ARTICLE **Communicated by Misha Tsodyks**

Spike-Driven Synaptic Plasticity: Theory, Simulation, VLSI Implementation

Stefano Fusi
INFN Sezione RM1, University of Rome "La Sapienza," 00185, Roma, Italy

Mario Annunziato *Physics Department, University of Pisa, 56100 Pisa, Italy*

Davide Badoni *INFN, Sezione RM2, University of Rome "Tor Vergata," 00133, Roma, Italy*

Andrea Salamon *Physics Department, University of Rome "La Sapienza," 00185, Roma, Italy*

Daniel J. Amit *INFN Sezione RM1, University of Rome "La Sapienza," 00185, Roma, Italy, and Racah Institute of Physics, Hebrew University, Jerusalem*

We present a model for spike-driven dynamics of a plastic synapse, suited for aVLSI implementation. The synaptic device behaves as a capacitor on short timescales and preserves the memory of two stable states (efficacies) on long timescales. The transitions (LTP/LTD) are stochastic because both the number and the distribution of neural spikes in any finite (stimulation) interval fluctuate, even at fixed pre- and postsynaptic spike rates. The dynamics of the single synapse is studied analytically by extending the solution to a classic problem in queuing theory (Takàcs process). The model of the synapse is implemented in aVLSI and consists of only 18 transistors. It is also directly simulated. The simulations indicate that LTP/LTD probabilities versus rates are robust to fluctuations of the electronic parameters in a wide range of rates. The solutions for these probabilities are in very good agreement with both the simulations and measurements. Moreover, the probabilities are readily manipulable by variations of the chip's parameters, even in ranges where they are very small. The tests of the electronic device cover the range from spontaneous activity (3–4 Hz) to stimulus-driven rates (50 Hz). Low transition probabilities can be maintained in all ranges, even though the intrinsic time constants of the device are short ($\sim$ 100 ms).

Synaptic transitions are triggered by elevated presynaptic rates: for low presynaptic rates, there are essentially no transitions. The synaptic device can preserve its memory for years in the absence of stimulation. Stochasticity of learning is a result of the variability of interspike intervals; noise is a feature of the distributed dynamics of the network. The fact

that the synapse is binary on long timescales solves the stability problem of synaptic efficacies in the absence of stimulation. Yet stochastic learning theory ensures that it does not affect the collective behavior of the network, if the transition probabilities are low and LTP is balanced against LTD.

1 Introduction

Material learning devices have the capability of changing their internal states in order to acquire (learn) and store (memorize) information about the statistics of the incoming flux of stimuli. In the case of neural devices, this information is stored in the structure of couplings between neurons. Hence, developing a learning neural network device means designing a plastic synapse whose synaptic efficacy can be rapidly modified to acquire information and, at the same time, can be preserved for long periods to keep this information in memory. These two requirements are particularly important for on-line unsupervised learning, in which the network must handle an arbitrary stream of stimuli, with no a priori knowledge about their structure and relevance. In the absence of an external supervisor, the synaptic efficacy should be either updated (even just a little) by each stimulus that is presented, or there will be no trace of any of them. Most discussion of learning dynamics to date has focused on the structure of a developing analog synaptic variable as a function of an incoming set of stimuli (Grossberg, 1969; Sejnowski, 1976; Bienenstock, Cooper, & Munro, 1982). Those must be extended before a material implementation can be contemplated.

Developing an electronic plastic synapse that can be embedded in a large learning neural network implies analog VLSI (aVLSI) technology. Since the number of synapses can be as large as the square of the number of neurons, the design of the plastic synapse should follow two principles: economy in the number of transistors, to reduce chip area, and low power consumption. The neuromorphic approach, introduced by Mead (1989), has been adopted for the construction of a wide range of neural systems that emulate biological counterparts. Since these devices exploit the inherent physics of analog transistors to produce computation, they use less power and silicon area than their equivalent digital devices (Sarpeshkar, 1998). However, this approach has been applied mainly to specialized systems (e.g., sensory devices), with no learning. The lack of a simple synaptic device that could couple the neuromorphic analog neurons was probably one of the main limiting factors.

There are two common approaches to implementing synaptic devices. First, the synaptic weights are stored on a digital device (typically a RAM) off-chip: the analog neurons read the synaptic value via digital-to-analog converters (DAC). Such memory can have arbitrary analog depth and can be preserved for indefinitely long periods. It is not, however, an acceptable solution for implementation of large-scale networks of analog neurons

because most of the space would be occupied by the DAC and ADC converters. In the second approach, the memory of the synaptic efficacies is stored as the voltage across a capacitor (Elias, Northmore, & Westerman, 1997; Hafliger, Mahowald, & Watts, 1996): the device implementing the synapse can preserve a continuous set of values, but the memory is volatile and is lost on timescales of RC, where C is the capacitance and R is the leakage resistance. A promising solution in this direction is emerging in the form of synaptic devices using floating-gate technology (see, e.g., Diorio, Hasler, Minch, & Mead, 1998). These devices can preserve the synaptic efficacy with good analog depth and for times of the order of years (the problems related to the updating procedure, such as high voltages, are probably going to be solved).

Here we chose an alternative solution, which is technically quite simple and is well within the scope of familiar technology. It is based on the theoretical observation that the performance of the network is not necessarily degraded if the analog depth of the synapse is limited, even to the extreme. Most of the effort to achieve large analog depth has been motivated by the drive to implement learning rules based on the hypotheses: that closely spaced, stable synaptic efficacies are essential for the proper functioning of the network and that on short timescales it is possible to acquire information about a stimulus, by small changes in the synaptic efficacy induced by neural activities elicited by that stimulus (Hopfield, 1982). As a matter of fact, another line of study always had it that associative memory could operate with binary synapses. In fact, if stimuli are coded sparsely (i.e., the fraction of neurons each stimulus activates is low), such networks have also optimal performance (Willshaw, 1969; Tsodyks, 1990; Amit & Fusi, 1994). Some empirical evidence that efficacies may indeed be binary, is beginning to accumulate (Petersen, Malenka, Nicoll, & Hopfield, 1998). It is this option that was chosen in the development of the analog device since the binary dynamic memory element has been the natural choice in the entire information industry for the last half-century.

Networks with this kind of synapses operate as palimpsests (Nadal, Toulouse, Changeux, & Dehaene, 1986; Parisi, 1986; Amit & Fusi, 1994); that is, old stimuli are forgotten to make room for more recent ones. There is a sliding window of retrievable memory; only part of the past is preserved. The width of the window, the memory span, is related to the learning and the forgetting rates: fast rates imply short windows. It also depends on the fraction of synapses that change following each presentation: fractions close to 1 imply fast learning and forgetting, but the memory span is reduced. Note, by the way, that deep analog synapses do not share the palimpsest property, and when their memory is overloaded, they forget everything (Amit, Gutfreund, & Sompolinsky, 1987; Amit, 1989).

Every stimulus makes a subdivision of the entire pool of synapses in the network; only active neurons are involved in changing synapses. Moreover, not all the synapses that connect neurons activated by a given stimulus

must be changed in a given presentation, or at all, in order to acquire information about the stimulus to be learnt. The learning process requires a mechanism that selects which synaptic efficacies are to be changed following each stimulation. In the absence of an external supervisor, a local, unbiased mechanism could be stochastic learning: at parity of conditions (equal pair of activities of pre- and postsynaptic neurons), the transitions between stable states occur with some probability; only a randomly chosen fraction of those synapses that might have been changed by a given stimulus is actually changed upon presentation of that stimulus.

If stimuli are sparsely coded; transition probabilities are small enough, and long-term potentiation (LTP) is balanced against long-term depression (LTD), then optimal storage capacity is recovered, even for two state synapses (Amit & Fusi, 1994). Moreover, small transition probabilities lead to slow learning and hence to the generation of internal representations that are averages over many nearby stimuli, that is, class representatives. Fast learning generates representatives faithful to the last stimulus of a class. Slow learning also corresponds to observations in inferotemporal cortex (Miyashita, 1993). On the other hand, the particular model chosen (two states for the synapse and linear interspike behavior of synapse and neuron) simplifies significantly the hardware implementation. It also makes possible an analytical treatment of the synaptic dynamics. On the network level, sparse coding and low transition probabilities make possible a detailed analysis of the learning process (Brunel, Carusi, & Fusi, 1998).

What we present here is a study of the properties of the stochastic dynamics in a prototypical case. We focus on the dependence of the synaptic transition probabilities on the activities of the two neurons it connects. It is not possible to describe such dynamics in a general case. We consider this study in analogy with the studies of the single integrate-and-fire (IF) neuron, as carried out, for example, by Ricciardi (1977; Tuckwell, 1988). This study has clarified the spike emission statistics of a leaky IF neuron with an afferent current of gaussian distribution. Clearly, the spiking dynamics of the neuron does not have to follow the RC circuit analogy (see, e.g., Gerstein & Mandelbrot, 1964), nor is the current in general gaussian. In the same sense we will choose a specific dynamical process for the synapse—the one also implemented in aVLSI, as well as specific processes for the pre- and postsynaptic activities. The synaptic dynamics is triggered by presynaptic spikes, which arrive in a Poisson distributed process characterized by the rate; the postsynaptic neuron is an IF neuron with a linear leak and a "floor" (Fusi & Mattia, 1999), driven by a current that is a sum of a large number of Poisson processes. The synapse is described by an analog variable that makes jumps up or down whenever, upon presentation of a stimulus, a spike appears on the presynaptic neuron. The direction of the jump is determined by the level of depolarization of the postsynaptic neuron. The dynamics of the analog synaptic variable is similar to the BCM rule (Bienenstock et al., 1982), except that our synapse maintains its memory on long timescales in the absence

of stimulation. Between spikes, the synaptic variable drifts linearly up, toward a ceiling (down, toward a floor) depending on whether the variable is above (below) a synaptic threshold. The two values that delimit the synaptic variable are the stable synaptic efficacies.

Since stimulation takes place on a finite (short) interval, stochasticity is generated by both the pre- and postsynaptic neuron. Transitions (LTP, LTD) take place on presynaptic bursts, when the jumps accumulate to overcome the refresh drifts. The synapse can preserve a continuous set of values for periods of the order of its intrinsic time constants. But on long timescales, only two values are preserved: the synaptic efficacy fluctuates in a band near one of the two stable values until a strong stimulation produces a burst of spikes and drives it out of the band and into the neighborhood of the other stable value.

