J Neurosci* 21: 3476–3482
- Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391: 892–896
- Turrigiano GG, Nelson SB (2000) Hebb and homeostasis in neuronal plasticity. *Curr Opin Neurobiol* 10: 358–364
- Umekiya M, Chen N, Raymond LA, Murphy TH (2001) A calcium-dependent feedback mechanism participates in shaping single NMDA miniature EPSCs. *J Neurosci* 21: 1–9
- van Ooyen A (2001) Competition in the development of nerve connections: a review of models. *Network* 12: R1–47
- van Rossum MCW, Bi GQ, Turrigiano GG (2000) Stable Hebbian learning from spike-timing-dependent plasticity. *J Neurosci* 20: 8812–8821
- von der Malsburg C (1973) Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik* 14: 85–100
- von der Malsburg C (1979) Development of ocularity domains and growth behaviour of axon terminals. *Biol Cybern* 32: 49–62
- Watanabe S, Hoffman DA, Migliore D, Johnston M (2002) Dendritic  $K^+$  channels contribute to spike-timing dependent long-term potentiation in hippocampal pyramidal neurons. *Proc Natl Acad Sci, USA* 99: 8366–8371
- Yao H, Dan Y (2001) Stimulus timing-dependent plasticity in cortical processing of orientation. *Neuron* 32: 315–323
- Zhang LI, Tao HW, Holt CE, Harris WA, Poo MM (1998) A critical window for cooperation and competition among developing retinotectal synapses. *Nature* 395: 37–44