

Criticality of avalanche dynamics in adaptive recurrent networks

Anna Levina^{a,*}, Udo Ernst^b, J. Michael Herrmann^c

^aGraduate School 'Identification in Mathematical Models', Göttingen University and MPI for Dynamics and Self-Organization, Bunsenstr. 10, 37073 Göttingen, Germany

^bGroup for Neural Theory, Ecole Normale Supérieure, 3 rue d'Ulm, F-75014 Paris, France

^cBernstein Center for Computational Neuroscience Göttingen and Institute for Nonlinear Dynamics Göttingen University, Bunsenstr. 10, 37073 Göttingen, Germany

Available online 3 November 2006

Abstract

In many studies of self-organized criticality (SOC), branching processes were used to model the dynamics of the activity of the system during avalanches. This mathematical simplification was also adopted when investigating systems with a complicated connection topology including recurrent and subthreshold interactions. However, none of these studies really analyzed whether this convenient approximation was indeed applicable. In present paper we study the correspondences between avalanches generated by branching processes and by a fully connected neural network. The benefit from the analysis is not only the justification of such correspondence but also a simple learning rule, which allows self-organization of the network towards a critical state as recently observed in slice experiments. © 2006 Elsevier B.V. All rights reserved.

Keywords: Neuronal avalanches; Neural adaptation; Self-organized criticality; Critical branching

1. Introduction

Self-organized criticality (SOC) attracted attention of many scientist over the past 20 years, since Bak and Tang [2] discovered the connection between SOC and the appearance of power-law distributions. SOC allows a complex global behavior in systems with rather simple units. In neuronal systems SOC is beneficial because it transfers system to the most sensitive state, when perturbations may lead to the responses on all time and size scales.

Neural networks with long-range connectivity develop scale-independent avalanches of activity if the synaptic efficiencies have suitable values. For homogeneous and temporally constant connectivity the synaptic strengths need to be accurately fixed at a value that depends on network-size in order to obtain a power-law distribution of firing events [5]. The critical synaptic strength depends on network size and deviations lead to a relative surplus of either small or very large avalanches. Later is has been

shown [3] that the theoretically predicted regimes are indeed realized in slices of rat cortex and in neuronal cultures. The exponent of the observed power law also coincided with the theoretical value of $-\frac{3}{2}$. In a recent paper [8] we have studied the effect of short-time synaptic plasticity and have shown that the combined effect of synaptic facilitation and depression drastically increases the parameter range where critical behavior is observed and allow thus to refer to the system as a self-organized critical one.

In many studies [1,3,7,9] of SOC authors used branching processes to model dynamics of the system activity during avalanches. In some cases [2] this approach is justified by the specific system topology (for example in case of sandpile model, where each site can activate only neighbors). The branching process approximation was derived from the statistical evidence that in most cases each site in a recurrent network can be active only once during an avalanche. Consequently, recurrent activations could be neglected. Avalanches were then described by considering only the average number of sites activated by the relaxation of one site. However, the network structure in cortical

*Corresponding author.

E-mail address: anna@chaos.gwdg.de (A. Levina).

slices is much more complicated than in networks where the branching process approximation is accurate. Thus, it is questionable if this approximation is valid. Especially taking into account that in branching models all interactions which do not lead to immediate firing are discarded.

Branching processes have been used to describe and model neuronal dynamics in [4]. The authors observed, that branching processes approximation fits the data they obtained from multielectrode recordings, but the question of adaption was not addressed and mechanisms underlying applicability of such approximation were not discussed. However, it is not fully understood why branching processes should describe well dynamics of such complicated system as a cortical slice.

The present paper gives mathematical proof why branching processes can be used to describe neuronal avalanches. It also accounts for another experimentally observed effect. If in the experiment the dynamics of the network is moved to another regime by application of a blocking inhibition agent picrotoxin [3] the recovery occurs on a time scale of hours. This slow adaptation is not incompatible with the short-time synaptic dynamics in [8] but suggests rather a slow adaptation process. In order to derive a learning rule towards critical behavior we study a formal version of the network dynamics, so-called branching processes where a learning rule is easily derived and theoretically justified.