Stochastic learning solves the theoretical problems related to performance of the network, but small transition probabilities require bulky devices (Badoni, Bertazzoni, Buglioni, Salina, Amit, & Fusi, 1995). The dynamics of our synapse makes use of the stochasticity of the spike emission process of pre- and postsynaptic neurons. Hence, on the network level, the new synapse exploits the collective dynamical behavior of the interconnected neurons (Amit & Brunel, 1997; van Vreeswijk & Sompolinsky, 1997; Fusi & Mattia, 1999).

In section 2 we present the model of the synapse and describe the details of the dynamics. In section 3 the transition probabilities are estimated analytically by extending the theory of the Takàcs queuing process and are studied as a function of the pre- and postsynaptic neuron rates. In section 4 we present the aVLSI implementation of the modeled synapse, and in section 5 we test the hardware device and show that stimulations induce stochastic transitions between the two stable states and the transition probabilities that can be obtained satisfy the requirements of the stochastic learning theory. In section 6 we discuss extensions of the approach to other synaptic dynamics, the role of the synapse in a network, and the correspondence of the implemented synapse with neurobiological observation.

2 Synaptic Dynamics

The synaptic dynamics is described in terms of the analog internal variable $X(t)$, which in turn determines the synaptic efficacy $J(t)$, defined as the change in postsynaptic depolarization, or afferent current, induced by a single presynaptic spike. The synaptic efficacy has two values: it is potentiated ($J(t) = J_1$) when the internal variable X is above the threshold θ_X and depressed ($J(t) = J_0$) if $X < \theta_X$. X is restricted to the interval $[0, 1]$ and obeys

$$\frac{dX(t)}{dt} = R(t) + H(t); \tag{2.1}$$

at the boundaries (0 and 1), the right-hand side vanishes if it attempts to drive X outside the allowed interval. Hence X= 0 and 1 are reflecting barriers for the dynamics of X.

$R(t)$ is a refresh term, responsible for memory preservation. $H(t)$ is the stimulus-driven Hebbian learning term, which is different from 0 only upon arrival of a presynaptic spike. Between spikes, the synaptic dynamics, dominated by $R(t)$, is:

$$R(t) = -\alpha\Theta\left(-X(t) + \theta_X\right) + \beta\Theta\left(X(t) - \theta_X\right), \tag{2.2}$$

which is a constant drift down of rate α, or up with rate β, depending on whether X is below or above a synaptic threshold θ_X (Θ is the Heaviside function).

The Hebbian term depends on variables related to the activities of the two neurons connected by the synapse: (1) the spikes from the presynaptic neuron and (2) the instantaneous depolarization V_{post} of the postsynaptic neuron. Presynaptic spikes trigger the Hebbian dynamics. Each spike induces a jump in X whose value depends on the instantaneous depolarization of the postsynaptic neuron. The depolarization is indirectly related to the activity of the postsynaptic neuron (see the appendix) and is available at the synaptic contact. The capacitance of the neuron is exploited to compute this quantity, which is smooth and provides a finite support for coincidence between the presynaptic spikes and the postsynaptic activity. (See section 6 for possible extensions.)

In the general case, the Hebbian term can be written as

$$H(t) = \sum_k F(V_{post}(t_k^{pre}))\delta(t - t_k^{pre}). \tag{2.3}$$

The sum is over all presynaptic spikes; spike number k arrives at the synapse at t_k^{pre}. The value of V_{post} upon arrival of the presynaptic spike determines the jump, $F(V_{post}(t_k^{pre}))$, in the synaptic variable. The sign of F determines whether the synapse is potentiated or depressed at the event. This form of the Hebbian term, for the analog synaptic variable, is similar to the BCM rule (Bienenstock et al., 1982).

In the case studied and implemented here, we have chosen a particular form for F. If the pre-synaptic spike arrives when $V_{post} > \theta_V$, then the synaptic internal state is pushed up by a, toward the potentiated state ($X \to X+a$). If it arrives while $V_{post} < \theta_V$, then $X \to X-b$ ($b > 0$). This can be summarized formally as

$$F(V_{post}) = a\Theta(V_{post}(t) - \theta_V) - b\Theta(\theta_V - V_{post}(t)). \tag{2.4}$$

The dynamics of equation 2.1 is illustrated in Figure 1 by simulating a synapse connecting two spiking neurons that linearly integrate a gaussian

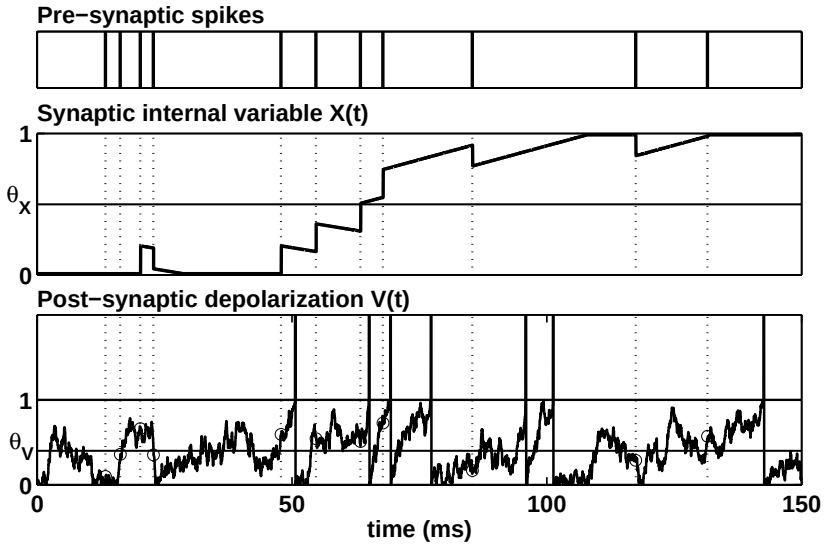


Figure 1: Synaptic dynamics. Time evolution of the simulated internal variable $X(t)$ describing the synaptic state (center), presynaptic spikes (top) and depolarization $V(t)$ of postsynaptic neuron (bottom). In the intervals between successive presynaptic pulses, $X(t)$ is driven toward one of the two stable states 0, 1 depending on whether $X(t)$ is below or above the threshold θ_X. At the times of presynaptic spikes, indicated by dotted vertical lines, $X(t)$ is pushed up ($X \to X+a$) if the depolarization $V(t)$ (circles) is above the threshold θ_V, and down ($X \to X-b$) if $V(t) < \theta_V$.

current. The choice of the linear integrate-and-fire (LIF) neuron, with a reflecting barrier at zero depolarization, is also a natural candidate for aVLSI implementation (Mead, 1989). Moreover, networks composed of such neurons have qualitatively the same collective behavior as networks of conventional RC IF neurons, which exhibit a wide variety of characteristics observed in cortical recordings (Fusi & Mattia, 1999).

2.1 Stable Efficacies and Stochastic Transitions. In the absence of stimulation, the pre- and postsynaptic neurons have low (spontaneous) rates, and typically the Hebbian term is zero. Hence the dynamics is dominated by the refresh term $R(t)$, which stabilizes the synaptic internal variable at one of the two extreme values, 0 and 1. In the interval $(0, 1)$, we have: If $X(t) > \theta_X (X(t) < \theta_X)$, then a positive current β (negative $-\alpha$) drives $X(t)$ up (down) until it stops at the reflecting barrier 1 (0) (see Figure 1). 1 and 0 are stable states, unless a large fluctuation drives the synaptic efficacy across (below from above, or above from below) the threshold θ_X.

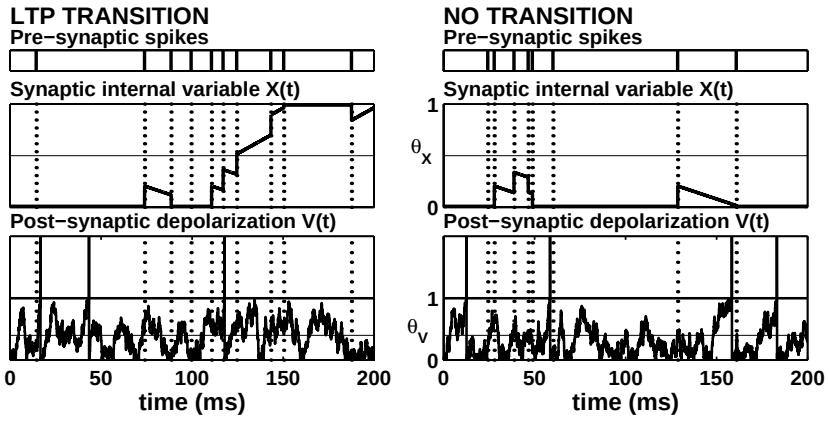


Figure 2: Stochastic LTP. Pre- and postsynaptic neurons have the same mean rate and the synapse starts from the same initial value ($X(0) = 0$). (Left) LTP is caused by a burst of presynaptic spikes that drives $X(t)$ above the synaptic threshold. (Right) At the end of stimulation, X returns to the initial value. At parity of activity, the final state is different in the two cases. See Figure 1 for a definition of the three windows.

To illustrate the stochastic nature of the learning mechanism, we assume that the presynaptic spike train is Poissonian, while the afferent current to the postsynaptic neuron is gaussian and uncorrelated with the presynaptic process.

The synaptic dynamics depends on the time statistics of the spike trains of the pre- and postsynaptic neurons. During stimulation, synapses experience two types of transitions between their two stable states:

- *Long-term potentiation*. Both pre- and postsynaptic neurons emit at a high rate. The frequent presynaptic spikes trigger many jumps in $X(t)$, mostly up ($X \to X + a$), since in order to emit at a high rate, the postsynaptic neuron spends much of the time near the emission threshold ($\theta = 1$), and hence above the threshold θ_V.
- *Long-term homosynaptic depression*. The presynaptic neuron emits at a high rate, while the postsynaptic neuron emits at a low (spontaneous) rate. The many jumps triggered are mostly down ($X \to X - b$), because the postsynaptic neuron is mostly at low depolarization levels.

