Reducing Spike Train Variability: A Computational Theory Of Spike-Timing Dependent Plasticity *

Sander M. Bohte [a]   Michael C. Mozer [b]

[a] CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
[b] Dept. of Computer Science, University of Colorado, Boulder, USA

Abstract

Experimental studies have observed synaptic potentiation when a presynaptic neuron fires shortly before a postsynaptic neuron, and synaptic depression when the presynaptic neuron fires shortly after. The dependence of synaptic modulation on the precise timing of the two action potentials is known as *spike-timing dependent plasticity* or *STDP*. We derive STDP from a simple computational principle: synapses adapt so as to minimize the postsynaptic neuron's response variability to a given presynaptic input, causing the neuron's output to become more reliable in the face of noise.

Using an objective function that minimizes response variability and the biophysically realistic spike-response model of Gerstner (2001), we simulate neurophysiological experiments and obtain the characteristic STDP curve along with other phenomena including the reduction in synaptic plasticity as synaptic efficacy increases. We compare our account to other efforts to derive STDP from computational principles, and argue that our account provides the most comprehensive coverage of the phenomena. Thus, reliability of neural response in the face of noise may be a key goal of unsupervised cortical adaptation.

---

1 Introduction

Experimental studies have observed synaptic potentiation when a presynaptic neuron fires shortly before a postsynaptic neuron, and synaptic depression when the presynaptic neuron fires shortly after (Markram, Lübke, Frotscher, & Sakmann, 1997; Bel1, Han, Sugawara, & Grant, 1997; Zhang, Tao, Holt, Harris, & Poo, 1998; G.-q. Bi & Poo, 1998; Debanne, Gähwiler, & Thompson, 1998; Feldman, 2000; Sjöström, Turrigiano, & Nelson, 2001; Nishiyama, Hong, Mikoshiba, Poo, & Kato, 2000). The dependence of synaptic modulation on the precise timing of the two action potentials, known as *spike-timing dependent plasticity* or *STDP*, is depicted in Figure 1. Typically, plasticity is observed only when the presynaptic and postsynaptic spikes occur within a 20–30 ms time window, and the transition from potentiation to depression is very rapid. The effects are long lasting, and are therefore referred to as long-term potentiation (LTP) and depression (LTD). An important observation is that the relative magnitude of the LTP component of STDP decreases with increased synaptic efficacy between presynaptic and postsynaptic neuron, whereas the magnitude of LTD remains roughly constant (G.-q. Bi & Poo, 1998). This finding has led to the suggestion that the LTP component of STDP might best be modeled as additive, whereas the LTD component is better modeled as being multiplicative (Kepecs, van Rossum, Song, & Tegner, 2002). For detailed reviews of STDP, see (G.-q. Bi & Poo, 2001; Roberts & Bell, 2002; Dan & Poo, 2004).

Because these intriguing findings appear to describe a fundamental learning mechanism in the brain, a flurry of models have been developed that focus on different aspects of STDP. A number of studies focus on biochemical models that explain the underlying mechanisms giving rise to STDP: (Senn, Markram, & Tsodyks, 2000; G.-Q. Bi, 2002; Karmarkar, Najarian, & Buonomano, 2002; Saudargiene, Porr, & Wörgötter, 2004; Porr & Wörgötter, 2003). Many researchers have also focused on models that explore the consequences of a STDP-like learning rules in an ensemble of spiking neurons, i.e. (Gerstner, Kempter, van Hemmen, & Wagner, 1996; Kempter, Gerstner, & van Hemmen, 1999; Song, Miller, & Abbott, 2000; van Rossum, Bi, & Turrigiano, 2000; Kempter, Gerstner, & van Hemmen, 2001; Izhikevich & Desai, 2003; Abbott & Gerstner, 2004; Burkitt, Meffin, & Grayden, 2004; Shon, Rao, & Sejnowski, 2004; Legenstein, Naeger, & Maass, n.d.), a comprehensive review of the different types and conclusions can be found in Porr and Wörgötter (2003). Finally, a recent trend is to propose models that provide
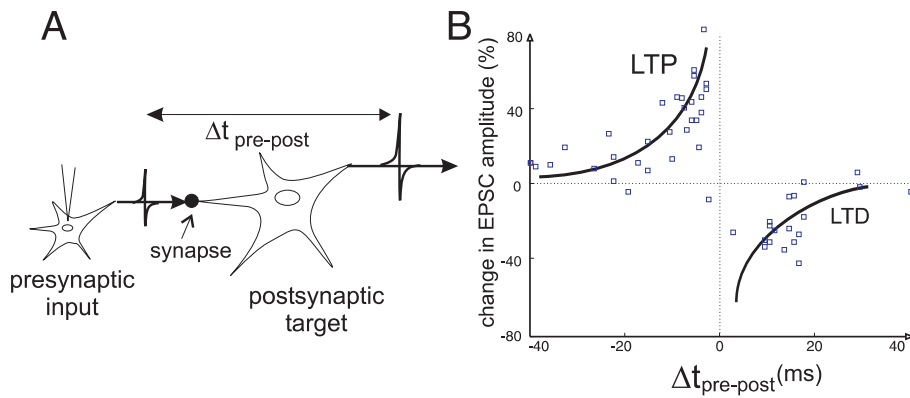
Figure 1: (A) Measuring STDP experimentally: presynaptic and postsynaptic spike pairs are repeatedly induced at a fixed interval $\Delta t_{pre-post}$, and the resulting change to the strength of the synapse is assessed; (B) change in synaptic strength after repeated spike pairing as a function of the difference in time between the presynaptic and postsynaptic spikes: a presynaptic before postsynaptic spike induces LTP, postsynaptic before presynaptic LTD (data redrawn from Zhang *et al.*, 1998). We have superimposed an exponential fit of LTP and LTD.

fundamental computational justifications for STDP. This article proposes a novel justification and we explore the consequences of this justification in detail.

Most commonly, STDP is viewed as a type of asymmetric Hebbian learning with a temporal dimension. However, this perspective is hardly a fundamental computational rationale, and one would hope that such an intuitively sensible learning rule would emerge from a first-principle computational justification.

Several researchers have tried to derive a learning rule yielding STDP from first principles. Dayan and Häusser (2004) show that STDP can be viewed as an optimal noise-removal filter for certain noise distributions. However, even small variation from these noise distributions yield quite different learning rules, and the noise statistics of biological neurons are unknown. Similarly, Porr and Wörgötter (2003) propose an unsupervised learning rule based on the correlation of bandpass-filtered inputs with the derivative of the output, and show that the weight change rule is qualitatively similar to STDP.

Hopfield and Brody (2004) derive learning rules that implement ongoing network self-repair. In some circumstances, a qualitative similarity to STDP is found, but the shape of the learning rule depends on both network architecture and task. Eisele (private communication) has shown that an STDP-like learning rule can be derived from the goal of maintaining the relevant connections in a network.

Rao and Sejnowski (1999, 2001) suggest that STDP may be related to prediction, in particular to temporal difference (TD) learning. They argue that STDP emerges when a neuron attempts to predict its membrane potential at some time t from the potential at time $t - \Delta t$. As Dayan (2002) points out however, temporal difference learning depends on an estimate of the prediction error, which will be very hard to obtain. Rather, a quantity that might be called an "activity difference" can be computed, and the learning rule is then better characterized as a "correlational learning rule between the stimuli, and the *differences* in successive outputs" (Dayan, 2002) (see also Porr and Wörgötter (2003), Appendix B). Furthermore, Dayan argues that for true prediction, the model has to show that the learning rule works for biologically realistic timescales. The qualitative nature of the modeling makes it unclear whether a quantitative fit can be obtained. Lastly, the derived difference rule is inherently instable, as it does not impose any bounds on synaptic efficacies; also, STDP emerges only for a narrow range of Δt values.

4

Chechik (2003) relates STDP to information theory via maximization of mutual information between input and output spike trains. This approach derives the LTP portion of STDP, but fails to yield the LTD portion. Nonetheless, an information-theoretic approach is quite elegant and has proven valuable in explaining other neural learning phenomena (e.g., Linsker, 1989).

The account we describe in this paper also exploits an information-theoretic approach. We are not the only ones to appreciate the elegance of information-theoretic accounts. In parallel with a preliminary presentation of our work at the NIPS 2004 conference, two quite similar information-theoretic accounts also appeared (Bell & Parra, 2005; Toyoizumi, Pfister, Aihara, & Gerstner, 2005). It will be easiest to explain the relationship of these accounts to our own once we have presented ours.

The computational approaches of Chechik, Dayan & Häusser, and Porr & Wörgötter are all premised on a rate-based neuron model that disregards the relative timing of spikes. It seems quite odd to argue for STDP using neural firing *rate*: if spike timing is irrelevant to information transmission, then STDP is likely an artifact and is not central to understanding mechanisms of neural computation. Further, as Dayan and Häusser (2004) note, because STDP is not quite additive in the case of multiple input or output spikes that are near in time (Froemke & Dan, 2002), one should consider interpretations that are based on individual spikes, not aggregates over spike trains.

In this paper, we present an alternative theoretical motivation for STDP from a spike-based neuron model that takes the specific times of spikes into account. We conjecture that a fundamental objective of cortical computation is to achieve *reliable* neural responses, that is, neurons should produce the identical response—both in the number and timing of spikes—given a fixed input spike train. Reliability is an issue if neurons are affected by noise influences, because noise leads to variability in a neuron's dynamics and therefore in its response. Minimizing this variability will reduce the effect of noise and will therefore increase the informativeness of the neuron's output signal. The source of the noise is not important; it could be intrinsic to a neuron (e.g., a time-varying threshold) or it could originate in unmodeled external sources that cause fluctuations in the membrane potential uncorrelated with a particular input.

We are not suggesting that increasing neural reliability is the *only* objective of learning. If it were, a neuron would do well to shut off and give no response regardless of the input. Rather, reliability is but one of many

objectives that learning tries to achieve. This form of unsupervised learning must, of course, be complemented by other unsupervised, supervised and reinforcement learning objectives that allow an organism to achieve its goals and satisfy drives. We return to this issue below and in our conclusions section.

We derive STDP from the following computational principle: synapses adapt so as to minimize the variability in the timing of the spikes of the postsynaptic neuron's output in response to given presynaptic input spike-trains. This variability reduction causes the response of a neuron to become more deterministic, and less sensitive to noise, which provides an obvious computational benefit.

In our simulations, we follow the methodology of neurophysiological experiments. This approach leads to a detailed fit to key experimental results. We model not only the shape (sign and time course) of the STDP curve, but also the fact that potentiation of a synapse depends on the efficacy of the synapse—it decreases with increased efficacy. In addition to fitting these key STDP phenomena, the model allows us to make predictions regarding the relationship between properties of the neuron and the shape of the STDP curve. The detailed quantitative fit to data makes our work unique among first-principle computational accounts.

