Spike-Timing Dependent Plasticity, Learning Rules

Walter Senn[1], Jean-Pascal Pfister[1]

AUTHOR COPY

Synonyms

STDP, Spike-timing-dependent synaptic plasticity, Spike-dependent synaptic learning rules

Definition

Biological phenomenon. Spike-timing dependent plasticity (STDP) in its narrow sense refers to the change in the synaptic strength as a result of electrically eliciting pairs of action potentials ('spikes') with a fixed time difference between the pre- and post-synaptic action potentials (Markram et al., 1997; Bi and Poo, 1998; Sjostrom et al., 2001). STDP is typically observed for synapses between hippocampal or cortical pyramidal neurons in slices of juvenile rodents, and the spike pairings are repeated 50-100 times with various frequencies, e.g. 1 or 10 Hz. This protocol induces a change in the amplitude of a single excitatory postsynaptic potential (EPSP) which is plotted against the spike time difference $\Delta t = t_{\text{post}} - t_{\text{pre}}$ between the postsynaptic spike and the presynaptic spike (Fig. 1). The change takes in many cases a few minutes to be expressed and lasts at least for the duration of the experiment. Typically, when the presynaptic spike precedes the postsynaptic spike by roughly 10 ms, the synapse is potentiated; if the presynaptic spike follows the postsynaptic spike, the synapse is depressed (for reviews see Bi and Poo (2001); Senn (2002); Sjostrom et al. (2008); Sjöström and Gerstner (2010)).
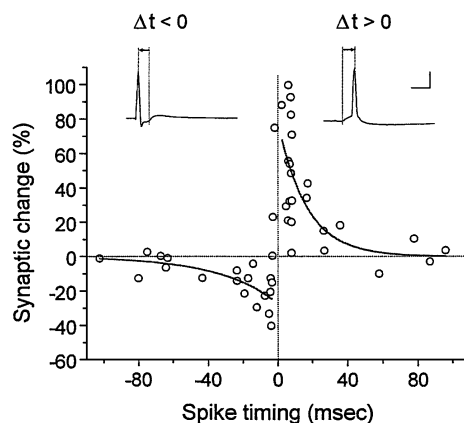
Figure 1. *Change of the EPSP amplitudes as a function of the time difference $\Delta t = t_{\text{post}} - t_{\text{pre}}$ between the post- and pre-synaptic spike. Pairing protocol: 60 spike pairs at 1 Hz. Inset: postsynaptic action potential, relative to the time of the presynaptic spike (vertical line). Scale bars: $10\,ms$, $50\,mV$. Figure from Bi and Poo (2001).*

Learning rules. In a computational context, STDP refers to plasticity rules that depend on the timing of pre- and postsynaptic spikes and that are involved in various learning scenarios for neuronal networks. These learning rules either emphasize the link to the biophysics underlying the synaptic modification

---

[1]Department of Physiology, University of Bern, Switzerland; senn@pyl.unibe.ch, pfister@pyl.unibe.ch

(Senn et al., 2001; Shouval et al., 2002; Karmarkar and Buonomano, 2002; Rubin et al., 2005; Graupner and Brunel, 2012), or are minimalistic with respect to a biological implementation (Kempter et al., 2001; Song and Abbott, 2001), or are derived from the maximization of a utility function (Pfister et al., 2006; Toyoizumi et al., 2007; Florian, 2007; Urbanczik and Senn, 2009; Friedrich et al., 2011). The learning rules are studied in the context of supervised, unsupervised, or reinforcement learning. When evaluated from the performance point of view, learning rules that are mathematically derived from an optimization principle are superior over STDP rules designed to fit a given set of experimental data (Frémaux et al., 2010). Interestingly, biological plausibility and computational relevance may go together when considering 2-compartment neurons with synapses on a dendritic tree (Urbanczik and Senn, 2014).

Detailed Description

STDP models come in different flavors, emphasizing more the phenomenology, the biophysics, or the computational aspects. As learning rules, their primary focus is on enabling computations rather than on reproducing synaptic plasticity data. An excellent and comprehensive review to STDP models, starting with the basic pair-based STDP models (Fig. 2A) and including also functional consequences, is found in the Scholarpedia article by Sjöström and Gerstner (2010). Here we highlight the properties of third-order STDP models and focus on gradient rules.

