Journal of Computational Neuroscience 13, 167–186, 2002

Spike-Based Synaptic Plasticity and the Emergence of Direction Selective Simple Cells: Simulation Results

N.J. BUCHS AND W. SENN
Physiological Institute, University of Bern, Bühlplatz 5, CH-3012 Bern
buchs@cns.unibe.ch
wsenn@cns.unibe.ch

Received February 12, 2002; Revised May 21, 2002; Accepted May 24, 2002

Action Editor: Jonathan D. Victor

Abstract. Direction selectivity (DS) of simple cells in the primary visual cortex was recently suggested to arise from short-term synaptic depression in thalamocortical afferents (Chance F, Nelson S, Abbott L (1998), J. Neuroscience 18(12): 4785–4799). In the model, two groups of afferents with spatially displaced receptive fields project through either depressing and non-depressing synapses onto the V1 cell. The degree of synaptic depression determines the temporal phase advance of the response to drifting gratings. We show that the spatial displacement and the appropriate degree of synaptic depression required for DS can develop within an unbiased input scenario by means of temporally asymmetric spike-timing dependent plasticity (STDP) which modifies both the synaptic strength and the degree of synaptic depression. Moving stimuli of random velocities and directions break any initial receptive field symmetry and produce DS. Frequency tuning curves and subthreshold membrane potentials akin to those measured for non-directional simple cells are thereby changed into those measured for directional cells. If STDP is such that down-regulation dominates up-regulation the overall synaptic strength adapts in a self-organizing way such that eventually the postsynaptic response for the non-preferred direction becomes subthreshold. To prevent unlearning of the acquired DS by randomly changing stimulus directions an additional learning threshold is necessary. To further protect the development of the simple cell properties against noise in the stimulus, asynchronous and irregular synaptic inputs are required.

Keywords: synaptic plasticity, spike-timing dependent plasticity, synaptic depression, direction selectivity, activity-dependent development, simple cell

1. Introduction

The temporally asymmetric spike-timing dependent plasticity (STDP) which was recently observed in neocortex and other brain regions shows a remarkable sensitivity to the timing between pre- and postsynaptic spikes (Markram et al., 1997; Zhang et al., 1998; Bi and Poo, 1998; Feldman, 2000). According to these experiments, the synaptic efficacy is upregulated if the synapse is activated 10 ms before the occurrence of a postsynaptic action potential, while it is downregulated if the synapse is active 10 ms after the postsynaptic action potential. This high sensitivity, on the other hand, is opposed to the high irregularity of the neuronal firing and the stochastic nature of synaptic transmission. In vivo spike trains show an inter spike interval (ISI) variability which is in the same order of the ISIs themselves (Softky and Koch, 1993), typically larger than 20 ms. In the visual system these neurons are often driven by noisy stimuli which introduce further variability. How then is it possible to meaningfully extract and encode visual features by Hebbian mechanisms if

the sensitivity of the synaptic rule is below the noise level of the neural code? The paradox becomes even more striking when one considers that simple cells in the primary visual cortex (V1) are maximally direction selective for slowly moving stimuli crossing the receptive field (of roughly 1 deg) within one second (Saul and Humphrey, 1992b; Hawken et al., 1996). How can the cells learn to discriminate between stimuli extending over one second if the learning window itself is less than 100 ms. How can cells acquire direction selectivity (DS), if in a natural environment moving stimuli are constantly changing their direction and velocity?

To investigate these questions we modeled a single simple cell in the primary visual cortex with a realistic number of afferents projecting from the lateral geniculate nucleus (LGN) through either depressing or non-depressing synapses. We implemented a stochastic model of synaptic transmission with short-term depression resulting from vesicle depletion. The degree of synaptic depression is represented by the probability of vesicle discharge, P_{dis}, given a spike and a ready releasable vesicle. Direction selectivity in our model is based on short-term synaptic depression in thalamocortical connections as suggested by Chance et al. (1998). As a consequence we consider STDP acting on the absolute synaptic strength as well as on the degree of synaptic depression. Visual stimuli consisted of light gratings moving across the receptive field (RF) with randomly sampled velocities and directions.

The simulations show that for a directionally unbiased set of moving stimuli, in the presence of irregularly spiking LGN neurons and stochastic synaptic transmission, it is in fact possible to acquire the simple cell direction selectivity (DS) with similar properties as observed in vivo. The large number of independent synaptic release sites play a key role in obtaining these results. Although a single synapse is not able to reliably decide whether the presynaptic activity in general leads or lags the postsynaptic activity, a large population may do so. A large number of afferents, together with the asynchrony of the spike trains, also gives the system robustness against noise in the stimulus. The simulations further confirm the importance of the competitive aspect of STDP stressed by Song et al. (2000), which in our case leads to the stable formation of the simple cell RF with realistic response frequencies under different stimulation protocols. However, additional nonlinearities are necessary to stabilize the acquired RF properties against direction changes. We show that in a natural environment, with stimuli moving at different speeds in different directions, the selectivity can only be stabilized if we impose a learning threshold to the postsynaptic activity which prevents unlearning of any previously acquired selectivity. Without such a threshold, the simple cell will switch its preferred direction if a large number of stimuli move in the opposite direction.

The basic mechanism making our simple cell directionally selective is the temporal phase advance of the synaptic response attributed to synaptic depression (Chance et al., 1998). Combining depressing and non-depressing afferents with a spatial offset will produce coincident synaptic input to the simple cell if stimulated in the preferred direction, but asynchronously timed input if stimulated in the non-preferred direction. Synaptic depression and the corresponding phase advance only occurs if P_{dis} is large (say 0.8), and it is virtually not present if P_{dis} is small (say 0.1). Hence, a developmental model explaining the specific spatial arrangement and the different response characteristics of the two groups of synaptic afferents needs to modify both the absolute synaptic strength and the degree of depression encoded by P_{dis}. In fact, both types of synaptic plasticity are found in neocortex (Feldman, 2000; Markram and Tsodyks, 1996, respectively). We show that an activity-dependent modification of these two synaptic parameters may turn a non-directional model simple cell into a directional simple cell, each characterized by the corresponding temporal tuning curve and the subthreshold membrane potential responses. The LGN activity driving the described synaptic plasticity may either come from prenatal waves of retinal activity, or from postnatal activity before or after eye opening (cf. Discussion). Additional experimental support for the required forms of STDP in adult V1 comes from recent in vivo recordings in cats showing clear short-term plasticity of the RF which depends on the millisecond timing between repeated pairings of stimuli (Gao et al., 2001). In these experiments, repeatedly presenting two spatially displaced visual stimuli with a delay of roughly 10 ms shifted the RF of the cortical cell towards the location of the second stimulus, suggesting a functional relevance of STDP in the primary visual cortex (see also Djupsund et al., 2001).

Models for simple cell DS have a long history, but the question which of the various directional mechanisms eventually dominates the simple cell response characteristics is still controversial. One line of models follows a feedforward scheme which is based on spatially dislocated excitatory/inhibitory (Watson and

Ahumada, 1985; Borst and Egelhaaf, 1989; Heeger, 1993) or excitatory/excitatory afferents (Adelson and Bergen, 1985; Chance et al., 1998) with different temporal response properties. Another, more recent line of models stresses the importance of recurrent feedback connections in generating DS (Suarez et al., 1995; Maex and Orban, 1996; Mehta, 2000). Models investigating developmental aspects of directional simple cells exploit the lagged/non-lagged paradigm of LGN cell responses (Feidler et al., 1997; Wimbauer et al., 1997), network effects (Nagano and Fujiwara, 1979; Rao and Sejnowski, 2000), or simple cell inhibitory feedback (Mehta, 2000). The present model is of the feedforward type and focuses on computational aspects of STDP and short-term synaptic depression. To this end we neglect cortico-cortical connections and the heterogeneity of the LGN latencies. In reality, the different putative directional mechanisms might well work together.

An early version which explains the emergence of DS by a pure redistribution of synaptic efficacy through modification of the vesicle discharge probability, without adapting the absolute synaptic weights, appeared as an extended abstract (Buchs and Senn, 2001).

2. Model and Methods

Our model encompasses the visual pathway from the retinal input through the LGN to the activity of a simple cell in the primary visual cortex. Special attention is devoted to the synaptic mechanisms between the LGN neurons and the simple cell.

2.1. Receptive Field of LGN Neurons and the Simple Cell

As a visual stimulus we consider either stationary counterphase gratings or drifting gratings moving with different speeds from the left or the right through the visual field. This stimulus drives the retinal ganglion cells which themselves project onto LGN cells. To describe the transfer function from the light stimulus to the LGN output we follow Maex and Orban (1996) in the choice of the spatio-temporal linear filter and the subsequent half rectification. We use the analytical expression derived in the 'analysis paper' (Senn and Buchs, 2002) to calculate the instantaneous Poisson firing rate of the on- and off-center LGN cells. In response to a drifting sinewave grating a LGN cell at retinal position x^i (measured in degrees) fires with an instantaneous Poisson rate $f^i_{pre}(t) = \max\{\pm A(C)A_\circ(k, \Omega)\cos(kx^i - \Omega t), f^{back}\}$. The quantities $C, k/2\pi$, and $\Omega/2\pi$, describe the luminance contrast, the spatial frequency (in cycles/deg), and the temporal frequency (in Hz) of the drifting grating, respectively. The function $A(C)$ represents a logarithmic fit of the contrast frequency curve from Ohzawa et al. (1985), and $A_\circ(k, \Omega)$ captures the frequency dependency of the linear spatio-temporal LGN filter (for an explicit definition see the analysis paper). On- and off-center LGN cells are distinguished by the choice of the sign $+$ or $-$ in front of the amplitude factors, respectively. To account for the refractory time of LGN cells, spikes obtained from the Poisson spike generator are discarded if they followed the previous spike within 3 ms. Unless stated differently, we choose a stimulus contrast of $C = 0.5$, a spatial frequency of $k/2\pi = 1$ cycles/deg, different temporal frequencies $\Omega/2\pi$, and a background rate of $f^{back} = 5$ Hz. With these parameters, the maximal LGN firing rate is 60 Hz. For the subthreshold voltage experiments we use counterphase gratings leading to LGN responses of the form $f^i_{pre}(t) = \max\{\pm A\cos(kx^i)\cos(\Omega t), f^{back}\}$.

The present description of the LGN firing rates assumes periodic full field stimulations. Since transient deviations from the periodic response are only observed during the first cycle we apply the same formulas to our generalized stimulation scenario with 4 consecutive cycles drifting in randomly chosen directions. To simplify the analysis we consider gratings with only optimal spatial frequency and optimal orientation. The reasoning is that suboptimal gratings, and probably also natural stimuli which are composed of a superposition of many such local suboptimal gratings, are less efficient in driving the simple cell, and hence are less efficient in inducing RF modifications. In turn, optimal gratings with opposite drifting directions are most efficient in annihilating each others RF modifications. Hence, to study RF development we concentrate on the dominant stimuli, and vary their features along one dimension, the temporal frequency of the gratings. To isolate the effect of short-term plasticity and STDP on thalamocortical synapses we neglect cortico-cortical input onto the model simple cell, although it represents perhaps 90% of the whole input (Ahmed et al., 1994). Moreover, feedforward input is modeled monosynaptically, despite that inhibitory input is mediated by cortical interneurons and hence is at least disynaptic.

