Toward a Biophysically Plausible Bidirectional Hebbian Rule

Grzywacz, Norberto M.; Burgi, Pierre-Yves

How to cite

GRZYWACZ, Norberto M., BURGI, Pierre-Yves. Toward a Biophysically Plausible Bidirectional Hebbian Rule. In: Neural computation, 1998, vol. 10, n° 3, p. 499–520. doi: 10.1162/089976698300017629

This publication URL: https://archive-ouverte.unige.ch/unige:17433
Publication DOI: 10.1162/089976698300017629

ARTICLE Communicated by Ralph Linsker

Toward a Biophysically Plausible Bidirectional Hebbian Rule

Norberto M. Grzywacz
Pierre-Yves Burgi
Smith-Kettlewell Eye Research Institute, San Francisco, CA 94115, U.S.A.

Although the commonly used quadratic Hebbian–anti-Hebbian rules lead to successful models of plasticity and learning, they are inconsistent with neurophysiology. Other rules, more physiologically plausible, fail to specify the biological mechanism of bidirectionality and the biological mechanism that prevents synapses from changing from excitatory to inhibitory, and vice versa. We developed a synaptic bidirectional Hebbian rule that does not suffer from these problems. This rule was compared with physiological homosynaptic conditions in the hippocampus, with the results indicating the consistency of this rule with long-term potentiation (LTP) and long-term depression (LTD) phenomenologies. The phenomenologies considered included the reversible dynamics of LTP and LTD and the effects of *N*-methyl-*D*-aspartate blockers and phosphatase inhibitors.

1 Introduction

Hebb (1949) postulated a synaptic mechanism for learning where increases in synaptic strength are dependent on concurrent activity in pre- and postsynaptic cells. Stent (1973) further proposed that negative correlation in pre- and postsynaptic activity reduces synaptic strength.

One of the most successful mathematical formulations of this Hebb-Stent bidirectional rule is the covariance rule (Sejnowski, 1977). This and related rules are the cornerstone of many proposed unsupervised learning schemes (von der Malsburg, 1973; Oja, 1982; Linsker, 1986; Miller, 1994; Burgi & Grzywacz, 1997). In this rule, changes in synaptic weights are proportional to the covariance between the firing rates of the pre- and postsynaptic cells. There is now neurophysiological evidence supporting a form of synaptic plasticity that is governed by something like the covariance rule (Stanton & Sejnowski, 1989; Xie, Berger, & Barrionuevo, 1992; Dudek & Bear, 1993; Frégnac, Burke, Smith, & Friedlander, 1994). However, application of this rule leads to two major physiological inconsistencies. First, the synaptic weight of any given synapse can become positive or negative. This situation is biologically unlikely because synapses are either excitatory or inhibitory.[1]

[1] A possible exception is the $GABA_A$ synapses, which are excitatory early in devel-

Second, the covariance rule predicts synaptic potentiation in the absence of pre- and postsynaptic activities. Computationally, potentiation without activity may not make sense, since quiescence in the brain most often means "no neural data" rather than "neural evidence for" jointly weak activity in two cells. And neurophysiological experiments using low-level presynaptic activity in conjunction with postsynaptic hyperpolarization indicate long-term depression (LTD) rather than long-term potentiation (LTP) (Xie et al., 1992).

A more biophysically plausible bidirectional rule has been proposed: the BCM rule (Bienenstock, Cooper, & Munro, 1982; Yang & Faber, 1991; Intrator & Cooper, 1992). This rule proposes the existence of an activity-dependent threshold, such that presynaptic activities above and below this threshold lead to LTP and LTD, respectively. Although the BCM rule solves the second problem above, without modification, this rule cannot prevent synapses from changing polarity. Imposing hard bounds on the synaptic weights is possible, but this solution would leave unspecified what biological process sets these bounds. The BCM rule does not even specify what biological mechanism controls the varying threshold. This criticism should not be taken as an attack on the BCM rule, since it was designed as a computational, not biophysical, rule. The criticism only underscores the necessity of a more biophysical model. In section 5, we even point out that our rule is in some sense a biophysically plausible instantiation of the BCM rule.

The main goal of this article is to present a new physiological bidirectional Hebbian rule, which has bounded nonnegative synaptic weights and does not change these weights in the absence of pre- and postsynaptic activity. The intention is not to model the exact biophysical details of synaptic plasticity, but to focus on mechanisms of synaptic stability during quiescence and bidirectionality. To achieve this, we restricted the underlying biophysics describing the new Hebbian mechanism to three simple constraints: (1) the overall response of the postsynaptic cell is the linear weighted sum of all presynaptic activities; (2) this response produces a feedback messenger, which at a given synapse is proportional to the local concentration of a postsynaptic agent; (3) enzymes controlling synaptic strength are activated at the presynaptic site by the evoked feedback messenger. This model's essence is thus in the details of these enzymes, namely, constraint 3. The other constraints were chosen for computational simplicity to be linear (section 5). Furthermore, whether the enzymatic complex in constraint 3 is pre- or postsynaptic is irrelevant (section 5).

To evaluate the neurophysiological and computational properties of the

opment but become inhibitory later on (Cherubini, Rovira, Gaiarsa, Corradetti, & Ari, 1990; Walton, Schaffner, & Barker, 1993). However, even these synapses do not change polarity more than once, the changes are always in the same direction (from excitatory to inhibitory), and they always occur regardless of the network. Therefore, these changes do not have the flexibility required by the covariance rule.

new rule, we apply a set of physiological and pharmacological tests to our model. Here, the tests are for one synapse, corresponding to the case of homosynaptic modifications.[2] The results of these tests are compared qualitatively with similar experimental conditions in the hippocampus. The new rule and this comparison appeared in abstract form elsewhere (Grzywacz & Burgi, 1994).

2 Model

In this model (see Figure 1A), a feedback messenger activates presynaptic enzymes controlling synaptic strength. These enzymes work by controlling the active and inactive states of a gating molecule. For computational simplicity, it is assumed that the synaptic weight (w) is proportional to the active state's concentrations. (For ease of understanding, in Figure 1A, this active state is labeled w.) A possible interpretation of the active state is that it facilitates synaptic-vesicle docking (more on this in section 3). Let us suppose that at a given synapse, the maximal possible synaptic weight (proportional to the maximal concentration of the gating molecules) is w_0, that is, the pool of gating molecules is finite. In this case, $w_0 = w + w^*$, where w^* is proportional to the inactive state's concentration. Let us denote the presynaptic activities and synaptic strengths (transmitter release per unit of presynaptic activity) of a set of synapses onto a single postsynaptic cell by I_j and w_j, respectively, where $1 \leq j \leq p$ and p is the number of synapses. The overall response of the postsynaptic cell is

$$R = \sum_{j=1}^{p} w_j I_j \tag{2.1}$$

(an assumption shared by most other models; for a review, see Churchland & Sejnowski, 1992). We assume the concentration of a postsynaptic agent (C) at a given postsynaptic site to be proportional to the multiplication of R by the concentration of neurotransmitter ($w_j I_j$) in the synaptic cleft of that synapse. Because, for simplicity, it is also assumed that the concentration of feedback messenger ($[m_j]$) produced by synaptic activity is proportional to the local concentration of C, one gets

$$[m_j] = \alpha R\, w_j I_j, \tag{2.2}$$

where α is a constant.