During stimulation, the synapses move up and down. Following the removal of the stimulus, the synaptic efficacy may return to its initial state, or it may make a transition to another state. Figure 2 shows two cases at parity of spike rates of pre- and postsynaptic neurons. In one case (left) a fluctuation drives the synaptic efficacy above threshold; when the stimulus is removed,

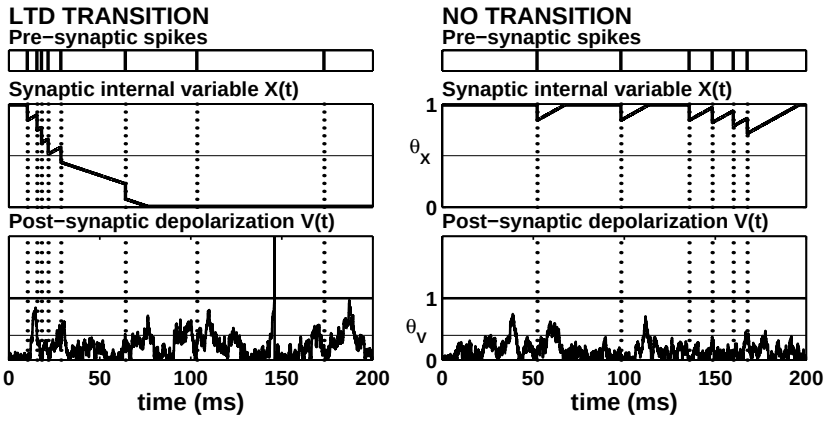


Figure 3: Stochastic LTD. As in the case of LTP, pre- and postsynaptic neurons have the same mean rate, and the synapse starts from $X(0) = 1$. (Left) A presynaptic burst provokes LTD. (Right) At the end of stimulation, X returns to the initial value. Conventions as in Figure 1.

X is attracted to the high state: LTP has occurred. In the second case (right), when the stimulus is removed, X is below threshold and is attracted by the refresh to the initial state. There is no transition. The stochasticity of LTD transitions is exemplified in Figure 3. At low presynaptic rates, there are essentially no transitions, irrespective of the postsynaptic rate.

The empirical transition probability is defined as the fraction of cases, out of a large number of repetitions of the same stimulation conditions, in which at the end of the stimulation, the synapse made a transition to a state different from its original state.

3 Estimating Expected Transition Probabilities

3.1 General Considerations. Given the synaptic parameters and the activity of the two neurons connected by the synapse during the stimulation, the transition probabilities contain most of the information relevant for the learning process. These probabilities depend on the following factors:

- *The duration of the stimulation*. Transition probabilities increase with presentation time T_p. At small transition probabilities, the dependence is linear.
- *The structure and the intensity of stimulation*. The presynaptic neuron acts as a trigger. At low firing rates, the probability of having a transition is negligible. The activity of the postsynaptic neuron is related to the statistics of the depolarization and, hence, the probability of having either a potentiating or a depressing jump.

- *Temporal patterns in the spike trains emitted by the two neurons.* Synchronization can affect the transition probability. We will consider such cases in the determination of the parameters of the VLSI synapse (see section 5).
- *The parameters governing the synaptic dynamics* (see equations 2.1–2.3). The two refresh currents (α, β), the jumps induced by the presynaptic spikes (a, b) and the threshold θ_X.

We study the dependence of LTP and LTD transition probabilities on the mean emission rates of the pre- and postsynaptic neurons, (ν_{pre} ν_{post}). The activities of the two neurons are read in different ways. For the presynaptic neuron, we take the probability per unit time that a spike is emitted, which is the mean rate. For the postsynaptic neuron, we use the probability that the depolarization is below (or above) the threshold when the presynaptic spike arrives. In order to determine this probability, we have to specify the framework in which the synaptic device operates. The appendix describes the scenario used. We point out in section 6 that this is not a restrictive feature of the framework.

3.2 Synaptic Dynamics as a Generalized Takács Process. We assume that presynaptic spikes are emitted in a Poisson process with rate ν_{pre}. The value (size and direction) of successive jumps is determined by the independently distributed random variables $V_{post}(t_k^{pre})$ and $V_{post}(t_{k+1}^{pre})$ of equation 2.3. The independence of these variables is a good approximation when ν_{pre} is not too high, because the mean interval between t_k^{pre} and t_{k+1}^{pre} is longer than the typical correlation time of the Wiener process driving the depolarization of the postsynaptic neuron. Yet the transition probabilities may be quite sensitive to these correlations, even if the correlation time of the Wiener process is relatively short. In fact, transitions are usually caused by bursts of presynaptic spikes, for which the time intervals between successive spikes are short, irrespective of the mean rate. Here, for simplicity, we ignore these correlations and assume that the random variables $V_{post}(t_k^{pre})$ are independent. This position would be analogous to the neglect of possible temporal correlations in the afferent current to an IF neuron (Ricciardi, 1977).

In this case, $X(t)$ is a Markov process, similar to the virtual waiting time in Takács single-server queue (e.g., see Cox & Miller, 1965). The process is somewhat generalized, to take into account that:

- There are two possible drifts, one up and one down, depending on where $X(t)$ is relative to its threshold.
- There are two reflecting barriers.
- Jumps can be down as well as up.

Those generalizations can be made, and we proceed to describe the resulting equations for the distribution.

The probability density function of $X(t)$ consists of two discrete probabilities, $P_0(t)$ that $X(t) = 0$ and $P_1(t)$ that $X(t) = 1$, and a density $p(X, t)$ for $X(t) \in (0, 1)$. The density $p(X, t)$ evolves according to two equations, one for each subinterval $]0, \theta_X[$ and $]\theta_X, 1[$. $p(\theta_X, t) = 0$ for any t, since θ_X is an absorbing barrier (see, e.g., Cox & Miller, 1975):

$$\frac{\partial p(X,t)}{\partial t} = \alpha \frac{\partial p(X,t)}{\partial^{(+)} X} + \nu[-p(X,t) + A(X,t)] \qquad \text{if } X \in]0, \theta_X[$$
$$\frac{\partial p(X,t)}{\partial t} = -\beta \frac{\partial p(X,t)}{\partial^{(-)} X} + \nu[-p(X,t) + B(X,t)] \qquad \text{if } X \in]\theta_X, 1[.$$

The first term on the right-hand side of each equation is the drift due to the respective refresh currents, $-\alpha$ and β. The terms proportional to ν are due to presynaptic spikes: $\nu p(X, t)$ is the fraction of processes leaving X at time t, per unit time, $\nu A(X, t)$ is the probability of arriving at X in the lower subinterval, and $\nu B(X, t)$ is the same probability in the upper subinterval. The expressions for A and B are, respectively,

$$\begin{aligned} A &= Q_a[P_0(t)\delta(X-a) + \Theta(X-a)p(X-a,t)] \\ &\quad + Q_b\Theta(-X+1-b)p(X+b,t) \\ B &= Q_a\Theta(X-a)p(X-a,t) + Q_b[P_1(t)\delta(X-(1-b)) \\ &\qquad + \Theta(-X+1-b)p(X+b,t)]. \end{aligned}$$

In A, the term in square brackets accounts for the $+a$ jumps, starting from $X = 0$ (first term) or from $X = X - a$ (second term). The last term accounts for b jumps (we assume that $b < 1 - \theta_X$). Q_b $(= 1 - Q_a)$ is the probability that the depolarization of the postsynaptic neuron is below the threshold θ_V; it is related to the firing rate of the postsynaptic neuron (see the appendix). The interpretation of B above θ_X is analogous.

The two discrete probabilities P_0 and P_1 are governed by:

$$\frac{dP_0(t)}{dt} = -\nu Q_a P_0(t) + \alpha p(0,t) + \nu Q_b \int_0^b p(x,t)dx$$
$$\frac{dP_1(t)}{dt} = -\nu Q_b P_1(t) + \beta p(1,t) + \nu Q_a \int_{1-a}^1 p(x,t)dx.$$

They describe processes that leave the stable states or arrive at them driven by spikes, the first and third term, respectively, on the right-hand side of each equation, and processes that arrive at each of the two stable states driven by the refresh drift, middle term in each equation. These equations are completed by the normalization condition, which restricts the process

to the interval [0, 1]:

$$P_0(t) + \int_0^1 p(x,t)dx + P_1(t) = 1.$$

3.3 LTP and LTD Transitions Probabilities. From the solution of the Takács equations, one derives the transition probabilities per presentation for LTP and LTD. To obtain the LTP probability, P_{LTP}, one assumes that the synapse starts at $X(0) = 0$, which corresponds to the initial condition $p(X, 0) = 0$, $P_1(0) = 0$ and $P_0(0) = 1$. If the stimulation time is T_p, we have:

$$P_{LTP} = \int_{\theta_X}^1 p(x, T_p)dx + P_1(T_p).$$

Similarly, for the depression probability per presentation, P_{LTD}, the initial condition is the high value—$p(X, 0) = 0$, $P_0(0) = 0$, $P_1(0) = 1$—and

$$P_{LTD} = \int_0^{\theta_X} p(x, T_p)dx + P_0(T_p).$$

3.4 LTP/LTD Probabilities: Theory.

3.4.1 The Determination of Q_a. In order to compute P_{LTP} and P_{LTD} as a function of the mean rates of two neurons connected by the synapse, we solve numerically the Takács equation for each pair of frequencies ν_{pre}, ν_{post}. To do this, one must calculate Q_a and Q_b, which depend on the distribution of the depolarization of the (postsynaptic) neuron, $p(v)$. In general, this distribution is not uniquely determined by the neuron's emission rate. For an IF neuron, for example, the rate is a function of two variables: the mean and the variance of the afferent current. For the linear IF neuron, given the mean and variance, one can compute explicitly $p(v)$ (Fusi & Mattia, 1999).

Hence, we use a stimulation protocol that allows a one-to-one correspondence between the output rate and the distribution of the depolarization. This is achieved by assuming that the afferent current to the neuron is a sum of a large number of independent Poisson spike processes (see the appendix). Figure 4 exhibits the characteristics of the resulting distribution. On the left is $p(v)$, for $0 < v \leq \theta$ for the relevant ν_{post} (between 1 and 60 Hz). For low spike rates, $p(v)$ is concave, and the probability for a depolarization near spike emission threshold $\theta(= 1)$, as well as for an upward jump (above $\theta_V(= 0.7)$), is low. As ν_{post} increases, the distribution becomes convex and increasingly uniform.

Given $p(v)$ at some value of ν_{post}, one computes $Q_a = \int_{\theta_V}^{\theta} p(v)$, which determines whether the synaptic jump is down or up. Figure 4 (right) shows Q_a as a function of ν_{post}. In conclusion, with the parameters α, β, a, b, θ_X, and Q_a, Q_b and fixing the stimulation time T_p the probabilities for LTP and LTD can be computed for any pair ν_{pre}, ν_{post}.