Before delving into the details of our approach, we attempt to give a basic intuition about the approach. Noise in spiking neuron dynamics leads to variability in the number and timing of spikes. Given a particular input, one spike train might be more likely than others, but the output is nondeterministic. By the response variability minimization principle, adaptation should reduce the likelihood of these other possibilities. To be concrete, consider a particular experimental paradigm. In Zhang et al. (1998), a presynaptic neuron is identified with a weak synapse to a post neuron, such that this presynaptic input is unlikely to cause the postsynaptic neuron to fire. However, the postsynaptic neuron can be induced to fire via a second presynaptic connection. In a typical trial, the presynaptic neuron is induced to fire a single spike, and with a variable delay, the postsynaptic neuron is also induced to fire (typically) a single spike. To increase the likelihood of the observed postsynaptic response other response possibilities must be suppressed.

With presynaptic input preceding the postsynaptic spike, the most likely alternative response is no output spikes at all. Increasing the synaptic connection weight should then reduce the possibility of this alternative response. With presynaptic input following the postsynaptic spike, the most likely al-

ternative response is a second output spike. Decreasing the synaptic connection weight should reduce the possibility of this alternative response. Because both of these alternatives become less likely as the lag between pre and post spikes is increased, one would expect that the magnitude of synaptic plasticity diminishes with the lag, as is observed in the STDP curve.

Our approach to reducing response variability given a particular input pattern involves computing the gradient of synaptic weights with respect to a differentiable model of spiking neuron behavior. We use the Spike Response Model (SRM) of Gerstner (2001) with a stochastic threshold, where the stochastic threshold models fluctuations of the membrane potential or the threshold outside of experimental control. For the stochastic SRM, the response probability is differentiable with respect to the synaptic weights, allowing us to calculate the gradient that reduces response variability with respect to the weights. Learning is presumed to take a gradient step to reduce the response variability. In modeling neurophysiological experiments, we demonstrate that this learning rule yields the typical STDP curve. We can predict the relationship between the exact shape of the STDP curve and physiologically measurable parameters, and we show that our results are robust to the choice of the few free parameters of the model.

Many important machine learning algorithms in the literature seek local optimizers. It is often the case that the initial conditions – which determine which local optimizer will be found – can be controled to avoid unwanted local optimizers. For example, with neural networks, weights are initialized near the origin; large initial weights would lead to degenerate solutions. And K-means has many degenerate and suboptimal solutions; consequently, careful initialization of cluster centers is required.

In the case of our model's learning algorithm, the initial conditions also avoid the degenerate local optimizer. These initial conditions correspond to the original weights of the synaptic connections and are constrained by the specific methodology of the experiments that we model: the subthreshold input must have a small but nonzero connection strength, and the suprathreshold input must have a large connection strength ($< 10\%$, $> 70\%$ probability of activating the target respectively). Given these conditions, the local optimizer that our learning algorithm discovers is an extremely good fit to the experimental data.

In parallel with our work, two other groups of authors have proposed explanations of STDP in terms of neurons maximizing an information-theoretic measure for the Spike-Response Model (Bell & Parra, 2005; Toyoizumi et al.,

2005). Toyoizumi et al. (2005) maximize the mutual information of the input and output between a pool of presynaptic neurons and a single postsynaptic output neuron whereas Bell and Parra (2005) maximize sensitivity between a pool of (possibly correlated) presynaptic neurons and a pool of postsynaptic neurons. Bell and Parra (2005) use a causal SRM model and do not obtain the LTD component of STDP. As we will show, when the objective function is minimization of (conditional) response variability, obtaining LTD critically depends on a stochastic neural response. In the derivation of Toyoizumi et al. (2005), LTD – which is very weak in magnitude– is attributed to the refractoriness of the spiking neuron (via the autocorrelation function), where they use questionably strong and enduring refractoriness. In our framework, refractoriness suppresses noise in the neuron after spiking, and we show that in our simulations strong refraction in fact diminishes the LTD component of STDP. Furthermore, the mathematical derivation of Toyoizumi et al. (2005) is valid only for an essentially constant membrane potential with small fluctuations, a condition which is clearly violated in experimental conditions studied by neurophysiologists. It is unclear whether the derivation would hold under more realistic conditions.

Neither of these approaches thus far succeeds in quantitatively modeling specific experimental data with neurobiologically-realistic timing parameters, and neither explains the relative reduction of STDP as the synaptic efficacy increases as we do. Nonetheless, these models make an interesting contrast to ours by suggesting a computational principle of optimization of information transmission, as contrasted with our principle of neural response variability reduction. Experimental tests might be devised to distinguish between these competing theories.

In Section 2 we describe the sSRM, and in Section 3 we derive the minimal entropy gradient. In Section 4 we describe the STDP experiment which we simulate in Section 5. We conclude with Section 6.

2 The Stochastic Spike Response Model

The Spike Response Model (SRM), defined by Gerstner (2001), is a generic integrate-and-fire model of a spiking neuron that closely corresponds to the behavior of a biological spiking neuron and is characterized in terms of a small set of easily interpretable parameters (Jolivet, Lewis, & Gerstner, 2003; Paninski, Pillow, & Simoncelli, 2005). The standard SRM formulation de-
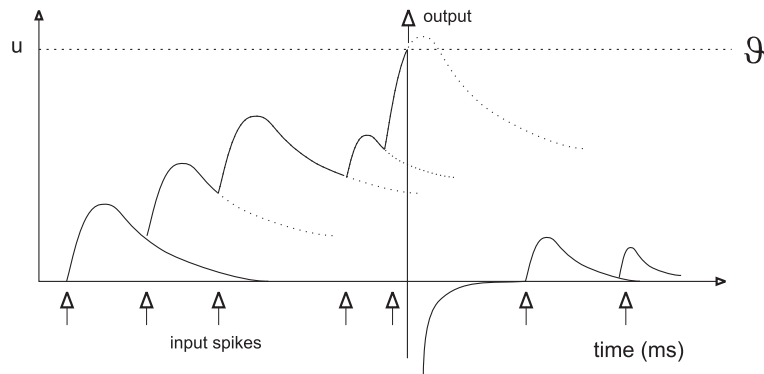
Figure 2: Membrane potential $u(t)$ of a neuron as a sum of weighted excitatory PSP kernels due to impinging spikes. Arrival of PSP's marked by arrows; once the membrane potential reaches threshold it is reset and a reset function η is added to model the recovery effects of the threshold.

scribes the temporal evolution of the membrane potential based on past neuronal events, specifically as a weighted sum of postsynaptic potentials (PSP's) modulated by reset and threshold effects of previous postsynaptic spiking events. The general idea is depicted in Figure 2; formally (following Gerstner, 2001), the membrane potential $u_i(t)$ of cell i at time t is defined as:

$$u_i(t) = \sum_{f_i \in \mathcal{G}_i^t} \eta(t - f_i) + \sum_{j \in \Gamma_i} w_{ij} \sum_{f_j \in \mathcal{G}_j^t} \epsilon(t|f_j, \mathcal{G}_i^t), \tag{1}$$

where Γ_i is the set of inputs connected to neuron i, $\mathcal{G}_i^t$ is the set of times prior to t that a neuron i has spiked, with firing times $f_i \in \mathcal{G}_i^t$; w_{ij} is the synaptic weight from neuron j to neuron i, $\epsilon(t|f_j, \mathcal{G}_i^t)$ is the PSP in neuron i due to an input spike from neuron j at time f_j given postsynaptic firing history $\mathcal{G}_i^t$, and $\eta(t - f_i)$ is the refractory response due to the postsynaptic spike at time f_i.

To model the postsynaptic potential ϵ in a leaky-integrate-and-fire neuron, a spike of presynaptic neuron j emitted at time f_j generates a postsynaptic current $\alpha(t)$ for a presynaptic spike arriving at f_j for $t > f_j$. In the absence of postsynaptic firing this kernel (following Gerstner and Kistler (2002), Eqs (4.62)–(4.56), pp. 114–115) can be computed as:

9

$$\epsilon(t|f_j) = \int_{f_j}^{t} \exp\left(-\frac{s-f_j}{\tau_m}\right) \alpha(s-f_j)\, ds, \tag{2}$$

where τ_m is the decay time of the postsynaptic neuron's membrane potential. Consider an exponentially decaying postsynaptic current $\alpha(t)$ of the form

$$\alpha(t) = \frac{1}{\tau_s} \exp\left(-\frac{t}{\tau_s}\right) \mathcal{H}(t), \tag{3}$$

(Figure 3A), where τ_s is the decay time of the current and $\mathcal{H}(t)$ is the Heaviside function. In the absence of postsynaptic firing, this current contributes a postsynaptic potential of the form

$$\epsilon(t|f_j) = \frac{1}{1 - \tau_s/\tau_m} \left[\exp\left(-\frac{(t-f_j)}{\tau_m}\right) - \exp\left(-\frac{(t-f_j)}{\tau_s}\right) \right] \mathcal{H}(t-f_j), \tag{4}$$

with current decay time-constant τ_s and decay time-constant τ_m.

When the postsynaptic neuron fires after the presynaptic spike arrives—at some time $\hat{f}_i$ following presynaptic spike at time f_j— the membrane potential is reset, and only the remaining synaptic current $\alpha(t')$ for $t' > \hat{f}_i$ is integrated in Equation (2). Following Gerstner, 2001 (Section 4.4, Equation 1.66), the PSP that takes such postsynaptic firing into account can be written as

$$\epsilon(t|f_j, \hat{f}_i) = \begin{cases} \epsilon(t|f_j) & \hat{f}_i < f_j, \\ \exp\left(-\frac{(f_j - \hat{f}_i)}{\tau_s}\right) \epsilon(t|f_j) & \hat{f}_i \geqslant f_j. \end{cases} \tag{5}$$

This function is depicted in Figure 3B, for the cases when a postsynaptic spike occurs both before and after the presynaptic spike. In principle, this formulation can be expanded to include the postsynaptic neuron firing more than once after the onset of the postsynaptic potential. However, for fast current decay times τ_s, it is only useful to consider the residual current input for the first postsynaptic spike after onset and assume that any further postsynaptic spiking is modeled by a postsynaptic potential reset to zero from that point on.