Phenomenological STDP models

The simplest online model that phenomenologically reproduces the basic STDP curve (Fig. 1) separately induces long-term potentiation (LTP) or long-term depression (LTD), by either a pre-post or post-pre coincidence detector, respectively. The key feature is that pre- and postsynaptic spikings are each tracked by a leaky integrator, the so-called synaptic eligibility traces, while LTP and LTD is triggered proportionally to these traces at the times of the post- and presynaptic spikes, respectively (Fig. 2A).

The triplet model The simple STDP model which depends on pairs of spikes (pre-post and post-pre) correctly predicts the weight change only for a restricted number of protocols. If potentiation is assumed to be governed by triplets of spikes (pre-post-post) instead of pairs of spikes, a much broader class of experimental data can be captured (Pfister and Gerstner, 2006). This so-called triplet model can be expressed as a sum of a depression term (Fig. 2 B2) and a triplet term where at the time of the postsynaptic spike the weight change is proportional to the product of a postsynaptic and a presynaptic eligibility trace (Fig. 2 B1).

The triplet model becomes especially relevant when the repetition frequency of the pre-post pairs increases. The pair-based model predicts a decrease of potentiation as a function of the pairing frequency. But in the visual cortex (L5→L5 pyramidal neurons, Sjostrom et al. (2001)) potentiation increases with increasing repetition frequency, and this is well reproduced by the triplet model (Fig. 2C; it is also qualitatively captured by the early STDP model by Senn (2002)).

This triplet model has also interesting computational properties. Under the assumption of independent pre- and postsynaptic Poisson firing rate (Pfister and Gerstner, 2006), the expected weight change predicted by the triplet model is consistent with the Bienenstock-Cooper-Munro (BCM) learning rule (Bienenstock et al., 1982) which elicits input selectivity, i.e. the output neuron becomes strongly responsive to one given (rate-based) input pattern and much less to all the other ones. Furthermore, if the independent Poisson assumption is relaxed such that output firing rate depends on the presynaptic spike timings, the triplet rule becomes sensitive to third-order spiking correlations in the input, thereby generalizing the BCM learning rule to spiking-correlated patterns (Gjorgjieva et al., 2011).

Extended models A next important extension of STDP models takes account of the modulation of plasticity by the postsynaptic voltage (Clopath et al., 2010; Clopath and Gerstner, 2010). This unifying model is formulated in terms of the postsynaptic voltage time course and presynaptic spikes. It can explain the widest set of STDP experiments, including burst-induced synaptic plasticity and those experiments that reveal the dependence on the postsynaptic voltage, as e.g. in Artola et al. (1990) and Sjostrom et al. (2001). This voltage-dependent model can also be seen as an extension of the triplet model where the postsynaptic eligibility trace in the potentiation term is replaced by a low-pass filter of the postsynaptic voltage. The triplet model, in turn, can be seen as a simplified version of the model by Senn et al. (2001). This latter model also depends on triple events (pre-post-post) for the induction of long-term plotentiation, but the pre-post-post ordering is important while in the triplet model both pre-post-post as well as post-pre-post events lead to potentiation.

