IDIAP RESEARCH REPORT

Optimal Spike-Timing Dependent Plasticity for Precise Action Potential Firing in Supervised Learing

Jean-Pascal Pfister [a] Taro Toyoizumi [b]
David Barber [c] Wulfram Gerstner [a]

IDIAP–RR 05-88

[a] Laboratory of Computational Neuroscience, School of Computer and Communication Sciences and Brain-Mind Institute, EPFL, CH-1015 Lausanne
[b] Department of Complexity Science and Engineering, Graduate School of Frontier Sciences, The University of Tokyo
[c] IDIAP Research Institute, Rue du Simplon 4, CH-1920 Martigny, Switzerland

IDIAP Research Institute **www.idiap.ch**
Rue du Simplon 4 P.O. Box 592 1920 Martigny – Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

Optimal Spike-Timing Dependent Plasticity for Precise Action Potential Firing in Supervised Learing

Jean-Pascal Pfister Taro Toyoizumi
David Barber Wulfram Gerstner

Abstract. In timing-based neural codes, neurons have to emit action potentials at precise moments in time. We use a supervised learning paradigm to derive a synaptic update rule that optimizes via gradient ascent the likelihood of postsynaptic firing at one or several desired firing times. We find that the optimal strategy of up and down regulating synaptic efficacies depends on the relative timing between presynaptic spike arrival and desired postsynaptic firing. If the presynaptic spike arrives before the desired postsynaptic spike timing, our optimal learning rule predicts that the synapse should become potentiated. The dependence of the potentiation on spike timing directly reflects the time course of an excitatory postsynaptic potential. However, our approach gives no unique reason for synaptic depression under reversed spike-timing. In fact, the presence and amplitude of depression of synaptic efficacies for reversed spike timing depends on how constraints are implemented in the optimization problem. Two different constraints, i.e., control of postsynaptic rates or control of temporal locality, are studied. The relation of our results to Spike-Timing Dependent Plasticity (STDP) and reinforcement learning is discussed.

Optimal Spike-Timing Dependent Plasticity for Precise Action Potential Firing in Supervised Learning

Jean-Pascal Pfister, Taro Toyoizumi[*], David Barber[†], Wulfram Gerstner
Laboratory of Computational Neuroscience,
School of Computer and Communication Sciences
and Brain-Mind Institute,
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne
{jean-pascal.pfister, taro.toyoizumi}@epfl.ch
david.barber@idiap.ch, wulfram.gerstner@epfl.ch

Abstract

In timing-based neural codes, neurons have to emit action potentials at precise moments in time. We use a supervised learning paradigm to derive a synaptic update rule that optimizes via gradient ascent the likelihood of postsynaptic firing at one or several desired firing times. We find that the optimal strategy of up- and downregulating synaptic efficacies depends on the relative timing between presynaptic spike arrival and desired postsynaptic firing. If the presynaptic spike arrives before the desired postsynaptic spike timing, our optimal learning rule predicts that the synapse should become potentiated. The dependence of the potentiation on

[*]Current address: Department of Complexity Science and Engineering, Graduate School of Frontier Sciences, The University of Tokyo.

[†]Current address: IDIAP, Rue du Simplon 4, Case Postale 592, CH-1920 Martigny.

spike timing directly reflects the time course of an excitatory postsynaptic potential. However, our approach gives no unique reason for synaptic depression under reversed spike-timing. In fact, the presence and amplitude of depression of synaptic efficacies for reversed spike timing depends on how constraints are implemented in the optimization problem. Two different constraints, i.e., control of postsynaptic rates or control of temporal locality, are studied. The relation of our results to Spike-Timing Dependent Plasticity (STDP) and reinforcement learning is discussed.

1 Introduction

Experimental evidence suggests that precise timing of spikes is important in several brain systems. In the barn owl auditory system, for example, coincidence detecting neurons receive volleys of temporally precise spikes from both ears (Carr and Konishi 1990). In the electrosensory system of mormyrid electric fish, medium ganglion cells receive input at precisely timed delays after electric pulse emission (Bell et al. 1997). Under the influence of a common oscillatory drive as present in the rat hippocampus or olfactory system, the strength of a constant stimulus is coded in the relative timing of neuronal action potentials (Hopfield 1995; Brody and Hopfield 2003; Mehta et al. 2002). In humans precise timing of first spikes in tactile afferents encode touch signals at the finger tips (Johansson and Birznieks 2004). Similar codes have also been suggested for rapid visual processing (Thorpe et al. 2001), and for the rat's whisker response (Panzeri et al. 2001).

The precise timing of neuronal action potentials also plays an important role in Spike-Timing Dependent Plasticity (STDP). If a presynaptic spike arrives at the synapse *before* the postsynaptic action potential, the synapse is potentiated; if the timing is reversed the synapse is depressed (Markram et al. 1997; Zhang et al. 1998; Bi and Poo 1998; Bi and Poo 1999; Bi and Poo 2001). This biphasic STDP function is reminiscent of a temporal contrast or temporal derivative filter and suggests that STDP is sensitive to the temporal features of a neural code. Indeed, theoretical studies have shown that, given a biphasic STDP function, synaptic plasticity can lead to a stabilization of synaptic weight dynamics (Kempter et al. 1999; Song et al. 2000; Kempter et al. 2001; van Rossum et al. 2000; Rubin et al. 2001) while the neuron

remains sensitive to temporal structure in the input (Gerstner et al. 1996; Roberts 1999; Kempter et al. 1999; Kistler and van Hemmen 2000; Rao and Sejnowski 2001; Gerstner and Kistler 2002a).

While the relative firing time of pre- and postsynaptic neurons, and hence temporal aspects of a neural code, play a role in STDP, it is, however, less clear whether STDP is useful to *learn* a temporal code. In order to further elucidate the computational function of STDP, we ask in this paper the following question: What is the ideal form of a STDP function in order to generate action potentials of the postsynaptic neuron with high temporal precision?

This question naturally leads to a supervised learning paradigm - i.e., the task to be learned by the neuron is to fire at a *predefined desired firing time* t^{des}. Supervised paradigms are common in machine learning in the context of classification and prediction problems (Minsky and Papert 1969; Haykin 1994; Bishop 1995), but have more recently also been studied for spiking neurons in feedforward and recurrent networks (Legenstein et al. 2005; Rao and Sejnowski 2001; Barber 2003; Gerstner et al. 1993; Izhikevich 2003). Compared to unsupervised or reward-based learning paradigms, supervised paradigms on the level of single spikes are obviously less relevant from a biological point, since it is questionable what type of signal could tell the neuron about the 'desired' firing time. Nevertheless, we think it is worth addressing the problem of supervised learning - firstly as a problem in it's own right, and secondly as a starting point of spike base reinforcement learning (Xie and Seung 2004; Seung 2003). Reinforcement learning in a temporal coding paradigm implies that certain sequences of firing times are rewarded whereas others are not. The "desired firing times" are hence defined *indirectly via the presence or absence of a reward signal*. The exact relation of our supervised paradigm to reward-based reinforcement learning will be presented in section 4. Section 2 introduces the stochastic neuron model and coding paradigm which are used to derive the results presented in section 3.

2 Model

2.1 Coding Paradigm

In order to explain our computational paradigm, we focus on the example of temporal coding of human touch stimuli (Johansson and Birznieks 2004), but

the same ideas would apply analogously to the other neuronal systems with temporal codes mentioned above (Carr and Konishi 1990; Bell et al. 1997; Hopfield 1995; Brody and Hopfield 2003; Mehta et al. 2002; Panzeri et al. 2001). For a given touch stimulus, spikes in an ensemble of N tactile afferents occur in a precise temporal order. If the same touch stimulus with identical surface properties and force vector is repeated several times, the relative timing of action potentials is reliably reproduced whereas the spike timing in the same ensemble of afferents is different for other stimuli (Johansson and Birznieks 2004). In our model, we assume that all input lines, labeled by the index j with $1 \leq j \leq N$ converge onto one or several postsynaptic neurons. We think of the postsynaptic neuron as a detector for a given spatio-temporal spike patterns in the input. The full spike pattern detection paradigm will be used in Section 3.3. As a preparation and first steps towards the full coding paradigm we will also consider the response of a postsynaptic neuron to a single presynaptic spike (Section 3.1) or to one given spatio-temporal firing pattern (Section 3.2).

2.2 Neuron Model

Let us consider a neuron i which is receiving input from N presynaptic neurons. Let us denote the ensemble of all spikes of neuron j by $x_j = \{t_j^1, \ldots, t_j^{N_j}\}$ where t_j^k denotes the time when neuron j fired its k^{th} spike. The spatio-temporal spike pattern of all presynaptic neurons $1 \leq j \leq N$ will be denoted by boldface $\mathbf{x} = \{x_1, \ldots, x_N\}$.

A presynaptic spike elicited at time t_j^f evokes an excitatory postsynaptic potential (EPSP) of amplitude w_{ij} and time course $\epsilon(t - t_j^f)$. For the sake of simplicity, we approximate the EPSP time course by a double exponential

$$\epsilon(s) = \epsilon_0 \left[\exp\left(-\frac{s}{\tau_m}\right) - \exp\left(-\frac{s}{\tau_s}\right) \right] \Theta(s) \tag{1}$$

with a membrane time constant of $\tau_m = 10$ ms and a synaptic time constant of $\tau_s = 0.7$ ms which yields an EPSP rise time of 2 ms. Here $\Theta(s)$ denotes the Heaviside step function with $\Theta(s) = 1$ for $s > 0$ and $\Theta(s) = 0$ else. We set $\epsilon_0 = 1.3$ mV such that a spike at a synapse with $w_{ij} = 1$ evokes an EPSP with amplitude of approximately 1 mV. Since the EPSP amplitude is a measure of the strength of a synapse, we refer to w_{ij} also as the efficacy (or "weight") of the synapse between neuron j and i.