The simple cell RF is implemented by a push-pull mechanism with excitatory on-center and inhibitory off-center afferents located at each site within the

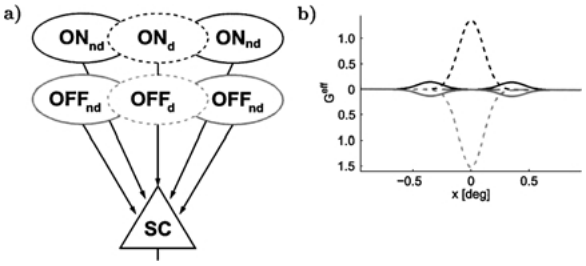

Figure 1. Initial synaptic arrangement. (**a**) Schematic drawing of the initial spatial arrangement of on- and off-center LGN cells projecting through excitatory (bright ovals) and inhibitory (dark ovals) synapses, respectively, onto the model simple cell in V1. The middle clusters of afferents show synaptic depression (dashed line, index d). (**b**) The initial effective synaptic strength $\bar{G}^{eff}$ of the different types of (depressing/non-depressing, excitatory/inhibitory) afferents. For better visualization, the inhibitory conductances from the off-center cells are plotted downwards. (Dark dashed line: $\bar{G}_d^{eff,E}(x) = \rho_d^E(x)\bar{G}_d^E(x)$; dark solid line: $\bar{G}_{nd}^{eff,E}(x)$; bright dashed line: $\bar{G}_d^{eff,I}(x)$; bright solid line: $\bar{G}_{nd}^{eff,I}(x)$. The strength of the depressing synapses (dashed lines) is scaled down by a factor of 0.5.)

simple cell RF. Both types of afferents are modeled as monosynaptic connections from the LGN to the simple cell. In an initial state the simple cell RF is symmetric. Afferents located in the center of the simple cell RF project through depressing synapses and are either excitatory or inhibitory, depending on whether they are on- or off-center, respectively (Fig. 1a). Similarly, on- and off-center afferents located in the RF surround project through non-depressing excitatory and inhibitory synapses onto the simple cell, respectively. For simplicity we only consider optimally oriented gratings and, since the structure of the simple cell RFs extends along this optimal orientation, we restrict ourselves to a one-dimensional simple cell RF. The different clusters of excitatory on-center and inhibitory off-center afferents are modeled by Gaussian distributions specifying the effective synaptic strength. This effective synaptic strength, $\bar{G}^{eff}(x)$, is characterized by the product of the LGN cell density, $\rho(x)$, associated with the retinal position x, and the strength of an individual synaptic connection, $\bar{G}(x)$ (Fig. 1b, see also Section 2 in the analysis paper). The fact that the fixed LGN cell density is itself Gaussian makes the development of the simple cell RF much more robust against random stimulus fluctuations compared to a flat LGN density (cf. point 5 in the discussion). The current experimental literature provides evidence neither for nor against a spatial subdivision of the simple cell RF into depressing and non-depressing synapses. Our simulations, however, show that an initial symmetric subdivision of the RF is necessary if strong DS is to evolve from the same type of STDP applied to depressing and non-depressing synapses (cf. point 7 in the discussion).

In our simulations we spatially distributed 800 afferents among each of the 6 clusters in Fig. 1b such that the densities in each cluster became Gaussian. The width of the on- and off-cluster in the middle was $\sigma_d = 0.15$ deg and the width of the surrounding clusters positioned around $x_{nd} = \pm 0.35$ deg was $\sigma_{nd} = 0.15$ deg. In most of our simulations we had a total number of 4800 afferents from the LGN onto the simple cell. This number may be too high considering that the total number of synapses onto the simple cell is in the range of tens of thousands and only 10% of these are assigned to thalamocortical afferents (Ahmed et al., 1994). On the other hand, each synaptic bouton may have multiple release sites and each release site several releasable vesicles, each contributing to the stochasticity of the postsynaptic current. Tests with a total of 600 synapses gave similar qualitative results, but the chance for developing DS was lowered (see Results).

2.2. *Synaptic Transmission and Short-Term Depression*

To consider a realistic degree of jitter in the synaptic transmission we modeled a stochastic neurotransmitter release, and hence follow each of the (4800) vesicles of the thalamocortical connections individually. For the non-depressing synapses the probability of neurotransmitter release, P_{rel}, is fixed to 0.5. For depressing synapses P_{rel} is activity-dependent because the vesicle recovery process has a relatively slow time constant of $\tau_{rec} = 150$ ms. Immediately after vesicle discharge, the readily releasable pool is depleted. It is stochastically refilled by a Poisson process, i.e. at any time step dt (=1 ms) the vesicle recovers with probability dt/τ_{rec}. Formally, $P_{rel} = P_{dis}P_v$, where P_v is the probability of a vesicle being recovered, and P_{dis} is the vesicle discharge probability (given a spike and a recovered ready releasable pool). This is a stochastic version of the depressing synapse model of Tsodyks and Markram (1997), see also Senn et al. (2001) and Section 3.1 in the analysis paper. P_{dis} is modified by the synaptic learning rule and is initially set to 0.03.

2.3. *Simple-Cell Model and Receptor Dynamics*

The simple cell in V1 is modeled as an integrate-and-fire neuron receiving excitatory and inhibitory input

from on- and off-center LGN cells, respectively, as described above. Both types of input project through depressing and non-depressing synapses, and the subthreshold membrane potential is governed by

$$\tau_m \frac{dV}{dt} = V_{rest} - V + G^E_{nd}(V_E - V) + G^E_d(V_E - V) + G^I_{nd}(V_I - V) + G^I_d(V_I - V). \quad (1)$$

The membrane time constant is $\tau_m = 30$ ms, the resting potential $V_{rest} = -70$ mV, and the reversal potential for the excitatory (inhibitory) synapses is $V_E = 0$ mV ($V_I = -100$ mV). When the membrane potential V reaches the threshold value of -52 mV, an action potential is fired, and the potential is reset to $V_{reset} = -58$ mV. To take into account the absolute refractory time we clamp the membrane potential after each spike for 3 ms at V_{reset}. The G's represent the activity-dependent (dimensionless) total synaptic conductances corresponding to the four types of connections. (Recall that $G = gR_m$ and $\tau_m = C_m R_m$, where g represents the corresponding total synaptic conductances, R_m the total membrane resistance, and C_m the total membrane capacitance.)

The dynamics of an individual synaptic conductance is modeled by a first order kinetics with instantaneous response to a synaptic release. For instance, for an off-center LGN cell at location x^i projecting through a depressing synapse, the synaptic conductance $G^{I,i}_d$ increases by $\bar{G}^{I,i}_d$ immediately after each synaptic release, and then decays exponentially towards zero with time constant $\tau_G = 2$ ms. The total synaptic conductance of these inhibitory depressing synapses is the sum of the individual conductances, $G^I_d = \sum_i G^{I,i}_d$. The individual synaptic peak conductances before their modification are chosen to be $\bar{G}^{E/I,i}_d = 0.2$ and $\bar{G}^{E/I,i}_{nd} = 0.01$. These synaptic conductances were calibrated to obtain a realistic time course for the simple cell subthreshold membrane potential in response to stationary gratings (Fig. 8a).

2.4. Synaptic Modifications

The implementation of the temporally asymmetric synaptic learning rule is a reduced version of the one in Senn et al. (2001), which was designed to fit the experimental data from neocortical layer 5 pyramidal cells (Markram et al., 1997). We applied the rule not only to the discharge probability P_{dis} of the depressing synapses, but also to the synaptic strengths $\bar{G}^{E,i}_{d/nd}$ of the individual excitatory depressing and non-depressing connections. Roughly speaking, the rule upregulates these synaptic parameters at each postsynaptic spike, depending on the previous presynaptic activity, and it downregulates them at each presynaptic spike, depending on the previous postsynaptic activity. To keep track of the pre- and postsynaptic activities we introduce variables C^i_{pre} and C^i_{post}, alluding to some pre- and postsynaptic calcium concentrations. The variables are instantaneously increased by 1 at each presynaptic *release* and postsynaptic spike, respectively, and otherwise decay towards zero with time constants $\tau^C_{pre} = 20$ ms and $\tau^C_{post} = 80$ ms (Fig. 2). To take into account the strong nonlinear increase of the synaptic change as a function of the firing rate around 8 Hz observed by Markram et al. (1997) we introduce secondary messenger S and apply postsynaptic thresholds θ for the induction of the synaptic modifications (cf. Senn et al., 2000). Again, the variables S^i_{pre} and S^i_{post} are increased by 1 immediately after a presynaptic release and a postsynaptic spike, respectively, and they

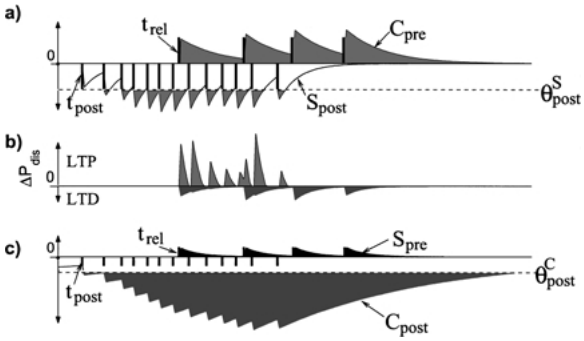

Figure 2. The synaptic modifications according to Eq. (2). (**a**) LTP is specified by an internal presynaptic variable (C_{pre}, upper line) which is increased at the time of a presynaptic release (t_{rel}) and decays with a time constant of $\tau^C_{pre} = 20$ ms, and an internal postsynaptic variable (S_{post}, lower line, clipped downwards) which is increased at the time of a postsynaptic spike (t_{post}) and decays with a shorter time constant $\tau^S_{post} = 10$ ms. Only the hatched area between S_{post} and the learning threshold (θ^S_{post}, dashed line) enters in the Hebbian LTP term. (**b**) The effective synaptic modification. Above the line: the product of the two hatched areas in (a), C_{pre} times $[S_{post} - \theta^S_{post}]^+$, specifying the LTP term in Eq. (2). Below the line: the product of the two hatched areas in (c), S_{pre} times $[C_{post} - \theta^C_{post}]^+$, specifying the corresponding LTD term. (**c**) LTD is specified by an internal postsynaptic variable (C_{post}, lower line, clipped downwards) which is increased at the time of a postsynaptic spike (t_{post}) and decays with the long time constant $\tau^C_{post} = 80$ ms, and an internal presynaptic variable (S_{pre}, upper line) which is increased at the time of a presynaptic spike (t_{rel}) and decays with $\tau^S_{pre} = 10$ ms. Only the hatched area between C_{post} and the learning threshold (θ^C_{post}, dashed line) enters in the Hebbian LTD term shown in (b).

