A Mathematical Approach to Self-Organized Criticality in Neural Networks

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch–Naturwissenschaftlichen Fakultäten
der Georg–August–Universität zu Göttingen

vorgelegt von
Anna Levina
aus
Sankt-Petersburg, Russland

Göttingen 2008

D7
Referent: Prof. Dr. Manfred Denker
Koreferent: Prof. Dr. Theo Geisel
Tag der Mündlichen Prüfung:

Contents

Chapter 1

Introduction

Power-law distributions are frequently found in the sciences that study "complex" systems. They are observed, for instance, in the population sizes of cities, in the sizes of earthquakes, the frequencies of words in a text, in neurological data, in the severity of violent conflicts, and in the migration of albatrosses. The multitude of examples includes in particular problems such as the function of the brain, ecosystems, and society, which are of vital interest to all of us.

It is natural to think that the prevalence of spatio-temporal power-law statistics is not accidental. On the contrary, it is most likely that there exists an underlying pattern. The pursuit of such a pattern attracted many scientists; as a result, there exist dozens of model algorithms to generate power-law distributions [109]. However, most algorithms lack a unifying concept, which would not only provide power-law distributions, but also explain their role and highlight the similarity of the systems where they appear. A concept which filled this gap was proposed by Bak, Tang and Wiesenfeld [8] and is called *self-organized criticality* (SOC). The idea of SOC is indeed very simple: it proposes that the system has a critical state as an attractor. It is well known in statistical physics that near a phase transition observable of the system has a power-law distribution. However, in equilibrium systems, the critical point is reached only by tuning a control parameter precisely, which makes it rather rare to be found stable in freely developing systems. Thus, SOC proposes to have a look at non-equilibrium systems instead.

Starting with the first article in 1987, more than 8000 articles have been published on SOC, and more than 1000 over the last two years (data from Google scholar). Regardless of the number of publications, there are still many questions to be answered; some have not even been asked. Here we are giving answers to some of these questions. We shall also be posing further ones.

We are mainly interested in the adequate modeling and understanding of the recent

evidence of self-organized criticality in neuronal data [14]. In the slices of cortical tissue experimentalists observed epochs of neuronal activity, called *avalanches*, separated by periods of rest. The empirical distribution of the avalanche sizes falls off as a power-law and is stable over many hours of recording. The old and well-known models of SOC (sandpile [8], stick-slip [54] or Zhang [154] models) are not well-suited to describe the neuronal realm: they allow only for local interactions and have very simple units, which can substitute neurons only in a rather vague sense. There have only been a attempts to model neuronal avalanches with neuron-like elements. One of the first models [53] appeared before experimental data became available, and was able to predict the critical exponent as well as some extracritical phenomena. However, the critical distribution is only obtained after fine-tuning the connectivity parameters, which makes it unsuitable to explain the neuronal observations. Another study obtained the critical distribution by defining a specific network connectivity [130]. While this is an interesting possibility, there have not been sufficiently many observations so far to accept or reject the hypothesis of the presence of such structures in cortical networks. This thesis focuses on the question of obtaining a neural network model, which will be self-organized critical only by the synaptic dynamics and not relying on specific network connectivity.

Combining the mathematical background and physical environment, we elaborate the problem from different perspectives. We planned to start with studying mathematical models of SOC, but were surprised to find that mathematical treatment of SOC is still in its infancy. Thus, the most basic questions needed to be answered first.

We decided to investigate the simplest model, which still captures important neural characteristics. Namely, we consider a fully connected network of simple integrators, which for specific values of the connectivity parameters is shown to reach the critical state [53]. We start by formulating the mathematical model of the network in the language of dynamical systems, which seems to be the most natural thing, when trying to access the macroscopic behavior (such as the avalanches dynamics) of large systems from the microscopic dynamical evolution. We consider questions of existence and properties of the invariant measure. In the only study of SOC from the dynamical systems perspective [20] existence, uniqueness, and ergodicity of an invariant measure are only conjectured.

Having the invariant measure at hand, we obtain the tool to study avalanches in the neural network more closely. We use this tool to investigate the relation between avalanche dynamics and the branching process. In the physical studies of SOC, it is taken for granted that the branching process captures all important traits of the models, providing an accurate mean-field approximation. By rigorous mathematical reasoning, we study the validity of such approximation, and the speed of convergence for the large networks limit. Among other results, it allows us to analytically obtain the power-law exponent of the critical distribution.

Benefiting from our mathematical considerations, we will propose a new plasticity rule, which allows the initial network to self-organize to a critical state from any initial conditions. Such a plasticity rule is assigned to account for the experimentally observed effect of slow adaptation towards criticality in a cortical network, after it is shifted away from the critical state by chemical alteration.

An important alternative approach to achieve SOC in a neural network is to make use of short-term changes in the connection strength. We study, how biologically more detailed modeling of the interactions, so called *dynamical synapses* [101], influences the avalanche size distribution in a spiking neural network. The concept of dynamical synapses takes into account the finite availability of the neuronal resources. It assumes that with each transmission, a certain fraction of resources is utilized and then slowly restored when the cell is at rest. Our intuition about the network with such synaptic dynamics is as follows:

- If the network was only sparsly active and avalanches are small, synaptic resources are almost fully restored, which leads to a large avalanche.
- If large avalanches occurred recently, synaptic resources are drastically decreased and only small avalanches are possible.

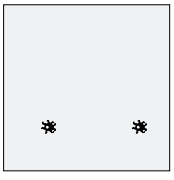
Thus, we hypothesize that the critical state is a stable attractor in such a network. We check this hypothesis in chapter 5. We also consider the analytical treatment of the network by means of statistical mechanics.

Our aim is to obtain criticality by neuronal dynamics regardless of the network structure. However, it is not possible to study the network without selecting a specific connectivity. To show that our results do not depend on the particular choice of connections, we also consider structural modifications of the initial network. We study small-world networks, random networks and networks with nearest-neighbor connectivity.

Regarding our criticism of already existing models for being not suitable to model neuronal processes, biological plausibility is an important issue. Unfortunately, the more detailed a model is, the less one is able describe it analytically. As a compromise, we first study properties of the very simple model analytically, and then study how realistic details change its behavior. We concentrate in particular on the influence of the neuronal leakage and presence of a subpopulation of inhibitory neurons.

Another topic of special interest is the modification of the classical dynamical synapses, so-called *facilitatory synapses* [103]. We study the effect of this modification on the network dynamics.

This thesis is structured into two parts: chapters 2–5 focus more on the physical aspects of our work that are of relevance to computational neuroscience. They can



be read separately from the chapters 6 and 7, in which the mathematically rigorous results are presented.

In **chapter 1**, we present selected topics of the neurophysiological structure of the cortical networks, as well as the main ideas and most prominent models of self-organized criticality. In **chapter 2**, we discuss the simple model of the neuronal network that exhibits criticality. For this classical model, we present some known results [53] and our new findings. **Chapter 3** is devoted to the long-term plasticity rule and the approximation of the network dynamics by the branching processes. In this chapter, we mainly discuss the applications of the mathematical treatment, which is presented in detail in chapter 7. In **chapters 4** and **5**, we elaborate on the short-term synaptic plasticity, depressive and facilitatory synapses. We show how such plasticity is sufficient to obtain SOC neural network. **Chapter 6** contains the mathematical description of the network from chapter 2 and finally, we derive in **chapter 7** the mathematical connection between avalanches and branching processes. The last chapter contains a summary of the main results of this thesis and an outlook.

Chapter 2

Fundamentals

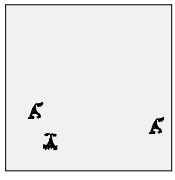
This chapter summarizes basic concepts from neuroscience and the theory of self-organized criticality (SOC) used throughout this thesis. We start with the physiological description of neurons and their connections [79, 119, 11]. The second part of this chapter is devoted to the description of self-organized criticality [74, 7, 51]. We describe the most prominent model of SOC and selected results. We introduce the experimental findings [14] which gave evidence that cortical networks are self-organized critical. These findings motivated a large part of the research presented here.

2.1 Neurophysiological basics

2.1.1 Single neuron

There are two types of cells in the central nervous system: nerve cells or neurons, and glia cells. The latter are more numerous (about 90% of all cells) and serve mainly to form a support system for neurons. Although glia cells have recently been found to contribute also in learning and neuromodulation [55, 146] we will only concentrate on the nerve cells.

Neurons are cells that are specialized in information transmission and computation. They communicate by sending and receiving electrical pulses, so-called *action potentials*. Neurons maintain a potential inside the cell membrane of about -65 mV relative to the extracellular medium. The membrane of a neural cell is a lipid bilayer, selectively permeable for ions. Potential outside the cell is conventionally defined to be 0 mV. At rest, the voltage drop across the membrane is non-zero. The *membrane potential* is the electrical potential difference across a cell's membrane.



The non-zero membrane potential and the ion-concentration gradient is kept by numerous channels and ion pumps build into the membrane. Passive channels allow for the influx of Na^+ and the outflux of K^+ along the concentration gradient. Pumps balance this flows by moving K^+ ions in, and Na^+ ions out, of the cell.

Voltage-dependent ion channels play an important role in the dynamics of the membrane potential. They are responsible for the generation of action potentials. When a neuron is depolarized sufficiently, voltage-gated Na^+ channels open and the resulting influx of Na^+ compel even more channels to open by a positive feedback. Large influx of cations lead to a strong depolarization. At the same time, slower voltage gated K^+ channels open and K^+ ions rush out of the cell. This leads to repolarization, which in turn closes Na^+ channels. The entire cycle lasts only a few milliseconds and the peak amplitude of a spike typically reaches about 100 mV. After every spike, the voltage is reset to approximately the same value, which is typically below the resting potential. This effect is called afterhyperpolarization. A few milliseconds after every spike, it is impossible to evoke another action potential even by a very strong stimulation. This period is called the absolute refractory time.

The variability in the initial depolarization, required for spike generation, is not very large. Thus, models usually assume, that a specific threshold can be set, such that crossing the threshold level indicates a spike. Nevertheless, variability in the threshold and speed of the action potential onset have a large influence on the response properties of the neuron [107].

Depending on their location within the brain and their function in information processing, neurons have a different histology. However, in most of the neurons one can distinguish the following parts: the dendrites, the soma and the axon. *Dendrites* constitute a wide branching tree of links between neurons, which allow a neuron to receive signals from other cells. The *soma*, or the cell body, is the central part of a neuron, where signals arrive from dendrites are collected and processed. It is the metabolic center of the neuron and it contains nuclei. The *axon* is a long projection of a neuron, that conducts electrical signals away from the soma. The axon is connected by *synapses* to dendrites of other neurons.

2.1.2 Synaptic interaction

Synapses allow for the interaction between different neurons. The signal transfer is generally unidirectional: from the axon of one neuron to the dendrites of another neuron. Synapses can be either chemical or electrical. Chemical synapses release a specific chemical messenger – the *neurotransmitter*. Electrical synapses allow direct ion flux between cells and are very fast; therefore, they are particularly important for the neuronal synchronization [40], but they are rare between the principal cells

of the mammalian brain.

We will concentrate on chemical synapses here. Usual chemical synapses can be found at dendritic spines and are formed by membrane swellings on both sides of the contact. Following the direction of the signal transfer, the axon side of the synapse is called the *presynaptic terminal*, whereas the dendritic side is referred to as the *postsynaptic terminal*. Neurotransmitters are stored in the presynaptic side in spherical vesicles.

As an action potential arrives at the presynaptic terminal, depolarization triggers the opening of Ca^{2+} channels. The influx of Ca^{2+} into the presynaptic side leads to the fusion of vesicles and to the release of neurotransmitters into the synaptic cleft. Neurotransmitters cause the opening of ion channels of the postsynaptic neuron by binding to postsynaptic receptors. The resulting *postsynaptic currents* change the ionic concentration. They either raise (depolarize) the potential of the postsynaptic neuron, bringing it closer to threshold (*excitatory postsynaptic potentials*, EPSP), or lower it, bringing the neuron further away from the threshold and therefore delaying the firing (*inhibitory postsynaptic potential*, IPSP) [79].

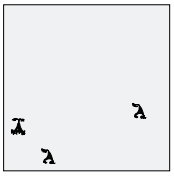
The mechanism of synaptic transmission, as sketched above, has many consequent steps, which lead to delays in the information transmission. Experimentally observed delays vary from 0.3 ms to several milliseconds.

Neuronal networks are sparsly connected. The probability of two neurons on the distance less than the connection range to be connected is between 1% and 25%. At the same time, the number of synapses per neuron ranges from several thousands to several dozens of thousands.

2.1.3 Plasticity

The absolute value of the postsynaptic potential, which is also called *synaptic strength*, varies between different synapses of one neuron. Moreover, the strength of the individual synapse is not constant, but changes depending on the previous activity of that synapse, the activity of the neuron and the activity of the postsynaptic neuron. Synaptic changes are divided into two classes according to the time scale on which they occur: the *long-term plasticity* and the *short-term plasticity*.

The long-term potentiation (LTP, increase of the synaptic strength) and the long-term depression (LTD, decrease of the synaptic strength) are changes which persist up to hours or more. Such effects can be induced by changes in the state of neurotransmitter receptors [21, 127], changes in the distribution of ion channels [144] and morphological changes of neurons [125], however most importantly by neuronal activity.



LTP was first phenomenologically described in 1945 by Donald Hebb [67], who proposed the concept that "cells that fire together, wire together". This means, that nearly simultaneous activity of pre- and postsynaptic neurons increases the synaptic weight (*Hebb rule*):

$$\Delta W_{ij} = k(r_i \cdot r_j),$$

where W_{ij} is the synaptic weight between neuron i and j, and r_j and r_j are instantaneous firing rates of the neurons i and j.

The refined version of Hebbian learning was found in an experimental setup where two cells were forced to fire in a different temporal order [18]. It was discovered that, if a presynaptic cell fires before a postsynaptic cell (in agreement with a possible causal relation), then their synaptic strength is increased, and vice versa.

The short-term synaptic plasticity induces changes of synaptic strength which last for milliseconds. It is often defined by the metabolic properties of the neuron such as depletion of the synaptic resources [101], residual Ca^{2+} in the cell [155], or could be also induced by neuromodulators [32]. In the first case, depression or facilitation of the synaptic connection depends on the frequency of stimulation of the synapse. Therefore, synapses which are the subjects of such a dynamics are referred to as *frequency-dependant* or *dynamic synapses*. An overview of the synaptic plasticity can be found in many text books; for example, it is given in [61].

2.1.4 Neuronal models

The mathematical treatment of the neuronal dynamics presupposes the construction of appropriate neuron models. The spectrum of existing models varies from highly detailed models [100] to the very simple and abstract models with two discrete states [70]. Detailed models are good to obtain a realistic picture of the neuronal dynamics in a simulation, whereas simple models highlight general principles and allows one to perform analytical computations and predictions. Nevertheless, oversimplification may lead to wrong results.

If the morphology of the neuron is neglected, then the model describes a single voltage variable V. The dynamics of the voltage can be expressed in terms of currents and ion fluxes. The *integrate-and-fire* (IF) model provides the simplest phenomenological description of the voltage dynamics. It was proposed at the beginning of the 20^{th} century [84] and has been well studied and often used in the neural models since then.

The IF model describes the neuron as an electrical circuit, which consist of a capacitor with the capacitance C_m in parallel with a resistor with leak conductivity g_l representing the passive open ion channels, driven by a current $I(t)$. Thus, the

voltage dynamics is described by

$$C_m \frac{dV(t)}{dt} = g_l \left(V_{\text{rest}} - V(t)\right) + I(t), \tag{2.1}$$

where V_{rest} represents the resting potential of approximately -65mV. When the potential difference at the capacitor exceeds a certain threshold θ, the generation of an action potential is started and the capacitor is reset to the resting potential V_r.

Despite its simplicity, the model reproduces many aspects of the response properties of a real neuron if the input current does not vary too fast [44, 120]. Further development leads to improved models that include quadratic integrate-and-fire models [25, 85], refractory periods, dynamic thresholds or time constants [23, 75, 95]. In this thesis, the most simple integrate-and-fire model without leak (i.e. $g_l = 0$) will be used.

2.2 Self-organized criticality

There has been a long history of attempts to propose physical theories to explain a seemingly heterogeneous variety of phenomena by a few general principles. Self-organized criticality (SOC) is certainly a prominent example of a theory with such a unifying purpose. It was proposed to explain the emergence of complexity.

The term "self-organized criticality" emphasizes two aspects of the system behavior. *Self-organization* is used to describe the ability of certain nonequilibrium systems in the absence of control or manipulation by an external agent to develop specific structures and patterns. The word *criticality* is used in order to emphasize the similarity with phase transitions: a system stays at the border of stability and chaos. The concept was coined by Bak et al. [8] who proposed the first SOC model and discovered the connection between SOC and the appearance of the power-law distributions.

Self-organized criticality has been used to model phenomena as diverse as the piling of granular media [57], plate tectonics [64], forest fires [97], stick–slip motion [54], and electric power system blackouts [30]. It has also recently become appealing to biologists [3, 14, 28] .

Critical behavior has been shown to bring about optimal computational capabilities [88], optimal information transmission [14], storage of information [65], and sensitivity to sensory stimuli [37, 46, 81].

In many of the above-mentioned instances, the applicability self-organized criticality is largely disputed. Among them is, for example, one of the earliest exam-

ples of a natural phenomenon described by SOC: the magnitude of earthquakes. Most authors accept the SOC model. However, if the SOC claims are true, then earthquakes are unpredictable, which makes some authors skeptical about the concept [41, 150, 151].

There are also some scientists who criticize the generality claim of the self-organized criticality theory [58, 137]. Their scepsis is based on the observation, that nowadays, all models and natural phenomena that have a power-law statistics of any observable are immediately interpreted as SOC. Although, sometimes there exists much more simple and natural explanation of power-law statistics [12].

Nevertheless, it is without doubt that SOC is an important tool to understand more about appearance of power-laws and the phenomena of self-organization. Even if it should not be considered as a complete description of every phenomenon, it is still an important first step in comprehending complexity.

One example of a system which may be a good candidate for SOC was found in the cortical slices [14, 15] and later in other neural systems [104]. This new discovery revived the interest in self-organized criticality and also motivated this work.

In the following parts, we will first briefly introduce the most important SOC models and some of their implications. Then we will remark on mathematical contributions to the studies of self-organized criticality. Finally we will describe the experiments of Beggs and Plenz.

2.2.1 Sandpile model

In this section, we describe the sandpile model which was proposed by Bak in collaboration with Tang and Wiesenfeld [8, 9]. It attracted great interest as the first and clearest example of self-organized criticality.

The model is inspired by a real pile of sand. Grains of the model "sand" are dropped into the system and are lost at the boundaries, allowing the system to reach a stationary state that balances input and output. In the limit of infinitely slow input, the system displays a highly fluctuating, scale-invariant avalanche-like pattern of activity.

Formally, the sandpile model is defined on the d-dimensional lattice of size L^d. The variable $z(\mathrm{r})$ describes the "energy" of the site $\mathrm{r} \in \mathbb{R}^d$, and e_j, $j = 1, \ldots, d$ are basis vectors on the lattice. The dynamics of z obeys the following rules:
1) If for all sites of the latice r , $z(\mathrm{r_k}) < z_c$, then one site r_i is selected randomly and one grain of sand is dropped there

$$z(\mathrm{r}_i) \to z(\mathrm{r}_i) + 1. \tag{2.2}$$

2) If $z(\mathrm{r}_i) \geq z_c$, then site r_i relaxes by the following rule:

$$\begin{aligned} z(\mathrm{r}_i) &\rightarrow z(\mathrm{r}_i) - 2d \\ z(\mathrm{r}_i \pm e_j) &\rightarrow z(\mathrm{r}_i \pm e_j) + 1 \quad \text{for } j = 1, \ldots, d. \end{aligned}$$

On the boundary, energy dissipates from the system and there is no other form of dissipation in the system. At each moment of time, only one of the rules is used. This creates an infinite separation of timescales of the internal relaxation (rule 2) and the external input (rule 1).

An avalanche is then a sequence of relaxations that directly follows the addition of one grain. The system reaches a stationary state that is characterized by a power-law distribution of avalanches sizes [8, 77, 99]

$$P(s) \sim s^{-\gamma}.$$

Dhar [48] has shown that the sandpile model described above is Abelian, i.e. the final state of the system is independent of the order in which rules are applied to different sites. One consequence of this is that the order of the redistributions during an avalanche does not affect the final structure of the avalanche [4, 10, 49, 50, 118].

Many techniques have been applied to the analysis of sandpile models. Examples include mean-field studies [59, 139, 140, 141], renormalization group methods [115, 142], branching processes [118], invasion percolation [123] and damage analysis [17].

One of the natural questions is whether a real sandpile exhibits SOC. There were a number of experimental studies with different types of sand and different mechanism of sand addition [73, 106, 122] but they all did not show any SOC. Instead, self-organized criticality was found in ricepiles for a specific type of rice [57].

2.2.2 Mathematical approach to SOC

Despite the great interest in self-organized criticality among physicists, only a very limited number of mathematicians have studied the phenomenon of SOC by rigorous methods. One of the rare examples is a group spread between Bielefeld and Nice of B. Cessac, Ph. Blanchard, and T. Krueger. They studied the Zhang model [154] of cellular automata on the d-dimensional lattice.

In a series of articles [20, 33, 34, 35, 36, 145] they proposed a mathematical model of the self-organized critical system using methods and tools from dynamical systems theory, ergodic theory and statistical physics. They have shown that the Zhang model of SOC is a hyperbolic dynamical system with singularities. They also have analyzed the transport dynamics and related it to the Lyapunov spectrum. This

establishes an unexpected relation between the structure of the (fractal) attractor [33] on which the dynamics resides and the transport properties. They have also shown that one can construct Gibbs measures (in the sense of Sinai-Ruelle-Bowen [22, 124, 126]) which are directly related to avalanches distributions.

After the Sinai-Ruelle-Bowen measure is found, it is possible to use techniques from statistical physics of critical phenomena (Lee Yang zeros [87, 149]) to analyze the behavior of avalanche distributions when the system size tends to infinity. The authors have shown that this method allows detection of a bias in the numerics which leads to spurious critical exponents [36]. Finally, using methods from quantum field theory they have studied a stochastic partial differential equation modeling transport in SOC models [34]. They have shown that a perturbative method requires one to handle all terms in the series and they have been able to extract two free parameters in the theory, that can be related to the transport and the scale behavior of Lyapunov exponents.

Throughout their work, Cessac *et al.* [34, 36] assumed the existence of the unique Sinai-Ruelle-Bowen measure on the extended state space; based on this conjecture, they build their thermodynamical formalism. In our study, some results concerning such a measure is provided (for a different system). In this sense, our work is complementary to the approach of Cessac *et al.*

2.2.3 Neuronal avalanches

In this section one example of a self-organized critical system, namely neuronal avalanches, is described. The discovery of the critical avalanches in the brain motivated this work on SOC and renewed general interest in the field.

First experiments were done in cultured and acute cortical slices by John Beggs and Dietmar Plenz [14, 15, 116, 129]. Cultures were planted on a multielectrode array and local field potential signals were recorded from the 64 electrodes of the array during a long period of time (on a time scale of hours).

A local field potential (LFP) is a signal which reflects the sum of all synaptic activity within a volume of tissue [26]. The precise origin of the signal is not yet clear, and it was found that phenomena that are unrelated to synaptic events also contribute to the LFP [27]. Nevertheless, LFP characterizes the cumulative activity in the slice better than spikes. It also often gives a larger, than spikes recorded from a small number of electrodes, correlation with the behavior in *in vivo* experiments.

In Fig. 2.1, the extraction of the filtered LFP signal from the cortical slice is shown. The first filtering stage extracts the LFPs from the recorded signal, which are then thresholded to obtain the binary signal on each electrode. The data is then organized in 4 ms bins. After such processing, the data consists of short intervals of

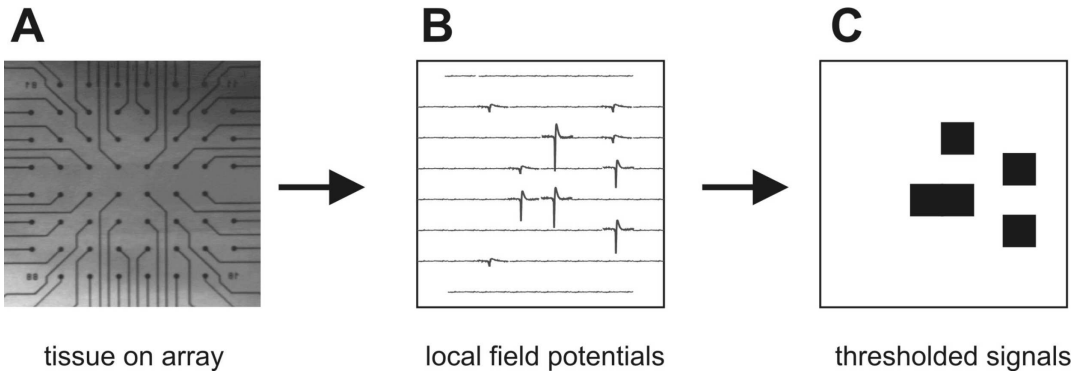


Figure 2.1: Extraction of LFPs from a multielectrode array.

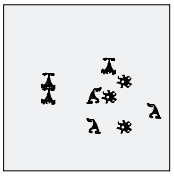
activity, when one or more electrodes detected LFPs above the threshold, separated by longer periods of silence. When the activity periods are studied in fine temporal resolution, it is possible to see the spread of activity in the slice. A set of consecutively active frames is called an avalanche. The size of an avalanche is defined as the number of electrodes which were active during the avalanche. Avalanche sizes turn out to follow the power-law distribution with the exponential cutoff at the size of the multielectrode array. The distribution is stable over many hours. The exponent of $-3/2$ characterizes both acute cortical slices and cultures.

The application of the inhibition antagonist *picrotoxin* makes the network more exitable and changes the distribution of the avalanche sizes to bimodal. After a few days, the network returns to the critical state despite pharmacological influence (Fig. 2.2). This implies that there are some long-term regulatory mechanisms which tune the network to criticality.

Propagation pathways similar to the ones described in Ref. [143] form significant similarity clusters [15]. It has also been observed that activity at one electrode can be followed by activity at any other electrode, which allows to rule out the hypothesis of a synfire chain [52] nature of the avalanches.

In the cortex, the emergence of power-law distribution of avalanche sizes with slope $-3/2$ depends on an optimal concentration of dopamine [129] and on the balance of excitation and inhibition [14, 129], which suggests that particular parameters must be appropriately "tuned" [13]. This provides additional evidences in favor of the hypothesis that there are mechanisms in the cortex which lead to robust and local self-organization towards the critical state.

Despite the complex relationship between LFP waveform and underlying neuronal processes, neuronal avalanches have also been recently identified in vivo in the normalized LFPs extracted from ongoing activity in awake macaque monkeys [114]. This proves that critical avalanches are not an artifact of the neuronal cultures and slices and also present in the brain of the behaving animals.



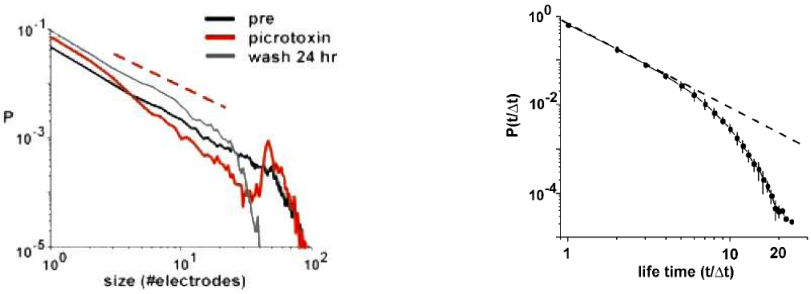


Figure 2.2: Avalanche sizes and durations distribution. (left) Black: avalanche size distribution in unperturbed system, red: avalanche size distribution after application of picrotoxin, gray: relaxation of distribution after washing out of the reagent, dashed line: power-law with exponent -3/2. (right) Distribution of avalanche durations, dashed line: power-law with exponent -2. *Pictures are taken from [14]*

Until recently, avalanches were observed only in LFP recordings, however two publications appeared last year, in which critical avalanches were measured in spikes. In the dissociated cultures of rat hippocampal neurons and intact leech ganglia, avalanche sizes have a power-law distribution with slope $-3/2$ [104]. In cortical culture with an array of 500 electrodes, a power law with an exponent of approximately -2.1 was observed [13].

2.2.4 Models of neuronal avalanche

There were some attempts to model self-organized criticality in neural networks and other neural systems. In this section we briefly describe the most prominent of them.

Among the earlier models, the study of Herz and Hopfield [68, 71] deserves special mention. They tried to connect the reverberations in a neural network to the power law distribution of earthquake sizes. They considered leaky integrate-and-fire neurons with a constant input current on a lattice and were mainly interested in synchronization of such a network in the non-leaky case.

The first example of a globally coupled system that shows criticality was presented by Eurich, Herrmann and Ernst [53]. In this study, a simple model of a fully connected neural network of non-leaky integrate-and-fire neurons was investigated. The results predicted the critical exponent as well as different dynamical phenomena, which were later observed experimentally [14, 15]. The model and its implications will be discussed in chapter 3.

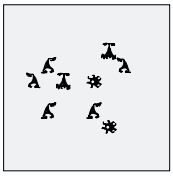
After the experiments by Beggs and Plenz, further models have been proposed. Among them is, for example, the model by Teramae and Fukai [130]. The authors studied a network of Izhikevich neurons [72] organized into synfire chains [52]. The specific wiring leads to the stable avalanche propagation and a power-law statistics of the avalanche sizes.

Another approach was used by Rohrkemper and Abbot [1] who concentrate on the development of the connectivity structure and do not consider the synaptic plasticity. They used the activity-dependent neurite outgrowth model [111] to obtain a network structure from the initially disconnected neurons on a two-dimensional layer. The authors claim that the resulting network produces a stable power-law behavior, although the plots presented in the article do not show a clear power-law distribution.

SOC was also obtained from a different perspective in the works of de Arcangelis *et.al* [45, 113]. The authors considered the network on a grid with an anti-hebbian learning rule. Their model is intended to explain the power-law distribution of the EEG spectrum.

Some authors claim that the power-law statistics or $1/f$ scaling in the neural data may be explained independently of critical states. An alternative explanation of the $1/f$ frequency scaling arises from the filtering properties of extracellular media [12].

In summary, we demonstrated examples, illustrating that ideas of SOC have found large interest in the scientific community. However, neither the mathematical understanding, nor the explanation of SOC in neurobiological experiments, is sufficiently studied.



Chapter 3

Static synapses

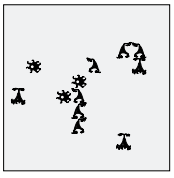
In this chapter we discuss the model of the neural network proposed in 2002 by Eurich et al. [53]. That study presented the first example of a globally coupled system that exhibits criticality. It also predicted the critical exponent as well as some extra-critical dynamical phenomena, which were later observed in the experimental research [14].

The simplicity of the model allows to derive a formula for the distribution of avalanche sizes analytically. However, its dynamics still poses a large repertoire which is accessed by changing the only parameter of the model. One of the unique features of this study is that the distribution of avalanche sizes is also found for a finite-size system, in contrast to most other researches, where analytical results are obtained only in the thermodynamical limit. In chapter 7 the mathematical treatment of this model is presented.

3.1 Description of the model

Consider a set of N identical threshold elements characterized by a state variable $h \geq 0$ which corresponds to the neuronal membrane potential. The system is initialized by arbitrary values $h_i \in [0, \theta)$ $(i = 1, \ldots, N)$, where θ is the threshold above which elements become unstable and relax.

In the model, time is measured in discrete steps, $t = 0, 1, 2, \ldots$. 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 elements. If $h_i(t+1)$ is larger than θ, periodic boundary conditions are applied. Hence, the membrane potential of the i^{th} element



at time $t+1$ is computed by

$$h_i(t+1) = \begin{cases} h_i(t) + I_i^{\text{ext}}(t) + I_i{}^{\text{int}}(t) & \text{if } h_i(t) + I_i^{\text{ext}}(t) + I_i{}^{\text{int}}(t) < 1, \\ h_i(t) + I_i^{\text{ext}}(t) + I_i{}^{\text{int}}(t) - 1 & \text{otherwise.} \end{cases}$$

where without loss of generality θ is set to be equal one. For the external input $I_i^{\text{ext}}(t)$, one element is randomly chosen from a uniform distribution on $1, \ldots, N$, and a constant amount of energy $\Delta h \in (0, 1]$ is added to the element's membrane potential. The external input is considered to be delivered slowly compared to the internal relaxation dynamics, i. e. it occurs only if no element has exceeded the threshold in the previous time step. This corresponds to an infinite separation of the time scales of external driving and avalanche dynamics that have been suggested as a necessary condition for SOC [128, 141, 140].

The external input can formally be written as

$$I_i^{\text{ext}}(t) = \delta_{ri}(t)\, \delta_{\xi(t-1),0}\, \Delta h, \tag{3.1}$$

where $r \in \{1, 2, \ldots, N\}$ indicate the element chosen to get external input, $\xi(t-1)$ is the number of suprathreshold elements in the previous time step, and $\delta_{\cdot\cdot}$ is the Kronecker delta.

The internal input $I_i^{\text{int}}(t)$ is given by $I_i{}^{\text{int}}(t) = \xi(t-1)\alpha/N$, where α/N is the coupling strength between the elements. We assume connections to be excitatory, that is, $\alpha > 0$.

At some time t_0 the element receiving external input exceeds threshold and an avalanche starts, $\xi(t_0) = 1$. The system is globally coupled, such that during an avalanche all elements receive internal input, including the unstable elements themselves. An example of the avalanche is shown in Fig 3.1.

The avalanche duration $D \geq 0$ is defined to be the smallest integer for which the stopping condition $\xi(t_0 + D) = 0$ is satisfied. The avalanche size, L, is given by $L = \sum_{k=0}^{D-1} \xi(t_0 + k)$. The model [53] allows to calculate the probability $P(L, N, \alpha)$ of an avalanche of size $L \geq 0$ in the regime $0 \leq L \leq N$ in a system consisting of N elements with coupling parameter α. Avalanche size distributions can alternatively be described by a conditional probability $P_0(L, N, \alpha) = P(L, N, \alpha | L > 0)$, which is related to $P(L, N, \alpha)$ via

$$P_0(L, N, \alpha) = \frac{P(L, N, \alpha)}{1 - P(0, N, \alpha)}\,. \tag{3.2}$$

Due to the global coupling of the elements, there are no spatial boundary conditions to be specified in the model.

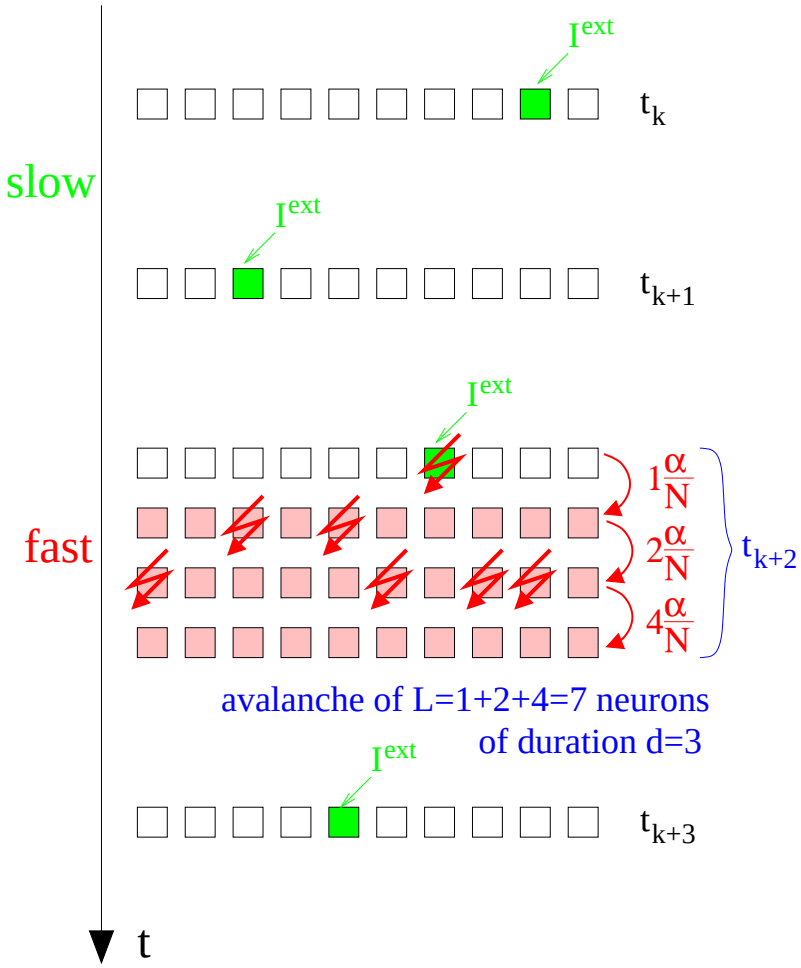


Figure 3.1: Example of an avalanche. A network of 10 neurons is depicted, each square denote one neuron, each raw is a state of the network at the fixed moment of time. Green color marks the neurons which receive the external input. As long as there are no super-threshold neurons in the network the time scale of the external input is taken for the time discretization. Red sparks mark supra-threshold neurons. Light red color denotes internal input. The firing happens on the fast time-scale and during the avalanche discretization step is taken to be equal to the synaptic delay.

3.2 Simulations and analysis

Depending on the connection strength α, the network can produce a rich repertoire of behaviors. If $\alpha \geq 1$, then the network accumulates energy with each avalanche and once there is an avalanche of a size of at least N the activity will never terminate. If $1 > \alpha > 1 - \Delta h$ then it is possible for a neuron to fire more than once during the avalanche and multi-peaked distribution is observed.

In Fig. 3.2 examples of avalanche size distributions are shown for various values of $\alpha < 1$. For small values of α, subcritical avalanche-size distributions are observed. This regime is characterized by a negligible number of avalanches that extend to system size. If $\alpha = \alpha_{\text{cr}}$ the system has an approximate power-law avalanche distribution with the slope close to $-3/2$ for avalanche sizes almost up to the size of the system where an exponential cut-off takes place. Avalanche size distributions become non-monotonous when α is above the critical value α_{cr}.

If the "distance" $d(P(\alpha, N), P_{\text{pl}})$ from the observed distribution $P(\alpha, N)$ to a "ideal" power-law distribution P_{pl} with the slope $-3/2$ is defined, numerical analysis allows to find the parameter α_{cr} which minimizes this distance. There is some arbitrariness in the selection of the metrics, authors of [53] selected rather conventionally the symmetric version of the Kullback-Leibler divergence, which is defined as

$$d(\alpha) = \sum_L (P(L, \alpha, N) - P_{\text{pl}}(L))(\log(P(L, \alpha, N)) - \log(P_{\text{pl}}(L))). \tag{3.3}$$

In general, Kullback-Leibler divergence is not a metric, because it does not satisfy triangular inequality, but it is non-negative, symmetric and equals to zero if and only if distributions coincide. The argument α at the minimum of the divergence is found to depend on N as,

$$\alpha_c(N) \approx 1 - N^{-\mu} \text{ with } \mu = 0.5 \pm 0.01\,. \tag{3.4}$$

One of the important advantages of this model is the possibility to perform the analytical calculation of the avalanche size distribution. Namely,

$$P_0(L, \alpha, N) = L^{L-2} \binom{N-1}{L-1} \left(\frac{\alpha}{N}\right)^{L-1} \left(1 - L\frac{\alpha}{N}\right)^{N-L-1} \frac{N(1-\alpha)}{N - (N-1)\alpha}. \tag{3.5}$$

The conditions required for such calculations are discussed and checked in chapter 7. The exact formula allows to find critical connectivity strength in the thermodynamical limit, which appears to be $\alpha_c = 1$. Thus it is also possible to find a limiting power-law exponent. We take first the limit of the large system size $N \to \infty$ and then limit $\alpha \to 1$, because in the opposite case the limit does not exist.

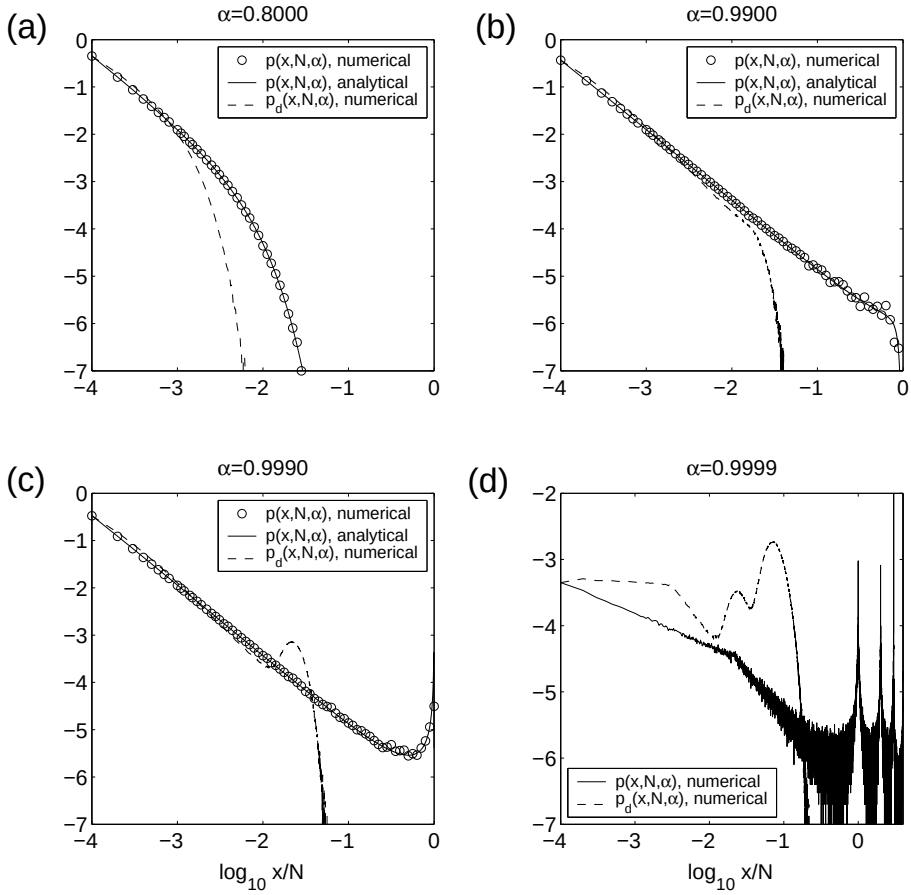
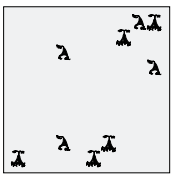


Figure 3.2: Probability distributions of avalanche sizes, $P_0(x, N, \alpha)$, and avalanche durations, $p_d(x, N, \alpha)$, in the subcritical (a; $\alpha = 0.8$), critical (b; $\alpha = 0.99$), supra-critical (c; $\alpha = 0.999$), and multi-peaked (d; $\alpha = 0.99997$) regime. (a-c) Solid lines and symbols denote the analytical and the numerical results for the avalanche size distributions, respectively. In (d), the solid line shows the numerically calculated avalanche size distribution. The dashed lines in (a-d) show the numerically evaluated avalanche duration distributions. In all cases, the presented curves are temporal averages over 10^7 avalanches with $N = 10000$, and $\Delta h = 0.022$. *Figure and caption taken from [53].*



We obtain

$$\gamma(L) = \lim_{\alpha\to 1}\lim_{N\to\infty} \log\frac{p(L,N,\alpha)}{p(L+1,N,\alpha)} \Big/ \log\frac{L}{L+1}\,. \tag{3.6}$$

$\gamma(L)$ becomes constant for the large L:

$$\lim_{L\to\infty}\gamma(L) = -3/2.$$

In the case $\alpha = 1$, system is conservative and then a power-law behavior with an exponent of $-3/2$ is reached in the regime of large avalanche sizes. Therefore, in the thermodynamical limit the critical state is obtained in the conservative system. This appears natural in the context of the discussion about SOC and conservation laws. In a sequence of articles numerous authors found arguments against existing non-conservative SOC models [42, 63, 82, 24] and proposed variants that were claimed to be critical in non-conservative cases [110, 93]. For Abelian sandpile (i.e. system where dynamics does not depend on the order of nodes update) it is proven, that conservation is necessary for SOC [136], but so far there is no final point in this discussion, though all observed SOC systems are conservative.