The reset response $\eta(t)$ models two phenomena. First, a neuron can be in a refractory period: it simply cannot spike again for about a millisecond after a spiking event. Second, after the emission of a spike, the threshold of the neuron may initially be elevated and then recover to the original value (Kandel, Schwartz, & Jessell, 2000). The SRM models this behavior as
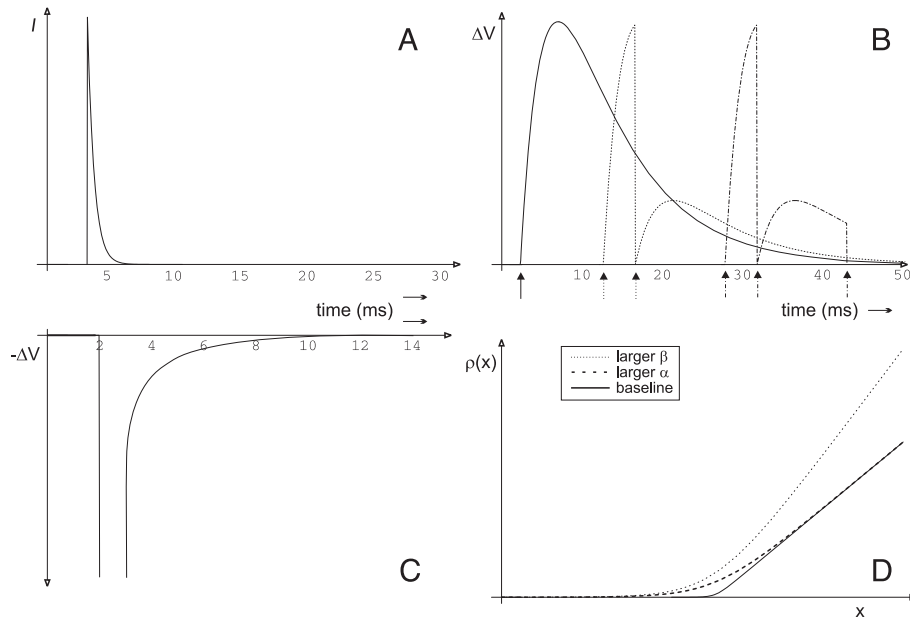
10

Figure 3: (A) $\alpha(t)$ function: synaptic input modeled as exponentially decaying current. (B) Postsynaptic potential due to a synaptic input in the absence of postsynaptic firing (solid line), and with postsynaptic firing once and twice (dotted respectively dashed lines, postsynaptic spikes indicated by arrow) (C) reset function $\eta(t)$, (D) spike probability $\rho(u)$ as a function of potential u for different values of α and β parameters.

negative contributions to the membrane potential (Equation 1): with $s = t - \hat{f}_i$ denoting the time since the postsynaptic spike, the refractory reset function is defined as (Gerstner, 2001):

$$\eta(s) = \begin{cases} \mathcal{U}_{abs} & 0 < s < \delta_r \\ \mathcal{U}_{abs} \, \exp\left(-\frac{s+\delta_r}{\tau_r^f}\right) + \mathcal{U}_r \, \exp\left(-\frac{s}{\tau_r^s}\right) & s \geq \delta_r, \end{cases} \quad (6)$$

where a large negative impulse $\mathcal{U}_{abs}$ models the absolute refractory period, with duration δ_r; the absolute refractory contribution smoothly resets via a fast decaying exponential with time constant τ_r^f. The term $\mathcal{U}_r$ models the slow exponential recovery of the elevated threshold with time constant τ_r^s. The function η is depicted in Figure 3C.

We made a minor modification to the SRM described in (Gerstner, 2001) by relaxing the constraint that $\tau_r^s = \tau_m$, and also by smoothing the absolute refractory function (such smoothing is mentioned in Gerstner (2001) but is not explicitly defined). In all simulations, we use $\delta_r = 1$ms, $\tau_r^s = 3$ms, and $\tau_r^f = 0.25$ ms (in line with estimates for biological neurones, Kandel et al. (2000), the smoothing parameter was chosen to be fast compared to τ_r^s).

The SRM we just described is deterministic. Gerstner (2001) introduces a stochastic variant of the SRM (sSRM) by incorporating the notion of a stochastic firing threshold: given membrane potential $u_i(t)$, the probability of the neuron firing at time t is specified by $\rho\big(u_i(t)\big)$. Herrmann and Gerstner (2001) find that for a reasonable escape-rate noise model of the integration of current in real neurons, the probability of firing is small and constant for small potentials, but around a threshold ϑ, the probability increases linearly with the potential. In our simulations, we use such a function:

$$\rho(v) = \frac{\beta}{\alpha} \left\{ \ln[1 + \exp\big(\alpha \, (\vartheta - v)\big)] - \alpha \, (\vartheta - v) \right\}, \quad (7)$$

where α determines the abruptness of the constant-to-linear transition in the neighborhood of threshold ϑ, and β determines the slope of the linear increase beyond ϑ. This function is depicted in Figure 3D for several values of α and β. We remark that we also conducted simulation experiments with sigmoidal and exponential density functions, and found no qualitative difference in the results.

12

3   Minimizing Conditional Entropy

We now derive the rule for adjusting the weight from a presynaptic input neuron j to a postsynaptic neuron i, so as to minimize the entropy of i's response given a particular spike train from j.

A spike train is described by the set of all times at which a neuron i emitted spikes within some interval between 0 and T, denoted $\mathcal{G}_i^T$. We assume the interval is wide enough that the occurrence of spikes outside the interval do not influence the state of a neuron within the interval (e.g., through threshold reset effects). This assumption allows us to treat intervals as independent of each other. The set of input spikes received by neuron i during this interval is denoted $\mathcal{F}_i^T$, which is just the union of all output spike trains of connected presynaptic neurons j: $\mathcal{F}_i^T = \bigcup \mathcal{G}_j^T \ \forall j \in \Gamma_i$.

Given input spikes $\mathcal{F}_i^T$, the stochastic nature of neuron i may lead not only to the observed response $\mathcal{G}_i^T$, but a range of other possibilities. Denoting the set of possible responses Ω_i, where $\mathcal{G}_i^T \in \Omega_i$. Further, let binary variable $\sigma(t)$ denote the state of the neuron in the time interval $[t, t + \Delta t)$, where $\sigma(t) = 1$ means the neuron spikes and $\sigma(t) = 0$ means no spike. A response $\xi \in \Omega_i$ is then equivalent to $[\sigma(0), \sigma(\Delta t), \dots, \sigma(T)]$.

Given a probability density $p(\xi)$ over all possible responses ξ, the differential entropy of neuron i's response conditional upon input $\mathcal{F}_i^T$ is then defined as:

$$h(\Omega_i | \mathcal{F}_i^T) = - \int_{\Omega_i} p(\xi) \log\big(p(\xi)\big) \, d\xi. \tag{8}$$

According to our hypothesis, a neuron adjusts its weights so as to minimize the conditional response variability. Such an adjustment is obtained by performing gradient descent on the weighted likelihood of the response, – which corresponds to the conditional entropy – with respect to the weights:

$$\Delta w_{ij} = -\gamma \frac{\partial h(\Omega_i | \mathcal{F}_i^T)}{\partial w_{ij}}, \tag{9}$$

with learning rate γ.

In this section, we compute the right hand side of Equation (9) for an sSRM neuron. Substituting the entropy definition of Equation (8) into Equa-

tion (9), we obtain:

$$
\begin{aligned}
\frac{\partial h(\Omega_i|\mathcal{F}_i^T)}{\partial w_{ij}} &= -\frac{\partial}{\partial w_{ij}} \int_\Omega p(\xi) \log\big(p(\xi)\big)\, d\xi \\
&= -\int_{\Omega_i} p(\xi) \frac{\partial \log\big(p(\xi)\big)}{\partial w_{ij}} \Big(\log\big(p(\xi)\big) + 1\Big)\, d\xi. \qquad (10)
\end{aligned}
$$

We closely follow Xie and Seung (2004) to derive $\frac{\partial \log\big(p(\xi)\big)}{\partial w_{ij}}$ for a differentiable neuron model firing at times $\mathcal{G}_i^T$. First, we factorize $p(\xi)$:

$$
p(\xi) = \prod_{t=0}^{T} P(\sigma(t)|\{\sigma(t'), \forall t' < t\}). \qquad (11)
$$

The states $\sigma(t)$ are conditionally independent as the probability for a neuron i to fire during $[t, t + \Delta t)$ is determined by the spike probability density of the membrane potential:

$$
\sigma(t) = \begin{cases} 1 & \text{with probability} \quad p_i = \rho_i(t)\Delta t, \\ 0 & \text{with probability} \quad p_i = 1 - p_i(t), \end{cases}
$$

with $\rho_i(t)$ shorthand for the spike probability density of the membrane potential, $\rho\big(u_i(t)\big)$; this equation holds for sufficiently small $\sigma(t)$ (see also Xie and Seung (2004) for more details).

We note further that:

$$
\frac{\partial \ln(p(\xi))}{\partial w_{ij}} \equiv \frac{1}{p(\xi)} \frac{\partial p(\xi)}{\partial w_{ij}}
$$

and

$$
\frac{\partial \rho_i(t)}{\partial w_{ij}} = \frac{\partial \rho_i(t)}{\partial u_i(t)} \frac{\partial u_i(t)}{\partial w_{ij}}, \qquad (12)
$$

it is straightforward to derive:

$$
\begin{aligned}
\frac{\partial \log(p(\xi))}{\partial w_{ij}} &= \int_{t=0}^{T} \frac{\partial \rho_i(t)}{\partial u_i(t)} \frac{\partial u_i(t)}{\partial w_{ij}} \frac{\left(\sum_{f_i \in \mathcal{F}_i^T} \delta(t - f_i) - \rho_i(t)\right)}{\rho_i(t)}\, dt, \\
&= -\int_{t=0}^{T} \rho_i'(t)\, \epsilon(t|f_j, f_i)\, dt + \sum_{f_i \in \mathcal{F}_i^T} \frac{\rho_i'(f_i)}{\rho_i(f_i)}\, \epsilon(f_i|f_j, f_i), \quad (13)
\end{aligned}
$$

14

where $\rho_i'(t) \equiv \frac{\partial \rho_i(t)}{\partial u_i(t)}$ and $\delta(t - f_i)$ is the Dirac delta, and we use that in the sSRM formulation,

$$\frac{\partial u_i(t)}{\partial w_{ij}} = \epsilon(t|f_j, f_i).$$

The term $\rho_i'(t)$ in Equation (13) can be computed for any differentiable spike probability function. In the case of Equation (7),

$$\rho_i'(t) = \frac{\beta}{1 + \exp(\alpha(\vartheta - u_i(t))}.$$

Substituting our model for $\rho_i(t), \rho_i'(t)$ from Equation (7) into Equation (13), we obtain

$$\frac{\partial \log(p(\xi))}{\partial w_{ij}} = -\beta \int_{t=0}^{T} \frac{\epsilon(t|f_j, f_i)}{1 + \exp[\alpha(\vartheta - u_i(t))]} dt + \tag{14}$$

$$+ \sum_{f_i \in \mathcal{G}_i^T} \frac{\epsilon(f_i|f_j, f_i)}{\alpha \left\{ \ln\big(1 + \exp[\alpha(\vartheta - u_i(f_i))]\big) - \alpha(\vartheta - u_i(f_i)) \right\}(1 + \exp[\alpha(\vartheta - u_i(f_i))])}.$$

Equation (14) can be substituted into Equation (10), which when integrated provides the gradient-descent weight update that implements conditional entropy minimization (Equation (9)).

The hypothesis under exploration is that this gradient-descent weight update yields STDP. Unfortunately, an analytic solution to Equation (10) (and hence Equation (9)) is not readily obtained. Nonetheless, numerical methods can be used to obtain a solution.

We are not suggesting a neuron performs numerical integration of this sort in real time. It would be preposterous to claim biological realism for an instantaneous integration over all possible responses $\xi \in \Omega_i$, as specified by Equation (10). Consequently, we have a dilemma: What use is a computational theory of STDP if the theory demands intensive computations that could not possibly be performed by a neuron in real time? This dilemma can be circumvented in two ways. First, the resulting learning rule might be cached in some form through evolution so that the computation is not necessary. That is, the solution—the STDP curve itself—may be built into a neuron. As such, our computational theory provides an argument for why neurons have evolved to implement the STDP learning rule. Second, the specific response produced by a neuron on a single trial might be considered to

15

be a sample from the distribution $p(\xi)$, and the integration in Equation (10) can be performed by a sampling process over repeated trials; each trial would produce a stochastic gradient step.

