Stable Hebbian Learning from Spike Timing-Dependent Plasticity

M. C. W. van Rossum,[1] **G. Q. Bi,**[2] **and G. G. Turrigiano**[1]

[1]*Brandeis University, Department of Biology, Waltham, Massachusetts 02454-9110, and* [2]*University of California at San Diego, Department of Biology, La Jolla, California 92093-0357*

We explore a synaptic plasticity model that incorporates recent findings that potentiation and depression can be induced by precisely timed pairs of synaptic events and postsynaptic spikes. In addition we include the observation that strong synapses undergo relatively less potentiation than weak synapses, whereas depression is independent of synaptic strength. After random stimulation, the synaptic weights reach an equilibrium distribution which is stable, unimodal, and has positive skew. This weight distribution compares favorably to the distributions of quantal amplitudes and of receptor number observed experimentally in central neurons and contrasts to the distribution found in plasticity models without size-dependent potentiation. Also in contrast to those models, which show strong competition between the synapses, stable plasticity is achieved with little competition. Instead, competition can be introduced by including a separate mechanism that scales synaptic strengths multiplicatively as a function of postsynaptic activity. In this model, synaptic weights change in proportion to how correlated they are with other inputs onto the same postsynaptic neuron. These results indicate that stable correlation-based plasticity can be achieved without introducing competition, suggesting that plasticity and competition need not coexist in all circuits or at all developmental stages.

Key words: Hebbian plasticity; synaptic weights; synaptic competition; activity-dependent scaling; temporal learning; stochastic approaches

Changes in the synaptic connections between neurons are widely believed to contribute to memory storage, and the activity-dependent development of neuronal networks. These changes are thought to occur through correlation-based, or Hebbian, plasticity, but the precise plasticity rules remain unclear. In general, such learning rules should allow synaptic inputs to change in strength, depending on their correlation with postsynaptic firing or with the activity of other inputs (Sejnowksi, 1977). Second, they should generate a stable distribution of synaptic weights. Finally, to account for activity-dependent development, they should generate competition between the inputs of a neuron, so that strengthening some inputs weakens others (Shatz, 1990; Miller, 1996).

Unconstrained Hebbian plasticity does not generate a stable weight distribution, because once an input is strengthened its correlation with the postsynaptic activity increases. This leads to further potentiation and the synaptic weights grow to infinitely large values. Analogously, once depressed, synapses decrease to zero. These problems are usually fixed by constraining the learning rules, for instance by keeping the sum of weights constant. Alternatively, the postsynaptic activity can be used to adjust the threshold for potentiation (Bienenstock et al., 1982; Kirkwood et al., 1996), to regulate neuronal excitability (Desai et al., 1999), or to scale all of a neuron's synaptic weights (Turrigiano et al., 1998). All these mechanisms stabilize postsynaptic activity and introduce competition, but the choice of constraint influences strongly the behavior of the model (Miller and MacKay, 1994).

Recently, plasticity known as spike timing-dependent plasticity (STDP) has been observed (Bell et al., 1997; Markram et al., 1997; Bi and Poo, 1998). If a synaptic event precedes the postsynaptic spike, the synapse is potentiated. If it follows the postsynaptic spike, the synapse is depressed. STDP rules have been implemented in several modeling studies (Blum and Abbott, 1996; Ger-

stner et al., 1996; Eurich et al., 1999; Kistler and van Hemmen, 2000). Interestingly, because of strong competition between inputs, the postsynaptic firing rate in these models is independent of synaptic input rate (Song and Abbott, 2000). However, as commonly implemented, these learning rules are unstable and require hard bounds on the synaptic weights. The synaptic weights are driven to these bounds and obtain a bimodal distribution, which is unlikely to reflect the weight distribution in biological neurons.

Here we present an intrinsically stable STDP learning rule. It incorporates the experimental observation that potentiation is weaker for strong synapses (Debanne et al., 1996, 1999; Bi and Poo, 1998). This learning rule generates a stable, unimodal, positively skewed distribution of synaptic weights that closely resembles the distribution of quantal amplitudes measured from central neurons (Bekkers et al., 1990; Turrigiano et al., 1998). The weights depend on their correlation with other inputs, so that learning occurs through cooperation between inputs. Finally, competition is almost absent, and must be introduced independently by implementing activity-dependent synaptic scaling (Turrigiano et al., 1998; O'Brien et al., 1998; Turrigiano, 1999). The scaling does not change the shape or the stability of the weight distribution.

MATERIALS AND METHODS

Experimental foundation of the plasticity rules. In this section we extract the parameters of the model from the experimental data presented in Bi and Poo (1998). Synapses can be potentiated and depressed by pairing a synaptic event with a postsynaptic spike. If the synaptic event occurs before the postsynaptic spike, the synapse will be potentiated; if the postsynaptic spike precedes the synaptic event the synapse will be depressed (Markram et al., 1997; Zhang et al., 1998; Bi and Poo, 1998). The amount of conductance change decreases approximately exponentially with the time difference between the synaptic event and the postsynaptic spike, δt. Such a synaptic modification window is illustrated in Figure 1. Although in Bi and Poo (1998) the time window was slightly asymmetric (the time constant for the exponential function was 34 ± 13 msec for depression and 17 ± 9 msec for potentiation), this asymmetry is not essential, and we use in our model the same time constant for depression and potentiation, $\tau_{STDP} = 20$ msec.

There is experimental evidence that the amount of change also depends on the initial synaptic size (Debanne et al., 1996, 1999; Bi and Poo, 1998). The relative amount of depression is independent of synaptic weight, whereas the relative amount of potentiation decreases for stronger synapses (Figs. 1B, 2A). We describe this by assuming that the amount of potentiation is inversely proportional to the weight (Kistler and van Hemmen, 2000).

In the data of Bi and Poo (1998), synaptic events and postsynaptic spikes

were paired 60 times. To deduce how much the synapse changes after one single pairing, we divide the weight change by 60, implicitly assuming independence between the pairing events. Denoting with w the synaptic conductance in Siemens, we describe potentiation as $w \rightarrow w + w_p$, and depression as $w \rightarrow w + w_d$. The plasticity rules are:

$$w_p = c_p e^{-\delta t/\tau_{STDP}}$$

$$w_d = -c_d w e^{\delta t/\tau_{STDP}},$$

where c_d is the average amount of relative depression after one pairing, $c_d = 0.003$; and c_p is the average amount of potentiation after one pairing, $c_p = 7$ pS.

It will turn out that fluctuations in the amount of depression and potentiation are important. That is, the amount of conductance change is a noisy quantity. To describe this, we first tried an additive noise model in which a random conductance value was added to the weights after every pairing. For this model, weak synapses show strong fluctuations, and strong synapses show small fluctuations (Fig. 2*B*). Because this does not seem to correspond to the data, we rejected this model. Rather, the fluctuations in the relative change seem roughly constant, which implies that the fluctuations are multiplicative (Fig. 2*C*). We thus arrive at the following plasticity rules:

$$w_p = (c_p + \nu w)e^{-\delta t/\tau_{STDP}} \qquad (1)$$

$$w_d = (-c_d w + \nu w)e^{\delta t/\tau_{STDP}}, \qquad (2)$$

where ν is a Gaussian random variable with zero mean, its SD, σ, is extracted from the data. Assuming that no other noise sources were present in the measurement of the conductance changes, we find $\sigma = 0.015$. Equations 1 and 2 are the basis for this study.

The experimental data on the size dependence of the potentiation can also be fitted with a straight line in Figure 2*A* as was done in Bi and Poo (1998). That is, $w_p = -c_p' w \log(w/w_{max})$, where w_{max} is the conductance above which a potentiation protocol actually leads to depression of the synapse. For this plasticity rule runaway learning toward infinite weights is surely impossible. The equilibrium distribution that results from these rules looks similar to the one in Figure 3*A* in the central region, but a smaller second peak appears in the distribution around $w = 0$. It is not clear whether this has a biological analog. On the one hand it could be an artifact caused by our limited experimental knowledge of Equation 1 at small w. Alternatively, it might be a useful biological mechanism corresponding to silent synapses or weak synapses that can be pruned.

