Spike-timing dependent plasticity for evolved robots.

Ezequiel A. Di Paolo
School of Cognitive and Computing Sciences
University of Sussex, Brighton BN1 9QH, U.K.
Email: ezequiel@sussex.ac.uk
Fax: 44-1273-671320

Abstract

Plastic spiking neural networks are synthesized for phototactic robots using evolutionary techniques. Synaptic plasticity asymmetrically depends on the precise relative timing between presynaptic and postsynaptic spikes at the millisecond range and on longer-term activity-dependent regulatory scaling. Comparative studies have been carried out for different kinds of plastic neural networks with low and high level of neural noise. In all cases, the evolved controllers are highly robust against internal synaptic decay and other perturbations. The importance of the precise timing of spikes is demonstrated by randomising the spike trains. In the low neural noise scenario weight values undergo rhythmic changes at the mesoscale due to bursting, but during periods of high activity they are finely regulated at the microscale by synchronous or entrained firing. Spike train randomisation results in loss of performance in this case. In contrast, in the high neural noise scenario, robots are robust to loss of information in the timing of the spike trains, demonstrating the counterintuitive results that plasticity which is dependent on precise spike timing can work even in its absence provided the behavioural strategies make use of robust longer-term invariants of sensorimotor interaction. A comparison with a rate-based model of synaptic plasticity shows that under similarly noisy conditions asymmetric spike-timing dependent plasticity achieves better performance by means of efficient reduction in weight variance over time. Performance also presents negative sensitivity to reduced levels of noise, showing that random firing has a functional value.

Keywords: Evolutionary robotics, spiking neural networks, spike-timing dependent plasticity, activity-dependent synaptic scaling, neural noise, robustness.

Short Title: Spike-timing dependent plasticity for evolved robots

1 Introduction

Synthetic approaches to the design of autonomous robots aim, amongst other things, at providing minimal integrated models of brain mechanisms in an embodied and situated platform. There is a vast distance between these models and actual brains and yet simple biologically-inspired controllers giving rise to adaptive, lifelike robot behaviour can turn out to be very valuable despite, or perhaps because of, their simplicity. These models may capture some interesting aspect of brain organisation or, in their functioning, they may reveal unusual or unexpected properties of known mechanisms. In contrast, contemporary work in computational neuroscience, it all its sophistication, often lacks such a whole-agent dimension. The properties of single neurons in specific subsystems are typically modelled and studied under idealised conditions (e.g., random inputs, uncorrelated or with known correlations, etc.) and it is hard to see how the sensorimotor loops might ever be closed, valuable though this work is. Recent studies in evolutionary robotics have aimed at harnessing the power of automatic evolutionary design to try to cross the gap between these two modes of research. So far, these studies have been mainly exploratory, drawing inspiration from neuroscience to enrich the building blocks used for evolutionary design – but the potential is there for feeding useful information back to neuroscience. On this issue see a recent review by Ruppin (2002). One example of this kind of research is the work by Husbands and colleagues using gaseous diffusion of neuromodulators as part of their evolved robot controllers (Husbands et al., 1998).

Recent work in evolutionary robotics has began to explore the use of spike-based neural controllers (Floreano & Mattiussi, 2001). Spiking neural networks possess a number of attractive features. They have comparatively greater computational power than similar networks of threshold sigmoidal gates (Maass, 1997). They can support a variety of functional specificity from rate-based codes to structured codes based on the timing of action potentials (Gerstner et al., 1997). They can perform novel kinds of computation such the recognition of temporal patterns using transient synchrony (Hopfield & Brody, 2001) and real-time computation without stable states in high-dimensional "liquids" of transient activity, (Maass et al., 2002). Their complexity makes evolutionary robotics an appropriate tool of design and exploration.

Here the exploration continues by addressing the evolvability and properties of plastic spiking neural networks where synaptic plasticity depends on the precise timing of spikes. Parameters regulating plasticity in light-seeking robots are evolved in simulation. It is found that not only such controllers are evolvable, but also that they produce a rich variety of behaviours and desirable properties such as sensorimotor and synaptic robustness. Two series of experiments have been carried out, one where neurons are modelled with low levels of noise and another with significant levels of neural noise. In both cases, the evolved controllers produce regular patterns of neural activity. However, despite the precise spike timing needed to activate the plastic rules, only controllers with low neural noise seem to rely on relative timing information. High neural noise controllers are generally quite robust to jitter and spike train randomisation suggesting counterintuitively that, under certain conditions, spike-timing dependent synaptic rules can work very well even spike timing is disrupted. However, these controllers retain some unique properties even in the presence of noise, as shown by a comparison with rate-based evolved plastic networks.

There are two kinds of motivations for this work. The first is, as suggested, exploratory. To the best of our knowledge this is the first attempt to control an integrated robot using spike-timing dependent plasticity in combination with an evolutionary strategy for design. It is an open question whether there are any benefits in the use of these mechanisms, either in terms of better adaptability or complexity of performance. These questions we will try to answer at least partially by the end of the paper. Such motivation runs in parallel with another which is less clearly realised in the current paper because of its preliminary nature; this is to inform computational neuroscience not with a model that is detailed in its microstructure but grossly simplified in terms of its relevance to whole behaving agents, but the other way around; i.e., a model simple in design where sensory stimuli correlate to motor activity through environmental coupling in situated robot.

2 Spike-timing dependent plasticity (STDP)

FIGURE 1 here

Experimental neuroscientific evidence suggests that the degree and direction of change in the strength of a synapse subject to repeated pairings of pre- and postsynaptic action potentials depends on their relative timing (Markram et al., 1997; Bi & Poo, 1998). See (Bi & Poo, 2001) for a review. Synaptic modification depends on whether the pre- and postsynaptic spikes are separated in time in less than a critical window of the order of a few tens of milliseconds. In most cases studied, if a presynaptic spike precedes the postsynaptic spike the synapse is potentiated, whereas the opposite relation leads to depression of the synapse. This results in a temporally asymmetric plasticity rule (figure 1) that deserves the name Hebbian because of its tendency to strengthen causal correlations between spikes. There is empirical evidence, however, for non-Hebbian plasticity of this kind (Abbott & Nelson, 2000; Bi & Poo, 2001). Many theoretical studies have concerned themselves with this rule of plasticity and its desirable properties such as a trend towards inherent stability in weight distribution and neural activity, unlike purely rate-based Hebbian rules which often require additional constrains (Kempter et al., 1999; Song et al., 2000; Rubin et al., 2001). One possible expression for this rule is:

$$\Delta w = \begin{cases} w_{max} A^{+} \exp(-s/\tau^{+}) & \text{if } s > 0 \\ -w_{max} A^{-} \exp(s/\tau^{-}) & \text{if } s < 0 \end{cases}$$

where $s = to - ti$ is the time difference between a postsynaptic and presynaptic spike and A^{+} and A^{-} are positive constants. Other filters may be used instead of the exponential decay, but this form is particularly suitable for implementation in an evolutionary robotics context as will be shown in the next section.

One of the key concerns when studying rules for synaptic plasticity is their regulatory properties. Hebbian learning on its own leads to runaway processes of potentiation and cannot account for the stability of neural function. Additional elements, such as the directional damping of synaptic change (Rubin et al., 2001) or longer-term stabilizing regulation based on postsynaptic activity (Horn et al., 1998; Turrigiano, 1999) may come to the rescue. These can lead to unsaturated distributions of synaptic strengths in the first case and to regulated neuronal firing in the second and will also be investigated in this work.

Although STDP is a topic that has drawn much attention recently, most theoretical studies have concerned themselves with the properties of the temporally asymmetric plastic rule. There are, however, a few hypotheses about its functional role. For instance, Abbott and Blum (1996), show in a general model how firing patterns in a neural array (such as a receptive field), where neurons fire preferably at certain input values in a sequence of inputs, can by means of temporal asymmetry in plasticity lead to prediction of the inputs in a sequence through repeated presentation. This is because the synapses of neurons that fire in succession are strengthened from those that fire first to those that fire later (and are depressed in the opposite direction). Empirical evidence in the experience-dependent change in skewness in place fields in the rat hippocampus supports the findings of this model, (Mehta et al., 1997, 2000). Related to this, Yao and Dan (2001) have found that repetitive pairing of visual stimuli at two different orientations induced a shift in orientation tuning in cat cortical neurons depending on the relative timing of presentation and compatible with STDP.

Other related functional implications have also been suggested. Rao and Sejnowski (2001) suggest that STDP could be involved in implementing some form of temporal difference learning (Sutton, 1988) and show this in a model of input spike prediction; and Chechik (2002) has recently compared theoretical rules of plasticity derived from the principle of information maximisation of relevant input with empirical rules to conclude that temporal asymmetry can increase input information to near optimal levels.

This kind of functionality is hard to compare with the results obtainable from the present work on a simple robotic task, as it is more likely to play a significant roles when sensory surfaces or arrays are included in the robot model as well as something equivalent to receptive fields. Because of the constraints put by the number of evolutionary evaluations, such elements are not included in this initial study but will be of central importance in the future.

3 Methods

3.1 Robots and task

Since we are interested in exploring a novel mechanism for robot control, the chosen task is at this stage deliberately simple so as to facilitate comparisons with alternative approaches. Simulated robots are evolved to perform phototaxis on a series of light sources. Robots have circular bodies of radius $R_0 = 4$ with two motors and two light sensors. The angle between sensors is of 120 degrees but a small random displacement between -5 to 5 degrees is added at the start of each evaluation. Motors can drive the robot backwards and forwards in a 2-D unlimited arena.

The neural network consists of six nodes, fully-connected except for self-connections. Neurons can be either excitatory or inhibitory and this is set genetically. Trials with larger number of neurons have been carried out successfully, but not systematically studied.

The whole system is simulated using an Euler integration method with a time step of 1 ms (25% of the minimum timescale). Robots are run for two independent evaluations, each consisting on the sequential presentation of two distant light sources. Only one source is presented at a time for a relatively long period T_S chosen randomly for each source from the interval [7.5 s, 12.5 s], (each evaluation consists therefore of an average of 2×10^4 update cycles). The initial distance between robot and new source is randomly chosen from [60, 80], the angle from $[0, 2\pi)$ and the source intensity from [3000, 5000]. The intensity decays in inverse proportion to the square of the distance to the source.

The simulated robots use photoreceptors that are activated by the light intensity corresponding to their current position if the light source is directly visible (i.e., an angle of acceptance of 180 degrees). This intensity is multiplied by the sensor gain equal for both sensors (genetically set from range [0.1, 50]) and clipped for values beyond of maximum of 20. A spike train is generated using a Poisson process with variable rate (maximum 200 Hz) by linearly transforming the sensor value into the instantaneous firing frequency. The Poisson spike trains coming from the left and right sensors are fed into neurons n2 and n3 respectively Additionally, uniform noise is present in the sensors (and motors) with range 0.2 (before scaling by gains) – this results in spikes that fire randomly with very low probability when the sensor is not stimulated.

Two motors control the robot wheels. Each motor is controlled by two neurons, one that drives it forwards and the other one backwards, using a spike-based leaky integrator. The left motor is controlled by neurons n0 (forward) and n4 (backward), and the right by neurons n1 (forward) and n5 (backward).

A population of 30 robots is evolved using a generational GA with truncation selection. In the plastic scenarios described below initial weights are randomly chosen at the start of each evaluation from the interval $[0, w_{max}]$ while the parameters for the plasticity windows and scaling constants are evolved. In the non-plastic scenario, synaptic strengths are encoded genetically. Other genetically set parameters include sensor and motor gains, motor decay constant and whether neurons are inhibitory or excitatory. All parameters are encoded in a real-valued genotype, each gene assuming a value within [0, 1]; each parameter is linearly scaled to the corresponding range of values, except for sensor and motor gains which are scaled exponentially. Only vector mutation (Beer, 1996) is used with a standard deviation of vector displacement of 0.5 (maximum genotype length is 220), genetic boundaries are reflective. Fitness is calculated according to:

$$F = \frac{(1 - M^2)}{T_S} \int f dt; \quad f = 1 - \frac{d}{D_i}$$

if the current distance to the source d is less than the initial distance D_i, otherwise $f = 0$. M measures the average difference in activity between the motors divided by the motor gain.

$$M = \frac{0.125}{T_S} \int \frac{(M_L - M_R)}{M_G} dt.$$

Near optimal fitness will be obtained by robots approaching a source of light rapidly and with minimal integrated angular movement.

3.2 Neural controller

An integrate-and-fire model with reversal is used for the neural controller. The time evolution of the membrane potential V of a neuron is given by:

$$\tau_m \frac{dV}{dt} = V_{rest} - V + g_{ex}(t)(E_{ex} - V) + g_{in}(t)(E_{in} - V)$$

where τ_m is the membrane time constant (range [10 ms, 40 ms]), v_{rest} = -70 mV is the rest potential, the excitatory and inhibitory reversal potentials are respectively E_{ex} = 0 mV and E_{in} = -70 mV.

A noisy threshold value, V_{thres}, is given by a normal distribution with a genetically set mean value (range [-65 mV, -50 mV]) and a deviation of 1 mV. When the membrane potential reaches this threshold, an action potential is fired and V is reset to V_{rest}. An absolute refractory time of 4 ms prevents the neuron from firing another spike within this period.

Every time a spike arrives to neuron j from an excitatory presynaptic neuron i the excitatory conductance of j is increased by the current value of the synaptic strength ($w_{ij}(t)$): $g_{ex}(t) \rightarrow g_{ex}(t) + w_{ij}(t)$. The inhibitory conductance g_{in} is similarly affected by spikes coming from inhibitory neurons. Conductances otherwise decay exponentially:

$$\tau_{ex} \frac{dg_{ex}}{dt} = -g_{ex}; \quad \tau_{in} \frac{dg_{in}}{dt} = -g_{in}$$

with τ_{ex} and τ_{in} genetically set for each neuron from the range [4 ms, 8 ms].

The current motor value is stored in a variable $M_{L,R}$ which is directly translated into the left and right velocities respectively.

$$\tau_{mot} \frac{dM_{L,R}}{dt} = -M_{L,R} + M_G(\sum \delta(t - t^{(f)}_{forw}) - \delta(t - t^{(f)}_{back}))$$

with τ_{mot} genetically set from the range [40 ms, 100 ms] and M_G from [0.1, 50]. Both motors have the same value for their gains and decay constants. This approach marks a difference from previous work on the evolution of spiking controllers which have used a neural rate estimation method for driving the motor (Floreano & Mattiussi, 2001).