3.1 Numerical Computation

In this section, we describe the procedure for numerically evaluating Equation (9) via Simpson's integration (Hennion, 1962).

This integration is performed over the set of possible responses Ω_i (Equation (10)) within the time interval $[0 \ldots T]$. The set Ω_i can be divided into disjoint subsets Ω_i^n which contain exactly n spikes: $\Omega_i = \bigcup \Omega_i^n \; \forall \, n$.

Using this breakdown,

$$
\begin{aligned}
\frac{\partial h(\Omega_i | \mathcal{F}_i^T)}{\partial w_{ij}} &= -\int_{\Omega_i} g(\xi)\big(\log\big(g(\xi)\big)+1\big)\frac{\partial \log\big(g(\xi)\big)}{\partial w_{ij}}d\xi, \\
&= -\sum_{n=0}^{n=\infty} \int_{\Omega_i^n} g(\xi)\big(\log\big(g(\xi)\big)+1\big)\frac{\partial \log\big(g(\xi)\big)}{\partial w_{ij}}d\xi. \quad (15)
\end{aligned}
$$

It is illustrative to walk through the alternatives. For $n = 0$, there is only one response given the input. Assuming the probability of $n = 0$ spikes is p_0, the $n = 0$ term of Equation (15) reads:

$$
\frac{\partial h(\Omega_i | \mathcal{F}_i^T)}{\partial w_{ij}} = p_0\big(\log\big(p_0\big)+1\big) \int_{t=0}^{T} -\rho_i'(t)\,\epsilon(t|f_j, f_i)dt. \quad (16)
$$

The probability p_0 is the probability of the neuron not having fired between $t = 0$ and $t = T$ given inputs $\mathcal{F}_i^T$ resulting in membrane potential $u_i(t)$ and hence probability of firing at time t of $\rho(u_i(t))$:

$$
p_0 = S[0, T] = \exp\Big(-\int_{t=0}^{T} \rho\big(u_i(t)\big)dt\Big), \quad (17)
$$

which is equal to the survival function S for a non-homogenous Poisson process with probability density $\rho(u_i(t))$ for $t = [0 \ldots T]$. (We use the inclusive/exclusive notation for S: S(0,T) computes the function excluding the endpoints, S[0,T] is inclusive.)

For $n = 1$, we must consider all responses containing exactly one output spike: $\mathcal{G}_i^T = \{f_i^1\}, f_i^1 \in [0, T]$. Assuming that neuron i fires only at time f_i^1

with probability $p_1(f_i^1)$, the $n = 1$ term of Equation (15) reads:

$$\frac{\partial h(\Omega_i|\mathcal{F}_i^T)}{\partial w_{ij}} = \int_{f_i^1=0}^{f_i^1=T} p_1(f_i^1)\Big(\log\big(p_1(f_i^1)\big) + 1\Big) \Big[\int_{t=0}^{T} -\rho_i'(t)\,\epsilon(t|f_j, f_i^1)dt +$$

$$\frac{\rho_i'(f_i^1)}{\rho_i(f_i^1)}\,\epsilon(f_i^1|f_j, f_i^1)\Big]df_i^1. \tag{18}$$

The probability $p_1(f_i^1)$ is computed as:

$$p_1(f_i^1) = S[0, f_i^1]\,\rho_i(f_i^1)\,S(f_i^1, T), \tag{19}$$

where the membrane potential now incorporates one reset at $t = f_i^1$:

$$u_i(t) = \eta(t - f_i^1) + \sum_{j\in\Gamma_i} w_{ij} \sum_{f_j\in\mathcal{F}_j^t} \epsilon(t|f_j, f_i^1)$$

For $n = 2$, we must consider all responses containing exactly two output spikes: $\mathcal{G}_i^T = \{f_i^1, f_i^2\}$ for $f_i^1, f_i^2 \in [0, T]$. Assuming that neuron i fires at f_i^1 and f_i^2 with probability probability $p_2(f_i^1, f_i^2)$, the $n = 2$ term of Equation (15) reads:

$$\frac{\partial h(\Omega_i|\mathcal{F}_i^T)}{\partial w_{ij}} = \int_{f_i^1=0}^{f_i^1=T} \int_{f_i^2=f_i^1}^{f_i^2=T} p_2(f_i^1, f_i^2)\Big(\log\big(p_2(f_i^1, f_i^2)\big) + 1\Big) \Big[\int_{t=0}^{T} -\rho_i'(t)\,\epsilon(t|f_j, f_i^1, f_i^2)dt +$$

$$+ \quad \frac{\rho_i'(f_i^1)}{\rho_i(f_i^1)}\,\epsilon(f_i^1|f_j, f_i^1, f_i^2) + \frac{\rho_i'(f_i^2)}{\rho_i(f_i^2)}\,\epsilon(f_i^2|f_j, f_i^1, f_i^2)\Big]df_i^1\,df_i^2. \tag{20}$$

The probability $p_2(f_i^1, f_i^2)$ can again be expressed in terms of the survival function:

$$p_2(f_i^1, f_i^2) = S[0, f_i^1]\,\rho_i(f_i^1)\,S[f_i^1, f_i^2]\rho_i(f_i^2)S(f_i^2, T), \tag{21}$$

with $u_i(t) = \eta(t - f_i^1) + \eta(t - f_i^2) + \sum_{j\in\Gamma_i} w_{ij} \sum_{f_j\in\mathcal{F}_j^t} \epsilon(t|f_j, f_i^1, f_i^2)$.

This procedure can be extended for $n > 2$ following the pattern above. In our simulation of the STDP experiments, the probability of obtaining zero, one or two spikes already accounted for 99.9% of all possible responses, adding the responses of three spike ($n = 3$) accounted for all possible responses got this number up to ≈ 99.999, which is close to the accuracy of our numerical computation. In practice, we found that taking into account $n = 3$ already

had no significant contribution to computing Δw, and we did not compute higher order terms as the cumulative probability of these responses was below our numerical precision. For the results we present later we only used terms $n \leq 2$, we demonstrate that this is sufficient in Appendix A.

In this section, we have replaced an integral over possible spike sequences Ω_i with an integral over the time of two output spikes, f_i^1 and f_i^2, which we compute numerically.

4 Simulation Methodology

We modeled in detail the experiment of Zhang et al. (1998) involving asynchronous co-stimulation of convergent inputs. In this experiment, depicted in Figure 4, a *postsynaptic* neuron is identified that has two neurons projecting to it, one weak (*subthreshold*) and one strong (*suprathreshold*). The subthreshold input results in depolarization of the postsynaptic neuron, but the depolarization is not strong enough to cause the postsynaptic neuron to spike. The suprathreshold input is strong enough to induce a spike in the postsynaptic neuron. Plasticity of the synapse between the subthreshold input and the postsynaptic neuron is measured as a function of the timing between subthreshold and postsynaptic neurons' spikes ($\Delta t_{pre-post}$) by varying the intervals between induced spikes in the subthreshold and the suprathreshold inputs ($\Delta t_{pre-pre}$). This measurement yields the well-known STDP curve (Figure 1b).

In most experimental studies of STDP, the postsynaptic neuron is induced to spike not via a suprathreshold neuron, but rather by depolarizing current injection directly into the postsynaptic neuron. To model experiments that induce spiking via current injection, additional assumptions must be made in the Spike Response Model framework. Because these assumptions are not well established in the literature, we have focused on the synaptic input technique of Zhang et al. (1998). In Section 5.1, we propose a method for modeling a depolarizing current injection in the spike-response model.

The Zhang et al. (1998) experiment imposes four constraints on a simulation. (1) The suprathreshold input alone causes spiking $> 70\%$ of the time. (2) The subthreshold input alone causes spiking $< 10\%$ of the time. (3) Synchronous firing of suprathreshold and/or subthreshold inputs cause LTP if and only if the postsynaptic neuron fires. (4) The time constants of the excitatory PSPs (EPSPs) —τ_s and τ_m in the sSRM—are in the range of
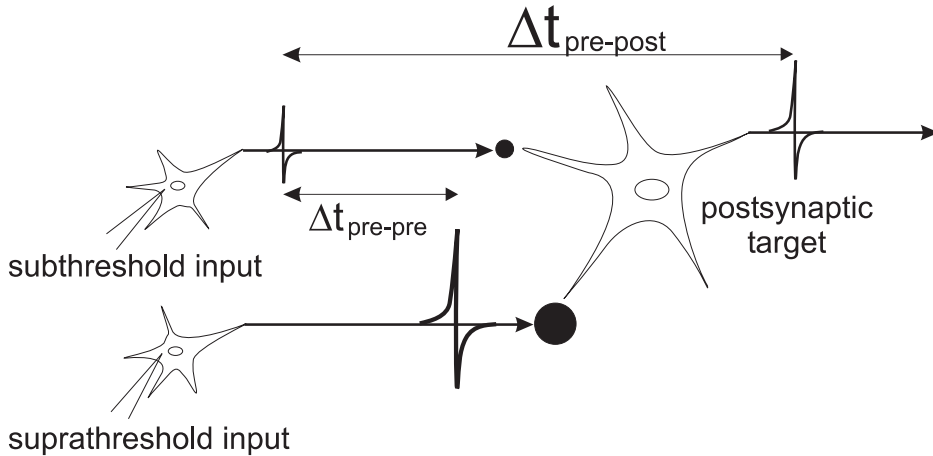
Figure 4: Experimental setup of Zhang et al. (1998)

1–5ms and 7–15ms respectively. These constraints remove many free parameters from our simulation. We do not explicitly model the two input cells; instead, we model the EPSPs they produce. The magnitude of these EPSPs are picked to satisfy the experimental constraints: in most simulations, unless reported otherwise, the suprathreshold EPSP (w_{supra}) alone causes a spike in the post on 85% of trials, and the subthreshold EPSP (w_{sub}) alone causes a spike on fewer than 0.1% of trials. In our principal recreation of the experiment (Figure 5), we added normally distributed variation to w_{supra} and w_{sub} to simulate the experimental selection process of finding suitable supra-subthreshold input pairs according to: $w'_{supra} = w_{supra} + N(0, \sigma_{supra})$ and $w'_{sub} = w_{sub} + N(0, \sigma_{sub})$ (we controlled the random variation for conditions outside the specified firing probability ranges). Free parameters of the simulation are ϑ and β in the spike-probability function (α can be folded into ϑ), and the magnitude (u_r^s, u_{abs}) and time constants ($\tau_r^s, \tau_r^f, \Delta_{abs}$) of the reset. We can further investigate how the results depend on the exact strengths of the subthreshold and suprathreshold EPSP's.

