Learning Input Correlations through Nonlinear Temporally Asymmetric Hebbian Plasticity

R. Gütig,[1]* **R. Aharonov,**[2]* **S. Rotter,**[1] and **Haim Sompolinsky**[2,3]

[1]Institute of Biology III, University of Freiburg, 79104 Freiburg, Germany, and [2]Interdisciplinary Center for Neural Computation and [3]Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel

Triggered by recent experimental results, temporally asymmetric Hebbian (TAH) plasticity is considered as a candidate model for the biological implementation of competitive synaptic learning, a key concept for the experience-based development of cortical circuitry. However, because of the well known positive feedback instability of correlation-based plasticity, the stability of the resulting learning process has remained a central problem. Plagued by either a runaway of the synaptic efficacies or a greatly reduced sensitivity to input correlations, the learning performance of current models is limited. Here we introduce a novel generalized nonlinear TAH learning rule that allows a balance between stability and sensitivity of learning. Using this rule, we study the capacity of the system to learn patterns of correlations between afferent spike trains. Specifically, we address the question of under which conditions learning induces spontaneous symmetry breaking and leads to inhomogeneous synaptic distributions that capture the structure of the input correlations. To study the efficiency of learning temporal relationships between afferent spike trains through TAH plasticity, we introduce a novel sensitivity measure that quantifies the amount of information about the correlation structure in the input, a learning rule capable of storing in the synaptic weights. We demonstrate that by adjusting the weight dependence of the synaptic changes in TAH plasticity, it is possible to enhance the synaptic representation of temporal input correlations while maintaining the system in a stable learning regime. Indeed, for a given distribution of inputs, the learning efficiency can be optimized.

Key words: Hebbian learning; spike-timing-dependent plasticity; synaptic updating; symmetry breaking; unsupervised learning; infomax; activity-dependent development

Introduction

Correlation-based plasticity has long been proposed as a mechanism for unsupervised experience-based development of neuronal circuitry, particularly in the cortex. However, the specifics of a biologically plausible model of plasticity that can also account for the observed synaptic patterns have remained elusive. Two major issues are stability and competition (Miller and MacKay, 1994; Miller, 1996; Abbott and Nelson, 2000; Song et al., 2000; van Rossum et al., 2000; Rao and Sejnowski, 2001; van Ooyen, 2001). If maps, such as ocular dominance maps, emerge from initially random (but statistically homogeneous) synaptic configurations by a Hebbian mechanism (but see Crowley and Katz, 2000), this would imply that there is an inherent instability in the dynamics of synaptic learning that destabilizes an initially homogeneous synaptic pattern. However, this raises the question as to what mechanism prevents synapses from growing to unrealistic

values when driven by unstable dynamics. The emergence of inhomogeneous synaptic patterns also requires a competition mechanism that makes some synapses decrease their efficacies as other synapses grow in strength. Such competition is absent in the most naive Hebb rule, which contains only a mechanism for synaptic enhancement. Recent experiments have led to an important refinement of correlation-based or Hebbian learning, by showing that activity-induced synaptic changes can be temporally asymmetric with respect to the timing of presynaptic and postsynaptic action potentials with a precision of down to tens of milliseconds. Causal temporal ordering of presynaptic and postsynaptic spikes induces synaptic potentiation, whereas the reverse ordering induces synaptic depression (Levy and Steward, 1983; Debanne et al., 1994, 1998; Magee and Johnston, 1997; Markram et al., 1997; Bi and Poo, 1998, 2001; Zhang et al., 1998; Feldman, 2000; Sjöström et al., 2001).

In this work, we address the question of whether temporally asymmetric Hebbian (TAH) plasticity rules provide an adequate mechanism for unsupervised learning of input correlations. Two models of TAH plasticity have been studied recently that differ in the way that they implement the weight dependence of the synaptic changes and the boundaries of the allowed range of synaptic efficacies. The additive model (Abbott and Blum, 1996; Gerstner et al., 1996; Eurich et al., 1999; Kempter et al., 1999, 2001; Roberts, 1999; Song et al., 2000; Levy et al., 2001; Câteau et al., 2002) assumes that changes in synaptic efficacies do not scale with synaptic strength, and the boundaries are imposed as hard constraints. This model retains inherently unstable dynamics while exhibiting strong competition between afferent synapses. Be-

cause this model yields binary synaptic distributions, its ability to generate graded representations of input features is restricted. Moreover, because of the strong competition, patterns in the synaptic distribution can emerge that do not reflect patterns of correlated activity in the input. On the other hand, the multiplicative model (Kistler and van Hemmen, 2000; van Rossum et al., 2000; Rubin et al., 2001) assumes linear attenuation of potentiating and depressing synaptic changes as the corresponding upper or lower boundary is approached. This model results in stable synaptic dynamics. However, because of reduced competition, all synapses are driven to a similar equilibrium value, even at moderately strong input correlations. Thus, neither the additive nor the multiplicative model provides a satisfactory scenario for a robust learning rule that implements a synaptic storage mechanism of temporal structures in the inputs. Here, we introduce a nonlinear TAH Hebbian (NLTAH) model, a novel generalized updating rule that allows for continuous interpolation between the additive and multiplicative models. We demonstrate that by appropriately scaling the weight dependence of the updating, it is possible to learn synaptic representations of input correlations while maintaining the system in a stable regime. Preliminary results have been published previously in abstract form (Aharonov et al., 2001; Gütig et al., 2001).

Materials and Methods

Temporally asymmetric Hebbian plasticity. We describe TAH plasticity as a change in the synaptic efficacy w between a pair of cells, where the range of w is normalized to [0, 1]. A single pair of presynaptic and postsynaptic action potentials with time difference $\Delta t \equiv t_{post} - t_{pre}$ induces a change in synaptic efficacy Δw given by:

$$\Delta w = \begin{cases} -\lambda f_-(w) \times K(\Delta t) & \text{if } \Delta t \leq 0 \\ \lambda f_+(w) \times K(\Delta t) & \text{if } \Delta t > 0. \end{cases} \quad (1)$$

The temporal filter $K(\Delta t) = \exp(-|\Delta t|/\tau)$ (Song et al., 2000; van Rossum et al., 2000; Rubin et al., 2001) implements the spike-timing dependence of the learning. The time constant τ of the exponential decay determines the temporal extent of the learning window. Following experimental measurements (Bi and Poo, 1998), we let $\tau = 20$ msec throughout this paper. The learning rate λ, $0 < \lambda \ll 1$, scales the magnitude of individual weight changes. The temporal asymmetry of the learning is represented by the opposite signs of the weight changes for positive and negative time differences. The updating functions $f_+(w)$, $f_-(w) \geq 0$, which are in general weight dependent, scale the synaptic changes and implement synaptic potentiation for causal time differences ($\Delta t > 0$), and depression otherwise. Here, we introduce a family of nonlinear updating functions in which the weight dependence has the form of a power law with a non-negative exponent μ:

$$f_+(w) = (1 - w)^\mu \text{ and } f_-(w) = \alpha w^\mu, \quad (2)$$

with $\alpha > 0$ denoting a possible asymmetry between the scales of potentiation and depression. Figure 1 A shows the updating curves (Eq. 2) for several values of μ. For $\mu = 0$, the updating functions are independent of the current synaptic efficacy, and the rule recovers the additive TAH learning model. This model requires that weights, which would have left the allowed range after an updating step, are clipped to the appropriate boundary (0 or 1). The case $\mu = 1$ corresponds to the multiplicative model, in which the updating functions linearly attenuate positive and negative synaptic changes as a synapse approaches the upper or lower boundary of the allowed range. Intermediate values of the updating parameter μ determine the range of the boundary effects on the changes in w. Note that any non-zero μ, given a sufficiently small learning rate, automatically prevents the synaptic efficacies from leaving the allowed range [0, 1], thereby preventing the runaway problem of synaptic efficacies and removing the necessity of artificially clipping synaptic weights. Figure 1 B provides an illustrative example of the effects of the parameter μ on a sequence of synaptic weight changes (see legend for details).
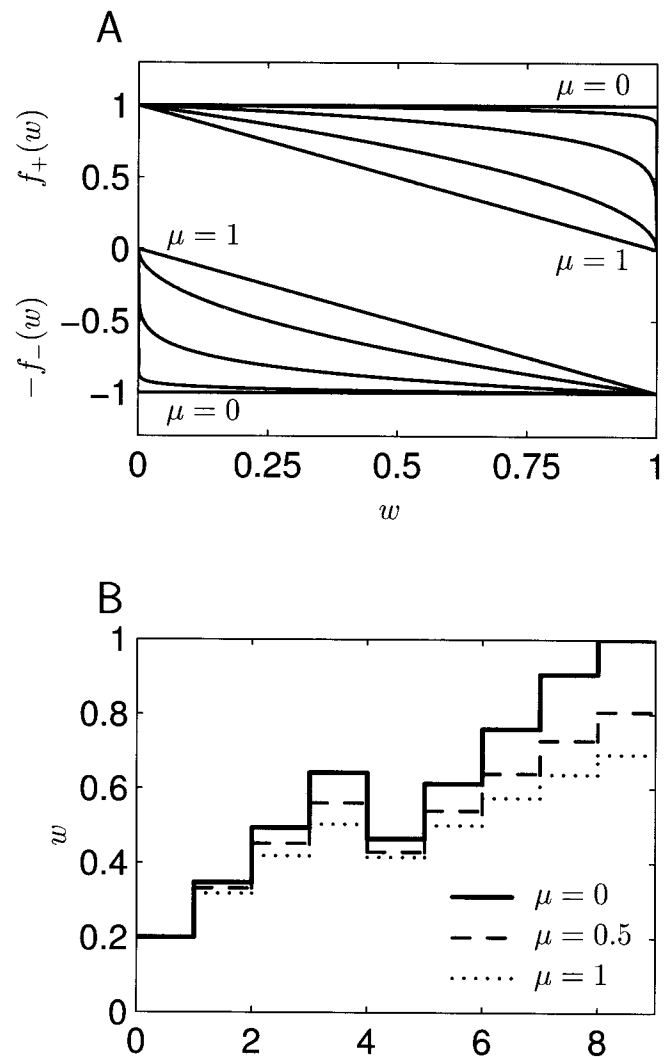

Figure 1. Non-linear TAH. *A*, Effect of the parameter μ on the updating functions $f_+(w)$ (top half) and $f_-(w)$ (bottom half) for $\mu = 0, 0.02, 0.15, 0.5, 1$ ($\alpha = 1$). As μ increases, the curves change from the constant additive updating curves ($\mu = 0$, horizontal lines at 1 and −1) to the multiplicative updating functions with linear weight dependence ($\mu = 1$, straight lines with slope −1). *B*, Illustration of the effect of μ on synaptic changes. At each step, one pair of presynaptic and postsynaptic spikes induces a change in the efficacy w according to Equations 1 and 2. To elucidate the weight dependency of the updating, the magnitude of the time difference between the spikes Δt is the same in all steps, namely $-\tau$ at step 4 and τ at all other steps. For illustrative purposes, we assume that the spike pairs are sufficiently distant so that the weight change is effected only by the current spike pair. Furthermore, the effect of a single pair is unrealistically magnified ($\lambda = 0.4$). For additive updating ($\mu = 0$), all potentiating changes are of equal magnitude, because Δt is constant here and the updating is independent of the synaptic efficacy. Note, however, that the last additive efficacy change would have resulted in an efficacy of >1, and hence the efficacy is clipped to 1. The depression in step 4 is larger exactly by a factor of $\alpha = 1.2$. In contrast, when $\mu > 0$, although the contribution to the synaptic change from the time dependency is constant, the actual change in synaptic efficacy is not. Because potentiation is scaled by $(1 - w)^\mu$, as the synapse becomes stronger the same time difference induces a smaller change. This scaling effect is more pronounced for larger values of μ (compare the cases of $\mu = 0.5$ and $\mu = 1$).

Following previous work (Kempter et al., 1999; Song et al., 2000; Rubin et al., 2001; but see van Rossum et al., 2000), the plasticity effects of individual spike pairs are assumed to sum independently: given a postsynaptic spike, each synapse is potentiated according to Equations 1 and 2 by pairing the output spike with all preceding synaptic events. Conversely, a synapse is depressed when a presynaptic event occurs, using all pairs that the synaptic event forms with preceding output spikes.

Mean synaptic dynamics. Because in general the spike times of the

presynaptic and postsynaptic neurons are stochastic, the dynamics of synaptic changes are also a stochastic process. However, if the learning rate λ is small, the noise accumulated over an appreciable amount of time is small relative to the mean change in the synaptic efficacies, called the synaptic drift. This drift, denoted as $\dot{w}$, is the mean rate of change of the synaptic efficacy. Using Fokker–Planck mean field theory, the synaptic drifts are described in terms of the correlations between the presynaptic and postsynaptic activity (Kempter et al., 1999, 2001; Kistler and van Hemmen, 2000; Rubin et al., 2001). We consider a pair of stationary presynaptic and postsynaptic processes described by the pulse trains $\rho^{\mathrm{pre}}(t) = \Sigma_k \delta(t - t_k^{\mathrm{pre}})$ and $\rho^{\mathrm{post}}(t) = \Sigma_k \delta(t - t_k^{\mathrm{post}})$, with mean rates $r^{\mathrm{pre}} = \langle \rho^{\mathrm{pre}} \rangle$ and $r^{\mathrm{post}} = \langle \rho^{\mathrm{post}} \rangle$ and the raw cross-correlation function:

$$\Gamma_{\mathrm{pre,post}}(\Delta t) = \langle \rho^{\mathrm{pre}}(t)\rho^{\mathrm{post}}(t + \Delta t) \rangle_t. \qquad (3)$$

The angular brackets denote averaging over time t while keeping the time lag Δt between the two spike trains fixed. This cross-correlation is the probability density for the occurrence of pairs of presynaptic and postsynaptic spikes with time difference Δt. Using this probability density, the synaptic drift $\dot{w}$ is given by integrating the synaptic changes Δw (Eq. 1) over the time differences Δt weighted by the probabilities $\Gamma_{\mathrm{pre,post}}(\Delta t)$:

$$\dot{w} = -\lambda f_-(w) \int_{-\infty}^{0} d\Delta t K(\Delta t)\Gamma_{\mathrm{pre,post}}(\Delta t)$$

$$+ \lambda f_+(w) \int_{0}^{\infty} d\Delta t K(\Delta t)\Gamma_{\mathrm{pre,post}}(\Delta t). \qquad (4)$$

The integral in the first term represents the synaptic depression that stems from all input–output correlations with negative time lag (i.e., acausal correlations). These correlations are filtered by the temporal window $K(\Delta t)$ of the learning. This contribution from the spike-timing dependence of the learning is multiplied by the weight-dependent scale of depressing synaptic changes $f_-(w)$. Conversely, the second term represents the potentiating drift originating from causal input–output correlations, which is scaled by $f_+(w)$. Note that the weight-dependent scales $f_\pm(w)$ are evaluated outside the time integrals, because when λ is small, w does not change appreciably during the time scale of the temporal filter of learning. In summary, the dynamic evolution of the synaptic weights depends on the properties of the correlation between the presynaptic and postsynaptic activity $\Gamma_{\mathrm{pre,post}}(\Delta t)$, which in turn depend on the details of the spike generation mechanism of the postsynaptic cell as well as on the statistics of the afferent inputs (Kuhn et al., 2003).