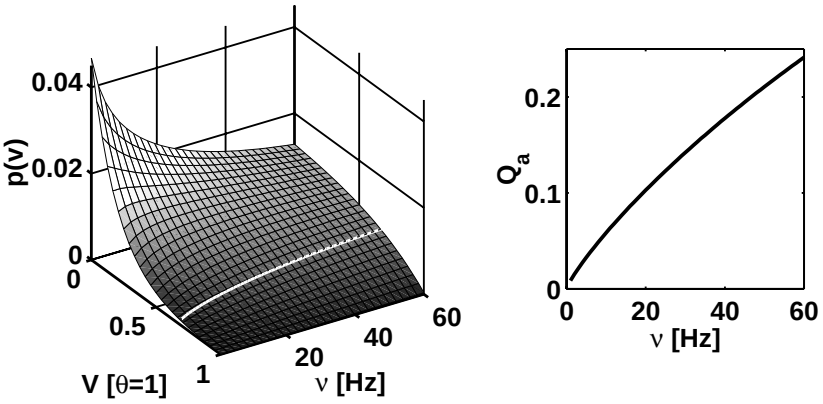


Figure 4: Distributions of the postsynaptic depolarization, $p(v)$ (left) and Q_a (probability for upward jump) (right) versus $\nu(=\nu_{post})$, in a stimulation procedure described in the appendix. For low spike rates, $p(v)$ is concave, and the probability for a depolarization near spike emission threshold $\theta(=1)$ is low. As ν_{post} increases, the distribution becomes convex and increasingly uniform. The white line on the surface is $p(v)$ at the upward jump threshold $\theta_V(=0.7)$. $Q_a = \int_{\theta_V}^{\theta} p(v)$.

3.4.2 Parameter Choices. The sets of parameters of the synaptic dynamics are the same as in the hardware tests (see section 5). $T_p = 500$ ms, as in a typical neurophysiological experiment. The rate of spontaneous activity has been set at 4 Hz and the typical stimulation rate at 50 Hz. We have chosen two sets of synaptic parameters to check that behavior can be manipulated (see Table 1). For each set we looked for a regime with the following

Table 1: Two Sets of Parameters of the Synaptic Dynamics.

	a	b	α	β
Synaptic parameters: Set 1				
DQ	0.150 ± 0.009	0.117 ± 0.012	0.00670 ± 0.00054	0.0184 ± 0.0018
PQ	498 ± 30 mV	395 ± 25 mV	22.1 ± 1.7 mV/ms	62.4 ± 6.0 mV/ms
Synaptic parameters: Set 2				
DQ	0.170 ± 0.007	0.138 ± 0.020	0.00654 ± 0.00035	0.0185 ± 0.002
PQ	562 ± 25 mV	457 ± 40 mV	21.6 ± 1.0 mV/ms	61.3 ± 6.5 mV/ms

Notes: The power supply voltage (the distance between the two stable synaptic efficacies) is 3.3 V. It sets the scale for the dimensionless units (DQ quantities) used in the extended Takács equations and determined by the measuring procedure. PQ are the physical quantities. The synaptic threshold X_θ is 0.374 (1236 mV) in both cases. The static regulations for the two refresh currents α and β are the same for the two sets (except for electronic noise).

characteristics:

- **Slow learning**. LTP at stimulation ($\nu_{pre} = \nu_{post} = 50$ Hz) is 0.03 for set 1 and 0.1 for set 2. LTD for neurons with anticorrelated activity ($\nu_{pre} = 50$ Hz, $\nu_{post} = 4$ Hz) is at 10^{-3} and 10^{-2}, respectively. This means that approximately 30 (10) repetitions of the same stimulus are required to learn it (Brunel et al., 1998). Smaller probabilities can be achieved for the same rates, but exhaustive hardware tests become too long.
- **LTP/LTD balancing**. We have aimed at a region in which the total number of LTP and LTD transitions are balanced in a presumed network. For random stimuli of coding level f, balancing implies $P_{LTD} \simeq f P_{LTP}$. In our case, $f = 0.03$ for set 1 and 0.1 for set 2.
- **Stability of memory**. In spontaneous activity the transition probabilities should be negligible.

In Figure 5 we plot the surfaces $P_{LTP}(\nu_{pre}, \nu_{post})$ and $P_{LTD}(\nu_{pre}, \nu_{post})$ for parameter set 1. The ν_{pre},ν_{post} space divides in three regions:

1. ν_{pre} and ν_{post} both high; P_{LTP} is much higher than P_{LTD}. The stronger the stimulation, the higher the transition probabilities. P_{LTP} drops sharply when one of the two rates goes below 10 Hz and becomes negligible around the spontaneous rates—about 5×10^{-8}.
2. ν_{pre} high, and ν_{post} around spontaneous levels. LTD dominates, even though stimulus-driven P_{LTD} probability at ($\nu_{pre} = 50$ Hz, $\nu_{post} = 4$ Hz) is a factor of 30 smaller than the stimulus-driven P_{LTP} ($\nu_{pre} = \nu_{post} = 50$ Hz). This guarantees the balance between potentiation and depression. P_{LTD} decays very rapidly in the direction of lower ν_{pre} and more slowly in the direction of increasing ν_{post}.
3. ν_{pre} low: Both LTP and LTD transition probabilities are small. This is the region in which the synapse remains unchanged on long timescales.

A network with such synapses should perform well as an associative memory (Amit & Fusi, 1994; Brunel et al., 1998).

Figure 6 is another representation of the phenomenology discussed above: the color map is the surface of the asymptotic probability P_{asy} of having a potentiated state following a large number of presentations of the same

Figure 6: *Facing page*. Asymptotic "learning rule" P_{asy} (see equation 3.1) versus (ν_{pre}, ν_{post}). If P_{LTP} and P_{LTD} are less than 10^{-3}, their value under stimulation P_{asy} is set to 0.5; there is no clear tendency. In the upper right corner (both neurons activated by stimulus), LTP is dominant; in the upper left corner, LTD dominates (homosynaptic LTD). The bright wedge (green) corresponds to equal probabilities. In the lower region, transition probabilities are negligible. The parameters are the same as in Figure 5.

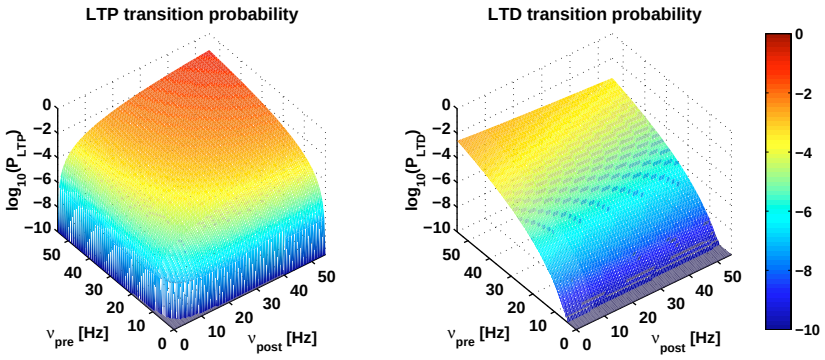
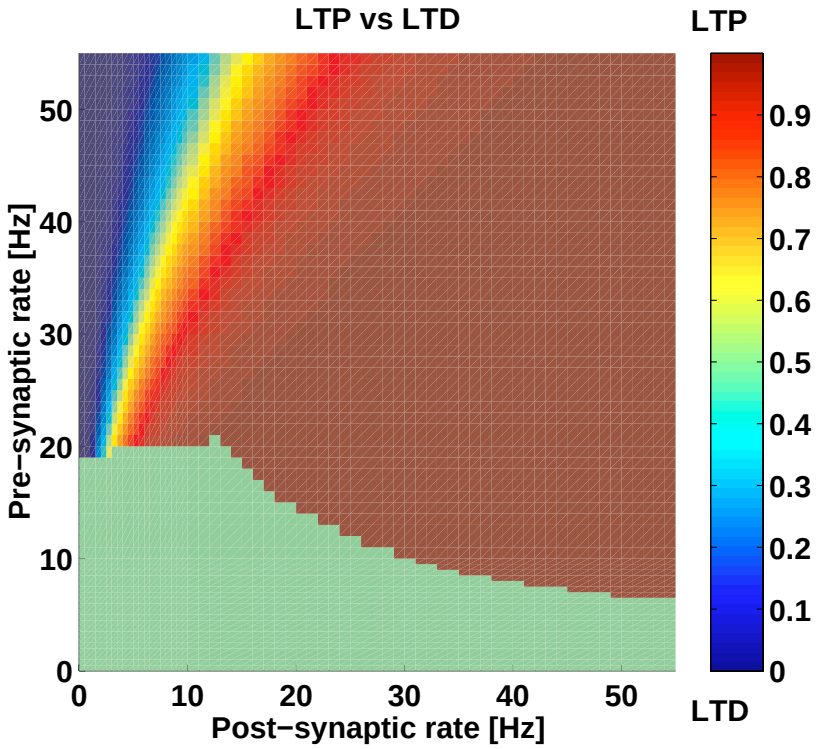


Figure 5: (Left) LTP transition probabilities (log scale). (Right) LTD versus pre- and postsynaptic neuron rates (parameter set 1, Table 1).



pattern of activities, given by:

$$P_{asy}(\nu_{pre}, \nu_{post}) = \frac{P_{LTP}(\nu_{pre}, \nu_{post})}{P_{LTP}(\nu_{pre}, \nu_{post}) + P_{LTD}(\nu_{pre}, \nu_{post})}. \tag{3.1}$$

It is near 1 if the synapse tends to be potentiated, near 0 if it tends to be depressed, and around 0.5 if the two transition probabilities are about equal. When both transition probabilities are 10^{-3} times their value in the presence of stimulation ($P_{LTP} < 10^{-4}$, $P_{LTD} < 10^{-6}$), we set the surface height to 0.5, to indicate that there is no clear tendency.

Figure 6 illustrates the way the synapse encodes activity information, or the "learning rule." For each pair of rates (ν_{pre}, ν_{post}), P_{asy} determines whether the synapse will be potentiated or depressed. If the stimulus activates the two neurons connected by the synapse, the synaptic efficacy tends to be potentiated. If the presynaptic neuron is firing at high rates and the postsynaptic neuron has low activity (below 5–20 Hz, depending on the presynaptic neuron activity), the synapse tends to be depressed.