To obey one of the main constraints of the model—that no synaptic change should occur in the absence of activity—we postulate a molecular

[2] Heterosynaptic tests using a simplified version of the rule appear elsewhere (Burgi & Grzywacz, in press).

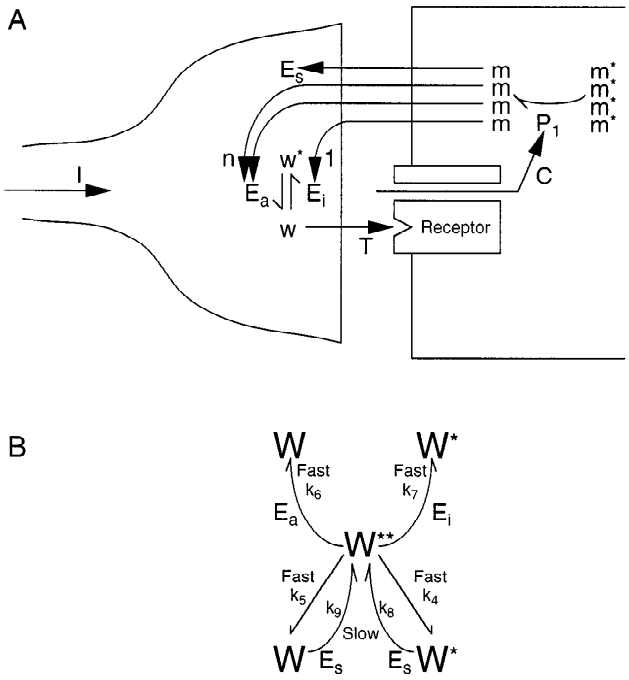

Figure 1: Schematic representation of the new model. (A) The presynaptic input (I) causes the release of the transmitter (glutamate in the hippocampus), whose quantity (T) depends on a particle that determines synaptic strength (w). Specifically, T is proportional to the multiplication of I and w. A fraction of the T molecules of the transmitter attaches to special receptors (NMDA receptors in the hippocampus), which, if the postsynaptic site is depolarized, allow the entrance of an agent C (Ca^{2+} in the hippocampus) into it. This agent activates an enzyme (P_1), which produces a messenger (m) that can propagate to the presynaptic terminal and activate an enzymatic switch (E_s). When this switch is on, it allows the flow between the active (quantified by w) and inactive (quantified by w^*) states of the particle determining synaptic strength. This flow is such that low concentrations of m cause the inactivation of the particle (with a stoichiometry of 1) through enzyme E_i, whereas high concentrations of m cause the activation of the particle (with a stoichiometry of n) through enzyme E_a. (B) The switch E_s promotes the transitions $w,w^* \rightharpoonup w^{**}$, where w^{**} is a transient metastate, while the transitions $w^{**} \rightharpoonup w,w^*$ depend on the messenger through the inactivating (E_i) and activating (E_a) enzymes. Rates of the thermal decays are indicated by k_4 to k_5.

switch (see Figure 1A; see also Lisman, 1989). It would operate presynaptically and would have to be turned on by the messenger to allow for synaptic changes. A possible implementation of this switch is

$$E^*_{s,j} + m_j + E_{s,j} \xrightarrow{k_1} m_j + 2E_{s,j}, \tag{2.3}$$

with the thermic reactions

$$E_{s,j} \underset{k_3}{\overset{k_2}{\rightleftharpoons}} E^*_{s,j}, \tag{2.4}$$

where $E_{s,j}$ and $E^*_{s,j}$ are the active and inactive states of the switch, respectively, and k_1, k_2, and k_3 are the rate constants of the reactions. What equation 2.3 expresses is an autocatalytic switch. The more active switch molecules there are (higher concentrations of $E_{s,j}$), the easier is to activate switch molecules. This is similar to Lisman's self-phosphorylating enzymes (Lisman, 1989), which become active when phosphorylated. Lisman's notion was that when activated, these enzymes would tend to become all phosphorylated due to their positive-feedback reaction. Consequently, the behavior of such enzymes would tend to be all or none, ideal for a switch. To Lisman's switch notion, we add only the thermal reactions in equation 2.4 to allow the switch to be turned off. (For it to be rapidly turned off, one must assume $k_3 \ll k_2$.) From the chemical equations 2.3 and 2.4, one can express the variations of the concentrations, $[E_{s,j}]$, of the switches as

$$\frac{d[E_{s,j}]}{dt} = -k_2[E_{s,j}] + ([E_{s,0}] - [E_{s,j}])(k_3 + k_1[m_j][E_{s,j}]), \tag{2.5}$$

where $[E_{s,0}] = [E_{s,j}] + [E^*_{s,j}]$ is the maximal concentration of active switches.

The feedback messenger activates two presynaptic enzymes, which control synaptic strength by shifting the balance of the gating molecule toward either its inactive or active states. One enzyme, E_i, reduces synaptic strength in proportion to its current concentration and is activated by the feedback messenger with a stoichiometry of 1. The other enzyme, E_a, requires $n > 1$ molecules of the messenger to increase synaptic strength up to a maximal value. The postsynaptic agent C can thus control LTP and LTD through the different stoichiometries. For high C, the large-stoichiometry enzyme dominates, causing LTP, whereas for low C, the other enzyme dominates, causing LTD.

The switch allows back-and-forth transitions between the active and inactive states of the gating molecule, depending on the amount of messenger. To achieve this goal, we propose that the switch promotes the transitions from these states to a transient unstable metastate (see Figure 1B). This metastate decays enzymatically back to the active and inactive states with rate constants that depend on the messenger (see Figure 1B). By writing the equations corresponding to the enzymatic reactions schematically shown

in Figure 1B and by assuming that the decay from the metastate is much faster than the arrival to it, one can express the variation of w_j as follows (see the appendix for details):

$$\frac{dw_j}{dt} = \frac{-k_9[E_{s,j}](k_4+k_7\beta[m_j])w_j+k_8[E_{s,j}](k_5+k_6\gamma[m_j]^n)(w_0-w_j)}{k_4+k_5+k_6\gamma[m_j]^n+k_7\beta[m_j]}, \quad (2.6)$$

where k_4 and k_5 are rate constants of thermic reactions; k_6, k_7, k_8, and k_9 are rates of enzymatic reactions (see Figure 1B); and β and γ are constants linking m_j to the enzymes (see the appendix).

Because equations 2.1 and 2.2 have no dynamics and because equation 2.5 essentially represents an on-off switch, the important dynamics are in equation 2.6. The LTP-LTD phenomenology is introduced explicitly in the E_i and E_a reactions modeled by this equation. It says that when the switch is on ($E_{s,j}$ significantly larger than zero), weight changes can occur and depend on the present weight (w_j) and the feedback messenger ($[m_j]$). If the weight is near zero ($w_j \approx 0$) or near saturation ($w_j \approx w_0$), then the synaptic weight tends to increase or decrease, respectively. For large amounts of feedback messenger ($k_7\beta[m_j] \ll k_6\gamma[m_j]^n$), the tendency will be for synaptic weight to increase. For intermediate amounts ($k_7\beta[m_j] \gg k_6\gamma[m_j]^n$), weights will tend to fall. Finally, for low amounts, the switch will turn off, preventing weight changes. The denominator of equation 2.6 imposes bounds on how fast synaptic changes can occur.