Simulation details. To analyze the consequences of the above plasticity rules we simulate a single cell receiving random inputs. We use a leaky integrate and fire neuron with: 100 MΩ input resistance, 20 msec time constant, −60 mV resting potential, and −50 mV firing threshold. After firing the membrane potential resets to the resting potential. The neuron receives input from 25 inhibitory and 100 excitatory synapses. The inhibitory synapses have a reversal potential of −70 mV, and a time constant of 5 msec. The inhibitory synapses are not plastic but are fixed at a conductance of 2000 pS and are stimulated with Poisson trains of 20 Hz. The excitatory synapses have a reversal potential of 0 mV and a time constant of 5 msec. The excitatory synapses also receive Poisson input. In some cases correlation across synaptic inputs is introduced. Correlation is implemented by randomly distributing N Poisson trains among the inputs. Every time step the Poisson trains are redistributed (Destexhe and Paré, 1999). For every synaptic event there is a chance $1/N$, that it is shared by another synapse. This yields a cross-correlation coefficient between trains of $C(\Delta t) = 1/N\delta(\Delta t)$.

Plasticity was implemented as follows: when a synaptic event occurs after a postsynaptic spike, the synapse is depressed according to Equation 2. We assume that all plasticity events are independent, but one also needs to specify the behavior when there are multiple synaptic events at the same input. We assume that at a given synapse, only the first synaptic event after a given spike depresses the synapse; subsequent synaptic events do not depress the synapse further before another postsynaptic spike occurs. Potentiation was implemented analogously: only the first postsynaptic spike after the synaptic event leads to potentiation of the synapse, according to Equation 1. In the simulations we use $c_p = 1$ pS.

Alternatively, one can assume that all pairing events cause depression and potentiation, as was assumed in most other studies (Kempter et al., 1999; Song et al., 2000). For Poisson trains one can show that this only changes the rate of change, the equilibrium state is the same in both implementations. We choose our implementation because it is consistent with the calculations presented in the appendix.

The plasticity rules are independent of the presynaptic frequency, which is likely to be an over-simplification (Markram et al., 1997). However, this assumption is not essential for our argument: frequency-dependent potentiation or depression would shift the mean synaptic weight depending on stimulation frequency, but stability and competition would not be altered.

The parameters used in the simulation reflect a typical cultured neuron; such neurons have few synapses with large conductances. In slice or in vivo the number of inputs runs in the thousands, and we expect that the plasticity rules will yield synapses with correspondingly smaller mean weights, indeed, such scaling could be accomplished by activity-dependent scaling (see below) and does not qualitatively change our results.

Activity-dependent scaling. Here we show how activity-dependent scaling is incorporated in the model. As the precise mechanism behind activity-dependent scaling is not known, we present only a possible implementation. Activity-dependent scaling is a mechanism that adjusts the synaptic weights to regulate the postsynaptic activity. The postsynaptic activity is measured with a slow-varying sensor, $a(t)$. It increases with every postsynaptic spike, and decays exponentially between spikes:

$$\tau \frac{da(t)}{dt} = -a(t) + \sum_i \delta(t - t_i),$$

where t_i are the spike times. The biological time constant of the activity sensor τ is unknown, but is expected to be slow, we use $\tau = 100$ sec. Activity-dependent scaling scales the weights to prevent too low or too high activity levels. The scaling is thought to be multiplicative and independent of presynaptic activity (Turrigiano et al., 1998; Turrigiano, 1999). A simple implementation would be to update the weights every time step according to:

$$\frac{dw(t)}{dt} = \beta w(t)[a_{goal} - a(t)],$$

where a_{goal} is the desired postsynaptic activity, set to 20 Hz, and β is a constant determining the strength of the scaling. This mechanism scales the synapses towards the activity goal, but because the plasticity rules Equations 1 and 2 also pull on the weights, in the end a residual deviation between the actual activity and its goal value remains. Such errors are prevented by using an "integral controller", (Riggs, 1970),

$$\frac{dw(t)}{dt} = \beta w(t)[a_{goal} - a(t)] + \gamma w(t) \int_0^t dt'[a_{goal} - a(t')], \qquad (3)$$

where γ is another constant. As the second term accumulates the error, this term will in the long run be dominant if the goal value is not reached. As a result the steady-state activity level will eventually become equal to its goal value.

As with any feedback system, oscillations readily occur. These oscillations can arise as follows: when the activity does not have the desired value, the weights are slowly adjusted, and the activity moves back to its desired value. However, as the activity sensor has a delay, the weights can be overcompensated and overshoot. This leads to oscillations in the activity. Because such oscillation are not known to occur, the parameters were adjusted to prevent them. Suitable parameter values can be calculated from control theory (Riggs, 1970). The parameters do not require a sensitive adjustment. In our simulations we use a strength $\beta = 4 \times 10^{-5}$/sec/Hz and γ is 10^{-7}/sec²/Hz. These values give slow scaling without oscillations. Our arguments would not change if oscillations would occur, because we are interested in the equilibrium.

The scaling can be incorporated into the weight evolution equation (see Appendix). The scaling shifts the point where potentiation and depression are balanced, thus adjusting the mean weight while approximately preserving the shape of the distribution, consistent with experimental observations (see Discussion).

RESULTS

Experimental foundation

We base our model on two experimental observations. The first is STDP: it has been observed that synapses can be potentiated and depressed by pairing a synaptic event with a postsynaptic spike. If the synaptic event occurs ~50 msec or less before the postsynaptic spike, the synapse will be potentiated; if the spike precedes the synaptic event, the synapse will be depressed (Markram et al., 1997; Bi and Poo, 1998). The amount of weight change is approximately exponential in the time between the synaptic event and the postsynaptic spike. The resulting synaptic modification window is plotted in Figure 1.

The second essential ingredient is that the amount of synaptic change depends on the synaptic size. For STDP protocols it was observed that the relative amount of depression is independent of the initial synaptic size, whereas relatively potentiation is larger for weak synapses than strong synapses. In Figure 2*A* we plot such data from experiments on cultured neurons (Bi and Poo, 1998). A similar observation was made in hippocampal slices (Debanne et al., 1996, 1999). Possibly closely related to this, it has been observed that the amount of potentiation and depression depends on the history of synaptic stimulation (Yang and Faber, 1991; Ngezahayo et al., 2000). Including the weight dependence in the plasticity rules has drastic consequences for the weight distribution, the stability of the plasticity, and synaptic competition.
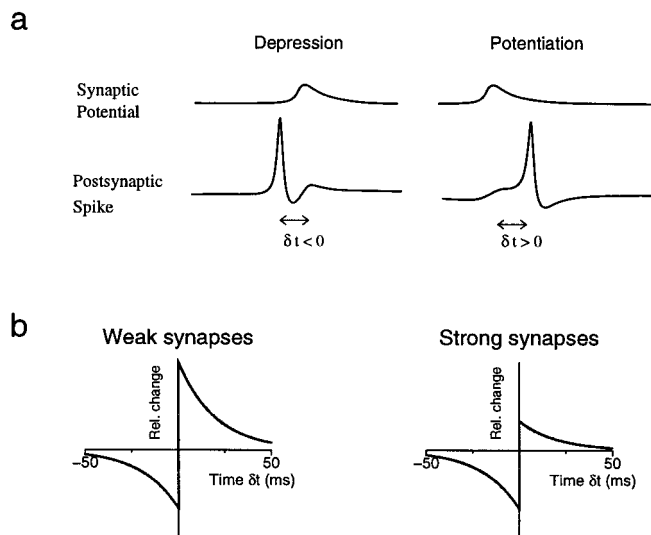
Figure 1.   Spike timing-dependent plasticity. *a*, Synapses are potentiated if the synaptic event precedes the postsynaptic spike. Synapses are depressed if the synaptic event follows the postsynaptic spike. *b*, The time window for synaptic modification. The relative amount of synaptic change is plotted versus the time difference between synaptic event and the postsynaptic spike. The amount of change falls off exponentially as the time difference increases. In addition, the amount of potentiation decreases for stronger synapses, whereas the relative amount of depression is independent of synaptic size.