STDP. The properties of plastic windows (figure 1) are evolved for each synapse in the neural network controller. Following (Song et al., 2000), synaptic change is implemented using two recording functions per synapse $P^-(t)$ and $P^+(t)$. Every time a spike arrives at the synapse the corresponding $P^+(t)$ is incremented by A^+, and every time the postsynaptic neuron fires the corresponding $P^-(t)$ is decremented by A^-. Otherwise, these functions decay exponentially with time constants τ^- and τ^+ respectively. $P^-(t)$ is used to decrease the synaptic strength every time the presynaptic neuron fires: $w_{ij}(t) \rightarrow w_{ij}(t) + w_{max}P^-(t)$. Analogously, $P^+(t)$ is used to increase the synaptic strength every time the postsynaptic neuron fires: $w_{ij}(t) \rightarrow w_{ij}(t) + w_{max}P^+(t)$. In all the cases studied in the first series of experiments the maximum synaptic strength is $w_{max} = 1$. This method facilitates the computational implementation of STDP by eliminating the need of keeping track of spike trains or calculating other response functions which could be more costly.

The values for A^+ and A^-, and τ^+ and τ^- are genetically set per synapse from the ranges [0.0001, 0.05] and [10 ms, 40 ms] respectively. In all the experiments reported here the plastic windows are Hebbian, that is, spikes arriving before a postsynaptic action potential always potentiate a synapse and those arriving after always depress it. Experiments relaxing this constraint, i.e., allowing anti-Hebbian or purely potentiating or depressing windows, have also been carried out successfully, but are not reported here.

Activity-dependent synaptic scaling (ADS). Some of the mechanisms used by neurons to regulate their firing rate homeostatically are thought to affect all incoming synapses scaling them up or down independently of the presynaptic activity, (Turrigiano, 1999). If the postsynaptic activity is above a certain target, excitatory synapses are scaled down, otherwise, they are scaled up, thus preventing sustained levels of activity that are too high or too low. Following (van Rossum et al., 2000) excitatory synapses are modified according to:

$$\tau_{ads}\frac{dw_{ij}}{dt} = w_{ij}(z_{goal} - z_j)$$

where $z_{goal} = 50$ Hz and τ_{ads} is genetically set from the range [1 s, 10 s]. The firing rate z_j of a neuron is estimated by a leaky integration of the spike train:

$$\tau_z\frac{dz_j}{dt} = -z_j + \sum \delta(t - t^{(f)})$$

where $t^{(f)}$ are the times when the neuron emits a spike (the sum runs over all previous spikes) and $\tau_z = 100$ ms.

In real neurons, this is a mechanism that acts over long timescales (over hundreds to thousands of seconds) (Turrigiano et al., 1998), but due to computational limitations (the very long evaluation runs that would be required) the chosen timescale ($\sim \tau_{ads}$) is faster than this but still significantly slower than the rest of the timescales in the system. Even though the above mechanism acts on excitatory synapses, in the current context it has also been applied when the presynaptic neuron is inhibitory by multiplying the right hand side above by -1. A similar homeostatic mechanism has been successfully implemented in robots capable of adapting to sensorimotor disruptions not previously experienced (Di Paolo, 2000).

Directional damping. Synaptic weights are constrained within the range [0, 1]. This can be achieved simply by a stop condition at the boundaries or by means of damping factors that vanish as the weight value approaches a boundary. The choice can have important consequences. No damping leads to a bimodal distribution of weights under random stimulation (Song et al., 2000), where most weights adopt the minimum or maximum values in the range, but few values in-between. The same happens with purely positional damping, i.e., factors that slow down weight change near the boundaries, but depending only on the current weight value. A biologically plausible alternative is directional damping whereby if a weight value is near a boundary, changes that push this value towards the boundary are slowed down, but changes that push it away from the boundary are not. The equilibrium weight distribution in this case tends to be unimodal and centred around the point where potentiation and depression equilibrate, (Rubin et al., 2001). Directional damping is supported empirically by the observation that spike-driven potentiation is more pronounced than the expected linear variation at synapses of relatively low initial strength in cultured hippocampal cells (Bi & Poo, 1998). It was also observed that the mean fractional negative change was constant over a wide range of initial weights, corresponding to the a damping linear factor for absolute depression equal to the current weight value.

Linear directional, or multiplicative, damping is simply implemented by transforming a weight change (as resulting from STDP or ADS or both): $\Delta w_{ij} \rightarrow (1 - w_{ij})\Delta w_{ij}$ if $\Delta w_{ij} > 0$ and $\Delta w_{ij} \rightarrow w_{ij}\Delta w_{ij}$ if $\Delta w_{ij} < 0$ for $w_{ij} \in [0, 1]$.

Neural noise. Different sources of neural noised have been modelled. At any given time, Gaussian noise with zero mean and 1 mV deviation is applied to the value of the firing threshold. This is the only source of neural noise in the first set of experiments. Additionally, for the second set, the refractory period is randomly set every time step using a uniform distribution ([2 ms, 4 ms] for cases of short refractory period, [4 ms, 8 ms] for long refractory period). Background noise is modelled as an incoming Poisson train to every neuron with a frequency of 10 Hz, and spontaneous firing has also been modelled using a baseline 10 Hz Poisson process for each neuron, but subject to refraction.

Synaptic decay. In order to test the robustness of the evolved controllers to perturbations in their internal configuration, synaptic weights are allowed to decay exponentially to 0 with a time constant that can be as fast as 100 ms. Synaptic decay is not affected by directional damping and is not applied during evolution but only during behavioural tests.

Poisson filters and randomised delays. In order to test the reliability of the evolved controllers on the precise timing of spikes, the simple expedient of filtering the output of a neuron with a Poisson process emitting random spikes at the same instantaneous rate has been used. Information about firing rate is conserved, but precise spike-timing is disrupted. Because, the rate z is estimated using only previous spikes, it is only possible to approximate the instantaneous firing rate of the neuron in this manner. It is expected, however, that if controllers rely heavily on firing rates, the disruption in performance should not be too strong. A more sophisticated method consist in introducing artificial random delays in the firing time of single or multiple neurons. This is done by keeping a short sub-train corresponding to the last T ms of activity and swapping the current fire state of a neuron with a randomly selected state in the sub-train thus conserving the number of spikes. This is similar in objective to tests *in vivo* on honeybee odour discrimination demonstrating the role of synchronised neural assemblies (Stopfer et al., 1997). These tests are applied only after evolution.

CTRNN. Control runs using rate-based continuous-time, recurrent neural networks (Beer, 1990) have been performed. These are defined by:

$$\tau_i \frac{dV_i}{dt} = -V_i + \sum_j w_{ji} z_j + I_i; \quad z_j = \frac{1}{1 + \exp[-(V_j + b_j)]},$$

where, V_i represents the membrane potential of neuron i, τ_i the decay constant (range [0.4 s, 4 s]), b_i the bias (range [-3, 3]), z_i the firing rate, w_{ij} the strength of synaptic connection from node i to node j (range [-8, 8]), and I_i the degree of sensory perturbation for sensory nodes. A plastic version of this controllers has also been used and is described in more detail later.

4 Results: Low neural noise

FIGURE 2 here

Robots using spiking controllers were successfully evolved for four different scenarios: 1. No plasticity (evolution of fixed weights), 2. STDP No Damping (evolution of STDP windows without directional damping, random initial weights), 3. STDP (evolution of STDP windows with directional damping, random initial weights), and 4. STDP+ADS (evolution of STDP windows and ADS with directional damping, random initial weights).

Evolvability. Five 100-generation independent runs were made for each of the four scenarios above. It was qualitatively found that, contrary to expectations, the more complex case (in term of the dimensions of the search space and the additional features of the mechanism), that is STDP+ADS, was easier to evolve, particularly during the initial generations, than the simpler case of no plasticity. This is observed in figure 2 where the mean population fitness, averaged over the five runs, is plotted for these cases (error bars indicate standard deviation). There is little observable distinction between using STDP with or without damping (not shown) but there is a significant difference between STDP+ADS and no plasticity. However, the long term fitness of the best individual of the last generation is not significantly different between the cases (figure 3a). These were obtained by running for each case the best individual of 5 runs for 10 independent evaluations. For comparison purposes, a CTRNN controller has been evolved and tested under the same conditions and is also shown in the figure.

FIGURE 3 here

Synaptic decay. Plastic controllers maintain their functionality dynamically as a consequence of their own activity. In order to assess their reliability a disruptive decay of synapses was introduced as described above. Figure 3b shows results for the three plastic set-ups (again using 5 independent runs for each and testing the best individual 10 times). It is apparent that STPD+ADS controllers are able to perform quite reliably even for decay times of up to 250 ms, while controllers using STDP with or without damping are unable to maintain their performance. It is possible to explain this as a consequence of the compensatory nature of the ADS mechanism, which is able to alter synapses as a consequence of longer term changes in neural activity in ways that tend to maintain this activity and therefore the functionality of the controller.

Behavioural strategies. Evolved robots show a rich variety of behavioural strategies and not all can be discussed here. Practically all of the observed strategies are active, involving scanning behaviour, which is not surprising given that sensors saturate rather easily. Figure 4a shows a trajectory for an STDP+ADS robot. Unlike what is commonly observed using rate-based neural controllers, robots are often able to come to a full stop near a source of light (but not facing it) – this can be observed in the inset showing the distance to the light source. During these periods there is very little neural activity, until a random spike from the sensors triggers a round of activity and the robot starts scanning again (see arrow in figure 4b which shows the network activity for the same controller). In the absence of light robots scan their surroundings using this mechanism.

FIGURE 4 here

Timing dependence. Applying a Poisson filter test to all the neurons in a controller, whereby rate information is conserved but not spike timing, results in total loss of fitness ($< 1\%$). This was tested in different controllers in the four scenarios with the same result. A more detailed study involves the application of random delays to single neurons. As explained above, this method conserves the number of spikes but randomises the timing within a sub-train of a given size corresponding to the last T ms in the simulation. Figure 5 shows how fitness is affected for a best individual in one STDP+ADS run when random delays are applied to all the neurons and to single neurons (similar results were obtained for the other classes). There is a sharp reduction in fitness in the first case for the randomisation of sub-trains as short as 2 ms, but applying the test to single neurons shows that the controller is crucially dependent on the precise timing of only 3 of the six neurons (corresponding in this case to one motor and the two sensor neurons, but to different neurons in other cases).

FIGURE 5 here

Synaptic dynamics and internal regulation. Weight values also exhibit rich dynamical behaviour. Figure 6a shows three of the synaptic weights affecting neuron 0 for the same run corresponding to figure 4 (all other weights behave similarly). The other neurons show similar qualitative dynamics at this timescale, consisting of rhythmic periods of activity (during which the robot rotates left and right) punctuated by silent periods (during which the robot gently comes to a full stop facing away from the light). While neurons are firing, STDP drives the weights always to a same area within the range, the rest of the time ADS takes control and increases the synaptic efficacies so as to compensate for the lack of firing activity in the neuron. This has the effect of hyper-sensitizing the whole network, so that even a single spike arriving from the sensors is often capable of triggering a new round of activity and the robot starts moving again.

FIGURE 6 here

The dynamics of the microstructure are also interesting. Figure 6b shows how a particular synaptic strength is regulated during a period of activity lasting about 0.6 s. Its value is maintained nearly constant. The insets show the pattern of firing of the pre- and postsynaptic neurons which undergoes a process of synchronisation that persist while at least one sensor is active and finally de-synchronises and stops. The plasticity window shown in the figure indicates that synchronous firing translates into net potentiation which is compensated by ADS resulting in an equilibrium (just after activity stops, ADS depresses the weight due to the inertia of the frequency estimation before re-sensitizing the network). This regulatory pattern has been observed in all of the cases studied for the STDP+ADS class.

The neurons that drive the left motor (n0 and n4) fire at a same frequency, yet it is possible to observe that their phase relation changes non-randomly so as to produce a peak in motor output roughly towards the middle of the activity round. The right motor does something similar just after the left motor peaks, although its driving neurons (n1 and n5) are not clearly entrained, (data not shown).

5 Results: Noisy neurons

A second series of experiments using noisy neurons were run for the cases of no plasticity, STDP (with damping) and STDP+ADS (five 400-generation independent runs each). In addition to small threshold noise, each neuron included a noisy refractory period and either 10 Hz Poisson baseline firing or 10 Hz Poisson extra input to each neuron.

One undesirable feature of the first set of experiments are the high frequencies of firing utilised, reaching up to 200 Hz. Although this is unrealistically high, it was not expected that with a very small network all the evolved properties would be equally plausible biologically. There may be, however, a problem in that the low-noise networks often seem to fire at the maximum possible frequency, i.e. a typical interspike interval equal to the absolute refractory period. This feature may indeed be undesirable, so the following modifications were made. In the case of ADS, z_{goal} was lowered from 50 Hz to 40 Hz. The range for firing threshold was reduced from [-65 mV, -50 mV] to [-60 mV, -50 mV], the maximum frequency of the sensor input trains was reduced from 200 Hz to 100 Hz, and the range of synaptic strengths was reduced from [0, 1] to [0, 0.5] (synaptic scaling was modified accordingly). The inhibitory reversal potential E_{in} was changed from -70 mV to -80 mV. The average refractory period (now uniformly distributed) was of 6 ms (instead of 4 ms) although successful trials were also performed for shorter average refractory periods. The average duration of single light source presentation was changed from 10 s to 20 s and the range of possible motor gains from [0.1, 50] to [1, 20]. These modifications succeeded in producing networks with more plausible frequencies (range 10 Hz to 80 Hz).

Robustness. Again test were carried out for robustness against synaptic decay. Figure 7 shows that these results are not significantly altered by the addition of neural noise. Controllers can also cope with sensorimotor disturbances such as asymmetric modification of sensor and motor gains (not shown). Networks seem to be a bit less robust against synaptic decay, but still STDP+ADS controllers cope better than STDP only controllers (figure 7b).

Neural noise. Figure 8 shows the response of three independently evolved individuals for each set to variations in the amount of spontaneous background firing. In all these cases, controllers were evolved with a Poisson background firing for each neurons of 10 Hz (corresponding to the vertical line). Fitness values have been obtained for 20 runs per point and normalised for 10 Hz. The response to increased level of random firing is, as expected, a decreased level of performance. Some controllers are more robust than others but the trend in clear in all of them. Interestingly, decreasing the level of neural noise also results in worse controllers (again with different degrees of sensitivity). This implies that random firing is being used functionally. Similar results were observed for controllers evolved with a 10 Hz Poisson input.

