Spike-timing dependent plasticity: Input-rate normalization for Poisson Inputs

Zeph Landau[1,2]
Kenneth D. Miller[1−3]
[1]Sloan Center for Theoretical Biology
[2]Keck Center for Integrative Neuroscience,
[3]Departments of Physiology and Otolaryngology,
University of California, San Francisco 94143-0444
email: landau@phy.ucsf.edu, ken@phy.ucsf.edu

DRAFT: For submission to *Neural Computation*

Abstract

1 Introduction

A well-established idea in neuroscience is that many phenomena of learning and development involve Hebb-like (correlation-based) activity-dependent synaptic modification: modification rules that lead to the result that "neurons that fire together, wire together" (*e.g.*, **???**). Recent studies have found that, in many systems, this modification shows a surprisingly precise dependence on the timing of pre- and postsynaptic action potentials. If presynaptic action potentials precede postsynaptic action potentials by no more than about 50 msec, synapses are strengthened (referred to as long-term potentiation, LTP); while if presynaptic action potentials follow postsynaptic spikes over a similar interval, synapses are weakened (long-term depression, LTD). This has been observed at synapses between excitatory neurons in multiple systems: in neocortical slices, in connections between layer 5 pyramidal cells (**?**) and from layer 4 neurons to layer 2/3 pyramidal cells (**?**); in hippocampal slices (**?**) and cell cultures (**?**); and in *in vivo* studies of tadpole tectum (**?**). (A similar precise dependence on spike timing, but with opposite signs of modifications, has also been observed in the cerebellar-like electrosensory lobe of electric fish (**?**); we shall not deal with this case in this paper).

This form of plasticity has been termed spike-timing-dependent plasticity, or STDP (**?**). The results on STDP may be summarized by saying that, at many synapses, plasticity depends crucially on causal relationships: presynaptic neurons that are likely to have causally contributed to a postsynaptic spike are strengthened, while those that could not have causally contributed are weakened.

Recent simulation studies (???) have demonstrated that STDP can provide the normalizing or competitive element that has long been known, both experimentally (?) and theoretically (???), to be required in order for Hebb-like plasticity to explain a variety of phenomena in neural development. While a Hebb-like rule ensures that "neurons that fire together wire together", in the absence of further constraints, multiple, different coactive input patterns could all wire together onto the same postsynaptic neuron. Biologically, however, there appears to be a competitive process in which, when one coactive input pattern succeeds in innervating a given postsynaptic neuron, other coactive patterns are excluded. For example, nearby retinal ganglion cells fire together in both the left eye and the right eye, yet the visual thalamus and the input layer of visual cortex show ocular dominance segregation: individual neurons become innervated by inputs driven by the left eye or the right eye, but not both, a process that has been shown to depend on activity-dependent competition between inputs serving the two eyes (????????).

Figure 1: The STDP modification function. The change in the strength of a synapse due to a single pre- and postsynaptic action potential pair is proportional to $F(\Delta t)$. Δt is the time of the presynaptic spike minus the time of the postsynaptic spike. In this figure, F is expressed as a percentage of the maximum allowed synaptic strength.

STDP obviously incorporates the Hebb-like element of plasticity. The recent simulation studies (**??**) showed that it could simultaneously explain the competitive element, provided that the LTD part of the STDP rule was stronger than the LTP part, in the following sense: a presynaptic input that fires independently of the firing of the postsynaptic cell should show net LTD. These studies also offered an intuitive explanation, as follows.

We suppose the STDP rule is described by a function $F(\Delta t)$ (Fig. 1), as follows: the change in the strength of a synapse due to a single pre- and postsynaptic action potential pair, in which the presynaptic spike occurs a time Δt after the postsynaptic spike, is proportional to $F(\Delta t)$. The requirement that LTD be dominant is expressed as $\int dt\, F(t) < 0$. Let $\rho^{\mathrm{pre}}(t)$ be the spike train of a given presynaptic neuron, and $\rho^{\mathrm{post}}(t)$ the postsynaptic cell's spike train. Then, for a fixed set of weights, the expected change in the synaptic weight between the two cells should be proportional to

$$\int dt\, d\tau\, \rho^{\mathrm{post}}(t)\rho^{\mathrm{pre}}(t+\tau)F(\tau) \propto \int d\tau\, C(\tau)F(\tau)$$

where $C(\tau) \propto \int dt\, \rho^{\mathrm{post}}(t)\rho^{\mathrm{pre}}(t+\tau)$ is the correlation function between the pre- and postsynaptic spike trains. Given $\int dt\, F(t) < 0$, a positive peak in $C(\tau)$ for $\tau < 0$ is required for the synapse to be strengthened; such a positive peak represents the strength of the causal influence of the presynaptic action potential in causing postsynaptic action potentials. When the overall set of weights to the postsynaptic neuron is too strong, the neuron is regularly driven to spike by the mean input it receives, and so the influence of any one input's spikes on postsynaptic firing is small; therefore, all weights show net LTD. This continues until the mean input is insufficient to cause the cell to regularly fire, so that most spiking is driven by positive fluctuations in the input rather than the mean input. In this fluctuation-driven or "balanced" regime, postsynaptic spiking is irregular **??** and individual inputs have a strong influence on postsynaptic spiking. Thus, the total weight to a neuron is driven down to a level in which its spiking occurs in this fluctuation-driven regime. Should the total weight sink too low, the peak in C will grow even larger and all weights will increase, again until the "right" level of input is attained. In the vicinity of this balanced level of input, competition between inputs takes place for control of the postsynaptic cell's spiking takes place.