Biophysical STDP models Another class of STDP models explains the synaptic modifications as a nonlinear function of the postsynaptic calcium concentration. The question whether the postsynaptic calcium alone can capture the characteristical STDP curve of Fig. 1 (see Shouval et al. (2002) versus Karmarkar and Buonomano (2002)) has been affirmed by taking into account of the calcium dynamics (Rubin et al., 2005) or additional nonlinearities (Graupner and Brunel, 2012). Functionally, these threshold-nonlinearities are very similar to the ones imposed on the pre- and postsynaptic eligibility traces introduced in the phenomenological models (Senn et al., 2001; Clopath et al., 2010). Yet, by starting with individual protein kinetics a biophysical model may explain how these nonlinearities arise (Rubin et al., 2005), see also Spike Timing-Dependent Plasticity (STDP), Biophysical Models.
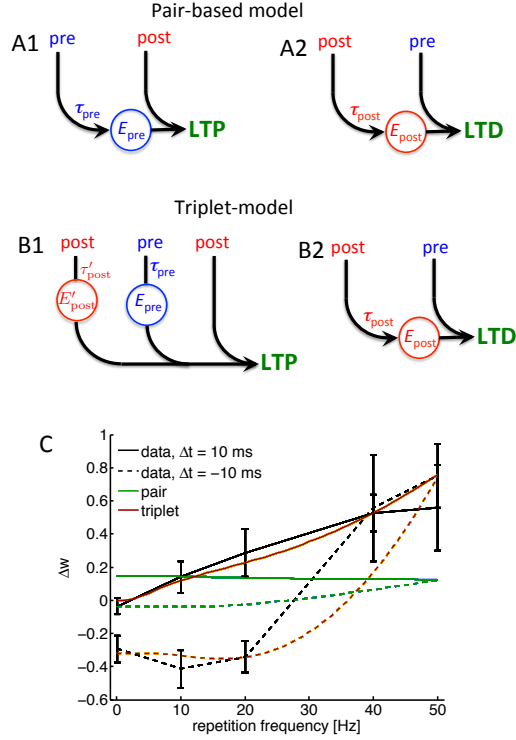
Gradient-based STDP learning rules

By their nature, the phenomenological and biophysical STDP models are not directly designed as synaptic learning rules that solve an explicit learning task. When canonical target functions for the learning can be defined, such as in the supervised and reinforcement learning scenario, spike-timing dependent learning rules can be derived from gradient procedures that maximize/minimize these functions. A very convenient neuron model suited for a theory of learning is the escape rate neuron. Indeed since it allows to explicitly quantify the probability for a given postsynaptic spike train as a function of the afferent synaptic strengths w_j, the likelihood of a given spike train is differentiable with respect to w_j, see e.g. Pfister et al. (2006). This neuron stochastically emits spikes with instantaneous firing rate $\rho(u)$ that is an increasing function of the instantaneous membrane potential $u(t)$. The latter is itself a sum of the postsynaptic potentials (PSP's) weighted by the synaptic strengths, $u(t) = \sum_j w_j \text{PSP}_j(t)$, optionally subtracted with a reset kernel after a postsynaptic spike.

Supervised learning In the supervised learning scenario, the target function can be defined as a distance between the desired postsynaptic spike train, $S_{\text{post}}^{\text{cl}}(t) = \sum_i \delta(t - t_i^{\text{post}})$, that is clamped as an output to the neuron, and the spike trains that would be generated by the neuron itself. If we pick out a specific synapse, the presynaptic eligibility trace $E_{\text{pre}}(t)$ is again obtained by the leaky integration of the presynaptic spike train $S_{\text{pre}}(t) = \sum_i \delta(t - t_i^{\text{pre}})$. Typically, the integration time constant τ_{pre} is equal to the membrane time constant, and hence this trace can also be identified with the postsynaptic potential induced by that synapse, $E_{\text{pre}}(t) = \text{PSP}(t)$. The gradient rule that maximizes the log-likelihood of reproducing the clamped target spike trains is then obtained as (Fig. 3A, Pfister et al. (2006))

$$E_{\text{pre}}(t) \quad = \quad \int_{-\infty}^{t} S_{\text{pre}}(\tilde{t}) e^{-\frac{t-\tilde{t}}{\tau_{\text{pre}}}} d\tilde{t} \tag{1}$$

$$\dot{w}(t) \quad = \quad \eta \frac{\rho'}{\rho} \left(S_{\text{post}}^{\text{cl}}(t) - \rho(u(t)) \right) E_{\text{pre}}(t), \tag{2}$$

