



Long term memory: Encoding and storing strategies of the brain

Stefano Fusi

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

Abstract

Plastic material devices, either artificial or biological, should be capable of rapidly modifying their internal state to acquire information and, at the same time, preserve it for long periods (the stability–plasticity dilemma). Here we compare, in a simple and intuitive way, memory stability against noise of two different strategies based, respectively, on fully analog devices that accumulate linearly small changes and on systems with a limited number of stable states and threshold mechanisms. We show that the discrete systems are more stable, even with short inherent time constants, and can easily exploit the noise in the input to control the learning rate. We finally demonstrate the strategy by discussing a model of a biologically plausible spike-driven synapse. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Synaptic plasticity; Long term memory; Learning

1. Introduction

Material (artificial or biological) learning devices, like the synapses, 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 stimulations. In a realistic situation, the stimulations carrying relevant information are separated by long time intervals of noisy input which tends to erase the memory of the previously acquired information. Moreover the interference of novel stimulations with already acquired older ‘memories’ may give rise to memory loss (e.g. the oldest stimulations are forgotten to make room for the new ones). This is also known as the

E-mail address: fusi@cns.unibe.ch (S. Fusi).

stability–plasticity dilemma: the memory should be stable against irrelevant inputs (e.g. noise) for long periods and, at the same time, the internal state should be rapidly modified to acquire the information conveyed by the relevant inputs. This dilemma becomes particularly arduous when dealing with material memory devices that do not allow arbitrarily large time constants or parameters fine tuning, especially if the devices are small (e.g. it is reasonable to assume that permanent changes can not be arbitrarily small).

Here we show one possible encoding and storing strategy that solves this dilemma and we exemplify it by discussing a model of a spike-driven learning synapse. The strategy is based on the assumption that information to be coded is *redundant*: e.g. for the synapses this means that many cells on the dendritic tree carry similar information. We compare two possible scenarios: in the first each synapse is described in terms of one continuous internal variable x . In the absence of any stimulation, the value encoded by x is preserved forever. In the second, the synapse is discrete on long time scales: it has only a limited number of attracting stable states: when x drifts away from one of them, a recalling force drives it back to the closest stable state. To make a change permanent, the internal variable should cross some threshold, to be then attracted towards a different stable state. Let K be the number of stable states and Δx the minimal distance between two stable states.

2. Preserving information: The stability problem

We now consider the current generated by the synaptic inputs as the relevant variable. We assume that it is approximately the linear sum of many input neuronal activities a_i multiplied by the corresponding weights J_i , which, in turn, depend on the internal state of the synapses. Let I_N be the current induced by N neurons that encode the same information, i.e. that are activated in the same way by a generic stimulus ($a_i = a$ for $i = 1, \dots, N$):

$$I_N = \frac{1}{N} \sum_{i=1}^N J_i a_i = \frac{a}{N} \sum_{i=1}^N J_i.$$

If we start from the fully analog synaptic values and we clamp them to the closest stable states (see Fig. 1), the error on I_N goes as $\sim 1/(K\sqrt{N})$. If N is large enough (the code is redundant), the error becomes negligible and there is no relevant loss of information, which would be the only disadvantage of the discrete code. This is a known property of some neural networks (see e.g. [6]).

However, memory preservation is much more stable in the case of discrete synapses since the effects of noise do not accumulate. Let δt be the typical “response” time of the synaptic device, i.e. the time interval during which any change of an internal variable is established: the noise induces small jumps δx with probability p , either upwards or downwards, once every δt . The ratio $p/\delta t$ can also be seen as the rate of events that can induce permanent changes (e.g. the spikes). Let τ be the time constant of the recalling force: no matter how far x gets from one stable state, in a time of the order of τ , it is

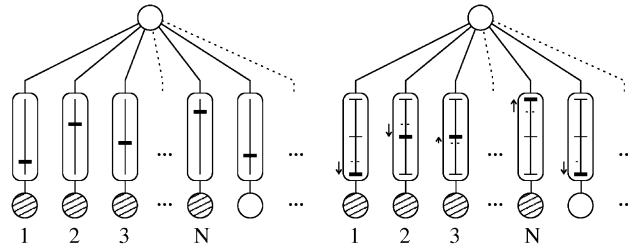


Fig. 1. Clipping synaptic efficacies: passing from fully analog synapses (left) to three-state synaptic efficacies (right) does not degrade much of the memory. The input neurons (below) are arranged in such a way that the first N neurons are driven by a generic stimulus to the same activity level. These neurons carry the same information (redundancy) for that specific stimulus. The efficacies are different because other uncorrelated stimuli, activating different subsets of neurons, had been previously encoded. When clipped to the closest stable state, the synapses are pushed up and down and the final “error” on the afferent current I_N , generated by N neurons, is equivalent to a noise whose amplitude scales as $1/\sqrt{N}$.

driven back to the closest stable state. For the fully analog synapse, after time T , the mean displacement is of the order of $\delta x \sqrt{pT/\delta t}$. Hence, to have an error of Δx , one has to wait a time of the order of:

$$T_{\Delta x} \sim \delta t \left(\frac{\Delta x}{\delta x} \right)^2 p^{-1}.$$

If the internal variable x hits one of the boundaries, this time is even shorter [4]. For the discrete synapse, the same error Δx is produced when a fluctuation drives the internal variable across the threshold. This happens with a probability $\sim (p\tau/\delta t)^h$ per τ where $h = \theta/\delta x$ is the number of jumps required to reach θ . Hence

$$T_{\Delta x} \sim \delta t \frac{\tau}{\delta t} \left(\frac{p\tau}{\delta t} \right)^{-h}, \tag{1}$$

which can be much longer than the time of the fully analog synapse, especially if p is small. It can be so long, that x practically never hits the boundaries (see Section 4). The best case is when h is maximal, i.e. when the synapse is binary. The same behavior could be obtained in the analog case by adding an extra device that triggers permanent modifications only if some threshold is crossed. However, there is accumulating experimental evidence that the single synaptic contacts are actually binary on long time scales [5].

3. Acquiring information: The plasticity problem

It was rather intuitive and well known that discreteness can increase stability without necessarily degrading memory performance. What was less clear is whether this is still true in case of on-line learning, when discrete synapses are updated after every stimulus presentation. Actually discreteness can be advantageous also in this

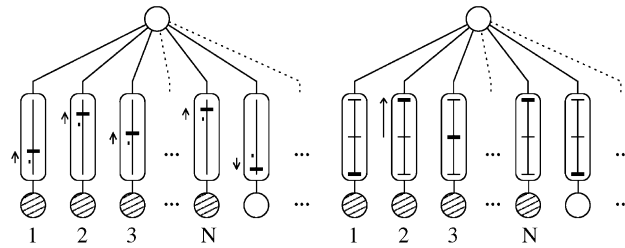


Fig. 2. Updating synaptic efficacies. The scheme is described in Fig. 1. Upon the presentation of a generic stimulus, the analog synapses (left) are potentiated by $\Delta = \Delta x/4$. Since the N synapses see the same pre- and post-synaptic activity they are all updated in the same way. The same change in I_N can be obtained in the discrete case by modifying only a fourth of the N synapses (synapse #2 in the figure). This can be obtained with a stochastic selection mechanism that updates each synapse with probability $q = 1/4$. Interestingly the presentation of a generic pattern interferes with the memory of other uncorrelated patterns in the same way in the two scenarios. Indeed, if f is the fraction of neurons activated by a different stimulus, the final change in its current would be $fN\Delta$ in the analog case and $fqN\Delta x$ in the discrete case. For a more general analytical study see [1].

case. Since the code is redundant, there is no need to modify all the synapses. If the fraction of synapses that are changed following each stimulation is small (slow learning), it is possible to better redistribute the synaptic ‘memory’ resources among the different patterns of stimulation and actually recover the optimal storage capacity even with binary synapses [1]. Slow learning is usually difficult because it is rather unlikely that the minimal change Δ inducible by the input is arbitrarily small. After M repetitions of the same signal, the minimal change of I_N would be $M\Delta$. In the discrete case, the noise superposed to the stimulations can turn in our favor by providing a triggering signal which selects in a local and unbiased way a small fraction of synapses to be changed. With the threshold mechanism of the discrete case, the input, at parity of signal, can induce or not a permanent change, i.e. a transition to a different stable state. In this case the minimal change in I would be $Mq\Delta x$, where q is the transition probability for each synapse. $q\Delta x$ can be much smaller than Δ and the average number of synapses changed after each repetition can be even < 1 (see Fig. 2). This scheme has the very attractive feature that it transfers part of the updating process outside the device (e.g. embedded in the input): q is not necessarily related to the intrinsic dynamics of the system. This can be a much better strategy, especially for small devices with short time constants.

4. Spike-driven synaptic plasticity

To demonstrate how the load of generating low probability events can be transferred outside the device, we discuss a model of a bistable ($K = 2$) spike-driven learning synapse which has been recently introduced [3]. The transitions between the two states are activity dependent and stochastic, even without any intrinsic noise source in the synaptic device. The synapse exploits the fluctuations in the inter-spike intervals,

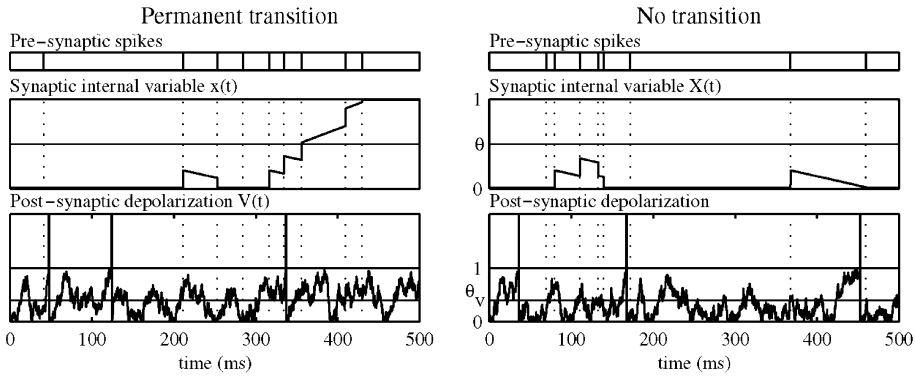


Fig. 3. Simulation of stochastic LTP. Pre- and post-synaptic neurons have the same mean rate and the synapse starts from the same initial value. At parity of activity (signal), the final state is different in the two cases.