This model would suit perfectly to model neuronal avalanches [14], if it would be really self-organized in a sense compatible with biological neurons. But actually, to obtain the critical regime one has to tune the connectivity parameter very precisely. Another drawback is, that the parameter tuning must become more precise with increasing system size to sustain the same quality of fit to the power-law distribution. It is hard to expect such a precision of connectivity tuning from neuronal systems. Thus there is a necessity to extend the model, in order to achieve self-organized criticality in a way that is compatible with biological observations. In the next chapter we are going to introduce one of such models.

3.3 Abelian distribution

In this and the following section several topics which were not considered in the paper of Eurich et al. [53] are discussed. We prove that the distribution, introduced in the paper is indeed probability distribution. We also derive its mean value, which will be used later in chapters 5 and 6.

Definition 3.3.1. Let $N \in \mathbb{N}, \alpha < 1$. Then Abelian distribution is defined for each $0 \le L \le N$ by

$$P_{\alpha,N}(L) = C_{\alpha,N} L^{L-2} \binom{N-1}{L-1} \left(\frac{\alpha}{N}\right)^{L-1} \left(1 - L\frac{\alpha}{N}\right)^{N-L-1}, \tag{3.7}$$

where $C_{\alpha,N}$ is a normalizing constant.

We call this distribution Abelian, because of the Abelian sums which are present in its description. It was found in [53], that avalanche size in the fully connected neural network of size N with connectivity parameter α has an Abelian distribution.

Lemma 3.3.1.

$$C_{\alpha,N} = \frac{N(1-\alpha)}{N-(N-1)\alpha}. \tag{3.8}$$

Proof.

$$\sum_{L=1}^{N} C_{\alpha,N} L^{L-2} \binom{N-1}{L-1} \left(\frac{\alpha}{N}\right)^{L-1} \left(1-L\frac{\alpha}{N}\right)^{N-L-1} = 1. \tag{3.9}$$

We introduce a new variable $x = \alpha/N$ and get

$$\sum_{L=1}^{N} L^{-1} \binom{N-1}{L-1} (Lx)^{L-1} (1-Lx)^{N-L-1} = \frac{1}{C_{\alpha,N}}, \tag{3.10}$$

which is equivalent to

$$\sum_{L=1}^{N-1} L^{-1} \binom{N-1}{L-1} (Lx)^{L-1} (1-Lx)^{N-L-1} = \frac{1}{C_{\alpha,N}} - \frac{(Nx)^{N-1}}{N(1-Nx)}. \tag{3.11}$$

We can compute the sum on the left side of (3.11)

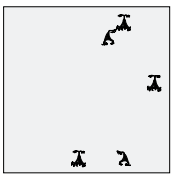
$$\begin{aligned} &\textstyle\sum_{L=1}^{N-1} L^{-1} \binom{N-1}{L-1} (Lx)^{L-1} (1-Lx)^{N-L-1} = \\ &\textstyle\sum_{L=1}^{N-1} L^{-1} \binom{N-1}{L-1} (Lx)^{L-1} \sum_{m=0}^{N-L-1} (-1)^m \binom{N-L-1}{m} (Lx)^m \end{aligned} \tag{3.12}$$

Introducing $k = L-1$ we can rewrite the sum in the left side of (3.13) as a polynomial in x

$$\sum_{i=0}^{N-2} x^i \sum_{k=0}^{i} (-1)^{i-k} \binom{N-1}{k} \binom{N-k-2}{i-k} (k+1)^{i-1} = \sum_{i=0}^{N-2} P_i(N) x^i, \tag{3.13}$$

where $P_i(N)$ is a polynomial of the degree i. If $i = 0$ then $P_0(N) = 1$. Consider $i > 0$ to identify uniquely the polynomial $P_i(N)$ it is sufficient to find its values in $i+1$ different points $N = 1, \ldots, i+1$. If $k > N-1$ then $\binom{N-1}{k} = 0$. For $N < i+2$ we realize that if $k < N-1$ then $\binom{N-k-2}{i-k} = 0$, hence

$$P_i(N) = (-1)^{i-k} \binom{N-1}{N-1} \binom{-1}{i-k} N^{i-1} = N^{i-1} \text{ for N} = 1, \ldots, \text{i} + 1 \text{ and i} > 0$$



which means that $P_i(N) = N^{i-1}$ for any N and $i > 0$. Then the left side of (3.13) is

$$1 + \sum_{i=1}^{N-2} x^i N^{i-1} = 1 + x\frac{1-(Nx)^{N-2}}{1-Nx}. \tag{3.14}$$

Inserting (3.14) into (3.11) we get a final equation to determine $C_{\alpha,N}$

$$1 + x\frac{1-(Nx)^{N-2}}{1-Nx} = \frac{1}{C_{\alpha,N}} - \frac{(Nx)^{N-1}}{N(1-Nx)}.$$

Hence,

$$\frac{1}{C_{\alpha,N}} = 1 + x\frac{1-(Nx)^{N-2}}{1-Nx} + \frac{(Nx)^{N-1}}{N(1-Nx)} = \frac{(1-Nx)N + Nx}{N(1-Nx)} = \frac{N-(N-1)\alpha}{N(1-\alpha)}.$$

□

Theorem 3.3.2. *If ξ has a Abelian distribution with parameters α, N, then*

$$E\xi = \frac{N}{N-(N-1)\alpha}. \tag{3.15}$$

Proof. From (3.7) and Lemma 3.3.1 we have

$$E\xi = \sum_{L=1}^{N} L^{L-1}\binom{N-1}{L-1}\left(\frac{\alpha}{N}\right)^{L-1}\left(1-L\frac{\alpha}{N}\right)^{N-L-1}\frac{N(1-\alpha)}{N-(N-1)\alpha}.$$

We have to prove that

$$\sum_{L=1}^{N} L^{L-1}\binom{N-1}{L-1}\left(\frac{\alpha}{N}\right)^{L-1}\left(1-L\frac{\alpha}{N}\right)^{N-L-1} = \frac{1}{1-\alpha}. \tag{3.16}$$

Introducing the new variable $x = \alpha/N$ we can rewrite this equation as

$$\sum_{L=1}^{N}\binom{N-1}{L-1}(Lx)^{L-1}(1-Lx)^{N-L-1} = \frac{1}{1-Nx}. \tag{3.17}$$

Transforming sum in (3.17) we obtain

$$\sum_{L=1}^{N-1}\binom{N-1}{L-1}(Lx)^{L-1}(1-Lx)^{N-L-1} + (Nx)^{N-1}(1-Nx)^{-1} = \frac{1}{1-Nx}. \tag{3.18}$$

which is equivalent to

$$\sum_{L=1}^{N-1} \binom{N-1}{L-1} (Lx)^{L-1} (1-Lx)^{N-L-1} = \sum_{i=0}^{N-2} (Nx)^i. \tag{3.19}$$

Both left and right sides of the equation (3.19) are polynomials in x of the degree $N-2$. Hence to prove that equation (3.19) is identity it is sufficient to prove that coefficients of x^i are equal on the both sides for every i. In other words we have to show that

$$\sum_{k=1}^{i} (-1)^{i-k} \binom{N-1}{k} \binom{N-k-2}{i-k} (k+1)^i = N^i. \tag{3.20}$$

Again, both sides of (3.20) are polynomials of N of the degree i. It is sufficient to prove that both sides of (3.20) are equal for $i+1$ different points. We can select these points to be $N = 1, \ldots, i+1$.

If $k > N-1$ then $\binom{N-1}{k} = 0$. Due to $N < i+2$ if $k < N-1$ then $\binom{N-k-2}{i-k} = 0$, hence the only non-zero item of the sum is the one corresponding to $k = N-1$, but

$$(-1)^{i-k} \binom{N-1}{N-1} \binom{-1}{i-k} N^i = N^i.$$

□

3.4 Finite-size scaling

In order to study the critical distribution in a finite system it is beneficial to find the explicit dependence of the observables on the system size. The equation that summarizes such dependence is referred to as *finite-size scaling*. In this section we develop scaling relations for the avalanche size distribution of the network 3.1 in the critical state. We use classical procedures described in Refs. [78, 38, 39, 29].

In Fig. 3.3 (left) distributions of the avalanche sizes are plotted for several values of N. The large-L cutoff scales with N, which is necessary condition to obtain a truly critical distribution in the limit. Assume, that

$$P_c(L, N) = \mathrm{const}_N L^{-\gamma}. \tag{3.21}$$

Denote the critical distribution of the avalanche sizes in the network of N neurons by $P_c(L, N)$. Criticality is defined as before by the distance to the power-law distribution. In the critical state the connectivity parameter α scales with N as

$$\alpha = 1 - N^{\mu}, \tag{3.22}$$

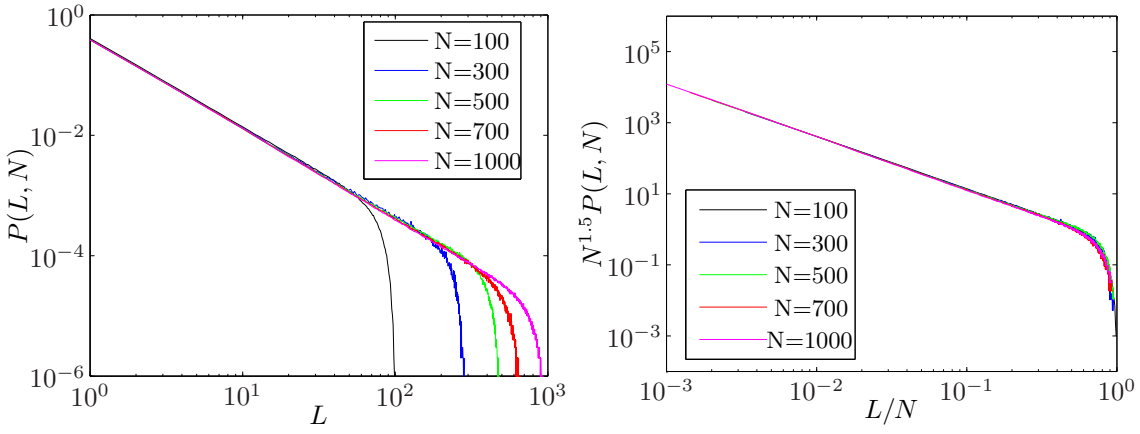


Figure 3.3: Critical distribution of the avalanche size for different network sizes. (left) avalanche size distribution without final-size scaling; (right) axes scaled by $1/N$ and $N^{1.5}$. N=100, 300, 500, 700, and 1000

where $\mu \approx -0.5$ [53]. Thus, for the critical distribution the average avalanche size is given by

$$\langle L \rangle = \frac{N}{N-(N-1)\alpha} \sim \frac{N}{N-(N-1)(1-N^{-0.5})} \sim \sqrt{N}. \tag{3.23}$$

This results differs from the sandpile models, where $\langle L \rangle \sim N^2$ for undirected models, and $\langle L \rangle \sim N$ for directed models [48, 78].

The finite-size scaling ansatz reads

$$P_c(L,N) = N^{-\beta} g\left(\frac{L}{N^\nu}\right), \tag{3.24}$$

where g is a universal scaling function and $\beta > 0$ and $\nu > 0$ are critical indices. The constant ν describes how the cutoff size scales with system size. We assume that $g \in C(\mathbb{R}_+)$ and $g(0) \in \mathbb{R}$. Let us rewrite Eq. 3.24 in the form of Eq. 3.21

$$P_c(L,N) = N^{-\beta}\left(\frac{L}{N^\nu}\right)^{-\frac{\beta}{\nu}} f\left(\frac{L}{N^\nu}\right) \approx L^{-\frac{\beta}{\nu}} f\left(\frac{L}{N^\nu}\right). \tag{3.25}$$

Therefore

$$\frac{\beta}{\nu} = \gamma = \frac{3}{2}. \tag{3.26}$$

To find a second equation, relating ν and β we can use the average avalanche size

scaling (3.23). We can express $\langle L \rangle$ by Eq. 3.24

$$\langle L \rangle = \sum_{L=1}^{\infty} LP(L,N) / \sum_{L=1}^{\infty} P(L,N) = \int_1^{\infty} LN^{-\beta} g\left(\frac{L}{N^{\nu}}\right) dL \tag{3.27}$$

$$= N^{2\nu-\beta} \int_{1/N^{\nu}}^{\infty} \tilde{L} g(\tilde{L}) d\tilde{L} \sim N^{2\nu-\beta} \quad \text{for} \quad N \to \infty. \tag{3.28}$$

Thus we have a system of the simple equations

$$\frac{\beta}{\nu} = 3/2 \tag{3.29}$$

$$2\nu - \beta = 1/2, \tag{3.30}$$

which is solved by $\nu = 1$, $\beta = 1.5$.

To check whether the scaling ansatz 3.24 is correct for the avalanche size distribution we "descale" distributions for different values of N and plot them on the same graph. In Fig. 3.3 we see that after descaling all distributions fall onto one curve. This shows that the obtained finite-size scaling is a good approximation for the influence of the finiteness of the system on the avalanche size distribution. It also justifies conditions on smoothness of scaling function g.

The obtained finite-size scaling relation

$$P_c(L,N) = N^{-\frac{3}{2}} g\left(\frac{L}{N}\right), \tag{3.31}$$

allows to justify the numerical prediction of the power-law exponent for the finite systems. It also shows that with an increasing system size the power-law cut-off is increasing, leaving in the limit of $N \to \infty$ the perfect power-law distribution.

Chapter 4

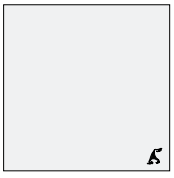
Long term plasticity

In this chapter we present an extension of the model from the previous chapter, which enables the connection strength to adapt. By this means the network becomes self-organized critical. The adaptation of the synaptic strength follows a homeostatic learning rule, which is appropriate to achieve a robust criticalization and which is present in biological neurons [138]. Some of the results from this chapter are published in [89].

The ability of cortical networks to self-regulate is implied by experiments of Beggs and Plenz [14]. They observed that by application of the inhibition blocking agent picrotoxin the dynamics of the network is moved away from the critical state. During the following hours the critical state is gradually recovered. Such slow adjustment is not compatible with the short time synaptic dynamics in [90, 91], which we will discuss in chapters 5 and 6, but suggests a slow adaptation process that we will study in this chapter.

Commonly it is assumed that criticality is intimately connected to the occurrence of power laws [8, 92, 99]. This fact is usually referred to as the empirical identification of SOC [108, 14, 3]. While the detection of a power law involves temporal and spatial sampling on a global scale, self-organization in a complex system should rely on local mechanisms. The main difficulty lies in the description of the critical state. We propose a local way to describe criticality and based on this we derive a learning rule which genuinely introduces SOC into the neuronal model. The learning rule is based on the trade-off between the reliability of the signal transmission and minimization of the resources spent on it. It turns out that for each neuron the best local strategy is to cause by each emitted spike on average one other neuron to fire.

For any initial condition the system self-tunes to a critical state, where avalanche sizes have a power-law distribution with a critical exponent close to the biologically observed value of about $-3/2$.



4.1 Branching processes

In many studies [14, 86, 93] of self-organized criticality branching processes were used as a mean-field approximation to the model dynamics of the system activity during avalanches. In some cases [5] this approach is justified by the specific system topology. For instance, in the case of the sandpile model only immediate neighbors can be activated. Then the branching process approximation is 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. However, neural networks that resemble cortical slices are much more complicated than networks where the branching process approximation is exact. Thus it is questionable whether the branching processes are a valid approximation. To answer this question, we studied analytically the deviation of the neural network model from a branching process. In the present chapter we present a sketch of the proof. Chapter 8 deals with the exact formulation of the problem and the complete proof.

The Galton-Watson branching process was first introduced in 1874 to explain the disappearance of British family names [147]. Since then, they are extensively studied in mathematics [6, 66] and applied in biology and physics [80]. It was observed [15], that branching processes approximation fits the data obtained from multielectrode recordings, but the question of adaption was not addressed as well as mechanisms underlying applicability of such approximation were not discussed.

Branching processes describe a hierarchical structure, where units produce offspring with a certain probability that does not depend on the unit or on the offspring of earlier generations. Here is, roughly speaking, a difference to the present neural model, where the effectiveness of a spike in activating other units depends on whether these have been firing recently. The intuitive idea in the following is to show that in the limit of a large network this distinction becomes irrelevant.

Recall that an avalanche is an interval of network activity which is framed by two intervals without activity. Between avalanches the network is driven by noise. Let us reduce the model, excluding the dynamics between the avalanches. In order to do so we assume, that the external input is mixing and provides a uniform distribution of the membrane potentials before each avalanche.

To describe the evolution of an avalanche we consider the random variables ξ_t that represents the number of neurons which are active at time $t = 1, 2, \ldots$ An avalanche starts always with one neuron firing, thus $\xi_1 = 1$. Stochasticity in the process is brought by the randomness of the initial conditions. After the initial conditions are determined the ξ_t are defined deterministically for all t.

Let us give a strict definition of the branching processes.

Definition 4.1.1. Let $M_0, M_1, M_2, \ldots$ be a Markov process, M_i takes values in $\mathbb{N}_0$, $M_0 = 1$. Let $\mathcal{P}$ be the distribution of M_1 ($P(M_1 = i) = p_i$). Let $P(M_{i+1} = 0|M_i = 0) = 1$ and $(M_{i+1}|M_i = k)$ distributed as a sum of k independent random variables distributed as M_1. Then $M_0, M_1, M_2, \ldots$ is a Galton-Watson branching process.

The branching parameter α_{br} is an average number of the offsprings of one species.

$$\alpha_{\text{br}} = E\left(\xi_2 \,|\xi_1 = 1\right) = E\xi_2. \tag{4.1}$$

The process $\{\xi_t\}_{t=1}^{\infty}$ is obviously not a branching process. It is even not a Markov chain, because $\xi_t \leq N - \sum_{s=1}^{t-1} \xi_s$ and therefore ξ_t depends not only on the preceding value ξ_{t-1}. However, in the limit of large network size we can prove the following theorem.

Theorem 4.1.1. *Assume that before an avalanche, the vector of all membrane potentials (except for the neuron which gets external input) is uniformly distributed in $[0, 1)^N$, let α be the connectivity parameter of the network and ξ_t the number of neurons that fire in the t^{th} time step of the avalanche. Then the distribution of the random sequence $\{\xi_t\}_{t=1}^{\infty}$ converges in distribution to a Galton-Watson branching process with the branching parameter α, as N goes to infinity.*

To apply the theorem to neural networks we should check that the membrane potentials prior to an avalanche are distributed uniformly in $[0, 1)$. In Fig. 4.1, we visualize the 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.

The theorem (proof is given in chapter 8) allows us to use results from the theory of branching processes. We formulate one of them here in a form suitable for the particular case of the branching processes which are limits of the neural avalanche processes.

Theorem 4.1.2. *[112]: Let $\{\xi_t\}_{t=1}^{\infty}$ be a branching process, and α is a branching parameter, then*

$$P\left(\sum_{i=1}^{\infty} \xi_i = L\right) = d\left(\frac{a}{2\pi\alpha f''(a)}\right)^{\frac{1}{2}} \alpha^{-L} L^{-\frac{3}{2}} + O(\alpha^{-L} L^{-\frac{5}{2}}),\ L \to \infty. \tag{4.2}$$

In the critical case, when $\alpha = 1$, the avalanche size $\xi = \sum_{t=1}^{\infty} \xi_t = r$ has the probability distribution:

$$P(\xi = r) = C r^{-\frac{3}{2}} + O(r^{-\frac{5}{2}}),\ r \to \infty.$$

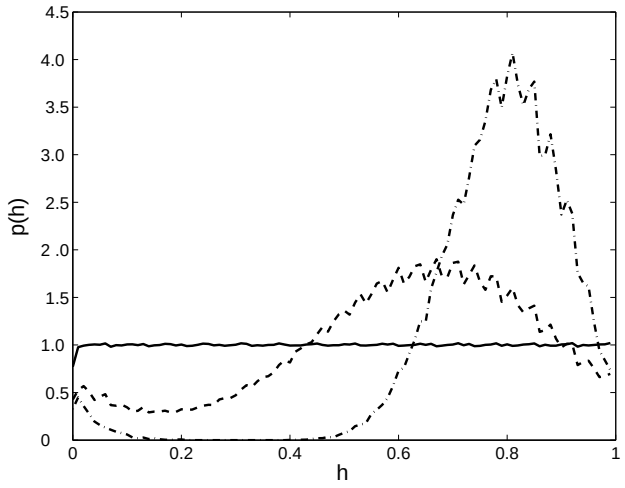


Figure 4.1: The probability density of the membrane potentials immediately before an avalanche is triggered. Initially all potentials are set to $h_i = 0.5$. The snapshots after 100 (dashed-dotted), 1000 (dashed), and 10^5 (solid) avalanches illustrate the ergodicity of the state space.

Therefore, the neural network can be critical only if the branching ratio is equal to 1, and in this case the power-law exponent is uniquely given by $-3/2$.

4.2 Self-organization in a network with long term plasticity

We derive a synaptic plasticity rule, which is intuitively governed by a conservation principle. While correlational learning rules in excitatory networks tend to increase the average network activity here instead the objective is chosen to achieve a constant activity, which is realized by minimizing the functional

$$E(t) = \frac{1}{2}\left(\sum_i S_i(t) - \sum_i S_i(t-1)\right)^2, \tag{4.3}$$

where $S_i(t) = 1$ if neuron i fires at time t and $S_i(t) = 0$ otherwise. The averages are taken over all neurons in the system. In a continuous time system appropriate temporal averages have to be used instead of the reference to the previous time step. In order to arrive at a learning rule for a single neuron we condition Eq. (4.3) on the activity of the firing of the neuron and adjust the synapses so that the firing event causes on average one firing in the postsynaptic neuron j. Let us consider

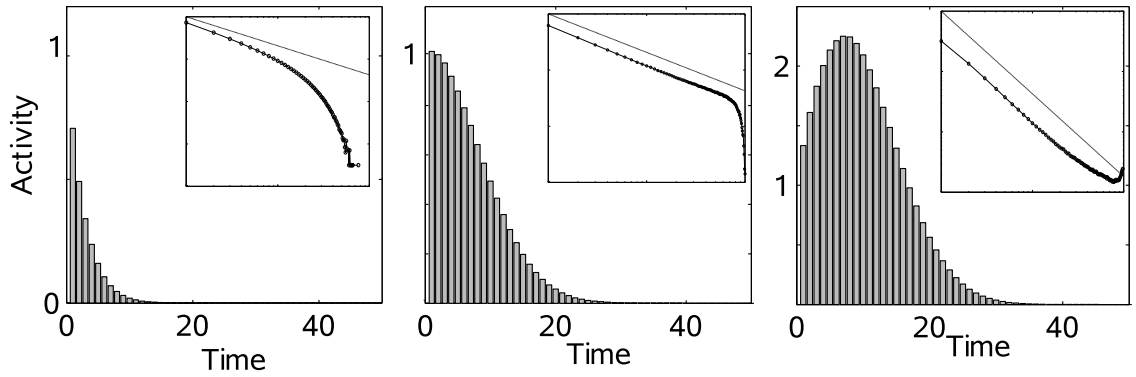


Figure 4.2: Avalanche width as a function of the time after avalanche onset. The values are obtained in a network of $N = 100$ neurons with constant synaptic efficacies and are averaged over 10^5 avalanches without a conditioning on the total avalanche size. Insets depict the corresponding avalanche size distribution. The results for the subcritical case ($\alpha = 0.7$) are given on the left, in the center figure the critical case is presented ($\alpha = 0.9$), and on the right the supercritical regime ($\alpha = 0.98$) is displayed.

the situation at the beginning of an avalanche, then (4.3) becomes

$$E_i(t) = \begin{cases} \frac{1}{2}(\sum_j S_j(t) - 1)^2 & \text{if } S_i(t-1) = 1 \\ 0 & \text{otherwise .} \end{cases} \tag{4.4}$$

In a globally coupled network the total activity $\sum_i S_i(t)$ can be calculated locally from the internal input at the next time step, but also in a massively but not fully connected network a reasonable estimate of $\sum_i S_i(t)$ is available. In a continuous-time system the aspect of causation is usually expressed by a windowed temporally asymmetric learning rule [18].

To characterize the activity of the network from a different perspective, we simulated the neuronal network from chapter 3 and plotted the histogram of the avalanche shapes for different connectivity parameters in Fig. 4.2. Each bar of the histogram represents the average number of neurons that fired at a certain step of the avalanche starting with the second one. Since an avalanche always starts with the activation of a single neuron, the first bar is the average number of neurons activated by a single neuronal discharge. In a branching process it can be considered as the branching ratio. One can clearly see, that for the critical case the branching ratio is approximately equal to one. For the sub- and supracritical cases the branching ratio is respectively smaller or larger than one.

In section 4.1, we showed that the avalanche dynamics can be described by a branching process with the branching parameter α. A critical state in the network cor-

responds to a critical branching process, i.e. in the case that each firing neuron on average entails the firing of one other neuron. This suggests proposing a learning rule which drives the network towards its critical state: Assume the neuron starts the avalanche by first crossing the threshold due to external input trigger $l = \sum_i S_i(t)$ neurons to fire. Then the synaptic weights in the network should increase if $l < 1$, and decrease if $l > 1$. This consideration leads to the learning rule

$$\alpha^{t+1} = \alpha^t + \varepsilon(1 - l), \tag{4.5}$$

where $\varepsilon \ll 1$ is a learning rate. The simulated networks are naturally finite. In Ref. [53] finite size-effects have been shown to be relevant also for relatively large networks and the considerations in section 4.1 also suggest the presence of finite size effects. In the simulations the maximal number of neurons responding to a single spike is bounded. This leads to an underestimation of the branching parameter. To compensate for this effect we use a $1/N$ correction. In the limit $N \to \infty$ this correction is vanishing and the branching parameter is approaching one. As a result of this learning scheme, the synaptic weights converge to a value α_c which was identified as the critical value for finite system sizes in chapter 3, [53]. Fig. 4.3 (left) shows a typical example of the adaptation dynamics.

To improve the performance and to make the model more realistic we consider also independent synapses, i.e. each neuron has its own parameter of synaptic strength. In this case independent learning for different synapses is applicable, i.e.

$$\alpha_i^{t+1} = \varepsilon\alpha_i^t + (1 - l) - 1/N. \tag{4.6}$$

In Fig. 4.3 (left) one can see, that during learning all weights are slowly converging to the critical value, computed for a case of the single connectivity parameter. The dynamics of the avalanche size distribution in the course of learning is shown in Fig. 4.3 (right).

To further investigate the learning dynamics we estimate the best matching power-law exponent and the mean-squared deviation from the best matching power-law at different stages of learning. A network was initiated with random connectivity parameters which are uniformly distributed on $[0, 1]$. Every 10^5 steps the learning process was interrupted and the network continued to operate with fixed connectivity to collect statistics of the avalanche size distribution. Results of this procedure are plotted in Fig. 4.4. The left part of the figure shows, that the exponent approaches the critical value of $-3/2$. The right side shows the deviation from the estimated power law. After $4 \cdot 10^5$ learning steps the optimal performance is reached and the distribution of the best matched power-law exponents and mean-squared deviations is stationary with mean value 0.006 and standard deviation 0.001. To check the stationarity we used the augmented Dickey-Fuller test, and the obtained statistics enable us to reject the hypothesis of nonstationarity at the 95% confidence

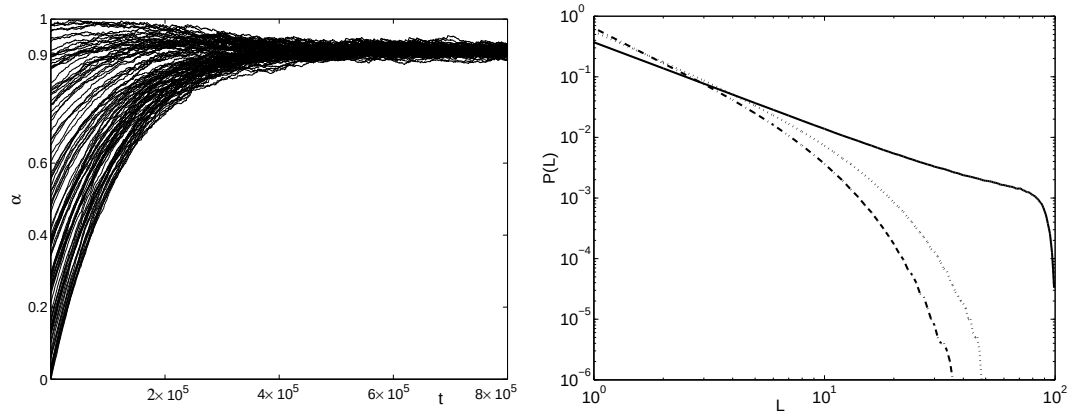


Figure 4.3: (Left) Learning dynamics of individual synapses. The lines represent the trajectories of the synaptic strengths of a single neuron for up to $8 \cdot 10^5$ steps in order to demonstrate the initial learning phase. The weights are initialized by a uniform distribution. A network of size $N = 100$ has been trained for $2 \cdot 10^6$ steps with learning rate $\varepsilon = 10^{-4}$. (Right) Evolution of the avalanche size distribution before learning (dashed), after 10^5 avalanches (dotted), and after $2 \cdot 10^6$ avalanches (solid).

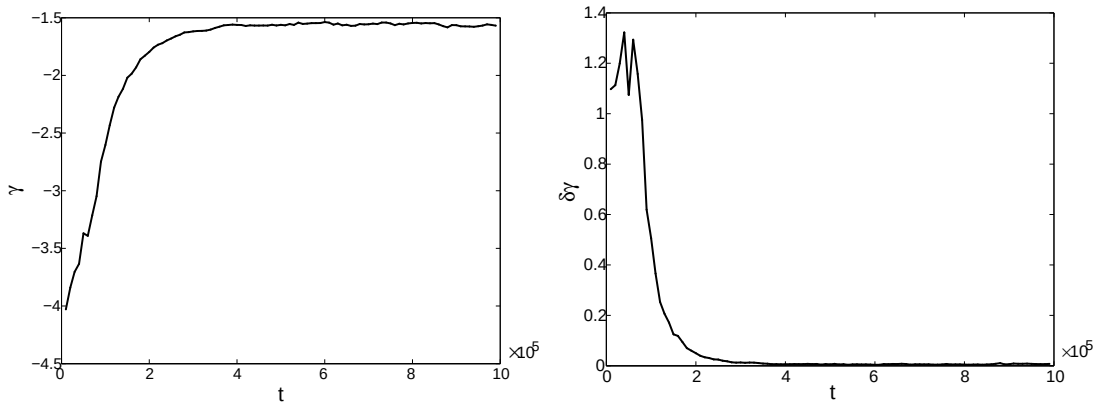
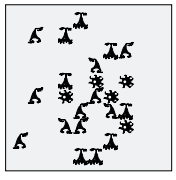


Figure 4.4: (Left) Exponent of the power-law fit to the avalanche size distribution in the course of the learning. (Right) Time evolution of the fitting error. The fit is performed for avalanche sizes $L/N \leq 2/3$. Other parameters are same as in Fig. 4.3.



level.

The neuronal network equipped with such a learning rule is indeed self-organized critical. There is no need to tune synaptic parameters externally, and the avalanche dynamics together with the slow learning drives the system to a critical state independent of the initial conditions.

We have further considered the question whether networks with partial connectivity show the same behavior as reported above. We conduct a simulation, where each neuron was connected randomly to one tenth of the other neurons. For compensation the synapses were scaled by a factor of ten. The thus rescaled network with static synapses exhibits critical behavior for synapses strengths which are indistinguishable from the results for the fully connected case as can be expected from simple statistical considerations. If the synapses are allowed to adapt, the behavior is as well in perfect agreement with the fully connected case.

In the biological context the question of the physiological realizability of the learning rule is relevant. It is known that homeostatic regulation is present e.g. in mechanisms for regulating the firing rate by synaptic scaling in the rat hippocampus [138]. These mechanisms couple correlational and normalizing effects and are considered to bring about homeostatic regulation [81]. In a biologically plausible combination of homeostatic learning and coincidence detection the criticalizing mechanisms could turn out to be realizable. However, recording techniques are not yet developed enough to provide experimental evidences for or against realization of our leaning rule.

Chapter 5

Depressing synapses

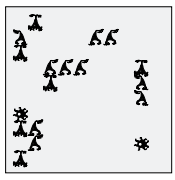
As we showed in chapter 3 a network with static synapses reaches a critical state for the precisely tuned connectivity parameter. One can say that self-organization in this model extends only to the distribution of the membrane potentials and not to the connection strength.

When power-law statistic of event sizes, indicating the critical state, were found in the recordings from cortical slices [14, 15], the necessity to find new models emerged. Connections between cortical neurons are subject to numerous changes due to short and long-term synaptic plasticity. Thus the model, which is employed to mimic neuronal self-organization, should also include self-organization of the synaptic strength.

In chapter 4 we studied a possible long-term plasticity rule which drives the system to the critical state. In this chapter we concentrate on the influence of short term plasticity in the connections between pyramidal neurons [101] on the network dynamics. We show that, due to such plasticity mechanism, the system is guided to the SOC state.

We derive an analytical expression for the average coupling strength among neurons and average inter-spike intervals, and prove that average synaptic strength allows to predict the parameters of the avalanche size distribution. We show that theoretically found values are in a good agreement with the values obtained from network simulation.

An analytical form of the self-consistency equation will be used later in this chapter to study the thermodynamical limit, and to discover that in the large network self-organization leads to the critical state for all parameter values above a certain threshold. We also study the influence of different types of connectivity, the neuronal leakage, and the introduction of an inhibitory population.



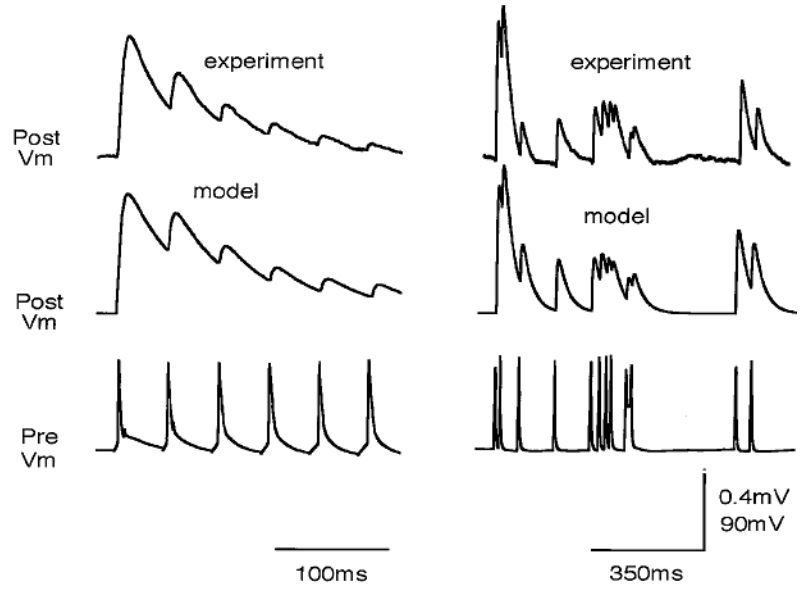


Figure 5.1: Functional synaptic model.(left) Postsynaptic potential generated by a regular spike train (bottom), at a frequency of 23 Hz measured experimentally (top; average more than 50 sweeps), and computed with the model (middle). (right) Same as on the left for irregular spike train. *Picture is taken from [102].*

5.1 Model of the synapse

In this section we briefly describe the biological background and the model for the special type of the short-term synaptic plasticity, the so called *dynamic synapses*. First observations of it were done by Thomson Deuchars[131], who found that synaptic transmissions in the neocortex depend on the frequency of presynaptic activity. By dual whole-cell patch clamp recording a more detailed picture of such frequency dependence was obtained [101, 102, 135].

For the phenomenological model of the synaptic dynamics the following simplified representation of the synapse was proposed. The synaptic connection in the model is characterized by its absolute amount of "resources", which can be partitioned into three states: effective, inactive, and recovered. Each presynaptic action potential activates a certain fraction of resources available in the recovered state, which then quickly inactivates with a time constant of a few milliseconds and recovers with a time constant of about 1 sec. This model could reflect various possible biophysical mechanisms of synaptic depression, such as receptor desensitization [47] or depletion of synaptic vesicles [105].

Repeating the study [102] the dynamics of the synaptic resources in different states

can be written as

$$\begin{aligned} \frac{dR}{dt} &= \frac{I}{\tau_{rec}} \\ \frac{dE}{dt} &= -\frac{E}{\tau_{inact}} + U_{SE} R \delta(t - t_{AP}) \\ I &= 1 - R - E, \end{aligned} \tag{5.1}$$

where E (effective), I (inactive), and R (recovered) are the fraction of resources in the corresponding state. Time scales of inactivation and recovery are given by τ_{rec} and τ_{inact} respectively. The action potential arrives at time t_{AP} and leads to the instantaneous shift to the effective state of a fraction, U_{SE} (utilization of synaptic efficacy parameter), of synaptic resources available in the recovered state. The postsynaptic response to the action potential is proportional to the amount of synaptic resources in the effective state.

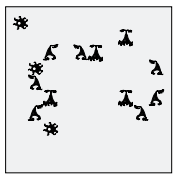
We are attributing such synapses as depressing, because higher frequency of firing leads to the smaller effect of each spike. The steady state response to a periodic firing depends on the frequency of spikes. Thus, depressing synapses are one of the mechanisms connecting the rate code and the precise timing code.

By tuning the parameters of the model U_{SE}, τ_{inact}, and τ_{rec} the authors found very close resemblance to the recordings of paired neurons [102]. In Fig. 5.1 traces obtained from simulation are compared to the averaged over many trial postsynaptic potential.

It is easy to see that in the system of equations (5.1) variables are redundant and their number can be reduced to just two. We will go further and take into account that by the fit of the model to the experimental measurements the inactivation constant was found to be of an order of 3 ms. As we are going to model events, which are typically separated by much longer time intervals, we took that inactivation happens instantaneous. Thus we are left with only one variable, the amount of synaptic resources in the recovered state.

5.2 Network with depressing synapses

Similar to the static model 3.1, we consider a fully connected network of N simple integrate-and-fire neurons. Each neuron is characterized by a membrane potential $0 < h_i(t) \leq \theta$, $i = 1, \ldots, N$, where θ is a threshold. The neurons receive external inputs by a random process $\zeta \in \{1, \ldots, N\}$ which with the rate τ_s selects a neuron $\zeta(t) = i$ whose membrane potential h_i is advanced by an amount I^{ext}. Thus τ_s determines the frequency of the external input. On average one neuron receives external input every $N\tau_s$ milliseconds.



Each neuron integrates inputs until it reaches the threshold. A soon as $h_i(t) > \theta$ the neuron emits a spike which is delivered to all postsynaptic neurons at a fixed delay $\tau_d \ll \tau_s$. The membrane potential is then reset by $h_i(t_{\rm sp}^+) = h_i(t_{\rm sp}) - \theta$. Suprathreshold activity is communicated to other neurons along neural connections of strength $uJ_{ij}(t)$. Where $J_{ij}(t)$ denotes the amount of synaptic resources available at time t in a synapse between neurons j and i. And u is a fraction of the resources which are used for synaptic transmission. As the first neuron reaches the threshold by receiving external input an avalanche of neural activity is triggered. An avalanche is defined same as in chapter 3.

The dynamics of the membrane potential is thus described by the following equation

$$\dot{h}_i = \delta_{i,\zeta(t)} I^{\rm ext} + \frac{1}{N} \sum_{j=1}^{N} u J_{ij} \delta \left(t - t_{\rm sp}^j - \tau_d \right). \tag{5.2}$$

In Eq. (5.2) τ_s presents implicitly as a rate of the external input. Later, for the discrete version of the model we take τ_s as a discretization time step.

The variables J_{ij} are subject to the following dynamics

$$\dot{J}_{ij} = \frac{1}{\tau_J} \left(\frac{\alpha}{u} - J_{ij} \right) - u J_{ij} \delta \left(t - t_{\rm sp}^j \right), \tag{5.3}$$

which describes the amount of available neurotransmitter in the corresponding synapse [101]. The dynamics of J_{ij} is composed of two parts. First, a synapse decreases its strength when activated because it depletes of the resources of synaptic transmitters. Second, it slowly recovers while the neuron is silent. Namely, if a spike arrives at the synapse, the available transmitter is diminished by a fraction u, while in the absence of spikes the synapse recovers and the amount of the resources approaches its a maximal value $\frac{\alpha}{u}$ at a slow time scale τ_J, which depends on τ_s as

$$\tau_J = \nu N \cdot \tau_s. \tag{5.4}$$

Thus if a neuron spikes rarely the synaptic strength uJ_{ij} approaches α.

Without loss of generality and for the simplicity of notation we use $\theta = 1$. If $\theta \neq 1$ then to preserve the meaning of parameters we scale correspondingly the maximal synaptic strength $\alpha_\theta = \alpha\theta$. After such scaling all computations hold same as in the case $\theta = 1$.