The dependent variable of the simulation is $\Delta t_{pre-pre}$, and we measure the time of the post spike to determine $\Delta t_{pre-post}$. In the experimental protocol, a pair of inputs is repeatedly stimulated at a specific interval $\Delta t_{pre-pre}$ at a low frequency of 0.1Hz. The weight update for a given $\Delta t_{pre-pre}$ is measured by comparing the size of the EPSC before stimulation and (about) half an hour after stimulation. In terms of our model, this repeated stimulation can be

19

considered as drawing a response ξ from the stochastic conditional response density $p(\xi)$. We estimate the expected weight update for this density $p(\xi)$ for a given $\Delta t_{pre-pre}$ using Equation 9 by approximating the integral by a summation over all time-discretized output responses consisting of 0, 1, or 2 spikes. Note that performing the weight update computation like this implicitly assumes that during the synaptic efficacies in the experiment do not change much during repeated stimulation; since long-term synaptic changes require the synthesis of e.g. proteins this seems a reasonable assumption, also reflected in the half hour or so that the experimentalists wait after stimulation before measuring the new synaptic efficacy.

5   Results

Figure 5A shows an STDP curve produced by the model, obtained by plotting the estimated weight update of Equation (9) against $\Delta t_{pre-post}$ for fixed supra and subthreshold inputs. Specifically, we vary the difference in time between subthreshold and suprathreshold inputs (a pre-pre pair) and we compute the expected gradient for the subthreshold input w_{sub} over all responses of the postsynaptic neuron via Equation (9). We thus obtain a value for Δw for each $\Delta t_{pre-pre}$ data point, we then compute $\Delta w_{sub}(\%)$ as the relative percentage change of synaptic efficacy: $\Delta w_{sub}(\%) = \Delta w / w_{sub} \times 100\%$ [1]. For each $\Delta t_{pre-pre}$, the corresponding value $\Delta t_{pre-post}$ is determined by calculating for each input pair the average time at which the postsynaptic neuron fires relative to the subthreshold input. Together, this results in a set of $(\Delta t_{pre-post}, \Delta w_{sub}(\%))$ data points. The continuous graph in Figure 5A is obtained by repeating this procedure for fixed supra and subthreshold weights, and connecting the resultant points.

In Figure 5B, the supra and subthreshold weights in the simulation are randomly perturbed for each pre-pre pair, to simulate the fact that in the experiment, different pairs of neurons are selected for each pre-pre pair, leading inevitably to variation in the synaptic strengths. Mild variation of the input weights yields the "scattering" datapoints of the relative weight changes similar to the experimentally observed data.

Clearly, the mild variation we apply is only small relative to the observed

---

[1] we set the global learning rate γ in Equation (9) such that the simulation curve is scaled to match the neurophysiological results. In all other experiments where we use relative percentage change $\Delta w_{sub}(\%)$ the same value for γ is used.

in vivo distributions of synaptic weights in the brain (e.g., Song, Sjöström, Reigl, Nelson, and Chklovskii (2005)). However, Zhang *et al.* (1998) did not sample randomly from synapses in the brain but rather selected synapses that had a particularly narrow range of initial EPSPs to satisfy the criteria for "supra-" and "subthreshold" synapses (see also Section 4). Hence the experimental variance was particularly small (see Figure 1e of Zhang et al, 1998), and our variation of the size of the EPSP is in line with the observed variations in the experimental results of Zhang *et al.* (1998).

The model produces a good quantitative fit to the experimental data points (triangles), especially compared to other related work as discussed in the introduction, and robustly obtains the typical LTP/LTD time windows associated with STDP.

In figure 5C, we show our model fit compared to the models of Toyoizumi et al. (2005) and Chechik (2003). Our model obtained the lowest sum squared error (1.25 vs 1.63 and 3.27^2 respectively, see Appendix B for methods). This despite the lack of data in the region $\Delta t_{\text{pre-post}} = 0 \ldots 10 ms$ in the Zhang *et al.* (1998) experiment, where difference in LTD behavior is most pronounced.

The qualitative shape of the STDP curve is robust to settings of the spiking neuron model's parameters, as we will illustrate shortly. Additionally, we found that the type of spike-probability function ρ (exponential, sigmoidal, or linear) is not critical.

Our model accounts for an additional finding that has not been explained by alternative theories: The relative magnitude of LTP decreases as the efficacy of the synapse between the subthreshold input and the postsynaptic target neuron increases; in contrast, LTD remains roughly constant (G.-q. Bi & Poo, 1998), Figure 6A shows this effect in the experiment of Bi and Poo (1998), and Figure 6B shows the corresponding result from our model. We compute the magnitude of LTP and LTD for the peak modulation (i.e., $\Delta t_{pre-post} = -5$ for LTP and $\Delta t_{pre-post} = +5$ for LTD) as the amplitude of the subthreshold EPSP is increased. The model's explanation for this phenomenon is simple: As the synaptic weight increases, its effect saturates, and a small change to the weight does little to alter its influence. Consequently, the gradient of the entropy with respect to the weight goes toward zero. Sim-

---

[2]Of note for this comparison is that our spiking neuron model uses a more sophisticated difference of exponentials (Equation 4) to describe the EPSP, whereas the spiking neuron models in Toyoizumi et al. (2005) and Chechik (2003) use a single exponential. These other models might be improved using the more sophisticated EPSP function.
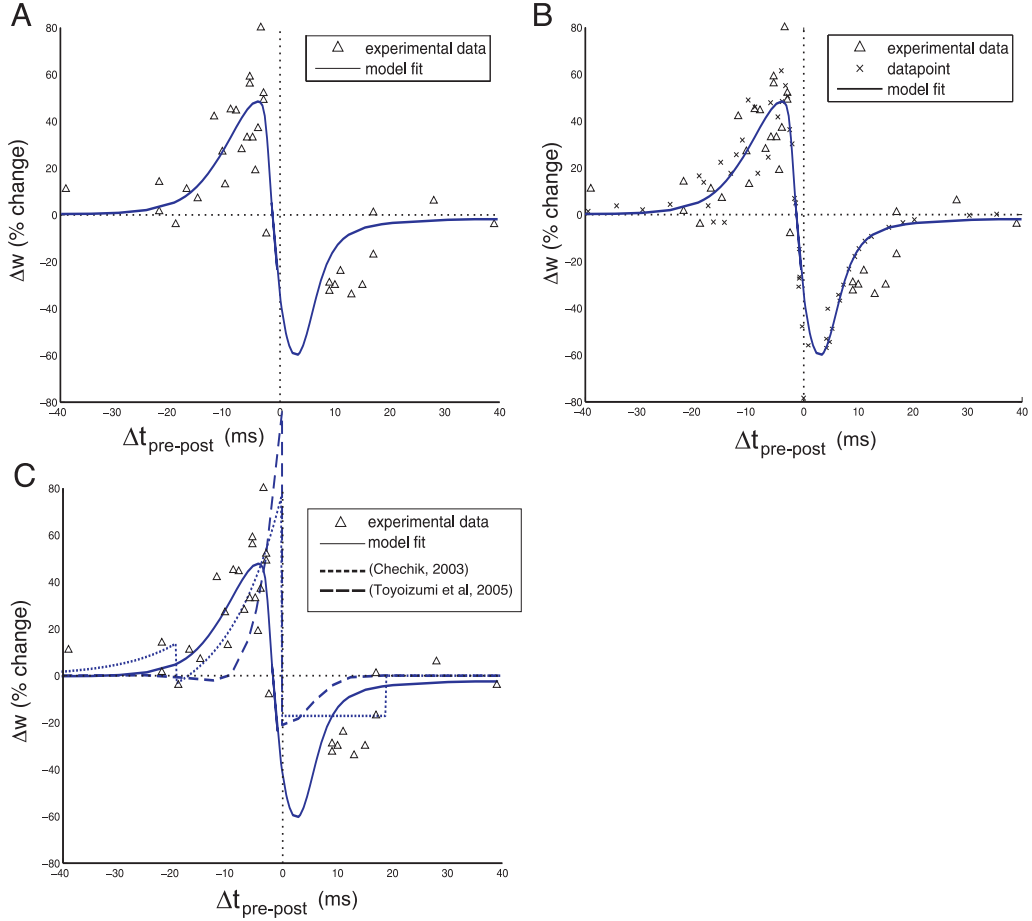
21

Figure 5: (A) STDP – experimental data (triangles) and model fit (solid line). (B) Added simulation datapoints with perturbed weights w'_{supra} and w'_{sub} (crosses). STDP data redrawn from Zhang et al. (1998). Model parameters: $\tau_s = 2.5ms$, $\tau_m = 10ms$, sub and suprathreshold weight perturbation in (B): $\sigma_{supra} = 0.33\,w_{supra}$, $\sigma_{sub} = 0.1\,w_{sub}$. (C) Model fit compared to previous generative models (redrawn from Chechik, 2003, short dashed line; Toyoizumi et. al., 2005, long dashed line). Free parameters of Chechik (2003) and Toyoizumi et al. (2005) were fit to the experimental data as described in Appendix B.
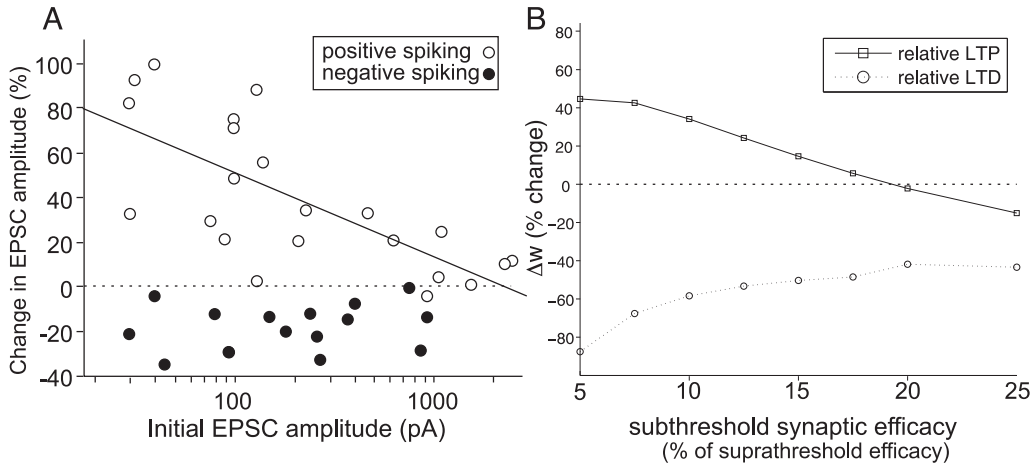
Figure 6: Dependence of LTP and LTD magnitude on efficacy of the sub-threshold input. (A) Experimental data redrawn from Bi & Poo (1998); (B) Simulation result.

ilar saturation effects are observed in gradient-based learning methods with nonlinear response functions, such as back propagation.

