

# Emerging evolutionary features in noise driven STDP networks?

Zsolt Palotai

Department of Information Systems,  
Eötvös Loránd University,  
Pázmány Péter sétány 1/C,  
Budapest, Hungary H-1117.  
E-mail: zspalotai@vnet.hu

Gábor Szirtes

Department of Information Systems,  
Eötvös Loránd University,  
Pázmány Péter sétány 1/C,  
Budapest, Hungary H-1117.  
E-mail: gszirtes@inf.elte.hu

András Lőrincz

Department of Information Systems,  
Eötvös Loránd University,  
Pázmány Péter sétány 1/C,  
Budapest, Hungary H-1117.  
E-mail: lorincz@inf.elte.hu

**Abstract**—In this paper we study the emergent structure of networks in which spike-timing dependent synaptic plasticity is induced only by external random noise. We show that such noise driven Hebbian networks are able to develop a broad range of network structures, including scale-free small-world networks. The development of such network structures may provide an explanation of the role of noise and its interplay with Hebbian plasticity. We also argue that this model can be seen as a unification of the famous Watts-Strogatz and preferential attachment models of small-world and scale-free nets. Our results may support Edelman's idea on that the development of central nervous system may have evolutionary components.

## I. INTRODUCTION

In the last few years spike-timing dependent synaptic plasticity (STDP) (see e.g. [1], [2] and references therein), which is an extension of the classical Hebbian learning mechanism, has been the subject of intensive research. It is broadly accepted in the machine learning literature [3], [4], [5] and is strongly supported in neuronal modelling [6], [7] that spike based encoding can be efficient in compression, allows for sparse representation, low energy consumption and that it can be robust against noise. The last property seems to be indispensable knowing the stochastic behavior of the neurons and of the external environment. But if noise should be suppressed, how come that a great part of the signals propagating through several brain regions experienced in different species (ranging from frogs to primates) is considered to be internally generated noise [8], [9]? What can be the reason for counteracting the perfect information processing and transmission? One possible role of noise in the nervous system is provided by the recognition that noise can enhance the response of nonlinear systems to weak signals, via a mechanism known as stochastic resonance (see, e.g., [10]). However, noisy functioning may have additional roles. For example, it has been shown that synaptic background activity may promote distinguishing very similar inputs [11]. It has been also demonstrated [12] that strict conditions on stability of Hebbian mechanisms can be released by introducing random external noise instead of maintaining competition among neurons over the input sets. In this paper we address the question whether noise may have any impact on *structural* changes.

One of the most exciting findings in recent scientific research is that many complex interactive systems possess a surprising *structural* and functional property: the emergence of scale-free small-world networks (SFNs) of the building blocks. Such SFNs may be found in distinct fields ranging from metabolic reaction chains to social relation systems (see, e.g. [13] and references therein). One may find SFNs in neurobiology as well. For example, the completely mapped neural network of the nematode worm *C. elegans* [14] is considered to form a small-world network [15]. Intriguingly, most systems featuring scale-free small world interactions, are considered selective or evolutionary systems. An outstanding example is the Internet, which displays this network structure at the hardware level of servers and also at the level of web pages [16], [17], [18]. This fascinating self-organizing system has inspired several studies and models. The original model of the the World Wide Web (WWW) by Watts and Strogatz [19] explored random restructuring of the links among a finite number of 'nodes'. Barabási and his colleagues introduced the concept of *preferential attachment* to model the WWW [20], [16]. The idea has been extended to other types of networks [17] and the focus has been put on the search of general mechanisms underlying the development of these distinct connection systems.

In the following, we examine what network structures may emerge in a simplistic neural system fed by noise and subject to STDP-type Hebbian learning. In what follows, such noise driven Hebbian networks will be called NDH nets for short. We shall find that under certain conditions, the NDH network searches the input space uniformly on every scale for novel information.