To illustrate how the synaptic dynamics depend on the frequency of firing we plotted in Fig. 5.2 the dynamics of synaptic resources. We generate the spike train with different frequencies of spikes. When the firing frequency is small enough (not more then one spikes in 50 time-steps) then the synapse has enough time to recover to the maximal value. If the frequency is larger, then the synaptic strength at firing

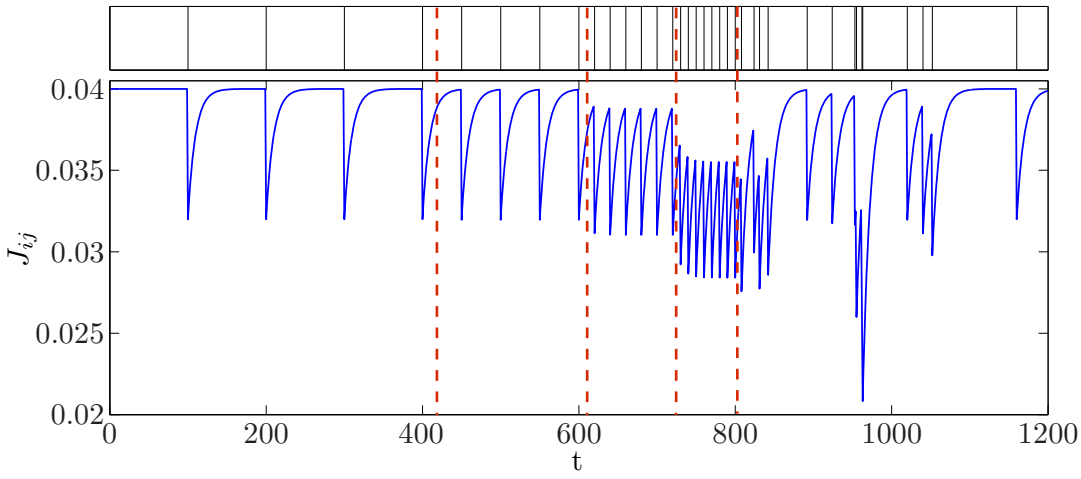


Figure 5.2: Amount of synaptic resources. On the top panel the spike train of the neuron is plotted, on the bottom the amount of the synaptic resources J_{ij}. Dashed vertical lines separate different firing frequencies of the neuron j. From left to right: one spike in 100, 50, 20, 10 time steps. The last interval correspond to a Poisson spike-train with a rate of one spike per 30 time steps. $N = 100$, $\alpha = 0.8$, $u = 0.2$, $\nu = 10$.

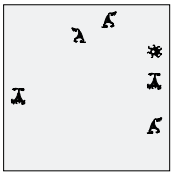
is smaller then maximally possible [101]. In the last block of Fig. 5.2 a neuron fires a Poisson spike-train and then the synaptic dynamics is rather irregular.

The simple intuition about how the described synaptic dynamics (5.3) can lead to the critical state is as follows: when a network is overly active and avalanches are large, synaptic connections are weakened by the exhaustion of synaptic resources. On the other hand when avalanches are small, synapses have time to recover and gain enough strength to trigger a large avalanche.

5.3 Simulation results

In this section we present some results obtained by the network simulation. We study distributions of avalanche sizes and durations as well as the behavior of synaptic parameters.

In Fig. 5.3, we show examples of avalanche size distributions for various values of α. For small values of α, subcritical avalanche-size distributions are observed. This regime is characterized by a negligible number of avalanches that extend to the system size. Near α_{cr} the system has an approximate power-law distribution for avalanche sizes almost up to the size of the system where an exponential cut-off is observed. Avalanche size distributions become non-monotonous when α is well above the critical value α_{cr}.



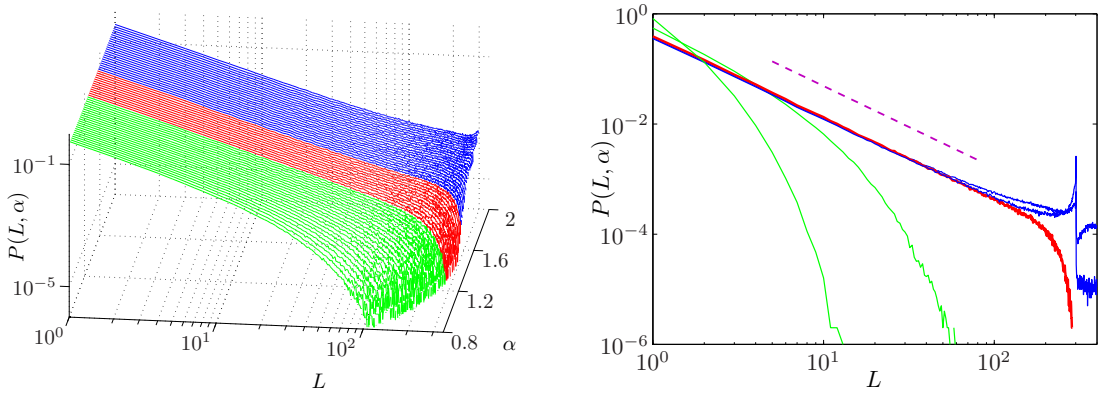


Figure 5.3: Distribution of avalanche sizes for different values of α. (left) At $\alpha < 1.3$ the distribution is subcritical (plotted green). It becomes critical in an interval around $\alpha = 1.4$ (plotted red). For $\alpha > 1.6$ the distribution is supercritical (plotted blue). (right) Characteristic examples of all three kinds of distributions with the same color-code. Results are obtained for $N = 300$, $\nu = 10$, $u = 0.2$, $I^{\text{ext}} = 0.025$. [91]

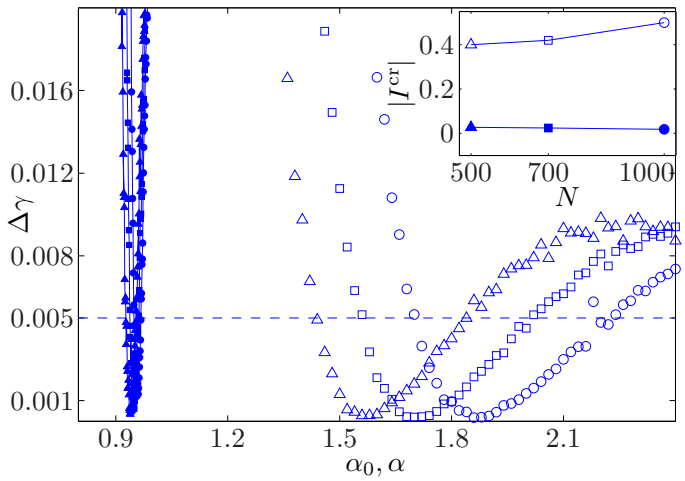


Figure 5.4: The range of connectivity parameters where critical events extend to the system size. The mean-squared deviation $\Delta\gamma$ from the best-matching power-law is plotted in dependence of α_0 for static synapses and in dependence of α for depressing synapses: Empty circles, squares and triangles stand for networks with dynamic synapses and system sizes $N = 1000$, 700, and 500, respectively. Filled symbols represent the static model. The inset shows the lengths of the parameter intervals where the deviation from the best-matching power-law is smaller than 0.005. Symbols are the same as in the main figure.[91]

To compare the obtained results with the static model from the chapter 3 we plotted in Fig. 5.4 the deviations from the best-matching power-law for different connectivity parameters and system sizes for the static and the depressing model. We took the same number of avalanches for both models and scaled it with the square of the system size. To qualify the difference between the models we defined a threshold describing a 'very good fit' and compared the parameter regions which deliver a critical distribution with at least this quality. The difference is one order of magnitude, see inset in Fig. 5.4. The result does not depend on a particular choice of the threshold as long as it is not too large. In a contrast to the static case, the critical region with dynamic synapses apparently increases with the system size. This is a reason to believe that in the thermodynamical limit the network is critical for a substantial fraction of the connectivity parameters.

As in experimental papers [14], we also considered the duration of avalanches. In Fig. 5.5 distribution of avalanche durations are plotted for different maximal synaptic strength α. Similar to the experiments of Beggs and Plenzs [14] the distributions deviate strongly from a power-law for avalanche durations larger than ten and consist mainly of the exponential cut-off, cf. Fig. 5.5 (left). To qualify the distribution we measured the slope of the power-law by fitting the region of the short avalanche durations, where power-law is still an applicable hypothesis. We have done this to check for similarity to the experimental finding, where the measured exponent is found to be equal to -2. We plotted the deviation of avalanche size distribution from the nearest power-law $\Delta\gamma_l$ versus the slope of the duration distribution γ_d Fig. 5.5 (right). Thereby we can see which duration distribution is observed in which regime. At criticality, i.e. when the deviation of the avalanche size distribution from a power-law is minimal, this exponent is approximately equal to -2 in perfect agreement with the experiments.

To summarize, one of the important things that we have seen is that the maximal synaptic strength α determines the behavior of the network. Also there is a large interval of parameters α such that avalanche size distribution follows a power law.

5.4 Statistics of synaptic parameters

In the previous section we described the general statistics of the dynamical regimes produced by the model network. Here we discuss more specific statistics, which allow later to access properties of the model analytically.

A first prominent choice of study is the synaptic strength. In Fig. 5.6 (left) we plotted the typical behavior of the single synapse strength over time in different regimes: subcritical, critical and super-critical. In the right part of Fig. 5.6 the resulting distribution of the synaptic resources in the moment right before a spike

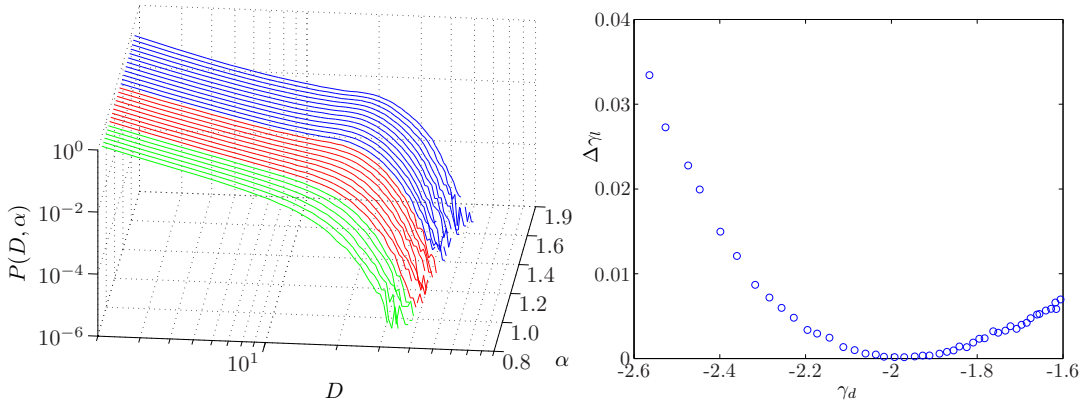


Figure 5.5: Avalanche duration distribution for different values of the connectivity parameter α (left) and power-law exponent (right) fitted in the beginning of the duration distribution as a function of the goodness of the power-law fit to the avalanche size distribution $N = 100$, $\nu = 10$, $u = 0.2$. The color code in the left graph are taken from the Fig. 5.3 (green – subcritical distribution of avalanche sizes, red – critical, and blue – supercritical)

is plotted. If one were to multiply the values of J_{ij} in this plot by the parameter u it will give the distribution of the synaptic strength.

Another characteristic of neuronal activity, which is often used both in biological and theoretical studies is the inter-spike interval (ISI). It is defined as the time between two successive spikes of one neuron. Separation of time-scales makes avalanche durations to be much shorter than inter-avalanche intervals. As it was mentioned before, for the simulation and description of inter-avalanche periods we used discretization with the time-step equal to a rate of the external input, i.e. each moment of time one neuron in the network is receiving external input. Therefore, ISI cannot be negative and it is approximately equal to zero if the neuron spikes twice during an avalanche.

The distribution of the inter-spike intervals reveals the regularity of the spike train of a neuron. For example, if inter-spike intervals are independent and exponentially distributed, then the spike train is Poisson, and all previous history of a spike-train contains no information about a following inter-spike interval. In the depressing synapses model the situation is different, as one can see in Fig. 5.7 (left) the inter-spike interval distributions are of a Gaussian-like shape for all connectivity parameters. ISI can not be negative and it is approximately equal to zero if the neuron spikes twice during an avalanche.

Already from the ISI distributions one can determine a supercritical regime, which is distinguished by a non-zero probability of the zero interspike interval. This is different in the static model, where a neuron can spike twice during an avalanche

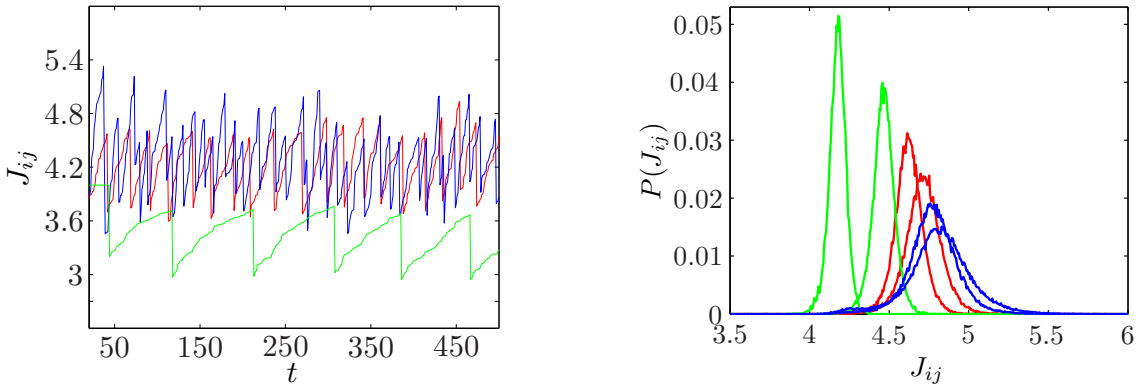


Figure 5.6: Dynamics and distribution of synaptic resources. (left) Dynamics of J_{ij} for different values of the connectivity parameter α: green $\alpha = 0.8$ subcritical, red $\alpha = 1.4$ critical, blue $\alpha = 2.0$ supracritical. (right) Distribution of the synaptic resources at spike-time for different parameters $\alpha = 0.8, 1.0, 1.4, 1.6, 1.8, 2.0$, with the same color-code. $N = 300$, $\nu = 10$, $u = 0.2$.

only in the very supracritical multi-piked regime. Basis for this difference is the variance of the synaptic strength in the dynamic synapses model, which is especially large for the supracritical regime (Fig. 5.6). Thus, dynamic synapses may allow for avalanches of sizes larger than the system size without having a multi-piked distribution. Moreover, due to the inhomogeneity of the synaptic weights, it is possible that a neuron spikes twice in the course of an avalanche of size smaller than N. The typical inter-spike interval is more sensitive to changes of the maximal synaptic strength α when $\alpha < \alpha_{\text{crit}}$, c.f. Fig. 5.7 (left),.

To get an insight in the behavior of the inter-spike interval, we considered also inter-avalanche interval (IAI), i.e. the time between two successive avalanches. In Fig. 5.7 (right) we plotted this distribution for different connectivity parameters. The graph reveals no difference in distributions of IAIs between the three different regimes (Fig. 5.7 (right)). The linear shape of the graphs in the log-lin scale reveals the exponential nature of the distributions. Namely, as it will be shown later IAIs follow a geometric distribution, which is a discrete version of the exponential distribution. Therefore, if one considers the avalanche as one synchronous event (like in the concept of population spikes [96, 134]) the behavior of the network on the slow time scale is similar to the behavior of a Poisson neuron. However the similarity disappears when one examines the temporal structure of an avalanche more closely. Later in this chapter inter-avalanche and inter-spike intervals will be discussed in more detail.

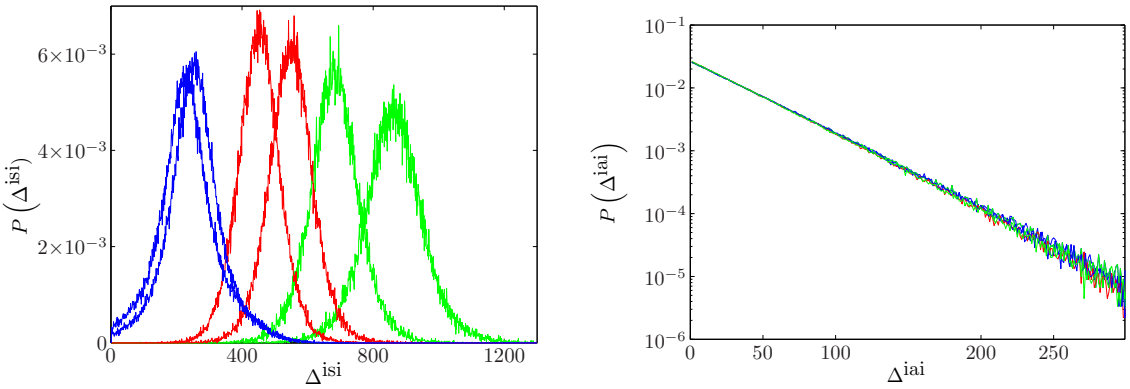


Figure 5.7: The inter-spike (left) and inter-avalanche (right) interval distribution for different parameters α. $N = 100$, $\nu = 10$, $u = 0.2$. Color-code on the graph: green – subcritical, red – critical, and blue – supercritical. For simplicity of notation IAI and ISI are measured by discrete steps of size τ.

5.5 Mean-field approximation

The aim of this section is to relate the parameter α to the distribution of avalanche sizes. Intuitively, it is clear that the synaptic strength is the most important factor for the avalanche size. We would connect α to the synaptic strength and show that the latter allows to describe the avalanche size distribution.

We simplify our task, by considering not the distribution of synaptic strengths, but only its mean $u\langle J_{ij}\rangle$. Numerical simulations suggest that the distribution of J_{ij} is rather narrow bell-shaped (Fig. 5.6), thus first moment exists for any α.

We employ mean-field theory, which is a common approach for studying the dynamics of large globally coupled system. The main idea of this method is to consider a single element of the system as being subject to a local field produced by the combined effect of interactions with the rest of the system, instead of taking into account the full complexity of mutual interactions.

How does the average synaptic strength relate to the avalanche size distribution? To address this question we studied the dependence of the power-law fit to the avalanche size distribution on the average synaptic strength. We simulated the network for different parameters α and plotted slope of the best matched power-law as a function of $u\langle J_{ij}\rangle$ (Fig. 5.8).

Imagine the simplification of the dynamic synapses network, where synaptic strength is fixed to be equal to $\alpha_0 := u\langle J_{ij}\rangle$. Such a simplified network is exactly the static network studied in chapter 3. To understand the connection between the models we simulate the dynamic model and extract $u\langle J_{ij}\rangle$. Then we simulate the static model

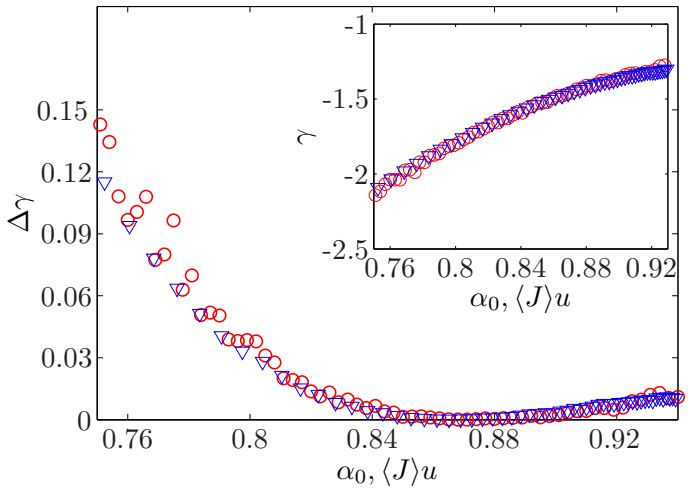


Figure 5.8: Deviation from the closest power-law distribution against the mean synaptic strength for static and dynamic synapses. Circles stand for the static model and triangles for depressing synapses. The inset shows the estimated power-law exponents in both cases. $N = 100,\ \nu = 10,\ u = 0.2$

with synaptic strength α_0 and compare the parameters of the closest power-law for both networks. As shown in Fig. 5.8 there is almost no difference in the parameters of the power-law fit. Thus, to estimate the avalanche-size distribution we can use the distribution obtained for the static model (Eq. 3.5).

Thus the main question is: how to determine the average synaptic strength $u\langle J_{ij}\rangle$ from the maximal synaptic strength α. In order to calculate the average synaptic strength analytically, we consider in addition the neural inter-spike intervals Δ^{ISI}. It also has a bell-shaped distribution (Fig. 5.7) with a finite first moment. On the one hand, if the inter-spike intervals are short then the synapses have a short time to recover and the average synaptic strength resides at a low level. On the other hand, large synaptic strengths lead to long avalanches and to large inputs to neurons during the avalanches, which tends to shorten inter-spike intervals. This trade-off determines the effective synaptic strengths and the inter-spike intervals which are realized by the dynamics of the network. In order to express this reasoning more formally, we solve the dynamical equations (5.2) and (5.3). By this we assume that there is no trend in the changes of the synaptic strength and ISI, i.e. that mean values are stationary.

The idea of the following computation is as follows: first we derive a couple of independent equations connecting $\langle\Delta^{\mathrm{ISI}}\rangle$ and $\langle J_{ij}\rangle$:

$$\langle J_{ij}\rangle = G\left(\langle\Delta^{\text{isi}}\rangle\right), \tag{5.5}$$
$$\langle\Delta^{\text{isi}}\rangle = F\left(\langle J_{ij}\rangle\right). \tag{5.6}$$

Then we compose a self-consistency equation and solve it numerically and graphically, which is equivalent to solving the system of equations (5.5,5.6).

From (5.3) we have for the dynamics of the synaptic resources

$$\dot{J}_{ij} = \frac{1}{\tau_J}\left(\frac{\alpha}{u} - J_{ij}\right) - uJ_{ij}\delta\left(t - t^j_{\text{sp}}\right). \tag{5.7}$$

Let the average inter-spike interval $\langle\Delta^{\text{isi}}\rangle$ be known. Solving Eq. 5.7 between two consecutive spikes of neuron j yields

$$J^{b,2}_{ij} = \frac{\alpha}{u}\left(1 - \left(1 - \frac{u}{\alpha}J^{a,1}_{ij}\right)e^{-\left(t^j_2 - t^j_1\right)/\tau_J}\right), \tag{5.8}$$

where $J^{a,1}_{ij}$ and $J^{b,2}_{ij}$ are, respectively, the synaptic strengths after a spike of neuron j at time t^j_1 and before the subsequent spike at time t^j_2. By the definition of the time-constant τ_J (5.4)

$$\frac{t^j_2 - t^j_1}{\tau_J} = \frac{\Delta^{\text{isi}}}{\nu N}.$$

Within a short interval containing the spike, J_{ij} decreases by a fraction u such that $J^{a_1}_{ij} = (1-u)\,J^{b_1}_{ij}$. The Eq. 5.8 then can be rewritten as

$$J^{b,2}_{ij} = \frac{\alpha}{u}\left(1 - \left(1 - \frac{u}{\alpha}(1-u)J^{b,1}_{ij}\right)e^{-\frac{\Delta^{\text{isi}}}{\nu N}}\right). \tag{5.9}$$

The average amount of the synaptic resources $\langle J_{ij}\rangle$ is equal to the mean synaptic strength $E_s\left(J^{b,s}_{ij}\right)$ before a spike from neuron j. Here E_s is a spike triggered averaging, taken just before the spike is emitted, i.e. $E_s\left(J^{b,s}_{ij}\right)$ is the average synaptic resources of the neuron before a spike, averaged over all spikes of the synapse connecting neurons i and j.

Stochasticity in the synaptic strength is brought by stochasticity in the inter-spike intervals, which are in turn defined by the external input. Thus, we can consider all averages with respect to the distribution of the initial membrane potentials and the external input.

All synapses and neurons in the network are governed by the same set of equations and connectivity structure is symmetric. Lets assume that external input is mixing (for the static network it is checked in chapter 7). This assumption allows us to make temporal averaging instead of averaging among the neurons. Thus the average amount of the synaptic resources is the same for all neurons. At stationarity

$\left\langle J_{ij}^{b,1} \right\rangle = \left\langle J_{ij}^{b,2} \right\rangle$ therefore one can find the average synaptic strength by solving the equation

$$\langle J_{ij} \rangle = \frac{\alpha}{u} - \left(\frac{\alpha}{u} - (1-u) \langle J_{ij} \rangle \right) e^{-\frac{\langle \Delta^{\text{isi}} \rangle}{\nu N}}. \tag{5.10}$$

This allows to write the Eq. 5.5 expressing $\langle J_{ij} \rangle$ as a function of $\left\langle \Delta^{\text{isi}} \right\rangle$,

$$\langle J_{ij} \rangle = \frac{\alpha}{u} \frac{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}}{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle} + u e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}}. \tag{5.11}$$

In order to obtain the Eq. 5.6, we are now going to establish a relation between $P\left(\Delta^{\text{isi}}\right)$ and the inter-avalanche interval distribution $Q\left(\Delta^{\text{iai}}\right)$. As it is shown rigorously for the static synapses in chapter 8, the neuronal membrane potentials before an avalanche are uniformly distributed on the interval $[\epsilon_N, \theta]$, where ϵ_N is a lower bound of $h_i(t_{\text{sp}}) - \theta$ with $\epsilon_N \to 0$ for $N \to \infty$. Under these conditions, $Q(\Delta^{\text{iai}})$ has a geometric distribution (derivation in the Appendix A.3)

$$Q\left(\Delta^{\text{iai}}\right) = \frac{I^{\text{ext}}}{\theta - \epsilon_N} \left(1 - \frac{I^{\text{ext}}}{\theta - \epsilon_N} \right)^{\Delta^{\text{iai}}}. \tag{5.12}$$

Let κ_j be the number of avalanches between two spikes of the neuron j. A mean-field approximation relates the averages of the distributions of inter-spike and inter-avalanche intervals

$$\left\langle \Delta^{\text{isi}} \right\rangle = \langle \kappa \rangle \langle \Delta^{\text{iai}} \rangle. \tag{5.13}$$

The average inter-avalanche interval is given by

$$\langle \Delta^{\text{iai}} \rangle = \frac{\theta - \epsilon_N}{I^{\text{ext}}}. \tag{5.14}$$

In order to determine the average number of avalanches between two spikes of each neuron, we compute the time to reach the threshold by accumulating external inputs and synaptic input from other neurons. Let neuron j fire during the current avalanche, therefore its membrane potential is reset by the value of the threshold θ. Then during the inter-avalanche interval Δ^{iai}, each moment of time the neuron can get the external input I^{ext} with the probability $1/N$. Therefore, during one inter-avalanche interval the neuron accumulates the on average external input of size $I^{\text{ext}} \frac{\langle \Delta^{\text{iai}} \rangle}{N}$. During an avalanche the neuron on average gets the internal input of a size $u \langle J_{ji} \rangle \langle L \rangle$, where $\langle L \rangle$ is the mean avalanche size. The neuron accumulated on average half of the internal input from both avalanches during those it fired, therefore one can compute the number of avalanches between two spikes of the neuron by computing the number of pairs "inter-avalanche interval and following

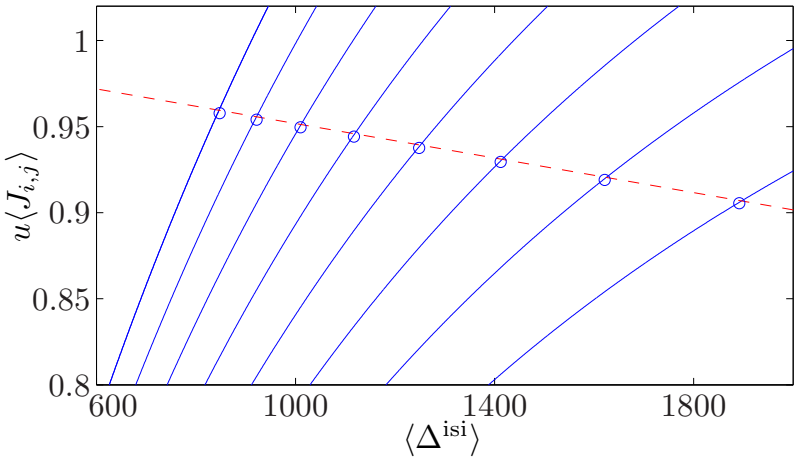


Figure 5.9: Average synaptic strength and inter-spike interval for different initial conditions. The dashed line is the graph of (5.20) and the solid lines are obtained from (5.11) for $\alpha = 1.3, \ldots, 2.0$ in steps of 0.1 (from right to left). Circles show $u\langle J_{ij}\rangle$ and $\langle\Delta^{\text{isi}}\rangle$ obtained from simulations of the network with matching α. Further parameters are $N = 500,\ \nu = 10,\ u = 0.2$.

avalanche", which is required to accumulate input equal to θ. Namely

$$\langle\kappa\rangle = \frac{\theta}{\frac{u\langle J_{ji}\rangle}{N}\langle L\rangle + I^{\text{ext}}\frac{\langle\Delta^{\text{iai}}\rangle}{N}}. \tag{5.15}$$

The distribution of avalanche sizes can be computed analytically for a network with static synapses of strength α_0 [53]:

$$P(L,\alpha_0,N) = L^{L-2}\binom{N-1}{L-1}\left(\frac{\alpha_0}{N}\right)^{L-1}\left(1-L\frac{\alpha_0}{N}\right)^{N-L-1}\frac{N(1-\alpha_0)}{N-(N-1)\alpha_0}. \tag{5.16}$$

In the case of dynamic synapses we use a mean-field approximation and set $\alpha_0 = u\langle J_{ji}\rangle N$ in (5.16). This allows us to compute $\langle L\rangle$ as a function of $u\langle J_{ji}\rangle$.

Combining (5.13), (5.14), and (5.16) we obtain a relation between the inter-spike interval and the average synaptic strength:

$$\langle\Delta^{\text{isi}}\rangle = \frac{\theta-\epsilon_N}{I^{\text{ext}}}\frac{\theta}{\frac{u\langle J_{ji}\rangle}{N}\langle L\rangle\left(u\langle J_{ji}\rangle\right)+\frac{\theta-\epsilon_N}{N}}. \tag{5.17}$$

Another way to compute $\langle\kappa\rangle$ is to imagine neurons as points on a circle of perimeter θ, the distance from zero counted clockwise representing the membrane potential. Then all synaptic interactions are just rotations of the circle by a certain angle.

A spike emitted at crossing of point θ and external input is then a rotation of one selected point. In such a setting an avalanche does not change the order of neurons on the circle. External input can change the sequence of the neurons, but on average it preserves the order. Thus, a neuron can fire a second time only after all other neurons have fired; hence on average it fires every $N/\langle L\rangle$ avalanches, i.e.

$$\langle\kappa\rangle = \frac{N}{\langle L\rangle}. \tag{5.18}$$

In the Appendix A.1 we show that both approaches give the same results as $N \to \infty$. Given the distribution of L one can compute the mean of L, and as shown in secton 3.3

$$\langle L(\alpha_0)\rangle = \frac{N}{N-(N-1)\alpha_0}. \tag{5.19}$$

Thus

$$\left\langle \Delta^{\text{isi}} \right\rangle = \frac{\theta-\epsilon_N}{I^{\text{ext}}}(N-\langle uJ_{ij}\rangle(N-1)). \tag{5.20}$$

Summarizing the preceding computations, we obtained two equations connecting the average inter-spike interval and the average synaptic strength. The combination of them gives a self-consistency equation

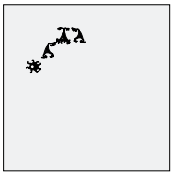
$$\langle J_{ij}\rangle = G[F(\langle J_{ij}\rangle)], \tag{5.21}$$

which can be solved numerically.

We plot in Fig. 5.9 the mean synaptic strength as a function of the average inter-spike interval, which is given by the Eq. (5.11) $\langle J_{ij}\rangle\left(\left\langle \Delta^{\text{isi}}\right\rangle\right) = G\left(\left\langle\Delta^{\text{isi}}\right\rangle\right)$ and an inverse of function relating the average inter-spike interval to the mean synaptic value Eq. (5.20) $\left\langle\Delta^{\text{isi}}\right\rangle^{-1}\left[\langle J_{ij}\rangle\right] = F^{-1}\left(\langle J_{ij}\rangle\right)$. The intersections of these curves give the solution of the Eq. (5.21) and therefore determine the state of the system for any given α. This solution is unique because F and G are increasing functions. For comparison, the results of network simulations for corresponding values of α are also displayed in Fig. 5.9. Simulated quantities are in good agreement with theoretical averages, which further justifies approximations used to obtain the self-consistency equation. The stationary distribution is less sensitive to changes of the parameter α near the critical value of the synaptic strength than further away from it. This brings about the large critical region for the depressing synapses model.

5.5.1 Thermodynamical limit

As it was already indicated in Fig. 5.4, the "critical parameters interval" apparently grows with the system size. In this section we give a solid basis for such growth and discuss the behavior of the system as N goes to infinity.



First let us recall both parts of the self-consistency equation (5.5)

$$\begin{aligned}\langle \Delta^{\text{isi}} \rangle &= \frac{\theta - \epsilon_N}{I^{\text{ext}}} \left(N - \langle u J_{ij} \rangle (N-1) \right), \\ u\langle J_{ij} \rangle &= \alpha \frac{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}}{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle} + u e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}}.\end{aligned}$$

Explicitly, we can write

$$\frac{\alpha \left(1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle} \right)}{1 - (1-u) e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}} = \frac{N}{N-1} - \frac{I^{\text{ext}} \langle \Delta^{\text{isi}} \rangle}{N-1}. \tag{5.22}$$

In the limit $N \to \infty$ one should scale the external input by $I^{\text{ext}} \sim N^{-w}$ with $w > 0$. We will not explicitly solve the self-consistency equation, but consider the approximation of the solution. As the system size growth, the average inter-spike interval scales with N^{λ}. Let us now distinguish the following cases:

1. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1+\epsilon}$ and $\epsilon > w$ then the right hand side (r.h.s.) of (5.22) approaches $-\infty$, while the l.h.s. is larger than 0.

2. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1+\epsilon}$ and $0 < \epsilon \leq w$ then the r.h.s. of (5.22) tends to 1, while the l.h.s. approaches α, hence a solution is only possible if $\alpha = 1$ and in this case $u\langle J_{ij} \rangle \to 1$.

3. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1-\epsilon}$ and $\epsilon > 0$ then the r.h.s. of (5.22) tends to 1, while the l.h.s. approaches 0.

4. If $\langle \Delta^{\text{isi}} \rangle \sim N$, we can assume that $\langle \Delta^{\text{isi}} \rangle = cN + o(N)$. From (5.22) we find that in the limit the self-consistency equation takes the form

$$\frac{\alpha - \alpha e^{-\frac{c}{\nu}}}{1 - (1-u) e^{-\frac{c}{\nu}}} = 1. \tag{5.23}$$

   This equation has a unique solution for any $\alpha > 1$

$$c = -\nu \left(\ln(\alpha - 1) - \ln(\alpha - 1 + u) \right). \tag{5.24}$$

In all cases where a solution exists (2 and 4) we find $u\langle J_{ij} \rangle \to 1$, which is the critical connectivity for the network with static synapses in the limit $N \to \infty$. Therefore, in the thermodynamic limit the network with dynamic synapses becomes critical for any $\alpha \geq 1$.

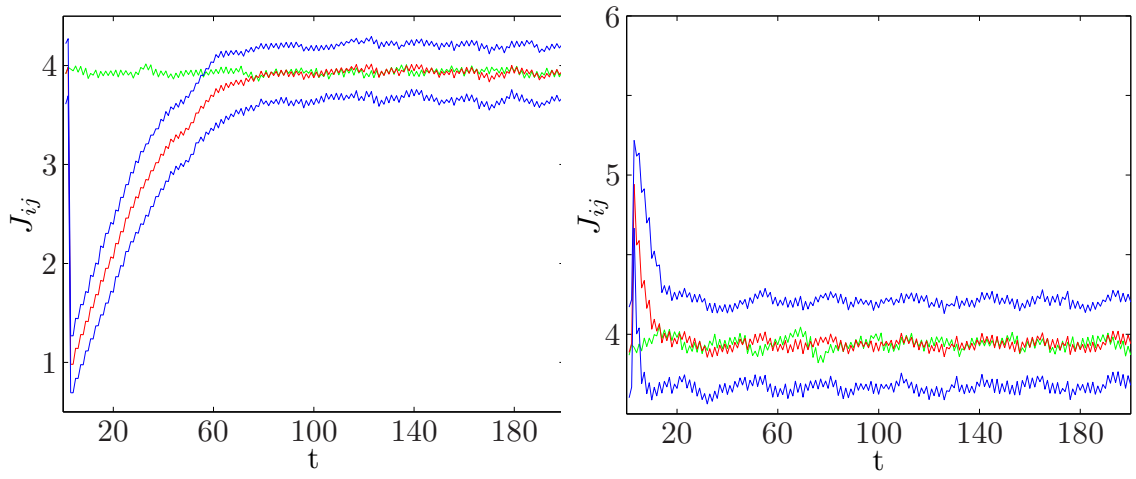


Figure 5.10: Reaction of synaptic strength on perturbation. At time $t = 0$ perturbation of a size -0.03 (left) or 0.01 (right) is applied to all synapses. Red traces - relaxation after the perturbation, green - unperturbed network. All traces are averages over all synapses in the network. Standard deviation bounds are plotted in blue. $N = 100,\ \nu = 10,\ u = 0.2,\ \alpha = 1.0$

5.5.2 Stability of the solutions

Here we discuss the stability of the solution of the self-consistency equation (5.5). We apply a perturbation ΔJ to all synapses at time t_p such that for each i, j: $\widetilde{J_{ij}} = J_{ij} + \Delta J$. If $\Delta J > 0$ a perturbation results in the temporary increase of the avalanche size. Then the average size $\langle \tilde{L} \rangle$ of a few avalanches following t_p will be larger than $\langle L \rangle$. At the same time the inter-avalanche intervals depend only on the external input and the distribution of membrane potentials, therefore such a perturbation would not change subsequent inter-avalanche intervals compared to the unperturbed network. Let us examine how the average number of avalanches between two subsequent spikes of the neuron changes after the perturbation

$$\tilde{\kappa} = \frac{N}{\langle \tilde{L} \rangle} < \kappa.$$

Thus the inter-spike intervals decrease on average.

$$\widetilde{\Delta^{\text{isi}}} = \tilde{\kappa}\widetilde{\Delta^{\text{iai}}} = \tilde{\kappa}\Delta^{\text{iai}} < \kappa\Delta^{\text{iai}} = \Delta^{\text{isi}}.$$

The change of the inter-spike interval influences the average synaptic activity. The average synaptic strength at the time of the second spike after perturbation can be

found from (5.3),

$$\begin{aligned}\langle \tilde{J}_{ij}^2 \rangle &= \frac{\alpha}{u} - \left(\frac{\alpha}{u} - (1-u)\left(\langle J_{ij}\rangle + \Delta J\right)\right) e^{-\frac{\widetilde{\Delta^{\mathrm{isi}}}}{\nu N}} \\ &= \frac{\alpha}{u} - \left(\frac{\alpha}{u} - (1-u)\langle J_{ij}\rangle\right) e^{-\frac{\widetilde{\Delta^{\mathrm{isi}}}}{\nu N}} + (1-u)\Delta J e^{-\frac{\widetilde{\Delta^{\mathrm{isi}}}}{\nu N}} \\ &< \frac{\alpha}{u} - \left(\frac{\alpha}{u} - (1-u)\langle J_{ij}\rangle\right) e^{-\frac{\widetilde{\mathrm{isi}}}{\nu N}} + (1-u)\Delta J < \langle J_{ij}\rangle + \Delta J.\end{aligned}$$

Therefore, positively perturbed synapses decrease their strength to return to the equilibrium state. The same arguments show that negatively perturbed synapses increase their strength after perturbation to counteract it.

We also verified the above arguments by numerical simulation of the network perturbation. In Fig. 5.10 the reaction to the external perturbation is plotted. The graph shows how the average synaptic strength in the simulated system is driven back to the fixed point by a few spikes. Thus the solution of (5.5) is indeed stable for any α.

5.6 Diverse connectivities

For the analytical considerations and simulations presented in previous sections we considered fully connected networks. This was not only for the reason of analytical simplification, but also to separate effects coming from the internal neural adaptation from those determined by specific connectivity. In the literature there are examples showing how the networks of simple neurons equipped with specific connectivity can achieve critical behavior. For example, in the article of Lin and Chen [92] the neurons are represented by the simple random number generator and only the connectivity influences the avalanche size. The authors found that the small-world connectivity is then the condition of criticality in the network.

The topic of this section is to describe, how the combination of specific connectivity and depressive synapses shapes the statistics of the network dynamics. We first introduce some different connectivity patterns and then compare them. To start with, let us consider a simple random connectivity scheme. Each neuron is connected to $p_1 N$ randomly selected neurons, where $p_1 < 1$ is connection fraction, which is the same for all neurons. In the brain it is assumed that the average connectivity is around 10% [79]. We take $p_1 \sim 0.2$ to decrease a simulation size, but results do not depend on a particular choice of p_1 if it is large enough. Such a random network is statistically homogeneous and connected if p is large enough. Therefore, the mean-field approximation is still valid and the critical regime is specified by rescaling α. Namely, the critical α is approximately equal to $\alpha_N^{\mathrm{cr}}/p_1$, where α_N^{cr} is obtained from the critical parameter region of the fully connected network of

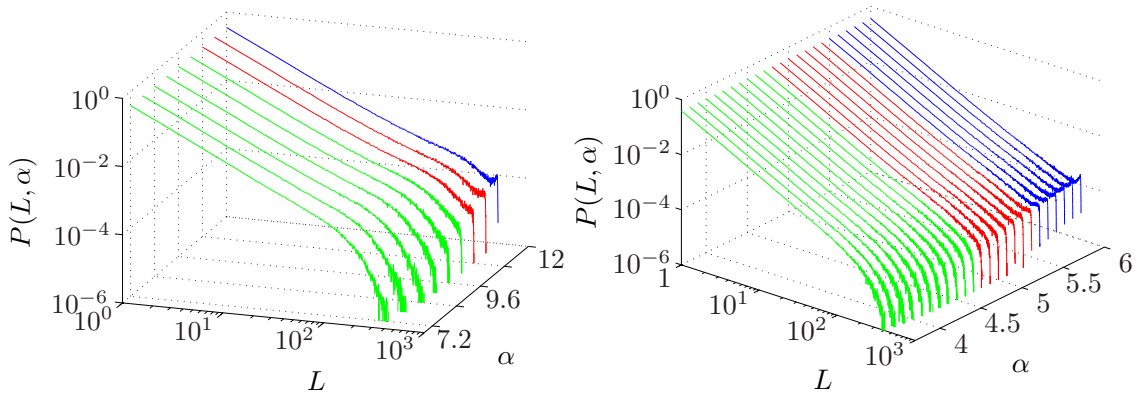


Figure 5.11: (left) The distribution of avalanche sizes in the network with nearest-neighbor connections. To determine the "most critical" distribution the threshold for the least-squares deviation was set at 0.01. (right) The distribution of avalanche sizes in the network with the small-world connections. To determine the "most critical" distribution the threshold for the least-squares deviation was set at 0.005. $N = 500$, $\nu = 10$, $u = 0.2$, average connectivity is 20%.

size $p_1 \times N$. A slightly different connection scheme is a random network without restriction on the degree of each node. For every pair of neurons the probability to form a connection is fixed and equal to p_2. One have to take p_2 to be not too small, namely $p_2 \gg 1/N$, which guarantees that the network is still connected. As above we take $p_2 \sim 0.2$. There is room for a significant difference from the previous model and therefore response of the network to spikes from different neurons may differ.

