Biol. Cybern. 87, 459–470 (2002)
DOI 10.1007/s00422-002-0356-8

Biological Cybernetics

Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates

Stefano Fusi

Institute of Physiology, University of Bern, Bühlplatz 5, 3012 Bern, Switzerland

Received: 18 May 2002 / Accepted: 15 July 2002

Abstract. Synaptic plasticity is believed to underlie the formation of appropriate patterns of connectivity that stabilize stimulus-selective reverberations in the cortex. Here we present a general quantitative framework for studying the process of learning and memorizing of patterns of mean spike rates. General considerations based on the limitations of material (biological or electronic) synaptic devices show that most learning networks share the palimpsest property: old stimuli are forgotten to make room for the new ones. In order to prevent too-fast forgetting, one can introduce a stochastic mechanism for selecting only a small fraction of synapses to be changed upon the presentation of a stimulus. Such a mechanism can be easily implemented by exploiting the noisy fluctuations in the pre- and postsynaptic activities to be encoded. The spike-driven synaptic dynamics described here can implement such a selection mechanism to achieve slow learning, which is shown to maximize the performance of the network as an associative memory.

1 Introduction: cortical reverberations

In a growing number of neurophysiological experiments in which a primate performs a memory task, neurons are observed to have elevated spike rates in the delay periods between successive visual stimuli (for recent reviews see, e.g., Miyashita and Toshirio 2000; Wang 2001). The delay-activity distribution across the recorded cells is automatically triggered by specific visual stimuli, and it lasts for long periods after the removal of the sensory stimulus (up to 30 s; Fuster 1995). In the inferotemporal cortex the sustained activity is stimulus specific, that is, each visual stimulus evokes a characteristic pattern of delay activity. When unfamiliar, novel stimuli are presented, the recorded neurons may respond with elevated selective rates to the stimulus, but they do not show any sustained delay activity (Miyashita 1993). This is an indication that the delay-activity phenomenon is established only after many presentations of the same visual stimuli. A typical example of recorded delay activity is shown in Fig. 1. In the absence of any experimental evidence of the contrary, we assume that the spike activities of different neurons are asynchronous, and that all the information about the identity of the stimulus is encoded in the mean spike frequency of every cell.

1.1 The attractor picture

The experimental findings described above have been interpreted as an expression of the cortical interactions between large numbers of neurons, and a comprehensive picture has been suggested in the framework of attractor neural networks (Amit 1995; Amit and Brunel 1997a,b). In this framework the selective sustained activity is not a single-cell property, but rather the result of a feedback mechanism that maintains a reverberating activity in the absence of the sensory stimulus. During the delay period, those neurons that have been driven to high spike rates by the visual stimulus and that are coupled by strong-enough synaptic connections excite one another in such a way that the enhanced activity is stably self-sustained until the arrival of the next visual stimulus. The dynamics of this population of neurons, and in particular of the feedback mechanism, is governed by the set of all the synaptic connections and efficacies. This set stores passively all the possible delay activity distributions in response to different stimulations: the visual stimulus selects one of the potential responses by determining the initial state of activation of the population. Following the removal of the stimulus, the network dynamics is attracted towards one of the stable patterns of activity (attractors), which represents the response of the network. Since these responses are expressed in terms of spike-rate variations, they can be communicated actively to other areas for further processing.

Correspondence to: (e-mail: fusi@cns.unibe.ch,
Tel.: +41-31-6318778, Fax: +41-31-6314611)

The formation of the suitable synaptic structure for stabilizing the delay activity distributions discussed in Amit (1995) is probably one of the best exemplifications of the Hebbian paradigm (Hebb 1949). The repeated imposition of a pattern of reverberating activity to an assembly of neurons eventually leads to a synaptic structure that makes the reverberating activity stable, even in the absence of the sensory stimulus that triggered it. In other words, when a cell takes part in the collective dynamics that activates another cell (during the imposition of the visual stimulus), its efficacy is increased, since this change goes in the direction of further stabilizing the imposed reverberating activity. This is equivalent to a covariance-based updating rule for the synaptic efficacy: when the two neurons are simultaneously active the synapse should be potentiated (Sejnowski 1977; Hopfield 1982).

2 An analytic, quantitative framework for realistic learning

Attractor formation was our starting point for developing a general framework that would provide a simple and general description of a wide class of biologically plausible prescriptions for updating the synaptic connections. This framework has been introduced in Amit and Fusi (1992, 1994), and relies on basic constraints that are likely to pose limitations on any type of material (biological or artificial) synaptic device. The guiding principle that dictated the assumptions proved to be surprisingly powerful for drawing general conclusions about synaptic dynamics. The assumptions are:

1. Locality in time (online learning) and in space: the only information available to the synapse is the current activity of the two neurons it connects and the previous synaptic state. The synapse is supposed to acquire information from every single presentation, and no temporary storage is available for a later update in which the information about more than one stimulus is available at the same time.
2. All internal variables describing the synaptic state are bounded.
3. Long-term modifications of the synaptic internal variables cannot be arbitrarily small.