decay both with a time constant of $\tau^S_{pre/post} = 10$ ms. The synaptic modification is composed of two Hebbian terms describing long-term potentiation (LTP) and long-term depression (LTD) of the specific parameter. LTP is the product of the presynaptic calcium concentration times the thresholded postsynaptic messenger concentration, and LTD is the product of the thresholded postsynaptic calcium concentration times the presynaptic messenger concentration (Fig. 2). For the modification of the discharge probability we therefore have

$$\frac{dP^i_{dis}}{dt} = r^{up}_P \left(P^{max}_{dis} - P^i_{dis}\right) C^i_{pre} \left[S^i_{post} - \theta^S_{post}\right]^+ - r^{dn}_P P^i_{dis} S^i_{pre} \left[C^i_{post} - \theta^C_{post}\right]^+. \quad (2)$$

The saturation factors $(P^{max}_{dis} - P^i_{dis})$ and P^i_{dis} in front of the Hebbian terms ensure that P_{dis} is not driven beyond the physiological regime. We put $[x]^+ = \max\{x, 0\}$, $P^{max}_{dis} = 1$, and the rates are set to $r^{up}_P \equiv r^{up,E}_{P,d} = 2.5$ and $r^{dn}_P \equiv r^{dn,E}_{P,d} = 0.25$. The adaptation of the excitatory synaptic strengths, $\bar{G}^{E,i}_{d/nd}$, follows the same rule (2) with P^i_{dis} replaced by $\bar{G}^{E,i}_{d/nd}$. The corresponding parameter values are $\bar{G}^{max}_{d/nd} = 1/0.1$, $r^{up,E}_{G,d} = 0.5$, $r^{up,E}_{G,nd} = 2$, $r^{dn,E}_{G,d} = 0.9$, and $r^{dn,E}_{G,nd} = 0.25$.

The synaptic modifications introduced so far concerned excitatory connections only, and previous simulations show that these are in fact sufficient to explain the emergence of DS (see Buchs and Senn (1999), and the analysis paper). However, intracellular recordings of subthreshold membrane potentials of simple cells reveal the presence of strong inhibitory afferents, which show a distinct temporal behavior for directional and non-directional cells (Jagadeesh, 1993, see Fig. 8). To account for these differences we also subject the inhibitory afferents to STDP. Since we model inhibition monosynaptically, synaptic plasticity for these afferents has to be modeled by the reverse learning rule. This would effectively correspond to the original rule operating on excitatory thalamocortical synapses which project through inhibitory interneurons onto the simple cell. Hence, modeling direct thalamocortical inhibition, we modify P^i_{dis} and $\bar{G}^{I,i}_{d/nd}$ of the inhibitory cells according to (2), but with the reversed sign in front of both terms, and with the corresponding saturation factors interchanged. The rates are $r^{up,I}_{P,d} = 0.5$ and $r^{dn,I}_{P,d} = 2$, $r^{up,I}_{G,d} = 0.15$, $r^{up,I}_{G,nd} = 0.2$, and $r^{dn,I}_{G,d} = 12.5$, $r^{dn,I}_{G,nd} = 5$, and $\bar{G}^{max}_{d/nd}$ is the same for excitatory and inhibitory synapses. The main criterion for the choice of the learning parameters was to achieve the subthreshold membrane potential of directional cells (Fig. 8d). As a guideline we set our parameters so that LTP dominates the modification of P_{dis} (since the initial P_{dis} is small), while LTD dominates the modification of the synaptic strength. The specific values of the learning parameters are less crucial.

To test the necessity of the different elements in the rule (2) we have also simulated a minimal version of a temporally asymmetric learning rule. This reduced version is obtained from (2) by setting the thresholds θ^S_{post} and θ^C_{pre} to zero (thereby discarding the brackets in (2)), and by considering an instantaneous decay of the secondary messenger (thus replacing S^i_{post} with $\delta(t - t_{post})$ and S^i_{pre} with $\delta(t - t_{rel})$). Since in this reduced version the thresholds were missing, however, DS could not be maintained when the stimulus directions were randomized (see Fig. 6 and see Buchs and Senn, 1999).

3. Results

3.1. Emergence of Direction Selectivity

To illustrate the emergence of DS we started with a symmetric distribution of the LGN afferents as shown in Fig. 1. The synaptic efficacies are tuned such that for a sinewave stimulation the model simple cell shows the strongest response when the stimulus is centered over the RF. If the light grating moves from left to right, excitatory synapses from on-center LGN cells which are located in the left half of the RF are activated immediately before the occurrence of the postsynaptic spikes (raster plot in Fig. 3a, markers slightly left of the vertical lines) and therefore upregulate their synaptic strengths ($\bar{G}^E_{d/nd}$), while those in the right half of the RF are activated after the postsynaptic spikes (markers slightly right of the vertical lines), and therefore downregulate. The membrane potential is driven by the currents of four groups, on-and off-center LGN afferents, with either depressing or non-depressing synapses (Fig. 3a, lower traces). Repeated stimulations with gratings from the left induces a shift in the RF as shown on the left axis of Fig. 3a and in Fig. 4a.

The same rule also modifies the vesicle discharge probability of the depressing excitatory synapses (P^E_{dis}), which is a measure of the amount of depression. Since we have chosen small initial values for P_{dis} to mimic the subthreshold membrane potential of non-directional cells, P_{dis} increases during the learning process within the whole RF. Mathematically, this is due to the factor P^E_{dis} in front of the LTD term in Eq. (2), making this

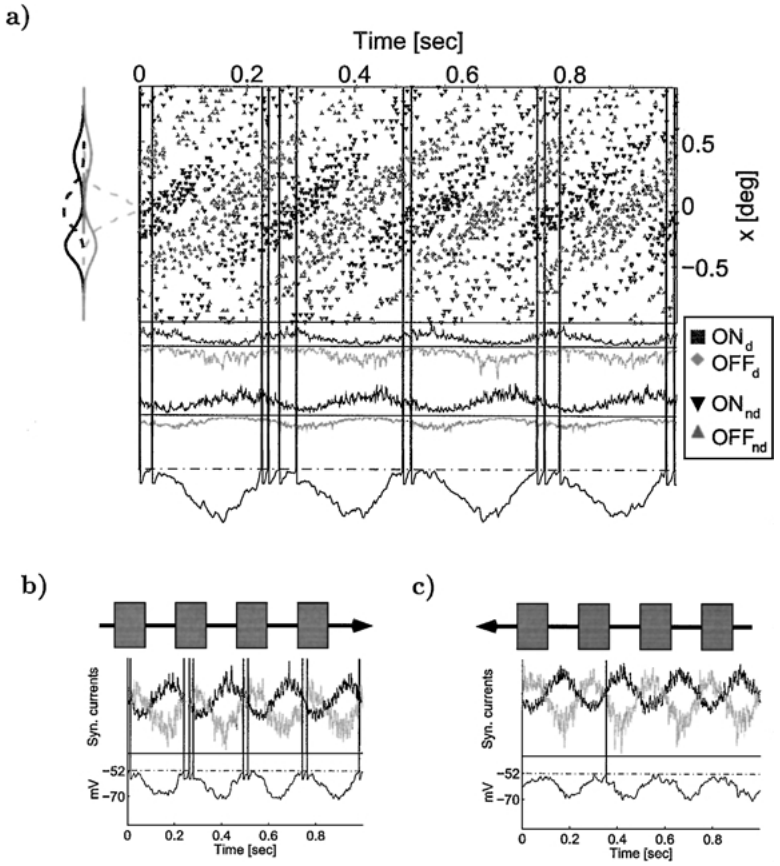

Figure 3. (**a**) Spike raster plot in response to a rightward moving grating (rightward running upwards). Each mark represents a release event of a synapse (■, ♦: depressing excitatory, inhibitory; ▼, ▲: non-depressing excitatory, inhibitory synapses). Releases of only each 30th synapse are shown. Left inset: effective synaptic strength of the different groups, same scaling and line code as in Fig. 1b. Lowest trace: postsynaptic membrane time course with attached action potentials (vertical lines running through the raster plot and indicating the postsynaptic spike times). Middle traces: current time courses generated by the synaptic subgroup corresponding to the legend on the right, with inhibitory currents plotted downwards. Excitatory synapses with synaptic releases immediately before a post-synaptic action potential (markers slightly left of the vertical lines) increase their strength and their probability of discharge, while excitatory synapses activated after a postsynaptic action potential (markers slightly left of the vertical lines) decrease these parameters. The rule for inhibitory synapses is reversed. (b, c) Responses after repeated stimulations with a rightwards drifting grating. (**b**) Upper traces: Total current from the depressing (grey) and non-depressing (dark) synapses induced by a grating ($\Omega/2\pi = 4$ Hz, $k/2\pi = 1$ deg^{-1}) moving in the preferred direction. Lower traces: Due to the phase advanced response of the depressing synapses the postsynaptic membrane potential is repeatedly driven above threshold (dashed dotted line) and generates action potentials (vertical lines). (**c**) Same as in (b) but for a grating drifting in the opposite direction. The response of the depressing synapses falls exactly into the trough of the non-depressing ones and the membrane potential is too low to reliably generate spikes. (For comparison with in vivo recordings see e.g. Fig. 1A in Carandini and Ferster, 2000.)

term vanish in the initial state. Due to the temporal structure of the stimulus, on the other hand, the dominating LTP term is stronger on the left half of the RF, inducing the spatial asymmetry of the P_{dis} distribution shown in Fig. 4b. Modifying the corresponding parameters of the inhibitory synapses, $\bar{G}^I_{d/nd}$ and P^I_{dis}, respectively, by the analogous rule (2), but with reversed signs, induces asymmetries in the same direction. Since inhibitory pathways are activated through off-center LGN cells, their activation is phase shifted by half a cycle and the synapses left from the RF center are activated *after* the postsynaptic activity (see raster plot Fig. 3). Hence, for inhibitory afferents the reversed rule achieves the same effects as the original rule for excitatory synapses (Fig. 4).