To study more biologically realistic network models we constructed small-world networks. A procedure for creating small-world connectivity structure has been proposed by Watts and Strogatz [148]:
1) Start with a one-dimensional ring-lattice with N sites. Each lattice site is connected to its $2k$ nearest neighbors by a bond so that for N interacting sites, or neurons, we have $2Nk$ bonds.
2) Randomly choose two sites of the lattice and place a bond between them. Self-connections and duplicate links are excluded. Then one of the bonds going to a neighbor site of one of the end points of new bond is removed.
3) Repeat step 2 until the number of "re-wired" bonds is the fraction ω of all bonds of the original lattice, i.e. ωkN.

If $\omega = 0$ then the network is a regular nearest-neighbor network and if $\omega = 1$ then network is purely random. In a standard small-world network $\omega \sim 1\%$.

We compare now these four types of connectivity (nearest neighbors, two random networks and small-world connectivity, with respect to the avalanche size distributions they provide). For comparison we selected parameters such that the average

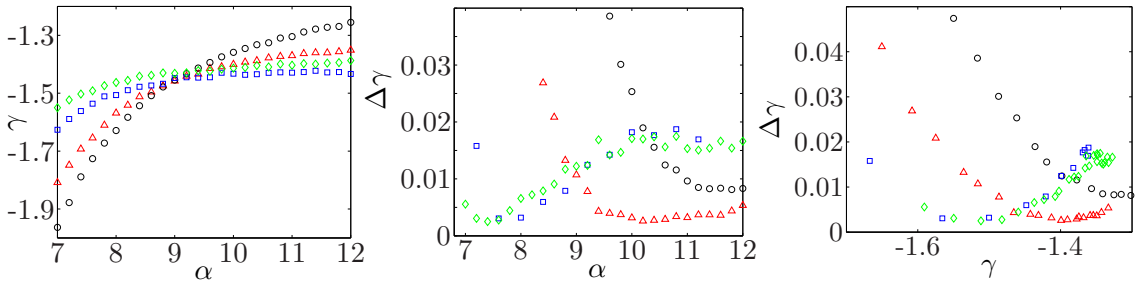


Figure 5.12: Comparison of different connectivity schemes. (left) Exponent of the best matching power-law for different α, (middle) deviation from the best-matched power-law for different α, (right) deviation from the best-matched power-law in dependence of the exponent of the best-matched power-law. Different symbols and colors stay for different connectivity: black circles - nearest neighbors, red triangular - small-world, blue squares - random connectivity with fixed number of connection per neuron, green diamonds - random connectivity. $N = 500$, $\nu = 10$, $u = 0.2$, average connectivity 20%.

connectivity is the same in all networks, i.e. we took $p = p_1 = p_2 = 2k/N$ which on average lead to pN connections per neuron. In Fig. 5.11 we plotted the distribution of avalanche sizes for nearest neighbors connectivity and small-world connectivity. In case of the nearest neighbor connectivity there exists no parameter which brings about critical distributions i.e. when the small avalanches are distributed according to a power-law, the distribution is already not monotonous. This is also clear from the Fig. 5.11 (left).

To understand the mechanism that leads to such a difference let us imagine the avalanche to start at some neuron n and each neuron is connected to $2k$ neighboring neurons. At the next time step only some of its neighbors, i.e. neurons with index $n-k, \ldots, n+k$ can fire and communicate internal input to all of their neighbors. Thus each avalanche creates a set of neurons with a large common input, similar to the fully-connected network. On the other hand the synaptic strength should be scaled with N/k and therefore common input by absolute value is much larger then in the fully connected network, which provides larger synchronization and non-monotonicity of distribution in the parameter interval, where one would expect to find a critical distribution. In contrast to the nearest neighbors connection, the small-world network produces a critical region that is even larger than a fully connected network.

In Fig. 5.12 we compared all four networks with respect to the parameters of the distributions they generate. On the left side the best-matching power-law exponent is plotted as a function of the maximal synaptic strength α. All the obtained curves have a very similar shape, but the nearest neighbors curve is the steepest. In the

middle the deviation from the best matched power-law is plotted. The distribution is critical if the deviation from the nearest power-law is small enough, i.e. $\Delta\gamma < \theta_{\text{pl}}$. Independent of the particular choice of small θ_{pl} the critical region corresponding to the random network with a fixed number of connections per neuron is smaller than the critical region corresponding to the unrestricted random network which is turn smaller than the critical region corresponding to the small-world network.

In the right side of Fig. 5.12 the deviation from the best-matching power-law exponent versus the best matched power-law exponent is plotted. This graph allows to determine which exponent the avalanche-size distribution has when a power-law is indeed the good approximation for it.

Summarizing the comparison above we conclude: If the network is sparse enough ($k \ll N/2$) the nearest-neighbor connectivity does not generate critical behavior. Distributions produced by two types of a random network are very similar. The best matching power-law in the critical regime also has the same exponent, as it is shown in Fig. 5.12 (right). But the unrestricted random network has a critical parameters interval which is slightly larger (Fig. 5.12 (middle)). The small-world network generates a power law with a smaller critical exponent and has the largest critical parameters region.

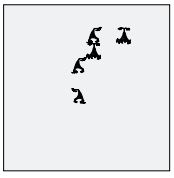
5.7 Leakage

In the real neuron there is always flux of the ions through the passive channels in the membrane (section 2.1). This results in the homeostatic regulation of the membrane potential. Such regulation is referred to as *leakage* because it leads to forgetting of the previously occurred synaptic inputs. And if the neuron is depolarized and close to firing, then leakage can delay a spike.

There are examples where the introduction of an arbitrarily small leakage to the model neurons destroys the results obtained for a leak-free system, cf. e.g. [60, 132]. Thus, robustness of the model with respect to leakage is clearly an important issue. We have performed numerical simulations that unambiguously show the structural stability of our model w.r.t. non-zero leak. It is sufficient for leak-stability that the loss currents are compensated by an active process which we include in the model by a simple adaptive threshold. Let us introduce leakage with time scale τ and a slightly subthreshold current in addition to the dynamics of the membrane potential:

$$\dot{h}(t) = -\frac{1}{\tau}h(t) + C + I^{ext} + I^{int}.$$

C is the synaptic current summarizing a regulatory mechanism. It allows the neuron to stay close to the threshold, I^{ext} is a previously described external input and I^{int}



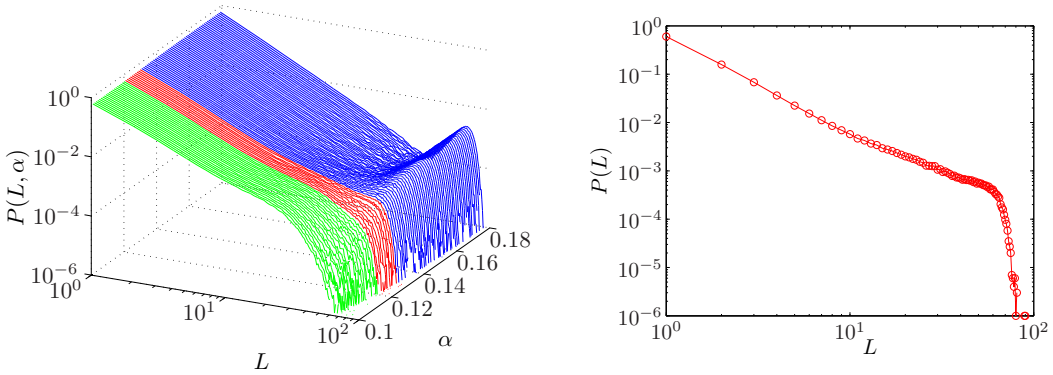


Figure 5.13: Distribution of avalanche sizes for different values of the connectivity parameter α (left) and for the critical case ($\alpha = 0.125$, right). Further parameters are $N = 100$, $C = 0.98$, $\tau = 40ms$.

is the synaptic input. Throughout this thesis time was measured in the units of synaptic delays, which for neurons was found to be about 1 ms [94].

Fig. 5.13 shows the distribution of avalanche sizes for different maximal synaptic strengths. The critical region is determined by the condition that the mean square deviation from the best-matching power-law is not larger than 0.05. The input current keeps all neurons near the threshold; therefore the critical regime is achieved already at smaller maximal synaptic resources. The qualitative behavior of the model, however, does not change when a moderate leakage is introduced. The power-law exponent becomes larger in magnitude reaching approximately 1.9. This is in agreement with experimental data on avalanches found in spike measurements, where the exponent is observed to be around two [13]. Strong leakage suppresses power-laws in our model, but this happens only at leakages exceeding those studied in Ref. [45], which was so far the only model with leaky neurons. Should stronger leaks be observed experimentally in neural systems showing SOC, then our model would need to be extended by additional biological mechanisms.

5.8 Inhibition

So far we were considering networks constructed purely of excitatory neurons, but in the cortical networks there are approximately 10-20% inhibitory neurons [2]. It is known that inhibition often leads to synchronization in neuronal networks, which may destroy a critical state. In this section, we show how the introduction of an inhibitory population changes the behavior of the network.

There are two different ways to include inhibition. The first one follows the so-called Dale's principle [43, 79], which is believed to hold true for many neurons in the brain. It states that each neuron releases only one synaptic transmitter.

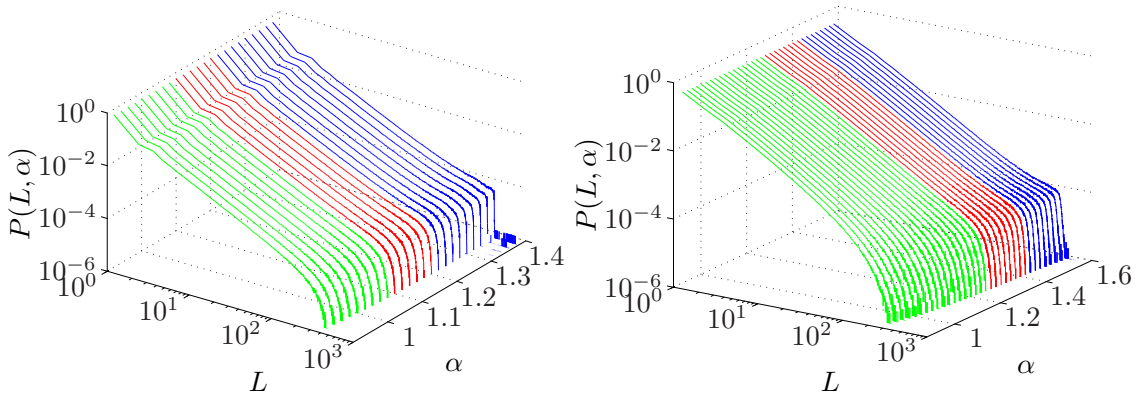


Figure 5.14: Network with 20% inhibitory connection with (left) and without (right) Dale's principle. The avalanche size distribution is plotted for different α. Green lines - subcritical distributions, red lines - critical, and blue lines - supercritical distributions. To determine the critical distribution the threshold for the least-squares deviation was set at 0.005. For "Dale's principle network" avalanches smaller then six are not considered in the estimation of the power-law exponent. $N = 500,\ \nu = 10,\ u = 0.2$.

This means that each neuron is either inhibitory or excitatory. To generate such a network we selected randomly 20% of the neurons and all their synapses were then acting inhibitory, i.e. if inhibitory neuron j is firing, then all post-synaptic neurons i are getting input of $-uJ_{ij}$. In the model a membrane potential is always larger then 0. Therefore if $h_i(t) + I^{\text{int}} < 0$ i take $h(t+1) = 0$.

There are evidences [121] that Dale's principle fails to describe some neurons in the cortex. Thus we consider also another way to add inhibition, which allows some synapses of one neuron to contain inhibitory neurotransmitter and others excitatory. To generate such a network we selected randomly 20% of all synapses which were then acting inhibitory.

In Fig. 5.14 distributions of avalanche size plotted for the networks with and without Dale's principle for different parameters α. In "Dale's network" there is a probability of 0.2 that neuron activated by the external input is inhibitory and then obviously the avalanche will not continue. Hence, as clearly seen in Fig. 5.14, for small avalanche sizes "Dale's network" statistics do not follow a power-law distribution. However, if we concentrate only on avalanches larger than five the effect of Dale's principle is not present any more. Both networks have a power-law exponent of approximately 1.5 in the critical regime and possess a large critical interval. The exponent for Dale's network is a slightly smaller ≈ 1.47, whereas for a simple random inhibitory network exponent is ≈ 1.52. Thus, we conclude that inhibitory population does not change the results obtained for the excitatory network.

Chapter 6

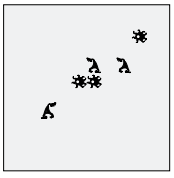
Facilitatory synapses

In the previous chapter we have already seen, how depressing synapses turn the simple network model into self-organized critical. But for some synapses in the brain, presynaptic activity has not only a depressing influence, but also can enhance the synaptic transmission. A model of such activity dependent facilitation and justifying experiments are discussed in the first section.

In this chapter we discuss the influence of facilitation on the network dynamics. We show, that criticality is achieved in the network with facilitatory synapses in a self-organized way. We present a mean-field analysis of the network activity. First results about the network with facilitatory synapses were already published [90].

The heuristic basis of the influence of facilitatory synapses is the following: On the one hand, when avalanches are large and neurons often fire, synaptic depression is strong and leads to the reduced activity and therefore, smaller avalanches. On the other hand, rare firing leads to full recovering of the synapses and that is when the facilitatory part of synaptic activity plays an important role. This brings the network to a certain balanced state, which is characterized by the presence of rare large avalanches in the background activity of the small avalanches.

The introduction of facilitation into the synaptic dynamics changes the transition to criticality. For depressing synapses this transition is continuous, which corresponds to a second order phase transition, whereas the facilitatory dynamics makes the transition by a "discontinuous jump", thus a first order phase transition is observed. Furthermore, the dependence of the mean synaptic strength and the network dynamics in general on the maximal synaptic strength α is not as simple as it is in the case of depressing synapses.



6.1 Model of facilitatory synapses

The model with depressing synapses described in section 5.1 can well account for the synaptic responses between pyramidal neurons in mammalian cortex, which have according to experimental findings no facilitatory component in the dynamics. However it is prominent in the connection between pyramidal neurons and inhibitory inter-neurons [131]. Usually short term facilitation is modeled by introducing a "facilitation factor" which is increased at each spike by a certain amount and decays between spikes [98, 153].

To add facilitation in the model introduced in the section 5.1 one can assume, that the value of U_{SE} is increased by a certain amount due to each presynaptic spike [103, 133]. The dynamics of U_{SE} can be interpreted as calcium accumulation in the presynaptic terminal. Calcium entering a cell through ion channels at spike arrival is responsible for the release of neurotransmitter [16]. The simplified scheme assumes that an action potential causes a fraction of all U calcium channels to open. After the spike channels are closing with a time constant τ_{facil}. Thus the corresponding kinetic equation reads

$$\frac{dU_{SE}}{dt} = -\frac{U_{SE}}{\tau_{facil}} + U(1 - U_{SE})\delta(t - t_{sp}). \tag{6.1}$$

In Fig. 6.1 one example of the response to a periodic spike-train is shown. A response to presynaptic spikes grows at the beginning of the experiment and then reaches a stationary state. The excitatory postsynaptic potential (EPSP) in the stationary regime depends on the frequency of spikes, in the bottom panel the dependence of stationary response on the frequency of presynaptic firing is plotted. The theoretical prediction (solid line) describes the experimental observations (circles) rather well.

6.2 Network model

The network with facilitatory synapses is described by the Eq. 5.2 and Eq. 5.3 together with the additional equation describing the dynamics of the fraction of synaptic resources used for spike. The third equation is analog to Eq. 6.1 and summarizes the facilitatory mechanisms in the synaptic connections.

Consider a set of N integrate-and-fire neurons characterized by a membrane potential $h_i \geq 0$, and two parameters for each synapse: $J_{ij} \geq 0$, $u_{ij} \in [0, 1]$. The parameter J_{ij} characterizes the number of available resources on the presynaptic side of the connection from neuron j to neuron i. Each spike leads to the usage of a portion of the resources of the presynaptic neuron. Hence, at the next synaptic event less neurotransmitter will be available i.e. activity will be depressed. Between

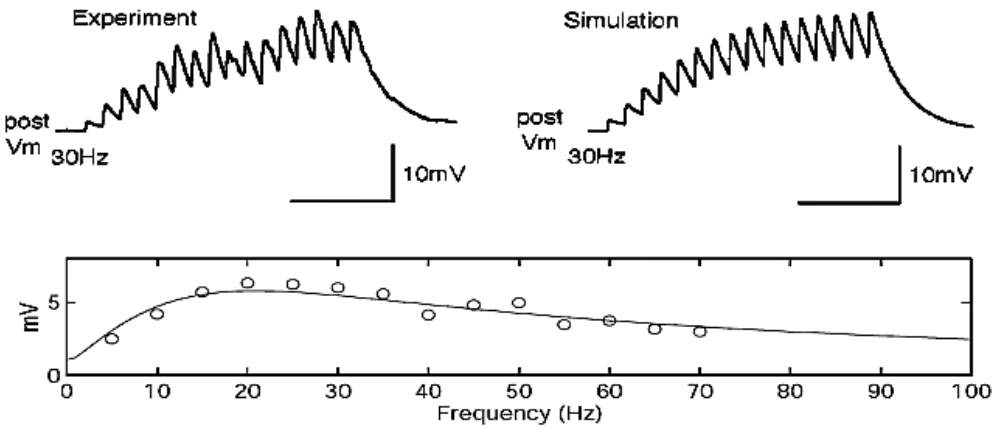


Figure 6.1: Facilitatory synapses. (top left) Average excitatory postsynaptic potential (EPSPs) (4 sweeps) recorded in interneuron (30 Hz), (top right) Simulated synaptic response (30 Hz). The postsynaptic potential is computed by using a passive membrane mechanism. On the bottom: Stationary level of EPSPs as a function of the presynaptic frequency for facilitatory synapses. Open circles – experimental results for one of the recorded synaptic connections between pyramidal neuron and inhibitory interneuron. Solid line - model results. *Pictures are taken from [103, 133].*

spikes resources are slowly recovering on a timescale τ_J.

The parameter u_{ij} denotes the actual fraction of resources on the presynaptic side of the connection from neuron j to neuron i, which will be used in a synaptic transmission. When a spike coming from neuron i arrives at the postsynaptic side j, it causes an increase of u_{ij}. Between spikes, u_{ij} slowly decreases towards the minimal value u_0 on a timescale τ_u.

The combined effect of J_{ij} and u_{ij} results in the facilitation and depression of the synapses. The interplay of the time scales τ_u and τ_J defines the stationary synaptic strength.

The dynamics of a membrane potential h_i consists of the integration of excitatory postsynaptic currents over all synapses of the neuron and the slowly delivered random input. The timescale of the external input τ_s is much larger then the synaptic delay τ_d, which is the time between emission of a spike by the presynaptic neuron and the arrival of the EPSP at the postsynaptic neuron.

When the membrane potential exceeds the threshold θ, the neuron emits a spike and h_i resets by extracting the threshold value, i.e. $h_i(t_{\rm sp}^+) = h_i(t_{\rm sp}) - \theta$. Superthreshold activity is communicated to other neurons along neural connections of strength $u_{ij}(t)J_{ij}(t)$. The dynamics of the membrane potential can be described by

following equation

$$\dot{h}_i = \delta_{i,\zeta(t)} I^{\text{ext}} + \frac{1}{N} \sum_{j=1}^{N} u_{ij}(t_{\text{sp}}) J_{ij}(t_{\text{sp}}) \delta\left(t - t^j_{\text{sp}} - \tau_d\right) \tag{6.2}$$

The variables J_{ij} is governed by the same equation as in the depressing model

$$\dot{J}_{ij} = \frac{1}{\tau_J}\left(\frac{\alpha}{u_0} - J_{ij}\right) - u_{ij} J_{ij} \delta\left(t - t^j_{\text{sp}}\right). \tag{6.3}$$

The variables u_{ij} are subject to the dynamics

$$\dot{u}_{ij} = \frac{1}{\tau_u}\left(u_0 - u_{ij}\right) + (1 - u_{ij}) u_0 \delta\left(t - t^j_{\text{sp}}\right). \tag{6.4}$$

Here u_0 denotes the minimal fraction of the synaptic resources used for a spike and α/u_0 is the maximal amount of the synaptic resources. Synaptic recovery happens on a slow time scale of the external input. The relation between the time constants is the following

$$\tau_J = \nu_J N \tau_s, \tag{6.5}$$
$$\tau_u = \nu_u N \tau_s, \tag{6.6}$$

and $1 < \nu_J,\ \nu_u \ll N$. We will take $\nu_J = \nu_u = \nu$, for any other relation between the time-constants there are appropriate values u_0 and α such that the network is in the critical state.

If the interval between avalanches is large and synapses have enough time to recover to the limit value, then $J_{ij} = \alpha/u_0$, $u_{ij} = u_0$ and therefore synaptic efficacy is equal to α. Short inter-spikes intervals (ISI) change both synaptic parameters: J_{ij} is getting smaller and u_{ij} larger. Without additional consideration it is not possible to predict the cumulative effect of short ISIs. The maximal possible synaptic strength is equal to α/u_0, but this maximum is not realizable, because fully recovered synaptic resources require a long inter-spike interval, which in turn would reduce the fraction of resources used for a spike.

To visualize the dynamics of the synaptic resources we plotted in Fig 6.2 the amount of neurotransmitter in the synapse J_{ij}, the fraction of it which will be used for a spike u_{ij} and synaptic efficacy $J_{ij}u_{ij}$ for different firing frequencies. Note, that "jumps" in the graph of J_{ij}, u_{ij}, and $u_{ij}J_{ij}$ are happening after the spike.

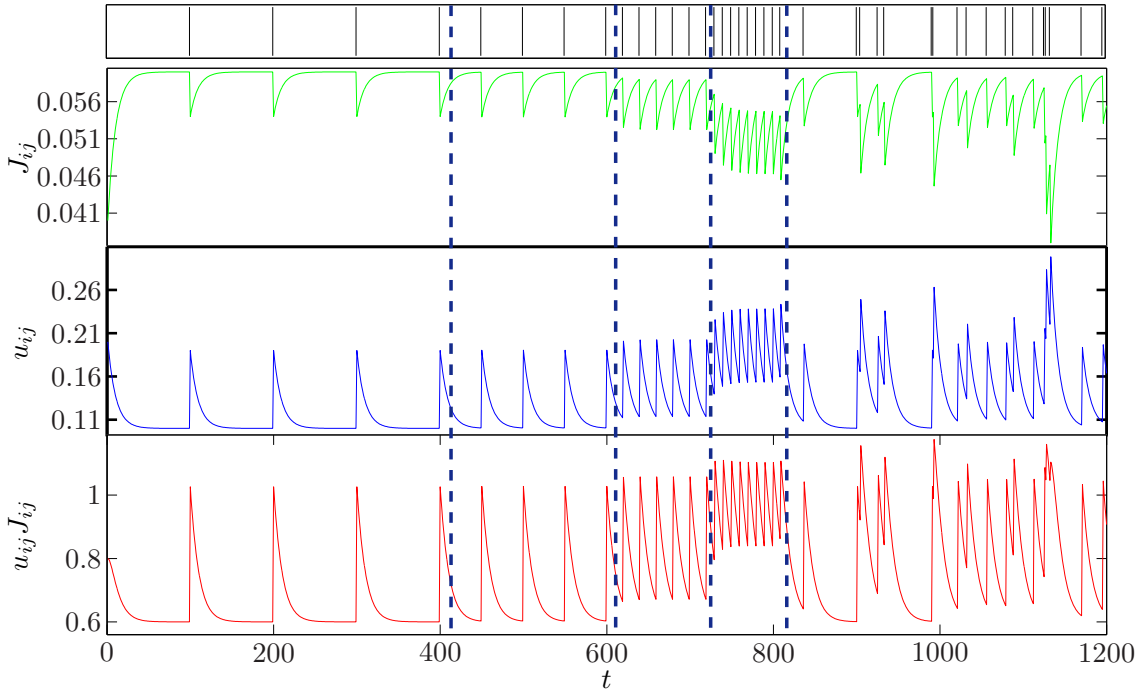


Figure 6.2: Amount of synaptic resources. On the top panel the spike train of the neuron is shown (black), on the next panel the amount of the synaptic resources J_{ij} is plotted in blue, on the third panel the fraction of the synaptic resources which will be used for a spike u_{ij} is plotted in green, on the bottom panel the synaptic efficacy is plotted in red. Dashed vertical lines separate different firing frequencies. From left to right: one spike in 100, 50, 20, 10 time-steps. The last interval corresponds to a Poisson spike-train with the intensity of 1 spike in 30 time-steps. $N = 300$, $\alpha = 0.6$, $\nu = 10$, $u_0 = 0.1$.

6.3 Simulation results

First, let us have a look on the distribution of the avalanche size for different values of the parameter α. As it is seen in Fig. 6.3 a sharp transition in the distribution happens at the value $\alpha \approx 0.54$ from the clearly subcritical state into the vicinity of the critical region. To characterize the obtained distributions in more details the parameters of the fitted power law are displayed, Fig. 6.4. Note that for a network of 500 units, the absolute critical exponent is smaller than in the large system limit where $\gamma = -1.5$ and that the step size has been drastically reduced in the vicinity of the phase transition to exclude possibility of the numerical artifact. Clearly, an error level above 0.1 indicates that the fitted curve is far from the original distribution, which is therefore not the critical distribution.

To characterize the width of the critical region in Fig. 6.5 the mean square deviation of the avalanche sizes distribution from the power law for different α and N is plotted. For comparison the results obtained from the simulation of the static network are also plotted. The value of the threshold is rather arbitrary, but as it is shown in Appendix A.2 the distributions which have a deviation from the power law of less than 0.005 are statistically a power-law. The length of the critical interval for the parameter α is plotted in the inset. It is three times smaller than in the depressing case, but still 10 times larger then for static synapses. It is also important to remark, that the critical interval is increasing with the system size, therefore in the limit the system might get critical for a substantial part of the parameters space.

Additionally, the distributions of the synaptic parameters u_{ij} and J_{ij} are shown in Fig. 6.6. Both have bell-shaped distributions with a variance much smaller than a mean values. This is important for the next section where we describe the mean-field approximation. Large values of α lead to a more often firing of each neuron in the network. When neuron fires more often the value of u_{ij} increases, whereas the value of J_{ij} decreases. This explains why the mean of u_{ij} is smaller and the mean of J_{ij} is larger for large values of α and vice versa.

6.4 Mean-field analysis

In this section we present the mean-field analysis of the network model with facilitatory synapses. Following the lines of analysis in section 5.5 we start by deriving a set of equations describing the dependency between the average synaptic strength and the inter-spike interval. After solving the self-consistency equation we will study the bifurcation arising from the introduction of the facilitation in the model.

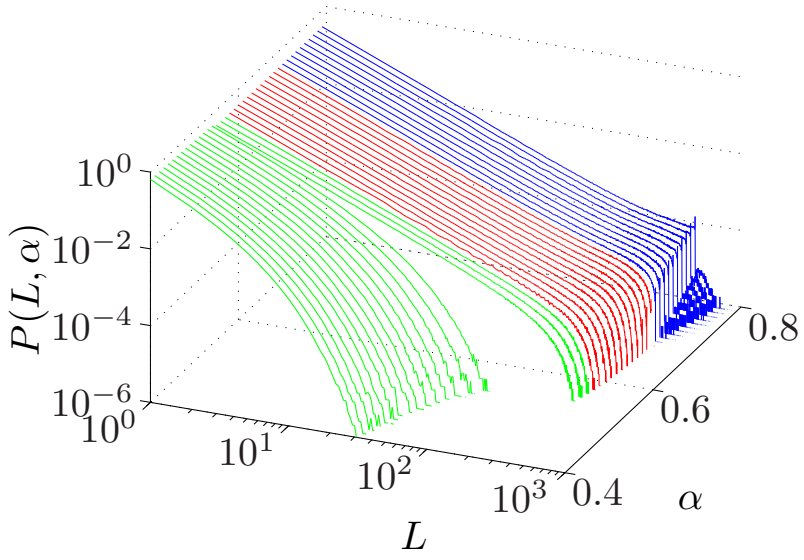


Figure 6.3: Distribution of avalanche sizes for different values of α. At $\alpha < 0.54$ the distribution is subcritical (plotted green) . It becomes critical in an interval around $\alpha = 0.65$ (plotted red). For $\alpha > 0.69$ the distribution is supercritical (plotted blue). $N = 500$, $\nu = 10$, $u_0 = 0.1$.

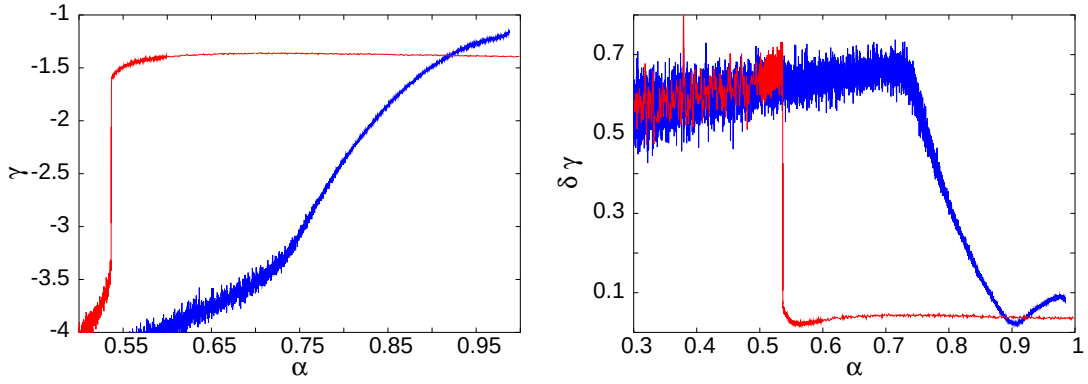


Figure 6.4: The best matching power-law exponent (left) and the mean squared deviation from the best matching power-law (right). The red line represents the present model, while the blue line stands for the model with static synapses. Average synaptic efficacy α ranges from 0.3 to 1.0 with a step size of 0.001. For the fit, avalanches of a size larger than 1 and smaller than $N/2$ have been used. Presented curves are temporal averages over 10^7 avalanches with $N = 200$, $u_0 = 0.1$, $\nu = 10$. *Figures are taken from [90].*

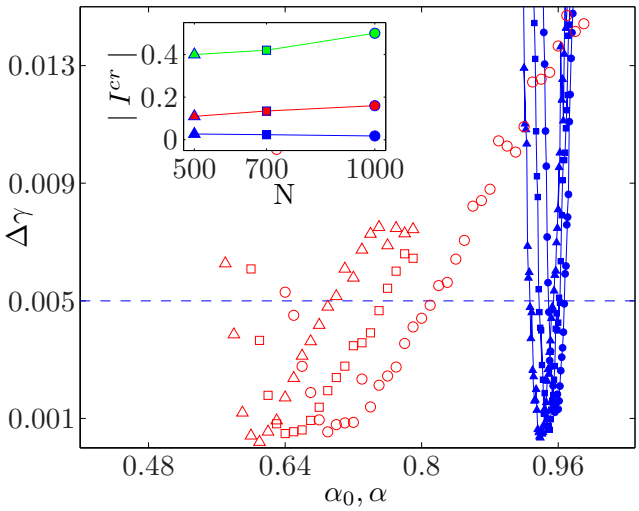


Figure 6.5: The range of connectivity parameters where critical dynamics extends to the system size. The mean-squared deviation from the best-matching power-law is plotted in dependence of α_0 for static synapses and α for facilitatory synapses: red circles, squares and triangles stand for networks with facilitatory synapses and system sizes $N = 1000$, 700, and 500, respectively. The blue symbols represent the static model. The inset shows the lengths of the parameters intervals where the deviation from the best-matching power-law is smaller than 0.005, in green the results from the depressing synapses model are shown. Symbols in the inset are the same as in the main figure.

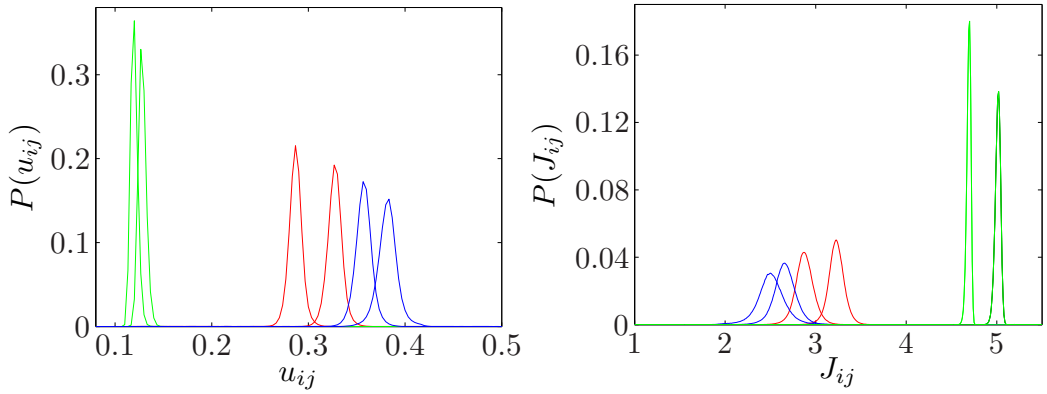


Figure 6.6: The distribution of the synaptic resources J_{ij} and their usage u_{ij} for different α. Green: subcritical state $\alpha = 0.4$, 0.53, red: critical state $\alpha = 0.54$, 0.6, blue: super-critical state $\alpha = 0.7$, 0.75. $N = 300$, $u_0 = 0.1$, $\nu = 10$.

α	$\langle u_{ij}\rangle\langle J_{ij}\rangle$	$\langle u_{ij}J_{ij}\rangle$	$\mid \langle u_{ij}\rangle\langle J_{ij}\rangle$ - $\langle u_{ij}J_{ij}\rangle \mid$
0.4	0.431	0.436	0.005
0.55	0.9045	0.911	0.006
0.8	0.957	0.960	0.003

Table 6.1: Difference between the true averaged synaptic strength and the synaptic strength obtained from the independence hypothesis. $N = 300$, $u_0 = 0.1$, $\nu = 10$.

We consider again the averages of the following dynamical quantities: the amount of synaptic resources J_{ij}, the fraction of resources used for a spike u_{ij}, and the inter-spike interval Δ^{isi}. The average synaptic strength is equal to $\langle u_{ij}J_{ij}\rangle$ which is in general not the same as $\langle u_{ij}\rangle\langle J_{ij}\rangle$. As it is shown in the Table 6.1 the differences between the true value of the average synaptic efficacy $\langle u_{ij}J_{ij}\rangle$ and the approximate value $\langle u_{ij}\rangle\langle J_{ij}\rangle$ is comparably small, though it is obvious that u_{ij} and J_{ij} are not independent because they both depend on the size of the preceding inter-spike interval. We use approximation instead of the exact expression for the the average synaptic strength because it allows us to do the analytical computations.

We continue according to the following plan: First we derive a couple of equations connecting $\langle\Delta^{\text{isi}}\rangle$ and $\langle u_{ij}\rangle\langle J_{ij}\rangle$:

$$\langle u_{ij}J_{ij}\rangle \approx \langle u_{ij}\rangle\langle J_{ij}\rangle = G\left(\langle\Delta^{\text{isi}}\rangle\right), \tag{6.7}$$

$$\langle\Delta^{\text{isi}}\rangle = F\left(\langle u_{ij}J_{ij}\rangle\right). \tag{6.8}$$

Given the system of the equations (6.7,6.8) it is easy to construct a self-consistency equation. Which we later solve numerically, obtaining the stationary values of $\langle u_{ij}J_{ij}\rangle$ and $\langle\Delta^{\text{isi}}\rangle$.

To start with, let us recall the equations governing the synaptic dynamics of the network

$$\dot{u}_{ij} = \frac{1}{\tau_u}(u_0 - u_{ij}) + (1 - u_{ij})u_0\delta\left(t - t^j_{\text{sp}}\right), \tag{6.9}$$

$$\dot{J}_{ij} = \frac{1}{\tau_J}\left(\frac{\alpha}{u_0} - J_{ij}\right) - u_{ij}J_{ij}\delta\left(t - t^j_{\text{sp}}\right). \tag{6.10}$$

The behavior of the variable u_{ij} does not depend on the value of J_{ij}, therefore first we derive the equation for the dependency of u_{ij} on Δ^{isi} and then use it to derive the equation for J_{ij} analogously.

Between two spikes of the neuron j only the relaxation dynamics affects the variable

u_{ij}, thus it is easy to solve the equation (6.9):

$$u_{ij}^{b,2} = u_0 + \left(u_{ij}^{a,1} - u_0\right) e^{-\left(t_2^j - t_1^j\right)/\tau_J} \tag{6.11}$$

where $u_{ij}^{a,1}$ and $u_{ij}^{b,2}$ are the synaptic strengths after a spike of neuron j at time t_1^j and before the subsequent spike at t_2^j respectively. By the definition of time-constants (6.5,6.6)

$$\frac{t_2^j - t_1^j}{\tau_u} = \frac{\Delta^{\text{isi}}}{\nu N}. \tag{6.12}$$

Within a short interval containing the spike, u_{ij} increases by $(1 - u_{ij}^{b,2})u_0$ such that $u_{ij}^{a_1} = u_{ij}^{b,1} + u_0\left(1 - u_{ij}^{b,1}\right)$. The Eq. 6.12 can be rewritten as

$$u_{ij}^{b,2} = u_0 + \left(u_{ij}^{b,1} - u_0 u_{ij}^{b,1}\right) e^{-\frac{\Delta^{\text{isi}}}{\nu N}}. \tag{6.13}$$

At stationarity $\left\langle u_{ij}^{b,1}\right\rangle = \left\langle u_{ij}^{b,2}\right\rangle$ therefore by substituting the inter-spike interval in the Eq. 6.13 by the average $\langle\Delta^{\text{isi}}\rangle$ one can find the average synaptic strength by solving the equation

$$\langle u_{ij}\rangle = u_0 + \langle u_{ij}\rangle(1 - u_0)e^{-\frac{\langle\Delta^{\text{isi}}\rangle}{\nu N}}. \tag{6.14}$$

This allows to express $\langle u_{ij}\rangle$ in terms of $\left\langle\Delta^{\text{isi}}\right\rangle$,

$$\langle u_{ij}\rangle = \frac{u_0}{1 - (1 - u_0)e^{-\frac{\langle\Delta^{\text{isi}}\rangle}{\nu N}}} = G_1\left(\langle\Delta^{\text{isi}}\rangle\right). \tag{6.15}$$

Now we can analogously write the expression for $\langle J_{ij}\rangle$ as

$$\langle J_{ij}\rangle = \frac{\alpha}{u_0}\frac{1 - e^{-\frac{1}{\nu N}\langle\Delta^{\text{isi}}\rangle}}{1 - e^{-\frac{1}{\nu N}\langle\Delta^{\text{isi}}\rangle} + \langle u_{ij}\rangle e^{-\frac{1}{\nu N}\langle\Delta^{\text{isi}}\rangle}} = G_2\left(\langle\Delta^{\text{isi}}\rangle\right). \tag{6.16}$$

Combining Eq. 6.16 and Eq. 6.15 we obtain the first part of the self-consistency equation

$$\langle u_{ij}J_{ij}\rangle \approx \langle u_{ij}\rangle\langle J_{ij}\rangle = G_1\left(\langle\Delta^{\text{isi}}\rangle\right) G_2\left(\langle\Delta^{\text{isi}}\rangle\right) = G\left(\langle\Delta^{\text{isi}}\rangle\right) \tag{6.17}$$

The dependence of the inter-spike interval on the average synaptic strength is not affected by the more complicated synaptic dynamics. Therefore, the second part of the self-consistency equation is the same as in the depressing model (section 5.5)

$$\left\langle\Delta^{\text{isi}}\right\rangle = \frac{\theta - \epsilon_N}{I^{\text{ext}}}(N - \langle u_{ij}J_{ij}\rangle(N-1)) = F\left(\langle u_{ij}J_{ij}\rangle\right). \tag{6.18}$$

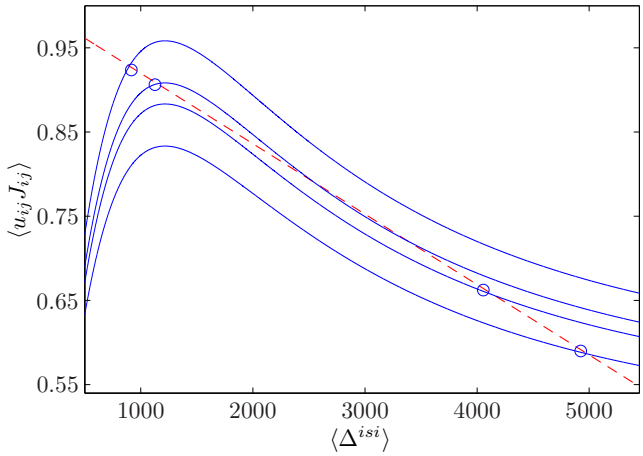


Figure 6.7: Average synaptic strength and inter-spike interval for different initial conditions. The dashed line corresponds to Eq. 6.18 and the solid lines are obtained from Eq. 6.17 for $\alpha = 0.5, 0.535, 0.55, 0.7$. Circles show $\langle u_{ij}J_{ij}\rangle$ and $\langle\Delta^{\text{isi}}\rangle$ obtained from simulations of the network with matching α. Further parameters are $N = 300$, $\nu = 10,\ u_0 = 0.1$.