Integrate-and-fire neuron. To study the implications of the above NLTAH plasticity model in a biologically motivated spiking neuron, we simulate a leaky integrate-and-fire neuron, with parameters similar to those of Song et al. (2000). The membrane potential of the neuron is described by:

$$C_{\mathrm{m}}\frac{dV}{dt} = \frac{V_{\mathrm{rest}} - V}{R_{\mathrm{m}}} + g_{\mathrm{exc}}(t)(E_{\mathrm{exc}} - V) + g_{\mathrm{inh}}(t)(E_{\mathrm{inh}} - V),$$

with membrane capacitance $C_{\mathrm{m}} = 200$ pF, membrane resistance $R_{\mathrm{m}} = 100$ MΩ, resting potential $V_{\mathrm{rest}} = -70$ mV, and excitatory and inhibitory synaptic reversal potentials $E_{\mathrm{exc}} = 0$ mV and $E_{\mathrm{inh}} = -70$ mV, respectively. Whenever the membrane potential exceeds a threshold of -54 mV, an action potential is generated and the neuron is reset to the resting potential with no refractory period. Modeling synaptic conductance dynamics by α-shaped response functions, excitatory and inhibitory conductances are given by:

$$g_{\mathrm{A}}(t) = \bar{g}_{\mathrm{A}} \sum_{j=1}^{N_{\mathrm{A}}} w_j(t) \sum_{t_j < t}(t - t_j)\exp(-(t - t_j)/\tau_{\mathrm{A}}), \quad A = \mathrm{exc, inh},$$

respectively, where the t_j values are the spike times of synapse j and $\tau_{\mathrm{exc}} = \tau_{\mathrm{inh}} = 5$ msec. The values $\bar{g}_{\mathrm{exc}} = 30$ nS and $\bar{g}_{\mathrm{inh}} = 50$ nS were chosen such that the total charge injected per spike (at the threshold potential) is $Q_{\mathrm{exc}} = 0.04$ pC and $Q_{\mathrm{inh}} = 0.02$ pC, respectively. While the efficacies w of the $N = N_{\mathrm{exc}} = 1000$ excitatory synapses are plastic and governed by the

TAH learning rule, all $N_{\mathrm{inh}} = 200$ inhibitory efficacies are held fixed at -1. In the numerical simulations, the integrate-and-fire neuron is driven by Bernoulli (i.e., zero-one) processes defined over discrete time bins of duration $\Delta T = 0.1$ msec, approximating Poisson spike trains with a stationary rate r. For the inhibitory inputs, $r = 10$ Hz. All equilibrium synaptic distributions obtained with this model neuron result from an initially uniform synaptic state with all efficacies set to 0.5. In each case, the learning process (learning rate, $\lambda = 0.001$) is simulated until the shape of the synaptic distribution ceases to change.

Linear Poisson neuron. To investigate analytically the properties of the TAH learning rule, we consider in addition a linear Poisson neuron (Kempter et al., 2001). The spiking activity of this neuron $\rho^{\mathrm{post}}(t)$ is a realization of a Poisson process with the underlying instantaneous rate function:

$$R^{\mathrm{post}}(t) = \frac{1}{N} \sum_{j=1}^{N} w_j(t)\rho_j^{\mathrm{pre}}(t - \epsilon), \qquad (5)$$

where, as before, the N presynaptic input spike trains and the output spikes are characterized by a series of δ pulses [i.e., $\rho_j^{\mathrm{pre}}(t) = \Sigma_k \delta(t - t_k^{\mathrm{pre},j})$ and $\rho^{\mathrm{post}}(t) = \Sigma_k \delta(t - t_k^{\mathrm{post}})$]. The parameter $0 < \epsilon \ll \tau$ denotes a small constant delay in the output. Because this delay is small compared with the temporal window of learning, we approximate $\exp(-\epsilon/\tau) \approx 1$ throughout this work. As before, $w_j(t) \in [0, 1]$ denotes the efficacy of the jth synapse. Except for Figures 5 and 9, in which we investigate the large N limit, we let $N = 100$ throughout this work.

In Figure 2A, we numerically simulate the linear Poisson neuron receiving uncorrelated Poisson input spike trains. Generating the spike arrival times in continuous time (down to machine precision), the postsynaptic process defined in Equation 5 is implemented by generating a postsynaptic spike with probability w_i/N, whenever a presynaptic spike arrives at a synapse (i) of the neuron.

Mean synaptic dynamics for the linear Poisson neuron. For the integrate-and-fire neuron, there is no simple exact expression relating the correlations between the presynaptic and postsynaptic spike trains to the system parameters such as the rates and input correlations. However, because of the linear summation of inputs in the linear Poisson neuron (Eq. 5), this model permits the expression of the input–output correlations $\Gamma_{\mathrm{pre,post}}(\Delta t)$ in closed form. Considering the case that all input spike trains have a common rate r, we obtain from Equations 3 and 5 that the correlation of the activity at synapse i with the output activity is:

$$\Gamma_{i,\mathrm{post}}(\Delta t) = \langle \rho_i^{\mathrm{pre}}(t)\rho^{\mathrm{post}}(t + \Delta t) \rangle_t$$

$$= \frac{1}{N} \sum_{j=1}^{N} w_j \langle \rho_i^{\mathrm{pre}}(t)\rho_j^{\mathrm{pre}}(t + \Delta t - \epsilon) \rangle_t.$$

Substituting the above in Equation 4, and rearranging the terms, we obtain the drift of the ith synapse:

$$\dot{w}_i = -\lambda f_-(w_i)\frac{1}{N} \sum_{j=1}^{N} w_j \tau r^2 \left(1 + \int_{-\infty}^{0} d\Delta t \frac{1}{\tau}K(\Delta t)\Gamma_{ij}^o(\Delta t - \epsilon)\right)$$

$$+ \lambda f_+(w_i)\frac{1}{N} \sum_{j=1}^{N} w_j \tau r^2 \left(1 + \int_{0}^{\infty} d\Delta t \frac{1}{\tau}K(\Delta t)\Gamma_{ij}^o(\Delta t - \epsilon)\right),$$

where we define the normalized cross-correlations between the input spike trains by:

$$\Gamma_{ij}^o(t') = \frac{\langle \rho_i^{\mathrm{pre}}(t)\rho_j^{\mathrm{pre}}(t + t') \rangle_t}{r^2} - 1. \qquad (6)$$

We denote the integrated normalized cross-correlations appearing in the above drift equation by:

$$C_{ij}^\pm = \int_{0}^{\infty} d\Delta t \frac{1}{\tau}K(\Delta t)\Gamma_{ij}^o(\pm\Delta t - \epsilon). \qquad (7)$$

These matrices are the effective between-input correlations for positive and negative time lags. If $C_{ij}^+ > 0$, the activity at synapse j temporally follows that at synapse i, such that its contribution to the postsynaptic activity results in a potentiating drift on synapse i. Conversely, if $C_{ij}^- > 0$, the activity at synapse j precedes that at synapse i and contributes to its depression. Note that the effective correlations C_{ij}^+ are zero if the ith and jth input spike trains are uncorrelated. Finally, the synaptic drifts can be written as:

$$\dot{w}_i = \frac{\lambda \tau r^2}{N}\left[-\Delta f(w_i)\sum_{j=1}^{N} w_j + f_+(w_i)\sum_{j=1}^{N} w_j C_{ij}^+ - f_-(w_i)\sum_{j=1}^{N} w_j C_{ij}^-\right],$$

(8)

with $\Delta f = f_- - f_+$. Note that the first term in Equation 8 describes competition between the synapses when $\Delta f > 0$; independently from the input correlations, the amount of induced depression on a given synapse w_i is large when other synapses w_j are strong. The second term represents the cooperative increase of the synaptic weights inherent in TAH learning. In contrast, the last term denotes depressive synaptic interactions stemming from negative time correlations in the input activity.

Generating correlated inputs. We consider input spike trains with rate r and instantaneous correlations defined by:

$$\Gamma_{ij}^o(\Delta t) = \frac{1}{r}c_{ij}\delta(\Delta t),$$

(9)

where $\delta(t)$ is the Dirac-δ function and c_{ij} is non-negative. In this case,

$$C_{ij}^- = 0 \text{ and } C_{ij}^+ = \frac{c_{ij}}{\tau r}.$$

(10)

The backward effective correlations C_{ij}^- vanish because the argument of $\Gamma_{ij}^o(-\Delta t - \epsilon)$ in Equation 7 is never 0. Recall that for a Poisson process $\rho(t)$ with rate r, the raw autocorrelation is $\langle \rho(t)\rho(t + \Delta t)\rangle = r^2 + r\delta(\Delta t)$. Hence, the normalized autocorrelation is $c_{ii} = 1$ and the between-input correlations are $c_{ij} \leq 1$, with equality only if the two spike trains i and j are identical. In the numerical simulations, we generate populations of correlated spike trains by conditioning the binwise spike probabilities at time bin T on the activity of a common reference Bernoulli spike train $X_0(T)$ with the binwise spike probability $r\Delta T$. To obtain a positive pairwise correlation coefficient of $0 \leq c = \mathrm{Cov}(X_i(T), X_j(T))/\sqrt{\mathrm{Var}(X_i(T))\mathrm{Var}(X_j(T))}$ between two spike trains $X_i(T)$ and $X_j(T)$, the conditional probabilities $\vartheta = P(X_k(T) = 1|X_0(T) = 1)$ and $\varphi = P(X_k(T) = 1|X_0(T) = 0)$ for $k = i, j$ are determined by:

$$\vartheta = r\Delta T + \sqrt{c}(1 - r\Delta T), \quad \varphi = r\Delta T(1 - \sqrt{c}).$$

(11)

This choice of ϑ and φ for all spike trains within the correlated group guarantees that the spike trains have rates r and an instantaneous pairwise correlation coefficient c (see Appendix). For small bin sizes, this process mimics the instantaneously correlated Poisson point processes defined above (Eq. 9). We will also consider the case of delayed correlations of the form $\Gamma_{ij}^o(\Delta t) = r^{-1}c_{ij}\delta(\Delta t - D_{ij})$. These correlations are obtained by shifting the instantaneously correlated input spike trains relative to each other by a time delay D_{ij}.

Measuring the performance of learning rules. A natural way to measure the performance of a learning rule is to quantify its ability to imprint the statistical features of the neuronal input onto the distribution of the learned synaptic weights. One measure of this ability is the mutual information between the neuronal inputs and the synaptic weights. However, direct calculation of the mutual information in cases in which the number of synaptic weights is large is computationally not feasible. Instead, we use here a related quantity that measures the effect of a small change in the statistics of the input on the learned synaptic weights. We denote the features of an ensemble of neuronal inputs by the vector $\Phi = (\Phi_1, \ldots, \Phi_R)$, where the Φ_i parameterize specific input features (e.g., mean strength of the inputs or temporal correlations between different inputs).

Given these features, we calculate the $N \times R$ susceptibility matrix χ_{ij}, the elements of which are:

$$\chi_{ij} \equiv \frac{\partial w_i}{\partial \Phi_j}.$$

(12)

The ijth element measures the amount of change in the ith synaptic efficacy that is incurred by a small change in the jth input feature, Φ_j. A global sensitivity measure S is constructed from this matrix by calculating:

$$S \equiv \frac{1}{2}\ln|\det(\chi^T\chi)|,$$

(13)

where $\det(\cdot)$ denotes the determinant. The average sensitivity S_{avg} is defined as $\langle S\rangle_\Phi$, where the average $\langle\cdot\rangle$ is taken over the distribution of the feature vector Φ. The rational for calculating S is that it is closely related to the mutual information between the input features and the weight distribution. Specifically, if the mapping from the feature space to the synaptic weight space induced by the learning dynamics is invertible, maximizing S_{avg} is equivalent to maximizing the mutual information (Bell and Sejnowski, 1995; Shriki et al., 2001) in the limit of a small learning rate λ. In this work, we focus on the equilibrium properties of the TAH learning rule (i.e., the weight distributions that result after the learning dynamics have converged to a stable stationary state). Therefore, χ_{ij} is evaluated at the fixed point solution $w^\star$ of the drift equations in the linear Poisson neuron, Equation 8 (see Appendix). The possibility of using the analytic expressions for χ_{ij} in calculating the sensitivity S is the main advantage of using it as a measure of performance of various learning models.

Results

To understand learning phenomena in biological nervous systems in terms of neural network function, it is crucial to bridge the gap between the microscopic mechanisms that implement experience-based changes in neuronal signaling pathways and the macroscopic properties of the learning system composed of these pathways. In this paper, we focus on two general goals of learning that can be defined at the network level and also investigate the importance of the updating parameter μ of the learning rule in these contexts. First, we consider the question of how a network can develop a functional connectivity architecture, as for example in ocular dominance columns. As noted in the Introduction, this type of learning task typically requires the synaptic learning dynamics to be competitive, to allow segregation between initially homogeneous synaptic populations. Moreover, it is important that the learning process is robust in the sense that the learned synaptic patterns faithfully reflect meaningful features in the neuronal input activity, rather than being dominated by contributions from random noise. Therefore, we study here how the interplay between competition and stability in TAH plasticity affects the learned synaptic distributions. In the second part of Results, we turn to the conceptually different learning task of imprinting information about the input activity of a neuron into the respective synaptic efficacies. In this context, the sensitivity of the learning dynamics to features in the neuronal input becomes crucial. Thus, using the sensitivity measure introduced in Materials and Methods, the second part of Results concentrates on a quantitative evaluation of the performance of different TAH learning rules.

The emergence of synaptic patterns by symmetry breaking in TAH learning

One of the basic requirements for the activity-driven formation of cortical maps is the ability of the learning to generate spatially inhomogeneous synaptic patterns from a population of synapses

with statistically homogeneous inputs. The emergence of such symmetry breaking is an essential property of current cortical plasticity models (Miller, 1996). In this section, we study the conditions under which the TAH learning models introduced above exhibit symmetry breaking and, hence, qualify as candidate models for the development of functional maps. Moreover, because the learning dynamics may also lead to symmetry breaking that overrides the correlation structure of the afferent activity, it is important to ask what learning rules ensure a faithful representation of the input activity within the learned synaptic connections. We address these questions in three basic types of homogeneous afferent activities, that differ with respect to the correlation structure of the input spike trains: uncorrelated inputs, uniformly correlated inputs, and uniformly correlated subpopulations without correlations between the subpopulations ("correlated subgroups"). Before treating these specific cases, we highlight the general features of the synaptic learning dynamics in a population of synapses with statistically homogeneous input activities. These results apply to all three cases of homogeneous populations of inputs.

Dynamics of a population of synapses with homogeneous inputs
To study the symmetry breaking in the synaptic patterns, we consider the learning dynamics in cases in which the input statistics are spatially homogeneous. This means that each input obeys the same spike statistics and has the same pattern of correlations with the other inputs. This assumption implies that the presynaptic rates r_i (where i denotes the index of the different afferents) are all equal. Likewise, the total sum of the correlations that each input has with the rest of the inputs is the same. In particular, the mean effective causal correlations, C_0:

$$C_0 = \frac{1}{N} \sum_{j=1}^{N} C_{ij}^+, \qquad (14)$$

is the same for all input channels i, and similarly for the backward correlations (Kempter et al., 2001).

To understand the implications of spatial homogeneity in the presynaptic inputs on the learning dynamics, it is useful to concentrate on the linear Poisson neuron model (Eqs. 5, 8). For convenience, we assume that all correlations between input spikes are instantaneous (see Materials and Methods, Eq. 10).