Figure 2. *Phenomenological STDP models.* **(A)** *Simplest model reproducing Fig. 1.* **(A1)** *Each presynaptic spike stepwise increases a presynaptic eligibility trace E_{pre} that otherwise exponentially decays to 0 with time constant τ_{pre} ($\approx 17\,ms$ to fit the LTP in Fig. 1, see also Eq. 1). LTP is induced by each postsynaptic spike proportionally to the amount of E_{pre} available at that time.* **(A2)** *LTD is induced by each presynaptic spike proportionally to E_{post} that low-pass filters the postsynaptic spiking with time constant τ_{post} ($\approx 34\,ms$ to fit the LTD in Fig. 1). Note that the post-pre chain is itself acausal and does not appear in the gradient-based learning schemes represented in the subsequent figures.* **(B)** *Triplet rule.* **(B1)** *In the triplet model, LTP is induced at the time of the postsynaptic spike and is proportional to the product $E_{\mathrm{pre}}E'_{\mathrm{post}}$.* **(B2)** *In the triplet model, LTD is induced by pairs of spikes as in B1.* **(C)** *Weight change as a function of the repetition frequency of the pre-post pairs (solid lines, $t_{\mathrm{post}} - t_{\mathrm{pre}} = 10\,ms$) and the post-pre pairs (dashed lines , $t_{\mathrm{post}} - t_{\mathrm{pre}} = -10\,ms$). The triplet model (brown) fits well the data from (Sjostrom et al., 2001) (black) while the pair-based model (green) cannot.*

where η is some small learning rate. Here, $\rho' = \rho'(u(t))$ is the derivative of the escape rate ρ with respect to u, evaluated at t.

Interestingly, by expressing Eq. 2 as a sum of a potentiation and a depression term, we note that potentiation depends on 3 factors (the postsynaptic spike, the presynaptic eligibility trace and a nonlinear function of the postsynaptic membrane potential $\rho'(u)/\rho(u)$) and depression on 2 factors (the presynaptic eligibility trace and $\rho'(u)$). This learning rule has close similarities to the phenomenological voltage-triplet rule developed by Clopath and Gerstner (2010), with a correspondance highlighted in more details in Brea et al. (2013). A purely 2nd-order, phenomenological STDP rule has also been shown to perform well for supervised learning (Ponulak and Kasiński, 2010), but it assumes that target spikes and self-generated spikes can be distinguished by a synapse.
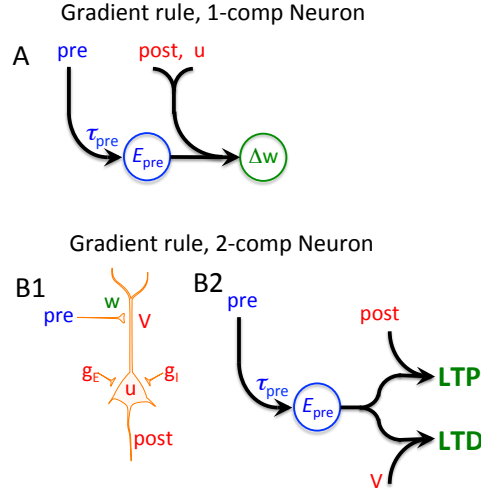
Figure 3. *Gradient-based SDTP for supervised learning (A,B) and unsupervised learning (B).* **(A)** *Such a gradient-based learning rule that has the goal to reproduce the timing of given ('clamped') output spike needs to take account of the postsynaptic membrane potential* u *beside the* pre- *and* post-*synaptic spikes (Eqs 1 & 2).* **(B1)** *In a biological version,* u *is only slightly 'nudged' by excitatory and inhibitory conductances* g_E *and* g_I. *The strength of synapses on the dendrites are adapted such that the dendritic potential* V *converges to the nudged somatic potential* u. **(B2)** *The corresponding gradient rule yields LTP that does only depend on the pre-post spike timings, and LTD that depends only on the pre-synaptic spike time (captured by* E_{pre}) *and the local dendritic voltage (*V, *see Eqs 1 & 3).*