The combination of the Eq. 6.17 and the inverse of the Eq. 6.18 gives rise to the self-consistency equation

$$\frac{N}{N-1} - \frac{I^{\text{ext}}\langle\Delta^{\text{isi}}\rangle}{(\theta-\epsilon_N)(N-1)} = G\left(\langle\Delta^{\text{isi}}\rangle\right). \tag{6.19}$$

The numerical solution of Eq. 6.19 allows to find the steady state of the system for any given parameter α. In Fig. 6.7 the graphical solution of the self-consistency equation is depicted together with the simulation results. Predicted by the mean-field approximation values of the inter-spike intervals and the synaptic strength are the intersections between the curves for the correspondent α and the line, representing the inverse of the Eq. 6.18 which does not depend on α. The circles are simulation results for the correspondent α. Simulated quantities are in the good agreement with theoretical averages, which allows to justify approximations used to obtain self-consistency equation.

The curves with more than one intersection points with the line are of the particular interest, they are studied in the next section.

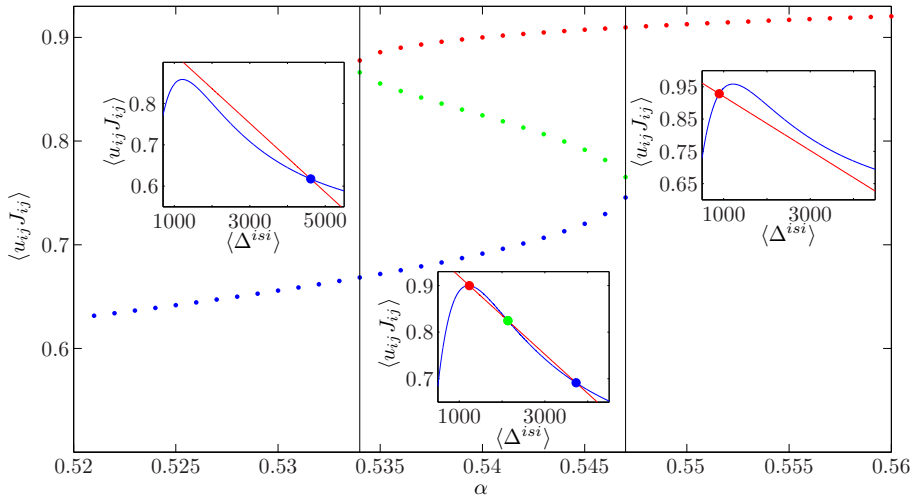


Figure 6.8: Bifurcation diagram. Solution of the self-consistency equation $\langle u_{NJ}J_{NJ}\rangle$ plotted versus parameter α. In the inset red lines are the graphs of Eq. 6.18 and the blue lines are obtained from Eq. 6.17 for a value of α from the correspondent parameter interval. $N = 300,\ \nu = 10,\ u_0 = 0.1$.

6.4.1 Bifurcation of the steady state parameters

In this section we describe the mechanism, that brings about the discontinuity in the distributions of avalanche sizes. As we have seen in Fig. 6.7, for some values of α there are three intersection points of the graph of function F (6.17) and inverse of G (6.18). For such values of α there are two possible states of the network. The third state is unstable and therefore not realizable in the network where the stochastic dynamics plays a role of a strong noise which takes the system immediately away from the unstable fixed point.

Depending on the parameter α there are different regimes in the network dynamics:

1. If $\alpha < \alpha_c$ there is only one solution of the self-consistency equation.
2. If $\alpha = \alpha_c$ then there is the saddle-node bifurcation point, which is characterized by the appearance of one stable and one unstable branches. The new stable branch is not a continues prolongation of the previously existing branch.
3. If $\alpha^c > \alpha > \alpha_c$ there are three solutions of the self-consistency equation, two of them are stable. The network "selects" the branch depending on the initial conditions. For the simulations we always took the same initial conditions, which created the sharp transition between two states of the network.
4. If $\alpha = \alpha^c$ there is a second bifurcation point. The initial stable branch and the unstable branch disappear.

5. If $\alpha > \alpha^c$ there is the unique solution of the self-consistency equation.

All regimes are shown in Fig. 6.8. For the network of $N = 300$ neurons with $u_0 = 0.1$. The bifurcation points are at $\alpha_c = 0.534$, $\alpha^c = 0.547$.

6.4.2 Stability analysis

The bifurcation diagram in the previous section makes the stability claims, which have to be checked. We use simple perturbation analysis to study the stability of the fixed points. On the first sight perturbation may influence synaptic strength (for instance, by changing the maximal amount of the synaptic resources), or inter-spike interval (by changing the strength of the external input) or both. But on the second glance, perturbation of the ISI is immediately translated in the changes of the synaptic strength by Eq. 6.16, 6.15. Thus we consider only perturbations in the synaptic strength.

In the most interesting case 3 there are three fixed points, two of which are stable. Consider, that parameter α satisfy $\alpha^c > \alpha > \alpha_c$ and at some time t_p a perturbation ΔJ is applied to all synapses such that for each i, j: $\widetilde{J_{ij}} = J_{ij} + \Delta J$. The effect of such perturbation can depend on the current state of the system and the value of ΔJ. Perturbation of the parameter u_{ij} leads to a qualitatively similar results.

In Fig. 6.9 one of the possible outcomes is shown. Before the perturbation network is in the "active" state (synaptic strength is large), negative perturbation compel the network to switch to the "resting" state. The individual dynamics of the synaptic parameters plotted in Fig. 6.10 (left). If $\Delta J < 0$ a perturbation results in the temporary decrease of the avalanche size. This lead to the increase in the synaptic resources J_{ij} but decrease of u_{ij}, such that there product in total decrease. The changes of the synaptic strength and the inter-spikes interval during the transient period are shown in Fig. 6.10 (right).

In Fig. 6.11 the probability of switch is plotted as a function of ΔJ. For this graph 25 realization of the network in the active state was perturbed with the same $\Delta J < 0$. After the network returned to the stationary state the average synaptic strength was measured. As there is a stochastic influence on the network dynamics, the result of the perturbation can not be described deterministically, thus for each ΔJ the probability that network converges to "active" or "resting" state is plotted. If $|\Delta J|$ is large enough (in Fig 6.11 if $\Delta J < -0.21$), then network switches without fail, if $|\Delta J|$ is small then steady state is preserved after the perturbation. Between these extreme values of ΔJ, both outcomes are possible.

As unstable fixed point closer to the "active" state, perturbation which is required to switch from the "active" state is smaller, then one needed to leave the "resting"

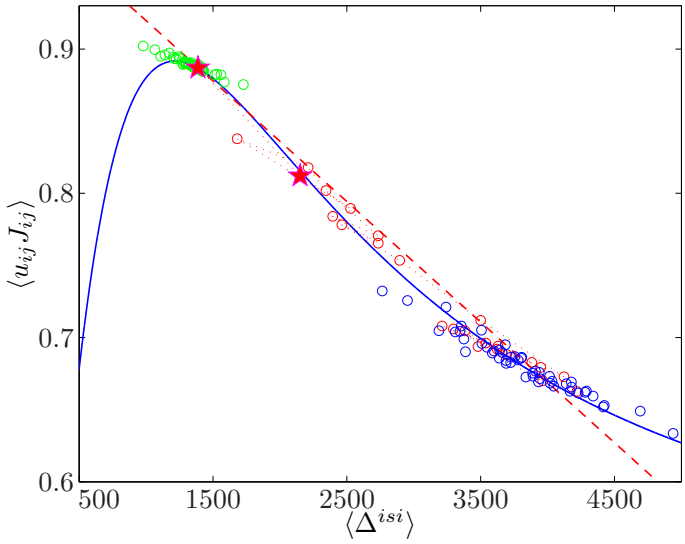


Figure 6.9: Switch of the stable fixed point by the strong perturbation. Each circle depicts strength of one synapse and ISI for it. Initial state of the synapse plotted in green. At the point marked with star the perturbation is applied, resulting in another state also marked with star. Red circles denote a transient period, until the network stabilizes. The new stable state is depicted by blue circles. Dotted line shows the course of the transient. Dashed red and solid blue lines are same as in Fig. 6.7 $\alpha = 0.535$, $\Delta J = -0.5$, $N = 300$, $\nu = 10$, $u_0 = 0.1$.

state Fig. 6.11. If parameter α increases, unstable fixed point moves towards the "resting" state, resulting in the decrease of stability.

6.4.3 Thermodynamical limit

In Fig. 6.3 we have seen, that the critical region grows with the system size at least for relatively small N. Such tendency is similar to what is observed in the depressing model and it makes us believe that the tendency will preserve and in the thermodynamical limit the critical region will be infinite. In this section we check such conjecture. The derivation mainly follows the line of the analogous proof for the depressing synapses from section 5.5.1.

First we recall equations which compose the self-consistency equation:

$$\begin{aligned} \langle u_{ij} \rangle &= \frac{u_0}{1 - (1 - u_0) e^{-\frac{\langle \Delta^{\mathrm{isi}} \rangle}{\nu N}}}, \\ \langle J_{ij} \rangle &= \frac{\alpha}{u_0} \frac{1 - e^{-\frac{1}{\nu N} \langle \Delta^{\mathrm{isi}} \rangle}}{1 - e^{-\frac{1}{\nu N} \langle \Delta^{\mathrm{isi}} \rangle} + \langle u_{ij} \rangle e^{-\frac{1}{\nu N} \langle \Delta^{\mathrm{isi}} \rangle}}, \end{aligned}$$

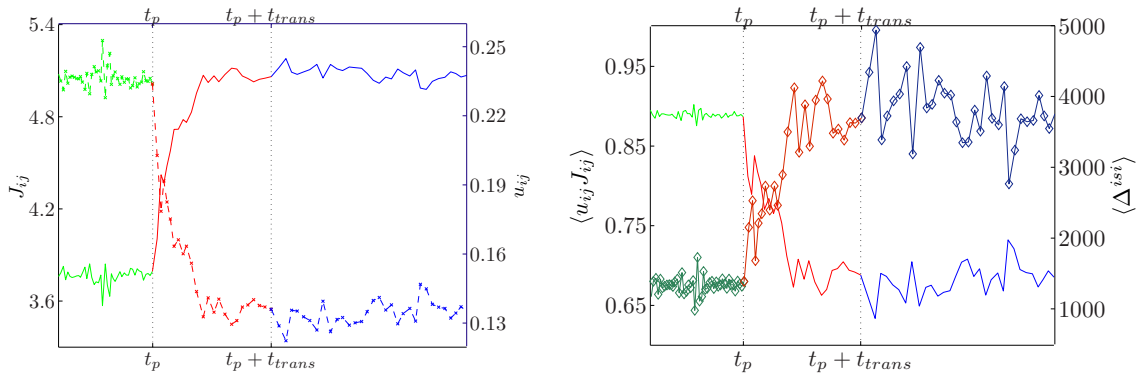


Figure 6.10: Evolution of the synaptic parameters and ISI after the perturbation. (left) Dashed graph with crosses corresponds to the right axes, solid-line graph plotted on the left axes. (right) Graph with diamants plotted on the right axes, graph without symbols on the left axes. Color code corresponds to Fig. 6.9: green part depicts initial stationary state, red part - transient period, and blue part new stationary state. $\alpha = 0.535$ $N = 300$, $\nu = 10$, $u_0 = 0.1$.

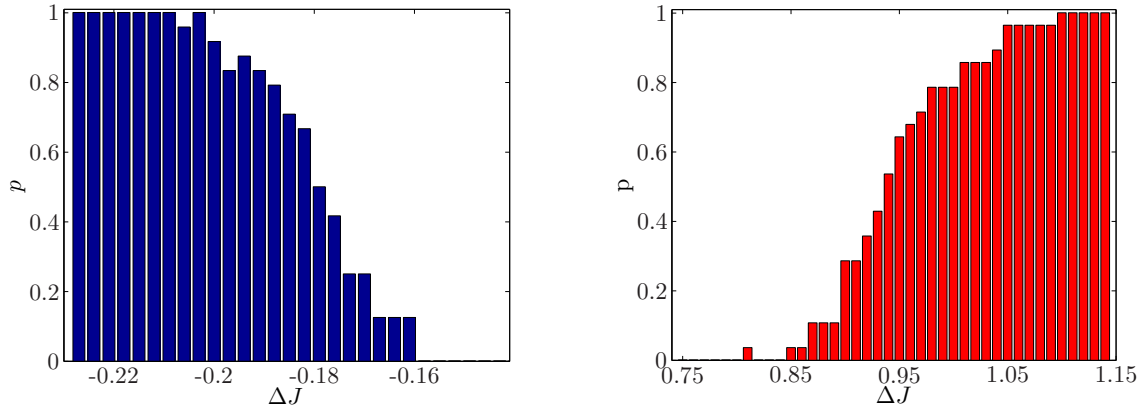


Figure 6.11: Outcome of perturbation. Probability of switching to another state after the perturbation as a function of ΔJ. (left) if the network was initially in the "active" state, (right) if the network was initially in the "resting" state. Histograms are computed based on 25 trials. $\alpha = 0.535$, $N = 300$, $\nu = 10$, $u_0 = 0.1$.

$$\left\langle \Delta^{\text{isi}} \right\rangle = \frac{\theta - \epsilon_N}{I^{\text{ext}}}(N - \langle u_{ij} J_{ij} \rangle (N-1)).$$

An explicit form of the self-consistency equation reads

$$\frac{\alpha \langle u_{ij} \rangle}{u_0} \frac{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}}{1 - e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle} + \langle u_{ij} \rangle e^{-\frac{1}{\nu N}\langle \Delta^{\text{isi}} \rangle}} = \frac{N}{N-1} - \frac{I^{\text{ext}} \langle \Delta^{\text{isi}} \rangle}{(\theta - \epsilon_N)(N-1)}. \tag{6.20}$$

In the limit $N \to \infty$ one should scale the external input by $I^{\text{ext}} \sim N^{-w}$ with $w > 0$. We will not explicitly solve the self-consistency equation, but consider the approximation of the solution. As the system size growth, the average inter-spike interval will scale with N. Let us now distinguish the following cases:

1. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1+\epsilon}$ and $\epsilon > w$,
   then the right hand side (r.h.s.) of Eq. 6.20 approaches $-\infty$, while $\langle u_{ij} \rangle$ approaches u_0 and the l.h.s. is larger than 0.

2. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1+\epsilon}$ and $\epsilon < w$,
   then the right hand side (r.h.s.) of Eq. 6.20 approaches 1 and the l.h.s tends to α. Therefore solution exists only for $\alpha = 1$.

3. If $\langle \Delta^{\text{isi}} \rangle \sim N^{1-\epsilon}$ and $\epsilon > 0$,
   then the r.h.s. of Eq. 6.20 tends to 1, whereas $\langle u_{ij} \rangle$ approaches 1 and $\langle J_{ij} \rangle$ tends to 0, therefore l.h.s. approaches 0.

4. If $\langle \Delta^{\text{isi}} \rangle \sim N$,
   we can assume that $\langle \Delta^{\text{isi}} \rangle = cN + o(N)$. Then $\langle u_{ij} \rangle$ is approaching

   $$\langle u_{ij} \rangle_{\text{lim}} = \frac{u_0}{1 - (1-u_0)e^{-\frac{c}{\nu}}}, \tag{6.21}$$

   whereas $\langle J_{ij} \rangle$ is tending to

   $$\langle J_{ij} \rangle_{\text{lim}} = \frac{\alpha}{u_0} \frac{1 - e^{-\frac{c}{\nu}}}{1 - e^{-\frac{c}{\nu}} + \langle u_{ij} \rangle_{\text{lim}} e^{-\frac{c}{\nu}}}. \tag{6.22}$$

   r.h.s. of (6.20) tends to 1, thus c is a solution of the equation

   $$\langle u_{ij} \rangle_{\text{lim}} \frac{\alpha}{u_0} \frac{1 - e^{-\frac{c}{\nu}}}{1 - e^{-\frac{c}{\nu}} + \langle u_{ij} \rangle_{\text{lim}} e^{-\frac{c}{\nu}}} = 1. \tag{6.23}$$

   Lets $z = e^{-\frac{c}{\nu}}$, Eq. 6.23 can be solved with respect to z

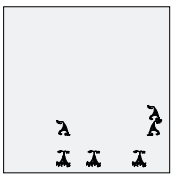
   $$z_{\pm} = \frac{2 - u_0 - \alpha \pm \sqrt{-4u_0 + 4u_0^2 + \alpha^2}}{2(1 - u_0}. \tag{6.24}$$

As $u_0 < 1$, if α is sufficiently large, namely

$$\alpha > 2\sqrt{u_0 - u_0^2} \tag{6.25}$$

then $z \in \mathbb{R}$. And z_+ is positive therefore, there exist a solution of Eq. 6.23.

Thus we showed that in all the cases, where solution exists (ii, iv), $\langle u_{ij} J_{ij} \rangle = 1$ and therefore, network is in the critical state. We also proved that solution exists for all large enough values of α.



Chapter 7

Dynamical systems formalism

Despite the constant attention from physicists to the mathematical theory of self-organized criticality it is still in infants. One rare example of the research on strict mathematical basis of SOC is the works of Krueger, Cessac, and Blanchard [19, 34]. Their advances in the research of the Zang model are described in section 2.2.2.

In the present chapter we propose the mathematical model of the neural network, which was introduced in chapter 3. The network is characterized by the local interactions, which lead to events widely extended in size and duration. To describe such a phenomena the dynamical systems formalism is used. It is a natural approach to try to access the macroscopic behavior of large systems from the microscopic dynamical evolution. The macroscopic behavior at stationarity is characterized by a probability measure, which one has to extract from the microscopic evolution.

The structure of this chapter will be as follows: first we introduce the model and discuss its simple properties, then we find the invariant measure and in the end we prove that the system posses the property of the topological transitivity. Where it is possible we try to give also the physical interpretation alongside with definitions and theorems.

7.1 Definitions

Let $X = [0, \infty]^N$, $X_s = \{\mathbf{x} \in X | x_i < 1, i = 1, \ldots, N\}$. Let $\mathcal{A}$ be a measurable space of sequences $a = \{a_1, a_2, a_3, \ldots, \}$, $a_i \in \{1, \ldots, N\}$ equipped with a probability measure ν. Let σ be a left shift, $\sigma : \mathcal{A} \to \mathcal{A}$, $a = \{a_1, a_2, a_3, \ldots, \} \in \mathcal{A}$, then $\sigma(a) = \{a_2, a_3, \ldots\}$.

Define $G : \{1, \ldots, N\} \to \mathcal{X}$, where $\mathcal{X}$ is a set of all hyperbolic maps $X_s \to X$. To mimic a neural network behavior we will consider the particular case where G is

defined as

$$G_k(\mathbf{x}) = \mathbf{x} + e_k\delta,$$

for $1 \leq k \leq N$, where e_k is a k-s basis vector in $\mathbb{R}^N$ and $0 < \delta < 1$ is a parameter of the model.

Let $Z : X \to \{0,1\}^N$, $\mathbf{x} \mapsto \mathbf{z}$, $Z_i(\mathbf{x}) = z_i = I_{[1,\infty)}(x_i)$, where $I_{[a,b]}$ is an indicator function of the interval $[a,b]$.

$$F(\mathbf{x}) = \mathbf{x} + \frac{\alpha}{N} e \operatorname{card}\{i | z_i = 1\} - \mathbf{z},$$

where $e = \sum_{k=1}^{N} e_k$, $\alpha \in (0,1)$ is a parameter of the model, referred in physics as a connection strength. If $\mathbf{y} = F(\mathbf{x})$ then

$$y_i = x_i + \frac{\alpha}{N} \operatorname{card}\{l | z_l = 1\} - z_i. \tag{7.1}$$

The physical model that allowes for the interpretation of the described variables and functions is given in section 3. Namely, the parameter N denote the number of neurons in the network, the coordinate x_i of the variable $\mathbf{x}$ denote the membrane potential of the neuron i. Thus the space X_s is a space of all stable configurations (i.e., configurations when there is no neuron to fire). Members of the sequence $a \in \mathcal{A}$ are indices of neurons to get the external input, and the sequence in total defines the order of external inputs during the life of the network. The mapping G_l brings the external input to the neuron l. The mapping F describes one step of the avalanche: reset of all super-threshold neurons and delivery of the internal input to all neurons. Parameter δ denotes the strength of the external input to the network and α/N denotes the strength of connections between neurons.

Assumption 1. We will assume that the external input $\delta \in \mathbb{R} \setminus \mathbb{Q}$ is smaller than the synaptic input $\alpha/N \in \mathbb{R} \setminus \mathbb{Q}$, and for any $a, b \in \mathbb{R} \setminus \mathbb{Q}$, $a\delta + b\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

The assumptions of irrationality and rational independence will be important for the proof of topological transitivity later in this chapter. Keeping in mind the physical interpretation of the model it is natural to assume that the parameters α and δ are irrational with probability one, because in the real world they are rather arbitrary and have no specific constrains to be rational. The last assumption that $\delta < \alpha/N$ is done to shorten the proofs. It is not changing anything in the logic of them, but allows to consider a smaller number of similar cases. From the physical point of view this means that we consider the strongly connected network with the weak external input.

We define a measurable map $T : \mathcal{A} \times X_s \to \mathcal{A} \times X_s$ as follows

$$T(a, \mathbf{x}) = \left(\sigma a, F^{k(a,\mathbf{x})} \circ G_{a_1}(\mathbf{x})\right), \tag{7.2}$$

where $k(a, \mathbf{x}) \geq 0$ is a minimal integer such as $F^{k(a,\mathbf{x})} \circ G_{a_1}(\mathbf{x}) \in X_s$.

Remark 7.1.1. *For any $\boldsymbol{x} \in X_s$ and $a \in \mathcal{A}$ there exists $k(a, \boldsymbol{x}) < \infty$ such that, $T(a, \boldsymbol{x}) = (\sigma a, F^{k(a,\boldsymbol{x})} \circ G_{a_1}\boldsymbol{x})$.*

Proof. Denote $F(\mathbf{x})_i$ the i^{th} coordinate of the vector $F(\mathbf{x})$. Then for any $\mathbf{x} \in X_s$ we have

$$\sum_{i=1}^{N} F(\mathbf{x})_i = \sum_{i=1}^{N} x_i + \frac{\alpha}{N} N \sum_{i=1}^{N} z_i - \sum_{i=1}^{N} z_i \leq \sum_{i=1}^{N} x_i - 1 + \alpha$$

and $F(\mathbf{x})_i > 0$ for any $i = 1, \ldots, N$. Therefore sum of all coordinates of $\mathbf{x}$ is decreasing by at least $1 - \alpha > 0$ with each iteration of F, but all coordinates stay larger than zero. Define k_M as

$$k_M = \min \left\{ n \in \mathbb{N} \middle| n > \frac{\sum_{i=1}^{N} x_i}{1 - \alpha} \right\}.$$

If $k \geq k_M$ then $\sum_i F^{k_M}(x)_i < 1$ and hence for all $i \geq N$, $0 < x_i < 1$ therefore $F^k(\mathbf{x}) \in X_s$. □

The map T summarizes the full dynamics of the network. If $k(a, \mathbf{x})$ is larger than zero then there is an avalanche happening in the network. The $k(a, \mathbf{x})$ represent the duration of the avalanche. The remark above proves that in dissipative system (when $\alpha < 1$) all avalanches have a finite duration. Now we will introduce the avalanche size, which is described as the number of the coordinates of vector $\mathbf{x}$ which were larger than one during the avalanche time.

Definition 7.1.1. Let $\mathbf{x} \in X$ and $a \in \mathcal{A}$ be fixed, $T = F^{k(a,\mathbf{x})} \circ G_a$ defined as before and $k(a, \mathbf{x}) > 0$, then avalanche size $\xi = \xi(a, \mathbf{x})$ is defined as

$$\xi(a, \mathbf{x}) = \sum_{j=0}^{k(a,\mathbf{x})} \sum_{i=1}^{N} Z_i \left(F^j \circ G_a \mathbf{x} \right). \tag{7.3}$$

Define projections $\pi_X : \mathcal{A} \times X_s \to X_s$, $\pi_X(a, \mathbf{x}) = \mathbf{x}$ and $\pi_{\mathcal{A}} : \mathcal{A} \times X_s \to \mathcal{A}$, $\pi_{\mathcal{A}}(a, \mathbf{x}) = a$.

Proposition 7.1.2. *Let $x \in X_s$, $a \in \mathcal{A}$ and let $\alpha + \delta < 1$. Then for any $i \leq N$,*

$$\sum_{j=1}^{k(a,\boldsymbol{x})} Z_i(F^j \circ G_{a_1}\boldsymbol{x}) \leq 1.$$

Proof. Let us assume that the statement of the proposition does not hold, then the set

$$I_c = \left\{ i \leq N \,\middle|\, \sum_{j=1}^{k(a,\mathbf{x})} Z_i(F^j \circ G_{a_1}\mathbf{x}) \geq 2 \right\} \neq \emptyset.$$

Define for each $i \in I_c$, $n_i \leq k(a,\mathbf{x})$ such that $\sum_{j=1}^{n_i-1} Z_i(F^j \circ G_{a_1}\mathbf{x}) \leq 1$ but $\sum_{j=1}^{n_i} Z_i(F^j \circ G_{a_1}\mathbf{x}) \geq 2$. Define $n \leq k(a,\mathbf{x}) = \min_{i\in I_c} n_i$ and $m = \operatorname{argmin}_{i\in I_c} n_i$. Then for any $l \leq N$,

$$\sum_{j=1}^{n-1} Z_l(F^j \circ G_{a_1}\mathbf{x}) \leq 1, \tag{7.4}$$

$$\sum_{j=1}^{n} Z_m(F^j \circ G_{a_1}\mathbf{x}) \geq 2. \tag{7.5}$$

Let $\mathbf{y} = F^{n-1} \circ G_{a_1}(\mathbf{x}) \in X$. Subtracting the Ineq. 7.4 for $l = m$ from the Ineq. 7.5 we obtain that $Z_m(F^n \circ G_{a_1}\mathbf{x}) \geq 1$. Therefore $y_m > 1$

$$1 < y_m = x_m + \delta \cdot \delta_{m,a_1} + \alpha \frac{\sum_{l=1}^{N}\sum_{j=1}^{n-1} Z_l(F^j \circ G_{a_1}\mathbf{x})}{N} - 1 < x_i + \delta + \alpha - 1 < x_i < 1,$$

hence we come to the contradiction to the fact that $I_c \neq \emptyset$.

□

Remark 7.1.3. *Proposition 7.1.2 has a physical meaning that esch neuron can fire only once during an avalanche.*

Proposition 7.1.4. *Let $x \in X_s$ and let l satisfy $\#\left\{i : x_i > 1 - \alpha\frac{l}{N}\right\} < l$ and let $\alpha + \delta < 1$ then*

$$\xi(a,\boldsymbol{x}) = \sum_{j=0}^{k(a,x)} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a \boldsymbol{x}\right) < l.$$

Proof. Let $\mathbf{y} = T(a,\mathbf{x}) = F^{k(a,\mathbf{x})} \circ G_a\mathbf{x}$, for any $m \leq k(a,\mathbf{x})$ define the avalanche size after m steps as,

$$\xi_m = \sum_{j=0}^{m} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right). \tag{7.6}$$

We are going to prove that if $\xi(a,\mathbf{x}) = \xi_{k(a,\mathbf{x})} \geq l$ then there exit at least l distinct indices $(i_n)_{n=1}^{l}$ such that $x_{i_n} > 1 - \alpha\frac{l}{N}$.

If $\xi_{k(a,\mathbf{x})} \geq l$ then there exists $n < k(a,\mathbf{x})$ such that $\xi_n < l$ and $\xi_{n+1} \geq l$. By the Prop. 7.1.2 for any $i \leq N$, $\sum_{j=0}^{m} Z_i\left(F^j \circ G_a\mathbf{x}\right) \leq 1$. Combining it with the

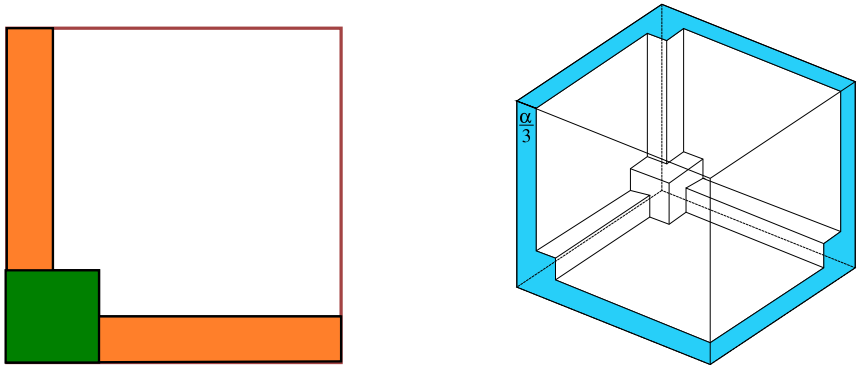


Figure 7.1: Set D. (left) for $N = 2$, set D is uncolored, C_1 is green, C_2 is orange. (right) $N = 3$ set D is uncolored.

definition of ξ_m (7.6) we find that there are ξ_n distinct indices $q_1, \ldots, q_{\xi_n}$ such that $\sum_{j=0}^{n} Z_{q_s}\left(F^j \circ G_a \mathbf{x}\right) = 1$, for all $s = 1, \ldots, \xi_n$. For these indices

$$0 < \pi_X\left(T^n \circ G_a \mathbf{x}\right) = x_{j_s} + \alpha \xi_n / N - 1$$

therefore $x_{q_s} > 1 - \alpha \xi_n / N > 1 - \alpha l / N$. There are $\xi_{n+1} - \xi_n > 0$ indices k_t such that for any $t = 1, \ldots, \xi_{n+1} - \xi_n$,

$$\sum_{j=0}^{n} Z_{k_t}\left(F^j \circ G_a \mathbf{x}\right) = 0, \tag{7.7}$$

$$\sum_{j=0}^{n+1} Z_{k_t}\left(F^j \circ G_a \mathbf{x}\right) \neq 0. \tag{7.8}$$

Due to (7.7) the sets of indices $\{q_i, \ldots, q_s\}$ and $\{k_1, \ldots, k_t\}$ are not intersecting. We define a set B as $B = \{q_i, \ldots, q_s\} \bigcup \{k_1, \ldots, k_t\}$. From (7.8) it follows that for any $t \leq \xi_{n+1} - \xi_n$

$$1 < \pi_X\left(T^{n+1} \circ G_a \mathbf{x}\right)_{k_t} = x_{k_t} + \alpha \xi_n / N,$$

hence $x_{k_t} > 1 - \alpha \xi_n / N > 1 - \alpha l / N$. Therefore, for all indices $i \in B$, $x_i > 1 - \alpha l / N$ and $\#(B) = \xi_{n+1} \geq l$. □

Remark 7.1.5. *Proposition 7.1.4 allows to define an avalanche size directly from the configuration x without performing the iterations of mapping F.*

7.2 The measure support

Definition 7.2.1. Let for any $a \in \mathcal{A}$, $D_a \subset X_s$ be a set

$$D_a = \bigcap_{i=1}^{\infty} \pi_X(T^i(a, X_s)), \tag{7.9}$$

and $\hat{D}_a$ be a subset of X_s on which $\pi_X\left(T(\sigma^n a, \cdot)\right)$ acts bijective for any $n \in N$ i.e.

$$\hat{D}_a = \Big\{ \mathbf{x} \in X_s \,\Big|\, \forall\, n \in \mathbb{N} \quad \pi_X\left(T(\sigma^n a, \mathbf{x})\right) \in \hat{D}_a, \\ \text{and} \quad \exists\, \mathbf{y} \in \hat{D}_a \quad \pi_X\left(T(\sigma^n a, \mathbf{y})\right) = \mathbf{x} \Big\} \tag{7.10}$$

Remark 7.2.1. *It is easy to see that for any $a \in \mathcal{A}$, $\hat{D}_a$ is a subset of D_a. Later in this this section we will show that also the opposite inclusion holds.*

It is clear that $\hat{D}_a \neq X_s$. For example, if $\mathbf{x}_i < \min(\alpha/N, \delta)$ there exists no such $b \in \mathcal{A}$, $\mathbf{y} \in X_s$ that $\pi_X T(b, \mathbf{y}) = \mathbf{x}$.

Let us define a collection of sets C_i such as

$$C_i = \left\{ \mathbf{x} \in X_s \,\middle|\, \#\left\{ j \leq N | x_j < \frac{i\alpha}{N} \right\} \geq i \right\}. \tag{7.11}$$

Remark 7.2.2. *If $\boldsymbol{x} \in X_s \setminus C_l$ then there exists at least $N - l + 1$ distinct indices $k_1, \ldots, k_{N-l+1}$ such that $x_{k_j} > \alpha l/N$.*

$$\#\left\{ i \,\middle|\, x_i > \alpha \frac{l}{N} \right\} \geq N - l + 1 \tag{7.12}$$

Theorem 7.2.3. *For any $a \in A$*

$$\hat{D}_a = X_s \setminus \bigcup_{i=1}^{N} C_i.$$

Proof. Denote $D := X_s \setminus \bigcup_{i=1}^{N} C_i$.
We will first prove that T is invariant on D and then that it is also invertable on D in the sense of definition 7.10.
1. Let us fix any $\mathbf{x} \in D$ and any $a \in \mathcal{A}$. We have to prove that $\pi_X(T(a, \mathbf{x})) \in D$. In other words, we have to prove that for all $i = 1, \ldots, N$, $\mathbf{y} = \pi_X\left(T(a, \mathbf{x})\right) \notin C_i$. To prove this we find a set of distinct indices $L_i := \{l_1, \ldots, l_{N-i+1} |\ \forall m,\ \ l_m \leq N\}$ such that for any $m \in L_i$, $y_{l_m} > \alpha i/N$. Consider two cases $k(a, \mathbf{x}) = 0$ and $k(a, \mathbf{x}) > 0$:

If $k(a,\mathbf{x}) = 0$ then for each j, $x_j \le y_i$, hence for each i, if $\mathbf{x} \notin C_i$ then $\mathbf{y} \notin C_i$. Therefore, we have to consider only the case $k(a,\mathbf{x}) > 0$. By (7.3) we can define the avalanche size $\xi(a,x) = \xi$. We will distinguish three cases $\xi < i$, $\xi = i$, and $\xi > i$.

a) Let $\xi < i$. We know that $\mathbf{x} \notin C_{i-\xi}$, then by Rem. 7.2.2 there exist $N-(i-\xi)+1$ distinct indices $k_1, \dots, k_{N-(i-\xi)+1}$ such that $x_{k_j} > \alpha(i-\xi)/N$.

We have defined $\sum_{j=0}^{k(a,\mathbf{x})} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right) = \xi$, hence there exist at least $N-\xi$ distinct indices r_s such that $\sum_{j=1}^{N} Z_{r_s}\left(F^j \circ G_a\mathbf{x}\right) = 0$.

$$\#\left(\left\{k_1, \dots, k_{N-(i-\xi)+1}\right\} \bigcap \left\{r_1, \dots, r_{N-\xi}\right\}\right) \ge N - i + 1,$$

therefore there exist distinct indices $(l_m)_1^{N-i+1}$ such that for all $m \le N-i+1$, $\sum_{j=0}^{k(a,\mathbf{x})} Z_{l_m}\left(F^j \circ G_a\mathbf{x}\right) = 0$ and $x_{l_m} > \alpha(i-\xi)/N$.

$$\begin{aligned} y_i &= x_i + \frac{\alpha}{N} \sum_{j=0}^{k} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right) - \sum_{j=0}^{k} Z_i\left(F^j \circ G_a\mathbf{x}\right) \\ &= x_i + \alpha\frac{\xi}{N} - \sum_{j=0}^{k} Z_i\left(F^j \circ G_a\mathbf{x}\right), \end{aligned} \tag{7.13}$$

hence

$$y_{l_m} = x_i + \alpha\frac{\xi}{N} > \alpha\frac{i-\xi}{N} + \alpha\frac{\xi}{N} = \alpha\frac{i}{N} \tag{7.14}$$

b) Let $\xi = i$. Again, as $\sum_{j=0}^{k(a,\mathbf{x})} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right) = \xi = i$ there exists a set of $N-i$ distinct indices $K_{N-i} = \{k_1, \dots, k_{N-i}\}$ such that for any $m \in K_{N-i}$, $\sum_{j=0}^{k(a,\mathbf{x})} Z_{k_m}\left(F^j \circ G_a\mathbf{x}\right) = 0$, and for them by (7.13, 7.14) we have $y_{k_m} > \alpha i/N$.

Consider x_{a_1}: As $\xi > 0$, we have $x_{a_1} + \delta > 1$ and $\sum_{j=0}^{k(a,\mathbf{x})} Z_{a_1}\left(F^j \circ G_a\mathbf{x}\right) = 1$. Consequently

$$\begin{aligned} y_{a_1} &= x_{a_1} + \frac{\alpha}{N} \sum_{j=0}^{k(a,\mathbf{x})} \sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right) - \sum_{j=0}^{k(a,\mathbf{x})} Z_{a_1}\left(F^j \circ G_a\mathbf{x}\right) \\ &= x_{a_1} + \delta + \alpha\frac{i}{N} - 1 > \alpha\frac{i}{N} \end{aligned}$$

We have $a_1 \notin K_{N-i}$, thus $L_i = \{a_1\} \cup K_{N-i}$

c) Let $\xi > i$. As above, we find set of indices $K_{N-\xi}$, such that for any $m \in K_{N-i}$, $\sum_{j=0}^{k(a,\mathbf{x})} Z_{k_m}\left(F^j \circ G_a\mathbf{x}\right) = 0$, and $y_{k_m} > \alpha\xi/N > \alpha i/N$.

Applying Rem. 7.1.4, we find that there are at least $\xi - i$ indices j_s, such that for

any $s \leq \xi - i$, $x_{j_s} > 1 - \alpha(\xi - i)/N$. From Prep. 7.1.2 $\sum_{j=0}^{k} Z_{j_s}(F^j \circ G_a \mathbf{x}) < 2$. On the other hand if $\sum_{j=0}^{k} Z_{j_s}(F^j \circ G_a \mathbf{x}) < 1$, then

$$y_{j_s} = x_{j_s} + \alpha\frac{\xi}{N} > 1 - \alpha\frac{\xi - i}{N} + \alpha\frac{\xi}{N} = 1 + \alpha\frac{i}{N} > 1,$$

which is impossible. Therefore $\sum_{j=0}^{k} Z_{j_s}(F^j \circ G_a \mathbf{x}) = 1$. Hence for these indices

$$y_{j_s} = x_{j_s} + \alpha\frac{\xi}{N} - 1 > 1 - \alpha\frac{\xi - i}{N} + \alpha\frac{\xi}{N} - 1 = \alpha\frac{i}{N}.$$

Sets of indices $\{k_1, \ldots, k_{N-\xi+1}\}$ and $\{j_1, \ldots, j_{\xi-i}\}$ are not intersecting. Hence we found $L_i = \{k_1, \ldots, k_{N-\xi+1}\} \cup \{j_1, \ldots, j_{\xi-i}\}$.

2. Now we prove that for any $\mathbf{x} \in D$ and any $b \in \mathcal{A}$ there exists the preimage of $\mathbf{y} \in D$ such that $\pi_X(T(b, \mathbf{y})) = \mathbf{x}$. Let $b = cb_1b_2\ldots$. If $\mathbf{y} = \mathbf{x} - \delta e_c \in D$ then $\mathbf{y}$ is a desired preimage.

Let $\mathbf{y}' = \mathbf{x} - \delta e_c \notin D$. We have $\mathbf{x} \in D$ thus for all $i \leq N$, $x_i > \alpha/N$ and $\delta < \alpha/N$ hence $\mathbf{y} \in X_s$ and therefore there exists l such that $\mathbf{y}' \in C_l$. Thus there exists a set of indices $I_l = \{i_1, \ldots, i_l\}$ such that for any $h \leq l : y_{i_h} < \alpha\frac{l}{N}$. If $c \notin I_l$ then $\mathbf{x} \in C_l$, which contradict to the fact that $\mathbf{x} \in D$. Therefore there exists j, such that

$$x_c > \alpha\frac{j}{N} \quad \text{and} \quad x_c - \delta e_c < \alpha\frac{j}{N}.$$

We have $\delta < \alpha/N$, hence such index j is unique and $j = l$.

Let us define a transformation $S : X_s \to X$ such that

$$S(\mathbf{x}) = \mathbf{x} - \alpha e + e_{m(\mathbf{x})}, \tag{7.15}$$

where $m(\mathbf{x}) = \text{argmin}_{i \leq N, i \neq c}(x_i, x_c - \delta)$. Define $\mathbf{y}$ as:

$$y = S^l(\mathbf{x}) - \delta e_c \tag{7.16}$$

In the following we will show that $\mathbf{y} = S^l(\mathbf{x}) - \delta e_c$ is a desired preimage.

In Fig. 7.2 the action of the mapping S is shown for the 2-d system. After some number of iterations of S for both initial conditions the image is found to be contained in D. After this time the iterating procedure is stopped and final value $\mathbf{y}$ is obtained. In the following we show that for any initial condition $\mathbf{x}$ for which $\mathbf{y}' = \mathbf{x} - \delta e_c \notin D$, such procedure gives a single solution $\mathbf{y}$ and it is a preimage of $\mathbf{x}$

At first let us show that $\mathbf{y} \in D$. As we know $\mathbf{y}' \in C_l$. Denote $H = \{h_1 \ldots h_l\}$ a set of coordinates such that for any $h \in H$ $y'_h < \alpha\frac{l}{N}$. Furthermore, let for any $i < l - 2 : h_i < h_{i+1}$, and $h_l = c$. The set H contains the indices of the l smallest

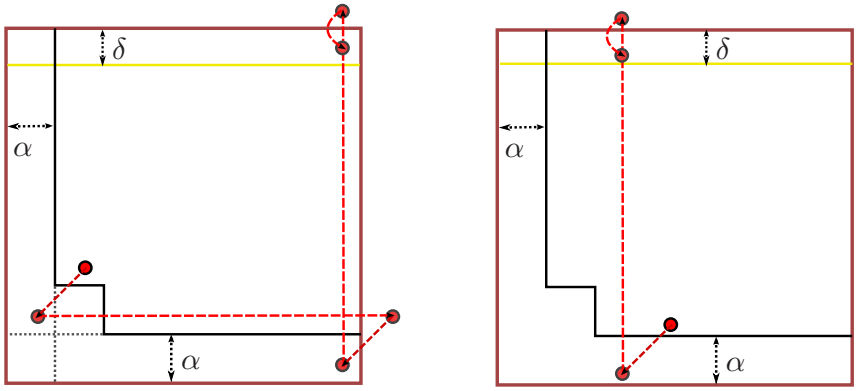


Figure 7.2: Procedure of the preimage identification for $N = 2$. $\mathbf{y}' \in C_2$, then $l = 2$ (left) and $\mathbf{y}' \in C_1$, then $l = 1$ (right). Arrows indicate the sequence of images of $\mathbf{x}$ (depicted by red circles) under the action of S.

coordinates of $\mathbf{y}'$. Since $\mathbf{x} \in D$, yields for any $i < l$, $h_i > \alpha\frac{i}{N}$. As H is odered and contains the indices of smallest coordinates of $\mathbf{y}'$, we have $m(\mathbf{x}) = h_1$.