The important consequence of the spatial homogeneity across the presynaptic inputs is that the product of the effective correlation matrix C_{ij}^+ with a homogeneous vector of synaptic efficacies $\mathbf{w}_o = (w_o, w_o, \ldots, w_o)$ remains homogeneous, because $C^+ \mathbf{w}_o = NC_0 \mathbf{w}_o$. Hence, the homogeneous synaptic state is an eigenvector of the effective correlation matrix C^+ with eigenvalue NC_0. The existence of this homogeneous eigenvector is important for the synaptic learning, because it means that in a homogeneous synaptic state $\mathbf{w}_o$, all synapses experience identical drifts (Eq. 8). Moreover, if there is a $w_o = w^\star$ such that the synaptic drifts become zero, the learning dynamics have a steady-state solution, with $w_i = w^\star$ for all synapses. We call a solution in which all of the learned synapses are equal a homogeneous solution. Indeed, we show in the Appendix that for all non-zero updating parameter μ values in our model (Eq. 2), there exists a steady-state homogeneous solution of the learning ($\dot{w}_i = 0$ in Eq. 8), with $w^\star$ being the solution of the equation:

$$\frac{f_-(w^\star)}{f_+(w^\star)} = \alpha \left(\frac{w^\star}{1 - w^\star} \right)^{\mu} = 1 + C_0. \qquad (15)$$

This equation expresses the weight-dependent balance between depression and potentiation that is controlled by the mean effective correlation C_0.

Although the homogeneous synaptic steady state always exists, it may be unstable with respect to small perturbations of the synaptic efficacies, driving them into inhomogeneous states. Because of the important functional consequences of this emergence of inhomogeneous synaptic patterns at the network level, it is important to understand the features of the learning dynamics that give rise to this phenomenon of symmetry breaking. Therefore, we analyze the effects of small deviations of the synaptic efficacies from the homogeneous synaptic steady-state $\mathbf{w}^\star$. For each synapse w_i, we denote a corresponding small deviation from the homogeneous solution by $\delta w_i = w_i - w^\star$ and express its temporal evolution as a function of all deviations δw_j. As we show in the Appendix, this temporal evolution is determined by three separate contributions:

$$\dot{\overline{\delta w}}_i = \lambda \tau r^2 \left[-g_o \delta w_i - \frac{\Delta f(w^\star)}{N} \sum_{j=1}^{N} \delta w_j + \frac{f_+(w^\star)}{N} \sum_{j=1}^{N} C_{ij}^+ \delta w_j \right],$$

where $g_o = w^\star f_+(w^\star) \frac{d}{dw}[f_-(w)/f_+(w)]_{w=w^\star} = \alpha \mu w^{\star \mu}/(1 - w^\star) > 0$ (Appendix, Eqs. 20, 21). The first term is a local stabilizing term. It counteracts individual deviations from the homogeneous solution, maintaining the synaptic efficacies at the same value $w^\star$. To understand the origin of this stabilizing term in the learning dynamics, we consider the effect of a single synaptic deviation δw_i on the balance between depression and potentiation. If a synapse is strengthened by a deviation $\delta w_i > 0$, the resulting scale of potentiation $f_+(w^\star + \delta w_i)$ decreases, whereas the scale of depression $f_-(w^\star + \delta w_i)$ increases (Eq. 2). Conversely, a weakening deviation $\delta w_i < 0$ shifts the balance between potentiation and depression in favor of potentiation. Because this stabilizing drift stems from the weight dependence of the ratio $f_-(w^\star)/f_+(w^\star)$ (Appendix), it is not present in the additive model ($\mu = 0$), where the $f_\pm$ values themselves are constant. The second term is proportional to the net drift $-\Delta f(w^\star) = f_+(w^\star) - f_-(w^\star)$. This drift is negative, because at the homogeneous solution, $f_-(w^\star) > f_+(w^\star)$ when depression balances the potentiating correlations (see Eq. 15, and recall that $C_0 > 0$). The negative drift is multiplied by the total perturbation $\Sigma_j \delta w_j$, which denotes the change in the output rate attributable to the changes in the synaptic efficacies. Thus, this term represents the competition between the synapses. This competition results from the fact that strengthening the efficacy of any synapse increases the output rate, thereby increasing the frequency of occurrence of net negative drift in all of the synapses. It is important to note that this competition is acting between all synapses, unrelated to the correlation structure in the afferent input.

Finally, the last term is a cooperative term. Synapses that are positively correlated cooperate to elevate their weights. This cooperation is driven by the potentiating component of the TAH learning and depends on the pattern of correlations among the input channels. We emphasize that the cooperativity in the synaptic learning in general does not originate from a possible advantage of correlated synapses to drive a potentially nonlinear spike generator of the postsynaptic cell, but rather already occurs because of an inherently increased probability of correlated synapses to precede postsynaptic spikes, even when nonlinear cooperative effects in the spike generator are absent.

The stability of the homogeneous synaptic steady state results

from the interplay between the stabilizing, the competitive, and the cooperative drifts in the learning dynamics. As we derive in the Appendix, perturbations of the steady state that slightly change all weights by the same amount δw (homogeneous perturbations) decay to zero with time and, hence, do not destabilize the learning of a homogeneous synaptic distribution. In contrast, inhomogeneous perturbations (i.e., perturbations in which the deviations of the synaptic efficacies from $w^\star$ are not identical) can grow exponentially through the learning dynamics and drive the system into inhomogeneous synaptic states. In the Appendix, we specifically show that the homogeneous synaptic state becomes unstable if the largest real part of all inhomogeneous eigenvalues (eigenvalues corresponding to inhomogeneous eigenvectors) of the effective correlation matrix C^+ is sufficiently large. Denoting this eigenvalue by NC_1, we find that when:

$$\mathcal{H} = C_1 f_+(w^\star) - g_o > 0, \qquad (16)$$

the homogeneous state is unstable. This inequality means that symmetry breaking occurs whenever the cooperation between synapses C_1 is strong enough to outweigh the stabilizing term g_o. Note that the competition coefficient Δf does not enter directly into the stability criterion. This is because the competition term is proportional to the total weight value, and hence is not sensitive to inhomogeneous perturbations that do not change this value. Nevertheless, this term has a crucial role in the stability of the learning, because it suppresses the homogeneous growth of all synapses. As is shown in the Appendix, Equation 16 implies that the homogeneous solution is always stable in the multiplicative model ($\mu = 1$).

Although this analysis was performed using the plasticity equations of the linear Poisson neuron, it is qualitatively valid as well for other neuron models, as we show for specific cases. Below we study how the emergence of symmetry breaking (i.e., transitions from homogeneous to inhomogeneous synaptic distributions) depends on the nonlinearity of the TAH dynamics, namely the parameter μ, as well as on the asymmetry between depression and potentiation α, and on the size of the synaptic population N.

Uncorrelated inputs: linear neuron
In this section, we investigate the synaptic distributions that result from the TAH learning process when the postsynaptic neuron is driven by independent Poisson spike trains of equal rate r. For this input regime, it has been found in an integrate-and-fire neuron that additive learning ($\mu = 0$) breaks the symmetry of the statistically identical presynaptic inputs and leads to a bimodal weight distribution (Song et al., 2000; Rubin et al., 2001). However, it was shown by Rubin et al. (2001) that multiplicative learning ($\mu = 1$) leads to a unimodal distribution of synapses. As shown in the preceding section, these qualitatively different learning behaviors originate in the stabilizing effect of the weight dependence of the synaptic changes on the homogeneous synaptic state. Here, we study the generalized nonlinear TAH rule with arbitrary $\mu \in [0, 1]$.

In the uncorrelated case, $c_{ij} = \delta_{ij}$, and hence $C_{ij}^+ = \delta_{ij}/\tau r$. Both its homogeneous and inhomogeneous eigenvalues normalized by N are:

$$C_0 = C_1 = \frac{1}{\tau r N}. \qquad (17)$$

Thus, the only cooperation in the learning dynamics stems from the positive feedback induced by the correlation of each synapse

with its own contribution to the postsynaptic activity. The effects of this self-correlation on the learning dynamics decrease inversely to the effective size of the presynaptic population $\tau r N$ (i.e., the expected number of spikes that arrive within the learning time window).

By inserting the above expression for C_0 into Equation 15, we obtain the steady-state efficacy $w^\star$ of the synaptic population when the learned synaptic state is homogeneous (see Appendix, Eq. 19). In this case, the output rate of the linear neuron is given by this steady-state efficacy times the rate of the presynaptic inputs r (compare Eq. 5). Figure 2A depicts the output rate of the postsynaptic neuron as a function of the presynaptic input rate r, for $\alpha = 1.05$. We focus on this value of α here, because we want to compare the nonlinear rules with the additive rule. In the latter case, α must be close to 1; otherwise, practically all synapses will become zero (see Appendix). For $\mu = 1$ (multiplicative TAH), the efficacy $w^\star$ is fairly independent of r and, hence, the output rate grows linearly with the input rate. However, if μ is sufficiently small, $w^\star$ decreases inversely with the input rate, resulting in the output rate being nearly constant.

To study the regime in which the synaptic learning dynamics break the symmetry of the uncorrelated input population, we substitute Equation 17 into Equations 15 and 16, computing the homogeneous solution $w^\star$ (Eq. 19) and the regime of its stability. Figure 3A depicts the critical contour lines according to the stability condition (Eq. 16). Each line traces the critical combination of the parameters μ and $\tau r N$ for a fixed value of α, such that $\mathcal{H} = 0$. Outside the corresponding contour ($\mathcal{H} < 0$), the homogeneous synaptic state is stable, and thus learning generally results in all synapses having the same efficacy. In contrast, inside the contour line, the learning dynamics induce symmetry breaking.

Figure 3A shows how the outcome of TAH learning depends on the effective size of the presynaptic population. For a sufficiently small $\tau r N$, the relative contribution of each input channel to the postsynaptic activity is large and, hence, the resulting strong positive feedback drives all synapses to a stable homogeneous state near the upper boundary (Fig. 3B, squares). In contrast, as $\tau r N$ is increased, the effect of a single synapse on the postsynaptic activity decreases. Therefore, for a sufficiently large $\tau r N$, the stabilizing force induced by the weight dependence of the synaptic changes dominates the learning dynamics for any non-zero μ, resulting in a stable homogeneous synaptic state (Fig. 3B, triangles). In between the two extremes of small and large $\tau r N$, there is a regime of intermediate effective population sizes for which symmetry breaking may occur, with the synaptic population segregating into a strong and a weak group. Such a case is shown in Figure 3B (circles).

Importantly, Figure 3A demonstrates that as the number of afferents N increases, the regime of values of μ for which the homogeneous solution is unstable shrinks to zero. The inset of Figure 3A shows the value of μ at the border between stability and instability of the homogeneous solution, as a function of the effective population size. It is apparent that this μ decreases linearly with $1/(\tau r N)$ when $\tau r N$ is large (also see Appendix). Hence, for any sizable degree of weight dependence and large synaptic populations, symmetry breaking does not occur.

In the purely additive TAH model, synaptic changes do not scale at all with the efficacy of a synapse, and the weights have to be constrained by an additional clipping to prevent unrealistic synaptic growth. As a result, the additive learning dynamics do not possess stationary synaptic states in the above sense that the individual synaptic drifts become zero. Instead, synapses with
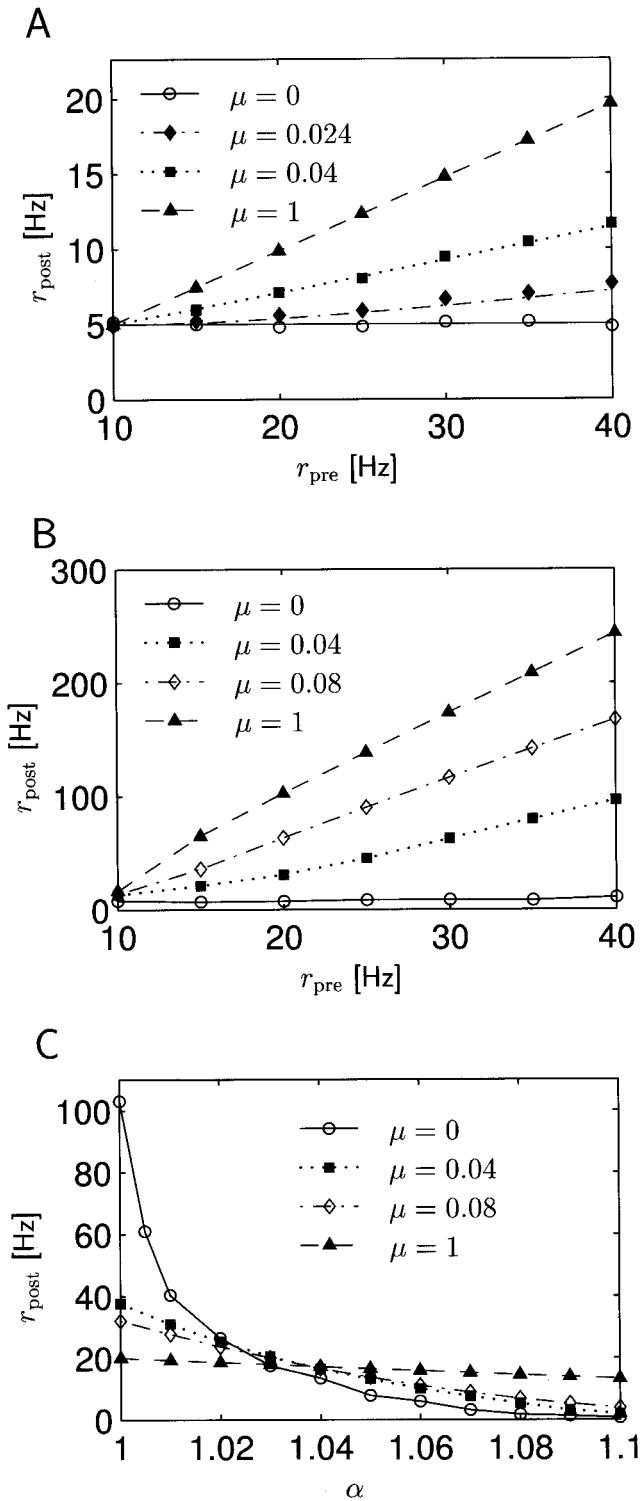
Figure 2. Firing rate responses of a neuron driven by uncorrelated Poisson input processes. *A*, Output firing rate of the linear Poisson neuron as a function of the input rate for selected values of $\mu =$ 0, 0.024, 0.04, 1, with N = 100 and $\alpha =$ 1.05. Solid lines show analytical results derived from the homogeneous synaptic state (Eq. 19) (which is stable for all values $\mu > 0$ and r shown here) or for additive TAH ($\mu = 0$) from the ratio of strong synapses given by Equation 25 (see Results and Appendix). In both cases, the output rate is given by the input rate multiplied with the corresponding mean synaptic efficacy. Plot symbols depict the output rates of a numerically simulated spiking linear Poisson neuron (Materials and Methods) with parameters as in the analytical calculation and a learning rate $\lambda = 0.001$ (except for $\mu = 0$, where we used $\lambda = 0.003$). *B*, *C*, Results from the integrate-and-fire neuron with N = 1000. All results refer to the neuron after convergence of the learning process. *B*, The output firing rate of the neuron as a function of the input rate for $\alpha = 1.05$. *C*, The output firing rate as a function of α for an input rate of 10 Hz.
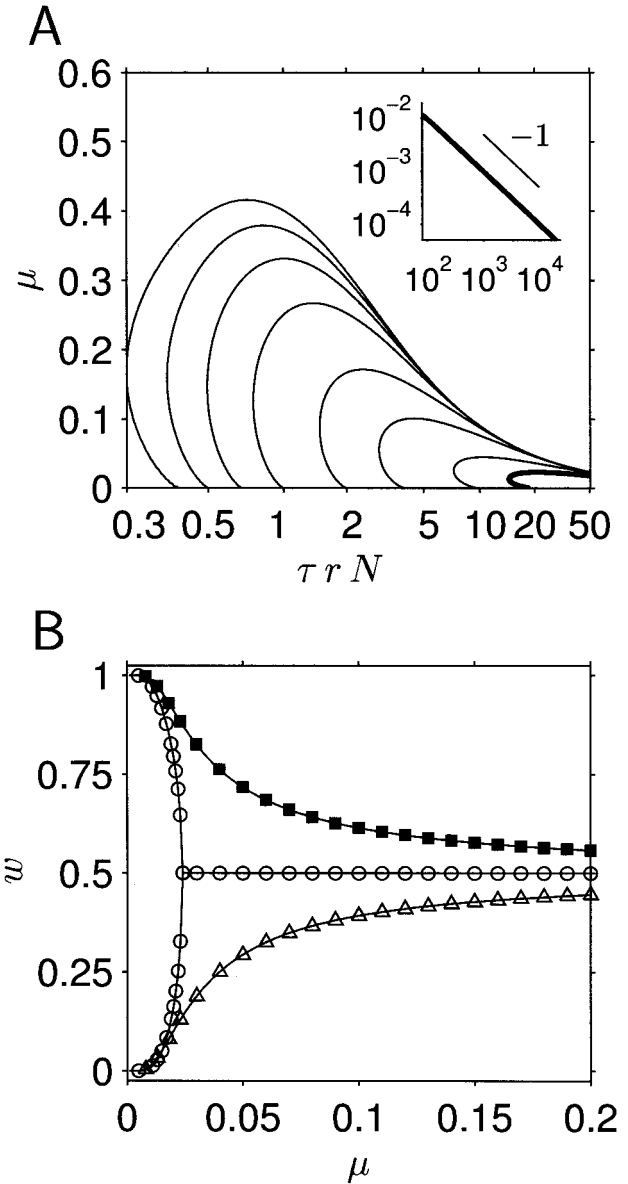
Figure 3. Symmetry breaking in the linear neuron driven by uncorrelated Poisson input processes. *A*, The critical contour lines of the stability criterion $\mathcal{H} = 0$ (Eq. 16) for $\alpha =$ 1.05, 1.1, 1.25, 1.5, 2, 2.5, 3, 3.5 (from right to left, thick line corresponds to $\alpha =$ 1.05). The homogeneous solution is stable outside the contour and unstable in its interior. Inset, Critical contour for large πN and $\alpha =$ 1.05. *B*, Equilibrium synaptic weights as a function of μ for πN = 10 (squares), 20 (circles), and 200 (triangles), with $\alpha =$ 1.05. Plot symbols mark the efficacy values obtained from simulating the mean field learning dynamics of the linear neuron with N = 100 and the input rate r = 5, 10, 100 Hz, adjusted to obtain the desired value of πN. As μ is decreased, each simulation is initialized with the synaptic efficacies of the equilibrium state obtained for the previous value of μ, plus a perturbation vector ranging from −0.001 to 0.001. For small and large values of πN (squares and triangles), the homogeneous solution is stable for all μ (except the vanishing regime of very small $\mu < \mu_{crit}$ in the large πN case). For πN = 20 (circles), the synaptic population splits into two groups (54 strong and 46 weak synapses) when μ crosses the corresponding critical contour in *A*. The solid lines depict the analytically obtained values of the equilibrium weights. For πN = 20 (circles), the line was obtained by numerically solving the analytical expression for a bimodal synaptic state with the same split ratio that was obtained in the simulations.

