The trouble with weight-dependent STDP

Dominic Standage and Thomas Trappenberg

Abstract— We fit a weight-dependent STDP rule to the classic data of Bi and Poo (1998), showing that this rule leads to very slow learning in a simulation with an integrate-and-fire neuron. The slowness of learning is explained by an inequality between the range of initial weights in the data and the largest relative potentiation. We show that slow learning can be overcome with an increased learning rate, but that this approach leads to rapid forgetting in the presence of realistic levels of background spiking. We offer a potential solution to this problem, but moreover, our study demonstrates that weight-dependent STDP rules, commonly used in neural simulations, are not supported by weight-dependent STDP data. We discuss the implications of this finding for several interpretations of weight-dependent plasticity and STDP more generally, and recommend directions for further research.

I. INTRODUCTION

Activity-dependent change in synaptic strength, or synaptic plasticity, is widely believed to provide the basis for learning and memory, as originally proposed by Hebb [1]. Countless physiological data reveal Hebbian plasticity, and similarly vast numbers of neural simulations show associative learning from activity-dependent rules. While early experiments showed potentiation [2] and depression [3] of synaptic strength as a function of pre-synaptic firing rates, more recent experimental methods reveal similar changes based on the timing of pre- and post-synaptic activity [4], [5], [6]. Experiments showing spike-time dependent plasticity (STDP) reveal that the direction (potentiation or depression) and magnitude of synaptic change depend on the order and latency of pre- and post-synaptic activity, within a time window on the order of tens of milliseconds [6], [7].

Weight-dependent STDP rules provide the additional constraint that changes in synaptic strength or *weight* are a function of the initial strength of a synapse. Such rules have been widely studied [8], [9], [10], [11], largely because they provide a cap on synaptic strength while leading to asymptotic weight distributions that compare favourably with experimental data [9] and that are believed to preserve the statistics of input activity [10]. As such, weight-dependent STDP rules are now commonly used in modelling studies [12], [13].

While several experimental studies have shown weight-dependent plasticity [6], [14], [15], [16], only Bi and Poo's (1998) study [6] has done so under the STDP pairing protocol. In recent work [17] we fit a weight-dependent STDP rule to the data of Bi and Poo. Here, we use this rule in a neural simulation, showing that self-organisation of weights and associative learning are achieved, but that learning is

Dominic Standage and Thomas Trappenberg, Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada (email: {standage,tt}@cs.dal.ca).

extremely slow. While an increased learning rate accelerates the learning process, it also leads to rapid forgetting in the presence of biologically realistic background spiking.

Our study shows that the common interpretation of weight-dependent STDP data has biologically unrealistic consequences. These consequences result because weight-dependent plasticity rules assume that weights may traverse the full range of initial values shown in weight-dependent studies (Bi and Poo's weight-dependent data are shown in Figure 1B). This assumption is inconsistent with Bi and Poo's data, and also with weight-dependent data resulting from other plasticity protocols [14], [16]. Given the common inclusion of weight-dependent terms in learning rules, this issue is an important one, and poses questions about experimental methods, data and their interpretation.

II. A WEIGHT-DEPENDENT STDP RULE FIT TO DATA

For simplicity, we assume the weight- and spike-time-dependencies are independent. The general form of a weight-dependent STDP rule may be expressed as

$$\Delta w_{\{p,d\}} = k f_{\{p,d\}}(w) e^{-c_{\{p,d\}}\Delta t}, \qquad (1)$$

where $\Delta w_{\{p,d\}}$ is the change in weight (p for potentiation, d for depression), $\Delta t = t_{post} - t_{pre}$ is the time between pre- and post-synaptic firing, c captures the timescale of the STDP window, and k is a learning rate.

A log-linear fit to Bi and Poo's weight-dependent STDP data, hypothesized by Bi and Poo and reproduced in Figure 1B, yields the following form of $f(w)$:

$$f_{\{p,d\}}(w) = (a_{\{p,d\}} - b_{\{p,d\}} \log w) w \ \theta(w),$$

where $\theta(w)$ is the Heaviside function, keeping weights positive, and the parameters a and b are different for potentiation and depression.