Let us further suppose that the postsynaptic neuron i receives an additional input $I(t)$ that could either arise from a second group of neurons or from intracellular current injection. We think of the second input as a 'teaching' input that increases the probability that the neuron fires at or close to the desired firing time $t^{\rm des}$. For the sake of simplicity we model the teaching input as a square current pulse $I(t) = I_0\Theta(t - t^{\rm des} + 0.5\Delta T)\Theta(t^{\rm des} + 0.5\Delta T - t)$ of amplitude I_0 and duration ΔT. The effect of the teaching current on the membrane potential is

$$u_{\rm teach}(t) = \int_0^\infty k(s)I(t-s)ds \tag{2}$$

with $k(s) = k_0 \exp(-s/\tau_m)$ where k_0 is a constant that is inversely proportional to the capacitance of the neuronal membrane.

In the context of the human touch paradigm discussed in section 2.1, the teaching input could represent some preprocessed visual information ('object touched by fingers starts to slip now'), feedback from muscle activity ('strong counterforce applied now'), cross-talk from other detector neurons in the same population ('your colleagues are active now'), or unspecific modulatory input due to arousal or reward ('be aware - something interesting happening now').

In the context of training of recurrent networks (e.g. Rao and Sejnowski 2001), the teaching input consists of a short pulse of an amplitude that guarantees action potential firing.

The membrane potential of the postsynaptic neuron i (Spike Response Model, Gerstner and Kistler 2002b) is influenced by the EPSPs evoked by all afferent spikes of stimulus $\mathbf{x}$, the 'teaching' signal and the refractory effects generated by spikes t_i^f of the postsynaptic neuron

$$u_i(t|\mathbf{x}, y_t^i) = u_{\rm rest} + \sum_{j=1}^N w_{ij} \sum_{t_j^f \in x_j} \epsilon(t - t_j^f) + \sum_{t_i^f \in y_t^i} \eta(t - t_i^f) + u_{\rm teach}(t) \tag{3}$$

where $u_{\rm rest} = -70$ mV is the resting potential, $y_t^i = \{t_i^1, t_i^2, \ldots, t_i^F < t\}$ is the set of postsynaptic spikes that occurred before t and t_i^F always denotes the last postsynaptic spike before t. On the right-hand side of Eq. (3), $\eta(s)$ denotes the spike-afterpotential generated by an action potential. We take

$$\eta(s) = \eta_0 \exp\left(-\frac{s}{\tau_m}\right) \Theta(s) \tag{4}$$

where $\eta_0 < 0$ is a reset parameter that describes how much the voltage is reset after each spike; for the relation to integrate-and-fire neurons see (Gerstner and

Kistler 2002b). The spikes themselves are not modeled explicitly but reduced to formal firing times. Unless specified otherwise, we take $\eta_0 = -5$ mV.

In a deterministic version of the model, output spikes would be generated whenever the membrane potential u_i reaches a threshold ϑ. In order to account for intrinsic noise, and also for a small amount of 'synaptic noise' generated by stochastic spike arrival from additional excitatory and inhibitory presynaptic neurons which are not modeled explicitly we replace the strict threshold by a stochastic one. More precisely we adopt the following procedure (Gerstner and Kistler 2002b). Action potentials of the postsynaptic neuron i are generated by a point process with time dependent stochastic intensity $\rho_i(t) = g(u_i(t))$ that depends non-linearly upon the membrane potential u_i. Since the membrane potential in turn depends on both the input and the firing history of the postsynaptic neuron, we write:

$$\rho_i(t|\mathbf{x}, y_t^i) = g(u_i(t|\mathbf{x}, y_t^i)). \tag{5}$$

We take an exponential to describe the stochastic escape across threshold, i.e, $g(u) = \rho_0 \exp\left(\frac{u-\vartheta}{\Delta u}\right)$ where $\vartheta = -50$ mV is the formal threshold, $\Delta u = 3$ mV is the width of the threshold region and therefore tunes the stochasticity of the neuron, and $\rho_0 = 1/\text{ms}$ is the stochastic intensity at threshold (see Fig. 1). Other choices of the escape function g are possible with no qualitative change of the results. For $\Delta u \to 0$, the model is identical to the deterministic leaky integrate-and-fire model with synaptic current injection (Gerstner and Kistler 2002b).

We note that the stochastic process, defined in Eq. (5) is similar to, but different from a Poisson process since the stochastic intensity depends on the set y_t of the previous spikes of the postsynaptic neuron. Thus the neuron model has some 'memory' of previous spikes.

2.3 Stochastic generative model

The advantage of the probabilistic framework introduced above via the noisy threshold is that it is possible to describe the probability density[1] $P_i(y|\mathbf{x})$ of

[1]For the sake of simplicity, we denoted the set of postsynaptic spikes from 0 to T by y instead of y_T.

an entire spike train[2] $Y(t) = \sum_{t_i^f \in y} \delta(t - t_i^f)$ (see appendix A for details):

$$\begin{aligned} P_i(y|\mathbf{x}) &= \left(\prod_{t_i^f \in y} \rho_i(t_i^f | \mathbf{x}, y_{t_i^f}) \right) \exp\left(- \int_0^T \rho_i(s|\mathbf{x}, y_s) ds \right) \\ &= \exp\left(\int_0^T \log(\rho_i(s|\mathbf{x}, y_s)) Y(s) - \rho_i(s|\mathbf{x}, y_s) ds \right) \end{aligned} \tag{6}$$

Thus we have a generative model that allows us to describe explicitly the likelihood $P_i(y|\mathbf{x})$ of emitting a set of spikes y for a given input $\mathbf{x}$. Moreover, since the likelihood in Eq. (6) is a smooth function of its parameters, it is straightforward to differentiate it with respect to any variable. Let us differentiate $P_i(y|\mathbf{x})$ with respect to the synaptic efficacy w_{ij}, since this is a quantity that we will use later on:

$$\frac{\partial \log P_i(y|\mathbf{x})}{\partial w_{ij}} = \int_0^T \frac{\rho_i'(s|\mathbf{x}, y_s)}{\rho_i(s|\mathbf{x}, y_s)} [Y(s) - \rho_i(s|\mathbf{x}, y_s)] \sum_{t_j^f \in x_j} \epsilon(s - t_j^f) ds. \tag{7}$$

where $\rho_i'(s|\mathbf{x}, y_s) = \frac{dg}{du}|_{u=u_i(s|\mathbf{x}, y_s)}$.

In this paper, we propose three different *optimal models* called A, B and C (cf. Table 1). The models differ in the stimulation paradigm and the specific task of the neuron. In section 3, the task and hence the optimality criteria are supposed to be given explicitly. However, the task in model C could also be defined indirectly by the presence or absence of a reward signal as discussed in section 4.1. The common idea behind all three approaches (A-C) is the notion of optimal performance. Optimality is defined by an objective function L that is directly related to the likelihood formula of Eq. (6) and that can be maximized by changes of the synaptic weights. Throughout the paper, this optimization is done by a standard technique of gradient ascent:

$$\Delta w_{ij} = \alpha \frac{\partial L}{\partial w_{ij}} \tag{8}$$

with a learning rate α. Since the three models correspond to three different tasks, they have a slightly different objective function. Therefore, gradient ascent yields slightly different strategies for synaptic update. In the following we start with the simplest model with the aim to illustrate the basic principles that generalize to the more complex models.

[2]Capital Y is the spike train generated by the ensemble (lower case) y.

3 Results

In this section we present synaptic updates rules derived by optimizing the likelihood of postsynaptic spike firing at some desired firing time $t^{\rm des}$. The essence of the argument is introduced in a particularly simple scenario, where the neuron is stimulated by one presynaptic spike and the neuron is inactive except at the desired firing time $t^{\rm des}$. This is the raw scenario that is further developed in several different directions.

Firstly, we may ask the question of how the postsynaptic spike at the desired time $t^{\rm des}$ is generated: (i) the spike could simply be given by a supervisor. As always in maximum likelihood approaches, we then optimize the likelihood that this spike could have been generated by the neuron model (i.e., the generative model) given the known input; (ii) the spike could have been generated by an strong current pulse of short duration applied by the supervisor (teaching input). In this case the a priori likelihood that the generative model fires at or close to the desired firing time is much higher. The two conceptual paradigms give slightly different results as discussed in scenario A.

Secondly, we may, in addition to the spike at the desired time $t^{\rm des}$ allow for other postsynaptic spikes generated spontaneously. The consequences of spontaneous activity for the STDP function are discussed in scenario B. Thirdly, instead of imposing a single postsynaptic spike at a desired firing time $t^{\rm des}$, we can think of a temporal coding scheme where the postsynaptic neuron responds to one (out of M) presynaptic spike patterns with a desired output spike train containing several spikes while staying inactive for the other $M-1$ presynaptic spike patterns. This corresponds to a pattern classification task which is the topic of scenario C.

Moreover, optimization can be performed in an unconstrained fashion or under some constraint. As we will see in this section, the specific form of the constraint influences the results on STDP, in particular the strength of synaptic depression for *post-before-pre* timing. To emphasize this aspect, we discuss two different constraints. The first constraint is motivated by the observation that neurons have a preferred working point defined by a typical mean firing rate that is stabilized by homeostatic synaptic processes (Turrigiano and Nelson 2004). Penalizing deviations from a target firing rate is the constraint that we will use in scenario B. For very low target firing rate, the constraint reduces to the condition of 'no activity' which is the constraint implemented in scenario

A.

The second type of constraint is motivated by the notion of STDP itself: changes of synaptic plasticity should depend on the relative timing of pre- and postsynaptic spike firing and not on other factors. If STDP is to be implemented by some physical or chemical mechanisms with finite time constants, we must require the STDP function to be local in time, i.e., the amplitude of the STDP function approaches zero for large time differences. This is the temporal locality constraint used in scenario C. While the unconstrained optimization problems are labeled with the subscript u (A_{u}, B_{u}, C_{u}), the constrained problems are marked by the subscript c (A_{c}, B_{c}, C_{c}) (c.f Table 1).

