Biol. Cybern. 87, 344–355 (2002)
DOI 10.1007/s00422-002-0350-1

Biological Cybernetics

Beyond spike timing: the role of nonlinear plasticity and unreliable synapses

Walter Senn

Physiological Institute, University of Bern, Buehleplatz 5, 3012 Bern, Switzerland

Received: 22 April 2002 / Accepted: 23 May 2002

Abstract. Spike-timing-dependent plasticity (STDP) strengthens synapses that are activated immediately before a postsynaptic spike, and weakens those that are activated after a spike. To prevent an uncontrolled growth of the synaptic strengths, weakening must dominate strengthening for uncorrelated spike times. However, this weight-normalization property would preclude Hebbian potentiation when the pre- and postsynaptic neurons are strongly active without specific spike-time correlations. We show that nonlinear STDP as inherent in the data of Markram et al. [(1997) Science 275:213–215] can preserve the benefits of both weight normalization and Hebbian plasticity, and hence can account for learning based on spike-time correlations and on mean firing rates. As examples we consider the moving-threshold property of the Bienenstock–Cooper–Munro rule, the development of direction-selective simple cells by changing short-term synaptic depression, and the joint adaptation of axonal and dendritic delays. Without threshold nonlinearity at low frequencies, the development of direction selectivity does not stabilize in a natural stimulation environment. Without synaptic unreliability there is no causal development of axonal and dendritic delays.

1 Introduction

One of the remarkable features of sensory systems is their capability to adapt to the characteristics of the environment. Why, for instance, do we prefer to see moving objects with a specific velocity bandwidth? The answer is, in most cases such preferences are governed by the statistics of the natural scenes. To biophysically realize the match between input statistics and feature-extraction capabilities, the nervous system may operate at different levels and time scales, from genetically guided development to activity-dependent fine tuning of neuronal connections. Spike-timing-dependent synaptic plasticity (STDP) is a possible candidate to explain activity-dependent changes in the adult primary visual cortex (area V1) which are expressed on timescales of minutes (Yao and Dan 2001; Fu et al. 2002), but may also play an important role in the developing cortex (Hubel and Wiesel 1962; Katz and Shatz 1996). The main characteristics of STDP are that the synaptic strength is upregulated if the presynaptic spike precedes the postsynaptic spike within 5–50 ms, and downregulated if it follows the postsynaptic spike within a similar time range (Markram et al. 1997; Bi and Poo 1998; Feldman 2000). In the present review we highlight four different features of our learning rule which go beyond the standard properties of models of STDP: the inherent nonlinearity, the learning threshold, the relation to short-term depression, and the synaptic unreliability. We will also see that the statistics of the stimulus distribution is not always the dominant factor in the development of the synaptic structure. While this is true for the development of delay lines where the temporal structure of the stimuli is mapped onto the transmission delays of the neurons, the specific velocities of moving stimuli are not necessarily encoded in the synaptic structure of direction-selective simple cells. Rather, the velocity selectivity of these cells strongly depends on the parameters of the synaptic learning rule. Before going into the details of these phenomena we briefly comment on the four additional features of the learning rule.

1.1 Nonlinear synaptic plasticity

An important requirement of STDP is that long-term depression (LTD) dominates long-term potentiation (LTP), thereby preventing an uncontrolled increase of the postsynaptic activity through normalization of the synaptic weights (Kempter et al. 1999; Song et al. 2000). When the pre- and postsynaptic spike times are not correlated, this normalization property implies that LTP

Correspondence to: e-mail: wsenn@cns.unibe.ch,
Tel.: +41-31-6318721, Fax: 41-31-6314611

is wiped out by LTD, and that the net LTD becomes more dominant with increasing pre- and postsynaptic frequencies. Unfortunately, this is just the opposite of the Hebb rule, according to which it is the net LTP that should increase with the pre- and postsynaptic frequencies. Hence, additional nonlinearities are required if STDP should be relevant for both encoding information represented in a spike correlation code and a mean rate code without spike correlations (Amit 1997; Shadlen and Newsome 1998). For instance, storing spatial patterns in an associative memory requires that LTP dominates at elevated pre- and postsynaptic activities, while only elevated presynaptic activity should induce LTD (see, e.g., Fusi et al. 2000). Similarly, the development of orientation columns based on a mean rate code, as described in the original Bienenstock–Cooper–Munro (BCM) rule, can only be explained if synapses are strengthened at elevated pre- and postsynaptic frequencies. Recent in vivo experiments which include the postsynaptic depolarization as a further key variable in the induction of synaptic plasticity show that STDP may be consistent with these requirements (Sjöström et al. 2001). To realize the two main features – synaptic normalization (achieved through the dominance of LTD at high pre- and postsynaptic frequencies) and classical Hebbian potentiation (achieved through the dominance of LTP at lower, but still elevated frequencies) – the rule for the synaptic modification must qualitatively change as a function of the pre- and postsynaptic activities. The algorithm designed to reproduce the data of Markram et al. (1997) exhibits nonlinearities which in fact imply such a qualitative change.

1.2 Learning thresholds

Another nonlinearity, again shared by the cited algorithm, is the threshold for the induction of synaptic modifications. Repetitive pairing of pre- and postsynaptic spikes with 10-ms delay only induced LTP if the pairing frequency was above 5 Hz (cf. Sjöström et al. 2001). Such a threshold property is important to protect acquired synaptic structures against random uncorrelated activity. For instance, direction selectivity (DS) of simple cells in the area V1 can only develop in a natural input scenario – with stimuli moving with equal probability in different directions – if STDP is endowed with a learning threshold. As we show, without such a threshold any acquired change will be canceled if by chance subsequent stimuli move in the opposite directions. A learning threshold must also be present in the adult brain if natural stimulation should not always change the orientation selectivity of cells. Repetitive strong stimulations with orientationally biased stimuli can overcome this threshold and induce shifts in the orientation selectivity of V1 cells (Yao and Dan 2001). The importance of learning thresholds is well known for the formation of Hebbian ensembles with multiple patterns and realistic noise levels (Fusi et al. 1999).

1.3 Modification of short-term depression

The development model for the direction-selective simple cells we consider is based on synaptic short-term depression (Chance et al. 1998) and its activity-dependent modification. This exploits the fact that the response of depressing synapses are phase advanced with respect to that of nondepressing synapses. STDP acting on both the synaptic strength and the degree of short-term depression (Markram and Tsodyks 1996) may induce the appropriate spatial displacement of depressing and nondepressing synapses from which DS emerges. Our analysis reveals that the time constant of synaptic depression ($\sim$200 ms) and the time window for STDP ($\sim$40 ms) both lead to a maximal DS for stimuli with temporal frequencies of 0.5–16 Hz. This agrees well with recordings from simple cells in V1 (Saul and Humphrey 1992). In contrast to our expectations, however, the development of this frequency band is independent of the distribution of the stimulus frequencies. One may speculate that the synaptic parameters (i.e., the 200-ms and 40-ms values) are tuned by another mechanism to allow simple cells to extract motion in the behaviorally relevant velocity bandwidth.

1.4 Synaptic unreliability

Unreliable synapses play a role in development similar to that of random mutations in evolution. Transmission failures allow the specifical reward of connections which successfully contributed to the postsynaptic firing, while excluding synapses from this reward which were not active and hence did not contribute. As an example we consider the development of the receptive field (RF) of a cell in V1. Such cells have a wide subthreshold integration field where stimuli generate slow depolarization, and a narrow discharge field where stimuli may evoke postsynaptic spikes (Bringuier et al. 1999). This receptive field structure may correspond to the anatomical finding that lateral input from distal V1 cells tend to project on the apical dendritic tree of layer V pyramidal cells, while feedforward and nearby connections project closer to the soma (Nieuwenhuys 1994). How can such an arrangement develop in an activity-dependent manner? The answer is STDP and synaptic unreliability. The synaptic position along the dendritic tree is intimately related to the latency of the corresponding synaptic response. Assuming a specific spatiotemporal statistics of sensory input patterns, we show that STDP and unreliable transmission leads to an appropriate selection of axonal delays and dendritic latencies, with the observed synaptic distribution along the dendritic tree.