Bi and Poo controlled spike timing in their weight-dependent experiment by limiting their spike latencies Δt. We used the midpoints of these latencies in our weight-dependent fit ($\Delta t_p = 10, \Delta t_d = 17.5$) and assumed that for a given Δt, the largest absolute values for $\Delta w/w$ represent synapses with the smallest initial weights. We therefor assign an initial weight $w = 30$pA to the largest STDP data (open symbols in Figure 1A) and only include these data in our spike-time fit for c. The resulting parameter values are $a_{\{p,d\}} = \{208, -54\}$, $b_{\{p,d\}} = \{26.4, 3.5\}$, $c_{\{p,d\}} = \{0.054, 0.042\}$. Because these parameters were fit to data expressed in percentage terms, and because we make the common assumption that each of the 60 pairings in Bi and Poo's experiment contributed equally to synaptic change [9], we use a learning rate of $k = 1/6000$.
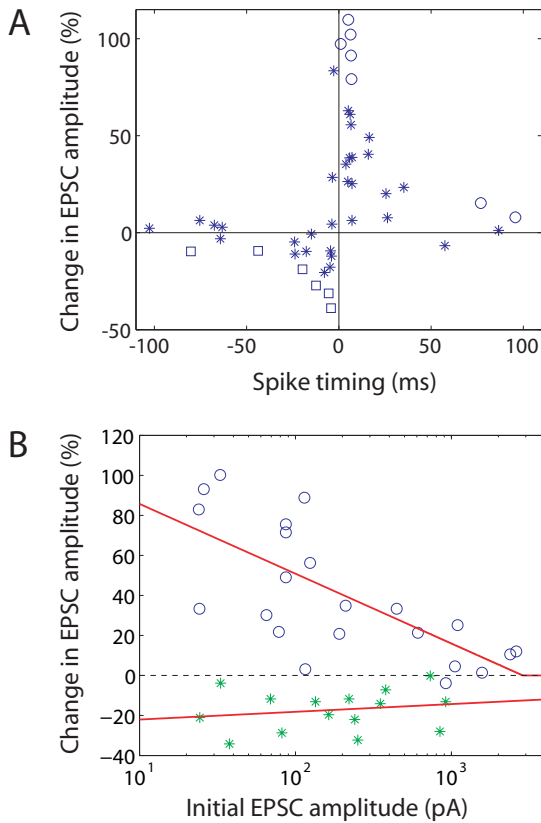
Fig. 1. (A) Reproduction of Bi and Poo's (1998) spike-time dependent data. Circles and squares represent estimates of $w = 30$ for potentiation and depression respectively. (B) Log-linear fit to Bi and Poo's weight-dependent STDP data, imposing a maximum weight where the upper solid line meets the x-axis.
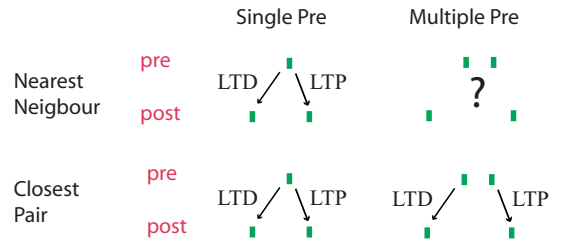

Fig. 2. Cartoons depict spike interaction models described in the text. It is unclear how the *nearest neighbour* model extends to the case where multiple pre-synaptic spikes fall between two post-synaptic spikes. We define the *closest pair* model for this case.
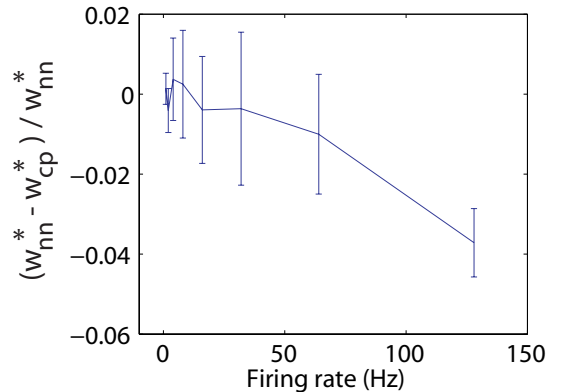

Fig. 3. The relative difference between mean weights w^* under *nearest neighbour* (analytic) and *closest pair* (numeric) spike interactions (see text). There is a very small (but significant) relative difference at high rate. Error bars show standard deviations.

III. STDP AND SPIKE INTERACTIONS

STDP rules are fit to data from experiments in which pairings of pre- and post-synaptic spikes occur at low frequency, typically 0.5-5Hz. The long inter-spike intervals (ISI's) of these spike trains effectively isolate spike pairings from the effects of previous and subsequent spikes. How to apply STDP rules to spike trains more realistic than temporally isolated pairings of pre- and post-synaptic spikes is the subject of great research interest [20], [21], [22], [23], [16], [24]. Here, we use a model we call *closest pair* interactions, where only the pair of pre-synaptic spikes closest to the surrounding post-synaptic spikes contributes to plasticity, as shown in Figure 2.