## II. PRELIMINARIES

### A. Model description

We assume that the network is *sustained* by inputs with no spatio-temporal structure; that is the input is random noise. Our models consist of  $N$  number of simplified integrate-and-fire-like 'neurons' or nodes. The dynamics of the internal

activity is written as

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

for  $i = 1, 2, \dots, N$ . ( $N$  was 200 in our simulations.) Variable  $x^{(ext)} \in (0, 1)^N$  denotes the randomly generated input from the environment,  $a_i$  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  $i$ . If  $\Delta t = 1$  then we have a discrete-time network and each parameter has a time index, or if  $\Delta t$  is infinitesimally small then Eq. 1 becomes a set of coupled differential equations. Neuron  $j$  outputs a spike (neuron  $j$  ‘fires’) when  $a_j$  exceeds a certain level, the threshold parameter  $\theta$ . Spiking means that the output of the neuron  $a_j^s$  (superscript  $s$  stands for ‘spiking’) is set to 1. After firing,  $a_j$  is set to zero at the next time step for the discrete-time network. For the continuous version of Eq. 1,  $a_j$  is set to zero after a very small time interval. For simplicity, no refractor 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 which is still plausible from a neurobiological point of view. Note that if  $\Delta t = 1$  and the threshold is set to zero (i.e., if a neuron receives any excitation then it fires and is reset to zero) then Eq. 1 represents ‘binary neurons’ without temporal integration. This can be seen as the simplest model within our framework. Also, if the threshold is kept and if  $a_i$  is set to zero before each time step, irrespective if the  $i^{th}$  neuron fires or not, then the original model of McCullough and Pitts [21] is recovered.

Synaptic strengths are modified as follows:

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

where  $K$  is a kernel function which defines the influence of the temporal activity correlation on synaptic efficacy,  $t_i, t_j$  the spiking times of neuron  $i$  and  $j$ , respectively and  $a_i^{t_i, s}$  is the firing activity of neuron  $i$  at time  $t_i$ . Initially all  $w_{ij}$  is set to a small identically and independently distributed random number.  $\Delta w_{ij}/\Delta t$  may be taken over discrete or over infinitesimally small time intervals. Some possible kernels are depicted in Fig. 1. The kernel is a function of the time differences. Because, in our case, the input is noise with no temporal correlation, only the ratio of the positive (strengthening) and the negative (weakening) areas of the kernel function ( $r_{A^+/A^-}$ ) should count. Temporal grouping and reshaping of the kernel would not modify our results as long as the aforementioned ratio is kept constant and the input is pure noise. For this special case, the difference between the two kernel types of Fig. 1 does not have much impact on the temporal evolution of our model network. It should be noted that including inputs with spatiotemporal structure and other, known details of synaptic plasticity mechanisms, this kernel shape independence will not hold. Our only constraint on the kernel, namely the constraint that  $r_{A^+/A^-} < 1$ , is required to constrain weights. This constraint redistributes weight strengths.

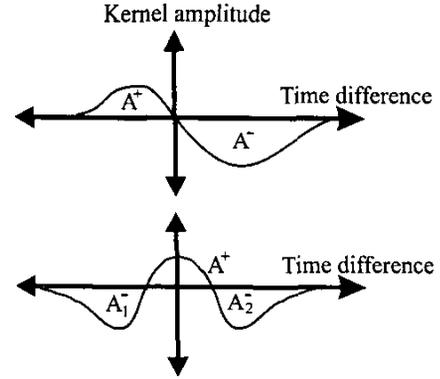


Fig. 1. **Kernel functions.** Two temporal kernels as a function of time difference between spiking time of neuron  $i$  and  $j$  ( $t_i - t_j$ ). Relevant parameter of the shape for noise-sustained systems is the ratio ( $r_{A^+/A^-}$ ) of the areas (sums of positive and negative parts/components) of the kernel,  $A^+$  and  $A^-$ , respectively ( $r_{A^+/A^-} = A^+/A^-$ ).

### B. Gauges of evaluation

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* [22]. This measure is more appropriate for weighted networks than the characteristic path length ( $L$ ) and the clustering coefficient ( $C$ ) of Watts and Strogatz [19] useful for networks with equal weights. The method of [22] 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  $w_{il} > 0$ ),  $\epsilon_{kj}^i$  is the inverse of *shortest distance* 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  $L$ .

Many different networks belong to the same structural family regarded as ‘small-worlds’. Their most characteristic feature is that they are efficient locally and globally, too. While local and global connectedness are useful tools to characterize a network architecture, it is worth investigating the degree distributions of the incoming and outgoing connections as well [13]. They may provide information about the scaling of different properties of the given structure, like the change of the diameter as a function of the number of nodes. One particular subfamily of small-world nets can also be characterized as ‘scale-free’ networks, because their most

significant properties scale according to power-law with the connection number distribution. Most scale-free nets are also small-worlds, provided that connection strength is not too sparse and basically no part of the network is isolated.