Synaptic weight distribution after prolonged random stimulation

We ask how the synaptic weights of a neuron evolve when they are subject to the plasticity rules sketched in Figure 1. We first study the case when the neuron receives random synaptic input. Consider the distribution of synaptic weights, $P(w)$. A single bin in this distribution describes the probability that synapses have a weight w. The synapses in this bin are continuously potentiated and depressed because of ongoing coincidences of presynaptic and postsynaptic spikes. Because of the size dependence of the plasticity, strong synapses experience a net depression, whereas weak synapses experience a net potentiation. This confines the synaptic weights. After a while the distribution reaches an equilibrium, at which the individual synapses still change, but the distribution is stationary.

The above picture turns out to be correct in simulations. We simulate an integrate and fire neuron receiving random synaptic input. The presynaptic signal is provided by 100 excitatory synapses stimulated with independent Poisson trains. The synapses are subject to the plasticity rules sketched in Figure 1 and given in Equations 1 and 2. The parameters of the plasticity rules are based on physiological data (see Materials and Methods). As fluctuations in the amount of synaptic change induced by potentiation and depression are important for the shape of the resulting weight distribution, a multiplicative noise model is part of the plasticity rules.

The distribution of synaptic weights after prolonged stimulation is shown in Figure 3*A*. The analysis in the Appendix shows explicitly that this equilibrium distribution is independent of the initial distribution. The resulting distribution is very stable and does not require any fine tuning of parameters. The distribution is unimodal and has a positive skew. This is similar to the distribution found in quantal synaptic current measurements and synaptic staining studies (O'Brien et al., 1998; Turrigiano et al., 1998). For comparison we plot in Figure 3*B* a distribution of quantal amplitudes, as observed from a single cultured cortical pyramidal neuron held at -70 mV (Turrigiano et al., 1998). In the Discussion, we expand on their similarity.

In the equilibrium state, the synaptic weights continuously make small random jumps, but their movement is confined. The mean
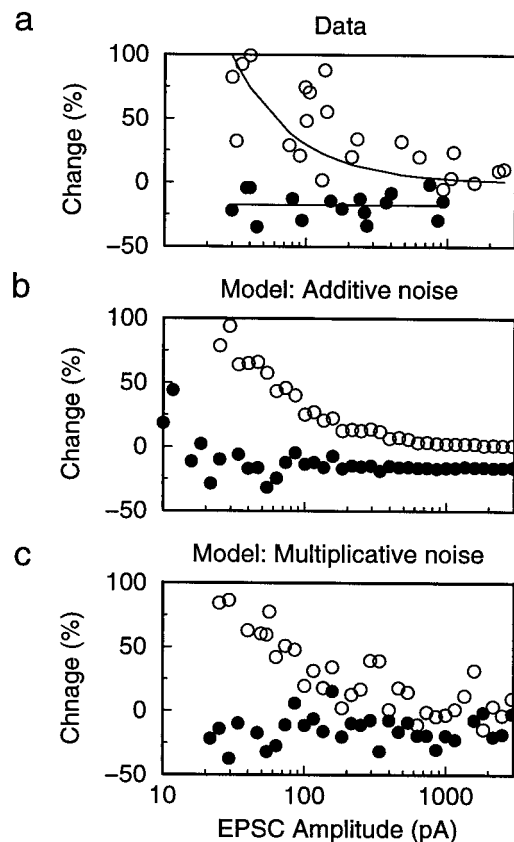

Figure 2.   The weight dependence of the STDP conductance change. *a*, The data from Bi and Poo (1998) describing the relative synaptic change as a function of the initial synaptic size. Potentiating (*open circles*) and depressing (*filled circles*) pairings were repeated 60 times. The depression data are fitted to a constant; the potentiation data are inversely proportional to the synaptic size. *b*, Additive noise model: the data is simulated by applying the plasticity rule 60 times. After every synaptic change a random conductance value is added. The random conductance is drawn from a Gaussian distribution with zero mean and SD of 8 pA. This description of the noise was rejected. *c*, Simulation of the data using a multiplicative noise model, in which the noise in the conductance change is weight dependent. Multiplicative noise gives a better description of the spread in the data.

synaptic weight is approximately located where the confining force vanishes, and the weight experiences no net depression or potentiation. There are two contributions to the confining force: (1) the probability to cause a postsynaptic spike increases linearly with the weight of the synapse. Strong synapses therefore have a larger probability of being potentiated, whereas the probability for being depressed is independent of synaptic strength (see Appendix). Thus, once potentiated, strong synapses have an even higher chance for subsequent potentiation. This is a destabilizing force, and if this were the only force present, weights would run off to infinity. (2) However, stronger synapses experience a smaller conductance change when potentiated. This constitutes the second force. Because potentiation decreases with increasing weight, but depression does not, this force is stabilizing. When the two forces are combined, the stabilizing force wins, and the stable distribution shown in Figure 3*A* results.

The analytical treatment presented in the Appendix gives an accurate description of the distribution found in the simulations. It describes how the weights evolve by combining the weight dependence of potentiation and the probability that a synapse of given weight will be potentiated and depressed. This yields the weight distribution (Fig. 3*A*, *solid line*). The analysis also shows explicitly that the destabilizing force plays only a minor role. If this force is completely turned off, as is easily done in the analytical expressions, weak and strong synapses have an equal probability for potentiation. Yet, this hardly changes the shape of the distribution
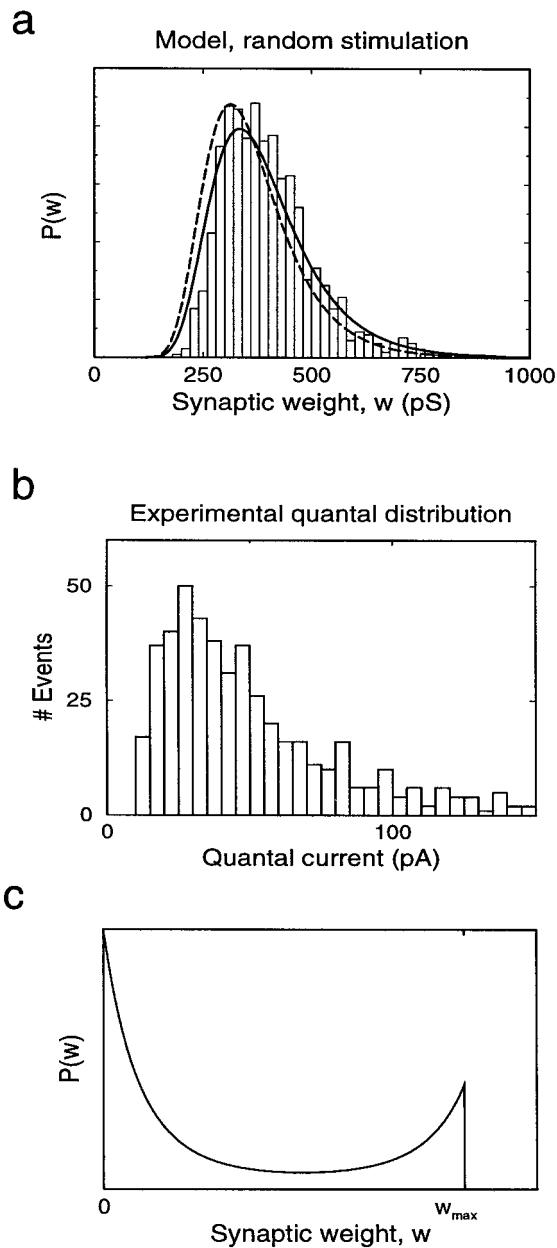
a

Model, random stimulation

b

Experimental quantal distribution

c

Figure 3. *a,* The equilibrium distribution of the synaptic weights of a neuron after prolonged synaptic stimulation with uncorrelated Poisson trains. *Histogram,* Distribution of weights from a simulation of an integrate and fire neuron. *Solid line,* Analytical prediction from Equation 14, with W_{tot} extracted from Figure 7. *Dashed line,* Analytical prediction when strong synapses do not have an enhanced probability for potentiation. Simulation parameters: 20 Hz input rate, postsynaptic firing rate ~25 Hz, weights were averaged over 10 runs. *b,* Experimental quantal amplitude distribution as observed from a single cultured cortical pyramidal neuron. *c,* When potentiation and depression do not depend on the weight, as was assumed in many other models, a bimodal weight distribution results. Limits on the minimal and maximal weight have to be imposed ($w_{min} = 0$ and w_{max}); the weights cluster at these limits.