FIGURE 8 here

Spike timing. The introduction of different sources of neural noise disrupts the very regular firing patterns obtained in the first series of experiments. Still, it is possible to observe in some cases that during periods of activity, the noisy controllers still seem to fire with some regularity. In order to observe this more clearly some paired patterns and their covariograms (cross-correlograms corrected for time-dependent shifts in frequency) are shown in figure 9 for two STDP+ADS controllers, one of them exhibiting bursting behaviour. Covariograms are roughly proportional to the probability of the two neurons firing after the corresponding shift in time. Around 40 seconds of data are used in their calculation (see Appendix). In both cases it is possible to observe distinct peaks and valleys in the covariation of the two spike trains for some values of the time-shift (x-axis). The peaks correspond to very short shifts (a few ms) in time even though there are no peaks corresponding to zero shift (synchrony). The smooth lines give an idea of the interval of significance ($\pm\sigma_N$) given the variance of each spike train (see Appendix). Some of the peaks and valleys clearly cross this interval implying that the timing regularity in the patterns is significant. Still, the question remains whether such regularity has a functional value or whether it is epiphenomenal.

FIGURE 9 here

Disruption of spike timing. To answer the last question we repeat the experiments on disruption of spike trains done in the first series. On application of a Poisson filter performance in the first series dropped to practically 0%. Figure 10a shows the effect of applying the same filters to noisy controllers for each case (5 independent runs, 20 evaluations, each). It is clear that even though there is a reduction in performance, the effect is not nearly as dramatic as before. The case of STDP+ADS is even quite robust. And on observation, it was found that the disrupted robots did perform phototaxis, only less efficiently.

FIGURE 10 here

These results are supported by the randomisation of spike trains (figure 10b). In sharp contrast to figure 5 there is a considerable degree of robustness for disruption corresponding to randomising the sub-trains of up to 10 ms, and then a slow decay for longer sub-trains. The average result for STDP+ADS means that up to 60 ms of spike timing information can be shuffled without significant loss of fitness. This indicates that noisy controllers are very unlikely to be using any information contained in the precise timing of spikes. Covariograms and sample trains for the same controllers and neurons shown in figure 9 can be seen in figure 11 corresponding to spike train randomisation with a sub-train size of 40 ms. No noticeable peak stands out of the estimated interval of significance.

FIGURE 11 here

Analysis of one strategy. A series of different behavioural strategies have been observed. Most of them involved periods of high neuronal activity in the whole network, punctuated by periods of low activity and random background firing. Figure 12 shows one strategy for an STDP bursting controller. The robot performs an unusual approach to the light source by travelling "backwards" and *avoiding* sensor activation by occluding the light with its own body. As long as the occlusion persists and the source is within sensor range, the robot will travel in the general direction of the source. This situation corresponds to periods of low and random firing and is marked by the thin segments in the trajectory. Eventually, one of the sensors comes into contact with the light – no matter which sensor it is, effectively the same round of neural activity is triggered by this event. This causes the robot to start moving at a higher speed in an arc that lasts until light is occluded once again (thick segments in the trajectory). The strategy works by taking advantage of an invariant geometrical relation: if the robot travels backwards (or alternatively if the sensors are at the back of an occluding volume) then, as long as the periods of sensor activation are *minimised*, the robot will eventually reach the source of light. The robot approaches the light by keeping its sensors oriented within the cone of shadow caused by its own body.

FIGURE 12 here

All neurons in this particular case are excitatory (although typically controllers are composed of a combination of neural types) and similar in their properties except neuron n0, corresponding to the left motor forward control, which has a longer membrane decay constant τ_m. As a result when the all the neurons are firing, this neuron is firing at a significantly lower rate resulting in faster speed on the left side in the backward direction and thus the robot is able to describe the arc that will eventually produce a shadow on both sensors.

The explanation is supported by a simple experiment. If the sensors are positioned diametrically opposed and at 90 degrees to the axis of the motors (so that one sensor will always be active), the robot fails to perform (less than 0.1 % of the original fitness), however, if one removes the sensor pointing in the direction of movement and leaves only the other sensor at equal angles between the motors the robot is still able to perform phototaxis, although with lower fitness (around 50 %) due to the longer time it takes to reach the position of the source.

The strategy also explains why performance degrades when the level of neural noise is reduced (figure 8). The robot relies for its approach on the low level of noisy activity that keeps it travelling in the direction of the source while the sensors remain inactive. Given the self-correcting nature of the trajectory, it does not matter whether such random activity causes the robot to deviate a little, as this will only trigger neural activation, and a corrective arc trajectory that will last until sensor readings drop once again, eventually resulting in a course regulation. It also seems plausible now that even the application of Poisson filters and jitter in the spike trains will not have a very strong effect in the performance, as shown above (figure 10).

It must be stressed that this is one of the many strategies found in the second series of experiments. Other strategies resemble more traditional forms of phototaxis with an active oscillatory (but not cycloidal) element particularly in the case of STDP+ADS where rhythmic bursting is often observed.

Synaptic dynamics. Figure 13a shows the weight dynamics for one run in an STDP controller. Except for a few slow-changing weights, most of the synapses settle very rapidly into a stable value. The distribution of values covers the whole range. Figure 13b shows for the same run a detail of the change in one particular weight together with the pre and postsynaptic trains. The synaptic strength is quite stable. Both neurons fire with a very similar frequency and there's almost a one-to-one correspondence in the number of spikes.

As in the case with low neural noise, endogenous bursting is very common for STDP+ADS controllers. The corresponding weight dynamics are similar as well. This is shown in figure 14a where the bursting of one neuron and one corresponding incoming synaptic strength are shown. Notice that sensor activation (not shown to scale) does not drive the bursts of activity, but that these are endogenously generated by the oscillatory dynamics created by the ADS balancing. Interestingly, the pattern is altered when the sensor is up and this produces the overall effect of a retardation in the phase of oscillation. Figure 14b shows a detail with both the pre- and postsynaptic spike trains for the same synapse.

FIGURE 13 and 14 here

6 Comparison with rate-based synaptic plasticity

It is not surprising that if spike-timing is perturbed by the presence of noise, then the controller will rely more heavily on other network properties or sensorimotor invariants. However, the rules that determine weight change remain temporally asymmetric and so dependent to some extent on spike-timing and this is the paradoxical result. In order to investigate if there is any additional feature granted by the use of STDP in the noisy case, a comparison is made with a rate-based model of synaptic plasticity. A CTRNN controller is used in which synapses undergo weight change depending on the pre- and postsynaptic firing rates.

Designing a fair comparison is not straightforward. Simple Hebbian rules of the form $dw_{ij}/dt = \eta_{ij} z_i z_j$ have been unable by themselves to produce a single well-performing controller over a dozen attempts. However, extended rules or combination of them are known to do so with some ease (Floreano & Urzelai, 2000). A general expression fo such rules is:

$$\frac{dw_{ij}}{dt} = \eta_{ij}(A_0 + A_1 z_i + A_2 z_j + A_3 z_i z_j),$$

where all parameters (A_k, η_{ij}) are subject to evolutionary change. Even when weights are initialised randomly, if firing rates (z_i) have an initial short-range random distribution around their mean value (0.5), the above parameters can determine in great part the initial direction of weight change, thus facilitating the task of evolution by providing an innate bias for the development of the neural network configuration. This is not what happens with STDP controllers as neurons are randomly initialised as firing or non-firing at $t = 0$. In other words, all the cases of STDP controllers studied here start with random weight initial values *and* random initial weight derivatives.

Since firing rates in the CTRNN are initialised around the middle of the range the above condition can be achieved by modifying the plastic rule to:

$$\frac{dw_{ij}}{dt} = \eta_{ij}(A_1(z_i - 0.5) + A_2(z_j - 0.5) + A_3(z_i - 0.5)(z_j - 0.5)),$$

The initial direction of weight change in this case is also random, and evolution must be able to build a bootstrapping process whereby neural properties and environmental interaction play a stronger role in the shaping of the controller. These are the conditions under which STDP controllers are evolved. Damping factors are also applied.

Maximum weight derivatives are made of the same order (half the weight range can be covered in one second of simulated activity) and in some trials even faster than in the STDP case. This condition results in $A_k \in [-1, 1]$ and $\eta_{ij} \in [-3, 3]$. Noise is also simulated in the rate-based model by perturbing neuron rates with the addition of a uniformly distributed random variable of range 0.1. Poisson input noise was simulated by modifying the input currents: $I_i \rightarrow I_i + \sqrt{I_i}\rho$ with ρ a normally distributed random variable with unit stardard deviation and zero mean (Tuckwell, 1988). A non-noisy scenario was also study where these conditions are not applied; the differences were not significant.

It was found that rate-based controllers evolved under the same conditions as the STDP networks significantly underperformed in the phototaxis task. Figure 15 shows the average fitness of 5 independent runs both for noisy and non-noisy scenarios compared with STDP controllers with neural noise both under normal conditions and with spike-train randomisation (20 ms).

FIGURE 15 and 16 here

Inspection of the synaptic dynamics shows that the network is too slow to settle into a stable condition. This is shown in figure 16 where a comparison is made for two weights representative of the fastest and slowest convergence, both for one rate-based and one STDP noisy controller during 10 independent runs.

STDP controllers are able to rapidly define a direction of weight change depending on the relation between the plastic rules and the neural properties. The randomness in weight derivative lasts only of a few tens of milliseconds (it cannot be appreciated in the figure) whereas rate-based plastic controllers take much longer to settle into a given range (if they settle at all). Figure 17a shows the average reduction in variance across trials for all weights in the two cases corresponding to figure 16. Even though this variance takes into account all the weights (and consequently may overestimate the variance of those with higher functional significance) it is clear that the difference, both in transient and longer term behaviour, is significant.

It is also interesting to compare the weight dynamics for an STDP controller with neural noise both when functioning normally and when a Poisson filter is applied (total loss of timing information). A squared-difference matrix $Dw_{ij}(t) = (w_{ij}(t) - wp_{ij}(t))^2$ is recorded and averaged over 10 trials ($wp_{ij}(t)$ correspond to the weights for a Poisson-filtered run). Figure 17b shows this quantity for all the weights and for the average across all weights (thick line). It is clear that the long-term behaviour of the weight matrix does not depend strongly on timing information as weights tend to approach the same value as the Poisson-filtered network. This is interesting as it suggests that, under noisy conditions, evolution is able to find networks which rely more strongly on the timing properties of the neurons and plastic rules to determine long-term weight distribution.

FIGURE 17 here

The performance of the rate-based plastic controllers improves to levels comparable with STDP controllers with the presentation of more light sources per evaluation (higher that 5 instead of the 2 presented in the STDP case) and if initial weight derivatives are permitted to be set genetically.

7 Discussion

Despite the exploratory nature of this work, it is possible to assert that the richness of behaviour in plastic spiking neural networks, not often found in other controllers, makes them extremely interesting candidates for further testing in adaptive behaviour research. It is perhaps inevitable that a synthetic design method should be used to approach their complexity in integrated agents with closed sensorimotor loops. This paper has shown that it can be done successfully for a simple task. The suitability of these mechanisms in other scenarios, and ultimately their potential for more complex cognitive performance, remains to be seen. In particular, it will be of much interest to explore more closely the functionality proposed for STDP (section 2) by using a more appropriate neural architecture. One of the main disadvantages of the approach is obviously the extra computational cost involved in the longer evaluations (roughly 200 times the equivalent of CTRNN controllers). But this cost has passed from being prohibitive a few years ago to being acceptable nowadays if the benefits justify it.

The first series of evolved controllers demonstrate the use of some uncommon mechanisms (in a robotics context) such as the precise timing information of spike trains. Neural networks undergo rhythmic periods of activity during which pairs of neurons start un-correlated, then they reach a highly entrained state, and finally they lose their entrainment and become inactive. Such periods can be triggered by a sensory stimulus or if the inactivity has lasted long enough, even by single random spikes coming from the sensors, owing to the excitability that builds up thanks to the ADS mechanism which acts as an adaptive balancer of synaptic input. During these periods synaptic strengths are kept nearly constant and despite firing in strongly entrained mode motor neurons maintain enough variety to coordinate their relative timing and achieve functionally useful movement. Single neuron randomisation of spike trains has revealed that not all neurons are crucial in allowing the network to make use of timing information.

The second series introduced a more plausible scenario with various sources of neural noise. In sharp contrast to the first series, robot performance for noisy controllers degraded little or nothing at all on application of Poisson filters or randomisation of spike trains, indicating that despite the nature of the plastic rules that drive the controllers in the STDP and STDP+ADS cases, and despite some evidence of regularity in spike times, these controllers do not need to make use of precise timing information in order to work properly. The dependence on precise timing in the low noise series would indeed seem to fit with the nature of the plastic rules. However, as most direct evidence for STDP originates from cell culture studies (Bi & Poo, 2001), it is still an open question how these mechanisms will operate on a behaving animal – particularly if they underly learning processes that need to operate at a very different timescale, (Mehta et al., 2002). In the second series, the same plastic rules are at work and yet precise timing seems not to be essential, even though it is naturally present to some degree in the unperturbed robot. In a sense, regularity in spike timing in this case is epiphenomenal and STDP can function *without* it.

It is a well-known principle of evolutionary design that reliable aspects of the performance evaluation may be taken advantage of by evolution, and that making those aspects unreliable produces solutions that are robust to their variation. This principle guides the minimal simulation design strategy (Jakobi, 1997). In the current context, the presence of neural noise may have made it easier for evolution to find solutions that rely on neural firing rates while ignoring noisy spike timing. This is supported by the evidence in figure 17b where the weights of STDP controllers both with and without Poisson filters converge to a same distribution.

Although non-plastic rate-based controllers will work optimally for this task (Beer & Gallagher, 1992), evolving purely plastic rate-based controllers under the same conditions results in poorer performance than STDP controllers due to lack of efficiency in reaching a stable weight distribution. Temporally asymmetric plasticity is much more stable and efficient, even in the presence of noise, and able to 'develop' a controller faster (which is an implicit fitness requirement in this case).