### III. RESULTS

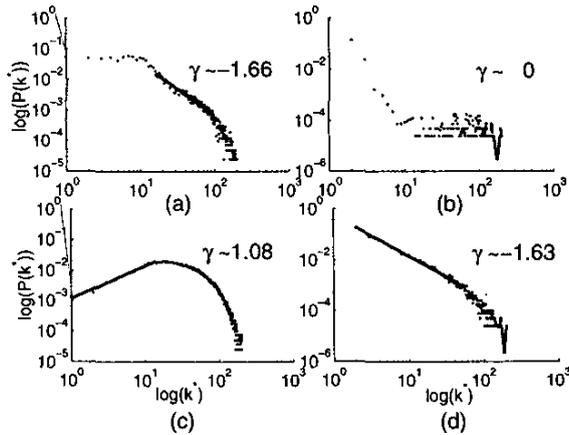


Fig. 2. **Log-log plots for different parameters.** The four diagrams display typical distributions for parameters (a):  $r_{A^+/A^-} = 0.1$   $r_{ex} = 0.3$ , (b):  $r_{A^+/A^-} = 0.1$   $r_{ex} = 0.6$ , (c):  $r_{A^+/A^-} = 0.6$   $r_{ex} = 0.3$  and (d):  $r_{A^+/A^-} = 0.6$   $r_{ex} = 0.75$ . Cases (a) and (d) are arbitrary examples from the power law region.

Figure 2 summarizes our findings in different parameter regions and starting the network from  $\mathbf{W} \equiv \mathbf{0}$ . The figure displays the emergence of scale free nets as a function of the excitation level  $r_{ex}$ , the average ratio of neurons receiving excitation from the environment, and the ratio of the area of potentiation to the area of depression ( $r_{A^+/A^-}$ ) in kernel  $K$ . The length of the scale-free regions was determined by first plotting the distribution of the sum of the weights of outgoing connections (averaged over 20 runs, each run contains 10000 samples) for every parameter set studied. Results are depicted on loglog plot. Supposing a power-law distribution ( $P(k^*) \approx k^{*\gamma} e^{-k^*/\xi}$ , where  $k^*$  denotes the discretized values of the connection strength), a linear fitting was made to approximate  $\gamma$ . The width of the scale-free region was estimated by the length of the region with power-law distribution relative to the full length covered on the log scale. Maximum error of the linear fit was set to  $10^{-3}$  STD. That is, for 100 discretization points, the width of a region spreading an order of magnitude on the loglog plot is equal to 0.5. Figure 3 shows the corresponding connection matrices. While case (c) 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. However, cases (a) and (d) show strong clustering in a rather sparse structure and therefore correspond to scale-free small world networks characterized by their  $\gamma$  values ( $-1.66$  and  $-1.63$ , respectively). Figure 3 depicts the corresponding connectivity matrices.

With the help of the above introduced connectivity length measures we studied also the emerging network structures as

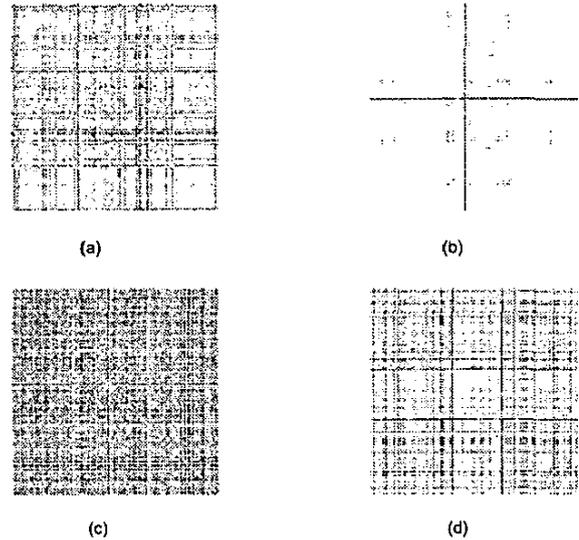


Fig. 3. **Connectivity matrices.** The four diagrams display connectivity matrices corresponding to the cases in Fig. 2. Cases (a) and (d) are arbitrary examples from the power law region, these subfigures approximate scale-free connectivity structures.

a function of the following parameters: (i) the magnitude of the external excitation and (ii) the strengthening–weakening area ratio ( $r_{A^+/A^-}$ ) of kernel  $K$ . It can be seen that many connection weights have been vanished and it has made possible to talk about ‘subgraphs’ with local connectivity. As an extreme case of the general model, the binary neuron model was also investigated and no important difference were found.

