

# Balancing Feed-Forward Excitation and Inhibition via Hebbian Inhibitory Synaptic Plasticity

Yotam Luz<sup>1</sup>, Maoz Shamir<sup>1,2\*</sup>

**1** Department of Physiology and Neurobiology, Ben-Gurion University of the Negev, Beer-Sheva, Israel, **2** Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva, Israel

## Abstract

It has been suggested that excitatory and inhibitory inputs to cortical cells are balanced, and that this balance is important for the highly irregular firing observed in the cortex. There are two hypotheses as to the origin of this balance. One assumes that it results from a stable solution of the recurrent neuronal dynamics. This model can account for a balance of *steady state* excitation and inhibition without fine tuning of parameters, but not for *transient* inputs. The second hypothesis suggests that the feed forward excitatory and inhibitory inputs to a postsynaptic cell are already balanced. This latter hypothesis thus does account for the balance of transient inputs. However, it remains unclear what mechanism underlies the fine tuning required for balancing feed forward excitatory and inhibitory inputs. Here we investigated whether **inhibitory synaptic plasticity is responsible for the balance of transient feed forward excitation and inhibition**. We address this issue in the framework of a model characterizing the stochastic dynamics of temporally anti-symmetric Hebbian spike timing dependent plasticity of feed forward excitatory and inhibitory synaptic inputs to a single post-synaptic cell. Our analysis shows that inhibitory Hebbian plasticity generates 'negative feedback' that balances excitation and inhibition, which contrasts with the 'positive feedback' of excitatory Hebbian synaptic plasticity. As a result, this balance may increase the sensitivity of the learning dynamics to the correlation structure of the excitatory inputs.

**Citation:** Luz Y, Shamir M (2012) Balancing Feed-Forward Excitation and Inhibition via Hebbian Inhibitory Synaptic Plasticity. PLoS Comput Biol 8(1): e1002334. doi:10.1371/journal.pcbi.1002334

**Editor:** Peter E. Latham, Gatsby Computational Neuroscience Unit, University College London, United Kingdom

**Received:** February 14, 2011; **Accepted:** November 16, 2011; **Published:** January 26, 2012

**Copyright:** © 2012 Luz, Shamir. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was supported by THE ISRAEL SCIENCE FOUNDATION (grant No. 722/10). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: shmaoz@bgu.ac.il

## Introduction

### Balance of feed forward excitation and inhibition in the cortex

There is a striking difference between the number of synaptic contacts received by a typical cortical cell, which is about 3,000 to 10,000, and the required number of about 30 excitatory inputs to bring the cell to its firing threshold [1,2]. For example, inputs from 1,000 excitatory presynaptic cells firing at a medium rate of 10 spikes/s will yield  $100 \pm 10$  (mean  $\pm$  standard deviation) input spikes every 10 ms. At this high level of input, the postsynaptic cell is expected to saturate its firing rate. In addition, whereas *in vitro* experiments have shown that cortical cells fire relatively regularly, *in vivo* recordings reveal a highly irregular neural response [1]. However, fluctuations in the excitatory postsynaptic current are expected to be negligible relative to their mean, and therefore they themselves cannot account for the irregular firing of the postsynaptic cell, as observed *in vivo*.

The prevalent explanation for these seemingly contradictory findings is that excitatory and inhibitory inputs to cortical cells are balanced; i.e., the mean excitatory and inhibitory inputs to the cell are on the same order of magnitude. In terms of the balance hypothesis, the firing of the postsynaptic cell in the above example is not determined by the mean input of 100 excitatory postsynaptic potentials every 10 ms, which will be canceled by the mean inhibitory input, but rather by the fluctuations that are an order of magnitude smaller. Thus, the firing rate of the postsynaptic cell will not saturate and will be characterized by a high degree of variability.

There is, however, some confusion as to the nature and origin of the balanced state. One approach to this problem suggested that this balance results from recurrent neural dynamics [3]–[7]. Tsodyks and Sejnowski [3] showed that feedback (i.e., recurrent interactions) in finite chaotic networks can produce the desired balance. However, this solution requires strong interactions between the neurons and assumes a high probability of synaptic failure. Van Vreeswijk and Sompolinsky [6,7] showed analytically that the balance of *steady state* excitation and inhibition can be obtained as a stable solution to the network dynamics, and requires no special fine tuning of parameters due to the feedback from the network dynamics. These studies focused on the balancing of the lateral inputs via feedback of recurrent interactions and ignored the feed forward inputs to the system. However, empirical findings have shown there is a fast balance of *transient* inputs to a barrel cortical cell [8,9]. An alternative model suggested by Newsome and Shadlen [2], argues for the importance of balancing *feed forward* excitation and inhibition inputs in feed forward (propagating) neural networks. However, it remains unclear what mechanism underlies the fine tuning required for balancing feed forward excitatory and inhibitory inputs.

### Spike-timing-dependent plasticity of excitatory synapses

**Experimental characterization.** The overwhelming majority of scientific works on synaptic plasticity has focused on excitatory synapses. It is generally believed that synaptic plasticity is the basis for learning and memory. According to Hebb's rule [10], which

## Author Summary

One of the longstanding enigmas in neuroscience is the origin of inherent neural noise. It has been suggested that this noise results from a careful balance of excitatory and inhibitory inputs to neurons. Obtaining this balance requires fine tuning of the relative strengths of the inhibitory and excitatory inputs to each cell. However the mechanism that enables this fine tuning of parameters remains unclear. We suggest that a balance of excitatory and inhibitory inputs can be achieved via a process of unsupervised learning of the inhibitory synaptic strengths. We find that whereas Hebbian learning induces strong positive feedback on excitatory inputs that acts as a force that pulls them towards their boundaries, Hebbian learning of inhibition induces negative feedback. This negative feedback acts to balance the average excitatory input to the cell and makes the learning process less sensitive to the statistics of the inhibitory inputs. Surprisingly, this balance increases the sensitivity of learning to the statistics of the excitatory inputs. Thus, the balance of feed-forward excitation and inhibition emerges as a natural outcome of Hebbian learning applied to the inhibitory inputs to the cell.

remains the foundation of current assumptions on the nature of learning and memory, the interaction strength between two cells that are co-activated will facilitate synaptic efficacy. This rule has been extended to the temporal domain, where it is known as spike-timing-dependent plasticity (STDP). In many cases, the following causal relationship is assumed to exist: an excitatory synapse undergoes long-term potentiation if presynaptic firing precedes postsynaptic firing, and long-term depression is induced when the temporal order of firing is reversed [11–20]; this relationship is termed ‘temporally asymmetric Hebbian spike timing dependent plasticity’, Figure 1A (but see also [21,22] for examples of temporally asymmetric *anti* Hebbian plasticity, Figure 1B).

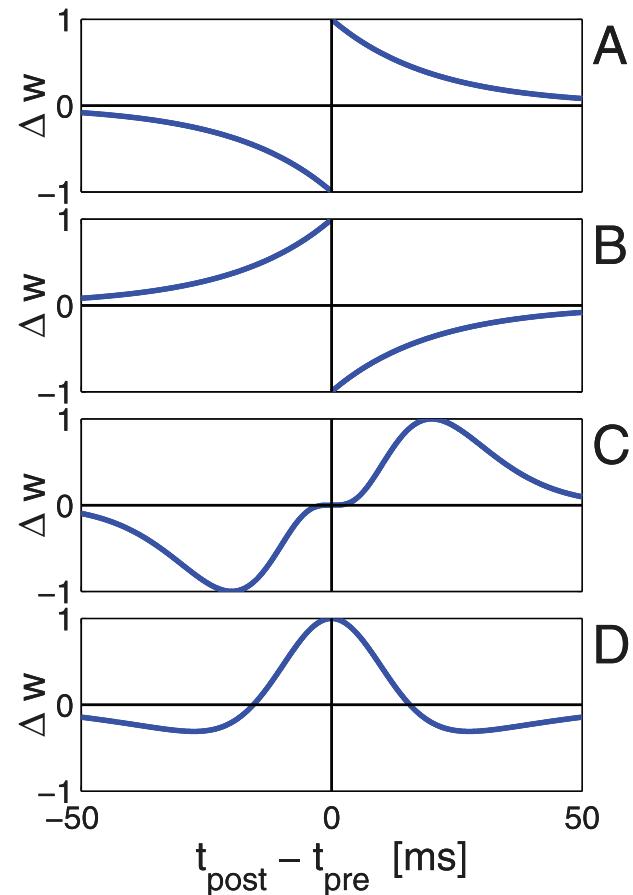
**Theoretical characterization.** Considerable theoretical effort has been devoted to studying the characteristics of STDP learning of excitatory synapses [23–43]. The canonical STDP rule shown in Figure 1A induces positive feedback in the following sense. If a certain synaptic weight is large, a presynaptic spike is more likely to elicit firing of the postsynaptic cell, following firing of the presynaptic cell. This will strengthen the synaptic weight, according to the STDP rule, which, in turn, will increase the likelihood of eliciting firing of the postsynaptic cell following firing of the presynaptic cell. On the other hand, if the synaptic weight is weak, then pre- and post-firing will be uncorrelated and the learning process will sample the STDP curve randomly. If the total depression (area under the acausal branch: post firing before pre) is larger than the total potentiation (area under the causal branch), then the synapse will be further weakened. Theoretical studies have shown that this positive feedback generates a bimodal distribution of excitatory synaptic weights. Note that this bimodal distribution of synaptic weights will exist for a limited range of parameters. Other choices of parameters will cause all the weights to cluster around either their upper or lower boundary. Further investigation has shown that a unimodal distribution can be obtained by scaling the amount of plasticity with the synaptic weight [40,41,42,43].

Theoretical studies have also explored the development of neuronal response properties [28,29,36,39,42]. In the absence of a reward signal, the STDP rule acts as an unsupervised learning algorithm. In unsupervised learning, the postsynaptic cell ‘learns’ salient features of the statistics of the presynaptic cell’s activity,

such as the correlation structure. Correlated synaptic inputs from a large group of cells are more likely to cause the postsynaptic cell to fire and, hence, strengthen their synaptic weights, whereas the STDP rule induces competition between different groups of correlated presynaptic cells. The origin of correlated activity of presynaptic cells in primary sensory regions reflects shared preferences for external stimuli. These results have been used to explain the development of ocular dominance columns. Recent theoretical studies have also managed to investigate learning of recurrent excitation [44–47].