2 Nonlinear spike-timing-dependent plasticity

2.1 Model of short-term depression and its modification

We designed an algorithm for synaptic modifications based on in vitro recordings from pairs of interconnected

layer V pyramidal cells in somatosensory cortex of young rats. The recordings include (i) the classical STDP experiments from Markram et al. (1997), and (ii) an additional experiment expressing intrinsic nonlinearities of STDP (Senn et al. 1997, 2001b). The algorithm is based on a stochastic synaptic model with short-term depression caused by vesicle depletion. The dynamic variable is given by the probability of a vesicle being recovered, P_v. Given a recovered vesicle, a presynaptic spike will discharge this vesicle with probability P_{dis}. The total vesicle release probability takes into account the probability for vesicle recovery and its discharge, $P_{rel} = P_v P_{dis}$. Once a vesicle is discharged, P_v is set to 0, from where it stochastically recovers to 1 with a time constant τ_{rec} of 150–800 ms. This is a probabilistic version of the synaptic depression model in Tsodyks and Markram (1997), with the discharge probability (P_{dis}) corresponding to the use of synaptic efficacy (U_{SE}). High P_{dis} (close to 1) implies a strong transient response to an increase in the presynaptic firing rate followed by strong depression, while low P_{dis} (close to 0) implies a weak response, but also weak depression.

According to the findings on synaptic redistributions (Markram and Tsodyks 1996) we were assuming that it is P_{dis} which is subject to long-term modification. However, other experiments do only report modifications of the absolute synaptic strength. In fact, if G represents the absolute synaptic strength, the postsynaptic current evoked by presynaptic test impulses, given at 1 Hz for instance, has an amplitude proportional to the product GP_{dis}. This is because the vesicles recover between the interspike intervals, $P_v \approx 1$, and we get $P_{rel} \approx P_{dis}$. Isolated test impulses therefore cannot distinguish between a modification of the absolute synaptic strength and a modification of the discharge probability. Although the present algorithm is formulated in terms of a modifications of P_{dis}, we will assume in the following applications that also the absolute synaptic strength G is subject to the same modifications.

To account for the nonlinearities present in the experimental data we constructed the following scheme: Postsynaptic NMDA receptors (NMDARs) can make transitions from a recovered state either into an upregulating state triggered by neurotransmitter release (via binding of glutamate), or into a downregulating state triggered by a postsynaptic spike (via removing of the Mg^{2+} block), from where they recover with a time constant τ_N in the range of several tens of milliseconds. The fraction of receptors in the respective states are denoted by N_{rec}, N_{up}, and N_{dn}. Up- and downregulating secondary messengers, S_{up} and S_{dn}, with a slower decay time constant τ_S, are activated proportionally to the amount of N_{up} and N_{dn}, whenever there is a postsynaptic spike or a presynaptic release, respectively. In terms of kinetic schemes the cross-gating of the NMDARs and the activation of the secondary messengers reads

$$N_{dn} \overset{1\cdot post}{\underset{\tau_N}{\leftarrow}} N_{rec} \quad \overset{1\cdot rel}{\underset{\tau_N}{\rightarrow}} N_{up}, \text{ and}$$

$$\bar{S}_{up} \overset{N_{up}\cdot post}{\underset{\tau_S}{\leftarrow}} S_{up}, \bar{S}_{dn} \overset{N_{dn}\cdot rel}{\underset{\tau_S}{\rightarrow}} S_{dn} \ ,$$

where $\bar{S}_{up}$ $(= 1 - S_{up})$ and $\bar{S}_{dn}$ $(= 1 - S_{dn})$ are the up- and downregulating secondary messengers in the recovered state. The labels 'x·post' and 'x·rel' indicate that the transitions take place at the time of a postsynaptic spike and a presynaptic release, respectively, and that the fraction x of the available source is moved to the new state. As a third event, following a postsynaptic spike, P_{dis} is upregulated proportionally to $[S_{up} - \Theta_{up}]^+$, the amount of upregulating secondary messenger above a certain threshold. Similarly, following a presynaptic release, P_{dis} is downregulated proportionally to $[S_{dn} - \Theta_{dn}]^+$. By introducing multiplicative saturation which restricts P_{dis} to the unit interval, one obtains

$$\tau_P \frac{dP_{dis}}{dt} = (1 - P_{dis}) \sum_{t_{post}} [S_{up} - \Theta_{up}]^+ - P_{dis} \sum_{t_{rel}} [S_{dn} - \Theta_{dn}]^+ \ , \quad (1)$$

where τ_P is a slow learning time constant (of the order of minutes, see Senn et al. 2001b).

The rationale behind this scheme is that a running mean of the presynaptic activity is stored in N_{up}, and that the postsynaptic activity, via S_{up}, reads out this stored presynaptic activity and upregulates the synaptic parameter. Similarly, the postsynaptic activity is stored in N_{dn}, and the presynaptic activity, via S_{dn}, reads out the stored postsynaptic activity and downregulates the synaptic parameter. The sequence of three events – the modification of NMDARs, the activation of secondary messenger, and the change of P_{dis} with each presynaptic release or postsynaptic spike – result in an algorithm that is of third order in the pre- and postsynaptic spike frequencies (see Eq. 2 below).

2.2 Experimental basis of the nonlinear STDP algorithm

The algorithm in Sect. 2.1 was designed to reproduce the following three experiments (see also Markram et al. 1997; Senn et al. 2001b):

1. Pre- and postsynaptic spike trains of 10 Hz, paired with ±10 ms lag (five spikes per train, repeated ten times every 4 s), induced a slow, long-lasting increase of P_{dis} by 20% in the pre–post case and a decrease by 20% in the post–pre case. A lag of ±100 ms did not induce any change. This is the standard property of STDP also shared by other experiments and algorithms (Gerstner et al. 1996; Kempter et al. 1999; Kistler and van Hemmen 2000; Song et al. 2000), and is modeled by the cross-gating of NMDARs.
2. Paired pre- and postsynaptic spike trains with a delay of 2 ms of the postsynaptic train (five spikes per train, repeated ten times every 4 s) induced a change in P_{dis} which monotonically increases up to 150% with a pairing frequency of 2–40 Hz. This frequency dependence of LTP naturally arises in the algorithm from the memory effect of the NMDA states and the accumulated secondary messenger. The crucial nonlin-

earity is the 'learning threshold' around 5 Hz (Fig. 1a, b), and it is implemented by the thresholds applied to the secondary messengers (although the experiment can only tell about an LTP threshold, we assumed a similar threshold for LTD). Standard models of STDP would predict a linear increase with the number of spike pairs and therefore fail to reproduce this experiment.

3. Pre- and postsynaptic spike trains, paired at 20 Hz with a delay of 2 ms of the postsynaptic train (repeated ten times every 4 s), did *not* show a monotonic increase of the change in P_{dis} as the number of spikes in the paired trains increased from 2 to 20. Rather, the induced change remained independent, or even decreased with the number of paired spikes (Senn et al. 2001b). It is this result, together with the effect of synaptic depression and the frequency dependence of LTP, which justifies the third-order algorithm.

When applying the algorithm to a stimulation protocol with 50 spikes paired at 0.1 Hz and a ±5-ms lag, only 2% LTP and 4% LTD (Senn et al. 2001b) were obtained. This is a consequence of the threshold nonlinearity revealed by experiment 2, and contrasts with the strong LTP obtained by extracellular stimulation of inputs to L2/3 pyramidal cells at the same low pairing frequency (Feldman 2000). As shown by Sjöström et al. (2001), the discrepancy might arise from the fact that the size of the postsynaptic response should be above 2.3 mV (at least for L5 pyramidal neurons in V1 of young rats) to reliably induce LTP. The saturation of LTP with the number of paired spikes in a 20-Hz train (experiment 3) could be explained by an effective time constant of the LTP integrator of roughly 50ms. The fact that LTP actually slightly decreases with the number of paired APs (experiment 3) but not with frequency (experiment 2), however, reveals that STDP is highly nonlinear: On the one hand, there is a long internal LTD time constant which leads to an annihilation of LTP by LTD after many pre–post spike pairs (experiment 3). On the other hand, the threshold nonlinearity around 5Hz suggests that there is an even longer internal time constant for LTP, and this seems to contradict the increasing dominance of LTD. Moreover, in the presence of synaptic depression and the long LTD time constant, it is difficult to explain the strong frequency-dependent