Fig. 4 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. The diagram highlights the emerging small-world properties, i.e., small local connectivity values (high clustering coefficients) for case (d) and the opposite for case (c). Global connectivity lengths were about the same for the original and for the randomized networks.

### IV. DISCUSSION AND OUTLOOK

#### A. On the generality of the model

Consider that discrete time steps of our model corresponds to short refractory periods. Then our model with the constraint on the ratio of the potentiation/depression kernel area falls under the category of ‘input-restricted’ STDP models as described by Burkitt and his colleagues [23]. Their analysis shows if the output action potential 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 presented STDP based system can be related to well-established rate-based Hebbian learning models: As it has been mentioned, our constraint on the kernel

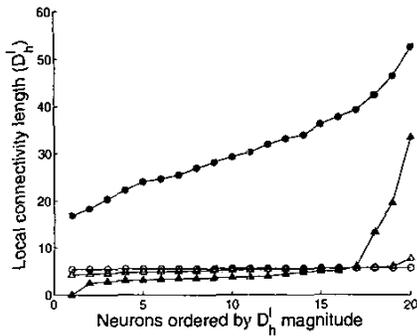


Fig. 4. Local connectivity length distances in ascending order. For better visualization not all data points are marked and the points are connected with solid lines. Lines with upward triangle markers represent STDP learning in NDH networks. Lines with circles represent randomized networks: the same set of weights are randomly redistributed amongst the nodes of the network. Lines with empty markers: weight set of non-scale-free-like NDH net of case (c) (Fig. 3). Lines with solid markers: weight set of scale-free like NDH net (case (d) of Fig. 3). **Note 1:** The global harmonic mean distances for the original and for the randomized networks are about the same: in case (c) of Fig. 3  $D_h^g \approx D_h^{gr} \approx 5.5$ , whereas in case (d) of Fig. 3  $D_h^g \approx D_h^{gr} \approx 10$ . **Note 2:** The scale-free-like network (case (d)) and the non-scale-free-like one (case (c)) achieve about the same local connectivity length distances, but the latter has a much larger number of non-vanishing connections (Fig. 3(c)).

(i.e.,  $r_{A^+ / A^-} < 1$ ) redistributes synaptic weights. Empirical data indicate that indeed, there are mechanisms to redistribute weight strengths; potentiation for weak synapses is favored whereas strong synapses tend to be depressed (see, e.g., [24], [25], [12]). The model is also able to selectively potentiate higher-rate synaptic inputs without embedding any specific competitive mechanism. As a consequence, in spite of our simple set up, the emerging results seem to be fairly general.

#### B. Relation of noise driven Hebbian nets to other models

Although this paper is intended only to show some experimental (simulation) results on noise induced network structures of simplified neuron models, the results can be related to other, well-known mechanisms, too. In the following we show that under some (strong) constraining assumptions, our model can be transformed to the model of Barabási et al [16], the model of preferential attachment. The following assumptions are made to enable the above-mentioned transition:

- Let us suppose that at  $t = 0$  there are  $N$  nodes, from which only  $n$  nodes ( $n \ll N$ ) have at least one connection to other nodes.
- Let the changes in activity and connection strength be discrete by choosing both the weakening and strengthening step of the kernel to be of unit strength.
- Spikings of the cloud of  $(N - n)$  isolated nodes can be considered independent and the spiking probability is small. For such isolated nodes, only the external input, the second part of the right hand side of Eq. 1, counts. Furthermore, the coincidence of spiking of two isolated neurons is negligibly small if the temporal kernel is short. At any time instant, when a neuron of the isolated cloud fires the nodes of the connected set may fire or not.

If no coincidence occurs then there will be no change in the network. However, such coincidences are much more likely given the connectivity structure between the neurons of the connected set. This is so, because if one neuron fires then there is a chain of firing amongst these neurons. In turn, the development of new connections between two isolated neurons is not likely, whereas isolated neurons tend to develop new connections toward the connected sub-net.

- In contrast to the cloud, the activity of the connected neurons is strongly dependent on the spiking activity of the 'neighbors'. If firing starts in the connected cloud of neurons then the first term of the right hand side of Eq. 1 will dominate the resulting firing chain. Input initiates the firing chain, whereas recurrent excitation from other nodes control that chain. In turn, the probability of firing can be taken as (approximately) proportional to the recurrent activity, controlled by the incoming connection distribution.
- Having established a connection between two nodes, it is kept steady and may not change by time. This is a strong assumption, which is tacitly assumed by the original model of preferential attachment, too.