As we mentioned earlier, other theories have had difficulty reproducing the typical shape of the LTD component of STDP. In (Chechik, 2003), the shape is predicted to be near uniform, and in Toyoizumi et al. (2005) the shape depends on the autocorrelation. In our stochastic Spike Response Model this component arises due to the stochastic variation in the neural response: in the specific STDP experiment, reduction of variability is achieved by reducing the probability of multiple output spikes. To argue for this conclusion, we performed simulations that make our neuron model less variable in various ways, and each of these manipulations results in a reduction in the LTD component of STDP. In Figures 7A and 7B, we make the threshold more deterministic by increasing the values of α and β in the spike probability density function. In Figure 7C, we increase the magnitude of the refractory response η, which will prevent spikes following the initial post-synaptic response. And finally, in Figure 7D, we increase the efficacy of the suprathreshold input, which prevents the postsynaptic neuron's potential from hovering in the region where the stochasticity of the threshold can induce a spike. Modulation of all of these variables make the threshold more deterministic, and decrease LTD relative to LTP.
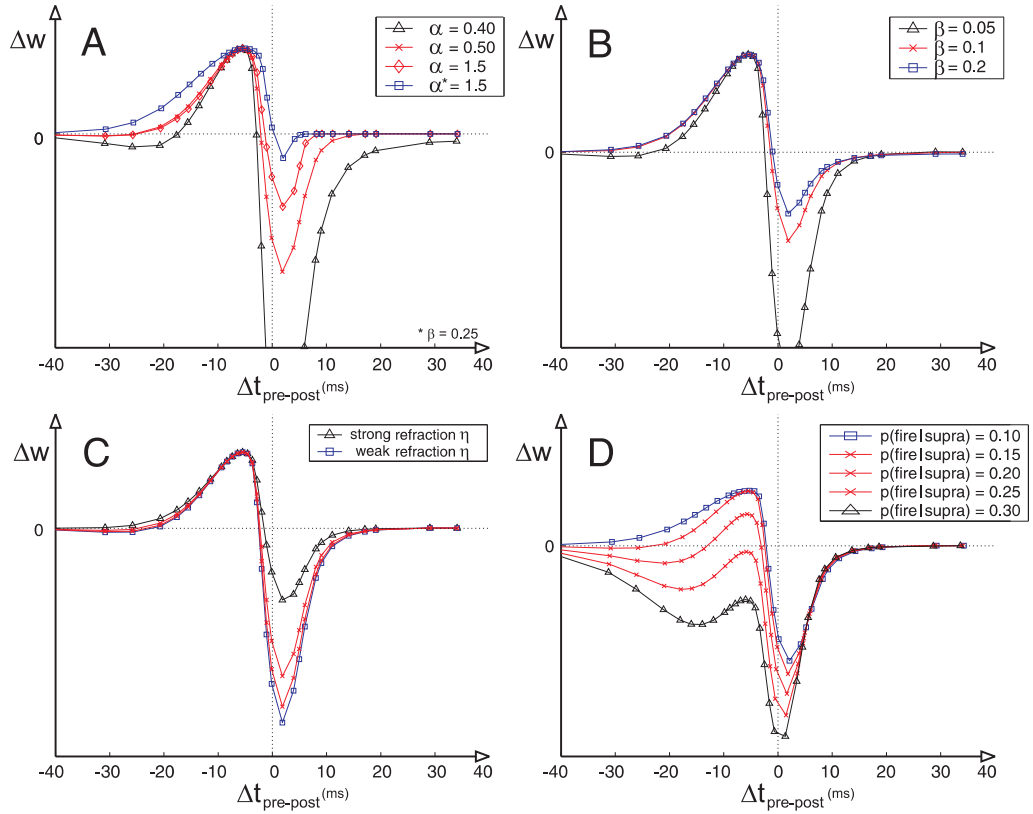
Figure 7: Dependence of relative LTP and LTD on (A) the parameter α of the stochastic threshold function, (B) the parameter β of the stochastic threshold function, (C) the magnitude of refraction, η, and (D) efficacy of the suprathreshold synapse, expressed as $p(\text{fire}|\text{supra})$, the probability that the postsynaptic neuron will fire when receiving only the suprathreshold input. Larger values of $p(\text{fire}|\text{supra})$ correspond to a weaker suprathreshold synapse. In all graphs, the weight gradient for individual curves is normalized to peak LTP for comparison purposes.
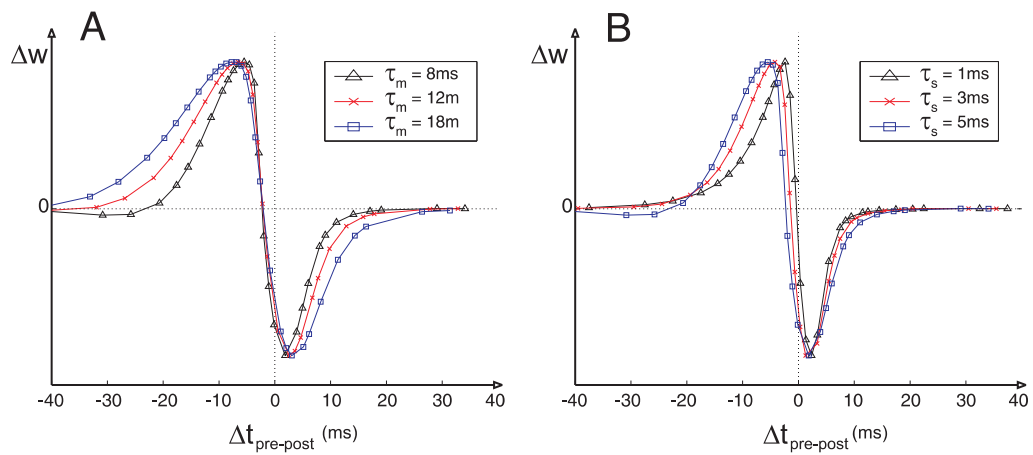
Figure 8: Influence of time constants of the sSRM model on the shape of the STDP curve. (A) varying the membrane potential time-constant τ_m (B) varying the EPSP rise time constant τ_s. In both figures, the magnitude of LTP and LTD have been normalized to 1 for each curve to allow for easy examination of the effect of the manipulation on temporal characteristics of the STDP curves.

Our simulation results are robust to biologically-realizable variation in the parameters of the sSRM model. For example, time constants of the EPSPs can be varied with no qualitative effect on the STDP curves. Figures 8A and 8B show the effect of manipulating the membrane potential decay time τ_m and the EPSP rise time τ_s, respectively. Note that manipulation of these time constants does predict a systematic effect on STDP curves. Increasing τ_m increases the duration of both the LTP and LTD windows, whereas decreasing τ_s leads to a faster transition from LTP to LTD. Both predictions could be tested experimentally by correlating time constants of individual neurons studied with the time course of their STDP curves.

5.1   Current Injection

We mentioned earlier that in many STDP experiments, an action potential is induced in the postsynaptic neuron not via a suprathreshold presynaptic input, but via a depolarizing current injection. In order to model experiments using current injection, we must characterize the current function and its effect on the postsynaptic neuron. In this section, we make such a proposal framed in terms of the Spike Response Model, and report simulation results using current injection.

We model the injected current $\mathcal{I}(t)$ as a rectangular step function:

$$\mathcal{I}(t) = \mathcal{H}(t - f_I)\, \mathcal{I}_c\, \mathcal{H}(t - [\Delta_I - f_I]), \tag{22}$$

where the current of magnitude $\mathcal{I}_c$ is switched on at $t = f_I$ and off at $t = f_I + \Delta_I$. In the Zhang et al (1998) experiment, Δ_I is 2ms, a value we adopted for our simulations as well.

The resulting postsynaptic potential, ϵ_c is:

$$\epsilon_c(t) = \int_0^t \exp\left(-\frac{s}{\tau_m}\right) \mathcal{I}(s)\, ds. \tag{23}$$

In the absence of postsynaptic firing, the membrane potential of an integrate-and-fire neuron in response to a step current is (Gerstner, 2001):

$$\epsilon_c(t|f_I) = \mathcal{I}_c(1 - \exp[-(t - f_I)/\tau_m]). \tag{24}$$

In the presence of postsynaptic firing at time $\hat{f}_i$, we assume—as we did previously in Equation (5)—a reset and subsequent integration of the residual

current:

$$
\begin{aligned}
\epsilon_c(t|\hat{f}_i) \;=\; & \mathcal{H}(\hat{f}_i - t) \int_0^t \exp\!\left(-\frac{s}{\tau_m}\right) \mathcal{I}(s)\, ds \;+ \\
& \mathcal{H}(t - \hat{f}_i) \int_{\hat{f}_i}^t \exp\!\left(-\frac{s}{\tau_m}\right) \mathcal{I}(s)\, ds.
\end{aligned}
\tag{25}
$$

These ϵ_c kernels are depicted in Figure 9 for a postsynaptic spike occurring at various times $\hat{f}_i$. In our simulations, we chose the current magnitude $\mathcal{I}_c$ to be large enough to elicit spiking of the target neuron with probability greater than 0.7.

Figure 10A shows the STDP curve obtained using the current injection model for the exact same model parameter settings used to produce the result based on a suprathreshold synaptic input (depicted in Figure 5A) superimposed on the experimental data STDP obtained by depolarizing current injection from Zhang et al (1998). Figure 10B additionally superimposes the earlier result on the current injection result, and the two curves are difficult to distinguish. As in the earlier result, variation of model parameters has little appreciable effect on the model's behavior using the current injection paradigm, suggesting that current injection vs. synaptic input makes little difference on the nature of STDP.

6 Discussion

In this paper, we explored a fundamental computational principle, that synapses adapt so as to minimize the variability of a neuron's response in the face of noisy inputs, yielding more reliable neural representations. From this principle—instantiated as entropy minimization—we derived the STDP learning curve. Importantly, the simulation methodology we used to derive the curve closely follows the procedure used in neurophysiological experiments (Zhang et al., 1998): assuming variation in of sub- and suprathreshold synaptic efficacies from experimental pair-to-pair even recovers the noisy scattering of efficacy changes. Our simulations furthermore obtain an STDP curve that is robust to model parameters and details of the noise distribution.