3 Methods

Model simulations were performed on a SPARC II workstation, using C++. To solve the differential equations, we used the fifth-order Runge-Kutta method with adaptive step size control (Press, Teukolsky, Vetterling, and & Flannery, 1992). The equations' initial conditions were $w_j = w_0/2$ (the middle of w_j's range) and $E_{s,j} = E_{s,0}k_3/(k_2 + k_3)$, (the steady-state value of $E_{s,j}$ in the absence of inputs).

The simulations involved solving numerically equations 2.1, 2.2, 2.5, and 2.6. A direct count revealed that these equations have 16 parameters: p, α, β, γ, k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9, $[E_{s,0}]$, w_0, and n. However, not all the parameters were independent. By substituting equation 2.2 for $[m_i]$ in equation 2.5, one sees that α appears multiplied by k_1. Similarly, by substituting equation 2.2 for $[m_i]$ in equation 2.6, one sees that α and α^n always appear multiplied by $k_6\beta$ and $k_7\gamma$, respectively. Moreover, β and γ do not appear in any other form. Therefore, without loss of generality, we set $\alpha = \beta = \gamma = 1$, effectively eliminating these three parameters. For related reasons, the parameters $[E_{s,0}]$ and w_0 could also be set to 1, which is akin to forcing $[E_{s,j}]$ and w_j to lie between 0 and 1. Because $[E_{s,i}]$ always appeared multiplied by

k_8 and k_9 in equation 2.6, the absolute amplitude of this variable is irrelevant. As for w_j, its absolute value is irrelevant, because plots of synaptic modification are of relative, rather than absolute, modulations. Finally, k_4, k_5, k_6, and k_7 are not mutually independent, since inspection of equation 2.6 (with $[E_{s,0}] = 1$) reveals that k_4, k_5, $k_7[m_i]$, and $k_6[m_i]^n$ all have the same physical dimension, which is arbitrary, since it cancels in a ratio. Hence, without loss of generality, we set $k_6 = 1$.

The only parameters that were independent in the model and thus were not automatically set to 1 were p, k_1, k_2, k_3, k_4, k_5, k_7, k_8, k_9, and n. In the simulations of the model, we set $p = 1$ not because it was dependent on other parameters, but because we wished to simulate the homosynaptic condition. As for the other parameters, we chose their values after a cursory exploration of the parameter space to yield simulations that resembled the hippocampus data (to restrict the parameter space, we attempted successfully to model the data with $k_4 = k_5$ and $k_8 = k_9$). Variations of each these parameters by factors of three up and down yielded data that were quantitatively but not qualitatively different, demonstrating the robustness of the model. The values of the parameters throughout the article were $k_1 = 10$, $k_2 = 0.5$, $k_3 = 0.001$, $k_4 = k_5 = 1$, $k_7 = 100$, $k_8 = k_9 = 0.33$, and $n = 2$.

These parameter values are reasonable and do not violate the assumptions used to derive the model's equation 2.6. The assumption that the decay from the metastate is much faster than the arrival to it is automatically satisfied in this equation. Multiplying k_4, k_5, k_6, and k_7 by any common factor leaves the equation unchanged. Therefore, one can choose them to ensure an arbitrarily fast decay (see Figure 1B) without modifying the simulations. However, although the absolute values of k_4, k_5, k_6, and k_7 are irrelevant, their relative values matter. We chose the decay through E_i (governed by k_7) to be only 100 times faster than the other decays, which is biochemically reasonable (Mathews & van Holden, 1990). In contrast, the absolute values of k_8 and k_9 matter, since they set the maximal rate constants of synaptic changes. Their associated time constants correspond to the order of minutes, which is slow[3] but not biochemically unreasonable (Mathews & van Holden, 1990). Finally, besides rate constants, equation 2.6 also has the parameter n. The choice $n = 2$ means that two feedback-messenger molecules are needed to activate one E_a enzyme, another biochemically reasonable assumption.

[3] Slowness implies that E_s promotes a high energy-of-activation transition. This would be the case if, for instance, our particles corresponded to (the "bulky") vesicles and the enzymes to proteins involved in vesicle turnover and membrane docking. If such an interpretation were correct, it would point to an interesting relationship between our model and the synaptic-depression-gain-control ideas of Abbott, Varela, Sen, & Nelson (1997). Synaptic depression may involve depletion of vesicles immediately available for transmitter release (Takeuchi, 1958; Thies, 1965).

The parameters of the switch equation (2.5) are also reasonable. Similar to k_8 and k_9, the time constant parameter for turning off the switch, $1/k_2$, is also of the order of minutes. In turn, k_3 corresponds to a very slow process. It was included only to prevent the active state of the switch from disappearing during quiescence and thus to allow the switch to turn on.[4] As for k_1, this is not strictly speaking a rate constant, since k_1's dimension is that of the inverse of the multiplication of time by concentration squared (see equation 2.5). Because $k_1[m_j][E_{s,j}]$ sums k_3, it is useful to think of $k_1[m_j][E_{s,j}]$ as the real rate constant. This "rate constant" and thus, k_1, becomes relevant for the model only when $k_1[m_j][E_{s,j}] \geq k_3$. Hence, k_1's role is to set the threshold above which the feedback messenger starts turning on the switch. A good working definition of the switch's turning on is when $[E_{s,j}]$ crosses $E_{s,0}/2$. In this case, if the threshold is that value of m_j for which $E_{s,j} = 0.5E_{s,0}$ at steady state, then from equations 2.5 and $k_3 \ll k_2$, the threshold occurs when $m_j \approx 2k_2/(k_1E_{s,0})$.

The stimuli in the model simulations were chosen as follows. In LTP and LTD physiological experiments, stimuli consist of trains of short current pulses. These pulses are delivered to elicit action potentials, which are necessary to activate the synapses. However, we are not modeling action potentials, and our synapses can function with continuous stimuli. Hence, we use continuous stimuli with amplitudes I_j. The unit of the amplitude is hertz, and its numerical value is equal to the numerical value of the frequency of the current pulses.

Among the experimental conditions simulated, two key ones include blockade of *N*-methyl-*D*-aspartate (NMDA) receptors and the effects of calyculin A, a phosphatase inhibitor. In hippocampus, blockade of NMDA receptors results in the reduction of the Ca^{2+} concentration in the postsynaptic site (Jahr & Stevens, 1987; Mayer, MacDermott, Westbrook, Smith, & Barker, 1987; Ascher & Nowak, 1988). We propose to identify Ca^{2+} with the model's postsynaptic agent C (see the definition after equation 2.1). Consequently, in the model, NMDA antagonists should result in a reduction in the concentration of the feedback messenger. The effect of NMDA blockade is thus modeled by reducing parameter α in equation 2.2. (The percentage of NMDA blockade in Figure 5 is the percentage of α reduction.) To model the effects of calyculin A, one must identify the term in the switch equation (see equation 2.5) that would correspond to the phosphatase. As explained before equation 2.5, and in more detail, in section 4, for the hippocampus, we identify the first and second terms of the right-hand side of equation 2.5 with the phosphatase and an autophosphorylating enzyme respectively. Therefore, to model the effect of calyculin A, k_2 is set to 0 in equation 2.5.