This latter constraint does not seem to be realistic in any model. There is no reason that for a growing connection structure should remain steady for old connections. Note, however, that random rewiring of old connections can give rise to scale-free network structure, too. In fact, this rewiring mechanism is the original model of Watts and Strogatz [19]. As it was noted at the very beginning (see Section II) our model has an intrinsic weight redistributing property originated by the constraint that  $r_{A^+ / A^-} < 1$ . In turn, the incremental growing of the connected sub-net (by connecting new isolated neurons) and the weight redistributing property of NDH nets can be seen as the synthesis of the preferential attachment mechanism with continuous new entries in the model of preferential attachment [16] and the rewiring mechanism of the model of Watts and Strogatz [19]. That is, constraining our model lead to a combination of two models both generating small-world structures. Nonetheless, by means of numerical simulations we have shown that our model can produce such connection structures without the explicit requirement on growing, and without a direct mechanism of weight rewiring.

#### C. Remarks on evolutionary systems

Interestingly, all the listed examples, one way or the other, usually are also regarded as evolutionary systems. In our particular case, the obtained results can also be interpreted in an evolutionary context by reconsidering Edelman's alternative neuronal group selection theory [27] about the fundamental role of selection during and after development of the nervous system. According to Edelman, a theory to describe a system's temporal change can be considered as 'selectionist', if it includes the following components:

- source of diversification leading to variants,

- a means for encounter with an environment not initially categorized,
- a means for differential amplification over some period of time of those variants in a population that have greater adaptive value.

It is no surprise that a system with these features falls into the class of evolutionary systems as far as we look at the system as a whole. In the nervous systems, there are at least two types of temporal changes serving the first requirement: Diversification can occur via the emergence of redundant connectivity during development *and* via the modification of synaptic efficacy during life-time learning. The main thesis of this paper is to demonstrate how diversification can be realized by noise under STDP rules. The second requirement is fulfilled if the pool of the not yet seen input patterns is not limited.<sup>1</sup>

Now, we can argue that noise in the nervous system has an important role: Noise has no spatiotemporal structure. Thus, obviously it cannot induce ‘learning’ in general sense. However, noise with STDP — according to our results — gives rise to a search mechanism, which scans at all scales simultaneously. Search in a scale-free manner can be the most efficient if no structural formation is known in advance. The searching feature is robust: The noise generated structure is changing rapidly; results depicted in the figures are averaged over several runs. The continuous change induced by noise can be interpreted in the following way. The noise *together* with the proportionally expressed LTD and LTP mechanisms yields a continuous sparsification and regeneration of the connections. LTP ‘chooses’ sound patterns (see, e.g., [26]), whereas LTD helps to ‘forget’ those patterns and maintains a competition amongst patterns. synchronous patterns or pattern series are quickly learned by NDH nets and approximately stable connectivity patterns may emerge. Noise, in this case, may modify the connectivity strengths and search may be performed ‘around’ an average stable connectivity pattern. Also, the noise may help the system to escape from local minima. Noisy Hebbian learning, under the condition  $r_{A+}/A- < 1$ , is able to simultaneously learn correlations and make selections among the discovered structures or patterns. Consider the noiseless case: The rank of the developing connectivity structure cannot grow beyond the rank of the correlation matrix, which can be less than the number of neurons. Our results suggest that the noise in NDH nets serves to scan the remaining part of the space – the space with dimension of ‘number of neurons minus the rank of the correlation matrix’ – for novel correlations uniformly on every scale.

## V. CONCLUSION

As far as other evolving networks are considered, the profound implication of our result is that *local* (Hebbian) learning rules may be sufficient to form and maintain an efficient network in terms of information flow. In this aspect our model differs from existing models, such as the model on

<sup>1</sup>Considerations about the third requirement are beyond the scope of the present study.

preferential attachment [16], the global optimization scheme [28], and also from the original Watts and Strogatz model.

In summary, we have demonstrated that small-world architecture with scale-free domains may emerge in sustained networks under STDP Hebbian learning rule without any other specific constraint on the evolution of the net. According to our results, evolution and plasticity of neural networks may be maintained by noise randomly generated within the central nervous system. 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.

## ACKNOWLEDGEMENTS

This work was partially supported by the Hungarian National Science Foundation, under Grant No. OTKA 32487.

