Noise induced structures in STDP networks

Gábor Szirtes, Zsolt Palotai, András Lőrincz [*]

Department of Information Systems, Eötvös Loránd University, Pázmány Péter sétány 1/C, Budapest, Hungary H-1117

Abstract

In this paper we study the emergent structures in networks with spike-timing dependent synaptic plasticity that are subject to external noise. We show that scale-free small-worlds can emerge in such noise driven Hebbian networks. This may explain the interplay between noise and Hebbian plasticity. We also argue that this model can be seen as a unification of the Watts-Strogatz and preferential attachment models of scale-free small-worlds.

Key words: small world, Hebbian learning, central nervous system, scale-free network

1 Introduction

In the last few years spike-timing dependent synaptic plasticity (STDP) (see e.g. [9] and references therein), which is an extension of the classical Hebbian learning mechanism, has been the subject of intensive research. It is supposed that spike based encoding is robust against noise. On the other hand, it is also supposed that a great part of the signals is noise due the high variance of the spike generating mechanisms (see e.g. [5]). Although several *functional* roles have been assigned to the ambiguous presence of noise (see, e.g., [6,10]), in this paper the possible *structural* role is explored. We study what network structures in simplified neural systems may emerge via noise induced STDP.

[*] corresponding author

Email addresses: gszirtes@inf.elte.hu (Gábor Szirtes,), zspalotai@vnet.hu (Zsolt Palotai,), lorincz@inf.elte.hu (András Lőrincz).

2 The model

Our network is made of N simplified integrate-and-fire-like 'neurons'. The network is *excited* by noise of *no* spatio-temporal structure. The dynamics of the internal activity is as follows:

$$\frac{\Delta a_i}{\Delta t} = \sum_j w_{ij} a_j^s + x_i^{(ext)}, \tag{1}$$

for $i = 1, 2, \ldots, N$. (N was 200 in our simulations.) Variable $x^{(ext)} \in (0,1)^N$ denotes the randomly generated external input, a_i is the internal activity of neuron i, w_{ij} is ij^{th} element of matrix $\mathbf{W}$, i.e., the connection strength from neuron j to neuron i. For $\Delta t = 1$, this is a discrete-time network, whereas in the limit ($\Delta t \rightarrow 0$) Eq. 1 becomes a set of coupled differential equations. Neuron j outputs a spike when a_j exceeds the threshold parameter θ. Spiking (denoted by superscript s) means that the output of the neuron a_j^s is set to 1. After spiking, a_j is set to zero. For simplicity, no refractory time is assumed. Amount of excitation received by neuron i from neuron j is $w_{ij} a_j^s$. Equation 1 describes the simplest form of 'integrate–and–fire' network models. Note that if $\Delta t = 1$ and the threshold is set to zero (i.e., if a neuron receives any excitation then it outputs a spike and its activity is reset to zero) then Eq. 1 represents 'binary neurons' without temporal integration. This is the simplest model within our framework. Synaptic strengths are modified as follows:

$$\frac{\Delta w_{ij}}{\Delta t} = \sum_{(t_i, t_j)} K(t_j - t_i) a_i^{t_i, s} a_j^{t_j, s}, \tag{2}$$

where kernel function K defines the synaptic efficacy as a function of the difference of the spiking times t_i, t_j. For inputs with no temporal correlation, only the ratio of the positive, i.e. strengthening, and negative, i.e., weakening areas (r_{A^+} and r_{A^-}, respectively) of the kernel function matters. Connection strength remain bounded only if $r_{A^+/A^-} < 1$. Temporal reshaping of the kernel would not modify the results.

In the first place, we have been interested in the emerging local and global connectivity structure of $\mathbf{W}$. The structure can be best described by a weighted graph, in which 'nodes' stand for neurons, while 'edges' or 'directed edges' denote the connections among them. We applied the so called connectivity length measure based on the concept of *network efficiency* [8]. This measure is more appropriate for weighted networks than the characteristic path length (L) and the clustering coefficient (C) of Watts and Strogatz [11]. The method of [8] is applicable for describing both global *and* local properties. *Local efficiency* between nodes i and j is $\epsilon_{ij} = 1/d_{ij}$, where d_{ij} corresponds to the

shortest path length throughout all of the possible paths from neuron j to i, where the path length between each connected pair of vertices is the inverse of the weight between them. The local harmonic mean *distance* for node i is defined as $D_h^l(i) = \frac{n^{(i)}(n^{(i)}-1)}{\sum_{j,k} \epsilon_{kj}^i}$ where $n^{(i)}$ is the number of neurons in subgraph $G^{(i)}$ (subgraph $G^{(i)}$ consists of all nodes l around neuron i with $w_{il} > 0$), ϵ_{kj}^i is the inverse of *shortest path length* between nodes k and j in $G^{(i)}$. The inverse of $D_h^l(i)$ describes how good the local communication is among the first neighbors of node i with node i removed. Small $D_h^l(i)$ means that node i has high clustering coefficient. The *global distance* is a good approximation of L. The mean *global distance* in the network is defined as $D_h^g = \frac{N(N-1)}{\sum_{i \neq j} \epsilon_{ij}}$ and provides a measure for the *size* (or the diameter) of the network.