## Spike-timing-dependent plasticity of inhibitory synapses

In contrast to the considerable number of empirical studies on excitatory synaptic plasticity, much less is known about inhibitory plasticity. Nevertheless, evidence for STDP of inhibitory synapses is beginning to emerge [48–56]. Woodin et al. [49] found that in hippocampal cultures and acute hippocampal slices, inhibitory synapses are potentiated if pre- and post-spikes are paired to fire within about 20 ms of each other and are depressed when the time difference is larger than about 20 ms, irrespective of the order of firing (e.g., similar to the STDP curve illustration in Figure 1D). In



**Figure 1. Illustration of different types of STDP curves.** The synaptic weight modification as a function of a  $\Delta t = t_{\text{post}} - t_{\text{pre}}$  caricature of four cases. **A** Hebbian temporally asymmetric STDP, e.g., Bi and Poo [11] for an excitatory synapse. **B** Anti-Hebbian temporally asymmetric STDP, e.g., [21,22]. **C** Hebbian temporally asymmetric STDP - inhibitory (entorhinal) cortical synapse, e.g., Haas et al. [50]. **D** Hebbian temporally symmetric STDP - inhibitory hippocampal synapse, e.g., Woodin et al. [49].

doi:10.1371/journal.pcbi.1002334.g001

the entorhinal cortex, Haas et al. [50] reported a temporally asymmetric Hebbian STDP rule, in which an inhibitory synapse is potentiated if the inhibitory presynaptic cell fires 5–15 ms before the postsynaptic cell and is depressed if the postsynaptic cell fires 5–15 ms before the presynaptic cell, Figure 1C.

Here we investigate the hypothesis that the temporally asymmetric Hebbian STDP of inhibitory synapses is responsible for the balance of transient feed forward excitation and inhibition. This issue is addressed in the framework of a stochastic dynamical model for learning feed forward inputs to a single postsynaptic cell, and then illustrated using numerical simulations or the learning dynamics of feed forward connections onto an integrate and fire neuron. We start by analyzing the stochastic learning dynamics of a single inhibitory synapse. Then we turn to investigate the learning dynamics of a population of inhibitory feed forward synapses. Finally we study a model that incorporates learning of both feed forward excitatory and inhibitory inputs to a single postsynaptic cell.

## Results

### Learning of a single inhibitory synapse

Considerable theoretical attention has been devoted to the study of learning a single excitatory synapse, see e.g., [41,43], and in particular the distribution of resultant synaptic weights. Two forces work to shape this distribution. The first is the positive feedback of learning an excitatory synapse (see above) that pushes the synaptic weights to their high and low saturation boundaries and hence contributes to a bimodal distribution. The second is the weight dependence of the STDP rule, which is able to decrease the strength of the positive feedback close to the saturation boundaries and hence contributes to a unimodal synaptic distribution. Following the results of Haas et al. [50] (but see also Woodin et al. [49]), we studied a family of temporally asymmetric STDP rules for the inhibitory synaptic weight  $w^I$  of the form:

$$\Delta w^I = \pm \lambda f_{\pm}(w^I) K(|t|) \quad (1)$$

where  $w^I \in [0,1]$  is the dynamic variable that describes the synaptic strength,  $\Delta w^I$  is the change in the synaptic strength following pre (–) or post (+) synaptic firing,  $t$  is the time difference between the pre- and post-synaptic firing,  $\lambda$  is the learning rate, and  $f(\cdot)$  and  $K(\cdot)$  are the weight dependence and temporal filter of the STDP rule, respectively. For convenience we adopted a notation similar to that of [42]. Equation (1) defines the synaptic change due to a single pair of pre-post spikes. We shall assume that the STDP rule is additive with respect to all pairs of pre-post spike times. Equation (1) describes a temporally asymmetric STDP rule, as reported in Haas et al. [50]. The temporal filter,  $K(t)$ , can be modeled by a decaying exponent with a characteristic timescale of about 20 ms (see Figure 1A, e.g., [11]), or by a gamma distribution, similar to the results of Haas et al. [50] (see Figure 1C). For concreteness throughout the paper in all of our numerical simulations we used  $K(t) = e^{-|t|/\tau}$  with  $\tau = 20\text{ms}$ . Note that the STDP rule is temporally asymmetric and not antisymmetric due to the different scalings of depression and potentiation in  $w^I$ , which is expressed in the  $f_{\pm}(w^I)$  dependence of the synaptic update rule. The structure of  $f_{\pm}(w^I)$  is somewhat less clear from the empirical literature; thus, for convenience of analysis, we adopted the formulation in [42], which generalizes, e.g., [23], [41] (but see [27]):

$$f_+(w^I) = (1 - w^I)^{\mu} \quad (2)$$

$$f_-(w^I) = \alpha (w^I)^{\mu} \quad (3)$$

where  $\mu \in [0,1]$  is a parameter that characterizes the weight dependence of the STDP rule. Following equation (1), changes in the synaptic weight,  $w^I$ , occur only at times where either pre or post synaptic cells have fired:

$$w^I(t + \delta_t) = w^I(t) + \lambda f_+(w^I) \hat{\rho}_{\text{fire}}^{\text{post}} \left( \in [t, t + \delta_t) \right) \sum_{j=1}^{\infty} K(t - t_j^{\text{pre}}) \quad (4)$$

$$+ \lambda f_-(w^I) \hat{\rho}_{\text{fire}}^{\text{pre}} \left( \in [t, t + \delta_t) \right) \sum_{j=1}^{\infty} K(t - t_j^{\text{post}})$$

where  $\hat{\rho}_{\text{fire}}^{\text{post/pre}} \left( \in [t, t + \delta_t) \right)$  is a stochastic variable which is one if the post/pre fired at time interval  $[t, t + \delta_t)$  and zero otherwise;  $\{t_j^{\text{post/pre}}\}$  are the spike times of the post/pre synaptic neuron, respectively; the summation is over past times:  $t_j^{\text{post/pre}} < t$ . Note that the summation over all past spike times results from our assumption that the synaptic update rule, equation (1), is additive with respect to all pre-post spike time pairs. Taking the short time limit,  $\delta_t \rightarrow 0$ , yields

$$\lim_{\delta_t \rightarrow 0} \frac{w^I(t + \delta_t) - w^I(t)}{\delta_t} = \frac{dw^I(t)}{dt} \quad (5)$$

$$\lim_{\delta_t \rightarrow 0} \frac{1}{\delta_t} \hat{\rho}_{\text{fire}}^{\text{pre}} \left( \in [t, t + \delta_t) \right) = \rho_{\text{pre}}^I(t) \quad (6)$$

where  $\rho_{\text{pre}}^I(t) = \sum_j \delta(t - t_j^{\text{pre}})$  describes the spike train of the inhibitory pre-synaptic neuron in terms of a series of delta function pulses at the spike times of the cell,  $\{t_j^{\text{pre}}\}_{j=1}^{\infty}$  (the summation is over all the spike times). Similarly  $\rho_{\text{post}}(t)$  describes the post synaptic spike train. We obtain:

$$\dot{w}^I = \lambda f_+(w^I) \mathcal{L}^+(t) - \lambda f_-(w^I) \mathcal{L}^-(t) \quad (7)$$

$$\mathcal{L}^+(t) = \int_{-\infty}^t dt' \rho_{\text{post}}(t') \rho_{\text{pre}}^I(t) K(t - t') \quad (8)$$

$$\mathcal{L}^-(t) = \int_{-\infty}^t dt' \rho_{\text{post}}(t') \rho_{\text{pre}}^I(t) K(t - t') \quad (9)$$

In the limit of slow learning rate, the synaptic weight,  $w^I$ , is relatively fixed over long periods of time,  $\mathcal{O}(1/\lambda)$ , during which the right hand side of equation (7) is sampled by the dynamics such that we can neglect its fluctuations around its mean in the limit of  $\lambda \rightarrow 0$ . This approximation yields deterministic dynamic equations for the mean synaptic weights:

$$\dot{w}^I = \lambda f_+(w^I) L^+ - \lambda f_-(w^I) L^- \quad (10)$$

$$L^- = \int_{-\infty}^t dt' \langle \rho_{\text{post}}(t') \rho_{\text{pre}}^I(t) \rangle K(t-t') \quad (11)$$

$$L^+ = \int_{-\infty}^t dt' \langle \rho_{\text{post}}(t) \rho_{\text{pre}}^I(t') \rangle K(t-t') \quad (12)$$

where  $\langle \cdot \rangle$  denotes averaging with respect to the distribution of the neural firing, for a given fixed synaptic weight,  $w^I$ . To proceed with the analysis we need to specify the cross-correlation function between the pre-synaptic input and the post-synaptic response, and in particular its dependence on the synaptic weight,  $w^I$ . However, the calculation of dependence of the temporal structure of the pre-post firing probability on the synaptic weight, even in the simple case of an integrate and fire neuron is a first-passage time problem [57,58], which is not a trivial task. Recently, Ostojic et al [59] succeeded in analyzing the cross-correlation function between two integrate and fire neurons assuming the synaptic coupling is sufficiently weak such that the firing rate of the postsynaptic cell can be approximated by a linear function of the presynaptic input. Similarly, in this section, we assume that the synaptic coupling is sufficiently weak such that we can approximate the postsynaptic firing rate by a linear function of the presynaptic input (see also [43,60,61]), yielding

$$\langle \rho_{\text{post}}(t) \rho_{\text{pre}}^I(t') \rangle = \begin{cases} r_{\text{post}} r_{\text{pre}} (1 - w^I \gamma(t-t')), & t > t' \\ r_{\text{post}} r_{\text{pre}}, & t < t' \end{cases} \quad (13)$$

where  $r_{\text{pre/post}}$  is the pre/post synaptic mean firing rate; and the function  $\gamma(t-t')$  describes the decrease in the conditional firing rate of the postsynaptic neuron at time  $t$  following an inhibitory input spike at time  $t' < t$ .

Substituting equation (13) into equation (10) one obtains:

$$\dot{w}^I = -\lambda r_{\text{pre}} r_{\text{post}} \bar{K} \left\{ \Delta f(w^I) + w^I \frac{\gamma \bar{K}}{\bar{K}} f_+(w^I) \right\} \quad (14)$$

where  $\Delta f(w^I) \equiv f_-(w^I) - f_+(w^I)$ , and  $\bar{X} \equiv \int X(t) dt$ . This model highlights the following three major differences between the dynamics of inhibitory and excitatory synapses.

1. The temporally asymmetric Hebbian STDP rule, equation (1), yields a *negative feedback*, which is characterized by a unimodal distribution for the inhibitory synapses. This contrasts with the temporally asymmetric Hebbian STDP rule for an excitatory synapse, which yields a positive feedback and allows for bi-stable solutions.
2. As there is only one stable fixed point for the drift velocity, which is stable for all  $\mu \in [0,1]$ , there is no theoretical need for  $\mu \neq 0$ , which was introduced to weaken the positive feedback of the excitatory STDP. Hence, we can take  $\mu = 0$ .
3. The relative strength of the depression needs to be weaker than the potentiation; i.e., in the  $\mu = 0$  case, the ratio  $\alpha$  of the area under the acausal and the causal branches of the STDP curve needs to be  $\alpha < 1$ , to prevent decay of all inhibitory synapses to zero.