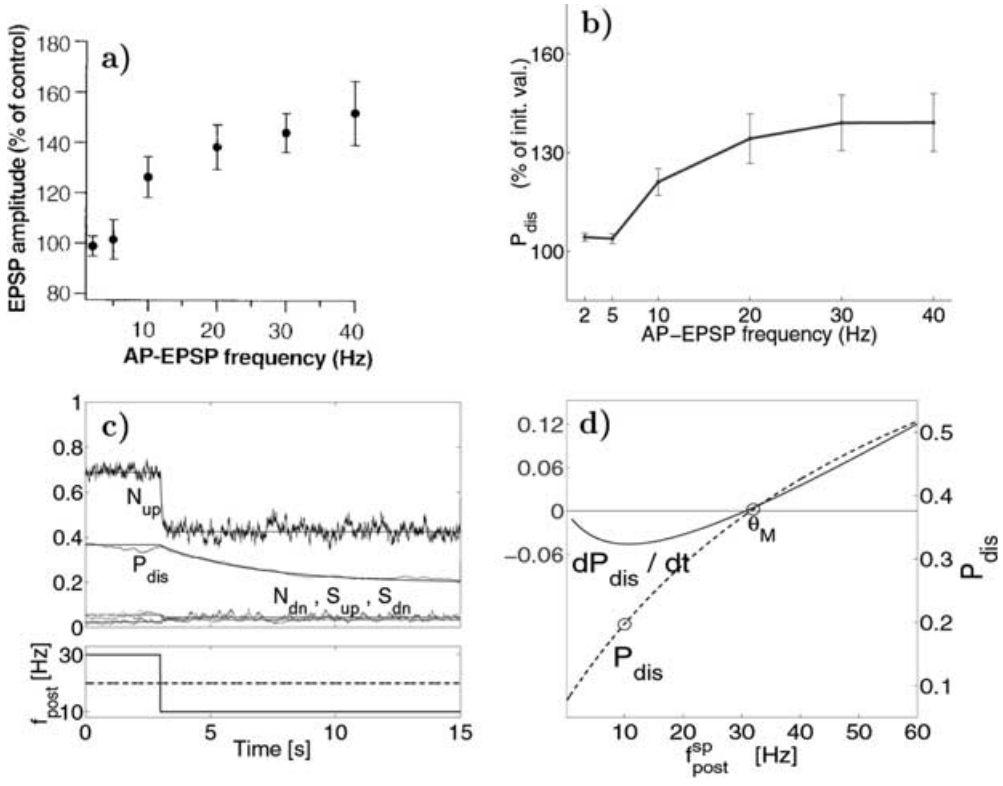

Fig. 1. Nonlinear spike-timing-dependent plasticity (STDP): learning threshold and activity-dependent time window. **a, b** The modification of the probability of vesicle discharge, ΔP_{dis}, monotonically increases with the pairing frequency, showing a threshold nonlinearity (the 'learning threshold') around 5 Hz. Experimental data reproduced from Markram et al. (1997) (**a**), and the corresponding simulation results (**b**). **c, d** STDP may reproduce the Bienenstock–Cooper–Munro (BCM) rule for uncorrelated pre- and postsynaptic Poisson spike trains by means of an activity-dependent effective time window. **c** A step decrease of the postsynaptic firing rate from 30 to 10Hz (*lower panel*) brings P_{dis} down from 0.38 to 0.2 (*upper panel*, with additional variables described in the text). **d** The final value of P_{dis}, with a presynaptic spike frequency fixed at 20 Hz, monotonically increases with the postsynaptic spike frequency (*dashed line*, *right ordinate*, with the *circles* representing the steady-state values of P_{dis} in **c**). For postsynaptic frequencies below 30 Hz LTD dominates, for frequencies above 30 Hz LTP dominates, as indicated by the temporal derivative of P_{dis} (*solid line*, *left ordinate*). With a fixed time window for STDP, either LTP or LTD would dominate over the whole range of postsynaptic frequencies, depending on whether the integral over the time window is positive or negative. Adapted from Senn et al. (2001b)

increase of LTP (experiment 2). The third-order scheme with the postsynaptic spike being twice engaged in the induction of LTP is one way to satisfy these opposing requirements. This third-order nonlinearity also implies that both features – classical Hebbian plasticity based on mean firing rates and weight normalization based on spike (anti)correlations – may be inherent in STDP.

2.3 STDP for spike triples

The cross-gating of the NMDARs implies further nonlinearities for the coincidence detection of the pre- and postsynaptic events. As shown by recent experiments on L2/3 connections in cortical slices of rats, the modification induced by successive pre- and postsynaptic spike pairs is not equal to the sum of the modifications induced by the individual spike pairs (Froemke and Dan 2002). A triple of post–pre–post spikes with a (extracellularly evoked) presynaptic spike, symmetrically surrounded by a postsynaptic spike 20 ms before and 20 ms after, for instance, favors the modification induced by the first post–pre spike pair, and hence leads to LTD. In turn, in a pre–post–pre spike triple, LTD induced by the second post–pre pair is suppressed and LTP caused by the first pre–post spike pair is dominant (Fig. 2a). Such suppression effects can originate from different mechanisms, such as a reduced dendritic calcium influx for the second postsynaptic spike, and synaptic depression, respectively. But it is also a natural consequence of the NMDAR cross-gating alone since both the pre- and postsynaptic events are competitively acting on the recovered NMDARs. A pre–post spike pair following a postsynaptic spike is less effective in triggering LTP because of the reduced N_{up} state, while a post–pre pair following a presynaptic release is less effective in triggering LTD because of the reduced N_{dn} state (Fig. 2a).

There is a slight qualitative difference between our model and the one proposed by Froemke and Dan (2002). In the present case the effect of the second spike pair, the pre–post pair within the post–pre–post triple, say, is reduced because of the first post–pre interval. In the Froemke and Dan model it is reduced because of the post–post interval: The LTD reduction factor in their case decreases with the time of the second postsynaptic spike. The contribution of LTP, which also decreases with the second postsynaptic spike time, can therefore be balanced out by the LTD triggered by the preceding spike time interval. As a result, the region where LTP occurs within the post–pre–post quadrant has a horizontal part in the boundary (see Fig. 3b in Froemke and Dan 2000), which does not arise in the corresponding region of our model (see the lower right 'LTP' triangle in Fig. 2a). Unfortunately, the existing experimental data does not allow to argue either for or against one of the two alternative models.

2.4 STDP for Poisson spike trains

We now consider a scenario where the pre- and postsynaptic neurons are firing with Poisson frequencies f_{pre} and f_{post}, respectively. Taking the means and neglecting the cross-gating of the NMDAR, (1) roughly translates to