(Fig. 3*A, dashed line*) indicating the minor role of the destabilizing force in our model.

Stability: comparison to other models

How does this model compare to other STDP models? In most other models (but see Kistler and van Hemmen, 2000), potentiation and depression change the synaptic weight by a fixed amount, independent of the synaptic weight (Blum and Abbott, 1996; Gerstner et al., 1996; Amarasingham and Levy, 1998; Kempter et al., 1999; Song et al., 2000). The typical shape of the weight distribu-

tion for such a model is shown in Figure 3*C.* Note that, depending on the parameters, the synapses split into two groups of either weak or strong synapses, despite the absence of any structure in the input.

The behavior can be understood from the force terms introduced above, which determine the net potentiation and depression that a certain weight experiences. Because here potentiation and depression have an identical weight dependence, the stabilizing force vanishes. Left is the destabilizing force. This small but important force causes strong synapses to get even larger as they will have a higher probability of inducing a spike. This positive feedback will cause the weights to run off to infinite values. Therefore, a hard limit on the maximal weight is required for these models. Similarly, weights below a certain threshold will be depressed till they hit the lower bound. In the Appendix we show how also for these models the synaptic weight distribution can be calculated.

In these models the weight distribution is sensitive to small perturbations of the parameters and the destabilizing force (see Appendix). In contrast, in our model the effect of the destabilizing force is small. The distribution is dominated by the differential weight dependence of potentiation and depression, overruling the positive feedback and stabilizing the distribution.

Correlated input potentiates synapses

Above we determined the weight distribution reached after long random stimulation. From a functional point of view this is a rather dull situation: no memory is stored, and all inputs obtain on average the same weight. The question arises how memory patterns are impressed and stored in the model. In contrast to conventional forms of long-term potentiation (LTP) and long-term depression (LTD), in our STDP scheme synapses are not strengthened by increasing their input rates. A different stimulation frequency changes the rate at which depression and potentiation occur, but will not effectively change the synaptic weight.

An effective way to store memories is to introduce correlation among inputs (Oja, 1982; Song et al., 2000). When inputs are correlated, the probability for potentiation is larger because synaptic events will often occur simultaneously and induce postsynaptic spikes. Weak synapses piggyback on the strong ones (Zhang et al., 1998). However, the probability for depression is unaltered (see Appendix). This is illustrated in Figure 4 where four groups of inputs with varying amount of correlation are presented to the neuron. Correlation shifts the distribution towards higher conductance values as the balance between potentiation and depression shifts. The shape of the distribution remains qualitatively the same, and the stability is maintained. The effect of the correlation is twofold: first, most of the postsynaptic spikes will be triggered by correlated inputs, and, second, these inputs will be potentiated. The mean weight is proportional to the correlation. In models without weight-dependent potentiation, inputs with correlations above a certain threshold obtain maximal weight, and inputs with less correlation obtain essentially zero weight. There is a sharp transition between the two.

Lack of competition

In many models of constrained Hebbian plasticity there is strong competition between synapses: enhancement of one synapse leads to depression of the other synapses (Miller, 1996; Song et al., 2000). In our model there is practically no competition: the synaptic weights are insensitive to changes in the other inputs. To demonstrate this, we simulate the following situation (Fig. 5*A*): the postsynaptic neuron receives two groups of inputs. Initially both groups are uncorrelated, and as a result both groups have identical mean weights. Next, strong correlation within the first group is introduced, and this potentiates these synapses as described above. The mean conductance of this group and the postsynaptic firing rate increase. If there were competition present between the synapses, this stimulation should lead to a reduction of the weight of
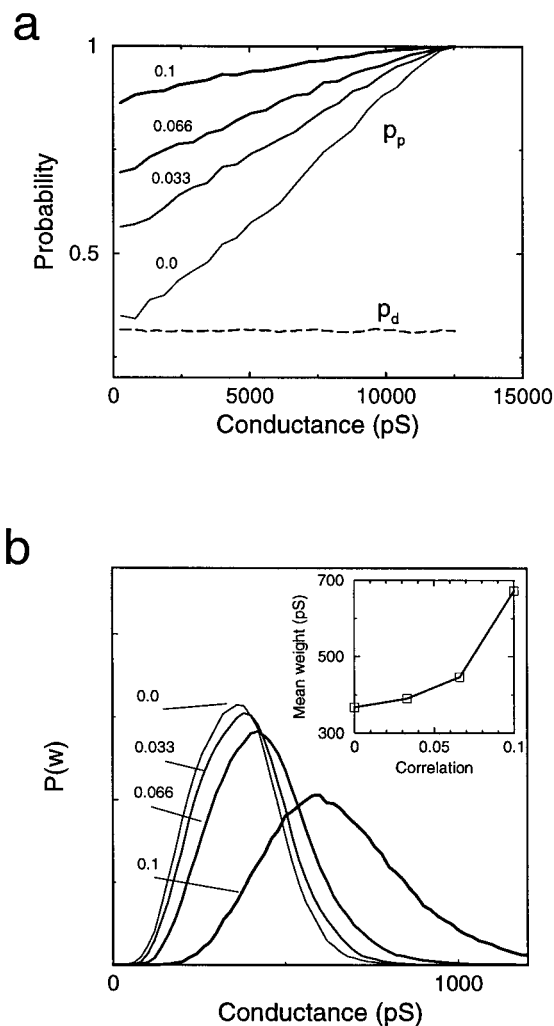
a

b

Figure 4. Effect of correlation in the inputs on the synaptic weights. The inputs consisted of four groups of 25 synapses having different amounts of correlation within the group (correlation coefficients: 0, 0.033, 0.066, 0.1). *a*, The probability for inducing potentiation, p_p and depression p_d vs. the weight. The probability for inducing potentiation is increased when correlations between inputs are present, whereas the probability for inducing depression is unaltered. The *labels* indicate the correlation coefficient. *b*, The weight distributions of the different groups. The different amounts of correlation lead to the coexistence of multiple weight distributions. The weights of the more strongly correlated groups are larger. The *inset* shows the mean conductance of the different groups as a function of the correlation.

synapses in the other group. The weights of the other group of synapses are, however, hardly affected, as is illustrated in Figure 5*A*. Thus, there is little competition.

In STDP models in which potentiation and depression are independent of the synaptic weight, there is strong competition. The reason is as follows. In STDP the potentiation mainly occurs if the input has caused the spike, in other words, the inputs compete for the postsynaptic spike. When one input starts driving the postsynaptic spikes and its weight increases, the other inputs will become less correlated with the postsynaptic spikes, and these inputs will effectively be depressed (see Appendix).

The competition in these models is so strong that there is a limited regime in which increasing the input rate is counteracted by the reduction of the synaptic weights, causing the postsynaptic firing frequency to be almost independent of the input rate (Song et al., 2000). In our model, the potentiation and depression of synapses is limited. As a result it is not very sensitive to changes in the total input. Competition and output rate normalization are virtually absent, and the output rate follows the input rate.
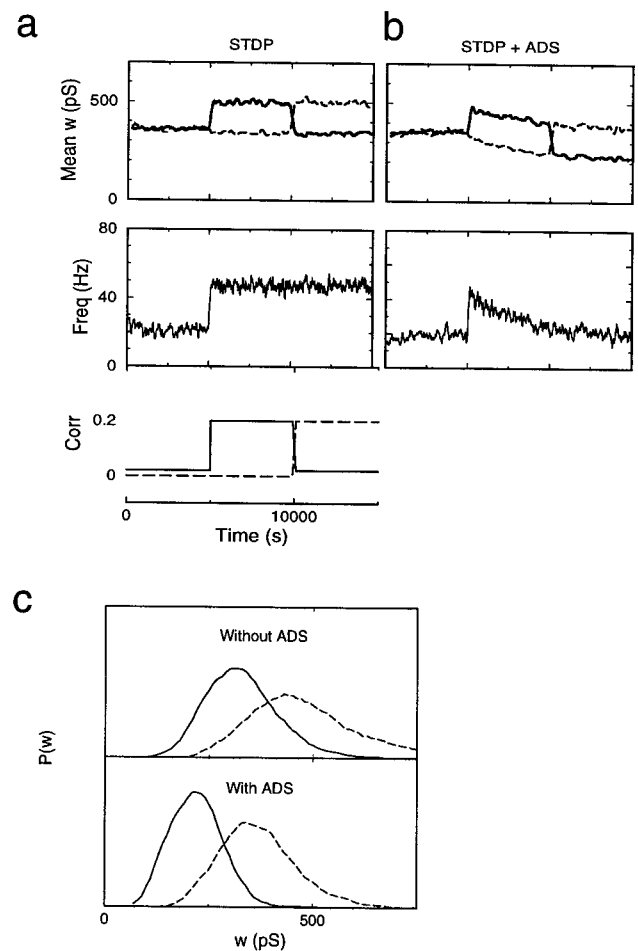
a   b