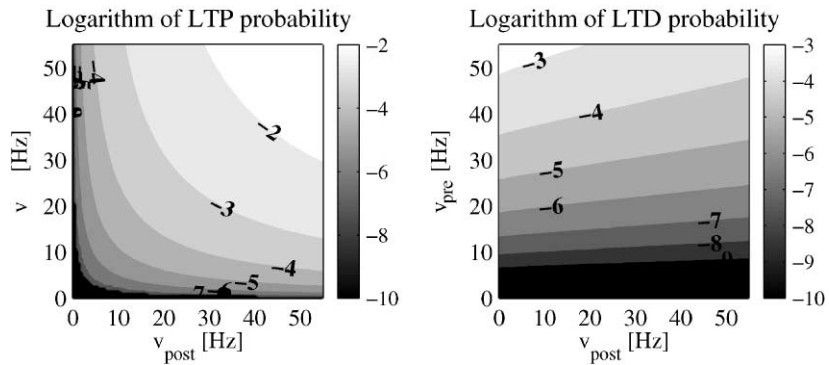


Fig. 4. Contour plots of LTP and LTD probabilities (q) on log scale vs pre- and post-synaptic neuron rates for a 500 ms stimulation. LTP occurs when pre- and post-synaptic rates are both high. Around the white plateau, P_{LTP} drops sharply and becomes negligible for spontaneous rates. The strong non-linearity allows to discriminate easily between relevant signals and background noise.

which are the results of the collective dynamics of the network. This noise is always superposed to the signal (pre- and post-synaptic mean frequencies) during the stimulation and is different from synapse to synapse. Each pre-synaptic spike drives the internal state x either up or down depending on whether the post-synaptic depolarization is above or below the threshold θ_V . LTP/LTD might occur or not at parity of mean pre-synaptic and post-synaptic activities (see Fig. 3). In this case p (see Eq. (1)) is the probability of coincidence of two events (e.g. a pre-synaptic spike and high depolarization) and hence can be very small. In Fig. 4 we show that the stochastic transitions between stable states are easily manipulable. In the presence of noise (low, spontaneous activity), the time to wait for a transition can be of the order of years, even if the longest time constant τ is of the order of 100 ms, whereas under stimulation (higher frequencies) the transition probabilities are easily controllable in the range

10^{-2} – 10^{-1} , as expected from Eq. (1). Extensive simulations of the learning process in networks of integrate-and-fire neurons connected by the proposed synapse are presented in [2].

We believe that this strategy based on the combination of discreteness and external stochasticity is a good general strategy for storing variables on long time scales and it is likely to underlie the basic mechanisms of many other biological small systems. Moreover this analysis shows that synaptic models in which single events (e.g. single spikes) modify permanently the synaptic efficacy can be hardly used as long term memory devices since the information acquired during the stimulation would be erased in a short time by the spontaneous activity.

References

- [1] D.J. Amit, S. Fusi, Learning in neural networks with material synapses, *Neural Comput.* 6 (1994) 957–982.
- [2] P. Del Giudice, M. Mattia, Long and short term synaptic plasticity and the formation of working memory: a case study, *Neurocomputing* 38–40 (2001) 1175–1180, this issue.
- [3] S. Fusi, M. Annunziato, D. Badoni, A. Salamon, D.J. Amit, Spike-driven synaptic plasticity: theory, simulation, VLSI implementation, *Neural Comput.* 12 (2000) 2227–2258.
- [4] G. Parisi, A memory which forgets, *J. Phys. A* 19 (1986) L617.
- [5] C.C.H. Petersen, R.C. Malenka, R.A. Nicoll, J.J. Hopfield, All-or-none potentiation at CA3-CA1 synapses, *Proc.Natl.Acad.Sci.* 95 (1998) 4732.
- [6] H. Sompolinsky, The theory of neural networks: the Hebb rule and beyond, in: L. van Hemmen, I. Morgenstern (Eds.), *Heidelberg Colloquium on Glassy Dynamics*, Springer, 1987.



Stefano Fusi was born in 1968 in Florence, Italy. He received his master degree in physics from the university of Roma in 1992. He had been working as a researcher in the National Institute of Nuclear Physics (INFN, Roma) from 1993 to 1999 and received a Ph.D. in physics from the Hebrew University of Jerusalem in 1999. He is currently working in the Institute of Physiology of Bern. His research interests include long-term synaptic plasticity, in vivo experiments on behaving monkeys, neuromorphic VLSI hardware and analytical studies of networks of spiking neurons.