It is hard to draw general implications for more complex scenarios from the non-reliance of noisy controllers on spike timing. What this exploration does show is that it is *possible* for STDP rules to work in an integrated sensorimotor system even in the absence of precise timing information (which seems paradoxical) while still retaining advantanges over rate-based models evolved under the same conditions. The important question is in what situations, if any, analogous robustness may be found in natural systems. To answer this, behavioural studies are needed, but for the moment autonomous robotics modelling may provide some initial insights. The next step will be to approach more complex tasks where the evolutionary approach has already proved successful; task requiring different levels of memory, orientation to stimuli, and selective attention such as visual shape discrimination (Harvey et al., 1994), delayed response (Jakobi, 1997) or Beer's minimal cognitive systems (Beer, 1996; Slocum et al., 2000) with the addition of sensory arrays and neural fields.

Acknowledgements

Thanks to Jianfeng Feng for comments and advice and for first suggesting STDP might be worth looking into. Many thanks also to Eytan Ruppin and Gal Chechick comments and good advice. The review process for this paper was handled by the Editor in Chief. The author wishes to acknowledge the support of the Nuffield Foundation, (grant no. NAL/00274/G).

Appendix: Estimation of σ_N^2

Here we calculate the variance σ_N^2 for a covariogram of two independent spike trains $S_1(t)$ and $S_2(t)$ with respective time-dependent averages $R_1(t)$ and $R_2(t)$ and variances $\sigma_1^2(t)$ and $\sigma_2^2(t)$. We follow Brody's (1999) development for many identical recorded trials by adapting it to a single sufficiently long trial.

First we estimate the $R_1(t)$ and $R_2(t)$ by performing simple sliding-window averages of the spike trains with windows of size $2t_d$. The corresponding variances are estimated using the same windows. In our case $t_d = 20$ ms.

The covariogram is defined as:

$$C(\tau) = \sum_{t=-\infty}^{\infty} (S_1(t) - R_1(t))(S_2(t+\tau) - R_2(t+\tau))$$

Independence of the trains is the null hypothesis. The expectation $E(C(\tau))$ in this case is zero. To calculate the corresponding variance σ_N^2, we obtain the variance of each term in the sum yielding:

$$\sigma_N^2(\tau) = \sum_{t=-\infty}^{\infty} \sigma_1^2(t)\sigma_2^2(t+\tau) + \sigma_1^2(t)R_2(t+\tau) + R_1(t)\sigma_2^2(t+\tau),$$

where we have used that the variance of a series x is $E(x^2) - E(x)^2$, and the variance of the product of two independent series x and y is therefore $E(x^2)E(y^2) - E(x)^2E(y)^2 = (\sigma_x^2 + R_x^2)(\sigma_y^2 + R_y^2) - R_x^2 R_y^2$.

The bands plotted in figures 9 and 11 correspond to the $\pm\sigma_N$ interval for each τ. To facilitate comparisons, all the covariograms have been normalised by a factor that makes the value of the autocovariogram for $\tau = 0$ equal to 1 – i.e., by dividing each series i by $\sum(S_i(t) - R_i(t)).^2$. The same factor has been applied correspondingly to the terms above in the calculation of σ_N^2.

References

Abbott, L. F., & Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cerebral Cortex*, **6**, 406–416.

Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature Neurosci*, **3**, 1178–1183.

Beer, R. D. (1990). *Intelligence as Adaptive Behavior: An Experiment in Computational Neuroscience*. San Diego: Academic Press.

Beer, R. D. (1996). Toward the Evolution of Dynamical Neural Networks for Minimally Cognitive Behavior. In Maes, P., Mataric, M. J., Meyer, J.-A., Pollack, J. B., & Wilson, S. W. (Eds.), *From Animals to Animats 4: Proceedings of the Fourth Internationall Conf on Simulation of Adaptive Behavior*, pp. 421 – 429. Cambridge, MA: MIT Press.

Beer, R. D., & Gallagher, J. C. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, **1**, 91–122.

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

Bi, G. Q., & Poo, M. M. (2001). Synaptic modifications by correlated activity: Hebb's postulated revisited. *Ann Rev Neurosci*, **24**, 139–166.

Brody, C. D. (1999). Correlations without synchrony. *Neural Computation*, **11**, 1537–1551.

Chechik, G. (2002). Spike-timing dependent plasticity and relevant mutual information maximization. In preparation.

Di Paolo, E. A. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In Meyer, J.-A., Berthoz, A., Floreano, D., Roitblat, H., & Wilson, S. (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on the Simulation of Adaptive Behavior* Paris, France. Cambridge MA: MIT Press.

Floreano, D., & Mattiussi, C. (2001). Evolution of spiking neural controllers for autonomous vision-based robots. In Gomi, T. (Ed.), *Evolutionary Robotics IV*. Springer Verlag.

Floreano, D., & Urzelai, J. (2000). Evolutionary Robots with on-line self-organization and behavioral fitness. *Neural Networks*, **13**, 431 – 443.

Gerstner, W., Kreiter, A. K., Markram, H., & Herz, A. V. M. (1997). Neural codes: firing rates and beyond. *Proc. Natl. Acad. Sci. USA*, **94**, 12740–12741.

Harvey, I., Husbands, P., & Cliff, D. (1994). Seeing the light: artificial evolution, real vision. In Cliff, D., Husbands, P., Meyer, J.-A., & Wilson, S. (Eds.), *From Animals to Animats 3, Proc. of 3rd Intl. Conf. on Simulation of Adaptive Behavior*, pp. 392 – 401. MIT Press, Cambridge, Mass.

Hopfield, J. J., & Brody, C. D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *Proc. Natl. Acad. Sci. USA*, **98**, 1282–1287.

Horn, D., Levy, N., & Ruppin, E. (1998). Memory maintenance via neuronal regulation. *Neural Computation*, **10**, 1–18.

Husbands, P., Smith, T., Jakobi, N., & O'Shea, M. (1998). Better living through chemistry: Evolving GasNets for robot control. *Connection Science*, **10**, 185–210.

Jakobi, N. (1997). Evolutionary Robotics and the Radical Envelope-of-Noise Hypothesis. *Adaptive Behavior*, **6**, 325–368.

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

Maass, W. (1997). Networks of spiking neurons: the third generation of neural network models. *Neural Networks*, **10**, 1656–1671.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. Submitted to Neural Computation.

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

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

Mehta, M. R., Barnes, C. A., & McNaughton, B. L. (1997). Experience-dependent asymmetric expansion of hippocampal place fields. *Proc. Natl. Acad. Sci. USA*, **94**, 8918–8921.

Mehta, M. R., Quirk, M. C., & Wilson, M. A. (2000). Experience-dependent asymmetric shape of hippocampal receptive fields. *Neuron*, **25**, 707–715.

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

Rubin, J., Lee, D. D., & Sompolinsky, H. (2001). Equilibrium properties of temporally asymmetric Hebbian plasticity. *Phys Rev Lett*, **86**, 364–367.

Ruppin, E. (2002). Evolutionary autonomous agents: A neuroscience perspective. *Nature Reviews Neuroscience*, **3**, 132–141.

Slocum, A., Downey, D., & Beer, R. D. (2000). Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention. In Meyer, J., Berthoz, A., Floreano, D., Roitblat, H., & Wilson, S. (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*, pp. 430 – 439. Cambridge, MA: MIT Press.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neurosci*, **3**, 919–926.

Stopfer, M., Bhagavan, S., Smith, B. H., & Laurent, G. (1997). Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. *Nature*, **390**, 70–74.

Sutton, R. S. (1988). Learning to predict by the method of temporal differences. *Machine Learning*, **3**, 220–224.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology. Vol 2.* Cambridge University Press.

Turrigiano, G. G. (1999). Homeostatic plasticity in neuronal networks: The more things change, the more they stay the same. *Trends Neurosci.*, **22**, 221–227.

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., & Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, **391**, 892–896.

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

Yao, H., & Y., D. (2001). Stimulus timing-dependent plasticity in cortical processing orientation. *Neuron*, **32**, 315–323.

List of figures

squared difference in weights between normal and Poisson-filtered conditions for a STDP controller with neural noise. Each line corresponds to one synapse and is obtained as the average of 10 runs. The thick line is the average for all weights.

Tables and figures

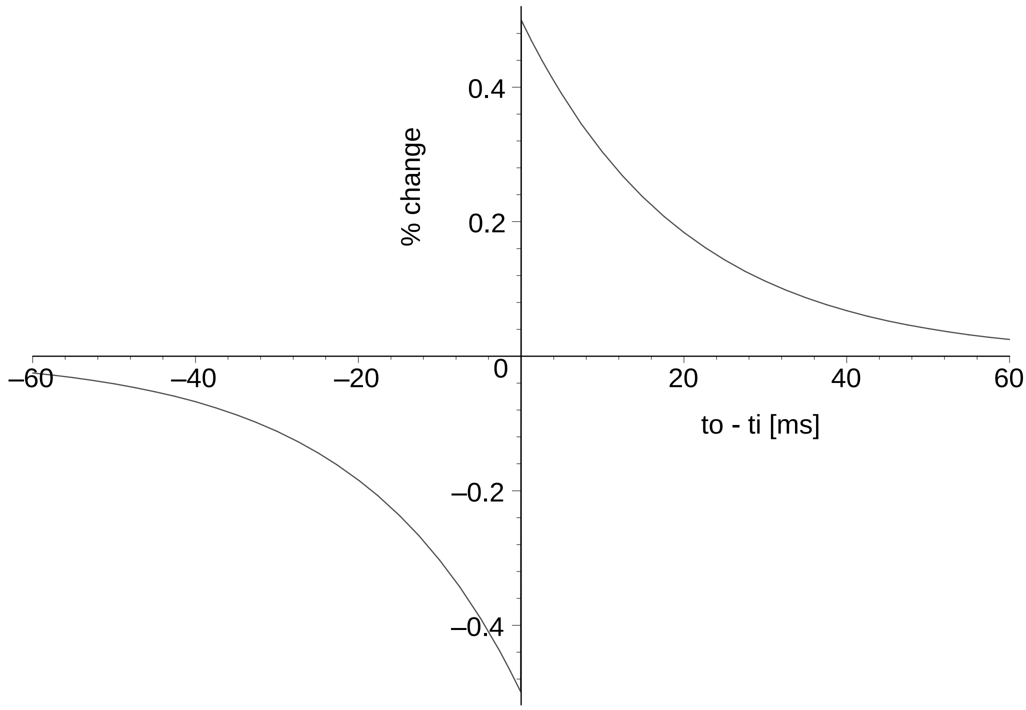


Figure 1:

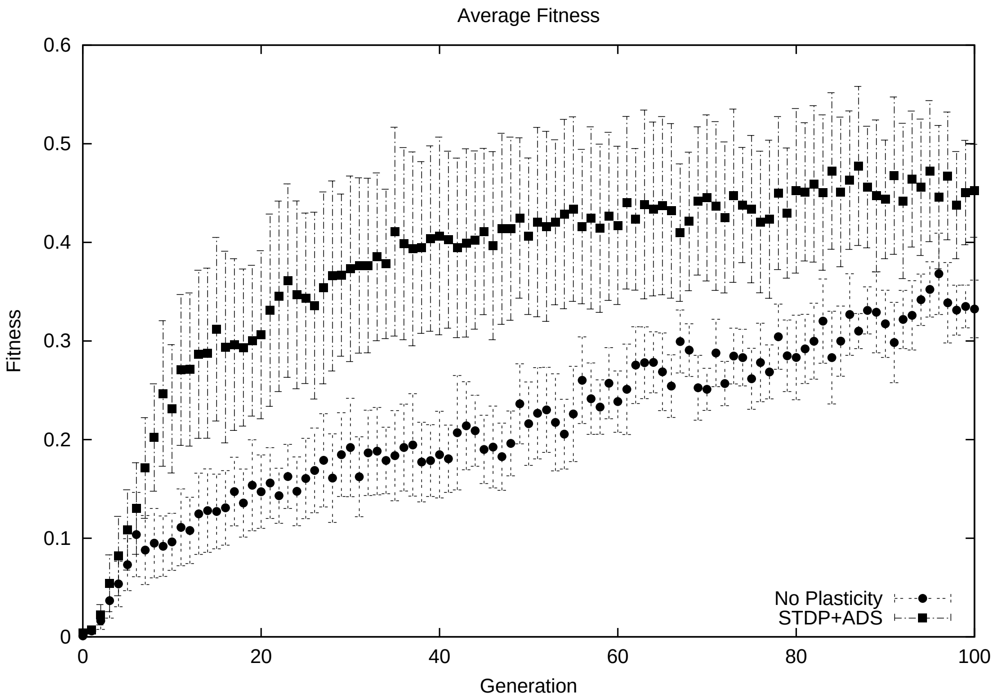
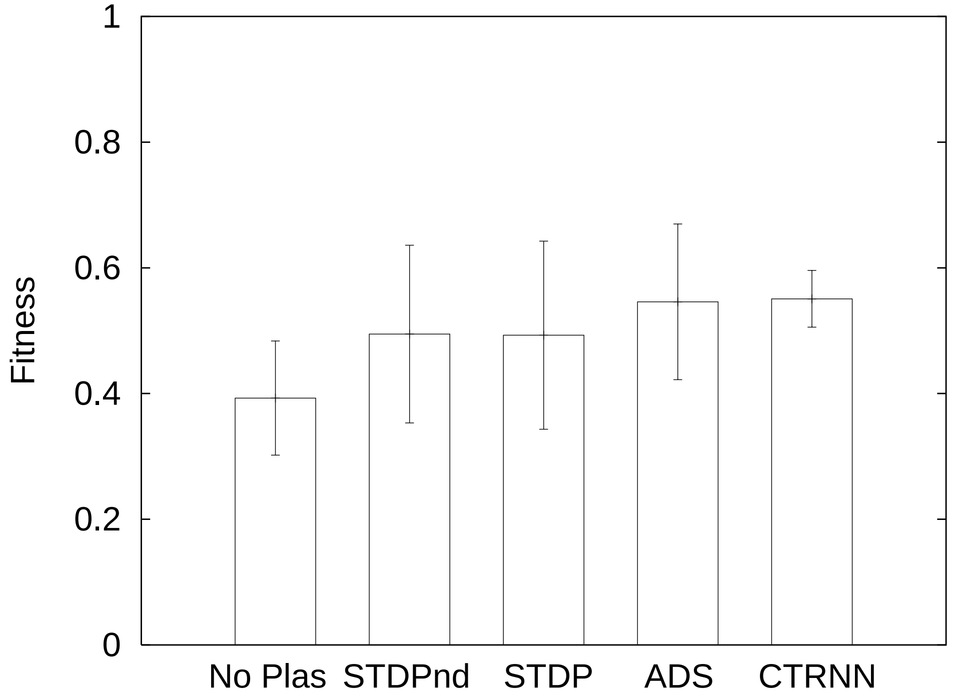
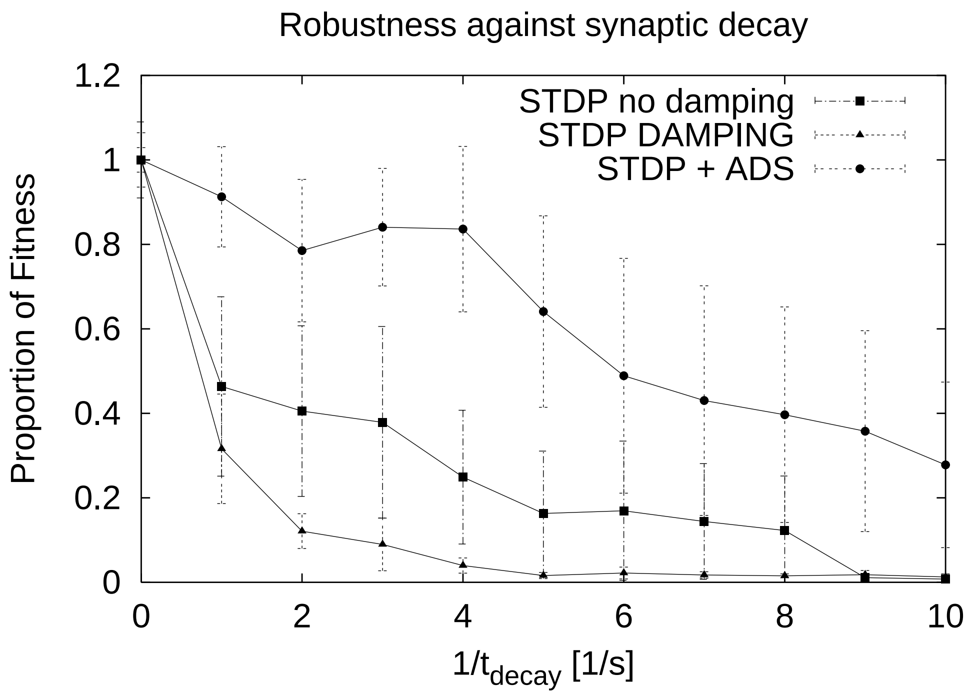