It is also worth studying the distributions of the incoming and outgoing connections [1] as they may provide information about the scaling properties of the structure. The network is called 'small-world' if it is efficient locally *and* globally. It is called 'scale-free' if the distribution of connection strengths corresponds to a *power-law* distribution. Most scale-free nets are also small-worlds, provided that connection strength is not too sparse and no part of the network is isolated.
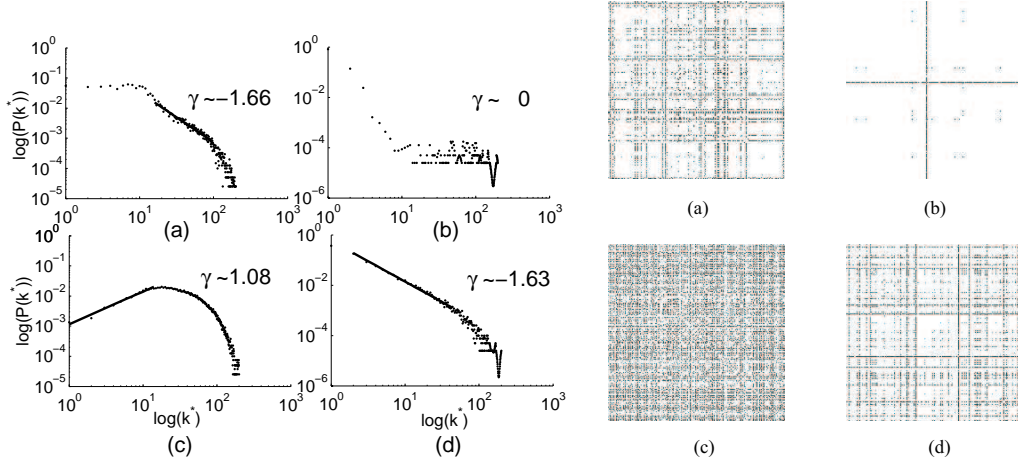
3   Results

Figure 1. Log-log plots for different parameters and the corresponding connectivity matrices
The four diagrams display typical distributions for parameters (a): $r_{A+/A-} = 0.1$ $r_{ex} = 0.3$, (b): $r_{A+/A-} = 0.1$ $r_{ex} = 0.6$,(c): $r_{A+/A-} = 0.6$ $r_{ex} = 0.3$ and (d): $r_{A+/A-} = 0.6$ $r_{ex} = 0.75$. Cases (a) and (d) are arbitrary examples from the power law region. The darker dots denote higher weights in the matrices.

Figure 1 summarizes our findings in different parameter regions. The distri-

bution of the weighted connections is plotted for different values of excitation level r_{ex} and parameter $r_{A+/A-}$. The corresponding connection matrices are also shown. While case (c) of Fig. 1 resembles a random structure, case (b) seems to be a winners-take-most network, in which only a few neurons dominate over the total amount of the connection strength. Cases (a) and (d) show strong clustering in a rather sparse structure and therefore correspond to scale-free small world networks characterized by their γ values (-1.66 and -1.63, respectively). Fig. 2 displays the comparison of the resulting structures of case (c) and (d) with a random net in which the same weights of the dynamic network have been randomly assigned to different node pairs. While the global connectivity length was almost the same for the Hebbian network and its corresponding random net in case (d) ($D_h^g \approx D_h^{gr} \approx 10$), the *local* connectivity values in the Hebbian network were much smaller indicating high clustering.

That is, for network (d), connectivity structure is sparse but information flow is still efficient. Tolerance against external noise r_{ex} was also studied. Increasing r_{ex}, the average value changes barely but suddenly drops to zero. For randomized networks, this drop is preceded by a monotonous increase of about an order of magnitude (Fig. 2).
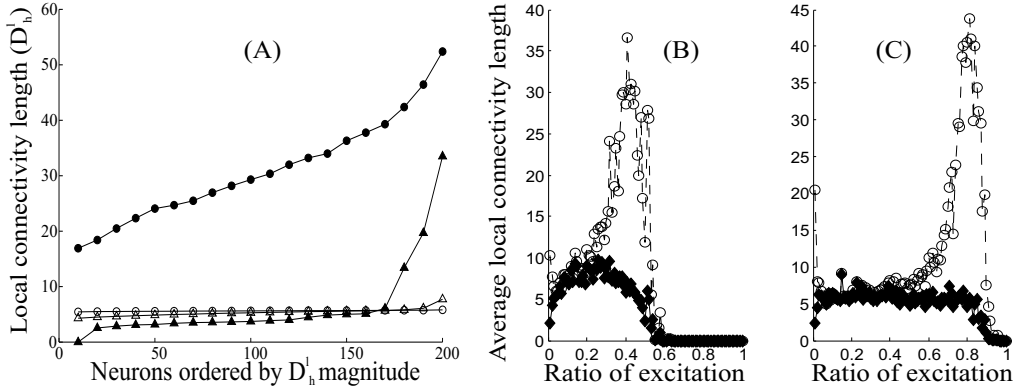