Application of S does not change the order of the coordinates except of the smallest coordinate. Let for some $g < l$, $\mathbf{y}^g = S^g(\mathbf{x})$, and $i = m(\mathbf{y}_g)$, then $y_i^g = y_i^{g-1} - \frac{\alpha}{N} + 1 > \alpha\frac{l}{N}$. Thus for any $g \le l$, $m\left(S^g(\mathbf{x})\right) = h_g$, and we can write the coordinates of the vector $\mathbf{y}$ explicitly

$$y_i = \begin{cases} x_i(t) - \alpha\frac{l}{N} + 1 & \text{if } i \in H \text{ and } i \neq c, \\ x_c(t) - \alpha\frac{l}{N} + 1 - \delta & \text{if } i = c, \\ x_i(t) - \alpha\frac{l}{N} & \text{otherwise.} \end{cases}$$

This implies that for any $i : 0 < y_i < 1$, which mean that $\mathbf{y} \in X_s$. Now we show that $g = l$ is the first entrance in X_s of the map $S^g(\mathbf{x})$ among all $g > 0$. I.e. we should proof that for any $l > g > 0 : \mathbf{y}^g = S^g(\mathbf{x}) \notin X_s$. Let $i = m\left(S^g(\mathbf{x})\right)$, thus $i = h_g$ and $x_i > \alpha\frac{i}{N}$. Then

$$y_i^g = x_i - \alpha\frac{i}{N} + 1 > 1.$$

Now we show that $\pi_X\left(T(b, \mathbf{y})\right) = \mathbf{x}$. We have $\tilde{\mathbf{y}} = G_c(\mathbf{y})_c = y_c + \delta = x_c(t) - \alpha\frac{l}{N} + 1$, but $x_c > \alpha\frac{l}{N}$, thus $k(b, \mathbf{y}) > 0$.

$$F(\tilde{\mathbf{y}}) = \tilde{\mathbf{y}} + \frac{\alpha}{N} e \operatorname{card}\{i | z_i = 1\} - \mathbf{z} = \sum_{i, \tilde{y}_i > 1} \tilde{\mathbf{y}} + \frac{\alpha}{N} - e_i,$$

thus $F(\tilde{\mathbf{y}}) = S^{-\operatorname{card}\{i | z_i = 1\}}(\mathbf{y})$. For all $i < l$, $S^{-i}(\mathbf{y}) \notin X_s$ and $S^{-l}(\mathbf{y}) \in X_s$, therefore $F^{k(b,\mathbf{y})} = S^{-l}(\mathbf{y}) = \mathbf{x}$.

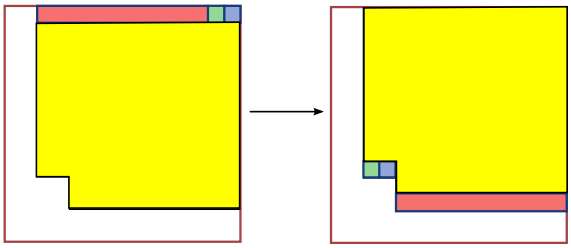


Figure 7.3: Image of $\pi_X T(a, \mathbf{x})$ for different $\mathbf{x} \in D$. Each area on the left has the same color as its image on the right.

□

Proposition 7.2.4. *For any $\boldsymbol{x} \in X_s$ and $a \in \mathcal{A}$ there exists $n \in \mathbb{N}_0 < \infty$ such that $\pi_X(T^n(a, X)) \in D$*

Proof. Let $\mathbf{x} \in C_1$. $G_a\mathbf{x} > \mathbf{x}$ and $\mathbf{x} < e$, hence there exists m, such that for $(b, \mathbf{y}) = T^m(a, \mathbf{x})$, $T(b, \mathbf{y}) = F^{k(b,y)} \circ G_b(b, \mathbf{y})$ and $k(b, y) \geq 1$. For this $(b, \mathbf{y})$ the avalanche size ξ is not smaller than 1. Repeating the reasoning of Theorem 7.2.3 1.b we prove that $\pi_X(T(b, \mathbf{y})) \notin C_1$.

Let $\mathbf{x} \in X \setminus \bigcup_{i=1}^{n-1} C_i$ and $T = F^k(a, \mathbf{x}) \circ G_a$, where $k(a, \mathbf{x}) \geq 1$, then $T(a, \mathbf{x}) \in X \setminus \bigcup_{i=1}^{n} C_i$. In the proof of the Theorem 7.2.3 1 we used only fact that $\mathbf{x} \notin C_j$, $j < i$ to prove that $T(a, \mathbf{x}) \notin C_i$. Hence, knowing that $\mathbf{x} \notin \bigcup_{i=1}^{n-1} C_i$ we can repeat the same reasoning to prove that $T(a, \mathbf{x}) \notin C_l$ for any $l \leq n$.

□

Remark 7.2.5. *The proposition above provides the inclusion $D_a \subset D$ and combined with the Theorem 7.2.3 proves that if there exists measure on X, which is invariant under T, then it will be concentrated on D.*

Proposition 7.2.6. *Let $\boldsymbol{x} \in D$, $a = (a_1, \ldots) \in \mathcal{A}$, $\alpha + \delta < 1$, $\boldsymbol{z} = \pi_X(T(a, \boldsymbol{x}))$ then the following statements are equivalent:*

1. $k(a, \boldsymbol{x}) \geq 1$
2. $\boldsymbol{z} - \delta e_j \in C_l$ *for* $j = a_1$ *and* $l = \xi(a, \boldsymbol{x})$.

Proof. *"1⇒2":* Let

$$S = \left\{ j \,\middle|\, \sum_{i=0}^{k(a,\mathbf{x})} Z_j \left(F^i \circ G_a \mathbf{x} \right) \neq 0 \right\},$$

$|S| = \xi(a, \mathbf{x})$, $a_1 \in S$. For any $j \in S$, $j \neq a_1$

$$\begin{aligned} \mathbf{z}_j &= \mathbf{x}_j - 1 + \alpha\frac{\xi}{N} < \alpha\frac{\xi}{N}, \\ \mathbf{z}_{a_1} &= \mathbf{x}_{a_1} + \delta - 1 + \alpha\frac{\xi}{N} < \alpha\frac{\xi}{N} + \delta. \end{aligned}$$

Therefore, $\mathbf{z} - \delta e_{a_1} \in C_\xi$.
"2⇒1": As $\mathbf{z} - \delta e_j \in C_l$ it is clear that $\mathbf{x} \neq \mathbf{z} - \delta e_j$. But $\mathbf{z} = \pi_X(T(a, \mathbf{x})) \neq \mathbf{x} + \delta e_j = G_a(x)$, thus $k(a, \mathbf{x}) > 0$ □

Theorem 7.2.7.

1. *If $\boldsymbol{x} \in D$, and $a = \{a_1, a_2 \ldots\}, b = \{b_1, b_2, \ldots\} \in \mathcal{A}$ such that $a_1 \neq b_1$, then $\pi_X(T(a, \boldsymbol{x})) \neq \pi_X(T(b, \boldsymbol{x}))$.*

2. *For any $\boldsymbol{x} \in D$, there exist distinct $\boldsymbol{y}^1, \ldots, \boldsymbol{y}^n$ such that $\boldsymbol{x} = \pi_X(T(a^i, \boldsymbol{y}^i))$, for $a^i = (i, a_2, a_3, \ldots)$.*

3. *For any $\boldsymbol{x} \in D$ and $a \in \mathcal{A}$ there exists one and only one $\boldsymbol{y} \in D$, such that $\pi_X\left(T(a, \boldsymbol{y})\right) = \boldsymbol{x}$.*

Proof. 1. Let $\mathbf{y} = \pi_X(T(a, \mathbf{x})) = \pi_X(T(b, \mathbf{x}))$, then

$$\begin{aligned} y_{a_1} &= x_{a_1} + \delta + \frac{\alpha}{N}\sum_{j=0}^{k(a,\mathbf{x})}\sum_{i=1}^{N} Z_i\left(F^j \circ G_a\mathbf{x}\right) - \sum_{j=0}^{k(a,\mathbf{x})} Z_i\left(F^j \circ G_a\mathbf{x}\right) \\ &= x_{a_1} + \frac{\alpha}{N}\sum_{j=0}^{k(a,\mathbf{x})}\sum_{i=1}^{N} Z_i\left(F^j \circ G_b\mathbf{x}\right) - \sum_{j=0}^{k(a,\mathbf{x})} Z_i\left(F^j \circ G_b\mathbf{x}\right). \end{aligned}$$

Hence

$$x_{a_1} + \delta + \alpha\frac{c_1}{N} - d_1 = x_{a_1} + \alpha\frac{c_2}{N} - d_2,$$

where $c_1, c_2, d_1, d_2 \in \mathbb{N}_0$. Such equation has no solutions, because α and δ are rationally independent.
2. We showed in the proof of the Theorem 7.2.3 (2) that for any $\mathbf{x} \in D$, $a \in \mathcal{A}$ and $b \in \mathcal{A}$, $b = b_1, a_1, a_2, \ldots$ there exists $\mathbf{y} \in D$, such that $T(b, \mathbf{y}) = (a, \mathbf{x})$. So it is only left to prove that if $b \neq c$, and $\pi_X(T(b, \mathbf{y}^1)) = \pi_X(T(c, \mathbf{y}^2)) = \mathbf{x}$ then $\mathbf{y}^1 \neq \mathbf{y}^2$. If this condition is not fulfilled, then for some $\mathbf{y} \in D$ there exists $b \neq c$, such that $\pi_X(T(b, \mathbf{y})) = \pi_X(T(c, \mathbf{y}))$, which contradicts to the proven statement 1.
3. Let $\pi_X\left(T(a, \mathbf{y}^1)\right) = \pi_X\left(T(a, \mathbf{y}^2)\right) = \mathbf{x}$. Without loss of generality we can assume that $a_1 = 1$. Let $k(a, \mathbf{y}^1) = k_1$ and $k(a, \mathbf{y}^2) = k_2$. As it was proved in Prop. 7.2.6, if $\mathbf{x} - \delta e_1 \in D$, then $k_1 = k_2 = 0$ and $y_1 = \mathbf{x} - \delta e_1 = \mathbf{y}_2$. If $\mathbf{x} - \delta e_1 \notin D$, then by

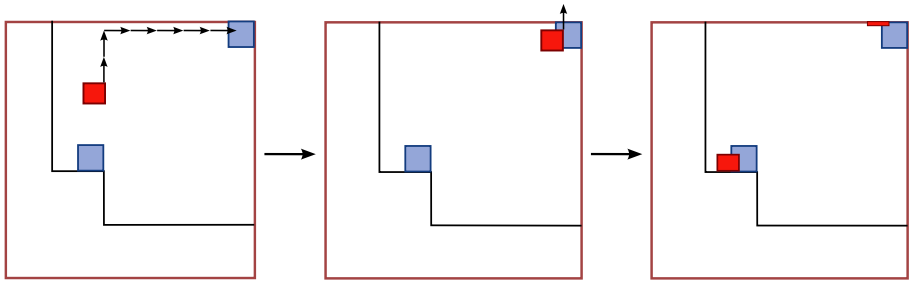


Figure 7.4: Action of the combination of the procedures **a** and **b** on a small cube depicted in red. Q_δ and Q_i are depicted in blue.

the Prop. 7.2.6 $\xi(a, \mathbf{y}^1) = \xi(a, \mathbf{y}^2) =: \xi$. For any $j \leq N$

$$x_j = y_j^1 + \alpha\frac{\xi}{N} - \sum_{j=0}^{k_1} Z_i\left(F^j \circ G_a \mathbf{y}^1\right) = y_j^2 + \alpha\frac{\xi}{N} - \sum_{j=0}^{k_2} Z_i\left(F^j \circ G_a \mathbf{y}^2\right). \quad (7.17)$$

If $\sum_{j=0}^{k_1} Z_i\left(F^j \circ G_a \mathbf{y}^1\right) = \sum_{j=0}^{k_2} Z_i\left(F^j \circ G_a \mathbf{y}^2\right)$ then by Eq. 7.17 $y_j^1 = y_j^2$. Otherwise, without loss of generality $\sum_{j=0}^{k_1} Z_i\left(F^j \circ G_a \mathbf{y}^1\right) > \sum_{j=0}^{k_2} Z_i\left(F^j \circ G_a \mathbf{y}^2\right)$. Then by the Prop.7.1.2 $\sum_{j=0}^{k_1} Z_i\left(F^j \circ G_a \mathbf{y}^1\right) = 1$ and $\sum_{j=0}^{k_2} Z_i\left(F^j \circ G_a \mathbf{y}^2\right) = 0$. Thus $y_j^1 - 1 = y_j^2$, but $\mathbf{y}^1, \mathbf{y}^2 \in D$ and therefore $y_j^1 < 1$, $y_j^2 > 0$ which is impossible. Hence we showed that for any $j \leq N$, $y_j^1 = y_j^2$. □

Proposition 7.2.8. *[53]* $\mathrm{Vol}(X_s \setminus D) = \alpha$

7.3 Existence and properties of an invariant measure

Theorem 7.3.1. *Let μ be a uniform Bernoulli measure on $\mathcal{A}$ and λ a Lebesgue measure on D. Then $\mu \times \lambda$ is a measure on $\mathcal{A} \times D$, which is invariant under the transformation T.*

Proof. We have to prove that for any $a \in \mathcal{A}$ and $B \subset D$, $(\mu \times \lambda)(a, D) = (\mu \times \lambda)(T^{-1}(a, B))$. $\pi_{\mathcal{A}}\left(T^{-1}(a, B)\right) = \{g_1, \ldots, g_N\}$, where $g_i = \{i, a_1, a_2, \ldots\}$. Let $B_1, \ldots, B_N \subset D$ satisfy $T(g_i, B_i) = (a, B)$. Then $T^{-1}(a, B) = \bigcup_{i=1}^{N}(g_i, B_i)$.

We prove that for any $i \leq N$, $\lambda(B_i) = \lambda(B)$. Without loss of generality $i = 1$. Let us cut D by the hyperplane $x_1 = 1 - \delta$, hence we define $D_1 = \{\mathbf{x} \in C | x_1 > 1 - \delta\}$, $D_2 = C \setminus D_1$. On D_2, $\pi_X T(b_1, \cdot)$ acts as a shift, and therefore preserves the measure λ.

We can decompose D_1 into a union of parallelepipeds

$$D_{i_1,i_2,\dots i_N} = \{x \in D | x_1 > 1-\delta, 1-\alpha\frac{i_2+1}{N} < x_2 < 1-\alpha\frac{i_2}{N}, \dots \\ 1-\alpha\frac{i_N+1}{N} < x_N < 1-\alpha\frac{i_N}{N}, i_1, \dots, i_N \in \mathbb{N}_0\}.$$

On each $D_{i_1,i_2,\dots i_N}$, $\pi_X T(b_1, \cdot)$ acts as a shift and by Theorem 7.2.3 (statement 3) for any $l \neq m$, $\pi_X T(b_1, D_{i_l}) \bigcap \pi_X T(b_1, D_{i_m}) = \emptyset$, hence $\pi_X T(b_1, \cdot)$ preserves the measure λ.

□

Theorem 7.3.2. *For almost all (with respect to a uniform Bernoulli measure) sequences $a \in \mathcal{A}$, $\pi_X T(a, \boldsymbol{x})$ is topologically transitive on D.*

It is sufficient to prove that for any rectangular parallelepiped $Q_1 \subset D$ and cube $Q_2 \subset D$ and almost any $a \in \mathcal{A}$, there exists n such that $\pi_X T^n(a, Q_1) \bigcap Q_2 \neq \emptyset$. Without loss of generality we can consider only cubes with the side length smaller than δ.

Plan of the proof:

1. First we will show that for any point $\mathbf{x} \in D$ there exists a finite sequence $a^m = \{a_1, \dots, a_m\}$ such that for any sequence a beginning with a^m, all coordinates of the point $\pi_X T^m(a, \mathbf{x}) \in D$ have any predetermined residuals mod δ with any predetermined precision.

2. Then we will introduce a partial order on D and show that there exists finite sequence which leads to the mapping of any given point in D to the vicinity of any of the minimum with respect to this partial order.

3. Having proved both previous points, we can show that for any two rectangular parallelepipeds there exists the final transformation, such that the image of the first will intersect with the second.

Before starting the proof we will recall the conditions, which were posed on the parameters of the model α and δ:

i) α and δ are irrational and rationally independent,

ii) $\alpha + \delta < 1$,

iii) $\delta < \alpha/N$,

iv) $\alpha\frac{N+1}{N} < 1$.

The last condition is new, it does not pose much of physical constrains, but helps in the proof

Lemma 7.3.3. *Consider $\varepsilon > 0$ and any point $\boldsymbol{z} \in [0,\delta)^N$, then for any point $\boldsymbol{x} \in D$ there exists a sequence $a(\boldsymbol{x},\boldsymbol{z}) = (a_i)_{i=1}^M$ such that*

$$\left|\pi_X\left(T^M(a(\boldsymbol{x},\boldsymbol{z}),\boldsymbol{x})\right) - \boldsymbol{z}\right| < \varepsilon \pmod{\delta} \tag{7.18}$$

and $M = M(N,\varepsilon,\alpha,\delta)$ does not depend on $\boldsymbol{x}$ and $\boldsymbol{z}$.

In the proof of this lemma we will use the following Minkowski theorem. Let $\|w\|$ denote the distance from w to the nearest integer to w , i.e.

$$\|w\| = \min_{n\in\mathbb{N}} |w-n|. \tag{7.19}$$

Theorem 7.3.4. [31] *If θ is irrational and α can not be represented in a form $\alpha = m\theta + n$ for any $m,n \in \mathbb{N}$ then there exist infinitely many integer q such as*

$$|q|\|q\theta - \alpha\| < \frac{1}{4}$$

Proof of the Lemma 7.3.3. **a.** Denote Q_δ the cube $[1-\delta,1]^N$. For any point $\mathbf{x} \in D$ we construct a sequence $b = (b_i)_{i=1}^{M'}$ such that for any $a \in \mathcal{A}$, $\pi_X T^{M'}(ba,\mathbf{x}) \in Q_\delta$ and $M' \le N\lfloor N/\delta \rfloor$, where $\lfloor w \rfloor$ is the largest integer which is smaller than w and $ba = (b_1,\ldots,b_{M'},a_1,\ldots)$.

$$b = \overbrace{1\ldots1}^{i_1}\overbrace{2\ldots2}^{i_2}\ldots\overbrace{N\ldots N}^{i_N},$$

where $i_k = \left\lfloor \frac{1-x_k}{\delta} \right\rfloor$.

For any $j \le M' = \sum_{k=1}^N i_k$ we have $T^j(b,\mathbf{x}) = G_{\sigma^j(b)}(T^{j-1}(b,\mathbf{x}))$ and the addition of δ does not change the residuals modulus δ, therefore

$$\pi_X T^{M'}(a,\mathbf{x}) \equiv \mathbf{x} \pmod{\delta}. \tag{7.20}$$

b. Define for every $i \le N$

$$Q_i = [\alpha-\delta,\alpha)^{i-1} \times [\alpha,\alpha+\delta) \times [\alpha-\delta,\alpha)^{N-i}. \tag{7.21}$$

Q_i is a square, laying in the δ-neighborhood of C_N. If $\mathbf{x} \in Q_\delta$ then for any $i \le N$ $x_i + \alpha > x_i + \delta > 1$. Let $a_1 = i$ then by the conditions ii) and iii)

$$\begin{aligned} G_a(\mathbf{x}) &= (x_1,\ldots,\mathbf{x}_{i-1},\mathbf{x}_i+\delta,x_{i+1},\ldots,x_N) \notin D, \\ F\circ G_a(\mathbf{x}) &= \left(x_1+\frac{\alpha}{N},\ldots,\mathbf{x}_i+\delta+\frac{\alpha}{N}-1,\ldots,x_N+\frac{\alpha}{N}\right) \notin D, \end{aligned}$$

$$F^2 \circ G_a(\mathbf{x}) \;=\; (x_1 + \alpha - 1, \ldots, \mathbf{x}_i + \delta + \alpha - 1, \ldots, x_N + \alpha - 1) \in D.$$

Therefore $T(a, \mathbf{x}) = F^2 \circ G_a(\mathbf{x}) \in Q_i$. For any $j, k \leq N$,

$$\pi_X T(a, \mathbf{x})_j - \pi_X T(a, \mathbf{x})_k \equiv x_j - x_k \pmod{\delta} \tag{7.22}$$

and for any j

$$\pi_X T(a, \mathbf{x})_j \equiv x_j + \alpha - 1 \pmod{\delta}. \tag{7.23}$$

Consider any $\mathbf{x} \in D$. By $\mathbf{a}$ we have a sequence $b = (b_i)_{i=1}^{M'}$ which "transports" $\mathbf{x}$ to the cube Q_δ. As it is shown in $\mathbf{b}$ concatenating the sequence b and one element sequence i we obtain $c = bi$, such that for any $a \in \mathcal{A}$ $\pi_X T^{M'+1}(ca, \mathbf{x}) \in \mathbb{Q}_i$.

Notation. Henceforth we refer to $\mathbf{y} = \pi_X T^{M'}(ba, \mathbf{x}) \in Q_\delta$ as *result of application of a procedure* $\boldsymbol{a}$ *to* $\boldsymbol{x}$. We call $c = bi$ *the combination of the procedures* $\boldsymbol{a}$ *and* $\boldsymbol{b}$

A combination of the procedures $\mathbf{a}$ and $\mathbf{b}$ applied to $\mathbf{x}$ keeps the differences between coordinates of $\mathbf{x}$ fixed $\pmod{\delta}$, but changes the residuals of each coordinate by $\alpha - 1 \pmod{\delta}$ Eq. (7.20, 7.22, 7.23). Fig. 7.4 illustrates the action of such a combination on a cube for $N = 2$.

c. Let $\mathbf{x} \in Q_k$ and let L be the smallest integer such that $x_k + L\delta \geq 1$. Consider $c = (c_j)_{j=1}^{L}$ is such that for any $j < L$, $c_j = k$ then, for any $a \in \mathcal{A}$

$$\mathbf{y} = T^{L-1}(ca, \mathbf{x}) = (x_1, \ldots, \mathbf{x}_{k-1}, x_k + (L-1)\delta, x_{k+1} \ldots, x_N)\,. \tag{7.24}$$

By the condition iv) on α and δ

$$\begin{aligned} G_k(\mathbf{y}) \;&=\; (x_1, \ldots, \mathbf{x}_{k-1}, x_k + L\delta, x_{k+1} \ldots, x_N) \notin D, \\ F \circ G_k(\mathbf{y}) \;&=\; \left(x_1 + \frac{\alpha}{N}, \ldots, x_k + L\delta + \frac{\alpha}{N} - 1, \ldots, x_N + \frac{\alpha}{N}\right) \in D. \end{aligned} \tag{7.25}$$

Thus we have procedure $\mathbf{c}$

$$\pi_X \left(T^L(ca, \mathbf{x})\right) = \left(x_1 + \frac{\alpha}{N}, \ldots, x_k + L\delta + \frac{\alpha}{N} - 1, \ldots, x_N + \frac{\alpha}{N}\right). \tag{7.26}$$

Procedure $\mathbf{c}$ applied to $\mathbf{x}$ allows to keep all residuals but x_k fixed, whereas residual of x_k is changed by $\frac{\alpha}{N} - 1 \pmod{\delta}$.

Let us consider all coordinates of $\mathbf{x}$ relative to the first coordinate $\pmod{\delta}$ i.e. consider the vector $V(\mathbf{x}) = (x_2 - x_1 \pmod{\delta}, \ldots, x_N - x_1 \pmod{\delta}) =: \tilde{\mathbf{x}}$. Define $\tilde{\mathbf{z}} := (z_2 - z_1 \pmod{\delta}, \ldots, z_N - z_1 \pmod{\delta})$. In the following we show how to combine procedures $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ to approximate $\tilde{\mathbf{z}}$ by $\tilde{\mathbf{y}} = V\left(\pi_X\left(T^{M'}(a, \mathbf{x})\right)\right)$ such that for any $1 < k \leq N$

$$|\tilde{y}_k - \tilde{z}_k| \pmod{\delta} < \varepsilon/2.$$

Afterwards, by iteration of procedures $\mathbf{a}$ and $\mathbf{b}$, which does not change the relative

residuals we archive that

$$|y_1 - z_1| \pmod{\delta} < \varepsilon/2$$

which completes the proof of the lemma.

The combination of **a**, **b**, and **c** increases the given coordinate by the value $\gamma = \frac{\alpha}{N} - 1 \pmod{\delta}$. The coordinate that will be increased is determined in the procedure **b** by choosing an appropriate cube Q_i. By the property i) $\gamma \in \mathbb{R} \setminus \mathbb{Q}$ and γ is rationally independent on δ. Let us fix the $k \leq N$. By the iteration of the procedures **a**, **b** and **c** we obtain for any initial $\tilde{x}_k$ and any $l \in \mathbb{N}$, $V_k\left(\pi_X\left(T^{M'}(a, \mathbf{x})\right)\right) = \hat{x}_k + l\gamma$. Therefore the problem is reduced to the following task: find $L \in \mathbb{N}$ such that for any $\tilde{x}_k$ there exists $l < L$ such that

$$|\tilde{x}_k + l\gamma - \tilde{z}_k| \pmod{\delta} < \varepsilon/2. \tag{7.27}$$

We divide the left side of (7.27) by δ and denote $q = (\tilde{x}_k - \hat{z}_k)/\delta$ and $\phi = \gamma/\delta$. By the property i) $\phi \in \mathbb{R} \setminus \mathbb{Q}$. Then the inequality (7.27) takes the form $\|q + l\phi\| < \varepsilon/2$, where $\|\cdot\|$ is defined by (7.19).

To find a bound on l, which is uniform for all $q \in \mathbb{R}$ it is enough to find bound on l for $q \in [0, 1]$. There exists a set of $K = \left[\frac{4}{\varepsilon}\right] + 2$ rational numbers $0 < q_1 < \ldots < q_K < 1$ such as for any $j < K$, $q_{j+1} - q_j < \varepsilon/2$ and $|1 - q_K + q_1| < \varepsilon/2$. By the Minkowski theorem 7.3.4 for any $j \leq K$ there exists $l_j \in \mathbb{N}$ such that $\|q_j + l_j\phi\| < \varepsilon/2$. Let us select $L(\varepsilon, \alpha, \delta) = \max_{j \leq K} l_j$. For any $q \in [0, 1]$ we can find a q_j such that $\|q_j - q\| < \varepsilon/4$ and then

$$\|q + l_j\phi\| < \|q_j + l_j\phi\| + \|q_j - q\| < \varepsilon/2.$$

The constant L depends only on δ, α and ε and does not depend on the specific $\mathbf{z}$ and $\mathbf{x}$. Therefore to approximate each $\tilde{z}_k$ one needs at maximum $(N-1)L$ consequent applications of **a**, **b** and **c** which results in a sequence a^1 of length $M_1 \leq 3NL\left(\left\lfloor\frac{1}{\delta}\right\rfloor + 1\right)$, such that for any $k \leq N$ and any $a \in \mathcal{A}$

$$\left|V_k\left(\pi_X\left(T^{M_1}(a^1 a, \mathbf{x})\right)\right) - \tilde{z}_k\right| < \varepsilon/2 \pmod{\delta}. \tag{7.28}$$

Thus we conclude the first part of the proof.

Now, by the compostion of **a** and **b** we find a sequence a^2 of length M_2 such that for any $a \in \mathcal{A}$

$$\left|\pi_X\left(T^{M_1+M_2}(a^1 a^2 a, \mathbf{x})\right)_1 - z_1\right| < \varepsilon/2 \pmod{\delta}. \tag{7.29}$$

The combination of the procedures **a** and **b** keeps $\tilde{\mathbf{x}}$ fixed $\pmod{\delta}$, but changes the residuals of each coordinate by $\gamma = \alpha - 1 \pmod{\delta}$ and γ is rationally independent of δ. Thus we can precisely repeat the procedure described above, with the difference

that we combine now only **a** and **b** and not **a**, **b**, and **c**.

Analogously to the previous reasoning, one needs not more than L iterations, and therefore $M_2 < 2L\left(\left\lfloor\frac{1}{\delta}\right\rfloor + 1\right)$. Hence we found the desired sequence $a(\mathbf{x}, \mathbf{z}) = (a_i)_{i=1}^{M} = a^1 a^2$ and $M = M_1 + M_2$

$$M < 3NL(\delta, \alpha, \varepsilon)\left(\left\lfloor\frac{1}{\delta}\right\rfloor + 1\right) + 2L(\delta, \alpha, \varepsilon)\left(\left\lfloor\frac{1}{\delta}\right\rfloor + 1\right),$$

which is independent on $\mathbf{x}$ and $\mathbf{z}$

□

For any $\mathbf{x}, \mathbf{y} \in D$, we say that $\mathbf{x} > \mathbf{y}$ if $x_i > y_i$ for all $i \leq N$. Thus we defined a partial order on D. Consider $\mu \in Cl(D)$, where $Cl(D)$ is a closure of the set D,

$$\mu = \left(\alpha\frac{1}{N}, \alpha\frac{2}{N}, \ldots, \alpha\right) \tag{7.30}$$

and let $(\mu_i)_{i=1}^{N!}$ be the set of all vectors composed by permutations of coordinates of μ.

Proposition 7.3.5. *For any $\boldsymbol{x} \in D$ there exists $k \leq N!$ such that $\boldsymbol{x} > \mu_k$ and there are no $\boldsymbol{y} \in D$ such that there exists an $l \leq N! : \boldsymbol{y} \leq \mu_l$.*

Proof. Let S be a permutation which orders the coordinates of $\mathbf{x}$, i.e. $x_{S(1)} \leq x_{S(2)} \leq \ldots \leq x_{S(N)}$ and let $\mathbf{z} = (x_{S(1)}, x_{S(2)}, \ldots, x_{S(N)})$. We have $\mathbf{x} \in D$, hence $\mathbf{z} \in D$. It is enough to prove that $\mathbf{z} > \mu$. If it is not true then there exists k such that $z_k \leq \mu_k = \alpha\frac{k}{N}$. Thus, for any $i \leq k$, $z_i \leq \alpha\frac{k}{N}$ which mean that $\mathbf{z} \in C_k$ and therefore $\mathbf{z} \notin D$. Contradiction.

If there exists $\mathbf{y} \in D$ such that $\mathbf{y} \leq \mu_l$ we can analogously order the coordinates of $\mathbf{y}$ and μ_l and come to the same contradiction. □

Lemma 7.3.6. *For any point $\boldsymbol{x} \in D$ and any $l \leq N!$ there exists a sequence $a(\boldsymbol{x}) = (a_i)_{i=1}^{M(N,\delta)}$ such that for any $a \in \mathcal{A}$, $\boldsymbol{y} = \pi_X\left(T^{M(N,\delta)}(a(\boldsymbol{x})a, \boldsymbol{x})\right) < \mu_l + \delta e$. Where $e = (1, \ldots, 1) \in \mathbb{R}^N$ and $M(N, \delta)$ does not depend on $\boldsymbol{x}$ or l.*

Proof. Let $\mathbf{x} \in Q_1$. If $\mathbf{x} \in D \setminus Q_1$ then by the procedure described in the Lemma 7.3.3 we can find a sequence a^1 of the length of M_1, such that for any $a \in \mathcal{A}$: $\pi_X\left(T^{M_1}(a^1 a, \mathbf{x})\right) \in Q_1$.

In the following we show how to find a sequence a^2 such that $|T^{M_2}(a^2, \mathbf{x}) - \mu_1| < \delta$, where

$$\mu_1 = \left(\alpha, \frac{\alpha(N-1)}{N}, \frac{\alpha(N-2)}{N}, \ldots, \frac{\alpha}{N}\right). \tag{7.31}$$

μ_1 is one of the possible permutations $(\mu_i)_{i=1}^{N!}$. For all others permutations the procedure is completely analogous up to the order of coordinates.

Notation.

1. In the following, we consider all coordinates with a precision δ, i.e. when we write $\mathbf{x} \simeq (r_1, r_2, \ldots, r_N)$, it means that for any $i \leq N$, $x_i \in (r_i - \delta, r_i]$.

2. Denote $a_{i,\delta}$ the sequence $\{i, \ldots, i\}$ of a length L such that $x_i + (L-1)\delta < 1$, $x_i + L\delta \geq 1$.

3. We write that the sequence $a = (a_j)_{j=1}^{R}$ is *applied* to $\mathbf{x}$ and that $\mathbf{y}$ is a result of the application if $\mathbf{y} = \pi_X\left(T^R(a, \mathbf{x})\right)$.

Now we prove by induction that for any $k \leq N$ there exists a finite sequence a^k such that for $\mathbf{y}^k = \pi_x\left(T^{M_k}(a^k, \mathbf{x})\right)$ the following condition holds:

$$\mathbf{y}^k = \left(\frac{\alpha k}{N} + \delta, \frac{\alpha(k-1)}{N} + \delta, \ldots, \frac{\alpha}{N} + \delta, \frac{\alpha(N+1)}{N}, \ldots, \frac{\alpha(N+1)}{N}\right)$$

First, let us prove the base for $k = 1$, $k = 2$. Initially, $\mathbf{x} \in Q_1$, hence

$$\mathbf{x} \simeq (\alpha + \delta, \alpha, \ldots, \alpha).$$

Applying $a_{1,\delta}$ to $\mathbf{x}$ we get as in Eq. (7.25)

$$\mathbf{x}_1 = \pi_X\left(T\left(\sigma^{L-1}a_{1,\delta}a, T^{L-1}\left(a_{1,\delta}, \mathbf{x}\right)\right)\right) = \pi_X\left(F \circ G_1\left(T^{L-1}\left(a_{1,\delta}, \mathbf{x}\right)\right)\right),$$

where L is a length of $a_{1,\delta}$.

$$\mathbf{x}_1 \simeq \left(\frac{\alpha}{N} + \delta, \frac{\alpha(N+1)}{N}, \ldots, \frac{\alpha(N+1)}{N}\right).$$

Applying $a_{2,\delta}$ to $\mathbf{x}_1$ we get

$$\mathbf{x}_2 \simeq \left(\alpha + \delta, \frac{\alpha(N-1)}{N} + \delta, \ldots, \frac{\alpha(N-1)}{N}\right).$$

Applying again $a_{1,\delta}$

$$\mathbf{x}_3 \simeq \left(\frac{\alpha}{N} + \delta, \alpha + \delta, \alpha, \ldots, \alpha\right),$$

and then $a_{2,\delta}$ to $\mathbf{x}_3$ we obtain

$$\mathbf{x}_4 \simeq \left(\frac{2\alpha}{N} + \delta, \frac{\alpha}{N} + \delta, \frac{\alpha(N+1)}{N}, \ldots, \frac{\alpha(N+1)}{N}\right).$$

Now we prove the inductive step: for any $k < N$ if

$$\mathbf{z} \simeq \left(\overbrace{\frac{\alpha k}{N} + \delta, \frac{\alpha(k-1)}{N} + \delta, \ldots, \frac{\alpha}{N} + \delta}^{k \text{ coordinates}}, \overbrace{\frac{\alpha(N+1)}{N}, \ldots, \frac{\alpha(N+1)}{N}}^{N-k \text{ coordinates}} \right),$$

then by a finite number of operations of the form $a_{i,\delta}$ one can get

$$\mathbf{y} \simeq \left(\overbrace{\frac{\alpha(k+1)}{N} + \delta, \frac{\alpha k}{N} + \delta, \ldots, \frac{\alpha}{N} + \delta}^{k+1 \text{ coordinates}}, \overbrace{\frac{\alpha(N+1)}{N}, \ldots, \frac{\alpha(N+1)}{N}}^{N-k-1 \text{ coordinates}} \right).$$

To do so, let us first apply $a_{k+1,\delta}$

$$\mathbf{z}^2 = \left(\alpha + \delta, \frac{\alpha(N-1)}{N} + \delta, \ldots, \frac{\alpha(N-k+1)}{N} + \delta, \frac{\alpha(N-k)}{N}, \ldots, \frac{\alpha(N-k)}{N} \right),$$

and then consequently apply $a_{1,\delta}$,

$$\mathbf{z}^3 = \left(\overbrace{\frac{\alpha}{N} + \delta, \alpha + \delta, \frac{\alpha(N-1)}{N} + \delta, \ldots, \frac{\alpha(N-k)}{N} + \delta}^{k \text{ coordinates}}, \overbrace{\frac{\alpha(N-k+1)}{N}, \ldots, \frac{\alpha(N-k+1)}{N}}^{N-k \text{ coordinates}} \right).$$

Then $a_{2,\delta}$

$$\mathbf{z}^3 = \left(\overbrace{\frac{2\alpha}{N} + \delta, \frac{\alpha}{N} + \delta, \alpha + \delta, \ldots, \frac{\alpha(N-k+1)}{N} + \delta}^{k \text{ coordinates}}, \overbrace{\frac{\alpha(N-k+2)}{N}, \ldots, \frac{\alpha(N-k+2)}{N}}^{N-k \text{ coordinates}} \right),$$

and so forth until $a_{k,\delta}$, resulting in

$$\mathbf{z}^{k+1} = \left(\overbrace{\frac{\alpha k}{N} + \delta, \frac{\alpha(k+1)}{N} + \delta, \ldots, \frac{\alpha}{N} + \delta}^{k \text{ coordinates}}, \overbrace{\alpha, \ldots, \alpha}^{N-k \text{ coordinates}} \right).$$

By the application of $a_{k+1,\delta}$ on $\mathbf{z}^{k+1}$ we obtain $\mathbf{y}^{k+1}$.

In total there are at maximum $k+2$ iterations of $a_{i,\delta}$ needed for each induction step, which results in a sequence of maximum $(k+2)\left(\left[\frac{1}{\delta}\right]+1\right)$ symbols. Through

N iteration of such procedure the final sequence $a^f = (a_i)_{i=1}^{M(N,\delta)}$ is acquired. It has

$$M(N,\delta) \leq \sum_{k=1}^{N} (k+2)\left(\left[\frac{1}{\delta}\right]+1\right) + N_1 \tag{7.32}$$

symbols, where N_1 is the number of symbols required to transport any $\mathbf{x}^0 \in D$ to the cube (procedures **a.** and **b.**) Q_1, $N_1 \leq 2N\left(\left[\frac{1}{\delta}\right]+1\right)$. □

Lemma 7.3.7. *For any points $\boldsymbol{x}, \boldsymbol{z} \in D$ and any $\varepsilon > 0$ there exists a finite sequence $a^1 = (a_i)_{i=1}^{M}$ such that for any $a \in \mathcal{A}$ Euclidean distance $\rho\left(\boldsymbol{z}, \pi_X(T^M(a^1 a, \boldsymbol{x}))\right) < \varepsilon$ and $M = M(N,\varepsilon,\alpha,\delta)$ does not depend on $\boldsymbol{x}$ and $\boldsymbol{z}$.*

Proof. By the Prep. 7.3.5 there exists a $j < N!$ such that $\mathbf{z} > \mu_j$. We can assume that $\mathbf{x} \in Q_\delta$ (otherwise we apply to $\mathbf{x}$ the procedure **a** from the Lemma 7.3.3). As it is shown in the Lemma 7.3.6 from any point $\mathbf{x} \in D$ it is possible to reach δ-neighborhood of μ_j by finite, independent on $\mathbf{x}$ sequence a^1 of length M^1.

In the Lemma 7.3.6 the sequence a^1 to reach the δ-neighborhood of μ_j is found explicitly and it does not depend on the precise choice of $\mathbf{x} \in Q_\delta$. Consequently we can find a vector $\mathbf{d} = (d_1, \ldots, d_N)$ such that

$$\mathbf{x} - \pi_X\left(T^{M^1}(a^1, \mathbf{x})\right) \equiv \mathbf{d} \pmod{\delta}. \tag{7.33}$$

As it is proved in the Lemma 7.3.3, for any points $\mathbf{x}, \mathbf{z}' \in D$ it is possible to reach the ε-neighborhood of $\mathbf{z}'$ (mod δ) in the finite number of steps, i.e. there exists a sequence a^2 of the length M^2 such that for any $k \leq N$

$$\left|\pi_X(T^{M^2}(a^2, \mathbf{x}))_k - z'_k\right| < \varepsilon \pmod{\delta}. \tag{7.34}$$

Let us construct such a sequence for the point $\mathbf{z}' = \mathbf{z} + \mathbf{d}$. Then

$$\pi_X\left(T^{M^2+M^1}(a^2a^1, \mathbf{x})\right) = \pi_X\left(T^{M^1}(a^1, T^{M^2}(a^2a^1, \mathbf{x}))\right) \equiv T^{M^2}(a^2a^1, \mathbf{x}) - \mathbf{d} \pmod{\delta}.$$

Thus by Eq. 7.34

$$\left|\pi_X\left(T^{M^2+M^1}(a^2a^1, \mathbf{x})\right) - \mathbf{z}\right| < \varepsilon \pmod{\delta}.$$

If $\mathbf{z}$ is in the δ-neighborhood of μ_j then after applying both sequences a^1 and a^2 the image of $\mathbf{x}$ is in the ε-neighborhood of $\mathbf{z}$. Otherwise, because $\mathbf{z} < \mu_j$ we can find the sequence a^3 of length $M^3 < N\left(\left[\frac{1}{\delta}\right]+1\right)$ which does not change the residuals and makes the difference between the image of $\mathbf{x}$ and $\mathbf{z}$ smaller than ε. □

Now we have all the necessary tools to prove the Theorem 7.3.2.

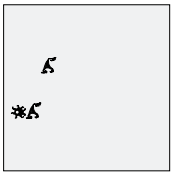
Proof of the theorem 7.3.2.
We have to prove that for almost all $a \in A$ and for any open subsets of D: $U, V \subset D$, there exists an $n \in N$ such that $\pi_X\left(T^n(a, U)\right) \cap V \neq \emptyset$. Let us fix any point $\mathbf{x} \in U$. V is an open set, therefore there exists a cube $Q_V \subset V$ with sides parallel to the sides of the unit cube. Let $\mathbf{z}$ be the center of the cube Q_V and ε be the half of its side length.