While this intuitive explanation is satisfying, one would like to be able to say something more precise about when such normalization will occur, the level at which weights will be normalized, and the outcome of the competition. In general, however, computing the correlation function $C(\tau)$, and thus the expected change in the weight of a given input, for a nonlinear spiking model is difficult (and linear models will not show the normalizing behavior). In this paper, we demonstrate a means of attacking the problem that allows us to draw some general conclusions about the behavior of an STDP rule given uncorrelated Poisson inputs. Rather than directly compute the expected change in a given weight, we focus on the difference in the expected changes in any pair of weights. This allows us to prove several general results. First, we show, independent of details of the spiking model, that the STDP rule leads all (but at most one) weights to go to the maximum allowed value or to zero (the minimum allowed value); this behavior was observed in the previous simulations, but not explained. Second, we show, for a simple integrate-and-fire spiking model, that the competition bounds the sum of the firing rates of the inputs that attain maximal weights.

2 Two Models

We shall describe two models: a weight based model and a transmission based model. Both models consider the scenario of n presynaptic cells, $x_1, \ldots x_n$ connected to 1 postsynaptic cell y. We denote by $x_i(t)$ (resp. $y(t)$) the spike train for x_i (resp. y), with $x_i(t) = 1$ if x_i fired at time t and $x_i(t) = 0$ otherwise. We denote by τ^k the firing time of the k^{th} postsynaptic spike.

Weight Based Model.

In the weight based model we allow the synaptic strength between presynaptic and postsynaptic cell to change in a spike timing dependent way. We denote by w_i $(0 \leq w_i \leq 1)$ the weight of the synaptic connection between x_i and y, and $\mathbf{w} = (w_1, w_2, \ldots w_n)$ the corresponding weight vector. To complete the model we will need to specify two things:

a) *A "spiking neuron model" for the postsynaptic cell y.* The firing of y is some function of the input spike trains $x_i(t)$ and the weights w_i.

b) *The spike timing dependent learning rule.* We will denote the learning function by F (*e.g.*, Fig. 1), so that if the postsynaptic cell fires at time τ^k (i.e. $y(\tau^k) = 1$) then w_i changes by

$$\sum_{\substack{t:\ x_i(t) = 1 \\ t > \tau^{k-1}}} F(t - \tau^k)$$

Transmission Based Model

In the transmission based model we allow the probability of spike transmission between presynaptic and postsynaptic cell to change in a spike timing dependent way. Here we denote by q_i $(0 \leq q_i \leq 1)$ the probability that a spike from presynaptic cell x_i is transmitted to the postsynaptic cell y. We denote by $\mathbf{q} = (q_1, q_2, \ldots q_n)$ the corresponding probability vector. We implement transmission by defining a random variable $s_i(t)$ which is independent for all t and has the property that $s_i(t) = 1$ with probability

q_i and is zero otherwise. In this way, transmitted spikes occur at those times t when $x_i(t)$ and $s_i(t)$ are both equal to one. This is equivalent to the condition that $x_i(t)s_i(t) = 1$ and we shall call the function $x_i(t)s_i(t)$ the *transmitted spike train.* To complete the model we again need to specify two things:

a) *A "spiking neuron model" for the postsynaptic cell* y. The firing of y is some function of the transmitted spike trains $x_i(t)s_i(t)$

b) *The spike timing dependent learning rule.* We will denote the learning function by F (*e.g.*, Fig. 1), so that if the postsynaptic time τ^k (i.e. $y(\tau^k) = 1$) then q_i changes by

$$\sum_{\substack{t:\ x_i(t)s_i(t)=1 \\ t>\tau^{k-1}}} F(t-\tau^k)$$

In this way, the change of $\mathbf{q}$ depends only on the transmitted spikes.

We note that if $x_i(t)$ is Poisson with rate r_i then the transmitted spike train is Poisson with rate $r_i q_i$.

For either model we will need to specify the statistics of the presynaptic cell activity, i.e. the statistics of the spike trains $x_i(t)$. Sections **??**,4, and 6 consider when the statistics of $x_i(t)$ are uncorrelated Poisson of rate i. In section 7 we provide results for the case when the statistics of $x_i(t)$ are spatially correlated.

3 Stable Weight and Stable Probability Vectors

What follows in this section is described for the weight based model. Very similar considerations apply to the transmission based models by replacing $\mathbf{w}$ by $\mathbf{q}$ and $x_i(t)$ by $x_i(t)s_i(t)$.

In considering the behavior of the model we are quickly led to consider the probability density function $p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0]dt$ denoting the infinitesimal probability of the presynaptic cell x_i firing at time t and there are no postsynaptic spikes in the interval $[t, 0)$, *given* the vector of weights $\mathbf{w}$ and the knowledge that the postsynaptic cell y fires at time $\tau^k = 0$.

Knowledge of this probability density function yields the expected dynamics of the change of weights. Specifically, if the current weight vector is given by $\mathbf{w}$, then the expected weight change of w_i, which we shall denote by $E_i(\mathbf{w})$, is given by integrating the probability density function against the learning function F:

$$E_i(\mathbf{w}) = \int_{-\infty}^{\infty} F(t)p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0]dt. \quad (1)$$

(The above expression has ignored boundary conditions; the expression for $E_i(\mathbf{w})$ changes for w_i very close to 0 or 1. We address this issue in section 9.) A significant understanding of this model would be achieved if we could find those weight vectors $\mathbf{w}$ about which the system tends to settle. To make this notion precise we define a stable weight vector $\mathbf{w}$.