[4] If $k_3 = 0$ and there are no inputs to the system for a long time (namely, $[m_j] = 0$), then the k_2 term will drive $[E_{s,j}]$ to zero. In this case, the switch cannot turn on, that is, its derivative remains at zero (see equation 2.5).

4 Results

This section presents numerical analysis of the model in the homosynaptic condition, that is, equations 2.1, 2.2, 2.5, and 2.6, with $p = 1$ (see footnote 2).

In the simulations, the phenomena of homosynaptic LTD and LTP can be elicited by applying weak and strong electrical stimulation to the synaptic input, respectively. Three consecutive 1-minute-long 50-Hz trains cause increasing LTP, as shown in Figure 2A. The corresponding LTD effect is shown in Figure 2B using three consecutive 2-minute-long 2-Hz trains. After each train, the excitatory postsynaptic potential (EPSP) slope (taken to be proportional to synaptic gain) changes with respect to baseline level. The simulated LTP and LTD have transient and long-lasting components. The transience is due to the time it takes for the switch to turn off after the stimulus stops (see Figure 2C). For LTP, the slope increase following each subsequent electrical stimulation diminishes as synaptic weight approaches its upper saturating level. Similarly, for LTD, the slope decrease diminishes as synaptic weights approach zero. These effects are visible in both the transient and long-lasting components. Both the LTP and LTD are accompanied by transient "turning on" of the enzymatic switch (see Figure 2C). Although the switch rapidly turns on in both cases, it is faster for the LTP condition (which thus saturates more quickly) than for the LTD condition, because larger amounts of messenger are involved in the former condition. Reversibility of LTP and LTD is illustrated in Figure 3, where a sequence of inputs alternating between strong (3-minute-long 60-Hz trains) and weak (15-minute-long 4-Hz) electrical stimulation causes a sequence of responses alternating between LTP and LTD, respectively.

Changes in the EPSP slope resulting from electrical stimulation are assessed through simulations of a range of frequencies (0–50 Hz). The changes for the long-lasting components (asymptotic) and the peak amplitude of the transient component are shown in Figures 4A and 4B, respectively. Five solid curves are shown in these figures. These curves correspond to different impulse duration (1, 2, 5, and 15 minutes) and to infinite duration (step function). (The infinite-duration condition yields the same curve for the long-lasting and transient components. We obtain it analytically after setting the derivatives in equations 2.5 and 2.6 to zero.) The strength of LTP and LTD varies with stimulus duration in such a manner that the crossing point between LTD and LTP always occurs at the same input frequency (with the parameters used, this frequency is 20 Hz). This crossing point is not absolute but depends on the quantity m_j, which itself is dependent on synaptic strength and, thus, prior activity. In general, the higher the synaptic strength, the easier LTD becomes, and vice versa (see the discussion after equation 2.6).

To test whether our model is further compatible with basic LTP-LTD phenomenology in hippocampus, two more experimental conditions have been simulated. The first involves the blockade of LTP and LTD with NMDA

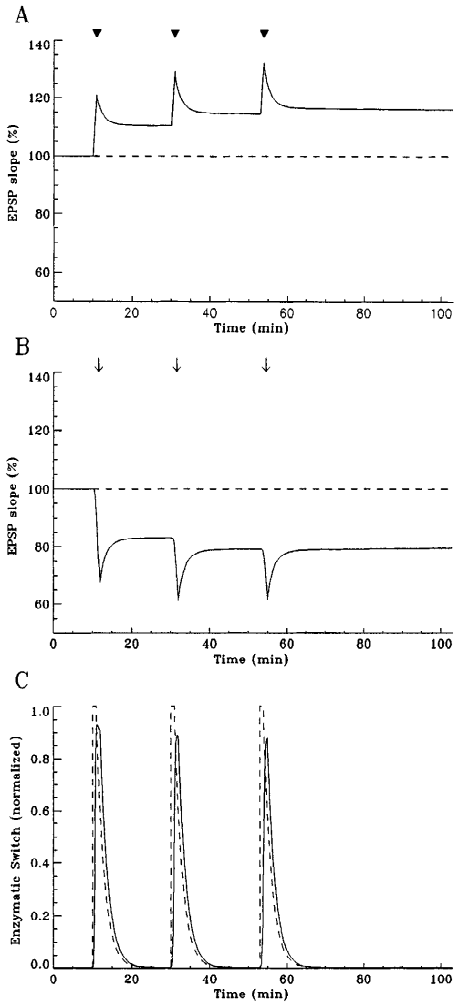

Figure 2: Simulation of LTP and LTD. (A) Three consecutive strong inputs (1-minute-long 50-Hz trains, indicated by the filled triangles) caused increasing LTP. This LTP had transient and long-lasting components. (B) In contrast, three consecutive weak inputs (2-minute-long 2-Hz trains indicated by the arrows) caused increasing LTD. This LTD also had transient and long-lasting components. (C) Both the LTP and LTD were accompanied by transient "turning on" of the enzymatic switch (solid lines for LTD and dashed lines for LTP).

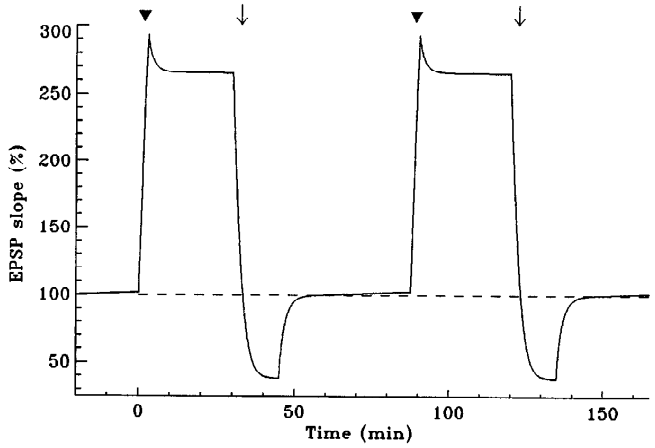

Figure 3: Reversibility of LTP and LTD. A sequence of inputs alternating between strong (3-minute-long 60-Hz trains) and weak frequencies (15-minute-long 4-Hz trains) caused a sequence of responses alternating between LTP and LTD.

antagonists. In hippocampus, the application of NMDA antagonists results in a reduction of the local concentration of postsynaptic Ca^{2+}, and thus in our model, in a reduction of the concentration of the feedback messenger. As the amount of messenger diminishes, the LTD term in equation 2.6 becomes more predominant. This effect can be verified in Figure 5A, where the crossing point for LTD and LTP shifts toward higher frequencies as the NMDA conductance is reduced. Therefore, for sufficient NMDA blockade, LTP disappears. Moreover, the amount of LTD falls and eventually disappears (see Figure 5B). This is because with little feedback messenger, the enzyme E_i and the switch (see Figure 1) do not work.