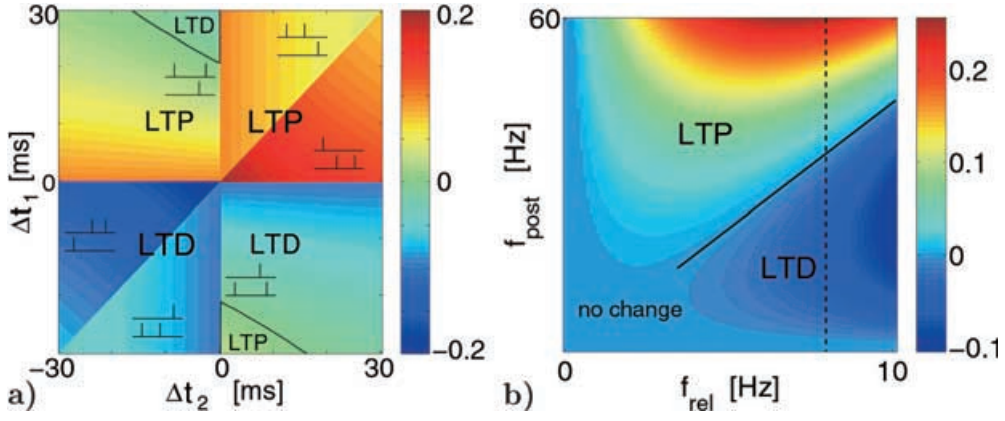

Fig. 2. Encoding spike timings and mean firing rates with nonlinear STDP. **a** Change of P_{dis} as a function of one postsynaptic spike and two presynaptic releases (above the main diagonal, with $\Delta t_i = t^0_{\mathrm{post}} - t^i_{\mathrm{rel}}$, $i = 1, 2$), and of one presynaptic release and two postsynaptic spikes (below the main diagonal, with $\Delta t_i = t^i_{\mathrm{post}} - t^0_{\mathrm{rel}}$, $i = 1, 2$). The key observation is that in the pre–post–pre case LTP is in general dominant (*top left*), even if the post–pre interval is shorter than the previous pre–post interval (as indicated by the sketch showing the pre- and postsynaptic spike timings), while LTD is in general dominant in the post–pre–post case (*bottom right*), even when the pre–post interval is shorter than the previous post–pre interval (*inset sketch*). Synaptic depression will further favor LTP in the pre–post–pre case. The simulations of (1) qualitatively reproduce the results of Froemke and Dan (2002) (for details see text). Parameter values: $\tau_N = 30$ ms, $\tau_S \gg \tau_N$, and $\Theta_{\mathrm{up}} = \Theta_{\mathrm{dn}} = 0$ (justified by the extracellular stimulations). **b** Change of P_{dis} as a function of the synaptic release rate and the postsynaptic firing rate according to (2). As required by the Hebbian theory, LTP dominates for large pre- and postsynaptic frequencies, and LTD dominates for large pre- but small postsynaptic frequencies. In the case of synaptic depression the release rate is a saturating function of the presynaptic firing rate, as indicated in the text. The *dashed vertical line* corresponds to the trace of $\mathrm{d}P_{\mathrm{dis}}/\mathrm{d}t$ shown in Fig. 1d. The slope of the separation line between LTP and LTD (*dotted line*) is given by ratio $\rho_{\mathrm{dn}} P_{\mathrm{dis}}/\rho_{\mathrm{up}}(1 - P_{\mathrm{dis}})$ (= 4 in the present case). We set $\Theta_{\mathrm{up}} = \Theta_{\mathrm{dn}} = 10$

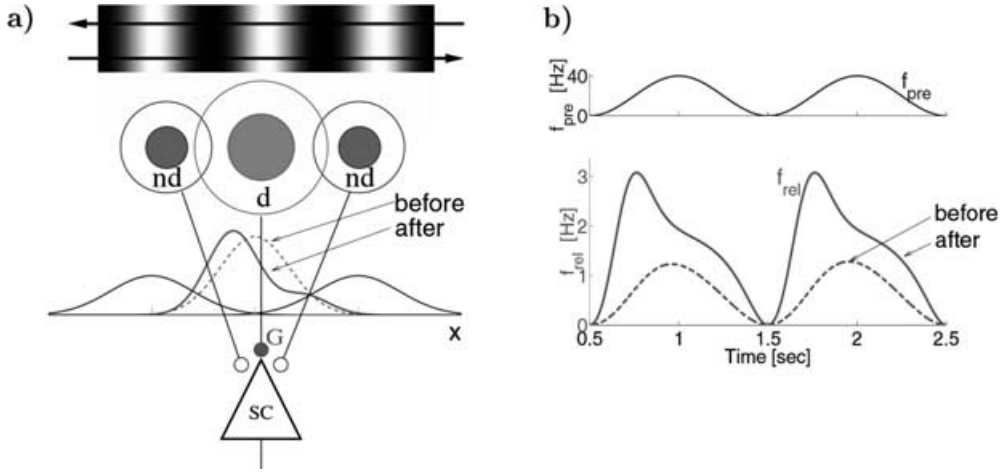

Fig. 3a,b. Developmental model for direction selectivity. **a** An initially nondirectional simple cell (SC) in V1 with symmetrical receptive field is composed of depressing synapses in the center *d* and nondepressing in the surround (*nd*). The *dashed curve* and the two *flanking curves* represent the effective synaptic strength (defined by the synaptic density times the synaptic strength) of the depressing and nondepressing synapses, respectively. When driven by moving gratings with randomly chosen directions and velocities (snapshot of stimulus shown on *top*), STDP applied to the strength (G) of the depressing synapses breaks the symmetry and shifts the effective synaptic strength to one side (in this case to the left, lines *before* and *after*; based on the analytical calculations in Senn and Buchs 2002). **b** Rate of neurotransmitter release of a depressing synapse in response to a 1-Hz-modulated presynaptic firing rate (*top trace*) before (*lower line*) and after (*dashed line*) random stimulations. The temporal phase advance of the release rate develops because STDP also modifies P_{dis} and therefore changes the degree of short-term depression ($\tau_{\mathrm{rec}} = 0.5$ s; P_{dis} before: 0.05, after: 0.5)

$$\frac{\mathrm{d}P_{\mathrm{dis}}}{\mathrm{d}t} \approx \rho_{\mathrm{up}}(1-P_{\mathrm{dis}})f_{\mathrm{post}}[f_{\mathrm{rel}}f_{\mathrm{post}} - \Theta_{\mathrm{up}}]^{+} - \rho_{\mathrm{dn}}P_{\mathrm{dis}}f_{\mathrm{rel}}[f_{\mathrm{post}}f_{\mathrm{rel}} - \Theta_{\mathrm{dn}}]^{+} \quad , \tag{2}$$

where $\rho_{\mathrm{up/dn}}$ determines the speed of LTP and LTD, respectively, and f_{rel} is the rate of neurotransmitter release given by the product of P_{rel} times the presynaptic firing rate f_{pre}. In the present case of synaptic depression, the average release probability is itself a function of the presynaptic firing rate, and one gets $f_{\mathrm{rel}} = P_{\mathrm{dis}}f_{\mathrm{pre}}/(1+\tau_{\mathrm{rec}}P_{\mathrm{dis}}f_{\mathrm{pre}})$. The NMDAR cross-gating would reduce the LTP and LTD terms both by the same factor of the form $1/(1+f_{\mathrm{rel}}+f_{\mathrm{post}})$ (see Senn et al. 2001b).

Equation (2) is a of third order in the rates, with the postsynaptic firing rate entering with a square in the LTP term, but not in the LTD term. As a consequence, LTD dominates at small f_{post}, while at higher f_{post} it is LTP which eventually dominates, as required by the Hebbian theory (Figs. 1d and 2b). Similarly, since the f_{rel} enters with a square in the second term, LTD may overcome LTP at high presynaptic frequencies (and weak synaptic depression). The imbalance towards LTD may be even enhanced if the postsynaptic frequency is small (Fig. 2b). This is exactly what is needed if the separation between two uncorrelated Hebbian ensembles should increase. The dominance of LTD over LTP at high presynaptic frequencies is also required for the synaptic weight normalization mechanism. Such a normalization will particularly affect those synapses which show both high presynaptic frequency and a high P_{dis}, and therefore those which are most responsible for the postsynaptic response.

To fully account for the Hebbian theory it is important that – over an intermediate range of pre- and postsynaptic frequencies – the synaptic modification is an increasing function of both these frequencies. Although according to the simplified rule (2) the synaptic change is monotonically decreasing with f_{rel} (and hence with f_{pre}), this can be different when taking the threshold nonlinearity in the LTD term into account: if the presynaptic activity remains below the LTD threshold (Θ_{dn}), only the LTP term remains which is in fact monotonically increasing with the pre- and postsynaptic rate (Fig. 2b). Hence, by carefully choosing the parameters in the presented algorithm, it is possible to encode activity patterns represented in a mean rate code (patterns based on spike correlations), and make use of the normalization property. To intuitively understand these multiple properties it may be helpful to interpret the third-order rule (2) as a second-order rule for a STDP with activity- and weight-dependent time windows. For a simplified model of STDP which directly incorporates higher order nonlinearities see Abarbanel et al. (2002).

As an example for coding with mean firing rates we show that the algorithm reproduces the main characteristics of the BCM rule originally developed to explain ocular deprivation experiments (Bienenstock et al.1982). Neglecting the thresholds in (2), we obtain for stationary Poisson frequencies

$$P_{\mathrm{dis}} \approx \frac{1}{1+\frac{\rho_{\mathrm{dn}}f_{\mathrm{rel}}}{\rho_{\mathrm{up}}f_{\mathrm{post}}}} \quad . \tag{3}$$

This formula shows that P_{dis} is in fact monotonically increasing with f_{post} (see Fig. 1d, solid line). When starting from any value of f_{post} (30Hz in Fig. 1c, d) and step-like increasing or decreasing f_{post}, the value of P_{dis} will tend with a slow time constant (τ_P) to a new equilibrium point. In the case of an upward jump it is LTP, and in case of a downward jump it is LTD which dominates (Fig. 1d). Since the same situation holds for the new equilibrium, the separation between LTP and

LTD now moves to the new value of f_{post}. This is the moving-threshold property of the BCM rule. Note that with only second-order terms in (2) no dependence of P_{dis} on the pre- and postsynaptic mean frequencies is obtained in steady-state conditions. A similar curve for the synaptic modification in response to random spike timings, as shown here for dP_{dis}/dt, is experimentally obtained by Sjöström et al. (2001). For an application of the synaptic modifications based on (3) to perceptual learning, see Adini et al. (2002).