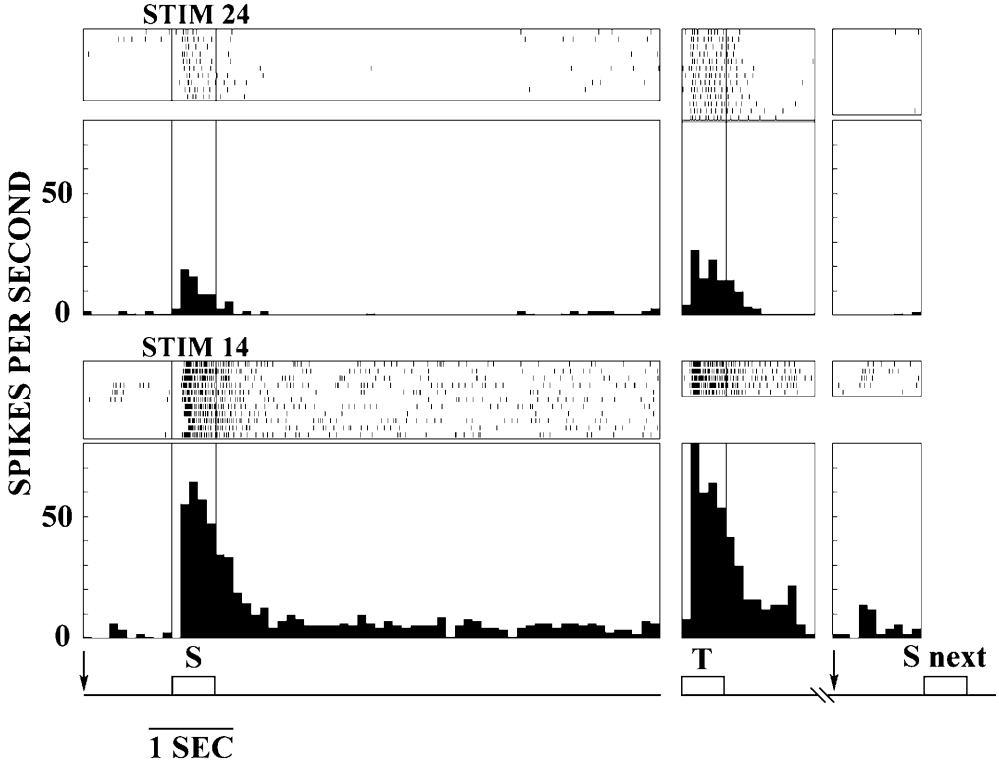

Fig. 1. Stimulus-selective delay activity in the cortex. The monkey performs a delayed-match-to-sample task in which it has to compare a visual sample stimulus (*S*) to a visual test stimulus (*T*) and respond differently depending on whether the two stimuli are the same or different. The plot shows the response of a cell to two different stimuli (*STIM 24* and *STIM 14*). The *rasters* show the spikes emitted by the cell in different trials, and the peristimulus time histogram shows the mean spike rate across all the repetitions of the same stimulus as a function of time. Note that the trial intervals are sorted and reorganized according to their corresponding stimulus identity. Consequently, the number and order of the rasters for the sample and test stimuli are not aligned: the data shown during and following the test stimulus are combined across match and nonmatch conditions. The visual stimulus S triggers a sustained delay activity in response to stimulus 14, but not to stimulus 24. The information about the last stimulus seen is propagated up to the presentation of the next visual stimulus. The mechanism underlying the selective delay activity seems to be automatic, i.e., not effected by the task. Indeed the sustained activity is triggered also in the intertrial interval, between the test stimulus and the sample of the next trial, where there is no need to hold in memory the identity of the last stimulus seen. See also Amit et al. (1997) for a description of many other features of the delay activity (figure adapted from Yakovlev et al. 1998)

2.1 The palimpsest property

Under the above assumptions, any network of neurons exhibits the "palimpsest" property (Nadal et al. 1986; Parisi 1986; Amit and Fusi 1992, 1994): old stimuli are automatically forgotten to make room for the most recent ones. The memory span is limited to a sliding window containing a certain number of stimuli that induced synaptic modifications. Within this window, recent stimuli are best remembered while stimuli outside the memory window are completely forgotten, as if they had never been seen by the network. The width of the sliding window depends on how many synapses are changed following each presentation: if this number is small the network is slow to acquire information from the stimuli to be learnt, but the memory span is large. Otherwise, if the fraction of synapses that are changed upon each stimulus presentation is large, the network learns quickly but the memory span is quite limited. This constraint can be so tight that it might seriously compromise the functioning of the network as an associative memory. In Sect. 2.1.1 we illustrate the palimpsest property when the network learns uncorrelated, random patterns of activity.

2.1.1 The problem of fast forgetting. Learning can be seen as a stochastic process when the stimuli to be stored are random (Heskes and Kappen 1991; Amit and Fusi 1992): each stimulus imposes a specific activity level to the two neurons connected by the synapse. If these stimuli are random and uncorrelated, each synapse will see a random sequence of pairs of activities. We denote the activity of neuron i by the variable χi_i^t, where ξ_i^t denotes a generic stimulus shown at time t. In our case ξ_i^t represents the mean spike rate, but in principle it can be any quantity related to what should be encoded by the synapse (e.g., it might be the degree of correlation between the spike activities of the pre- and postsynaptic neurons). To simplify the analysis we assume that the set of stable internal synaptic states is discrete (but it can be arbitrarily large). As a consequence the presentation of a sequence of uncorrelated stimuli induces a random walk among the stable synaptic values which can be described as a Markov process. More formally, the probability $\mathbf{M}_{KJ}$ that a synapse makes a transition from the internal stable state K to the stable state J is given by:

$$\mathbf{M} = \sum_{\xi_{\text{pre}},\xi_{\text{post}}} p(\xi_{\text{pre}}, \xi_{\text{post}})\mathbf{Q}(\xi_{\text{pre}}, \xi_{\text{post}}) \tag{1}$$

where the sum extends over all the pairs of activities that induce long-term synaptic modifications. For each pair of activities, $p(\xi_{\text{pre}}, \xi_{\text{post}})$ is the probability that a stimulus imposes the activities ξ_{pre} and ξ_{post} to the pair of neurons connected by the synapse, and $\mathbf{Q}(\xi_{\text{pre}}, \xi_{\text{post}})$ is a matrix of binary values that encodes the learning rule, i.e., how the neural activities modify the internal synaptic state: if $Q^{KJ}(\xi_{\text{pre}}, \xi_{\text{post}}) = 1$, then a transition from state K to state J occurs whenever the specific pair of activities ξ_{pre} and ξ_{post} is imposed by the stimulus to the pre- and postsynaptic neurons. Otherwise, when $Q^{KJ}(\xi_{\text{pre}}, \xi_{\text{post}}) = 0$, the transition cannot take place. If a particular pair of activities leaves the synapse unchanged, then $\mathbf{Q}$ is equal to the identity matrix (the synapse remains in the initial state and only the diagonal terms are nonzero). If the pair of activities potentiates or depresses the synapse by inducing a transition to one of the two neighboring states, then the matrix has the following form (top: long-term potentiation, LTP; bottom: long-term depression, LTD):

$$\mathbf{Q} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Low indexes (J, K) correspond to depressed synaptic states. Notice that each row contains only one nonzero term.

We now focus on the memory trace left by a generic stimulus $\boldsymbol{\xi}^1$, followed by the presentation of $p-1$ other uncorrelated random patterns. We want to know whether the final, current synaptic matrix still preserves any dependence on the pattern of activity imposed by stimulus $\boldsymbol{\xi}^1$. As shown below, the memory trace of the most recent stimuli $\boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^p$ is stronger than the one of $\boldsymbol{\xi}^1$. Hence if $\boldsymbol{\xi}^1$ can be retrieved from memory, then all the other $p-1$ patterns can be recalled a fortiori, and p provides an estimate of the number of patterns that can be stored in the synaptic matrix. The conditional distribution function $\rho_J^p(\xi_{\text{pre}}^1, \xi_{\text{post}}^1)$ that a synapse is in state J following the presentation of p patterns, the first of which imposed $\xi_{\text{pre}}^1, \xi_{\text{post}}^1$ on the synapse, satisfies the equation

$$\rho_J^p(\xi_{\text{pre}}^1, \xi_{\text{post}}^1) = \sum_{K=1}^{n_s} \rho_K^1(\xi_{\text{pre}}^1, \xi_{\text{post}}^1)(\mathbf{M}^{p-1})_{KJ} \ , \tag{2}$$

where the index K runs over all the stable n_s synaptic states. The initial distribution ρ_K^1 induced by the presentation of stimulus $\boldsymbol{\xi}^1$ is modified $p-1$ times by the successive random stimuli. We implicitly hypothesize that all the patterns of activity induced by the p stimuli have the same statistics and are not correlated with the first pattern $\boldsymbol{\xi}^1$.

It is reasonable to assume that there is always a sequence of synaptic transitions, on any given synapse, that can bring the synapse from any one of its stable states to any other state. This type of dynamics is

ergodic and hence, for a large number of presentations of random patterns, there is an asymptotic distribution:

$$\rho_J^p(\xi_{\text{pre}}^1, \xi_{\text{post}}^1) \to \rho_J^\infty \tag{3}$$

which is independent of $\xi_{pre}^1, \xi_{post}^1$. When the distribution of the synapses becomes too close to the asymptotic distribution, the memory trace of the first pattern $\boldsymbol{\xi}^1$ fades away, and eventually disappears. This is another expression of the palimpsest property. The number of patterns of which there is still some memory trace depends essentially on the forgetting rate, i.e., the convergence rate to the asymptotic distribution.

2.1.2 A tight constraint on storage capacity. The synaptic matrices that would be obtained after the presentation of a very large number of statistically independent random patterns (in principle an infinite number) would be independent of the initial condition determined by the presentation of the first, oldest pattern $\boldsymbol{\xi}^1$. Following each stimulus presentation, the synaptic matrix still changes, but the statistics remains constant and is only determined by the asymptotic distribution of (3). In order to preserve a dependence (and hence some memory) on $\boldsymbol{\xi}^1$, the difference between the synaptic matrix $\mathbf{W}(p)$ following the presentation of p patterns and any of the asymptotic synaptic matrixes $\mathbf{W}(\infty)$ should be large enough. This is just a necessary condition because this memory trace expresses only a generic dependence of the synaptic matrix on the structure of $\boldsymbol{\xi}^1$, and such a dependence could be too small for the network dynamics to retrieve information about stimulus $\boldsymbol{\xi}^1$. In more formal terms we should impose

$$\Delta W = \left| \sum_{i,j} [W_{ij}(p) - W_{ij}(\infty)] \right| > \delta \tag{4}$$

where δ depends on the details of the network neural dynamics. If the network is composed of N neurons, N is large, and the number of synapses scales as N^2 (the best case), then ΔW is estimated by

$$\Delta W \sim N^2 \left| \sum_J [\rho_J^p(\xi_{\text{pre}}^1, \xi_{\text{post}}^1) - \rho_J^\infty] W^J \right|$$

where now the sum extends over all the stable synaptic states $1, \ldots, n_s$ that correspond to the synaptic efficacies $W^1, \ldots, W^{n_s}$. We implicitly hypothesized that the statistics of the stimuli is "translation invariant," i.e., it does not depend on the specific pair of neurons that we are considering. What follows can be easily extended to the case in which the statistics of the stimuli is more complex.

Since we assume that the process is ergodic, there will be an n_0 such that all the elements of $\mathbf{M}^{n_0}$ ($\mathbf{M}$ raised to the power n_0) are greater than 0. n_0 is the mean number of presentations that are needed to visit all the stable synaptic states. If the synapse is discrete, and only transitions between neighboring synaptic states are permitted, then n_0 corresponds to the number of stable states n_s. In general $n_0 \leq n_s$. If the synapse is deeply analog (i.e., the number of stable states is large), n_0 is proportional to the full range of variability of the synaptic efficacy $W^{\max} - W^{\min}$ divided by the minimal change in the synaptic efficacy that a single stimulus presentation can induce.

For the ergodic theorem (see, e.g., Shiryaev 1984) we have, for any pair (K, J) of states, that $\mathbf{M}^p$ converges to its limit geometrically:

$$\left| (\mathbf{M}^p)_{KJ} - (\mathbf{M}^\infty)_{KJ} \right| < (1 - \Lambda_{\min})^{[p/n_0]-1} \tag{5}$$

where $\Lambda_{\min}$ is the minimal value of $\mathbf{M}^{n_0}$, and depends in a complicated way on the smallest nonzero components of $\mathbf{M}$ and on n_0. The inequality (5) sets an upper bound for the difference between ρ_K^p and the asymptotic, memoryless distribution ρ_K^∞. The necessary condition of (4) for retrieving the oldest pattern ξ^1 can be rewritten as follows:

$$\Delta W \sim N^2 \left| \sum_{J,K} [\rho_K^1(\xi_{\text{pre}}^1, \xi_{\text{post}}^1)(\mathbf{M}^{p-1})_{KJ} - \rho_K^1(\xi_{\text{pre}}^1, \xi_{\text{post}})(\mathbf{M}^\infty)_{KJ}] W^J \right| > \delta$$

We used the explicit expression of the conditional distribution ρ^p given in (2). If we now replace the difference between the two matrices $(\mathbf{M}^{p-1} - \mathbf{M}^\infty)$ with its maximum given by (5), we obtain

$$\Delta W < N^2 (1 - \Lambda_{\min})^{\left[\frac{p}{n_0}\right]-1} \sum_J W^J \sum_K \rho_K^1(\xi_{\text{pre}}^1, \xi_{\text{post}}^1)$$

$$< N^2 n_s W^{n_s} (1 - \Lambda_{\min})^{\left[\frac{p}{n_0}\right]-1}$$

where W^{n_s} is the maximal efficacy. The final condition for retrieving pattern ξ^1 now reads

$$\Delta W \sim W^{n_s} n_s N^2 (1 - \Lambda_{\min})^{[p/n_0]-1} > \delta$$

which poses a constraint on the maximum number of patterns p that can be stored in the synaptic matrix:

$$p < \frac{-n_0 \log(n_s N^2/\delta)}{\log(1 - \Lambda_{\min})} \sim \frac{n_0 \log(N\sqrt{n_s})}{\Lambda_{\min}} \tag{6}$$

This constraint is extremely tight and very general. Only a number of patterns that scale as $\log N$ can be retrieved, which makes the network a very inefficient associative memory. Note that if one of the parameters such as n_0 or $\Lambda_{\min}$ becomes N-dependent, then one can extend the memory span. For instance, if the number of synaptic states increases with N (as in the case of the Hopfield (1982) model), n_0 and $1/\Lambda_{\min}$ provide two extra N-dependent factors which in some cases destroy the palimpsest behavior. In those cases the storage capacity is mainly limited by the interference between the memory traces of different stimuli, and not by memory fading. For a wide class of models, as soon as the

maximum storage capacity is surpassed, the network suddenly becomes unable to retrieve any of the memorized patterns. This is also known as the blackout catastrophe (see, e.g., Amit 1989). In order to prevent this, one has to stop learning at the right time, or the network must be able to forget. Hence the palimpsest property, if forgetting is not too fast, can be a desirable feature of the learning process.

2.1.3 The solution: the stochastic selection mechanism. The expression for p (Eq. 6) contains the directions to solve the problem of the logarithmic constraint. Indeed $\Lambda_{\min}$ depends on the fraction of synapses that make a transition to a different stable state (contained in the structure of **M**), which, in turn, depends on the statistical properties of the stimuli (i.e., the fraction of neurons that are driven to activities that induce long-term synaptic changes) and on the inherent synaptic dynamics which in our case is described by the **Q** matrix (see Eq. 1). Decreasing the fraction of synapses that are changed upon each presentation can dramatically increase the memory span. What the network needs is a mechanism that selects which synaptic efficacies have to be changed following each stimulation. In the absence of an external supervisor that could perform this task, a possible local, unbiased mechanism is stochastic learning: at parity of conditions (activities of the pre- and postsynaptic neurons) the transitions between stable states occur with some probability. This means that only a fraction of randomly chosen synapses are changed upon each stimulus presentation. This stochastic selection would correspond to transforming the binary **Q** matrix, which expresses a deterministic learning rule (given a specific pair of pre- and postsynaptic activities the synapse always behaves in the same way), into the matrix of the probabilities of making a transition from one state to another state given a specific pair of pre- and postsynaptic activities. In this way the elements of **Q** can be arbitrarily small and $\Lambda_{\min}$ can tend to zero, hence increasing the memory span. See Sect. 2.2 for an example of this.

2.2 The standard learning scenario

The scenario depicted in Sect. 2.1.3 has been studied analytically in (Amit and Fusi 1992, 1994; Battaglia and Fusi 1995; Fusi 1995; Kahn et al. 1995; Lattanzi et al. 1997; Brunel et al. 1998). Learning can be described as a random walk of the internal synaptic variable among the stable states. Every stimulus imposes a pattern of neuronal activity, and a randomly selected fraction of the synapses make a transition to a different stable state. Such a system can be described as a Markov process, and the final distribution of the synaptic efficacies can be computed as an explicit function of the patterns of activity imposed by the stimuli. From the analysis it turns out that in general there are at least three elements that characterize online updating rules for material synapses:

1. The memory trace left by each stimulus presentation. This represents the way the synapse encodes the activity of the pre- and postsynaptic neurons. It depends also on the initial distribution of the synaptic states.
2. The decay of the memory trace due the palimpsest property. New stimuli use resources that had been previously allocated to other stimuli, and, in doing so, erase the memory trace of the oldest stimuli. The forgetting rate depends on the statistics of the stimuli and on the inherent transition probability of the synapse.
3. The interference between memory traces corresponding to different stimuli: the simultaneous presence in the synaptic matrix of memory traces corresponding to different patterns within the memory span generates noise that might prevent the system from recalling the stored patterns correctly.

We now illustrate the meaning of these three ingredients with a simple example. Assume that the synapse has only two stable states ($n_s = 2$), corresponding to two efficacies: $W^1 = 0, W^2 = W$. In this case the transition matrix **M** is 2×2 and contains only two independent terms: (i) the probability of making a transition to the potentiated state and (ii) the probability of jumping into the depressed state (see below for the expression of **M**).

Each stimulus selects randomly a mean fraction f of neurons and drives them to a state of elevated activity, the same for all the active neurons. In order to estimate the retrievable memory trace of the first stimulus presented ($\boldsymbol{\xi}^1$), we introduce the classical signal-to-noise ratio (indicated as S/R; see, e.g., Amit 1989): the signal S expresses the distance between the distribution of the total synaptic input across all neurons that should be active, and the corresponding distribution across the neurons that should stay quiescent when the pattern of activity $\boldsymbol{\xi}^1$ is imposed on the network. Quantitatively S is defined as the difference between the averages of these two distributions. The noise R represents the mean width of the two distributions. A high S/R would allow the network to retrieve from memory the pattern of activity $\boldsymbol{\xi}^1$ that is embedded in the synaptic matrix, and make it a stable point of the collective dynamics.

The synaptic updating rule is as follows: when the two neurons connected by the synapse are both active, the synapse makes a transition to the potentiated state with probability $q^{\mathrm{P}}_{\mathrm{AA}}$ and a transition to the depressed state with probability $q^{\mathrm{D}}_{\mathrm{AA}}$. The other transitions occur with probabilities denoted with a similar notation (e.g., $q^{\mathrm{D}}_{\mathrm{IA}}$ is the probability that depression occurs when the presynaptic neuron is active and the postsynaptic neuron is inactive). These probabilities correspond to the terms of the $\mathbf{Q}(\xi_{\mathrm{pre}}, \xi_{\mathrm{post}})$ matrix introduced in (1) and, again, encode the learning rule. For simplicity we dropped one of the state indexes K, J (the one corresponding to the initial state) since the synapse is bistable, and we expressed the dependence on the activities ξ_{pre} and ξ_{post} in two subscript indexes. With this notation, the most general Markov matrix has the following form (see Eq. 1):

$$\mathbf{M} = f^2 \begin{pmatrix} 1-q^{\mathrm{P}}_{\mathrm{AA}} & q^{\mathrm{P}}_{\mathrm{AA}} \\ q^{\mathrm{D}}_{\mathrm{AA}} & 1-q^{\mathrm{D}}_{\mathrm{AA}} \end{pmatrix} + f(1-f) \begin{pmatrix} 1-q^{\mathrm{P}}_{\mathrm{AI}} & q^{\mathrm{P}}_{\mathrm{AI}} \\ q^{\mathrm{D}}_{\mathrm{AI}} & 1-q^{\mathrm{D}}_{\mathrm{AI}} \end{pmatrix} + f(1-f) \begin{pmatrix} 1-q^{\mathrm{P}}_{\mathrm{IA}} & q^{\mathrm{P}}_{\mathrm{IA}} \\ q^{\mathrm{D}}_{\mathrm{IA}} & 1-q^{\mathrm{D}}_{\mathrm{IA}} \end{pmatrix} + (1-f)^2 \begin{pmatrix} 1-q^{\mathrm{P}}_{\mathrm{II}} & q^{\mathrm{P}}_{\mathrm{II}} \\ q^{\mathrm{D}}_{\mathrm{II}} & 1-q^{\mathrm{D}}_{\mathrm{II}} \end{pmatrix}$$

where f^2 is the probability that the pre- and postsynaptic neurons are active (AA), $f(1-f)$ is the probability that the postsynaptic neuron is active and the presynaptic neuron is inactive (IA), and so on for all the four combinations of activities.

For such a system, after the presentation of p patterns, the signal corresponding to the oldest pattern is (for the derivation of this expression in a simpler case, see Amit and Fusi 1994)

$$S = Wf\lambda^{p-1}[(1-c_0)(q^{\mathrm{P}}_{\mathrm{AA}} - q^{\mathrm{P}}_{\mathrm{IA}}) + c_0(q^{\mathrm{D}}_{\mathrm{IA}} - q^{\mathrm{D}}_{\mathrm{AA}})] \qquad (7)$$

where c_0 is the initial fraction of potentiated synapses and λ is the smallest eigenvalue of the Markov transition matrix. λ is given by

$$\lambda = 1 - (q^{\mathrm{P}}_{\mathrm{AA}} + q^{\mathrm{D}}_{\mathrm{AA}})f^2 - (q^{\mathrm{P}}_{\mathrm{IA}} + q^{\mathrm{D}}_{\mathrm{IA}} + q^{\mathrm{P}}_{\mathrm{AI}} + q^{\mathrm{D}}_{\mathrm{AI}})f(1-f) - (q^{\mathrm{P}}_{\mathrm{II}} + q^{\mathrm{D}}_{\mathrm{II}})(1-f)^2$$

which is essentially 1 minus the sum of all the transition probabilities, each multiplied by the corresponding probability of occurrence of a specific pair of activities (e.g., the probability of both pre- and postsynaptic neurons being active is f^2). Let c_∞ be the asymptotic fraction of potentiated synapses that one would get after an infinite number of presentations of different stimuli. The interference noise term depends mostly on c_∞ and not on the conditional distributions that still preserve memory of the first pattern ξ^1 (Amit and Fusi 1994):

$$R \sim W\sqrt{c_\infty \frac{f}{N}}$$

where N is the number of neurons in the network.

From these formulae it is clear that the most efficient way of storing patterns of activities (i.e., when S/R is maximal) is when the Hebbian term $q^{\mathrm{P}}_{\mathrm{AA}}$ dominates over $q^{\mathrm{P}}_{\mathrm{IA}}$, the transition probabilities are low ($\lambda \sim 1$), and the coding level f of the stimuli tends to zero with N. The transition probabilities should scale in such a way that all terms in λ tend to zero with f at the same rate as $q^{\mathrm{P}}_{\mathrm{AA}}f^2$ (e.g., $q^{\mathrm{P}}_{\mathrm{IA}} \sim q^{\mathrm{P}}_{\mathrm{AA}}f$) or faster. This would correspond to a scenario in which learning is slow – i.e., the stimuli have to be repeatedly presented to the network in order to be learnt – and the updating rule ensures the balance between the mean number of potentiations and the mean number of depressions. In such a case the network performs extremely well as an associative memory (e.g., it recovers the optimal storage capacity in terms of information that can be stored in the synaptic matrix), even if the synaptic efficacy is binary (two stable states only). The number of different patterns that can be stored and retrieved from memory without errors can be as large as $N^2/(\log N)^2$, if the mean fraction f of active neurons scales as $\log N/N$ (Amit and Fusi 1994). Slow learning also allows automatic prototype extraction from a class of stimuli that evoke similar, correlated patterns of activity (Fusi 1995; Brunel et al. 1998).

Increasing the number n_{s} of stable synaptic states does not improve the performance of the network much, provided that in the patterns of activity to be stored large populations of neurons encode the same information (e.g., they have the same mean spike rate). In the case analyzed here this assumption is definitely true, since there are only two levels of activity imposed by the stimuli: neurons are either active or inactive. All the active and inactive neurons are read by the postsynaptic neuron through two classes of synaptic efficacies, and the total synaptic current can be written as

$$I_i = \frac{1}{N}\sum_{j\in \mathrm{A}} W_{ij}\xi_j + \frac{1}{N}\sum_{j\in \mathrm{I}} W_{ij}\xi_j = \frac{\xi_{\mathrm{A}}}{N}\sum_{j\in \mathrm{A}} W_{ij} + \frac{\xi_{\mathrm{I}}}{N}\sum_{j\in \mathrm{I}} W_{ij}$$

If the two groups of synapses are large enough, the two sums are equivalent to an analog weight that has a number of states as large as the number of neurons that encode the same activity (when each synaptic efficacy is binary). For a wide class of network models, if p patterns can be stored with binary synapses, no more than $(n_{\mathrm{s}}-1)p$ patterns can be memorized with a multistable synapse (Amit and Fusi 1994). In small networks, the performance improvement that can be achieved by increasing n_{s} can be more dramatic.

Interestingly, the dependence on the initial fraction c_0 of potentiated synapses c_0 can change the behavior of the network qualitatively. By changing this parameter one can even implement the phenomenon of primacy, in which the first stimulus seen by the network has a stronger memory trace than the subsequent stimuli (Kahn et al. 1995).

Finally, a few considerations about the coding level f of the patterns are mentioned. The optimal case is achieved when both f and the transition probabilities q tend to 0 with N, as explained above. However there are many intermediate cases that are studied in Amit and Fusi (1994) in which the network can perform well as an associative memory, even if only the condition of sparseness of the stimuli is imposed. However, in this case the number p of random patterns with a mean fraction f of active neurons that can be stored and successfully retrieved does not surpass $1/f$. In the case of stochastic learning p can be as large as $1/f^2$.

3 Synaptic dynamics

Stochastic learning provides a good tool for studying quantitatively the process of attractor formation without specifying the details of the synaptic dynamics. However, to proceed further, it is important to have a detailed

model that provides a link between the transition probabilities and the synaptic dynamics driven by the neural activities. This implies the identification of the source of noise that drives the stochastic selection mechanism. The solution we propose is suggested by the analysis of cortical recordings: the spike trains recorded in vivo are quite irregular (see, e.g., Softky and Koch 1993), and the interspike interval variability is directly accessible by the synapse. It is actually possible to exploit this source of stochasticity to achieve stochastic transitions in a fully deterministic synaptic model.

3.1 The model

The synaptic dynamics is described in terms of an internal variable $X(t)$. X is restricted to the interval $[0, 1]$ and, inside the permitted interval, obeys

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = -\alpha\Theta\Big(-X(t)+\theta_X\Big) + \beta\Theta\Big(X(t)-\theta_X\Big) + H(t) \ , \tag{8}$$

where Θ is the Heaviside function. The first two terms make the upper and the lower bound the only two stable states of X when $H(t)$, the stimulus driven Hebbian learning term, is 0. The internal state decays at a rate α to the lower bound $X = 0$ when $X < \theta_X$. Otherwise X is attracted towards the upper bound $X = 1$ with a constant drift β. The Hebbian term depends on variables related to the activities of the two neurons connected by the synapse. Presynaptic spikes trigger temporary modifications in the synaptic variable: each spike induces a jump in X whose value depends on the instantaneous depolarization of the postsynaptic neuron V_{post}. It is upwards ($X \to X + a$) if the depolarization is high ($V_{\mathrm{post}} > V_{\mathrm{H}}$), and downwards if the depolarization is low($V_{\mathrm{post}} < V_{\mathrm{L}}$). The depolarization is indirectly related to the activity of the postsynaptic neuron. A single reading of the instantaneous depolarization does not contain much information about the postsynaptic mean rate. However, the required information is distributed across several neurons that are driven to the same activity by the stimulus. This means that even a single instantaneous reading of the depolarization of a population of cells contains all the information about the mean spike rate and many other statistical properties of the activity of these neurons (see the Appendix).

3.2 Stochastic transitions

To illustrate the stochastic nature of the learning mechanism, we assume that the presynaptic spike train is Poisson distributed, while the afferent current to the postsynaptic neuron is a noisy, Gauss-distributed stochastic process. Such a situation is meant to mimic what happens in vivo during the presentation of a visual stimulus. The heavy bombardment of synaptic inputs that determines the activity of the postsynaptic neuron is here emulated by a Gaussian-distributed current (see, e.g., Tuckwell 1988).

The synaptic dynamics depends on the detailed statistics of the spike trains of the pre- and postsynaptic neurons. During stimulation the synapses move temporarily up and down, driven by the presynaptic spikes. Following the removal of the stimulus, the synaptic efficacy may return to its initial state, or it may make a transition to another state. Then, in the presence of spontaneous activity, the internal state of the synapse continues to change upon the arrival of each spike, but the probability of crossing the threshold θ_X becomes negligible and the synapse fluctuates around one of the stable states. Figure 2 shows two cases, during a typical stimulation, at parity of mean spike rates of pre- and

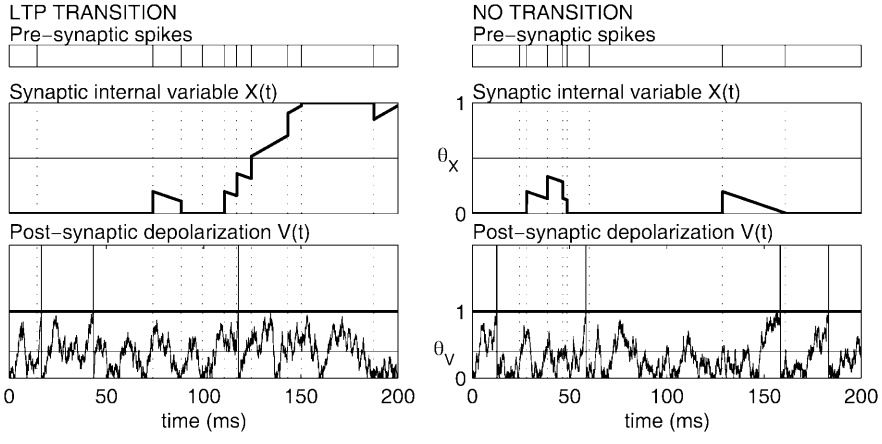

Fig. 2. Stochastic long-term potentiation (*LTP*): pre- and postsynaptic neurons fire at the same mean rate and the synapse starts from the same initial value [$X(0) = 0$] in both cases illustrated in the *left* and *right panels*. In each panel are plotted as a function of time (from *top* to *bottom*): the presynaptic spikes, the simulated synaptic internal variable $X(t)$, and the depolarization $V(t)$ of an integrate-and-fire post-synaptic neuron (the threshold for emitting a spike is 1). *Left panels*: LTP is caused by a burst of presynaptic spikes that drives $X(t)$ above the synaptic threshold. *Right panels*: At the end of stimulation, X returns to the initial value. Even though the mean firing rates are identical, the final state is different in the two cases (figure reproduced from Fusi et al. 2000a)

postsynaptic neurons: in one case (left panels in Fig. 2), a fluctuation drives the synaptic efficacy above threshold and, when the stimulus is removed, X is attracted to the high state: LTP has occurred; in the second case (right panels), when the stimulus is removed, X is below threshold and is attracted by the refresh to the initial state (no transition occurred). In the two cases the statistics of the activity to be encoded – the mean spike frequency – is the same, but the realization of the stochastic process that generated the pre and postsynaptic activities is different.