Unconstrained scenarios		Constrained scenarios	
A_{u}	Postsynaptic spike imposed $L^{A_{\mathrm{u}}} = \log(\rho(t^{\mathrm{des}}))$	A_{c}	No activity $L^{A_{\mathrm{c}}} = L^{A_{\mathrm{u}}} - \int_0^T \rho(t)dt$
B_{u}	Postsynaptic spike imposed + spontaneous activity $L^{B_{\mathrm{u}}} = \log(\bar{\rho}(t^{\mathrm{des}}))$	B_{c}	Stabilized activity $L^{B_{\mathrm{c}}} = L^{B_{\mathrm{u}}} - \frac{1}{T\sigma^2}\int_0^T(\bar{\rho}(t)-\nu_0)^2dt$
C_{u}	Postsynaptic spike patterns imposed $L^{C_{\mathrm{u}}} = \log\left(\prod_i P_i(y^i\|\mathbf{x}^i)\prod_{k\neq i} P_i(0\|\mathbf{x}^k)^{\frac{\gamma}{M-1}}\right)$	C_{c}	Temporal locality constraint $L^{C_{\mathrm{c}}} = L^{C_{\mathrm{u}}}$, $P_{\Delta\Delta'} = a\delta_{\Delta\Delta'}\left(\Delta - \tilde{T}_0\right)^2$

Table 1: Summary of the optimality criterion L for the three unconstrained scenarios (A_{u}, B_{u}, C_{u}) and the three constrained scenarios (A_{c}, B_{c}, C_{c}). The constraint for scenario C is not included in the likelihood function $L^{C_{\mathrm{c}}}$ itself, but rather in the deconvolution with a matrix P that penalizes quadratically the terms that are non-local in time. See appendix C for more details.

3.1 Scenario A: "one postsynaptic spike imposed"

Let us start with a particularly simple model which consists of one presynaptic neuron and one postsynaptic neuron (c.f. Fig. 2A). Let us suppose that the task of the postsynaptic neuron i is to fire a single spike at time t^{des} in response to the input which consists of a single presynaptic spike at time t^{pre}, i.e. the input is $x = \{t^{\mathrm{pre}}\}$ and the desired output of the postsynaptic neuron is $y = \{t^{\mathrm{des}}\}$. Since there is only a single pre- and a single postsynaptic neuron involved, we

drop in this section the indices j and i of the two neurons.

3.1.1 Unconstrained scenario A_{u}: 1 spike at t^{des}

In this subsection, we assume that the postsynaptic neuron has not been active in the recent past, i.e. refractory effects are negligible. In this case, we have $\rho(t|\mathbf{x}, y_t) = \rho(t|\mathbf{x})$ because of the absence of previous spikes. Moreover, since there is only a single presynaptic spike (i.e. $\mathbf{x} = \{t^{\mathrm{pre}}\}$), we write $\rho(t|t^{\mathrm{pre}})$ instead of $\rho(t|\mathbf{x})$.

Since the task of the postsynaptic neuron is to fire at time t^{des}, we can define the optimality criterion $L^{A_{\mathrm{u}}}$ as the log-likelihood of the firing intensity at time t^{des}, i.e.

$$L^{A_{\mathrm{u}}} = \log\left(\rho(t^{\mathrm{des}}|t^{\mathrm{pre}})\right) \tag{9}$$

The gradient ascent on this function leads to the following STDP function:

$$\Delta w^{A_{\mathrm{u}}} = \alpha \frac{\partial L^{A_{\mathrm{u}}}}{\partial w} = \alpha \frac{\rho'(t^{\mathrm{des}}|t^{\mathrm{pre}})}{\rho(t^{\mathrm{des}}|t^{\mathrm{pre}})} \epsilon(t^{\mathrm{des}} - t^{\mathrm{pre}}) \tag{10}$$

where $\rho'(t|t^{\mathrm{pre}}) \equiv \frac{dg}{du}|_{u=u(t|t^{\mathrm{pre}})}$. Since this optimal weight change $\Delta w^{A_{\mathrm{u}}}$ can be calculated for any presynaptic firing time t^{pre}, we get a STDP function which depends on the time difference $\Delta t = t^{\mathrm{pre}} - t^{\mathrm{des}}$ (c.f. Fig. 2B). As we can see directly from Eq. (10), the shape of the potentiation is exactly a mirror image of an EPSP. This result is independent of the specific choice of the function $g(u)$.

The drawback of this simple model becomes apparent, if the STDP function given by Eq. 10 is iterated over several repetitions of the experiment. Ideally, it should converge to an optimal solution given by $\Delta w^{A_{\mathrm{u}}} = 0$ in Eq. (10). However, the optimal solution given by $\Delta w^{A_{\mathrm{u}}} = 0$ is problematic: for $\Delta t < 0$, the optimal weight tends towards ∞ whereas for $\Delta t \geq 0$, there is no unique optimal weight ($\Delta w^{A_{\mathrm{u}}} = 0$, $\forall w$). The reason of this problem is, of course, that the model describes only potentiation and includes no mechanisms for depression.

3.1.2 Constrained scenario A_{c}: "no other spikes than at t^{des}"

In order to get some insight of where the depression could come from, let us consider a small modification of the previous model. In addition to the fact that the neuron has to fire at time t^{des}, let us suppose that it should not

fire anywhere else. This condition can be implemented by an application of Eq. (6) to the case of a single input spike $x = \{t^{\rm pre}\}$ and a single output spike $y = \{t^{\rm des}\}$. In terms of notation we set $P(y|x) = P(t^{\rm des}|t^{\rm pre})$ and similarly $\rho(s|x,y) = \rho(s|t^{\rm pre}, t^{\rm des})$ and use Eq. (6) to find:

$$P(t^{\rm des}|t^{\rm pre}) = \rho(t^{\rm des}|t^{\rm pre}) \exp\left[-\int_0^T \rho(s|t^{\rm pre}, t^{\rm des})ds\right] . \tag{11}$$

Note that for $s \leq t^{\rm des}$, the firing intensity does not depend on $t^{\rm des}$, hence $\rho(s|t^{\rm pre}, t^{\rm des}) = \rho(s|t^{\rm pre})$ for $s \leq t^{\rm des}$. We define the objective function $L^{A_{\rm c}}$ as the log-likelihood of generating a single output spike at time $t^{\rm des}$, given a single input spike at $t^{\rm pre}$. Hence, with Eq. (11):

$$\begin{aligned} L^{A_{\rm c}} &= \log(P(t^{\rm des}|t^{\rm pre})) \\ &= \log(\rho(t^{\rm des}|t^{\rm pre})) - \int_0^T \rho(s|t^{\rm pre}, t^{\rm des})ds \end{aligned} \tag{12}$$

and the gradient ascent $\Delta w^{A_{\rm c}} = \alpha \partial L^{A_{\rm c}}/\partial w$ rule yields

$$\Delta w^{A_{\rm c}} = \alpha \frac{\rho'(t^{\rm des}|t^{\rm pre})}{\rho(t^{\rm des}|t^{\rm pre})} \epsilon(t^{\rm des} - t^{\rm pre}) - \alpha \int_0^T \rho'(s|t^{\rm pre}, t^{\rm des})\epsilon(s - t^{\rm pre})ds \tag{13}$$

Since we have a single postsynaptic spike at $t^{\rm des}$, Eq. (13) can directly be plotted as a STDP function. In Fig. 3 we distinguish two different cases. In Fig. 3A we optimize the likelihood $L^{A_{\rm c}}$ in the absence of any teaching input. To understand this scenario we may imagine that a postsynaptic spike has occurred spontaneously at the desired firing time $t^{\rm des}$. Applying the appropriate weight update calculated from Eq. (13) will make such a timing more likely the next time the presynaptic stimulus is repeated. The reset amplitude η_0 has only a small influence.

In Fig. 3B we consider a case where firing of the postsynaptic spike at the appropriate time was made highly likely by a teaching input of duration $\Delta T = 1$ ms centered around the desired firing $t^{\rm des}$. The form of the STDP function depends on the amount η_0 of the reset. If there is no reset $\eta_0 = 0$, the STDP function shows strong synaptic depression of synapses that become active *after* the postsynaptic spike. This is due to the fact that the teaching input causes an increase of the membrane potential that decays back to rest with the membrane time constant τ_m. Hence the window of synaptic depression is also exponential with the same time constant. Qualitatively the same is true, if we include a weak reset. The form of the depression window remains

the same, but its amplitude is reduced. Only for strong reset to or below resting potential the effect inverses. A weak reset is standard in applications of integrate-and-fire models to in vivo data and is one of the possibilities to explain the high coefficient of variation of neuronal spike trains in vivo (Bugmann, Christodoulou, and Taylor 1997; Troyer and Miller 1997).

A further property of the STDP functions in Fig. 3 is a negative offset for $|t^{\rm pre} - t^{\rm des}| \to \infty$. The amplitude of the offset can be calculated for $w \simeq 0$ and $\Delta t > 0$, i.e. $\Delta w_0 \simeq -\rho'(u_{\rm rest}) \int_0^\infty \epsilon(s) ds$. This offset is due to the fact that we do not want spikes at other times than $t^{\rm des}$. As a result, the optimal weight $w^\star$ (i.e. solution of $\Delta w^{A_{\rm u}}$), should be as negative as possible ($w^\star \to -\infty$ or $w^\star \to w^{\rm min}$ in the presence of a lower bound) for $\Delta t > 0$ or $\Delta t \ll 0$.

3.2 Scenario B: "spontaneous activity"

The constraint in Scenario A_c of having strictly no other postsynaptic spikes than the one at time $t^{\rm des}$ may seem artificial. Moreover, it is this constraint which leads to the negative offset of the STDP function discussed at the end of the previous paragraph. In order to relax the constraint of "no spiking", we allow in scenario B for a reasonable spontaneous activity. As above, we start with an unconstrained scenario B_u before we turn to the constrained scenario B_c.

3.2.1 Unconstrained scenario $B_{\rm u}$: maximize the firing rate at $t^{\rm des}$

Let us start with the simplest model which includes spontaneous activity. Scenario $B_{\rm u}$ is the analog of the model $A_{\rm u}$, but with two differences.

