

Strategies to associate memories by unsupervised learning in neural networks

E. J. Agnes, B. E. P. Mizusaki, R. Erichsen Jr. and L. G. Brunnet

Instituto de Física, Universidade Federal do Rio Grande do Sul, Brazil.

Abstract. In this work we study the effects of three different strategies to associate memories in a neural network composed by both excitatory and inhibitory spiking neurons, which are randomly connected through recurrent excitatory and inhibitory synapses. The system is intended to store a number of memories, associated to spatial external inputs. The strategies consist in the presentation of the input patterns through trials in: i) ordered sequence; ii) random sequence; iii) clustered sequences. In addition, an order parameter indicating the correlation between the trials' activities is introduced to compute associative memory capacities and the quality of memory retrieval.

Keywords: unsupervised learning; spiking neurons; homeostasis; STDP.

PACS: 87.18.Sn, 87.19.lg, 87.19.lj, 87.19.lv, 87.19.lw

INTRODUCTION

Memories, and everything that is processed by the brain, are associated with the connections among neurons. It is well established that learning includes mechanisms based on Hebb's hypothesis [1], which consists basically on potentiating and depressing connections between neurons with correlated and uncorrelated activities, respectively. The main mechanism of synaptic modification that is responsible for the emergence of associative memory in an unsupervised way is Spike-Timing-Dependent Plasticity (STDP) [2, 3], which is a hebbian-like rule. Additionally to the associative memory mechanism, the cells present homeostasis, which has been studied in theoretical works [4, 5].

It is known that, with some predefined connections between neurons and without synaptic plasticity [6, 7], a neural network of spiking neurons has a certain capacity to store memories. But how these specific connection matrices are acquired in an unsupervised way is yet unknown, and here we introduce three distinct methods to do so.

LEARNING

We have used the Izhikevich model [8] with Regular Spiking (RS) parameters for the 320 excitatory neurons and Fast Spiking (FS) for the 80 inhibitory neurons. The learning process was based on the work by Liu and Buonomano [5], where a trial, τ , is defined as the network response after a spatial input. All the synaptic modifications are applied after each trial, since the time window of a trial is less than 150 ms, which could match synaptic plasticity time scales.

As used in ref. [5], the probability of connection from an excitatory to another excitatory neuron, $P(exc. \rightarrow exc.)$, was set to 0.12. For the other connections, we used

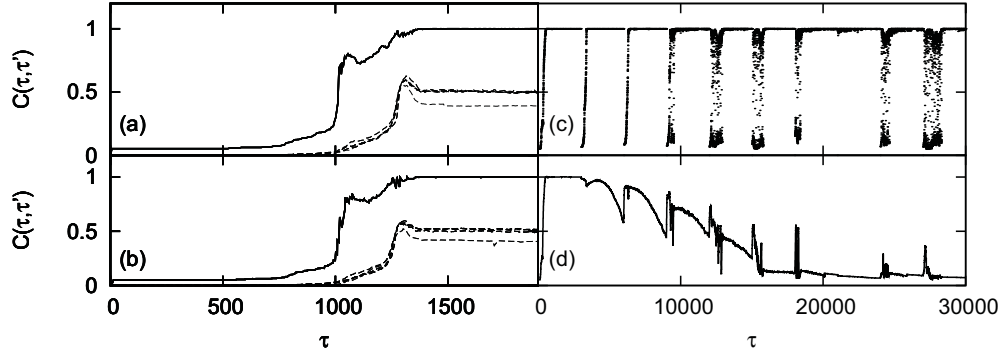


FIGURE 1. Order parameter, $C(\tau, \tau')$, versus learning trials, τ . In (a) and (b), the ordered sequence and the random sequence, respectively. The correlation in the clustered order in (c) and (d). Details of τ' on the text.

$P(exc. \rightarrow inh.) = 0.2$ and $P(inh. \rightarrow exc.) = 0.2$. Each input pattern consists on a randomly chosen set of 20 excitatory and 10 inhibitory neurons that fire within the first 10 ms of the trial. Initial synaptic weights were set to $W_{EE} = 0.04$ nS, $W_{EI} = 0.008$ nS and $W_{IE} = 0.1$ nS. Excitatory synapses were modified according to both homeostatic and STDP rules used in ref. [5], with the same parameter values. Short-term plasticity (STP) was implemented as described in ref. [8], with depression in all synapses.