Average Fitness
Fitness
Generation
No Plasticity
STDP+ADS
0
0.1
0.2
0.3
0.4
0.5
0.6
0
20
40
60
80
100

Figure 2:

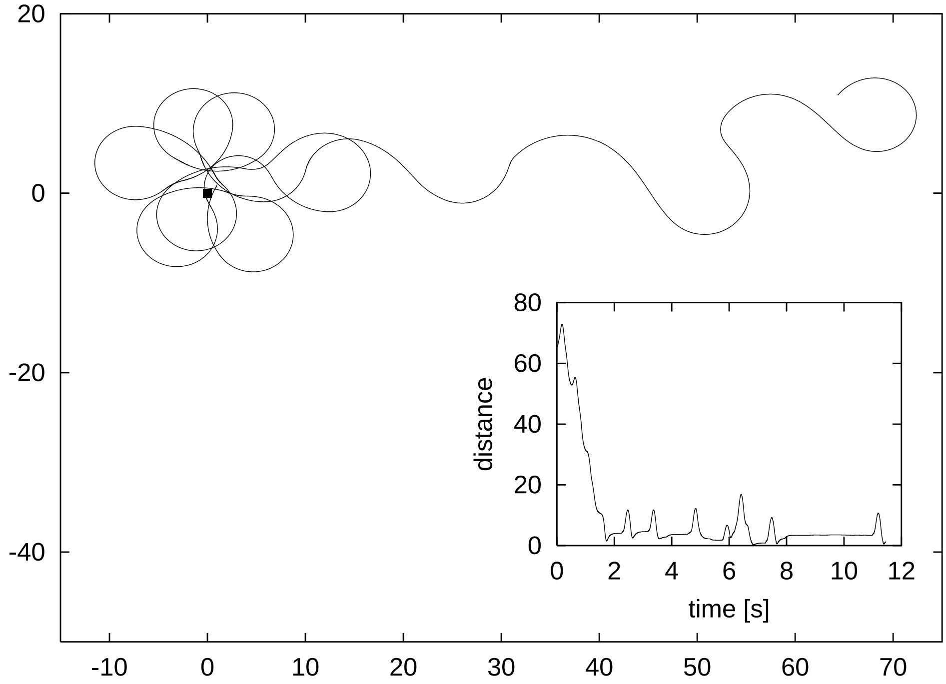

1
0.8
0.6
0.4
0.2
0
Fitness
No Plas
STDPnd
STDP
ADS
CTRNN

(a)

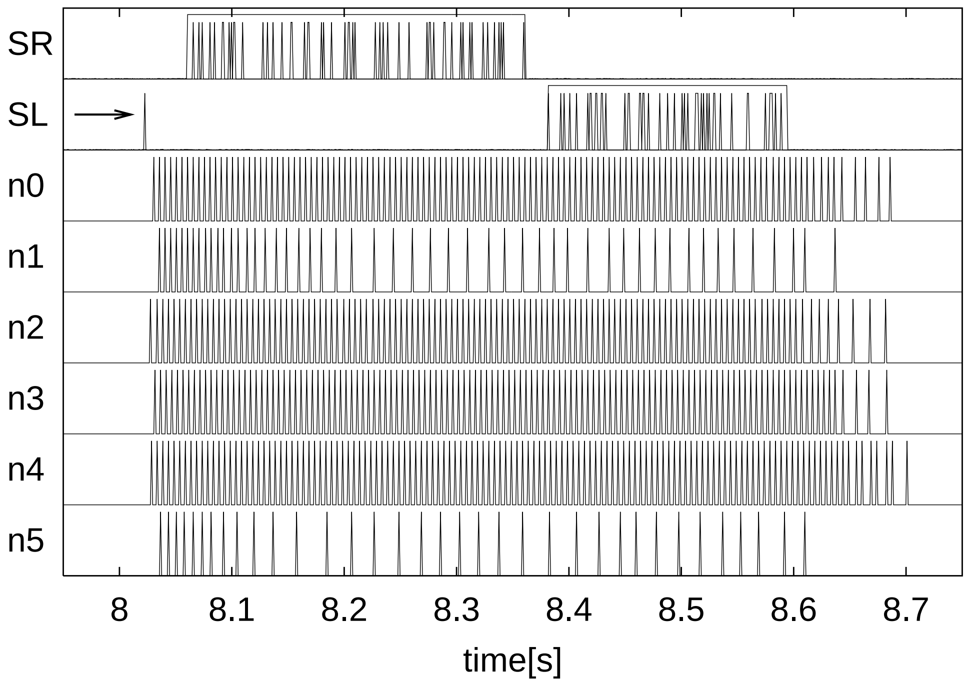

Robustness against synaptic decay
1.2
1
0.8
0.6
0.4
0.2
0
Proportion of Fitness
STDP no damping
STDP DAMPING
STDP + ADS
0
2
4
6
8
10
1/t_decay [1/s]

(b)

Figure 3:

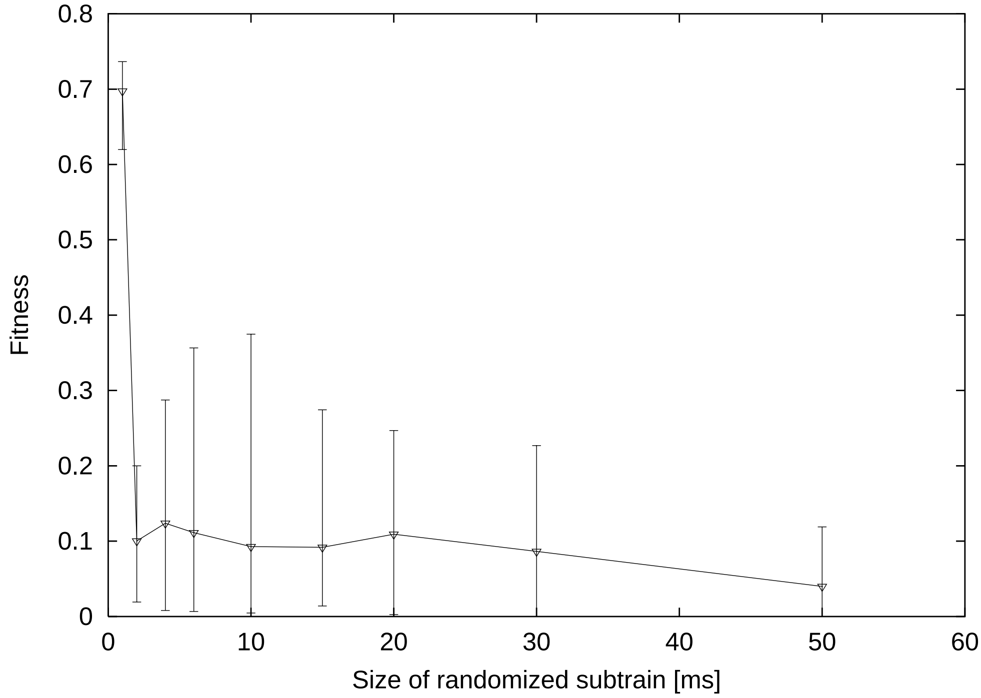

20
0
-20
-40
-10
0
10
20
30
40
50
60
70
80
60
40
20
0
distance
0
2
4
6
8
10
12
time [s]

(a)


SR
SL
n0
n1
n2
n3
n4
n5
8
8.1
8.2
8.3
8.4
8.5
8.6
8.7
time[s]

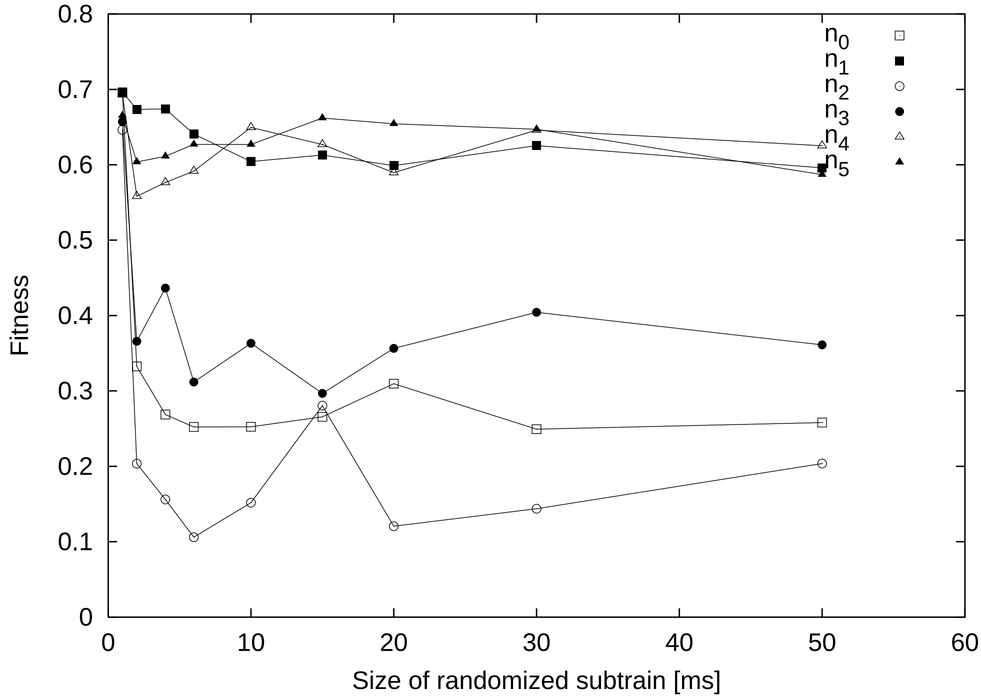
(b)

Figure 4:



0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Fitness
0
10
20
30
40
50
60
Size of randomized subtrain [ms]

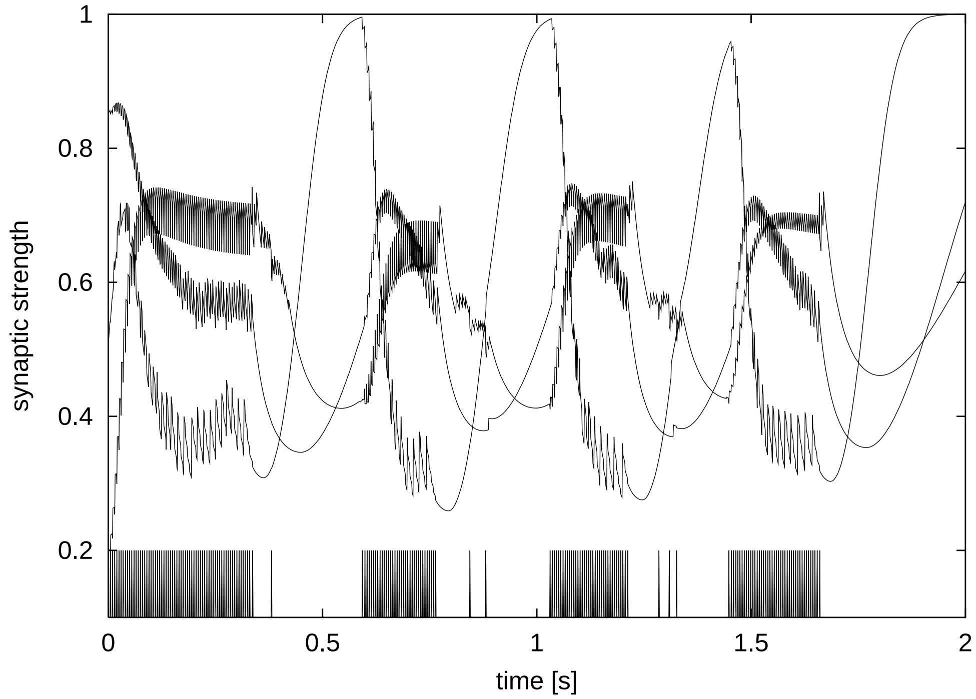
(a)