First, we include spontaneous activity in the model. Since $\rho(t|\mathbf{x}, y_t)$ depends on the spiking history for any given trial, we have to define a quantity which is independent of the specific realizations y of the postsynaptic spike train.

Secondly, instead of considering only one presynaptic neuron, we consider $N = 200$ presynaptic neurons, each of them emitting a single spike at time $t_j = j\delta t$, where $\delta t = 1$ ms (see Fig. 4A). The input pattern will therefore be described by the set of delayed spikes $\mathbf{x} = \{x_j = \{t_j\}, j = 1, \ldots, N\}$. As long as we consider only a single spatio-temporal spike pattern in the input, it is always possible to relabel neurons appropriately so that neuron $j+1$ fires after neuron j.

Let us define the instantaneous firing rate $\bar{\rho}(t)$ that can be calculated by averaging $\rho(t|y_t)$ over all realizations of postsynaptic spike trains:

$$\bar{\rho}(t|\mathbf{x}) = \langle \rho(t|\mathbf{x}, y_t) \rangle_{y_t|\mathbf{x}} \,. \tag{14}$$

Here the notation $\langle \cdot \rangle_{y_t|\mathbf{x}}$ means taking the average over all possible configuration of postsynaptic spikes up to t for a given input $\mathbf{x}$. In analogy to a Poisson process, a specific spike train with firing times $y_t = \{t_i^1, t_i^2, \ldots, t_i^F < t\}$ is generated with probability $P(y_t|\mathbf{x})$ given by Eq. (6). Hence, the average $\langle \cdot \rangle_{y_t|\mathbf{x}}$ of Eq. (14) can be written as follows (see appendix B for numerical evaluation of $\bar{\rho}(t)$):

$$\bar{\rho}(t|\mathbf{x}) = \sum_{F=0}^{\infty} \frac{1}{F!} \int_0^t \cdots \int_0^t \rho(t|\mathbf{x}, y_t) P(y_t|\mathbf{x}) dt_i^F \ldots dt_i^1. \tag{15}$$

Analogously to the model A_{u}, we can define the quality criterion as the log-likelihood $L^{B_{\mathrm{u}}}$ of firing at the desired time t^{des}:

$$L^{B_{\mathrm{u}}} = \log(\bar{\rho}(t^{\mathrm{des}}|\mathbf{x})). \tag{16}$$

Thus the optimal weight adaptation of synapse j is given by

$$\Delta w_j^{B_{\mathrm{u}}} = \alpha \frac{\partial \bar{\rho}(t^{\mathrm{des}}|\mathbf{x})/\partial w_j}{\bar{\rho}(t^{\mathrm{des}}|\mathbf{x})} \tag{17}$$

where $\frac{\partial \bar{\rho}(t|\mathbf{x})}{\partial w_j}$ is given by

$$\frac{\partial \bar{\rho}(t|\mathbf{x})}{\partial w_j} = \bar{\rho}'(t|\mathbf{x}) \epsilon(t - t_j) + \left\langle \rho(t|\mathbf{x}, y_t) \frac{\partial}{\partial w_j} \log P(y_t|\mathbf{x}) \right\rangle_{y_t|\mathbf{x}}, \tag{18}$$

$\frac{\partial}{\partial w_j} \log P(y_t|\mathbf{x})$ is given by Eq. (7) and $\bar{\rho}'(t|\mathbf{x}) = \left\langle \frac{dg}{du}|_{u=u(t|\mathbf{x},y_t)} \right\rangle_{y_t|\mathbf{x}}$.

Figure 4B shows that, for our standard set of parameters, the differences to scenario A_{u} are negligible.

Figure 5A depicts the STDP function for various values of the parameter Δu at a higher postsynaptic firing rate. We can see a small undershoot in the *pre-before-post* region. The presence of this small undershoot can be understood as follows: enhancing a synapse of a presynaptic neuron that fires too early would induce a postsynaptic spike that arrives before the desired firing time and therefore, because of refractoriness, would prevent the generation of a spike at the desired time. The depth of this undershoot decreases with the stochasticity of the neuron and increases with the amplitude of the refractory period (if $\eta_0 = 0$, there is no undershoot). In fact, correlations between

pre- and postsynaptic firing reflect the shape of an EPSP in the high-noise regime, whereas they show a trough for low noise (Poliakov et al. 1997; Gerstner 2001). Our theory shows that the *pre-before-post* region of the optimal plasticity function is a mirror image of these correlations.

3.2.2 Constrained scenario $B_{\rm c}$: "firing rate close to ν_0"

In analogy to model $A_{\rm c}$ we now introduce a constraint. Instead of imposing strictly no spikes at times $t \neq t^{\rm des}$, we can relax the condition and minimize deviations of the instantaneous firing rate $\bar{\rho}(t|\mathbf{x}, t^{\rm des})$ from a reference firing rate ν_0. This can be done by introducing into Eq. (16) a penalty term $P_{\rm B}$ given by

$$P_{\rm B} = \exp\left(-\frac{1}{T}\int_0^T \frac{(\bar{\rho}(t|\mathbf{x}, t^{\rm des}) - \nu_0)^2}{2\sigma^2} dt\right). \tag{19}$$

For small σ, deviations from the reference rate yields a large penalty. For $\sigma \to \infty$, the penalty term has no influence. The optimality criterion is a combination of a high firing rate $\bar{\rho}$ at the desired time under the constraint of small deviations from the reference rate ν_0. If we impose the penalty as a multiplicative factor and take as before the logarithm, we get:

$$L^{B_{\rm c}} = \log\left(\bar{\rho}(t^{\rm des}|\mathbf{x}) P_{\rm B}\right) \tag{20}$$

Hence the optimal weight adaptation is given by

$$\Delta w_j^{B_{\rm c}} = \alpha \frac{\partial \bar{\rho}(t^{\rm des}|\mathbf{x})/\partial w_j}{\bar{\rho}(t^{\rm des}|\mathbf{x})} - \frac{\alpha}{T\sigma^2}\int_0^T (\bar{\rho}(t|\mathbf{x}, t^{\rm des}) - \nu_0) \frac{\partial}{\partial w_j}\bar{\rho}(t|\mathbf{x}, t^{\rm des}) dt. \tag{21}$$

Since in scenario B each presynaptic neuron j fires exactly once at time $t_j = j\delta t$ and the postsynaptic neuron is trained to fire at time $t^{\rm des}$, we can interpret the weight adaptation $\Delta w_j^{B_{\rm c}}$ of Eq. (21) as a STDP function $\Delta w^{B_{\rm c}}$ which depends on the time difference $\Delta t = t^{\rm pre} - t^{\rm des}$. Fig. 5 shows this STDP function for different values of the free parameter σ of Eq. (19). The higher the standard deviation σ, the less effective is the penalty term. In the limit of $\sigma \to \infty$, the penalty term can be ignored and the situation is identical to that of scenario $B_{\rm u}$.

3.3 Scenario C: "pattern detection"

3.3.1 Unconstrained scenario C_{u}: "spike pattern imposed"

This last scenario is a generalization of the scenario A_{c}. Instead of restricting the study to a single pre- and postsynaptic neuron, we consider N presynaptic neurons and M postsynaptic neurons (see Fig. 6). The idea is to construct M independent *detector neurons*. Each detector neuron $i = 1, \ldots, M$, should respond best to a specific prototype stimulus, say $\mathbf{x}^i$, by producing a desired spike train y^i, but should not respond to other stimuli, i.e. $y^i = 0$, $\forall \mathbf{x}^k$, $k \neq i$ (see Fig. 6). The aim is to find a set of synaptic weights that maximize the probability that neuron i produces y^i when $\mathbf{x}^i$ is presented and produces no output when $\mathbf{x}^k$, $k \neq i$ is presented. Let the likelihood function $L^{C_{\mathrm{u}}}$ be

$$L^{C_{\mathrm{u}}} = \log \left(\prod_{i=1}^{M} P_i(y^i|\mathbf{x}^i) \prod_{k=1, k\neq i}^{M} P_i(0|\mathbf{x}^k)^{\frac{\gamma}{M-1}} \right) \tag{22}$$

where $P_i(y^i|\mathbf{x}^i)$ (c.f Eq. (6)) is the probability that neuron i produces the spike train y^i when the stimulus $\mathbf{x}^i$ is presented. The parameter γ characterizes the relative importance of the patterns that should not be learned compared to those that should be learned. We get

$$L^{C_{\mathrm{u}}} = \sum_{i=1}^{M} \log(P_i(y^i|\mathbf{x}^i)) + \gamma \left\langle \log(P_i(0|\mathbf{x}^k)) \right\rangle_{\mathbf{x}^k \neq \mathbf{x}^i} \tag{23}$$

where the notation $\langle \cdot \rangle_{\mathbf{x}^k \neq \mathbf{x}^i} \equiv \frac{1}{M-1} \sum_{k \neq i}^{M}$ means taking the average over all patterns other than $\mathbf{x}^i$. The optimal weight adaptation yields

$$\Delta w_{ij}^{C} = \alpha \frac{\partial}{\partial w_{ij}} \log(P_i(y^i|\mathbf{x}^i)) + \alpha\gamma \left\langle \frac{\partial}{\partial w_{ij}} \log(P_i(0|\mathbf{x}^k)) \right\rangle_{\mathbf{x}^k \neq \mathbf{x}^i} \tag{24}$$

The learning rule of Eq. (24) gives the optimal weight change for each synapse and can be evaluated after presentation of all pre- and postsynaptic spike patterns, i.e. it is a "batch" update rule. Since each pre- and postsynaptic neuron emit many spikes in the interval $[0, T]$, we can not directly interpret the result of Eq. (24) as a function of the time difference $\Delta t = t^{\mathrm{pre}} - t^{\mathrm{des}}$ as we did in scenario A or B.