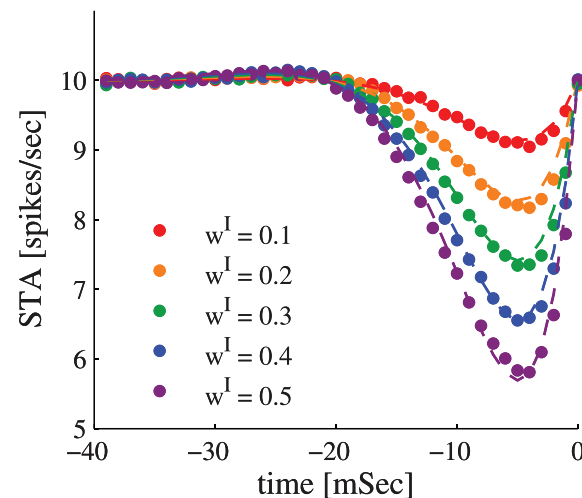
**Numerical simulations of inhibitory STDP.** To test our results beyond the analysis of the above simplified model we performed numerical simulations of the learning dynamics of a feed forward inhibitory synapse to a conductance based integrate

and fire postsynaptic neuron (see Methods for details). Figure 2 shows the spike triggered average firing rate of a single presynaptic inhibitory cell as a function of time relative to the firing of the postsynaptic cell (negative times imply pre fired before post), for different fixed values (i.e., without learning) of the synaptic coefficient strength  $w^I = 0.1, 0.2, 0.3, 0.4,$  and  $0.5$  in red, orange, green, blue, and purple circles, respectively. The dashed lines show fits of the form  $STA(t) = r_{\text{pre}}(1 - w^I \gamma(t))$  with  $\gamma(t) = a_\gamma \sin(\omega_\gamma |t|) e^{-|t|/\tau_\gamma}$ . For short times preceding the postsynaptic firing rate the conditional mean firing rate of the inhibitory presynaptic neuron is less than its marginal mean ( $r_{\text{pre}} = 10$  spike/s in the specific example of Figure 2). This decrease is approximately linear in the synaptic weight; whereas for long times, the spike triggered average converges to  $r_{\text{pre}}$ . Hence, equation (13) provides a fair description of the pre-post correlations.

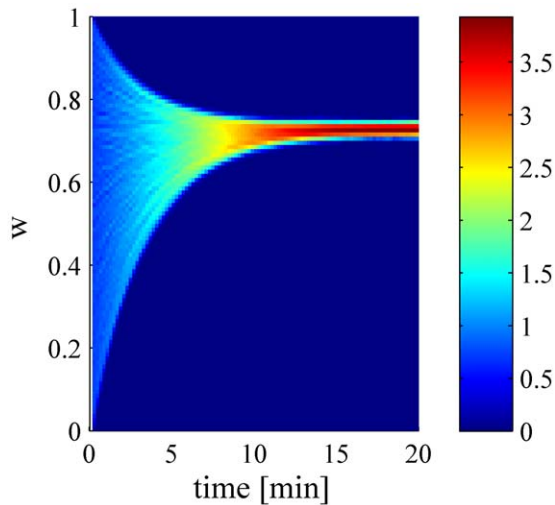
We simulated the learning dynamics of a single inhibitory synapse, keeping the rest of the inhibitory and excitatory inputs to the cell fixed. Figure 3 shows the temporal evolution of the empirical distribution of the synaptic weight. The empirical distribution was obtained by averaging over 1999 realizations of the stochastic learning dynamics of a single inhibitory synapse with uniformly distributed initial conditions. As expected from theory, the synaptic weight converges to a single fixed point,  $w_0^I$ , regardless of initial conditions or noise realization. Figures 4A and B show the dependence of the asymptotic synaptic weight,  $w_0^I$ , on different parameters of the learning dynamics. The solid red lines show the theoretical prediction; i.e., the fixed point of equation (14), that was calculated using the function  $\gamma(t)$  that was obtained from the fit to the spike triggered average, Figure 2. The dashed blue lines show the fixed point solution to equation (14) using an optimized value for  $\gamma \bar{K} / \bar{K}$  to best fit the simulation results.

### Learning a feed forward inhibitory synaptic population

Before deriving the full model for studying the learning dynamics of both feed-forward excitatory and inhibitory synapses,



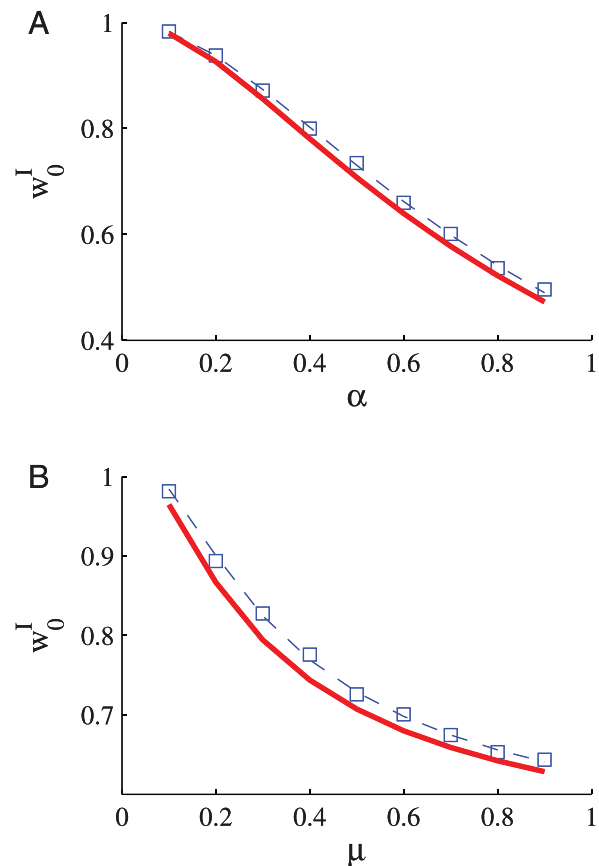
**Figure 2. Spike triggered average of inhibitory presynaptic cell.** The conditional mean firing rate of the inhibitory presynaptic cell given the postsynaptic cell has fired at time  $t=0$ , is plotted as function of time, for different values for the strength of the presynaptic weight  $w^I = 0.1, 0.2, 0.3, 0.4,$  and  $0.5$  in red, orange, green, blue and purple circles, respectively. The dashed lines show the fits of the form  $STA(t) = r_{\text{pre}}(1 - w^I \gamma(t))$  with  $\gamma(t) = a_\gamma \sin(\omega_\gamma |t|) e^{-|t|/\tau_\gamma}$ . The parameter  $\omega_\gamma$  was set to match the zero crossing point of  $\gamma(t)$ , and we optimized the fit over the parameters  $a_\gamma$  and  $\tau_\gamma$ . doi:10.1371/journal.pcbi.1002334.g002



**Figure 3. The dynamics of the synaptic weight distribution.** The probability density of the synaptic weight,  $P(w^I, t)$  is shown in color code as a function of time. The range of values of  $w^I$ ,  $[0, 1]$ , was divided into one hundred equally sized bins, and the probability of having a value in a corresponding bin of size of  $1/100$  was estimated numerically. The color scale is shown in terms of  $\log(1 + \text{Pr}(w^I, t))$ . The stochastic learning dynamics of a single inhibitory synapse was simulated using an integrate and fire model (see Methods). The probability density was estimated from the simulations by averaging over 1999 repeats with different realizations for the noise (stochasticity of the presynaptic neurons' firing) and with initial conditions that were uniformly spaced in the interval  $(0, 1)$ . Here we used  $\alpha=0.5$ ,  $\mu=0.5$ , and  $\lambda=10^{-3}$ . doi:10.1371/journal.pcbi.1002334.g003

it is instructive to first study the artificial case of learning only the inhibitory inputs. We model a population of  $N_I$  inhibitory synaptic weights,  $\{w_j^I\}_{j=1}^{N_I}$ , from  $N_I$  presynaptic inhibitory neurons projecting onto a single postsynaptic cell. Let us denote by  $\rho_i^I(t)$  the spike train of the  $i$ th presynaptic inhibitory neuron:  $\rho_i^I(t) = \sum_{j=1}^{\infty} \delta(t - t_j^i)$  where  $\{t_j^i\}_{j=1}^{\infty}$  are the spike times of cell  $i$  (we shall omit the subscript pre hereafter). As noted above, the cross-correlation function between the pre and postsynaptic cells is an important quantity that affects the neural dynamics. In the previous section we calculated the cross-correlation function using a linear approximation to an 'exact' model. Instead, here and in the following section we will use an exact solution of a simplified linear model of a more abstract neuron. For analytical tractability the postsynaptic response,  $\rho_{\text{post}}(t)$ , is modeled to be a delayed linear sum of its inputs:  $\rho_{\text{post}}(t + \epsilon) = I_{\text{exc}} - \frac{1}{N_I} \sum_{i=1}^{N_I} w_i^I \rho_i^I(t)$ ; where  $\epsilon$  is small and positive to ensure causality, and  $I_{\text{exc}}$  represents a constant excitatory input to the cell. For simplicity we assume that the correlations are instantaneous,  $\langle \rho_i^I(t) \rho_j^I(t') \rangle = r^2 + c_{ij} \delta(t - t')$ . We shall further assume that the statistics of the input neuron responses are isotropic; i.e., no input neuron is statistically special. This assumption implies that: 1) The mean firing rate of all inhibitory presynaptic neurons is equal,  $\langle \rho_i^I \rangle = r$ . 2) The correlation structure of each input neuron with the rest of the input population (up to a permutation of indices) is the same. In particular, the correlation of a single input neuron with the total response of the population,  $\sum_j c_{ij}^H$ , is equal for all input neurons; hence, the uniform vector  $\mathbf{v}_0 = (1, 1, 1, \dots, 1)$  is an eigenvector of the matrix  $\mathbf{c}$ . We obtain:

$$\dot{w}_i^I = \lambda f_+(w_i^I) \mathcal{L}_i^+(t) - \lambda f_-(w_i^I) \mathcal{L}_i^-(t), \quad i \in \{1, \dots, N_I\} \quad (15)$$



**Figure 4.** The asymptotic synaptic weight,  $w_0^I$ , in learning dynamics of a single inhibitory synapse is shown as a function of **A** the ratio between potentiation and depression,  $\alpha$  (for  $\mu=0.5$ ) **B** the parameter  $\mu$  (for  $\alpha=0.5$ ). The open blue squares show the results obtained in simulating the stochastic learning dynamics using an integrate and fire postsynaptic neuron (see Methods). The solid red line shows the fixed point of equation (14), calculated using the function  $\gamma(t)$  that was obtained from the fit to the spike triggered average, Figure 2. The dashed blue line shows the fixed point solution to equation (14) using a value for  $\bar{\gamma} \bar{K} / \bar{K}$  that was optimized to best fit the simulation results. doi:10.1371/journal.pcbi.1002334.g004

$$\mathcal{L}_i^+(t) = \int_{-\infty}^t dt' \rho_{\text{post}}(t') \rho_i^I(t') K(t - t') \quad (16)$$

$$\mathcal{L}_i^-(t) = \int_{-\infty}^t dt' \rho_{\text{post}}(t') \rho_i^I(t') K(t - t') \quad (17)$$

As above, in the limit of slow learning rate,  $\lambda \rightarrow 0$ , we can neglect the fluctuations of the synaptic weights around their mean, yielding

$$\frac{1}{\lambda r^2 \tau} \dot{w}_i^I = -\Delta f(w_i^I) \left( W_{\text{exc}} - \frac{1}{N_I} \sum_{j=1}^{N_I} w_j^I \right) - f_+(w_i^I) \frac{1}{N_I} \sum_{j=1}^{N_I} C_{ij}^H w_j^I \quad (18)$$

where  $\tau = \int_0^{\infty} K(t) dt$ ,  $W_{\text{exc}} = I_{\text{exc}}/r$ , and  $C_{ij}^H = \frac{c_{ij}}{\tau r^2}$  is a non-negative symmetric matrix. From the assumption of isotropy the uniform vector  $\mathbf{v}_0 = (1, 1, 1, \dots, 1)$  is an eigenvector of the correlation matrix,  $\mathbf{C}^H \mathbf{v}_0 = N_I C_0^H \mathbf{v}_0$ , with eigenvalue  $N_I C_0^H = \sum_{j=1}^{N_I} C_{ij}^H$ .