Finally, we note that the LTP probability, which sets the upper bound on the stability of memory, is 5×10^{-8} (5×10^{-7} for parameter set 2). In other words, the mean time one has to wait until the synapse makes a spontaneous transition and loses its memory is approximately 100 days for parameter set 1 and 10 days for parameter set 2. This time depends on the typical activation rates, the spontaneous rates, and the time constants of the synaptic dynamics. For parameter set 2, it is relatively short because the probabilities are high. Still it is 10^7 times longer than the longest time constant of the synaptic dynamics. For lower spontaneous rates, one can achieve mean times of years. For example, with a spontaneous rate of 2 Hz, it is 4 years for parameter set 1, and 1.2 years for set 2.

4 Hardware Implementation

The mechanism has been implemented in aVLSI (Badoni, Salamon, Salina, & Fusi, 2000) using standard analog CMOS 1.2μ technology. The synapse occupies about $90\mu \times 70\mu$.

The schematic design of the synapse is depicted in Figure 7. The pre- and postsynaptic neurons are simulated in this study. In the network implementation, they are a modified version of Mead's neuron (Mead, 1989; Fusi, Del Giudice, & Amit, in press; Badoni et al., 2000). There are four functional blocks.

4.1 Capacitor C. The **capacitor** C acts as an analog memory element: the amount of charge on the capacitor is directly related to the internal synaptic state $X(t)$:

$$X(t) = \frac{U(t) - U_{\min}}{U_{\max} - U_{\min}},$$

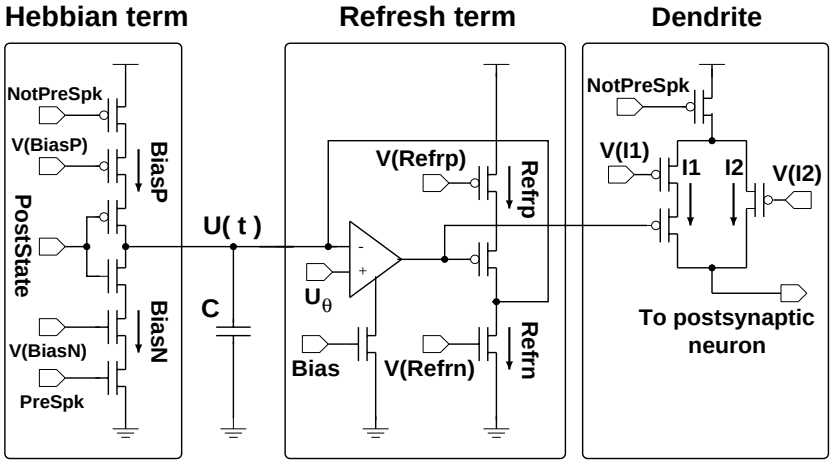


Figure 7: Schematic of the synaptic device. The voltage across the capacitor C is the internal variable $X(t)$. The refresh block implements $R(t)$ and provides the refresh currents. The activity-dependent Hebbian block implements $H(t)$. The dendrite block produces the synaptic transmission when a presynaptic spike arrives (the devices corresponding to the pre- and postsynaptic neurons are not shown, they are simulated). See the text for a description of each block.

where $U(t)$ is the voltage across the synaptic capacitor and can vary in the band delimited by $U_{\min} \sim 0$ and $U_{\max} \sim V_{dd}$ (the power supply voltage).

4.2 Hebbian Block. The **Hebbian block** implements $H(t)$, which is activated only on the arrival of a presynaptic spike. The synaptic internal state jumps up or down, depending on the postsynaptic depolarization. The input signal PostState determines the direction of the jump and is the outcome of the comparison between two voltage levels: the depolarization of the postsynaptic neuron $V(t)$ and the threshold θ_V. If the depolarization is above the threshold, then PostState is ~ 0; otherwise it is near the power supply voltage V_{dd}.

In the absence of presynaptic spikes, the two MOSFETs controlled by NotPreSpk and PreSpk are open (not conducting), and no current flows in the circuit. During the emission of a presynaptic spike, both of these MOSFETs are closed, and the PostState signal controls which branch is activated. If PostState $= 0$ (the postsynaptic depolarization is above the threshold θ_V), the three topmost MOSFETs of the upper branch are closed, and the current flows from V_{dd} to the capacitor. The charging rate is determined by BiasP, which is the current flowing through the upper MOSFET and determined by $V(BiasP)$ (this MOSFET is the slave of a current mirror). Analogously,

when PostState = 1, the lower branch is activated, and the current flows from the capacitor to the ground.

Hence, the jumps a and b (see equation 2.3) are given by

$$a = \frac{BiasP\ \Delta t}{C(U_{\max} - U_{\min})}, \qquad b = \frac{BiasN\ \Delta t}{C(U_{\max} - U_{\min})},$$

where Δt is the duration of the presynaptic spike.

4.3 Refresh Block. The **refresh block** implements $R(t)$, which charges the capacitor if the voltage $U(t)$ is above the threshold U_θ (corresponding to θ_X) and discharges it otherwise. It is the term that tends to damp any small fluctuation that drives $U(t)$ away from one of the two stable states.

If the voltage across the capacitor is below the threshold U_θ ($U(t) < U_\theta$), the voltage output of the comparator is $\sim V_{dd}$, the pMOS controlled by $V(Refrp)$ is not conducting, while the nMOS controlled by $V(Refrn)$ is closed. The result is a current flow from the capacitor to the ground.

If $U(t) > U_\theta$, the comparison produces the minimal output voltage as the output of the comparator, the lower pMOS is closed, and the current flows from V_{dd}, through the upper pMOS, to the capacitor. Since the nMOS is always closed, part of the current coming from V_{dd} flows to the ground. The net effect is that the capacitor is charged at a rate proportional to the difference between the current determined by $V(Refrp)$ (Refrp) and the current flowing through the lower MOS (Refrn) (see Figure 7).

The two refresh currents α and β of equation 2.2 are, respectively:

$$\alpha = \frac{Refrn}{C(U_{\max} - U_{\min})}, \qquad \beta = \frac{Refrp - Refrn}{C(U_{\max} - U_{\min})}.$$

See Badoni et al. (2000).

4.4 Dendrite Block. The **dendrite block** implements the synaptic efficacy. The postsynaptic potential evoked by a presynaptic spike (EPSP) depends on the synaptic internal state $X(t)$. In our implementation, if $X(t)$ is below the threshold X_θ, the EPSP is J_0; otherwise, it is J_1.

This mechanism is implemented as follows: The current is injected only upon the arrival of a presynaptic spike (the topmost MOSFET acts as a switch that is closed while the presynaptic neuron is emitting a spike). The output of the refresh block comparator determines the amount of current injected in the postsynaptic neuron. The lower MOSFET in the left branch acts as a switch whose state depends on the outcome of the comparison performed by the refresh term between $X(t)$ and θ_X. If $X(t) > \theta_X$, the switch is closed, and the current injected is the sum of the two currents I_1 and I_2. Otherwise, only the MOSFET controlled by V_{I2} is conducting, and the total current is I_2. Hence the EPSPs J_0 and J_1 can be regulated through the control

currents I_1 and I_2:

$$J_0 = I_2 \Delta t / C_n$$
$$J_1 = (I_1 + I_2) \Delta t / C_n, \tag{4.1}$$

where C_n is the capacity of the postsynaptic neuron and Δt is the spike duration.

5 Testing the Electronic Synapse

5.1 The Testing Setup. We present measured LTP and LTD probabilities for the electronic synapse, for two sets of parameters of the synaptic dynamics, as a function of the spike rates of the pre- and postsynaptic neurons. The main objectives are to test whether the behavior of the electronic synapse is well described by the model of section 2 and to check whether the desired low LTP and LTD transition probabilities and the balance between them can be achieved and manipulated with reasonable control currents.

We use the same setup for measuring the synaptic parameters of the chip as for the measurements of the transition probabilities. The only difference is the choice of special stimulation protocols for the parameter determination, much as in neurobiology (Markram, Lubke, Frotscher, & Sakmann, 1997).

5.2 Measuring the Synaptic Parameters. In the parameter measurement protocol, the presynaptic neuron emits a regular train of spikes, while the postsynaptic depolarization is kept fixed (for LTP, it is constantly above the threshold θ_V). If the interspike interval of the regular train is T, then the arrival of a presynaptic spike (at $t = 0$) generates a triangular wave in $X(t)$:

$$X(t) = X(0) + a - \alpha t.$$

We assume that $a < \theta_X$. If the rate is low ($T > a/\alpha$), the synaptic internal state is reset to 0 by the refresh current α, before the arrival of the next spike. If $T < a/\alpha$, the effects of successive spikes accumulate and eventually drive the synaptic potential across the synaptic threshold, θ_X, to the potentiated state. The transition occurs when

$$na - (n-1)T\alpha \geq \theta_X, \tag{5.1}$$

where n is the minimal number of presynaptic spikes needed to make the synaptic potential cross the synaptic threshold θ_X.

If one increases progressively the period T, at some point the minimal number of pulses n will change, and one extra spike will be needed in order to cross the threshold. At this point the relation (5.1) becomes an equality:

$$na - (n-1)T\alpha = \theta_X. \tag{5.2}$$

By increasing or decreasing the period T, we can collect a set of such special points. For each special period T_k, we have a corresponding n_k: for $T < T_k$, n_k presynaptic pulses are needed to cross the threshold, while for $T > T_k$ the minimal number of spikes is $n_k + 1$. Actually two of these special points, for example, n_i and n_j, determine α and a:

$$\alpha = \theta_X \frac{n_i - n_j}{n_i(n_j - 1)T_j - n_j(n_i - 1)T_i},$$
$$a = \theta_X \frac{(n_j - 1)T_j - (n_i - 1)T_i}{n_i(n_j - 1)T_j - n_j(n_i - 1)T_i}. \tag{5.3}$$

An analogous procedure would provide an estimate of β and b.

The synapse is initially set to $X = 0$. A high-frequency train of pulses is generated for the presynaptic neuron: $T = T_{\min} = 11$ ms for a and α and $T_{\min} = 2$ ms for b and β. The difference is due to the fact that at parity of frequency, LTD requires more spikes to cross the threshold (for example, for set 1, $\theta_X = 2.49a$, while $(1 - \theta_X) = 5.35b$).