Ideally, we would like to write the total weight change of the optimal rule

given by Eq. (24) as a sum of contributions

$$\Delta w_{ij}^{C} = \sum_{\substack{t^{\mathrm{pre}} \in x_j^i \\ t^{\mathrm{des}} \in y^i}} \Delta W^{C_{\mathrm{u}}}(t^{\mathrm{pre}} - t^{\mathrm{des}}), \tag{25}$$

where $\Delta W^{C_{\mathrm{u}}}(t^{\mathrm{pre}} - t^{\mathrm{des}})$ is a STDP function and the summation runs over all pairs of pre- and postsynaptic spikes. The number of pairs of pre- and postsynaptic spikes with a given time shift is given by the correlation function which is best defined in discrete time. We assume time steps of duration $\delta t = 0.5$ ms. Since the correlation will depend on the presynaptic neuron j *and* the postsynaptic neuron i under consideration, we introduce a new index $k = N(i-1)+j$. We define the correlation in discrete time by its matrix elements $C_{k\Delta}$ that describe the correlation between the presynaptic spike train $X_j^i(t)$ and the postsynaptic spike train $Y^i(t - T_0 + \Delta\delta t)$. For example, $C_{3\Delta} = 7$ implies that 7 spike pairs of presynaptic neuron $j = 3$ with postsynaptic neuron $i = 1$ have a relative time shift of $T_0 - \Delta\delta t$. With this definition, we can rewrite Eq. (25) in vector notation (see appendix C.1 for more details):

$$\Delta \mathbf{w}^{C} \stackrel{!}{=} C \Delta \mathbf{W}^{C_{\mathrm{u}}} \tag{26}$$

where $\Delta \mathbf{w}^{C} = (\Delta w_{11}^{C}, \ldots, \Delta w_{1N}^{C}, \Delta w_{21}^{C}, \ldots, \Delta w_{MN}^{C})^T$ is the vector containing all the optimal weight change given by Eq. (24) and $\Delta \mathbf{W}^{C_{\mathrm{u}}}$ is the vector containing the discretized STDP function with components $\Delta W_{\Delta}^{C_{\mathrm{u}}} = \Delta W^{C_{\mathrm{u}}}(-T_0 + \Delta\delta t)$ for $1 \leq \Delta \leq 2\tilde{T}_0$ with $\tilde{T}_0 = T_0/\delta$. In particular, the center of the STDP function (i.e. $t^{\mathrm{pre}} = t^{\mathrm{des}}$) corresponds to the index $\Delta = \tilde{T}_0$. The symbol $\stackrel{!}{=}$ expresses the fact that we want to find $\Delta \mathbf{W}^{C_{\mathrm{u}}}$ such that $\Delta \mathbf{w}^{C}$ is as close as possible to $C \Delta \mathbf{W}^{C_{\mathrm{u}}}$. By taking the pseudo-inverse $C^{+} = (C^T C)^{-1} C^T$ of C, we can invert Eq. (26) and get

$$\Delta \mathbf{W}^{C_{\mathrm{u}}} = C^{+} \Delta \mathbf{w}^{C} \tag{27}$$

The resulting STDP function is plotted in Fig. 8A. As it was the case for the scenario A_{u}, the STDP function exhibits a negative offset. In addition to the fact the postsynaptic neuron i should not fire at other times than the ones given by y^i, it should also not fire whenever pattern $\mathbf{x}^k$, $k \neq i$ is presented. The presence of the negative offset is due to those two factors.

3.3.2 Constrained scenario C_c: "temporal locality"

In the previous paragraph, we obtained a STDP function with a negative offset. This negative offset does not seem realistic because it implies that the STDP function is not localized in time. In order to impose temporal locality (finite memory span of the learning rule) we modify Eq. (27) in the following way (see appendix C.2 for more details):

$$\Delta \mathbf{W}^{C_c} = (C^T C + P)^{-1} C^T \Delta \mathbf{w}^C \tag{28}$$

where P is a diagonal matrix which penalizes non-local terms. In this paper, we take a quadratic suppression of terms that are non-local in time. With respect to a postsynaptic spike at $t^{\rm des}$, the penalty term is proportional to $(t-t^{\rm des})^2$. In matrix notation, and using our convention that the postsynaptic spike corresponds to $\Delta = \tilde{T}_0$, we have:

$$P_{\Delta\Delta'} = a\delta_{\Delta\Delta'} \left(\Delta - \tilde{T}_0\right)^2 \tag{29}$$

The resulting STDP functions for different values of a are plotted in Fig. 8B. The higher the parameter a, the more non-local terms are penalized, the narrower is the STDP function.

Fig. 9A shows the STDP functions for various number of patterns M. No significant change can be observed for different numbers of input patterns M. This is due to the appropriately chosen normalization factor $1/(M-1)$ in the exponent of Eq. (22).

The target spike trains y^i have a certain number of spikes during the time window T, i.e. they set a target value for the mean rate. Let $\nu^{\rm post} = \frac{1}{TM}\sum_{i=1}^{M}\int_0^T y^i(t)dt$ be the imposed firing rate. Let w_0 denote the amplitude of the synaptic strength such that the firing rate $\bar{\rho}_{w_0}$ given by those weights is identical to the imposed firing rate: $\bar{\rho}_{w_0} = \nu^{\rm post}$. If the actual weights are smaller than w_0, almost all the weights should increase whereas if they are bigger than w_0, depression should dominate (c.f Fig 9B). Thus the exact form of the optimal STDP function depends on the initial weight value w_0. Alternatively, homeostatic process could assure that the mean weight value is always in the appropriate regime.

In Eqs. (25) and (26) we imposed that the total weight change should be generated as a sum over *pairs* of pre- and postsynaptic spikes. This is an assumption which has been made in order to establish a link to standard

STDP theory and experiments where spike pairs have been in the center of interest (Gerstner et al. 1996; Kempter et al. 1999; Kistler and van Hemmen 2000; Markram et al. 1997; Bi and Poo 1998; Zhang et al. 1998). It is, however, clear by now that the timing of spike pairs is only one of several factors contributing to synaptic plasticity. We therefore asked how much we miss if we attribute the 'optimal' weight changes calculated in Eq. (24) to spike-pair effects only. To answer this question we compared the optimal weight change Δw_{ij}^{C} from Eq. (24) with that derived from the pair-based STDP rule $\Delta w_{ij}^{\rm rec} = \sum_{t^{\rm pre}\in x_j^i} \sum_{t^{\rm des}\in y^i} \Delta W^{C_{\rm c}}(t^{\rm pre}-t^{\rm des})$ with or without locality constraint, i.e. for different values of the locality parameter ($a = 0, 0.04, 0.4$): see Fig. 10. More precisely, we simulate $M = 20$ detector neurons, each one of them having $N = 400$ presynaptic inputs, so each subplot of Fig. 10 contains 8000 points. Each point in a graph corresponds to the optimal change of one weight for one detector neuron (x axis) compared to the weight change of the very same weight due to pair based STDP (y axis). We found that in the absence of a locality constraint the pair-wise contributions are well correlated with the optimal weight changes. With strong locality constraints the quality of the correlation drops significantly. However, for a weak locality constraint that corresponds to a STDP function with reasonable potentiation and depression regimes, the correlation of the pair-based STDP rule with the optimal update is still good. This suggests that synaptic updates with a STDP-function based on pairs of pre- and postsynaptic spikes is close to optimal in the 'pattern detection' paradigm.

4 Discussion

4.1 Supervised vs Unsupervised and Reinforcement Learning

Our approach is based upon the maximization of the probability of firing at desired times $t^{\rm des}$ with or without constraints. From the point of view of machine learning, this is a supervised learning paradigm implemented as a maximum likelihood approach using the spike response model with escape noise as a generative model. Our work can be seen as a continuous-time extension of the maximum likelihood approach proposed in Barber (2003).

The starting point of all supervised paradigms is the comparison of a de-

Unconstrained scenarios		Constrained scenarios	
A_{u}	*pre-before-post* LTP $\sim$ EPSP	A_{c}	*post-before-pre* LTD (or LTP) $\sim$ spike afterpot.
B_{u}	*pre-before-post* LTP/LTD $\sim$ reverse correlation	B_{c}	*post-before-pre* LTD $\sim$ increased firing rate
C_{u}	*pre-before-post* LTP $\sim$ EPSP LTD $\sim$ background patterns	C_{c}	*post-before-pre* LTD $\sim$ background patterns $\sim$ temporal locality

Table 2: Main results for each scenario.

sired output with the actual output a neuron has, or would have, generated. The difference between the desired and actual output is then used as the driving signal for synaptic updates in typical model approaches (Minsky and Papert 1969; Haykin 1994; Bishop 1995). How does this compare to experimental approaches? Experiments to STDP have been mostly performed in vitro (Markram et al. 1997; Magee and Johnston 1997; Bi and Poo 1998). Since in typical experimental paradigms firing of the postsynaptic neuron is enforced by strong pulses of current injection, the neuron is not in a natural 'unsupervised' setting; on the other hand, the situation is also not fully supervised, since there is never a conflict between the desired and actual output of a neuron. In one of the rare in vivo experiments to STDP (Frégnac et al. 1988; Frégnac et al. 1992), the spikes of the postsynaptic neuron are also imposed by current injection. Thus, a classification of STDP experiments in terms of supervised, unsupervised, or reward based, is not as clearcut as it may seem at a first glance.