From the assumption of isotropy, a homogeneous solution,  $w_i^I = w_0^I, \forall i \in \{1, \dots, N\}$ , to the dynamics exists and obeys

$$\frac{1}{\lambda_I^2 \tau} \dot{w}_0^I = -\Delta f(w_0^I)(W_{\text{exc}} - w_0^I) - f_+(w_0^I)C_0^I w_0^I \quad (19)$$

with  $\Delta f(w^I) = f_-(w^I) - f_+(w^I)$ . The fixed point equation for the homogeneous solution,  $w_i^I = w_0^I, \forall i \in \{1, \dots, N\}$ , is given by

$$\frac{f_-(w_0^I)}{f_+(w_0^I)} \equiv \alpha \left( \frac{w_0^I}{1 - w_0^I} \right)^\mu = 1 - C_0^I \frac{w_0^I}{W_{\text{exc}} - w_0^I} \quad (20)$$

The left hand side of equation (20),  $\alpha \left( \frac{w_0^I}{1 - w_0^I} \right)^\mu$ , starts from 0 at  $w_0^I = 0$ , increases monotonically in the range of  $w_0^I$ ,  $\left( \log \left( \frac{w_0^I}{1 - w_0^I} \right) \right)^\mu > 0, \forall w_0^I \in (0, 1)$ , and diverges to infinity as  $w_0^I \rightarrow 1$ . The right hand side of equation (20) decreases monotonically  $\left( 1 - C_0^I \frac{w_0^I}{W_{\text{exc}} - w_0^I} \right) < 0, \forall w_0^I < W_{\text{exc}}$ , starts from the value of one at  $w_0^I = 0$ , crosses zero at  $w_0^I = \frac{W_{\text{exc}}}{1 + C_0^I}$  and is continuous in the range  $[0, W_{\text{exc}})$ . Hence, equation (20) has a unique solution,  $w_0^{I*}$ , in the range of  $w_0^I \in (0, \min\{1, W_{\text{exc}}\}) \Rightarrow (W_{\text{exc}} - w_0^{I*}) > 0$ , implying a net positive input to the postsynaptic cell. For  $w_0^I < w_0^{I*}$  the temporal derivative of the homogeneous solution, equation (19), will be positive, and for  $w_0^I > w_0^{I*}$  it will be negative. Hence, the uniform solution is stable to fluctuations in the uniform direction. For  $W_{\text{exc}} < 1$ , equation (20) may have an additional solution with  $w_0^I > W_{\text{exc}}$ . This solution is not physical, because  $W_{\text{exc}} - w_0^I > 0$  represents the case where the net input to the postsynaptic cell is inhibitory. A neuron with net inhibitory input will not fire and there will be no learning.

To study the stability of the homogeneous solution to general perturbations, we consider an arbitrary (though small) deviation from the homogeneous solution,  $\delta w_i^I = w_i^I - w_0^I$  (note that we omitted the  $*$  in the notation of the uniform solution). To first order in the deviations one obtains:

$$\begin{aligned} \frac{1}{\lambda_I \tau^2} \delta \dot{w}_i^I &= -g_0 \delta w_i^I + \Delta f(w_0^I) \frac{1}{N_I} \sum_{j=1}^{N_I} \delta w_j^I \\ &\quad - f_+(w_0^I) \frac{1}{N_I} \sum_{j=1}^{N_I} C_{ij} \delta w_j^I \equiv - \sum_{j=1}^{N_I} M_{ij} \delta w_j^I \end{aligned} \quad (21)$$

$$g_0 = (I_{\text{exc}} - w_0^I) \frac{d}{dw_0^I} \Delta f(w_0^I) + C_0^I \frac{d}{dw_0^I} f_+(w_0^I) \quad (22)$$

$$= \mu (I_{\text{exc}} - w_0^I) (\alpha w_0^I + (1 - w_0^I)^{\mu-1}) - \mu C_0^I w_0^I (1 - w_0^I)^{\mu-1}$$

At the homogenous fixed point, equation (20), one obtains

$$g_0 = \alpha \mu (W_{\text{exc}} - w_0^I) \frac{(w_0^I)^{\mu-1}}{1 - w_0^I} \quad (23)$$

Hence,  $g_0 > 0$ . The eigenvalues  $\{M_n\}$  of the stability matrix  $\mathbf{M}$

obey

$$M_0 = +g_0 - \Delta f(w_0^I) + f_+(w_0^I)C_0^I \quad (24)$$

$$M_n = +g_0 + f_+(w_0^I)C_n^I, \quad (n \neq 0) \quad (25)$$

where  $C_n^I$  is an eigenvalue of  $\mathbf{C}^{\text{II}}/N_I$ , and  $C_0^I$  is the specific eigenvalue in the uniform direction. For stability against fluctuations in the uniform direction, see above. At orthogonal directions  $M_n > 0, \forall n \neq 0$  since  $C_n^I \geq 0$  due to the positivity of the correlation matrix  $\mathbf{C}^{\text{II}}$ . Hence, due to the negative feedback of Hebbian learning of inhibition, the uniform solution is always stable.

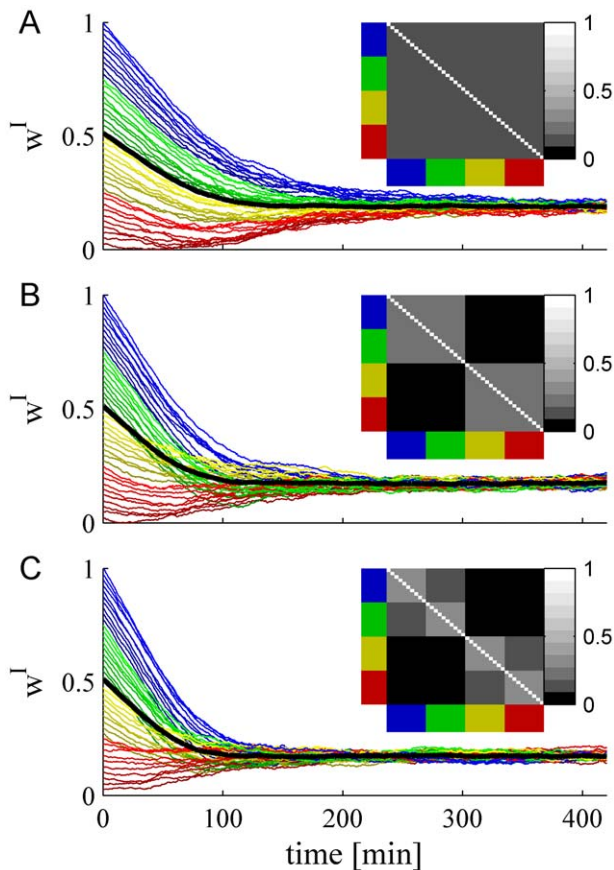
The STDP learning rule is an unsupervised learning rule and as such can learn salient features of the statistics of its inputs. The input statistics are expressed in the learning dynamics, equation (18), by the effective interactions between the synapses generated via the input correlations,  $C_{ij}^{\text{II}}$ , and the learning dynamics. Such sensitivity to input statistics may be manifested in solutions to the fixed point of the synaptic dynamics, equation (18), that reflect the correlations' structure of the input population. However, the homogenous solution, in which  $w_i^I = w_0^I, \forall i \in \{1, \dots, N\}$ , always exists from the assumption of isotropy and is stable. Thus, unlike the learning dynamics of excitatory synapses, temporally asymmetric Hebbian learning stabilizes the homogeneous solution.

**Numerical simulations of inhibitory STDP.** Numerical simulations corroborate the claim that the negative feedback of inhibitory plasticity stabilizes the homogeneous solution. Figures 5A, B and C show two examples of the stochastic learning dynamics of a population of  $N_I = 40$  inhibitory synapses (the values of the excitatory synaptic weights were held fixed). The weight of every synapse is depicted as a function of time. In the example in Figure 5A the system is *homogeneous* with uniform correlation structure between all inhibitory presynaptic neurons, with a correlation coefficient of  $cc = 0.1$  (see inset). After a transition period, which scales linearly with the learning rate, all memory of their initial conditions are lost and the system converges to a uniform solution. Figure 5B shows an example of a *non-homogeneous* system, where the inhibitory presynaptic population is composed of two sub-populations of equal size with a correlation coefficient of  $cc = 0.2$  between cells from the same sub-population and a correlation coefficient of  $cc = 0$  between cells from different sub-populations (see inset). The different sub-populations are depicted by different hues of red and yellow versus green and blue and are also distinguished by the range of their initial conditions. Nevertheless, as can be seen from the figure, the homogeneous solution remains stable, in line with the above analysis. Moreover, because Hebbian learning of inhibition induces negative feedback, the non-uniform correlations accelerate the convergence to a uniform solution (compare Figure 5A and 5B). Figure 5C shows another example of a *heterogenous* population with a more elaborate correlation structure (see inset), yet the homogeneous solution of the STDP dynamics remains stable.