This model is equivalent to the *nearest neighbour* model suggested in [23] when exactly one pre-synaptic spike falls between two post-synaptic spikes. These two interaction models are depicted in Figure 2. Because the analysis in [23] does not consider spike trains beyond this simplified case, we compare asymptotic weights calculated with their analytical method to asymptotic weights from simulations under closest-pair interactions.

To this end, we numerically approximate the means of equilibrium weight distributions (equilibrium weights w^*) for *closest pair* interactions, comparing them to analytic values

for the *nearest neighbour* case.

In the simulations, post-synaptic spikes were time-locked to Poisson-distributed pre-synaptic spikes at $\Delta t = 4$ms. Weights were averaged over 5000 pairings following 5000 equilibrating pairings at spike rates $r = \{1, 2, 4, 8, 16, 32, 64, 128\}$Hz. Analytic calculations for the *nearest neighbour* case with the same spiking statistics were calculated in [17]. The relative difference between numeric and analytic asymptotic weights for these spike interaction models is shown in Figure 3. There is a systematic difference that increases with increasing spike rate, but the relative difference is very small.

IV. SIMULATIONS WITH A NEURON

To test the effectiveness of Equation 1 for associative learning, we drive an integrate-and-fire (IF) node with two groups of synapses. Before learning, synchronous input from one group is strong enough to sporadically drive the node in combination with background activity from all synapses. The other group cannot drive the node. We test whether repeated synchronous activation of the high-strength group allows self-organisation of these weights, and whether synchronous activation of both groups (following self-organisation) is sufficient for the low-strength group to 'piggyback' the high-strength group. If so, the low-strength group will have learned to drive the node by associativity.

A. Model and parameters

The membrane potential of the IF node is described by

$$\tau_m \frac{dv(t)}{dt} = V_{rest} - v(t) + G_e(V_e - v(t)) + G_i(V_i - v(t))$$

with membrane time constant $\tau_m = 20$ ms and reversal potentials $V_{rest} = -70$mV, $V_e = 0$mV and $V_i = -70$mV where subscripts e and i refer to excitatory and inhibitory potentials respectively. When the membrane potential crosses threshold Θ, it is reset to -60mV with an absolute refractory period of 2mS. A relative refractory period is implemented by increasing $\Theta = -54$mV by $\gamma = 20$mV when the neuron fires, after which γ decays exponentially with half width 10mS. Excitatory and inhibitory conductances G_e and G_i are described (as in [25]) by

$$\tau_e \frac{dG_e(t)}{dt} = -G_e(t) + \tau_e \Sigma_j \ g_j \ \delta(t - t_j^f)$$

and

$$\tau_i \frac{dG_i(t)}{dt} = -G_i(t) + \tau_i \Sigma_j \ g_j \ \delta(t - t_j^f),$$

where $\tau_e = \tau_i = 5$ms, δ refers to the Dirac delta function and t_j^f is the time of firing of inputs mediated by conductance synapses g_j (excitatory for G_e and inhibitory for G_i).

EPSC's as conductances: There is no obvious way to translate Bi and Poo's EPSC measurements to synaptic conductances. To begin with, we do not know the amplitudes of corresponding EPSP's and so cannot determine conductances with Ohm's law. More problematic is the variation in these data ($\sim 25 - 2500$pA) and the size of the larger EPSC's. Because our objective is to study the implications these data for STDP rules, we cannot simply ignore the larger currents. Rather than use current synapses (independent of v), we use a linear transformation, allowing a range of 30-3000pA to correspond to 10-150pS. A maximum conductance of 150pS is used in [25]. Inhibitory synapses are not subject to plasticity and are uniformly allocated conductances of 500pS (also from [25]).

The simulation was run with 1000 excitatory synapses and 200 inhibitory synapses. Poisson background firing with mean 10Hz was mediated by all synapses throughout the simulation, and synchronous inputs were given a uniformly distributed random jitter of 1-5ms, superimposed on the Poisson activity. We initialised the high-strength synapses (trainers) to uniformly distributed random values between 800 and 1200pA. The low-strength synapses (piggybackers) were initialised to 50pA, providing sub-threshold input. All other excitatory synapses were given uniformly distributed random values between 500 and 900pA. Under our transformation to conductances, these weights are too weak for background activity to drive the node, but are able to elicit occasional firing when large populations fire synchronously on top of background activity. Axonal delays AD were assigned uniformly distributed random values (to the nearest integer) from $AD = 4 - 17$ms for excitatory synapses and $AD = 3 - 6$ms for inhibitory synapses. Equation 1 was applied under the *closest-pair* spike interaction model.

