Dynamical and stochastic aspects of self-organized critical behavior in sandpiles

Martin Rypdal

A dissertation for the degree of Philosophiae Doctor

UNIVERSITY OF TROMSØ
Faculty of Science
Department of Mathematics and Statistics

October 2008

Acknowledgements:

I thank my collaborators Boris Kruglikov and Kristoffer Rypdal who have contributed essentially to the work presented in this thesis. Boris has been my supervisor for many years, and I have truly appreciated his guidance and excellent teaching. I would also like to thank the members of the Sophus Lie Seminar for providing an inspiring scientific environment in our department. In particular Einar Mjølhus, Per Jakobsen, Valentin Lychagin, Tor Flå and Marius Overholt. Boris Kozelov, Odd Erik Garcia and Tanja Zivkovic of the Complex Systems Group in Tromsø have also contributed to my work through collaboration and useful discussions. I finally want to mention Yakov Pesin who visited Tromsø in 2005. His lectures on dimension theory and multi-fractal formalism for dynamical systems were inspiring and useful.

List of papers that constitute the thesis:

1. B. Kruglikov, M. Rypdal, *Dynamics and entropy in the Zhang model of SOC*, Journal of Statistical Physics **122**, 975 (2006)

2. B. Kruglikov, M. Rypdal, *Entropy via multiplicity*, Discrete and Continuous Dynamical Systems **16**, 395 (2006)

3. B. Kruglikov, M. Rypdal, *A piece-wise affine contracting map with positive entropy*, Discrete and Continuous Dynamical Systems **16**, 393 (2006)

4. M. Rypdal, K. Rypdal, *A stochastic model of the toppling activity in self-organized critical systems*, Preprint: arXiv:0710.4010v2 (2008), (Accepted by the New Journal of Physics).

5. M. Rypdal, K. Rypdal, *Modeling of temporal fluctuations in sandpiles*, Preprint: arXiv:08073416v1(2008), (Accepted by Physical Review E).

Other papers which are a part of this work, but not directly incorporated in the manuscript are:

6. K. Rypdal, B. Kozelov, S. Ratynskaia, M. Rypdal, B. Klumov, C. Knapek, *Scale-free vortex cascade emerging from random forcing in a strongly coupled system*, New Journal of Physics **10** (2008).

7. K. Rypdal, M. Rypdal, T. Zivkovic, B. Kozelov, *Complexity in astro and geospace systems: the turbulence versus SOC controversy*, AIP Conference Proceedings **932**, 203 (2007).

Contents

Chapter 1

Introduction

In 1987 Per Bak, Chao Tang and Kurt Wiesenfeld (BTW) constructed a computer model of the cellular-automaton type, which has become widely known as the BTW sandpile model. Its appeal is based on the assumption that a large class of natural systems exhibit some common dynamical characteristics, which are shared by the dynamics and statistics of avalanches arising in a pile of sand as sand grains are slowly and randomly fed from the top of the pile. One property that these systems have in common is that the attractors of the dynamics give rise to scale-invariant properties of the statistical distribution of certain global variables. This property implies that one can draw some parallels to the behavior of equilibrium systems at critical points.

It is believed that the attractors of sandpile models generally have full basins of attraction, and that statistical properties of the asymptotic dynamics to some extent are independent of the detailed algorithm describing the interaction between neighboring sand grains. Thus the "critical state" is attained without fine tuning of parameters that we know is required to bring *equilibrium systems* to criticality. The dynamics observed in the BTW model was named Self-Organized Criticality (SOC), and in the enthusiasm that followed the first BTW articles, SOC was proposed as the governing dynamical mechanism in a large class of complex natural systems where scaling laws have been found empirically.

The view held by many scientists working in the field of natural complexity is that SOC represents a class of dynamical phenomena in the same way as for instance turbulence or chaos. This means that SOC should be loosely characterized by a set of mathematical properties, which enables us to recognize and identify the SOC properties when examining observation data from natural systems. However, contrary to the case of turbulence or chaos, the defining properties of SOC systems are not yet understood in a satisfactory way. In fact, when a natural system is proposed as an example of SOC, there is often a significant lack of mathematical evidence, and the arguments presented are often of a rather philosophical character.

It can also be problematic to distinguish SOC from other classes of dynamical behavior. In recently published articles [13, 14] evidence has been presented indicating that a large class of different complex systems exhibits global observables which are distributed according to the so called Bramwell-Holdsworth-Pinton (BHP) distribution. Among the quantities claimed to follow the BHP distributions are the toppling activity in the BTW sandpile and fluctuations of the global dissipation rate for bounded turbulent flows, and this has been taken as evidence that SOC and turbulence share some fundamental dynamical mechanisms. It turns out that the toppling activity in the BTW sandpile is *not* BHP distributed, but nevertheless, the work of Bramwell *et al.* is a good example of how little we actually understand of SOC dynamics. If we had better fundamental knowledge of the governing processes in sandpile models, then we would not base a comparison of SOC and turbulence on the single-time statistics of one observable quantity.

On the other hand, it is difficult to compare two complex systems which are described within different mathematical frameworks, and consequently a large number of different approaches to a description of SOC have been attempted. Scientists who want to compare SOC with turbulence have a tendency to place sandpile models into a framework where quantities have direct analogies in turbulence theory, for instance by defining Reynolds numbers for sandpiles. Inversely, there exist so-called shell models of turbulence where the energy cascades can be interpreted as avalanches.

The scaling properties of the self-organized critical states have also tempted scientist to compare SOC phenomena with phase transitions in equilibrium systems. Although this approach may have provided some useful analogies, the mathematical formalism of statistical physics is not very useful for attacking sandpiles. For instance, Ruelle's thermodynamic formalism for lattice systems is only effective when the equilibrium state is characterized by specified local interactions and only simple restrictions on the set of legal configuration. In sandpiles the set of recurrent configurations has a very complex structure which cannot be captured by local interactions, and from a mathematical point of view, the mechanism leading to self-organized criticality is fundamentally different from phase transitions. Other approaches inspired by statistical mechanics, such as the dynamical driven renormalization group calculations, have been more successful, but providing little general dynamical insight.

The main problem is that we obtain poor descriptions of SOC when enforcing frameworks which are specifically designed for completely different phenomena. It is actually quite evident that successful descriptions of sandpile dynamics must rely more on the specific properties of these models. At the same time it is important that the approaches do not become too specific. An example of this is Dhar's Δ-formalism [23], which depends crucially on the so-called abelian property, a non-generic property only found in particular sandpile models.

Two different frameworks

In this thesis we propose to investigate SOC within two different frameworks: *dynamical systems* and *stochastic differential equations*. We claim that both of these approaches are sufficiently general to include a large class of sandpile systems and to allow general comparison with other complex phenomena, yet sufficiently specific to capture non-trivial elements of SOC.

The dynamical systems approach allows us to relate SOC to the more established notion of chaos. In particular, well defined dynamical invariants can be computed, and these can be compared with other dynamical systems. An example is the dynamical entropy of sandpile models. The BTW and Zhang models [3, 57] both

have dynamical entropy equal to

$$\frac{\log N}{1 + \langle \omega \rangle / \langle \tau \rangle},$$

where N is the number of sites on the model's lattice, $\langle \omega \rangle$ is the average waiting time between avalanches and $\langle \tau \rangle$ is the average avalanche duration. This entropy is always positive, which implies chaos. But for slowly driven sandpiles (no feeding of sand grains during avalanches) the avalanche duration is much greater than the waiting time, and the entropy goes to zero as the number of sites on the lattice grows. For strongly driven sandpiles the entropy remains positive, illustrating a fundamental difference between strongly and weakly driven sandpiles.

The other approach presented in this thesis involves modeling the so-called toppling activity of sandpiles by a stochastic differential equation on the form

$$dX(t) = f(X)\, dt + \sigma \sqrt{X(t)}\, dB_H(t)\,.$$

Here $B_H(t)$ represents a stochastic process known as a fractional Brownian motion. This model is used to compute several critical exponents in the BTW and Zhang models. Moreover, we show explicitly how the stochastic differential equations can be used to illustrate essential differences between the toppling activity in a slowly driven sandpile, the toppling activity in a strongly driven sandpile, and the fluctuations of kinetic energy in a turbulent fluid.

The structure of the thesis

In the first part of the thesis we formulate the BTW and Zhang models as dynamical systems following a construction of Blanchard, Cessac and Krüger [6, 7, 8]. This involves modeling the random driving mechanism using a Bernoulli shift and describing the complete dynamics on an extended phase space $\Sigma_N^+ \times M$, where Σ_N^+ is the set of all infinite words over an alphabet consisting of the sites in the lattice. The emphasis of our investigation will be calculation of dynamical invariants that are useful for the characterization of chaos, such as dynamical entropy, Lyapunov exponents, attractor dimension and symbolic representations. Three research articles constitute chapters 3, 4, and 5. Chapter 3 focuses on the description of the Zhang model [57] as a dynamical system. Chapters 4 and 5 deal with specific mathematical results and ideas that where discovered through the work on the Zhang model. Chapter 2 is an introduction which contains definitions, elementary concepts and concrete examples essential for understanding the technical articles that follow.

The second part of the thesis presents the stochastic theory for the toppling activity. It consists of two research articles (chapters 7 and 8) and an introduction (chapter 6). The introduction includes a derivation of the stochastic model for the

so-called random neighbor model [20], a discussion of its generalization to the BTW and Zhang models and a brief survey of the required background on self-similar stochastic processes and stochastic differential equations.

Part 1:

Sandpile models and dynamical systems

Chapter 2

Basic dynamical properties of the BTW and Zhang models

Abstract: This chapter is an introduction to the first part of the thesis. Through concrete examples and calculations we show how the BTW and Zhang models can be formulated as dynamical systems. For the BTW model we preform symbolic coding in some simple situations. Characterization of sandpile models through dynamical invariants is discussed.

2.1 SOC and dynamical systems

We begin by formulating the BTW sandpile model as a dynamical system. Later in this chapter the formalism will be extended to the Zhang model.

2.1.1 The BTW model as a dynamical system

The BTW model is defined on a d-dimensional hyper-cubic lattice $\Lambda \subset \mathbb{Z}^d$ with $N = L^d$ sites (or nodes). Every site $i \in \Lambda$ is associated with a non-negative integer $z_i \in \{0, 1, 2, \dots\}$, which we can think of as the number of grains in position i of the lattice. The collection $z = \{z_i \,|i \in \Lambda\}$ is called the configuration of the sandpile, and the set of configurations $\{0, 1, 2, \dots\}^N$ is called the configuration space. A configuration is called stable if $z_i < 2d$ for all $i \in \Lambda$ and non-stable otherwise. The set of stable configurations is denoted M.

The time evolution of configurations is defined in two parts: If site i is overcritical it will topple. This means that if $z_i \geq 2d$, then site i looses one grain to each of its nearest neighbors on the lattice. Sites on the boundary of the lattice have less than $2d$ neighbors, so the total number of grains on the lattice is not conserved if there are overcritical boundary sites.

Remark 2.1 *The neighbors of i are defined as those sites j for which $d_\Lambda(i, j) = 1$, where d_Λ is the "Manhattan metric" on Λ.*

All the overcritical sites in a configuration will topple simultaneously in a time step, and the configuration evolves according to the rule $z \mapsto f(z)$, where $f : \{0, 1, 2, \dots\}^N \to \{0, 1, 2, \dots\}^N$ is a map defined by

$$(fz)_i = z_i - 2d\,\Theta(z_i - 2d) + \sum_{d_\Lambda(i,j)=1} \Theta(z_j - 2d)\,. \tag{2.1}$$

Here Θ is the unit step function defined by

$$\Theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 1 \end{cases}.$$

The notations and definitions become clearer by considering a concrete example:

Example: The simplest non-trivial sandpile we can consider is $d = 1$ and $N = 2$. In this case each configuration is an ordered pair $z = (z_1, z_2)$, where z_1 and z_2 are non-negative integers. There are four stable configurations: $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$. It is easy to see from equation 2.1 that $f(z) = z$ if z is stable. Through the

dynamics of the BTW model (which we have not defined completely yet), the only non-stable configurations that can be reached from a stable initial configuration are $(2,0)$, $(0,2)$, $(2,1)$ and $(1,2)$. Under f, these configurations are mapped as follows:

$$(2,1) \mapsto (0,2)\,,\ (1,2) \mapsto (2,0)\,,\ (0,2) \mapsto (1,0) \text{ and } (2,0) \mapsto (0,1)\,.$$

The BTW model (as given by equation 2.1) is defined so that grains "fall off" the lattice when there are overcritical sites on the boundary, and because of this we always reach a stable configuration after sufficiently many iterations of f. To initiate new avalanches we need to feed grains to the lattice, and this is usually done by adding (dropping) a grain to a randomly chosen site whenever the configuration is stable. Since this part of the dynamics involves randomness, we must eventually introduce an underlying probability space. But first we give a purely topological description.

The "trick" is to prefix all dropping positions in an infinite squence

$$\mathbf{t} = (t_1 t_2 t_3 \cdots),\ t_i \in \Lambda\,.$$

The first time the configuration is stable a grain is dropped onto site t_1, the next time the configuration becomes stable a grain is dropped onto site t_2, and so on. Equivalently, we can drop a grain in the position corresponding to the first entry of $\mathbf{t}$ whenever the configuration is stable, and then let $\mathbf{t} \mapsto \sigma(\mathbf{t})$, where

$$\sigma(t_1 t_2 t_3 \cdots) = (t_2 t_3 t_4 \cdots)\,.$$

The dynamics of the BTW sandpile can then be described by a single map:

$$\hat{f}(\mathbf{t}, z) = \begin{cases} (\sigma(\mathbf{t}), z + e_{t_1}) & \text{if } z \text{ is stable} \\ (\mathbf{t}, f(z)) & \text{if } z \text{ is non-stable} \end{cases}, \tag{2.2}$$

where e_i denotes the configuration with $(e_i)_j = \delta_{ij}$.

Since we are interested in the asymptotic properties of the dynamics, we can restrict our attention to the set of configurations which can be reached more than once in an orbit.[1] These configurations are called recurrent, and we denote set of recurrent configurations by $\mathcal{R}$. Our dynamical system becomes

$$\hat{f} : \Sigma_N^+ \times \mathcal{R} \to \Sigma_N^+ \times \mathcal{R}\,,$$

where $\hat{f}$ is given by equation 2.2. Here Σ_N^+ denotes the set of all infinite words over the alphabet Λ.

[1] This is a finite set.

2.1.2 Symbolic coding of the BTW model

For a positive integer n we let Σ_n^+ denote the set of all infinite sequences in the symbols $\{1, 2, \ldots, n\}$, and we let $\sigma : \Sigma_n^+ \to \Sigma_n^+$ denote the shift map defined in the previous section. A topological Markov chain is the restriction of the map σ to a particular shift-invariant subset of Σ_n^+. This subset is defined by a directed graph where the vertices are $\{1, 2, \ldots, n\}$. It consists of those sequences $(s_1 s_2 s_3 \cdots) \in \Sigma_N^+$ with the following property:

$$\forall k \geq 1 : \text{the graph has an arrow from } s_k \text{ to } s_{k+1}\,.$$

The directed graph can be represented by a $n \times n$ matrix $A = ||a_{ij}||$, where $a_{ij} = 1$ if the graph has an arrow from i to j, and $a_{ij} = 0$ otherwise. The subset of Σ_n^+ defined by the graph (or matrix) is denoted by Σ_A^+, and the restriction of σ to Σ_A^+ is denoted σ_A^+.

Remark 2.2 *The reason why topological Markov are so interesting, is that they serve as good models for a large class of chaotic dynamical systems. In fact, many properties of such systems can be proved by constructing a continuous change of variables (toplological conjugacy) which transform the system to a topological Markov chain.*

We now use an example to illustrate how the BTW model can be coded to a topological Markov chain:

Example: We continue to look at the case $d = 1$ and $N = 2$. There are seven recurrent configurations: $(1, 0)$, $(0, 1)$ $(0, 2)$, $(2, 0)$, $(2, 1)$, $(1, 2)$ and $(1, 1)$. We denote them by $z_1 \ldots, z_7$ in the order that they are written. Our phase space is then $\Sigma_2^+ \times \mathcal{R}$, where $\mathcal{R} = \{z_1, \ldots, z_7\}$.

We partition the phase space into fourteen subsets

$$[i] \times \{z_j\},\ i \in \{1, 2\}\,,\ j \in \{1, \ldots, 7\}\,,$$

where $[i] := \{\mathbf{t} \in \Sigma_2^+ \,|\, t_1 = i\}$ denotes the cylinder of the symbol i. Let

$$Z_1 = [1] \times \{z_1\}\,,\ Z_2 = [2] \times \{z_1\}\,,\ Z_3 = [1] \times \{z_2\}\,,\ \text{and so on.}$$

The next step is to determine the images of the partition elements Z_i. Consider the set $Z_1 = [1] \times \{z_1\}$. We know that $f(z_1) = z_4$, and secondly (since z_1 is a stable configuration) the shift map σ is applied. Since $\sigma_2^+([1]) = \Sigma_2^+$, we have

$$\hat{f}(Z_1) = \Sigma_2^+ \times \{z_4\} = \Big([1] \times \{z_4\}\Big) \cup \Big([2] \times \{z_4\}\Big) = Z_7 \cup Z_8\,.$$

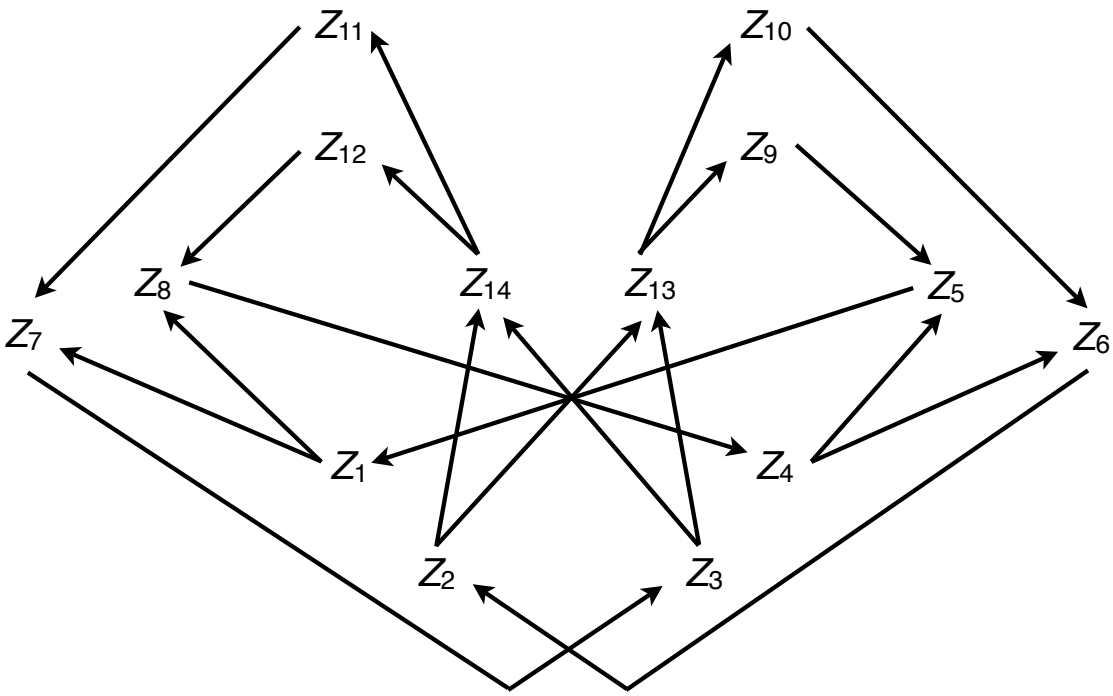


Figure 2.1: The Markov graph defined by the dynamics of the BTW model for $d = 1$ and $N = 2$.

In this case the partition element Z_1 is mapped to two different partition elements, Z_7 and Z_8. This is because z_1 is a stable configuration, so that a dropping is preformed and σ is applied. If we consider a partition element where the corresponding configuration is non-stable, then there is no shift, and the partition element is only mapped to one other partition element. For instance,

$$\hat{f}(Z_5) = \hat{f}([1] \times \{z_3\}) = [1] \times \{f(z_3)\} = [1] \times \{(1,0)\} = Z_1 \,.$$

We construct a directed graph by drawing an arrow from the symbol i to the symbol j if $f(Z_i) \cap Z_j \neq \emptyset$. This means that there will be two arrows from partition elements corresponding to stable sites and one arrow from partition elements corresponding to non-stable sites. The complete graph is shown in figure 2.1 . The matrix $A = ||a_{ij}||$ is given by the general formula

$$a_{ij} = \begin{cases} 1 & \text{if } \hat{f}(Z_i) \cap Z_j \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \,.$$

If we carry out all the computations we get:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix} . \tag{2.3}$$

Remark 2.3 *Topology on Σ_N^+ and Σ_A^+ can be defined by metrics on the form*

$$d(\mathbf{t}, \mathbf{s}) = a^{-k} \ , \ k = \min\{l \,|\, t_l \neq s_l\} \, ,$$

where $a > 0$ is some positive real number. The number a is not essential for the topology defined by these metrics, but other properties, like the Hausdorff dimensions of subsets depend essentially on a.

With topologies in place, the equivalence between the maps $\hat{f}$ and σ_A^+ can be described as a topological conjugacy. This means that there exists a homeomorphism $g : \Sigma_N^+ \times \mathcal{R} \to \Sigma_A^+$ such that $g \circ \hat{f} = \sigma_A^+ \circ g$. Such a coding map always exists for the BTW model, and in general it is given by the rule $g(\mathbf{t}, z) = \mathbf{s}$, where $\mathbf{s} = (s_1 s_2 s_3 \cdots) \in \Sigma_A^+$ is the sequence defined by $\hat{f}^n(\mathbf{t}, z) \in Z_{s_n}$ $\forall n \geq 1$.

2.1.3 The natural invariant probability measure

Usually the BTW model is defined such that whenever the configuration is stable, a site $i \in \Lambda$ is chosen randomly with respect to the uniform probability distribution on Λ, i.e. site $i \in \Lambda$ has probability $p_i = N^{-1}$ of being chosen. The choice of dropping position at one time step is independent of the choice of dropping position at previous times, so the uniform probability distribution $p_1 = \cdots = p_N = N^{-1}$ extends to a uniform Bernoulli measure on Σ_N^+. The probability that a dropping sequence begins with $i_1 i_2 \cdots i_k$ is equal to $p_{i_1} p_{i_2} \cdots p_{i_k} = N^{-k}$. Written as a probability measure, this becomes

$$\mu_{\mathrm{Ber}}([i_1 \cdots i_k]) = p_{i_1} \cdots p_{i_k} = N^{-k} \, ,$$

where $[i_1 \cdots i_k] = \{\mathbf{t} \,|\, t_1 = i_1, \ldots, t_k = i_k\}$ is the cylinder of the finite sequence $i_1, \ldots, i_k$.

The uniform Bernoulli measure μ_{Ber} induces a Markov measure μ on the topological Markov chain Σ_A^+. This measure can be described by a transition probability matrix $\Pi = ||\pi_{ij}||$. This matrix has the same dimensions as the matrix A, and its entries represent the probabilities of transitions from Z_i to Z_j.

If $a_{ij} = 0$, then $\pi_{ij} = 0$, since the probability of an impossible transition must be zero. Also, if row i of A only contains one non-zero element, then $\pi_{ij} = 1$. If row i of A has more than one non-zero element, then the partition element Z_i corresponds to a stable configuration. This implies that the shift map is applied, so if $Z_i = [j] \times \{z\}$, then

$$\hat{f}(Z_i) = \hat{f}([j] \times \{z\}) = \Sigma_N^+ \times \{f(z)\} = \bigcup_{j' \in \Lambda} [j'] \times \{f(z)\} \,.$$

Since $\hat{f}(Z_i)$ intersects N different partition elements, row i of A has N different 1-s. Each of these correspond to a different position of dropping, and since each position has the same probability, the corresponding entries in Π are equal to N^{-1}. This gives the following formula:

$$\pi_{ij} = \begin{cases} 0 & \text{if } a_{ij} = 0 \\ 1 & \text{if } a_{ij} = 1 \text{ and } \sum_j a_{ij} = 1 \\ N^{-1} & \text{if } a_{ij} = 1 \text{ and } \sum_j a_{ij} = N \end{cases} . \tag{2.4}$$

Let $Q = N \times |\mathcal{R}|$, so that A and Π are $Q \times Q$ matrices. If we choose an initial partition element Z_i with respect to a probability distribution $q = (q_1, \ldots, q_Q)^T$, then the probability of $\hat{f}(\mathbf{t}, z) \in Z_j$ is

$$\begin{aligned} \text{Prob}[\hat{f}(\mathbf{t}, z) \in Z_j] &= \sum_{i=1}^{Q} \text{Prob}[(\mathbf{t}, z) \in Z_i] \cdot \text{Prob}[\hat{f}(\mathbf{t}, z) \in Z_j | (\mathbf{t}, z) \in Z_i] \\ &= \sum_{i=1}^{Q} q_i \, \pi_{ij} = (\Pi^T q)_j = (\mathcal{L} q)_j \,, \end{aligned}$$

where $\mathcal{L} = \Pi^T$. By repeating the argument we see that $\text{Prob}[\hat{f}^n(\mathbf{t}, z) \in Z_j] = (\mathcal{L}^n q)_j$, and thus a stable probability vector P is reached as the limit

$$P = \lim_{n \to \infty} \mathcal{L}^n q \,.$$

If this limit exists, then P must be a normalized egienvector of $\mathcal{L}$.

If the matrix A is mixing, which simply means that there exists an integer k such that all the elements of A^k are positive, then the Perron-Frobenius theorem

ensures that $\mathcal{L}$ has a unique positive and normalized eigenvector P and that $\mathcal{L}^n q \to P$ independently of q. In other words, the asymptotic probability distribution of partition elements is independent of the initial distribution.

Through the probability vector P we can now define the Markov measure μ:

$$\mu([s_1 \dots s_n]) = P_{s_1}\, \pi_{s_1 s_2} \cdots \pi_{s_{n-1} s_n}\,.$$

The corresponding probability meausre $g_*^{-1}\mu$ on $\Sigma_N^+ \times \mathcal{R}$ is such that its projection to Σ_N^+ coincides with the uniform Bernoulli measure μ_{Ber}.

Example: We continue illustrate the theory by looking at the simple example $d = 1, N = 2$. In the last section we found the matrix A, and from equation 2.4 we see that

$$\Pi = \frac{1}{2}\begin{pmatrix}
0&0&0&0&0&0&1&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1\\
0&0&0&0&0&0&0&0&0&0&0&0&1&1\\
0&0&0&0&1&1&0&0&0&0&0&0&0&0\\
2&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&2&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&2&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&2&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&2&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&2&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&2&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&2&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&1&0&0
\end{pmatrix}.$$

In this example the matrix A is not mixing. However, the matrix $\mathcal{L} = \Pi^T$ has a unique positive and normalized eigenvector P corresponding to the eigenvalue $\lambda = 1$. A simple computation shows that

$$P = \frac{1}{24}\,(2,2,2,2,2,2,2,2,1,1,1,1,2,2)^T\,,$$

which means that the partition elements Z_9, Z_{10}, Z_{11} and Z_{12} have probability $1/24$, and the other partition elements have probability $1/12$. In particular, the configurations $(2,1)$ and $(1,2)$ have probability $1/12$, and the five other configurations have probability $1/6$.

Notice that all stable configurations have the same probability, in this case $1/6$. This is always true for the BTW model, a mathematical result first proved by Dhar [23]. We will return to this result in section 2.1.8.