positive drifts are held at the upper boundary, whereas synapses with negative drifts saturate at the minimum allowed efficacy. Our treatment of the additive model in the Appendix shows that the numbers of synapses gathering at the upper and lower bound-

aries critically depend on the ratio of depression and potentiation α, as well as on the effective population size τrN. As in the nonlinear TAH learning model, small effective synaptic populations ($\tau rN < 1/(2(\alpha - 1))$) will lead to all synapses saturating at the upper boundary because of the strong positive feedback. However, as τrN increases beyond a critical value, the synaptic population breaks into two groups, one of which remains saturated at the upper boundary while the other, losing the competition, saturates at the lower boundary. The ratio of synapses saturating at the top boundary is $n_{\mathrm{up}} = 1/2\tau rN(\alpha - 1)$ (Appendix). Because this ratio is inversely proportional to the input rate r, the output rate of the postsynaptic neuron becomes independent of the input rate, as shown in Figure 2A.

Uncorrelated inputs: integrate-and-fire neuron
We now turn to the behavior of TAH learning in the integrate-and-fire neuron driven by uncorrelated inputs. Figure 2B shows the output rate of this neuron model versus the input rate for different values of μ. As the figure demonstrates, the output-rate normalization quickly deteriorates as μ departs from the additive model and synaptic changes become dependent on the efficacy of the synapse. Figure 2C demonstrates that the sensitivity of the output rate to the parameter α rapidly diminishes as μ increases. Comparing A and B of Figure 2 shows the qualitative similarity between the output rate responses of the linear Poisson and the integrate-and-fire model neurons. Note that we have not attempted to match the overall scale of the output rates in the two models. The output rate of the linear neuron can be arbitrarily changed by a gain factor without affecting any other results.

Figure 4 displays the histograms of the equilibrium distributions of learned synaptic efficacies as a function of the updating parameter μ. Recovering the behavior of additive (Song et al., 2000) and multiplicative (Rubin et al., 2001) updating models for $\mu = 0$ and $\mu = 1$, respectively, the plot reveals the transition between these models for intermediate values of μ. Specifically, it shows the emergence of symmetry breaking as μ approaches zero.

As expected from the analysis of the linear neuron, we find that in the integrate-and-fire neuron also, the critical value of μ at which the synaptic distribution becomes bimodal decreases as the effective population size τrN increases. Increasing the rate of the input processes from 10 Hz (Fig. 4A) to 40 Hz (Fig. 4B) lowers the first occurrence of a bimodal weight distribution from $\mu_{\mathrm{crit}} = 0.023$ to $\mu_{\mathrm{crit}} = 0.017$. The inset in each panel depicts the equilibrium weight distribution for the intermediate value of $\mu = 0.019$, showing a clearly bimodal distribution for the 10 Hz input (Fig. 4A) and a clearly unimodal distribution for the 40 Hz input (Fig. 4B). Moreover, as expected from the equations describing the homogeneous steady state in the linear neuron (Eqs. 15 and 17), the synaptic efficacy of the homogeneous state at a given μ decreases when the input rate increases.

It is interesting to note the close similarity in the μ dependence of the learned synaptic distributions in the linear and the integrate-and-fire neurons. For example, in both cases, the critical μ for symmetry breaking is close to 0.023 for input rates of 10 Hz [compare Fig. 4A with Fig. 3B (circles)]. This is despite the fact that the two models have very different spike generators and different sizes of synaptic populations. The reason for this similarity is that the input–output correlations in the integrate-and-fire neuron with 1000 synapses turn out to match in magnitude the corresponding correlations of the linear neuron with 100 synapses (data not shown).
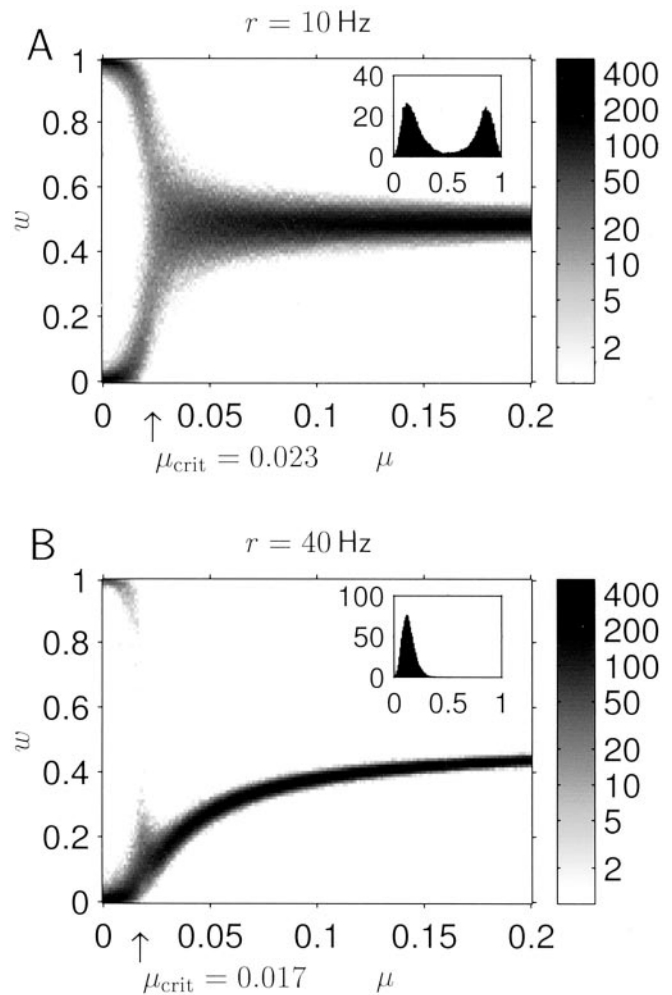

Figure 4. Symmetry breaking in the integrate-and-fire neuron driven by uncorrelated Poisson input processes. *A, B,* Histograms of the equilibrium synaptic distributions in logarithmic gray-scale as a function of the updating parameter μ for N = 1000 and α = 1.05. The input rate is 10 Hz in *A* and 40 Hz in *B*. The arrow marks the critical μ where the first bimodal distribution occurs. Insets, The synaptic distributions at μ = 0.019. Note that whereas the histograms shown in the main figures are single realizations of the equilibrium synaptic distributions for each μ, the histograms shown in the insets were obtained by averaging 30 different readouts of the converged synaptic weights taken at consecutive intervals of 500 sec.

In summary, for uncorrelated inputs and biologically realistic sizes of the presynaptic population, N, on the order of thousands, and for rates on the order of $\geq$10 Hz, the regime in μ and α in which symmetry breaking between uncorrelated inputs as well as output rate normalization occur is extremely narrow. Thus, the learning behavior changes qualitatively as soon as synaptic plasticity becomes weight dependent.

Uniformly correlated inputs
We briefly discuss here the case in which the presynaptic inputs have positive uniform instantaneous correlations, namely that for all $i \neq j$, c_{ij} (Eq. 9) are equal. This situation may, for instance, occur when the entire presynaptic pool of a neuron is driven by a common source. Treating the behavior of the linear Poisson neuron, we show in the Appendix that positive uniform correlation increases the value of the synaptic efficacy in the homogeneous synaptic steady state. Moreover, the uniform correlation does not alter the $1/(\tau rN)$ dependence of the destabilizing drifts. As a result, in nonadditive learning, when the effective synaptic pop-
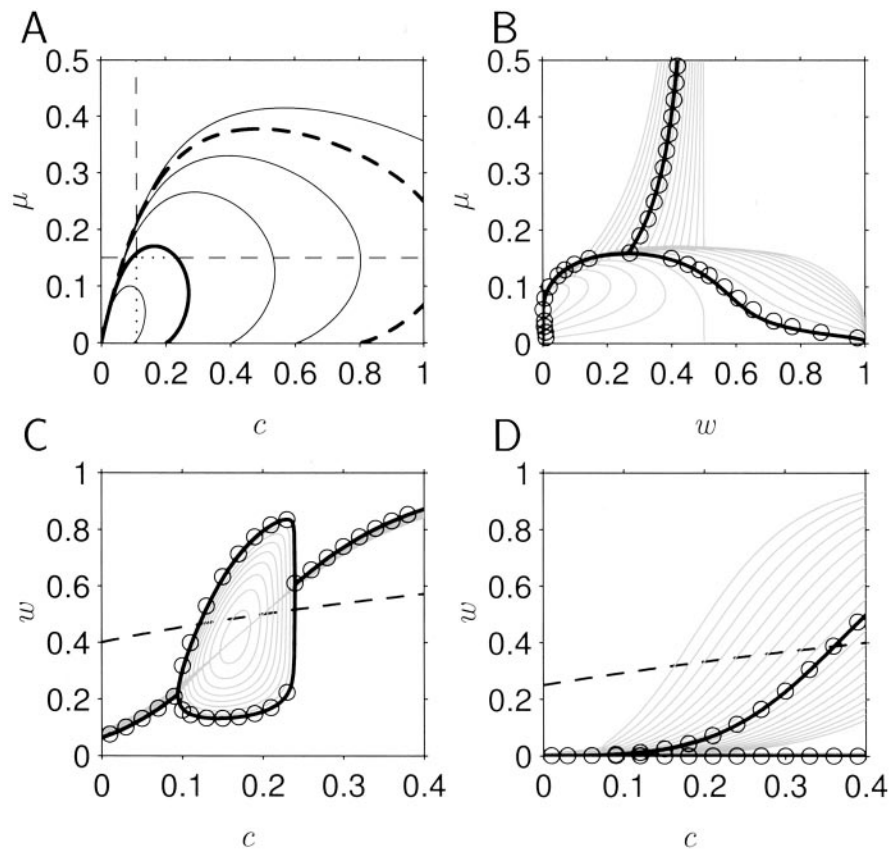
$$C_0 = C_1 = \frac{c}{\tau r M}.$$

Comparing these expressions with their respective values in the case of N uncorrelated inputs (Eq. 17), we note that the learning behavior in both input scenarios is equivalent when N is identified with M/c, the number of correlated subgroups divided by the strength of the within-group correlation. The stability of the homogeneous synaptic steady state in a large network comprising M correlated synaptic subgroups, each with within-group correlation c, behaves as in an uncorrelated network of finite size M/c. Thus, in the limit of a large presynaptic population, the correlation strength c scales the effective number of presynaptic inputs from M for $c = 1$ to infinity for $c = 0$. Importantly, the largest inhomogeneous eigenvector is such that when the homogeneous solution loses stability, the symmetry is broken between the correlated subgroups and not within each subgroup.

The nature of the synaptic pattern that emerges once the homogeneous synaptic state loses stability depends on the number of afferent subgroups. Here we focus on the example of two equally sized subgroups (i.e., $M = 2$). A similar scenario, which is motivated by the problem of the activity-driven development of ocular dominance maps, has recently been studied by Miller and MacKay (1994) and Song and Abbott (2001). The regimes of symmetry breaking in which the learned synaptic efficacies segregate according to the two correlated input groups are depicted in Figure 5A (this figure is equivalent to Fig. 3A, with c replacing $2/N$). Thus, symmetry breaking between two correlated subgroups can occur in nonadditive TAH learning models even when the number of presynaptic inputs N is large. This is demonstrated in Figure 5B (solid black line), which plots the learned synaptic efficacies as μ as varied, with c held fixed at 0.11. As is evident from Figure 5, for this level of correlation, symmetry breaking occurs below a fairly high value of $\mu \approx 0.15$. Note that in contrast to our treatment of the uncorrelated inputs, here we do not use α close to 1 but rather set it to a generic value of $\alpha = 1.5$.