B. Associative learning under Equation 1

The IF node was first driven by Poisson activity alone, establishing a baseline membrane potential. This activity was followed by a period of self-organisation by the trainer synapses, driven by periodic synchronous activity at 10Hz. Periodic 10Hz activity provides a sufficiently long ISI to isolate neural activity in each cycle from the effects of the previous cycle. Following this activity, the piggybackers were simultaneously activated with the trainers for the associative task, again at 10Hz. The node was subsequently driven by Poisson activity for comparison with the pre-training response to noise.

Figure 4A shows the trajectories of three trainer synapses and three piggyback synapses over the course of the simulation. The lower solid curve depicts a fast trainer synapse with a small initial weight [$AD = 4$ms, $w(t = 0) = 800$pA], the upper solid curve shows a slow trainer with a large initial weight [$AD = 17$ms, $w(t = 0) = 1200$pA] and the middle curve shows an intermediate case [$AD = 9$ms, $w(t = 0) = 1000$pA]. Early during self-organisation, the trainer group is not strong enough to regularly fire the node, but in combination with background activity, input synchrony generates enough post-synaptic firing for weights to increase until the group regularly drives the node.

Post-synaptic firing during early and late self-organisation is depicted by the upper and lower insets of the figure respectively. All three weights initially increase because they all contribute to post-synaptic activity. As faster weights gain in strength, EPSC onset of slower weights occurs after post-synaptic firing, subjecting them to LTD. As expected, this effect is first exhibited by the slow synapse and later the by mid-latency synapse. The former is effectively inactivated, with final value $w(t = 4 \cdot 10^5 \text{ms}) = 2.4$pA. The latter assumes an intermediate value [$w(t = 4 \cdot 10^5 \text{ms}) = 1009$pA] determined by an ongoing cycle of 'promotion' and 'relegation' to and from the effective input group. This cycle is governed by the interaction of background firing and post-synaptic activity, and the overlap between the jitter in synchronous input and axonal delay. The fast synapse reaches an asymptotic value close to the log rule cut-off.

The dashed curves in Figure 4A show the trajectories of three piggyback synapses. The arrow at the top of the figure shows onset of associative learning. At this point, these small synapses are climbing toward equilibrium weights determined by their background firing and the now-periodic post-synaptic firing generated by the trainer group. With the onset of associative activity, the fast piggyback synapse follows the fast trainer synapse toward an asymptotic weight, the mid-latency synapse assumes a value similar to that of the mid-latency trainer synapse, and the slow piggyback synapse is effectively inactivated.