2.1.4 Characterization of chaos

As we have seen, the symbolic coding $g : \Sigma_N^+ \times \mathcal{R} \to \Sigma_A^+$ gives us a complete topological and metric (statistical) description of the dynamics of the BTW model. In principle we can compute all desired information about the dynamics by determining the matrix A and its properties. However this is not possible in practice since the size of A will grow very rapidly as we increase N. On the other hand, there is some general information about the dynamics which can be deduced without knowing the detailed form of the matrix A.

As we have already mentioned, topological Markov chains serve as prototypes of chaotic dynamics, and since the BTW model is conjugated to σ_A^+ it is natural to believe that it exhibits chaotic behavior. On the other hand, chaos is characterized by a rapid loss of memory along orbits, contrary to the long range correlations we expect to see along avalanches in sandpiles. In order to resolve this apparent paradox we will quantify the degree of chaos in the BTW model. The results show that sandpile dynamics can be described as *weakly* chaotic, with predictability on time scales comparable to the durations of system size avalanches.

Before we start our analysis of the chaotic properties of the BTW model we will briefly discuss the dynamical invariants usually used to describe chaotic behavior.

Chaos is usually characterized by sensitivity to initial conditions. This means that in any small neighborhood in phase space we can find two initial conditions which orbits diverge essentially from each other after a finite number of iterations. This property is rather imprecise, so several mathematical invariants have been developed to quantify this effect.

An example is the Lyapunov exponents. Let M be a manifold (not to be confused with the set of stable configurations in the BTW model) and $G : M \to M$ a smooth map. For any $x \in M$ and any tangent vector $v \in T_x M$ one defines

$$\chi(x, v) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log ||d_x f^n(v)|| \,.$$

It is very easy to show that for each $x \in M$ there exists only $\dim(M)$ different values of χ. These numbers are called the Lyapunov exponents of the map G along the orbit starting at x. They are denoted $\chi_k(x)$. Often there exists a natural invariant probability measure on M for which the numbers $\chi_k(x)$ are constant almost everywhere. In these situations we have a set of x-independent exponents. A positive Lyapunov exponent indicates that the map G (on average) has expanding properties in at least one (varying) direction. This expansion causes sensitivity to initial conditions and hence chaos.

Another characterization of chaos is given by dynamical entropies. There are two kinds of entropy, metric and topological. Just as the Lyapunov exponents, the metric entropy of a map G depends on an invariant probability measure μ. If

$\xi = \{Z_i\}$ is a partition of the phase space, then the quantity

$$H_\mu(\xi) = -\sum_i \mu(Z_i) \log \mu(Z_i)$$

is a measure of the (logarithmic) amount of uncertainty we are faced with in trying to guess which partition element a point (chosen randomly with respect to μ) belongs to. We then use the dynamical system to refine the partition by constructing

$$\xi_n = \xi \vee G^{-1}(\xi) \vee \cdots \vee G^{-n+1}(\xi)\,,$$

where $G^{-1}(\xi)$ denotes the partition consisting of the sets $G^{-1}(Z_i)$ and $\vee$ denotes the join of partition, i.e. $\{Z_i\} \vee \{X_i\} = \{Z_i \cap X_j\}$. Knowing the position of a point with respect to this partition is the same as knowing the ξ-address of an all the points in an orbit segment of length n. Hence $H_\mu(\xi_n)$ must be greater than or equal to $H_\mu(\xi)$. In particular, if the dynamics is chaotic and unpredictable, then $H_\mu(\xi_n)$ should grow with n. We define

$$h_\mu(G;\xi) = \overline{\lim_{n\to\infty}} \frac{1}{n} H_\mu(\xi_n)\,.$$

This quantity measures the exponential asymptotic growth rate of $\exp H_\mu(\xi_n)$ as $n \to \infty$. The metric entropy of G with respect to the measure μ is defined as $h_\mu(G) = \sup_\xi h_\mu(G;\xi)$, where the supremum is taken over all "good" partitions of the phase space. This supremum is attained for any partition ξ for which the diameter of ξ_n tends to zero as $n \to \infty$. A dynamical system $G : M \to M$ is chaotic with respect to the invariant measure μ if $h_\mu(G) > 0$.

Remark 2.4 *For smooth maps there is a close relationship between the metric entropy and the set of positive Lyapunov exponents. In fact, if G is a $C^{1+\epsilon}$ diffeomorphism and μ a smooth invariant probability measure, then positive metric entropy is equivalent to the existence of a positive Lyapunov exponent. This follows from the Pesin entropy formula*

$$h_\mu(G) = \int \sum_{\chi_i(x)>0} \chi_i(x)\, d\mu(x)\,,$$

where the sums are taken over the positive Lyapunov exponents $\chi_i(x)$. If a "good" invariant measure μ is not smooth, for instance if it is supported on a fractal attractor, then it is typically absolutely continuous on the so called unstable manifolds. These measures are called SRB measures, and for them, the Ledrappier-Young formula states

$$h_\mu(G) = \int \sum_{\chi_i(x)>0} \chi_i(x)(d_\mu^i(x) - d_\mu^{i-1}(x))\, d\mu(x)\,,$$

where the numbers $d^i_\mu(x)$ are pointwise fractal dimensions in the point x of the the measure μ on the i-th leaf of the unstable manifold. For non-invertible smooth maps the equality in Pesin's formula does not always hold, but we have the Margulis-Ruelle inequality

$$h_\mu(G) \leq \int \sum_{\chi_i(x)>0} \chi_i(x)\, d\mu(x)\,.$$

Although metric entropy is one of the best existing tools for characterization of chaos, it is important to keep in mind that it depends essentially on the choice of invariant measure. Any chaotic system has an infinite number of periodic orbits, and if we choose a measure supported on one of these orbits, then the metric entropy is zero. We can only assign physical meaning to the notion of metric entropy if the corresponding invariant measure reflects the "right" statistical properties of the dynamics. Such measures are characterized by the property that time-averages equal ensemble averages, meaning that we for any continuous function ϕ have

$$\frac{1}{n}\sum_{k=0}^{n-1} \phi(G^k x) \to \int \phi\, d\mu\,, \tag{2.5}$$

for μ_{Leb}-almost all initial conditions x. These "good" measures are called *natural* invariant measures (or SRB measures if the dynamics is defined by a hyperbolic diffeomorphism). It is important not to confuse the natural invariant measures with ergodic measures. The famous Birkhoff ergodic theorem only asserts that equation 2.5 holds for μ-almost all initial conditions x.

Whenever a dynamical system admits a natural invariant measure μ, then this measure captures all the relevant statistical properties of the dynamics, and then $h_\mu(G) > 0$ implies chaotic dynamics.

Another (and closely related) tool for characterization of chaos is topological entropy. Contrary to metric entropy, topological entropy does not depend on an invariant measure, and it is defined as long as the phase space X is a metrizable topological space. Given a metric d (compatible with the topology) we define iterated balls

$$B_{\epsilon,n}(x) = \{y \in X \,|\, d(f^k x, f^k y) < \epsilon \text{ for } k = 0, \ldots, n-1\}\,.$$

The set $B_{\epsilon,n}(x)$ consists of the points in phase space whose orbit segments of length n lie within an ϵ-tube of the orbit segment $x, f(x), \ldots, f^{n-1}(x)$ in X^n. Let $N(\epsilon, n)$ be the number of such balls needed to cover the space space X. In other words, $N(\epsilon, n)$ is the number of ϵ-distinguishable orbit segments of length n. In chaotic systems, this number grows exponentially with n, and topological entropy is defined as the exponential growth rate:

$$h_{\text{top}}(G) = \lim_{\epsilon\to 0} \overline{\lim_{n\to\infty}} \frac{1}{n} \log N(\epsilon, n)\,.$$

The topological entropy depends on the topology only, i.e. the choice of (compatible) metric d is irrelevant. For continuous maps the topological entropy is related to the metric entropy via the variational principle:

$$h_{\text{top}}(G) = \sup_{\mu} h_{\mu}(G)\,,$$

where the supremum is taken over all invariant Borel probability measures μ. The variational principle tells us that positive topological entropy implies that the metric entropy is positive for some invariant measure. For a large class of dynamical systems the supremum is attained, in which case we have a measure of maximal entropy. Typically the measure of maximal entropy is unique, and it often coincides with the natural invariant measure.

Finally we mention that there exits other topological characterizations of chaos. First, the set of periodic points is typically dense in the attractor. Also, the number of periodic points of period n grow exponentially with n, usually with a rate equal to the topological entropy. Other typical properties include topological transitivity in the attractor (there exists a dense orbit on the attractor) and topological mixing, which is a stronger property than transitivity.

2.1.5 Chaos in the BTW model

We know that the BTW model is topologically conjugated to a topological Markov chain $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$. Topological entropy, growth in the number of periodic points, transitivity and topological mixing are all invariant under this conjugacy, and hence it suffices to study the map σ_A^+ in order to calculate these invariants. Fortunately this is an easy task given the transition matrix A. In particular, the topological entropy is equal to the logarithm of spectral radius of A. The reason for this is that the number of sequence segments of length n starting with the symbol i and ending with the symbol j is equal to the matrix entry $(A^n)_{ij}$. This number represents the number of (n, ϵ)-balls (for $\epsilon = a^{-1}$) needed to cover Σ_A^+. This means that $N(n, \epsilon)$ is equal to $||A^n||_1$, where $|| \cdot ||_1$ denotes the matrix norm defined as the sum of the absolute values of all the entries. In finite dimensions we can always change to an equivalent matrix norm which is arbitrarily well approximated by the spectral radius $\rho(A)$. Hence $N(n, \epsilon) \sim \rho(A)^n$, and $h_{\text{top}}(\hat{f}) = h_{\text{top}}(\sigma_A^+) = \log \rho(A)$.

There is also a simple formula which relates the metric entropy of the Markov measure μ to the matrix Π:

$$h_{\mu}(\sigma_A^+) = -\sum_{ij} P_i \pi_{ij} \log \pi_{ij}\,.$$

Using equation 2.4 we see that

$$\sum_i P_i \Big(-\sum_j \pi_{ij} \log \pi_{ij} \Big) = \sum_i P_i\, \Gamma_i = \langle \Gamma \rangle\,,$$

where $\Gamma_i = \log N$ if the partition element Z_i corresponds to a stable site, and $\Gamma_i = 0$ otherwise.

It is also easy to see that $\langle \Gamma \rangle = \log(N) \cdot \text{Prob}[z \in M]$ where $\text{Prob}[z \in M]$ is the probability that a configuration is stable. This probability can be expressed via avalanche observables. Under the dynamics, any stable configuration will become unstable after some time. When this happens an avalanche is initiated. The avalanche terminates when the configuration is stable again. The duration of this avalanche, i.e. number of subsequent time steps for which the configuration is unstable, is denoted τ. After an avalanche there is a quiet period before new avalanche starts. The duration of the quiet period is called the waiting time and it is denoted ω. Since configurations are stable during quiet periods and unstable during avalanches, it is clear that the probability that configuration is stable is given by the ratio of the mean waiting time to the mean duration of a combined quiet period and avalanche. That is:

$$\text{Prob}[z \in M] = \frac{\langle \omega \rangle}{\langle \omega \rangle + \langle \tau \rangle} = \frac{1}{1 + \langle \tau \rangle / \langle \omega \rangle} \, .$$

Hence

$$h_\mu(\hat{f}) = \frac{\log N}{1 + \langle \tau \rangle / \langle \omega \rangle} \, . \tag{2.6}$$

Example: For $d = 1$ and $N = 2$, the transition matrix given by equation 2.3 has spectral radius $\rho(A) = \sqrt{2}$. Hence the topological entropy of the system is $h_{\text{top}}(\hat{f}) = \frac{1}{2} \log 2$. We can also compute the metric entropy $h_\mu(\hat{f})$ by using equation 2.6. We have seen that there are three stable configurations, each with probability 1/6. This implies that that the probability of a stable configuration is 1/2, and thus $h_\mu(\hat{f}) = \frac{1}{2} \log 2$.

In the example above we saw that the topological entropy was equal to the metric entropy $h_\mu(\hat{f})$. This is always the case, and it means that μ always is a measure of maximal entropy. When the matrix A is mixing this result is very easy to prove by using a general result which states that the measures of maximal entropy on topological Markov chains are Markov measures [32]. In fact, there is an algorithm for constructing these measures, and by following this procedure one easily proves the claim. Thus we have that

$$h_{\text{top}}(\hat{f}) = h_\mu(\hat{f}) = \frac{\log N}{1 + \langle \tau \rangle / \langle \omega \rangle} \, .$$

Note that $N(\epsilon, n)$ grows like $\sim \exp(n/\xi)$, where

$$\xi = \frac{1 + \langle\tau\rangle/\langle\omega\rangle}{\log N} .$$

The number ξ represents an important time scale in the BTW model. On time scales comparable with ξ (or longer), the BTW model behaves chaotically.

2.1.6 Decay of correlations

Another way of characterizing the loss of predictability in time is using auto-correlation functions. Suppose that $\phi : \Sigma_N^+ \times \mathcal{R} \to \mathbb{R}$ is a function which is constant on the partition elements $Z = [j] \times \{z\}$. Then ϕ can be viewed as a function $\phi : \Sigma_A^+ \to \mathbb{R}$, which is constant on cylinders of length one, i.e. $\phi(\mathbf{s}) = s_1$. We can then think of ϕ as a vector with Q entries, $\phi = (\phi_1, \ldots, \phi_Q)^T$. The auto-correlation function with respect to the function ϕ is

$$C_\phi(n) = \langle \phi \cdot (\phi \circ (\sigma_A^+)^n) \rangle - \langle \phi \rangle^2 ,$$

and when the averages are taken with respect to the natural invariant measure μ, we get

$$\begin{aligned} C_\phi(n) &= \sum_{s_1,\ldots,s_n} P_{s_1} \pi_{s_1 s_2} \cdots \pi_{s_{n-1} s_n} \phi_{s_1} \Big[\phi_{s_n} - \sum_j P_j \phi_j \Big] \\ &= \sum_{s_1} P_{s_1} \phi_{s_1} \sum_{s_n} \Big[\phi_{s_n} - \sum_j P_j \phi_j \Big] (\Pi^n)_{s_1 s_n} \\ &= \sum_{s_1} P_{s_1} \phi_{s_1} (\Pi^n \psi)_{s_1} = \langle \phi \cdot \Pi^n \psi \rangle , \end{aligned}$$

where $\psi = \phi - \langle \phi \rangle$.

Define the $Q \times Q$ matrix $J = [P, \ldots, P]^T$ where all rows equal to the probability vector P. Since $\langle \psi \rangle = \sum_i P_i \psi_i = 0$ we have $J\psi = 0$. Thus we can write

$$C_\phi(n) = \langle \phi \cdot (\Pi^n - J)\psi \rangle .$$

Let $v = (1, \ldots, 1)^T \in \mathbb{R}^Q$. Since the sum over each row of Π^n and each row of J is equal to one, we have $\Pi^n v = Jv = v$. If we assume that the matrix A is mixing, then the Perron-Frobenius theorem ensures that the matrix Π has a unique maximal eigenvalue $\lambda_1 = 1$, which has multiplicity one and a positive corresponding eigenvector which is proportional to v. The matrix J has rank one, so it is easy to express the eigenvalues of $\Pi^n - J$ in terms of the eigenvalues of Π. If $\lambda_1 = 1, \lambda_2, \ldots, \lambda_Q$ is the spectrum of Π, then $\lambda_1 = 0, \lambda_2^n, \ldots, \lambda_Q^n$ are the eigenvalues

of $\Pi^n - J$. Indeed, if R is the coordinate change matrix that brings Π to its Jordan block form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & B_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_m \end{pmatrix},$$

then

$$RJR^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

and

$$\Pi^n - J = R^{-1} \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & B_1^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_m^n \end{pmatrix} R\,.$$

If the matrix Π is semi-simple the auto-correlation function can be written as

$$C_\phi(n) = \sum_{i=2}^{Q} c_i(\phi)\, \lambda_i^n\,,$$

where $c_i(\phi)$ are some constants depending on the function ϕ. Since the Perron-Frobenius theorem ensures that $|\lambda_i| < 1$ for $i \neq 1$, we have exponential decay of the auto-correlation function. In particular, the slowest mode of correlation decay is given by the second largest eigenvalue (in absolute value) λ_2. The same is true even if the matrix Π is not semi-simple. In this case we can bring Π to a Jordan block form.

2.1.7 Return maps

The dynamical system associated with the BTW model can be significantly simplified if we consider the return maps to the set of stable recurrent configurations. This means that we compress the duration of avalanches to one time step by defining maps

$$F_i(z) = \hat{f}^{m(i,z)}(z + e_i)\,,$$

where $m(i,z) = \min\{n > 0 \,|\, \hat{f}^n(z + e_i) \in M\}$. This defines a dynamical system $F : \Sigma_N^+ \times \mathcal{R} \to \Sigma_N^+ \times \mathcal{R}$ by the formula

$$F(\mathbf{t}, z) = (\sigma_N^+ \mathbf{t}, F_{t_1}(z))\,.$$

Remark 2.5 *For some natural phenomena exhibiting SOC the avalanches occur on time scales which are much shorter than the characteristic time scale of the driving mechanism. For these systems the formulation of the return maps is not just a mathematical trick, but a physically relevant reformulation of the model.*

As with the map $\hat{f}$, the dynamical system defined by F admits symbolic coding with respect to a partition with elements on the form $Z_i = [j] \times \{z\}$. For the return map F there is random dropping in every time step, so the matrix A has N 1-s in all rows. The matrix Π has exactly N non-zero elements in each row, all equal to $1/N$. Hence the metric entropy is

$$h_\mu(F) = \sum_{ij} P_i \pi_{ij} \log \pi_{ij} = \sum_i P_i \log N = \log N \,.$$

Since $A(1, \ldots, 1)^T = N(1, \ldots, 1)^T$, the maximal eigenvalue of A is equal to N, and hence the topological entropy is $h_{\text{top}}(F) = \log N$.

2.1.8 Dhar's theory for return maps in the BTW model

In this section we briefly discuss some important results of D. Dhar [23]. He used the so called Abelian property of the BTW model to show that if we restrict to the set of stable recurrent configurations, then each map F_i is invertible. This means that for a given a partition element $Z' = [i'] \times \{z'\}$ and map F_i, there exists a unique partition element $Z = [i] \times \{F_i^{-1}(z)\}$ whose pre-image under F_i intersects Z'. Thus there are only N different partition elements that can be mapped to Z', and the column of A corresponding to the partition element Z' has exactly N 1-s. It follows that the the matrix $\mathcal{L} = \Pi^T$ has N non-zero elements in each row, each of them equal to $1/N$. Then the $P = (1/N, \ldots, 1/N)$ satisfies $\mathcal{L}P = P$. In other words: All stable recurrent configurations have the same probability.

If the matrix A is mixing, then the Perron-Frobenous theorem ensures that P is the only invariant probability vector and that $\mathcal{L}^n q \to P$ (for positive q) with exponential speed. However, there is no general result that ensures mixing of the matrix A, and there are examples where $\mathcal{L}^n q$ does not converge to P for all q. On the other hand, in all of these examples there is coincidence between the time average of almost all initial conditions and the ensemble averages taken with respect to the Markov measure given by P.

Another important result of Dhar is that the number of recurrent configurations is equal to the absolute value of the determinant of a $N \times N$ matrix Δ. This matrix is defined by $\Delta_{ii} = -2d$, $\Delta_{ij} = 1$ if i and j are nearest neighbors in the lattice and $\Delta_{ij} = 0$ otherwise.

One of the applications of this result is construction of the matrix A using a computer. The idea is simply to start with a configuration which we know is

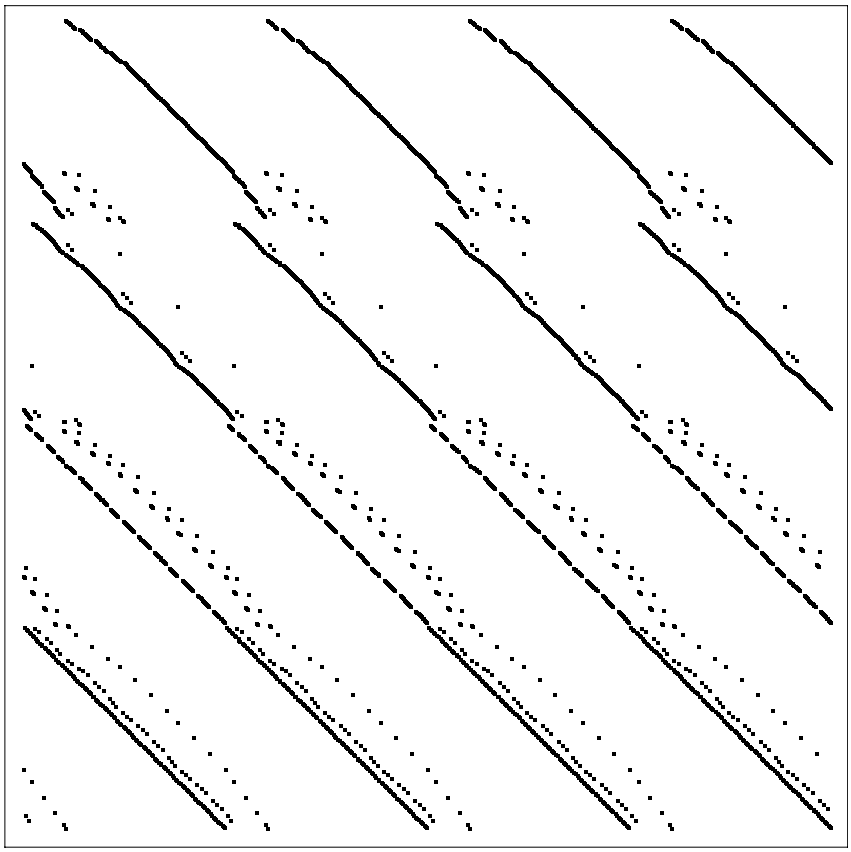


Figure 2.2: The 1-s in the matrix A for the BTW model with $d = 2$ and $N = 4$.

recurrent (for instance the configuration with all entries equal to $2d - 1$). We can simulate the dynamics of the model until we have collected $|\det \Delta|$ different stable configurations. We then know that we have all the stable recurrent configurations, and we can apply the maps F_i to each of them to construct A. The spectrum of A will provide us with important information about the correlations between avalanches.

Example: We consider the map F for the case $d = 2$ and $N = 4$. In this case the Δ-matrix is

$$\Delta = \begin{pmatrix} -4 & 1 & 1 & 0 \\ 1 & -4 & 0 & 1 \\ 1 & 0 & -4 & 1 \\ 0 & 1 & 1 & -4 \end{pmatrix}.$$

Thus the number of recurrent configurations is $|\det \Delta| = 192$. This gives $Q = 4 \cdot 192 = 768$ partition elements on the form $Z = [i] \times \{z\}$, and the matrix A

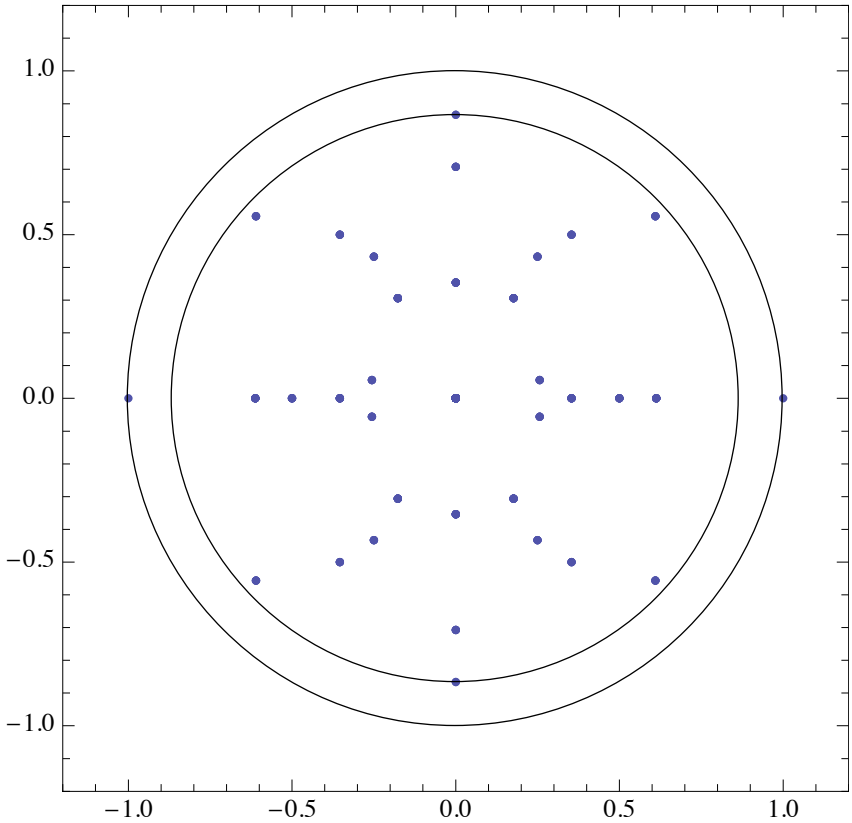


Figure 2.3: The spectrum of the matrix Π for the BTW model with $d = 2$ and $N = 4$.

has dimensions 768×768. We construct A by computing $F_i(z)$ for all $z \in \mathcal{R}$ and $i = 1, \ldots, N$. The non-zero elements of the matrix A are shown i figure 2.2. From the matrix A it is easy to construct the matrix Π, and the spectrum of Π is shown in figure 2.3. There are two eigenvalues with absolute value equal to one, 1 and -1. This implies that the matrix is not mixing. In fact the set of stable recurrent configurations can be divided into two disjoined sets ($\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$) which are of equal cardinality. The dynamics of the model alternates between these two sets with period 2, giving rise to the eigenvalue -1.

Due to the alternation between the two sets $\mathcal{R}_1$ and $\mathcal{R}_2$ some observables $\phi(\mathbf{s}) = \phi(s_1)$ (those who have components in the direction of the eigenvector corresponding to the eigenvalue -1) can give rise to auto-correlation functions that do not decay to zero. However, this correlation vanishes if we consider the map F^2 rather than F. Hence we can view this as a trivial correlation. For observables $\phi(\mathbf{s}) = \phi(s_1)$ that are not affected by this periodicity, the rate of decay in the autocorrelation function is no slower than $\sim r^n$, where $r = \max\{|\lambda| \,|\, \lambda \in \mathrm{Sp}(\Pi) \setminus \{-1, 1\}\}$. In our example we compute that $r = \sqrt{3}/2$.

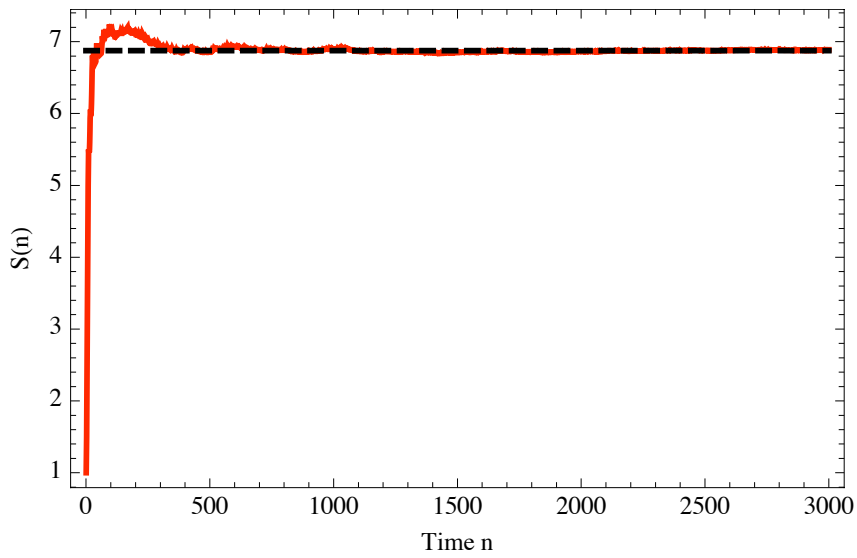


Figure 2.4: Convergence of the time average (for total number of grains) to the ensemble average 55/8 for the BTW model with $d = 2$ and $N = 4$.