The fixed point, equation (20), for the  $\mu = 0$ , is given by

$$w_0^I = \frac{W_{\text{exc}}}{1 + \frac{C_0^I}{1 - \alpha}} \quad (26)$$

Figure 6 shows the asymptotic value of the learned synaptic uniform weights  $w_0^I$  as a function of the strength of the excitatory input to the postsynaptic cell,  $W_{\text{exc}}$ . The deviations from the linear relation at low levels of excitatory input result from the non-linearity



**Figure 5. The stochastic learning dynamics of a population of  $N_I=40$  pre synaptic inhibitory neurons.** Each trace shows the dynamics of a single synaptic weight. We color cells 1–10 in blue, 11–20 in green, 21–30 in yellow, and 31–40 in red. The firing rate statistics of the inhibitory neurons followed Poisson statistics with a mean rate of  $r_{pre} = 10$  spikes/s. Initial conditions were distributed evenly from 1 to 0 for cells 1 to 40, respectively, i.e.,  $w_1^I(t=0) > w_2^I(t=0) > \dots > w_{40}^I(t=0)$ . Thus, cells from the ‘blue’ population have higher initial conditions than cells from the ‘green’ population and so on. The different hues of each color distinguish the cells on the basis of their initial conditions. The thick black line shows the population average of the synaptic weights. Panels A, B and C differ in the correlation structure of the pre-synaptic neurons, shown in the inset. **A Homogeneous population** with uniform correlations. The correlation coefficient between all inhibitory cell pairs was  $cc=0.1$  (see Methods). **B Heterogeneous population.** The population of  $N_I=40$  inhibitory neurons was composed of two homogenous sub-populations of 20 cells each (sub-population one: blue and green, sub-population two: yellow and red). We used a correlation coefficient of  $cc=0$  between all cells from different sub-population and uniform correlation coefficient of  $cc=0.2$  between cells from the same sub-population (see Methods). **C Heterogeneous population.** The population of  $N_I=40$  inhibitory neurons was composed of four homogenous sub-populations of 10 cells each. The different colors distinguish the cells belonging to the different populations. The correlation coefficient within each sub-populations was  $cc=0.3$ . The correlation coefficient between cells in the blue sub-populations and the green sub-populations, and pairs from yellow and red was  $cc=0.1$ . All other correlation coefficients were zero. Here we used  $\alpha=0.85$ ,  $\mu=0$ , and  $\lambda=2 \cdot 10^{-4}$ . doi:10.1371/journal.pcbi.1002334.g005

of the integrate and fire neuron used in the simulations. In particular, note that the linear interpolations (dashed lines) do not cross the abscissa at the origin, but rather at a positive value, representing the effect of a positive threshold in the  $I-v$  curve (i.e., input current vs. output firing rate) of the postsynaptic neuron.

Figure 7 shows the asymptotic value of the learned synaptic uniform weights  $w_0^I$  as a function of the strength of the uniform correlation coefficient between the inhibitory presynaptic cells.

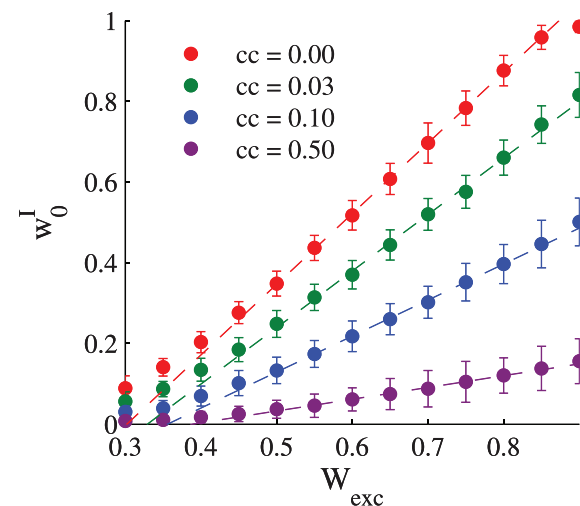
### Learning feed forward inhibitory and excitatory ‘synaptic populations’

We now turn to generalize the above formalism to study the effect of Hebbian inhibitory synaptic plasticity in the framework of a simplified model for learning  $N_E$  excitatory synapses,  $\{w_j^E\}_{j=1}^{N_E}$ , and  $N_I$  inhibitory synapses,  $\{w_i^I\}_{i=1}^{N_I}$ , constituting feed-forward input to a single postsynaptic cell. We denote by  $\rho_i^{E/I}(t)$  the spike train of the  $i$ th excitatory/inhibitory neuron,  $\rho_i^{E/I} = \sum_{j=1}^{\infty} \delta(t-t_j^{E/I,i})$  where  $\{t_j^{E/I,i}\}_{j=1}^{\infty}$  are the spike times of the cell. As above, the postsynaptic response,  $\rho_{post}(t)$ , is modeled to be a delayed linear sum of its inputs:

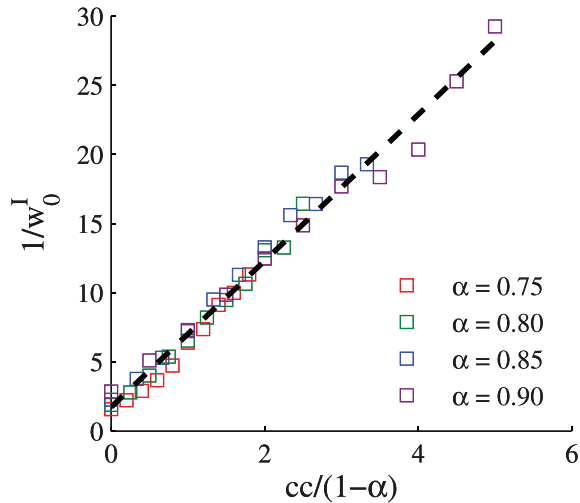
$$\rho_{post}(t+\epsilon) = \frac{1}{N_E} \sum_{j=1}^{N_E} w_j^E \rho_j^E(t) - \frac{1}{N_I} \sum_{i=1}^{N_I} w_i^I \rho_i^I(t) \quad (\text{where } \epsilon \text{ is small and positive}).$$

We further assume that the system is isotropic, the mean firing rates are equal for all presynaptic neurons,  $\langle \rho_j^X \rangle = r$  ( $\forall i \in \{1, \dots, N_X\}, X \in \{I, E\}$ ), correlations are instantaneous,  $\langle \rho_i^X(t) \rho_j^X(t') \rangle = r^2 + c_{ij}^{XX} \delta(t-t')$ , and inhibitory and excitatory inputs are uncorrelated. The excitatory synapses also follow temporally asymmetric Hebbian spike timing dependent plasticity according to equations (1)–(2) with  $\alpha_E > 1$  and  $\mu_E \in [0, 1]$ . In the limit of a slow learning rate the mean field synaptic dynamics are given by

$$\frac{1}{\lambda r^2 \tau} \dot{w}_i^E = -\Delta f(w_i^E) \left( \frac{1}{N_E} \sum_{j=1}^{N_E} w_j^E - \frac{1}{N_I} \sum_{j=1}^{N_I} w_j^I \right) + f_+(w_i^E) \frac{1}{N_E} \sum_{j=1}^{N_E} C_{ij}^{EE} w_j^E \quad (27)$$



**Figure 6. Balance of excitation and inhibition.** The asymptotic value of the uniform synaptic coefficient,  $w_0^I$  is shown as a function of the total excitatory input to the cell, for different levels of uniform correlations between the presynaptic inhibitory neuron population,  $cc=0, 0.03, 0.1$ , and  $0.5$  from top to bottom. The strength of the excitatory synapses were uniform and were held fixed during each simulation. In this simulation we used  $\alpha=0.9$ ,  $\mu=0$ , and the correlations between the inhibitory cells were uniform. The dashed lines show linear regression lines for comparison. The regression was computed using only the points with  $w_{exc}$  that were in the range of  $[0.5, 0.85]$ . doi:10.1371/journal.pcbi.1002334.g006



**Figure 7. The effect of correlations on the excitation and inhibition balance.** The asymptotic value of the uniform synaptic coefficient,  $w_0^I$  is shown as a function of the level of the homogeneous correlations in the firing of the inhibitory presynaptic population, for different values of  $\alpha = 0.75, 0.8, 0.85,$  and  $0.9$  for red, green, blue, and purple, respectively. The dashed line shows the linear interpolation of the data,  $(w_0^I)^{-1} = a \frac{cc}{1-\alpha} + b$ ;  $a = 5.3$  and  $b = 1.7$ . In this simulation we used  $\mu = 0$ , the correlations between the inhibitory cells were uniform and were varied in the range  $cc = 0, 0.05, \dots, 0.5$ , for every value of  $\alpha$ . doi:10.1371/journal.pcbi.1002334.g007

$$\frac{1}{\lambda r^2 \tau} \dot{w}_i^I = (1 - \alpha_I) \left( \frac{1}{N_E} \sum_{j=1}^{N_E} w_j^E - \frac{1}{N_I} \sum_{j=1}^{N_I} w_j^I \right) - \frac{1}{N_I} \sum_{j=1}^{N_I} C_{ij}^{II} w_j^I \quad (28)$$

where  $C_{ij}^{XX} = \frac{C_{ij}^{XX} - r^2}{r^2 \tau}$ ,  $X \in \{I, E\}$ . Note that equation (28) uses  $\mu = 0$  for the inhibitory synapses. The homogeneous solution,  $w_i^I = w_0^I, \forall i$  and  $w_j^E = w_0^E, \forall j$ , to the fixed point equations of the dynamics always exists, and obeys:

$$w_0^I = \frac{1}{1 + \frac{C_0^I}{1 - \alpha_I}} w_0^E \quad (29)$$

$$0 = -f_-^E(w_0^E)(w_0^E - w_0^I) + f_+^E(w_0^E)((1 + C_0^E)w_0^E - w_0^I) \quad (30)$$

where from our assumption of isotropy the uniform vector is an eigenvector of the non-negative symmetric matrix  $\mathbf{C}^{XX}$  ( $X = E, I$ ) with the corresponding eigenvalue  $N_I C_0^X = \sum_j C_{ij}^{XX}$ . Fluctuation analysis around the homogeneous fixed point yields:

$$\begin{aligned} \frac{1}{\lambda r^2 \tau} \delta \dot{w}_i^E &= -g_0^E \delta w_i^E - \Delta f(w_0^E)(\delta w_0^E - \delta w_0^I) \\ &+ f_+(w_0^E) \frac{1}{N_E} \sum_{j=1}^{N_E} C_{ij}^{EE} \delta w_j^E \end{aligned} \quad (31)$$

$$\frac{1}{\lambda r^2 \tau} \delta \dot{w}_i^I = (1 - \alpha_I)(\delta w_0^E - \delta w_0^I) + \frac{1}{N_I} \sum_{j=1}^{N_I} C_{ij}^{II} \delta w_j^I \quad (32)$$

$$g_0^E = (w_0^E - w_0^I) \frac{df_-^E(w_0^E)}{dw_0^E} - ((1 + C_0^E)w_0^E - w_0^I) \frac{df_+^E(w_0^E)}{dw_0^E} \quad (33)$$

$$= 2(w_0^E - w_0^I) \frac{df_-^E(w_0^E)}{dw_0^E} \geq 0$$

Let us denote by  $\mathbf{v}_n^X$  the eigenvectors of  $\mathbf{C}^{XX}$  ( $X = E, I$ ) with the corresponding non negative eigenvalue  $N_X C_n^X$ ; note that  $\mathbf{v}_0^X$  is the uniform  $N_X$  dimensional vector. The eigenvectors of the full stability matrix,  $\mathbf{M}$ , which is given by equation  $\frac{d}{dt} \begin{pmatrix} \delta \mathbf{w}^E \\ \delta \mathbf{w}^I \end{pmatrix} = -\mathbf{M} \begin{pmatrix} \delta \mathbf{w}^E \\ \delta \mathbf{w}^I \end{pmatrix}$ , are of the form  $a_1 \begin{pmatrix} \mathbf{v}_0^E \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ \mathbf{v}_0^I \end{pmatrix}$ , and  $\begin{pmatrix} \mathbf{v}_n^E \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{v}_n^I \end{pmatrix}$  for  $n \neq 0$ .