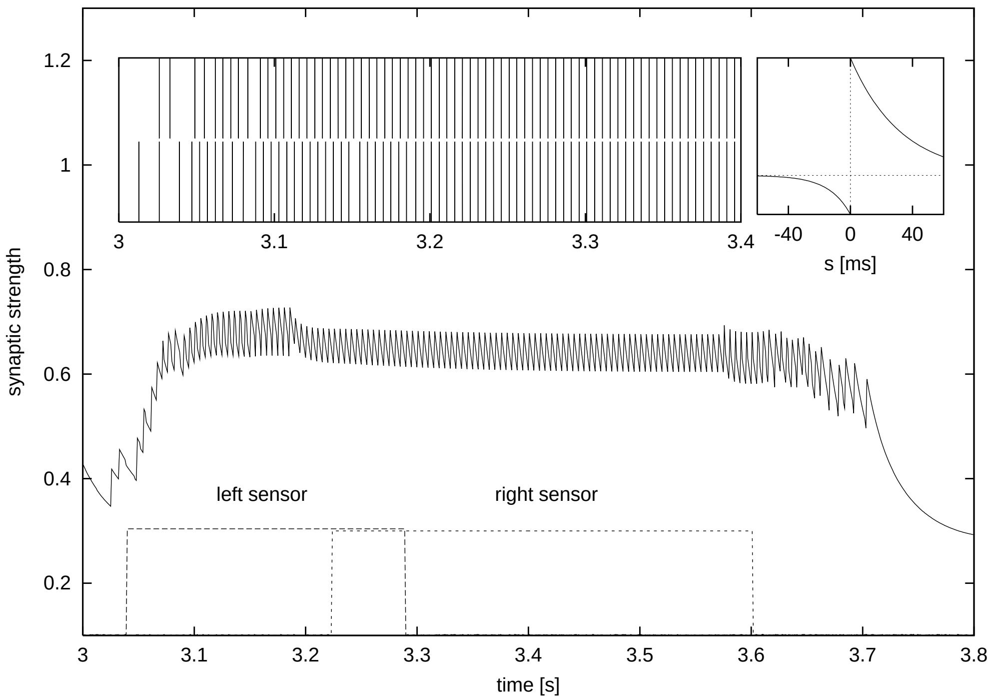
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Fitness
n_0
n_1
n_2
n_3
n_4
n_5
0
10
20
30
40
50
60
Size of randomized subtrain [ms]

(b)

Figure 5:

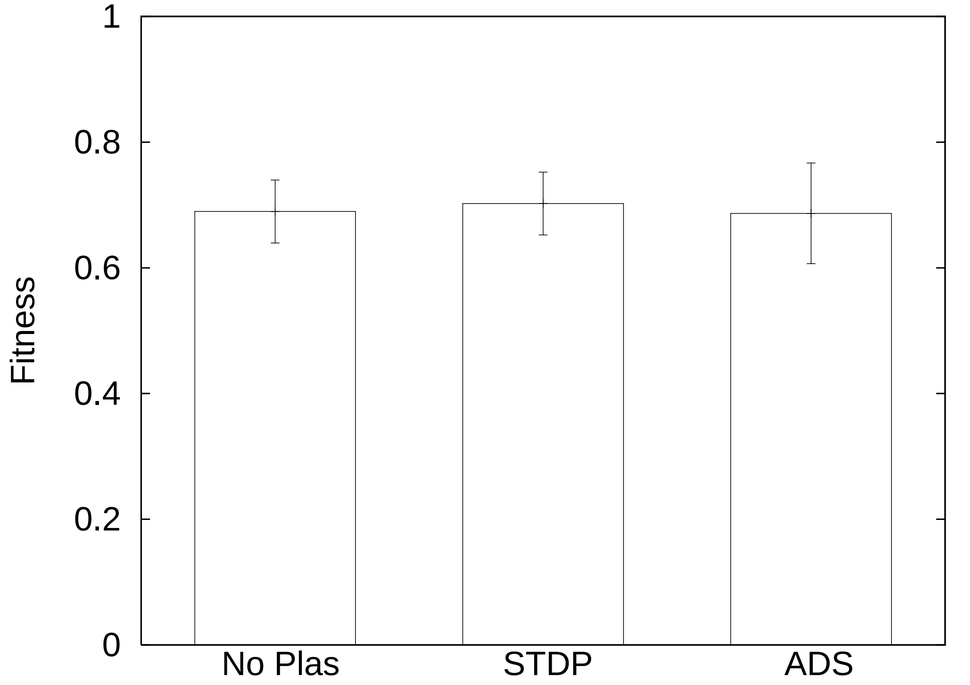

1
0.8
0.6
0.4
0.2
synaptic strength
0
0.5
1
1.5
2
time [s]

(a)


1.2
1
0.8
0.6
0.4
0.2
synaptic strength
3
3.1
3.2
3.3
3.4
-40
0
40
s [ms]
left sensor
right sensor
3
3.1
3.2
3.3
3.4
3.5
3.6
3.7
3.8
time [s]

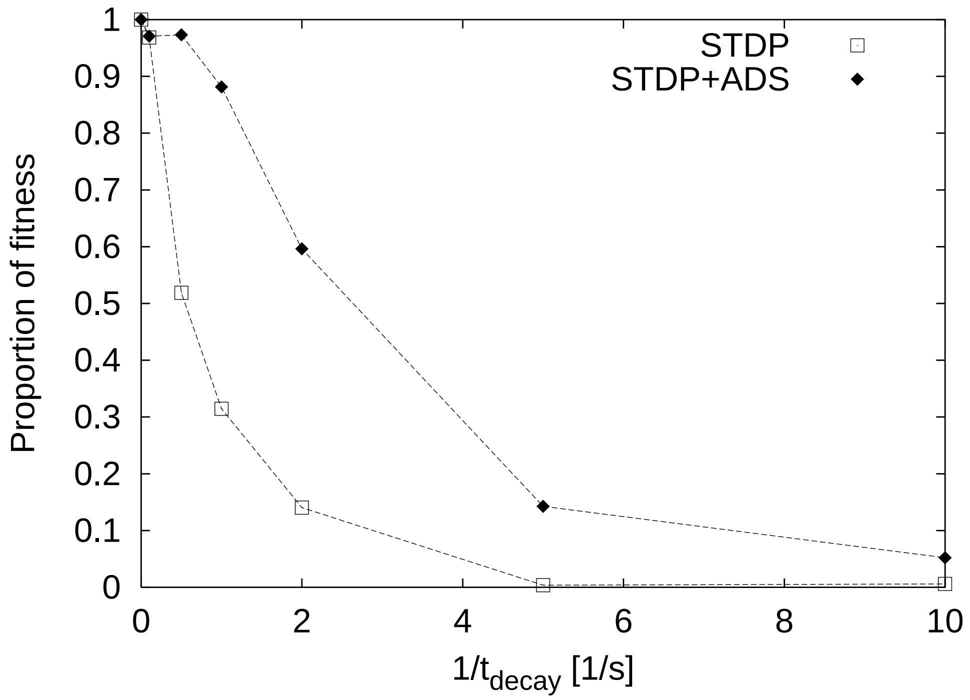
(b)

Figure 6:



1
0.8
0.6
0.4
0.2
0
Fitness
No Plas
STDP
ADS

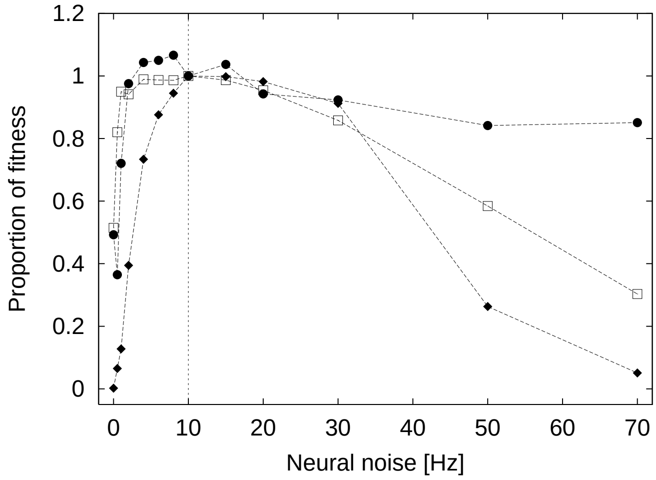
(a)



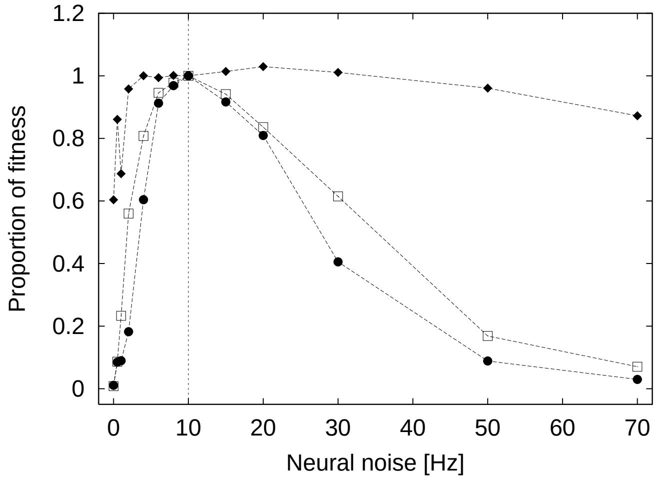
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Proportion of fitness
0
2
4
6
8
10
1/t$_{decay}$ [1/s]
STDP
STDP+ADS

(b)

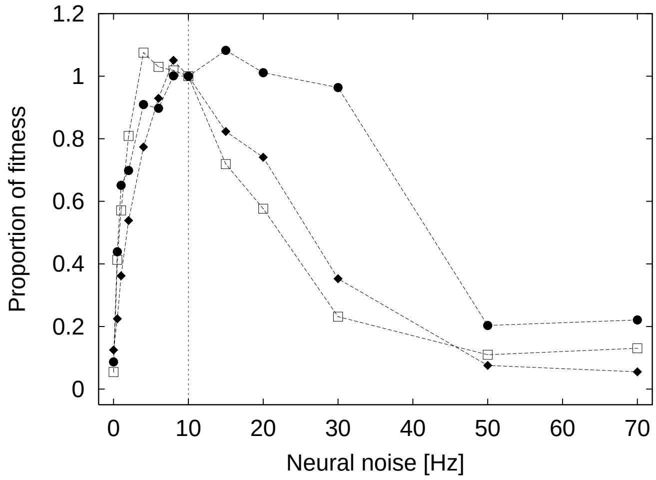
Figure 7:



(a) No Plasticity



(b) STDP



(c) STDP+ADS

Figure 8:

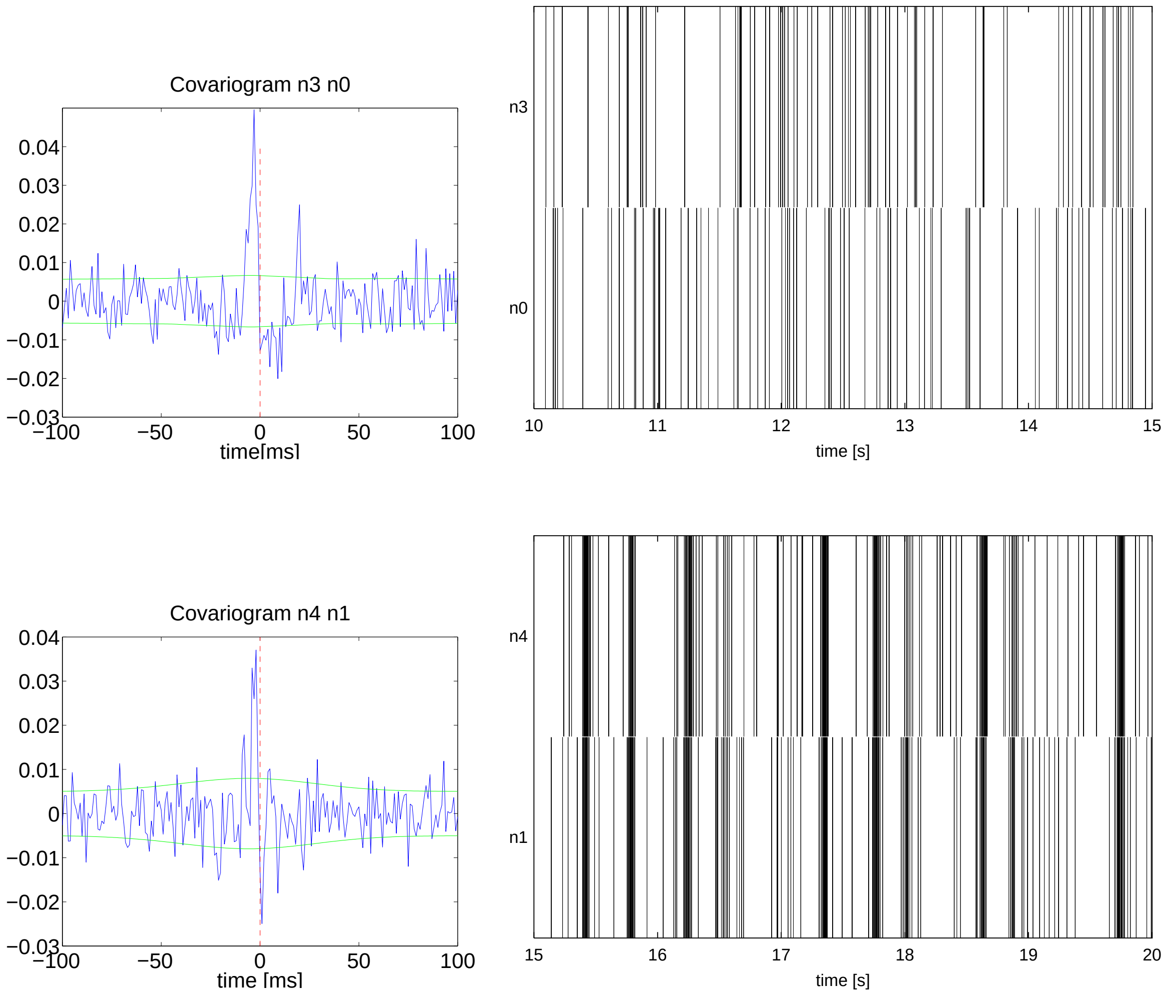
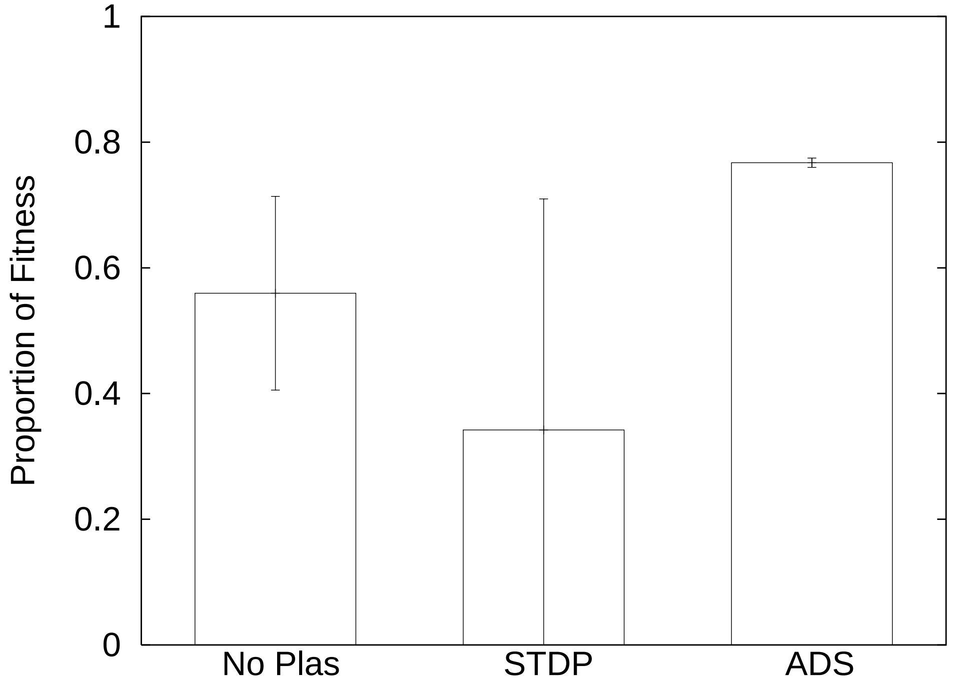
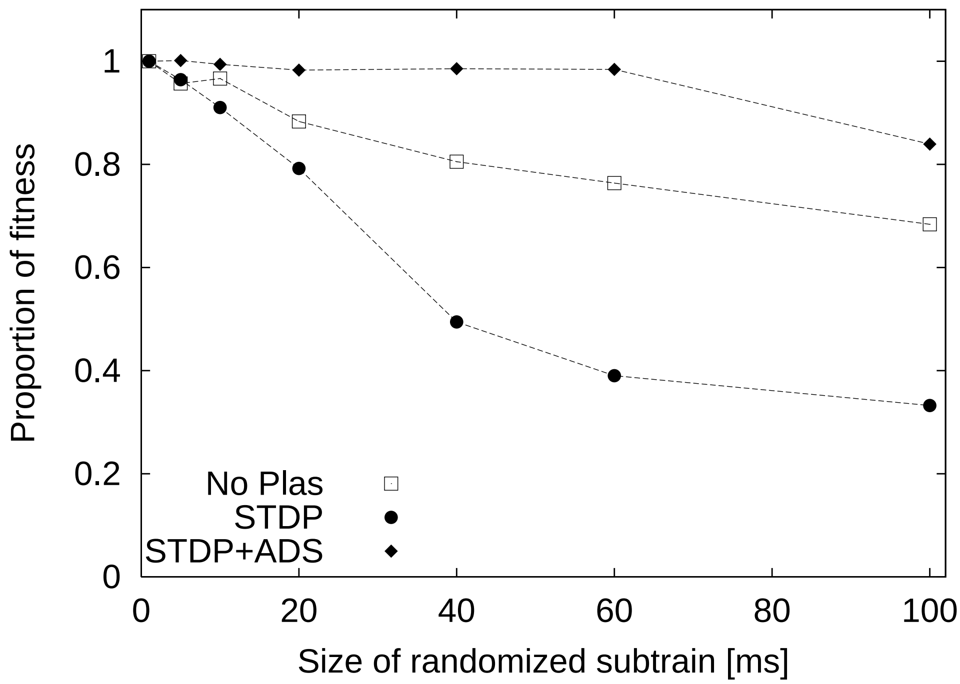