Arguably, clamping the postsynaptic spike train S_{post}^{cl} is also biologically unfeasible as it would require that the membrane potential u is ∞ at the time of a target spike and $-\infty$ else, conflicting with the evaluation of ρ and ρ' at the synaptically generated value of u. A biologically motivated alternative is to separate the spike-generating voltage from the synaptically induced voltage and consider a somatic and dendritic membrane potential, u and V, that are interpreted as a 'teacher' (u) and 'student' (V) potential, respectively (Fig. 2, B1; Urbanczik and Senn (2014)). The soma receives conductance-based synaptic input that represents a teaching signal, and the postsynaptic spike train S_{post} is stochastically generated in the 'free' run, i.e. according to a firing intensity $\rho(u)$ that is affected by this teaching input. Without teaching input, the somatic membrane potential is just the attenuated dendritic voltage, $u = \alpha V$, where α represents some dendritic attenuation factor, and the instantaneous somatic firing is therefore $\rho(\alpha V)$. But if the somatic teaching input is turned on, the somatic voltage typically differs from the 'dendritic prediction', $u \neq \alpha V$. Learning is driven by the 'prediction error' measured in terms of the firing rates, $\rho(u) - \rho(\alpha V)$. It reduces this error by adapting the synaptic strengths of the dendritic 'student inputs'. At the synaptic location on the dendrite, the somatic rate $\rho(u)$ can be sampled by the backpropagating spikes S_{post}. The learning rule (2) now translates to the biological version,

$$\dot{w}(t) = \eta \frac{\rho'}{\rho} \left(S_{post}(t) - \rho(\alpha V(t)) \right) E_{pre}(t), \tag{3}$$

that can operate all the time, without need for clamping (Fig. 2, B2).

Crucially, after learning the teaching input driving the synaptic plasticity (3) can be turn off or on, without affecting the somatic voltage and hence without inducing additional weight changes. This is a consequence of the conductance-based teacher input that itself only changes the membrane potential if it deviates from the reversal potential defined that teaching input (Urbanczik and Senn, 2014). The rule shares other interesting biological features. When the backpropagation is hampered, say due to insufficient

dendritic depolarization, S_{post} is thinned out at the synaptic site and a putative LTP turns into LTD, as observed for synapses on the distal apical tree of cortical pyramidal neurons (Sjostrom and Hausser, 2006). Similarly, when the dendritic depolarization V is enhanced without additional postsynaptic spikes, LTD dominates as observed for these same cells (Sjostrom et al., 2004).

Unsupervised learning The learning rule (2) in the free run is itself not suited for unsupervised learning since averaging $\dot{w}$ across trials cancels out to 0 at each point in time. However, if we consider a 2-dimensional sheet of 2-compartmental neurons as described in Figure 2B, with Mexican-hat shaped somato-somatic connections, the somatic potential u is nudged away from V^* and the somatic firing in average does not anymore reflect the dendritic drive, $\langle S_{\text{post}} \rangle \neq \rho(V^*)$. In this case, the lateral connectivity induces a soft winner-take-all dynamics in the network that becomes a spike-based self-organizing feature map (Urbanczik and Senn, 2014). When the dendrites of these neurons are supplied by spatio-temporal spike patterns via plastic synapses governed by the rule (3), the feature map learns to cluster the spike patterns according to their similarity.

Another form of a gradient-based unsupervised learning that maximizes the mutual information between the pre- and postsynaptic spike trains was also shown to share classical STDP features while being able to develop receptive field properties (Toyoizumi et al., 2007). In the unsupervised setting, functional properties have also been shown for phenomenological STDP models in forming auditory maps (Gerstner et al., 1996), cortical columns (Song and Abbott, 2001), direction selective neurons in the visual cortex (Buchs and Senn, 2002), or receptive fields similarly as described in the BCM-theory (Gjorgjieva et al., 2011).
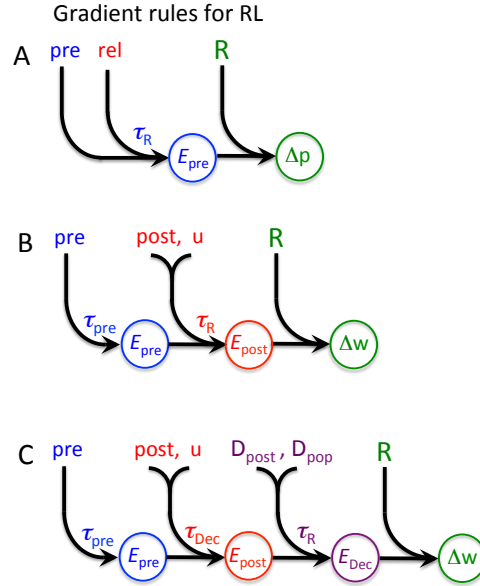
Reinforcement learning In reinforcement learning (RL), the putative synaptic weight changes induced by the pre- and postsynaptic activities are first low-pass filtered, and when a binary reward signal $R = \pm 1$ is applied, the changes accumulated until this time are multiplicatively modulated by R and turned into a real synaptic weight change. The phenomenological STDP model shown in Figure 2 has also been adapted to this reinforcement learning scenario where it is referred to as R-STDP (Izhikevich, 2007; Legenstein et al., 2008). However, R-STDP is shown to be problematic since for each stimulus class the expected reward must be 0 (Frémaux et al., 2010). This is because the integral over the STDP curve (Fig. 1) in general deviates from 0 and hence learning with $\langle R \rangle \neq 0$ would cause a weight drift. No weight drift is present for the rule in Eq. 2 in the free run, nor for Eq. 3 in the absence of somatic teaching conductances. These latter rule translates to the RL rule schematized in Figure 4B,