As it is proved in the Lemma 7.3.7, for any $\mathbf{y} \in D$ there exists a finite sequence $a^1 = (a_i)_{i=1}^L$ such that for any $k \leq N$, $|\pi_X\left(T^M(a^1, \mathbf{y})\right)_k - \mathbf{z}_k| < \varepsilon$ and hence $\pi_X\left(T^M(a^1, \mathbf{y})\right) \in Q_V$ and $L < M(N, \varepsilon, \delta, \alpha)$.

Thus for any sequence $a_1, \ldots, a_j$ there exists a sequence $b_1, \ldots, b_L, \ldots$, such that $\pi_X\left(T^{j+M}(ab, \mathbf{x})\right) \in Q_V$, where $ab = (a_1, \ldots, a_j, b_1, \ldots, b_L, \ldots)$ and $L < M(N, \varepsilon, \delta, \alpha)$.

Consider the sequence $a \in \mathcal{A}$. If $\mathcal{A}$ is endowed with the uniform Bernoulli measure, then with the probability one any finite subsequence will be found infinitely often in a. Therefore, for almost all $a \in \mathcal{A}$ the conditions of the theorem will hold true, and there exists an n such that $\pi_x\left(T^n(a, U)\right) \cap V \ni \pi_X\left(T^n(a, \mathbf{x})\right)$.

□



Chapter 8

Branching processes approximation

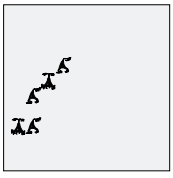
This chapter focuses on the description of the network dynamics during an avalanche. In physics, for such a description one uses a mean-field approximation [56, 76, 152], which regards each element of the system as a subject of a local field, instead of considering the full complexity of interactions.

For the SOC systems branching processes have served as a mean-field approximation already for twenty years. However, the accuracy of this approximation is still judged by the value of the critical exponent only and usually only in the limit of large networks. Thus, no studies have been conducted on the speed of convergence or any similarity between avalanches produced by branching processes and the system dynamics.

In the present chapter we are trying to fill this shortcoming by studying the example of the neural network proposed in chapter 3. We formulate the mathematical model of an avalanche and prove the convergence in distribution to a certain branching process. In chapter 4 we apply mathematical results presented in the current chapter to develop the long-term plasticity rule, which drives the network to the critical state from any initial condition.

8.1 Mathematical model of an avalanche

In the previous chapter the model of the network state and its dynamics is developed for all stages of its evolution. Here we want to concentrate only on the dynamics within an avalanche. To do so we consider only such $a \in \mathcal{A}$ and $\mathbf{x} \in D_N \subset [0,1)^N$, that $\pi_X(T(a,x)) = F^k \circ G_a(\mathbf{x})$ and $k(a,\mathbf{x}) > 0$. Let $X_s^N = [0,1)^N$. The phase space



is symmetric, therefore, without loss of generality $a_1 = N$ and $1 > x_N > 1 - \delta$. Denote by $\hat{D}_N$ the subset of D_N and $\overline{D}_{N-1} \subset X_s^{N-1}$ such as

$$\hat{D}_N = \{\mathbf{x} \in D \,|\, 1 > x_N > 1 - \delta\}, \tag{8.1}$$

$$\overline{D}_{N-1} = \left\{\mathbf{y} \in X_s^{N-1} \,|\, \forall x_N \in (1-\delta, 1), (y_1, \ldots, y_{N-1}, x_N) \in \hat{D}_N\right\}. \tag{8.2}$$

For the simplicity of notations, we consider the $(N+1)$-dimensional system, $D_{N+1} \subset X_s^{N+1}$, and $\overline{D}_N$ subspaces.

Suppose, that the distribution on $\overline{D}_N$ is induced by the distribution on D_N, and $\mathcal{A}$ is endowed with the uniform Bernoulli measure. As it is shown in the Theorem 7.3.1 the Lebesgue measure on D_{N+1} is the only $T-$invariant measure, thus $\mathbf{x}$ is uniformly distributed on D_{N+1}, hence $\mathbf{y} = (x_1, \ldots, x_N)$ is uniformly distributed on $\overline{D}_N$. Note, that $\overline{D}_N$ is not equal to a domain of the N-dimensional system D_N, but it is contained in D_N.

$$\overline{D}_N = X_s \setminus \bigcup_{i=1}^{N} C_i^N,$$

where

$$C_i^N = \left\{\mathbf{x} \in X_s \,\middle|\, \#\left\{j \le N \,\middle|\, x_j < \frac{i\alpha}{N+1}\right\} \ge i\right\} \subset \left\{\mathbf{x} \in X_s \,\middle|\, \#\left\{j \le N \,\middle|\, x_j < \frac{i\alpha}{N}\right\} \ge i\right\}.$$

We study the random variable describing the number of coordinates i such that $x_i > 1$ for each iteration of F. Namely, for any $l \le k(a, \mathbf{x})$ denote

$$\xi_l(\mathbf{x}) = \sum_{i=1}^{N} z_i(F^k(\mathbf{x})), \tag{8.3}$$

where $z_i(\mathbf{x}) = 1$ iff $x > 1$ and $z_i(\mathbf{x}) = 0$ othervise. If $l > k(a, \mathbf{x})$ then $\xi_l = 0$. The length of the avalanche is given then by $L = \sum_{l=1}^{k(a,\mathbf{x})} \xi_l(\mathbf{x})$.

For any $\mathbf{y} \in \overline{D}_N$, if $\alpha + \delta < 1$ then the sequence $(\xi_i)_{i=1}^{\infty}$ is fully defined and independent on the x_{N+1}. This follows from the Prop. 7.1.2, which states that if $\alpha + \delta < 1$ then each coordinate exceeds one at maximum ones.

Before considering $\mathbf{y}$ to be uniformly distributed on $\overline{D}_N$, we prove some results for $\mathbf{y}$ uniformly distributed in the unit cube $[0, 1)^N$ (which we address as *simplified phase-space*). We show that in such a setup the sequence $(\xi_i)_{i=1}^{\infty}$ can be approximated by the sequence generated by a certain class of stochastic processes, namely Galton-Watson branching processes. Afterwards, we prove that in the limit $N \to \infty$, the

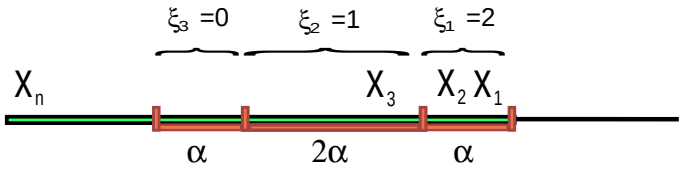


Figure 8.1: Example of one realization of the stochastic process $(\xi_{k,N})_{k=0}^{\infty}$. Coordinates of x are plotted in the descending order.

results hold also for $\mathbf{y}$ uniformly distributed in $\overline{D}_N$.

8.2 Simplified phase-space model

Let $N \in \mathbb{N}$, $\alpha \in (0,1]$, $\alpha_N = \frac{\alpha}{N}$ and $(x_{i,N})_{i=1}^{N}$ be i.i.d. uniformly distributed on $[0,1]$. We recursively construct the random sequence $(\xi_{k,N})_{k=0}^{\infty}$ as follows

$$\xi_{0,N} = 1, \quad \xi_{1,N} = \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N,1]}(x_{j,N}),$$
$$\xi_{k+1,N} = \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{k} \xi_{j,N},\, 1-\alpha_N \sum_{l=0}^{k-1} \xi_{l,N}]}(x_{j,N}). \tag{8.4}$$

It is easy to see, that $\xi_{k,N}$ defined here coincides with $\xi_k(\mathbf{x})$ defined by Eq. 8.3. In order to derive some properties of the large N limit of the sequence $(\xi_{k,N})_{k=0}^{\infty}$, we construct another sequence $\left(\xi^*_{k,N}\right)_{k=0}^{\infty}$ in a similar manner but arising from different initial random variables. Namely, let $(x^*_{i,N})_{i=1}^{N(\omega)}$ be instances of change of a Poisson point process with the intensity N on the interval $[0,1]$. $N(\omega)$ is a random variable representing the number of events in the interval, it has a Poisson distribution with rate N. We define $\xi^*_{i,N}$ to be generated from the sequence $\left(x^*_{i,N}\right)_{i=1}^{\infty}$ analogously to previous definition (8.4):

$$\xi^*_{0,N} = 1, \;\; \xi^*_{1,N} = \sum_{j=1}^{N(\omega)} \mathbb{I}_{[1-\alpha_N,1]}\left(x^*_{j,N}\right),$$
$$\xi^*_{k+1,N} = \sum_{j=1}^{N(\omega)} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{k} \xi^*_{j,N},\, 1-\alpha_N \sum_{l=0}^{k-1} \xi^*_{l,N}]}\left(x^*_{j,N}\right). \tag{8.5}$$

By definition of a Poisson point process, for any interval I

$$P\left(\sum_{j=1}^{N(\omega)} \mathbb{I}_I\left(x_{j,N}^*\right) = k\right) = \frac{(N|I|)^k e^{(N|I|)}}{k!} \tag{8.6}$$

and for any disjoint intervals I_1 and I_2

$$\sum_{j=1}^{N(\omega)} \mathbb{I}_{I_1}\left(x_{j,N}^*\right) \text{ and } \sum_{j=1}^{N(\omega)} \mathbb{I}_{I_2}\left(x_{j,N}^*\right) \text{ are independent.} \tag{8.7}$$

8.3 Branching processes

Definition 8.3.1. Let $Z_0, Z_1, Z_2 \ldots$ be a Markov process, Z_i takes values in $\mathbb{N}_0$, $Z_0 = 1$. Let $\mathcal{P}$ be the distribution of Z_1 ($P(Z_1 = i) = p_i$). Let $P(Z_{i+1} = 0|Z_i = 0) = 1$ and $(Z_{i+1}|Z_i = k)$ distributed as a sum of k independent random variables distributed as Z_1. Then $Z_0, Z_1, Z_2 \ldots$ is a Galton-Watson branching process.

Denote by p_i the probability $P(Z_1 = i)$, and let $f(z)$ be a generating function of a distribution of Z_1. Denote $\sum_{i=0}^{\infty} Z_i$ by L and let $F(l)$ be a generating function of a distribution of L, i.e.

$$f(z) = \sum_{\nu=0}^{\infty} p_\nu z^\nu, \tag{8.8}$$

$$F(l) = \sum_{n=1}^{\infty} P_n l^n, \qquad \text{where } P_n = P(L = n). \tag{8.9}$$

Let ρ and α be the radii of convergence of the power series on the right sides of (8.8) and (8.9) respectively. Since $f(1) = 1$ and $F(1) = P(L < \infty) < 1$ we know $\rho, \alpha \geq 1$.

Lemma 8.3.1. [66] *Let*

$$G(l, z) = lf(z) - z$$

then $z = F(l)$ is unique analytical solution of

$$G(l, z) = 0$$

in a certain neighborhood of $(0, 0)$.

Lemma 8.3.2. *Let $a = F(\alpha)$, then $\alpha < \rho$*

Theorem 8.3.3. [112] *Suppose that $p_0 > 0$ and let d be the largest integer such that $p_r \neq 0$ implies that r is a multiple of d, $r = 1, 2, \ldots$. If $r - 1$ is not divisible by d, then we have $P(L = r) = 0$, while if $r - 1$ is divisible by d, then*

$$P(L = r) = d \left(\frac{a}{2\pi\alpha f''(a)} \right)^{\frac{1}{2}} \alpha^{-r} r^{-\frac{3}{2}} + O(\alpha^{-r} r^{-\frac{5}{2}}),\ r \to \infty.$$

In the most interesting for applications to SOC case the expectation of Z_1 is equal to one (then the branching process is called critical). In such case $a = \alpha = 1$ are satisfying the conditions of the theorem and then

$$P(L = r) = d \left(\frac{1}{2\pi f''(a)} \right)^{\frac{1}{2}} r^{-\frac{3}{2}} + O(r^{-\frac{5}{2}}),\ r \to \infty.$$

Lemma 8.3.4. *$(\xi^*_{i,N})_{i=1}^{\infty}$ is a Galton-Watson branching process.*

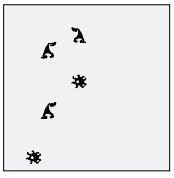
Proof. 1. First we should prove that $(\xi^*_{i,N})_{i=1}^{\infty}$ is a Markov chain. Using (8.6) and (8.7) we can compute:

$$\begin{aligned}
&P(\xi^*_{m+1,N} = k_{m+1} | \xi^*_{m,N} = k_m, \ldots, \xi^*_{1,N} = k_1) = \\
&= P\left(\sum_{j=1}^{N(\omega)} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{m} \xi^*_{j,N}, 1-\alpha_N \sum_{l=0}^{m-1} \xi^*_{l,N}]} \left(x^*_{j,N} \right) = k_{m+1} | \xi^*_{m,N} = k_m, \ldots, \xi^*_{1,N} = k_1 \right) \\
&= P\left(\sum_{j=1}^{N(\omega)} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{m} k_j, 1-\alpha_N \sum_{l=0}^{m-1} k_l]} \left(x^*_{j,N} \right) = k_{m+1} \right) = \frac{e^{-\alpha_N k_m} (\alpha k_m)^{k_{m+1}}}{k_{m+1}!} \\
&= P(\xi^*_{m+1,N} = k_{m+1} | \xi^*_{m,N} = k_m).
\end{aligned}$$

2. It remains to prove the "branching-property", namely, that $(\xi^*_{m+1,N} | \xi^*_{m,N} = k_m)$ is distributed as a sum of k_m independent copies of $\xi^*_{1,N}$

$$\begin{aligned}
P(\xi^*_{m+1,N} = k_{m+1} | \xi^*_{m,N} = k_m) &= P\left(\sum_{i=1}^{N(\omega)} \mathbb{I}_{[0, \alpha_N \cdot k_m]}(x^*_{i,N}) = k_{m+1} \right) \\
&= P\left(\sum_{j=1}^{m_k} \sum_{i=1}^{N(\omega)} \mathbb{I}_{[\alpha_N (j-1),\, \alpha_N \cdot j]}(x^*_{i,N}) = k_{m+1} \right).
\end{aligned}$$

By (8.6), $\sum_{i=1}^{N(\omega)} \mathbb{I}_{[\alpha_N (j-1),\, \alpha_N \cdot j]}(x^*_{i,N})$ has a Poisson distribution with the intensity α, and by the definition $\xi^*_{1,N} = \sum_{i=1}^{N(\omega)} \mathbb{I}_{[1-\alpha_N, 1]}(x^*_{i,N})$ has the same distribution. Also from (8.7) $\sum_{i=1}^{N(\omega)} \mathbb{I}_{[\alpha_N (j-1),\, \alpha_N \cdot j]}(x^*_{i,N})$ and $\sum_{i=1}^{N(\omega)} \mathbb{I}_{[\alpha_N (k-1),\, \alpha_N \cdot k]}(x^*_{i,N})$ are independent



for any $1 \leqslant i, j \leqslant N$ and $i \neq j$.

□

8.4 Approximation of the avalanche distribution

Lemma 8.4.1.

$$P(\xi_{1,N} = k_1, \ldots, \xi_{m,N} = k_m) = \binom{N}{k_1, \ldots, k_m} (1 - \alpha_N \sum_{i=0}^{m-1} k_i)^{N - \sum_{i=1}^{m} k_i} \prod_{i=1}^{m} (\alpha_N k_{i-1})^{k_i}.$$

Proof. The proof is by induction for m for arbitrary integers $k_1, \ldots, k_m$.
Base: $x_{i,N}$ is uniformly distributed on $[0, 1]$ hence $\xi_{1,N} = \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N, 1]}(x_{j,N})$ has a binomial distribution with success probability α_N and N trials, thus

$$P(\xi_{1,N} = k_1) = \binom{N}{k_1} \alpha_N^{k_1} (1 - \alpha_N)^{N - k_1}.$$

Inductive step: Let the inductive hypothesis holds for all $m \leq n$. We prove that it is also true for $m = n + 1$. Let $\xi_{n,N} = k_n, \ldots, \xi_{1,N} = k_1$, then by the definition of $\xi_{i,N}$ (Eq. 8.4):

$$\sum_{l=1}^{n} k_l = \sum_{r=1}^{n-1} \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{r} k_j, 1-\alpha_N \sum_{l=0}^{r-1} k_l]}(x_{j,N}) = \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{n-1} k_j, 1]}(x_{j,N})$$

Let

$$Y = \left\{ x_{i,N} \notin \left[1 - \alpha_N \sum_{l=0}^{n-1} k_j, 1 \right] \right\}, \tag{8.10}$$

then $\operatorname{card}(Y) = N - \sum_{l=0}^{n} k_j$. By definition,

$$\xi_{m+1,N} = \sum_{j=1}^{N} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{m} \xi_{j,N}, 1-\alpha_N \sum_{l=0}^{k-1} \xi_{l,N}]}(x_{j,N}) = \sum_{y \in Y} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{m} k_j, 1-\alpha_N \sum_{l=0}^{m-1} k_l]}(y).$$

For all $y \in Y$

$$P\left[\mathbb{I}_{[1-\alpha_N \sum_{l=0}^{n} k_j, 1-\alpha_N \sum_{l=0}^{n-1} k_l]}(y) = 1 \,\middle|\, \xi_{n,N} = k_n, \ldots, \xi_{1,N} = k_1 \right] = \frac{\alpha_N k_n}{1 - \alpha_N \sum_{l=0}^{n-1} k_j}.$$

Hence $\sum_{y \in Y} \mathbb{I}_{[1-\alpha_N \sum_{l=0}^{n} k_j, 1-\alpha_N \sum_{l=0}^{n-1}]}(y)$ has a binomial distribution with the success

probability $\alpha_N k_n/(1-\alpha_N\sum_{l=0}^{n-1}k_l)$ and $N-\sum_{l=0}^{n}k_j$ trials. Consequently,

$$\begin{aligned}
&P(\xi_{n+1,N}=k_{n+1}\mid \xi_{n,N}=k_n,\dots,\xi_{1,N}=k_1)=\\
&\quad=P\left(\sum_{j=1}^{N}\mathbb{I}_{[1-\alpha_N\sum_{l=0}^{n}k_j,1-\alpha_N\sum_{l=0}^{n-1}k_l]}(x_{j,N})\,\middle|\,\xi_{n,N}=k_n,\dots,\xi_{1,N}=k_1\right)\\
&\quad=P\left(\sum_{y\in Y}\mathbb{I}_{[1-\alpha_N\sum_{l=0}^{n}k_j,1-\alpha_N\sum_{l=0}^{n-1}k_l]}(y)\,\middle|\,\xi_{n,N}=k_n,\dots,\xi_{1,N}=k_1\right)\\
&\quad=\binom{N-\sum_{i=1}^{n}k_i}{k_{n+1}}\left(\frac{\alpha_N k_n}{1-\alpha_N\sum_{i=1}^{n}k_i}\right)^{k_{n+1}}\left(\frac{1-\alpha_N\sum_{i=1}^{n+1}k_i}{1-\alpha_N\sum_{i=1}^{n}k_i}\right)^{N-\sum_{i=1}^{n+1}k_i}.
\end{aligned}$$

By the induction hypothesis

$$P(\xi_{n,N}=k_n,\dots,\xi_{1,N}=k_1)=\binom{N}{k_1,\dots,k_n}\left(1-\alpha_N\sum_{i=0}^{n-1}k_i\right)^{N-\sum_{i=1}^{n}k_i}\prod_{i=1}^{n}(\alpha_N k_{i-1})^{k_i}.$$

Hence

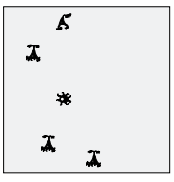
$$\begin{aligned}
&P(\xi_{n+1,N}=k_{n+1},\ \xi_{n,N}=k_n,\dots,\xi_{1,N}=k_1)=\\
&\quad=\binom{N-\sum_{i=1}^{n}k_i}{k_{n+1}}\binom{N}{k_1,\dots,k_m}\left(\frac{\alpha_N k_n}{1-\alpha_N\sum_{i=1}^{n-1}k_i}\right)^{k_{n+1}}\times\\
&\qquad\left(\frac{1-\alpha_N\sum_{i=1}^{n}k_i}{1-\alpha_N\sum_{i=1}^{n-1}k_i}\right)^{N-\sum_{i=1}^{n+1}k_i}\left(1-\alpha_N\sum_{i=0}^{n-1}k_i\right)^{N-\sum_{i=1}^{n}k_i}\prod_{i=1}^{n}(\alpha_n k_{i-1})^{k_i}\\
&\quad=\binom{N}{k_1,\dots,k_m,k_{n+1}}\left(1-\alpha_N\sum_{i=0}^{n}k_i\right)^{N-\sum_{i=1}^{n+1}k_i}\prod_{i=1}^{n+1}(\alpha_N k_{i-1})^{k_i}.
\end{aligned}$$

□

Thus we have also computed the probability to obtain a certain sequence $(\xi_{k,N})_{k=0}^{\infty}$ with a finite number of non-zero members:

$$\begin{aligned}
P^N_{k_1,\dots,k_m}&:=P(\xi_{1,N}=k_1,\dots,\xi_{m,N}=k_m,\xi_{m+1,N}=0)\\
&=\binom{N}{k_1,\dots,k_m}\left(1-\alpha_N\sum_{i=0}^{m}k_i\right)^{N-\sum_{i=1}^{m}k_i}\prod_{i=1}^{m}(\alpha_N k_{i-1})^{k_i}.
\end{aligned}$$

For a sequence $(\xi^*_{k,N})_{k=0}^{k=\infty}$ such probability can be easily computed analogously to



the Lemma 8.4.1.

$$P^{*,N}_{k_1,\dots,k_m} := P(\xi^*_{1,N} = k_1, \dots, \xi^*_{m,N} = k_m, \xi^*_{m+1,N} = 0) = \prod_{i=1}^{m+1} \frac{e^{-\alpha k_{i-1}} (\alpha k_{i-1})^{k_i}}{k_i!}.$$

Let us denote

$$k^0_m := \sum_{i=0}^{m} k_i,\ k^1_m := \sum_{i=1}^{m} k_i.$$

Lemma 8.4.2.

$$P^N_{k_1,\dots,k_m} - P^{*,N}_{k_1,\dots,k_m} = \frac{\prod_{i=1}^m (\alpha k_{i-1})^{k_i}}{k_1! \dots k_m!} e^{-\alpha k^0_m} \left(\frac{2{k^1_m}^2 - k^1_m}{2N} + O\left(\frac{{k^1_m}^4}{N^2} \right) \right) \tag{8.11}$$

Proof. We will use the following expansions:
First,

$$\begin{aligned} \frac{N!}{(N-k)!} &= N^k (1 - \frac{1}{N}) \dots (1 - \frac{k-1}{N}) = N^k \exp\left(\sum_{i=1}^{k-1} \ln\left(1 - \frac{i}{N}\right) \right) \\ &= N^k e^{-\frac{k(k-1)}{2N}} \exp\left(\sum_{i=1}^{k-1} \left(\frac{i}{N} + \ln\left(1 - \frac{i}{N}\right) \right) \right) \\ &= N^k e^{-\frac{k(k-1)}{2N}} \left(1 + O\left(\frac{k^3}{N^2} \right) \right). \end{aligned}$$

Second,

$$\begin{aligned} (N - k + 1) \ln\left(1 - \frac{k}{N}\right) &= (N - k + 1) \left(-\frac{k}{N} + \frac{1}{2}\frac{k^2}{N^2} + O\left(\frac{k^3}{N^3}\right) \right) \\ &= \left(-k + \frac{3k^2 - 2k}{2N} + O\left(\frac{k^3}{N^2}\right) \right) \end{aligned}$$

Third,

$$\left(1 - \frac{k}{N}\right)^{N-k+1} = e^{-k} \left(1 + \frac{3k^2 - 2k}{2N} + O\left(\frac{k^4}{N^2}\right) \right).$$

Thus,

$$\begin{aligned} P^N_{k_1,\dots,k_m} - P^{*,N}_{k_1,\dots,k_m} &= \frac{\prod_{i=1}^m (\alpha_N k_{i-1})^{k_i}}{k_1! \dots k_m!} \left(\frac{N!}{(N - k^1_m)!} (1 - \alpha_N k^0_m)^{N - k^1_m} - e^{-\alpha k^0_m} N^{k^1_m} \right) \\ &= \frac{\prod_{i=1}^m (\frac{\alpha}{N} k_{i-1})^{k_i}}{k_1! \dots k_m!} e^{-\alpha k^0_m} N^{k^1_m} \left(e^{-\frac{k^1_m (k^1_m - 1)}{2N}} \left(1 + O\left(\frac{{k^1_m}^3}{N^2} \right) \right) \right. \end{aligned}$$

$$\left(1+\frac{3{k_m^1}^2-2k_m^1}{2N}+O\left(\frac{{k_m^1}^4}{N^2}\right)\right)-1\Bigg)=$$
$$=\frac{\prod_{i=1}^m(\alpha k_{i-1})^{k_i}}{k_1!\dots k_m!}e^{-\alpha k_m^0}\left(\frac{2{k_m^1}^2-k_m^1}{2N}+O\left(\frac{{k_m^1}^4}{N^2}\right)\right)$$

□

Remark 8.4.3.

$$\left|P\left(\sum_{i=1}^\infty \xi_{i,N}=r\right)-P\left(\sum_{i=1}^\infty \xi^*_{i,N}=r\right)\right|<2^{r-1}C(r)\left|\left(\frac{2r^2-r}{2N}+O\left(\frac{r^4}{N^2}\right)\right)\right|, \tag{8.12}$$

here $C(r)=\max_{k_1+..+k_m=r,k_1>0\dots k_m>0}\frac{\prod_{i=1}^m(\alpha k_{i-1})^{k_i}}{k_1!\dots k_m!}e^{-\alpha k_m^0}$ *depends only on* r.

Proof. Let

$$K_r=\left\{\kappa=(k_i)_{i=1}^\infty \;\middle|\; \sum_{i=1}^\infty k_i=r,\ a_i\geq 0\ \forall i\in\mathbb{N},\ \text{if } k_i=0 \text{ then } k_j=0\ \forall j>i\right\},$$

then we have

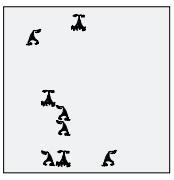
$$\operatorname{card}(K_r)=2^{r-1}.$$

$$\begin{aligned}\left|P\left(\sum_{i=1}^\infty \xi_{i,N}=r\right)-P\left(\sum_{i=1}^\infty \xi^*_{i,N}=r\right)\right| &\leq \sum_{\kappa\in K_r}\left|P^N_{k_1,\dots,k_m}-P^{*,N}_{k_1,\dots,k_m}\right|\\ &\leq 2^{r-1}\max_{\kappa\in K_r}\left|P^N_{k_1,\dots,k_m}-P^{*,N}_{k_1,\dots,k_m}\right|\\ &\leq 2^{r-1}\max_{\kappa\in K_r}\frac{\prod_{i=1}^m(\alpha k_{i-1})^{k_i}}{k_1!\dots k_m!}e^{-\alpha k_m^0}\left|\left(\frac{2r^2-r}{2N}+O\left(\frac{r^4}{N^2}\right)\right)\right|.\end{aligned}$$

□

Theorem 8.4.4. *Let* $f(x)=e^{\alpha(x-1)}$, $a>0$ *in the interior of the circle of convergence of* f *for which* $f'(a)=f(a)/a$ *and* $\gamma=a/f(a)$

$$P\left(\sum_{i=1}^\infty \xi_{i,N}=r\right)=$$
$$=\left(\frac{a}{2\pi\gamma f''(a)}\right)^{\frac{1}{2}}\gamma^{-r}r^{-\frac{3}{2}}+O\left(\gamma^{-r}r^{-\frac{5}{2}}\right)+\left|2^{r-1}C(r)\left(\frac{2r^2-r}{2N}+O\left(\frac{r^4}{N^2}\right)\right)\right|.$$



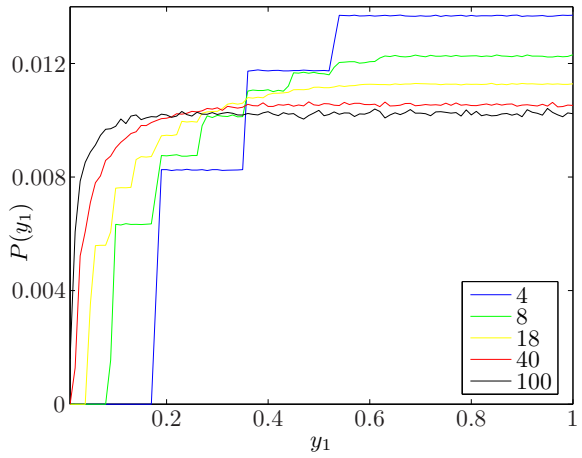


Figure 8.2: Empirical density of the y_1 for different dimensions N. All curves are obtained for $\alpha = 0.7$, $\delta = 0.02$

8.5 Full phase-space

In this section we show that each coordinate of $\mathbf{y} \in D_N$ converges in distribution to a uniform random variable on $[0, 1)$ in the limit of large system size $N \to \infty$. Due to S_N symmetry of the system, without loss of generality we consider only a distribution of the first coordinate of $\mathbf{y}$. In Fig. 8.2 we plot the empiric density of the y_1 for different system sizes.

Theorem 8.5.1. *Let y_1 be the first coordinate of the random vector $\boldsymbol{y} \in D_N$. Then y_1 converges in distribution to a random variable uniformly distributed on $[0, 1)$.*

Proposition 8.5.2. $\mathrm{Vol}(\overline{D}_N) = 1 - \alpha\frac{N}{N+1}$

Proof. Denote $\beta = \alpha\frac{N}{N+1}$, then $\frac{\beta}{N} = \frac{\alpha}{N+1}$. Using the results from [53]

$$\mathrm{Vol}(\overline{D}_N) = \mathrm{Vol}(D_N \text{ with parameter } \beta) = 1 - \beta.$$

□

Proof of the theorem 8.5.1. To understand the distribution of y_1 better, let us consider the probabilities $P(\mathbf{y}_1 \in I)$, where $I \subset [0, 1)$ is an interval.

$$\begin{aligned} P(\mathbf{y}_1 \in I) &= \frac{\mathrm{Vol}\left\{\mathbf{y} \in \overline{D}_N \mid y_1 \in I\right\}}{1 - \alpha\frac{N}{N+1}} \\ &= \frac{|I| - \mathrm{Vol}\left\{I \times [0,1)^{N-1} \setminus \bigcup_{i=1}^{N} \left(I \times [0,1)^{N-1} \cap C_i\right)\right\}}{1 - \alpha\frac{N}{N+1}}. \end{aligned}$$

Note, that we considered only such $\mathbf{x}$ that $x_{N+1} > 1-\delta > 1-\alpha$, therefore we have $I \times [0,1)^N \cap C_{N+1} = \emptyset$. Denote

$$\begin{aligned} C_i^1 &= \left\{ \mathbf{x} \in X_s \;\middle|\; \#\left\{ 1 < j \le N \;\middle|\; x_j < \alpha \frac{i}{N+1} \right\} \ge i \right\}, \\ C_i^2 &= \left\{ \mathbf{x} \in X_s \;\middle|\; \#\left\{ 1 < j \le N \;\middle|\; x_j < \alpha \frac{i}{N+1} \right\} = i-1, x_1 < \alpha \frac{i}{N+1} \right\}. \end{aligned}$$

$C_i^{1,2}$ is a partition of C_i, namely $C_i^1 \cup C_i^2 = C_i$ and $C_i^1 \cap C_i^2 = \emptyset$. For any $I \subset [0,1)$,

$$\mathrm{Vol}\left(C_i^1 \cap I \times [0,1)^{N-1}\right) = |I| \cdot \mathrm{Vol}\left(C_i^1\right), \tag{8.13}$$

and

$$\mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1 \cap I \times [0,1)^{N-1}\right) = |I| \cdot \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1\right).$$

Let us separate the following cases:
a. Let $I \subset \left(\alpha\frac{N}{N+1}, 1\right)$, then for any $i \le N+1$, $I \times [0,1)^{N-1} \cap C_i^2 = \emptyset$. Therefore from (8.13) we obtain

$$P(y_1 \in I) = |I| \frac{1 - \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1\right)}{1 - \alpha\frac{N}{N+1}},$$

and for any point $a \in (\alpha\frac{N}{N+1}, 1)$ it is possible to find a probability density function $f(a)$ of the random variable y_1 and

$$f^N(a) = \frac{1 - \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1\right)}{1 - \alpha\frac{N}{N+1}} = f_N^N. \tag{8.14}$$

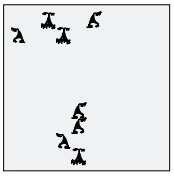
To find the volume of $\bigcup_{i=1}^N C_i^1$ we use, that it is equal to a volume of $\bigcup_{i=1}^N \overline{C}_i^1$, where

$$\overline{C}_i^1 = \left\{ \mathbf{x} \in [0,1)^{N-1} \,|\, \#\left\{ 1 \le j \le N-1 | x_j < \alpha \frac{i}{N+1} \right\} \ge i \right\}.$$

$\mathrm{Vol}\left(\bigcup_{i=1}^N \overline{C}_i^1\right)$ is equal to a volume of the C-set for $N-1$ dimensional system with a connectivity parameter $\gamma = \alpha\frac{N-1}{N+1}$. Therefore, by Prop. 8.5.2, we have

$$f_N^N = \frac{1 - \alpha\frac{N-1}{N+1}}{1 - \alpha\frac{N}{N+1}}. \tag{8.15}$$

b. Let $I \subset \left(\alpha\frac{N-1}{N+1}, \alpha\frac{N}{N+1}\right)$, then for any $i \le N$, $I \times [0,1)^{N-1} \cap C_i^2 = \emptyset$ and



$\mathrm{Vol}\left(I \times [0,1)^{N-1} \cap C_N^2\right) = |I| \cdot \mathrm{Vol}\,(C_N^2)$.

$$P(\mathbf{y}_1 \in I) = |I| \frac{1 - \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1\right) - \mathrm{Vol}\,(C_N^2) + \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1 \cap C_N^2\right)}{1 - \alpha\frac{N}{N+1}},$$

and for any point $a \in \left(\alpha\frac{N-1}{N+1}, \alpha\frac{N}{N+1}\right)$ the probability density $f^N(a)$ is defined and

$$f^N(a) = \frac{1 - \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1\right) - \mathrm{Vol}\,(C_N^2) + \mathrm{Vol}\left(\bigcup_{i=1}^N C_i^1 \cap C_N^2\right)}{1 - \alpha\frac{N}{N+1}} = f_{N-1}^N. \quad (8.16)$$

c. Analogously to **a** and **b** one can find the probability density for all points except the finite set $(a_i)_{i=1}^N$, $a_i = \alpha\frac{i}{N+1}$. The density is piecewise constant, for $k < N$ and for any $a : \alpha\frac{k}{N+1} < a < \alpha\frac{k+1}{N+1}$ we have $f(a) = f_k$. If $a > \alpha\frac{N}{N+1}$, $f^N(a) = f_N^N$. From the derivation procedure it is obvious that $f_N \geq f_{N-1} \geq \ldots \geq f_1$. Let us consider the limit density $f(a) = \lim_{N\to\infty} f^N(a)$. For any a

$$f(a) \leq \lim_{N\to\infty} f_N^N = \lim_{N\to\infty} \frac{1 - \alpha\frac{N-1}{N+1}}{1 - \alpha\frac{N}{N+1}} = 1. \quad (8.17)$$

However, by the dominated convergence theorem $\int_0^1 f(a)da = 1$, hence $f(a) = 1$ for almost all $a \in [0,1)$. Thus in the limit we obtain a uniform distribution on $(0,1]$.

□

Chapter 9

Conclusions

In this thesis, we studied different aspects of self-organized criticality (SOC). We were mainly interested in the following two issues:

- Appropriate modeling of self-organized criticality in neural systems. This topic became particularly interesting, as there is strong evidence [14, 129] for SOC in neuronal activity, recorded from the cortical slices.
- Mathematical modeling and study of SOC. This field is still rather untouched, which makes it even more attractive.

We started by introducing a simple model of a fully connected neural network in chapter 2, which serves as a basis for building up interesting extensions in the following chapters. This model has many merits: it is critical for some value of the connectivity parameter, it comprises larger similarity with neuronal systems than other SOC models, and last but not least, it is simple enough to allow for analytical treatment [53]. However, it also has a significant drawback: to obtain the critical state, one has to precisely tune the connectivity parameter. We introduced the results, which have been discussed in [53], and then presented our new findings on this model: we analytically derived the mean avalanche size and used this fact to find a finite-size scaling of the critical distribution.

We developed a mathematical model of this network in the form of the stochastic dynamical system $T(a, \mathbf{x})$, where $\mathbf{x} \in X_s = [0, 1)^N$ is a vector of membrane potentials of all N neurons in the network. The set $\mathcal{A}$ consists of all one-sided infinite sequences, and an $a \in \mathcal{A}$ denotes the sequence of external inputs. We analytically defined a set D, so that for any initial condition $\mathbf{x} \in X_s$ after some number of iterations M: $T(a, \mathbf{x})^M \in D$. On the set $D \times \mathcal{A}$, we found an invariant measure, which appears to be simply a product of a Lesbegue measure on D and any shift-invariant measure on $\mathcal{A}$. It turned out to be unique with respect to the choice of the measure

on $\mathcal{A}$. Then we studied the mixing properties of the mapping T. We proved that if the set $\mathcal{A}$ is endowed with the uniform Bernoulli measure, then for almost all $a \in \mathcal{A}$, the projection of the mapping T on D is topologically transitive.

The knowledge of the invariant measure allows us to simplify the description of avalanches by leaving out the external input component. An avalanche is thus a random sequence $(\xi_m)_{m=1}^{\infty}$ of integers, where ξ_m denotes the number of active neurons at time m. The stochasticity in this sequence is introduced by the randomness in the initial membrane potential vector $\mathbf{x}$. We considered finite-dimensional distributions of the first k members of the sequence ξ. We proved that they converge in the distribution to the sequence generated by a branching process with specific branching distribution as the system size tends to infinity. We analytically defined the difference between the probabilities of particular sequences generated by the systems dynamics and the branching process. This allowed us to use results from the theory of branching processes [66]. We showed that if the avalanche size grows slowly in comparison to the system size, then the network can only be critical in the limit of large networks for the specific connection strength $\alpha = 1$. The corresponding critical distribution has necessarily the exponent equal to $-\frac{3}{2}$.

Using the mathematical observation, we defined an extension to the model by incorporating a homeostatic learning rule. The idea of this rule is based on the tradeoff between the reliability of the synaptic transmission and the minimization of the energy expenses of each neuron. As a solution for this trade-off, we proposed that each firing neuron leads to one other spike in the network on average. As a result of such a plasticity rule, the network converges to a critical state from any initial conditions. The learning rule works very successfully, because in its essence it changes the branching parameter in the direction towards critical branching. Moreover, we proved that the avalanche size distribution coincides with the distribution of the number of offsprings in the branching process [89].

We also considered a short term synaptic plasticity: depressing and facilitatory synapses. We were able to show that if the simple network is endowed with a dynamic synapses plasticity, then the critical neuronal avalanches turn from an exceptional phenomenon into a typical and robust self-organized critical behavior [91]. We revealed that the previously studied mathematical model serves as a mean-filed approximation of the network with dynamic synapses. It allowed us to prove that, in the limit of large networks, the critical state is achieved from any realistic initial conditions. We also investigated different modifications of the network, and showed that SOC is robust with respect to the structural changes of the system.

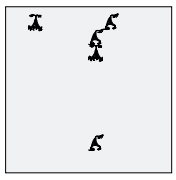
By modification of the dynamic synapses we obtained a network where the transition to criticality occurs by a first order phase transition [90]. The mean-field analysis provided an excellent theoretical prediction of the stationary state properties of the network. Extensive numerical studies were conducted to supplement the theoretical

treatment and have proved that approximations which are done to obtain analytical results are justified and do not change the essence of the dynamics.

Our study showed that it is possible to obtain a self-organized critical network with biologically realistic synaptic plasticity. On this basis, it is possible to construct more detailed models, which will capture also other properties of the local field potentials (LFP). Another interesting problem would be to find the explanation for the difference of the power-law exponents in spike recording and LFP recording [14, 13].

On the other hand, we created a basis to study SOC in neuronal network by the mathematically rigorous methods. There are many possibilities, by which to proceed. For example, the question of the ergodicity or other types of mixing (besides topological transitivity) in the avalanching dynamical system still remains unanswered. Preliminary numerical results are in favor of ergodicity, but no rigorous results are available so far.

All in all, we believe that this work is an important contribution to the understanding of critical phenomena in neural networks. We hope that it will motivate further research both in the study of neuronal models and in the mathematical approach to this exciting question.



Appendix A

A.1 Two methods to compute the average inter-spike interval

The two methods (5.15), (5.18) to compute $\left\langle \Delta^{\mathrm{ISI}} \right\rangle$ differ in the way to compute $\langle \kappa \rangle$. The first method comprises a more detailed way to perform the computations, which leads to the formula

$$\langle \kappa_1 \rangle = \frac{\theta}{\frac{u\langle J_{ji} \rangle}{N} \langle L \rangle + I^{\mathrm{ext}} \frac{\langle \Delta^{\mathrm{iai}} \rangle}{N}}. \tag{A.1}$$

The second formula appears more simple, but keeps also less details of the underlying processes

$$\langle \kappa_2 \rangle = \frac{N}{\langle L \rangle}. \tag{A.2}$$

Let us denote $\langle uJ_{ij} \rangle$ by α_0 The average avalanche size $\langle L \rangle$ is computed by Eq. 3.3.2

$$\langle L(\alpha_0) \rangle = \frac{N}{N - (N-1)\alpha_0} \tag{A.3}$$

and

$$\langle \Delta^{\mathrm{iai}} \rangle = \frac{\theta - \epsilon_N}{I^{\mathrm{ext}}}. \tag{A.4}$$

For the simplicity of computations we assume $\theta = 1$. In chapter 8 we show, that $\epsilon_N \to 0$, when $N \to \infty$. We consider large N and thus, we take $\epsilon_N = 0$. Incorporating this conditions into (A.1) and (A.3, A.4) we obtain for large N

$$\langle \kappa_1 \rangle = \frac{1}{\frac{\alpha_0}{N} \frac{N}{N-(N-1)\alpha_0} + \frac{1}{N}} = \frac{N\left(N - (N-1)\alpha_0\right)}{N - \alpha_0}. \tag{A.5}$$

Then the ratio $\langle\kappa_1\rangle/\langle\kappa_2\rangle$ is

$$\frac{\langle\kappa_1\rangle}{\langle\kappa_2\rangle} = \frac{N}{N-\alpha_0} \overset{N\to\infty}{\longrightarrow} 1. \tag{A.6}$$

The difference between $\langle\kappa_1\rangle$ and $\langle\kappa_2\rangle$ approaches $\alpha_0(1-\alpha_0)$. For the critical case $\alpha_0 \to 1$ and then the difference in the limit is equal zero. Also far from criticality it has no influence on the thermodynamical limit.