As soon as the synaptic voltage $X(t)$ crosses the threshold θ_X, the train is stopped and the number of spikes emitted by the presynaptic neuron since the beginning of the trial is stored in $n(T)$. Then T is increased (typically $\Delta T = 0.1$–0.2 ms) and the procedure is repeated, to obtain $n(T)$ for a given range of periods. The discontinuities of the curve are the special points (n_i, T_i). a_{ij} and α_{ij} are determined for each pair (i, j) of special points, using equations 5.3. The final estimate is given by the average of these values. The standard deviation provides an estimate of the fluctuations due to electronic noise.

The measured parameters and the corresponding standard deviations are reported in Table 1, in their dimensionless as well as physical values. The parameters entering the extended Takàcs equations are dimensionless. The corresponding physical quantities are obtained using the actual voltages for the threshold θ_X and the distance between the two stable synaptic efficacies equal to the power supply voltage. The parameters have been adjusted to satisfy the requirements listed in section 3. Note that the errors are larger for LTD parameters. This is due to the fact that more jumps are required to get to threshold, and the electronic noise accumulates as the square root of the number of jumps.

5.3 Measuring the Transition Probabilities. Poissonian spike trains are generated for the presynaptic neuron by the host computer. The probability of the distribution of depolarization of the LIF neuron is used to generate the up (down) jump probability Q_a (Q_b) for the postsynaptic neuron to be above (below) the threshold θ_V. (See the appendix.) For each presynaptic spike, an independent binary number ξ is generated: $\xi = 1$ with probability Q_a, 0 with probability Q_b (=$1 - Q_a$). This number determines the direction of the jump induced.

The synapse is set to one of its two stable values, and then the two neurons are stimulated (fixing their emission rates) for $T_p = 500$ ms. Following the stimulation, the host PC reads the resulting stable value of the synapse and establishes whether a transition has occurred. This procedure is repeated from 10^4 to 2×10^5 times (depending on the transition probability), and the transition probability is estimated by the relative frequency of occurrence of a transition.

The testing time is reduced by discarding trials in which the internal variable $X(t)$ is likely to remain throughout the trial, too far from the threshold to induce a transition. In order to identify these trials, we simulate the dynamics described by the ordinary differential equations of section 2 and take only the trials in which the simulated internal variable enters in a band around the threshold ($X(t) > X_\theta - 1.5a$ for LTP, and $X(t) < X_\theta + 1.5b$ for LTD). In the other trials, the electronic synapse is likely to remain unchanged, even if the parameters have not been estimated accurately or the simulated dynamics does not reproduce exactly the real dynamics. Hence they are safely counted as no-transition trials.

The upper and lower bounds of the confidence interval of the estimated probabilities are given by Meyer (1965):

$$P_{low/high} = \frac{Pn + \frac{k^2}{2} \mp k\left[P(1-P)n + \frac{k^2}{4}\right]^{1/2}}{n + k^2}, \tag{5.4}$$

where n is the total number of repetitions, P is the estimated probability, and k is the width of the confidence interval in sigmas. In our case, $k = 3$.

5.4 Results. The experimental transition probabilities are compared with those predicted by the theory of section 3 and with those measured in a computer simulation of the protocol used for the hardware test, for the two sets of parameters of Table 1. If the model of section 2 is adequate for describing the behavior of the electronic synapse, the computer simulation values should be in agreement with the measured experimental values. This question becomes more poignant given the uncertainties on the measured parameters, as expressed in Table 1. In what follows we use the central estimate for the parameters. On the other hand, if the theoretical approach, as expressed in the extended Takàcs equations, is a valid one, the numerical integration of the equations should give a good asymptotic estimate for both the experimental values and the simulated ones (for a large number of presentations).

Two sets of parameters have been chosen to show that beyond the quality of the model and of its solutions, the performance of the synaptic system can be manipulated rather effectively, despite the strict requirements imposed—that is, small, balanced transition probabilities for sparse coding stimuli. Note that the time constants of the refresh currents are the same for the two sets: $1/\alpha \simeq 150$ ms, and $1/\beta \simeq 50$ ms.

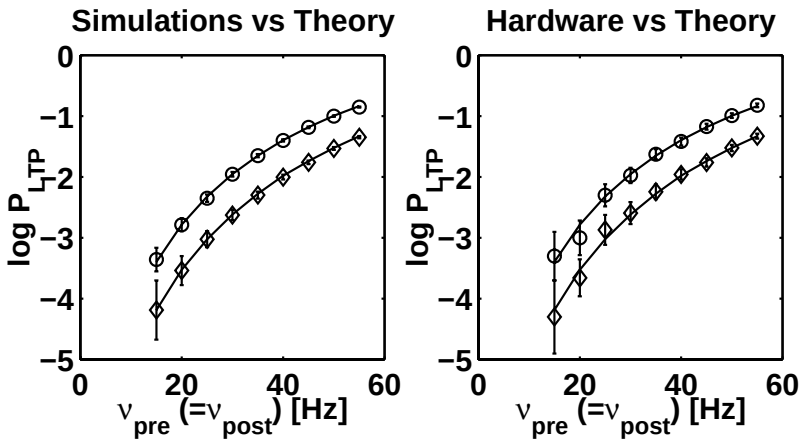


Figure 8: LTP probabilities versus presynaptic (= postsynaptic) rate, in a three-way comparison, for the two sets of parameters of Table 1. Lower curves and points: set 1; upper: set 2. (Left) Numerical solution of extended Takács equations (solid curves) and measurements in a simulation (diamonds and circles) (10^5 repetitions for the upper curves and 2×10^5 for the lower ones). (Right) Theoretical prediction (solid curves, as on left) and transition probabilities for the hardware device (from 10^4 to 4×10^4 presentations) (diamonds and circles). Error bars: confidence intervals (see equation 5.4).

The results are reported in Figures 8 and 9 for LTP and LTD, respectively. Each figure is a three-way comparison for two sets of parameters (see Table 1). Upper curves and points are for set 2; the lower ones are for set 1. The graphs on the left compare theory (extended Takács equations) and simulations; the ones on the right compare the same theory and probability measurements on the electronic device. To reduce the two-dimensional rate space to one, the probabilities for LTP are given for a typical stimulation situation in which both neurons have the same rate. The higher the rate is, the stronger is the stimulation. Instead, for LTD, the postsynaptic rate is set at a spontaneous level, 4 Hz, and the presynaptic rate is varied.

The theoretical predictions are almost identical to the simulations results, which implies that the numerical solution of the extended Takàcs equations provides a reliable estimate of the modeled dynamics, even for rather small probabilities (a range of five orders of magnitude). The theoretical predictions are also in good agreement with the measured experimental values. This is an indication that the device works as expected and is not sensitive to the uncertainties in the parameters.

Finally, Figures 10 and 11 present a comparison of the measured (chip) and theoretical (extended Takàcs) LTP and LTD transition probabilities (re-

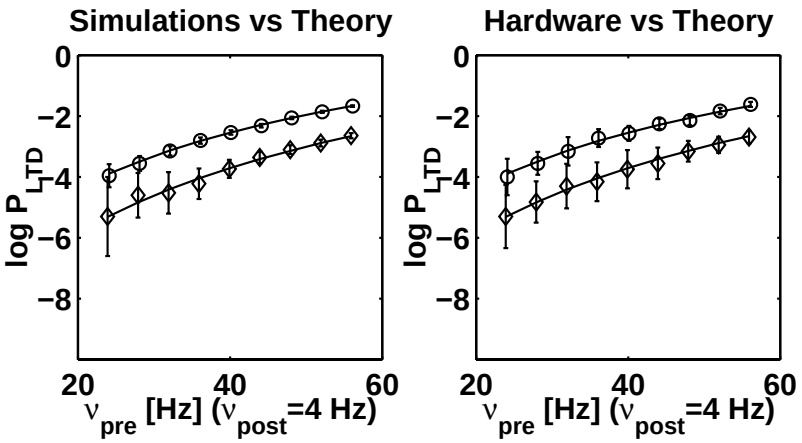


Figure 9: LTD probabilities versus presynaptic rate (postsynaptic rate is 4 Hz) in a three-way comparison theory-simulation-experiment, as in Figure 8.

spectively), over a two-dimensional (pre- and postsynaptic) rate space for parameter set 2. Each point of the surfaces represents the transition probability for a given pair of pre- and postsynaptic rates. It is coded by both color and elevation. The average percentage discrepancy between the measured and the predicted transition probabilities is about 11% for LTP and about 14% for LTD (the discrepancy between the measured value P_m and the prediction P_p is defined as $2|P_m - P_p|/(P_m + P_p)$). The fact that the relative discrepancies are larger for LTD is related to the fact that the uncertainties in the parameters are larger and that these probabilities are at least 10 times smaller.

6 Discussion

6.1 Balance Sheet. We have carried out an extensive test (about 1.5×10^6 presentations in all) in which most of the plane of the pre- and postsynaptic activities is explored. The agreement demonstrates that the dynamics of the synapse is well described by the model of section 2, and the simple synapse is a reliable, manipulable device, integrable on large scale.

6.2 The Scope of the Synaptic Dynamics. The particular realization of a plastic, synaptic device resolves the tension between the need to learn from every stimulus (fast) and the need to preserve memory on long timescales—to learn and forget slowly. The solution is conceptually simple and is implementable in aVLSI. The result satisfies the expectations quite well. The synapse has two dynamical regimes: an analog, volatile regime, in which

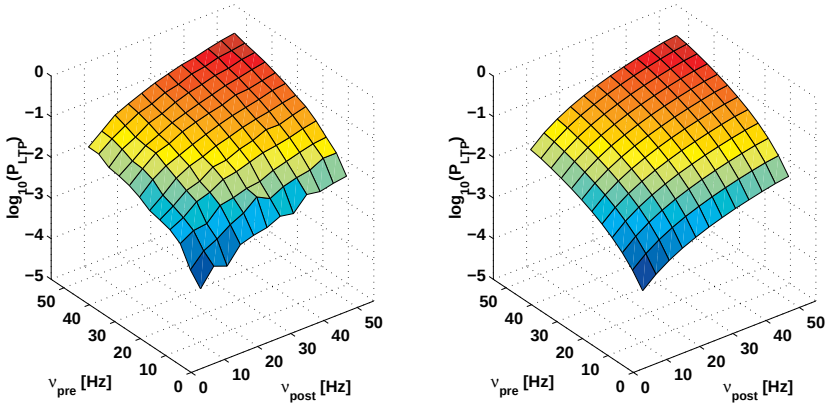


Figure 10: LTP transition probability surfaces (log scale). (Left) Experiment. (Right) Theory. The code is of both elevation and color. Number of repetitions for each point is 10^5. Average relative discrepancy is 11%.