Our results are critically dependent on the use of Gerstner's stochastic Spike Response Model, whose dynamics are a good approximation to those of a biological spiking neuron. The sSRM has the virtue of being characterized by parameters that are readily related to neural dynamics, and its dynamics
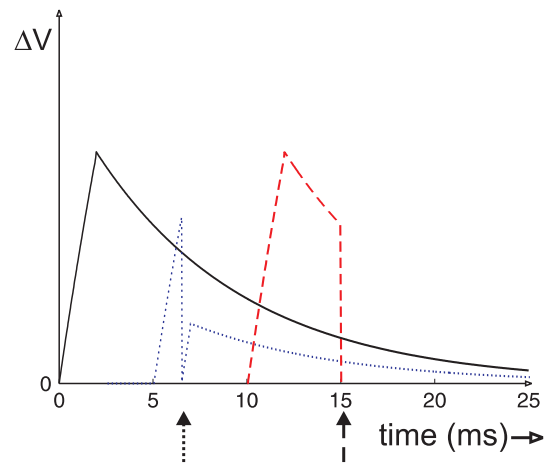
27

Figure 9: Voltage response of a spiking neuron for a 2ms current injection in the Spike Response Model. Solid curve: The postsynaptic neuron produces no spike, and the potential due to the injected current decays with the membrane time-constant τ_m. Dotted curve: The postsynaptic neuron spikes while the current is still being applied. Dashed curve: The postsynaptic neuron spikes after application of the current has terminated (moment of postsynaptic spiking indicated by arrows).
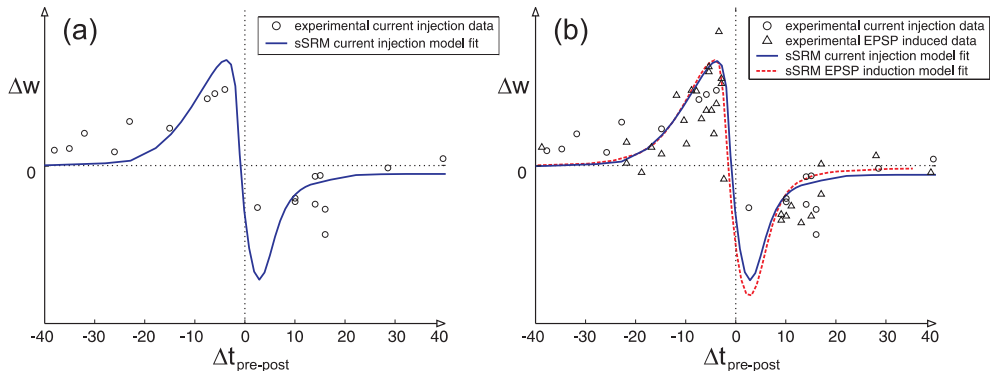
Figure 10: (A) STDP curve obtained for SRM with current injection (solid curve) compared with experimental data for depolarizing current injection (circles; redrawn from Zhang et al., 1998). (B) Comparing STDP curves for both current injection (solid curve) and suprathreshold input (dashed curve) models. The same model parameters are used for both curves. Experimental data redrawn from Zhang et al. (1998) for current injection (circles) and suprathreshold input (triangles) paradigms are superimposed.

are differentiable such that we can derive a gradient-descent learning rule that minimizes the response variability of a postsynaptic neuron given a particular set of input spikes.

Our model predicts the shape of the STDP curve and how it relates to properties of a neuron's response function. These predictions may be empirically testable if a diverse population of cells can be studied. The predictions include the following: First, the width of the LTD and LTP windows depend on the (excitatory) PSP time constants (Figure 7A,B). Second, the strength of LTD relative to LTP depends on the degree of noise in the neuron's response; the LTD strength is related to the noise level.

Our model also can characterize the nature of the learning curve for experimental situations that deviate from the boundary conditions of Zhang et al. (1998). In Zhang et al., the subthreshold and suprathreshold inputs produced postsynaptic firing with probability less than .10 and greater than .70, respectively. Our model can predict the consequences of violating these conditions. For example, when the subthreshold input is very strong or the suprathreshold input is very weak, our model produces strictly LTD, i.e., anti-Hebbian learning. The consequence of a strong subthreshold input is

29

shown in Figure 6B, and the consequence of a weak suprathreshold input is shown in Figure 7D. Intuitively, this simulation result makes sense because— in the first case— the most likely alternative response of the postsynaptic neuron is to produce more than one spike, and—in the second case—the most likely alternative response is no postsynaptic spike at all. In both cases, synaptic depression reduces the probability of the alternative response. We note that such strictly anti-Hebbian learning has been reported in relation to STDP-type experiments (Roberts & Bell, 2002).

For very noisy thresholds and for weak suprathreshold inputs, our model produces an LTD "dip" before LTP (Figure 7D). This dip is in fact also present in the work of Chechik (2003). We find it intriguing that this dip is also observed in the experimental results of Nishiyama et al. (2000). The explanation for this dip may be along the same lines as the explanation for the LTD window: given the very noisy threshold, the subthreshold input may occasionally cause spiking, and decreasing its weight would decrease response variability. This may not be offset by the increase due to its contribution to the spike caused by the suprathreshold input, as it is too early to have much influence. With careful consideration of experimental conditions and neuron parameters, it may be possible to reconcile the somewhat discrepant STDP curves obtained in the literature via our model.

In our model, the transition from LTP to LTD occurs at a slight offset from $\Delta t_{pre-post} = 0$: if the subthreshold input fires 1–2ms before the post-synaptic neuron fires (on average), then neither potentiation nor depression occurs. This 1–2ms offset is attributable to the current decay time constant, τ_s. The neurophysiological data are not sufficiently precise to determine the exact offset of the LTP/LTD transition in real neurons. Unfortunately, few experimental data points are recorded near $\Delta t_{pre-post} = 0$. However, the STDP curve of our model does pass through the one data point in that region (Figure 5A), so the offset may actually be a real phenomenon.

The main focus of the simulations in this paper was to replicate the experimental paradigm of Zhang et al. (1998), in which a suprathreshold presynaptic neuron is used to induce the postsynaptic neuron to fire. The Zhang et al. (1998) study is exceptional in that most other experimental studies of STDP use a depolarizing current injection to induce the postsynaptic neuron to fire. We are not aware of any established model for current injection within the SRM framework. We therefore proposed a model of current injection within the SRM framework in Section 5.1. The proposed model is an ideal abstraction of current injection that does not take into account

effects like current onset and offset fluctuations inherent to such experimental methods. Even with these limitations in mind, the current injection model produced STDP curves very similar to the ones obtained by the simulation of the suprathreshold input induced postsynaptic firing.

The simulations reported in this paper account for classical STDP experiments in which a single presynaptic spike is paired with a single postsynaptic spike. The same methodology can be applied to model experimental paradigms involving multiple presynaptic and/or postsynaptic spikes. However, the computation involved becomes nontrivial. We are currently engaged in modeling data from the multi-spike experiments of Froemke and Dan (2002).

We note that one set of simulation results we reported is particularly pertinent for comparing and contrasting our model to the related model of Toyoizumi et al. (2005). The simulations reported in Figure 7 suggest that noise in our model is critical for obtaining the LTD component of STDP, and that parameters that reduce noise in the neural response also reduce LTD. We found that increasing the strength of neuronal refraction reduces response variability and therefore diminishes the LTD component of STDP. This notion is also put forward in very recent work by Pfister, Toyoizumi, Barber, and Gerstner (2006), where an STDP-like rule arises from from a supervised learning procedure that aims to obtain spikes at a times specified by a teacher. The LTD component in this work also depends on the probability of stochastic activity.

In sharp contrast, Toyoizumi et al. (2005) suggest that neuronal refraction is responsible for LTD. Because the two models are quite similar, it seems unlikely that the models make opposite predictions and the discrepancy may be due to Toyoizumi et al. (2005)'s focus on analytical approximations to solve the mathematical problem at hand, limiting the validity of comparisons between that model and biological experiments in the process.

It is useful to reflect on the philosphy of choosing reduction of spike-train variability as a target function, as it so obviously has the degenerate but energy efficient solution of emitting no spikes at all. The usefulness of our approach clearly relies on the stochastic gradient reaching a local optimum in the likelihood space that does not always correspond to the degenerate solution. We compute the gradient of the input weights with respect to the conditionally independent sequence of response intervals $[t, t + \Delta]$. The gradient approach tries to push the probability of the responses in these intervals to either 0 or 1, irrespective of what the response is ("not firing" or

"firing"). We find that in the sSRM spiking neuron model this gradient can be towards either state of each response interval. which can be attributed to the monotonically increasing spike-probability density as a function of the membrane potential. This spike-probability density allows neurons to become very reliable by firing spikes only at very specific times, at least when starting from a set of input weights that, given the input pattern, is likely to induce a spike in the postsynaptic neuron.

The fact that the target function is the reduction of postsynaptic spike-train variability does predict that in the case of small inputs impinging on a postsynaptic target causing only occasional firing, the prediction would be that the average weight update due to this target function would reduce these inputs to zero.

We have modeled the experimental studies in some detail, beyond the level of detail achieved by other researchers investigating STDP. Even a model with an entirely heuristic learning rule has value if it obtains a better fit to the data than other models of similar complexity. Our model has a learning rule that goes beyond heuristic: the learning rule is derived from a computational objective. To some, this objective may not be as exciting as more elaborative objectives like information maximization. As it is, our model stands alone from the contenders in providing a first-principle account of STDP that fits experimental data extremely well. Might there be a mathematically sexier model? We certainly hope so, but it has not yet been discovered.

We reiterate the point that our learning objective is viewed as but one of many objectives operating in parallel. The question remains as to why neurons would respond so highly variable to fixed input spike-trains: a more deterministic threshold would eliminate the need for any minimization of response variability. We can only speculate that the variability in neuronal responses may also well serve these other objectives, such as exploitation/exploration in reinforcement learning, or the exploitation of stochastic resonance phenomena (e.g. Hahnloser, Sarpeshkar, Mahowald, Douglas, and Seung (2000)).

It is interesting to note that minimization of conditional response variability corresponds to one part of the equation that maximizes mutual information. The mutual information $\mathcal{I}$ between input X and outputs Y is defined as:

$$\mathcal{I}(X, Y) = \mathcal{H}(Y) - \mathcal{H}(Y|X).$$

Hence minimization of the conditional entropy $\mathcal{H}(Y|X)$ – our objective – along with the secondary unsupervised objective of maximizing the marginal entropy $\mathcal{H}(Y)$ together maximize mutual information. The first unsupervised objective is notoriously hard to compute (e.g. see Bell and Parra (2005) for an extensive discussion) whereas, as we have shown, the second objective — conditional entropy minimization — can be computed relatively easily via stochastic gradient descent. Indeed, in this light, it is a virtue of our model that we account for the experimental data with only one component of the mutual-information objective (taking the responses in the experimental conditions as the set of responses Y). The relatively simple nature of the experiments that uncovered STDP lack any interaction with other (output) neurons, and we may speculate that STDP may be the degenerate reflection of information maximization in the absence of such interactions. If subsequent work shows that STDP can be explained by mutual information maximization (without the drawbacks of existing work, such as the rate based treatment of (Chechik, 2003), or the unrealistic auto-correlation function and difficulty of relating to biological parameters of (Toyoizumi et al., 2005)), this work contributes in helping to tease apart the components of the objective that are necessary and sufficient for explaining the data.

Appendix A

To compute Δw, we stop at $n = 2$ as in the experimental conditions that we model the contribution of $n > 2$ spikes is vanishingly small: we find that the propability of three spikes occurring is typically $< 1e-5$, and the $n = 3$ term already did not contribute significantly, as shown for example in Figure 11.