When testing the DS after training with a grating moving again in the training direction, the spatial shift in the RF together with the temporal phase advance

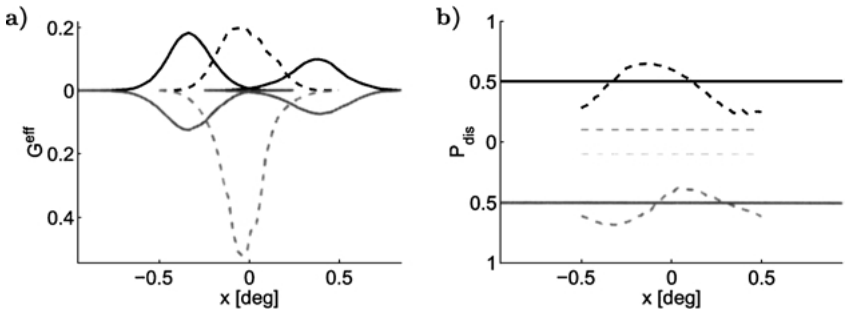

Figure 4. Synaptic distribution after learning. (**a**) Same as in Fig. 1b, but now after application of the drifting 'light' stimuli with drift velocity sampled from a Gaussian distribution (shown in Fig. 6b). The curves represent the effective synaptic strength $\bar{G}^{eff}$, averaged over 30 synapses, and binned together on the x-axis. (The strengths of the depressing synapses (dashed lines) are again scaled down by a factor of 0.5.) (**b**) The distribution of the vesicle discharge probability $P_{dis}^{E/I}$ for the excitatory and the inhibitory synapses (plotted up- and downwards, respectively) before (thin dashed lines at $P_{dis} = 0.03$) and after repeated stimulations (thick dashed lines). Straight lines at $P_{dis} = 0.5$ represent the unchanged discharge probabilities of the non-depressing excitatory and inhibitory synapses.

of the depressing synapses lead to a summation of the response of the two groups of afferents (Fig. 3b). If the grating moves in the opposite direction, however, the response of the depressing synapses falls into the trough originating from the non-depressing surround and the sum is smaller (Fig. 3c). Importantly, the synaptic rule adjusts the weights automatically in order to produce output firing rates in a physiological regime. In this regime the response to the non-preferred direction is reduced to single occasional spikes, while for the preferred direction the cell responds with firing rates between 10–20 Hz. This self-regulation is due to the dominance of downregulation in the rule for the synaptic strengths, implying that initially the simple cell response to both directions decays. As soon as the peak depolarization drops below threshold for the non-preferred direction, no further changes are induced by stimuli from that direction, and the overall decay eventually saturates. Once the rates in the learning equation (2) are tuned to get the appropriate dominance of LTD over LTP, the acquisition of the simple cell response property was very robust with respect to the modification of the neuronal parameters and the RF parameters.

The simple cell DS emerges from the interplay between the spatial RF asymmetry induced by modifying the synaptic strength, and the temporal phase advance of the depressing synapses induced by modifying the vesicle discharge probability. For a stimulus moving in the preferred direction, the spatial and temporal advances add, while they cancel each other for a stimulus moving in the non-preferred direction. The temporal and spatial shift as a function of the temporal frequency ($\Omega/2\pi$) of the stimulus is shown in Fig. 5, before and after repeated stimulations in different directions, but with fixed temporal frequency. Since P_{dis} is initially small, virtually no synaptic (short-term) depression and therefore no temporal phase advance is present (Fig. 5a, dashed line). This is different, however, when P_{dis} becomes large during the stimulation (Fig. 4b), thereby causing synaptic depression and a significant phase advance (Fig. 5a, full line). Similarly, the modification of $\bar{G}_d^E$ induces a shift in the effective synaptic strength of the depressing synapses opposite to the stimulus direction (Fig. 5b, full line). Interestingly, for stimuli extending over the entire width of the RF the asymptotic shift is independent of the grating velocity, but instead depends on the spatial derivative of the stimulus (this would be different for a moving narrow light bar, cf. Eq. (40) in the analysis paper). However, if we stimulate only with a finite number of cycles with the same temporal frequency, or with an infinite number of cycles with differently oriented gratings, the asymptotic shift cannot be reached, and spatial phase shifts evolve as shown in Fig. 5b. Note that the temporal and spatial shifts show a similar frequency dependency, implying that they effectively cancel in the non-preferred directions while they add in the preferred direction. For an explicit estimate of the two phase shifts as a function of the temporal frequency, and for further discussions of the parameter dependencies, we refer to the analysis paper.

To test the dependence of our results on the number of afferents we run all simulations with 600 instead of 4800 LGN cells. Although no qualitative differences were observed, the reduction increased the postsynaptic fluctuations and slightly impaired the proper

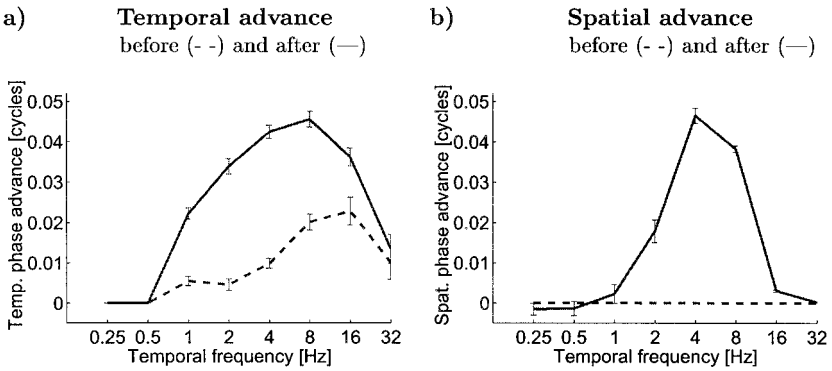

Figure 5. Temporal and spatial phase advance as a function of the temporal frequency $\Omega/2\pi$ of the drifting gratings. (**a**) The temporal phase advance (as a fraction of a full cycle) of the synaptic current consisting of the sum of all depressing (excitatory and inhibitory) synapses (cf. Fig. 3), before (dashed) and after learning (solid). The shift is represented by the phase of the first Fourier component of the synaptic response. The temporal phase advance mainly depends on the temporal frequency of the test grating, while it is only weakly dependent on the specific temporal frequency during training. (**b**) The spatial phase shift of the effective synaptic strength, $\bar{G}_d^{eff,E}(x)$, of the depressing excitatory synapses, before (dashed) and after stimulation with 40 cycles of a drifting grating with the corresponding temporal frequency. The optimal temporal frequency of roughly 4 Hz is determined by the parameters of the learning process and, in a first approximation, is given by $1/(2\pi\tau_L)$, where τ_L is the (effective) width of the learning function (see analysis paper). Note that the temporal phase advance (a), which is determined by the presynaptic firing rate and the parameters of the short-term depression, does also peak at roughly 2–8 Hz. Error bars represent standard deviations.

development of DS. When stimulating with a 4 Hz grating moving always from left to right, for instance, the average direction index was DI = 0.83 in the case of 600 LGN cells, compared to DI = 0.97 in the case of 4800 LGN cells (where the direction index is defined by $\mathrm{DI} = (\mathrm{P} - \mathrm{NP})/(\mathrm{P} + \mathrm{NP})$ with P and NP being the simple cell responses to the preferred and non-preferred directions).

3.2. *Symmetry Breaking for an Unbiased Set of Noisy Stimuli*

As we saw, the asymmetry in the learning rule induces RF shifts opposite to the stimulus direction, and the simple cell 'learns' to respond to motions from the direction it was 'trained' to. A more realistic scenario, however, would involve different stimulus velocities with different directions of motion. Without assuming a specific structure of the environment one would expect a normal distribution of velocities around zero. Our simulations show that even for such an unbiased set of stimuli DS emerges (Fig. 6a and b). Due to the initial RF symmetry the simple cell responds equally strongly to gratings moving with the same speed in opposite directions. However, a selectivity to either direction emerges after repeated presentation of drifting gratings with random velocities and directions sampled from a zero-mean Gaussian distribution (Fig. 6a), each presented for 4 cycles. This arises through a positive feedback mechanism which amplifies slight RF asymmetries: as soon as the response to one direction dominates, the RF shifts opposite to that direction, and this causes a further increase of the response to the same direction (see Section 3.4 of the analysis paper).

To test the dependence of this stability result on the details of the learning rule we implemented different forms of the rule with varying parameter values. It turned out that a crucial element ensuring the robustness against reversing stimulus directions are the thresholds imposed in the learning rule (2). Such learning thresholds (or, equivalently, strong nonlinearities in the learning curve) make a qualitative difference in unlearning. Without thresholds, any acquired selectivity can be unlearned if by chance several gratings moved in sequence in the non-preferred direction. In fact, for the reduced learning rule without thresholds and without secondary messengers, the selectivity only transiently emerged when the stimulus direction was randomly chosen (Fig. 6c). Unlearning of the acquired DS without additional nonlinearities is always possible since the simple cell still responds to the non-preferred direction and thereby counteracts the preferred synaptic modifications. With a learning threshold, however, the response to the non-preferred direction may fall significantly below this threshold if we assume that LTD dominates LTP. In this case the remaining responses

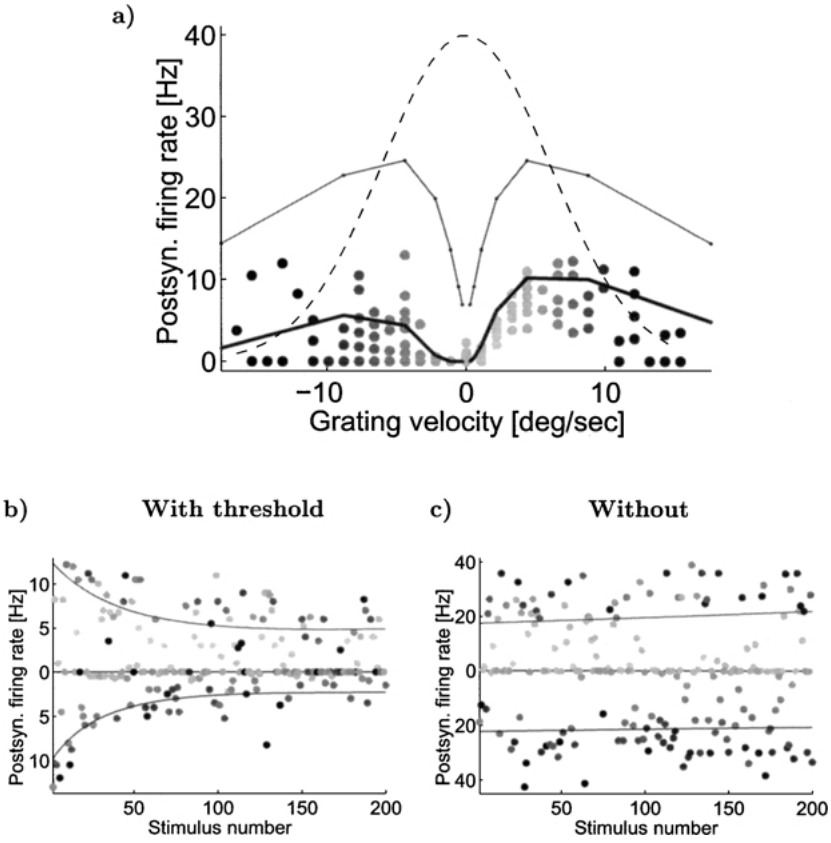

Figure 6. Unbiased input scenario requires a learning threshold. (**a**) Distribution of the postsynaptic response to gratings with random velocities, using the learning rule (2) with internal thresholds. According to the initial RF symmetry the simple cell response (dots, representing spike rates averaged over 4 cycles) does not depend on the stimulus direction (thin lines). After presenting moving gratings with velocities $\frac{\Omega}{k}$ (deg/sec) sampled from a Gaussian distribution (dashed line, mean = 0, SD = 6 deg/sec) the symmetry in the response is broken and the cell develops some DS for positive velocities (thick line, average of the individual samples). (**b**) Temporal evolution of the postsynaptic spike frequency for the same data shown in (a). The simple cell responses to rightward/leftward drifting gratings are plotted on the upward/downward ordinate. Due to the symmetry breaking of the RF, the responses to the two directions drop with different speeds. Once a directional preference is established, the threshold in the learning rule prevents unlearning. (**c**) If we use the reduced learning rule without learning thresholds and infinitely fast secondary messenger (see Section 2.4), the acquired DS cannot be maintained. Although the symmetric RF is unstable and therefore transiently some DS evolves, a few gratings moving in series in the non-preferred direction can again abolish this selectivity. The brightness of the dots encodes the grating velocity according to (a).

to the non-preferred directions are too weak to counteract the acquired selectivity, even during a very long stimulus sequences in the non-preferred direction.