Definition 1 *A* stable weight $\mathbf{w} = (w_1, w_2, \ldots w_n)$ *shall be a configuration of weights such that:*

1. *The expected change of weight w_i, $E_i(\mathbf{w})$, is zero for all i.*
2. *sign $(E_i(\mathbf{w}+\delta)) = -sign(\delta_i)$, where δ ranges over all vectors with $|\delta_i| \leq \epsilon$ for some small number ϵ.*

Condition 1 ensures that the expected weight change at a stable weight is zero. Condition 2 ensures that near a stable weight, the expected weight change moves the weights towards $\mathbf{w}$. In this way, stable weights are points of attraction fro the system; when the system nears a stable weight, the expected dynamics tend to keep it near teh stable weight for a while. Thus by describing characteristics of the stable weights we will be describing characteristics that often hold for the weights as the evolve in time.

As mentioned at the beginning of this section we can define all the above quantities for the transmission based model by replacing $\mathbf{w}$ by $\mathbf{q}$ and $x_i(t)$ by $x_i(t)s_i(t)$. For completeness, we include the definition of a stable probability vector:

Definition 2 *A* stable probability vector $\mathbf{q} = (q_1, q_2, \ldots q_n)$ *shall be a configuration of probabilities such that:*

1. *The expected change of weight q_i, $E_i(\mathbf{q})$, is zero for all i.*
2. *sign $(E_i(\mathbf{q}+\delta)) = -sign(\delta_i)$, where δ ranges over all vectors with $|\delta_i| \leq \epsilon$ for some small number ϵ.*

In section 4 we will consider the weight based model in the case of uncorrelated Poisson statistics for the input cells. Under only mild assumptions on the spiking neuron model and the spike timing dependent plasticity function (F), we show a stable weight $\mathbf{w}$ necessarily has at most one weight not located near the boundary. In sections 5- 7 we specialize to a simple integrate and fire model, and restrict the shape of the learning function F. We show that the spike timing dependent plasticity rule serves to normalize the input firing rate of the presynaptic cells.

4 Uncorrelated Poisson Input– Weight Based Model

We consider the weight based model described above where the inputs $x_i(t)$ are assumed to be uncorrelated Poisson, with rates $r_i > 0$. We define the following properties of the model:

Property 1 *$F(t)$ is non-negative, non-decreasing from $-\infty$ to 0.*

Property 2 *Consider the two situations:*

a) x_1 fires at time 0,

b) x_2 fires at time 0.

For $w_2 > w_1$, the spiking neuron model is such that for every $t > 0$ the probability that y has spiked in the time interval $[0, t]$ is bigger for situation b) than for situation a). That is

$$\int_0^t p[t = \tau^k, \tau^{k-1} < 0|\mathbf{w}, x_2(0) = 1]dt > \int_0^t p[t = \tau^k, \tau^{k-1} < 0|\mathbf{w}, x_1(0) = 1]dt.$$

This is a reasonable assumption for a spiking neuron model (e.g. it is true of an integrate and fire model) since one would expect the effect on the postsynaptic cell of the connection with the larger weight to be greater.

Theorem 1 *Suppose equation (1) holds exactly for $w_i \in [a, b]$ with $[a, b] \subset [0, 1]$* [1]*. Given a model with Properties 1 and 2, any stable weight $\mathbf{w}$ must have at least $n - 1$ of the w_i lie in the set $[0, a] \cup [b, 1]$. (I.e. at most one weight is not near the boundary).*

Proof: We proceed by contradiction. Thus suppose $\mathbf{w}$ stable and $a < w_1 < w_2 < b$ (we address the possibility of $w_1 = w_2$ at the end). By definition we have $E_i(\mathbf{w}) = 0$ for $i = 1, 2$. Explicitly

$$E_i(\mathbf{w}) = \int_{-\infty}^{\infty} F(t)p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0]dt \tag{2}$$

$$= \int_{-\infty}^{0} F(t)p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0] + \int_{0}^{\infty} F(t)p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0]$$

[1] In fact, because of the boundary conditions, this assumption will not hold precisely. We address this issue in section 9 and show that nevertheless the result of the theorem remains valid

Since the inputs are uncorrelated, $p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w},\ y(0) = 1] = p[x_i(t) = 1|\mathbf{w}]$ for $t > 0$ since the presynaptic cell is not affected by the past history of the postsynaptic cell. Thus the above becomes

$$E_i(\mathbf{w}) = \int_{-\infty}^{0} F(t)p[x_i(t) = 1,\ t > \tau^{k-1}|\mathbf{w}, \tau^k = 0]dt + r_i \int_0^\infty F(t)dt \qquad (3)$$

Set $F_0^\infty = \int_0^\infty F(t)dt$. Applying Bayes rule to the first term we have

$$E_i(\mathbf{w}) = \int_{-\infty}^{0} F(t)p[\tau^k = 0, t > \tau^{k-1}|\mathbf{w}, x_i(t) = 1]p[x_i(t) = 1|\mathbf{w}]/p[\tau^k = 0|\mathbf{w}]dt + r_i F_0^\infty$$

and thus

$$E_i(\mathbf{w}) = \frac{r_i}{R(\mathbf{w})} \int_{-\infty}^{0} F(t)p[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}, x_i(t) = 1]dt + r_i F_0^\infty \qquad (4)$$

where $R(\mathbf{w})$ is the firing rate of the postsynaptic cell. Setting the above equal to zero (the aforementioned condition for a stable weight), and dividing by r_i yields

$$\int_{-\infty}^{0} F(t)p[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}, x_i(t) = 1]dt = -R(\mathbf{w})F_0^\infty,$$

for $i = 1, 2$.