Covariogram n3 n0
0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−100
−50
0
50
100
time[ms]
n3
n0
10
11
12
13
14
15
time [s]
Covariogram n4 n1
0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−100
−50
0
50
100
time [ms]
n4
n1
15
16
17
18
19
20
time [s]

Figure 9:


1
0.8
0.6
0.4
0.2
0
Proportion of Fitness
No Plas
STDP
ADS

(a)


1
0.8
0.6
0.4
0.2
0
Proportion of fitness
No Plas
STDP
STDP+ADS
0
20
40
60
80
100
Size of randomized subtrain [ms]

(b)

Figure 10:

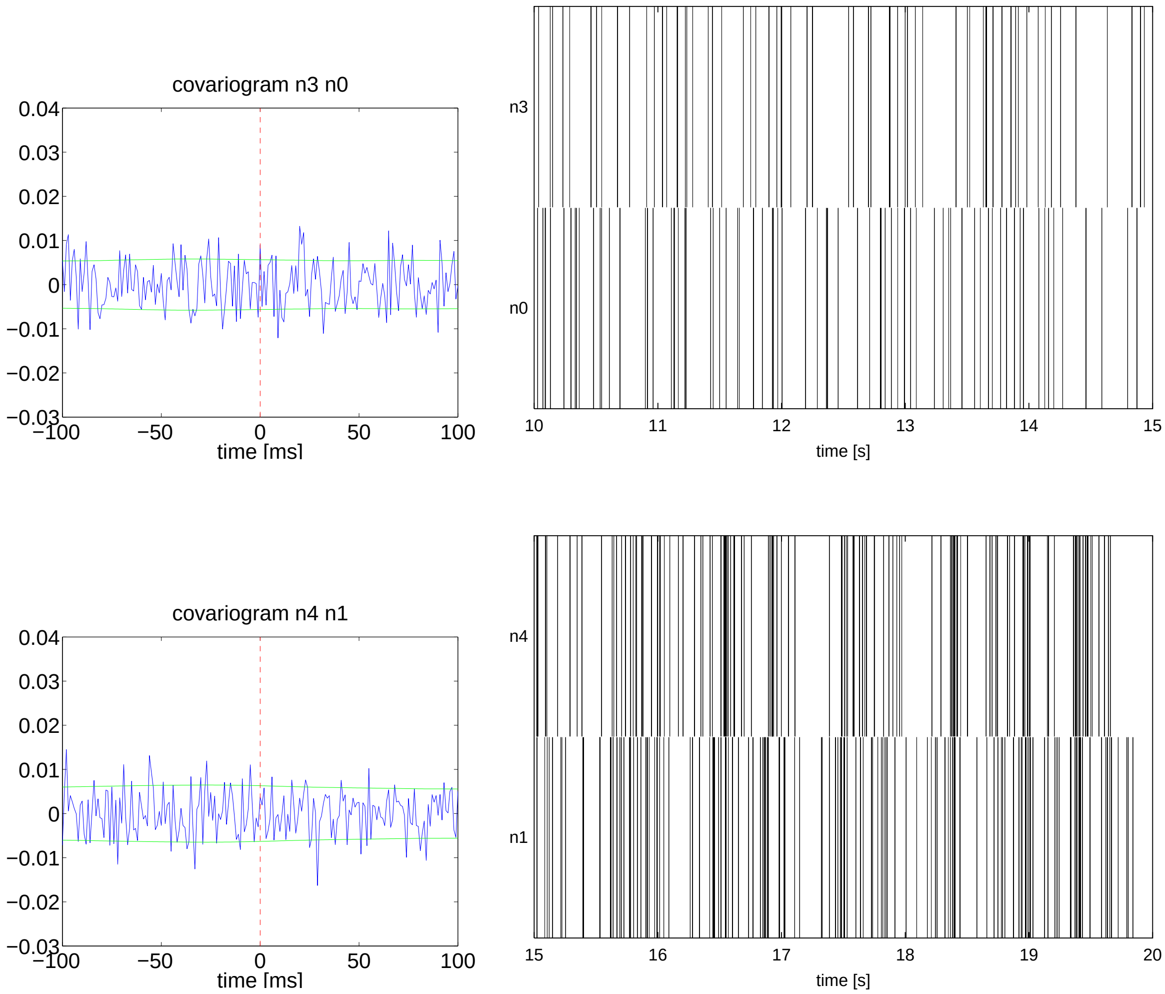

covariogram n3 n0
0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−100
−50
0
50
100
time [ms]
n3
n0
10
11
12
13
14
15
time [s]
covariogram n4 n1
0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
−100
−50
0
50
100
time [ms]
n4
n1
15
16
17
18
19
20
time [s]

Figure 11:

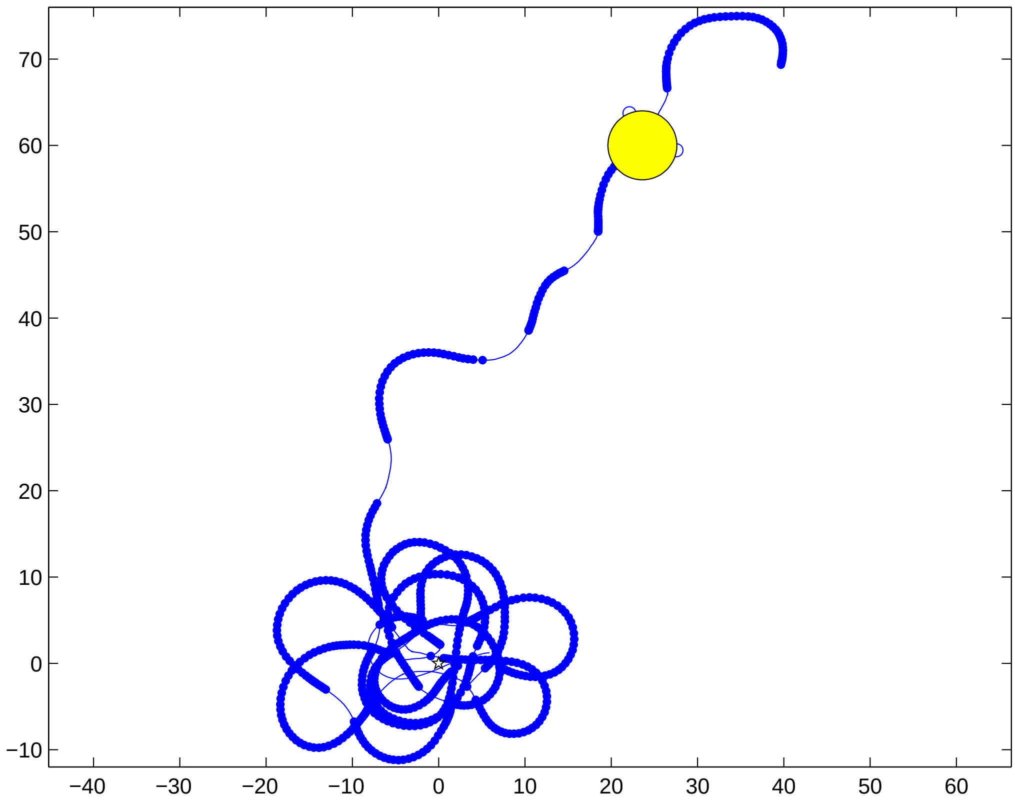
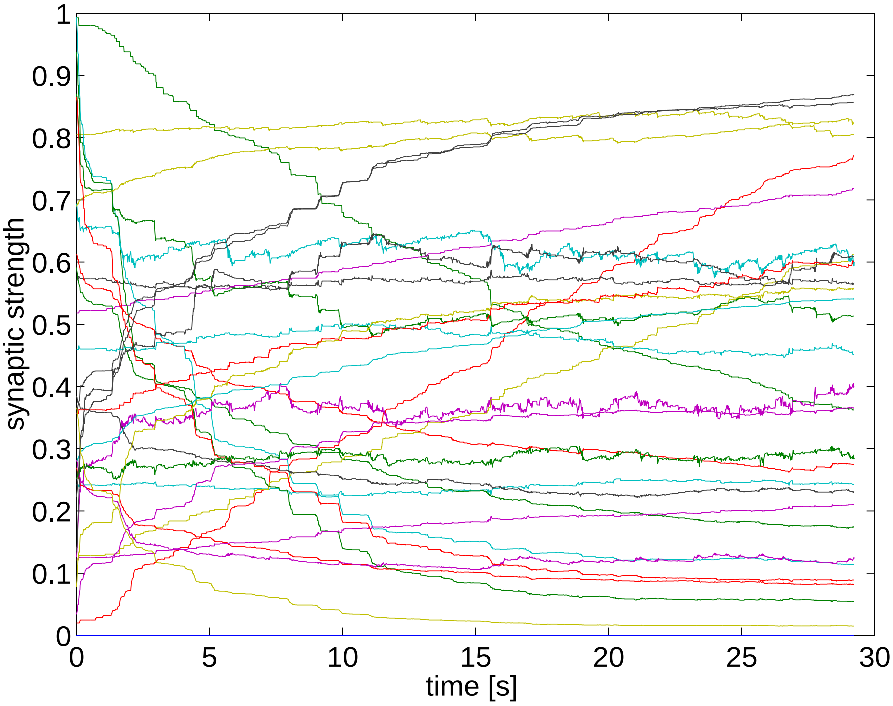
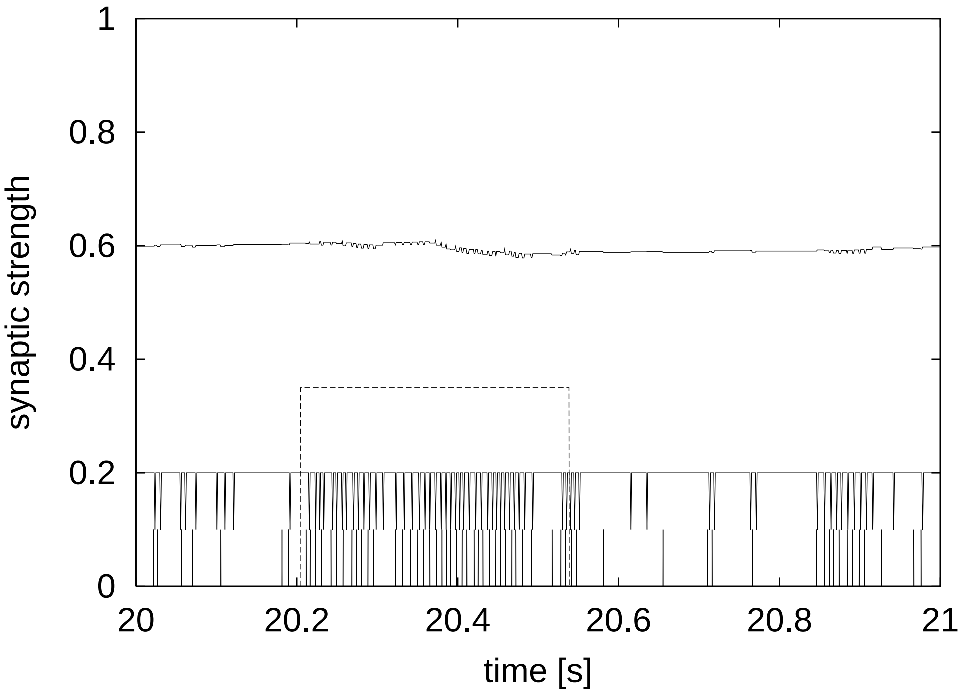