2. Neuronal network model

The correspondence between branching processes and recurrent networks is conveniently formulated in the following simplified model. The neural processing elements are N perfect, identical neural integrators i with membrane potentials $h_i \in [0, 1]$. The coupling is all-to-all with strength $\alpha = c/N$, and time is measured in discrete steps $t = 0, 1, 2, \dots$. The system is initialized by arbitrary values $h_i \in [0, \theta)$, where $\theta = 1$ is the firing threshold of the neurons. Depending on the state of the system at time t ,

the i th element receives external input $I_i^{\text{ext}}(t)$ or internal input $I_i^{\text{int}}(t)$ from other neural elements. The total input determines the activation \tilde{h}_i of a neuron i by

$$\tilde{h}_i(t+1) = h_i(t) + I_i^{\text{ext}}(t) + I_i^{\text{int}}(t).$$

From the activation $\tilde{h}_i(t+1)$, the membrane potential of the i th neuron at time $t+1$ is computed as

$$h_i(t+1) = \begin{cases} \tilde{h}_i(t+1) & \text{if } \tilde{h}_i(t+1) < 1, \\ \tilde{h}_i(t+1) - 1 & \text{if } \tilde{h}_i(t+1) \geq 1. \end{cases}$$

For the external input $I_i^{\text{ext}}(t)$, one element i is randomly chosen from a uniform distribution at each time step, and a constant amount $\Delta h > 0$ is added to h_i . As it was proven in [5], a particular choice of Δh is not important for resulting distributions. In simulations we used $\Delta h = 0.001$.

The external input is considered to be delivered slowly compared to the internal relaxation dynamics, i.e. it occurs only if no neurons fired in the previous time step. The internal input is given by $I_i^{\text{int}}(t) = M(t-1)\alpha$, where $M(t-1)$ is the number of neurons having fired in the previous time step. If a neuron receiving external input exceeds threshold at time t_0 , then an avalanche starts and we have $M(t_0) = 1$. The avalanche size L is then given by $L = \sum_{k=0}^{\infty} M(t_0+k)$. In [5] a closed expression for the resulting avalanche size distribution was obtained analytically. As it is shown on Fig. 1, this simple neural network the connectivity parameter α allows to obtain a power law distribution (Fig. 1b), as well as sub- and supercritical distributions.

3. Branching process approximation

In order to describe the dynamics of the introduced neural system as a branching process, we consider the random sequence $M(t)$. This allows us to utilize useful results from branching process theory for the exponent of

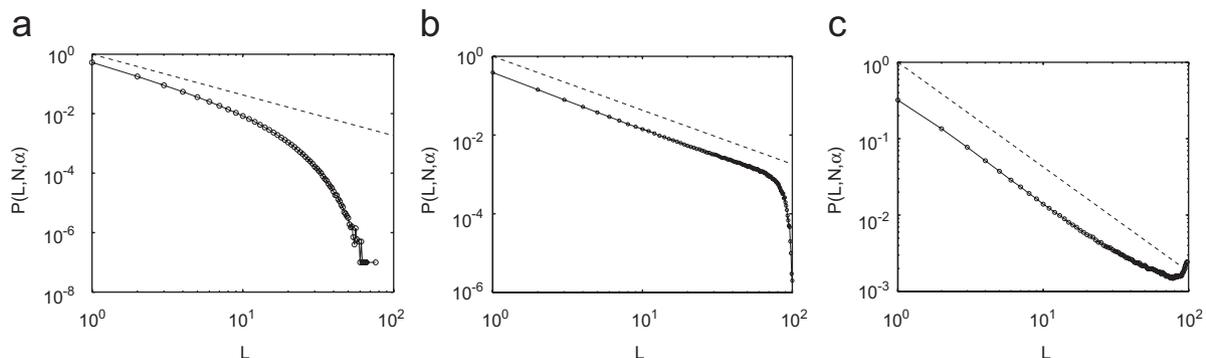


Fig. 1. Probability distributions of avalanche sizes $P(L, N, \alpha)$: (a) in the subcritical, $\alpha = 0.8$; (b) the critical, $\alpha = 0.9$, and (c) supercritical regime, $\alpha = 0.95$. In (a–c) the solid lines and symbols denote the numerical results for the avalanche size distributions, dashed lines show the best matching power-law. Here the curves are temporal averages over 10^6 , $N = 100$.

the power-law distribution of avalanche sizes. The Galton–Watson branching process was first introduced in 1874 to explain the disappearance of British family names [10]. Branching processes describe ‘family trees’ of units, where the offspring is determined independently of earlier generations by a probability distribution p_i , $i \in \mathbb{N}$, of descendants. Usually each unit is assumed to obey the same probability distribution, such that the sequence M_n of members of the generation n provides a sufficient statistics. More precisely, we can define the branching process by a sequence of random variables M_i taking values in \mathbb{N}_0 for all $i > 0$. We always assume that $M_0 = 1$. If $\{M_i\}_{i=0}^\infty$ is a Markov chain, $P(M_{i+1} = 0 | M_i = 0) = 1$ and $(M_{i+1} | M_i = k)$ distributed as a sum of k independent random variables distributed as M_1 . Then the sequence $M_0, M_1, M_2 \dots$ is called a Galton–Watson branching process.

Let us assume now that at the moment preceding an avalanche, the vector of all membrane potentials is uniformly distributed in $[0, 1]^N$, and M_n is the number of neurons firing in the n th time step of the avalanche. Then for large system sizes, the random sequence M_n approximates asymptotically a Galton–Watson branching process with branching parameter c .

In the following we will sketch the proof of this proposition: We first approximate a uniform distribution in $[0, 1]^N$ with a Poisson point process with rate N . So instead of the vector of membrane potentials $(h_i)_{i=1}^N$ we are considering the ordered vector $0 < h_1^* < h_2^* < \dots < h_{k(\omega)}^* < 1$ generated by a Poisson process on the interval $[0, 1]$ with rate N , $k(\omega)$ depends on the realization of a Poisson process (on average we will have same number of coordinates in both vectors).

Poisson approximation secures convergence of the finite-dimensional distributions to the corresponding uniform distribution, meaning that for any set of disjoint

intervals $I_1, \dots, I_m \subset [0, 1]$

$$\lim_{N \rightarrow \infty} P(\#\{i : h_i \in I_1\} = k_1, \dots, \#\{i : h_i \in I_m\} = k_m) - P(\#\{i : h_i^* \in I_1\} = k_1, \dots, \#\{i : h_i^* \in I_m\} = k_m) = 0.$$

For Poissonian membrane potentials $(h_i^*)_{i=0}^k$ we can explicitly check conditions from the definition of the branching process. For proving the Markov property, it has to be shown that

$$\begin{aligned} P(M_n = k_n | M_{n-1} = k_{n-1}, M_{n-2} = k_{n-2}, \dots, M_1 = 1) &= \frac{e^{ck_{n-1}} (ck_{n-1})^{k_n}}{k_n!} \\ &= P(M_n = k_n | M_{n-1} = k_{n-1}) \end{aligned}$$

and for the branching condition, respectively, that

$$\begin{aligned} P(M_{n+1} = k_{n+1} | M_n = k_n) &= \sum_{i_1 + \dots + i_{k_n} = k_{n+1}} \prod_{j=1}^{j=k_n} P(M_2 = i_j | M_1 = 1). \end{aligned}$$

While the former condition is obvious, the latter equality follows from simple combinatorial relation:

$$n^k = \sum_{i_{n-1}=0}^{k-i_{n-2}} \dots \sum_{i_2=0}^{k-i_1} \sum_{i_1=0}^k \frac{k!}{(k - i_{n-1} - \dots - i_1)! i_{n-1}! \dots i_1!}$$

To finally apply the mathematical result in the computational model, we must check if membrane potentials before avalanches are distributed uniformly in $[0, 1]$, which follows directly from the ergodicity of the external input. In Fig. 2 (left), we show results of the ‘mixing’ process which is induced by external input in between avalanches. The corresponding network was initiated with $h_i = 0.5$ for all i , and after some time the numerically computed distribution of the membrane potentials before avalanches turns out to be indeed uniform.

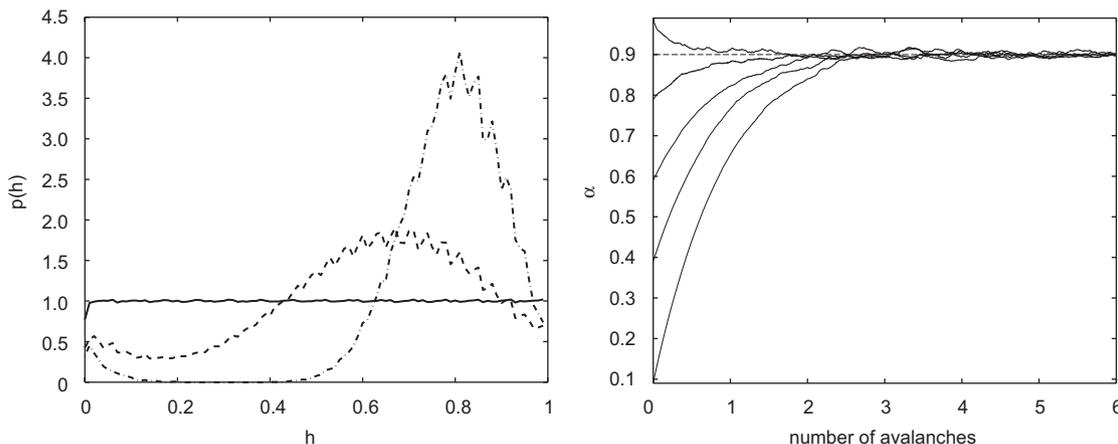


Fig. 2. (Left) Density function of the membrane potentials immediately before an avalanche is triggered. Three snapshots were taken at times when 100 (dashed-dotted line), 1000 (dashed line), and 10^5 (solid line) avalanches occurred, after the initialization of the system with all $h_i = 0.5$. (Right) Each trace shows how learning changes the mean synaptic weight during 6×10^5 avalanches when starting from different initial conditions. The dashed line depicts the critical value of the synaptic weights for 10^4 neurons, $\alpha_c \approx 0.9$.

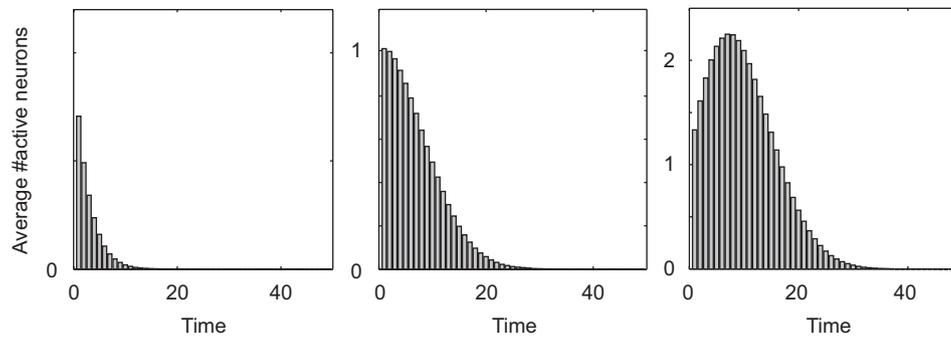


Fig. 3. Averaged avalanches shape. Avalanche shape averaged over 10^5 avalanches for subcritical connectivity parameter $\alpha = 0.7$ (left), critical $\alpha = 0.9$ (middle) and supracritical $\alpha = 0.98$ (right). $N = 100$.

4. Self-organization in the adaptive system

The proposition in the previous section shows that the avalanche dynamics generated by a fully connected neuronal network of large size N is equivalent to a branching process with parameter $c = \alpha N$, where α is a free parameter of the synaptic weights. This allows us to use results from the theory of branching processes stating that avalanche size distributions observe a power-law if and only if $c = \alpha N = 1$, and that the critical exponent has a value of $-\frac{3}{2}$ [6].

Simulating a neuronal network from Section 1 we plotted the histogram of the avalanche shapes for different connectivity parameters Fig. 3. Each bar of the histogram represents the average number of neurons that fired at the certain step of the avalanche. The first bar is a branching ratio calculated in the beginning of avalanche. For the critical case branching ratio is approximately equal to one (some deviation are due to the finite size effects). For the sub- and supracritical cases branching ratio is smaller or large than one, respectively.