From the point of view of neuroscience, paradigms of unsupervised or reinforcement are probably much more relevant than the supervised scenario discussed here. However, most of our results from the supervised scenario analyzed in this paper, can be reinterpreted in the context of reinforcement learning following the approach proposed by Xie and Seung (2004). To illustrate the link between reinforcement learning and supervised learning, we define a global reinforcement signal $R(\mathbf{x}, y)$ which depends on the spike timing of the presynaptic neurons $\mathbf{x}$ and the postsynaptic neuron y. The quantity optimized in reinforcement learning is the expected reward $\langle R \rangle_{\mathbf{x},y}$ averaged

over all pre- and postsynaptic spike trains, i.e.

$$\langle R \rangle_{\mathbf{x},y} = \sum_{\mathbf{x},y} R(\mathbf{x},y) P(y|\mathbf{x}) P(\mathbf{x}) \tag{30}$$

If the goal of learning is to maximize the expected reward, we can define a learning rule which achieves this goal by changing synaptic efficacies in direction of the gradient of the expected reward $\langle R \rangle_{\mathbf{x},y}$:

$$\langle \Delta w \rangle_{\mathbf{x},y} = \alpha \left\langle R(\mathbf{x},y) \frac{\partial \log P(y|\mathbf{x})}{\partial w} \right\rangle_{\mathbf{x},y} \tag{31}$$

where α is a learning parameter and $\frac{\partial \log P(y|\mathbf{x})}{\partial w}$ is the quantity we discussed in this paper. Thus the quantities optimized in our supervised paradigm reappear naturally in a reinforcement learning paradigm.

For an intuitive interpretation of the link between reinforcement learning and supervised learning consider a postsynaptic spike that (spontaneously) occurred at time t_0. If no reward is given, no synaptic change takes place. However, if the postsynaptic spike at t_0 is linked to a rewarding situation, the synapse will try to recreate in the next trial a spike at the same time, i.e., t_0 has the role of the desired firing time t^{des} introduced in this paper. Thus the STDP function with respect to a postsynaptic spike at t^{des} derived in this paper, can be seen as the spike timing dependence which maximizes the expected reward in a spike-based reinforcement learning paradigm.

4.2 Interpretation of STDP Function

Let us now summarize and discuss our results in a broader context. In all three scenarios, we found an STDP function with potentiation for *pre-before-post* timing. Thus this result is structurally stable and independent of model details. However, depression for *post-before-pre* timing does depend on model details.

In scenario A, we have seen that the behavior of the *post-before-pre* region is determined by the spike afterpotential (see table 2 for a result summary of the three models). In the presence of a teaching input and firing rate constraints, a weak reset of the membrane potential after the spike means that the neuron effectively has a depolarizing spike after potential (DAP). In experiments, DAPs have been observed by Feldman (2000), Markram et al. (1997) and Bi and Poo (1998) for strong presynaptic input. Other studies,

however, have shown that the level of depression does not depend on the postsynaptic membrane potential (Sjöström et al. 2001). In any case, a weak reset (i.e., to a value below threshold rather than to the resting potential) is consistent with the findings of other researchers that used integrate-and-fire models to account for the high coefficient of variation of spike trains in vivo (Bugmann et al. 1997; Troyer and Miller 1997).

In the presence of spontaneous activity (c.f. scenario B), a constraint on the spontaneous firing rate causes the optimal weight change to elicit a depression of presynaptic spikes that arrive immediately after the postsynaptic one. In fact, the reason of the presence of the depression in scenario B_c is directly related to the presence of a DAP caused by the strong teaching stimulus. In both scenarios A and B, depression occurs in order to compensate the increased firing probability due to the DAP.

In scenario C, it has been shown that the best way to adapt the weights (in a task where the postsynaptic neuron has to detect a specific input pattern among others) can be described as a STDP function. This task is similar to the one in Izhikevich (2003) in the sense that a neuron is designed to be sensitive to a specific input pattern, but different since our work does not assume any axonal delays. The depression part in this scenario arises from a locality constraint. We impose that weight changes are explained by a sum of pair-based STDP functions.

There are various ways of defining objective functions and we have used three different objective functions in this paper. The formulation of an objective function gives a mathematical expression of the 'functional role' we assign to a neuron. The functional role depends on the type of coding (temporal coding or rate coding) and hence on the information the postsynaptic neurons will read out. The functional role also depends on the task or context in which a neuron is embedded. It might seem that different tasks and coding schemes could thus give rise to a huge number of objective functions. However the reinterpretation of our approach in the context of reinforcement learning provides a unifying viewpoint: even if the functional role of some neurons in a specific region of the brain can be different from other neurons of a different region, it is still possible to see the different objective functions as different instantiations of the same underlying concept - the maximization of the reward, where the reward is task specific.

More specifically, all objective functions used in this paper maximized the

firing probability at a desired firing time t^{des} - reflecting the fact that in the framework of timing based codes, the task of a neuron is to fire at precise moments in time. With a different assumption on the neuron's role on signal processing, different objective functions need to be used. An extreme case is a situation, where the neuron's task is to avoid firing at time t^{des}. A good illustration is given by the experiments done in the electrosensory lobe (ELL) of the electric fish (Bell et al. 1997). These cells receive two sets of input: the first one contains the pulses coming from the electric organ while the second input conveys information about the sensory stimulus. Since a large fraction of the sensory stimulus can be predicted by the information coming from the electric organ, it is computationally interesting to subtract the predictable contribution and focus only on the unpredictable part of the sensory stimulus. In this context, a reasonable task would be precisely to ask the neuron not to fire at time t^{des} where t^{des} is the time where the predictable simulation arrives and this task could be defined indirectly via an appropriate reward signal. An objective function of this type would, in the end, reverse the sign of the weight change of the causal part (LTD for the *pre-before-post* region), and this is precisely what is seen experimentally (Bell et al. 1997).

In our framework, the definition of the objective function is closely related to the neuronal coding. In scenario C, we postulate that neurons emit a precise spike train whenever the "correct" input is presented and be silent otherwise. This coding scheme is clearly not the most efficient one. Another possibility is to impose postsynaptic neurons to produce a specific but different spike train for each input pattern and not only for the "correct" input. Such a modification of the scenario does not dramatically change the results. The only effect is to reduce the amount of depression and increase the amount of potentiation.

4.3 Optimality Approaches vs Mechanistic Models

Theoretical approaches to neurophysiological phenomena in general, and to synaptic plasticity in particular, can be roughly grouped into three categories: first, biophysical models that aim at explaining the STDP function from principles of ion channel dynamics and intracelluar processes (Senn et al. 2001; Shouval et al. 2002; Abarbanel et al. 2002; Karmarkar and Buonomano 2002); second, mathematical models that start from a given STDP function and an-

alyze computational principles such as intrinsic normalization of summed efficacies or sensitivity to correlations in the input (Kempter et al. 1999; Roberts 1999; Roberts and Bell 2000; van Rossum et al. 2000; Kistler and van Hemmen 2000; Song et al. 2000; Song and Abbott 2001; Kempter et al. 2001; Gütig et al. 2003); finally, models that derive 'optimal' STDP properties for a given computational task (Chechik 2003; Dayan and Häusser 2004; Hopfield and Brody 2004; Bohte and Mozer 2005; Bell and Parra 2005; Toyoizumi et al. 2005a; Toyoizumi et al. 2005b). Optimizing the likelihood of postsynaptic firing in a predefined interval, as we did in this paper, is only one possibility amongst others of introducing concepts of optimality (Barlow 1961; Atick and Redlich 1990; Bell and Sejnowski 1995) into the field of STDP. Chechik (2003) uses concepts from information theory, but restricts his study to the classification of stationary patterns. The paradigm considered in Bohte and Mozer (2005) is similar to our scenario B_{c}, in that they use a fairly strong teaching input to make the postsynaptic neuron fire. Bell and Parra (2005) and Toyoizumi et al. (2005a) are also using concepts from information theory, but they are applying them to the pre- and postsynaptic spike trains. The work of Toyoizumi et al. (2005a) is a clearcut *unsupervised* learning paradigm and hence distinct from the present approach. Dayan and Häusser (2004) use concepts of optimal filter theory, but are not interested in precise firing of the postsynaptic neuron. The work of Hopfield and Brody (2004) is similar to our approach in that it focuses on recognition of temporal input patterns, but in our approach we are in addition interested in triggering postsynaptic firing with precise timing. Hopfield and Brody emphasize the repair of disrupted synapses in a network that has previously acquired its function of temporal pattern detector.

Optimality approaches, such as ours, will never be able to make *s*trict predictions about the properties of neurons or synapses. Optimality criteria may, however, help to elucidate computational principles and provide insights about potential tasks of electrophysiological phenomena such as STDP.

Acknowledgments

This work was supported by the Swiss National Science Foundation (200020-103530/1 and 200020-108093/1). T.T was supported by the Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists

and a Grant-in-Aid for JSPS Fellows.

A Probability Density of a Spike Train

The probability density of generating a spike train $y_t = \{t_i^1, t_i^2, \ldots, t_i^F < t\}$ with the stochastic process defined by Eq. (5) can be expressed as follows:

$$P(y_t) = P(t_i^1, \ldots, t_i^F) R(t|y_t) \tag{32}$$

where $P(t_i^1, \ldots, t_i^F)$ is the probability density of having F spikes at times $t_i^1, \ldots, t_i^F$ and $R(t|y_t) = \exp\left(-\int_{t_i^F}^{t} \rho(t'|y_{t'})dt'\right)$ corresponds to the probability of having no spikes from t_i^F to t. Since the joint probability $P(t_i^1, \ldots, t_i^F)$ can be expressed as a product of conditional probabilities

$$P(t_i^1, \ldots, t_i^F) = P(t_i^1) \prod_{f=2}^{F} P(t_i^f | t_i^{f-1}, \ldots, t_i^1) \tag{33}$$

Eq. (32) becomes

$$\begin{aligned} P(y_t) &= \rho(t_i^1|y_{t_i^1}) \exp\left(-\int_0^{t_i^1} \rho(t'|y_{t'})dt'\right) \\ &\cdot \left\{\prod_{f=2}^{F} \rho(t_i^f|y_{t_i^f}) \exp\left(-\int_{t_i^{f-1}}^{t_i^f} \rho(t'|y_{t'})dt'\right)\right\} \exp\left(-\int_{t_i^F}^{t} \rho(t'|y_{t'})dt'\right) \\ &= \left(\prod_{t_i^f \in y_t} \rho(t_i^f|y_{t_i^f})\right) \exp\left(-\int_0^t \rho(t'|y_{t'})dt'\right) \end{aligned} \tag{34}$$