Finding all the stable recurrent configurations also implies that we easily can compute average quantities. For instance, the average value of the total number of grains in a stable recurrent configuration is 55/8. If we simulate the dynamics we can compare time averages

$$S_n = \frac{1}{n} \sum_{k=0}^{n-1} \phi(F^k(t, z))$$

with our ensemble average. The convergence of S_n to 55/8 is shown in figure 2.4.

2.2 The Zhang model

The Zhang model is similar to the BTW model, yet different in two essential aspects. The first difference is that the occupation number of a site is allowed to be any real number, not just integers. This means that a configuration $z = \{z_i\}_{i \in \Lambda}$ is a N-tuple in $\mathbb{R}^N$ rather than $\mathbb{Z}^N$. The occupation number of a site in the Zhang model is usually called the site's energy.

The other difference is in the relaxation rule. In the BTW model, a site always looses a fixed number of grains as it topples, independently of the actual occupation number (as long as it exceeds the threshold). In the Zhang model the amount of energy a site looses is a fixed proportion of its energy. As site i topples its energy

is reduced to ϵz_i and $(1-\epsilon)z_i$ is distributed in equal parts to its nearest neighbors. Here $\epsilon \in (0,1)$ is a fixed parameter of the model. This defines a map $f : \mathbb{R}^n \to \mathbb{R}^N$ given by

$$(fz)_i = z_i - (1-\epsilon)z_i\Theta(z_i - E_c) + \frac{1-\epsilon}{2d} \sum_{d_\Lambda(i,j)=1} z_j\Theta(z_j - E_c)\,.$$

The parameter $E_c > 0$ is called the critical threshold energy.

The random dropping is preformed in the same way as in the BTW model, and this can be described by extending the phase space to $\Sigma_N^+ \times \mathbb{R}^N$. As in the BTW model we define return maps $F_i : M \to M$, where $M = [0, E_c]^N$ is the set of stable configurations. This gives us the dynamical system $F : \Sigma_N^+ \times M \to \Sigma_N^+ \times M$, with

$$F(\mathbf{t}, z) = (\sigma_N^+(\mathbf{t}), F_{t_1}(z))\,.$$

Example: In this example we will find formulas for the maps F_i in the case $d = 1$, $N = 2$ and $E_c \geq (1+\epsilon)/(1-\epsilon)$.

Let $z = (z_1, z_2) \in M$ be a stable configuration. Avalanches start with the addition of 1 to a site. If this is added to the first site, then $z \mapsto z + e_1 = (z_1+1, z_2)$. This configuration is stable if $z_1 \leq E_c - 1$. Denote $M_{11} = \{z \in M \mid z_1 \leq E_c - 1\}$.

The configuration $z + e_1$ is non-stable for $z_1 > E_c - 1$. Applying the map f to $z + e_1$ we get

$$f(z + e_1) = \Big(\epsilon(z_1+1), z_2 + \frac{1-\epsilon}{2}(z_1+1)\Big)\,.$$

We have assumed that $E_c \geq (1+\epsilon)/(1-\epsilon) \geq \epsilon/(1-\epsilon)$, which implies that $\epsilon(E_c+1) \leq E_c$, and hence

$$\epsilon(z_1+1) \leq \epsilon(E_c+1) \leq E_c\,.$$

This means that the first component of $f(z+e_1)$ always is less than E_c. The second component

$$z_2 + \frac{1-\epsilon}{2}(z_1+1)$$

can be both greater than E_c and less than E_c, depending on z. It is less than E_c in the set

$$M_{12} = \Big\{z \in M \mid z_2 + \frac{1-\epsilon}{2}(z_1+1) \leq E_c\Big\}\,,$$

and for $z \in M_{12}$ we have

$$F_1(z) = \Big(\epsilon(z_1+1), z_2 + \frac{1-\epsilon}{2}(z_1+1)\Big)\,.$$

Applying the map f to the configuration above yields

$$\Big(\Big(\frac{1+\epsilon}{2}\Big)^2(z_1+1)+\frac{1-\epsilon}{2}\,z_2,\epsilon z_2+\epsilon\frac{1-\epsilon}{2}(z_1+1)\Big)$$

Since $E_c \geq (1+\epsilon)/(1-\epsilon)$ we have $(1-\epsilon)E_c \geq 1+\epsilon$, and hence the following estimate for the first component:

$$\begin{aligned}\Big(\frac{1+\epsilon}{2}\Big)^2(z_1+1)+\frac{1-\epsilon}{2}\,z_2 &\leq \Big(\frac{1+\epsilon}{2}\Big)^2(E_c+1)+\frac{1-\epsilon}{2}\,E_c\\ &= E_c+\frac{1+\epsilon}{4}\Big((1+\epsilon)-(1-\epsilon)E_c\Big)\leq E_c\,.\end{aligned}$$

The inequality $E_c \geq (1+\epsilon)/(1-\epsilon)$ is equivalent to

$$\epsilon \leq \frac{E_c-1}{E_c+1}\,,$$

so the second component can be estimated by

$$\begin{aligned}\epsilon z_2+\epsilon\frac{1-\epsilon}{2}(z_1+1) &\leq \frac{E_c-1}{E_c+1}E_c+\frac{1}{2}\frac{E_c-1}{E_c+1}\Big(1-\frac{E_c-1}{E_c+1}\Big)(E_c+1)\\ &= E_c-1\leq E_c\,.\end{aligned}$$

This shows that it is impossible to have non-stable configurations for more than two subsequent time steps in this example.

The map F_1 has the formula

$$F_1(z)=\begin{cases}T_{11}(z+e_1) & \text{for } z\in M_{11}=\Big\{z\in M\,|\,z_1+1\leq E_c\Big\}\\ T_{12}(z+e_1) & \text{for } z\in M_{12}=\Big\{z\in M\setminus M_{11}\,|\,\epsilon z_2+\frac{1-\epsilon}{2}(z_1+1)\leq E_c\Big\}\\ T_{13}(z+e_1) & \text{for } z\in M_{13}=M\setminus(M_{11}\cup M_{12})\end{cases}\,,$$

where

$$T_{11}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\,,\ T_{12}=\begin{pmatrix}\epsilon&0\\\frac{1-\epsilon}{2}&1\end{pmatrix}\ \text{and}\ T_{13}=\begin{pmatrix}\Big(\frac{1+\epsilon}{2}\Big)^2&\frac{1-\epsilon}{2}\\\epsilon\frac{1-\epsilon}{2}&\epsilon\end{pmatrix}\,.$$

The map F_2 is given by similar formulas, which are easy to derive using the symmetry in the model.

In the example above we saw that the maps $F_i : M \to M$ are piecewise affine. This is always true for the Zhang model. For each $i \in \Lambda$ there exists a finite partition $M = \cup_k M_{ik}$ (consisting of polyhedra $M_{ik} \subset M \subset \mathbb{R}^N$) such that the restrictions $F_i|_{M_{ik}}$ are affine maps on the form $z \mapsto T_{ik}(z+e_i) = T_{ik}z + T_{ik}e_i$. Here T_{ik} are $N \times N$ matrices.

2.2.1 Symbolic coding of the Zhang model

In the BTW model we used symbolic coding $g : \Sigma_N^+ \times \mathcal{R} \to \Sigma_A^+$ to calculate dynamical invariants such as topological and metric entropy. The topological Markov chain $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$ was also used to identify the natural invariant measure. The success of the coding technique for the BTW model suggests that one should attempt a similar construction for the Zhang model.

To do this we need to fix a partition of $\Sigma_N^+ \times M$. One natural choice is to take sets $Z = [i] \times M_{ik}$. We enumerate these sets $Z_1, Z_2, \ldots Z_Q$ and for each initial condition $(\mathbf{t}, z)$ we can construct a sequence $\mathbf{s} \in \Sigma_Q^+$ by the rule

$$F^n(\mathbf{t}, z) \in Z_{s_n} \, .$$

Let $g : \Sigma_N^+ \times M \to \Sigma_Q^+$ denote this map. The dynamics of the Zhang model has a symbolic representation $\Sigma := g(\Sigma_N^+ \times M) \subset \Sigma_Q^+$.

Unfortunately there are several problems with this construction. First, the map g need not be invertible. For a fixed a sequence $\mathbf{t} \in \Sigma_N^+$ there can be two distinct configurations $z, z' \in M$ such that $F_{t_n} \circ \cdots \circ F_{t_1}(z)$ and $F_{t_n} \circ \cdots \circ F_{t_1}(z')$ lie in the same partition element of the partition $M = \cup_k M_{t_{n+1}k}$ for all $n \geq 1$. Then $(\mathbf{t}, z)$ and $(\mathbf{t}, z')$ are mapped to the same sequence $\mathbf{s} \in \Sigma_Q^+$ under the coding map g. The second problem is that the set Σ typically has a more complicated structure than a topological Markov chain. It may not even be a closed set.

Remark 2.6 *We can always construct a topological Markov chain Σ_A^+ by defining a $Q \times Q$ matrix A by the rule $a_{ij} = 1$ if $F(Z_i) \cap Z_j \neq \emptyset$ and $a_{ij} = 0$ otherwise. By construction Σ is a subset of Σ_A^+, but we can not ensure equality. If $a_{ij} = 1$ and $a_{jk} = 1$, then there exists a sequence $\mathbf{s} \in \Sigma_A^+$ containing the segment (ijk), but this segment may not exist in any sequence in the set Σ. In other words, we might have*

$$\Big(F(Z_i) \cap Z_j\Big) \bigcap \Big(F^{-1}(Z_k) \cap Z_j\Big) = \emptyset \, .$$

This is impossible in the BTW model where each partition element only contains a single configuration.

2.2.2 Entropy in the Zhang model

Using that the map $\sigma_N^+ : \Sigma_N^+ \to \Sigma_N^+$ has topological entropy $\log N$, it is very easy to show that the topological entropy of the map $F : \Sigma_N^+ \times M \to \Sigma_N^+ \times M$ is at least $\log N$. We must therefore ask if there exist situations where $h_{\text{top}}(F) > \log N$. In other words: Do the singularities in the Zhang model produce additional entropy? As it turns out, this question is much more difficult to answer than one might imagine, and the mathematics constructed in order to understand the problem has

provided new general results on the on entropy in dynamical systems with singularities (see chapters 4 and 5). Here we give a brief introduction to some aspects of this topic:

It is not difficult to see that the modified map $F : [0,1] \times M \to [0,1] \times M$, with

$$F(x,z) = (Nx \bmod 1, F_{[x/N]}(z)),$$

($[\cdot]$ denotes integer value) has the same topological entropy as the original map F.

Remark 2.7 *The map $\sigma_N^+ : \Sigma_N^+ \to \Sigma_N^+$ can be replaced with an expanding map $x \to Nx \mod 1$ on the interval $[0,1]$. By expressing points $x \in [0,1]$ in N-digit expansions $0.i_1 i_2 i_3 \cdots$ we see that each point x corresponds to a sequence in Σ_N^+, and that the shift map σ_N^+ corresponds to multiplication with N (mod 1). The shift map σ_N^+ is not conjugated to the expanding map since there are real numbers in the interval $[0,1]$ which have more than one representation in the N-digit system. For instance, the numbers $0.0111\cdots$ and $0.100\cdots$ are equal. However, this does not effect the topological entropy.*

Using this modification the Zhang model becomes a piecewise affine map on $[0,1] \times \mathbb{R}^N \subset \mathbb{R}^{N+1}$ with domains of continuity on the form $[(i-1)/N, i/M] \times M_{ik}$. In each of these domains the map F is affine with a linear part given by a matrix with the following block form

$$d_{(x,z)}F = \begin{pmatrix} N & 0 \\ 0 & T_{ik} \end{pmatrix}.$$

Thus the linear parts of the map F^n are on the form

$$d_{(x,z)}F^n = \begin{pmatrix} N^n & 0 \\ 0 & T_{i_n k_n} \dots T_{i_1 k_1} \end{pmatrix},$$

and we see that $\log N$ is a Lyapunov exponent of the map F. From the definition of the Zhang model it is easy to see that the spectral radius of the matrices T_{ik} all are less than or equal to 1 (and any sufficiently long realized composition $T_{i_n k_n} \dots T_{i_1 k_1}$ has spectral radius less than 1). This implies that all the other Lyapunov exponents are negative. If one then applies the Ruelle inequality and the variational principle we get $h_{\text{top}}(F) \leq \log N$ (and hence $h_{\text{top}}(F) = \log N$). However, these general results do not always hold for piecewise affine maps. In fact, there exist piecewise affine contracting maps with positive entropy. An example of such a map is given in chapter 5, and even more examples are given in chapter 4.

There is a particular mechanism that can generate positive entropy in contracting affine maps with singularities. Even though the maps are contracting, and distances between points do not increase under each affine mapping, the angles between lines in phase space may grow with exponential speed. In the presence

of singularities, angular expansion may cause exponential orbit growth and positive entropy. Actually, the entropy can be bounded by so-called angular Lyapunov exponents which measure the asymptotic angular expansion rates. (See chapter 4 for details.) As a corollary, the entropy of a piecewise affine conformal contracting map is zero. Unfortunately, the Zhang model is not defined by piecewise conformal maps, and this approach fails to prove $h_{\text{top}}(F) = \log N$.

In the end we were forced to use a different technique, which only succeeds to prove that the entropy of the Zhang model is $\log N$ for generic ϵ and E_c. This result is presented in chapter 3.

Note that even though the BTW and Zhang model have the same topological entropy, the number of ϵ-distinguishable orbit segments grows faster in the Zhang model. For $E_c \gg 1$ and fixed ϵ we have $N(\epsilon, n) \sim N^n + w(n)$, where $w(n)$ grows at least linearly.

2.2.3 Removability of singularities

For some simple situations the question of entropy and coding is very easy to understand. This is when the singularities (lines of discontinuity) do not affect the asymptotic dynamics.

We define the attractor

$$\mathcal{A} = \pi_s\Big(\bigcap_{n\geq 0} F^n(\Sigma_N^+ \times M)\Big)\,.$$

Here $\pi_s : \Sigma_N^+ \times M \to M$ is the projection on to the configuration space M. Since we are interested in the asymptotic dynamics we can restrict our attention to the set $\Sigma_N^+ \times \mathcal{A}$. It can be shown that it is non-empty for sufficiently large E_c. Note that $\mathcal{A}$ can be written $\mathcal{A}_0 \cap \mathcal{A}_1 \cap \ldots$, where $\mathcal{A}_n$ is defined by

$$\mathcal{A}_n = \pi_s\Big(\bigcap_{k=0}^{n-1} F^n(\Sigma_N^+ \times M)\Big)\,.$$

Clearly these sets make up a nested sequence

$$\cdots \subset \mathcal{A}_n \subset \mathcal{A}_{n-1} \subset \cdots \subset \mathcal{A}_1 \subset \mathcal{A}_0 = M\,.$$

If the dynamics has a simpler description once we restrict to one of the sets $\mathcal{A}_n$, then this restriction can be made with out loosing any information about the asymptotic behavior of the model. This is especially useful if $\mathcal{A}_n$ is disjoined from the singularity set $\mathcal{S} = \cup_{ik}\partial M_{ik}$ for some $n \in \mathbb{N}$. In this case we say that singularities are removable, and it implies that the restriction of F to $\Sigma_N^+ \times \mathcal{A}$ is continuous. We can then use the variational principle combined with the Ruelle-Margulis inequality to show that $h_{\text{top}}(F) = \log N$. Another important consequence of removability of

singularities is that it ensures that F is semi-conjugated to a topological Markov chain.

Remark 2.8 *If the set $\mathcal{A}_n$ is disjoined from the singularity set we can construct coding by choosing a partition of $\Sigma_N^+ \times \mathcal{A}$ with elements on the form $Z = [i] \times K_j$, where $K_1, K_2, \ldots$ are the connected components of $\mathcal{A}_n$. With respect to this partition we have*

$$\Sigma = g(\Sigma_N^+ \times \mathcal{A}_n) = \Sigma_A^+ \,,$$

where $A = ||a_{kl}||$ is given by

$$a_{kl} = \begin{cases} 1 & \text{for } F(Z_k) \cap Z_l \neq \emptyset \\ 0 & \text{otherwise} \end{cases} .$$

The maps F_i are continuous on the connected sets K_j, and the continuous image of a connected set can only intersect one of the (disjoined) partition elements. Using this it is easy to show that the number of 1-s in each row of A is equal to N. Then the maximal eigenvalue of A is N, and $h_{\text{top}}(\sigma_A^+) = \log N$.

In the Zhang model there are examples when singularities are removable and the columns of A have a varying number of 1-s. Hence Dhar's result for Abelian sandpiles does not generalize to the Zhang model, not even when singularities are removable.

2.2.4 Characterization of the attractor

We have already explained several essential differences between the dynamics in the BTW and Zhang models. These become even clearer if one studies the geometry of the attractors in configuration space. For the BTW model attractors are simply equal to the sets of recurrent configurations, which are finite. In the Zhang model the attractors have a complex fractal structure (see figure 2.5). The set $\mathcal{A}$ is actually of full dimension for $E_c \gg 1$, but numerical simulations show that the support of the SRB measure has non-integer dimension.

We can construct the SRB measure by a sequence ν_n of probability measures

$$\nu_n = \frac{1}{n} \sum_{k=0}^{n-1} F_*^k (\mu_{\text{Ber}} \times \mu_{\text{Leb}}) \,.$$

where μ_{Leb} is the Lebesgue probability measure on M and μ_{Ber} is the Bernoulli measure on Σ_N^+. We assume that $\{\nu_n\}$ has a unique limit point ν (in the weak topology) and let ν_s denote its projection to M. Numerically, the measure ν_s can be

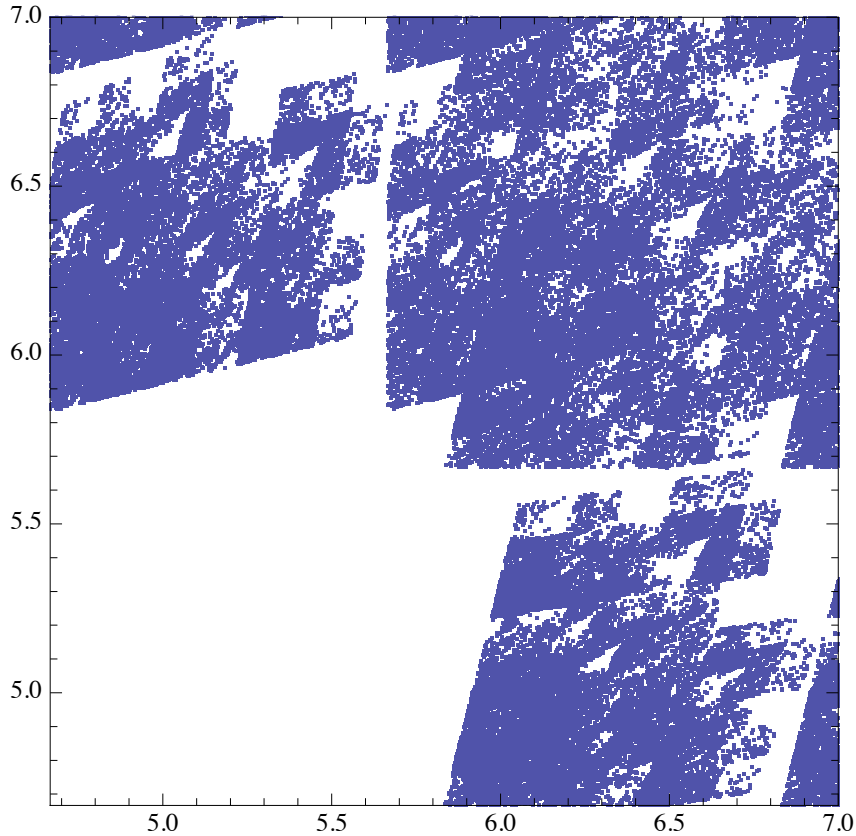


Figure 2.5: An orbit on the attractor in the configuration space M for $d = 1$, $N = 2$, $\epsilon = 2/3$ and $E_c = 7$.

approximated by simulating a long orbit and counting the number of configurations that belong to each cell of a fine partition. This numerical approximation can be used to calculate several characteristics of the attractor.

An especially interesting characteristic is the multi-fractal spectrum. This is defined as the function

$$\alpha \mapsto \dim_H(\mathcal{G}_\alpha)\,,$$

where $\mathcal{G}_\alpha = \{z \in M | d_{\nu_s}(z) = \alpha\}$ and

$$d_{\mu_s}(z) = \lim_{\delta \to 0} \frac{\log \nu_s(B_\delta(z))}{\log \delta}$$

is the point-wise dimension of the measure ν_s in the point z. Using a standard technique based on calculations of moments $\sum_i \nu_s(Y_i)^q$ for partitions $\{Y_i\}$ of varying diameter (see [44] for details) we can calculate this spectrum from our numerical approximation of ν_s. The result is shown in figure 2.6. The non-trivial multi-fractal spectrum illustrates the rich dynamical structure of the Zhang model.

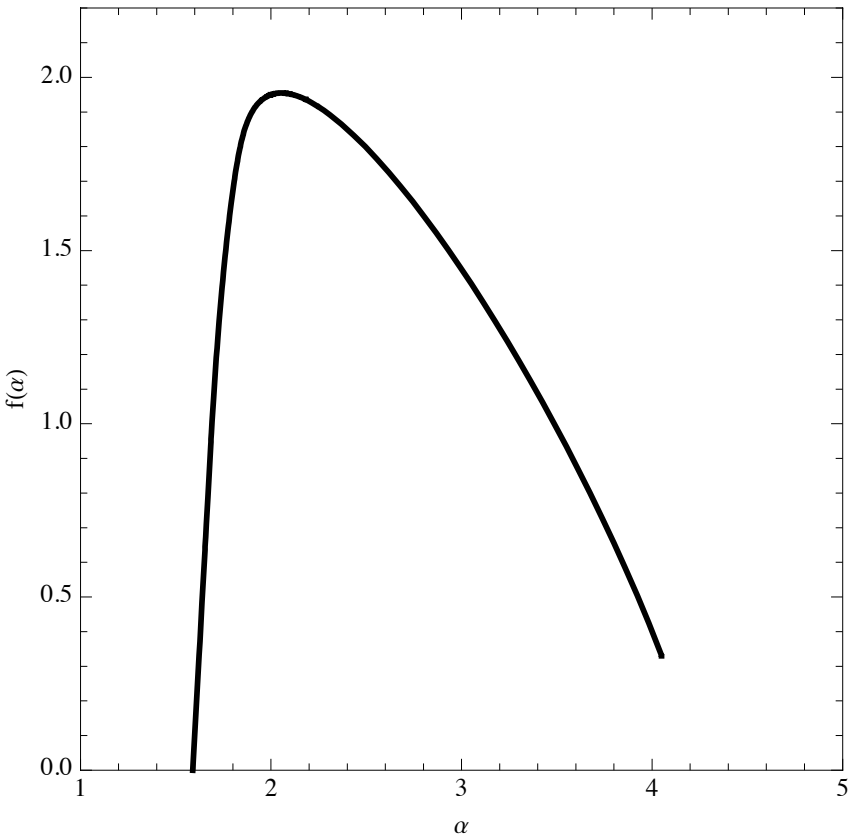


Figure 2.6: The multi-fractal spectrum $f(\alpha) = \dim_H \mathcal{G}_\alpha$ for the projection of the natural invariant measure onto the configuration space M. The parameters in this computation are $d = 1$, $N = 2$, $\epsilon = 2/3$ and $E_c = 7$.

Chapter 3

Dynamics and entropy in the Zhang model of SOC

Joint with B. Kruglikov

Abstract: We give a detailed study of dynamical properties of the Zhang model, including evaluation of topological entropy and estimates for the Lyapunov exponents and the dimension of the attractor. In the thermodynamic limit the entropy goes to zero and the Lyapunov spectrum collapses.

Refrence for this paper: B. Kruglikov, M. Rypdal, *Dynamics and entropy in the Zhang model of SOC*, Journal of Statistical Physics **122**, 975 (2006)

3.1 Introduction

In 1987 the concept of Self-Organized Criticality (SOC) was introduced by Bak, Tang and Wiesenfeld [3]. The attempt was to give an explanation of the omnipresence of fractal structures and power-law statistics in nature, and the claim was that certain physical systems can self-organize into stationary states, reminiscent of equilibrium system at the critical point, in the sense that one has scale invariance and long range correlations in space and time.

SOC is proposed as an explanation for variety of phenomena in nature, such as earthquakes, forest fires, stock markets and biological evolution [30]. However, most work has been devoted to the study of idealized "sandpile-like" computer models, such as the sandpile model [3], the abelian sandpile [23] and the Zhang model [57], that one believes exhibit SOC in the thermodynamic limit. Despite this effort, a satisfactory understanding of the model is not yet achieved. Through numerical investigation one has observed that in the thermodynamic limit, observables have power-law distributions. More precisely, the probability distribution function $P(s)$, of an observable s will have the form $P(s) \sim 1/s^{\tau_s}$ in the thermodynamic limit. There is not a widely agreed upon method for computing the SOC-exponents τ_s numerically, and due to the incomplete understanding of the dynamics of the models and lack of a formal treatment of the thermodynamic limit, it is difficult to properly explain the observed behavior. Hence it is not clear what the SOC-exponents really tell us about the dynamics of the SOC models.

In a series of papers by Cessac, Blanchard and Krüger [6, 7, 8] it was proposed that further understanding of SOC models can be achieved by studying the models in the framework of dynamical system theory. They showed how a particular model, the Zhang model, could be formulated as a dynamical system of skew-product type with singularities, where the randomness of the external driving is described by a Bernoulli shift, and the threshold relaxation dynamics is given by piecewise affine maps.

In this paper we give a detailed study of the dynamical system defined in [6]. We prove several basic properties, some of which are already stated in [6, 7, 8], before discussing fundamental dynamical properties. We observe that, depending on the parameters of the model, we can observe fundamentally different types of behavior. For low values of the threshold energy (critical energy), the dynamics can be relatively simple, since the singularities only effect the dynamics in a finite number of time-steps. In such situations we say that singularities are removable, and we show that the system permits symbolic coding whenever singularities are removable. We give examples of how symbolic coding can give a complete topological description of the dynamics as a topological Markov chain. Hence the dynamics is chaotic, but the essential dynamical invariants are all inherited from the Bernoulli shift factor. Moreover we can identify the physical invariant measure, and hence understanding

of the statistical properties is reduced to the theory of Markov chains.

As we increase the critical energy the role of the singularities becomes essential. Techniques, which are based on codings, are no longer applicable and a very interesting dynamics is observed. The dimensional characteristics of the attractor are also sensitive to the parameters of the system, as we show by generalizing the Moran's formula for IFS's. In addition, we observe the situation, when the dimension of the IFS-attractor increases to the maximum, while the support of the SRB-measure remains fractal.

To measure the complexity of the dynamics we study entropy and Lyapunov exponents. We show that the system is hyperbolic, with one positive exponent originating in the Bernoulli shift. However, due to the presence of singularities the Ruelle inequality and the Pesin formula are not directly applicable. We show that the metric entropy of any SRB-measure equals the topological entropy almost surely, and we evaluate the latter generalizing the technique developed by Buzzi [16, 17]. The result is that the Pesin formula and the variational principle hold a posteriori.

To give a satisfactory physical interpretation of the dynamics we rescale time to prevent infinitely slow driving of the system. We prove that for this physical system, the Lyapunov spectrum collapses completely and that the entropy goes to zero in the thermodynamic limit. This implies that the expanding (chaotic) properties are lost, so that we may expect power-laws statistics and long range correlation effects.