Figure 5, C and D, describes the behavior of the system as the within-group correlation is gradually turned on. As expected from the analysis of the uncorrelated input scenario, the substantial weight dependence of the synaptic changes induced when $\mu = 0.15$ (solid black lines), yields a stable homogeneous synaptic state if the within-group correlation is sufficiently weak. However, when the correlation reaches a critical value, the homogeneous state becomes unstable and the synaptic efficacies segregate into the two input groups, with the one winning the competition suppressing the other. As the correlation increases still further,

Figure 5. Symmetry breaking in the linear neuron driven by two equally sized correlated groups in the large N limit. *A*, The critical contour lines of the stability criterion $\mathcal{H} = 0$ (Eq. 16) for $\alpha = 1.25, 1.5, 2, 2.5, 3, 3.5$ (from left to right, thick solid line corresponds to $\alpha = 1.5$, thick dashed line to $\alpha = 3$) plotted as a function of the within-group correlation c. The homogeneous solution is stable outside a contour line and unstable in its interior. The lines depicted in *B–D* correspond to the equilibrium synaptic states obtained by numerically solving the analytical expressions for the two-group solutions with r = 10 Hz. Circles depict synaptic efficacies obtained by simulating the mean field learning dynamics of N = 1000 synapses. We set N = 1000 here to obtain a better approximation of the analytical results derived for the large N limit. Note that small discrepancies between the analytical curves and the simulated mean field equilibrium, stemming from the finiteness of N in the simulation, can be observed in *B. B*, Equilibrium synaptic weights (black line) as a function of μ for c = 0.11 (horizontal line in *A*) with $\alpha = 1.5$. Gray lines show the equilibrium weights for c = 0.05, 0.06, . . . , 0.2. *C, D*, Equilibrium synaptic weights (black lines) as a function of c for $\mu = 0.15$ (vertical line in *A*) with $\alpha = 1.5$ in *C* and $\alpha = 3$ in *D*. The gray lines in *C* show how the region in which the homogeneous solution is unstable vanishes as μ is increased from 0.15 to 0.1725 in steps of 0.0025. The gray lines in *D* show the change in the equilibrium synaptic weights when α is increased from 2 (largest group separation) to 4 (smallest group separation) in steps of 0.1. The dashed lines in *C* and *D* depict the equilibrium synaptic weights for the multiplicative case ($\mu = 1$).

ulation is sufficiently large, the homogeneous steady state remains stable for any positive uniform correlation strength. In fact, these correlations increase the stability of the homogeneous state (Appendix) and, hence, oppose the emergence of spontaneous symmetry breaking.

Correlated subgroups

We now consider afferent input activity to a neuron that is composed of M equally sized groups. These groups are defined by a uniform within-group correlation coefficient $c_{ij} = c > 0$ (compare Eq. 9) that is equal within all groups. For pairs of inputs belonging to different groups, the cross-correlation is zero. In this scenario, the M different presynaptic groups compete for control over firing of the postsynaptic neuron. We first treat the linear neuron and, for simplicity, focus on the case in which the overall number of presynaptic input channels N is large. In this limit, the homogeneous and largest inhomogeneous eigenvalues of C^+ normalized by N are:
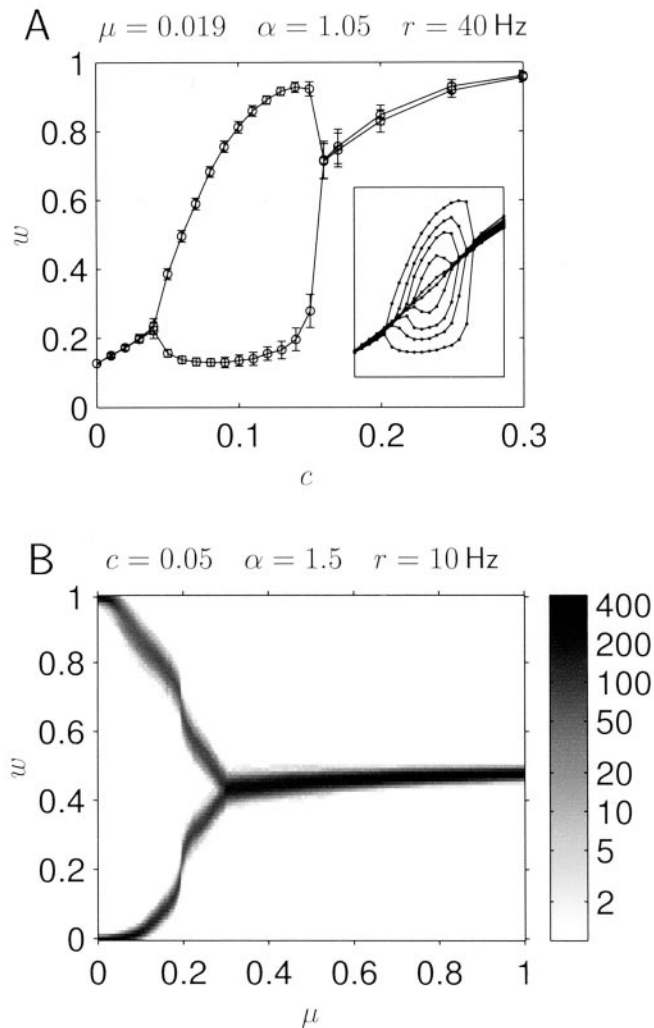
Figure 6. Symmetry breaking in the integrate-and-fire neuron driven by two equally sized correlated groups. Keeping $N = 1000$, each group consists of 500 synapses receiving Poisson inputs. A, The learned synaptic weights of the two input groups as a function of the within-group correlation c, for an input rate of 40 Hz, $\mu = 0.019$ and $\alpha = 1.05$. For each correlation, the group means (circles) are computed for 30 different realizations of the synaptic weight distribution, taken after convergence at successive intervals of 500 sec. Error bars indicate the corresponding SDs. The inset shows how the regime of symmetry breaking vanishes as μ is increased through $\mu = 0.019, 0.021, 0.022, 0.023, 0.024$. B, The learned distributions of synaptic weights as a function of μ for an input rate of 10 Hz, $c = 0.05$ and $\alpha = 1.5$. In all bimodal weight distributions depicted here, the splitting corresponds to the two input groups.
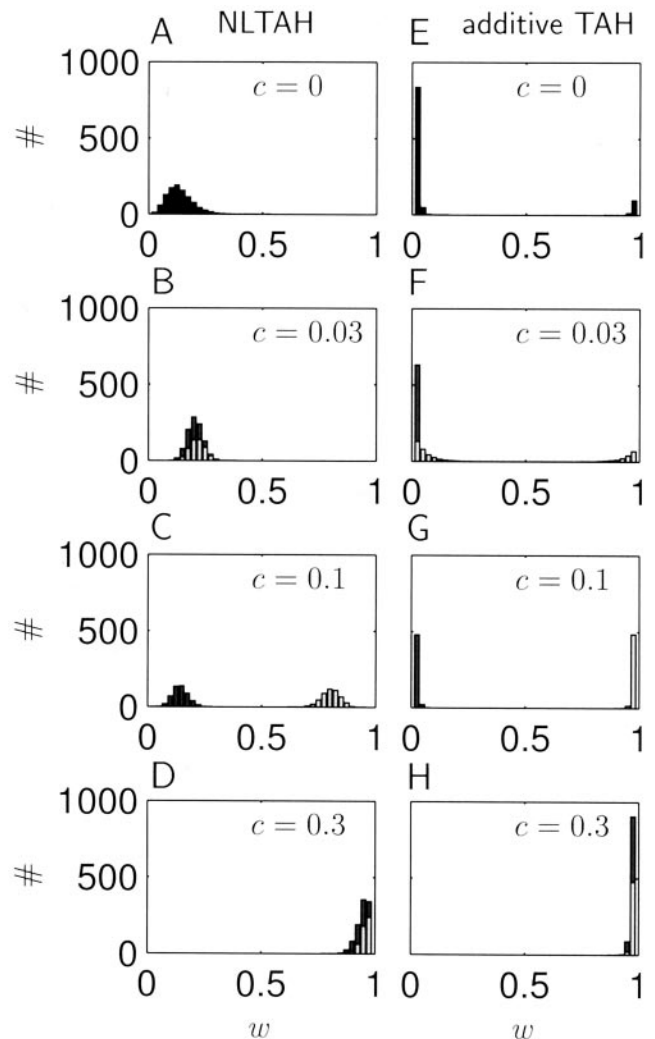
Figure 7. Symmetry breaking in nonlinear and additive TAH learning. The learned synaptic efficacies of the integrate-and-fire neuron driven by two equally sized correlated groups are depicted for selected within-group correlations $c = 0, 0.03, 0.1, 0.3$. Each input group consists of 500 synapses receiving Poisson inputs with a rate of 40 Hz (compare Fig. 6 A). The asymmetry parameter $\alpha = 1.05$. For $c = 0$ (A, E), the histogram of the total synaptic population is depicted. For $c > 0$, the light bars describe the distribution of the presynaptic input group with the higher mean, and the dark bars (stacked on top of the light ones) depict the distribution of the weaker group. All histograms were obtained by averaging 30 different readouts of the converged synaptic weights taken at consecutive intervals of 500 sec. A–D, Results from nonlinear TAH learning with $\mu = 0.019$. E–H, Results from additive TAH learning, $\mu = 0$. In E, a total of 112 synapses are in the upper mode. In F, the stronger group (light bars) has a bimodal distribution with 175 synapses (35%) in the upper mode.

another transition may occur at a higher value of c, above which the homogeneous synaptic state becomes stable again. The presence of this second transition (which is discontinuous) depends on the values of τr, the expected number of input spikes per synapse arriving within its learning time window, and the ratio of depression and potentiation α (Fig. 5, compare C and D). Importantly, for large values of μ, in particular in the multiplicative model ($\mu = 1$), the stabilizing force is so strong that the homogeneous synaptic state remains stable for all positive correlation strengths (Fig. 5C, D, dashed black lines; also see Appendix), and no segregation is possible.

The behavior described above for the linear neuron is reproduced qualitatively in simulations of the integrate-and-fire neuron, as shown in Figure 6A. To address the question of whether symmetry breaking in the integrate-and-fire neuron can also occur at higher values of μ, we follow the linear Poisson neuron

analysis shown in Figure 5A, which suggests that increasing the value of α extends the μ range of bimodal synaptic distributions. Figure 6B displays the learned synaptic distributions as a function of μ for a within-group correlation $c = 0.05$, with $\alpha = 1.5$ and $r = 10$ Hz. Similar to the linear neuron findings shown in Figure 5B, symmetry breaking occurs here in a wide regime of μ.

To emphasize important differences between symmetry breaking in nonlinear versus additive TAH learning, Figure 7 shows corresponding learned synaptic efficacies for selected cases of low, intermediate, and high within-group correlations. Figures 7A–D depicts learned weight distributions from Figure 6A for which $\mu = 0.019$. For each correlation, synaptic efficacies resulting from additive learning are depicted on the right (Fig. 7E–H). Except in Figure 7, A and B, where $c = 0$ (i.e., no input subgroups

are defined), the synaptic distribution of the subgroup with higher mean efficacy is depicted in light gray, whereas that of the subgroup with lower mean efficacy is displayed in dark gray.

Inspection of Figure 7, *A* and *B* versus *E* and *F*, shows that in the regime of low correlations, the learning behavior induced by the two types of plasticity is qualitatively different. While in nonlinear TAH learning (Fig. 7*A*,*B*), the homogeneous synaptic state is stable and all synapses distribute around the same mean efficacy, unstable additive learning induces symmetry breaking (Fig. 7*E*,*F*). Importantly, this symmetry breaking in general does not reflect the correlation structure in the afferent input. As shown in Figure 7*F*, when the within-group correlation is 0.03, the 500 synapses of the group winning the competition (light gray) split into two fractions of 325 versus 175 synapses, of which the larger fraction tends to zero efficacy and mixes with the efficacies of the losing input group. In contrast, in the nonlinear TAH model, unfaithful splitting of the weights occurs only for extremely small values of μ, of the order of $1/\tau r N$ (Fig. 7, compare *A* and *B* with *E* and *F*). This is because symmetry breaking within a uniformly correlated group does not occur for $\mu > 1/\tau r N$, and hence the weights of each subgroup remain the same.

For intermediate strengths of the within-group correlation, both learning rules induce symmetry breaking that faithfully reflects the structure of the input correlation, with the synaptic distributions of the two input groups well separated. This is shown in Figure 7, *C* and *G*, for a correlation of $c = 0.1$. Note, however, that whereas in additive learning the efficacies of both input groups reach the respective boundaries of the allowed range (i.e., are clipped to saturation), the weights resulting from nonlinear TAH learning do not saturate. As we show in the next section, this property of NLTAH plasticity enhances the sensitivity of the synaptic population to changes in the strength of the within-group correlation. Finally, when the within-group correlation is strong, in both types of learning all efficacies become large (Fig. 7*D*,*H*).

Clearly, the detailed quantitative properties of the learned synaptic patterns, as well as the parameter values at which symmetry breaking occurs, depend on the neuron model, and specifically on the spike generating mechanism. Nevertheless, the striking qualitative similarity in the findings from both neuron models investigated here suggests that the symmetry breaking induced by the within-group correlations is a general property of the nonlinear TAH rule with small but non-zero μ, independent of the specifics of the spike generator.

Synaptic representation of input correlations

In the previous section, we studied the emergence of symmetry breaking in homogeneous synaptic populations for different types of instantaneously correlated input activity. In this section, we study the more general issue of how information about the spatiotemporal structure of the afferent input is imprinted into the learned synaptic efficacies by TAH plasticity. Specifically, we investigate how the weight dependence of the synaptic changes affects the sensitivity of the learning to features embedded in the input spike trains.

An example of the associated phenomena is shown in Figure 8. Here we study the effect of weight dependence on the steady-state synaptic efficacies of the integrate-and-fire neuron receiving 1000 Poisson inputs that comprise a small subgroup of 50 correlated synapses ($c = 0.1$) while all other input cross-correlations are zero. In this scenario, the subgroup is statistically distinct from the rest of the synaptic population. The coherence of spikes within the subgroup increases the causal correlation of the mem-
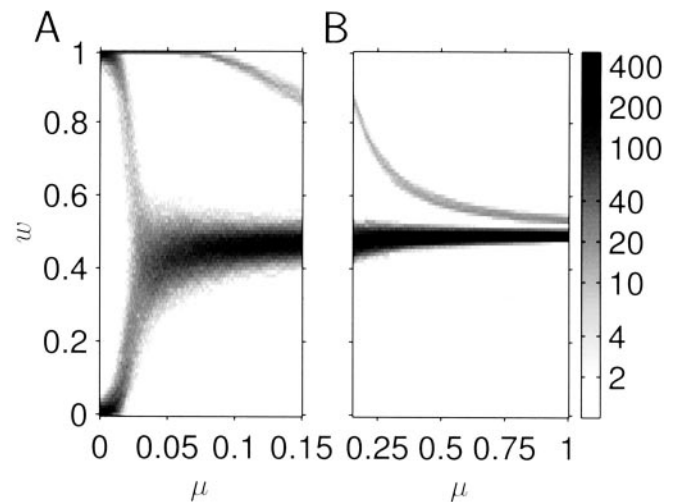

Figure 8. Effect of the updating parameter μ on the learned synaptic distribution of the integrate-and-fire neuron driven by a small correlated group and an uncorrelated background. The neuron is driven by 950 uncorrelated and 50 weakly correlated ($c = 0.1$) Poisson input processes with a rate of 10 Hz ($\alpha = 1.05$). Together, *A* and *B* cover the range of $\mu \in [0, 1]$, but note the difference in resolution.

ber synapses with the spiking activity of the postsynaptic neuron. Because of the ensuing cooperation between the correlated synapses, they grow stronger than those of the uncorrelated background. Figure 8 shows how the strength of the stabilizing drift induced by the weight dependence of the synaptic updating modulates the degree of separation between the two subpopulations. For decreasing values of μ, learning becomes increasingly affected by the correlation structure in the input, and the separation between the subgroup and the background is more pronounced. However, below a critical μ, the homogeneous state of the uncorrelated population loses stability and splits, resulting in a bimodal distribution of the background synapses. As a consequence, the representation of the afferent correlation structure in associated groups of synaptic efficacies is confounded by the mixing of the high-efficacy mode of the background with the subgroup of correlated synapses. This example raises the general problem of finding an optimal learning rule that, for a given type of input activity, compromises best between sensitivity and stability.