B Numerical Evaluation of $\bar{\rho}(t)$

Since it is impossible to numerically evaluate the instantaneous firing rate $\bar{\rho}(t)$ with the analytical expression given by Eq. (14), we have to do it in a different way. In fact, there are two different ways to evaluate $\bar{\rho}(t)$. Before going into the details, let us first recall that from the law of large numbers, the instantaneous firing rate is equal to the empirical density of spikes at time t:

$$\langle \rho(t|y_t) \rangle_{y_t} = \langle Y(t) \rangle_{Y(t)} \tag{35}$$

where $Y(t) = \sum_{t_i^f \in y_t} \delta(t - t_i^f)$ is one realization of the postsynaptic spike train. Thus the first and simplest method based on the r.h.s of Eq. (35) is to

build a PSTH by counting spikes in small time bins $[t, t+\delta t]$ over, say $K = 10'000$ repetitions of an experiment. The second, and more advanced method, consists in evaluating the l.h.s. of Eq. (35) by Monte-Carlo sampling: instead of averaging over all possible spike trains y_t, we generate $K = 10'000$ spike trains by repetition of the same stimulus. A specific spike train $y_t = \{t_i^1, t_i^2, \ldots, t_i^F < t\}$ will automatically appear with appropriate probability given by Eq. (6). The Monte-Carlo estimation $\tilde{\rho}(t)$ of $\bar{\rho}(t)$ can be written as

$$\tilde{\rho}(t) = \frac{1}{P} \sum_{m=1}^{P} \rho(t|y_t^m) \tag{36}$$

where y_t^m is the m^{th} spike train generated by the stochastic process given by Eq. (5). Since we use the analytical expression of $\rho(t|y_t^m)$, we will call Eq. (36) a semi-analytical estimation. Let us note that the semi-analytical estimation $\tilde{\rho}(t)$ converges more rapidly to the true value $\bar{\rho}(t)$ than the empirical estimation based on the PSTH.

In the limit of a Poisson process, i.e. $\eta_0 = 0$, the semi-analytical estimation $\tilde{\rho}(t)$ given by Eq. (36) is equal to the analytical expression of Eq. (14), since the instantaneous firing rate ρ of a Poisson process is independent of the firing history $y_t = \{t_i^1, t_i^2, \ldots, t_i^F < t\}$ of the postsynaptic neuron.

C Deconvolution

C.1 Deconvolution for Spike Pairs

With a learning rule such as Eq. (24), we know the optimal weight change Δw_{ij} for each synapse, but we still do not know the corresponding STDP function.

Let us first define the correlation function $c_k(\tau)$, $k = N(i-1)+j$ between the presynaptic spike train $X_j^i(t) = \sum_{t^{\text{pre}} \in x_j^i} \delta(t - t^{\text{pre}})$ and the postsynaptic spike train $Y^i(t) = \sum_{t^{\text{des}} \in y^i} \delta(t - t^{\text{des}})$:

$$c_k(\tau) = \int_0^T X_j^i(s) Y^i(s+\tau) ds, \qquad k = 1, \ldots, NM \tag{37}$$

where we allow a range $-T_0 \leq \tau \leq T_0$, with $T_0 \ll T$. Since the sum of the pair based weight change ΔW should be equal to the total adaptation of weights Δw_k, we can write:

$$\int_{-T_0}^{T_0} c_k(s) \Delta W(s) ds \stackrel{!}{=} \Delta w_k \qquad k = 1, \ldots, NM \tag{38}$$

If we want to express Eq. (37) in a matrix form, we need to descretize time in small bins δt and define the matrix element

$$C_{k\Delta} = \int_{\Delta\delta t - T_0}^{(\Delta+1)\delta t - T_0} c_k(s) ds \tag{39}$$

Now Eq. (37) becomes

$$\Delta \mathbf{w} \stackrel{!}{=} C \Delta \mathbf{W} \tag{40}$$

where $\Delta \mathbf{w} = (\Delta w_{11}, \ldots, \Delta w_{1N}, \Delta w_{21}, \ldots, \Delta w_{MN})^T$ is the vector containing all the optimal weight change and $\Delta \mathbf{W}$ is the vector containing the discretized STDP function, i.e. $\Delta W_\Delta = \Delta W(-T_0 + \Delta \delta t)$, for $\Delta = 1, \ldots, 2\tilde{T}_0$ with $\tilde{T}_0 = T_0/\delta t$.

In order to solve the last matrix equation, we have to compute the inverse of the non-square $NM \times 2\tilde{T}_0$ matrix C, which is known as the Moore-Penrose inverse (or the pseudo-inverse):

$$C^+ = (C^T C)^{-1} C^T \tag{41}$$

which exists only if $(C^T C)^{-1}$ exists. In fact, the solution given by

$$\Delta \mathbf{W} = C^+ \Delta \mathbf{w} \tag{42}$$

minimizes the square distance

$$D = \frac{1}{2}(C \Delta \mathbf{W} - \Delta \mathbf{w})^2 \tag{43}$$

C.2 Temporal Locality Constraint

If we want to impose a constraint of locality, we can add a term in the minimization process of Eq. (43) and define the following:

$$E = D + \frac{1}{2} \Delta \mathbf{W}^T P \Delta \mathbf{W} \tag{44}$$

where P is a diagonal matrix which penalizes non-local terms. In this paper, we take a quadratic suppression of terms that are non-local in time:

$$P_{\Delta\Delta'} = a \delta_{\Delta\Delta'} \left(\Delta - \tilde{T}_0\right)^2 \tag{45}$$

$\tilde{T}_0$ corresponds to the index of the vector $\Delta \mathbf{W}$ in Eqs. (40) and (44) for which $t^{\text{pre}} - t^{\text{des}} = 0$. Calculating the gradient of E given by Eq. (44) with respect to $\Delta \mathbf{w}$ yields

$$\nabla_{\Delta \mathbf{W}} E = C^T (C \Delta \mathbf{W} - \Delta \mathbf{w}) + P \Delta \mathbf{W} \tag{46}$$

By looking at the minimal value of E, i.e. $\nabla_{\Delta W} E = 0$, we have

$$\Delta \mathbf{W} = (C^T C + P)^{-1} C^T \Delta \mathbf{w} \tag{47}$$

By setting $a = 0$, we recover the previous case.

References

Abarbanel, H., R. Huerta, and M. Rabinovich (2002). Dynamical model of long-term synaptic plasticity. *Proc. Natl. Academy of Sci. USA 59*, 10137–10143.

Atick, J. and A. Redlich (1990). Towards a theory of early visual processing. *Neural Computation 4*, 559–572.

Barber, D. (2003). Learning in spiking neural assemblies. In *Advances in Neural Information Processing Systems 15*, pp. 149–156. MIT Press.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication*, pp. 217–234. MIT Press.

Bell, A. and T. Sejnowski (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation 7*, 1129–1159.

Bell, A. J. and L. C. Parra (2005). Maximising sensitivity in a spiking network. In L. K. Saul, Y. Weiss, and L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17*, pp. 121–128. Cambridge, MA: MIT Press.

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

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

Bi, G. and M. Poo (1999). Distributed synaptic modification in neural networks induced by patterned stimulation. *Nature 401*, 792–796.

Bi, G. and M. Poo (2001). Synaptic modification of correlated activity: Hebb's postulate revisited. *Ann. Rev. Neurosci. 24*, 139–166.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press.

Bohte, S. M. and M. C. Mozer (2005). Reducing spike train variability: A computational theory of spike-timing dependent plasticity. In L. K. Saul, Y. Weiss, and L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17*, pp. 201–208. Cambridge, MA: MIT Press.

Brody, C. and J. Hopfield (2003). Simple networks for spike-timing-based computation, with application to olfactory processing. *Neuron 37*, 843–852.

Bugmann, G., C. Christodoulou, and J. G. Taylor (1997). Role of temporal integration and fluctuation detection in the highly irregular firing of leaky integrator neuron model with partial reset. *Neural Computation 9*, 985–1000.

Carr, C. E. and M. Konishi (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci. 10*, 3227–3246.

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

Dayan, P. and M. Häusser (2004). Plasticity kernels and temporal statistics. In *Advances in Neural Information Processing Systems 16*. MIT Press.

Feldman, D. (2000). Timing-based LTP and LTD and vertical inputs to layer II/III pyramidal cells in rat barrel cortex. *Neuron 27*, 45–56.

Frégnac, Y., D. E. Shulz, S. Thorpe, and E. Bienenstock (1988). A cellular analogue of visual cortical plasticity. *Nature 333(6171)*, 367–370.

Frégnac, Y., D. E. Shulz, S. Thorpe, and E. Bienenstock (1992). Cellular analogs of visual cortical epigenesis. I plasticity of orientation selectivity. *Journal of Neuroscience 12(4)*, 1280–1300.

Gerstner, W. (2001). Coding properties of spiking neurons: reverse- and cross-correlations. *Neural Networks 14*, 599–610.

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

Gerstner, W. and W. K. Kistler (2002a). Mathematical formulations of Hebbian learning. *Biological Cybernetics 87*, 404–415.

Gerstner, W. and W. K. Kistler (2002b). *Spiking Neuron Models*. Cambridge University Press.

Gerstner, W., R. Ritz, and J. L. van Hemmen (1993). Why spikes? Hebbian learning and retrieval of time–resolved excitation patterns. *Biol. Cybern. 69*, 503–515.

Gütig, R., R. Aharonov, S. Rotter, and H. Sompolinsky (2003). Learning input correlations through non-linear temporally asymmetry hebbian plasticity. *J. Neuroscience 23*, 3697–3714.

Haykin, S. (1994). *Neural Networks*. Upper Saddle River, NJ: Prentice Hall.

Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature 376*, 33–36.

Hopfield, J. J. and C. D. Brody (2004). Learning rules and network repair in spike-timing-based computation networks. *Proc. Natl. Acad. Sci. USA 101*, 337–342.

Izhikevich, E. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks 14*, 1569–1572.