The second extra experimental condition deals with the mechanism for the maintenance of synaptic modifications. In our model, a switch is turned on by activity (to allow for synaptic changes) and turned off when there is no activity (so that synaptic changes get frozen). Although this switch could be implemented in many ways, Lisman (1989) advanced for it the most experimentally successful hypothesis so far. He essentially proposed that the turning on of the switch may be due to self-phosphorylating enzymes. Hence, a phosphatase would turn it off. In support, administration of calyculin A, a phosphatase inhibitor, eliminates LTD maintenance (Mulkey, Herron, & Malenka, 1993). (The mechanism for LTP maintenance seems to be less clear; Malinow, Madison, & Tsien, 1988.) Our model is consistent

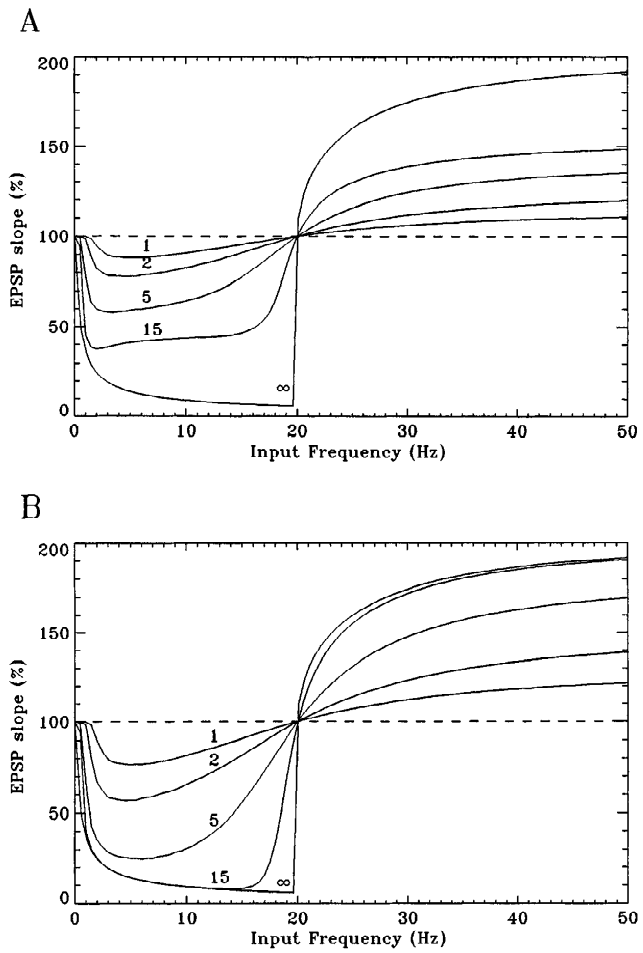

Figure 4: Frequency dependence. The amount of LTP and LTD as a function of input frequency and parametric on stimulus duration (1, 2, 5, 15, and ∞ minutes) were calculated. The calculations were performed for both the asymptotic behavior (A) and peak LTP and LTD (B). The amount of LTP and LTD increased with stimulus duration in such a manner that the crossing point between LTD and LTP always occurred at the same input frequency.

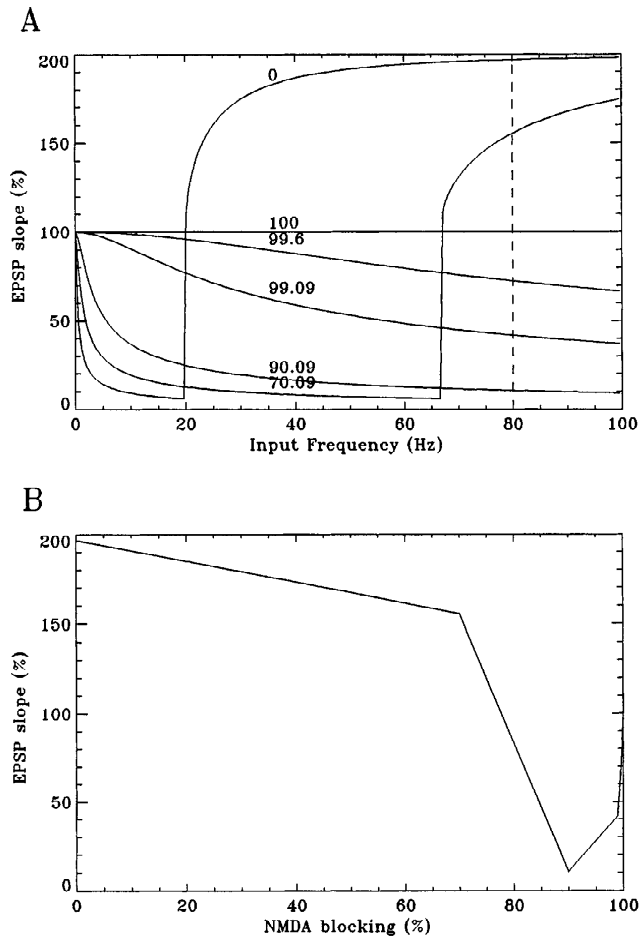

Figure 5: NMDA blockade. (A) When the NMDA conductances were reduced by the percentage indicated in the various curves, the curves shifted to the right. Hence, for sufficient NMDA blockade, LTP disappeared at reasonable input frequencies. Moreover, the amount of LTD fell, eventually disappearing. These changes were such that LTP disappeared before LTD as function of NMDA blockade. To illustrate this point more clearly, (B) shows synaptic changes as a function of the percentage of NMDA blockade for the 80-Hz input frequency, indicated by a dashed line in (A).

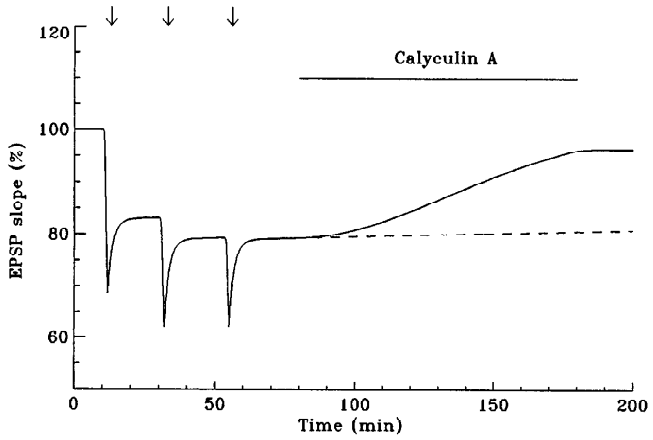

Figure 6: Administration of calyculin A, a phosphatase inhibitor, at the times indicated by the horizontal solid line reverted the LTD condition generated by three low-frequency inputs (2-minute-long 2-Hz trains). Under this drug, the synaptic strength returned to baseline even in the absence of any stimulus.

with this phenomenology, because the second term of the right-hand side of equation 2.5 implements an autocatalytic reaction and the first term, a breakdown of this reaction. Consequently, one can think of the second term as an autophosphorylating enzyme and of the first term as a phosphatase. A simple way to model the switch being permanently on upon application of a phosphatase inhibitor (such as calyculin A) is to set $k_2 = 0$ in equation 2.5. The resulting effect is shown in Figure 6, where LTD has been induced by three 2-minute-long low-frequency trains and was reset to near baseline by application of calyculin A.