The statistical properties we obtain hold for any SRB-measure. The existence of SRB-measures is in fact still an open problem. From the general theory of dynamical systems with singularities [33, 43, 53], we can give conditions that are sufficient for the existence of SRB-measures, but it is not known if these conditions hold for the majority of parameters. We expect this to be true (it was also conjectured in [6]) and derive statistical corollaries (some of which hold without the assumption of measure existence).

Apart from the physical importance of the Zhang model, it is interesting from a mathematical point of view. It can be described as a piecewise affine hyperbolic map of the form

$$F : \Sigma_N^+ \times M \to \Sigma_N^+ \times M, \quad ((t_0 t_1 t_2 \dots), x) \mapsto ((t_1 t_2 t_3 \dots), f_{t_0}(x)),$$

where Σ_N^+ is the set of right infinite sequences from finite alphabet and $\{f_i\}$ a collection of piece-wise affine non-expanding maps of M to itself. Previously piecewise affine expanding maps and piecewise isometries have been studied, but the contracting property of the relaxation dynamics provides difficulties. Therefore several methods are developed in this paper, which hold far beyond the framework of the Zhang model.

The structure of the paper is the following. In section 1 we describe the model, find bounds on avalanches size and time, which makes possible to use the Poincaré

return to reformulate the systems in a skew-product form. Then we describe, when degenerations occur (the original Zhang setting $\epsilon = 0$ is not the only possibility) and study the contraction property to conclude hyperbolicity of the model. In section 2 we introduce the concept of removability of singularities, which appears in the coding approach for the study of the model.

Section 3 is devoted to the study of measure entropy and Lyapunov spectrum and section 4 concerns the topological entropy. Since the standard theorems do not work in the presence of singularities, we develop a new technique for evaluation of these basic quantities. Section 5 briefly describes the dimension issues of the model.

In appendices A and B we provide bounds for the entropy and the dimension, which in the presence of singularities, overlaps and degenerations (this was designed for an application to the Zhang model) are new. The results are of interest in its own and can be read independently.

3.2 Basic properties of the Zhang model

In the Zhang model each site on the lattice is associated with a non-negative real number, which we call the energy of the site. The collection of energies is called an energy configuration, and can be represented as a point in N-dimensional space, where N is the number of sites in the lattice. If a configuration is unstable, the overcritical sites will lose some of their energy to their nearest neighbors, resulting in a new energy configuration. This transformation on $\mathbb{R}^N$ is denoted by f. If a configuration is stable, a site is chosen at random and an energy quantum $\delta = 1$ is added to this site. In [6] it was shown how the relaxation and random excitation can be formulated as a map of skew-product type on an extended phase-space. This extended phase-space has the configuration space as one factor, and the set of all possible sequences of excitations as the other factor. In [6] it was also shown how one can reformulate the dynamical system by considering the return maps to the set of stable configurations. This gives a simplification, in the sense that each avalanche is associated with an affine transformation. The set of stable configurations is partitioned into domains, where each domain corresponds to an avalanche.

3.2.1 Relaxation

Take $d, L \in \mathbb{N}$ and let $\Lambda \subset \mathbb{Z}^d$ be the cube $[1, L]^d$ of cardinality $N := L^d = |\Lambda|$. Let $\phi : \Lambda \to \Lambda' := \{1, \ldots, N\}$ be a bijection. We define a metric d_Λ on Λ by

$$d_\Lambda(\mathbf{k}, \mathbf{l}) = \sum_{1 \le n \le d} |k_n - l_n| \,,$$

and let $d_{\Lambda'} := \phi_* d_\Lambda$. In the following we omit primes when it is clear from the context that we are considering the metric space $(\Lambda', d_{\Lambda'})$. Elements of Λ will be

called sites. We say that sites i and j are nearest neighbors if $d_\Lambda(i,j) = 1$. The boundary $\partial\Lambda$ is defined as those sites $i \in \Lambda$ that have less than $2d$ nearest neighbors.

Fix parameters $E_c > 0$ and $\epsilon \in [0,1)$ and define $f : \mathbb{R}^N_{\geq 0} \to \mathbb{R}^N_{\geq 0}$ by

$$f(x)_i = x_i - \theta(x_i - E_c)(1-\epsilon)x_i + \frac{1-\epsilon}{2d} \sum_{d_\Lambda(i,j)=1} \theta(x_j - E_c)x_j \,,$$

where

$$\theta(a) = \begin{cases} 1 \text{ if } a > 0 \\ 0 \text{ if } a \leq 0 \end{cases} \quad .$$

Let $\|x\|_1 = \sum_{i=1}^N |x_i|$ be the 1-norm on $\mathbb{R}^N$.

Proposition 3.1 *For all $x \in \mathbb{R}^N_{\geq 0}$ we have*

$$\frac{1+\epsilon}{2}\|x\|_1 \leq \|f(x)\|_1 \leq \|x\|_1 \,,$$

and $\|f(x)\|_1 = \|x\|_1$ if and only if $x_i \leq E_c$ for all $i \in \partial\Lambda$. If there is $i \in \partial\Lambda$ such that $x_i > E_c$ then

$$\|f(x)\|_1 \leq \|x\|_1 - \frac{1-\epsilon}{2d}E_c \,.$$

Proof. Let $\{x_{i_k}\}_{k=1}^m$ be the entries of the vector x that are greater than E_c. Let n_{i_k} be the number of nearest neighbors of x_{i_k}. Then

$$\begin{aligned} \|f(x)\|_1 &= \sum_{i\in\Lambda} f(x)_i = \sum_{i\in\Lambda} x_i - (1-\epsilon)\sum_{k=1}^m x_{i_k} + \frac{1-\epsilon}{2d}\sum_{k=1}^m n_{i_k}x_{i_k} \\ &= \sum_{i\in\Lambda} x_i - (1-\epsilon)\sum_{k=1}^m (1 - \frac{n_{i_k}}{2d})x_{i_k} \,. \end{aligned}$$

The statement follows from the fact that we always have $d \leq n_{i_k} \leq 2d$, and $n_{i_k} = 2d$ if and only if $x_{i_k} \notin \partial\Lambda$. □

We say that a site $i \in \Lambda$ of the configuration x is relaxed if $x_i \leq E_c$, and excited if $x_i > E_c$. A configuration x is called stable if all sites are relaxed. The set of stable configurations is $M := [0, E_c]^N$. For each configuration x we define $m(x) = \min\{n \geq 0 \,|\, f^n(x) \in M\}$.

Proposition 3.2 *For all $x \in \mathbb{R}^N_{\geq 0}$ we have:*

$$m(x) \leq \frac{2dN}{1-\epsilon}\frac{\|x\|_1}{E_c}\Big(\frac{2d}{1-\epsilon} + 1\Big)^{\mathrm{diam}(\Lambda)/2} \,.$$

We need the following lemma ($[\cdot]$ denotes integer part):

Lemma 3.3 *For $x \in \mathbb{R}^N_{\geq 0}$ and $n \in \mathbb{N}$ let $\alpha_i(n,x)$ be the cardinality of the set $\{l \leq n \,|\, (f^l x)_i > E_c\}$. Let $\gamma = [2d/(1-\epsilon)] + 1$. If $d_\Lambda(i,j) = 1$, then $\alpha_j(n,x) \geq [\alpha_i(n,x)/\gamma]$.*

Proof. There is a finite increasing sequence $\{m_k\}$, such that $(f^{m_k}(x))_i > E_c$. We claim that on each interval $(m_k, m_{k+\gamma}]$ there is a number m such that $(f^m(x))_j > E_c$. In fact, in the opposite case

$$\left(f^{1+m_{k+\gamma-1}}(x)\right)_j \geq \gamma \frac{1-\epsilon}{2d} E_c > E_c \,.$$

Since $[0, \alpha_i(n,x)]$ contains $\beta = [\alpha_i(n,x)/\gamma]$ disjoined such intervals, we get $\alpha_j(n,x) \geq \beta$. Thus $\alpha_j(n,x) \geq [\alpha_i(n,x)/\gamma]$. □

Proof of Proposition 3.2. By applying inductively Lemma 3.3 we get:

$$\alpha_j(n,x) \geq \left[\frac{\alpha_i(n,x)}{\gamma^{d_\Lambda(i,j)}}\right]$$

In fact, if $j = j_0, j_1, \ldots j_k = i$ is a path with $d_\Lambda(j_s, j_{s+1}) = 1$ and

$$\left[\frac{\alpha_i(n,x)}{\gamma^k}\right] = t \,,$$

then $\alpha_{j_k}(n,x) \geq t\gamma^k, \alpha_{j_{k-1}}(n,x) \geq t\gamma^{k-1}, \ldots, \alpha_{j_0}(n,x) \geq t$. By Proposition 3.1

$$\alpha_j(n,x) \leq \frac{2d}{1-\epsilon} \frac{\|x\|_1}{E_c}$$

for $j \in \partial\Lambda$, so

$$\alpha_i(n,x) \leq \alpha(x) := \frac{2d}{1-\epsilon} \frac{\|x\|_1}{E_c} \gamma^{\mathrm{diam}(\Lambda)/2} \,.$$

for all $i \in \Lambda$ and all $n \in \mathbb{N}$. Suppose $f^m(x) \notin M$ for all $m \leq T$. If $T > N\alpha(x)$, then there must be a site $i \in \Lambda$ that is greater than E_c for more than $\alpha(x)$ different times. This is impossible so $m(x) \leq N\alpha(x)$. □

3.2.2 Random excitations

Define $\Sigma_N^+ = \Lambda^{\mathbb{N}}$ to be the set of right-infinite Λ-sequences and let $\sigma_N^+ : \Sigma_N^+ \to \Sigma_N^+$ be the left shift. We define a map $\hat{f} : \Sigma_N^+ \times \mathbb{R}^N_{\geq 0} \to \Sigma_N^+ \times \mathbb{R}^N_{\geq 0}$ by

$$\hat{f}(\mathbf{t}, x) = \begin{cases} (\sigma_N^+ \mathbf{t}, x + e_{t_0}) & \text{if } x \in M \\ (\mathbf{t}, f(x)) & \text{if } x \notin M \end{cases} ,$$

where $e_1, \ldots, e_N$ is the standard basis in $\mathbb{R}^N$. We denote points in $\Sigma_N^+ \times \mathbb{R}_{\geq 0}^N$ by $\hat{x} = (\mathbf{t}, x)$, and we define π_u and π_s to be the projections to Σ_N^+ and $\mathbb{R}_{\geq 0}^N$ respectively.

Proposition 3.4 *For all $\hat{x} \in \Sigma_N^+ \times \mathbb{R}_{\geq 0}^N$ it holds:*

$$\min\{m \geq 0 \,|\, \forall i \in \Lambda \, \exists m' \leq m : \, (\pi_s \circ \hat{f}^{m'}(\hat{x}))_i > E_c\} \leq n(E_c, \epsilon, \Lambda)\,,$$

where

$$n(E_c, \epsilon, \Lambda) = N(NE_c + 2)\Big(\Big[\frac{2d}{1-\epsilon}\Big] + 1\Big)^{\mathrm{diam}(\Lambda)}\,,$$

Proof. In $N[E_c] + 1$ time-steps, there must be an overcritical site. Since in the relaxation process there is always an overcritical site, then during arbitrary subsequent $N[E_c] + 2$ time-steps an exited site can be found. Hence after $N\xi(N[E_c] + 2)$ time-steps either all sites have been overcritical or there is a site that has been overcritical at least ξ times. However it follows from the proof of Proposition 3.2 that if one site is overcritical

$$\xi = \Big(\Big[\frac{2d}{1-\epsilon}\Big] + 1\Big)^{\mathrm{diam}(\Lambda)}$$

times, then all sites have been overcritical at least once. □

For $x \in \mathbb{R}_{\geq 0}^N$ and $i \in \Lambda$ we define $\tau(i, x) := \min\{n \in \mathbb{N} \,|\, f^n(x + e_i) \in M\}$. Proposition 3.2 assures us that this number is finite and

$$\max_{x \in M} \max_{i \in \Lambda} \tau(i, x) \leq \tau_m(E_c, \epsilon, \Lambda) = N^2\Big(1 + \frac{1}{NE_c}\Big)\Big(\frac{2d}{1-\epsilon} + 1\Big)^{\mathrm{diam}(\Lambda)/2+1}.$$

Thus we observe that neither $n(E_c, \epsilon, \Lambda)$ nor $\tau_m(E_c, \epsilon, \Lambda)$ are uniformly bounded in E_c, but there is the following alternative:

There exists a constant C_0, not depending on the energy E_c, such that either $n(E_c, \epsilon, \Lambda) \leq C_0$ or $\tau_m(E_c, \epsilon, \Lambda) \leq C_0$.

In fact, we can set $C_0 = 3N^2\big(\frac{2d}{1-\epsilon} + 1\big)^{\mathrm{diam}(\Lambda)+1}$. Thus we get that either relaxation happen sufficiently fast or all the sites keep being excited sufficiently often (uniformly in E_c).

But there does not exist such a bound uniform in ϵ or N.

3.2.3 Return maps

Let $\hat{x} = (\mathbf{t}, x) \in \Sigma_N^+ \times M$. For $n = 1, \ldots, \tau(t_0, x)$ define $C_n(\hat{x}) = \{i \in \Lambda \,|\, (\pi^s \circ \hat{f}^n \hat{x})_i > E_c\}$, and $A(\hat{x}) = \big(C_1(\hat{x}), \ldots, C_{\tau(t_0,x)}(\hat{x})\big)$. We call $A(\hat{x})$ the *avalanche* of the point $\hat{x}$. Let $\hat{M} := \Sigma_N^+ \times M$ and define an equivalence relation $\sim$ on $\hat{M}$ by

$$\hat{x} \sim \hat{y} \Leftrightarrow A(\hat{x}) = A(\hat{y})\,.$$

This gives a partition of $\hat{M}$. From the definition it is clear that $A(\hat{x})$ depends on t_0 and x only. Hence partition elements are of the form $[i] \times M_{ij}$, where

$$\forall i \in \Lambda : \bigcup_j M_{ij} = M$$

and $[i] = \{\mathbf{t} \in \Sigma_N^+ \,|\, t_0 = i\}$ is the cylinder of the symbol i. We see that for each $i \in \Lambda$, the domains $M_{i1}, M_{i2}, \ldots$ are separated by segments of at most $N!^{\tau_m} = \exp(\tau_m(E_c, \epsilon, \Lambda) \log N!)$ hyperplanes. Hence we have a finite number of domains $M_{i1}, \ldots, M_{iq_i}$ for each $i \in \Lambda$. By definition there is a unique avalanche for each partition element $[i] \times M_{ij}$. We denote this avalanche by A_{ij}. Its duration is $\tau_{ij} := \tau(i, x)$, for $x \in M_{ij}$, and define its size to be $s_{ij} = \sum_{n=1}^{\tau_{ij}} |C_n|$.

We define the piecewise continuous map $F : \hat{M} \to \hat{M}$ by

$$(\mathbf{t}, x) \mapsto (\sigma_N^+ \mathbf{t}, F_{t_0} x)$$

where $F_i(x) := f^{\tau(i,x)}(x + e_i)$. We define $F_{ij} := F_i|_{M_{ij}}$.

Remark 3.1 *From a mathematical point of view the formulation $(\hat{M}, F)$ is a simplification compared to $(\Sigma_N^+ \times \mathbb{R}_{\geq 0}^N, \hat{f})$. However, the duration of avalanches are suppressed so that all avalanches have the same duration. This is not satisfactory from a physical point of view, and hence we call $(\hat{M}, F)$ the mathematical model and $(\Sigma_N^+ \times \mathbb{R}_{\geq 0}^N, \hat{f})$ the physical model. We will later make a rescaling of time in the physical model, so that the driving does not become infinitely slow in the thermodynamic limit.*

For each $x \in \mathbb{R}_{\geq 0}^N$ we define a matrix $Q(x)$ by

$$Q_{kl}(x) = \begin{cases} \frac{1}{2d}\theta(x_l - E_c) & \text{if } d_\Lambda(k, l) = 1\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

and a diagonal matrix $J(x)$ by $J_{kl}(x) = (1 - (1 - \epsilon)\theta(x_l - E_c))\delta_{kl}$. Set

$$S(x) = J(x) + (1 - \epsilon)Q(x) \tag{3.1}$$

and observe that $f(x) = S(x)x$. Let $x(1) = x + e_i$ and $x(n) = f(x(n-1))$ for $n \in \{2, \ldots, \tau(t, x)\}$. Then $F_i(x) = L_i(x + e_i)$, where

$$L_i(x) = S(x(\tau(x, i))) \ldots S(x(1))\,.$$

If $x, y \in M_{ij}$, then $\tau(i, x) = \tau(i, y)$ and the same components of $x(n)$ and $y(n)$ are grater than E_c for each $n = 1, \ldots, \tau(t, x)$, so $L_i(x) = L_i(y)$. We define the linear map $L_{ij} := L_i(x)$ for $x \in M_{ij}$. We get $F_i|_{M_{ij}}(x) = L_{ij}(x + e_i)$.

Definition 3.1 *A sequence* $\{(i_n, j_n) \,|\, 1 \le n \le \theta\}$ *is said to be admissible if*

$$\bigcap_{n=1}^{\theta} (F_{i_{n-1}j_{n-1}} \circ \cdots \circ F_{i_1 j_1})^{-1}(M_{i_n j_n}) \neq \emptyset\,,$$

Theorem 3.5 *For all* i, j $\|L_{ij}\|_1 \le 1$. *Moreover for every constant* $c \in (0,1)$ *there is a number* $T \in \mathbb{N}$ *such that for every* $\theta > T$ *and admissible sequence* $\{(i_n, j_n) \,|\, 1 \le n \le \theta\}$ *it holds:*

$$\|L_{i_\theta j_\theta} \dots L_{i_1 j_1}\|_1 < c\,.$$

Proof. If A is an $N \times N$ matrix we let $C_k(A)$ be its k-th column. Observe that for any matrices A and B we have the following formula:

$$\|C_k(AB)\|_1 = \sum_l \|C_l(A)\|_1 B_{lk}\,. \tag{3.2}$$

By the construction: $\|C_k(S(x))\|_1 \le 1$ for all $k \in \Lambda$. Hence $\|L_{ij}\|_1 \le 1$.

To prove the second statement we note that for $\epsilon > 0$ the diagonal elements of the matrices $S(x)$ are non-zero and $\ge \epsilon$. Therefore

$$(S(x(m))S(x(m-1)))_{kl} \ge \epsilon \cdot \max\{S_{kl}(x(m)), S_{kl}(x(m-1))\}.$$

Moreover, $S_{kl}(x(m)) > 0$ if $x(m)_l > 0$ and $d_\Lambda(k,l) = 1$. It follows that any admissible product $L_{i_\theta j_\theta} \dots L_{i_1 j_1}$ of length $\theta \ge n(E_c, \epsilon, \Lambda)$ is positive. By Proposition 3.1 there must be at least one column such that the sum over this column is less than 1, for some factor $L_{i_t j_t}$ and hence for the whole product. Therefore the sum over each column of any admissible product of length $2n(E_c, \epsilon, \Lambda)$ must be less than 1. Let $c_0 < 1$ be the maximal norm of all admissible products of length $2n(E_c, \epsilon, \Lambda)$. For $k > k_0 := [\log c / \log c_0] + 1$ we have $c_0^k < c$ and hence $T = 2k_0 n(E_c, \epsilon, \Lambda)$ is the required number.

The above argument does not apply to the case $\epsilon = 0$, and a different proof must be given for this case (which actually works in general as well). Take $\hat{x} \in \hat{M}$ and let $x(t) \in M$ be the projection of its orbit to $\hat{M}$. Denote $S(x(t))$ by $S_t(\hat{x})$, and let

$$\tilde{S}_t(\hat{x}) = S_t(\hat{x}) \cdots S_0(\hat{x})\,.$$

We make the following claims:

1. There exists $\bar{n} \in \mathbb{N}$ such that for all $l, m \in \Lambda$ and all $\hat{x} \in \hat{M}$ there is $t \le \bar{n}$ such that $(\tilde{S}_t(\hat{x}))_{lm} \neq 0$.

2. For all $i \ge 0$ there exists $n_i \in \mathbb{N}$ such that $\|C_m(\tilde{S}_t(\hat{x}))\|_1 < 1$ for all $t \ge n_i$, $\hat{x} \in \hat{M}$ and all sites $m \in \Lambda$ with $d_\Lambda(m, \partial\Lambda) \le i$.

The second claim for $i = \frac{1}{2}\,\mathrm{diam}(\Lambda)$ implies the statement of the theorem.

To see the first claim we fix $\hat{x}$ and let $U \subset \Lambda^2$ be the subset of the pairs (l, m) with $(\tilde{S}_t)_{lm} = 0$ for all sufficiently large t. By inductively applying (3.2) we see that the columns for $\tilde{S}_t(\hat{x})$ are non-zero for all $t \geq 0$. So for all $\beta \in \Lambda$ there is $\alpha \in \Lambda$ such that $(\alpha, \beta) \in \Lambda^2 \setminus U$. Given sites α and β we choose t such that $\tilde{S}_t(\hat{x})_{\alpha\beta} \neq 0$. Consider now column α of the matrix $\tilde{S}_{t+1}(\hat{x}) = S_{t+1}(\hat{x})\tilde{S}_t(\hat{x})$. If α is stable, i.e. $x(t+1)_\alpha \leq E_c$, then $(S_{t+1}(\hat{x}))_{\alpha\alpha} > 0$ and $(\tilde{S}_{t+1}(\hat{x}))_{\alpha\beta} \neq 0$, so we just repeat the argument. But the site α can not be stable for more than $n(E_c, 0, \Lambda)$ iterations. Hence we can with no loss of generality choose t such that $x(t+1)_\alpha > E_c$. Then the column α of $S_{t+1}(\hat{x})$ has non-zero elements in all position that correspond to neighbors of α. Hence we obtain that $(\alpha', \beta) \in \Lambda^2 \setminus U$ for all α' with $d_\Lambda(\alpha', \alpha) = 1$. Any two points can be connected by a path of neighbors, so $U = \emptyset$, and the first claim follows. In fact, one can see that the bound $\bar{n}$ does not depend on a choice of $\hat{x}$ and satisfies: $\bar{n} \leq \mathrm{diam}(\Lambda) \cdot n(E_c, 0, \Lambda)$.

To prove the second claim let us note that if $\|C_k(\tilde{S}_t(\hat{x}))\|_1 < 1$, then $\|C_k(\tilde{S}_{t+1}(\hat{x}))\|_1 < 1$ because by (3.2): $\|C_k(AB)\| \leq \max_l \|C_l(A)\| \cdot \| C_k(B)\|$.

We will use induction on i starting from $i = 0$. Take $k \in \partial\Lambda$ and $t \leq \bar{n}$ such that $(\tilde{S}_t(\hat{x}))_{kk} \neq 0$. If $x(t+1)_k > E_c$, then $\|C_k(S_{t+1}(\hat{x}))\|_1 < 1$ and

$$\|C_k(\tilde{S}_{t+1}(\hat{x}))\|_1 = \sum_l \|C_l(S_{t+1}(\hat{x}))\|_1 (\tilde{S}_t(\hat{x}))_{lk} < 1\,,$$

and so we have the desired inequality. If $x(t+1)_k \leq E_c$, then $(S_{t+1}(\hat{x}))_{kk} = 1$ and hence $(\tilde{S}_{t+1}(\hat{x}))_{kk} \neq 0$. Then we repeat the argument. Since no site can be stable for more than $n(E_c, 0, \Lambda)$ successive time-steps we obtain the claim for $i = 0$ with $n_0 = n(E_c, 0, \Lambda) + \bar{n}$.

Consider now the case $i > 0$. For a site $m \in \Lambda$ with $d(m, \partial\Lambda) = i$, we take $l \in \Lambda$ with $d_\Lambda(l, m) = 1$ and $d_\Lambda(l, \partial\Lambda) = i - 1$. By the first claim we find some $t \leq \bar{n}$ such that $(\tilde{S}_t(\hat{x}))_{lm} \neq \emptyset$, and by the induction hypothesis for $t' \geq n_{i-1}$ we have:

$$\|C_l\big(S_{t+t'}(\hat{x}) \cdots S_{t+1}(\hat{x})\big)\|_1 < 1\,.$$

Using (3.2) we obtain $\|C_m\big(\tilde{S}_{t+t'}(\hat{x})\big)\|_1 < 1$. We can choose $n_i = n_{i-1} + \bar{n}$. $\square$

Lemma 3.6 *Let $E_c \geq \epsilon/(1-\epsilon)$. Then for any $\hat{x} \in \hat{M}$, $n \in \mathbb{N}$ and $i, j \in C_n(\hat{x})$ we have $d_\Lambda(i, j) \neq 1$.*

Proof. Take $\hat{x} \in M$ and let E_n be the maximal energy of a site in $C_n(\hat{x})$. Clearly $E_1 \leq E_c + 1$ and

$$E_{n+1} \leq \max\big\{\max\{\epsilon E_n, E_c\} + (1-\epsilon)E_n, E_c + 1\big\}\,.$$

From this we see by induction that

$$E_n \leq \max\big\{\frac{E_c}{\epsilon}, E_c + 1\big\} = \frac{E_c}{\epsilon}\,,$$

so $\epsilon E_n \leq E_c$ for all $n \in \mathbb{N}$, and this means that a site cannot be overcritical in two successive time-steps (for $\epsilon = 0$ the above argument does not work, but the statement holds obviously).

All avalanches start with a single site. Let $C_1(\hat{x}) = \{i\}$. Then $d(i,j) = 1$ for all $j \in C_2(\hat{x})$. This implies that any two elements of $C_2(\hat{x})$ can be connected with a path of length 2, so no two sites of $C_2(\hat{x})$ are nearest neighbors. If there exists a path of even length between two points in Λ, then all paths connecting these points are of even length. Therefore we can repeat the argument proving by induction that $d_\Lambda(i,j) \in 2\mathbb{Z}$ for all $i,j \in C_n(\hat{x})$. $\square$

Proposition 3.7 *The linear maps L_{ij} are all invertible whenever $\epsilon \geq 1/2$ or $\epsilon > 0$ and $E_c \geq \epsilon/(1-\epsilon)$. If we have $E_c \geq \epsilon/(1-\epsilon)$, then*

$$\det L_{ij} = \epsilon^{s_{ij}} \ .$$

Proof. Take arbitrary $x \in \mathbb{R}^N_{\geq 0}$. First we observe that since the sum over each column of $Q(x)$ is less than or equal to 1, we have $\|Q(x)v\|_1 \leq \|v\|_1$ for each $v \in \mathbb{R}^N$. This implies that

$$\begin{aligned} \|S(x)v\|_1 &= \|(J(x) + (1-\epsilon)Q(x))v\|_1 \\ &\geq \|J(x)v\|_1 - (1-\epsilon)\|Q(x)v\|_1 \\ &\geq (2\epsilon - 1)\|v\|_1 \ . \end{aligned}$$

If $\epsilon > 1/2$, then $S(x)v \neq 0$ for all $v \neq 0$, so we have invertibility.

For $\epsilon = 1/2$ the claim follows since in the above chain of inequalities at least one is strict if $v \neq 0$. In fact, if $J(x)v = \epsilon v$, then $v_i = 0$ for all relaxed sites i. We claim that the equality $Q(x)v = v$ is impossible. To see this denote by $\tilde{Q}$ the minor-matrix formed by the rows and columns of $Q(x)$, corresponding to exited sites, and denote by $\tilde{v}$ be the respective reduced vector. Then $\tilde{Q}\tilde{v} = \tilde{v}$.

Let U be the set of overcritical sites k with $v_k = \max v_l$ (we suppose it is positive, multiplying by -1 in the opposite case). Choose a boundary site $k \in U$, i.e. the number of neighbors l to k with $v_l = v_k$ is less than $2d$. Then:

$$v_k = \sum_l \tilde{Q}_{kl} v_l < v_k \sum_l \tilde{Q}_{kl} \leq v_k.$$

This contradiction yields the result.