Using the fixed point equations (29)–(30) one can show that the homogeneous solution is always stable to fluctuations in the homogeneous direction. Additionally, similar to the analysis of the previous section, the homogeneous solution is always stable to fluctuations in directions of modifying the inhibition,  $\begin{pmatrix} 0 \\ \mathbf{v}_n^I \end{pmatrix}$ . However, the homogeneous solution is not always stable with respect to fluctuations in non-homogeneous directions of the excitation. This point has been discussed at length in [28,29,42]. Essentially, as the positive feedback of the STDP dynamics of the excitatory synapses becomes strong (i.e., for small  $\mu_E$ ) the homogeneous solution of the excitatory synapses loses its stability and the learning dynamics becomes more sensitive to the correlation structure of its excitatory inputs. Specifically, the eigenvalues of the stability matrix  $\mathbf{M}$  in the directions of non-homogeneous fluctuations of the excitatory synapses,  $\begin{pmatrix} \mathbf{v}_n^E \\ 0 \end{pmatrix}$  ( $n \neq 0$ ), are:  $M_n = g_0^E - C_n^E = 2(w_0^E - w_0^I) \frac{df_-^E(w_0^E)}{dw_0^E} - C_n^E$ . The term  $g_0^E$  has to stabilize non-homogeneous fluctuations in the  $n$ th excitatory direction,  $C_n^E$ . For small  $g_0^E$  and sufficiently large  $C_n^E$  the homogeneous solution will lose its stability. In addition, note that here  $g_0^E$  is proportional to the deviation from the balance  $g_0^E \propto (w_0^E - w_0^I)$ , which is governed by the correlations between the inhibitory inputs  $w_0^E - w_0^I \propto \frac{C_0^I}{1 - \alpha_I + C_0^I}$ . As the correlations between the inhibitory neurons decrease the net input to the postsynaptic cell becomes more balanced and thus the homogeneous solution becomes less stable.

Nevertheless, since the learning dynamics of the inhibitory neurons, equation (28), is only sensitive to the mean *excitatory* input,  $1/N_E \sum_j w_j^E$ , and the directions of instability of the homogeneous solution are only in a heterogenous direction of the excitatory synapses and not the inhibitory, inhibition is still expected to remain uniform, obeying:

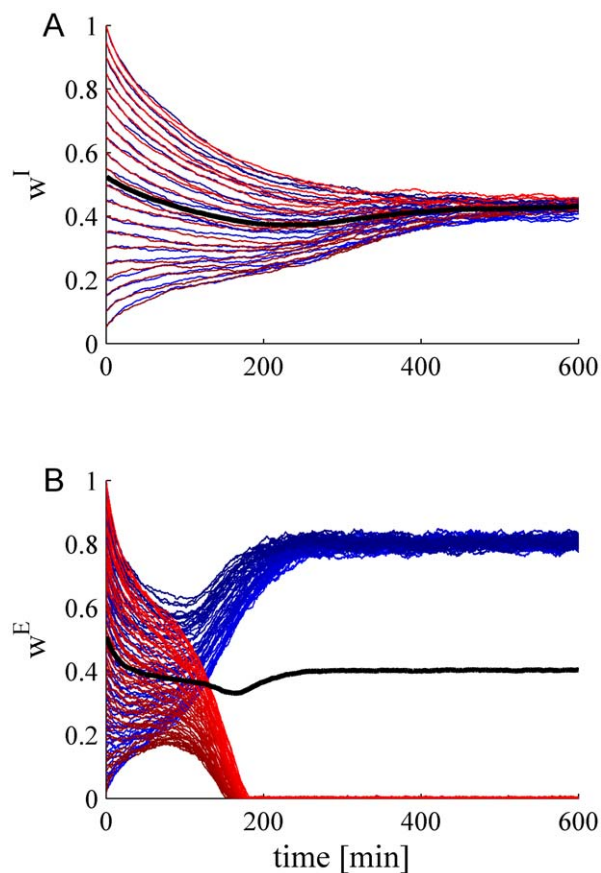
$$w_i^I = \frac{1}{1 + \frac{C_0^I}{1 - \alpha_I}} \frac{1}{N_E} \sum_{j=1}^{N_E} w_j^E \quad (34)$$

Hence, importantly, we find from equation (34) that the negative feedback of inhibitory plasticity works to *balance* the net excitatory input to the cell. In the absence of cross-correlations between inhibitory synapses,  $C_{ij}^{II} \propto \delta_{ij} \Rightarrow C_0^I \propto 1/N_I$ , our model predicts that Hebbian STDP dynamics converge to a complete balance of excitation and inhibition  $\left( \frac{1}{N_E} \sum_{j=1}^{N_E} w_j^E - \frac{1}{N_I} \sum_{j=1}^{N_I} w_j^I \right) \rightarrow 0$ , in



the limit of large  $N_I$ . In the presence of correlations, the net inhibitory input will scale linearly with the net excitatory input (for  $\mu_I=0$ ); however, with a coefficient that is less than one. In this case; i.e., in the presence of correlations, the deviation from exact balance will scale with the magnitude of the fluctuations. Thus, the fluctuations in the feed forward synaptic inputs to the cell will not be negligible relative to their mean.

**Numerical simulations of inhibitory and excitatory STDP.** We simulated the learning dynamics of a population of  $N_I=40$  inhibitory and  $N_E=120$  excitatory feed forward synapses onto a single integrate and fire post synaptic cell. Figure 8 shows an example in which every population (excitatory and inhibitory) is composed of two correlated subpopulations. As can be seen from the figure, although the correlation structure destabilizes the homogeneous solution to the excitatory synapses, Figure 8B, the homogeneous solution to the inhibitory synapses remains stable.

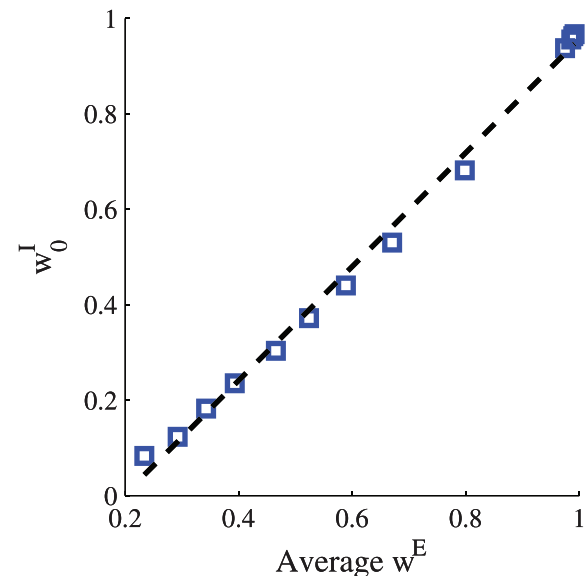


**Figure 8. The learning dynamics of a population of  $N_I=40$  inhibitory and  $N_E=120$  excitatory presynaptic neurons.** Each trace shows the temporal evolution of a single synaptic weight for: **A** Inhibitory population **B** Excitatory population. The firing rate statistics of the presynaptic neurons followed Poisson process statistics with a mean rate of  $r_{pre}=10$  spikes/s. Each group of excitatory and inhibitory populations was composed of two sub groups of equal sizes with a uniform correlation coefficient within each group of  $cc=0.1$  and a zero correlation coefficient between cell pairs from different sub groups (see Methods). The subgroups are distinguished by the different colors red and blue. There were no correlations between excitatory and inhibitory neurons. The thick black line shows the population average of the synaptic weights. Here we used  $\alpha_E=2$ ,  $\alpha_I=0.99$ ,  $\mu_E=\mu_I=0.1$ , and  $\lambda=10^{-4}$ .

doi:10.1371/journal.pcbi.1002334.g008

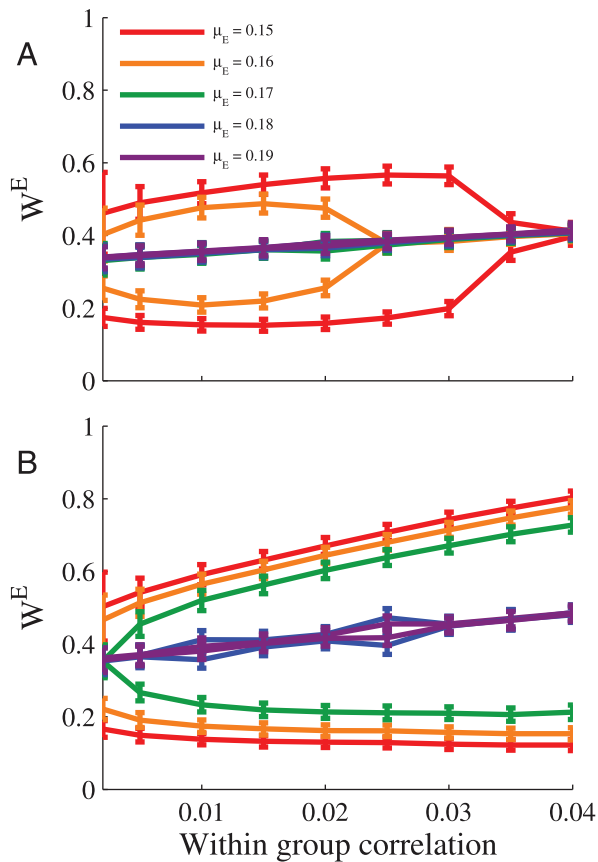
Figure 9 shows the dependence of the homogeneous inhibitory asymptotic weight on the asymptotic population average of the excitatory weights,  $\overline{w^E} = \frac{1}{N_E} \sum_j w_j^E$ . Note that since the integrate and fire neuron is not a 'linear' neuron, as was used for the analysis, an exact linear relation is not expected. For example, the linear regression line (dashed line) crosses the abscissa at positive  $\overline{w^E}$ , which is a manifestation of a threshold effect of the  $I-v$  (the relation between input current and output mean rate) curve of the postsynaptic neuron. Nevertheless, the inhibitory synaptic weights are uniform and increase monotonically with the mean excitatory input.