The contradiction stems from the fact that the left hand side depends on i while the right hand side does not. Properties 1 and 2 imply that

$$\int_{-\infty}^{0} F(t)p[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}, x_1(t) = 1]dt < \int_{-\infty}^{0} F(t)p[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}, x_2(t) = 1]dt \qquad (5)$$

(we include the calculation in Appendix ?), a contradiction since both quantities are equal to $-R(\mathbf{w})F_0^\infty$.

Thus $a < w_1 < w_2 < b$ an impossibility for a stable weight $\mathbf{w}$. To complete the proof we address the case $a < w_1 = w_2 < b$. Choose $\delta = (-\delta_1, \delta_2, 0, \ldots 0)$, $\delta_1, \delta_2 > 0$, and set $\mathbf{w}' = \mathbf{w} + \delta$. The stability of $\mathbf{w}$ implies

$$E_1(\mathbf{w}') > 0 \quad \text{and} \quad E_2(\mathbf{w}') < 0. \qquad (6)$$

By equation (5) we have

$$\begin{aligned} &\frac{1}{R(\mathbf{w}')} \int_{-\infty}^{0} F(t)P[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}', x_2(t) = 1]dt + F_0^\infty \\ > \ &\frac{1}{R(\mathbf{w}')} \int_{-\infty}^{0} F(t)P[\tau^k = 0,\ t > \tau^{k-1}|\mathbf{w}', x_1(t) = 1]dt + F_0^\infty \end{aligned}$$

and thus by equation (4),

$$E_2(\mathbf{w}')/r_2 > E_1(\mathbf{w}')/r_1.$$

Since $r_i > 0$, this contradicts equation (6) and the proof is complete. □

5 Integrate and Fire Model

We consider a simple integrate and fire model with membrane potential u. The neuron fires when $u(t)$ is above threshold Θ. After an output spike, the potential is reset to 0. We let τ^k represent the firing time of the kth postsynaptic spike, and for the weight based model (resp. transmission based model) we let t_i^k be the firing time of the kth spike (resp. transmitted spike) of the presynaptic cell x_i. For $\tau^k \leq t \leq \tau^{k+1}$, the change in membrane potential is described as

$$\frac{d}{dt}u(t) = \frac{1}{\tau_m}(V_{\text{rest}} - u(t)) + \sum_{i=1}^{n} \sum_{l:\tau^k \leq t_i^l \leq t} w_i(t)\delta(t - t_i^l).$$

That is, a spike of presynaptic neuron x_i at time t_i^l raises the potential by $w_i(t_i^l)$ where $w_i(t)$ is the value of the weight w_i at time t. Integrating this equation gives

$$u(t) = V_{\text{rest}}(1 - e^{-\frac{1}{\tau_m}(t-\tau^k)}) + \sum_{i=1}^{n} \sum_{l:\tau^k \leq t_i^l \leq t} w_i(t_i^l)e^{-\frac{1}{\tau_m}(t-t_i^l)}, \tag{7}$$

for $\tau^k \leq t \leq \tau^{k+1}$. Furthermore, on the assumption that no two inputs fire at precisely the same time, $u(\tau^{k+1}) \leq \Theta + \max_i(w_i) \leq \Theta + 1$, from which we find

$$\sum_{i=1}^{n} \sum_{l:\tau^k \leq t_i^l \leq t} w_i(t_i^l)e^{-\frac{1}{\tau_m}(t-t_i^l)} \leq (\Theta + 1) - V_{\text{rest}}(1 - e^{-\frac{1}{\tau_m}(t-\tau^k)}) \tag{8}$$

6 Input Rate Normalization

Considering the spike timing dependent learning rule for the postsynaptic firing at τ^k, we have

$$\Delta w_i = \sum_{l:\tau^{k-1} < t_i^l} F(t_i^l - \tau^k). \tag{9}$$

We consider a learning function F constrained as follows.

Property 3 *$F(t)$ is not identically zero and $0 \leq F(t) \leq Ce^{\frac{1}{\tau_m}t}$ for $t \leq 0$. (We make no assumption about $F(t)$, $t > 0$).*

We note that Property 3 can be viewed as a condition requiring that the tail of $F(-t)$ (as $t \to \infty$) decays as fast or faster than the tail of the EPSP in the simple integrate and fire model.

The next result gives an additional necessary condition for a weight vector $\mathbf{w}$ to be stable.