$$E_{\text{post}}(t) = \int_{-\infty}^{t} \frac{\rho'}{\rho} \left(S_{\text{post}}(\tilde{t}) - \rho(u(\tilde{t})) \right) E_{\text{pre}}(\tilde{t}) \, e^{-\frac{t-\tilde{t}}{\tau_{\text{R}}}} \, d\tilde{t} \tag{4}$$

$$\Delta w(T) = \eta R \, E_{\text{post}}(T), \tag{5}$$

where E_{pre} is given in (1) and for the 2-compartmental model the argument u of ρ and ρ' is replaced by V^*. The rule is shown to perform stochastic gradient ascent on the expected reward and has been studied in different applications (Xie and Seung, 2004; Pfister et al., 2006; Florian, 2007; Frémaux et al., 2010).

Stochastic gradient rules are not unique since the same gradient can be obtained from different estimators. The rule in Eq. 5, for instance, represents an estimator of the gradient of the expected reward, $\frac{\partial}{\partial w} \langle R \rangle = \langle R \frac{\partial}{\partial w} \log P_w(y|x) \rangle$, averaged across stimuli x, network activity y, time and reward. The reward R may depend on quantities downstream of x and y like the decision (or action) D that itself may stochastically depend on y. The reward $R(x, y)$ therefore is a stochastic function of (x, y) with conditional expectation $\langle R|x, y \rangle = \sum_D R(x, D) P(D|x, y)$. For a synapse that has only access to the pre- and postsynaptic activities (components of x and y), the samples $R(x, y)$ have a large variance and so will the samples $R(x, y) \frac{\partial}{\partial w} \log P_w(y|x)$ of the gradient estimate have. In contrast, $R(x, D)$ may be a deterministic

Figure 4. *Gradient-based SDTP for reinforcement learning (RL). With incorporating downstream quantities into the synaptic plasticity, learning becomes faster.* **(A)** *The simplest spike-based RL rule changes the presynaptic release probability* p *as a function of the presynaptic spike (*pre*) and the release (*rel*), low-pass filtered with a time constant* τ_R *corresponding to the typical reward delay (Seung, 2003).* **(B)** *The same synaptic modifications for supervised learning (Fig. 3) yields RL when low-pass filtered with* τ_R *and modulated with the delayed reward* R *(Eqs 1, 4 & 5).* **(C)** *As a decision is made by a population of neurons, synaptic updates should take account of the population decision signal* D_{pop}, *compare it with the single neuron decision* D_{post}, *low-pass filter the correlation between the two decision signals with* τ_R, *and only then implement the resulting weight change modulated by* R *(Eqs 1, 6-8). For an formal overview on these rules see also* [Reinforcement Learning in Cortical Networks](#).

function (or again a stochastic function with smaller variance) and the samples $R(x, D) \frac{\partial}{\partial w} \log P_w(D|x)$ of the same reward gradient $\frac{\partial}{\partial w}\langle R \rangle$ show a smaller variance. To calculate $\frac{\partial}{\partial w} \log P_w(D|x)$, however, a synapse needs to have access to D (beside the pre- and postsynaptic activities).