The sensitivity of excitatory plasticity to the statistical structure of the presynaptic input layer is illustrated in the example of Figure 8B: The homogenous solution loses its stability and the synaptic weights are segregated according to the correlation structure of two competing subgroups. Figure 10 shows the difference in the mean excitatory synaptic weight of each such subgroup as a function of within-group correlation coefficient (between-group correlations were zero). The sensitivity of the learning dynamics can be thought of as the degree in which the correlation structure is expressed in the resultant weights. As  $\mu_E$  is increased the difference between the two subgroups decreases, and the sensitivity vanishes. However, the learning dynamics is more sensitive to the correlation structure of the excitation with inhibitory plasticity (Figure 10B) than without inhibition (Figure 10A). Nevertheless we note that: 1. Although the effect



**Figure 9. Excitation inhibition balance.** The asymptotic value of the homogeneous inhibitory synaptic weight is shown as a function of the asymptotic mean excitatory synaptic weights,  $\overline{w^E}$ . Learning dynamics of a population of  $N_I=40$  inhibitory and  $N_E=120$  excitatory feed forward synapses onto a single integrate and fire post synaptic cell were simulated (see Methods). The firing rate statistics of the presynaptic neurons followed a Poisson process with uniform rates of 10 spikes/s. The firing of different inhibitory neurons were taken to be independent, whereas the excitatory neurons were modeled to have uniform correlations. To obtain different values for the mean excitatory input to the cell we varied the level of the uniform correlations between all presynaptic excitatory neurons  $cc_{exc}=0, 0.005, \dots, 0.06$ , from bottom to top. Here we used the following parameters  $\alpha_E=1.5$ ,  $\alpha_I=0.99$ , and  $\mu_E=\mu_I=0$ . The dashed line shows a linear regression line, for comparison.

doi:10.1371/journal.pcbi.1002334.g009



**Figure 10. Sensitivity of excitatory plasticity.** The learning dynamics of a population of  $N_I$  inhibitory and  $N_E=120$  excitatory presynaptic neurons was simulated. The firing rate statistics of the presynaptic neurons followed Poisson process statistics with a mean rate of  $r_{pre} = 10$  spikes/s. The inhibitory population was homogeneous and without correlations. The excitatory population was composed of two subgroups of equal size with a uniform correlation coefficient within each group,  $c_c$ , and a zero correlation coefficient between cell pairs from different subgroups (see Methods). The figure shows the mean synaptic weight of each excitatory subgroup ( $\pm$  standard deviation) at the end of the learning process, as a function of the within-group correlation strength for different values of  $\mu_E$ . **A** Without inhibition,  $N_I=0$ . **B** With learning of a homogeneous population of  $N_I=40$  inhibitory synapses. Here we used  $\alpha_E=1.5$ ,  $\alpha_I=0.99$ ,  $\mu_I=0$ , and  $\lambda=5\cdot 10^{-4}$ . The points on the graph represents the mean over the last 600 minutes, simulation time was 2400 minutes. doi:10.1371/journal.pcbi.1002334.g010

of increased sensitivity exists, it does not appear to be dramatic in the numerical simulations (note the values of  $\mu_E$ ). 2. The increased sensitivity results from the presence of the inhibition and not necessarily from their learning dynamics (examine the stability analysis of the homogeneous solution, above).

## Discussion

We have studied the computational effect of temporally asymmetric Hebbian plasticity of feed forward inhibition. Hebbian plasticity of inhibition generates negative feedback, in contrast to the positive feedback generated by Hebbian plasticity of excitation. This can be understood by the following intuitive explanation. If the feed forward inhibitory synapse is very strong, then it is less likely that a postsynaptic spike will follow a presynaptic spike. As a

result more pre-post spike pairs will fall on the acausal branch of the STDP learning curve than on the causal branch. This, in turn, will depress the strong synapse. On the other hand, if the synapse is weak, then pre and post spike times will be largely uncorrelated and the STDP dynamics will sample uniformly both branches of the STDP curve with equal probability. If the area under the acausal branch is smaller than the area under the causal branch,  $\alpha < 1$  (for  $\mu=0$ ), the weak synapse will potentiate (the case of  $\alpha > 1$  is not interesting since it reduces all inhibitory synapses to zero).

This negative feedback, in the case of temporally asymmetric Hebbian plasticity and instantaneous correlations as was studied here, implies that the inhibitory synaptic weight distribution is unimodal and that the homogeneous solution of learning a population of *inhibitory* synapses is stable, even when the homogeneous solution for the excitatory synapses loses its stability. However, in our analysis we focused on a simple case where there are no correlations between excitatory and inhibitory neurons. Incorporating such correlations to our model adds a term of the form  $\frac{1}{N_E} \sum_j C_{ij}^{IE} w_j^E$  to the right hand side of equation (28). This term will cause the homogeneous solution for the inhibition to cease to exist when the uniform solution to the excitation loses its stability. Yet the increased stability of the homogeneous solution of the inhibition suggests that the inhibitory feed forward input to a cell is expected to be more broadly tuned than the excitatory input.

In addition we found that inhibitory Hebbian plasticity works to balance the net excitatory inputs of the cell. Two terms govern the homogeneous fixed point of the inhibitory synapses. The first term results from the contribution of the product of the mean firing rates to the pre-post full-correlations (this is a generalization of the fixed point equation of equation (28) to the  $\mu \neq 0$  case):  $-\Delta f(w_0^I) (\bar{w}^E - w_0^I)$ , where  $\bar{w}^E = \frac{1}{N_E} \sum_j w_j^E$ . The second term results from the contribution of covariation in the firing:  $f_+(w_0^I) C_0^I w_0^I$ . The first term, works to balance the net inhibitory and excitatory inputs to the post-synaptic cell. In the absence of correlations between the pre-synaptic inhibitory neurons,  $C_{ij}^{II} \propto \delta_{ij}$ , the contribution of the covariance term will decay as  $1/N_I$ . In this case, neglecting the covariance term, the inhibitory fixed point will balance the net excitatory input:  $(\bar{w}^E - w_0^I) = 0$ , as long as the net excitatory input is not too large:  $\bar{w}^E < \frac{1}{1 + \sqrt{\alpha}}$ . Thus, in the absence of correlations, Hebbian STDP of inhibition will balance the excitation, even for the  $\mu > 0$  case.

In the presence of correlations, for the  $\mu=0$  case, inhibition will balance excitation in the sense that it will scale linearly with it, this balance is skewed towards excitation, as was reported, e.g., in Heiss et al. [9]. For  $\mu > 0$ , inhibition will not scale linearly with excitation. However, the deviation from exact balance (i.e.,  $\bar{w}^E - w_0^I$ ) is expected to scale (not necessarily linearly) with the magnitude of the fluctuations,  $C_0^I$ , for both the  $\mu=0$  and  $\mu > 0$  cases.

The balance has a twofold effect. First, balancing the mean excitation and inhibition inputs to the cell increases the relative contribution of the fluctuations to the cell's response. Note that we find that even if exact balance is not obtained the fluctuations are not expected to be negligible relative to the mean input. Second, the exact balance reduces the stability of the homogeneous solution to the learning dynamics to fluctuations of excitatory synaptic weights in a non-homogeneous direction. Hence, it increases the sensitivity of the STDP dynamics to the structure of the excitatory input.

Our analysis was performed using the framework of simplified models of postsynaptic neural response. Although these simplified

models provide a mere caricature of the rich response dynamics of a neuron to its synaptic inputs, they further our analytical understanding of the possible outcomes of Hebbian inhibitory synaptic plasticity, which cannot be achieved if the neural response dynamics need to be solved numerically. Furthermore, numerical simulations support our analytical results qualitatively beyond the framework of the simplified linear neuron model.

We introduced several additional simplifying assumptions to our model. First, we focused on temporally asymmetric Hebbian STDP for the inhibitory synapses. However, STDP rules may be highly variable and this variability may be manifested by a qualitative difference from our results. For example an anti-Hebbian STDP rule, such as depicted in Figure 1B, may generate positive feedback to the inhibitory plasticity, instead of the negative feedback reported here. Woodin et al. [49] reported a temporally *symmetric* STDP rule in which the learning rule acts as a coincidence detector of post- and pre-spikes, Figure 1D. This learning rule may act to shape the timing of postsynaptic cell spikes by ‘selecting’ the inhibitory inputs that fired in a specific time interval around the excitatory input. Such a mechanism would require a specific temporal correlations structure between the input inhibition and excitation.

Of particular interest is the ‘thalamocortical’ circuit, in which feed forward excitatory inputs arrive directly from thalamus to cortex whereas inhibitory inputs result from a relay of the thalamic input via local (cortical) interneurons. This network architecture manifests in a rich temporal and functional (similarity of preferred stimuli) correlation structure between excitation and inhibition, which may have a significant qualitative effect on learning dynamics. However, here we assumed a very basic correlations structure between the responses of the presynaptic neurons. The investigation of the different effects of various STDP rules and the elaborate spatial and temporal correlations structure of presynaptic neuron responses is beyond the scope of the current paper. Nevertheless, this work suggests a theoretical framework for addressing these issues.

## Methods

### Details of numerical simulations

**The leaky integrate and fire model.** We simulated learning dynamics of feed forward synaptic inputs into a single post synaptic integrate and fire cell. Following Song et al [37] and Gütiğ et al [42] the dynamics of the membrane potential of the postsynaptic cell,  $V(t)$ , obey:

$$C_m \frac{dV}{dt} = I_L + I_{\text{syn}} \quad (35)$$

$$I_L = \frac{1}{R_m} (V_{\text{rest}} - V) \quad (36)$$

$$I_{\text{syn}} = g_E (E_E - V) + g_I (E_I - V) \quad (37)$$

where  $C_m = 200 \text{ pF}$  is the membrane capacitance,  $R_m = 100 \text{ M}\Omega$  is the membrane resistance, the resting potential is  $V_{\text{rest}} = -70 \text{ mV}$ , and the excitatory and inhibitory reversal potentials are  $E_E = 0 \text{ mV}$  and  $E_I = -70 \text{ mV}$ , respectively. An action potential is fired once the membrane potential crosses the firing threshold of  $V_{\text{th}} = -54 \text{ mV}$ , after which the membrane potential is reset to the resting potential.

The synaptic conductances,  $g_E$  and  $g_I$  are given by

$$g_X(t) = g_X^0 \sum_{i=1}^{N_X} w_i^X(t) \sum_j [t - t_j]_+ e^{-(t-t_j)/\tau_X}, \quad X \in \{E, I\} \quad (38)$$

where  $[t]_+ = t$  for  $t > 0$  and 0 otherwise. For convenience we used  $\tau_E = \tau_I = 5 \text{ ms}$ .