Randomized grating velocities and drifting directions is one source of stachasticity. As another source of stochasticity we considered Gaussian noise (with zero mean and a 3 Hz standard deviation) in the instantaneous Poisson firing rates of the LGN cells. As expected, the noise in the firing rate is absorbed by the Poisson process and affects neither the development of DS nor the shape of the frequency response curves (cf. Fig. 7a). The noise robustness crucially depends on the fact that the individual spike trains themselves are stochastic. To reveal this functional role of the stochasticity we implemented a deterministic generation of the presynaptic spikes and the synaptic releases (by triggering a presynaptic spike whenever the time from the last spike exceeded the instantaneous presynaptic firing rate, and with a deterministic implementation of our depressing and non-depressing synapses, cf. Tsodyks and Markram (1997)). In the deterministic version we often observed high frequency oscillations of the total postsynaptic current in its rising phase (up to 100 Hz) due to a partial alignment of the presynaptic spikes

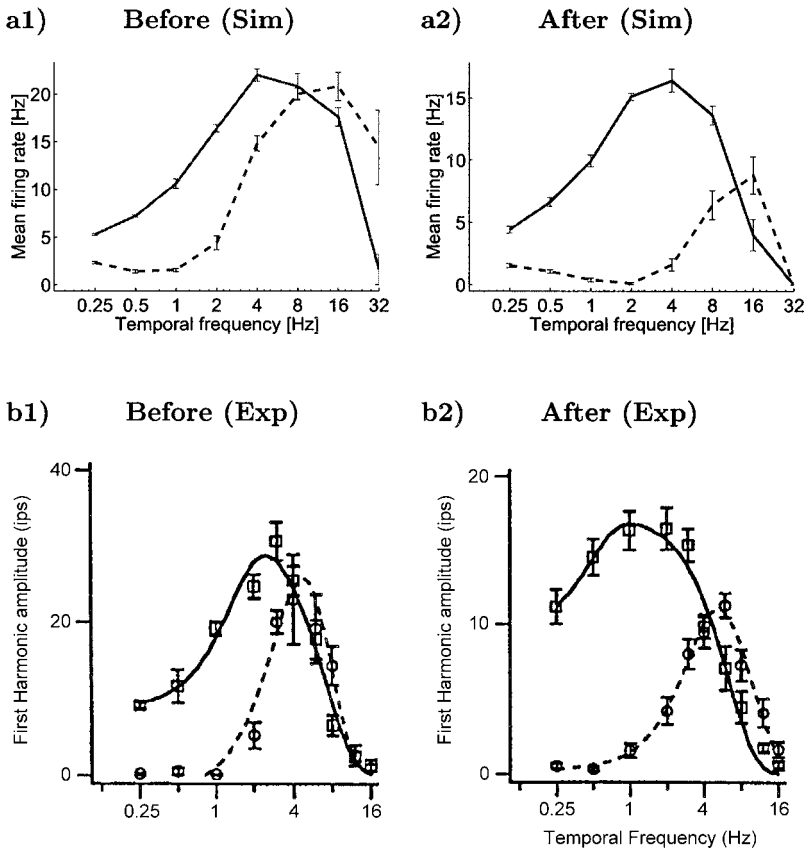

Figure 7. Activity-dependent shift of the simple cell tuning curves, observed in the simulations (**a**) and in vivo (**b**). (**a**) In an initial state, the RF is symmetric and the tuning curves for the two directions are identical (not shown). After presenting 50 grating stimuli with random drift velocities, the model simple cell develops weak DS (**a1**). Solid line: postsynaptic mean firing rate for the preferred direction as a function of the temporal frequency of the grating. Dashed line: same for the non-preferred direction. (**a2**) After additional 200 random stimulations the postsynaptic firing rate dropped (note the different scales on the ordinates) and the DS improved in a regime where the temporal and spatial phase shift is largest (cf. Fig. 5). Error bars represent standard errors of the mean. (**b1**) Tuning curve of a directional simple cell in layer 5 of area 17 (replotted from Saul and Humphrey, 1992b), showing the response amplitude in the preferred (solid) and non-preferred direction (dashed). (**b2**) Recordings from the same cell after 3 additional hours of stimulations show a drop in the overall response and in the response to the non-preferred direction similar to the simulations.

from the different afferents. This occurred even though the initial spike times were randomized and despite the fact that we used a purely feedforward architecture. When adding the Gaussian noise to the instantaneous firing rates of the LGN cells, however, these high frequency oscillations disappeared and the postsynaptic peak response drastically dropped. While the high frequency oscillation destroyed the simple cell DS, it was again restored by adding independent spike-time jitter. Hence, the same noise which is absorbed in the stochastic version may have a drastic effect on the simple cell response when the presynaptic spikes and the synaptic releases are deterministic.

3.3. From Non-Directional to Directional Simple Cells

Next we investigated the possibility that the learning rule may convert non-directional simple cells into directional cells. To test this possibility we started from an initially symmetric synaptic arrangement and determined the velocity tuning curves before and after stimulation with randomly sampled gratings. Depending on the parameter values (such as k, f^1, $r^{up/dn}$, θ_{post}), the original and emerging tuning curves are in fact qualitatively similar to those measured for non-directional and directional simple cells in the cat visual cortex (Saul

and Humphrey, 1992b). In our model it is important that the overall simple cell response decreases during the ongoing stimulations. Only when LTD dominates LTP, will the self-stabilization of the simple cell response occur (note the drop of the maximal firing rate of 20 Hz to 15 Hz in Fig. 7a).

The requirement that the simple cell responsiveness must decrease during the conversion of a directional cell to a non-directional cell is difficult to support experimentally. Nevertheless, there is one example of a simple cell in V1 which shows a similar decrease of the responsiveness and a similar shift of the temporal frequency tuning curve between the beginning and the end of its 3 hour long recording as it is typical in the model (Fig. 7a and b). In both model and data there is an increase in DS at temporal frequencies between 2–8 Hz. In the model this is because the spatial shift of the RF center caused by STDP and the temporal phase advance caused by synaptic depression are both maximal at temporal frequencies between 2–8 Hz (see Fig. 5). The reason for the activity-dependent shift of the tuning curves in the data is not established. We did not try to achieve a full quantitative fit of the data shown by this specific cell since the spectrum of measured tuning curves is rather large (see, for instance, Saul and Humphrey (1992b) and Hawken et al. (1996). Imposing a postsynaptic membrane time constant of $\tau_m = 50$ ms for the simple cell, for instance, would easily cut off the simple cell responses above 16 Hz (see e.g. Carandini et al., 2001). However, for synaptic depression time constants τ_{rec} below 200 ms it is not possible to achieve DS at low frequencies (0.25 Hz) which is as strong as that observed in vivo. Similar tuning curves as shown in Fig. 7a can be obtained by considering only excitatory synapses.

To further investigate the conversion of a non-directional to a directional simple cell we considered the modification of the time course of the subthreshold membrane potential caused by the synaptic plasticity. The response to counterphase gratings applied before and after the developmental process was very similar to the response of non-directional and directional simple cells in V1 (Fig. 8). The counterphase response before the developmental process showed virtually no distortion of the sinusoidal membrane potential oscillation (Fig. 8 a1). This is because in the initial configuration we assumed small vesicle discharge probabilities which makes synaptic depression virtually disappearing. Due to the linear behavior of the simple cell the counterphase responses predict well the response to the drifting gratings (a2), just as it is the case for non-directional simple cells in V1 (c1, c2). After random stimulation with drifting gratings, different temporal distortions are caused at different spatial phases (b1). The distortions are caused by the development of synaptic depression during the learning process and reproduce the advanced responses in the intracellular recordings from a direction selective simple cell in V1 (d1). If compared with the simulations in Chance et al. (1998) the distortion is less sharp, since in our case the synaptic afferents are continuously distributed in space (rather than concentrated in 3 separate positions). Interestingly, the temporal distortion of the total postsynaptic current is completely lost when considering drifting gratings (b2, d2). In our simulations this is due to the spatial averaging of differently timed synaptic responses: since the same temporal nonlinearities arise at different spatial positions, the total postsynaptic current obtained by sequentially activating these afferents is smoothed out. This is the remarkable linearity of simple cells emphasized by many works (Reid et al., 1987, 1991; Jagadeesh et al., 1993; for a review see Shapley, 1994). Despite this linearity, DS could be explained by non-linear synaptic depression emerging from an activity-dependent redistribution of the synaptic efficacy (through STDP acting on P_{dis}).

3.4. Strobe Rearing Prevents the Development of Direction Selectivity

Finally we show that the development of DS by means of an asymmetric synaptic learning rule is compatible with strobe rearing experiments. In these experiments cats were reared from birth to 8 months in a room illuminated only by a strobe lamp operating at 8 Hz (Pasternak et al., 1985; Cremieux et al., 1987; Humphrey and Saul, 1998). As observed by the different groups, DS in most simple cells could not properly evolve or was largely lost under such flickering illumination, although other features like orientation preferences remained intact. Further analysis suggests that the deficit in directional cells is because strobe rearing prevents the convergence of synaptic inputs with different response timing onto the same simple cell (Humphrey et al., 1998). Those remaining cells which still had differently timed afferents were less direction selective, and this can be explained with a reduced inseparability of the spatio-temporal RF (Humphrey and Saul, 1998).