Finally consider the last statement about the case $E_c \geq \epsilon/(1-\epsilon)$. It is proved by reducing the matrix $S(x)$. If $x_i \leq E_c$, then column $C_i(S(x))$ equals $(0,\ldots,0,1,0,\ldots 0)^T$, where the 1 is in the ith position. We can start the decomposition of $\det S(x)$ with column i, and hence we see that row i and column i can be removed from $S(x)$ without changing the determinant. We remove all rows and columns that correspond to relaxed sites. If $\rho(x)$ is the number of overcritical sites of x, we get a $\rho(x) \times \rho(x)$

matrix $S_{red}(x)$. If site k is overcritical then $J_{kk}(x) = \epsilon$. If $E_c \geq \epsilon/(1-\epsilon)$, then it follows from Lemma 3.6 that all nearest neighbors of k are relaxed. Hence column k of $Q(x)$ has only zero entries. This shows that $S_{red}(x) = \mathrm{diag}(\epsilon, \ldots, \epsilon)$. Then

$$\det S(x) = \det S_{red}(x) = \epsilon^{\rho(x)} ,$$

so $\det L_{ij} = \epsilon^{s_{ij}}$. □

Remark 3.2 *In the original model of Zhang one has $\epsilon = 0$ in which case $\det L_{ij} = 0$ if $L_{ij} \neq \mathbf{1}$. But it is not true that non-trivial kernels can occur for $\epsilon = 0$ only, contrary to what was stated in [6]. A simple counter-example is the case $N = 2$, $E_c = 1/3$ and $\epsilon = 1/3$. For $x_1 > 0$ and $2x_1 + 3x_2 < 1$ we have:*

$$F_1\Big(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\Big) = \frac{1}{9}\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} .$$

and so $\det L_{12} = 0$.

Having non-degenerate maps in the model is more convenient from the point of view of mathematical tools (though from a physical viewpoint it can make no big difference between degenerate and close-to-degenerate systems). Fortunately, degenerations occur only for a negligible set of parameters.

Theorem 3.8 *The maps L_{ij} are invertible for almost all (ϵ, E_c). In fact, they are invertible for the parameters complimentary to the set $\Xi \subset [0,1) \times (0,\infty)$, which consists of a finite set of vertical intervals for fixed d and N.*

Proof. Fix an avalanche A_{ij} and let L^{ϵ}_{ij} be the corresponding linear maps (we stress dependence on ϵ). These maps are the compositions of elementary matrices $S^{\epsilon}(x(\tau(x,i)))\ldots S^{\epsilon}(x(1))$, with the factors from (3.1)

$$S^{\epsilon}(x(t)) = \mathbf{1} + (\epsilon - 1)(\tfrac{dJ}{d\epsilon} - Q)(x(t))$$

being polynomial in ϵ and independent of the choice of $x = x(1) \in M_{ij}$. The condition $\det L^{\epsilon}_{ij} = 0$ is equivalent to $\det S^{\epsilon}(x(t)) = 0$ for some t. There are only finite number of possibilities for the matrix $S^{\epsilon}(x)$ (though a countable number for their compositions L^{ϵ}_{ij}, the length of which grow as $E_c \to 0$). Thus we obtain $k = k(d,N)$ polynomial equations $P_j(\epsilon) = 0$. Since $S^{\epsilon}(x) \to \mathbf{1}$ for all x as $\epsilon \to 1$ we see that $\lim_{\epsilon\to 1} P_j(\epsilon) = \det \mathbf{1} = 1$. This shows that the polynomial is not identically zero, so there can only be a finite number of roots. For each root ϵ_{ja} there is the maximal value E_c^{ia} of E_c (finite if $\epsilon_{ja} \neq 0$), where the corresponding matrix $S^{\epsilon}(x)$ can appear in the avalanche. Thus the set of degenerate systems is $\{(\epsilon, E_c) \,|\, \epsilon = \epsilon_{ja}, 0 < E_c < E_c^{ja}\}$. □

By proposition 3.7 Ξ does not intersect the set $\{\epsilon \geq 1/2\} \cup \{ E_c \geq \epsilon/(1-\epsilon)\}$.

3.3 Removability of singularities and coding

The map F may be considered as a piecewise affine map $F : I \times M \to I \times M$, where $I = [0,1]$ and $F(t,x) = (Nt \mod 1, F_{[Nt]}(x))$. The map $t \mapsto Nt \mod 1$ is not conjugated to σ_N^+ since the points $m/N^k \in I$ do not have unique representations in Σ_N^+. However the sets $\{m/N^k\} \times M \subset I \times M$ are singularities, and following the standard approach for piecewise affine maps, should be removed.

In some physical systems, like the Belykh family, the singularities propagate, intersecting themselves transversally. The Zhang model is not a general position system in this respect, because singularities $\{m/N^k\} \times M \subset I \times M = \hat{M}$ map into themselves, forming zero angle.

3.3.1 Construction of attractors

Define the (spatial) singularity set $S(F) = \cup_{ij} \partial M_{ij}$. Then $U = M \setminus S(F)$ consists of a collection of open connected sets $\mathcal{Z} = \{Z\}$. Let $U_0 := U$ and

$$U_n := \bigcup_{i \in \Lambda} F_i(U_{n-1}) \cap U \, .$$

We say that $x \in S(F)$ is a non-essential singularity of order m if there exists $\epsilon > 0$ and $m > 0$ such that

$$\mathrm{card}\{Z \in \mathcal{Z} \,|\, U_n \cap B_\epsilon(x) \cap Z \neq \emptyset\} \leq 1$$

for all $n > m$. Denote the set of non-essential singularities of order m by $NES(F;m)$ and let $NES(F) := \cup_{m \geq 0} NES(F;m)$ be the set of all non-essential singularities. Define $ES(F) = S(F) \setminus NES(F)$ to be the collection of essential singularities. Observe that there is a natural extension of F to $V_0 = U \cup NES(F)$. In the following we let F denote the extended map. As above we define

$$V_n := \bigcup_{i \in \Lambda} F_i(V_{n-1}) \cap V_0 \, .$$

Let $\mathcal{X} = \cap_{n \geq 0} V_n$ and $\mathcal{D} = \Sigma_N^+ \times \mathcal{X}$. Clearly $F(\mathcal{D}) = \mathcal{D}$. The set $\mathcal{Y} = \overline{\mathcal{X}}$ is called the *physical* (or spatial) attractor of F, and $\mathcal{A} = \Sigma_N^+ \times \mathcal{Y} = \overline{\mathcal{D}}$ is the extended attractor of F.

Proposition 3.9 *$F|_{\mathcal{D}}$ is continuous.*

Proof. The set $\mathcal{D}$ intersects non-essential singularities only. Hence we must show that if x is a non-essential singularity in $\mathcal{D}$, then the extension of each F_i to $NES(F)$ is continuous at the point x. Choose $m \in \mathbb{N}$ and $\varepsilon > 0$ such that $B_\varepsilon(x) \cap U_n$

intersects only one partition element $Z \in \mathcal{Z}$ for $n > m$ and let $y \in B_{\varepsilon/2}(x) \cap \mathcal{D}$. Then $B_{\varepsilon/2}(y) \subset B_\varepsilon(x)$ and so $B_{\varepsilon/2}(y) \cap U_n$ intersects the same partition element Z. So x and y are mapped by the same affine map $F_i|_Z$ for each $i \in \Lambda$. The claim follows. $\square$

In general, the map F does not have a continuous extension to $\mathcal{A}$, but only to $\mathcal{A} \setminus (\Sigma_N^+ \times ES(F))$. Actually, if $x \in ES(F) \cap \mathcal{Y}$ lies on the boundary of several continuity partitions for F_i, then there are several extensions of F to (i, x). Thus we can continuously extend F to $\mathcal{A}$ only when the essential singularities do not intersect the attractor (are removable).

3.3.2 Symbolic Coding

If the singularities can affect the dynamics only for a finite number of iterations, then the dynamics can be well approximated by a topological Markov chain.

Definition 3.2 *We say that singularities are removable if there exists $m \in \mathbb{N}$ such that $S(F) = NES(F, m)$.*

The physically most relevant observables $\phi : \hat{M} \to \mathbb{R}$ are those that are determined by avalanches. We say that ϕ is an *avalanche observable* if it is constant on continuity domains $[i] \times M_{ij}$.

Theorem 3.10 *If singularities are removable, then the map F is well-defined and continuous on $\mathcal{A}$ and there is a topological Markov-chain (Σ_A^+, σ_A^+) and a continuous semi-conjugacy $g : \mathcal{A} \to \Sigma_A^+$ such that for all $\hat{x}, \hat{y} \in \mathcal{A}$ and for all avalanche observables ϕ we have:*

$$g(\hat{x}) = g(\hat{y}) \Rightarrow \phi(F^n(\hat{x})) = \phi(F^n(\hat{y})) \; \forall n \geq 0 \, .$$

The Markov-chain is determined by a matrix A which has a maximal eigenvalue equal to N.

Remark 3.3 *It is clear that all properties related to distribution of avalanche size, duration, area, etc. are invariant under a semi-conjugacy such as this. Observe that for each avalanche observable ϕ on $\mathcal{A}$, there is a unique observable $\phi' : \Sigma_A^+ \to \mathbb{R}$ such that $\phi = \phi' \circ g$. Suppose we have a measure μ on $\mathcal{A}$, and let $\nu = g_* \mu$. If ϕ is an avalanche observable on $\mathcal{A}$, then the statistical properties of ϕ with respect to μ are equivalent to the statistical properties of ϕ' with respect to ν. In this (σ_A^+, Σ_A^+) is a good approximation to $F|_{\mathcal{A}}$. The coding gives estimates on entropy and growth of periodic points, but these estimates are asymptotically no better than what we get from the trivial semi-conjugacy $\hat{M} \to \Sigma_N^+$.*

Proof. Singularities are removable so there exists an integer $m \in \mathbb{N}$ such that $\pi_s \circ F^m(\hat{M})$ only intersects trivial singularities. Let $X_1, \ldots, X_s$ be the closure of the connected components of $\pi_s \circ F^m(\hat{M})$. F is well defined and continuous on these components. Let $Y_1, \ldots, Y_s$ be the intersections of the components $X_1, \ldots, X_s$ with $\mathcal{Y}$. We construct the partition $\mathcal{R} = \{[i] \times Y_k\}$ and enumerate it so that $\mathcal{R} = \{R_1, \ldots, R_r\}$, where $r = Ns$.

Let $A = \|a_{ij}\|$ be the $r \times r$ matrix defined by the rule: $a_{ij} = 1$ if $F(R_i) \cap R_j \neq \emptyset$, and $a_{ij} = 0$ otherwise. A sequence $R_{\omega_0} R_{\omega_1} \ldots$ is legal if $a_{\omega_{t-1}\omega_t} = 1$ for all $t \in \mathbb{N}$. Define $g : \mathcal{A} \to \Sigma_A^+$ by $g(\hat{x}) = (\omega_0 \omega_1 \ldots \omega_t \ldots)$, where $F^t(\hat{x}) \in R_{\omega_t}$. To prove that g is surjective it suffices to show that for each legal sequence $R_{\omega_0} R_{\omega_1} \ldots$, there is a point $\hat{x} \in \pi_s \circ F^m(\hat{M})$ such that $F^i(\hat{x}) \in R_{\omega_t}$ for all $i \in \mathbb{N}$. Note that each ω can be written as a pair (t, k), where $t \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, s\}$. Hence we can write

$$\bigcap_{n=0}^{\infty} F^{-n}(R_{\omega_n}) = \bigcap_{n=0}^{\infty} F^{-n}([t_n] \times Y_{k_n}) = \{\mathbf{t}\} \times \bigcap_{n=0}^{\infty} F_{t_0}^{-1} \circ \cdots \circ F_{t_{n-1}}^{-1}(Y_{k_n}) \,.$$

The continuous image of a connected set is connected, so for each $i = 1, \ldots, N$ and each $k = 1, \ldots, s$ there is a unique $l \in \{1, \ldots, s\}$ such that $F_i(X_k) \subset X_l$. This implies that we have a nested sequence

$$Y_0 \subset F_{t_0}^{-1}(Y_{k_1}) \subset F_{t_0}^{-1} \circ F_{t_1}^{-1}(Y_{k_1}) \subset \ldots$$

and hence the intersection is non-empty.

It is clear that $g_{\mathcal{R}}$ is continuous (see [49] for details). Since the partition $\mathcal{R}$ is a refinement of the continuity partition the conjugacy will be injective up to the classes of points that follow the same continuity domains. Hence if $\phi(F^n \hat{x}) \neq \phi(F^n \hat{y})$ for some avalanche observable ϕ and some $n \geq 0$, then $g(\hat{x}) \neq g(\hat{y})$. $\square$

Remark 3.4 *Suppose we modify the Zhang model by using a full shift (Σ_N, σ_N) as the excitation factor. It is then possible that the modified map F is injective on $\Sigma_N \times \mathcal{Y}$. Since we have strict attraction in the spatial factor after a fixed number of iterations it is clear that we can then obtain an injective coding, and hence a topological conjugacy. However, if we make this modification it is not clear that $\Sigma_N \times \mathcal{Y}$ equals the set*

$$\Omega = \overline{\bigcap_{n=-\infty}^{\infty} F^n(\Sigma_N \times M)} \,.$$

In fact if the maps $F_i|_{\mathcal{Y}}$ are all injective, then $F|_{\Omega}$ is invertible, but $F|_{\Sigma_N \times \mathcal{Y}}$ is typically non-invertible. The reason for this is that, due to contraction, a point $x \in \mathcal{Y}$ does not have preimages for all the maps F_i and F_i are invertible only on $F_i(\mathcal{Y}) \subset \mathcal{Y}$. So to obtain invertibility we must turn to the attractor Ω. From a physical point of view the spatial attractor is of the great interest, so it is desirable

to have an attractor which is a Cartesian product of the Bernoulli shift and the spatial attractor $\mathcal{Y}$.

We can always construct a coding of $F|_{\mathcal{D}}$ (even in non-removable case) by choosing a partition $\mathcal{R} = \{R_1, \dots R_r\}$, and taking $g_{\mathcal{R}} : \mathcal{D} \to \{1, \dots, r\}^{\mathbb{N}}$ to be the map sending a point $\hat{x} \in \mathcal{D}$ to the unique sequence $\omega \in \{1, \dots r\}^{\mathbb{N}}$ such that $F^t(\hat{x}) \in R_{\omega_t}$ for all $t \geq 0$. But there is no reason, however, to expect $g_{\mathcal{R}}(\mathcal{D})$ to be a topological Markov chain, cf. [6].

3.4 Metric properties

The natural volume on $\hat{M}$ is given by the product measure of the uniform Bernoulli measure on Σ_N^+ and the Lebesgue measure on M. By iterating this measure (and averaging) we can construct SRB-measures. However it can happen that the measures constructed are supported on essential singularities, where it is not possible to define the dynamics in such a way that the measure is invariant. Hence we must give some conditions to ensure the existence of SRB-measures. If there is an SRB-measure it is characterized by the fact that its projection to Σ_N^+ coincides with the uniform Bernoulli measure. From this it follows that any SRB-measure is a measure of maximal entropy. In situations where the system allows symbolic coding the SRB-measure corresponds to the Perry measure on the topological Markov chain Σ_A^+.

3.4.1 Existence and characterization of SRB-measures

Let $m = m^u \times m^s$, where $m^u = \mu_{\text{Ber}}$ is the uniform Bernoulli measure on Σ_N^+, and $m^s = \mu_{\text{Leb}}$ is the Lebesgue measure on M. We say that an invariant Borel probability measure μ on $\hat{M}$ has the SRB-property if there exists a measurable invariant set $G \subset \hat{M}$ such that

1. $m(G) > 0$
2. $m^u(\pi_u(G)) = 1$
3. All points $\hat{x} \in G$ are future generic with respect to μ, i.e.

$$\frac{1}{n} \sum_{t=0}^{n-1} \phi(F^t \hat{x}) \to \int \phi \, d\mu \,,$$

for all $\hat{x} \in G$ and all continuous functions $\phi : \hat{M} \to \mathbb{R}$.

For Axiom A attractors one can ensure the existence of measures for which the set of generic points has full Lebesgue measure and this is equivalent to saying that the canonical family of conditional measures on the unstable manifolds are absolutely continuous with respect to the Lebesgue measure. For non-invertible maps one can in general only expect the set of generic points to have positive measure and hence it is unreasonable to require that $m(G) = 1$. Condition 2 is (for physical reasons) important in the Zhang model. It means that the statistical properties do not depend on the choice of a generic sequence $\mathbf{t}$ of excitations. (This is always implicitly assumed in the numerical investigations of the Zhang model that can be found in the physical literature.) Moreover, condition 2 will be satisfied for the SRB-measures that can be constructed by iterating the measure m.

By a standard approach we can give conditions for existence of SRB-measures that hold if singularities are removable, but it is not known if these conditions hold in all non-removable situations.

Proposition 3.11 *Let (ϵ, E_c) does not belong to the negligible set Ξ of Theorem 3.8. If there exists $n \geq 0$, $C > 0$ and $q > 0$ such that*

$$\forall \delta > 0, \forall t \geq 0 : \, m\Big(F^{-t}\big(\Sigma_N^+ \times U_\delta(ES(F;n))\big)\Big) \leq C\delta^q \,,$$

then there exists a set $\mathcal{D} \subset \hat{M}$ (constructed in §3.3.1), which may intersect singularities, and a natural extension of F to $\mathcal{D}$ such that $F(\mathcal{D}) = \mathcal{D}$. Moreover the set $\mathcal{D}$ carries an F-invariant Borel probability measure with the SRB-property.

Remark 3.5 *Proposition 3.11 is a simple modification of the result of Schmeling and Troubetzkoy [53]. In their paper the conditions for existence are in general too restrictive for the Zhang model. In fact, in Example A of §3.7 we show a situation where the SRB-measure constructed in [53] does not exist, but we clearly have existence of a physically relevant measure. The reason for this paradox is that one in general remove all singular points on the construction of the attractor, even if there is a natural extension of F to the points of singularity. (Proposition 3.11 obviously applies to this example since $ES(F;5) = \emptyset$.)*

Proof. In [53] it is shown that a piecewise smooth map f with singularity set S has a measure, not supported on singularities, such that the set of generic points has positive Lebesgue measure. They require that the following conditions are satisfied:

1. The restrictions of f to each of its continuity domains are diffeomorphisms onto their image.

2. The second differentials $D^2 f_x$ does not grow too fast close to singularities. (See [53] for a more precise formulation.)

3. f is hyperbolic. In this context this means that there are constants $C > 0$ and $\lambda \in (0,1)$ such that for all $x \notin S$ there is a splitting of the tangent space at x into subspaces $E^+(x)$ and $E^-(x)$. There are cones $C^+(x)$ and $C^-(x)$ around $E^+(x)$ and $E^-(x)$ that are invariant under Df_x and Df_x^{-1} respectively. The angles between the $C^+(x)$ and $C^-(x)$ are bounded away from zero, and for all points x that do not intersect singularities in the first n iterations it holds:

$$\|D_x f^n(v)\| \geq C^{-1}\lambda^{-n}\|v\| \text{ for } v \in C^+(x)\,,$$

and

$$\|D_x f^n(v)\| \leq C\lambda^{n}\|v\| \text{ for } v \in C^-(x)\,.$$

4. There exists $C > 0$ and $q > 0$ such that $m(f^{-t}(U_\varepsilon(S))) \leq C\varepsilon^q$ for all $\varepsilon > 0$ and all $t \in \mathbb{N}$.

We apply this result to the map $F|_{\Sigma_N^+ \times (U_n \setminus ES(F;n))}$. The singularity set for this map is contained in $\Sigma_N^+ \times ES(F)$, so by assumption condition 4 is satisfied. Condition 1 follows from Proposition 3.7, condition two is obviously satisfied since F is piecewise affine and condition 3 follows from Theorem 3.5 with $E^+ = \mathbb{R}^1 \oplus 0$, $E^- = 0 \oplus \mathbb{R}^N$ and $C^\pm$ being the regular cones around them (actually Theorem 3.5 ensures hyperbolicity for some iterate F^T, which implies the claim).

In [53] the measures are constructed by iterating m, averaging and taking a weak limit. It is clear that, in the Zhang model, any measure obtained in this way will satisfy condition 2 in our definition of an SRB-measure. □

If an SRB-measure exists it can be characterized by a number of different properties. From a physical perspective it is reasonable to require that a relevant invariant measure should preserve the uniform Bernoulli structure on Σ_N^+. This corresponds to the Lebesgue measure on $[0,1]$ in the alternative formulation of the map F, and hence to absolutely continuous measure conditional measures on the unstable space $[0,1]$.

Proposition 3.12 *If μ is an SRB-measure on $\mathcal{D}$, then $\mu^u := (\pi_u)_*\mu$ is the uniform Bernoulli measure on Σ_N^+.*

Proof. There is a set $A = \pi_u(G)$ of full m^u-measure, such that all $\mathbf{t} \in A$ are generic with respect to μ^u. Take a continuous function $\phi : \Sigma_N^+ \to \mathbb{R}$. Then

$$\int \phi\, d\mu^u = \lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} \phi((\sigma_\Lambda^+)^t \mathbf{t}) = \int \phi\, dm^u\,,$$

where the left equality holds for $\mathbf{t} \in A$ and the right one for $\mathbf{t} \in B$ with $B \subset \Sigma_N^+$ a subset of full m^u-measure (from Birkhoff ergodic theorem). Since $A \cap B \neq \emptyset$, we get: $\int \phi\, d\mu^u = \int \phi\, dm^u$ for all continuous functions ϕ. □

3.4.2 Measures of maximal entropy

Suppose that there exists an invariant Borel probability measure μ on $\hat{M}$. Let $\mu^u := (\pi_u)_*\mu$ and let $\{\nu_{\mathbf{t}}\}$ to be the canonical family of conditional measures on the fibers $\pi_u^{-1}(\{\mathbf{t}\})$. By the Abramov-Rokhlin formula

$$h_\mu(F) = h_{\mu^u}(\sigma_N^+) + h_\mu(F|\sigma_N^+)\,,$$

where

$$h_\mu(F|\sigma_N^+;\mathcal{Q}) = \lim_{n\to\infty}\frac{1}{n}\int H_{\nu_{\mathbf{t}}}\Big(\bigvee_{k=0}^{n-1}(F_{t_{k-1}}\circ\cdots\circ F_{t_0})^{-1}(\mathcal{Q})\Big)\,d\mu^u(\mathbf{t})\,,$$

for a partition $\mathcal{Q}$ and

$$h_\mu(F|\sigma_N^+) = \sup_{\mathcal{Q}} h_\mu(F|\sigma_N^+;\mathcal{Q})\,,$$

The supremum is taken over all finite measurable partitions $\mathcal{Q}$ of $\hat{M}$. This formula was originally proved for product measures by Abramov and Rokhlin [1] and extended to arbitrary skew products by Bogenschultz and Crauel [11].

Below we use the notation $F_{\mathbf{t}}$ for the dynamics over a pre-fixed sequence $\mathbf{t} = (t_0t_1\dots)\in\Sigma_N^+$. By n-the iteration we mean the map $F_{\mathbf{t}}^n = F_{t_{n-1}}\circ\cdots\circ F_{t_0}$.

Theorem 3.13 *If μ is an invariant Borel probability measure on $\hat{M}$, then $h_\mu(F) = h_{\mu^u}(\sigma_N^+)$. In particular, for an SRB-measure $h_\mu(F) = \log N$.*

Proof. We prove the proposition by estimating $h_\mu(F|\sigma_N^+)$ from above. Let $\mathcal{Q}$ be a partition of M and

$$\mathcal{Q}_{t_0\dots t_{n-1}} := \bigvee_{k=0}^{n-1}(F_{t_{k-1}}\circ\cdots\circ F_{t_0})^{-1}(\mathcal{Q}) = \bigvee_{k=0}^{n-1}F_{\mathbf{t}}^{-k}(\mathcal{Q})\,.$$

Fix $\varepsilon>0$ and choose the partition $\mathcal{Q}$ such that $h_\mu(F|\sigma_N^+;\mathcal{Q})+\varepsilon \geq h_\mu(F|\sigma_N^+)$. The maps F_i are non-expanding so it follows from the Ruelle-Margulis inequality [32] that

$$\frac{1}{n}H_{\nu_{\mathbf{t}}}\Big(\bigvee_{k=0}^{n-1}F_{\mathbf{t}}^{-k}(\mathcal{Q})\Big)\longrightarrow h_{\nu_{\mathbf{t}}}(F_{\mathbf{t}}) = 0.$$

Moreover the convergence is μ^u-uniform and so the same holds for the integrals. Another way to see it is via the multiplicity notion of §3.5.2 (then $\mathcal{Q}$ should be subordinate to each continuity partition $\{M_{ij}\,|\,j=1,\dots,q_i\}$):

$$h_\mu(F|\sigma_N^+;\mathcal{Q}) \leq \lim_{n\to\infty}\frac{1}{n}\log\max_{|\mathbf{t}|\leq n}\mathrm{mult}(\mathcal{Q}_{t_0\dots t_{n-1}}\cap\mathrm{supp}(\nu_{\mathbf{t}})).$$

Therefore $h_\mu(F|\sigma_N^+)\leq\epsilon$. Let $\epsilon\to 0$. $\square$

Remark 3.6 *Theorem 3.13 is a partial case of Theorem 4 from [36].*

We say that an invariant measure μ is maximal if $h_\mu(F) = \sup_\nu h_\nu(F)$, where the supremum is taken over all invariant Borel probability measures on M. It follows from Theorem 3.13 that $h_\mu(F) \leq \log N$. So if $h_{\text{top}}(F) > \log N$, the variational principle fails (this can happen for piece-wise affine systems, see [36, 37]). But we show in §3.5.2 that the abnormal growth of $h_{\text{top}}(F)$ does not occur in the Zhang model, at least for generic values of parameters E_c, ϵ.

Corollary 3.1 *Any SRB-measure on $\mathcal{D}$ has entropy $\log N$ and is hence a maximal measure.*

Corollary 3.2 *Suppose singularities are removable and that μ is an SRB-measure on $\mathcal{A}$. Let $g : \mathcal{A} \to \Sigma_A^+$ be the semi-conjugacy constructed in the proof of Theorem 3.10. If (σ_A^+, Σ_A^+) is topologically transitive, then $g_*\mu$ is the Perry measure on Σ_A^+.*

Proof. A transitive topological Markov chain has a unique measure of maximal entropy. This measure is called the Perry measure [32]. □

3.4.3 Hyperbolic structure

There are several ways to define Lyapunov exponents for the Zhang model. The Zhang model can be represented as a piecewise affine map, where Bernoulli shift is represented as the expanding map $t \mapsto Nt \mod 1$ of the interval (see §3.5.2). Hence it is clear that there is one positive Lyapunov exponent $\chi_0^+ = \log N$. We define the other exponents by introducing the co-cycle $\mathcal{T} : \hat{M} \to GL(N, \mathbb{R})$, defined by $\mathcal{T}(\hat{x}) = L_{ij}$, where $\hat{x} \in [i] \times M_{ij}$. For $\hat{x} \in \hat{M}$ and $v \in \mathbb{R}^N \setminus \{0\}$ we define

$$\chi(\hat{x}, v) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log \frac{\|\mathcal{T}(F^{n-1}(\hat{x})) \dots \mathcal{T}(F(\hat{x}))\mathcal{T}(\hat{x})v\|}{\|v\|}.$$

It is a general fact that the function $\chi(\hat{x}, \cdot)$ takes at most N different values $\chi_1^-(\hat{x}) \geq \dots \geq \chi_N^-(\hat{x})$.