Figure 2. **Local connectivity length distances in ascending order**
(**A**) For better visualization every 10th data points are marked and the points are connected with a solid line. Lines with upward triangle markers: STDP learning. Lines with circles: same but randomly redistributed weights. Line with empty (solid) markers: noise driven Hebbian net of case (c) (case (d)). Random nets show similar local efficiencies, which is in contrast to the local efficiencies of the original nets. Global harmonic mean distances differ. For the original and for the randomized networks in case (c) $D_h^g \approx D_h^{gr} \approx 5.5$. For case (d) $D_h^g \approx D_h^{gr} \approx 10$. (**B**) and (**C**) Robustness against external excitation at different kernel constraints. (A): $r_{A+/A-} = 0.1$, (B): $r_{A+/A-} = 0.6$.

4

4 Conclusions

Considering the generality of our results, let us note that discrete time steps may correspond to short refractory periods. In this case our model with the only constraint on the ratio of the potentiation/depression kernel area falls under the category of 'input-restricted' STDP models as described in [4]. This analysis shows if the output spike interacts only with the n temporally nearest synaptic inputs, then in case of correlated inputs ('foreground synapses') and noise ('background synapses'), the learning dynamics is controlled by *stable* fixed-point. Some variance to our results may arise as a function of the refractory parameter. It is also important that the weight redistribution mechanism enabled by our constraint $r_{A^+/A^-} < 1$ is empirically observed: potentiation for weak synapses is favored whereas strong synapses tend to be depressed (see, e.g., [3]). Our model is also able to selectively potentiate higher-rate synaptic inputs without embedding any specific competitive mechanism: in spite of our simple set up, the emerging results seem to be fairly general. Although this paper is intended to study the role of noise in structural changes, the results can be related to other models describing the emergence of small worlds. With the following (strong) constraining assumptions, our model can be transformed to the 'preferential attachment' model of Barabási et al [2]:

- Let us suppose that at $t = 0$ there are n out of N nodes have at least one connection to other nodes.
- Let the changes in activity and connection strength be discrete.
- Spikings of the cloud of $(N - n)$ isolated nodes can be considered independent and for short temporal kernel spiking *coincidence* is negligibly small within the cloud as compared to the connected set. In turn, isolated neurons tend to develop new connections toward the connected sub-net.
- The activity of the connected neurons is strongly dependent on the spiking activity of the 'neighbors', which implies that spiking *probability* is (approximately) proportional to the recurrent activity controlled by the incoming connection distribution.
- The already established connections between nodes are kept steady and may not change by time, which is also assumed by the preferential attachment model.

This latter constraint is not realistic in any model. Note,however, that random rewiring of old connections can also give rise to scale-free network structure, as it was shown by Watts and Strogatz [11]. In turn, the incremental growing of the connected sub-net (by connecting new isolated neurons) and the weight redistribution property of our Hebbian nets can be seen as the synthesis of the preferential attachment mechanism and the rewiring mechanism. Nonetheless, we have demonstrated that our model can produce scale-free small world structures without the explicit requirement on growing, and without a direct

5

mechanism of weight rewiring. The profound implication of our results is that *local* (Hebbian) learning rules in a noise induced network may be sufficient to form and maintain an efficient structure in terms of information flow. In this aspect our model differs from existing models, such as the model on preferential attachment, the global optimization scheme [7], and also from the original Watts and Strogatz model. Considering the possible role of noise, we conjecture that the sustained nature of noise and the competition imposed by appropriate $r_{A+/A-}$ values are the two relevant components of plasticity and learning.

References

[1] R. Albert and A.L. Barabási, *Statistical mechanics of complex networks*, Reviews of Modern Physics **74** (2002), 47–91.

[2] A. L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), 509–512.

[3] G. Q. Bi and M. M. Poo, *Synaptic modifications in cultured hippocampal neurons: dependence on spike-timing, synaptic strength, and postsynaptic cell type*, Journal of Neuroscience **18** (1998), 10464–10472.

[4] A.N. Burkitt, H. Meffin, and D.B. Grayden, *Spike timing-dependent plasticity: The relationship to rate-based learning for models with weight dynamics determined by a stable fixed-point*, Neural Computation, to appear.

[5] D. Ferster, *Is neural noise is just a nuisance*, Science **273** (1996), 1812.

[6] N. Ho and A. Destexhe, *Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons*, Journal of Neurophysiology **84** (2000), 1488–1496.

[7] R. Ferrer i Cancho and R.V. Sol, *Optimization in complex networks*, arXiv:arch-ive.cond-mat/0111222, 2001.

[8] V. Latora and M. Marchiori, *Efficient behavior of small-world networks*, Physical Review Letters **87** (2001), no. 19.

[9] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes - exploring the neural code*, MIT Press, Cambridge, MA, 1996.

[10] M. C. W. van Rossum, G. Q. Bi, and G. G. Turrigiano, *Stable Hebbian learning from spike timing-dependent plasticity*, Journal of Neuroscience **20** (2000), no. 23, 8812–8821.

[11] D.J. Watts and S.H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature **393** (1998), 440–442.