To address this question, we need a quantitative measure for the performance of a learning rule in imprinting information about the input correlations onto the synaptic efficacies. Here we apply the sensitivity measure S (Eq. 13, Materials and Methods), which quantifies the sensitivity of the learned synaptic state to changes in features embedded in the input correlation structure. When S is high, small changes in the input features are picked up by learning and induce a large change in the learned synaptic efficacies. We emphasize that the goal of this performance measure is to quantify and compare general properties of different plasticity rules. It is therefore based only on the relationship between the afferent neuronal inputs and the learned synaptic efficacies. In particular, it avoids direct reference to the neuronal output activity.

We first illustrate the application of the sensitivity measure by considering a simple example in which the input feature to be represented by the learned synaptic efficacies is only one dimensional (i.e., a scalar quantity). Specifically, we apply S to the scenario discussed in the previous section, of two independent input groups with within-group correlation c. We investigate the be-
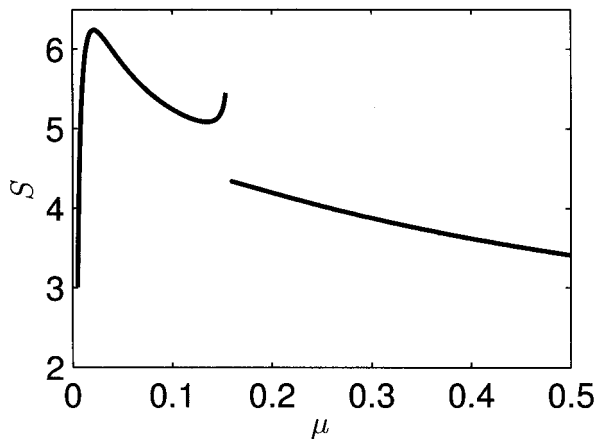
Figure 9. Sensitivity to the within-group correlation strength c in the linear neuron receiving input from two correlated subgroups. The sensitivity is plotted as a function of μ for correlation strength c = 0.11 and α = 1.5 (see Fig. 5B for parameter settings, large N limit).

havior of the linear Poisson neuron and quantify how the sensitivity of the learned synaptic distribution to the strength of the within-group correlation is affected by the weight dependence of the synaptic changes. We consider the sensitivity of the learning as a function of μ for a fixed correlation of c = 0.11. As shown in Figure 5, B and C, this correlation represents an intermediate correlation strength in the linear Poisson neuron treatment. Using the steady-state synaptic efficacies from Figure 5B, we compute S for values of μ between 0 and 0.5 (see Materials and Methods, Appendix). Figure 9 shows the resulting sensitivity curve. We note that each point quantifies the sensitivity of the learned synaptic weights to small changes in the correlation strength around c = 0.11.

As can be seen in Figure 5B, there are two qualitatively distinct regimes of synaptic distributions emerging from learning in this case. For high values of μ, no symmetry breaking takes place, and the correlation strength is represented by the common mean value of the synaptic efficacies. In this regime ($\mu \gtrsim 0.15$), S decreases monotonically with increasing μ (Fig. 9), because the higher weight dependence strengthens the confinement of the homogeneous synaptic state to the center range of the synaptic efficacies. For lower values of μ, symmetry breaking occurs, and the correlation strength is represented by the mean efficacy values of the two resulting groups. In this regime, S is nonmonotonous in μ. For very low μ, the synaptic efficacies are close to saturation at the boundaries and, hence, a change in the correlation strength cannot induce a large change in the efficacies. However, the centralizing drift induced by a large μ reduces the sensitivity. Thus, S has a maximum at an intermediate μ (in the present case around $\mu = 0.02$). Finally, at the transition between the regions of homogeneous and bimodal synaptic distributions ($\mu \approx 0.15$), sensitivity is large, because here a small change in c may cause an abrupt and large change in the synaptic efficacies, namely a bifurcation from a homogeneous to an inhomogeneous synaptic distribution. Note, however, that this transition region in μ is narrow.

We now turn to a richer input scenario in which the afferent correlation structure is inhomogeneous and the input feature space to be represented by the learned synaptic efficacies is high-dimensional. Specifically, we consider presynaptic activity in which each synapse receives spike inputs with a specific relative latency with respect to the remaining synaptic population. Such latency or delay-line scenarios have been studied previously in

the context of additive TAH learning (Gerstner et al., 1996; Song et al., 2000) and can, for instance, be motivated by their analogy to certain delay-line models in auditory processing (Jeffress, 1948).

We consider the input activity to consist of N time-shifted versions of one common Poisson spike train with rate r. Because the synaptic learning process depends on the relative timing of the input spikes, we fix one presynaptic input as reference, and treat the remaining N − 1 delays $\Delta = (\Delta_1, \ldots, \Delta_{N-1})$ as R = N − 1 dimensional vector of input features to be represented by the learned synaptic weights. Whereas the delays Δ fully specify the temporal correlation structure of the neuronal input activity, S measures the sensitivity of the learned synaptic efficacies to small independent changes in the individual delays. Because of the temporal sensitivity of TAH plasticity, it is intuitively clear that the learning dynamics will critically depend on the temporal scale of the relative delays. Although it is a natural choice to set this temporal scale through the SD of a Gaussian distribution from which the delays are drawn (Song et al., 2000; Aharonov et al., 2001; Gütig et al., 2001), we here apply the sensitivity measure to the simpler case in which we fix Δ such that the delays between the N inputs are uniformly spaced at a fixed delay $\sigma/(N-1)$ [i.e., $\Delta_i = i\sigma/(N-1)$ (i = 1, 2, ..., N − 1)]. We have checked that the qualitative behavior of S in the case of a fixed delay spacing is similar to that of S_{avg} (see Materials and Methods) obtained from averaging over an ensemble of Gaussian delay vectors with SD σ (Aharonov et al., 2001; Gütig et al., 2001).

We investigate here the behavior of the linear Poisson neuron. One important difference between the delay-line input scenario considered here and the input correlations treated above is that here non-zero cross-correlations between input spike trains also exist at negative time lags. Specifically, if the delays of the input activities of synapses i and j are given by Δ_i and Δ_j, respectively, and the additional delay of the postsynaptic neuron is ϵ (Eq. 5), the delay difference $\Delta_i - (\Delta_j + \epsilon)$ determines the temporal position of the sharp peak in the otherwise zero effective correlation between the two shifted Poisson inputs (Eq. 7). If this delay difference is negative, the output activity contributed by the jth synapse lags behind the input spikes at the ith synapse. Hence, the jth synapse contributes to the potentiation of synapse i, and the respective effective causal correlation C_{ij}^+ is positive. Correspondingly, in this case the backward effective correlation C_{ij}^- contributed by synapse j to the depression of synapse i is zero. Conversely, if the delay difference between the ith and jth input spike trains is positive, synapse i is depressed by the activity of synapse j, because C_{ij}^- becomes positive. In both cases, the magnitude of the effective correlation is scaled by the exponentially decaying time dependence of the learning rule (Eq. 1). The full expressions for the effective correlation matrices C_{ij}^+ and C_{ij}^- are given in the Appendix.

To calculate S for a given delay vector Δ, we numerically solve the drift equation of the synaptic learning (Eq. 8) for the synaptic steady state. Using the resulting learned synaptic distributions, we compute the susceptibility matrix χ (Eq. 12, Appendix), giving S (Eq. 13). Figure 10A shows the sensitivity S as a function of μ for different values of the temporal delay spacing σ. The curves clearly show an optimal weight dependence of the synaptic changes for which the sensitivity peaks. For larger values of μ, the performance of the learning deteriorates because the increasing confinement of the synaptic weights to the central range of efficacies restricts the sensitivity of the learning to changes in the input correlation structure. Conversely, for lower values of μ, the sensitivity is impaired because the synaptic efficacies are begin-
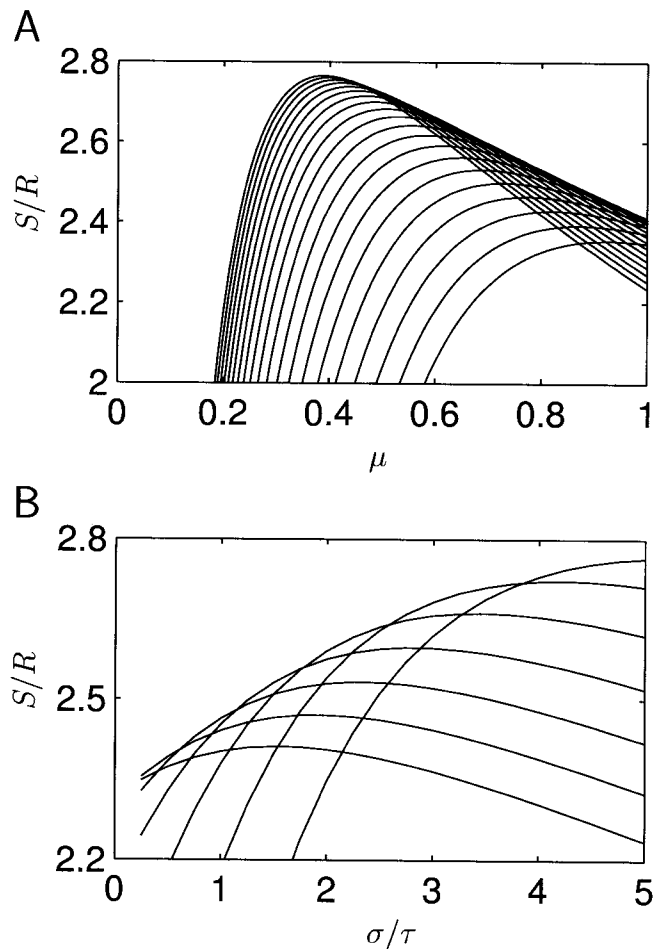
Figure 10. Sensitivity in the linear neuron evaluated at uniformly spaced input delays. The delay vector Δ covers the range from 0 to σ (see Results). The learning dynamics are simulated for the linear neuron, with N = 101 synapses driven by the delayed Poisson inputs with a rate of 10 Hz and α = 1.5. A, The sensitivity per input feature S/R as a function of μ for σ/τ = 0.25, 0.5, 0.75, . . . , 4.75, 5 (from bottom to top). The termination of the curves at low μ is a result of poor numerical convergence arising as the synaptic efficacies come close to the boundaries of the allowed range. B, The sensitivity per feature S/R as a function of the delay interval σ/τ for μ = 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 (from top to bottom).

ning to saturate at the boundaries of the allowed range as bimodal efficacy distributions emerge. The value of μ that optimally adjusts the weight dependence of the synaptic changes depends both on the system parameters and on the input correlations determined by the relative time delays between the inputs. Increasing σ (i.e., increasing the relative delays) weakens the effective correlations between the presynaptic inputs because of the exponentially decaying temporal extent of the learning rule (Eq. 7). Hence, a lower weight dependence of the synaptic changes (corresponding to a lower value of μ) is needed to pick up the correlations and allow sufficient sensitivity of the learning to the input delays. The effect of this change in the temporal extent of synaptic interactions on the learned efficacies is shown in Figure 11, which for each μ depicts all N synaptic efficacies for $\sigma = \tau$ (Fig. 11A) and $\sigma = 4\tau$ (Fig. 11B). Note that because of the equidistant delays, the relationship between the relative temporal position of a synapse within the presynaptic population and its steady-state efficacy is monotonic, with the leading synapse (Δ = 0) taking the largest weight. In the foreground, the corresponding sensitivity curves are shown. The plots clearly demonstrate that the saturation regime in which most synaptic weights accumulate
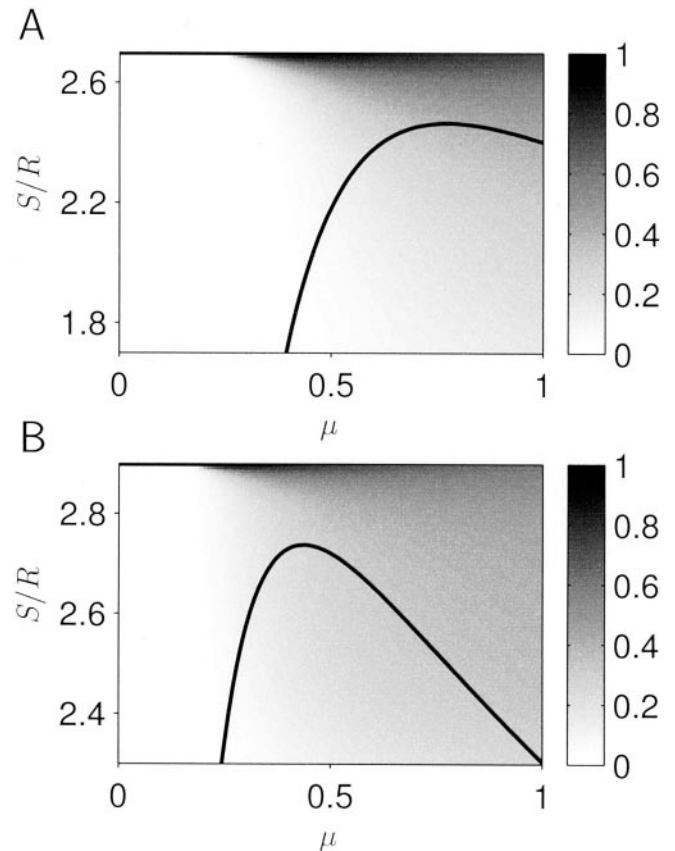

Figure 11. Mapping between synaptic delays and equilibrium weights. The parameters are identical to those used in Figure 10A, with σ/τ = 1 in A and σ/τ = 4 in B. The background depicts the learned synaptic efficacies as a function of μ. Each row i reflects the equilibrium weight w_i in gray-scale, with the top row corresponding to the leading synapse with zero delay and the bottom row to the last synapse with delay σ. The solid lines depict the sensitivity curves from Figure 10A.

at the boundaries of the allowed range (black and white) begins at higher values of μ when the temporal dispersion of the inputs is small (Fig. 11A) (i.e., the synaptic interactions are strong). The plot also reveals that in both cases for low values of μ, only the leading synapse remains at a high value. Finally, it can be seen that the peaks in the sensitivity curves approximately coincide with those values of μ for which the synaptic weights smoothly cover a large range of efficacies, as shown by the gradual change from dark to light values in the corresponding vertical cross sections.

Finally, we ask how the learning sensitivity depends on the statistics of the input delays for a fixed value of μ. To answer this question, Figure 10B shows S as a function of the delay-line spacing σ, demonstrating that S does not vary monotonically with σ, but rather has a maximum at an optimal temporal separation of the inputs. This is because tight spacing leads to strong effective correlations between the inputs, driving the synapses toward saturation. On the other hand, loose spacing reduces the effective correlations between the presynaptic inputs to the extent that the learning behaves essentially as if driven by an uncorrelated presynaptic population.