Figure 12:

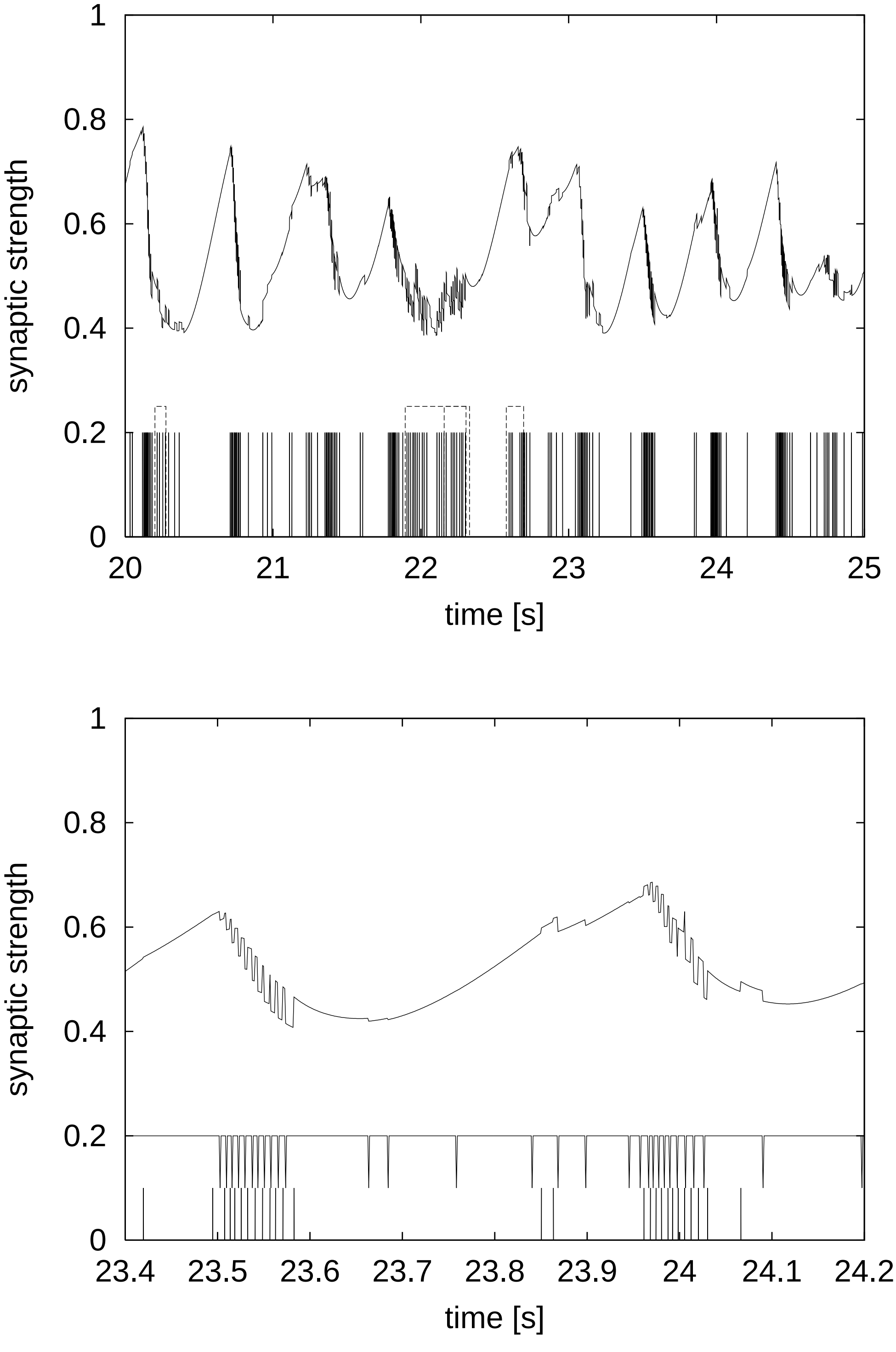


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
synaptic strength
0
5
10
15
20
25
30
time [s]



1
0.8
0.6
0.4
0.2
0
synaptic strength
20
20.2
20.4
20.6
20.8
21
time [s]

Figure 13:



synaptic strength
1
0.8
0.6
0.4
0.2
0
20
21
22
23
24
25
time [s]
synaptic strength
1
0.8
0.6
0.4
0.2
0
23.4
23.5
23.6
23.7
23.8
23.9
24
24.1
24.2
time [s]

Figure 14:

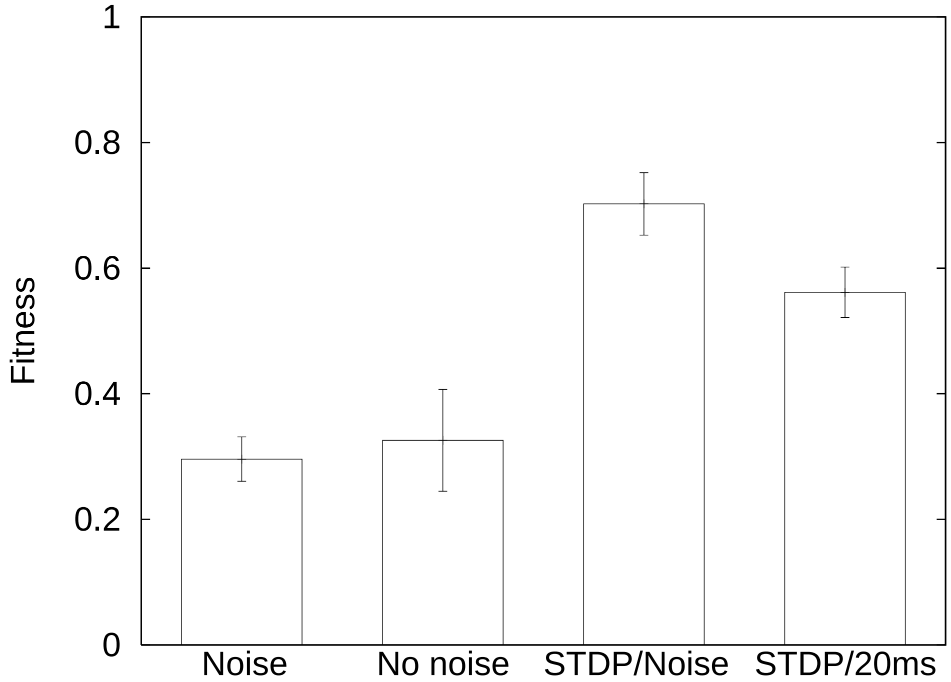


Figure 15:

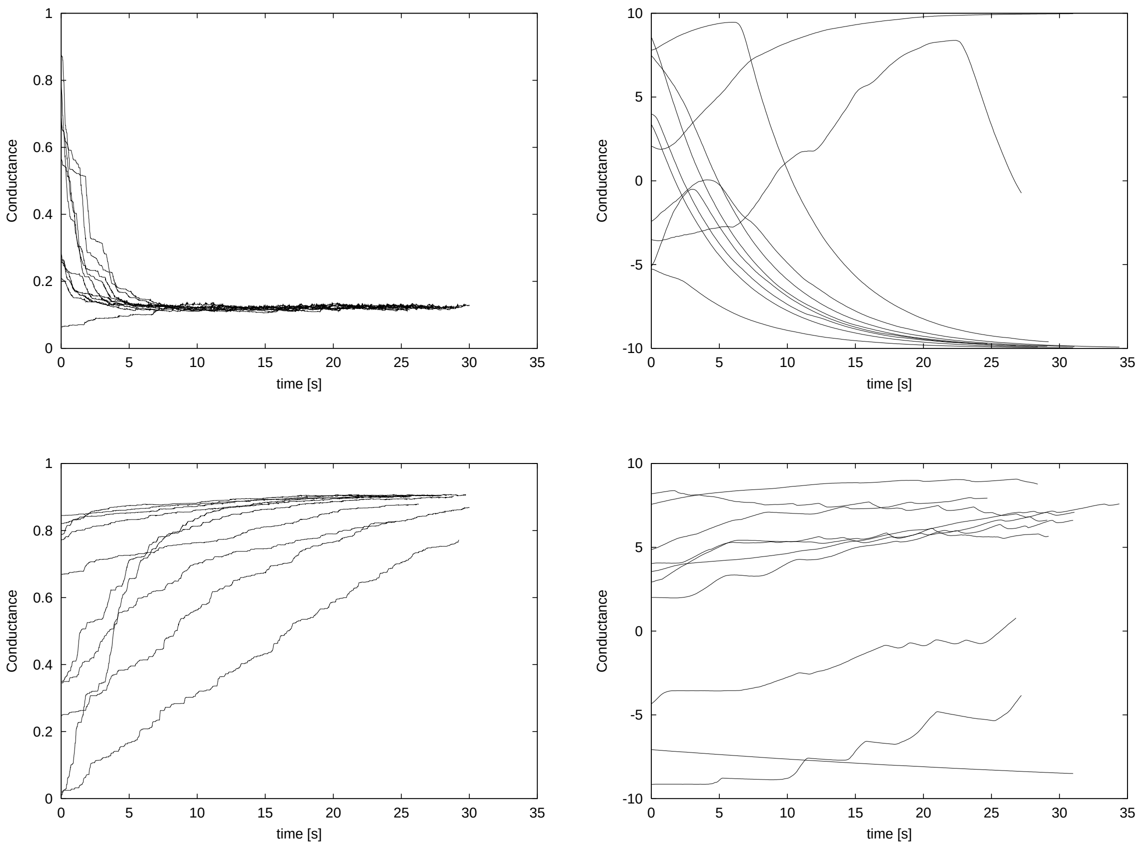
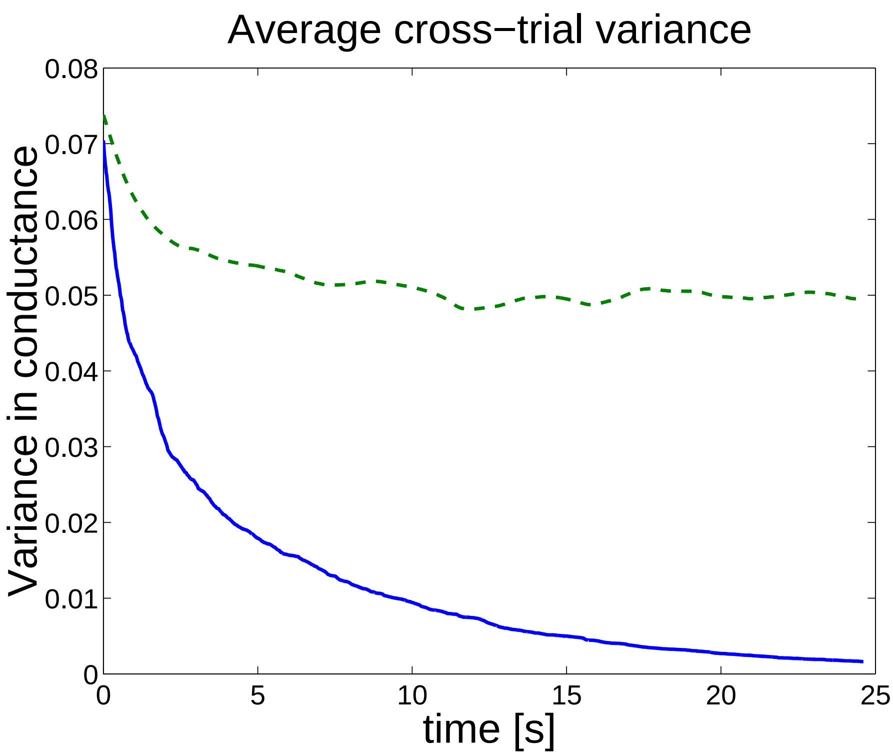
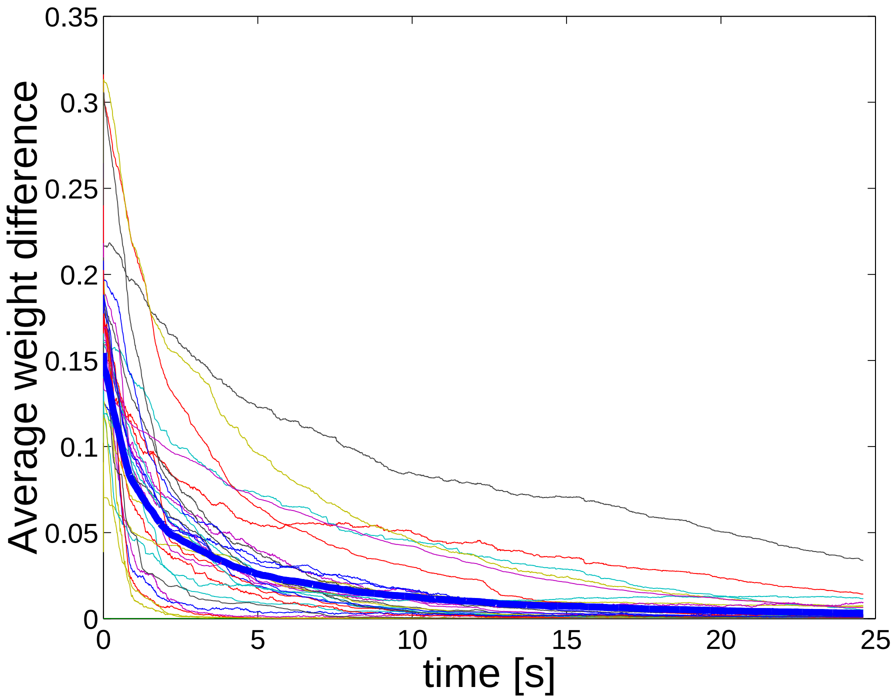


Figure 16:


Average cross−trial variance
Variance in conductance
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
20
25
time [s]

(a)


Average weight difference
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
time [s]

(b)

Figure 17: