

"Network Self-organization explains the distribution of synaptic efficacies in neocortex"

Pengsheng Zheng (Frankfurt Institute for Advanced Studies), Christos Dimitrakakis (EPFL), and Jochen Triesch (Frankfurt Institute for Advanced Studies)

Abstract—The information processing abilities of cortical circuits are thought to arise from their detailed connectivity structure, but this structure is notoriously hard to characterize. The distribution of synaptic strengths of local excitatory connections in the cortex is long-tailed, exhibiting a small number of synaptic connections of very large efficacy. These few very strong connections are comparatively stable and may allow the brain to form long-lasting memories. At present it is unclear, however, how these patterns of synaptic strength and stability arise. We show that these characteristics emerge through self-organization in a simple recurrent network model combining spike-timing-dependent plasticity with homeostatic forms of plasticity. Despite its simplicity the model achieves an excellent fit to data. Our results suggest that cortical circuits are shaped by a form of self-organization induced by the combined action of multiple forms of neuronal plasticity.

Presented at:



DevLeaNN 2011: A Workshop on Development and Learning in Artificial Neural Networks
Held at the ISC-PIF in Paris, France, on October 27, 28, 2011.

Edited by: T. Kowaliw, N. Bredeche, and R. Doursat

<http://devleann.iscpif.fr>

Permission to reprint for non-commercial use granted. Otherwise, all copyright retained by originating author.

Network Self-organization Explains the Distribution of Synaptic Efficacies in Neocortex

Pengsheng Zheng

Frankfurt Institute for Advanced Studies
Frankfurt am Main, Germany 60438
Email: zheng@fias.uni-frankfurt.de

Christos Dimitrakakis

INR 239, EPFL
Lausanne, Switzerland CH-1015
Email: christos.dimitrakakis@epfl.ch

Jochen Triesch

Frankfurt Institute for Advanced Studies
Frankfurt am Main, Germany 60438
Email: triesch@fias.uni-frankfurt.de

Abstract—The distribution of synaptic efficacies in neocortex has an approximately lognormal shape. Many weak synaptic connections coexist with few very strong connections such that only 20% of synapses contribute 50% of total synaptic strength. Furthermore, recent evidence shows that weak connections fluctuate strongly while the few strong connections are relatively stable, suggesting them as a physiological basis for long-lasting memories. It remains unclear, however, through what mechanisms these properties of cortical networks arise.

Here we show that lognormal-like synaptic weight distributions and the characteristic pattern of synapse stability can be parsimoniously explained as a consequence of network self-organization. We simulated a simple self-organizing recurrent neural network model (SORN) composed of binary threshold units. The network receives no external input or noise but self-organizes its connectivity structure solely through different forms of plasticity. Across a wide range of parameters, the network produces lognormal-like synaptic weight distributions and faithfully reproduces experimental data on synapse stability as a function of synaptic efficacy. Overall, our results suggest that the fundamental structural and dynamic properties of cortical networks arise from the self-organizing forces induced by different forms of plasticity.

I.

Recent evidence shows that the distribution of synaptic efficacies is highly skewed [1] and has an approximately lognormal distribution [2]. Only around 20% of synapses are responsible for 50% of total synaptic weight. This structure has been found to be highly dynamic, however, with synaptic contacts constantly being created and destroyed and sizes of dendritic spines fluctuating over time scales of hours and days [3]. This raises the question how stable long-term memories can be maintained in the presence of such strong synaptic fluctuations. Recent evidence has suggested a possible solution to the problem. The dynamics of dendritic spine size changes, which are closely related to synaptic efficacies, are such that small synapses tend to fluctuate relatively more than strong synapses [3].