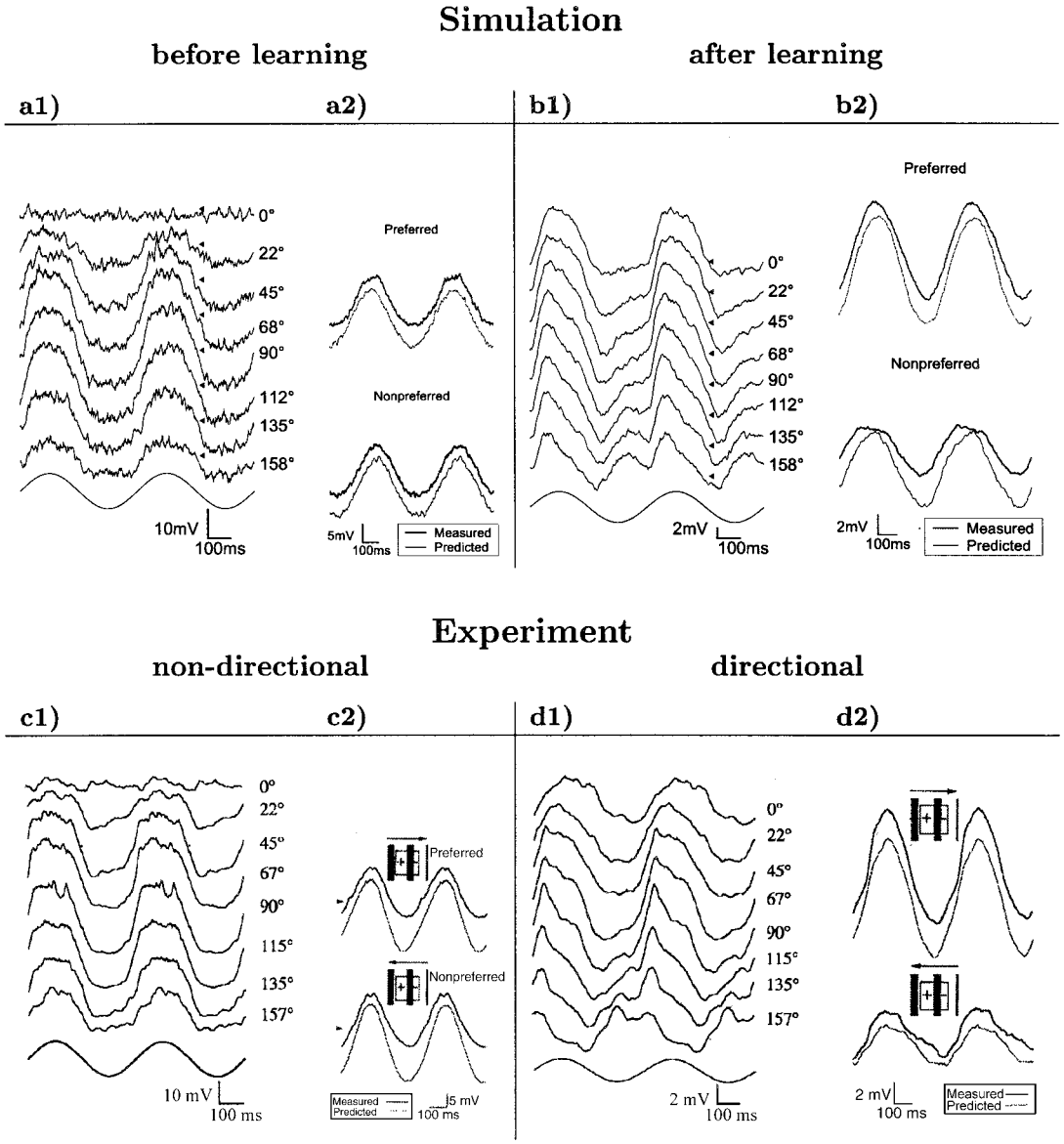

Figure 8. Comparison of subthreshold responses before and after learning with in vivo recordings in a non-directional cell in area 17 of the cat. (**a1**) Membrane potentials evoked by a contrast modulated stationary sinewave grating oscillating at 2 Hz, located at eight different spatial phases. Arrows indicate the level of the membrane resting potential, and the smooth trace at the bottom reflects the stimulus time course in the RF center. (**a2**) Top: Response to a drifting grating moving in the preferred direction with the same temporal frequency as the stationary grating (thick trace). The prediction from the appropriately shifted and scaled responses from the left is shown as a thin line below. Bottom: Same for the non-preferred direction. (**b**) Same as in a, but now after 300 presentation of drifting gratings with velocities sampled from a normal distribution around 0 deg/s. Due to the temporal nonlinearities the response predicted by the counterphase gratings does not perfectly coincide with the response to the drifting grating. (**c, d**) The original experiments for a non-directional (c) and directional (d) simple cell in the cat visual cortex (replotted from Jagadeesh, 1993, 1997). The nonlinearities in (b1) and (d1) of the membrane potential is attributed to the depressing synapses in the RF center.

To mimic strobe rearing we drove the LGN cells with drifting gratings illuminated only during 8 Hz flashes. Each flash generates an instantaneous LGN activity which is proportional to the instantaneous illuminance of the grating at the specific site and which decays with a time constant of 50 ms (the effective integration time of simple cells ranges from 44–100 ms, see Reid et al. (1992)). When applying such stimuli to the initial

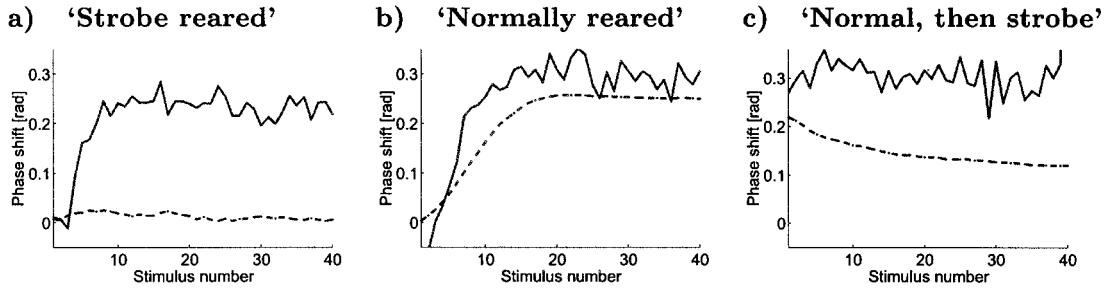

Figure 9. Strobe versus normal rearing. (**a**) During 'strobe rearing' (mimicked by illuminating drifting gratings during the short 8 Hz flashes) only the temporal phase advance develops (by uniformly increasing P_{dis}, solid line), without breaking the spatial symmetry of the RF (since, on the time scale of the learning window, the information about the direction of motion is lost, dashed line). This leads to a simple cell which is not direction selective. No symmetry breaking occurs even if the stimulating gratings move exclusively in one direction. (**b**) During 'normal rearing' the temporal phase advance of the depressing synapses (solid) and their spatial shift (dashed) both increase. (**c**) When applying the stroboscopic illumination scenario after 'normal rearing' (b), the acquired spatial phase shift of the depressing synapses fades out and DS is impaired.

symmetric arrangement with weak depression, the discharge probability increased uniformly, and no appropriate spatial synaptic structure in the synaptic strength emerged. This is because the time window of the learning rule is too narrow to capture the motion sampled every 125 ms by the flash. Moreover, the short flashes distort the presynaptic activity along a single LGN afferent in a step-like manner and, on the time scale of the learning window, the directional information is lost. Since the synaptic modification is proportional to the derivative of the presynaptic firing rate multiplied by the postsynaptic rate (Eq. (58) of the analysis paper) no directional information is contained in the final synaptic strengths.

Although no symmetry breaking occurs in our simulations for strobe rearing, the discharge probability may globally increase depending on whether the flash activity is strong enough to trigger the synaptic modifications. As a consequence, a temporal phase advance of the RF center may develop (Fig. 9a). Due to the lack of a spatial phase shift, however, strobe reared cells are not direction selective. For comparison, under normal rearing the temporal and spatial phase advance both co-evolve (Fig. 9b), resulting in a directional simple cell. When applying the strobe illumination to a normally reared, direction selective simple cell, the temporal phase advance is maintained since P_{dis} becomes uniformly large (Fig. 9c). However, since LTD dominates the modification of the synaptic strength, and since the strobe flashes break the correlations between the pre- and postsynaptic activity present in uniformly illuminated drifting gratings, the RF asymmetry acquired during normal rearing is again abolished (Fig. 9c). Hence, in agreement with the experiments, normally reared simple cells loose their DS when repeatedly exposed to strobe illumination.

Our model also allows for an interpretation of the behavioral experiments showing that strobe reared cats may become able to detect motion directions if the light intensity of the stimulus is increased by a factor of 10 from a normal level (Pasternak et al., 1985). To explain this phenomenon within our model we assume that some cells show an initial asymmetry in the spatial distribution of depressing and non-depressing afferents. If the flash induced activity is too weak to trigger synaptic modifications P_{dis} would remain small (0.03–0.1). In this case the usual light intensities do not generate a suitable phase advance and the simple cell remains unselective. However, with low P_{dis} synaptic depression sets in if the light intensity is increased strongly enough (because depression depends on the product $f_{pre}P_{dis}$), and thereby generates the required phase advance making the cell direction selective (simulations not shown, but see Fig. 2E in Chance et al. (1998), for the monotonic increase of the phase advance with stimulus contrast).

4. Discussion

We have presented a developmental model for the simple cell DS. The mechanism for DS we considered exploits the temporally nonlinear response characteristics arising from synaptic depression, as previously suggested by Chance et al. Synaptic depression of feedforward connections from the LGN to the primary visual cortex was recently observed in the cat (Stratford et al., 1996), and it is an open issue to what extent

DS is attributed to synaptic depression (Chance et al., 1998), intracortical connections (Heeger, 1993; Suarez et al., 1995), and different LGN latencies (Feidler et al., 1997). The present work shows that, in principle, synaptic depression together with STDP (Markram and Tsodyks, 1996; Markram et al., 1997) would have the power to convert non-directional simple cells into directional ones within a randomized stimulation scenario. Synaptic depression in the visual pathway is also supported by recent in vivo experiments on cross-orientation suppression in the primary visual cortex of the cat (Durand et al., 2001; Freeman et al., 2001). There is a wide range of behaviors seen in simple cells which can be explained by thalamocortical synaptic depression, as exposed in Carandini et al. (2002). Temporally asymmetric STDP in V1, on the other hand, is supported by recent experiments showing that the short-and long-term dynamics of RFs depend on the milliseconds delay between two spatially displaced flashing stimuli (Djupsund et al., 2001; Gao et al., 2001; Yao and Dan, 2001).