c

Figure 5. Competition between synaptic inputs and the effect of activity-dependent scaling (*ADS*). *a*, Behavior of model without ADS. *Bottom graph*, The neuron receives input from two sets of 50 synapses. Until time 5000 sec, both sets are uncorrelated. At 5000 sec the inputs within one set become strongly correlated, potentiating its weights. At 10,000 sec this set becomes again uncorrelated, whereas the other set becomes correlated, reversing the situation. *Middle graph*, The postsynaptic firing frequency jumps when the inputs become correlated. *Top graph*, The average synaptic weight for the two sets. The introduction of correlation potentiates the synapses, but changes in one group of synapses barely affect the other group. Competition is almost absent. *b*, Same situation but with activity-dependent scaling turned on. After the jump in firing rate the synapses are slowly scaled downward, until the activity is again at its goal value of 20 Hz. This introduces competition. Note the difference in time-scales between the slow competition and the much faster STDP. *c*, The corresponding weight distributions once equilibrium has been reached. The activity-dependent scaling scales the weights of both groups.

Activity-dependent scaling as a separate competition mechanism

The lack of competition in our model demonstrates that stable Hebbian learning is possible without competition. Nevertheless, competition is useful for developmental processes such as ocular dominance column plasticity, and output rate normalization is useful in situations when the input rate or the number of inputs undergoes large changes. Therefore we include activity-dependent scaling of synaptic weights in the model. Activity-dependent scaling is a homeostatic mechanism which, in reaction to changes in the postsynaptic activity, scales all synapses in an effort to keep the activity of the neuron within bounds. The scaling is multiplicative and does not seem to depend on presynaptic spike activity (Turrigiano et al., 1998). To implement activity-dependent scaling, we introduce a slow-varying sensor of activity. The weights are multiplicatively scaled if the readout of the activity sensor differs from some preset goal value (see Materials and Methods).

The scaling mechanism introduces competition between the

synapses (Fig. 5*B*). This is expected: if one synapse is potentiated, the postsynaptic activity rises, and the activity-dependent scaling kicks in to reduce all synaptic weights. The scaling works on long time scales, and in the end the goal level of activity is maintained. The shape of the weight distribution and its stability are not affected by the scaling (Fig. 5*C*). The competition is thus separated from the STDP.

Note that this additional plasticity rule updates the weights independent of the presynaptic rate, in contrast to the STDP. Thus, if there are two sets of synaptic inputs, one with a low rate and one with a high rate, the weights of the low rate inputs will mainly be governed by the activity-dependent scaling, whereas the high rate inputs will be ruled by the STDP.

DISCUSSION

Despite the importance of correlation-based plasticity in learning and development, the exact nature of the learning rules that operate in biological networks remains unclear. Here we have shown that a learning rule based closely on experimental data allows inputs to change in strength as a function of correlation, while generating and maintaining a stable distribution of synaptic weights. We use an STDP learning rule in which potentiation occurs when a postsynaptic spike follows a synaptic event, and depression occurs if a postsynaptic spikes precedes a synaptic event. In addition, this rule incorporates the experimental observation that the amount of potentiation decreases as the synapse strengthens (Debanne et al., 1996, 1999; Bi and Poo, 1998). In this weight-dependent STDP, the synaptic weights evolve into a unimodal, positively skewed distribution that closely resembles experimentally measured distributions of quantal amplitudes (Turrigiano et al., 1998) and of receptor number (O'Brien et al., 1998). The introduction of correlations between inputs increases synaptic strengths, but does not effect the shape and stability of the weight distribution. Weight-dependent STDP is intrinsically stable without requiring artificial constraints upon synaptic strengths.

The cause for instability in rate-based plasticity models is a destabilizing mechanism similar to the one in the STDP models: if a synapse is potentiated, the larger synapse causes a higher postsynaptic activity, which in turn potentiates the synapse even further. Weight-dependent potentiation could probably solve the problem of runaway learning for conventional LTP and LTD as well.

The shape of the synaptic weight distribution can be characterized by its mean, SD, and skew. In our model, the mean synaptic weight is determined by the balance point between potentiation and depression. Our analysis shows that this balance point is itself determined by two competing forces, one stabilizing and the other destabilizing. Because stronger synapses are more likely to evoke a postsynaptic spike, they are also more likely to be potentiated than weak synapses. This generates a destabilizing force that pushes synaptic strengths towards higher values. In models without the weight dependence of potentiation, this destabilizing force will tend to push synapses all the way to their upper and lower bounds. In our model, this destabilizing force is balanced by a reduction of the potentiation as synaptic weights increase. Because the amount of depression stays constant, for strong synapses depression will be larger than potentiation. This provides a brake on synaptic strength, constraining the weights of the synapses at central values.

The width of the synaptic weight distribution is strongly influenced by variations in the amount of potentiation and depression for different pairings of synaptic events and postsynaptic spikes. To make the model as realistic as possible, the magnitude of these fluctuations was extracted from the experimental data of Bi and Poo (1998). The magnitude may be overestimated because we assumed that all the measured noise arose from trial-to-trial fluctuations in the amount of potentiation or depression. Without this noise the simulated synaptic weight distribution has the same shape and overall behavior but is considerably narrower. Other factors could also contribute to a widening of the distribution, such as the presence of groups of inputs with different correlation levels (Fig. 4). The noise widens the weight distribution to values similar to those measured for quantal amplitudes, but many other factors could contribute to the width of these measured distributions. For example, the distribution of quantal amplitudes can be widened because of cable filtering (Spruston et al., 1993; Forti et al., 1997) and fluctuations in the transmitter content of vesicles (Frerking et al., 1995; Liu et al., 1999).

Another approach to assess the synaptic conductance distribution in central neurons is by using immunohistochemical methods to quantify the staining intensity of synaptic receptors. Using this method, the observed distributions of receptor staining are also unimodal and positively skewed (Nusser et al., 1997; O'Brien et al., 1998). As in our model, the shape of the distributions of both quantal amplitudes (Turrigiano et al., 1998; O'Brien et al., 1998) and of receptor staining (O'Brien et al., 1998) is preserved when synaptic strengths are scaled up or down in response to changes in activity.

A feature of STDP learning rules, with or without a weight dependence, is that inputs are potentiated as a function of their correlation on short time scales. This is because spikes in the postsynaptic neuron are chiefly caused by inputs correlated on short time scales. This contrasts with conventional rate-based Hebbian models in which the stimulation frequency determines which synapses get potentiated and in which short time scale correlations are not essential. Here, however, when precise timing does matter, correlations on short time scales are essential (Gerstner et al., 1996; Zhang et al., 1998; Kistler and van Hemmen, 2000; Song et al., 2000) as has been suggested for learning and memory (von der Malsburg, 1981).

An important feature of activity-dependent development in some CNS regions is competition between inputs onto a postsynaptic neuron (Shatz, 1990; Miller, 1996). Such competition allows some inputs to be retained, whereas other inputs are lost. In rate-based Hebbian models in which plasticity depends on the firing rate, synaptic normalization schemes are necessary to stabilize synaptic weights, and these schemes invariably introduce competition between synapses (Miller and MacKay, 1994). This has lead to the notion that competition is an inevitable consequence of stable Hebbian plasticity. STDP learning rules that do not include the weight dependence of potentiation also produce strong competition and, for instance, correlations in some inputs push other inputs to zero (Song et al., 2000). In contrast, weight-dependent STDP generates stable Hebbian plasticity without introducing much competition.