Proposition 3.14 *For all $\hat{x} \in \hat{M}$ the Lyapunov spectrum is:*

$$0 > \chi_1^-(\hat{x}) \geq \dots \geq \chi_N^-(\hat{x})$$

and for (ϵ, E_c) outside the negligible set Ξ from Theorem 3.8: $\chi_N^-(\hat{x}) > -\infty$.

Proof. From Theorem 3.5 we know that there exists $T \in \mathbb{N}$ and $c \in (0, 1)$ such that

$$\|\mathcal{T}(F^{T-1}(\hat{x})) \dots \mathcal{T}(F(\hat{x}))\mathcal{T}(\hat{x})\| \leq c$$

for all $\hat{x} \in \hat{M}$. It immediately follows that $\chi(\hat{x}, v) \leq T^{-1} \log c < 0$.

For $\epsilon \geq 1/2$ and arbitrary E_c or for $\epsilon > 0$ and $E_c \geq (1+\epsilon)/(1-\epsilon)$ all linear maps are invertible, and so for all $\hat{x} \in \hat{M}$ and all $v \in \mathbb{R}^N \setminus \{0\}$ we have $\chi(\hat{x}, v) \geq \log k$, where $k = \min_{ij} \min \mathrm{Sp}(L_{ij}) > 0$. $\square$

If there exists a unique SRB-measure, then it follows from the Osceledec theorem that there are numbers $\chi_1^-, \ldots, \chi_N^-$ such that $\chi_i(\hat{x}) = \chi_i^-$ for Lebesgue almost every $\hat{x} \in \hat{M}$. The numbers $\chi_0^+, \chi_1^-, \ldots, \chi_N^-$ are the Lyapunov exponents of the Zhang model. From Proposition 3.14 it follows that the Zhang model is hyperbolic in the sense that the Lyapunov spectrum consists of:

$$\chi_0^+ = \log N > 0 > \chi_1^- \geq \cdots \geq \chi_N^- > -\infty\,.$$

We see from Corollary 3.1 that the Pesin formula $h_\mu(F) = \chi^+$ holds for any SRB-measure. If there is no SRB-measure then the Lyapunov spectrum should be defined as functions on M:

$$\chi_i(x) = \int_{\Sigma_N^+} \chi_i(\mathbf{t}, x)\, d\mu_{\mathrm{Ber}}\,,$$

where μ_{Ber} is the uniform Bernoulli measure on Σ_N^+.

3.4.4 Entropy of physical vs. mathematical models

We can reformulate the Zhang system as the map $\hat{f} : \hat{B} \to \hat{B}$, where $\hat{B} = \cup_{i \geq 0} \hat{f}^i(\hat{M})$ is a compact $\hat{f}$-invariant subset of $\Sigma_N^+ \times \mathbb{R}^N$. We wish to compare this to the induced transformation $F : \hat{M} \to \hat{M}$ (cf. Remark 3.1).

If μ is an F-invariant Borel probability measure on $\hat{M}$, then there is an associated $\hat{f}$-invariant Borel probability measure $\hat{\mu}$ on $\hat{B}$ (and vice versa). Abramov's theorem ([15]) relates the entropies of both systems:

$$h_{\hat{\mu}}(\hat{f}) = h_\mu(F) \cdot \hat{\mu}(\hat{M}). \tag{3.3}$$

One does not need to assume ergodicity and can allow degenerations [22], as happens for the case of Zhang model. In ergodic situation by the recurrence theorem of Kac [15] for a μ-generic point $\hat{x} \in \hat{M}$:

$$\frac{1}{\hat{\mu}(\hat{M})} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \tau(F^k \hat{x}), \tag{3.4}$$

where $\tau(\hat{x}) = \tau(i, x)$ is the avalanche time initiated by addition of e_i to $x \in M$ (see §3.2.2). If $\mu = \mu_{\mathrm{SRB}}$ is a unique SRB-measure, the above point $\hat{x}$ can be chosen

Lebesgue generic. The resulting limit is the average avalanche time $\bar{\tau}$ (we discuss it in more details in §3.8.1-3.8.2) and we obtain:

$$h_{\hat{\mu}}(\hat{f}) = h_{\mu}(F)/\bar{\tau} \quad \text{resp.} \quad h_{\hat{\mu}_{\mathrm{SRB}}}(\hat{f}) = h_{\mu^u}(\sigma_N^+)/\bar{\tau}.$$

In general non-ergodic situation to get equality (3.4) we should integrate the terms in right-hand side and then we again obtain the average avalanche size $\langle\tau\rangle$, but now it is the space-average. Substituting this into (3.3) we get:

$$h_{\hat{\mu}}(\hat{f}) = h_{\mu}(F)/\langle\tau\rangle. \tag{3.5}$$

For SRB-measures this formula can be also obtained from Ledrappier-Young theorem [38], because we have only one positive Lyapunov exponent.

3.5 Topological entropy

To calculate the topological entropy of F we established in [36] a set of inequalities, using the technique developed by J. Buzzi [16], [17] for piecewise expanding maps and piecewise isometries, see Appendix A. The contraction in the maps F_{ij} provides difficulties, so several results were generalized to fit the framework of the Zhang model. It is not however true (as was widely believed) that the contraction does not contribute to topological (contrary to metric) entropy, the corresponding counter-example can be found in [37]. The Zhang model has a feature common to all such examples [36], namely angular expansion, but still for most values of the parameters this abnormal increase of the entropy does not occur.

3.5.1 Growth of the number of continuity domains

Let $\mathcal{P} = \{[i] \times M_{ij}\}$ be the partition of continuity for F, and enumerate the elements so that $\mathcal{P} = \{P_1, \dots P_r\}$. Let

$$[P_{a_0} \dots P_{a_{n-1}}] := \bigcap_{m=0}^{n-1} F^{-m}(P_{a_m}),$$

and

$$\mathcal{P}^n = \{[P_{a_0} \dots P_{a_{n-1}}] \neq \emptyset \,|\, a_m = 1, \dots, r\}.$$

We define the singularity entropy of F by

$$H_{\mathrm{sing}}(F) = \lim_{n\to\infty} \frac{1}{n} \log \mathrm{card}(\mathcal{P}^n).$$

Remark 3.7 *Define a map $g : \mathcal{A} \to \Sigma_r^+$ by letting $g(\hat{x})$ be the unique sequence $a_0 a_1 \dots$ such that $F^n(\hat{x}) \in P_{a_n}$. Then*

$$H_{\text{sing}}(F) = h_{\text{top}}(\sigma_r^+; g(\mathcal{A}))\,.$$

For piecewise affine expanding maps and piecewise isometries it is clear that this also equals the topological entropy, but due to the contraction, this is not obvious in the Zhang model. In addition, if singularities are not removable, the map g has discontinuities.

Definition 3.3 *Call a point $x \in S(F)$ an unstable singularity if for all i and $k \neq l$ we have: $\lim_{y\to x} F_{ik}(y) \neq \lim_{y\to x} F_{il}(y)$.*

Theorem 3.15 *If all singularities $S(F) \cap \mathcal{Y}$ are unstable, then $h_{\text{top}}(F) = H_{\text{sing}}(F)$.*

We need the following technical lemma:

Lemma 3.16 *If the singularities in $\mathcal{Y}$ are unstable, then there exists a constant $\gamma > 0$ such that for all $\delta > 0$, $x \in \mathcal{Y} \cap M_{ik}$ and $y \in \mathcal{Y} \cap M_{il}$, $k \neq l$, we have:*

$$d(x, y) < \delta \Rightarrow d(F_i(x), F_i(y)) > \delta\,.$$

Proof. Suppose that for all $\delta > 0$ there exists $x \in M_{ik} \cap \mathcal{Y}$ and $y \in M_{il} \cap \mathcal{Y}$ such that $d(x, y) < \delta$ and $d(F_{ik}(x), F_{il}(y)) \leq \delta$. There exist sequences $\{x_m\} \subset M_{ik}$ and $\{y_m\} \subset M_{il}$ such that $d(x_m, y_m) \to 0$ and $d(F_{ik}(x_m), F_{il}(y_m)) \to 0$. The sequence $\{x_m\}$ has a convergent subsequence $x_{m_n} \to z$. The point z lies in $S(F)$ and $y_{m_n} \to z$. By the continuity of the maps F_{ik} and F_{il} we have $F_{ik}(x_{m_n}) \to F_{ik}(z)$ and $F_{il}(y_{m_n}) \to F_{il}(z)$. Since the metric d is continuous on $M \times M$ we have

$$d(F_{ik}(z), F_{il}(z)) = \lim_{n\to\infty} d(F_{ik}(x_{m_n}), F_{il}(y_{m_n})) = 0\,.$$

Hence $F_{ik}(z) = F_{il}(z)$. Contradiction. □

Proof of Theorem 3.15. Let

$$[P_{a_0} \dots P_{a_{n-1}}] = [i_0 \dots i_{n-1}] \times K\,,$$

where $K \subset M$ is a convex polygon. Fix $\delta > 0$ and set $k = [\log 1/\delta]$. Let $z_1, \dots, z_{m(\delta)}$ be a δ-spanning set for K. Chose $\mathbf{t} \in \Sigma_N^+$ and define N^k sequences

$$\mathbf{s}_{r_0 \dots r_{k-1}} = (i_0 \dots i_{n-1} r_0 \dots r_{k-1} t_{n+k} t_{n+k-1} \dots) \in \Sigma_N^+\,.$$

Since the maps F_{ij} are contracting it is clear that the set $\{\hat{x} = (\mathbf{s}_{r_0 \dots r_{k-1}}, z_l)\}$ is a (n, δ)-spanning set for $[P_{a_0} \dots P_{a_{n-1}}]$, and since the minimum number of balls

needed to cover a convex polygon $K \subset M$ is bounded by $m(\delta) \le C_N \delta^{-N}$ we see that the number of (n,δ)-balls to cover $[P_{a_0} \dots P_{a_{n-1}}]$ is bounded by $m(\delta) N^k \cdot \#\{[P_{a_0} \dots P_{a_{n-1}}]\}$. Therefore we get an estimate for the number of (n,δ)-balls to cover $\hat{M}$ and so

$$h_{\mathrm{top}}(F) \le \lim_{n\to\infty} \frac{\log\bigl(C_\delta \delta^{-N} \operatorname{card}(\mathcal{P}^n)\bigr)}{n} = H_{\mathrm{sing}}(F)\,.$$

To see the opposite inequality let $A, B \in \mathcal{P}^n$ and take $\hat{x}_1 = (\mathbf{t}, x) \in A$ and $\hat{x}_2 = (\mathbf{s}, y) \in B$. Suppose that $A \ne B$ and that $t_0 = s_0, \dots, t_{n-1} = s_{n-1}$. Then there is $m < n$ such that $\pi_s \circ F^m(\hat{x}) \in M_{t_m k}$ and $\pi_s \circ F^m(\hat{y}) \in M_{t_m l}$, $k \ne l$. By Lemma 3.16 there is $\gamma > 0$ such that for all $\xi < \gamma$:

$$\max\{d(\pi_s \circ F^m(\hat{x}), \pi_s \circ F^m(\hat{y})), d(\pi_s \circ F^{m+1}(\hat{x}), \pi_s \circ F^{m+1}(\hat{y}))\} \ge \xi\,.$$

Therefore for δ sufficiently small, no $(n+1,\delta)$-ball can contain points of both A and B and the minimal $(n+1,\delta)$-spanning set has at least $\operatorname{card}(\mathcal{P}^n)$ elements. Then $h_{\mathrm{top}}(F) \ge H_{\mathrm{sing}}(F)$. □

Theorem 3.17 *For the Zhang models:* $h_{\mathrm{top}}(\hat{f}) = H_{\mathrm{sing}}(\hat{f})$, $h_{\mathrm{top}}(F) = H_{\mathrm{sing}}(F)$.

Proof. By Remark 3.7 all quantities are topological entropies. The corresponding systems in the second equality are the Poincaré return maps for the transformations of the first equality. The map $F : \hat{M} \to \hat{M}$ is already the return map for $\hat{f} : \hat{B} \to \hat{B}$ by the very construction.

To achieve the same claim for the symbolic system we extend the partition $\mathcal{P}$ of $\hat{M}$ to a partition $\tilde{\mathcal{P}}$ of $\hat{B}$, which on $\Sigma_N^+ \times (\mathbb{R}^N_{\ge 0} \setminus M) \cap \hat{B}$ equals the product of the standard partition $\Sigma_N^+ = \cup_i [i]$ and the partition of the spacial part by the hyperplanes $\{x_i = E_c\}$. Denote by $s \le r + N(2^N - 1)$ the number of elements of the new partition $\tilde{\mathcal{P}}$.

Let $\mathcal{B} \subset \hat{B}$ be the $\hat{f}$-invariant closure of $\mathcal{A}$ in $\Sigma_N^+ \times \mathbb{R}^N_{\ge 0}$. Define a map $\hat{g} : \mathcal{B} \to \Sigma_s^+$ by letting $\hat{g}(\hat{x})$ be the unique sequence $b_0 b_1 \dots$ such that $\hat{f}^n(\hat{x}) \in \tilde{\mathcal{P}}_{b_n}$. We wish to prove that

$$h_{\mathrm{top}}(\hat{f}) = H_{\mathrm{sing}}(\hat{f}) = h_{\mathrm{top}}(\sigma_s^+ | \hat{g}(\mathcal{B}))\,. \tag{3.6}$$

For this it is sufficient to check that the singularities $S(\hat{f}) \cap \pi_s(\mathcal{B})$ are unstable for $\hat{f}$ (the second equality follows from the definition).

Consider a singular point $x \in \mathbb{R}^N_{\ge 0}$. Let y, z tend to x by two different domains of the projected partition $\tilde{\mathcal{P}}$ in $\mathbb{R}^N_{\ge 0}$. Then we can subdivide $\{1, \dots, N\} = A \cup B \cup C$, where $y_i = E_c + 0$, $z_i = E_c - 0$ (with the obvious notations instead of limits) for $i \in A$, $y_i = E_c - 0$, $z_i = E_c + 0$ for $i \in B$ and y_i, z_i belong to the same side of E_c

for $i \in C$. Then for $\tilde{S} = S(y) - S(z)$ we have:

$$\tilde{S}_{ij} = \begin{cases} \epsilon - 1, & \text{if } i = j \in A, \\ 1 - \epsilon, & \text{if } i = j \in B, \\ \frac{1-\epsilon}{2d}, & \text{if } d_\Lambda(i,j) = 1, j \in A, \\ \frac{\epsilon-1}{2d}, & \text{if } d_\Lambda(i,j) = 1, j \in B, \\ 0 & \text{otherwise.} \end{cases}$$

We should check that $\tilde{S} \cdot v \neq 0$ for any vector $v \in \mathbb{R}^N_{\geq 0}$ with components $v_i = E_c$ for $i \in A \cup B$ (this implies that x is unstable). Note that other components v_i, $i \in C$ do not contribute to the product.

Let i be a site from A with not all neighbors from A (if the set A is empty consider B). Denote the number of A-neighbors of i by $k_A < 2d$ and the number of B-neighbors by $k_B \geq 0$. Then $(\tilde{S} \cdot v)_i = E_c(2d - k_A + k_B)\frac{\epsilon-1}{2d} \neq 0$.

Thus Theorem 3.15 implies (3.6). Moreover, the same reasons yield a more general statement. Namely, since $\hat{g}$ is a semi-conjugacy, we get:

$$h_{\text{top}}(\hat{f}|\mathcal{K}) = h_{\text{top}}(\sigma_s^+|\hat{g}(\mathcal{K})) \tag{3.7}$$

for any subset $\mathcal{K} \subset \mathcal{B}$ (recall that $\hat{g}$ may have discontinuities, but our arguments are not injured by this fact). This subset needs not to be invariant, and in this case we should use Bowen's definition of entropy [12]. Thus Katok's entropy formula [31] implies that $h_{\hat{\mu}}(\hat{f}) = h_{\hat{g}_*\hat{\mu}}(\sigma_s^+)$ for all $\hat{f}$-invariant measures $\hat{\mu}$, which are not supported on singularities.

If the system $(\hat{M}, F)$ possesses a measure μ of maximal entropy, then by the obtained result the second claim of the theorem follows from (3.5) and Kac's theorem [47]. In general, we can apply the above arguments to the partition of $\hat{M}$ by the subsets of equal return times and using the fact, that both returns of $\hat{f}$ and $\hat{g}$ have the same combinatorics, we get: $h_{\text{top}}(F) = h_{\text{top}}(\sigma_r^+|g(\mathcal{A}))$ (see, for instance, the loop equation approach [46]). □

3.5.2 Evaluation of topological entropy

It was predicted on the base of variational principle in [6] that topological entropy of the Zhang model is $h_{\text{top}}(F) = \log N$. However this principle does not apply because the map is not well-defined (continuously) on the whole space (or thanks to non-compactness if we remove the singularities). In fact, there can be no invariant measures on the non-singular part at all.

While we support the claim that $h_{\text{top}}(F) = \log N$, it will not be proved in full generality. We start with the asymptotic statement.

Consider the bifurcation diagram on the (E, ϵ) strip $(0, +\infty) \times [0, 1)$, where a point is critical if in its neighborhood dynamics of the Zhang model can experience

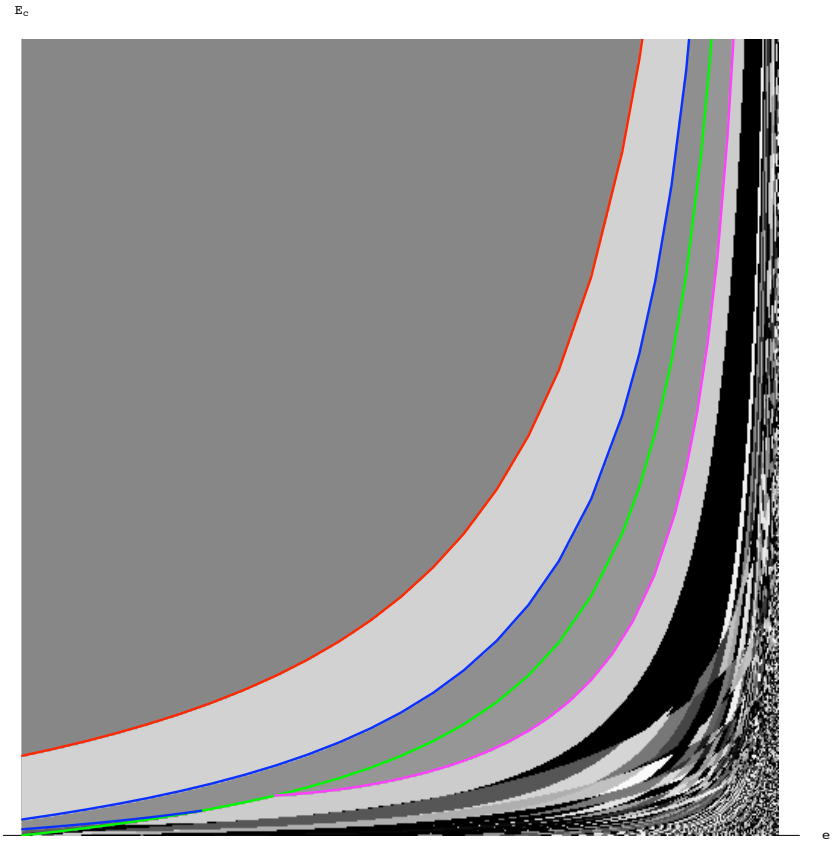


Figure 3.1: Shows the avalanche type domains on (ϵ, E_c)-bifurcation diagram for $N = 2$. The top domain is $E_c > \frac{1+\epsilon}{1-\epsilon}$. The line from infinity to the origin is $E_c = \frac{\epsilon}{1-\epsilon}$. We see infinitely many domains with different avalanches, accumulating to two of the axes.

avalanches of different types. Thus the strip is partitioned into different avalanche type domains. The partition depends on L, d. For $N = 2$ the diagram is shown on figure 3.1.

Note that for all L, d there is the top avalanche type domain representing the shortest avalanche time. For $N = 2$ it is given by the relation $E > \frac{1+\epsilon}{1-\epsilon}$. Also note that some domains have ϵ-projection strictly smaller than the interval $[0, 1)$. The next statement concerns only the top domain and the avalanche type domains that are adjacent to the line $\epsilon = 1$.

Theorem 3.18 *For the Zhang model:* $0 \leq h_{\text{top}}(F) - \log N \leq \theta(E, \epsilon)$, *where* $\lim_{E_c \to \infty} \theta(E_c, \epsilon) = 0$ *(ϵ fixed) and* $\lim_{\epsilon \to 1} \theta(E_c, \epsilon) = 0$. *In the latter case E_c changes accordingly with ϵ so that (E_c, ϵ) belongs to the same avalanche type domain (then $E_c \to \infty$, though this case differs from the former).*

We can remove the N-rational points from Σ_N^+ and then represent it as the

subset of $I = [0,1]$ with points n/N^k being deleted (this process does not change the topological entropy). In fact, we need to remove only points n/N for the others will be deleted by inverse iterations of the map of $I \times X$ with the formula $F(t,x) = (Nt, f_{[Nt]}(x))$. Thus we represent our system as a piece-wise affine partially hyperbolic map of the subset of $\mathbb{R}^{1+N}$. This is important for an application of the results from Appendix A.

The proof uses the notions of multiplicity and multiplicity entropy, due to G. Keller and J. Buzzi. Given a finite partition $\mathcal{Z} = \{Z_k\}$ of $\hat{M}$ we define

$$\mathcal{Z}^n = \{[Z_0 \dots Z_{n-1}] \neq \emptyset\},$$

where

$$[Z_0 \dots Z_{n-1}] = Z_0 \cap F^{-1}(Z_1) \cap \dots \cap F^{1-n}(Z_{n-1}).$$

Define multiplicity of a partition by

$$\mathrm{mult}(\mathcal{Z}) = \max_{\hat{x} \in \hat{M}} \mathrm{card}\{Z \in \mathcal{Z} \,|\, \overline{Z} \ni x\},$$

If $\mathcal{Z}$ is the continuity partition of the map F we often denote the multiplicity of the $\mathcal{Z}$ by $\mathrm{mult}(F)$. Then it is clear that $\mathrm{mult}(F^n) = \mathrm{mult}(\mathcal{Z}^n)$. The multiplicity entropy of F is (the limit exists by subadditivity, cf. [32])

$$H_{\mathrm{mult}}(F) = \sup_{\mathcal{Z}} \lim_{n\to\infty} \frac{1}{n} \log \mathrm{mult}(\mathcal{Z}^n),$$

and we see that if $\mathcal{Z}$ is the continuity partition, then

$$H_{\mathrm{mult}}(F) = \lim_{n\to\infty} \frac{1}{n} \log \mathrm{mult}(F^n).$$

It is clear that $H_{\mathrm{mult}}(F) = 0$ if the singularities are removable, so $h_{\mathrm{top}}(F) = \log N$ by the result in Appendix A. For big $E_c \gg 1$ the singularities are generally non-removable, but still we have the same effect asymptotically:

Proof of Theorem 3.18. We take $\theta(E_c, \epsilon) = H_{\mathrm{mult}}(F)$. Since the singularities $t = n/N \in I$ of the map $\sigma_N^+ : t \mapsto Nt$ do not intersect in inverse iterations, the multiplicity growth is only due to the spacial maps $F_i : M \to M$. Thus using the notation $F_{\mathbf{t}}$ for the dynamics over a prefixed sequence of excitations we obtain $H_{\mathrm{mult}}(F) = \sup_{\mathbf{t} \in \Sigma_N^+} H_{\mathrm{mult}}(F_{\mathbf{t}})$.

To show the first claim let us notice that when $\epsilon = \mathrm{const}$, but $E_c \to \infty$, then the avalanches map the critical part of the boundary $\{\exists i : x_i = E_c\} \subset \partial M$ far from ∂M, namely to the distance $\sim \gamma(\epsilon, L, d) E_c$, see Figures 3.16 for $N = 2$ and 3.17 for $N = 3$. To reach again the boundary and experience avalanche we need many shifts F_i by the basic vectors e_i. Thus the singularity can meet only after big number

of iterations. Since the initial picture of singularities has bounded multiplicity (for (ϵ, E_c) from the top avalanche type domain), the multiplicity decreases at least as k/E_c so that it vanishes in the limit.

To prove the second claim we use the inequality $H_{\text{sing}}(F) \leq \sum \rho_i(F)$ from Appendix A.3. If the avalanche type domain is fixed, the number of compositions of matrices $S(x)$ in one avalanche (see §3.2.3) is bounded. Every such a matrix tends to identity when $\epsilon \to 1$. Thus all the linear parts L_{ij} of avalanche maps F_{ij} tend to identity and so angular expansions $\rho_i(F)$ tend to zero. □

Notice that if E_c is fixed, but $\epsilon \to 1$, then the number of avalanches has unlimited grow and the previous argument do not work. However, due to estimates of §3.2.2 on the maximal avalanche length τ_m we conclude that $\theta(E_c, \epsilon) \to 0$ as either $E_c \to \infty$ or $\epsilon \to 1$, both quantities being related by the constraint $E_c \geq N(1-\epsilon)^{-\frac{1}{2}\operatorname{diam}(\Lambda)-\sigma}$ for some $\sigma > 0$ (we don't require, but allow $N \to \infty$ as well). This statement is stronger than in the theorem.

Now we are going to prove vanishing of $\theta(E_c, \epsilon)$ a.e. in the finite part.

Theorem 3.19 *For generic (E_c, ϵ) we have: $h_{\text{top}}(F) = \log N$.*

We are going to prove the statement not only for the Zhang model, but also for nearly Zhang models. By this we mean the following. The map F is a bundle over the Bernoulli shifts σ_N^+ with factors F_i being piece-wise affine partially contracting maps. That is $M = \cup_j M_{ij}$ and for $F_{ij} = F_i|_{M_{ij}}$ we have $F_{ij}(x) = L_{ij}(x) + b_{ij}$, $b_{ij} = L_{ij}e_i$. We are going to make arbitrary small generic perturbations of the matrices L_{ij} (still with spectrum within unit ball) and vectors b_{ij} (one should care that M_{ij} are mapped into M) and prove the statement for this modified system.

Due to round-off errors there is no much difference between the original and the perturbed systems in computer simulations. And from the point of view of the experiment such perturbations (instrument instability) are indispensable – look at [52] for the discussion of physical relevance of variation of parameters as the noise.

Proof of Theorem 3.19. We claim that for generic (E_c, ϵ) the singularities do not multiply. Actually, some intersections of singularities are deformable as we vary the parameters, just by the transversality reasons, but the other disappear with small perturbations.

Namely, for a multi-index $\sigma = (\alpha_1, \ldots, \alpha_k)$, $\alpha_s = (i_s, j_s)$ coding an orbit, denote $F_\sigma = F_{\alpha_k} \circ \cdots \circ F_{\alpha_1}$ the corresponding map along the orbit. The singularities of this map are: $\operatorname{Sing}(F_\sigma) = \cup_{r=1}^k F_{\sigma_{[r]}}^{-1} \operatorname{Sing}(F_{\alpha_r})$, where $\sigma_{[r]} = (\alpha_1, \ldots, \alpha_r)$.

If $z_0 \in \operatorname{Sing}(F_\sigma)$, then its orbit (with multi-possibilities due to singularities: mapping a singular point we extend the components of the map F_σ in various ways) meets several singularity planes, i.e. for some cuts $\sigma_s = \sigma_{[r_s]}$ of σ, $s = 1, \ldots, m$, we have the following system:

$$F_{\sigma_1}(z_0) = z_1, \ldots, F_{\sigma_m}(z_0) = z_1, \; l_{q_1}(z_1) = 0, \ldots, l_{q_m}(z_m) = 0, \tag{3.8}$$

where $l_{q_s}(z)$ are equations for the singularity hyperplanes of the corresponding map F_i (there are also inequalities, which we don't mention). When $m \le N$ there are occasions, when (3.8) has a solution continuously depending on (E_c, ϵ). However for $m > N$ this is no longer the case. In fact, considering nearly-Zhang models we see that for generic data the above system (3.8) is characterized by a collection of non-trivial polynomial in ϵ equations (for each E_c).