3 Learning thresholds and direction selectivity

The development of direction selective cells in V1 through STDP is an example where spike correlations play the dominant factor in the synaptic modification. The example also assigns a functional role to the threshold nonlinearity observed in the STDP data (Fig. 1).

3.1 Model of direction selectivity

We consider a feedforward model with a single simple cell in V1 and with depressing thalamocortical afferents in the RF center, flanked by nondepressing synapses in the RF surround (Fig. 3a; see also Buchs and Senn 2001, 2002; Senn and Buchs 2002). The effective synaptic strength is given by the product of the synaptic density (of Gaussian shape) times the absolute synaptic strength G. Such a fixed synaptic density turned out to be important in stabilizing the development of DS. We chose a randomized, directionally unbiased stimulation scenario with drifting four-cycle gratings. The grating velocities were sampled from a Gaussian distribution with mean zero and standard deviation 6°, and the spatial wavelength matched the RF width (Fig. 3a). STDP was applied to both the absolute synaptic strength (G) and the degree of short-term depression (P_{dis}). Since nondirectional simple cells do not show any temporal distortions of the subthreshold membrane potential in response to a stationary grating (Jagadeesh et al. 1993), we chose the initial parameters such that the effect of synaptic depression vanished ($P_{dis} = 0.05$; see Fig. 3b).

3.2 Emergence of direction selectivity in an unbiased input scenario

A large-scale simulation, with roughly 5000 independent stochastic release sites at depressing or nondepressing thalamocortical synapses, shows that the scenario described above leads to DS resembling to that of simple cells in V1 (Fig. 4a, b; Buchs and Senn). The basic mechanism underlying DS is that the response of the depressing afferents in the (spatially asymmetric) RF center is slightly advanced compared to the response in the nondepressing RF surround (cf.Fig. 3). With a leftwards shift of the RF center, for instance, the simple cell becomes selective to motion from left to right, because the responses from the two synaptic populations slightly overlap in time (Fig. 4c, d). Note that due to the spatial distribution of the depressing synapses, the temporal phase advance of the current from the whole population is much smaller than the phase advance of an individual synapse (as shown in Fig. 3b). Nevertheless, the direction indices of the model cell are in a realistic regime (up to 0.9).

In the course of the stimulations the symmetry of the RF is broken, although the directions of motion of the stimuli are equally distributed. This is because of a positive-feedback loop according to which a first rightwards-moving grating, say, shifts the RF slightly to the left, and this in turn enhances the response to a grating moving in the same direction. On the other hand, the response to the opposite direction is reduced and therefore too weak to neutralize the initial RF shift. Beside modifying the synaptic strength, STDP enhances P_{dis} from an initial value of 0.05 to roughly 0.5, causing synaptic depression with the required phase advance of the synaptic response (Fig. 3b). Although this modification is again asymmetric, the parameters of the learning rule are tuned such that, starting with the small initial values, P_{dis} always increases.

The threshold nonlinearity in the STDP is essential for the robust development of DS. Without such learning thresholds ($\Theta_{up} = \Theta_{dn} = 0$ in a slightly adapted implementation of STDP; see Buchs and Senn 2002), any acquired DS is abolished by a few gratings moving in sequence in a temporarily nonpreferred direction (Fig. 5b). A robust, self-organizing development of DS also requires the dominance of LTD over LTP. Such a dominance causes a decrease of the simple-cell response during repeated stimulations until it is too small to induce further synaptic modifications. Due to the symmetry breaking of the RF, the response to one of the stimulus directions decreases strongly and eventually saturates at a low average response level (Fig. 5a). This average response to the nonpreferred direction may even be below the learning thresholds around 5 Hz (lower curve in Fig. 5a). As a consequence, gratings moving in the nonpreferred directions are not able to induce any synaptic changes, not even after many repetitions. Only if the presynaptic activity would also drop, and the balance between LTP and LTD therefore shift in favor of LTP (see Eq. 3), would the selectivity to the new direction again develop. This is due to the moving-threshold property which, in the case of monocular deprivation experiments, leads to LTP at the nondeprived, originally nondominating eye (Hubel and Wiesel 1962).

3.3 Frequency tuning of simple cells, STDP, and synaptic depression

How can a narrow time window for STDP of roughly 40 ms explain the development of DS at such low frequencies as 1–4 Hz (Fig. 4a)? One part of the explanation is that the simple cell itself represents a band-pass filter with largest response at these frequencies. Another part is that a time window of width α induces the largest RF shift if

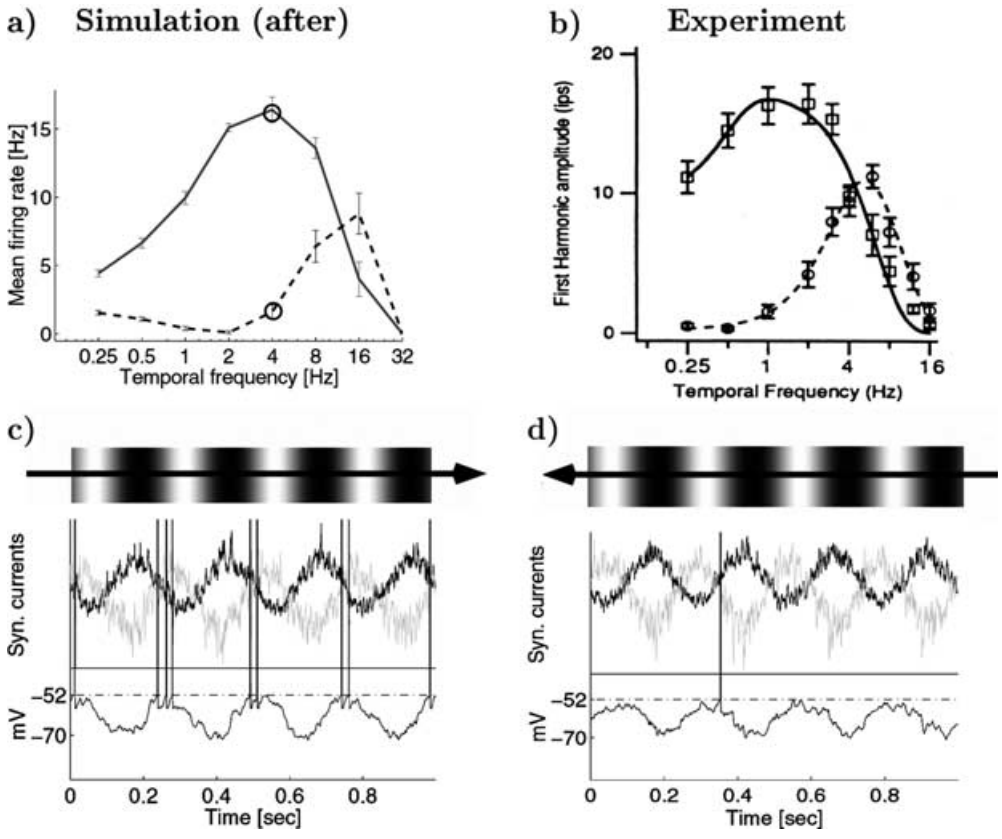

Fig. 4a–d. STDP generates direction selectivity for realistic temporal frequencies. **a** Response of the model simple cell to drifting gratings, after exposure to the unbiased stimulation scenario, showing the mean firing rate as a function of the temporal frequency for the preferred (*solid line*) and nonpreferred (*dashed line*) directions. Before the stimulation the two curves coincided. *Circles* correspond to the example shown in **c** and **d**. The optimal direction selectivity around 2 Hz results from three quantities, all being maximal around the same frequency: (i) the simple-cell response, (ii) the spatial phase advance induced by STDP (with a time window of 40 ms), and (iii) the temporal phase advance produced by synaptic depression (using a recovery time constant $\tau_{rec} = 0.5$ s), cf.Fig. 3. **b** An example of a temporal frequency tuning curve of a V1 cell (from Saul and Humphrey 1992). **c, d** Voltage train (*lower trace*), together with the total synaptic currents from the depressing and nondepressing synapses (*upper two traces*) in response to a 4-Hz grating moving in the preferred direction **c** and the nonpreferred direction **d**. The current from the depressing synapses (*trace with noisy downward peaks*) is slightly advanced for the preferred compared to the nonpreferred direction due to the summation of the spatial and temporal phase shifts. Adapted from Buchs and Senn (2002)