The cited experiments on synaptic depression and STDP in the primary visual cortex could be combined to investigate the effect of these mechanisms on the simple cell DS. If both cross-orientation suppression and DS rely on synaptic depression, masking with orthogonal gratings should impair the DS of a cell. If STDP is present in V1, repeated stimulation with unidirectional gratings should alter the cell's DS. Moreover, DS should be abolished when replacing the sine grating by a bar grating because the characteristic phase advance would not develop for a moving bar grating. This would be different, however, if DS is caused by different LGN latencies instead of synaptic depression.

To isolate the impact of synaptic depression onto DS we considered a purely feedforward architecture (Hubel and Wiesel, 1962) and neglect the different types of LGN latencies (Saul and Humphrey, 1992a). When adding LGN latencies distributed in the range of 0–150 ms to our model, DS still evolved. If after learning we turned off either the LGN latencies or the synaptic depression, DS was lost in both cases, showing that both mechanisms may jointly contribute, and with equal importance, to the simple cell DS. We emphasize that we do not make a statement about the origin of the neuronal activity driving the synaptic modifications. In fact, different species show different stages of simple cell DS at eye-opening (for a review and further references see Sur and Leamey (2001), for early works on kittens see Hubel and Wiesel (1963), and Albus and Wolf (1984)), and spontaneous waves of retinal activity before eye-opening (Meister et al., 1991, 1995) may generate LGN activity which may drive synaptic modifications in a similar manner as the visual input after eye-opening.

4.1. Basic Results

The important property of depressing synapses is that they cause a phase advanced response to a sinusoidally modulated stimulus. DS emerges if, in the stimulus direction, afferents with depressing synapses are spatially arranged behind non-depressing afferents. A stimulus activating in sequence non-depressing and depressing synapses therefore produces in-phase postsynaptic responses which add together. If the stimulus activates the synaptic populations in the reverse order however, the responses are slightly phase shifted, sum up less and trigger only weak simple cell responses (Fig. 3b and c). Although the phase advance of the total synaptic current summed over the spatially distributed population of depressing synapses is relatively small due the spatial averaging (Fig. 5a), this phase advance is still enough to produce DS with realistic direction indices (up to $\sim$0.9 for the optimal stimulus frequency, see Fig. 7, a2). However, the size of the possible phase advances is too small to explain the different response latencies observed in cortical cells (Saul and Feidler, 2002).

An activity-dependent development of the different phase shifts requires a rule which generates the correct spatial distribution of depressing and non-depressing afferents within the RF, and at the same time appropriately adjusts the degree of synaptic depression. These requirements are met by a temporally asymmetric STDP which acts on both, the synaptic strength and the degree of synaptic depression. This form of STDP imparts several properties to simple cells which may help them to develop DS in an unbiased input scenario with visual stimuli moving with various velocities and different directions. Such properties comprise the temporal tuning curve (Fig. 7) and the subthreshold membrance potential time course (Fig. 8). Depending on the rate of the synaptic modifications, the properties are either more akin to those of a non-directional or a directional simple cell. The model also offers an explanation of the impaired DS after exposure to stroboscopic illumination (Pasternak et al., 1985; Cremieux et al., 1987; Humphrey and Saul, 1998). Since the flash

period of a 8 Hz strobe light is beyond the time window of synaptic modification, a spatial RF asymmetry will either be lost or cannot develop (Fig. 9). If, in the model, DS is lost after exposure to strobe stimuli, it can be regained by increasing the contrast and light intensity of the moving stimuli.

4.2. The Learning Process

In order to stably reproduce the different experimental observations in the presence of the unbiased randomized stimulus scenario we had to introduce additional elements to our basic model. A careful analysis of these elements leads to different insights into the learning process, and raises further hypotheses about its biological implementation:

(1) *Symmetry breaking through positive feedback*. When exposed to an unbiased input scenario with gratings moving with equal probability in different directions, any initial symmetry in the simple cell RF will be broken. This is ascribed to the positive feedback loop induced by the Hebbian rule. Since the rule strengthens the synapses proportional to their contribution to the postsynaptic activity, a subsequent activation will lead to an increased response which in turn strengthens the same synapses (see the analysis paper for a formal description).

(2) *LTD > LTP implies self-regulation of simple cell responses.* Without additional regulatory mechanisms the above positive feedback loop would drive the synaptic strengths towards saturation level. In a more realistic input scenario with stimuli moving in different directions, as considered here, the synaptic growth is naturally bounded since a synapse is activated equally often before and after a postsynaptic spike. As recently shown in different papers (see e.g. Kempter et al., 1999; Song et al., 2000) this leads to a normalization of the synaptic strengths, provided the down-regulating branch in the rule dominates (in area) the up-regulating branch ('LTD > LTP'). In our simulations this dominance implies that the overall response to both directions decreases until the learning process ceases (see Fig. 6).

(3) *Preventing unlearning requires an internal synaptic threshold.* The above normalization property does not yet assure the stability of the learning process, since repeated presentation of stimuli moving in the non-preferred direction may unlearn the acquired DS (see Fig. 6b and Buchs and Senn, 1999). There is, however, a qualitative difference when applying a learning threshold to the postsynaptic activity in the two Hebbian terms of Eq. (2). Although the cell may still respond to the non-preferred direction, no synaptic changes are induced if this activity becomes too weak to drive the internal (post)synaptic variables above the learning threshold. Additional mechanisms may lead to a slow adaptation of this learning threshold such that the synapse may regain its plasticity after a period of low postsynaptic activity, e.g. induced by visual deprivation. Such a moving threshold was introduced in the BCM-theory (Bienenstock et al., 1982) which was also shown to be appropriate for learning DS for a natural class of moving stimuli (Blais et al., 2000; see also Feidler et al., 1997). We emphasize that the same learning threshold is also required to model the plasticity data of Markram et al. (1997) (see Senn et al., 2001). Interestingly, the third-order nonlinearity coming out from this data also implies a moving threshold property analogous to the BCM-theory (Senn et al., 2001). The necessity of a learning threshold to stabilize synaptic structures is well known in the context of working memory formation by attractor networks (Fusi et al., 2000).

(4) *Modification of both, synaptic strength and synaptic depression, produce stronger and more reliable DS.* In a previous study we showed that the modification of the degree of synaptic depression (P_{dis}) alone could explain the emergence of DS (Buchs and Senn, 2001). However, the emerged direction selectivity was sensitive to specific spatial distribution of P_{dis}, and sometimes DS failed to develop in the unbiased input scenario when by chance the stimulus directions were switched very frequently. When we co-modify both the discharge probability and the synaptic strength according to the same rule, the RF asymmetry is much stronger, and the development and performance of DS is more reliable. Both types of synaptic long-term modifications are indeed observed in neocortex; a change of the synaptic strength (Feldman, 2000), and a temporal redistribution of the synaptic efficacy (Markram and Tsodyks, 1996) as a function of the pre- and postsynaptic spike-time differences.

(5) *Gaussian synaptic densities make the RF development robust against parameter variations.* As an additional stabilization element we modeled fixed Gaussian spatial synaptic densities which significantly reduced the sensitivity of the results to variations in the model parameters. Formally, the effective synaptic strength, G^{eff}, is determined by the product of the

synaptic densities times the strength of the individual synapses. Whereas in our simulations the synaptic strength and the degree of depression were subject to synaptic modification, the initial Gaussian densities of the different types of LGN afferents remained fixed. Due to these densities, the quality of the acquired DS (in terms of realistic tuning curves) and the shape of the effective synaptic strengths after learning (Fig. 4a) were virtually indistinguishable for different runs and different sets of learning parameters and stimulus distributions. This was in strong contrast to the case when the spatial synaptic densities were rectangular. It is tempting to speculate that nature likewise assures the stability of the development by first forming broad Gaussian synaptic densities (via formation of new connections, see Lendvai et al. (1999) and Geinisman (2001)), making the system stable against input variations, on top of which activity-dependent synaptic modifications acts as a fine tuning.

(6) *STDP for inhibitory synapses?* In order to explain the appropriate development of the push-pull mechanism exerted by excitatory and inhibitory afferents (see the subthreshold membrance potential before and after learning, Fig. 8a and b), it is required that the inhibitory feedforward pathway is modified opposite to the excitatory pathway. Since, for reasons of simplicity, we simulated only monosynaptic inhibitory pathways from the LGN to the model simple cell, we had to postulate the reversed learning rule for these inhibitory synapses. In reality, however, inhibitory afferents are at least disynaptic with excitatory projections from the LGN. In such a disynaptic inhibitory pathway, the same opposite modifications are also obtained by applying the original rule to the excitatory thalamocortical synapses, while leaving the subsequent inhibitory synapses nonplastic. It would be interesting to test whether in fact thalamocortical excitatory synapses are subject to the standard form of STDP, and/or whether intracortical inhibitory synapses obey the reversed rule.

(7) *Different learning rules for depressing/non-depressing synapses may generate additional RF structure.* To assure the emergence of strong DS we had to assuming an initial symmetric synaptic structure, with depressing thalamocortical synapses surrounded by non-depressing ones (Fig. 1). This initial structure proved to be necessary since we applied the same learning rule to the strength of depressing and non-depressing synapses. If we would assume the same initial distribution for depressing and non-depressing synapses, both would experience a similar spatial phase shift and no DS would develop. As an alternative to a hypothetical initial RF structure one may postulate that the temporally asymmetric rule only modifies the strength of the non-depressing synapses, while the strength of the depressing synapses (and less critically P_{dis}) would be modified according to a symmetric Hebbian rule. In this case we expect that an appropriate RF structure would fully develop from an initially unstructured RF. This is because a RF shift would only develop for the non-depressing synapses while the depressing synapses would remain in the RF center, generating a RF structure similar to that shown in Fig. 4a. These simulations, however, remain to be done. There is no experimental data which would favor one or the other alternative. Although differences in the temporal response characteristics of the simple cell RF center and its surround are reported (see Orban et al. (1987), or Soodak et al. (1991) for LGN RF substructures) no structure of depressing and non-depressing synapses is proven. Similarly, it remains an open question whether different parameters of dynamic synapses are modified according to different rules (Markram et al., 1998).