## REFERENCES

- [1] G. Bi and M. Poo, “Synaptic modification of correlated activity: Hebb’s postulate revisited,” *Ann. Rev. Neurosci.*, vol. 24, pp. 139–166, 2001.
- [2] W. Gerstner and W. M. Kistler, *Spiking Neuron Models*. Cambridge University Press, August 2002.
- [3] W. Singer and C. M. Gray, “Visual feature integration and the temporal correlation hypothesis,” *Annual Review of Neuroscience*, vol. 18, pp. 555–586, 1995.
- [4] W. Maass, “Fast sigmoidal networks via spiking neurons,” *Neural Computation*, vol. 9, no. 2, pp. 279–304, 1997.
- [5] J. N. K. Bohte, S. M. and H. L. Poutre, “Spike-prop: backpropagation for networks of spiking neurons,” in *ESANN’2000*, 2000, pp. 419–425.
- [6] W. Gerstner, R. Kempfer, J. L. van Hemmen, and H. Wagner, “A neuronal learning rule for sub-millisecond temporal coding,” *Nature*, vol. 383, pp. 76–78, 1996.
- [7] S. Thorpe, A. Delorme, and R. V. Rullen, “Spike-based strategies for rapid processing,” *Neural Networks*, vol. 14, no. 6–7, pp. 715–725, 2001.
- [8] D. Ferster, “Is neural noise is just a nuisance,” *Science*, vol. 273, p. 1812, 1996.
- [9] K. D. Miller and T. W. Troyer, “Neural noise can explain expansive, power-law nonlinearities in neural response functions,” *Journal of Neurophysiology*, vol. 87, pp. 653–659, 2002.
- [10] N. Bibikov, “Addition of noise enhances neural synchrony to amplitude-modulated sounds in the frog’s midbrain,” *Hearing Research*, vol. 173, pp. 21–28, 2002.
- [11] N. Ho and A. Destexhe, “Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons,” *Journal of Neurophysiology*, vol. 84, pp. 1488–1496, 2000.
- [12] M. C. W. van Rossum, G. Q. Bi, and G. G. Turrigiano, “Stable Hebbian learning from spike timing-dependent plasticity,” *Journal of Neuroscience*, vol. 20, no. 23, pp. 8812–8821, 2000.
- [13] R. Albert and A. Barabási, “Statistical mechanics of complex networks,” *Reviews of Modern Physics*, vol. 74, pp. 47–91, 2002.
- [14] T. Achacoso and W. Yamamoto, *AY’s Neuroanatomy of C. Elegans for Computation*. Boca Raton, FL: CRC Press, 1992.
- [15] M. Marchioria and V. Latora, “Harmony in the small-world,” *Physica A*, vol. 285, pp. 539–546, 2000.
- [16] A. L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, 1999.
- [17] A. L. Barabási, R. Albert, and H. Jeong, “Scale-free characteristics of random networks: The topology of the world wide web,” *Physica A*, vol. 281, pp. 69–77, 2000.
- [18] J. Kleinberg, “Authoritative sources in a hyperlinked environment,” in *9th ACM-SIAM Symposium on Discrete Algorithms*, 1998, pp. 668–677.
- [19] D. Watts and S. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, pp. 440–442, 1998.
- [20] R. Albert, H. Jeong, and A. L. Barabási, “Diameter of the world wide web,” *Nature*, vol. 401, pp. 130–131, 1999.
- [21] W. McCullough and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bull. Math. Biophysics*, vol. 5, pp. 115–133, 1943.
- [22] V. Latora and M. Marchiori, “Efficient behavior of small-world networks,” *Physical Review Letters*, vol. 87, no. 19, 2001.

- [23] A. Burkitt, H. Meffin, and D. 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.
- [24] 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*, vol. 18, pp. 10464–10472, 1998.
- [25] D. L. Rubin, J. and H. Sompolinsky, "The equilibrium properties of temporally asymmetric hebbian plasticity," *Physical Review Letters*, vol. 86, no. 2, pp. 364–367, 2001.
- [26] D. Horn, N. Levy, I. Meilijson, and E. Ruppin, "Distributed synchrony of spiking neurons in a Hebbian cell assembly," in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K. Müller, Eds., vol. 12. Cambridge, Ma.: MIT Press, 2000, pp. 129–135.
- [27] G. M. Edelman, *Neural Darwinism: The theory of Neuronal Group Selection*. New York: Basic Books, 1987.
- [28] R. F. i Cancho and R. Sol, "Optimization in complex networks," arXiv:arch-ive.cond-mat/0111222, 2001.