the temporal frequency of the stimulus is roughly $\omega = 1/(2\pi\alpha)$. With $\alpha = 40$ ms this gives a frequency of $\omega \approx 4$ Hz. A last part explaining the specific frequency bandwidth is that the short-term synaptic depression produces a phase advance which is maximal for a temporal frequency of $\omega = \sqrt{1 + P_{dis} f_{pre} \tau_{rec}}\ /(2\pi\tau_{rec})$. For $P_{dis} = 0.5$, $\tau_{rec} = 0.5$ s, and a presynaptic mean firing rate $f_{pre} = 40$spikes/s, for instance, we obtain a maximal phase advance at $\omega \approx 1$ Hz, which again is in the range of observed best frequencies (Senn and Buchs 2002). If the cell were exposed to a statistics of gratings with doubled frequencies, a similar preference would develop. Hence, rather than the environment, it is the specific parameters of the system that determine the best frequency bandwidth of the emerging DS.

4 Synaptic unreliability as a delay selection mechanism

STDP may also guide the development of axonal delays (see, e.g., Gerstner et al. 1996). It is not clear, however, how STDP can control the joint development of axonal *and* dendritic delays. The goal is to select delay lines with total axonal and dendritic delays that match the typical spike time difference between the pre- and postsynaptic cells. Since STDP is based on the local time difference between the synaptic transmission and the backpropagated action potential, the time window for LTP has to match the total forward and backward dendritic delay. Only then will synapses be strengthened which support the firing of the postsynaptic cell. However, STDP may also strengthen synapses which exhibit the same preferred local time difference, but do not contribute to the firing of the postsynaptic cell because they were activated too early or too late. To prevent the strengthening of these 'blind passengers' an additional selection mechanism is required. Synaptic unreliability is a possible mechanism which allows STDP to pick out only those synapses which also take part in the postsynaptic firing.

As an example we consider foveal RFs of V1 cells in the cat. Such cells respond to stimuli within a diameter of $\sim 2^\circ$. The subthreshold spatial integration field extends over an area of $\sim 10^\circ$ and integrates lateral

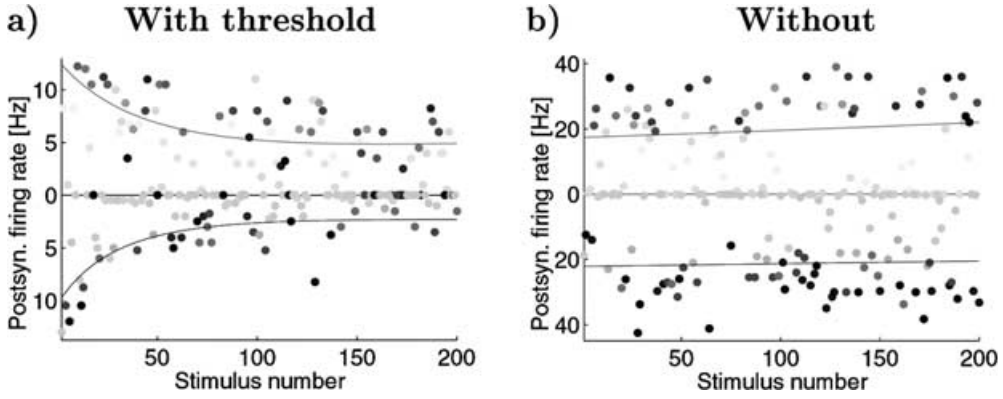

Fig. 5a,b. STDP in a natural environment requires learning thresholds. **a** Evolution of the simple-cell response to drifting (four-cycle) gratings with velocities sampled from a normal distribution with mean 0 and standard deviation 6 deg/s (and spatial frequency fixed to 1 cycle/deg). Responses to rightwards/leftwards-drifting gratings are plotted along the upward/downward ordinate, with *darker dots* corresponding to faster stimuli. Since LTD dominates LTP, the overall responses decay. The steady-state values of *lines* (defining the average responses to the stimuli in the two directions) reveal the development of stable direction selectivity to rightward-moving stimuli. **b** Same as in **a**, but without imposing learning thresholds for LTP and LTD. In this case, a few gratings moving in a nonpreferred direction may abolish any previously acquired direction selectivity, and no unique preference can develop. Adapted from Buchs and Senn (2002)

synaptic input over a cortical patch of ~1 cm. With an apparent speed for the horizontal propagation of 0.05–0.6 m/s the far-lateral inputs can be delayed with respect to the direct forward input from 20 to 200 ms (Bringuier et al. 1999). Is there a way for the postsynaptic cell to select a specific lateral delay within this range, say by guiding the growth of delay lines during development? In fact, STDP together with slow unbiased delay fluctuations can shift the average horizontal delay until the lateral synaptic input coincides with the expected forward input. An appropriate alignment of lateral and feedforward delay lines could explain DS for high-speed motions of 50–600 deg/s which run longitudinal to the preferred orientation (cf. also Chavane et al. 2000). There may also be considerable dendritic latency which, when measured from the centroid of the dendritic potential to the centroid of the forward-propagated somatic potential, can take values up to 40 ms (Agmon-Snir and Segev 1993). We formally show that in such a situation STDP may guide developmental shifts in the axonal and dendritic delays.

4.1 Shift of axonal delays at fixed spike correlations

Let us consider a single presynaptic cell (or a population of cells) A projecting through different delay lines onto neuron B, a layer V pyramidal cell, say. To explain the mechanism of axonal delay shifts we first assume a fixed spike-time correlation which does not depend on the timing of the input from A in B. Beside many uncorrelated spikes, for some stimuli cell A is assumed to fire before B, and these interspike intervals $\Delta t = t_A - t_B$ are Gaussian distributed with mean $\overline{\Delta t}$ (< 0) and variance ς^2. Let us further assume that the change of the synaptic strength G induced by a single pair of pre- and postsynaptic spikes is

$$\Delta G = \eta G \psi_\alpha(\Delta t_{\text{syn}}) \,, \tag{4}$$

where Δt_{syn} is the spike-time difference measured locally at the synapse, $\psi_\alpha(\Delta t)$ is the learning function shown in the inset of Fig. 6a with maximum at $\Delta t_{\text{syn}} = -\alpha$, and η is some small learning rate. The local spike-time difference, Δt_{syn}, is a function of Δt and the difference δ between the axonal delay and the backward dendritic delay:

$$\begin{aligned}\Delta t_{\text{syn}} &= (t_A + D_{\text{ax}}) - (t_B + D^{\text{b}}_{\text{den}}) \\ &= (t_A - t_B) + (D_{ax} - D^{\text{b}}_{\text{den}}) = \Delta t + \delta.\end{aligned}$$

From (4) one may derive the synaptic modifications for a whole population of delay lines characterized by different values for δ. Analysis of this population equation shows that, assuming LTD dominating LTP, the average delay difference $\overline{\delta} = \overline{D}_{\text{ax}} - \overline{D}^{\text{b}}_{\text{den}}$ changes proportionally to the derivative of the learning function ψ_α convolved with a Gaussian of width ς:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{\delta} \approx \sigma^2 \psi'_{\sqrt{\alpha^2+\varsigma^2}}(\overline{\Delta t} + \overline{\delta}) \,. \tag{5}$$

Here, $\psi_{\sqrt{\alpha^2+\varsigma^2}}$ is a stretched version of ψ_α with maximum at $-\sqrt{\alpha^2+\varsigma^2}$ instead of $-\alpha$. The learning rate of the average delay difference is roughly proportional to the variance σ^2 of the final delay-line distribution (Senn et al. 2001a). According to (5), the dynamics of $\overline{\delta}$ has an attracting fixed point determined by $\overline{\Delta t} + \overline{\delta} = -\sqrt{\alpha^2+\varsigma^2}$. Neglecting the delay of the backpropagating postsynaptic spike, $D^{\text{b}}_{\text{den}} = 0$, this fixed point condition translates to $t_A + \overline{D}_{\text{ax}} + \sqrt{\alpha^2+\varsigma^2} = t_B$. STDP implicitly changes the average axonal delay such that eventually the synaptic signal from A arrives at the dendritic site $\sqrt{\alpha^2+\varsigma^2}$ before the postsynaptic spike. Delay lines are strengthened if the delay fluctuations move the arrival time of the synaptic signal closer to α ms before the postsynaptic spike, and suppressed otherwise. This suppression is caused by the additional spikes of B which are uncorrelated to those of A.