The analysis was done with an order parameter which correlates the spike times of the excitatory neurons in two distinct trials. It is defined as

$$C(\tau, \tau') = \gamma^{-1} \sum_{i=1}^{320} \sum_{\{k,j\}} \exp \left[\frac{-(t_{i_k}^\tau - t_{i_j}^{\tau'})^2}{100} \right], \quad (1)$$

where $\gamma = MAX(320, S^\tau, S^{\tau'})$ is the normalization for the correlation value and $S^\tau = (1/320) \sum_i^{320} S_i^\tau$, a sum over only the 320 excitatory neurons. The sum over $\{k, j\}$ indicates that k th and j th spikes are close in time.

We used three different learning methods of spatial input presentation: i) ordered sequence; ii) random sequence; iii) clustered sequences. In the first strategy, each spatial pattern is presented in a ordered way through the trials. The second strategy consists in presenting the patterns in a random sequence and, in the third one, each pattern is presented successively for a defined number of trials, which was set to 3000 trials. The synaptic modifications are applied to excitatory connections, including a homeostatic plasticity and STDP, as described in ref. [5]. The homeostasis rule is used to increase the synaptic weights until the network presents a desired activity and thereafter to maintain a stable activity.

Figure 1 shows the evolution of the correlation $C(\tau, \tau')$, where $\tau' = \tau_{\mu=1}$ in (a) and (b) and $\tau' = \tau - 1$ in (c). A trial $\tau_{\mu=1}$ is the last trial - before τ - that the pattern $\mu = 1$ was the input pattern. In Fig. 1(d), the plot is the correlation of a trial between τ and $\tau + 1$ with the presentation of the pattern $\mu = 1$ without synaptic plasticity from homeostasis and STDP rules and the last trial with the presentation of pattern $\mu = 1$ with synaptic

plasticity. For $\tau < 3000$, the curve is the same as in Fig. 1(c) and for $\tau > 3000$, pattern $\mu = 2$ begins to be trained and the curve informs how much information about pattern $\mu = 1$ is being forgotten.

Full lines in graphs 1(a) and (b) correspond to τ which has as spatial input the pattern $\mu = 1$, indicating correlation when the same input is presented at different times. These lines converge to $C = 1$, showing that the spatiotemporal response from the network is equivalent when the same input is used. We tested the same with the other 4 patterns and the result is equivalent. Dashed lines correspond to correlation between trials with different input patterns and they do not converge to $C = 1$, indicating that different trained inputs evoke distinct spatiotemporal responses.

The plot in Fig. 1(c) shows that, using 30 patterns, the correlation converges to $C = 1$ for each one of them within a learning window of 3000 trials, but, as shown in Fig. 1(d), the network response to a learned memory vanishes as new memories begin to be stored.

CONCLUSIONS

We presented a neural network with unsupervised learning, simulated using the Izhikevich model and synaptic plasticity applied in trials. We defined three different ways of learning through trials - ordered, random and clustered sequences - and showed that the two first methods are equivalent for 5 patterns while the last method is robust for a large number of patterns but it presents a forgetting curve, which means that the memories are forgotten when new ones are trained. When the number of stored patterns increase in random and ordered sequences, some inputs do not develop spatiotemporal response (not shown). A detailed study of the network's behavior when increasing the number of input patterns is needed for a more accurate understanding of the problem.

ACKNOWLEDGMENTS

The authors acknowledge Brazilian funding agencies CNPq, Capes and Fapergs, also the Centro de Física Computacional - IF - UFRGS (CFCIF).

REFERENCES

1. J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation*. Addison-Wesley, Boston, 1991.
2. W. Gerstner, *et al.*, *Nature* **383**, 76–78 (1996).
3. C. Q. Bi, and M. M. Poo, *J. Neurosci.* **18**, 10464–10472 (1998).
4. G. G. Turrigiano, *Trends Neurosci.* **22**, 221–227 (1999).
5. J. K. Liu, and D. V. Buonomano, *J. Neurosci.* **29**, 13172–13181 (2009).
6. E. J. Agnes, R. Erichsen Jr., and L. G. Brunnet, *Phys. A* **391**, 843–848 (2012).
7. E. J. Agnes, R. Erichsen Jr., and L. G. Brunnet, *LNCS* **7552**, 145–152 (2012).
8. E. M. Izhikevich, and G. M. Edelman, *Proc. Natl. Acad. Sci. USA* **105**, 3593–3598 (2008).

Copyright of AIP Conference Proceedings is the property of American Institute of Physics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.