We have shown that the avalanche dynamics can be described by a branching process with parameter $c = \alpha N$. A critical state in the network corresponds to a critical branching process, i.e. to the case when each firing neuron on average entails the firing of one neuron. This allows to propose a learning rule which drives the network towards its critical state: Let the neuron starting an avalanche by first crossing the neural threshold θ due to external input trigger l neurons to fire. Then synaptic weights in the network should increase if $l < 1$, and decrease if $l > 1$. As a result of this learning, synaptic weights converge to a value α_c identified as the critical one for finite system sizes in [5]. Fig. 2 (right) shows a typical example of the adaptation dynamics. One can also consider a model with independent synapses, i.e. each neuron has its own parameter of synaptic strength. In this case independent learning for different synapses is applicable.

5. Conclusion

We have presented mathematical ground for utilization of the branching process approximation for recurrent

neuronal networks. This allowed us to construct the learning rule, which drives a network of neural elements into a critical state. We are going to extend the learning rule to be used in more complex neuronal networks, showing that it can induce critical behavior in systems beyond the simplistic model considered here. The presented results explain the experimentally observed fact of criticalization in neural systems and suggest mechanisms for the adaptation processes involved.

References

- [1] P. Alstrom, Mean-field exponents for self-organized critical phenomena, *Phys. Rev. A* 38 (9) (1988) 4905–4906.
- [2] P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: an explanation of $1/f$ noise, *Phys. Rev. Lett.* 59 (1987) 381–384.
- [3] J. Beggs, D. Plenz, Neuronal avalanches in neocortical circuits, *J. Neurosci.* 23 (2003) 11167–11177.
- [4] J. Beggs, D. Plenz, Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures, *J. Neurosci.* 24 (22) (2004) 5216–5229.
- [5] C.W. Eurich, M. Herrmann, U. Ernst, Finite-size effects of avalanche dynamics, *Phys. Rev. E* 66 (2002) 066137.
- [6] T.E. Harris, *The Theory of Branching Processes*, Springer, Berlin, 1963.
- [7] D.-S. Lee, K.-I. Goh, B. Kahng, D. Kim, Sandpile avalanche dynamics on scale-free networks, *Physica A* 338 (2004) 84–91.
- [8] A. Levina, J. M. Herrmann, T. Geisel, Dynamical synapses give rise to a power-law distribution of neuronal avalanches, *Adv. Neural Inf. Proc. Syst.* 18 (2005) 771–778.
- [9] S. Lise, H.J. Jensen, Transitions in nonconserving models of self-organized criticality, *Phys. Rev. Lett.* 76 (1996) 2326–2329.
- [10] H.W. Watson, F. Galton, On the probability of extinction of families, *J. Anthropol. Inst.* 138 (1874) 144.



Anna Levina received her diploma in Mathematics from Saint-Petersburg State University, Russia, in 2003. In 2004, she joined Graduate School ‘Identification in Mathematical Models’ of Göttingen University, where she is working on her PhD thesis in collaboration with Max-Planck Institute for Dynamics and Self-Organization, Göttingen. Her research interest include artificial neural networks, dynamical systems and stochastic processes.



Udo Ernst studied physics at the University of Frankfurt and at the Max-Planck Institut for Fluid Dynamics in Göttingen, where he got his doctorate. Udo Ernst was Research Assistant at the Institute for Theoretical Neurophysics at the University of Bremen, before he moved on to the Group for Neural Theory at Ecole Normale Supérieure in Paris. His research interests include collective phenomena in neural networks, gestalt perception and feature binding in probabilistic

models of primary visual cortex, and data analysis of multisensory recordings in primates.



J. Michael Herrmann received his Ph.D. in Computer Science from Leipzig University, Germany, in 1993. He has been a Postdoctoral Researcher at Leipzig University, at NORDITA, Copenhagen, at RIKEN's Laboratory for Information Representation in Japan, and at the Max Planck Institute for Fluid Dynamics in Göttingen. Presently, he is working at Göttingen University at the Institute for Nonlinear Dynamics. His research interest include artificial neural net-

works, dynamical systems, behavioral modeling, and autonomous robots.