6.1 Weight Based Model

Theorem 2 *Consider a post-synaptic firing model given by equation (7), and a spike timing dependent learning rule with Properties 1 and 3. For a weight vector* $\mathbf{w}$*, denote by* I_ϵ *the set of indices* i *for which* $w_i \in [1-\epsilon, 1]$*. Then for* $\mathbf{w}$ *to be stable, it is necessary that*

$$\sum_{i \in I_\epsilon} r_i \leq \frac{C}{|\int_0^\infty F(t)dt|}(1-\epsilon)^{-1}(\Theta+1).$$

Proof. By (8), for a postsynaptic spike at time τ^k with weight vector $\mathbf{w}$, we have for any ϵ,

$$\sum_{i \in I_\epsilon} \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} w_i(t_i^l) e^{-\frac{1}{\tau_m}(t-t_i^l)} \leq (\Theta+1) - V_{\text{rest}}(1-e^{-\frac{1}{\tau_m}(\tau^k-\tau^{k-1})}) \leq \Theta+1 \tag{10}$$

The above coupled with the fact that $w_i(t_i^l) \geq (1-\epsilon)$ yields

$$\sum_{i \in I_\epsilon} \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} e^{-\frac{1}{\tau_m}(t-t_i^l)} \leq (1-\epsilon)^{-1}(\Theta+1) \tag{11}$$

For $i \in I_\epsilon$, combining (9) and Property 3 about F yields

$$\begin{aligned} \Delta w_i &= \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} F(t_i^l - \tau^k) + \sum_{l:\ t_i^l > \tau^k} F(t_i^l - \tau^k) && (12) \\ &\leq C \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} e^{-\frac{1}{\tau_m}(\tau^k - t_i^l)} + \sum_{l:\ t_i^l > \tau^k} F(t_i^l - \tau^k), && (13) \end{aligned}$$

and thus

$$\begin{aligned} \sum_{i \in I_\epsilon} \Delta w_i &\leq C \sum_{i \in I_\epsilon} \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} e^{-\frac{1}{\tau_m}(\tau^k - t_i^l)} + \sum_{i \in I_\epsilon} \sum_{l:\ t_i^l > \tau^k} F(t_i^l - \tau^k) && (14) \\ &\leq C(1-\epsilon)^{-1}(\Theta+1) + \sum_{i \in I_\epsilon} \sum_{l:\ t_i^l > \tau^k} F(t_i^l - \tau^k) && (15) \end{aligned}$$

We conclude that

$$\sum_{i \in I_\epsilon} E_i(w_i) \leq C(1-\epsilon)^{-1}(\Theta+1) + (\sum_{i \in I_\epsilon} r_i) \int_0^\infty F(t)dt. \tag{16}$$

(The last term is the expected value of

$$\sum_{i \in I_\epsilon} \sum_{l:\ t_i^l > \tau^k} F(t_i^l - \tau^k)$$

since the spike train $x_i(t)$, $t \geq \tau^k$ remains Poisson with rate r_i given the knowledge that the postsynaptic cell fired at time τ^k.) Recall for $\mathbf{w}$ to be stable, it is necessary for $E_i(\mathbf{w}) = 0$ for all i and thus by (16) we arrive at the necessary condition

$$0 \leq C(1-\epsilon)^{-1}(\Theta + 1) + (\sum_{i \in I_\epsilon} r_i) \int_0^\infty F(t)dt. \tag{17}$$

Property 3 ensures $\int_{-\infty}^0 F(t)dt > 0$. Thus the first term of the right side of equation (3) is positive. Since $E_i(\mathbf{w}) = 0$ we conclude (again from 3) that $\int_0^\infty F(t)dt < 0$ and (17) becomes

$$\sum_{i \in I_\epsilon} r_i \leq \frac{C}{|\int_0^\infty F(t)dt|}(1-\epsilon)^{-1}(\Theta + 1).$$

□

6.2 Transmission Based Model

A similar result holds in the transmission based model.

Theorem 3 *Consider a post-synaptic firing model given by equation (7), and a spike timing dependent learning rule with Properties 1 and 3. For a probability vector* $\mathbf{q}$ *to be stable, it is necessary that*

$$\sum_i r_i q_i \leq \frac{C}{|\int_0^\infty F(t)dt|}(\Theta + 1).$$

Proof. By (8), for a postsynaptic spike at time τ^k with probability vector $\mathbf{q}$, we have

$$\sum_i \sum_{l:\ \tau^{k-1} \leq t_i^l \leq \tau^k} e^{-\frac{1}{\tau_m}(t - t_i^l)} \leq (\Theta + 1) - V_{\text{rest}}(1 - e^{-\frac{1}{\tau_m}(\tau^k - \tau^{k-1})}) \leq \Theta + 1 \tag{18}$$

The parallel calculations to equations (12) - (17) for the transmission based model yields the stated result. □

7 Spatially Correlated Inputs

We turn to the case of spatially correlated inputs; i.e. situations where groups of input are more likely to fire simultaneously. Specifically we consider the following:

Definition 3 *Let S^l, $l = 1, \dots, k$ be subsets of $\{1, 2, \dots n\}$ with $\cup_l S^l = \{1, 2, \dots n\}$, r_l positive real numbers. Let $z^l(t)$ be a Poisson spike train of rate r^l. By a spatially correlated set of inputs we shall mean a set of input spike trains*

$$x_i(t) = \sum_{j : i \in S^j} z^j(t).$$

The idea here is that each set S^l represents a common correlation among the inputs x_i, $i \in S^l$. We imagine that the occurence of certain events cause the inputs of S^l to fire together. This might describe the behavior of a system wher the inputs x_i, $i \in S^l$ when viewed as postsynaptic cells themselves, shared some common presynaptic input. We not that the S^l need not be disjoint, i.e. that a particular input x_i may be spatially correlated to several different sets of inputs (i.e. $i \in S^l$ for more than one l).

Notice when we let $S^l = \{l\}$, $1 \le l \le n$, the above spacially correlated set of inputs is just the uncorrelated model discussed up till now.