5 Discussion

We describe synaptic changes as being mediated by a retrograde messenger, which controls a particle's activation state presynaptically. This control is through manipulation of an enzyme that activates and another that inactivates this particle. These enzymes yield LTP or LTD depending on the concentration of the feedback messenger. The messenger also controls an enzymatic switch that impedes synaptic changes from occurring in the absence of pre- or postsynaptic activity. In turn, this messenger depends on the concentration of a postsynaptic agent, C, which is proportional to the multiplication of postsynaptic response and the level of incoming transmitter.

5.1 Properties of the Model and Data Accountability. In our bidirectional Hebbian rule, synapses have bounded nonnegative weights and do not change these weights in the absence of pre- and postsynaptic activity. These two properties are not simultaneously shared by any of the rules extant in the literature. Furthermore, the new rule explicitly states the biophysical mechanisms of bidirectionality, bounding, and synaptic stability.

From the computational perspective, it is interesting that our rule shares the key property of the BCM rule (Bienenstock et al., 1982; Yang & Faber, 1991; Intrator & Cooper, 1992). That rule has a variable threshold deciding between LTP and LTD as a function of a nonlinear statistic of the past activity. This form of the variable threshold has some useful computational properties, as shown by Intrator and Cooper (1992). As discussed after equation 2.6, our rule has a variable threshold deciding between LTP and LTD, which is a function of the actual value of the weight. The way the threshold varies with past activity is qualitatively similar in both rules.

From the mechanistic perspective, if one identifies C with Ca^{2+} entering the postsynaptic site via NMDA channels (Jahr & Stevens, 1987; Mayer et al., 1987; Ascher & Nowak, 1988) and the enzymatic switch with a system of phosphatase and self-phosphorylating enzymes, then our rule accounts for much of the LTP and LTD hippocampal phenomenology. Consistent with physiological data (Xie et al., 1992; Dudek & Bear, 1993; Heynen, Abraham, & Bear, 1996), the simulated LTP and LTD have transient and long-lasting components. Also in agreement with physiology, LTP and LTD are reversible (Heynen et al., 1996; Dudek & Bear, 1993; Mulkey et al., 1993), as illustrated in Figure 3. The strength of LTP and LTD varies with stimulus duration in such a manner that the crossing point between LTD and LTP always occurs at the same input frequency (Artola & Singer, 1993). Complete blockade of NMDA receptors prevents the LTP and LTD from happening (Morris, Anderson, Lynch, & Baudry, 1986; Goldman, Chavez-Noriega, & Stevens, 1990; Dudek & Bear, 1992). In addition, a bias in the amount of LTP and LTD is induced by either partial blockade of NMDA receptors (Cummings, Mulkey, Nicoll, & Malenka, 1996) or postsynaptic depolarization (simulations not shown, Artola, Bröcher, & Singer, 1990; Xie et al., 1992; Lin, Way, & Gean, 1993). A phosphatase inhibitor, calyculin A, eliminates the maintenance of LTD (Mulkey et al., 1993). And prior synaptic activity influences the induction of subsequent LTP and LTD systematically (Dudek & Bear, 1992; Huang, Colino, Selig, & Malenka, 1992; Wexler & Stanton, 1993).

5.2 Limitations of the Model. We modeled our Hebbian rule as if LTP and LTD are expressed presynaptically. Evidence for a presynaptic site of gain control comes from studies on the probability of synaptic quantal events (del Castillo & Katz, 1954). In these studies, such a probability increases as synapses are potentiated (Bekkers & Stevens, 1990; Malinow, 1991) and decreases when they are depressed (Stevens & Wang, 1994; Oliet, Malenka, & Nicoll, 1996). These changes in frequency of quantal events

are interpreted as reflecting changes in the release of transmitter, indicating presynaptic changes. However, because there is also evidence against an increase in transmitter release (Manabe & Nicoll, 1994) and some evidence for an increase in quantal amplitude in LTP (Kullman & Nicoll, 1992; Oliet, Malenka, & Nicoll, 1996), the presynaptic interpretation remains still somewhat controversial. We do not regard our presynaptic assumption as essential. The postsynaptic agent C could instead control activating and inactivating enzymes (such as our enzymes E_a and E_i) involved in the effectiveness of postsynaptic receptors. Similarly, the enzymatic switch complex could exist postsynaptically.

Another limitation of the model has to do with the explanation of potentiation transients (see Figure 2A). The mechanisms for these transients include synaptic facilitation, augmentation, and short-term potentiation (STP; Feng, 1941; Liley, 1956; Magleby & Zengel, 1976). In hippocampus, these mechanisms can be induced without LTP (Malenka & Nicoll, 1993). Nevertheless, as far as we know, there is no evidence against the postulated enzymatic switch contributing to part of the potentiation's transient as in our simulations. Furthermore, although STP can account for part of the potentiation transient, STP cannot account for LTD's transient. In contrast, the new bidirectional Hebbian rule can account for the experimentally observed LTD transients.

The LTP-LTD model also does not take into account nonlinear Ca^{2+} buffering (Holmes & Levy, 1990), voltage-dependent Ca^{2+} channels, or the details of glutamate and NMDA receptor functions (Madison, Malenka, & Nicoll, 1991; Clements, 1996). (As stated in section 1, the essence of this model is in the details of the E_a, E_i, and E_s enzymes. The other mechanisms in the model are chosen for computational simplicity to be linear.) Therefore, it is unrealistic to assume that the postsynaptic Ca^{2+} is proportional to the concentration of transmitter. Nevertheless, we confirmed that our qualitative conclusions hold for reasonable monotonic nonlinear relationships between the levels of incoming transmitter and Ca^{2+}.

5.3 Is There a Switch? Is an enzymatic switch a reasonable feature of the model? An argument against this switch comes from synaptic remodeling in cortex being widely thought to engage the same mechanisms as LTP-LTD in hippocampus (Artola & Singer, 1993; Hirsch & Gilbert, 1993; Kirkwood, Dudek, Gold, Aizenman, & Bear, 1993; Castro-Alamancos, Donoghue, & Connors, 1995). Because a deprived eye loses cortical representation (Wiesel & Hubel, 1963; Baker, Grigg, & von Noorden, 1974), this suggests that silent terminals can have weights changed. However, the deprived eye sends spontaneous signals that may be sufficient to turn the switch on. There is evidence that spontaneous maintained discharge from the eye has an important role in the normal development of visual cortical binocularity. Binocular injection of tetrodoxin, but not binocular visual deprivation, prevents segregation of the geniculocortical afferents serving the two eyes (Stryker &

Harris, 1986; Antonini & Stryker, 1993). Hence, deprivation results are not inconsistent with the switch, and there is evidence favoring it. As pointed out in section 4 for LTD, the switch's assumption receives direct support from experiments involving phosphatase inhibitors (Mulkey et al., 1993).[5] In addition, as discussed in section 5.2, at least part of the transience of LTP and LTD may be a consequence of the switch. Another piece of evidence in favor of a switch (though less direct) comes from experiments on the reversibility of LTP in the hippocampus. LTP is induced at synapses where the pre- and postsynaptic activities are correlated, and the signal for its induction is postsynaptic Ca^{2+} influx. Hence, the interruption of this influx through blockade of the NMDA receptor should induce LTD when the presynaptic activity is high. However, such an experiment fails to produce significant LTD (Goldman et al., 1990). Such a failure is consistent with our model for which lack of postsynaptic Ca^{2+} impedes the switch from turning on and consequently prevents any synaptic changes from occurring.