Learning is extremely slow under Equation 1: Self-organisation and associativity have been achieved, as synchronous input by either group consistently drives the node. Furthermore, Equation 1 has selected the fastest weights within the groups and shown the ability to recruit and discard
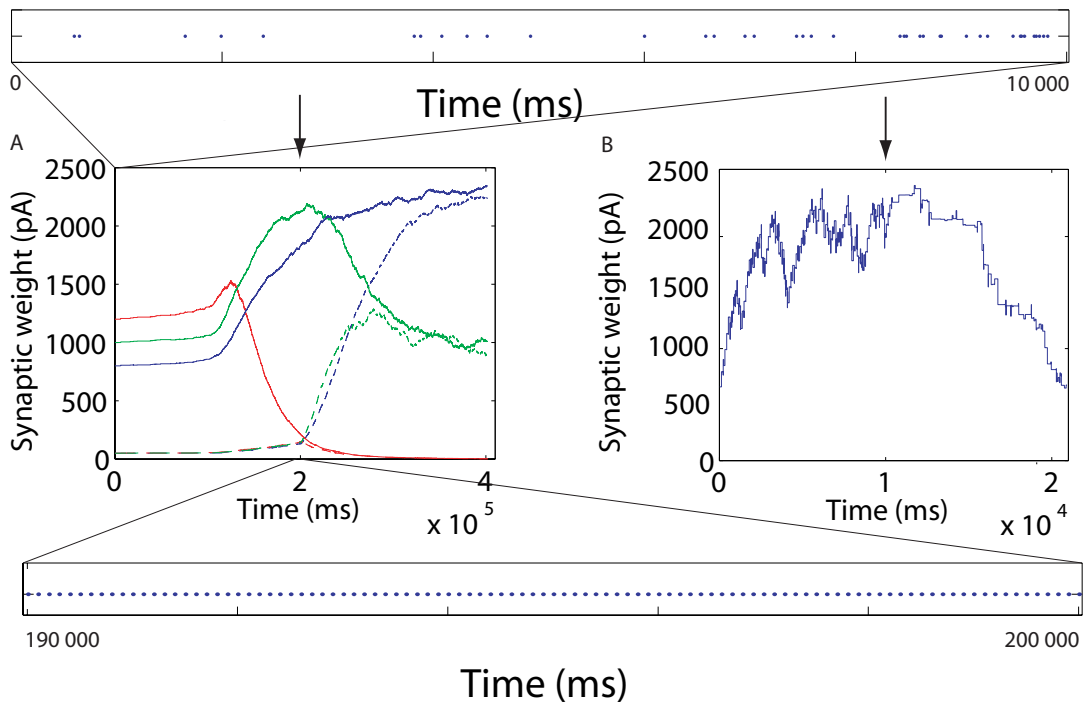
Fig. 4. (A) Self-organisation and associativity under Equation 1 with an integrate-and-fire node. Curves show weights over time for three synapses with large initial weights (trainers, see text) and three synapses with small initial weights (piggybackers, see text). Curves to the left of the arrows show weights as the trainers self-organise during periodic synchronous activity (10Hz, uniform-distributed random jitter 1-5ms) on top of Poisson background activity (mean 10Hz). Arrows depict the onset of associative learning when the piggyback synapses fire synchronously with the trainers, also at 10Hz. A range of 30-3000pA is linearly compressed to 10-150pS for use by the node (see text). The upper solid curves show the weights of three trainers with axonal delays of $AD = 17$, $AD = 9$ and $AD = 4$ms respectively, top to bottom on the left of the figures. During self-organisation, all synapses increase in weight until the faster trainers drive the node without the slower ones, rendering slower synapses subject to depression. The same effect is seen for the piggyback synapses during associative learning. Self organisation and associativity are achieved, but learning is very slow with learning rate $k = 1/6000$. The Fast ($AD = 4$ms) piggyback synapse is still converging to the log rule cut-off after 200,000ms (2000 pairings). (B) An increased learning rate ($k = 1/200$) accelerates learning, but background noise soon causes forgetting.

slower weights as and when needed, as shown in [25]. The mean membrane potential $\bar{v}$ in response to background noise has increased from $\bar{v} = -62.7$mV (standard deviation $sd = 1.1$) before learning to $\bar{v} = -60.7$mV ($sd = 1.2$) after learning, rendering the node more responsive to the timing of its inputs.

Unfortunately, the figure also shows that learning is extremely slow. The fast synapses are still experiencing net potentiation after 400,000ms. This interval accounts for $\sim$ 3000 pre-post pairings for the trainers and 2000 for the piggybackers. The slowness of learning highlights a major problem for weight-dependent STDP rules. Bi and Poo's data show that no synapse increases in strength by more than $\sim 100\%$, but weight-dependent STDP rules assume that synaptic strengths can span the full range of initial weights, a percentage increase of $\sim 10,000\%$.

C. Fast learning implies fast forgetting under weight-dependent STDP

To combat the slowness of learning shown in Figure 4A we raise k from $k = 1/6000$ to $k = 1/200$. With this learning rate, weights begin to asymptote after around 60 pre- and post-synaptic pairings, the same number of pairings used in Bi and Poo's experiments [6]. It is unclear what a learning rate $k \neq 1/6000$ represents, but an accelerated learning rate

is required due to the large variation in synaptic strength in Bi and Poo's weight-dependent data (Figure 1), allowing weights to make the transition from low to high values in a reasonable period of time.

Figure 4B shows the weight of a synapse mediating the same pre-synaptic spike statistics as the trainers did in the learning task, but over a much shorter timespan (10,000ms). The post-synaptic response is time-locked to $\Delta t = 2$ms after pre-synaptic activity, again superimposed on Poisson spiking at 10Hz. Learning is stopped at time $t = 10,000$ms (top arrow), after which pre- and post-synaptic spike trains are given independent Poisson distributions at 10Hz. The synapse forgets the learned association within 10s.

V. DISCUSSION AND CONCLUSIONS

It is possible to parameterize weight-dependent STDP rules and/or model neurons to accomplish various learning tasks, but our study shows unwanted consequences of these rules. The main consequence is due to the range of initial weights in the only weight dependent STDP data available [6], shown in Figure 1B. Weight-dependent STDP rules implicitly assume that a synapse can span this entire range of values, suggesting changes in synaptic efficacy around $10,000\%$. No synapse in these experiments, however,

changed in strength by more than around 100% (see Figure 1A,B).