For spatially correlated sets of inputs we have similar input normalization results.

7.1 Weight Based Model

Theorem 4 *Consider a weight based model with spatially correlated inputs, a post-synaptic firing model given by equation (7), and a spike timing dependent learning rule with Properties 1 and 3. For a weight vector $\mathbf{w}$, denote by I_ϵ the set of indices i for which $w_i \in [1 - \epsilon, 1]$. Then for $\mathbf{w}$ to be stable, it is necessary that*

$$\sum_{i \in I_\epsilon} r_i \le \frac{C}{|\int_0^\infty F(t)dt|}(1 - \epsilon)^{-1}(\Theta + l),$$

where l is the average number of precisely coactive inputs that trigger the post-synaptic spike.

Proof. The proof is identical to that of Theorem 2 with the exception that the $\Theta + 1$ that appears in equation (16) must be replaced by $\Theta + l$. The reason is that for uncorrelated inputs, we can assume that no two inputs fire at precisely the same time and therefore we can bound the total potential before firing to $\Theta + 1$ (see equation (8)). For correlated inputs, this is no longer the case and our bound for the total potential becomes

$\Theta +$ (the average number of precisely coactive inputs that trigger a postsynaptic spike.

The result then follows. □

7.2 Transmission Based Model

Theorem 5 *Consider a transmission based model with spatially correlated inputs, a post-synaptic firing model given by equation (7), and a spike timing dependent learning rule with Properties 1 and 3. For a probability vector* $\mathbf{q}$*to be stable, it is necessary that*

$$\sum_i r_i q_i \leq \frac{C}{|\int_0^\infty F(t)dt|}(\Theta + l).$$

where l *is the average number of precisely coactive inputs that trigger the post-synaptic spike.*

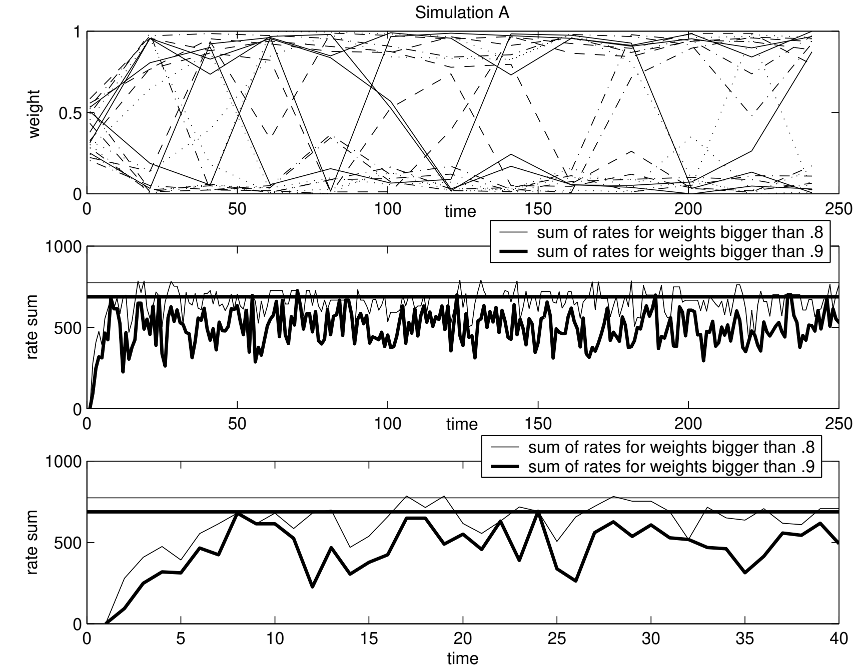
Proof The proof is the same as Theorem 3 with $\Theta + 1$ replaced by $\Theta + l$ (see the proof of Theorem 4 for the reason for this replacement). □