Such a stochastic selection mechanism shows that the importance of an internal threshold θ_X for the synaptic dynamics is at least twofold: (i) it stabilizes memory by ignoring all the fluctuations that do not drive the internal state across the threshold, and (ii) it provides a simple mechanism that allows the neural activity to select which synapses are to be modified. This mechanism is so simple and robust that it can be readily implemented in discrete electronics (Badoni et al. 1995; Del Giudice et al. 1998) or in analog VLSI (Fusi et al. 2000a; Chicca and Fusi 2001).

3.3 Transition probabilities

The 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. The spike-based synaptic dynamics described here produces the transition probabilities required to store an extensive number of patterns of mean rates. The stochastic process induced by the noisy pre- and postsynaptic activity on the internal synaptic variable X is known as a Takács process and can be studied using a density approach (Fusi et al. 2000a). The solution of the equations for the distribution of X gives the transition probabilities for any pair of pre- and postsynaptic spike rates, and for any stimulation time interval. For each postsynaptic frequency, the parameters μ and σ characterizing the input current are tuned to produce the desired mean firing rate ν_{post}, as explained in the Appendix. The relation between the mean spike frequency and the parameters that characterize the statistics of the input current depends on the details of the neuronal dynamics. In the case of simple integrate-and-fire neurons, this relation can be computed analytically (Ricciardi 1977; Fusi and Mattia 1999). Below we adopt the following model of an integrate-and-fire neuron: the depolarization V is the only dynamical variable. The neuron integrates linearly the synaptic input as long as V is below some threshold θ. When this threshold is crossed, the neuron emits a spike and is then reset to some value H, where it stays for the duration of a refractory period τ_{r}; then it starts to integrate the current again. In formal terms, the dynamics is governed by the following differential equation:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -\lambda + I(t)$$

where λ is a constant leakage and $I(t)$ is the total synaptic current. Such a neuron has a similar behavior to the one of the classical integrate-and-fire neuron with a leak proportional to the depolarization provided that V is limited from below by a rigid barrier (Fusi and Mattia (1999). Interestingly, the response function (the mean firing frequency as a function of the mean and variance of the input current) of these simple neurons can be fitted to that of real cortical pyramidal cells (A. Rauch, G. La Camera, H.-R. Luescher, W. Senn, S. Fusi, unpublished work, 2002).

The parameters of the current determine not only the mean firing rate, but also the distribution of the depolarization, and hence the probability that a presynaptic spike triggers an upwards (probability r_{a}) or a downwards jump (probability r_{b}). Given these two probabilities and the presynaptic rate ν_{pre}, the LTP/LTD probabilities [the $Q^{KJ}(\xi_{\text{pre}}, \xi_{\text{post}})$ of Eq. 1] can be computed numerically by calculating the probability that X is above/below the threshold θ_X at the end of the stimulation, when the initial state is 0/1.

These transition probabilities are plotted in Fig. 3. We assume that every stimulus brings a mean fraction f of cells to an activity level of 50 Hz; all the other neurons remain at a spontaneous level of 4 Hz. The LTP probability when the pre- and postsynaptic neurons are both active correspond to the q^{P}_{AA} introduced in Sect. 2.2 and is the highest transition probability. q^{P}_{IA} (presynaptic

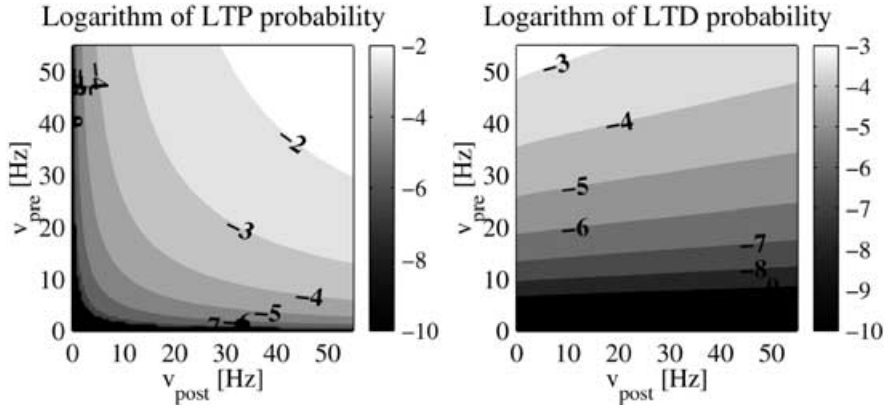

Fig. 3. Contour plots of LTP and long-term depression (*LTD*) probabilities (q) on a log scale vs pre- and postsynaptic neuron rates for a 500-ms stimulation. LTP occurs when pre and post synaptic rates are both high. Around the *white plateau*, P_{LTP} drops sharply and becomes negligible for spontaneous rates. The strong nonlinearity allows easy discrimination between relevant signals and background noise (figure adapted from Fusi et al. 2000a)

neuron at 50 Hz, postsynaptic neuron at 5 Hz) is clearly much smaller than q^{P}_{AA}, and $q^{D}_{IA} > q^{D}_{AA}$. All these relations go in the direction of maximizing the signal S (see Eq. 7) left by a single stimulus presentation. The parameters of the synapse are also tuned to ensure that q^{D}_{IA} is a fraction $f = 1/30$ of q^{P}_{AA}, to ensure the balance between potentiations and depressions. This is required to maximize the memory span. The highest probability q^{P}_{AA} is anyway small (0.03), which means that learning a pattern of activity would require many presentations ($\sim$30) of the same stimulus (slow learning). As one moves towards low, spontaneous activities, all the transitions probabilities (LTP and LTD) drop dramatically. This is necessary to prevent the network from forgetting in the presence of spontaneous neural activity. Interestingly, the mean time one has to wait until the synapse makes a spontaneous transition to another state is of the order of 100 days (or years, if the spontaneous activity is below 2 Hz), a period of time which is several orders of magnitude longer than the longest inherent time constant of the synaptic dynamics (the decay due to α occurs in a time of the order of $\sim$100 ms). This is another advantage of transferring the load of generating stochasticity to a system (the network) which is much bigger than a single synapse and hence offers a much larger state space. This permits the generation of rare events, as required for the synaptic transitions, without needing fine tuning of parameters or the resorting to biologically implausible long time constants.

4 Discussion

4.1 The problem of noise generation

Stochastic learning turned out to be a simple, unbiased mechanism for selecting the synapses that are to be updated upon a stimulus presentation. The mechanism has the great advantage of being localized in space: each synapse decides whether or not to change without knowing what the other synapses are doing, and nevertheless the average fraction of updated synapses is kept constant. Moreover slow learning can be easily achieved since the transition probabilities can be so low that the mean number of modified synapses is even smaller than one (Fusi 2001). However, this approach moves the problem to the generation of the proper noise to drive the synaptic dynamics. The fully deterministic spike-driven synaptic model described here exploits the variability in the neural activity to drive the stochastic mechanism: the source of randomness is in the spike-emission process of the neurons. The next issues, currently studied, are: (i) is it possible to generate the proper irregular spike activity with a deterministic network? and 2) would this irregular activity be good enough to drive the synaptic stochastic mechanism? It is known that deterministic networks of randomly connected integrate-and-fire neurons can generate highly irregular activity by exploiting the quenched, frozen disorder in the pattern of connectivity (van Vreeswijk and Sompolinsky 1996; Fusi et al. 2000b). There is preliminary evidence that this kind of noisy activity is actually good enough to generate small transition probabilities (Chicca and Fusi 2001).