More precisely, the set of (L_{ij}, b_{ij}) giving trivial polynomials has positive codimension and hence zero Lebesgue measure. Uniting these sets over all choices of multi-indices $(\sigma_1, \dots, \sigma_m)$ we see that the complement has full measure and dimension and so a generic perturbation yields the data (L_{ij}, b_{ij}) from it. Since non-trivial polynomials have only finite number of zeros, then for a generic nearly Zhang model and every E_c there is a countable subset of $\{0 \le \epsilon < 1\}$, so that the corresponding systems (3.8) have no solutions. This means that multiplicity of $F_{\mathbf{t}}^n$ does not grow with n and so $H_{\mathrm{mult}}(F_{\mathbf{t}}) = 0$.

Now let us look to the Zhang model. We restrict for simplicity of exposition to the case of two sites $N = 2$ (the other configurations involve more complicated combinatorics). In this case the linear parts of the affine maps are compositions of $L_1 = \mathbf{1} + (\epsilon - 1)A_1$ and $L_2 = \mathbf{1} + (\epsilon - 1)A_2$ with

$$A_1 = \begin{bmatrix} 1 & 0 \\ -1/2 & 0 \end{bmatrix}, \quad A_2 = J^{-1} A_1 J = \begin{bmatrix} 0 & -1/2 \\ 0 & 1 \end{bmatrix}, \text{ where } J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(we exclude the obvious matrix $\mathbf{1}$, see Example A for details). Notice that $\det L_1 = \det L_2 = \epsilon$.

Suppose that (3.8) has a continuous solution in some domain of (E_c, ϵ). Then it is algebraic in ϵ and linear in E_c (the latter is because the singularity lines within one avalanche type domain shift with velocity 1 in the direction of either $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ or $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Differentiating this by E_c we obtain:

$$L_{\rho_1}(z_0') = z_1', \dots, L_{\rho_m}(z_0') = z_m', \tag{3.9}$$

where $L_{\rho_i} = L_{\rho_{i,t_i}} \circ \dots \circ L_{\rho_{i,1}}$ for $\rho_i = (\rho_{i,1}, \dots, \rho_{i,t_i})$ is the linear part of F_{σ_i} (ρ_i is different from σ_i because dF_{σ_i} is a composition of several maps L_s). The points z_k, $1 \le k \le m$, are constrained to the singularity lines and we can suppose these are the lines $\{x_1 = E_c\}$ or $\{x_2 = E_c\}$ (all other singularities are mapped to them within one avalanche). Thus $z_k' = v_k + \psi_k w_k$, where $v_k = e_1$ or e_2 and $w_i = Jv_i$ is the other, ψ_k being an unknown scalar.

Let $\xi_k = L_{\rho_k}^{-1}(v_k)$, $\zeta_k = L_{\rho_k}^{-1}(w_k)$; these vectors depend meromorphically on ϵ. System (3.9) is solvable iff the affine lines $\xi_k + \psi_k \zeta_k$, $1 \le k \le m$, in $\mathbb{R}^2$ have a common point. We can suppose $m = 3$.

Denote by $\Omega(\xi,\eta) = \langle J\xi, \eta\rangle$ the standard symplectic form on $\mathbb{R}^2$. The above 3 lines intersect jointly iff

$$\Omega(\xi_1,\zeta_1)\Omega(\zeta_2,\zeta_3) + \Omega(\xi_2,\zeta_2)\Omega(\zeta_3,\zeta_1) + \Omega(\xi_3,\zeta_3)\Omega(\zeta_1,\zeta_2) = 0.$$

Dividing by $\prod_{k=1}^{3}\Omega(\xi_k,\zeta_k)$ and using the fact that $\Omega(\xi_k,\zeta_k) = \det(L_{\rho_k}^{-1}) = \epsilon^{-|\rho_k|}$ we get the equivalent equation:

$$\Omega(\eta_1,\eta_2) + \Omega(\eta_2,\eta_3) + \Omega(\eta_3,\eta_1) = 0. \tag{3.10}$$

Here $\eta_k = \epsilon^{-|\rho_k|}L_{\rho_k}^{-1}w_k = \tilde{L}_{\rho_k}w_k$, where $\tilde{L}_\tau = \tilde{L}_{\tau_1}\circ\cdots\circ\tilde{L}_{\tau_t}$ for $\tau = (\tau_1,\ldots,\tau_t)$ and $\tilde{L}_k = \epsilon^{-1}L_k^{-1} = \mathbf{1} + (\epsilon-1)\tilde{A}_k$ is the adjunct matrix for L_k, which gives

$$\tilde{A}_1 = \begin{bmatrix} 0 & 0 \\ 1/2 & 1 \end{bmatrix}, \quad \tilde{A}_2 = \begin{bmatrix} 1 & 1/2 \\ 0 & 0 \end{bmatrix}.$$

Now equation (3.10) holds iff there exist 3 not simultaneously zero numbers β_1,β_2,β_3 such that $\beta_1+\beta_2+\beta_3 = 0$ and $\beta_1\eta_1+\beta_2\eta_2+\beta_3\eta_3 = 0$. Since η_k is a polynomial matrix of degree $|\rho_k|$ and the products of A_t are always proportional to e_1 or e_2 (depending on the left-most factor) this last equation is never satisfied if the multi-indices ρ_k are different. □

Remark 3.8 *To support the usage of nearly-Zhang models note that the whole paradigm of SOC should allow generic perturbations of the data, for if there is a fine tuning of parameters, the model is unappropriate for physical explanation (of course, we should pass to the thermodynamic limit, but in practice this only means some large finite parameters).*

Our computer experiments did not expose any exponential growth of multiplicity in the Zhang model (though we see growth in complications of singularities), so we suggest that $H_{\mathrm{mult}}(F) = 0$ and hence $h_{\mathrm{top}}(F) = \log N$ always. In addition, by the above discussion we can disregard these exceptional values of (E_c,ϵ) even if there are any. This finishes discussion of topological entropy.

3.6 Geometry of the attractor

The construction of the spacial attractor $\mathcal{Y}$ can be interpreted as an iterated function system (IFS), where the maps F_i are not affine as usually considered, but piecewise affine. Hence one might expect that attractors $\mathcal{Y}$ are fractal, but with various size characteristics, like dimension and measure, depending on parameters E_c and ϵ.

3.6.1 Fractal structure

Computer experiments show that in certain cases the spacial attractor $\mathcal{Y}$ has fractal structure, see e.g. figures 3.2 and 3.3. We clearly see that $\mathcal{Y}$ consists of self-similar pieces. However the pieces overlap, making evaluation of the fractal dimension difficult. So we can provide only estimates of the attractor's size.

Nevertheless we observe from our experiments that Hausdorff dimension $\dim_{\mathrm{H}}(\mathcal{Y})$ and the Lebesgue measure $\mu_{\mathrm{Leb}}(\mathcal{Y})$ of the attractor grow piece-wise monotonically with E_c and ϵ. Thus the following effects occur in steps:

- The dimension and the measure of $\mathcal{Y}$ vanish.
- $\dim_{\mathrm{H}}(\mathcal{Y})$ is positive, while the measure is zero.
- Both $\dim_{\mathrm{H}}(\mathcal{Y})$ and $\mu_{\mathrm{Leb}}(\mathcal{Y})$ are positive.
- The attractor $\mathcal{Y}$ contains an interior point.

We will demonstrate the dimensional part in the next section, while we disregard the observation about the measure. The reason for this is that μ_{Leb} is not physically motivated and we should look for an SRB-measure.

The experiments show that such a measure exists and has support lying strictly inside $\mathcal{Y}$. (See figure 3.4 and figure 3.5. In figure 3.4 the attractor $\mathcal{Y}$ is shown for the case $N = 2$, $E_c = 20$ and $\epsilon = 2/3$, and figure 3.5 shows the orbit of a random initial condition. The latter corresponds to the support of the SRB-measure.) This is possible because the contraction rate of $f_{\mathbf{t}}$ is smaller for the exceptional sequences $\mathbf{t} \in \Sigma_N^+$, than for a generic one. Thus study of the IFS-attractor does not lead to conclusions about ergodicity or uniqueness of the SRB-measure. Still it provides an information about spacial distribution of the orbits in the Zhang dynamics.

3.6.2 Dimensional study of the attractor

The fractal properties of $\mathcal{Y}$ do not hold for all values of parameters (E_c, ϵ). An example where $\mathcal{Y}$ has integer dimension is shown in figure 3.4.

It was noted in [6] that Hausdorff (fractal) dimension of the attractor is about to increase as E_c grows. The arguments were the following: For bigger E_c the contraction rate decreases, so the theory of iterated function system (IFS) implies increasing of the Hausdorff dimension

$$\mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) = \dim_{\mathrm{H}}(\mathcal{Y})$$

as a function of E_c. While this seems to be true, the statement does not hold in precise sense. For instance, for $N = 2$, $E_c \in [\frac{1+\epsilon}{1-\epsilon}, \frac{2}{1-\epsilon}]$ the attractor is the set of 3

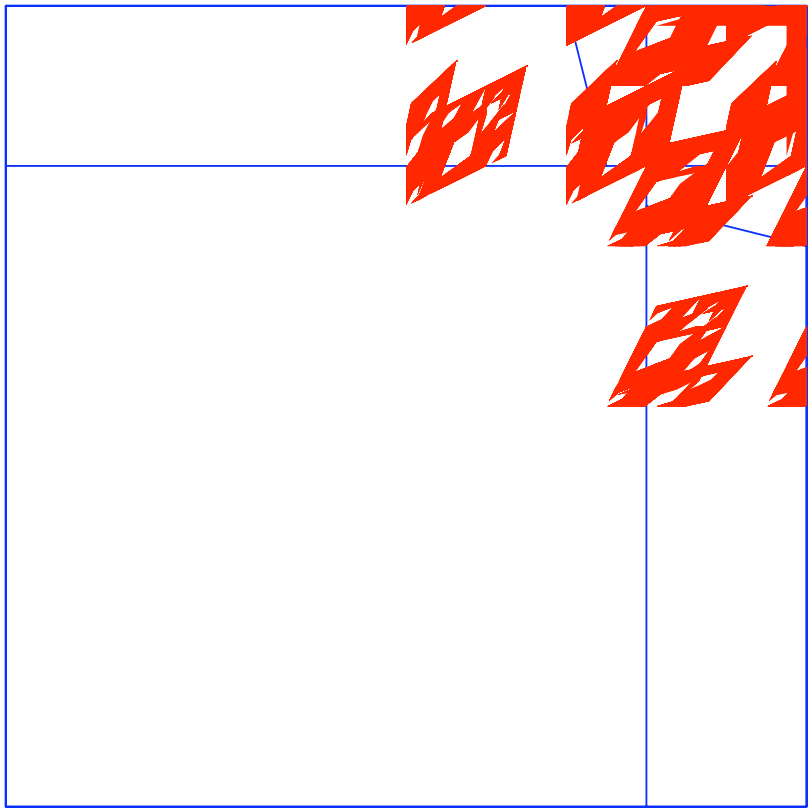


Figure 3.2: Shows the set U_{10} for $N = 2$, $E_c = 5$ and $\epsilon = 1/2$.

points, while it seems to have non-zero dimension for other parameters (computer simulations clearly show this).

The problem is with the framework of IFS, where usually only conformal maps are considered and certain regularity of their mapping graph and overlaps is assumed. However we will show validity of the claim in the asymptotic sense:

Theorem 3.20 *With fixed d, L and generic ϵ we have:* $\lim_{E_c \to \infty} \mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) = N$. *Moreover,* $\mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) = N$ *for big values* $E_c \gg 1/\epsilon$.

On the other hand for all $\epsilon \in [0, 1)$ it holds: $\lim_{E_c \to 0} \mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) = 0$.

In the above statement "generic" means both full Lebesgue measure and second Baire category. In fact, the equality holds for all ϵ outside a countable set. It seems though that the limit statement is valid for all ϵ. Thus we see that $\mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c)$ is not strictly monotone in E_c for its large values as one might expect from the arguments cited before the theorem.

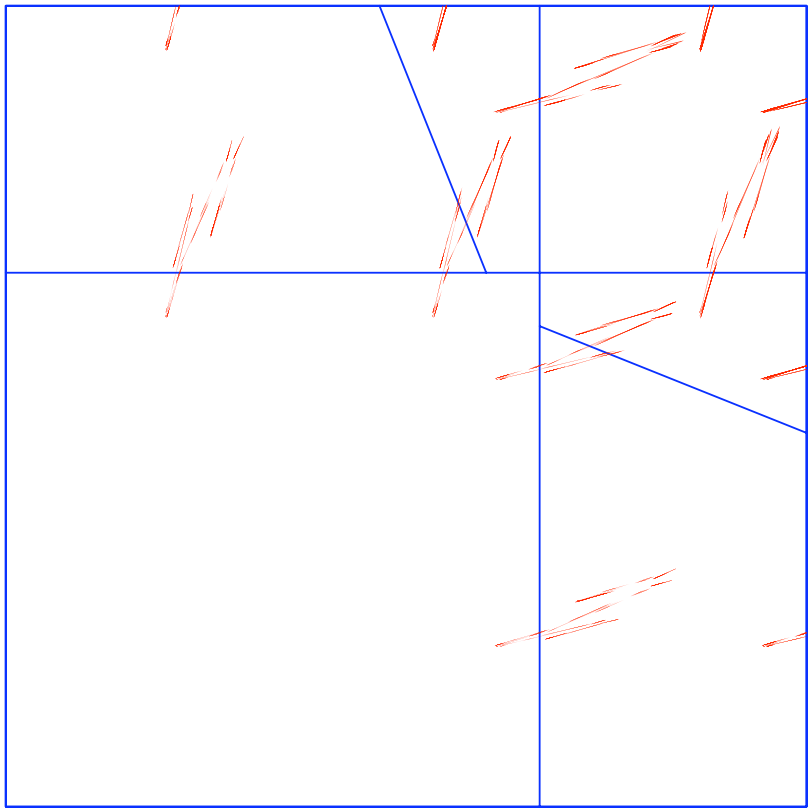


Figure 3.3: Shows the set U_{10} for $N = 2$, $E_c = 3$ and $\epsilon = 1/5$.

Proof. Consider first the statement about big energies $E_c \gg 1/\epsilon$. Let us start by demonstrating the idea of the proof on the example $N = 2$, see Figure 3.16.

The image of the vertical continuity domain M_{13} of height 1 adjusted to the right-top corner is a trapezium with the slope depending on ϵ (see (3.11) for the numeration of domains). It is thin – of constant length of horizontal section equal ϵ near its bottom side, but long – with diameter approximately equal $E_c \frac{1-\epsilon^2}{4}$. Thus if we shift $\sim C/\epsilon$ times this domain up and then all of the shifted images horizontally to the right, so that its first coordinate satisfies the inequality $E_c - 1 < x_1 \leq E_c$, then these shifts cover an open domain, including a unit square, in the vertical strip $K_1 = \{x_1 \in (E_c - 1, E_c]\}$ (note that we can leave a copy of the domain since this corresponds to the shift – dropping of energy $(x_1, x_2) \mapsto (x_1, x_2 + 1)$ on $M_{13} \subset K_1$ before the avalanche). An easy calculation shows that such shifts cover the whole upper part of K_1, strictly including the continuity domain $M_{13} \subset K_1$ adjusted to (E_c, E_c) and covering a vertical part of M_{12}. A similar scenario happen to the second coordinate. Thus in iterating the dynamics we will always have two continuity domains M_{13} and M_{23} adjusted to (E_c, E_c) and the adjacent parts of

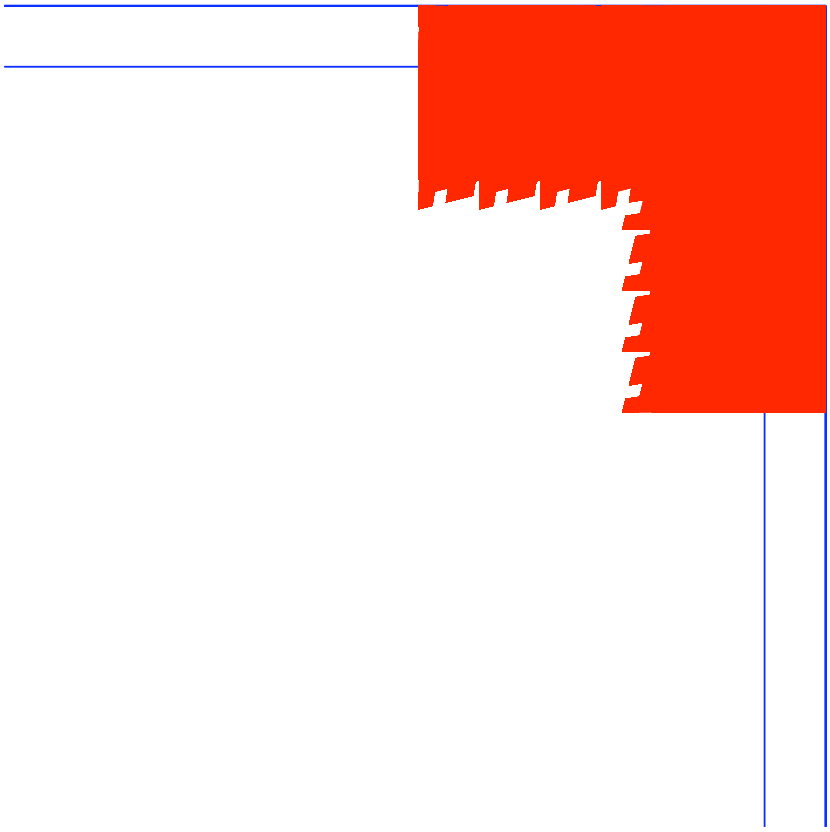


Figure 3.4: Shows the set $\mathcal{Y}$ in the top right corner of M for $N = 2$, $E_c = 20$ and $\epsilon = 2/3$. The dimension of $\mathcal{Y}$ is 2 in this example.

M_{12}, M_{22} lying in the attractor.

For general L and d we observe the same picture: With generic ϵ a continuity domain M_{ij} adjusted to the upper-most corner is mapped under avalanche-map F_{ij} to the trapezoid-like polyhedron with irrational slopes. Its shifts cover then an open domain in each of the strips $K_i = \{x_i \in (E_c - 1, E_c]\}$ and so after more shifts – the upper part of this strip, whence the statement.

We illustrate this process on Figure 3.17. The 3 domains adjusted to the corner (E_c, E_c, E_c) are mapped into interior of the cube and they have different irrational slopes (we picture them of zero thickness that corresponds to large values of $E_c \gg 1$), so that their shifts cover a big open domain near the faces adjusted to the above corner.

Consider now the second statement, $E_c \ll 1$. To estimate the fractal dimension from above we use the generalization of the Moran's formula from appendix B. It implies that if the IFS $f_1, \ldots, f_N$ satisfies $\|f_i\| \leq \delta$, then the Hausdorff dimension of the attractor admits the following estimate: $\mathfrak{D}_{\mathcal{Y}} \leq \log N\vartheta / \log \frac{1}{\delta}$, where ϑ is the maximal multiplicity of the continuity partitions for f_i.

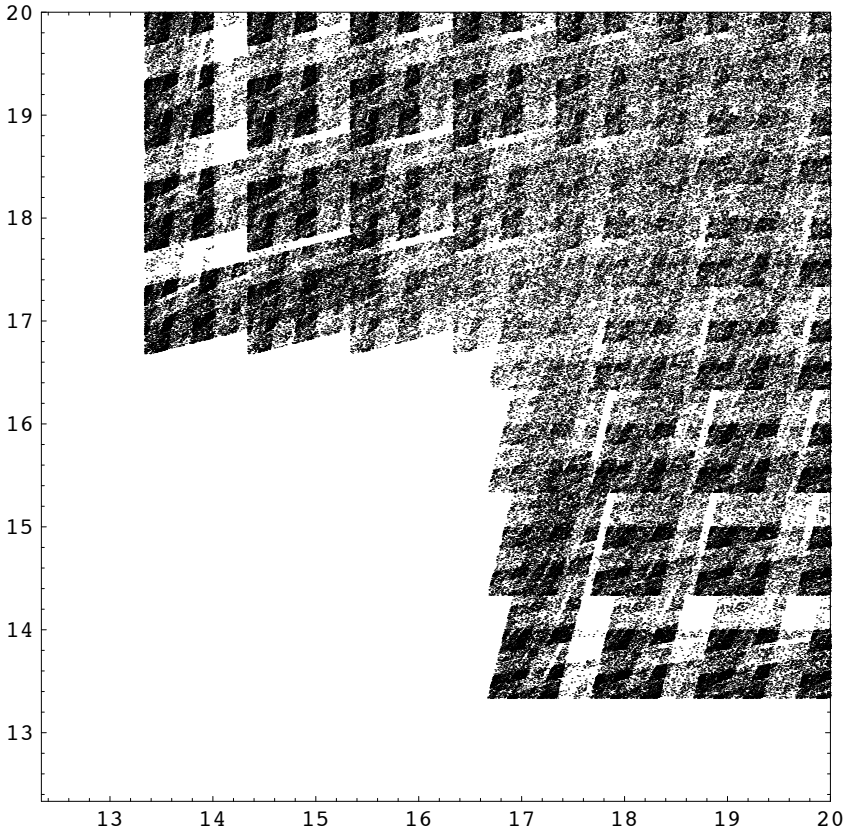


Figure 3.5: Shows an orbit of a randomly chosen point (i.e. the support of the SRB-measure) for $N = 2$, $E_c = 13$ and $\epsilon = 2/3$.

Now we claim that as $E_c \to 0$ we have: $\delta = \max \|f_i\| \leq (E_c)^\sigma$ for some $\sigma > 0$. To see this let us estimate the maximal duration of the avalanche $\bar{\tau}_m = \max_{(i,x)\in\Lambda\times M} \tau(i,x)$. This quantity tends to ∞ as $E_c \to 0$, but not as fast as $\tau_m(E_c, \epsilon, \Lambda) \sim C_1/E_c$ (see §3.2.2). Namely we state that $\bar{\tau}_m \sim C_2 \log 1/E_c$. Actually, if we drop energy 1 to an arbitrary site from a configuration in M, then in a finite E_c-independent time all the sites become overcritical. They remain overcritical, while the system does not loose a substantial amount of energy. During this process the total energy is dissipating in geometric progression with an average contraction rate $1 - \frac{1-\epsilon}{N} < 1$. So the duration of this stage has asymptotic $C_3 \log(1/E_c)$. The remaining time to finish the avalanche has a smaller asymptotic.

By §3.2.2 $n(E_c, \epsilon, \Lambda) < C_0$ for small E_c. Thus the proof of Theorem 3.5 implies that for certain E_c-independent constant $C_n = kC_0$ for each sequence of C_n steps in the avalanche process the product of the corresponding S-matrices will have norm $\leq c(\epsilon) < 1$, which is a uniform estimate in $E_c \ll 1$. The number of steps in one avalanche grows as $\log 1/E_c$. Therefore $\delta \leq C_4 c(\epsilon)^{C_n^{-1} \log 1/E_c} \leq \exp(-\sigma \log 1/E_c)$,

where $\sigma = \frac{1}{2}C_n^{-1}\log\frac{1}{c(\epsilon)}$ as was claimed in the estimate.

Next we claim that $\vartheta \le \varphi(E_c)$ with $\varphi = o((1/E_c)^{\upsilon})$ $\forall \upsilon > 0$. In fact, we described above the avalanche process for small energies. The first stage is finite and contributes only a bounded number of singularity hyperplanes. In its seconds stage all of the sites are excited, so there the corresponding number of singularity hyperplanes equals the duration. The last stage is shorter of time $\psi = o(\log 1/E_c)$, but the number of singularity hyperplanes grows faster, but still is bounded by $e^{\psi \log N!} = o((1/E_c)^{\upsilon})$ (see §3.2.3) for any $\upsilon > 0$.

Finally $\mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) \le \log N\vartheta/\log\frac{1}{\delta} \le \frac{C'+\log\varphi}{\sigma\log 1/E_c} \to 0$ as $E_c \to 0$. □

Corollary 3.3 *For big $E_c \gg 1/\epsilon$ the system $(\mathcal{A}, F)$ is not topologically transitive.*

Proof. Suppose $(\mathbf{t}, x)$ is a point with the dense orbit in $\mathcal{A}$. We know that for big E_c the Lebesgue measure of the spatial part $\mathcal{Y}$ of the attractor $\mathcal{A}$ is a positive number $\omega > 0$. It follows from Theorem 3.5 that under iterations with fixed excitation sequence $\mathbf{t}$ the volume of the spatial part M decreases in geometric progression with the number of avalanches. Thus after a finite number of steps it becomes less than ω. This iteration will be still a finite number of polyhedra, so that its closure does not coincide with $\mathcal{Y}$. Since it contains all the points $\pi_s(F^n(\mathbf{t}, x))$, we obtain contradiction. □

In the case $N = 2$ and $\epsilon > 1/2$, the value of E_c starting from which $\mathfrak{D}_{\mathcal{Y}}(\epsilon, E_c) = 2$ can be calculated precisely because even one shift of the sloped strip mentioned in the above proof overlaps with itself and is sufficient for obtaining an open domain in the attractor. This condition $\epsilon > 1/2$ together with $E_c \gg 1$ from the theorem ideologically coincide with the sufficient conditions for invertibility of the differentials of avalanche maps (Proposition 3.7). This makes an indication of a relation between this invertibility and fractality of the attractor in the spirit of Ledrappier-Young formula [38]. This latter is however unappropriate in our situation.

Note on the usage of the Ledrappier-Young formula. This formula, essentially used in [6] in the study of the Zhang model, cannot be used for the map $F : \mathcal{D} \to \mathcal{D}$ since this map is never invertible (in loc. cit. it was applied to F^{-1}). In addition to invertibility the Ledrappier-Young theorem is based on the SRB-property. For the map F^{-1} this property is equivalent to absolute continuity of the stable foliation for F w.r.t. the measure μ. If the measure has fractal support this cannot happen. Therefore all formulas based on this property may turn to be wrong. We demonstrate this in Example A of §3.7.

On the other hand, as we have just shown in the theorem, in thermodynamic limit $E_c \to \infty$ the fractality is lost and so the absolute continuity property is restored (but only for the geometric attractor, the support of SRB-measure is smaller!). This however does not help with non-invertibility of the factor (Σ_N^+, σ_N^+). Even

if we change this factor to invertible two-sided sequences (Σ_N, σ_N), the system remains non-invertible since not all points of the attractor (which can be quite fat) admit negative iterations (Remark 3.4). In addition, a new negative Lyapunov exponent $-\log N$ in the first factor appears and the formulas exploited in [6] become completely inadequate.

3.7 Examples

In the examples below we consider the one-dimensional Zhang model with two sites, $N = 2$.