This issue suggests several possibilities. One possibility concerns the standard computational interpretation of weight-dependent plasticity, in which there is an implicit assumption that weight-dependent data do not reflect saturated synapses. As an example, consider a synapse in Bi and Poo's weight-dependent experiment with an initial strength of 25pA. According to their data, this synapse would not have been potentiated by more than around 100% of control following their potentiation protocol. It would therefor be no larger than around 50pA after 60 pairings, much less than their largest initial weights. If all synapses have the same maximum, then for all but the largest synapses, an additional 60 pairings in Bi and Poo's experiment would have lead to weight changes on the same order as the 60 pairings they performed. The consequences of this assumption are far reaching. For instance, STDP data typically suggest a gradient of weight changes in either direction as a function of the time between pre- and post-synaptic spikes [6], [22]. In the case of potentiation, the standard interpretation dictates that given enough pairings, all weights will reach the same maximum, regardless of the latency between pre- and post-synaptic spikes.

A second possibility is that Bi and Poo's weight-dependent STDP data have been misinterpreted. Their potentiation data may reflect saturated synapses. If so, to be consistent with the data, weight-dependent STDP rules must implement intrinsic synaptic maxima rather than a global maximum. For instance, a synapse with an initial weight of around 25pA should have an intrinsic limit of around 50pA, regardless of the size of larger synapses. We are unaware of any weight-dependent Hebbian rule that operates under this premise, and the utility of such a rule is unclear. Certainly, large and small synapses would have different roles under this scheme.

These possibilities suggest that weight-dependent learning rules with a global maximum should be supported by weight-dependent data in which the largest relative change in weight is roughly equal to the relative difference between the largest and smallest initial weights. We may express this relationship as

$$max\left(\frac{w_i^{fin} - w_i^{init}}{w_i^{init}}\right) \approx \frac{w_{max}^{init} - w_{min}^{init}}{w_{min}^{init}}, \qquad (2)$$

where w_i^{init} and w_i^{fin} represent the weight of a synapse i before and after potentiation respectively, and w_{max}^{init} and w_{min}^{init} are the largest and smallest initial weights respectively. The data of Montgomery et al. (2001) [15] are roughly consistent with this equation for synapses with initial weights less than 50pA using a (non-STDP) pairing protocol.

The weight-dependent data of Debanne et al. (1999) [14] are intriguing in this regard. These authors paired pre-synaptic spikes with post-synaptic bursts in organotypic hippocampal slice cultures. Their CA3-CA1 data are reminiscent of Bi and Poo's in that the relative difference in initial weights is much greater than the normalized weight change, suggesting (as above) that these synapses have widely vary-

ing instrinsic maxima or that they have not reached saturation. Their CA3-CA3 synapses, however, appear to show the opposite effect (the left side of Equation 2 is greater than the right). Their combined data fit roughly with Equation 2. The degree to which plasticity data from different protocols may be related under computational rules is an open question and is currently receiving considerable attention [26], [23], [?].

We show in Section IV-B that a weight-dependent STDP rule fit to weight-dependent STDP data leads to very slow learning. As shown in Section IV-C, speeding up learning with an increased learning rate leads to rapid forgetting in the presence of noise, a problem previously discussed by Senn [27]. A solution to this problem may be grounded in pre- and post-synaptic mechanisms if we (quite happily) abandon the assumption that the effects of spike pairings sum linearly. Consider the insertion [28], [29] and removal [30], [31] of AMPA receptors (AMPAR's) in the post-synaptic membrane, correlated with plasticity in hippocampal regions CA1 and CA3 [32]. It seems likely that potentiation should be harder to initiate as the number of AMPAR's increases. Pre-synaptically, plasticity is expressed at least in part by changes in the probability of transmitter release [33]. Weight-dependence is implicit in such a mechanism, where synapses with a low probability of release have greater scope for potentiation, and synapses with a high probability of release have greater scope for depression. A convolution of pre- and post-synaptic plasticity mechanisms could protect learned associations if pre-synaptic plasticity were to precede post-synaptic plasticity in both directions. Under such a scheme, AMPAR's may not be inserted in the post-synaptic density until a sufficiently high probability of release has been achieved, providing an initial resistance to post-synaptic learning. In reverse, AMPAR's may not be removed from the post-synaptic density until the probability of release is sufficiently low, providing initial resistance to forgetting. Such a scheme requires further study.