Another advantage of transferring the load of generating stochasticity to the network dynamics is that the learning and forgetting speeds can be readily controlled by the statistics of the network activity. One simple control measure is the rate provoked by the stimuli, as discussed above. However, for the same mean frequencies, the synapse is rather sensitive also to the variability in the interspike intervals and to the degree of synchronization of the spike trains. This means that the network can easily and quickly switch from a single-shot learning modality (for highly correlated inputs) to slow learning (with uncorrelated inputs) and optimal storage capacity, without changing any inherent parameter of the synaptic dynamics (Chicca and Fusi 2001). This would fit nicely into a scenario in which the mean frequency encodes the information about the stimulus to be memorized, and the other statistical properties provide a triggering signal for learning (e.g., due to attention).

4.2 Biological plausibility

We showed that the network can perform well as an associative memory even if the analog depth of the synaptic efficacies is reduced to the extreme (bistable synapses). One might wonder whether biological synapses are discrete on long time scales. Such a discreteness is compatible with experimental data (see, e.g., Bliss and Collingridge (1993). Recently Peterson et al. (1998) provided preliminary experimental evidence that LTP is actually all or none, meaning that for each synapse only two efficacies can be preserved on long time scales.

A second issue concerns the protocols for inducing LTP and LTD. Recent experiments (see, e.g., Markram et al. 1997; Zhang et al. 1998) indicate that the precise timing of presynaptic spikes and postsynaptic action potentials can determine the directions of the synaptic change: LTP occurs when the presynaptic spikes precede postsynaptic spikes, and LTD when presynaptic spikes follow the postsynaptic spike within a short time window. The synaptic dynamics described here is compatible with such an experimental result: whenever a presynaptic spike precedes a postsynaptic action potential, the depolarization is likely to be near to the emission threshold (above the threshold V_H) and LTP is likely to be induced. When a presynaptic spike occurs just after the emission of a postsynaptic action potential, that model neuron is likely to be hyperpolarized (due to the spike after hyperpolarization), and the depolarization will tend to be below V_L. This produces LTD.

Of course the depolarization of our postsynaptic neuron should be considered as an effective variable: it is the only internal variable characterizing the neural dynamics of our simple integrate-and-fire neuron. Such a variable may not correspond directly to the depolarization of a complex biological neuron. There may instead be some internal variable which accounts also for other internal states and which is sensitive to particular time

intervals, before and after the postsynaptic spike. For instance, the effect of the postsynaptic action potential should definitely be incorporated in the model (in our model even the depolarization peak during the emission of the spike is totally ignored). However our study shows that it is easy to embed patterns of asynchronous activity in the synaptic matrix by reading the depolarization of the postsynaptic neuron, and our mechanism does not require strong nonlinearities as in a synaptic dynamics entirely based on spike timing (Senn 2002). Such a simple solution has probably not been overlooked by biology, and recent experiments have actually shown that there is an important dependence of LTP and LTD on the residual depolarization of the postsynaptic neuron (Sjöström et al. 2001).

4.3 Back to delay activity

A final issue is whether a network of integrate-and-fire neurons connected by the plastic synapses described here can reproduce the delay activity observed in cortical recordings and discussed in Sect. 1. The problem is complicated by the fact that the network activity induces synaptic changes, and the synaptic changes affect the network dynamics (for a review, see Del Giudice et al. 2002). The synaptic modifications can actually compromise the stability of the network dynamics. From recent work (Del Giudice et al. 2001; Amit and Mongillo (2002), it is becoming clear that the synaptic dynamics described here or a slightly modified version thereof is good enough to generate stable selective delay activity and to reproduce the phenomenology described in Fig. 1, provided that some extra regulatory mechanism for controlling the stability of the network dynamics is introduced. This mechanism might be short-term synaptic depression (Del Giudice et al. 2001), a progressive reduction of the external input coming from the sensory areas as the patterns of activity are learnt (Amit and Mongillo 2002), or, finally, an inherent property of the synaptic dynamics that automatically normalizes the statistics of the plastic weights (Fusi 2002).

Acknowledgements. Most of the work described in the first part of this manuscript was carried out in the group coordinated by D.J. Amit in Rome, in the context of the LANN initiative of INFN (see the cited works for more details). I would like to thank W. Gerstner, G. La Camera, S. Liu, A. Renart, and W. Senn: with their remarks they contributed to greatly improving a previous version of the manuscript.

Appendix: Mean spike rates vs depolarization distribution

The probabilities of occurrence r_a and r_b of the temporary modifications a and b control the direction in which the synapse is modified by the activity of the pre- and postsynaptic neurons. These probabilities depend on the statistics of the depolarization of the postsynaptic neuron under stimulation, and can be calculated analytically when the model of the neuron is simple. Here we focus on the simple model of an integrate-and-fire neuron with constant leak and a reflecting barrier at the resting potential (Fusi and Mattia 1999) which is described in Sect. 3.3. If such a neuron is injected with a Gaussian current characterized by its mean μ and variance σ^2, the stationary distribution of the depolarization $p(v)$ has a simple expression, and is given by (Fusi and Mattia 1999; Fusi et al. 2000a)

$$p(v) = \frac{\nu}{\mu}\left[\Theta(v-H)\left(1-e^{-\frac{2\mu}{\sigma^2}(\theta-v)}\right) + \Theta(H-v)\left(e^{-\frac{2\mu}{\sigma^2}H}-e^{-\frac{2\mu}{\sigma^2}\theta}\right)e^{\frac{2\mu}{\sigma^2}v}\right]$$

where ν is the mean firing frequency:

$$\nu = \left[\tau_r + \frac{\sigma^2}{2\mu^2}\left(e^{-\frac{2\mu\theta}{\sigma^2}}-e^{-\frac{2\mu H}{\sigma^2}}\right)+\frac{\theta-H}{\mu}\right]^{-1}$$

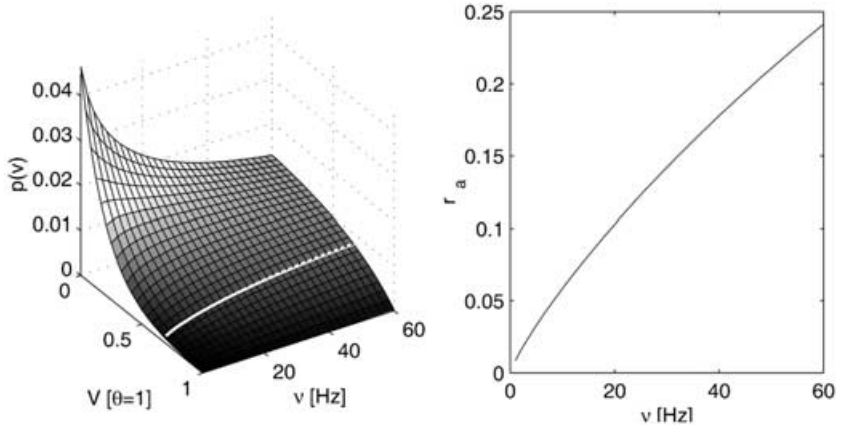

Fig. A1. Distributions of the postsynaptic depolarization, $p(v)$ (left), and r_a (probability of an upward jump) (right) vs $\nu(=\nu_{post})$ for the stimulation procedure described in the Appendix. In the case shown here the reset potential H has been chosen to be identical to the resting potential $V=0$. For low spike rates, $p(v)$ is concave and the probability of a depolarization near the spike-emission threshold $\theta(=1)$ is low. As ν_{post} increases, the distribution becomes convex, and increasingly uniform. The *white line* is the depolarization value which separates the regions corresponding to upwards (high V) and downwards (low V) jumps. Notice that in the case shown here, $V_H = V_L$. $r_a = \int_{V_H}^{\theta} p(v)$ (figure reproduced from Fusi et al. 2000a)

where τ_r is the absolute refractory period. r_a and r_b are given by the integral of $p(v)$ over the intervals $[V_H, \theta]$ and $[0, V_L]$, respectively. It is straightforward to compute these integrals analytically. The analytical expressions for $p(v)$ and v for the classical integrate-and-fire neuron with a leak proportional to the depolarization are also available (see, e.g., Brunel 2000).