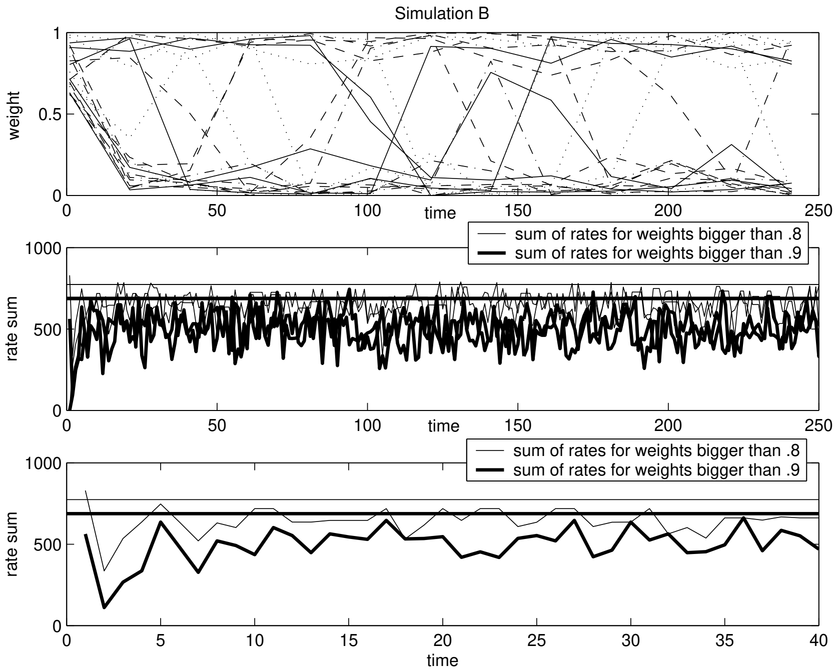
8 Simulations

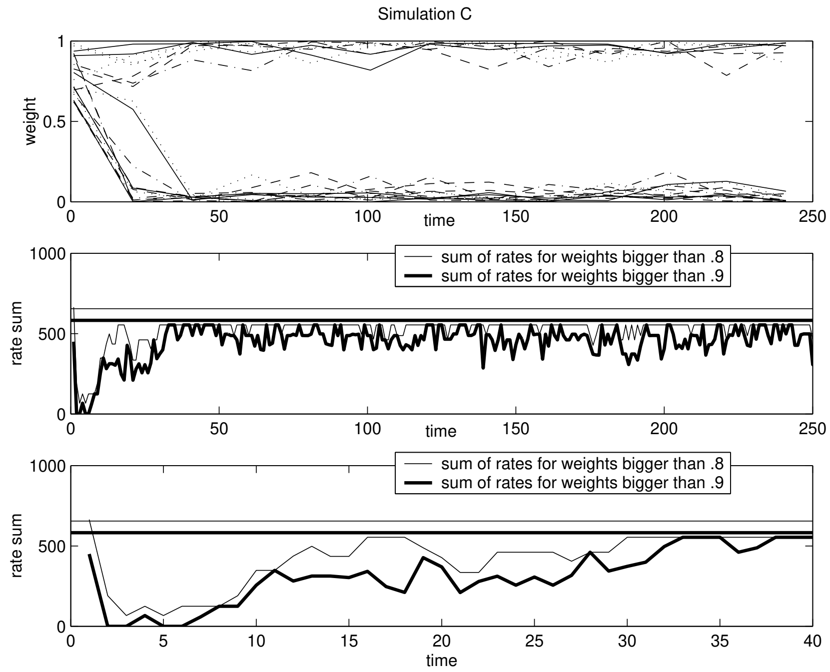
We now provide some simulations that exemplify the behavior described by theorems 1 and 2.

INSERT SIMULATION MODEL DESCRIPTION.

We present the results from three simulations; these results were typical across a range of parameters. Each simulation had 20 inputs. Simulation A had the mean rate of its inputs 80, the mean weight of its inputs .4 and the rate noise 24. Simulation B had the mean rate of its inputs at 100, the mean weights at the default of .8 and the rate noise at 30. Simulation C had the mean rate of its inputs at 80, the mean weights at the default of .8, and the rate noise at 24. In addition, simulation C had the resting potential of 0 (rather than the default -10) and a firing threshold of 10 (rather than the default of 2). For each of the three simulations, three graphs are presented. The first tracks each of the 20 weights over time (only four linestyles are used so that each linestyle is used for 5 different weights). The second plots the sum of the rates of those weights above .8 (thin line) and above .9 (thick line) over time. In addition, the second graph provides the straight line upper bound (both thin and thick) predicted by theorem 2. Finally the third graph is just an amplification of the initial time steps of the second graph showing the initial movement towards stabilization.


Simulation A
weight
time
rate sum
sum of rates for weights bigger than .8
sum of rates for weights bigger than .9
rate sum
sum of rates for weights bigger than .8
sum of rates for weights bigger than .9
time


Simulation B
weight
time
rate sum
sum of rates for weights bigger than .8
sum of rates for weights bigger than .9
rate sum
sum of rates for weights bigger than .8
sum of rates for weights bigger than .9
time



Theorems 1 and 2 discuss necessary conditions for stable weight configurations. As mentioned in section 3, we expect the dynamics to be "drawn" to stable configurations. In other words, when we look at the system over time, we expect it to often be near stable configurations and thus at those times to have characteristics similar to those of the stable configurations.

The ramifications of theorem 1

Theorem 1 tells us that the system will be drawn towards positions where at most one weight is not near the boundary (zero or one). This behavior is clearly demonstrated by the top graph of each simulation as each weight spends most of its time near zero or near one. In addition, when a weight does change from one extreme to the other, it does so quickly–not spending an extended amount of time hovering between zero and one.

The ramifications of theorem 2

Theorem 2 gives upper bounds for stable weights, for the sum of the rates of those weights above $(1-\epsilon)$. The middle graph in all three examples plots this weight sum for $(1-\epsilon) = .8$ and $(1-\epsilon) = .9$. The straight lines give the predicted upper bound. We see that the weight sum spends virtually no time above the upper bound and that furthermore, for these examples, this upperbound is reasonably "tight" in the sense that the actual values hover just below the bounds.

Theorem 2 is exemplified in a second way in the top graphs of each simulation. Notice how the number of weights near one does not vary much over time: an upward moving weight form zero towards one is almost always matched by a downward moving weight from one to zero. Theorem 2 bounds the sum of the rates for those weights above $(1-\epsilon)$; this bound is well approximated by the average rate times the number of weights above $(1-\epsilon)$. Thus we can think of theorem 2 as providing a loose upper bound for the number of weights above $(1-\epsilon)$.

9 Boundary effects and the ambiguity of the learning rule.

In both models, the changing parameters (w_i in the weight based model, q_i in the transmission based models) are bounded to the interval $[0, 1]$. Since this bounding is not a consequence of the learning rule, it will have an effect on the behavior of the model. Furthermore, there are several different ways to impose these bounds, each with potentially different effects on the model's behavior.