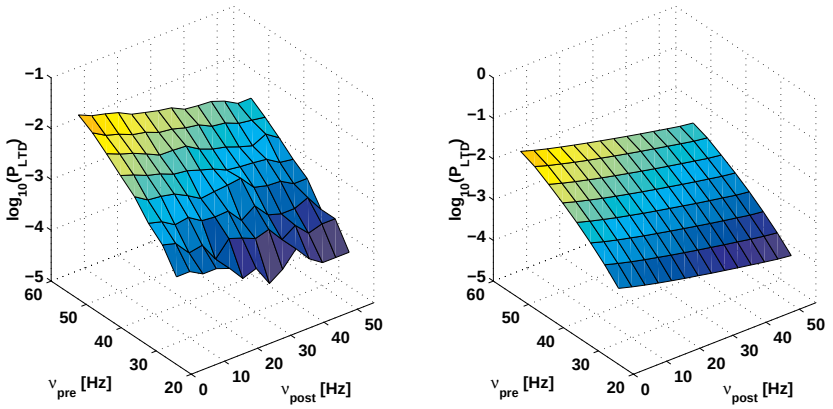


Figure 11: LTD transition probability surfaces. (Left) Experiment. (Right) Theory. As in Figure 10. Number of repetitions: 2×10^4 to 4×10^4.

it is modified little and the modifications decay in a relatively short time; and a discrete, binary efficacy, which is modified in a stochastic way and can preserve memory for very long times.

The synapse and the analysis presented in this study are a particular, though very rich, case of a wide range of situations that can, and should, be selected by confrontation with experiment and/or with computational

desiderata. The particular choices made were of two types: convenience of electronic implementation and amenability to detailed analysis. The first choices have suggested the dynamics of the synapse. The second have suggested the dynamics of the neurons connected by the synapse.

The synapse may have a different refresh mechanism (such as an exponential, rather than linear, decay). It may also use a different postsynaptic variable than the postsynaptic depolarization. For the mechanism to work, it must be related to a variable that smoothes the very brief instances in which postsynaptic spikes are emitted. An alternative to the depolarization could be some quantity that measures the time interval between spikes emitted by the pre- and the postsynaptic neurons. It would require some extra internal degree of freedom and hence some extra components in the chip, but the framework will still hold. The extended Takàcs description covers a wide range of cases and can go unmodified.

Simplifying assumptions were also made for the stimulation mechanism, mainly the fact that the neurons are LIF neurons and that the jumps, depending on the postsynaptic depolarization, are independently distributed. Both can be relaxed, at the price of losing the analytic tool of the Markov process. But simulation and experiment can be performed.

We reemphasize that the context in which the study has been conceived is similar to that of the IF neuron (Ricciardi, 1977; Fusi & Mattia, 1999), in which a motivated, simplifying choice is made for the neural dynamics and a simple, though generic, choice is made for the afferent current. Such an approach allows for a detailed study and for the development of intuitions and the generation of new questions. The simplified models can then be embedded in networks in which the assumptions may not be precisely satisfied and yet function satisfactorily. For the synapse, such a paradigmatic model has not been presented before.

6.3 Requisites at the Level of a Network. The synapse presented here is the result of an effort to extend the study of the properties of recurrent neural networks to include on-line learning (see also Badoni et al., 1995; Del Giudice, Fusi, Badoni, Dante, & Amit, 1998; Amit, Del Giudice, & Fusi, 1999). Sixty of these synapses have been embedded in a pilot aVLSI network of 21 (14 excitatory and 7 inhibitory; only excitatory-excitatory connections are plastic) neurons (Badoni et al., 2000; Fusi et al., in press), as well as in large-scale simulation experiments (Mattia & Del Giudice, in press).

The general framework developed for learning networks has convinced us that synapses have discrete and bounded values for their efficacies. This is suggested by considerations of implementability (see, e.g., Baldoni et al., 2000), as well as by considerations of performance as an associative memory. The latter are related to the fact that it should be a rather general requirement that an open learning memory have the palimpsest property than the "overload catastrophe." But for a network with discrete synapses to

perform well and have large enough memory, learning should be stochastic and slow. This puts a heavy load on the synapses, because implementing low transition probabilities for individual synapses requires cumbersome devices.

6.4 The Present Synapse in the Context of a Learning Network. The scheme has the very attractive (and novel) feature that it transfers the load of generating low-probability synaptic transitions to the collective dynamics of the network. In other words, not only memory is distributed but also functional noise.

The source of stochasticity is in the spike emission processes of the neurons, and small transition probabilities can be easily achieved because the LTP and LTD transitions are based on the (approximate) coincidence of events that are relatively rare (fluctuations in the presynaptic spike train and in the postsynaptic depolarization). These rare events are the result of the collective dynamics of the network, in which the pre- and postsynaptic neurons are embedded. Indeed, the time constants of the single neurons could not account for the long mean interspike intervals observed in cortical recordings for spontaneous activity, for example.

The statistics of recorded trains of spikes are good enough for the dynamics of our synapse (Softky & Koch, 1992). Networks composed of IF neurons show similar features (Amit & Brunel, 1997; van Vreeswijk & Sompolinsky, 1997; Fusi & Mattia, 1999). The next step will be to know whether the spikes generated by the simulated networks can drive the synaptic dynamics as expected. Preliminary results (Mattia & Del Giudice, in press) indicate that networks of LIF neurons, when stimulated, produce stochastic transitions. In those simulations, it is already clear that the learning process can be described as a random walk between the two stable states of the synapses, as in Brunel et al. (1998).

The behavior of the network can also be modified by setting the parameters appropriately. Varying the transition probabilities, it is possible to control the learning rate (in terms of repetitions of the same stimulus) and, consequently, the size of the sliding window of the memory span. This means that it is possible to have fast single-shot learning, with a correspondingly short memory, as well as slow learning, using the same device. In fact, one simple handle is the rates provoked by the stimuli.

6.5 The aVLSI Synapse and Biology. The synaptic device has been tested for biological timescales; that is, the rates are similar to those observed in associative cortex and the typical presentation time is 500 ms. The two parameters that determine the timescale of the synaptic dynamics are the two refresh currents, α and β. Varying these currents, one can obtain similar behavior on a wide range of timescales, presumably as fast as 10^3 times the biological scale (requiring a much faster data acquisition and stimulation system). It is tempting to speculate that one of the reasons for the

brain to operate at low speeds is also the lack of fast enough stimulation and acquisition systems.

Given that the ideas underlying the hardware implementation are rather general, it is natural to ask about their biological relevance. The question is not about a mapping onto a series of biochemical reactions that ultimately trigger the expression of synaptic change. The biological plausibility is related to observable consequences of spike activity—that is, to the experimental protocols that induce an activity-dependent LTP and LTD.

6.5.1 Discreteness of Synapses. It is known that the discreteness of the stable synaptic states is compatible with experimental data (see, e.g., Bliss & Collingridge, 1993). Recently Petersen et al. (1998) provided experimental evidence that long-term potentiation is all-or-none, meaning that for each synapse, only two efficacies can be preserved on long timescales. Models of networks with this type of synapses have been studied in detail (Amit & Fusi, 1994; Brunel et al., 1998).

6.5.2 Synaptic Dynamics. It appears that at low presynaptic rates, the synaptic changes are temporary, and after stimulation, the synaptic efficacy returns to the initial values (Markram et al., 1997). Sudden transitions to the potentiated or depressed state are observed when the rate is increased. This would be a nice corroboration of our scenario, in which a change in the internal variable that does not make X cross the threshold is damped.

Recent experiments (Markram et al., 1997; Zhang, Tao, Holt, Harris, & Poo, 1998; Bi & Poo, 1999) indicate that induction of LTP and LTD requires "precise timing" of presynaptic spikes and postsynaptic action potentials. LTP requires that presynaptic spikes precede postsynaptic spikes by no more than about 10 ms, and LTD that presynaptic spikes follow the postsynaptic spike within a short time window.

Our synapse is not fully compatible with these results, since it is the postsynaptic depolarization that determines whether the synapse is potentiated or depressed, rather than "precise timing." What is consistent is that whenever a presynaptic spike precedes a postsynaptic action potential, the depolarization is likely to be near the emission threshold, and therefore above the threshold θ_V, and LTP is likely to be induced. When a presynaptic spike occurs just after the emission of a postsynaptic action potential, that neuron is likely to be hyperpolarized, and the depolarization will tend to be below θ_V. This produces LTD. Where it fails is when the presynaptic spike precedes by more than 10 ms the postsynaptic one. Our synapse may still potentiate or depress; the biological one seems not to.

There may be different reasons for the remaining discrepancy. The depolarization of simple IF neurons may not directly map on the depolarization of a complex biological neuron. There may instead be some internal variable that accounts for other internal states and is sensitive to particular time intervals, before and after the postsynaptic spike.

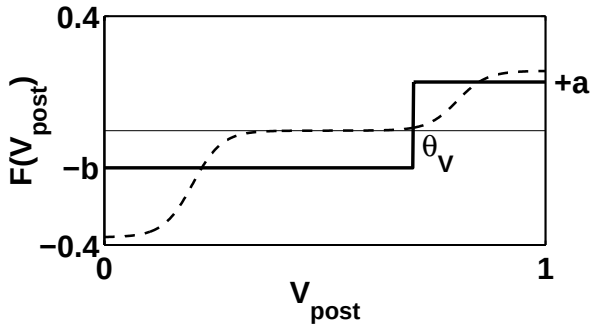


Figure 12: Alternative forms for the dependence of the analog synaptic jump (provoked by a presynaptic spike) F upon the postsynaptic depolarization, V_{post}. Full curve: (step function) is implemented in the present synapse—constant positive jumps for high depolarization, negative for low V_{post}. Dashed curve: an alternative, with a transition region of very small jumps (unable to produce any permanent modification)—produces a behavior similar to the "precise timing" plasticity observed in the experiments.