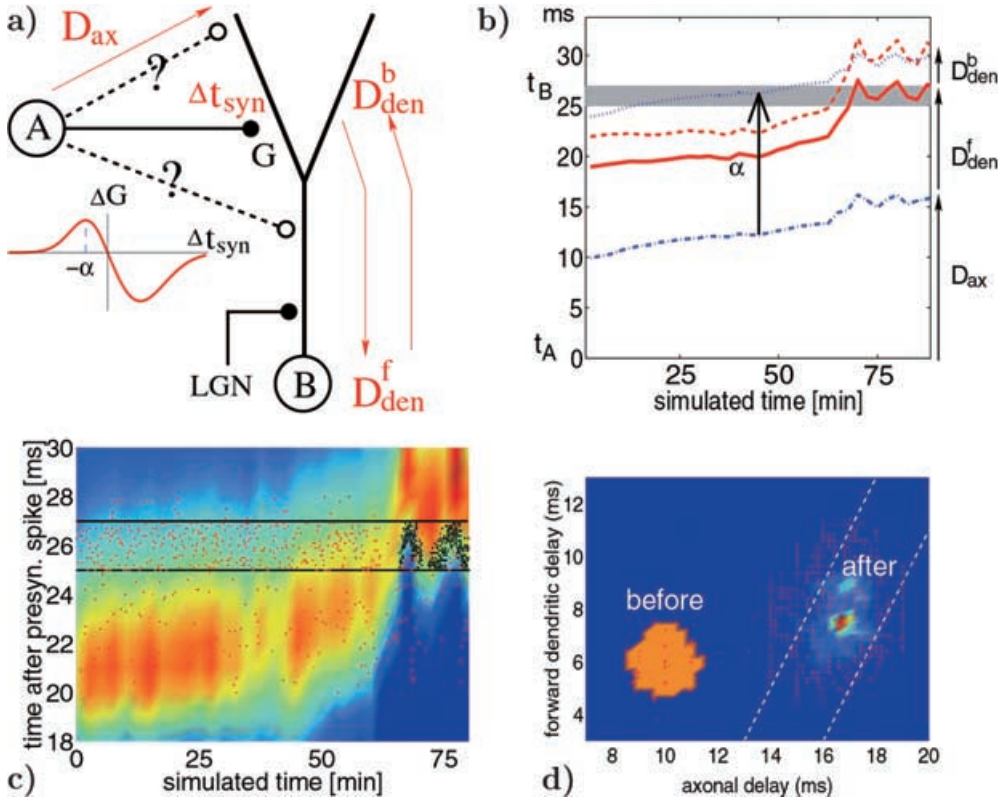

Fig. 6a–d. STDP and synaptic unreliability organize dendritic input locations. **a** The change of the synaptic strength (ΔG) depends on the local time difference between the pre- and postsynaptic signals at the synaptic site (Δt_{syn}), and the peak time of the learning window ($-\alpha$, cf.Eq. 4). It is determined by the firing time of A (t_{A}) and B (t_{B}), the average axonal delay (D_{ax}), and the average forward ($D^{\text{f}}_{\text{den}}$) and backward ($D^{\text{b}}_{\text{den}}$) dendritic delays. **b** 25–27 ms after the firing of A, the spontaneous spike rate of B is enhanced by additional subthreshold input from the lateral geniculate nucleus (LGN, *dashed horizontal region*). These spontaneous spikes 'attract' the input from A until it peaks at the time of the LGN input and eventually triggers by itself a spike in B. After this adaptation we have $-\Delta t_{\text{syn}} \approx D^{\text{f}}_{\text{den}} + D^{\text{b}}_{\text{den}} \approx \alpha$. The different *lines* are, from *bottom* to *top*, $t_{\text{A}} + D_{\text{ax}}$, $t_{\text{A}} + D_{\text{ax}} + D^{\text{f}}_{\text{den}}$, $t_{\text{A}} + D_{\text{ax}} + D^{\text{f}}_{\text{den}} + D^{\text{b}}_{\text{den}}$, and $t_{\text{A}} + D_{\text{ax}} + \alpha$. **c** The evolution of the depolarization induced by the input from A, together with the spike times of neuron B (*dots*) which are eventually triggered by A (*clustered dots* around 70min). **d** Synaptic unreliability makes axonal and (forward) dendritic delays unique (*left patch*, before stimulation; *right patch*, after stimulation). If the transmission were deterministic, synapses with the same Δt_{syn} ($= t_{\text{A}} + D_{\text{ax}} - t_{\text{B}} - D^{\text{b}}_{\text{den}}$) would be equally strengthened, irrespective of the time the signal peaks in B (*stripe between the two dashed lines*). We set $D^{\text{b}}_{\text{den}} = \frac{1}{3} D^{\text{f}}_{\text{den}}$. Adapted from Senn et al. (2001a)

Since for the correlated spikes, in the average, the response of A peaks in the soma of B at $t_{\text{A}} + \overline{D}_{\text{ax}} + \overline{D}^{\text{f}}_{\text{den}}$, cell A can only contribute to the firing of B if roughly $\overline{D}^{\text{f}}_{\text{den}} = \sqrt{\alpha^2 + \varsigma^2}$. In other words, if STDP should adapt the delay lines to support the firing of B, the average forward dendritic delay must roughly be the square-root sum of the learning window peak time, α, plus the jitter of the pre- and postsynaptic spike-time difference, ς. Note that $D^{\text{f}}_{\text{den}}$ is composed of the forward dendritic propagation delay plus the rise time of the postsynaptic somatic response triggered by A. For typical values of α and $D^{\text{f}}_{\text{den}}$ (both in the range 5–20 ms) the above relation is satisfied only if the jitter ς is small. If ς becomes larger, $\overline{\delta}$ shrinks according to (5). Hence, the noisier the spike-time difference from A to B is, the more that STDP will shorten the delay lines.

4.2 Joint shift of axonal and dendritic delays

If the backward dendritic delay, $D^{\text{b}}_{\text{den}}$, composed of the backpropagation time of the postsynaptic spike including the activation of additional processes involved in the induction of the synaptic modification is taken into account, STDP no longer selects a unique delay line. In fact, the learning rule cannot distinguish between delay lines with axonal and backward dendritic delays of the form $D_{\text{ax}} + \epsilon$ and $D^{\text{b}}_{\text{den}} + \epsilon$, because the argument $\Delta t_{\text{syn}} = (t_{\text{A}} + D_{\text{ax}}) - (t_{\text{B}} + D^{\text{b}}_{\text{den}})$ of the learning function depends on the difference between these delays only. A way to nevertheless choose between these delay lines is to introduce unreliable synaptic transmissions, and to modify the synaptic strength only if neurotransmitter is actually released. A synapse which then actively supports the generation of the postsynaptic spike will have a higher chance of being upregulated than a synapse which does not.

This is seen in the following scenario where cell B repeatedly receives subthreshold feedforward input from the lateral geniculate nucleus (LGN) some 25–27 ms after A fires. This input may cause a spontaneous spike elicited within the same time interval, and that spike in turn triggers STDP at the synapses from A. By these modifications, the axonal delays from A to B slowly shift

until the input from A reliably drives B within the prescribed interval (Fig. 6). The delay line with $t_A + D_{ax} + D^f_{den} \approx t_B$ is the 'fittest' in triggering postsynaptic spikes, and has the highest chance of being upregulated. Inappropriate delay lines with the same Δt_{syn} but with transmission failures would not profit from the 'correct' spike timing. Rather, due to LTD dominating LTP these delay lines will eventually be suppressed (Fig. 6d). The unique selection of the axonal and dendritic delays depends on the peak time of the learning function (α), and implicitly determines the optimal synaptic position on the dendritic tree.