μ and σ characterize the synaptic input and depend on the network interactions. We assume that v_{post} is changed by increasing or decreasing the average spike frequency of a subpopulation of presynaptic neurons (Fusi et al. 2000a). If the recurrent feedback of the postsynaptic neurons does not have a great effect on the network activity, then the parameters of the input current move along a linear trajectory in the (μ, σ^2) space. We chose μ as an independent parameter, and $\sigma^2 = J\mu + K$. In a network of excitatory and inhibitory neurons, in which in a spontaneous activity state the recurrent input is as large as the external input, we have that $J = J_E$ (the average coupling between excitatory neurons) and $K = v_0^I N_I J_I (J_I + J_E)$, where v_0 is the spontaneous activity of the N_I inhibitory neurons that are projecting to the postsynaptic cell (mean coupling J_I). r_a is plotted in Fig. A1. Since the external stimulus increases v_{post}, the distribution of the depolarization V changes in such a way that r_b decreases and r_a increases. Figure A1 exhibits the characteristics of the resulting distribution of V_{post}.

References

Amit DJ (1989) Modeling brain function. Cambridge University Press, New York

Amit DJ (1995) The Hebbian paradigm reintegrated: local reverberations as internal representations. Behav Brain Sci 18: 617–657

Amit DJ, Brunel N (1997a) Model of global spontaneous activity and local structured (learned) delay activity in cortex. Cereb Cortex 7: 237–252

Amit DJ, Brunel N (1997b) Dynamics of a recurrent network of spiking neurons before and following learning. Network 8: 373–404

Amit DJ, Fusi S (1992) Constraints on learning in dynamic synapses. Network 3: 443

Amit DJ, Fusi S (1994) Dynamic Learning in neural networks with material synapses. Neural Comput 6: 957

Amit DJ, Mongillo G (2002) Spike-driven synaptic dynamics generating working memory states, neural computation. Neural Comput (in press)

Amit DJ, Fusi S, Yakovlev V (1997) Paradigmatic attractor cell. Neural Comput 9: 1071–1093

Badoni D, Bertazzoni S, Buglioni S, Salina G, Amit DJ, Fusi S (1995) Electronic implementation of a stochastic learning attractor neural network. Network 6: 125–157

Battaglia FP, Fusi S (1995) Learning in neural networks with partially structured synaptic transitions. Network 6: 261

Bliss TVP, Collingridge GL (1993) A synaptic model of memory: long term potentiation in the hippocampus Nature 361: 31–39

Brunel (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci 8: 183–208

Brunel N, Carusi F, Fusi S (1998) Slow stochastic Hebbian learning of classes of stimuli in a recurrent neural network. Network 9: 123–152

Chicca E, Fusi S (2001) Stochastic synaptic plasticity in deterministic aVLSI networks of spiking neurons. In: Rattay F (ed) Proceedings of the World Congress on Neuroinformatics, Vienna, Austria, 24–29 September, pp 468–477

Del Giudice P, Mattia M (2001) Long and short-term synaptic plasticity and the formation of working memory: a case study. Neurocomputing 38: 1175–1180

Del Giudice P, Fusi S, Badoni D, Dante V, Amit DJ (1998) Learning attractors in an asynchronous, stochastic electronic neural network. Network 9: 183–205

Del Giudice P, Fusi S, Mattia M (2002) Modeling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. J Physiol (Paris) (in press)

Fusi S (1995) Prototype extraction in material attractor neural networks with stochastic dynamic learning. In: Rogers SK, Ruck DW (eds) Proceedings of SPIE 95, Orlando, Fla., pp 1027–1038

Fusi S (2001) Long term memory: encoding and storing strategies of the brain. Neurocomputing 38: 1223–1228

Fusi S (2002) Spike-driven synaptic plasticity for learning correlated patterns of asynchronous activity. In: Dorronsoro JR (ed) Proceedings of ICANN2002, Madrid, Spain, 27–30 August. (Lecture notes in computer science) Springer, Berlin Heidelberg New York

Fusi S, Mattia M (1999) Collective behavior of networks with linear (VLSI) integrate-and-fire neurons. Neural Comput 11: 643–662

Fusi S, Annunziato M, Badoni D, Salamon A, Amit DJ (2000a) Spike-driven synaptic plasticity: theory, simulation, VLSI implementation. Neural Comput 12: 2227–2258

Fusi S, Del Giudice P, Amit DJ (2000b) Neurophysiology of a VLSI spiking neural network. In: Amari S-I, Giles L, Gori M, Piuri V (eds) Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, 24–27 July, pp 121–126

Fuster JM 1995 Memory in the cerebral cortex. MIT Press, Cambridge, Mass.

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Wiley, New York

Heskes TM, Kappen B (1991) Learning processes in neural networks. Phys Rev A 44: 2718

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

Kahn PE, Wong KYM, Sherrington D (1995) A memory model with novel behaviour in sequential learning. Network 6: 415–427

Lattanzi G, Nardulli G, Pasquariello G, Stramaglia S (1997) Stochastic learning in a neural network with adapting synapses. Phys Rev E 56: 4567–4573

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

Miyashita Y (1993) Inferior temporal cortex: where visual perception meets memory. Annu Rev Neurosci 16: 245–263

Miyashita Y, Toshirio H (2000) Neural representation of visual objects: encoding and top-down activation. Curr Opin Neurobiol 10: 187–194

Nadal JP, Toulouse G, Changeux JP, 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 CCH, Malenka RC, Nicoll RA, Hopfield JJ (1998) All-or-none potentiation at CA3–CA1 synapses. Proc Natl Acad Sci USA 95: 4732–4737

Ricciardi LM (1977) Diffusion processes and related topics in biology. Springer, Berlin Heidelberg New York

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

Sejnowski TJ (1977) Storing covariance with nonlinearly interacting neurons. J Math Biol 4: 303

Shiryaev AN (1984) Probability. Springer, Berlin Heidelberg New York

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

Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. J Neurosci 1: 334–350

Tuckwell HC (1988) Introduction to theoretical neurobiology, vol 2. Cambridge University Press, Cambridge

Vreeswijk CA van, Sompolinsky H (1996) Chaos in neural networks with balanced excitatory and inhibitory activity. Science 274: 1724–1726

Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24: 455–463

Yakovlev V, Fusi S, Berman E, Zohary E (1998) Inter-trial neuronal activity in infero-temporal cortex: a putative vehicle to generate long term associations. Nat Neurosci 1: 310–331

Zhang LI, Tao HW, Holt CE, Harris WA, Poo M (1998) A critical window for cooperation and competition among developing retinotectal synapses. Nature 395: 37–44