Alternatively, a simple modification of the synaptic dynamics would allow reproducing the experimental results of Markram et al. (1997). For example, it is sufficient to extend the function F, expressing the value of the synaptic jump (see equation 2.4), upon the arrival of a presynaptic spike. The form used in our chip is depicted by the step function (full curve) in Figure 12. It takes a negative value $(-b)$ below some θ_V and a positive value (a) above. The second example in the figure depicts a case of soft transition: jump up for V_{post} near θ; down for V_{post} very low; in a region of intermediate V_{post}, the size of the jump is zero or very small. If the two regions in which the synaptic jumps are significant are concentrated near enough to the depolarization boundaries, the experimental results can be reproduced quantitatively. In other words, given that the experiment is carried out in a deterministic situation (Markram et al., 1997), there will be a sharp cutoff for LTP and LTD in the time difference of the arrival of the pre- and postsynaptic spikes.

But a synaptic mechanism for which both LTP and LTD have a sharp cutoff in the time difference between the arrival of pre- and postsynaptic spikes ("precise timing") is rather insensitive to input spike rates (Abbott & Song, 1998; Senn, Tsodyks, & Markram, in press), and it is not easy to encode information expressed in mean emission rates. Our synapse will operate in a regime based on time coincidence: although it is conceived to operate in an asynchronous regime (described in section 5), it is rather sensitive to spike synchronization.

Appendix: Depolarization and Firing Rate

The probability Q_b ($= 1 - Q_a$) that appears in the extended Takács equations of section 3 is the probability that the depolarization of the postsynaptic neuron is below the threshold θ_V. It is indirectly related to the firing rate of the postsynaptic neuron, and the relation depends on the neural dynamics and the structure of the input current.

A.1 Neural Dynamics. We chose a linear IF neuron, which integrates an afferent current and has a reflecting barrier at 0 (Fusi & Mattia, 1999). When the depolarization V crosses a threshold θ, a spike is emitted, and the depolarization is reset to 0 and forced to stay there during an absolute refractory period τ_{arp}. A constant negative current μ_n produces a linear decay in the absence of the afferent current $I(t)$. In the interval between the reflecting barrier and the emission threshold θ, the dynamics of V is:

$$\frac{dV}{dt} = I(t) - \mu_n.$$

Given the probability distribution function, $p_n(V)$, of the depolarization V, the probability Q_b is obtained as

$$Q_b = \int_{-\infty}^{\theta_V} p_n(v)dv.$$

For neurons with a reflecting barrier at 0, the lower bound in the above integral is 0. The probability distribution $p_n(V)$ can be computed analytically in terms of the mean μ and the standard deviation σ, per unit time, of the afferent current I. It is given by

$$p_n(V) = \frac{\nu(\mu,\sigma)}{\mu}\left[1 - \exp\left(-2\frac{\mu}{\sigma^2}(\theta - V)\right)\right],$$

where $\nu(\mu,\sigma)$ is the emission rate of the neuron for these μ and σ, given below. Note that the linear decay current μ_n has been absorbed in μ.

The emission rate of the neuron can also be expressed in terms of μ and σ:

$$\nu(\mu,\sigma) = \left[\tau_{arp} + \frac{\sigma^2}{2\mu^2}\left(\frac{2\mu\theta}{\sigma^2} - 1 + e^{\frac{-2\mu\theta}{\sigma^2}}\right)\right]^{-1}. \tag{A.1}$$

Thus, given the neural parameters ($\tau_{arp} = 2$ ms), if μ and σ^2 are linear functions of the stimulating spike rate, we have a unique relation between the rate of the postsynaptic neuron and the distribution of the depolarization.

A.2 Afferent Current. The current is produced by a large number of independently spiking excitatory and inhibitory neurons: $c_E = 240$ afferents

from excitatory neurons and $c_I = 60$ afferents from inhibitory neurons. The rates of the inhibitory neurons and 120 of the excitatory ones are fixed. They help take care of the level of spontaneous activity. $c_E^{ext} = 120$ excitatory afferents carry stimulation activity. The mean synaptic efficacies are chosen so that the neurons have spontaneous activity $\nu_E^0 = 4$ Hz when all 240 afferent excitatory neurons have that same rate and the 60 inhibitory neurons emit at $\nu_I^0 = 8$ Hz: $J_{E\to E} = 0.05\theta$ (equal for all excitatory synapses), $J_{I\to E} = 0.33\theta$. The afferent current to each neuron is well approximated by a gaussian, and its mean and variance per unit time can be estimated by using mean-field theory (as in Amit & Brunel, 1997):

$$\mu = -110.4 - \mu_n + 6\nu_{ext} \tag{A.2}$$

$$\sigma^2 = 68.34 + 0.37\nu_{ext}, \tag{A.3}$$

where the coefficients in the expression for σ^2 include a variability in the connectivity of 25%. The linear decay current μ_n is 27θ Hz.

Acknowledgments

We thank V. Dante for the first version of the DAC board for the static regulations; M. Mattia for the parameters of the linear IF neuron network; and P. Del Giudice for many useful discussions. This work has been supported in part by a grant from the Israel Fund for Basic Research.

References

Abbott, L. F., & Song, S. (1998). Temporally asymmetric hebbian learning, spike timing and neuronal response variability. In M. S. Kearns, S. Solla, & D. A. Cohn (Eds.), *Advances in neural information processing systems, 11*. Cambridge, MA: MIT Press.

Amit, D. J. (1989). *Modeling brain function*. Cambridge: Cambridge University Press.

Amit, D. J., & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex, 7*, 237.

Amit, D. J., Del Giudice, P., & Fusi, S. (1999). Apprendimento dinamico della memoria di lavoro: una realizzazione in elettronica analogica (in Italian). In *Frontiere della vita 3*, 599. Rome: Istituto della Enciclopedia Italiana.

Amit, D. J., & Fusi, S. (1994). Dynamic learning in neural networks with material synapses. *Neural Computation, 6*, 957.

Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1987). Statistical mechanics of neural networks near saturation, *Annals of Physics, 173*, 30.

Badoni, D., Bertazzoni, S., Buglioni, S., Salina, G., Amit, D. J., & Fusi, S. (1995). Electronic implementation of a stochastic learning attractor neural network. *NETWORK Computation in Neural Systems, 6*, 125.

Badoni, D., Salamon, A., Salina, A., & Fusi, S., (2000). *aVLSI implementation of a network with spiking neurons and stochastic learning*. In preparation.

Bi, G.-q., & Poo, M.-m. (1999). Activity induced synaptic modifications in hippocampal culture: Dependence on spike timing, synaptic strength and cell type. *J. Neurosci., 18*, 10464–10472.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience, 2*, 1, 32.

Bliss, T. V. P., & Collingridge, G. L. (1993). A synaptic model of memory: Long term potentiation in the hippocampus. *Nature, 361*, 31.

Brunel, N., Carusi, F., & Fusi, S. (1998). Slow stochastic Hebbian learning of classes of stimuli in a recurrent neural network. *NETWORK Computation in Neural Systems, 9*, 123.

Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes*. London: Methuen.

Del Giudice, P., Fusi, S., Badoni, D., Dante, V., & Amit, D. J. (1998). Learning attractors in an asynchronous, stochastic electronic neural network. *NETWORK Computation in Neural Systems, 9*, 183.

Diorio, C., Hasler, P., Minch, B. A., & Mead, C. (1998). Floating-gate MOS synapse transistors. In T. S. Lande (Ed.), *Neuromorphic systems engineering: Neural networks in silicon*. Norwell, MA: Kluwer.

Elias, J., Northmore, D. P. M., & Westerman, W. (1997). An analog memory circuit for spiking silicon neurons. *Neural Computation, 9*, 419.

Fusi, S., Del Giudice, P., & Amit, D. J. (in press). Neurophysiology of a VLSI spiking neural network. In *Proceedings of the IJCNN 2000*. Available online at: http://jupiter.roma1.infn.it.

Fusi, S., & Mattia, M. (1999). Collective behavior of networks with linear (VLSI) integrate-and-fire neurons. *Neural Computation, 11*, 633.

Gerstein, G. L., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal, 4*, 41.

Grossberg, S. (1969). On learning and energy-entropy dependence in recurrent and non recurrent signed networks. *Journal of Stat. Phys., 1*, 319.

Hafliger, P., Mahowald, M., & Watts, L. (1996). A spike based learning neuron in analog VLSI. In M. Hozer, M. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems, 9* (p. 692). Cambridge, MA: MIT Press.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent selective computational abilities. *Proc. Natl. Acad. Sci. USA, 79*, 2554.

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

Mattia, M., & Del Giudice, P. (in press). Asynchronous simulation of large networks of spiking neurons and dynamical synapses. *Neural Computation*.

Mead, C. (1989). *Analog VLSI and Neural System*. Reading, MA: Addison-Wesley.

Meyer, P. L. (1965). *Introductory probability and statistical applications*. Reading, MA: Addison-Wesley.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Ann. Rev. Neurosci., 16*, 245.

Nadal, J. P., Toulouse, G., Changeux, J. P., & Dehaene, S. (1986). Networks of formal neurons and memory palimpsests. *Europhys. Lett., 1*, 535.

Parisi, G. (1986). A memory which forgets. *J. Phys, A: Math. Gen., 19*, L617.

Petersen, C. C. H., Malenka, R. C., Nicoll, R. A., & Hopfield, J. J. (1998). All-or-none potentiation at CA3-CA1 synapses. *Proc. Natl. Acad. Sci., 95*, 4732.

Ricciardi, L. M. (1977). Diffusion processes and related topics on biology. Berlin: Springer-Verlag.

Sarpeshkar, R. (1998). Analog versus digital: Extrapolating from electronics to neurobiology. *Neural Computation, 10*, 1601.

Sejnowski, T. J. (1976). Storing covariance with nonlinearly interacting neurons. *J. Math. Biology, 4*, 303.

Senn, W., Tsodyks, M., & Markram, H. (in press). An algorithm for modifying neurotransmitter release probability based on pre- and post-synaptic spike timing. *Neural Computation*.

Softky, W. R., & Koch, C. (1992). Cortical cells should spike regularly but do not. *Neural Computation, 4*, 643.

Tsodyks, M. (1990). Associative memory in neural networks with binary synapses. *Modern Physics Letters, B4*, 713.

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

van Vreeswijk, C. A., & Sompolinsky, H. (1997). Chaos in neural networks with balanced excitatory and inhibitory activity. *Science, 274*, 1724.

Willshaw, D. (1969). Non-holographic associative memory. *Nature, 222*, 960.

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

Received April 28, 1999; accepted November 8, 1999.