Appendix

This appendix derives equation 2.6 from the enzymatic reactions shown in Figure 1B and the assumption that the decay from the metastate is much faster than the arrival time to it. Instead of using symbols for the concentrations of the active and inactive states of the gating molecule, we work directly with w and w^*, respectively, to simplify notation. These quantities are proportional to the above concentrations, but we absorb the constants of proportionality into the rate constants. Similarly, we use the variable w^{**} to quantify the metastate.

The kinetic diagrams for the reactions are:

$$\begin{aligned} E_s + w &\xrightarrow{k_9} E_s + w^{**} \\ E_s + w^* &\xrightarrow{k_8} E_s + w^{**} \\ w^{**} &\xrightarrow{k_5} w \\ w^{**} &\xrightarrow{k_4} w^* \\ E_a + w^{**} &\xrightarrow{k_6} E_a + w \\ E_i + w^{**} &\xrightarrow{k_7} E_i + w^*. \end{aligned} \tag{A.1}$$

From these diagrams, the differential equations for the metastate and active state are

$$\frac{dw^{**}}{dt} = -(k_4 + k_5 + k_6[E_a] + k_7[E_i])w^{**} + k_9[E_s]w + k_8[E_s]w^* \tag{A.2}$$

[5] Our model is not consistent with the failure of the phosphatase inhibitor, calyculin A, to eliminate LTP maintenance. One possibility is that two separate switches exist for LTP and LTD, with the former not involving a phosphatase.

and

$$\frac{dw}{dt} = -k_9 w[E_s] + (k_5 + k_6[E_a])w^{**}, \tag{A.3}$$

respectively. The assumption that the decay from the metastate is much faster than the arrival to it is akin to w and w^* changing on slow time scales to the metastate and to w^{**} remaining in equilibrium with w and w^*, adjusting on the fast time scale of its decay. Consequently, it is reasonable to regard w^{**} as in a steady state on the time scales over which w and w^* change, allowing us to make the approximation $dw^{**}/dt = 0$. (This approximation is similar to what is done in the derivation of the Michaelis-Menten equation; Walsh, 1977). Using this approximation, solving for w^{**} in equation A.2, and inserting the result into equation A.3, we get

$$\frac{dw}{dt} = \frac{-(k_4 + k_7[E_i])k_9[E_s]w + (k_5 + k_6[E_a])k_8[E_s]w^*}{k_4 + k_5 + k_6[E_a] + k_7[E_i]}. \tag{A.4}$$

Using $w^* = w_0 - w$, and substituting βm and γm^n for E_i and E_a, respectively, where β and γ are constants, we obtain equation 2.6. These last substitutions are valid under the hypothesis that enzymes E_i and E_a are available in large amounts and that their kinetics of activation are much faster than m's dynamics.

Acknowledgments

This work was supported by grants from the National Eye Institute (EY08921 and EY11170) and by the William A. Kettlewell Chair to N.M.G., by a grant from the Swiss National Fund for Scientific Research (8220-37180) to P.-Y.B., and by a core grant from the National Eye Institute to Smith-Kettlewell (EY06883).

References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science, 275*, 220–223.

Antonini, A., & Stryker, M. P. (1993). Development of individual geniculocortical arbors in cat striate cortex and effects of binocular impulse blockade. *J. Neurosci., 13*, 3549–3573.

Artola, A., Bröcher, S., & Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. *Nature, 347*, 69–72.

Artola, A., & Singer, W. (1993). Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *TINS, 16*, 480–487.

Ascher, P., & Nowak, L. (1988). The role of divalent cations in the *N*-methyl-*D*-aspartate responses of mouse central neurones in culture. *J. Physiol., 399*, 247–266.

Baker, F. H., Grigg, P., & von Noorden, G. K. (1974). Effects of visual deprivation and strabismus on the responses of neurons in the visual cortex of the monkey, including studies on the striate and prestriate cortex in the normal animal. *Brain Res., 66*, 185–208.

Bekkers, J. M., & Stevens, C. F. (1990). Presynaptic mechanism for long-term potentiation in the hippocampus. *Nature, 346*, 724–729.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *J. Neurosci., 2*, 32–48.

Burgi, P.-Y., & Grzywacz, N. M. (1997). Possible roles of spontaneous waves and dendritic growth for retinal receptive field development. *Neural Comput., 9*, 533–553.

Burgi, P.-Y., & Grzywacz, N. M. (in press). A biophysical model for the developmental time course of retinal orientation selectivity. *Vis. Res.*

Castro-Alamancos, M. A., Donoghue, J. P., & Connors, B. W. (1995). Different forms of synaptic plasticity in somatosensory and motor areas of the neocortex. *J. Neurosci., 15*, 5324–5333.

Cherubini, E., Rovira, C., Gaiarsa, J. L., Corradetti, R., & Ari, Y. B. (1990). GABA mediated excitation in immature rat CA3 hippocampal neurons. *Int. J. Devl. Neurosci., 8*, 481–490.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Clements, J. D. (1996). Transmitter timecourse in the synaptic cleft: Its role in central synaptic function. *Trends Neurosci., 19*, 163–171.

Cummings, J. A., Mulkey, R. M., Nicoll, R. A., & Malenka, R. C. (1996). Ca^{2+} signaling requirements for long-term depression in the hippocampus. *Neuron, 16*, 825–833.

del Castillo, J., & Katz, B. (1954). Quantal components of the end-plate potential. *J. Physiol., 124*, 560–573.

Dudek, S. M., & Bear, M. F. (1992). Homosynaptic long-term depression in area CA1 of hippocampus and effects of *N*-methyl-*D*-aspartate receptor blockade. *Proc. Natl. Acad. Sci. USA, 89*, 4363–4367.

Dudek, S. M., & Bear, M. F. (1993). Bidirectional long-term modification of synaptic effectiveness in the adult and immature hippocampus. *J. Neurosci., 13*, 2910–2918.

Feng, T. P. (1941). Studies on the neuromuscular junction. XXVI. The changes of the end plate potential during and after prolonged stimulation. *Chin. J. Physiol., 16*, 341–372.

Frégnac, Y., Burke, J. P., Smith, D., & Friedlander, M. J. (1994). Temporal covariance of pre- and post-synaptic activity regulates functional connectivity in the visual cortex. *J. Neurophysiol., 71*, 1403–1421.