Finally, there is another possibility regarding weight-dependent plasticity data. Our recent work [17] provides an alternative interpretation of Bi and Poo's weight-dependent data, showing that under several sets of assumptions, their data may reflect (non-weight dependent) plasticity among populations of synapses. This possibility suggests that tissue preparations and stimulation methods must be carefully considered when interpreting experimental data. Extracellular stimulation is commonly used to mimic pre-synaptic activity in STDP paradigms [7], but this method may stimulate hundreds if not thousands of synapses [34]. While intracellular stimulation of pre-synaptic neurons (used in Bi and Poo's experiments) eliminates the potential involvement of multiple pre-synaptic cells, neurons in culture (also used in Bi and Poo's experiments) typically make multiple contacts with the post-synaptic cell [35]. It is safe to say that single synapses do not mediate currents on the order of Bi and Poo's larger data ($\sim$ 2500pA). This issue overlaps with debates concerning pre- and post-synaptic contributions to plasticity [36], [37] and the degree to which plasticity is

a graded phenomenon [18], [32], [34]. Varying the number of pairings in STDP experiments would directly address the latter. STDP protocols are clearly promising for the investigation of weight-dependent plasticity, but the potentially confounding effects of populations of synapses must be carefully addressed.

REFERENCES

[1] D. O. Hebb, *The Organisation of Behaviour*. John Wiley, New York, 1949.

[2] T. V. P. Bliss and T. J. Lomo, "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path," *Journal of Physiology*, vol. 232, pp. 331–356, 1973.

[3] G. Lynch, T. Dunwiddie, and V. Gribkoff, "Heterosynaptic depression: a postsynaptic correlate of long-term potentiation," *Nature*, vol. 266, pp. 737–739, 1977.

[4] W. B. Levy and O. Steward, "Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus," *Neuroscience*, vol. 8, pp. 791–797, 1983.

[5] H. Markram, J. Lbke, M. Frotscher, and B. Sakmann, "Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps," *Science*, vol. 275, no. 5297, pp. 213–215, 1997.

[6] G. qiang Bi and M. ming Poo, "Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type," *Journal of Neuroscience*, vol. 18, pp. 10464–10472, 1998.

[7] G. M. Wittenberg and S. S.-H. Wang, "Malleability of spike-timing-dependent plasticity at the ca3ca1 synapse," *The Journal of Neuroscience*, vol. 26, no. 24, pp. 6610–6617, 2006.

[8] W. M. Kistler and J. L. van Hemmen, "Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic action potentials," *Neural Computation*, vol. 12, pp. 385–405, 2000.

[9] M. C. W. van Rossum, G. Q. Bi, and G. G. Turrigiano, "Stable hebbian learning from spike timing-dependent plasticity," *Journal of Neuroscience*, vol. 20, no. 23, pp. 8812–8821, 2000.

[10] R. Gutig, R. Aharonov, S. Rotter, and H. Sompolinsky, "Learning input correlations through nonlinear temporally asymmetric hebbian plasticity," *Journal of Neuroscience*, vol. 23, pp. 3697–3714, May 2003.

[11] A. Burkitt, H. Meffin, and D. B. Grayden, "Spike-timing-dependent plasticity: The relationship to rate-based learning for models with weight dynamics determined by a stable fixed point," *Neural Computation*, vol. 16, pp. 885–940, 2004.

[12] Y. Huang and Y. H. Wang, "Input synchrony and spike-timing dependent plasticity," *The European Physical Journal B*, vol. 31, pp. 385–390, 2003.

[13] R. Legenstein, C. Naeger, and W. Maass, "What can a neuron learn with spike-timing-dependent plasticity?" *Neural Computation*, vol. 17, pp. 2337–2382, 2005.

[14] D. Debanne, B. H. Gahwiler, and S. M. Thompson, "Heterogeneity of synaptic plasticity at unitary ca3-ca1 and ca3-ca3 connections in rat hippocampal slice cultures," *Journal of Neuroscience*, vol. 19, no. 24, pp. 10664–10671, 1999.

[15] J. M. Montgomery, P. Pavlidis, and D. V. Madison, "Pair recordings reveal all-silent synaptic connections and the postsynaptic expression of long-term potentiation," *Neuron*, vol. 29, pp. 691–701, 2001.

[16] H.-X. Wang, R. C. Gerkin, D. W. Nauen, and G.-Q. Bi, "Coactivation and timing-dependent integration of synaptic potentiation and depression," *Nature Neuroscience*, vol. 8, no. 2, pp. 187–193, 2005.

[17] D. Standage, S. Jalil, and T. Trappenberg, "Computational consequences of experimentally derived spike-time and weight dependent plasticity rules," 2007, under submission.

[18] C. C. Petersen, R. C. Malenka, R. A. Nicoll, and J. J. Hopfield, "All-or-none potentiation at ca3-ca1 synapses," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 4732–4737, 1998.