To investigate whether and how these properties can arise from self-organization induced by neuronal plasticity mechanisms, we extended a simple network model developed previously [4] by adding a structural plasticity. The self-organizing recurrent network (SORN) consists of binary threshold units (80% excitatory and 20% inhibitory). The network is composed of N^E excitatory and $N^I = 0.2 \times N^E$ inhibitory threshold units connected through weighted synaptic connections,

where W_{ij} is the connection strength from unit j to unit i . We distinguish connections from excitatory to excitatory units (W^{EE}), excitatory to inhibitory connections (W^{IE}) and inhibitory to excitatory connections (W^{EI}). For simplicity, connections between inhibitory units and self-connections of excitatory units are forbidden, and the W^{EI} and W^{IE} remain fixed at their random initial values. They have all-to-all topology and are drawn from the interval $[0, 1]$ and subsequently normalized such that the incoming connections to an inhibitory neuron sum up to one: $\sum_j W_{ij}^{IE} = 1$.

The network's activity state, at a discrete time t , is given by the binary vectors $x(t) \in \{0, 1\}^{N^E}$ and $y(t) \in \{0, 1\}^{N^I}$ corresponding to the activity of the excitatory and inhibitory units, respectively. The evolution of the network state is described by:

$$x_i(t+1) = \Theta \left(\sum_{j=1}^{N^E} W_{ij}^{EE}(t)x_j(t) - \sum_{k=1}^{N^I} W_{ik}^{EI}(t)y_k(t) + T_i^E(t) \right) \quad (1)$$

$$y_i(t+1) = \Theta \left(\sum_{j=1}^{N^E} W_{ij}^{IE}x_j(t) - T_i^I \right). \quad (2)$$

The T^E and T^I are threshold values for the excitatory and inhibitory units. They are initially drawn from a uniform distribution in the interval $[0, T_{max}^E]$ and $[0, T_{max}^I]$, respectively. The heaviside step function $\Theta(\cdot)$ constrains the activation of the network at time t to a binary representation: a neuron fires if the total drive it receives is greater than its threshold, otherwise it stays silent.

The network relies on several forms of plasticity: spike-timing dependent plasticity (STDP) of EE connections, synaptic scaling and structural plasticity of EE connections, and intrinsic plasticity regulating the thresholds of excitatory units.

The set of W^{EE} synapses learns via a causal STDP rule that strengthens the synaptic weight W_{ij}^{EE} by a fixed amount $\eta_{STDP} = 0.001$ whenever unit i is active in the time step following activation of unit j . When unit i is active in the time step preceding activation of unit j , W_{ij}^{EE} is weakened by the same amount:

$$\Delta W_{ij}^{EE}(t) = \eta_{STDP} (x_i(t)x_j(t-1) - x_i(t-1)x_j(t)). \quad (3)$$

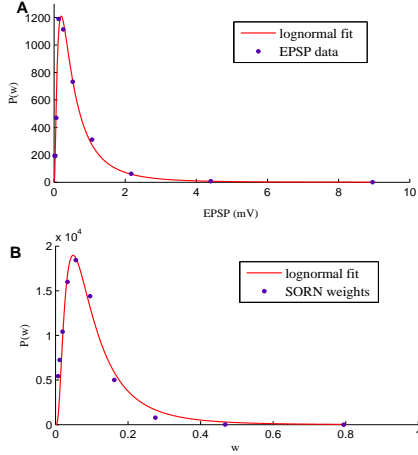


Fig. 1. Distribution of synaptic weights matches lognormal distribution of EPSPs in cortex. A: distribution of EPSP sizes from [2] and lognormal fit. B: distribution of weight strength in SORN and lognormal fit.

Synaptic normalization proportionally adjusts the values of incoming connections to an excitatory unit at each time step so that they sum up to one:

$$W_{ij}^{EE}(t) \leftarrow W_{ij}^{EE}(t) / \sum_j W_{ij}^{EE}(t). \quad (4)$$

An *intrinsic plasticity* rule maintains an identical average firing rate H_{IP} in every unit. To this end, a unit that has just been active increases its threshold while an inactive unit lowers its threshold by a small amount:

$$T_i^E(t+1) = T_i^E(t) + \eta_{IP} (x_i(t) - H_{IP}), \quad (5)$$

where $\eta_{IP} = 0.001$ is a small learning rate. We set the target rate to $H_{IP} = 0.1$ such that an excitatory unit spikes on average every 10th time step.