(8) *The width of the learning window limits the detectable stimulus velocities.* When 'training' the simple cell with bars moving in a single direction across a wide RF, the induced spatial shift of the synaptic strengths is roughly proportional to the stimulus velocities. A subsequent 'test' of the simple cell velocity tuning will give the strongest response for the same direction and velocity it was trained for (see Rao and Sejnowski (2000), for a similar scenario). This intuitive result, however, is only correct if the bar is narrow compared to the RF width (and the velocity is not too small compared to the RF width divided by the width of the learning window, see Eq. (40) in the analysis paper). If the moving stimulus extends over the full RF, as is the case for gratings, the RF shift is no longer a monotonic function of the velocity, but rather peaks at a specific value. At fixed spatial frequency, the maximal spatial phase shift is obtained for stimuli with temporal frequency $1/(2\pi\tau_L) \approx 4$ Hz, where $\tau_L = 40$ ms is the temporal width of the learning function (Fig. 5b, cf. also Fig. 4a and Eq. (37) in the analysis paper). Stimuli with the same spatial frequency but moving faster or slower generate a smaller spatial shift of the RF. Once the RF modification is accomplished, the same velocity which triggered a maximal RF shift is also the one which gives the strongest DS (compare Fig. 5b and 7 a2). This velocity is determined by the temporal width of the learning window, and not by the distribution of

the training velocities, as is the case for moving bars. Interestingly, the temporal frequencies producing the largest RF shifts are also the ones producing significant temporal phase advances by means of short-term synaptic depression (compare Fig. 5a and b). It appears that the parameters of the STDP and of the short-term synaptic depression are well adjusted to support a development of the simple cell DS in a realistic frequency regime.

(9) *Stochastic spike times prevent high frequency oscillations and make DS development robust against noise in the LGN.* In the Introduction we exposed the conundrum that the precision of the STDP exceeds the jitter of the individual spike times. The paradox is resolved by considering a large population of independent LGN afferents averaging out the effects of the spike jitter. In addition, the stochastic spike generation has the functional consequence of protecting the simple cell against high frequency oscillations. This is because with a deterministic spike generation there is a high chance that fast oscillations arise through transient alignments of spikes from different afferents. When imposing deterministic inter-spike intervals in our simulations, these oscillations can become so strong that they prevent the development of DS.

In summary, temporally asymmetric STDP attributed to depressing and non-depressing thalamocortical synapses may convert non-directional simple cells into directional simple cells, thereby generating qualitatively similar tuning curves and membrane potential responses as observed in vivo. The important features of the synaptic modifications ensuring a robust DS development in an unbiased stochastic input scenario are the learning thresholds imposed to the postsynaptic activity, the dominance of LTD over LTP, the co-modification of the synaptic depression and the synaptic strength (while keeping the synaptic densities fixed), and the stochastic spike generation of the LGN afferents.

Acknowledgments

This study was supported by the Swiss National Science Foundation (grant 21-57076.99) and the Silva Casa foundation. We would like to thank Alan Saul for helpful suggestions and stimulating discussions, Stefano Fusi and Jan Reutimann for critical comments and technical support, and Dave Thurbon for proof reading.

References

Adelson E, Bergen J (1985) Spatio-temporal energy models for the perception of motion. J. Opt. Soc. Am. A 2: 284–299.

Ahmed B, Anderson J, Douglas R, Martin K, Nelson J (1994) Polyneuronal innervation of spiny stellate neurons in cat visual cortex. J. Comp. Neurol. 341(1): 39–49.

Albus K, Wolf W (1984) Early post-natal development of neuronal function in the kitten's visual cortex: A laminar analysis. J. Physiol. 348: 153–185.

Bi G, Poo M (1998) Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. J. Neuroscience 18(24): 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. Neuroscience 2(1): 32–48.

Blais B, Cooper L, Shouval H (2000) Formation of direction selectivity in natural scene environments. Neural Computation 12(3): 1057–1066.

Borst A, Egelhaaf M (1989) Principles of visual motion detection. Trends in Neurosci. 12(8): 297–306.

Buchs N, Senn W (1999) Learning direction selectivity through adaptation of the probability of the vesicle release probability. In: Society for Neuroscience, Abstracts, No. 898.15.

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

Carandini M, Ferster D (2000) Membrane potential and firing rate in cat primary visual cortex. Journal of Neuroscience 20(1): 470–484.

Carandini M, Heeger D, Senn W (2001) Thalamocortical synaptic depression and the responsiveness of the primary visual cortex. Neuron, submitted.

Chance F, Nelson S, Abbott L (1998) Synaptic depression and the temporal response characteristics of V1 cells. J. Neuroscience 18(12): 4785–4799.

Cremieux J, Orban G, Duysens J, Amblard B (1987) Response properties of area 17 neurons in cats reared in stroboscobic illumination. J. Neuroscience 57: 1511–1535.

Djupsund K, Hayden B, Ham T, Dan Y (2001) Stimulus-timing dependent plasticity of intracortical connectivity in adult V1. In: Society for Neuroscience, Abstracts, No. 619.26.

Durand S, Freeman T, Mante V, Kiper D, Carandini M (2001) Cross-orientation suppression in cat V1 with very fast stimuli. In: Society for Neuroscience, Abstracts.

Feidler J, Saul A, Murthy A, Humphrey A (1997) Hebbian learning and the development of direction selectivity: The role of geniculate response timing. Network: Computation in Neural Systems 8: 195–214.

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.

Freeman T, Durand S, Kiper D, Carandini M (2001) Cross-orientation suppression in V1 following cortical adaptation. In: Society for Neuroscience, Abstracts.

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

Gao H, Fu Y-X, Shen K, Dan Y (2001) Spatial plasticity of receptive fields in the cat primary visual cortex. In: Society for Neuroscience, Abstracts, No. 821.62.

Geinisman Y, Berry R, Disterhoft J, Power J, Van der Zee E (2001) Associative learning elicits the formation of multiple-synapse boutons. J. Neuroscience 21(15): 5568–5573.

Hawken M, Shapley R, Grosof D (1996) Temporal-frequency selectivity in monkey visual cortex. Vis. Neurosci. 13: 477–492.

Heeger D (1993) Modeling simple-cell direction selectivity with normalized half-squared, linear operators. J. Neurophysiol. 70(5): 1885–1898.

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

Hubel D, Wiesel T (1963) Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. J. Neurophysiol. 26: 994–1001.

Humphrey A, Saul A (1998) Strobe rearing reduces direction selectivity in area 17 by altering spatiotemporal receptive-field structure. J. Neuroscience 80: 2991–3004.

Humphrey A, Saul A, Feidler J (1998) Strobe rearing prevents the convergence of inputs with different response timings onto area 17 simple cells. J. Neuroscience 80: 3005–3020.

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.

Jagadeesh B, Wheat H, Kontsevich L, Ferster D (1997) Direction selectivity of synaptic potentials in simple cells of the cat visual cortex. J. Neurophysiol. 78: 2772–2789.

Kempter R, Gerstner W, van Hemmen J (1999) Spike-based compared to rate-based Hebbian learning. In: MS Kearns, SA Solla, DA Cohn, eds., Advances in Neural Information Processing Systems, vol. 11, pp. 125–131.

Lendvai B, Stern E, Chen A, Svoboda K (2000) Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex in vivo. Nature 404: 876–881.

Maex R, Orban G (1996) Model circuit of spiking neurons generating directional selectivity in simple cells. J. Neurophysiol. 75: 1515–1545.

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

Markram H, Pikus D, Gupta A, Tsodyks M (1998) Potential for multiple mechanisms, phenomena and algorithms for synaptic plasticity at single synapses. Neuropharmacology 37: 489–500.

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

Mehta M (2000) From hippocampus to V1: Effect of LTP on spatio-temporal dynamics of receptive fields. In: J Bower, ed., Trends in Neuroscience, Proceedings of the CNS'99. pp. 905–911.

Meister M, Lagnado L, Baylor D (1995) Concerted signaling by retinal ganglion cells. Science 270: 1207–1210.

Meister M, Wong R, Baylor D, Shatz C (1991) Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. Science 252(5008): 939–943.

Nagano T, Fujiwara M (1979) A neural network model for the development of direction selectivity in the visual cortex. Biol. Cybernetics 32(1): 1–8.

Ohzawa I, Sclar G, Freeman R (1985) Contrast gain control in the cat's visual system. J. Neurophysiol. 54: 651–667.

Orban G, Gulyas B, Spileers W, Maes H (1987) Responses of cat striate neurons to moving light and dark bars: Changes with excentricities. J. Opt. Soc. Am. A 4(8): 1653–1665.

Pasternak T, Schumer R, Grizzi M, Movshon J (1985) Abolition of visual cortical direction selectivity affects visual behavior in the cat. Exp. Brain Res. 61: 214–217.

Rao R, Sejnowski T (2000) Predictive sequence learning in recurrent neocortical networks. In: S Solla, T Leen, K-R Muller, ed., Advances in Neural Information Processing Systems 12. MIT Press, pp. 164–170.

Reid R, Soodak R, Shapley R (1987) Linear mechanisms of directional selectivity in simple cells of cat striate cortex. Proc. Natl. Acad. Sci. USA 84(23): 8740–8744.

Reid R, Soodak R, Shapley R (1991) Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. J. Neurophysiol. 66(2): 505–529.

Reid R, Victor J, Shapley R (1992) Broadband temporal stimuli decrease the integration time of neurons in cat striate cortex. Vis. Neurosci. 9: 39–45.

Saul A, Feidler J (2002) Development of response timing and direction selectivity in cat visual thalamus and cortex. J. Neuroscience 22: 2945–2955.

Saul A, Humphrey A (1992a) Evidence of input from lagged cells in the lateral geniculate nucleus to simple cells in cortical area 17 of the cat. J. Neurophysiol. 68: 1190–1208.

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

Senn W, Buchs N (2002) Spike-based synaptic plasticity and the emergence of direction selective simple cells: Mathematical analysis. J. Computational Neuroscience 13: 167–185.

Senn W, Tsodyks M, Markram H (2001) An algorithm for modifying neurotransmitter release probability based on pre- and post-synaptic spike timing. Neural Computation 13(1): 35–68.

Shapley R (1996) Linearity and non-linearity in cortical receptive field. Ciba. Found. Symp. 184: 71–87.

Softky W, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. J. Neuroscience 13(1): 334–350.

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

Soodak R, Shapley R, Kaplan E (1991) Fine structure of receptive-field centers of X and Y cells of the cat. Vis. Neurosci. 6: 621–628.

Stratford K, Tarczy-Hornoch K, Martin K, Baninster N, Jack J (1996) Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. Nature 382: 258–261.

Suarez H, Koch C, Douglas R (1995) Modeling direction selectivity of simple cells in striate visual cortex within the framework of the canonical microcircuit. J. Neuroscience 15(10): 6700–6719.

Sur M, Leamey C (2001) Development and plasticity of cortical areas and networks. Nat. Rev. Neuroscience 2: 251–262.

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.

Watson A, Ahumada A (1985) Model of human visual-motion sensing. J. Opt. Soc. Am. A 2: 322–341.

Wimbauer S, Wenisch O, van Hemmen J, Miller K (1997) Development of spatiotemporal receptive fields of simple cells: II. Simulation and analysis. Biol. Cybernetics 77(6): 463–477.

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

Zhang L, Tao H, Holt C, Harris W, Poo M (1998) A critical window in the cooperation and competition among developing retinotectal synapses. Nature 395: 37–44.