Discussion

The understanding of activity-dependent refinement of neural networks has long been one of the central interests of synaptic learning studies. In this context, most investigations of unsupervised learning using correlation-based plasticity rules have been

conducted in the framework of additive plasticity models, which do not incorporate explicit weight dependence in the changes of synaptic efficacies. These simple models suffer from stability problems: either all synapses decay to zero or they grow without bound. An additional problem inherent to simple Hebbian models is the lack of robust competition. Indeed, it has been found that even the inclusion of synaptic depression mechanisms does not provide a robust source of synaptic competition, unless synaptic plasticity is fine-tuned to approximately balance the amount of potentiation and depression (Miller, 1996).

Recent studies of experimentally observed temporally asymmetric Hebbian learning rules have added two new ideas. One idea is that under these plasticity mechanisms, synapses compete against each other in controlling the time of firing of the target cell and, thus, engage in competition in the time domain. Although TAH learning rules are indeed inherently sensitive to temporal correlations between the afferent inputs, we have shown here that this sensitivity alone is not sufficient to resolve the problems associated with either stability or competition. In the additive model of TAH plasticity, hard constraints need to be imposed on the maximal and minimal synaptic efficacies to prevent the runaway of synaptic strength. In addition, as was shown here, in this model synaptic learning is competitive only when the ratio between depression and potentiation is fine-tuned, and even then the emergent synaptic patterns do not necessarily segregate the synaptic population according to the correlation structure in the neuronal input. The second idea is that TAH rules would exhibit novel behavior because of the role of the nonlinear spike-generation mechanism of the postsynaptic cell (Song et al., 2000). In fact, we have shown in this work that the qualitative features of TAH plasticity are strikingly insensitive to the nonlinear integration of inputs in the target cell (see also Kempter et al., 2001). For the parameter choices studied, the properties of the synaptic steady states in the integrate-and-fire neuron are qualitatively similar to those found in a linear input–output model for neuronal firing. Nevertheless, we note that there are substantial quantitative differences between the two models, particularly with respect to the parameters $\tau r N$ and c, which effectively determine the correlations between the presynaptic and postsynaptic spike trains. Although a quantitative analysis of these differences is beyond the scope of our work, such a study might reveal interesting insights into the quantitative effects of the details of the postsynaptic spike generator on the learned synaptic distributions. In addition, it is possible that the details of the spike-generation mechanism will affect the transient phase (i.e., the dynamics) of the synaptic learning process.

From the present work, we conclude that some of the underlying difficulties in correlation-based learning are alleviated by nonlinear plasticity rules such as the NLTAH rule. The nonlinear weight dependence of the synaptic changes provides a natural mechanism to prevent runaway of synaptic strength. As in additive TAH learning, synaptic competition is provided by the mixture of depression and potentiation. However, in NLTAH plasticity, the balance between depression and potentiation is maintained dynamically by adjusting the steady-state value of the synaptic efficacies. Indeed, we have shown that this competition is sufficient to generate symmetry breaking between two independent groups of correlated presynaptic inputs. However, for this to occur, the stabilizing drift induced by the weight dependence of the synaptic changes should not be too strong. In particular, the simple linear weight dependence (Eq. 2, with $\mu = 1$) assumed in the original multiplicative model is incapable of breaking the symmetry between competing input groups. In fact,

we have shown that with $\mu = 1$, the homogeneous synaptic state is stable for any pattern of homogeneous input correlations, provided there are no negative correlations in the afferent activity. The present power-law plasticity rule with $0 < \mu < 1$ provides a reasonable balance between the need for a stabilizing force and a potential for spontaneous emergence of synaptic patterns. Our study of symmetry breaking between two competing groups of correlated synapses is inspired by the activity-dependent development of ocular dominance selectivity. This scenario has also been studied recently by Song and Abbott (2001) using the additive version of TAH plasticity. In their model, achieving a faithful splitting between the two competing input groups with weak correlations requires relatively tight tuning of the depression to potentiation ratio, α.

One of the surprising results of our investigation is the possibility that when the correlation within input groups is made strong, the stability of the homogeneous synaptic state may be restored. We have shown that this apparently counterintuitive behavior, predicted by the analytical study of the mean synaptic dynamics of the linear Poisson neuron, is also seen in simulations of the full learning rule in the integrate-and-fire neuron. It would be interesting to explore possible experimental testing of this result, perhaps in the context of the development of ocular dominance. In this work, we have limited ourselves to correlated subpopulations of inputs with positive within-group correlations. However, in general, negative correlations are an additional potential source of competition (Miller and MacKay, 1994). Furthermore, we have not addressed the important issue of competition between synapses that target different cells. Lateral inhibitory connections between target neurons may provide a source of such competition.

The last part of the present work addresses situations with inhomogeneous input statistics. Different inputs are distinct in their temporal relationship to the rest of the input population. Here the issue is not whether a spatially modulated pattern of synaptic efficacies will form through TAH learning, but rather whether this pattern will efficiently imprint the information embedded in the input statistics. To quantify the imprinting efficiency of the learning rule, we introduced a new method for measuring learning rule sensitivity. In the present context, this measure quantifies the amount of information about the temporal structure in the inputs that a TAH rule can store. Using this method to study the novel class of NLTAH plasticity rules introduced here, we find that the optimal learning rule depends on the input statistics, in the present example on the characteristic time scale of the temporal correlations between the inputs. This finding suggests that biological systems may have acquired mechanisms for metaplasticity to adapt the learning rule to slow temporal changes in the input statistics. It should be pointed out that the sensitivity measure S focuses entirely on how the learned synaptic distribution changes as a result of changes in the correlation pattern among the input channels. It does not, however, address the problem of "readout," namely how the resulting changes in the synaptic distribution affect the firing pattern of the output cell. A measure that takes the postsynaptic spike train into account will in general depend on the details of the spike-generating mechanism rather than only capture the properties of the learning rule. In general, however, any readout mechanism will depend on the information that is available in the learned synaptic state. Hence, if the learning itself is insensitive to changes in the input features, the synaptic efficacies will fail to represent these changes and no readout mechanism will be able to extract

them. The sensitivity measure S therefore provides an upper bound on the learning performance of the full neural system (including readout). In summary, while quantitative claims about the optimality of specific learning rules have to consider specific readout mechanisms, our study of the general properties of the investigated plasticity rules provide general insights into the mechanisms that enable unsupervised synaptic learning to remain sensitive to input features during learning.

Present experimental results (Bi and Poo, 1998) based on the averaging of individual efficacy changes in different synapses suggest the possibility that indeed the ratio of depressing and potentiating synaptic changes increases in a stabilizing manner as synapses grow stronger (cf. van Rossum et al., 2000). However, available data do not provide conclusive evidence regarding the details of the weight dependence of the efficacy changes. Our work clearly demonstrates the importance of the weight dependence of the TAH updating rule. Synaptic learning rules that implement a stabilizing weight dependence of the type introduced in this work have several advantageous properties for the learning in neural networks. Specifically, our results predict that synaptic changes should be neither additive nor multiplicative, but rather should feature intermediate weight dependencies that could, for instance, result from a gradual saturation of the potentiating and depressing mechanisms. It will be interesting to see whether future experimental results will confirm such a prediction. In this context, it is also important to note that recent experiments and modeling studies reveal important nonlinearities in the accumulation of synaptic changes induced by different spike pairs (Castellani et al., 2001; Senn et al., 2001; Sjöström et al., 2001) as well as evidence for complex intrinsic synaptic dynamics that challenges the simple notion of a scalar synaptic efficacy (Markram and Tsodyks, 1996). The theoretical implications of these sources of nonlinearity and intrinsic dynamics remain to be explored.

Appendix

Generating correlated spike trains

We show here that two spike trains that are generated by conditioning their binwise spike probabilities on the activity of a common reference spike train $X_0(T)$ as described in Materials and Methods, have a pairwise correlation coefficient c. For clarity, we denote $X_i(T)$ simply by X_i. The pairwise correlation coefficient is defined by $\text{Cov}(X_i, X_j)/\sqrt{\text{Var}(X_i)\text{Var}(X_j)}$, where the covariance is $\text{Cov}(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$. Because X_i is either 0 or 1, $E[X_i] = P(X_i = 1)$, and therefore:

$$E[X_i] = P(X_i = 1)$$

$$= P(X_0 = 1)P(X_i = 1 | X_0 = 1) + P(X_0 = 0)P(X_i = 1 | X_0 = 0)$$

$$= r\Delta T \vartheta + (1 - r\Delta T)\varphi$$

$$= r\Delta T,$$

where in the final step we use Equation 11. Note that because $E[X_i] = r\Delta T$, the spike train X_i has rate r. Similarly:

$$E[X_i X_j] = P(X_i = 1, X_j = 1)$$

$$= P(X_0 = 1)P(X_i = 1, X_j = 1 | X_0 = 1) + P(X_0 = 0)P(X_i$$

$$= 1, X_j = 1 | X_0 = 0)$$

$$= r\Delta T \vartheta^2 + (1 - r\Delta T)\varphi^2$$

$$= (r\Delta T)^2 + c(r\Delta T)(1 - r\Delta T).$$

Finally, because $\text{Var}(X_i) = (r\Delta T)(1 - r\Delta T)$, the binwise correlation coefficient becomes:

$$\frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)\text{Var}(X_j)}} = \frac{(r\Delta T)^2 + c(r\Delta T)(1 - r\Delta T) - (r\Delta T)^2}{(r\Delta T)(1 - r\Delta T)} = c.$$

Homogeneous synaptic steady state for a homogeneous population of synapses

We derive the homogeneous synaptic steady-state solution by setting $\dot{w}_i = 0$ and $w_i = w^\star$ in Equation 8 with $C_{ij}^- = 0$:

$$\frac{\lambda \tau r^2}{N}\left[f_+(w^\star) \sum_{j=1}^{N} C_{ij}^+ w^\star - \Delta f(w^\star) \sum_{j=1}^{N} w^\star \right] = 0.$$

Using Equation 14, this yields:

$$f_+(w^\star)w^\star N C_0 - \Delta f(w^\star)w^\star N = 0,$$

which, discarding the trivial solution in which all synapses are zero, implies that at the homogeneous steady state:

$$f_+(w^\star)C_0 - \Delta f(w^\star) = f_+(w^\star)(1 + C_0) - f_-(w^\star) = 0.$$

$$(18)$$

From this, we find that the homogeneous solution is given by $w_i = w^\star$, where $w^\star$ is the solution to Equation 15. Because C_0 is positive, this equation has a unique solution $0 < w^\star < 1$ for any $\mu > 0$. Note, however, that in the additive model where $\mu = 0$, the ratio $f_-(w)/f_+(w) = \alpha$ and, hence, in general there is no homogeneous synaptic steady state unless $w^\star = 0$ or all weights are clipped to the upper boundary.

For uncorrelated input activity $C_{ij}^+ = \delta_{ij}/(\tau r)$ and hence $C_0 = 1/(\tau r N)$. Substituting C_0 into Equation 15, we obtain the homogeneous solution for this input scenario:

$$w^\star = \frac{1}{1 + \alpha^{1/\mu}\left(1 - \dfrac{1}{1 + \tau r N} \right)^{1/\mu}}.$$

$$(19)$$

Stability of the homogeneous synaptic steady state

We analyze the stability of the homogeneous synaptic steady state by deriving the time evolution of small perturbations $\delta w_i = w_i - w^\star$ of the synaptic efficacies w_i around the steady-state value $w^\star$. If these perturbations decay to zero, the homogeneous steady state is stable. For small perturbations, the time evolution is given by:

$$\dot{\delta w_i} = \sum_{j=1}^{N} \frac{\partial \dot{w}_i}{\partial w_j}\delta w_j.$$

Using the expression for the synaptic drifts from Equation 8, we obtain:

$$\dot{\delta w_i} = \frac{\lambda \tau r^2}{N}\left[-Ng_o\delta w_i - \Delta f(w^\star) \sum_{j=1}^{N} \delta w_j + f_+(w^\star) \sum_{j=1}^{N} C_{ij}^+ \delta w_j \right]$$

where:

$$g_o = w \left[\frac{d}{dw} \Delta f(w) - C_0 \frac{d}{dw} f_+(w) \right] \Bigg|_{w=w^*}$$

$$= w \left[\frac{d}{dw} f_-(w) - (1 + C_0) \frac{d}{dw} f_+(w) \right] \Bigg|_{w=w^*}$$

$$= \frac{w}{f_+(w)} \left[f_+(w) \frac{d}{dw} f_-(w) - f_-(w) \frac{d}{dw} f_+(w) \right] \Bigg|_{w=w^*}$$

$$= w f_+(w) \frac{d}{dw} \left[\frac{f_-(w)}{f_+(w)} \right] \Bigg|_{w=w^*} . \tag{20}$$

In the second step, we used Equation 15 to substitute for $1 + C_0$. Note that the factor g_o is proportional to the derivative of the ratio between the scales of negative and positive synaptic changes with respect to the weights. Hence, $g_o = 0$ in the additive model. For $\mu > 0$:

$$g_o = w(1 - w)^\mu \frac{d}{dw} \left[\alpha \left(\frac{w}{1 - w} \right)^\mu \right] \Bigg|_{w=w^*}$$

$$= \alpha \mu \frac{w^{*\mu}}{1 - w^*} \tag{21}$$

is positive because $0 < w^* < 1$. In matrix notation, the time evolution of a synaptic perturbation δw can be rewritten as:

$$\dot{\delta w} = \frac{\lambda \tau r^2}{N} J \delta w$$

with the matrix:

$$J_{ij} = -\delta_{ij} N g_o - \Delta f(w^*) + f_+(w^*) C_{ij}^+ .$$

If the eigenvalues of J are negative, all perturbations of the homogeneous state are attenuated by the learning dynamics and, hence, this synaptic state is stable. In contrast, if any eigenvalue of J is positive, a perturbation along the direction of the corresponding eigenvector will grow exponentially. The matrix J has a homogeneous eigenvector with eigenvalue $-N g_o - N \Delta f(w^*) + N C_0 f_+(w^*)$, which using Equation 18 reduces to $-N g_o$. Because this eigenvalue is always negative, the homogeneous component of any perturbation δw decays to zero with rate $\lambda \tau r^2 g_o$. In contrast, the temporal evolution of the strictly inhomogeneous component of δw (whose elements sum to zero) comprises a spectrum of rates that are determined by the various eigenvalues of J that correspond to inhomogeneous eigenvectors. The largest inhomogeneous eigenvalue of J is $f_+(w^*) N C_1 - N g_o$, where $N C_1$ denotes the largest inhomogeneous eigenvalue of C^+. Hence, the homogeneous synaptic state is stable if $f_+(w^*) N C_1 - N g_o < 0$, which gives the stability criterion stated in Results (Eq. 16). Inserting Equations 15 and 21 into the criterion 16, we obtain an upper bound for μ_{crit}, the largest value of μ for which the homogeneous solution is unstable:

$$\mu_{\text{crit}} < \frac{C_1}{1 + C_0} . \tag{22}$$

An important observation is that $C_1 \leq C_0$, and hence this bound is necessarily smaller than 1, implying that the homogeneous solution is always stable in the multiplicative model where $\mu = 1$. To see that $C_1 \leq C_0$, recall that because $N C_1$ is an eigenvalue of C^+, there is an eigenvector v such that $C^+ v = N C_1 v$. Specifically, for v_m, the largest component of v, this implies that $N C_1 v_m = \Sigma_{j=1}^N C_{mj}^+ v_j$, and hence, $N C_1 = \Sigma_{j=1}^N C_{mj}^+ (v_j/v_m) \leq \Sigma_{j=1}^N C_{mj}^+$. But because $\Sigma_{j=1}^N C_{mj}^+ = N C_0$ (Eq. 14), this yields $C_1 \leq C_0$.