Compared to the original SORN model, we introduce the structural plasticity which adds new synaptic connections between excitatory cells to the network. With probability $p_c = 0.1$ a new connection is added between a random pair of excitatory cells that are unconnected. The strength of this weight is set to 0.001. Here, the structural plasticity is introduced to simulate the new synapse generation process found in cortex [3].

We simulated SORN networks of different parameters and observed the resulting activity patterns and distributions of synaptic weights. Figure 1 shows the weight distribution of a 400 excitatory neuron network after 10000 time steps and compares it to data from [2]. Figure 2A illustrates the distribution of relative spine volume changes in cortex across one day from [3], and Figure 2B shows the distribution of synaptic weight changes in SORN. The similarities are striking. Numerical simulations show that, across a wide range of parameters, the network produces lognormal-like synaptic weight distributions and faithfully reproduces experimental data on synapse stability as a function of synaptic efficacy.

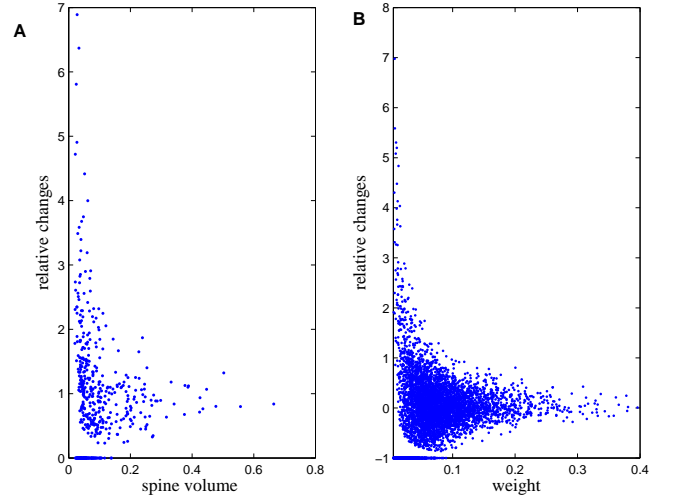


Fig. 2. Distribution of synaptic weight changes matches distribution of spine volume changes in cortex. A: distribution of relative spine volume changes across one day from [3]. B: distribution of synaptic weight changes in SORN.

Overall, our results suggest that the fundamental structural and dynamic properties of cortical networks arise from the self-organizing forces induced by different forms of plasticity. If our explanation of the origin of the statistics and dynamics of synaptic connections in the cortex is correct, then this has broad implications for our understanding of cortical circuits. Current computational models of local cortical circuits usually assume random network structure, sometimes with distance-dependent or layer-dependent connection probabilities[5]. We believe that the study of random networks where only connection statistics are matched to those in the brain, may be quite misleading when the goal is to understand processing in cortical circuits.

REFERENCES

- [1] D. Feldmeyer, J. Lübke, R. Silver, and B. Sakmann, "Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column," *Journal of Physiology*, vol. 538, pp. 803–822, 2002.
- [2] S. Song, P. Sjöström, M. Reihl, S. Nelson, and D. Chklovskii, "Highly non-random features of synaptic connectivity in local cortical circuits," *PLoS Biology*, vol. 3:e68, 2005.
- [3] N. Yasumatsu, M. Matsuzaki, T. Miyazaki, J. Noguchi, and H. Kasai, "Principles of long-term dynamics of dendritic spines," *The Journal of Neuroscience*, vol. 28, no. 50, pp. 13 592–13 608, 2008.
- [4] A. Lazar, G. Pipa, and J. Triesch, "SORN: a self-organizing recurrent neural network," *Frontiers in Computational Neuroscience*, vol. 3, no. 23, 2009, DOI 10.3389/neuro.10.023.2009.
- [5] S. Haeusler and W. Maass, "A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models," *Cerebral Cortex*, vol. 17, no. 1, pp. 149–162, 2007.