We discuss these issues here, showing that the essential effect on the model is mild in terms of the results described in this paper.

It will be helpful for the following to define the bounding function $b : (-\infty, \infty) \to [0, 1]$,

$$b(x) = \begin{cases} x & 0 \le x \le 1 \\ 0 & x \le 0 \\ 1 & x \ge 1 \end{cases}$$

We illustrate the essential issue with a simplified example. Suppose w_1 is very close to one, i.e. $w_1 = 1 - \delta$ for δ small. In addition, suppose $\tau^k = 0$ and x_1 has only fired at the three times $t_1, t_2, t_3,\ t_i \ge \tau^{k-1}$ with $F(t_1) = \delta/2$, $F(t_2) = 2\delta$, $F(t_3) = -\delta$. We present several possible ways of imposing the boundary constraint:

1. Compute Δw_1, add it to w_1 and then enforce the boundary conditions. In the example $\Delta w_1 = 3\delta/2$ and thus $new(w_1) = 1$.
2. Break $\sum F(t_i)$ into some chosen set of partial sums $S_1, S_2 \ldots$ ($\sum_i S_i = \Delta w_i$) and then set

   $$new(w_1) = \ldots b(b(b(w_1 + S_1) + S_2) + S_3) + \ldots$$

   In the above example

   (a) If $S_1 = F(t_1) + F(t_2)$, $S_2 = F(t_3)$ or if $S_i = F(t_i),\ i = 1, 2, 3$ then $new(w_i) = 1 - \delta$.

   (b) If $S_1 = F(t_2)$, $S_2 = F(t_3)$, $S_3 = F(t_1)$ then $new(w_1) = 1 - \delta/2$.

3. Step through time, for each t with $x_i(t) = 1$ set

$$new(w_1) = b(w_1 + \sum_{l:\tau^l < t} F(\tau^l - t)),$$

and for each $t = \tau^l$ set

$$new(w_1) = b(\sum_{t:\tau^{l-1} \leq t' \leq \tau^l} F(\tau^l - t'))$$

The above illustrates an inherent ambiguity in how to enforce the boundary conditions. For the simulations, we enforce the boundary conditions as in 3. above. In general, any choice affects equation (2). If w_i is close to 1 then equation (2) slightly overestimates $E_i(\mathbf{w})$ while if w_i is close to 0 then equation (2) slightly underestimates $E_i(\mathbf{w})$. In general however, the total change induced by a single spike is small relative to the interval $[0, 1]$. As such, the boundary constraint will only have a significant effect on the behavior of the model very near the boundary values of 0 and 1.

The affect on the previous results are:

> *Theorem 1:* Though the assumption that equation (1) holds for an interval $[a, b]$ is false, for a and b away from the boundary the equation is a very good approximation and *what should we say?*.
>
> *Theorems 2 and 4:* For Theorem 2 the important inequality appears in equation (11). For values of w_i near 1, the boundary effects only make this inequality more true. Thus the proof will only break down for ϵ close to 1 (w_i close to 0) where the bound is not useful anyway. The same is true for Theorem 4.
>
> *Theorems 3 and 5:* The proofs of these theorems are unaffected by the boundary considerations.

10 Discussion

Summarize results.

Discuss case in which STDP modifies release probability rather than maximum synaptic strength – actually, we now have results on this, include these in results as well.

Discuss connection to subtractive contraints in non-STDP Hebb rules – also shows behavior of all synapses going to max or min.

Discuss how if all weights started about equal that the behavior would point to the more correlated groups "winning".

Future directions: would like especially to say something more about the case of correlated inputs, obviously. Also, how good is our bound? Can we say anything about minimal as well as maximal total strength?

11 Appendix

Lemma 1 *Let $G(t)$ and $H(t)$ be non-negative differentiable functions from* 0 *to* ∞. *Suppose $H(t) > 0$ for $t \in (0, \infty)$, $H(0) = 0$, and $G(t)$ is not a constant function and is non-increasing from* 0 *to* ∞. *Then*

$$\int_0^\infty G(t)H'(t)dt > 0.$$

Proof. Integration by parts yields

$$\int_0^\infty G(t)H'(t)dt = G(t)H(t)|_{t=0}^\infty - \int_0^\infty G'(t)H(t)dt.$$

The first term on the right hand side is non-negative since $H(0) = 0$ and $\lim_{t\to\infty} G(t)$, $\lim_{t\to\infty} H(t)$ are both non-negative. $H(t) > 0$ and $G'(t) \leq 0$ for all $t \in (0, \infty)$ imply that the second integral is negative and the result follows. □

By setting $G(t) = F(-t)$, $H'(t) = p[t = \tau^k,\ \tau^{k-1} < 0|\mathbf{w}, x_2(0) = 1] - p[t = \tau^k, /\tau^{k-1} < 0|\mathbf{w}, x_1(0) = 1]$ the above Lemma yields:

$$\int_0^\infty F(-t)p[t = \tau^k,\ \tau^{k-1} < 0|\mathbf{w}, x_2(0) = 1] - p[t = \tau^k, /\tau^{k-1} < 0|\mathbf{w}, x_1(0) = 1] > 0.$$

Equation (5) follows from substituting $p[\tau^k = 0, \tau^{k-1} < -t|\mathbf{w}, x_i(-t) = 0]$ for $p[t = \tau^k,\ \tau^{k-1} < 0|\mathbf{w}, x_i(0) = 1]$ above and performing a change of variable $t \to -t$.