Competition can be introduced into weight-dependent STDP through an independent mechanism such as activity-dependent scaling. It is important to note that activity-dependent scaling is not needed to prevent runaway learning, but instead keeps the activity of the postsynaptic neuron within bounds as the input undergoes strong changes. This allows the activity-dependent scaling to be much slower than the STDP, as is suggested by experimental observations (Turrigiano et al., 1998). Our implementation of the scaling as an integral controller is well suited for this task because it is both slow and strong. The scaling mechanism literally scales the entire weight distribution up or down without qualitatively changing its shape, as was also observed experimentally (O'Brien et al., 1998; Turrigiano et al., 1998). Activity-dependent scaling introduces competition between the synapses because if some synapses are potentiated and the postsynaptic activity increases, all the synaptic weights will be scaled down.

Whereas strong competition is clearly important for some processes such as ocular dominance plasticity, in which inputs from one eye are retained, and inputs from the other eye are largely lost, it may not be desirable under all conditions and during all periods of development. In adult animals or in central circuits that code a continuous variable, such as direction, it may be advantageous to allow synaptic weights to change while retaining weak inputs. Such inputs could then be potentiated again if circumstances were to change, allowing the circuit greater flexibility. Our results demonstrate that stable Hebbian plasticity and synaptic competition are separable entities and suggest that learning rules may vary by region or developmental period to generate more or less competition.

APPENDIX

Derivation of the weight distribution

Apart from simulations, we present calculations that show how the synaptic weight distribution follows from the plasticity rules. The advantage of the analytical calculations is that although some approximations have to be made, the role of the various parameters in the model becomes clear and can be studied systematically.

Consider a neuron receiving uncorrelated Poisson inputs. Its synapses continuously undergo weights modifications according to the plasticity rules because of random coincidences of presynaptic and postsynaptic spikes. We denote the distribution of its synaptic weights with $P(w, t)$, where t denotes the time. A single bin in this distribution describes the probability that a synapse has a weight, w. Every time step the number of synapses in this bin can change because of potentiation and depression (Fig. 6). Collecting all terms that change the number of synapses in this bin, we have:

$$\frac{1}{\rho_{in}} \frac{\partial P(w, t)}{\partial t} = -p_p P(w, t) - p_d P(w, t)$$
$$+ p_p P(w - w_p, t) + p_d P(w + w_d, t), \quad (4)$$

where ρ_{in} is the presynaptic rate, assumed identical at all synapses, p_p is the probability that the synapse is potentiated, p_d is the probability that the synapse is depressed, the w_d and w_p describe how much the weight changes with depression and potentiation. The first two right-hand side terms in Equation 4 are loss terms decreasing the number of synapses with weight w. The last two terms are gain terms describing synapses with initially different weights acquiring new weight w because of either potentiation or depression.

For now we neglect the precise timing dependence of the plasticity. Instead, we assume that if the synaptic event occurs within a narrow time window t_w after a spike, the synapse is depressed an amount $w_d = -c_d w + vw$, see Equation 2 in Materials and Methods. And similar if the synaptic event occurs before the postsynaptic spike, the synapse is potentiated $w_p = c_p + vw$. In other words the exponential window is replaced by a square window of width t_w. The justification is that when averaged over many pairings, only the average amount of change is important. (This approximation introduces a small, negligible error in Eq. 7). In the simulations the exponential window is used.

By Taylor expanding $P(w - w_p)$ and $P(w + w_d)$, one obtains the Fokker–Planck equation (van Kampen, 1992):

$$\frac{1}{\rho_{in}} \frac{\partial P(w, t)}{\partial t} = -\frac{\partial}{\partial w}[A(w) P(w, t)] + \frac{1}{2} \frac{\partial^2}{\partial w^2}[B(w) P(w, t)],$$
$$(5)$$

with jump-moments A and B:

$$A(w) = -p_d c_d w + p_p c_p, \quad (6)$$

$$B(w) = p_d c_d^2 w^2 + p_p c_p^2 + (p_d + p_p) w^2 \sigma^2, \quad (7)$$

where σ^2 is the variance of the noise term v. This derivation requires that changes in w are small with respect to variations in $P(w, t)$, which is indeed the case. But to solve these equations we first need to know the probability for inducing potentiation p_p and the probability for inducing depression p_d.

The probability that a synaptic event causes a spike

First, we calculate the probability that a synaptic event depresses the synapse. This requires that the presynaptic event succeeds a postsynaptic spike within a short window. We use a simplified model: a non-leaky integrate and fire model. The cell receives background input from other synapses, described by a constant background current I_0. This current causes the neuron to fire regularly with an interspike interval $t_{isi} = V_{thr} C / I_0$, where V_{thr} is the threshold voltage relative to the resting voltage, and C is the membrane capacitance. We assume that the presynaptic signal is uncorrelated to other inputs and that the presynaptic signal is not
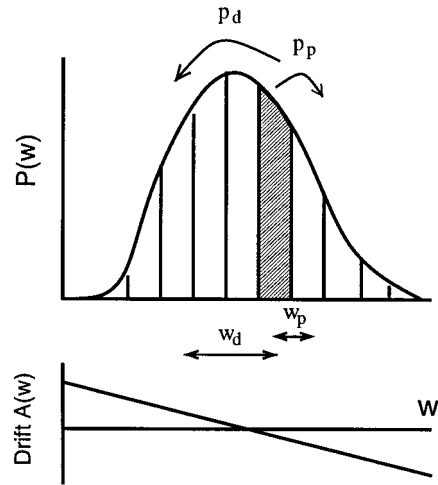

Figure 6. Diagram of the evolution of the synaptic weight distribution. Every time-step the number of synapses within a bin can change: with a probability p_d the synapse depresses an amount w_d, or with a probability p_p the synapse is potentiated an amount w_p. Similarly, synapses with initially a different weight can enter the bin. Finally, a steady state distribution is reached. *Bottom graph* shows the net weight change (drift), which determines whether on the average a weight will increase (positive drift) or decrease (negative drift).

affected by the spikes in the postsynaptic neuron (that is, no recurrent connections). At a random time the synaptic event arrives in the postsynaptic neuron. The postsynaptic spike, occurring earlier, is of course independent of this synaptic event. Therefore, the probability that the synaptic event occurs within a time window t_w after a spike is:

$$p_d = \frac{t_w}{t_{isi}}, \quad (8)$$

where it is implied that $t_w < t_{isi}$.

Next, we calculate the probability that a synaptic event potentiates the synapse, which requires that the postsynaptic spike comes after the synaptic event. Because the synaptic event can help to induce a spike, this is more complicated than the previous case (Kistler and van Hemmen, 2000). We calculate this as follows: the synaptic current is modeled with a brief square pulse of duration τ_{syn}, its amplitude is $w V_{syn}$, where w is the synaptic weight and V_{syn} is the synaptic drive (assumed constant). This synaptic current causes the membrane voltage to jump by an amount $\tau_{syn} w V_{syn} / C$. If after the jump the voltage is still below threshold, the interspike interval is shortened to $t'_{isi} = t_{isi} - \tau_{syn} w V_{syn} / I_0$. If, on the other hand, the neuron was already close to threshold it will spike (Fig. 7). One finds that the time between the synaptic event and the spike, δt, is distributed as,

$$P(\delta t) \quad = \frac{1}{t_{isi}} + \frac{V_{syn}}{V_{thr} C} w \quad (\text{if } \delta t < \tau_{syn})/(\text{if } \tau_{syn} < \delta t < t'_{isi}).$$
$$= \frac{1}{t_{isi}}$$

This describes an enhanced probability for small intervals between the synaptic event and a postsynaptic spike. The reason is that postsynaptic spikes likely follow the synaptic input. With $\tau_{syn} \ll t_w$, the probability that the synapse is potentiated, is

$$p_p = \int_0^{t_w} P(\delta t) d\delta t$$

$$= p_d (1 + w / W_{tot}), \quad (\text{but } p_p \leq 1), \quad (9)$$

with $W_{tot} = t_w I_0/(V_{syn}\tau_{syn})$. This probability is a sum of a constant term that describes random coincidence of presynaptic and postsynaptic spikes, and a term linear in w, which describes the enhanced probability that the synapse induces a spike. For synapses with zero weight, p_p equals the probability of inducing depression, p_d. The reason is that the postsynaptic neuron is unaffected by a tiny input, in other words, only random coincidences cause potentiation. The linear term depends on W_{tot}. The W_{tot} is the average current of all other inputs expressed as an instantaneous conductance. If the input to the cell is purely from excitatory synapses, one has:

$$W_{tot} = t_w \rho_{in} N \langle w \rangle, \tag{10}$$

where N is the number of synapses, and $\langle w \rangle$ is their average weight. This dependence of W_{tot} on the average weight shows that W_{tot} is a competition parameter, describing competition among inputs for a postsynaptic spike. As the total input W_{tot} increases, p_p gets smaller. Finally, for large synaptic conductances p_p reaches its upper limit of one. In that case the presynaptic event always induces a spike, a suprathreshold connection. In physiologically relevant situations such synapses are rare, and this upper limit can be ignored.