A.2 Estimation of the power-law parameters

Power-law distributions are often occurring in statistical observations, for example populations of cities, links to web sites, size of the earthquakes. Very often, such distributions are considered to be a power-law just by visual inspection and rather rarely parameters of distribution are estimated using proper statistical tools. In this section we discuss methods of power-law estimation and show how we estimated the parameters of the avalanche size distribution.

A quantity x obeys a power law if it is drawn from a probability distribution

$$p(x) \propto x^{-\gamma} \tag{A.7}$$

where γ is a parameter of the distribution known as the *exponent* or scaling parameter. In practice it is rarely possible to know for sure that the observation is drawn from a power-law distribution. However the observations might be consistent with the model A.7 at least for some interval of the distribution. Hence the ultimate goal is, given the observations, to specify the parameter γ and determine whether observations follow the power-law distribution with this exponent.

Taking logarithms of both sides of Eq. A.7 we see that the power-law distribution obeys $\ln p(x) = \gamma \ln x + \text{constant}$, implying that it follows a straight line on a doubly logarithmic plot. One way to probe for power-law behavior is, therefore, to construct a histogram representing the frequency distribution of x, and plot that histogram on double logarithmic axes. Using the least-squares linear regression on the logarithm of the histogram of x it is easy to obtain the estimation of the parameter γ. This method of exponent-estimation is very popular, because it is simple, fast and requires only a histogram.

Unfortunately, in many cases the usage of this method leads to large errors. One can easily find an evidence of them by trying to estimate the exponent of the data which is initially drawn from the power law. Results of such an estimation are shown in the Table A.1. The linear regression in the logarithmic scale gives rather poor results. The origin of such error are in the weighting of the deviations. In the least

Estimation method	Mean estimated γ	σ	bias
Linear	1.59	0.184	36%
Linear 5-points	2.50	0.045	0
Log-2 bins	1.77	0.038	29%
MLE	2.50	0.017	0

Table A.1: Different methods of power-law estimation. Sample results of parameter estimation using various methods for 10,000 samples of power-law distribution with $\gamma = 2.500$. Sample result based on 50 runs. Table is taken form [62]

square deviation the same weight is assigned for the deviation from the straight line for small values of the observable as for the large ones. But large values appear very rare and therefore their statistics in the histogram is strongly undersampled. This gives rise to the bias. There are many articles in different areas of research devoted to the bias induced by the linear regression estimation for example in [83, 69].

Very often in the real world power laws are present not in the form described in Eq. A.7, but with a cut-off. For example, a power-law distribution with exponential cut-off is given by

$$p(x) = C_{\gamma,\beta} x^{-\gamma} e^{-\beta x}, \tag{A.8}$$

where $C_{\gamma,\beta} = 1/\int_0^\infty x^{-\gamma} e^{-\beta x} dx$ is the normalization constant.

In the case when the estimated distribution has a cut-off the linear regression estimation is based only on a small values of the observable and, therefore, it does not feel the undersampling of the large x tail. It is also clear from the Table A.1, where the estimation which is based only on the first five bins of the histogram does not have any bias. In the case of neural avalanches a sharp cut-off is observed. For the estimation of the distribution parameters we used a linear regression for three reasons:

- It is fast and does not require storage of large amount of data
- Experimentalist use this method of estimation, therefore it is just reasonable to use the same to be able to compare results
- parameters estimated by linear regression are the same as the paremeters obtained by maximum likelihood estimation (MLE)

We discuss the last point in more details, but first we explain how to do the MLE estimation for the power law and especially for the power law with cut-off. The goal of the maximum-likelihood method is to obtain an estimate for the parameter of the probability distribution, which in this general case is γ. Because the experimental

outcomes are assumed to be statistically independent, the combined probability to find, in n measurements, the data $x_1, \ldots, x_n$ is given by:

$$L(\gamma) = \prod_{i=1}^{n} P(x_i \mid \gamma), \tag{A.9}$$

where L is the likelihood function. To avoid working with very large or small numbers that can cause computational inaccuracies, and to facilitate the analysis we use the logarithmic likelihood

$$\ln L(\gamma) = \sum_{i=1}^{n} \ln P(x_i \mid \gamma), \tag{A.10}$$

We now introduce γ^*, the value that maximizes L, which is known as the maximum-likelihood value. It is the best estimate for real γ and we want to calculate its value. We obtain γ^* by solving

$$\partial_\gamma \ln L(\gamma) = 0. \tag{A.11}$$

For the continuous power-law distribution the density is given by

$$p(x) = (\gamma - 1)x^\gamma \tag{A.12}$$

the logarithmic likelihood function then takes form

$$\ln L(\gamma) = n \ln(\gamma - 1) - \gamma \sum_{i=1}^{n} \ln x_i, \tag{A.13}$$

solution of (A.11) provide the following estimation of the power-law exponent

$$\gamma^* = 1 + n \left(\sum_{i=1}^{n} \ln x_i \right)^{-1}. \tag{A.14}$$

For the discrete power-law distribution normalization constant is given by Riemann zeta function

$$C_\gamma = \frac{1}{\zeta(\gamma)} = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}}. \tag{A.15}$$

Analogously to (A.12)(A.13)(A.14) one can find estimator for the exponent as a solution of the equation

$$-\sum_{i=1}^{n} \ln x_i = n \frac{\zeta'(\gamma)}{\zeta(\gamma)}. \tag{A.16}$$

However one can find solution of Eq. A.16 only numerically. The avalanche size distribution is obviously discrete therefore we will either use this estimation or

explicitly look for the maximization of the logarithmic likelihood function.

Normally, empirical data does not follow the power-law distribution for all its values, but only for some range $(x_{\min}, x_{\max})$. For $x > x_{\max}$ exponential or non-exponential cut-off is observed, and for $x < x_{\min}$ small-values deviations. For example, neuronal avalanches have a clear cut-off at the size a little smaller then the system size [14]. To incorporate this cut-offs in the estimation of the discrete power-law exponent one has to change the normalization constant, i.e.

$$\frac{1}{C\left(\gamma, x_{\min}, x_{\max}\right)} = \zeta(\gamma) - \sum_{k=1}^{x_{\min}} k^{-\gamma} - \zeta(\gamma, x_{\max}), \tag{A.17}$$

where $\zeta(\gamma, n)$ is the generalized Riemann zeta function. The bounds of the power-law region $x_{\min}$ and $x_{\max}$) can be extracted from the data.

It is not enough, to find the exponent of the best-matching power law, it is also necessary to check, whether the hypothesis H_0, that data is sampled from the power-law distribution is correct. Unfortunately, there is no method to do it. The only thing one can prove is that there are no reasons to claim, that data comes *not* from the power-law distribution, i.e. that H_0 hypotheses can not be rejected at some significance level. To do we define the "distance" between distributions and measure how far the empirical distribution of the data stays from the estimated distribution. In some cases, the distribution of such distance is know, which allows to make a test for the H_0 hypothesis. For non-normal data the most common distance is defined by the Kolmogorov-Smirnov statistic (KS) [117]. KS statistics is defined as the maximum difference between the cumulative distribution functions (CF) of the data and the fitted model

$$D = \max_{x_{\max} \geq x \geq x_{\min}} |S(x)P(x)|, \tag{A.18}$$

here $S(x)$ is the CF of the data for the observations with value between $x_{\min}$ and $x_{\max}$), and $P(x)$ is the CF for the power-law model that best fits the data on the interval $x_{\max} \geq x \geq x_{\min}$.

One of advantages of this statistics is formulated in the Glivenko-Cantelli theorem. Namely, if the sample comes from distribution $P(x)$, then D converges to 0 almost surely when the sample size goes to infinity. Kolmogorov also provided results about speed of this convergence and it is even possible to write explicitly the distribution of the KS statistics for the case when x_i are indeed drown from the hypothesized distribution. This allows to compute the p-value for the power-law hypothesis. P-value is a probability to get a statistics at least as extreme as observed given that the hypothesis H_0 holds. One rejects the null hypothesis if the p-value is smaller than or equal to the significance level.

α	γ_{LS}	$\Delta\gamma$	γ_{MLE}	KS-statistics	p-value
0.85	-2.01	0.022	-1.93	0.007	0.09
0.89	-1.376	0.002	-1.370	0.005	0.317
0.93	-1.20	0.01	-1.21	0.015	0.001

Table A.2: Check of the hypothesis of the power-law distribution. A network simulated for different values of connectivity parameter α_0. Basing on 10^5 avalanches an empiric distribution functions is found. MLE estimation of the power-law exponent γ_{MLE} and least-squares estimation γ_{LS} are found. The p-value is counted by the generation of the 1000 Monte-Karlo data sets.

Unfortunately, KS statistic assumes that the hypothesized distribution of data x is known which does not hold in our case. What we can obtain is the distance between the empirical distribution of data and the best-matched power law, which does not necessary coinside with the true underlying power law. Therefore, we have to use Monte-Carlo approach to find the p-value by generating number of power-law samples with the exponent γ. Then by comparing the distance between each of this generated samples and their best-matched power-laws one can get the p-value for the hypothesis H_0. Namely procedure is as follows:

1. Determine the best fit of the power law to the data, estimating both the scaling parameter γ and the cut-off parameters $x_{\min}$ and $x_{\max}$
2. Calculate the KS statistic for the goodness-of-fit of the best-fit power law to the data.
3. Generate a large number of synthetic data sets using the procedure above, fit to each the power-law, and calculate the KS statistic for each fit.
4. Calculate the p-value as the fraction of the KS statistics for the synthetic data sets whose value exceeds the KS statistic for the real data.
5. If the p-value is sufficiently small then the power-law distribution can be ruled out.

We performed this procedure for the avalanche sizes distribution of the network of size 100 in three different regimes: $\alpha_0 = 0.85$ subcritical regime, $\alpha_0 = 0.89$ critical regime, and $\alpha_0 = 0.93$ super-critical regime. Results of estimation are shown in the Table A.2. For the distributions, which are distinguished as critical by the thresholding of the linear method, it is not possible to reject hypothesis H_0 (p-value> 0.1). An estimation of the power-law exponents also gives similar results in both cases.

A.3 Distribution of the inter-avalanche intervals

Let membrane potentials of all neurons be independent and uniformly distributed on the interval $[\epsilon_N, \theta]$. We show, that in the limit $N \to \infty$ inter-avalanche intervals Δ^{iai} follow the geometric distribution,

$$G_p(n) = p(1-p)^{n-1}.$$

Let us first compute the probability $P(\Delta^{\text{iai}} = 1)$ of the event $\Delta^{\text{iai}} = 1$. Such event means that the neuron j which got the external input had the membrane potential $h_j \in (1 - I^{\text{ext}}, 1)$. The probability of it is equal to $p = I^{\text{ext}}/(1 - \epsilon_N)$.

Let B_L^N be the event that among L randomly chosen numbers between 1 and N there are no coinsiding ones. Denote by $\overline{B}_L^N$ the complement of B_L^N. Decompose $P(\Delta^{\text{iai}} = L)$ as

$$P\left(\Delta^{\text{iai}} = L\right) = P\left(\Delta^{\text{iai}} = L \,\middle|\, B_L^N\right) P\left(B_L^N\right) + P\left(\Delta^{\text{iai}} = L \,\middle|\, \overline{B}_L^N\right) P\left(\overline{B}_L^N\right). \quad \text{(A.19)}$$

We prove, that for any fixed L

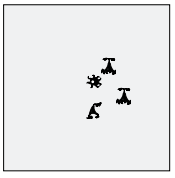
$$\begin{aligned} P\left(B_L^N\right) &\to_{N\to\infty} 1, \\ P\left(\Delta^{\text{iai}} = L \,\middle|\, B_L^N\right) &= G_p(L). \end{aligned}$$

We can write $P\left(B_L^N\right)$ explicitly

$$P\left(B_L^N\right) = \frac{N(N-1)\ldots(N-L+1)}{N^L} = \left(1 - \frac{1}{N}\right)\ldots\left(1 - \frac{L-1}{N}\right) \to_{N\to\infty} 1. \quad \text{(A.20)}$$

The event $A = \left\{\Delta^{\text{iai}} = L \,\middle|\, B_L^N\right\}$ mean, that first $L-1$ neurons selected for the external input had membrane potential smaller than $1 - I^{\text{ext}}$ and the last had the membrane potential larger than $1 - I^{\text{ext}}$. All the selected neurons were non-repeating, therefore all of them were uniformly distributed on $[\epsilon_N, \theta]$ and probability of the event A is given by the geometric distribution.

Thus we have proven that Δ^{iai} converges in distribution to the geometric distribution.



Bibliography

[1] L. F. Abbott and R. Rohrkemper. A simple growth model constructs critical avalanche networks. *Prog Brain Res.*, 165:13–9, 2007.

[2] M. Abeles. *Corticonics : neural circuits of the cerebral cortex*. Cambridge University Press, 1991.

[3] A. M. Alencar, S. V. Buldyrev, A. Majumdar, H. E. Stanley, and B. Suki. Avalanche dynamics of crackle sound in the lung. *Phys. Rev. Lett.*, 87(8):088101, Aug 2001.

[4] A. A. Ali and D. Dhar. Structure of avalanches and breakdown of simple scaling in the abelian sandpile model in one dimension. *Phys. Rev. E*, 52(5):4804–4816, Nov 1995.

[5] P. Alstrom. Mean-field exponents for self-organized critical phenomena. *Phys. Rev. A*, 38(9):4905–4906, 1988.

[6] K .B. Athreya and P. Jagers. *Classical and Modern Branching Processes*, volume 84 of *IMA*. Springer, 1997.

[7] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Springer Verlag, 1999.

[8] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of $1/f$ noise. *Phys. Rev. Lett.*, 59:381–384, 1987.

[9] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Phys. Rev. A*, 38:364–374, 1988.

[10] M. Barma and D. Dhar. Slow relaxation in a model with many conservation laws: Deposition and evaporation of trimers on a line. *Phys. Rev. Lett.*, 73(15):2135–2138, Oct 1994.

[11] M. F. Bear, B. W. Connors, and M. A. Paradiso. *Neuroscience*. Lippincott Williams & Wilkins, 2 edition, 2001.

[12] C. Bedard, H. Kroeger, and A. Destexhe. Does the 1/f frequency-scaling of brain signals reflect self-organized critical states? *Physical Review Letters*, 97:118102, 2006.

[13] J. Beggs. The criticality hypothesis: how local cortical networks might optimize information processing. *Proceedings of the Royal Society A.*, 2007.

[14] J. Beggs and D. Plenz. Neuronal avalanches in neocortical circuits. *J. Neurosci*, 23:11167–11177, 2003.

[15] J. Beggs and D. Plenz. Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *J. Neurosci*, 24(22):5216–5229, 2004.

[16] R. Bertram, A. Sherman, and E. F. Stanley. Single-domain/bound calcium hypothesis of transmitter release and facilitation. *Journal of Neurophysiology*, 75(5):1919–1931, 1996.

[17] A. Bhowal. Damage spreading in the 'sandpile' model of soc, 1997.

[18] G. Bi and M. Poo. Synaptic modification in cultured hippocampal neurons: Dependence on sike timing, synaptic strength and postsynaptic cell type. *J. of Neurosci.*, 18(24):10464–10472, 1998.

[19] M. Biskup, Ph. Blanchard, D. Gandolfo, and T. Krüger. Phase transition and critical behavior in a model of organized criticality. *Probability Theory and Related Fields*, 128(1):1 – 41, 2004.

[20] Ph. Blanchard, B. Cessac, and T. Krueger. What can one learn about self-organized criticality from dynamical systems theory ? *J. Stat. Physics*, 98:375, 2000.

[21] T. V. P. Bliss and G. L. Collingridge. A synaptic model of memory: long-term potentiation in the hippocampus. *A synaptic model of memory: long-term potentiation in the hippocampus*, 361:31 – 39, 1993.

[22] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes.in Math.* Springer-Verlag, 1975.

[23] R. Brette and W. Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94:3637–3642, 2005.

[24] H.-M. Bröker and P. Grassberger. Random neighbor theory of the olami-feder-christensen earthquake model. *Phys. Rev. E*, 56(4):3944–3952, Oct 1997.

[25] N. Brunel and P. E. Latham. Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural Comput.*, 15(10):2281 – 2306, 2003.

[26] G. Buzsáki. *Rhythms of the Brain*. Oxford University Press, 2006.

[27] G. Buzsáki, R. D. Traub, and T. Pedley. *Current Practice of Clinical Encephalography*, chapter The cellular synaptic generation of EEG, pages 1–11. Lippincott Williams and Wilkins, 3 edition, 2003.

[28] S. Camalet, X. Duke, F. Jülicher, and J. Prost. Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *Proc. Natl. Acad. Sci.*, 97:3183–3188, 2000.

[29] J. Cardy. *Scaling and Renormalization in Statistical Physics*. Number 5 in Cambridge Lecture Notes in Physics. Cambridge university press, 1996.

[30] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole. Evidence for self-organized criticality in a time series of electric power system blackouts. *IEEE TA on Circuits and Systems*, 51(9):1733–1740, 2004.

[31] J. W. S. Cassels. *An introduction to Diophantine approximation*. Number 35 in Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge University Press, 1957.

[32] D. Centonze, B. Picconi, P. Gubellini, G. Bernardi, and P. Calabresi. Dopaminergic control of synaptic plasticity in the dorsal striatum. *J Neurophysiol*, 94:4168–4177, 2005.

[33] B. Cessac. Some fractal aspects of self-organized criticality, 2004.

[34] B. Cessac, Ph. Blanchard, T. Krueger, and J.L. Meunier. Self-organized criticality and thermodynamic formalism. *Journal of Statistical Physics*, 115:1283, 2004.

[35] B. Cessac, Ph. Blanchard, and T. Krüger. Lyapunov exponents and transport in the zhang model of self-organized criticality. *Phys. Rev. E*, 64(1):016133, Jun 2001.

[36] B. Cessac and J. L. Meunier. Anomalous scaling and lee-yang zeros in self-organized criticality. *Phys. Rev. E*, 65(3):036131, Feb 2002.

[37] D. Chialvo. Are our senses critical? *Nature Physics*, 2:301–302, 2006.

[38] K. Christensen. *Self-organization in models of sandpiles, earthquakes, and fireflies*. PhD thesis, University of Aarhus, Denmark, 1992.

[39] K. Christensen and Z. Olami. Scaling, phase transitions, and nonuniversality in a self-organized critical cellular automaton model. *Phys. Rev. A*, 46:1829–1838, 1992.

[40] B. Connors and M. Long. Electrical synapses in the mammalian brain. *Annu, Rev. Neurosci.*, 27(393-418), 2004.

[41] A. Corral. Comment on "do earthquakes exhibit self-organized criticality?". *Physical Review Letters*, 95(15):159801, 2005.

[42] Á. Corral, C. J. Pérez, A. Díaz-Guilera, and A. Arenas. Self-organized criticality and synchronization in a lattice model of integrate-and-fire oscillators. *Phys. Rev. Lett.*, 74:118–121, 1995.

[43] H. Dale. Pharmacology and nerve-endings. *Proc. R. Soc. Med.*, 28:319–332, 1935.

[44] P. Dayan and L. Abbott. *Theoretical Neuroscience*. MIT Press, Cambridge, 2001.

[45] L. de Arcangelis, C. Perrone-Capano, and H. J. Herrmann. Self-organized criticality model for brain plasticity. *Phys. Rev. Lett.*, 96:028107(4).

[46] R. Der, F. Hesse, and G. Martius. Rocking stumper and jumping snake from a dynamical system approach to artificial life. *Adaptive Behavior*, 14(2):105–115, 2006.

[47] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski. An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Computation*, 6(1):14, 1994.

[48] D. Dhar. Self-organized critical state of sandpile automaton models. *Phys. Rev. Lett.*, 64(14):1613–1616, Apr 1990.

[49] D. Dhar and S. S. Manna. Inverse avalanches in the abelian sandpile model. *Phys. Rev. E*, 49(4):2684–2687, Apr 1994.

[50] D. Dhar and R. Ramaswamy. Exactly solved model of self-organized critical phenomena. *Phys. Rev. Lett.*, 63(16):1659–1662, Oct 1989.

[51] R. Dickman, M. A. Muñoz, A. Vespignani, and S. Zapperi. Paths to self-organized criticality. *Brazilian Journal of Physics*, 30:27 – 41, 03 2000.

[52] M. Diesmann, M.-O. Gewaltig, and A. Aertsen. Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402:529–533, 1999.

[53] C. W. Eurich, M. Herrmann, and U. Ernst. Finite-size effects of avalanche dynamics. *Phys. Rev. E*, 66:066137–1–15, 2002.

[54] H. J. S. Feder and J. Feder. Self-organized criticality in a stick-slip process. *Phys. Rev. Lett.*, 66:2669–2672, 1991.

[55] R. D. Fields and B. Stevens-Graham. New insights into neuron-glia communication. *Science*, 298:556–562, October 2002.

[56] H. Flyvbjerg, K. Sneppen, and P. Bak. Mean field theory for a simple model of evolution. *Phys. Rev. Lett.*, 71(24):4087–4090, Dec 1993.

[57] V. Frette, K. Christensen, A. M. Malthe-Sørenssen, J. Feder, T. Jøssang, and P. Meakin. Avalanche dynamics in a pile of rice. *Nature*, 397:49, 1996.

[58] R. Frigg. Self-organised criticality – what it is and what it isn't. *Studies In History and Philosophy of Science Part A*, 34(3):613–632, 2003.

[59] B. Gaveau and L. S. Schulman. Fluctuations in mean-field self-organized criticality. *Journal of Statistical Physics*, 74(3-4), 1994.

[60] W. Gerstner. Population dynamics of spiking neurons: Fast transients, asynchronous states, and locking. *Neural Computation*, 12:43–89, 2000.

[61] W. Gerstner and W. M. Kistler. *Spiking Neuron Models. Single Neurons, Populations, Plasticity.* Cambridge University Press, 2002.

[62] M. Goldstein, S. Morris, and G. Yen. Problems with fitting to the power-law distribution. *Europ. Phys. J. B*, 41:255–258, 2004.

[63] P. Grassberger. Efficient large-scale simulations of a uniformly driven system. *Phys. Rev. E*, 49(3):2436–2444, Mar 1994.

[64] B. Gutenberg and C. F. Richter. *Ann. Geophys.*, 9:1, 1956.

[65] C. Haldeman and J. Beggs. Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys. Rev. Lett.*, 94, 2005.

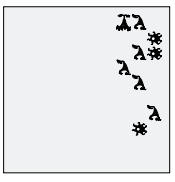
[66] T. E. Harris. *The theory of branching processes.* Springer, Berlin, 1963.

[67] D. O. Hebb. *Organization of Behavior: a Neuropsychological Theory.* Wiley, New York, 1949.

[68] A. V. M. Herz and J. J. Hopfield. Earthquake cycles and neural reverberations: collective oscillations in systems with pulse-coupled threshold elements. *Phys. Rev. Lett.*, 75:1222–1225, 1995.

[69] J. P. Hoogenboom, W. K. den Otter, and H. L. Offerhaus. Accurate and unbiased estimation of power-law exponents from single-emitter blinking data. *Journal of Chemical Physics*, 125:204713–204713–12, November 2006.

[70] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *PNAS*, 79(8):2554–2558, 1982.



[71] J. J. Hopfield and A. V. M. Herz. Rapid Local Synchronization of Action Potentials: Toward Computation with Coupled Integrate-and-Fire Neurons. *Proceedings of the National Academy of Sciences*, 92(15):6655–6662, 1995.

[72] E. M. Izhikevich. Which model to use for cortical spiking neurons? *IEEE Transactions on Neural Networks*, 15:1063–1070, 2004.

[73] H. M. Jaeger, C. Liu, and Sidney R. Nagel. Relaxation at the angle of repose. *Phys. Rev. Lett.*, 62(1):40–43, Jan 1989.

[74] H. J. Jensen. *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*. Cambridge Univ Pr, 1998.

[75] R. Jolivet, T. J. Lewis, and W. Gerstner. Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy. *J Neurophysiol*, 92:959–976, 2004.

[76] D. E. Juanico, C. Monterola, and C. Saloma. Dissipative self-organized branching in a dynamic population. *Phys. Rev. E*, 75, 2007.

[77] L. P. Kadanoff, S. R. Nagel, L. Wu, and S. Zhou. Scaling and universality in avalanches. *Phys. Rev. A*, 39(12):6524–6537, Jun 1989.

[78] L. P. Kadanoff, S. R. Nagel, L. Wu, and S. Zhou. Scaling and universality in avalanches. *Phys. Rev. A*, 39(12):6524–6537, Jun 1989.

[79] E. R. Kandel, J. H. Schwartz, and T. M. Jessell. *Principles of Neural Science*. Elsevier Science Publishing Co Inc, 3 edition, 1991.

[80] M. Kimmel and D. E. Axelrod. *Branching Processes in Biology*. Springer, 2002.

[81] O. Kinouchi and M. Copelli. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2:348–352, 2006.

[82] O. Kinouchi, S. T. R. Pinho, and C. P. C. Prado. Random-neighbor Olami-Feder-Christensen slip-stick model. *Phys. Rev. E*, 58:3997–4000, 1998.

[83] D. A. Koster, C. H. Wiggins, and N. H. Dekker. Multiple events on single molecules: Unbiased estimation in single-molecule biophysics. *Proceedings of the National Academy of Sciences*, 103(6):1750–1755, 2006.

[84] L. Lapicque. Recherches quantitatives sur lexcitationelectrique des nerfs traitee comme une polarisation. *J. Physiol. Pathol. Gen*, 9:620–635, 1907.

[85] P. E. Latham, B. J. Richmond, P. G. Nelson, and S. Nirenberg. Intrinsic dynamics in neuronal networks. i. theory. *J Neurophysiol*, 83(2):808–827, 2000.

[86] D.-S. Lee, K.-I. Goh, B. Kahng, and D. Kim. Sandpile avalanche dynamics on scale-free networks. *Physica A*, 338:84–91, 2004.

[87] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. ii. lattice gas and ising model. *Phys. Rev.*, 87(3):410–419, Aug 1952.

[88] R. A. Legenstein and W. Maass. Edge of chaos and prediction of computational performance for neural microcircuit models. *Neural Networks*, pages 323–333, 2007.

[89] A. Levina, U. Ernst, and J. M. Herrmann. Criticality of avalanche dynamics in adaptive recurrent networks. *Neurocomput.*, 70:1877–1881, 2007.

[90] A. Levina, J. M. Herrmann, and T. Geisel. Dynamical synapses give rise to a power-law distribution of neuronal avalanches. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 771–778. MIT Press, Cambridge, MA, 2006.

[91] A. Levina, J. M. Herrmann, and T. Geisel. Dynamical synapses causing self-organized criticality in neural networks. *Nature Physics*, 3:857–860, 2007.

[92] M. Lin and T.-L. Chen. Self-organized criticality in a simple model of neurons based on small-world networks. *Phys. Rev. E*, 71(016133), 2005.

[93] S. Lise and H. J. Jensen. Transitions in nonconserving models of self-organized criticality. *Phys. Rev. Lett.*, 76:2326–2329, 1996.

[94] J. E. Lisman, S. Raghavachari, and R. W. Tsien. The sequence of events that underlie quantal transmission at central glutamatergic synapses. *Nature Rev. Neurosci.*, 8:597–609, 2007.

[95] Y. H. Liu and X. J. Wang. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *J Comput Neurosci,*, 10(1):25–45, 2001.

[96] A. Loebel and M. Tsodyks. Computation by ensemble synchronization in recurrent networks with synaptic depression. *J. of Computational Neuroscience*, 13:111–124, 2002.

[97] B. D. Malamud, G. Morein, and D. L. Turcotte. Forest fires: An example of self-organized critical behavior. *Science*, 281(5384):1840 – 1842, 1998.

[98] A. Mallart and A. R. Martin. An analysis of facilitation of transmitter release at the neuromuscular junction of the frog. *J. Physiol.*, 193:679–697, 1967.

[99] S. S. Manna, L. B. Kiss, and J. Kertesz. Cascades and self-organized criticality. *Journal of Statistical Physics*, 61(3-4):923–932, 1990.

[100] H. Markram. The blue brain project. *Nature Rev. Neurosci*, 7:153–160, 2006.

[101] H. Markram and M. Tsodyks. Redistribution of synaptic efficacy between pyramidal neurons. *Nature*, 382:807–810, 1996.

[102] H. Markram and M. V. Tsodyks. Redistribution of synaptic efficacy: A mechanism to generate infinite synaptic input diversity from a homogeneous population of neurons without changing absolute synaptic efficacies. *J. Physiology (Paris)*, 90:229–232, 1996.

[103] H. Markram, Y. Wang, and M. Tsodyks. Differential signaling via the same axon from neocortical layer 5 pyramidal neurons. *PNAS*, 95:5323–5328, 1998.

[104] A. Mazzoni, F. D. Broccard, E. Garcia-Perez, P. Bonifazi, M. E. Ruaro, and V. Torre. On the dynamics of the spontaneous activity in neuronal networks. *PLoS ONE*, 2(5):e439, May 2007.

[105] B. L. McNaughton. Neuronal mechanisms for spatial computation and information storage. In L. Nadel, L. Cooper, P. Culicover, and R. M. Harnish, editors, *Neural connections, mental computation*, pages 285–350. Cambridge, MA: MIT Press, 1989.

[106] S. R. Nagel. Instabilities in a sandpile. *Rev. Mod. Phys.*, 64(1):321–325, Jan 1992.

[107] B. Naundorf, F. Wolf, and M. Volgushev. Unique features of action potential initiation in cortical neurons. *Nature*, 440(7087), 2006.

[108] M. E. J. Newman. Evidence for self-organized criticality in evolution. *Physica D*, 107:293–296, 1997.

[109] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46:323, 2005.

[110] Z. Olami, H. J. S. Feder, and K. Christensen. Self-organized criticality in a continous, nonconservative cellular automaton modeling earthquakes. *Phys. Rev. Lett.*, 68:1244–1247, 1992.

[111] A. Van Ooyen, J. Van Pelt, and M. A. Corner. Implications of activity dependent neurite outgrowth for neuronal morphology and network development. *J Theor Biol.*, 172(1):63–82, 1995.

[112] R. Otter. The multiplicative process. *Ann. Math. Statist.*, 20:248–263, 1949.

[113] G. L. Pellegrini, L. de Arcangelis, H. J. Herrmann, and C. Perrone-Capano. Activity-dependent neural network model on scale-free networks. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 76(1):016107, 2007.

[114] T. Petermann, M. Lebedev, M. Nicolelis, and D. Plenz. Neuronal avalanches in vivo. *Society for Neuroscience Abstracts*, 2006.

[115] L. Pietronero, A. Vespignani, and S. Zapperi. Renormalization scheme for self-organized criticality in sandpile models. *Phys. Rev. Lett.*, 72(11):1690–1693, Mar 1994.

[116] D. Plenz and T. C. Thiagarajan. The organizing principles of neuronal avalanches: cell assemblies in the cortex? *Trends in Neurosciences*, 30(3):101–110, 2007.

[117] W. H. Press, S.A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: the art of scientific computing*. Cambridge University Press, 2 edition, 1992.

[118] V. B. Priezzhev, D. V. Ktitarev, and E. V. Ivashkevich. Formation of avalanches and critical exponents in an abelian sandpile model. *Phys. Rev. Lett.*, 76(12):2093–2096, Mar 1996.

[119] D. Purves, G. Augustine, D. Fitzpatrick, L. Katz, A.-S LaMantia, and J. McNamara. *Neuroscience*. Sinauer, Sunderland, 1997.

[120] A. Rauch, G. La. Camera, H.R. Luscher, W. Senn, and S. Fusi. Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. *J Neurophysiol*, 90:1598–1612, 2003.

[121] M. Ren, Y. Yoshimura, N. Takada, S. Horibe, and Y. Komatsu. Specialized inhibitory synaptic actions between nearby neocortical pyramidal neurons. *Science*, 316(5825):758–761, 2007.

[122] J. Rosendahl, M. Vekić, and J. E. Rutledge. Predictability of large avalanches on a sandpile. *Phys. Rev. Lett.*, 73(4):537–540, Jul 1994.

[123] S. Roux and E. Guyon. Temporal development of invasion percolation. *Journal of Physics A Mathematical General*, 22:3693–3705, September 1989.

[124] D. Ruelle. *Thermodynamic formalism*. Reading. Addison-Wesley, 1978.

[125] M. M. Segal, E. E. Korkotian, and D. Murphy. Dendritic spine formation and pruning: common cellular mechanisms? *Trends neurosci.*, 23(2):53–57, 2000.

[126] Ya. G. Sinai. Gibbs measures in ergodic theory. *Russ. Math. Surveys*, 27(4):21–69, 1972.

[127] I. Song and Huganir R.L. Regulation of ampa receptors during synaptic plasticity. *Trends in Neurosciences*, 25(11):578–589, 2002.

[128] D. Sornette, A. Johansen, and I. Dornic. Mapping self-organized criticality onto criticality. *J. Phys. I*, 5:325–335, 1995.

[129] C. V. Stewart and D. Plenz. Inverted-u profile of dopamine-nmda-mediated spontaneous avalanche recurrence in superficial layers of rat prefrontal cortex. *J Neurosci.*, 26(31):8148–59, 2006.

[130] J. Teramae and T. Fukai. Local cortical circuit model inferred from power-law distributed neuronal avalanches. *J Comput Neurosci*, 22(3):301–312, 2007.

[131] A. M. Thomson and J. Deuchars. Temporal and spatial properties of local circuits in neocortex. *Trends Neurosci.*, 17(3):119–126, 1994.

[132] M. Timme, F. Wolf, and T. Geisel. Prevalence of unstable attractors in networks of pulse-coupled oscillators. *Phys. Rev. Lett.*, 89(15):154105, 2002.

[133] M. Tsodyks, K. Pawelzik, and H. Markram. Neural networks with dynamic synapses. *Neural Computation*, 10:821–835, 1998.

[134] M. Tsodyks, A. Uziel, and H. Markram. Synchrony generation in recurrent network with frequency-dependent synapses. *J. of Neuroscience*, 20:1–5, 2000.

[135] M. V. Tsodyks and H. Markram. Plasticity of neocortical synapses enables transitions between rate and temporal coding. In C. von der Malsburg et al, editor, *Lecture Notes in Comp. Sci.*, volume 1112, pages 445–450. Springer, 1996.

[136] T. Tsuchiya and M. Katori. Proof of breaking of self-organized criticality in a nonconservative abelian sandpile model. *Phys. Rev. E*, 61(2):1183–1188, Feb 2000.

[137] D. L. Turcotte. Self-organized criticality: Does it have anything to do with criticality and is it useful? *Nonlinear Processes in Geophysics*, 8:193–196, 2000.

[138] G. G. Turrigiano, K. R. Leslie, N. S. Desai, L. C. Rutherford, and S. B. Nelson. Activity-dependent scaling of quantal amplitude in neocortical pyramidal neurons. *Nature*, 391:892–896, 1998.

[139] M. Vergeles, A. Maritan, and J. R. Banavar. smean-field theory of sandpiles. *Phys. Rev. E*, 55(2):1998–2000, Feb 1997.

[140] A. Vespignani and S. Zapperi. Order parameter and scaling fields in self-organized criticality. *Phys. Rev. Lett.*, 78:4793–4796, 1997.

[141] A. Vespignani, S. Zapperi, and V. Loreto. Renormalization of nonequilibrium systems with critical stationary states. *Phys. Rev. Lett.*, 77:4560–4563, 1996.

[142] A. Vespignani, S. Zapperi, and L. Pietronero. Renormalization approach to the self-organized critical behavior of sandpile models. *Phys. Rev. E*, 51(3):1711–1724, Mar 1995.

[143] T. P. Vogels, K. Rajan, and L. F. Abbott. Neural network dynamics. *Annual Review of Neuroscience*, 28(1):357–376, 2005.

[144] G. Voglis and N. Tavernarakis. The role of synaptic ion channels in synaptic plasticity. *EMBO reports*, 7(11):11041110, 2006.

[145] D. Volchenkov, Ph. Blanchard, and B.Cessac. Quantum field theory renormalization group approach to self-organized criticality : the case of random boundaries. *International Journal of Modern Physics B*, 16(8):1171–1204, 2002.

[146] A. Volterra and C. Steinhäuser. Glial modulation of synaptic transmission in the hippocampus. *Glia*, 47(3):249–57, 2004.

[147] H. W. Watson and F. Galton. On the probability of extinction of families. *J. of the Anthropol. Inst.*, 138, 1874.

[148] D. Watts and S. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393:440–442, 1998.

[149] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. i. theory of condensation. *Phys. Rev.*, 87(3):404–409, Aug 1952.

[150] X. Yang, S. Du, and J. Ma. Do earthquakes exhibit self-organized criticality? *Physical Review Letters*, 92(22):228501, 2004.

[151] X. Yang, S. Du, and J. Ma. Yang et al. reply:. *Physical Review Letters*, 95(15):159802, 2005.

[152] S. Zapperi, Kent B. Lauritsen, and H. E. Stanley. Self-organized branching processes: Mean-field theory for avalanches. *Phys. Rev. Lett.*, 75(22):4071–4074, Nov 1995.

[153] J. E. Zengel and K. L. Magleby. Augmentation and facilitation of transmitter release. a quantitative description at the frog neuromuscular junction. *The Journal of General Physiology*, 80:583–611, 1982.

[154] Y.-C. Zhang. Scaling theory of self-organized criticality. *Phys. Rev. Lett.*, 63(5):470–473, Jul 1989.

[155] R. S. Zucker and W. G. Regehr. Short-term synaptic plasticity. *Annual Review of Physiology*, 64:355405, 2002.

Конец

Acknowledgements

There are many people who contributed to this work either directly, or by influencing my life. I find these two parts very interconnected, and would not dare to separate them here.

I would like to thank Theo Geisel, for giving me the possibility to work in this institute. For investing that much energy in keeping his group as nice as it is. And especially, for convincing me to do things which I never thought of doing myself, but which I now see as very important and useful.

I am grateful to Michael Herrmann for supervising my work, for raising many inspiring questions and for advising me how to answer them and giving me the freedom to do it my way. For explaining to me what "scientific work" is all about and for many interesting philosophical discussions and ideas about the scientific kitchen. And certainly for all this: The, the, a, an, the, a (if I tried to list all the articles, which Michael inserted in our publications, I would have to write another ten pages).

I thank Manfred Denker for supervising the mathematically rigorous part of my work, for letting me to turn into a halfway physicist, and at the same time for helping me to preserve traits of a mathematician. For showing how to keep my reasoning strict and clear. And for a wonderful sense of humor, which often cheered me up.

I am deeply grateful to Marc Timme for being around when I was justbits and pieces. For helping me with scientific writing and thinking. For showing me another way of thinking about the world, and for influencing my point of view in many questions.

I thank Fred Wolf for useful recommendations about scientific publications.

I thank Udo Ernst for interesting discussions and for our colloboration. I am grateful to Dietmar Plenz and John Beggs, for fruitful discussions and answers to my questions about their work.

I am indebted to Georg Martius for a lot of good advice on programming and on dealing with computers in general and for his wonderful program Simparex which allowed me to inhabit the full computer cluster with my simulations. Thank you, Denny Fliegner, for your presence and that of the cluster. And many thanks to York-Fabian Beensen for answering my constant computer-related questions and for confirming that "nice system administrator" is not an oxymoron.

I thank my teachers, Mikhail Gordin who was my diploma supervisor in Russia and introduced me into the science of mathematics, Alexander Golovanov who was for years my extra-curricular teacher who first showed me the beauty of mathematics.

Many thanks to Katharina Jeremias, Corinna Trautsch, and Regina Wunderlich for helping to make the organizational part of my life here as simple as it can be in Germany.

Great thanks to Sibylle Jane Kalz, who corrected important parts of this thesis in a in a very short time.

I thank Joachim Hass, Björn Naundorf, and Vincent David with whom I shared office space, and many nice conversations. I thank Vincent for helping me to solve many problems without leaving his chair and for always being open for discussions or chats.

I thank Raoul-Martin Memmesheimer for many interesting, funny, and motivating chats, conversation, and discussions, and for help with Mathematica.

I thank the Russian-speaking (and almost speaking) community at the Institute for a nice opportunity to exchange a few words in Russian sometimes. Thank you, Dmitry Bibitchkov, Vitaly Belik, Katja Fiedler,Peter Hiemeyer, Michael Herrmann, Tomas Kulvicius, Georg Martius, Michael Schnabel, Nataliya Shylo and Sasha Wolf.

I thank the above mentioned people as well as Lishma Anand, Oliver Bendix, Kai Bröking, Dirk Brockmann, Markus Butz, Hecke Degering, Ragnar Fleischmann, Harold Gutch, Jens Arnold Heide, Holger Hennig, Frank Hesse, Moritz Hiller, Min Huang, Matthias Ihrke, Sven Jahnke, Matthias Kaschube, Wolfgang Keil, Christoph Kirst, Christoph Kolodziejski, Michael Kreissl, Irene Markelic, Jakob Metzger, Alejandro Morales Gallardo, Lars Reichl, Holger Schanz, Fabio Schnittler Neves, Andreas Sorge, Kristin Stamm, Tatjana Tchumatchenko, Frank van Bussel, Steffen Wischmann, Florentin Wörgötter, for creating a unique and wonderful atmosphere in the institute, which made it much more then just a working place. For any of you, I have many things to mention, to be grateful for. Unfortunately, now that I am finishing a PhD and thus not yet writing a novel, I have to keep it reasonably short.

I thank Marc, Sarah, and Liesl for allowing me to live in their WG for a month and letting me feel at home there for a much longer time. For being good friends and for forcing me to start to speak German.

I thank my friend Rada Matic (Dakovic) without whom I would never have come here and definitely would not have survived the first year. I thank Olya (Olha Ivanyshyn) for being a good friend all these years. I thank Vitaly Belik for sharing pizzas, nice conversations and for being so supportive.

I thank my friends in Russia: Sonya, Tanya, Lisa, Masha, Shurik, Kifir, Lana, for their support and for making me feel that I can still come back home to them.

I thank my brother who is and always will be my best friend. My greatest thanks to my parents, who were courageous enough to let me go, who believed in me all the time, and supported me in everything. Without you, nothing would have been possible.

And Georg, thank you so much! You have made my decision to come here much more meaningful than I ever thought it could be.