For uncorrelated input activity, the above bound for μ_{crit} becomes:

$$\mu_{\text{crit}} < \frac{1}{1 + \tau r N} , \tag{23}$$

where we used C_0 and C_1 from Equation 17. Hence, for large $\tau r N$, the regime of μ for which symmetry breaking exists vanishes at least with $1/(\tau r N)$.

Additive TAH in the linear neuron: uncorrelated inputs

The drift of the ith synapse of a neuron receiving uncorrelated inputs and implementing the additive model is given by setting $\mu = 0$ in Equation 8 with $C_{ii}^+ = 1/(\tau r)$ and all other effective correlations equal 0:

$$\dot{w}_i = \lambda \tau r^2 (1 - \alpha) \left(\frac{1}{N} \sum_{j=1}^N w_j \right) + \frac{\lambda r w_i}{N} . \tag{24}$$

As explained above, this linear system has no steady state. Imposing the boundary conditions by clipping the efficacies results in all synapses taking the value of either 0 or 1. Thus, the learned synaptic distribution is fully described by n_{up}, the ratio of the synapses that are saturated at the upper boundary. For a ratio n_{up} to be consistent, the drift of a synapse with efficacy 0 must be nonpositive, whereas the drift of a synapse with efficacy 1 must be nonnegative. From imposing these conditions in Equation 24 we get:

$$\tau r^2 (1 - \alpha) n_{\text{up}} \leq 0 \text{ and}$$

$$\tau r^2 (1 - \alpha) n_{\text{up}} + \frac{r}{N} \geq 0 .$$

The first inequality implies $\alpha \geq 1$ if $n_{\text{up}} > 0$. It is important to note that the regime of $\alpha < 1$ would simply yield saturation of all efficacies at the upper boundary because all synapses experience a positive drift. The second condition yields:

$$n_{\text{up}} \leq \frac{1}{\tau r N(\alpha - 1)} .$$

However, it can be shown using methods similar to those of Rubin et al. (2001) that:

$$n_{\text{up}} = \frac{1}{2 \tau r N(\alpha - 1)} , \tag{25}$$

where if this quantity is >1, $n_{\text{up}} = 1$. Therefore, if $\alpha \leq 1 + 1/(2 \tau r N)$, all synapses will saturate at the upper boundary, whereas if $\alpha > 1 + 1/(2 \tau r)$, even a single synapse at the upper boundary will experience a negative drift, and hence no synapse will saturate at 1.

Moreover, the firing rate of the linear neuron is given by:

$$r^{\text{post}} = n_{\text{up}} r = \frac{1}{2 \tau N(\alpha - 1)} ,$$

which is independent of the input firing rate r [except for very low rates, where all synapses become strong (i.e., $n_{\text{up}} = 1$ and $r^{\text{post}} = r$)]. Thus, output rate normalization is a property of the linear neuron when the additive model is used.

Uniformly correlated inputs in the linear neuron

When the presynaptic inputs are uniformly correlated, namely $c_{ij} = c \geq 0$ for all $i \neq j$ ($c_{ii} = 1$), the effective correlation matrix $C_{ij}^+ = [c + \delta_{ij}(1 - c)]/(\tau r)$, and hence:

$$C_0 = \frac{1 + c(N - 1)}{\tau r N} \text{ and } C_1 = \frac{1 - c}{\tau r N}. \quad (26)$$

Because C_0 increases with c, the correlation increases the value of the synaptic efficacy in the homogeneous synaptic state $w^\star$ (Eq. 15). To see the effect of the correlation on the stability of the homogeneous solution, note that because C_1 decreases with c, the correlation decreases the value of $\mathcal{H}$ (Eq. 16), and hence increases the stability of the homogeneous solution. Moreover, because both C_0 and C_1 decrease with $1/(\tau r N)$, for large $\tau r N$ the critical μ (Eq. 22) approaches zero. Thus, for any $\mu > 0$, the homogeneous synaptic state is stable when the effective population is sufficiently large.

Computing the susceptibility matrix χ for the linear neuron

Here we compute the susceptibility matrix χ (Eq. 12, Materials and Methods) used in Results to evaluate the sensitivity measure S for the learning process in the linear Poisson neuron model. This matrix is obtained by the implicit function theorem. In the synaptic steady state ($w^\star$), the synaptic drifts are zero by definition, and hence:

$$\frac{d\dot{w}_k}{d\Phi_j}\bigg|_{w^\star} = 0 \ (k = 1 \ldots N, j = 1 \ldots R),$$

where R is the dimension of the space of input features. Using:

$$\frac{d\dot{w}_k}{d\Phi_j}\bigg|_{w^\star} = \frac{\partial \dot{w}_k}{\partial \Phi_j}\bigg|_{w^\star} + \sum_{i=1}^{n} \chi_{ij} \frac{\partial \dot{w}_k}{\partial w_i}\bigg|_{w^\star} = 0 \text{ with } \chi_{ij} = \frac{\partial w_i}{\partial \Phi_j}\bigg|_{w^\star}$$

and denoting:

$$M_{ij} = \frac{\partial \dot{w}_i}{\partial w_j}\bigg|_{w^\star} \quad M_{ij}^o = \frac{\partial \dot{w}_i}{\partial \Phi_j}\bigg|_{w^\star},$$

we obtain:

$$\chi = -M^{-1}M^o.$$

The matrices χ and M^o are of dimensions N by R, and M is an N by N matrix. Below we derive χ for the two input scenarios studied in Results.

Two correlated input groups
For the case of two correlated subgroups, the sensitivity to the within-group correlation is measured. Hence, the input feature is $\Phi = c$ with $R = 1$. Using Equation 8 with $C_{ij}^- = 0$, $C_{ij}^+ = c/(\tau r)$ if $i \neq j$ are in the same subgroup ($C_{ii}^+ = 1/(\tau r)$), and $C_{ij}^+ = 0$ otherwise, we derive the expressions for M and M^o:

$$M_{ii} = \frac{\tau r^2}{N}[(1 - w_i)^\mu - \alpha w_i^\mu] + \frac{rc}{N}(1 - w_i)^\mu + \frac{r}{N}(1 - c)(1 - w_i)^\mu$$

$$- \mu \frac{\tau r^2}{N} \left(\sum_{k=1}^{N} w_k\right)[(1 - w_i)^{\mu-1} + \alpha w_i^{\mu-1}]$$

$$- \mu(1 - w_i)^{\mu-1} \frac{r}{N}\left[(1 - c)w_i + \sum_{k \in \mathcal{G}_i} w_k\right]$$

$$M_{ij} = \frac{\tau r^2}{N}[(1 - w_i)^\mu - \alpha w_i^\mu] + \frac{rc}{N}(1 - w_i)^\mu \quad j \in \mathcal{G}_i$$

$$M_{ij} = \frac{\tau r^2}{N}[(1 - w_i)^\mu - \alpha w_i^\mu] \qquad\qquad j \notin \mathcal{G}_i$$

$$M_i^o = (1 - w_i)^\mu \frac{r}{N}\left[\sum_{k \in \mathcal{G}_i} w_k - w_i\right],$$

where $j \in \mathcal{G}_i$ if synapse j and i are in the same group.

Delayed Poisson inputs
For a neuron receiving time-shifted versions of a common Poisson spike train ρ_0^{pre}, the input processes are $\rho_i^{\text{pre}}(t) = \rho_0^{\text{pre}}(t - \Delta_i)$ ($i = 1, \ldots, N$), where Δ_i is the delay of the activity at synapse i. Substituting these into Equation 6, we obtain the effective correlation matrices (Eq. 7):

$$C_{ij}^\pm = \Theta(\mp(\Delta_i - \Delta_j - \epsilon))(\tau r)^{-1}\exp(\pm(\Delta_i - \Delta_j - \epsilon)/\tau),$$

(27)

where Θ denotes the Heaviside step function ($\Theta(x) = 1$ if $x \geq 0$ and $\Theta(x) = 0$ otherwise). In this case, the input features $\Phi = \Delta$ and $R = N - 1$. Based on Equations 8 and 27, we derive the expressions for M and M^o. For compactness, we define the matrix of relative delays $D_{ij} = (\epsilon + \Delta_j - \Delta_i)$, with $\Delta_N = 0$ corresponding to the delay of the reference spike train. Thus:

$$M_{ii} = \frac{\lambda}{N}\bigg[-\mu(1 - w_i)^{(\mu-1)}\bigg[\tau r^2 \sum_j w_j$$

$$+ r \sum_{j \neq i} w_j \Theta(D_{ij})\exp\left(-\frac{D_{ij}}{\tau}\right) + rw_i\bigg]$$

$$- \alpha\mu w_i^{(\mu-1)}\bigg[\tau r^2 \sum_j w_j$$

$$+ r \sum_{j \neq i} w_j \Theta(-D_{ij})\exp\left(\frac{D_{ij}}{\tau}\right)\bigg]$$

$$+ \tau r^2[(1 - w_i)^\mu - \alpha w_i^\mu] + (1 - w_i)^\mu r\bigg]$$

$$i = 1 \ldots N$$

$$M_{ij} = \frac{\lambda}{N}\bigg[\tau r^2[(1 - w_i)^\mu - \alpha w_i^\mu] - \alpha w_i^\mu r\Theta(-D_{ij})\exp\left(\frac{D_{ij}}{\tau}\right)$$

$$+ (1 - w_i)^\mu r\Theta(D_{ij})\exp\left(-\frac{D_{ij}}{\tau}\right)\bigg] \quad i, j = 1 \ldots N, i \neq j$$

$$M_{ii}^o = \frac{\lambda}{N}\bigg[(1 - w_i)^\mu\bigg[\frac{r}{\tau}\sum_{j \neq i} w_j\Theta(D_{ij})\exp\left(-\frac{D_{ij}}{\tau}\right)\bigg]$$

$$- \alpha w_i^\mu\bigg[-\frac{r}{\tau}\sum_{j \neq i} w_j\Theta(-D_{ij})\exp\left(\frac{D_{ij}}{\tau}\right)\bigg]\bigg] \quad i = 1 \ldots N - 1$$

$$M_{ij}^o = \frac{\lambda}{N}\bigg[(1 - w_i)^\mu\bigg[-\frac{r}{\tau}w_j\Theta(D_{ij})\exp\left(-\frac{D_{ij}}{\tau}\right)\bigg]$$

$$- \alpha w_i^\mu\bigg[\frac{r}{\tau}w_j\Theta(-D_{ij})\exp\left(\frac{D_{ij}}{\tau}\right)\bigg]\bigg]$$

$$i = 1 \ldots N, j = 1 \ldots N - 1, i \neq j,$$

where Θ denotes the Heaviside step function and all sums are taken over the N weights. The expressions are evaluated at the synaptic steady-state $w^\star$, which is obtained by numerically simulating the learning equations for all N synapses.

References

Abbott LF, Blum KI (1996) Functional significance of long-term potentiation for sequence learning and prediction. Cereb Cortex 6:406–416.

Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. Nat Neurosci 3:1178–1183.

Aharonov R, Gütig R, Rotter S, Aertsen A, Sompolinsky H (2001) Generalized synaptic updating in temporally asymmetric Hebbian learning. In: 10th Annual Computation and Neuroscience Meeting, p 119. San Francisco: Computational Neuroscience Organization.

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7:1129–1159.

Bi G-q, Poo M-m (1998) Activity-induced synaptic modifications in hippocampal culture: dependence on spike timing, synaptic strength, and cell type. J Neurosci 18:10464–10472.

Bi G-q, Poo M-m (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. Annu Rev Neurosci 24:139–66.

Castellani GC, Quinlan EM, Cooper LN, Shouval HZ (2001) A biophysical model of bidirectional synaptic plasticity: Dependence on AMPA and NMDA receptors. Proc Natl Acad Sci USA 98:12772–12777.

Câteau H, Kitano K, Fukai T (2002) An accurate and widely applicable method to determine the distribution of synaptic strengths formed by the spike-timing-dependent learning. Neurocomputing 44–46:343–351.

Crowley GC, Katz LC (2000) Early development of occular dominance colums. Science 290:1321–1324.

Debanne D, Gähwiler BH, Thompson SM (1994) Asynchronous pre- and postsynaptic activity induces associative long-term depression in area CA1 of the rat hippocampus *in vitro*. Proc Natl Acad Sci USA 91:1148–1152.

Debanne D, Gähwiler BH, Thompson SM (1998) Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. J Physiol (Lond) 507:237–247.

Eurich CW, Pawelzik K, Ernst U, Cowan JD, Milton JG (1999) Dynamics of self-organized delay adaptation. Phys Rev Lett 82:1594–1597.

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

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. Nature 384:76–78.

Gütig R, Aharonov R, Rotter S, Aertsen A, Sompolinsky H (2001) Generalized synaptic updating in temporally asymmetric Hebbian learning. In: Proceedings of the 28th Göttingen Neurobiology Conference, p 254. Stuttgart: Thieme.

Jeffress LA (1948) A place theory of sound localization. J Comput Physiol Psych 41:35–39.

Kempter R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. Phys Rev E 59:4498–4515.

Kempter R, Gerstner W, van Hemmen JL (2001) Intrinsic stabilization of output rates by spike-based Hebbian learning. Neural Comput 13:2709–2741.

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

Kuhn A, Rotter S, Aertsen A (2003) Higher-order statistics of input ensembles and the response of simple model neurons. Neural Comput 15:67–101.

Levy N, Horn D, Meilijson I, Ruppin E (2001) Distributed synchrony in a cell assembly of spiking neurons. Neural Netw 14:815–824.

Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neuroscience 8:791–797.

Magee JC, Johnston D (1997) A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. Science 275:209–213.

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 coincidence of postsynaptic APs and EPSPs. Science 275:213–215.

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

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

Rao RPN, Sejnowski TJ (2001) Spike-timing-dependent Hebbian plasticity as temporal difference learning. Neural Comput 13:2221–2237.

Roberts P (1999) Computational consequences of temporally asymmetric learning rules. I. Differential Hebbian learning. J Comput Neurosci 7:235–246.

Rubin J, Lee D, Sompolinsky H (2001) Equilibrium properties of temporally asymmetric Hebbian plasticity. Phys Rev Lett 86:364–367.

Senn W, Markram H, Tsodyks M (2001) An algorithm for modifying neurotransmitter release probability based on pre- and post-synaptic spike timing. Neural Comput 23:35–67.

Shriki O, Sompolinsky H, Lee D (2001) An information maximization approach to overcomplete and recurrent representations. In: Advances in neural information processing systems, Vol 13 (Dietterich TG, Becker S, Ghahramani Z, eds), pp 612–618. Cambridge, MA: MIT.

Sjöström PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32:1149–1164.

Song S, Abbott LF (2001) Cortical development and remapping through spike timing-dependent plasticity. Neuron 32:339–350.

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

van Ooyen A (2001) Competition in the development of nerve connections: a review of models. Network 12:R1–R47.

van Rossum MCW, Bi G-q, Turrigiano GG (2000) Stable Hebbian learning from spike timing-dependent plasticity. J Neurosci 20:8812–8821.

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