In Figure 7D we present the results of a simulation showing the probability for depression and potentiation. Although the synaptic time course, leak conductance and noise in the cell give rise to small correction terms (M. C. W. van Rossum, unpublished observations), Equations 8 and 9 are still qualitatively valid: p_d is independent of the weight, and p_p increases linearly with the weight and equals p_d for zero weight. Experimental verification of this law would be desirable.

The synaptic weight distribution

Using the above results for p_p and p_d, we have for the distribution of synaptic weights,

$$\frac{1}{\rho_{in}} \frac{\partial P(w,t)}{\partial t} = -\frac{\partial}{\partial w}[A(w)P(w,t)] + \frac{1}{2}\frac{\partial^2}{\partial w^2}[B(w)P(w,t)] \tag{11}$$

$$A(w) = p_d[-c_d w + (1 + w/W_{tot})c_p] \tag{12}$$

$$B(w) = p_d[c_d^2 w^2 + (1 + w/W_{tot})c_p^2 + (2 + w/W_{tot})w^2\sigma^2]. \tag{13}$$

This describes the evolution of the synaptic weight distribution under random stimulation. Of most interest is the steady-state solution, which corresponds to the equilibrium distribution a neuron obtains with random stimulation. It is independent of the initial distribution.

The steady-state solution is found by imposing that $\partial P/\partial t = 0$ and that the probability current $J(w) = A(w)P(w) - 1/2 \, \partial/\partial w [B(w)P(w)]$ vanishes. The resulting equation is easily solved numerically. An analytical solution is obtained if one assumes that the noise σ and W_{tot} are large, so that $B(w) \approx p_d(c_p^2 + 2w^2\sigma^2)$. In this case the distribution reads:

$$P(w) = N \frac{e^{\sqrt{2}\text{atan}\left(\frac{\sqrt{2}\sigma w}{c_p}\right)/\sigma}}{(2\sigma^2 w^2 + c_p^2)^{\frac{2\sigma^2 - c_p/W_{tot} + c_d}{2\sigma^2}}}, \tag{14}$$

where N normalizes the distribution such that $\int P(w)dw = 1$. This distribution is plotted in Figure 3A (solid line). It closely matches the simulation results. The steady-state solution is unimodal, and the mean weight is roughly located where potentiation and depression are balanced (here A crosses zero). The distribution peaks at $w = c_p/(c_d - c_p/W_{tot} + 2\sigma^2)$. Because W_{tot} depends on the average weight (Eq. 10), the distribution (Eq. 14) has to be solved self-consistently. In practice this poses no problem because the distribution depends only weakly on W_{tot}. Indeed, approximating $W_{tot} \to \infty$ only slightly changes the distribution (Fig. 3, dashed line). This means that the enhanced probability for inducing a spike for strong synapses is a minor effect. Finally, note that the distribution does

not vanish at zero conductance; this is hard to see in Figure 3, but is clear from Equation 14.

The evolution of the distribution can be compared to diffusion of particles (weights) in an external force field. In analogy with the diffusion equation, the A term is a force that the synapse experiences, its sign determines whether with the next event the weight will on the average increase or decrease (Fig. 6). The B term corresponds to a diffusion "constant" and determines the width of the distribution; it is determined by the amount of weight change and the noise. Although without the noise term, the distribution would still have a finite width and a positive skew, the noise broadens the distribution and due to its multiplicative character, the noise also enhances the positive skew of the distribution.

Application to other models

Other models of spike timing-dependent plasticity have not included the size dependence of the plasticity rules (Amarasingham and Levy, 1998; Kempter et al., 1999; Song et al., 2000). Also these models can be analyzed with our method. We show that they yield a dramatically different weight distribution. Following the notation of Song et al. (2000), we have, again neglecting the exponential timing dependence,

$$w_d = -A_- \tag{15}$$

$$w_p = A_+, \tag{16}$$

where A_+ is the amount of potentiation, and A_- is the amount of depression. This plasticity scheme requires slightly more potentiation than depression, that is, $A_+ = (1 - \epsilon)A_-$, where ϵ is a small, positive number. For small ϵ and large W_{tot}, one has for this model $A(w) = p_d(w/W_{tot} - \epsilon)A_-$ and $B(w) = 2p_d A_-^2$. There is no steady-state weight distribution unless a hard limit on the maximal weight, w_{max}, is explicitly imposed. The resulting distribution is:

$$P(w) = Ne^{(-\epsilon w + \frac{1}{2}w^2/W_{tot})/A_-} \quad (0 < w < w_{max}), \tag{17}$$

where N normalizes the distribution. This distribution matches the distributions observed in simulations (Song et al., 2000). It is plotted in Figure 3C with parameters: $w_{max} = 1$, $W_{tot} = 11$, $\epsilon = 0.05$, and $A_- = 0.005$. It is seen that some synaptic weights will cluster around zero weight. The reason is that for weak synapses the drive $A(w)$ is negative, pushing them toward smaller and smaller synaptic conductances. If one chooses $w_{max} > \epsilon W_{tot}$ the distribution is bimodal and has a second peak at the maximal weight. In that case $A(w)$ becomes positive for large w and weights for which $A(w)$ is positive are pushed towards w_{max}. When input correlations are present, the distribution becomes already bimodal for lower values of w_{max}.

The parameters are usually chosen such that the distribution is bimodal. This requires a balance between the potentiation and depression ratio, ϵ, and the competition parameter, w/W_{tot}. The weight distribution in these models is sensitive to small perturbations in this balance. This is seen in Figure 3C: the weight distribution plotted there would be symmetric if W_{tot} were 10, but a 10% change ($W_{tot} = 11$) causes already a considerably asymmetric distribution. Because of this strong dependence, Equation 17 needs to be solved self-consistently. Namely, through W_{tot} the distribution depends on the weights of the inputs (Eq. 10), but the weights are again given by the distribution. Solving Equation 17 for a range of input rates, shows that there is indeed a limited regime in which the postsynaptic firing frequency is almost independent of the input rate, as was seen in simulations (Song et al., 2000).