Johansson, R. and I. Birznieks (2004). First spikes in ensembles of human tactile afferents code complex spatial fingertip events. *Nature Neuroscience 7*, 170–177.

Karmarkar, U. and D. Buonomano (2002). A model of spike-timing dependent plasticity: one or two coincidence detectors. *J. Neurophysiology 88*, 507–513.

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

Kempter, R., W. Gerstner, and J. L. van Hemmen (2001). Intrinsic stabilization of output rates by spike-based Hebbian learning. *Neural Computation 13*, 2709–2741.

Kistler, W. M. and J. L. van Hemmen (2000). Modeling synaptic plasticity in conjunction with the timing of pre- and postsynaptic potentials. *Neural Comput. 12*, 385–405.

Legenstein, R., C. Naeger, and W. Maass (2005). What target functions can be learnt with spike-timing-dependent plasticity? *submitted*.

Magee, J. C. and D. Johnston (1997). A synaptically controlled associative signal for hebbian plastiticy in hippocampal neurons. *Science 275*, 209–213.

Markram, H., J. Lübke, M. Frotscher, and B. Sakmann (1997). Regulation of synaptic efficacy by coincidence of postysnaptic AP and EPSP. *Science 275*, 213–215.

Mehta, M. R., A. K. Lee, and M. A. Wilson (2002). Role of experience of oscillations in transforming a rate code into a temporal code. *Nature 417*, 741–746.

Minsky, M. L. and S. A. Papert (1969). *Perceptrons*. Cambridge Mass.: MIT Press.

Panzeri, S., R. Peterson, S. Schultz, M. Lebedev, and M. Diamond (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron 29*, 769–777.

Poliakov, A. V., R. K. Powers, and M. C. Binder (1997). Functional identification of the input-output transforms of motoneurones in the rat and cat. *J. Physiology 504*, 401–424.

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

Roberts, P. (1999). Computational consequences of temporally asymmetric learning rules: I. Differential Hebbian learning. *J. Computational Neuroscience 7*, 235–246.

Roberts, P. and C. Bell (2000). Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. *Computational Neuroscience 9*, 67–83.

Rubin, J., D. D. Lee, and H. Sompolinsky (2001). Equilibrium properties of temporally asymmetric Hebbian plasticity. *Physical Review Letters 86*, 364–367.

Senn, W., M. Tsodyks, and H. Markram (2001). An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. *Neural Computation 13*, 35–67.

Seung, S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron 40*, 1063–1073.

Shouval, H. Z., M. F. Bear, and L. N. Cooper (2002). A unified model of nmda receptor dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. USA 99*, 10831–10836.

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

Song, S. and L. Abbott (2001). Column and map development and cortical re-mapping through spike-timing dependent plasticity. *Neuron 32*, 339–350.

Song, S., K. Miller, and L. Abbott (2000). Competitive Hebbian learning through spike-time-dependent synaptic plasticity. *Nature Neuroscience 3*, 919–926.

Thorpe, S., A. Delorme, and R. Van Rullen (2001). Spike-based strategies for rapid processing. *Neural Networks 14*, 715–725.

Toyoizumi, T., J.-P. Pfister, K. Aihara, and W. Gerstner (2005b). Generalized bienenstock-cooper-munro rule for spiking neurons that maximizes information transmission. *Proc. National Academy Sciences (USA) 102*, 5239–5244.

Toyoizumi, T., J.-P. Pfister, K. Aihara, and W. Gerstner (2005a). Spike-timing dependent plasticity and mutual information maximization for a spiking neuron model. In L. K. Saul, Y. Weiss, and L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17*, pp. 1409–1416. Cambridge, MA: MIT Press.

Troyer, T. W. and K. Miller (1997). Physiological gain leads to high ISI variability in a simple model of a cortical regular spiking cell. *Neural Computation 9*, 971–983.

Turrigiano, G. and S. Nelson (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience 5*, 97–107.

van Rossum, M. C. W., G. Q. Bi, and G. G. Turrigiano (2000). Stable Hebbian learning from spike timing-dependent plasticity. *J. Neuroscience 20*, 8812–8821.

Xie, X. and S. Seung (2004). Learning in neural networks by reinforcement of irregular spiking. *Phys. Rev. E 69*, 041909.

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

Figure 1: **A**. Escape rate $g(u) = \rho_0 \exp\left(\frac{u-\vartheta}{\Delta u}\right)$. **B**. Firing rate of the postsynaptic neuron as a function of the amplitude I_0 of a constant stimulation current (arbitrary units). **C**. Interspike interval (ISI) distribution for different input currents.

Figure 2: **A**. Scenario A: a single presynaptic neuron connected to a postsynaptic neuron with a synapse of weight w. **B**. Optimal weight change given by Eq. (10) for the scenario $A_{\rm u}$. This weight change is exactly the mirror image of an EPSP.

Figure 3: Optimal weight adaptation for scenario $A_{\rm c}$ given by Eq. (7) in the case of a teaching signal (**A**) and in the absence of a teaching signal (**B**). The weight change in the *post-before-pre* is governed by the spike afterpotential $u_{AP}(t) = \eta(t) + u_{\rm teach}(t)$. The duration of the teaching input is $\Delta T = 1$ ms. The amplitude of the current I_0 is chosen so that $\max_t u_{\rm teach}(t) = 5$ mV. $u_{\rm rest}$ is chosen such that the spontaneous firing rate $g(u_{\rm rest})$ matches the desired firing rate $1/T$, i.e. $u_{\rm rest} = \Delta u \log \frac{\pi_0}{T} + \theta \simeq -60$ mV. The weight strength is $w = 1$.

Figure 4: Scenario B. **A**. $N = 200$ presynaptic neurons are firing one after the other at time $t_j = j\delta t$ with $\delta t = 1$ ms. **B**. The optimal STDP function of scenario $B_{\rm u}$.

Figure 5: **A**. The optimal STDP functions of scenario $B_{\rm u}$ for different level of stochasticity described by the parameter Δu. The standard value ($\Delta u = 3$ mV) is given by the solid line, decreased noise ($\Delta u = 1$ mV and $\Delta u = 0.5$ mV) are indicated by dot-dashed and dashed lines respectively . In the low-noise regime, enhancing a synapse which fires slightly too early can prevent the firing at the desired firing time $t^{\rm des}$ due to refractoriness. To increase the firing rate at $t^{\rm des}$ it is thence advantageous to decrease the firing probability some time before $t^{\rm des}$. Methods: For each value of Δu, the initial weight w_0 are set such that the spontaneous firing rate is $\bar{\rho} = 30$Hz. The amplitude of the STDP function is multiplied by Δu in order to make them comparable. Reset: $\eta_0 = -5$ mV. **B**. Scenario $B_{\rm c}$. Optimal STDP function for scenario $B_{\rm c}$ given by Eq. (21) for a teaching signal of duration $\Delta T = 1$ ms. The maximal increase of the membrane potential after 1 ms of stimulation with the teaching input is $\max_t u_{\rm teach}(t) = 5$ mV. Synaptic efficacies w_{ij} are initialized such that $u_0 = -60$ mV which gives a spontaneous rate of $\bar{\rho} = \nu_0 = 5$ Hz. Standard noise level: $\Delta u = 3$ mV.

Figure 6: Scenario C. N presynaptic neurons are fully connected to M post-synaptic neurons. Each postsynaptic neuron is trained to respond to a specific input pattern and not respond to $M-1$ other patterns as described by the objective function of Eq. (22).

Figure 7: Pattern detection after learning. **Top**. The *left* raster plot represents the input pattern the $i^{\rm th}$ neuron has to be sensitive to. Each line corresponds to one of the $N = 400$ presynaptic neurons. Each dot represents an action potential. The *right* figure represents one of the patterns the $i^{\rm th}$ neuron should not respond to. **Middle**. The *left* raster plot corresponds to 1000 repetitions of the output of neuron i when the corresponding pattern $\mathbf{x}^i$ is presented. The *right* plot is the response of neuron i to one of the pattern it should not respond to. **Bottom**. The *left* graph represents the probability density of firing when pattern $\mathbf{x}^i$ is presented. This plot can be seen as the PSTH of the middle graph. Arrows indicate the supervised timing neuron i learned. The *right* graph describes the probability density of firing when pattern $\mathbf{x}^k$ is presented. Note the different scales of vertical axis.

Figure 8: **A**. Optimal weight change for scenario $C_{\rm u}$. In this case, no locality constraint is imposed and the result is similar to the STDP function of scenario $A_{\rm c}$ (with $\eta_0 = 0$ and $u_{\rm teach}(t) = 0$) represented on Fig. 3. **B**. Optimal weight change for scenario $C_{\rm c}$ as a function of the locality constraint characterized by a. The stronger the importance of the locality constraint, the narrower is the spike-spike interaction. For **A** and **B**: $M = 20$, $\eta_0 = -5$ mV. The initial weights w_{ij} are chose so that the spontaneous firing rate matches the imposed firing rate.

Figure 9: **A**. Optimal STDP function as a function of the number of input patterns M. ($a = 0.04$, $N = 400$) **B**. Optimal weight change as a function of the weight w. If the weights are small *(dashed line)* potentiation dominates whereas if they are small *(dotted line)* depression dominates.

Figure 10: Correlation plot between the optimal synaptic weight change $\Delta \mathbf{w}^{\rm opt} = \Delta \mathbf{w}^{C_{\rm u}}$ and the reconstructed weight change $\Delta \mathbf{w}^{\rm rec} = C \Delta \mathbf{W}^{C_{\rm c}}$ using the temporal locality constraint. **A**. No locality constraint, i.e. $a = 0$. Deviations from the diagonal are due to the fact that the optimal weight change given by Eq. (24) can not be perfectly accounted for the sum of pair effects. The mean deviations are given by Eq. (43). **B**. A weak locality constraint ($a = 0.04$) almost does not change the quality of the weight change reconstruction. **C**. Strong locality constraint ($a = 0.4$). The horizontal lines arise since most synapses are subject to a few strong updates induced by pairs of pre- and postsynaptic spike times with small time shifts.