Instead of considering $R(x, y)$, the reward can even be seen as a stochastic function of only the presynaptic spikes and the synaptic releases, $R(x, \text{rel})$. This leads to a learning rule where synaptic releases that are correlated with subsequent rewards are made more likely by enhancing the corresponding release probability (Seung, 2003). But the variance of this reward gradient estimator can be reduced by taking account of the postsynaptic activity. In this way, more and more downstream information can be taken into account in the synaptic update, leading to learning rules that consider (A) only presynaptic spikes/releases and reward, (B) presynaptic spikes/releases, postsynaptic activity and reward, and (C) presynaptic spikes/releases, postsynaptic activities, single neuron and network decisions, and reward (Fig. 4). In these gradient estimators, the correlation between the synaptic parameter change and reward is progressively increased the more reward-relevant information the synapse exploits. In the case of only evaluating presynaptic spikes and releases, learning was claimed to mimic song acquisition in the zebra finch (Seung, 2003). When additionally evaluating the postsynaptic spikes and the membrane potential, the rule was shown to learn motor trajectories (Frémaux et al., 2010). When further evaluating the population decision the rule was shown to be successful in a complex sequential association task with delayed and scrambled rewards that are even hard to be learned by humans (Friedrich et al., 2011).

In population RL, the synaptic plasticity is modulated by the population decision that ultimately

leads to the reward signal (Urbanczik and Senn, 2009). The sign of the weight change should depend on whether the decision of the individual postsynaptic neuron D_{post} coincides with population decision D_{pop} formed by the majority of population neurons. These signals intrinsically depend on the neuronal code with which neurons and populations represent the possibly multivalued decisions and actions (Friedrich et al., 2014). In the simplest case of binary decisions, these signals may be set to 1 or -1, depending on whether the neuronal or population activity, low-pass filtered by τ_{Dec}, is above or below the corresponding decision threshold (Friedrich et al., 2011). The gradient-rule emerging from this reasoning reads as (cf. Fig. 4C)

$$E_{\text{post}}(t) \;=\; \int_{-\infty}^{t} \frac{\rho'}{\rho} \left(S_{\text{post}}(\tilde{t}) - \rho(u(\tilde{t})) \right) E_{\text{pre}}(\tilde{t})\, e^{-\frac{t-\tilde{t}}{\tau_{\text{Dec}}}}\, d\tilde{t} \tag{6}$$

$$E_{\text{Dec}}(t) \;=\; \int_{-\infty}^{t} D_{\text{post}}(\tilde{t})\, D_{\text{pop}}(\tilde{t})\, E_{\text{post}}(\tilde{t})\, e^{-\frac{t-\tilde{t}}{\tau_{\text{R}}}}\, d\tilde{t} \tag{7}$$

$$\Delta w(T) \;=\; \eta R\, E_{\text{Dec}}(T). \tag{8}$$

Intracellular recordings from dendrites during plasticity induction protocols has shown that SDTP also depends on dendritic NMDA-spikes (Gordon et al., 2006). This raises the question whether there are spike-timing dependent plasticity rules that take account of such dendritic spikes as well. There is in fact a class of gradient-based RL rules that incorporates the 'triple-spike timing' among the presynaptic, dendritic and postsynaptic spike sequence, including the dendritic and somatic voltage and the reward modulation, analogously to the 4-step cascade schematized in Figure 4C (Schiess et al., 2012).

Cross-References

Learning Rules: Overview

Spike Timing-Dependent Plasticity (STDP), Biophysical Models

Long Term Plasticity, Biophysical Models

Reinforcement Learning in Cortical Networks

Reward-Based Learning, Model-Based and Model-Free

Tempotron Learning

References

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

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

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

Bienenstock EL, Cooper LN, Munro PW (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience* 2:32–48.

Brea J, Senn W, Pfister JP (2013) Matching recall and storage in sequence learning with spiking neural networks. *J. Neurosci.* 33:9565–9575.

Buchs NJ, Senn W (2002) Spike-based synaptic plasticity and the emergence of direction selective simple cells: simulation results. *J Comput Neurosci* 13:167–186.