Intuitively it seems very unlikely that the gradient of the conditional response entropy is dominated by terms that are highly unlikely. This could only be the case if the gradient on the probability of getting three or more spikes would be much larger than the gradient on getting say two spikes. Given the setup of the model with an increasing probability of firing a spike as a function of the membrane potential, it is easy to see that changing a weight will change the probability of obtaining two spikes much more than the probability of obtaining three spikes. Hence the entropy gradient from components $n \leq 2$ will be (in practise: much) larger than the gradient for terms $n = 3, n = 4, \ldots$. As we remarked before, in our simulation of the experimental setup, the probability of obtaining three spikes given the input
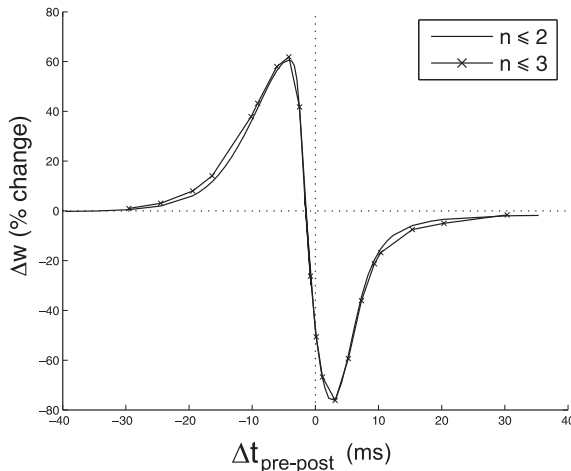
Figure 11: STDP graphs for Δw computed using the terms $n \leq 2$ (solid line) and $n \leq 3$ (crossed solid line).

was computed to be $< 1e - 5$: the overall probability was computed at up to $1e - 6$. The probability of $n = 4$ was below the precision of our simulation.

Appendix B

To compute the sum squared error when comparing the different STDP models in Section 5, we use linear regression in the free parameters to minimize the sum squared error between the model curves and the experimental data.

For m experimental datapoints $\{(\Delta t_1, \Delta w_1), \ldots (\Delta t_m, \Delta w_m)\}$ and model curve $\Delta w = f(\Delta t)$, we report for each model curve the sum squared error for those values of the free parameters that minimize the sum squared error E^2:

$$E^2 = \min_{\text{free parameters}} \sum_{i=1}^{m} \left(\Delta w_i - f(\Delta t_i) \right)^2.$$

For our model, linear regression is performed in the scaling parameter γ in Equation (9) that relates the gradient obtained with the model parameters mentioned to the weight change.

Where possible for the other models, we set model parameters to correspond to the values observed in the experimental conditions described in

Section 4.

For the model by Chechik (2003), the weight update is computed as the sum of a positive exponent and a negative damping contribution:

$$\Delta w = \gamma \left(\mathcal{H}(-\Delta t) \exp(\Delta t/\tau) - \mathcal{H}(t + \Delta)\mathcal{H}(-\Delta - t)K \right),$$

where Δt is computed as $t_{pre} - t_{post}$, K denotes the negative damping contribution that is applied over a time window Δ before and after the postsynaptic spike; and $\mathcal{H}()$ denotes the Heaviside function. The time constant τ is related to the decay of the EPSP and we set this value to the same value we use for our model: $10ms$. Linear regression to find the minimal sum squared error is performed on the free parameters γ, K, Δ.

In (Toyoizumi *et.al.*, 2005), the learning rule is the sum of two terms:

$$\Delta w = \gamma \left(\epsilon^2(\Delta t) - \mu_0(\phi \star \epsilon^2)(\Delta t) \right),$$

where $\epsilon(t)$ is the EPSP, modeled as $\Theta(t) \exp(-t/\tau)$, and $\mu_0(\phi \star \epsilon^2)(\Delta t)$ is a function of the autocorrelation function of a neuron $\phi \star \epsilon^2)(\Delta t)$, times the spontaneous neural activity in the absence of input, μ_0. The EPSP decay time constant used in (Toyoizumi *et.al.*, 2005) was already set to $\tau = 10ms$, and for the two terms in the sum we used the functions described by Figures 2A and 2B in (Toyoizumi *et.al.*, 2005). We performed linear regression to the one free parameter, γ.

Note that for this model, we obtain better LTD, and hence E^2, for larger values of μ_0 as those used in (Toyoizumi *et.al.*, 2005). However, then E^2 still remains worse than for the other two models, and the spontaneous neural activity becomes unrealistically large.

References

Abbott, L., & Gerstner, W. (2004). Homeostasis and learning through spike-timing dependent plasticity. In D. Hansel, C. Chow, B. Gutkin, & C. Meunier (Eds.), *Methods and models in neurophysics.*

Bel1, C. C., Han, V. Z., Sugawara, Y., & Grant, K. (1997). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature, 387,* 278–281.

Bell, A., & Parra, L. (2005). Maximising sensitivity in a spiking network. In L. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in neural information processing systems* (Vol. 17, pp. 121–128). MIT Press.

Bi, G.-Q. (2002, Dec). Spatiotemporal specificity of synaptic plasticity: cellular rules and mechanisms. *Biol. Cybern., 87,* 319-332.

Bi, G.-q., & Poo, M.-m. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci., 18*(24), 10464–10472.

Bi, G.-q., & Poo, M.-m. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Ann. Rev. Neurosci., 24,* 139–166.

Burkitt, A., Meffin, H., & Grayden, D. (2004). Spike timing-dependent plasticity: The relationship to rate-based learning for models with weight dynamics determined by a stable fixed-point. *Neural Computation, 16*(5), 885–940.

Chechik, G. (2003). Spike-timing-dependent plascticity and relevant mutual information maximization. *Neural Computation, 15,* 1481–1510.

Dan, Y., & Poo, M.-m. (2004, September). Spike timing-dependent plasticity of neural circuits. *Neuron, 44,* 23–30.

Dayan, P. (2002, March). Matters temporal. *TRENDS in Cognitive Sciences, 6*(3), 105–106.

Dayan, P., & Häusser, M. (2004). Plasticity kernels and temporal statistics. In S. Thrun, L. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems 16.* Cambridge, MA: MIT Press.

Debanne, D., Gähwiler, B., & Thompson, S. (1998). Long-term synaptic plasticity between pairs of individual ca3 pyramidal cells in rat hippocampal slice cultures. *J. Physiol., 507,* 237–247.

Feldman, D. (2000, July). Timing-based ltp and ltd at vertical inputs to layer ii/iii pyramidal cells in rat barrel cortex. *Neuron, 27,* 45-56.

Froemke, R., & Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature, 416,* 433–438.

Gerstner, W. (2001). A framework for spiking neuron models: The spike response model. In F. Moss & S. Gielen (Eds.), *The handbook of biological physics* (Vol. 4, p. 469-516).

Gerstner, W., Kempter, R., van Hemmen, J. L., & Wagner, H. (1996). A neural learning rule for sub-millisecond temporal coding. *Nature, 383,* 76–78.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models.* Cambridge, UK: Cambridge University Press.

Hahnloser, R. H. R., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., & Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature, 405,* 947–951.

Hennion, P. E. (1962, April). Algorithm 84: Simpson's integration. *Communications of the ACM, 5*(4), 208.

Herrmann, A., & Gerstner, W. (2001). Noise and the PSTH response to current transients: I. General theory and application to the integrate-and-fire neuron. *J. Comp. Neurosci., 11,* 135–151.

Hopfield, J., & Brody, C. (2004). Learning rules and network repair in spike-timing-based computation. *PNAS, 101*(1), 337–342.

Izhikevich, E., & Desai, N. (2003). Relating stdp to bcm. *Neural Computation, 15,* 1511–1523.

Jolivet, R., Lewis, T., & Gerstner, W. (2003). The spike response model: a framework to predict neuronal spike trains. In K. et al. (Ed.), *Proc. joint international conference icann/iconip 2003* (pp. 846–853). Springer, LNCS 2714.

Kandel, E. R., Schwartz, J., & Jessell, T. M. (2000). *Principles of neural science.* Mc Graw Hill.

Karmarkar, U., Najarian, M., & Buonomano, D. (2002). Mechanisms and significance of spike-timing dependent plasticity. *Biol. Cybern., 87,* 373-382.

Kempter, R., Gerstner, W., & van Hemmen, J. (1999). Hebbian learning and spiking neurons. *Phys. Rev. E, 59*(4), 4498–4514.

Kempter, R., Gerstner, W., & van Hemmen, J. (2001). Intrinsic stabilization of output rates by spike-based hebbian learning. *Neural Computation, 13,* 2709-2742.

Kepecs, A., van Rossum, M., Song, S., & Tegner, J. (2002). Spike-timing-dependent plasticity: common themes and divergent vistas. *Biol. Cybern., 87,* 446–458.

Legenstein, R., Naeger, C., & Maass, W. (n.d.). *What can a neuron learn*

with spike-timing-dependent plasticity? (submitted)

Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science*, *275*, 213–215.

Nishiyama, M., Hong, K., Mikoshiba, K., Poo, M.-m., & Kato, K. (2000, Nov). Calcium stores regulate the polarity and input specificity of synaptic modification. *Nature*, *408*, 584–588.

Paninski, L., Pillow, J., & Simoncelli, E. (2005). Comparing integrate-and-fire models estimated using intracellular and extracellular data. *Neurocomputing.* (to appear)

Pfister, J.-P., Toyoizumi, T., Barber, D., & Gerstner, W. (2006). Optimal spike-timing dependent plasticity for precise action potential firing. *Neural Computation*, *18*, 1318–1348.

Porr, B., & Wörgötter, F. (2003). Isotropic sequence order learning. *Neural Computation*, *15*(4), 831-864.

Rao, R., & Sejnowski, T. (1999). Predictive sequence learning in recurrent neocortical circuits. In T. L. S.Sollar & K. Muller (Eds.), *Advances in neural information processing systems 12* (pp. 164–170). Cambridge, MA: MIT Press.

Rao, R., & Sejnowski, T. (2001). Spike-timing-dependent plasticity as temporal difference learning. *Neural Computation*, *13*, 2221–2237.

Roberts, P., & Bell, C. (2002, Dec). Spike timing dependent synaptic plasticity in biological systems. *Biol. Cybern.*, *87*, 392–403.

Saudargiene, A., Porr, B., & Wörgötter, F. (2004). How the shape of pre- and postsynaptic signals can influence stdp: A biophysical model. *Neural Computation*, *16*, 595-625.

Senn, W., Markram, H., & Tsodyks, M. (2000). An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. *Neural Computation*, *13*, 35–67.

Shon, A., Rao, R., & Sejnowski, T. (2004). Motion detection and prediction through spike-timing dependent plasticity. *Network: Comput. Neural Syst.*, *15*, 179–198.

Sjöström, P., Turrigiano, G., & Nelson, S. (2001, December). Rate, timing, and cooperativity jointly determine cortical synpatic plasticity. *Neuron*, *32*, 1149–1164.

Song, S., Miller, K., & Abbott, L. (2000). Competitive hebbian learning through spiketime -dependent synaptic plasticity. *Nature Neuroscience*, *3*, 919–926.

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biology*, *3*(3), e68.

Toyoizumi, T., Pfister, J.-P., Aihara, K., & Gerstner, W. (2005). Spike-timing dependent plasticity and mutual information maximization for a spiking neuron model. In L. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in neural information processing systems* (Vol. 17, pp. 1409–1416). MIT Press.

van Rossum, R., Bi, G.-q., & Turrigiano, G. (2000). Stable hebbian learning from spike time dependent plasticity. *J. Neurosci.*, *20*, 8812–8821.

Xie, X., & Seung, H. (2004). Learning in neural networks by reinforcement of irregular spiking. *Physical Review E*, *69*(041909).

Zhang, L., Tao, H., Holt, C., Harris, W., & Poo, M.-m. (1998). A critical window for cooperation and competition among developing retinotectal synapses. *Nature*, *395*, 37–44.