Goldman, R. S., Chavez-Noriega, L. E., & Stevens, C. F. (1990). Failure to reverse long-term potentiation by coupling sustained presynaptic activity and *N*-

methyl-*D*-aspartate receptor blockade. *Proc. Natl. Acad. Sci. U.S.A.*, *87*, 7165–7169.

Grzywacz, N. M., & Burgi, P.-Y. (1994). A Hebbian rule consistent with physiology and producing sharp refinement of organization. *Abstr. Third Ann. Comput. Neural Syst. Meeting* (p. 38). Monterey, CA.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Heynen, A. J., Abraham, W. C., & Bear, M. F. (1996). Bidirectional modification of CA1 synapses in the adult hippocampus in vivo. *Nature*, *381*, 163–166.

Hirsch, J. A., & Gilbert, C. D. (1993). Long-term changes in synaptic strength along specific intrinsic pathways in the cat visual cortex. *J. Physiol.*, *461*, 247–262.

Holmes, W. R., & Levy, W. B. (1990). Insights into associative long-term potentiation from computational models of NMDA receptor-mediated calcium influx and intracellular calcium concentration changes. *J. Neurophysiol.*, *63*, 1148–1168.

Huang, Y.-Y., Colino, A., Selig, D. K., & Malenka, R. C. (1992). The influence of prior synaptic activity on the induction of long-term potentiation. *Science*, *255*, 730–733.

Intrator, N., & Cooper, L. N. (1992). Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Net.*, *5*, 3–17.

Jahr, C. E., & Stevens, C. F. (1987). Glutamate activates multiple single channel conductances in hippocampal neurones. *Nature*, *325*, 522–525.

Kirkwood, A., Dudek, S. M., Gold, J. T., Aizenman, C. D., & Bear, M. F. (1993). Common form of synaptic plasticity in the hippocampus and neocortex in vitro. *Science*, *260*, 1518–1521.

Kullman, D. M., & Nicoll, R. A. (1992). Long-term potentiation is associated with increases in quantal content and quantal amplitude. *Nature*, *357*, 240–244.

Liley, A. W. (1956). The quantal components of the mammalian end-plate potentials. *J. Physiol. Lond.*, *133*, 571–587.

Lin, J. H., Way, L.-J., & Gean, P.-W. (1993). Pairing of pre- and post-synaptic activities in hippocampal CA1 neurons induces long-term modifications of NMDA receptor-mediated synaptic potential. *Brain Res.*, *603*, 117–120.

Linsker, R. (1986). From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proc. Natl. Acad. Sci. USA*, *83*, 8390–8394.

Lisman, J. (1989). A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proc. Natl. Acad. Sci. USA*, *86*, 9574–9578.

Madison, D. V., Malenka, R. C., & Nicoll, R. A. (1991). Mechanisms underlying long-term potentiation of synaptic transmission. *Ann. Rev. Neurosci.*, *14*, 379–397.

Magleby, K. L., & Zengel, J. E. (1976). Augmentation: A process that acts to increase transmitter release at the frog neuromuscular junction. *J. Physiol. Lond.*, *257*, 449–470.

Malenka, R. C., & Nicoll, R. A. (1993). NMDA-receptor-dependent synaptic plasticity: Multiple forms and mechanisms. *Trends Neurosci.*, *16*, 521–527.

Malinow, R. (1991). Transmission between pairs of hippocampal slice neurons: Quantal levels, oscillations, and LTP. *Science*, *252*, 722–724.

Malinow, R., Madison, D. V., & Tsien, R. W. (1988). Persistent protein kinase activity underlying long-term potentiation. *Nature, 335*, 820–824.

Manabe, T., & Nicoll, R. A. (1994). Long-term potentiation: Evidence against an increase in transmitter release probability in the CA1 region of the hippocampus. *Science, 265*, 1888–1893.

Mathews, C. K., & van Holden, K. E. (1990). *Biochemistry*. Redwood City, CA: Benjamin / Cummings.

Mayer, M. L., MacDermott, A. B., Westbrook, G. L., Smith, S. J., & Barker, J. L. (1987). Agonist- and voltage-gated calcium entry in cultured mouse spinal cord neurons under voltage clamp measured using arsenazo III. *J. Neurosci., 7*, 3230–3244.

Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *J. Neurosci., 14*, 409–441.

Morris, R. G., Anderson, E., Lynch, G. S., & Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an *N*-methyl-*D*-aspartate receptor antagonist, AP5. *Nature, 319*, 774–776.

Mulkey, R. M., Herron, C. E., & Malenka, R. C. (1993). An essential role for protein phosphatases in hippocampal long-term depression. *Science, 261*, 1051–1055.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Math. Biol., 15*, 267–273.

Oliet, S. H. R., Malenka, R. C., & Nicoll, R. A. (1996). Bidirectional control of quantal size by synaptic activity in the hippocampus. *Science, 271*, 1294–1297.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C*. Cambridge: Cambridge University Press.

Sejnowski, T. J. (1977). Storing covariance with nonlinearly interacting neurons. *J. Math. Biol., 4*, 303–321.

Stanton, P. K., & Sejnowski, T. J. (1989). Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature, 339*, 531–533.

Stent, G. S. (1973). A physiological mechanism for Hebb's postulate of learning. *Proc. Natl. Acad. Sci. USA, 70*, 997–1001.

Stevens, C. F., & Wang, Y. (1994). Changes in reliability of synaptic function as a mechanism for plasticity. *Nature, 371*, 704–707.

Stryker, M. P., & Harris, W. A. (1986). Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci., 6*, 2117–2133.

Takeuchi, A. (1958). The long-lasting depression in neuromuscular transmission of frog. *Jpn. J. Physiol., 8*, 102–113.

Thies, R. E. (1965). Neuromuscular depression and the apparent depletion of transmitter in mammalian muscle. *J. Neurophysiol., 28*, 427–442.

von der Malsburg, C. (1973). Self organization of orientation selective cells in the striate cortex. *Kybernetik, 14*, 85–100.

Walsh, C. T. (1977). *Enzymatic reaction mechanisms*. San Francisco: W. H. Freeman.

Walton, M. K., Schaffner, A. E., & Barker, J. L. (1993). Sodium channels, $GABA_A$ receptors, and glutamate receptors develop sequentially on embryonic rat spinal cord cells. *J. Neurosci., 13*, 2068–2084.

Wexler, E. M., & Stanton, P. K. (1993). Priming of homosynaptic long-term depression in hippocampus by previous synaptic activity. *NeuroReport*, *4*, 590–594.

Wiesel, T. N., & Hubel, D. H. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.*, *26*, 1003–1017.

Xie, X., Berger, T. W., & Barrionuevo, G. (1992). Isolated NMDA receptor-mediated synaptic responses express both LTP and LTD. *J. Neurophysiol.*, *67*, 1009–1013.

Yang, X.-D., & Farber, D. S. (1991). Initial synaptic efficacy influences induction and expression of long-term changes in transmission. *Proc. Natl. Acad. Sci. USA*, *88*, 4299–4303.

Received October 22, 1996; accepted June 21, 1997.