Clopath C, Büsing L, Vasilaki E, Gerstner W (2010) Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat Neurosci* 13:344–352.

Clopath C, Gerstner W (2010) Voltage and Spike Timing Interact in STDP - A Unified Model. *Front Synaptic Neurosci* 2:25.

Florian RV (2007) Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. *Neural Comput* 19:1468–1502.

Frémaux N, Sprekeler H, Gerstner W (2010) Functional requirements for reward-modulated spike-timing-dependent plasticity. *J Neurosci* 30:13326–13337.

Friedrich J, Urbanczik R, Senn W (2011) Spatio-temporal credit assignment in neuronal population learning. *PLoS Comput Biol* 7:e1002092.

Friedrich J, Urbanczik R, Senn W (2014) Code-specific learning rules improve action selection by populations of spiking neurons. *Int. J. of Neural Systems* 24:1–17.

Gerstner W, Kempter R, van Hemmen JL, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383:76–81.

Gjorgjieva J, Clopath C, Audet J, Pfister JP (2011) A triplet spike-timing-dependent plasticity model generalizes the Bienenstock-Cooper-Munro rule to higher-order spatiotemporal correlations. *Proc. Natl. Acad. Sci. U.S.A.* 108:19383–19388.

Gordon U, Polsky A, Schiller J (2006) Plasticity compartments in basal dendrites of neocortical pyramidal neurons. *J. Neurosci.* 26:12717–12726.

Graupner M, Brunel N (2012) Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proc. Natl. Acad. Sci. U.S.A.* 109:3991–3996.

Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex* 17:2443–2452.

Karmarkar UR, Buonomano DV (2002) A model of spike-timing dependent plasticity: one or two coincidence detectors? *J. Neurophysiol.* 88:507–513.

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

Legenstein R, Pecevski D, Maass W (2008) A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLoS Comput. Biol.* 4:e1000180.

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

Pfister J, Toyoizumi T, Barber D, Gerstner W (2006) Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. *Neural Comput* 18:1318–1348.

Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity. *Journal of Neuroscience* 26:9673–9682.

Ponulak F, Kasiński A (2010) Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting. *Neural Comput* 22:467–510.

Rubin JE, Gerkin RC, Bi GQ, Chow CC (2005) Calcium time course as a signal for spike-timing-dependent plasticity. *J. Neurophysiol.* 93:2600–2613.

Schiess M, Urbanczik R, Senn W (2012) Gradient estimation in dendritic reinforcement learning. *J Math Neurosci* 2:2.

Senn W (2002) Beyond spike timing: the role of nonlinear plasticity and unreliable synapses. *Biol Cybern* 87:344–355.

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

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

Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 99:10831–10836.

Sjöström J, Gerstner W (2010) Spike-timing dependent plasticity. *Scholarpedia* 5:1362.

Sjostrom PJ, Hausser M (2006) A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron* 51:227–238.

Sjostrom PJ, Rancz EA, Roth A, Hausser M (2008) Dendritic excitability and synaptic plasticity. *Physiol. Rev.* 88:769–840.

Sjostrom PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32:1149–1164.

Sjostrom PJ, Turrigiano GG, Nelson SB (2004) Endocannabinoid-dependent neocortical layer-5 LTD in the absence of postsynaptic spiking. *J. Neurophysiol.* 92:3338–3343.

Song S, Abbott LF (2001) Cortical development and remapping through spike timing-dependent plasticity. *Neuron* 32:339–350.

Toyoizumi T, Pfister JP, Aihara K, Gerstner W (2007) Optimality model of unsupervised spike-timing-dependent plasticity: synaptic memory and weight distribution. *Neural Comput* 19:639–671.

Urbanczik R, Senn W (2009) Reinforcement learning in populations of spiking neurons. *Nat Neurosci* 12:250–252.

Urbanczik R, Senn W (2014) Learning by the dendritic prediction of somatic spiking. *Neuron* 8:521–528.

Xie X, Seung HS (2004) Learning in neural networks by reinforcement of irregular spiking. *Phys Rev E Stat Nonlin Soft Matter Phys* 69:041909.