5 Discussion

We showed that the nonlinearities present in our synaptic plasticity algorithm shares several interesting features of STDP and rate-based Hebbian plasticity, including the recent results on triple spike stimulations (Froemke and Dan 2002) and several properties of the BCM rule (bienenstock et al. 1982). Moreover, the predictions of the algorithm when applying pre- and postsynaptic Poisson firing rates are qualitatively confirmed by recent experiments showing the dominance of LTP over LTD at elevated postsynaptic frequencies (Sjöström et al. 2001). The dependency of the synaptic plasticity on the postsynaptic voltage revealed by these experiments is similar to the dependency on the model variable N_{up}: if, in analogy to hyperpolarizing the postsynaptic neuron, N_{up} is reset to zero between two successive pairs of spikes, LTP is prevented due to the threshold nonlinearity. On the other hand, similar to the effect of the postsynaptic depolarization, a single pair of pre–post spikes would trigger LTP if N_{up} would be artificially elevated. In fact, N_{up}, which models the fraction of NMDARs relieved from the Mg^{2+} block, is a monotonically increasing function of the postsynaptic membrane potential. It remains to be shown that LTP is induced by a high presynaptic spike rate and strong, in-vivo-like subthreshold depolarizations, while LTD is induced if the postsynaptic depolarization is only weak (see also Fusi et al. 2000).

There is, however, an important difference between the experimental results of Sjöström et al. (2001) and our algorithm. While the in vitro experiments did only reveal a threshold for LTP, such a LTD threshold is required for a stable development of DS in an ever-changing environment. Without LTD threshold, the activity induced by stimuli moving in the nonpreferred direction will destroy the synaptic structure required for DS. Our analysis also showed that the time window of STDP ($\sim$40 ms) supports the development of DS in a realistic temporal frequency regime, and that metaplasticity would be required to develop appropriate DS in an environment where the stimuli move faster. An important feature in our DS model is that STDP not only changes the synaptic strength, but also the time course of the short-term depression. Again, the time constant of synaptic depression ($\sim$200 ms) is tuned to support DS in a realistic frequency bandwidth, and an adaptation of this parameter would be needed to cope with other temporal frequencies.

Selection in biology is often based on some source of randomness. We showed that in the case of interneuronal delay lines, the selection of the appropriate axonal delay requires slow delay fluctuations, and the selection of the appropriate dendritic delay requires unreliable synaptic transmissions. Activity-dependent delay shifts governed by STDP may play an important role in shaping the subthreshold integration field of V1 cells which integrate over horizontal cortical distances of $\sim$1 cm with propagation delays up to 200 ms (Bringuier et al. 1999). A similar organizational principle might be at work at intercortical connections. The timing of these connections could implicitly be adjusted through delayed 'reward' signals projecting from an upstream cortical area back to a downstream area. This downstream area may then modify its response strength based on the timing of the reward signal in much the same way as the synaptic weight is modified by the timing of the backpropagating action potential. Such a mechanism, for instance, could support habit learning within the basal ganglia thalamocortical feedback loop (Senn et al. 2001a).

The sensitivity of the synaptic plasticity to the timing of individual pre- and postsynaptic spikes should not obscure the fact that it is only the spike-time correlation in a statistical sense which counts. In our DS simulations, for instance, the spike-time variability of an individual presynaptic spike train is larger than the width of the learning window. DS only develops because of the large number of independent LGN afferents. Asynchronous firing may even be necessary to protect the postsynaptic cell against resonances in the presynaptic spike times (Buchs and Senn 2002). Synaptic unreliability, moreover, is required if axonal and dendritic delays should consistently evolve. The view of the statistical nature of the spike-time code was formulated by von Neumann (1958), a contemporary of Hebb, as follows: ''It is therefore perfectly plausible that certain (statistical) relationships between such trains of pulses should also transmit information. ...instead of the precise system of markers where the position – presence or absence – of every marker counts decisively in determining the meaning of the message, we have here a system of notations in which the meaning is conveyed by the *statistical* properties of the message.''

Acknowledgements. This study was supported by the Swiss National Science Foundation (grant 3152-065234.01) and the Silva-Casa foundation. The author thanks Stefano Fusi, Henry Markram, and Misha Tsodyks for helpful discussions, Nissim Buchs and Martin Schneider for their simulations, and Jan Reutimann for proof reading.

References

Abarbanel HDI, Huerta R, Rabinovich MI (2002) Dynamical model of long-term synaptic plasticity PNAS 99: 10132–10137

Adini Y, Sagi D, Tsodyks M (2002) Context-enabled learning in the human visual system. Nature 415: 790–793

Agmon-Snir H, Segev I (1993) Signal delay and input synchronization in passive dendritic structures. J Neurophysiol 70: 2066–2085

Amit D (1997) The Hebbian paradigm reintegrated: local reverberations as internal representations. Behav Brain Sci 18: 617

Bi G, Poo M (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 18: 10464–10472

Bienenstock E, Cooper L, Munro P (1982) Theory for the development of neuron selectivity: orientation specifity and binocular interaction in visual cortex. J Neurosci 2: 32–48

Bringuier V, Chavane F, Glaeser L, Frégnac Y (1999) Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. Science 283: 695–699

Buchs N, Senn W (2001) Learning direction selectivity through spike-timing dependent modification of neurotransmitter release probability. Neurocomputing 38-40: 121–127

Buchs N, Senn W (2002) Spike-based synaptic plasticity and the emergence of direction selective simple cells: simulation results. J Comput Neurosci 13: 167–186

Chance F, Nelson S, Abbott L (1998) Synaptic depression and the temporal response characteristics of V1 cells. J Neurosci 18: 4785–4799

Chavane F, Monier C, Bringuier V, Baudot P, Borg-Graham L, Lorenceau J, Fregnac Y (2000) The visual cortical association field: a gestalt concept or a psychophysiological entity? J Neurophysiol 94: 333–342

Feldman D (2000) Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex. Neuron 27: 45–56

Froemke R, Dan Y (2002) Spike-timing-dependent synaptic modification induced by natural spike trains. Nature 416: 433–438

Fu Y-X, Djupsund K, Gao H, Hayden B, Shen K, Dan Y (2002) Temporal specifcity in the cortical plasticity of visual space perception. Science 296: 1999-2003

Fusi S, Annunziato M, Badoni D, Salamon A, Amit DJ (2000) Spike-driven synaptic plasticity: theory, simulation, VLSI implementation. Neural Comput 12: 2227–2258

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

Hubel D, Wiesel T (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol (Lond) 160: 106–154

Jagadeesh B, Wheat H, Ferster D (1993) Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. Science 262: 1901–1904

Katz L, Shatz C (1996) Synaptic activity and the construction of cortical circuits. Science 274: 1133–1138

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

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

Markram H, Tsodyks M (1996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. Nature 382: 807–810

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

Neumann J von (1958) The computer and the brain. Yale University Press, New Haven, Conn.

Nieuwenhuys R (1994) The neocortex. An overview of its evolutionary development, structural organization and synaptology. Anat Embryol (Berl) 190: 307–337

Saul A, Humphrey A (1992) Temporal frequency tuning of direction selectivity in cat visual cortex. Vis Neurosci 8: 365–372

Senn W, Tsodyks M, Markram H (1997) An algorithm for synaptic modification based on exact timing of pre- and post-synaptic action potentials. In: Gerstner W, Germond A, Hasler M, Nicoud J-D (eds) Artificial neural networks – ICANN'97 (Lecture notes in computer science, vol 1327), pp 121–126

Senn W, Schneider M, Ruf B (2001a) Activity-dependent development of axonal and dendritic delays, or, why synaptic transmission should be unreliable. Neural Comput 14: 583–619

Senn W, Tsodyks M, Markram H (2001b) An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. Neural Comput 13: 35–68

Senn W, Buchs N (2002) Spike-based synaptic plasticity and the emergence of direction selective simple cells: mathematical analysis. J Comput Neurosci, in press

Shadlen M, Newsome W (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18: 3870–3896

Sjöström P, Turrigiano G, Nelson S (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32: 1149–1164

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

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

Yao H, Dan Y (2001) Stimulus timing-dependent plasticity in cortical processing of orientation. Neuron 32: 315–323