The dependence of $A(w)$ on the weight is precisely the opposite to our model where $A(w)$ decreases with increasing weight stabilizing the weight distribution (Fig. 6). The stability of different learning rules can be analyzed with our method. If the potentiation depends as strong or more strongly on the weight than in our model, the weights will be stable. If the potentiation depends much more weakly on the weight than assumed here, the stability depends on parameters such as the threshold weight and postsynaptic
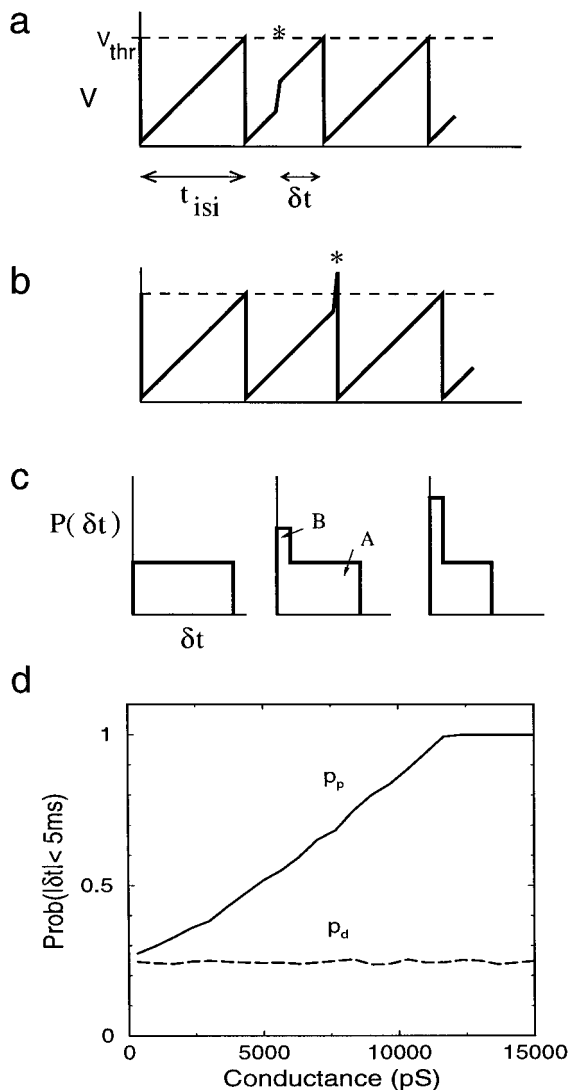
Figure 7. The probability for a synapse to induce potentiation. *a*, *b*, The membrane potential in the presence of a constant background current is plotted versus time. At the *asterisk* a synaptic event occurs. *a*, If this occurs long before the spike, the membrane potential makes a step, and the interspike interval is shortened. *b*, If the synaptic event occurs when the membrane potential is already close to threshold, an immediate spike is the result. *c*, Combining these two possible events gives the probability of the time between synaptic event and spike, plotted for three values of the synaptic weight *w*: 0, medium and large (left to right). *d*, The weight dependence of the probability to induce depression and potentiation. The probability for depression, p_d is independent of the synaptic weight. The probability for potentiation, p_p is linear in the weight, expressing the contribution of the synaptic event to the postsynaptic spike. In this simulation the weights are fixed (uniformly distributed). The probability that the synaptic event and the postsynaptic spike fall within a short time window is measured. This corresponds to the probability that the synapse would be depressed ($\delta t < 0$) or potentiated ($\delta t > 0$). The excitatory synapses were driven at 1 Hz, firing the cell at ~50 Hz.

firing frequency. The behavior of any model is determined by how $A(w)$ crosses zero, with positive or negative slope. In general we can distinguish two classes of learning rules. Assume that $B(w) \neq 0$ for any *w* and that $A(w)$ crosses zero only once, (1) if $A(w)$ crosses zero with negative slope the distribution is centered around the zero crossing as in our model, (2) but if $A(w)$ crosses zero with positive slope the weights are repelled from the zero crossing as happens above. This dichotomy does not seem to leave room for a stable model which intrapolates between the two classes.

Note added in proof. After completion of this study, we found that a similar approach to calculate the weight distribution has been followed by Rubin et al. (2000).

REFERENCES

Amarasingham A, Levy WM (1998) Predicting the distribution of synaptic strengths and cell firing correlations in a self-organizing, sequence prediction model. Neural Comput 10:25–57.

Bekkers JM, Richerson GB, Stevens CF (1990) Origin of variability in quantal size in cultured hippocampal neurons and hippocampal slices. Proc Natl Acad Sci USA 87:5359–5362.

Bell CC, Han VZ, Sugawara Y, Grant K (1997) Synaptic plasticity in cerebellum-like structure depends on temporal order. Nature 387:278–281.

Bi G-Q, Poo M-M (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, Mundo PW (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci 2:32–48.

Blum KI, Abbott LF (1996) A model of spatial map formation in the hippocampus of the rat. Neural Comput 8:85–93.

Debanne D, Gähwiler BH, Thompson SM (1996) Cooperative interactions in the induction of long-term potentiation and depression of synaptic excitation between hippocampal CA3-CA1 cell pairs in vitro. Proc Natl Acad Sci USA 93:11225–11230.

Debanne D, Gähwiler BH, Thompson SM (1999) Heterogeneity of synaptic plasticity at unitary CA1-CA3 and CA3-CA3 connections in rat hippocampal slice cultures. J Neurosci 19:10664–10671.

Desai NS, Rutherford LC, Turrigiano GG (1999) Plasticity in the intrinsic electrical properties of cortical pyramidal neurons. Nat Neurosci 2:515–520.

Destexhe A, Paré D (1999) Impact of network activity on the integrative properties of neocortical pyramidal neurons. J Neurophysiol 81:1531–1547.

Eurich CW, Pawelzik K, Ernst U, Cowan JD, Milton JG (1999) Dynamics of self-organized delay adaptation. Phys Rev Lett 82:1594–1598.

Forti L, Bossi M, Bergamaschi A, Villa A, Malgaroli A (1997) Loose-patch recordings of single quanta at individual hippocampal synapses. Nature 388:874–878.

Frerking M, Borges S, Wilson M (1995) Variation in GABA mini amplitude is the consequence of variation in transmitter concentration. Neuron 16:885–895.

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. Nature 383:76–78.

Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. Phys Rev E 59:4498–4514.

Kirkwood A, Rioult MC, Bear MF (1996) Experience-dependent modification of synaptic plasticity in visual cortex. Nature 381:526–528.

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.

Liu G, Choi S, Tsien RW (1999) Variability of neurotransmitter concentration and nonsaturation of postsynaptic AMPA receptors at synapses in hippocampal cultures and slices. Neuron 22:395–409.

Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213–215.

Miller KD (1996) Synaptic economics: competition and cooperation in synaptic plasticity. Neuron 17:371–374.

Miller KD, MacKay DJC (1994) The role of constraints in Hebbian learning. Neural Comput 6:100–126.

Ngezahayo A, Schachner M, Artola A (2000) Synaptic activity modulates the induction of bidirectional synaptic changes in adult mouse hippocampus. J Neurosci 20:2451–2458.

Nusser Z, Cull-Candy S, Farrant M (1997) Differences in synaptic GABA(A) receptor number underlie variation in GABA mini amplitude. Neuron 19:697–709.

O'Brien RJ, Kamboj S, Ehlers MD, Rosen KR, Fischbach GD, Kavanaugh MP, Huganir RL (1998) Activity-dependent modulation of synaptic AMPA receptor accumulation. Neuron 21:1067–1078.

Oja E (1982) A simplified neuron model as a principal component analyzer. J Math Biol 15:267–273.

Riggs DS (1970) Control theory and physiological feedback mechanisms. Baltimore: Williams and Wilkins.

Rubin J, Lee DD, Sompolinsky H (2000) Equilibrium properties of temporally asymmetric hebbian plasticity. (preprint cond-mat/0007392 at xxx.lanl.gov)

Sejnowksi TJ (1977) Storing covariance with nonlinearly interacting neurons. J Math Biol 4:303–321.

Shatz CJ (1990) Impulse activity and the patterning of connections during CNS development. Neuron 5:745–756.

Song S, Abbott LF (2000) Temporally asymmetric Hebbian learning, spike timing and neuronal response variability. In: Advances in neural information processing systems 11 (Kearns MS, Solla SA, and Cohn DA, eds) pp 69–75. Cambridge, MA: MIT.

Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat Neurosci 3:919–926.

Spruston N, Jaffe DB, Williams SH, Johnston D (1993) Voltage- and space-clamp errors associated with the measurement of electrotonically remote synaptic events. J Neurophysiol 70:781–802.

Turrigiano GG (1999) Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. Trends Neurosci 22:221–227.

Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 391:892–896.

van Kampen NG (1992) Stochastic processes in physics and chemistry. Amsterdam: North-Holland.

von der Malsburg C (1981) The correlation theory of the brain (Internal Report 81-2 Max Planck Institute Göttingen. Reprinted in Models of Neural Networks II, Domany E, van Hemmen JL, and Schulten K, eds). Berlin: Springer.

Yang XD, Faber DS (1991) Initial synaptic efficacy influences induction and expression of long-term changes in transmission. Proc Natl Acad Sci USA 88:4299–4303.

Zhang LI, Tao HW, Holt CE, Harris WA, Poo M (1998) A critical window for cooperation and competition among developing retinotectal synapses. Nature 395:37–44.