For the conductance, we introduced a scaling mechanism on the values used by Gütiğ et al [42] with the following rationale. The synaptic conductances,  $g_X^0$ , were scaled with a scaling factor,  $g_X^s$ , that decreased with the size of the population:  $g_X^0 = \bar{g}_X g_X^s$ . For the excitation Gütiğ et al [42] used  $g_E^0 = 30 \text{ nS}$  for 1000 excitatory synapses, and  $g_I^0 = 50 \text{ nS}$  and 200 inhibitory synapses. We used the same  $\bar{g}_E$  and  $\bar{g}_I$ , and in order to have the same average electric current in different synaptic population sizes we used  $g_E^s = 1000/N_E$  and  $g_I^s = 1000/N_I$  (the fact that there is 1000 in the numerator instead of 200 is explained below). To illustrate the above let us examine the case where the excitation  $N_E$  is taken to be 1000. In this case  $g_E^s$  will be 1 and it will match the working point of [42]. When inhibition  $N_I$  is taken to be 200, the scaling factor,  $g_I^s$  will be 5. The reasoning for this amplification is that in [42] the inhibitory synapses were held constant at their maximum value 1, whereas we are interested in a dynamic range for our inhibitory synapses to enable learning.

**Details of numerical simulations.** The synaptic spike trains to the integrate and fire neuron were simulated by Bernoulli processes (i.e., binary vectors - see below for details on generating these vectors) defined over discrete time bins of duration  $dt = 1 \text{ ms}$ . These vectors are then linearly filtered using a discrete convolution kernel in the shape of  $t \exp(-t/\tau_X)$  with limited length of  $10\tau_X$  (after which this kernel function is zero for all practical purposes) to generate the right side sum of equation (38).

Integration of the synaptic and leak currents in equation (38) to estimate the postsynaptic membrane potential,  $V(t)$ , was done using the Euler method with the same step size of  $dt = 1 \text{ ms}$ .

The firing rate statistics of all presynaptic cell activity throughout all our simulations followed Poisson statistics with stationary mean firing rates of  $r_{\text{pre}} = 10 \text{ spikes/s}$ . For the generation of instantaneously correlated Poisson point processes presynaptic activity (where applicable) we followed Gütiğ et al [42]. Using their defined mechanism to choose and generate the matrices for specific synaptic sub-group guarantees that within this sub-group the spike trains have the desired firing rates and instantaneous pairwise correlation coefficients.

**The learning rate.** For the learning rate,  $\lambda$ , equation (1), we used two approaches as explained below. For purposes of illustration, e.g., Figures 3, 5, and 8, where we were interested in showing the learning dynamics, we used a constant learning rate throughout the entire simulation. In those cases where we were only interested in obtaining the asymptotic value to which the synaptic weights converge we accelerated the learning dynamics by using the following learning rate approach. The simulation code was flexible to support a given vector of  $\lambda$  for each minute unit. Specifically we used the following formula to generate this vector:  $10^{-3} [1 + 15(1-t)^{10}]$ , where  $t \in (0,1]$ , is the ratio between the minute iteration time and the entire simulation time. Examining the behavior of this function shows that it starts from a value of  $16 \times 10^{-3}$  and decays significantly fast towards  $1 \times 10^{-3}$ , leaving the trailing 70% of the simulation with more or less the same learning rate of about  $1 \times 10^{-3}$ .

### Author Contributions

Conceived and designed the experiments: MS. Performed the experiments: YL. Analyzed the data: YL. Contributed reagents/materials/analysis tools: YL. Wrote the paper: MS YL.

## References

- Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci* 13: 334–50.
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18: 3870–96.
- Tsodyks M, Sejnowski T (1995) Rapid switching in Balanced Cortical Network Models. *Network: Comput Neural Syst* 6: 1–14.
- Amit DJ, Brunel N (1997) Model of global spontaneous activity and local structured delay activity during delay periods in the cerebral cortex. *Cereb Cortex* 7: 237–252.
- Amit DJ, Brunel N (1997) Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Comput Neural Syst* 8: 373–404.
- van Vreeswijk C, Sompolinsky H (1998) Chaotic balanced state in a model of cortical circuits. *Neural Comput* 10: 1321–71.
- van Vreeswijk C, Sompolinsky H (1996) Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274: 1724–6.
- Okun M, Lampl I (2008) Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat Neurosci* 11: 535–7.
- Heiss JE, Katz Y, Ganmor E, Lampl I (2008) Shift in the balance between excitation and inhibition during sensory adaptation of S1 neurons. *J Neurosci* 28: 13320–30.
- Hebb DO (1949) *The organization of behavior: a neuropsychological theory*. Wiley, New York.
- Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci* 18: 10464–72.
- Magee JC, Johnston D (1997) A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* 275: 209–13.
- 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* 507: 237–47.
- Zhang LI, Tao HW, Holt CE, Harris WA, Poo M (1998) A critical window for cooperation and competition among developing retinotectal synapses. *Nature* 395: 37–44.
- Markram H, Lobke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275: 213–215.
- Bi GQ, Poo M (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu Rev Neurosci* 24: 139–166.
- Feldman DE (2000) Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex. *Neuron* 27: 1–3.
- Dan Y, Poo MM (2006) Spike timing-dependent plasticity: from synapse to perception. *Physiol Rev* 86: 1033–48.
- Dan Y, Poo MM (2004) Spike timing-dependent plasticity of neural circuits. *Neuron* 44: 23–30.
- Caporale N, Dan Y (2008) Spike timing-dependent plasticity: a Hebbian learning rule. *Annu Rev Neurosci* 31: 25–46.
- Bell CC, Han VZ, Sugawara Y, Grant K (1997) Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* 387: 278–281.
- Zilberter M, Holmgren C, Shemer I, Silberberg G, Grillner S, et al. (2009) Input specificity and dependence of spike timing-dependent plasticity on preceding postsynaptic activity at unitary connections between neocortical layer 2/3 pyramidal cells. *Cereb Cortex* 19: 2308–20.
- Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. *Nat Neurosci* 3: 1178–1183.
- Iglesias J, Eriksson J, Grize F, Tomassini M, Villa A (2005) Dynamics of pruning in simulated large-scale spiking neural networks. *Biosystems* 79: 11–20.
- Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex* 17: 2443–2452.
- Izhikevich EM, Desai NS (2003) Relating STDP to BCM. *Neural Comput* 15: 1511–1523.
- Morrison A, Aertsen A, Diesmann M (2007) Spike-timing dependent plasticity in balanced random networks. *Neural Comput* 19: 1437–1467.
- Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. I. Input selectivity–strengthening correlated input pathways. *Biol Cybern* 101: 81–102.
- Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks. II. Input selectivity–symmetry breaking. *Biol Cybern* 101: 103–114.
- van Rossum MCW, Bi GQ, Turrigiano GG (2000) Stable Hebbian learning from spike timing-dependent plasticity. *J Neurosci* 20: 8812–8821.
- Kempster R, Gerstner W, van Hemmen JL (1999) Hebbian learning and spiking neurons. *Phys Rev E* 59: 4498–4514.
- Kempster R, Gerstner W, van Hemmen JL (2001) Intrinsic stabilization of output rates by spike-based Hebbian learning. *Neural Comput* 12: 2709–2742.
- Kistler WM, van Hemmen JL (2000) Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic action potentials. *Neural Comput* 12: 385–405.
- Zou Q, Destexhe A (2007) Kinetic models of spike-timing dependent plasticity and their functional consequences in detecting correlations. *Biol Cybern* 97: 81–97.
- Roberts P (1999) Computational consequences of temporally asymmetric learning rules. I. Differential Hebbian learning. *J Comput Neurosci* 7: 235–246.
- 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: 919–26.
- Perrinet L, Delorme A, Thorpe S (2001) Network of integrate-and-fire neurons using Rank Order Coding A: how to implement spike timing dependant plasticity. *Neurocomputing* 38: 817–822.
- Delorme A, Perrinet L, Thorpe S (2001) Network of integrate-and-fire neurons using Rank Order Coding B: spike timing dependant plasticity and emergence of orientation selectivity. *Neurocomputing* 38: 539–545.
- Câteau H, Fukai T (2003) A stochastic method to predict the consequence of arbitrary forms of spike-timing-dependent plasticity. *Neural Comput* 15: 597–620.
- Rubin J, Lee D, Sompolinsky H (2001) Equilibrium properties of temporally asymmetric Hebbian plasticity. *Phys Rev Lett* 86: 364–367.
- Gutig R, Aharonov R, Rotter S, Sompolinsky H (2003) Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J Neurosci* 23: 3697–3714.
- Morrison A, Diesmann M, Gerstner W (2008) Phenomenological models of synaptic plasticity based on spike timing. *Biol Cybern* 98: 459–78.
- Burkitt AN, Gilson M, van Hemmen JL (2007) Spike-timing-dependent plasticity for neurons with recurrent connections. *Biol Cybern* 96: 533–46.
- Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks III: Partially connected neurons driven by spontaneous activity. *Biol Cybern* 101: 411–26.
- Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2009) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV: structuring synaptic pathways among recurrent connections. *Biol Cybern* 101: 427–44.
- Gilson M, Burkitt AN, Grayden DB, Thomas DA, van Hemmen JL (2010) Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks V: self-organization schemes and weight dependence. *Biol Cybern* 103: 365–86.
- Holmgren CD, Zilberter Y (2001) Coincident spiking activity induces long-term changes in inhibition of neocortical pyramidal cells. *J Neurosci* 21: 8270–7.
- Woodin MA, Ganguly K, Poo MM (2003) Coincident pre- and postsynaptic activity modifies GABAergic synapses by postsynaptic changes in Cl<sup>-</sup> transporter activity. *Neuron* 39: 807–20.
- Haas JS, Nowotny T, Abarbanel HD (2006) Spike-timing-dependent plasticity of inhibitory synapses in the entorhinal cortex. *J Neurophysiol* 96: 3305–13.
- Akerman CJ, Cline HT (2007) Refining the roles of GABAergic signaling during neural circuit formation. *Trends Neurosci* 30: 382–9.
- Sun QQ (2007) The missing piece in the ‘use it or lose it’ puzzle: is inhibition regulated by activity or does it act on its own accord? *Rev Neurosci* 18: 295–310.
- Foeller E, Feldman DE (2004) Synaptic basis for developmental plasticity in somatosensory cortex. *Curr Opin Neurobiol* 14: 89–95.
- Feldman DE (2009) Synaptic mechanisms for plasticity in neocortex. *Annu Rev Neurosci* 32: 33–55.
- Gaiarsa JL, Caillard O, Ben-Ari Y (2002) Long-term plasticity at GABAergic and glycinergic synapses: mechanisms and functional significance. *Trends Neurosci* 25: 564–70.
- Xu C, Zhao MX, Poo MM, Zhang XH (2008) GABA(B) receptor activation mediates frequency-dependent plasticity of developing GABAergic synapses. *Nat Neurosci* 11: 1410–8.
- Tuckwell HC (1988) *Introduction to theoretical neurobiology: volume 2 nonlinear and stochastic theories*. Cambridge University Press.
- van-Kampen NG (2007) *Stochastic processes in physics and chemistry*. Elsevier 3rd edition.
- Ostojic S, Brunel N, Hakim V (2009) Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *J Neurosci* 29: 10234–53.
- Gerstner W (2000) Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural Comput* 12: 43–89.
- Plesser HE, Gerstner W (2000) Noise in integrate-and-fire neurons: from stochastic input to escape rates. *Neural Comput* 12: 367–84.