[19] D. H. O'Connor, G. M. Wittenberg, and S. S.-H. Wang, "Graded bidirectional synaptic plasticity is composed of switch-like unitary events," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9679–9684, 2005.

[20] P. J. Sjostrom, G. G. Turrigiano, and S. B. Nelson, "Rate, timing, and cooperativity jointly determine cortical synaptic plasticity," *Neuron*, vol. 32, pp. 1149–1164, 2001.

[21] G.-Q. Bi, "Spatiotemporal specificity of synaptic plasticity: cellular rules and mechanisms," *Biological Cybernetics*, vol. 87, pp. 319–332, 2002.

[22] R. C. Froemke and Y. Dan, "Spike-timing-dependent synaptic modification induced by natural spike trains," *Nature*, vol. 416, no. 6879, pp. 433–438, 2002.

[23] E. M. Izhikevich and N. S. Desai, "Relating stdp to bcm," *Neural Computation*, vol. 15, no. 7, pp. 1511–1523, 2003.

[24] R. C. Froemke, I. A. Tsay, M. Raad, J. D. Long, and Y. Dan, "Contribution of individual spikes in burst-induced long-term synaptic modification," *Journal of Neurophysiology*, vol. 95, pp. 1620–1629, 2006.

[25] S. Song, K. D. Miller, and L. F. Abbott, "Competitive hebbian learning through spike-timing-dependent synaptic plasticity," *Nature Neuroscience*, vol. 3, no. 9, pp. 919–926, 2000.

[26] H. Z. Shouval, M. F. Bear, and L. N. Cooper, "A unified model of nmda receptor-dependent bidirectional synaptic plasticity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 16, pp. 10831–10836, 2002.

[27] W. Senn, "Beyond spike timing: the role of nonlinear plasticity and unreliable synapses," *Biological Cybernetics*, vol. 87, pp. 344–355, 2002.

[28] P.-M. Lledo, X. Zhang, T. C. Sudhof, R. C. Malenka, and R. A. Nicoll, "Postsynaptic membrane fusion and long-term potentiation," *Science*, vol. 279, pp. 399–403, 1998.

[29] S.-H. Shi, Y. Hayashi, R. S. Petralia, S. H. Zaman, R. J. Wenthold, K. Svoboda, and R. Malinow, "Rapid spine delivery and redistribution of ampa receptors after synaptic nmda receptor activation," *Science*, vol. 284, pp. 1811–1816, 1999.

[30] H.-Y. Man, J. W. Lin, W. H. Ju, G. Ahmadian, L. Liu, L. E. Becker, M. Sheng, and Y. T. Wang, "Regulation of ampa receptormediated synaptic transmission by clathrin-dependent receptor internalization," *Neuron*, vol. 25, pp. 649–662, 2000.

[31] M. I. Daw, R. Chittajallu, Z. A. Bortolotto, K. K. Dev, F. Duprat, J. M. Henley, G. L. Collingridge, and J. T. R. Isaac, "Regulation of ampa receptormediated synaptic transmission by clathrin-dependent receptor internalization," *Neuron*, vol. 28, pp. 873–886, 2000.

[32] J. M. Montgomery and D. V. Madison, "Discrete synaptic states define a major mechanism of synapse plasticity," *Trends in Neurosciences*, vol. 27, no. 12, pp. 744–750, 2004.

[33] C. F. Stevens and Y. Wang, "Changes in reliability of synaptic function as a mechanism for plasticity," *Nature*, vol. 371, pp. 704–707, 1994.

[34] D. H. O'Connor, G. M. Wittenberg, and S. S.-H. Wang, "Dissection of bidirectional synaptic plasticity into saturable unidirectional processes," *Journal of Neurophysiology*, vol. 94, pp. 1565–1573, 2005.

[35] D. Debanne, B. H. Gahwiler, and S. M. Thompson, "Cooperative interactions in the induction of long-term potentiation and depression of synaptic excitation between hippocampal ca3-ca1 cell pairs in vitro," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pp. 11225–11230, 1996.

[36] N. J. Emptage, C. A. Reid, A. Fine, and T. V. Bliss, "Optical quantal analysis reveals a presynaptic component of ltp at hippocampal schaffer-associational synapses," *Neuron*, vol. 38, pp. 797–804, 2003.

[37] R. A. Nicoll, S. Tomita, and D. S. Bredt, "Auxiliary subunits assist ampa-type glutamate receptors," *Science*, vol. 311, pp. 1253–1256, 2006.