Example A: A computation shows that for $E_c \geq (1+\epsilon)/(1-\epsilon)$ we have six domains of continuity $[i] \times M_{ij}$, $i = 1, 2$. The domains M_{1j} are given by

$$\begin{aligned} M_{11} &= \{x \in [0, E_c]^2 \,|\, x_1 + 1 \leq E_c\} \\ M_{12} &= \{x \in [0, E_c]^2 \,|\, x_1 + 1 > E_c \text{ and } (1-\epsilon)(x_1+1)/2 + x_2 \leq E_c\} \\ M_{13} &= \{x \in [0, E_c]^2 \,|\, x_1 + 1 > E_c \text{ and } (1-\epsilon)(x_1+1)/2 + x_2 > E_c\} \end{aligned} \tag{3.11}$$

and the domains M_{2j} are symmetric to these. The maps F_{ij} are of the form $F_{ij}(x) = L_{ij}(x + e_i)$, where

$$L_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad L_{12} = \begin{bmatrix} \epsilon & 0 \\ \frac{1-\epsilon}{2} & 1 \end{bmatrix} \quad L_{13} = \begin{bmatrix} (\frac{1+\epsilon}{2})^2 & \frac{1-\epsilon}{2} \\ \epsilon\frac{1-\epsilon}{2} & \epsilon \end{bmatrix}$$

and

$$L_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad L_{12} = \begin{bmatrix} 1 & \frac{1-\epsilon}{2} \\ 0 & \epsilon \end{bmatrix} \quad L_{13} = \begin{bmatrix} \epsilon & \epsilon\frac{1-\epsilon}{2} \\ \frac{1-\epsilon}{2} & (\frac{1+\epsilon}{2})^2 \end{bmatrix}$$

The maps F_{11} and F_{21} correspond to avalanches of size 0, the maps F_{12} and F_{22} correspond to avalanches of size 1, and the maps F_{13} and F_{23} correspond to avalanches of size 2.

It was discovered in [6] that the physical attractor $\mathcal{Y}$ has the following simple structure:

$$\mathcal{Y} = \left\{ (\frac{1+\epsilon}{1-\epsilon}, \frac{\epsilon}{2-\epsilon}), (\frac{\epsilon}{2-\epsilon}, \frac{1+\epsilon}{1-\epsilon}), (\frac{1+\epsilon}{1-\epsilon}, \frac{1+\epsilon}{1-\epsilon}) \right\}.$$

for $E_c \in [\frac{1+\epsilon}{1-\epsilon}, \frac{2}{1-\epsilon}]$ We denote these points by a, b, c so that $\mathcal{Y} = \{a, b, c\}$. The maps $F_1|_{\mathcal{Y}}$ and $F_2|_{\mathcal{Y}}$ are permutations of $\mathcal{Y}$:

$$F_1|_{\mathcal{Y}} = \begin{pmatrix} a & b & c \\ b & c & a \end{pmatrix}, \; F_2|_{\mathcal{Y}} = \begin{pmatrix} a & b & c \\ c & a & b \end{pmatrix}.$$

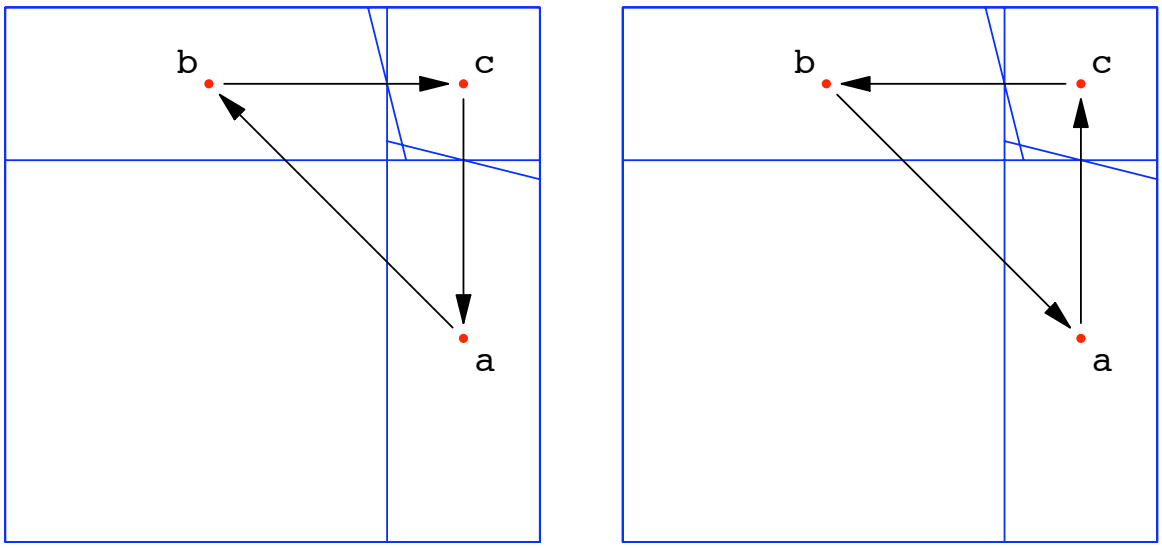


Figure 3.6: The figure shows the physical attractor $\mathcal{Y} = \{a, b, c\}$ and the maps $F_1|_{\mathcal{Y}}, F_2|_{\mathcal{Y}}$ for $E_c = 7/2$ and $\epsilon = 1/2$. The arrows on the picture to the left shows how the points of $\mathcal{Y}$ are mapped under F_1, and the picture on the right shows how the points are mapped under F_2.

Figure 3.6 shows the physical attractor $\mathcal{Y}$ and the maps $F_1|_Y$ and $F_2|_Y$ for $E_c = 7/2$ and $\epsilon = 1/2$. We construct a partition $\mathcal{R} = \{R_1, \ldots, R_6\}$ of $\Sigma_N^+ \times Y$ by

$$\begin{array}{lll} R_1 = [1] \times \{a\} & R_2 = [1] \times \{b\} & R_3 = [1] \times \{c\} \\ R_4 = [2] \times \{a\} & R_5 = [2] \times \{b\} & R_6 = [2] \times \{c\} \end{array} .$$

We let $A = \|a_{ij}\|$ be the 6×6 matrix where $a_{ij} = 1$ if $F(R_i) \cap R_j \neq \emptyset$ and $a_{ij} = 0$ otherwise. It is easy to verify that

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

It is clear that $g_{\mathcal{R}} : \mathcal{A} \to \Sigma_A^+$ is a topological conjugacy of the maps $F|_{\mathcal{A}}$ and σ_A^+. The matrix A is transitive and $\mathrm{Sp}(A) = \{-1, -1, 0, 0, 0, 2\}$. Hence $h_{\mathrm{top}}(F|_{\mathcal{A}}) = \log 2$.

If μ is the SRB-measure on $\mathcal{A}$, with $(\pi_u)_*\mu$ being the uniform Bernoulli measure on Σ_N^+, then $(g_{\mathcal{R}})_*\mu$ is the Perry measure on Σ_A^+. With respect to this measure it is

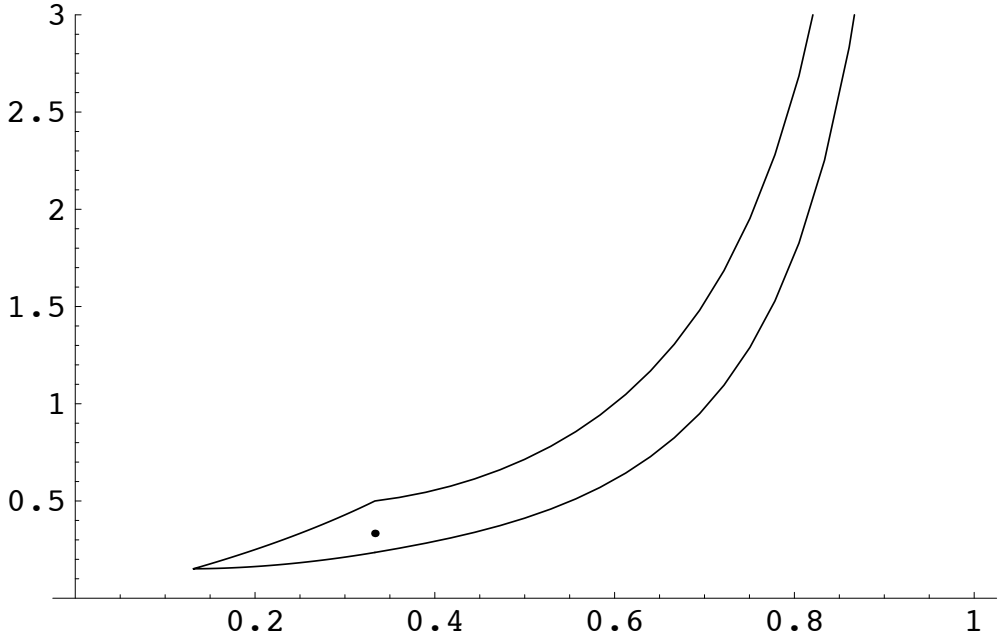


Figure 3.7: The region in the (ϵ, E_c)-plane where the avalanches are the same as for $E_c = \epsilon = 1/3$. The point $(1/3, 1/3)$ is shown in the interior of the region.

easy to see that the average avalanche size is $\overline{s}_0 = 1$. It then follows from that the sum of the negative Lyapunov-exponents is $\chi_1 + \chi_2 = \log \epsilon$. For instance we see that for $E_c = 11/2$ and $\epsilon = 2/3$ we have $h_\mu(F|_{\mathcal{A}}) > |\chi_1| + |\chi_2|$. So even though $F_1|_{\mathcal{Y}}$ and $F_2|_{\mathcal{Y}}$ are invertible, the Ruelle inequality (and therefore the Pesin formula) cannot be reversed.

We also remark that for $E_c = (1+\epsilon)/(1-\epsilon)$ we have $\mathcal{Y} \subset S(F)$, so the standard construction of SRB-measure will fail in this case. Still there is clearly a natural invariant measure.

The example with a trivial physical attractor can be generalized to all N for $d = 1$. Then $\mathcal{Y}$ consists of $N+1$ points $z_0, z_1, \ldots, z_N$ given by

$$z_n = \frac{1+\epsilon}{1-\epsilon}(1, 1, \ldots, 1) - e_n$$

where $e_0 = 0$ and $e_1, \ldots, e_N$ is the standard basis in $\mathbb{R}^N$.

Example B: For $E_c = \epsilon = 1/3$ the dynamics is very simple. The map F_1 has two domains of continuity: $M_{11} = \{(x, y) \in M \mid y \leq -2x/3 + 1/3\}$ and $M_{12} = M \setminus M_{11}$. The domains of continuity for F_2 are given by symmetry. The maps are given by

$F_{11}(x) = L_{11}(x + e_1)$ and $F_{12}(x) = L_{12}(x + e_1)$, where

$$L_{11} = \frac{1}{9}\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad L_{12} = \frac{2}{27}\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}.$$

We see that all four F_{ij} are mappings to the diagonal line in M. In fact, $F_{11}(M_{11}) = F_{21}(M_{21}) = [p_1, p_2]$, where $p_1 = (2/9, 2/9)$ and $p_2 = (1/3, 1/3)$. The interval is contained in $M_{12} \cap M_{22}$ since the two lines of singularity intersect the diagonal in the point $(1/5, 1/5)$. The images of M_{12} and M_{22} also coincide and is an interval $[p_1, p_3] \subset [p_1, p_2]$, where $p_3 = (22/81, 22/81)$. This shows that singularities are removable after one iteration. It is easy to see that for all $\mathbf{t} \in \Sigma_N^+$ the dynamics will contract to the fixed point $P = (4/17, 4/17)$. Hence the attractor of the system is $\mathcal{A} = \Sigma_N^+ \times \{P\}$. The dynamics on the attractor is of course conjugated to σ_N^+.

The example can easily be extended to a neighborhood of $(1/3, 1/3)$ in the (ϵ, E_c)-plane. In the region

$$\max\left\{\frac{1 - 2\epsilon + 13\epsilon^2}{5 + 12\epsilon - 13\epsilon^2}, \frac{\epsilon(7\epsilon^2 - 6\epsilon + 3)}{7\epsilon^3 - 10\epsilon^2 + 7\epsilon - 4}\right\} \leq E_c \leq \min\left\{\frac{\epsilon}{1 - \epsilon}, \frac{1 - 2\epsilon + 5\epsilon^2}{1 + 4\epsilon - 5\epsilon^2}\right\}$$

we have the same avalanches as for $E_c = \epsilon = 1/3$. This region is shown in figure 3.7. For all points in the region the maps depend continuously on ϵ and the lines of singularity depend continuously on ϵ and E_c. The condition for an atomic spacial attractor is that the images of the domains do not intersect the singularities. For $E_c = \epsilon = 1/3$ the images are bounded away from the lines of singularity, and hence there exists an open neighborhood of $(1/3, 1/3)$ in the (ϵ, E_c)-plane where the same holds, i.e. the attractor is of the form $\Sigma_N^+ \times \{P\}$ (where P is a point in M) and the dynamics is conjugated to σ_N^+.

Example C: Let us consider $E_c = 1/3$ and $\epsilon = 1/2$. In this case there are 28 domains of continuity, and 28 corresponding maps. A computer program is written to compute the sets U_n. The program uses exact calculations of the edges of the polygons that make up U_n, and hence it can be used to give rigorous "proof by computer" of removability of singularities. By using the program we obtain that singularities are removable. In fact $U_5 \cap \mathcal{S} = \emptyset$. The set U_5 consists of 13 connected components. Figure 3.8 shows the set U_5 and the lines of singularity, and figure 3.9 is a schematic illustration of how these connected components are situated with respect to the lines of singularity. The intersection of the connected components of U_5 with $\mathcal{Y}$ are denoted by $Y_1, \ldots, Y_{13}$. Then we construct the partition $\mathcal{R} = \{[i] \times Y_j\}$, and enumerate the elements so that $\mathcal{R} = \{R_1, \ldots, R_{26}\}$, where

$$\begin{array}{llll} R_1 = [1] \times Y_1 & R_3 = [1] \times Y_2 & \ldots & R_{25} = [1] \times Y_{13} \\ R_2 = [2] \times Y_1 & R_4 = [2] \times Y_2 & \ldots & R_{26} = [2] \times Y_{13} \end{array}.$$

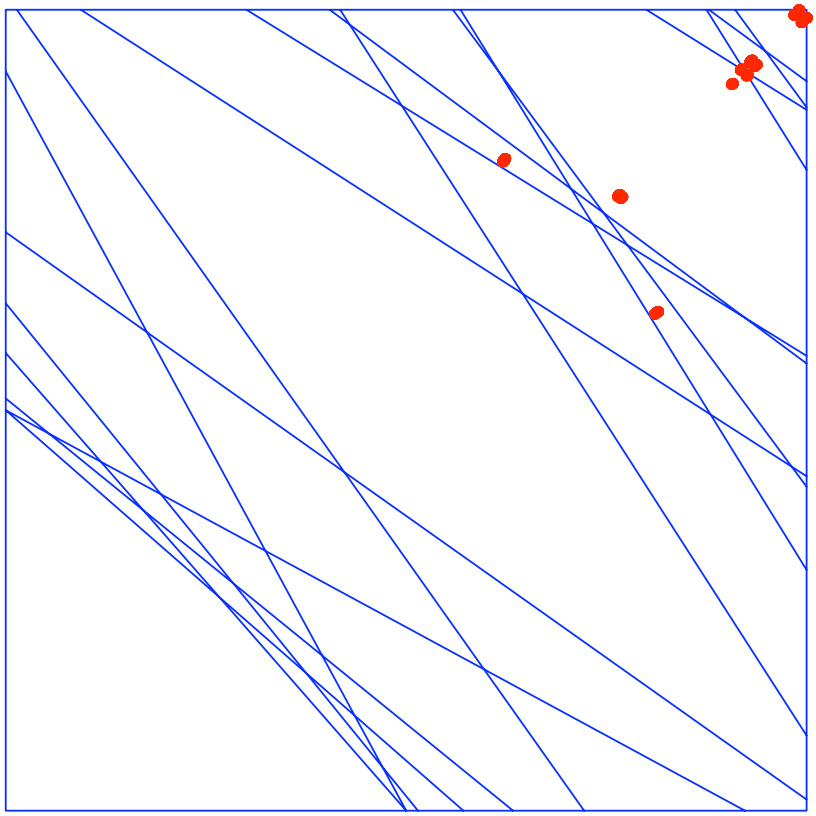


Figure 3.8: Shows the set U_5 for $N = 2$, $E_c = 1/3$ and $\epsilon = 1/2$. The points on the attractor are magnified in order to make them visible in the figure, and hence it looks as if they intersect singularities, but in fact they do not.

We construct the 26×26 matrix $A = \|a_{ij}\|$ by letting $a_{ij} = 1$ if $F(R_i) \cap R_j \neq \emptyset$, and $a_{ij} = 0$ otherwise. After making the computations, the matrix A becomes as shown in Figure 3.10. The black squares represent ones and white squares represent zeros. Direct computation shows that the matrix A is transitive.

Since singularities are removable the map $g_{\mathcal{R}} : \mathcal{A} \to \Sigma_A^+$ is an avalanche conjugacy between $F_{\mathcal{A}}$ and σ_A^+. The SRB-measure projects to the Perry measure on Σ_A^+, so it is possible to calculate properties such as average avalanche size. In this example a computation gives $\overline{s}_0 = 123/17$. The spectral radius of the matrix A is 2, and hence $h_{\text{top}}(\sigma_A^+) = \log 2$.

Example D: In the previous examples singularities are removable. This is however not the always the case. Figure 3.11 shows the set U_{20} for $N = 2$, $E_c = 7$ and $\epsilon = 1/2$. This is an example where singularities are non-removable. We will in the following show that singularities are non-removable for $E_c = (3+\epsilon)/(1-\epsilon)$.

Since $(3+\epsilon)/(1-\epsilon) > (1+\epsilon)/(1-\epsilon)$, the domains of continuity are given by

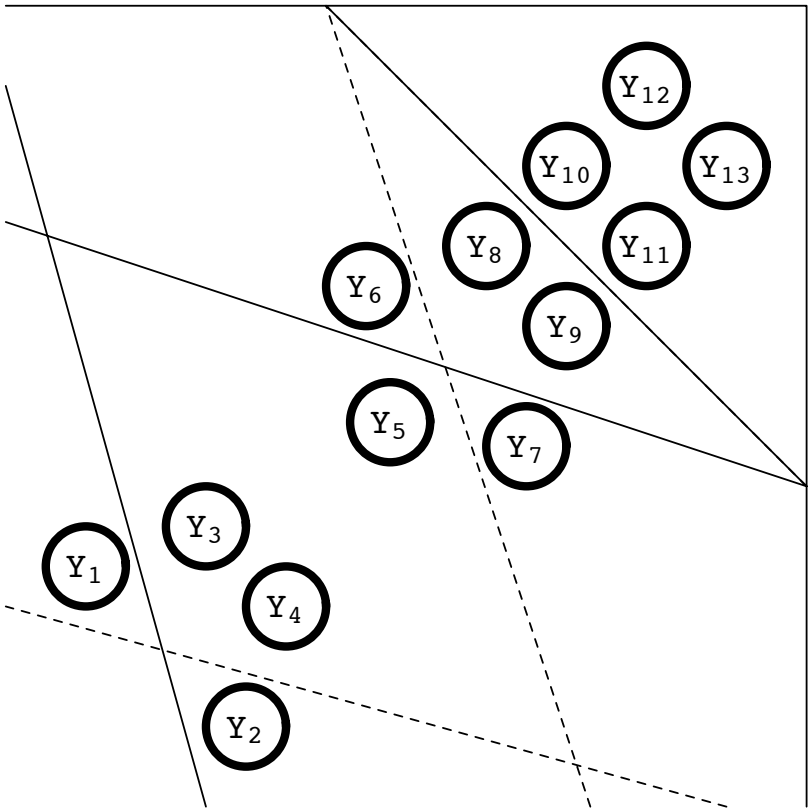


Figure 3.9: Shows how the 13 spatial partition elements are situated with respect to the lines of singularity.

the formulas presented in Example A. Take the point $p = (E_c, E_c - 1) \in M_{13}$. See Figure 3.12. Observe that p lies on the horizontal line $x_1 = E_c - 1$, and hence $p \in S(F)$. Clearly p is in the interior of M_{13} so

$$F_1(p) = F_{13}\Big(\frac{3+\epsilon}{1-\epsilon}, \frac{3+\epsilon}{1-\epsilon} - 1\Big) = \Big(\frac{3+\epsilon}{1-\epsilon} - 1, \frac{3+\epsilon}{1-\epsilon} - 2\Big) = p - (1,2)\,.$$

Denote $q_1 := F_1(p)$ and observe the it lies on the singularity line $x_2 = E_c - 1$. On the other hand q_1 is in the interior of M_{21}, so $F_2(q_1) = p - (1,1)$. Denote $q_2 := F_2(q_1)$. This point also lies in the interior of M_{21}. Let $q_3 := F_2(q_2) = p - (1,0)$. It is clear that $F_{11}(q_3) = p$, and in this sense p is a periodic point. However, $F_1(q_3)$ is not well defined since $q_3 \in \partial M_{12} \cap \partial M_{13}$.

In the following we let $\langle z_1, \ldots, z_n \rangle$ denote the open convex polygon with edges $z_1, \ldots, z_n$. Define $B_\delta(p) = \langle p, p + (0,\delta), p + (-\delta,\delta), p + (-\delta, 0) \rangle$. For small $\delta > 0$ we have $B_\delta(p) \in M_{13}$, and hence

$$F_1(B_\delta(p)) = F_{13}(B_\delta(p)) = \langle q_1, q_1 + \delta a, q_1 + \delta b, q_1 + \delta c \rangle$$

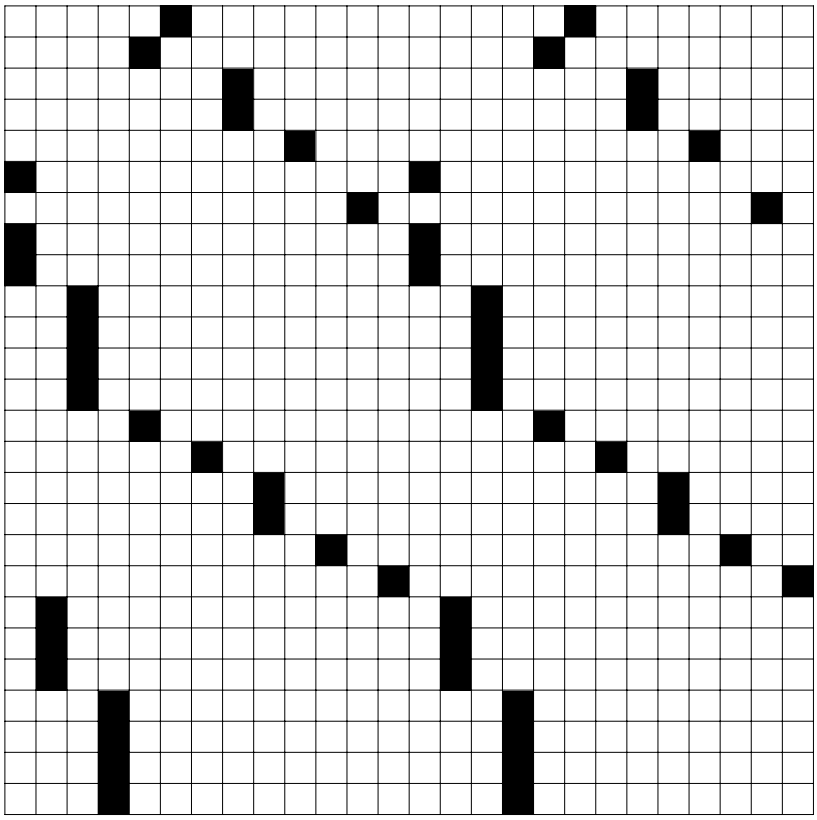


Figure 3.10: Shows the matrix A for the coding of Example C. Black squares are 1-s and white squares are 0-s.

where

$$a = \Big(\frac{1-\epsilon}{2}, \epsilon\Big)\,,\ b = \Big(\frac{1-4\epsilon-\epsilon^2}{4}, \frac{1+\epsilon}{2}\epsilon\Big)\,,\ c = \Big(-\big(\frac{1+\epsilon}{2}\big)^2, -\frac{1-\epsilon}{2}\epsilon\Big)\,.$$

The polygon $F_1(B_\delta(p))$ intersects the singularity line $x_1 = E_c - 1$. See Figure 3.12. A simple computation shows that

$$F_1(B_\delta(p)) \cap M_{11} = \langle q_1, q_1 + \delta a', q_1 + \delta b, q_1 + \delta c\rangle$$

where

$$a' = \Big(0, \big(\frac{2\epsilon}{1+\epsilon}\big)^2\Big)\,.$$

It then follows from the above discussion that

$$F^1 \circ F_2^2(F_1(B_\delta(p)) \cap M_{11}) = \langle p, p + \delta a', p + \delta b, p + \delta c\rangle\,.$$

It is then easy to verify that for all $\delta > 0$ there is $\gamma > 0$ such that $B_\gamma(p) \subset \pi_s \circ F^4(\Sigma_N^+ \times B_\delta(p))$. So for each $n \in \mathbb{N}$ there is $\delta_n > 0$ with $B_{\delta_n}(p) \subset U_n$. The

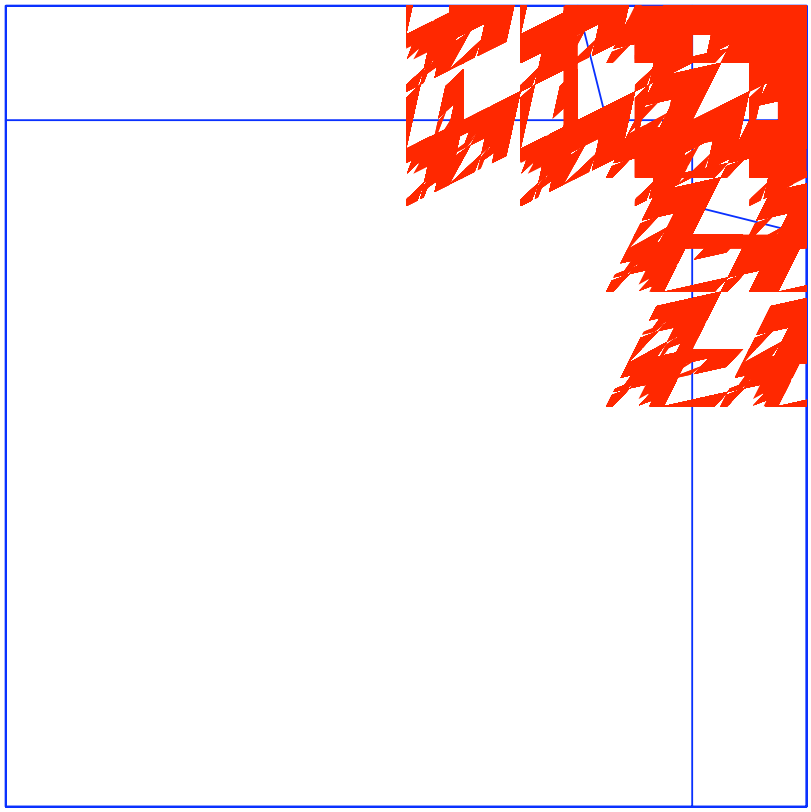


Figure 3.11: Shows the set U_{20} for $E_c = 7$ and $\epsilon = 1/2$. In this example singularities are non-removable.

image of $B_{\delta_n}(p)$ under F_1 intersects singularities, and its closure contains the point q_1, which thus is an essential singularity. Clearly the points p, q_2 and q_3 are also essential singularities.

3.8 Statistical properties

In order to evaluate the entropy and Lyapunov spectrum of the physical model in the thermodynamic limit we need to derive several estimates for the asymptotic behavior of observables like avalanche size, avalanche duration and "waiting-time" between avalanches. The results are derived using only the uniform Bernoulli measure on Σ_N^+, and hence hold for any SRB-measure and for time-averages.

In §3.8.3 we define the thermodynamic limit at the double limit $E_c \to \infty$, $L = \sqrt[d]{N} \to \infty$, contrary to [6] where the thermodynamic limit is defined as the limit $L \to \infty$ only. As $E_c \to \infty$ we must make a scaling of time in the physical model. Otherwise the influx of energy to the system will go to zero. With this new

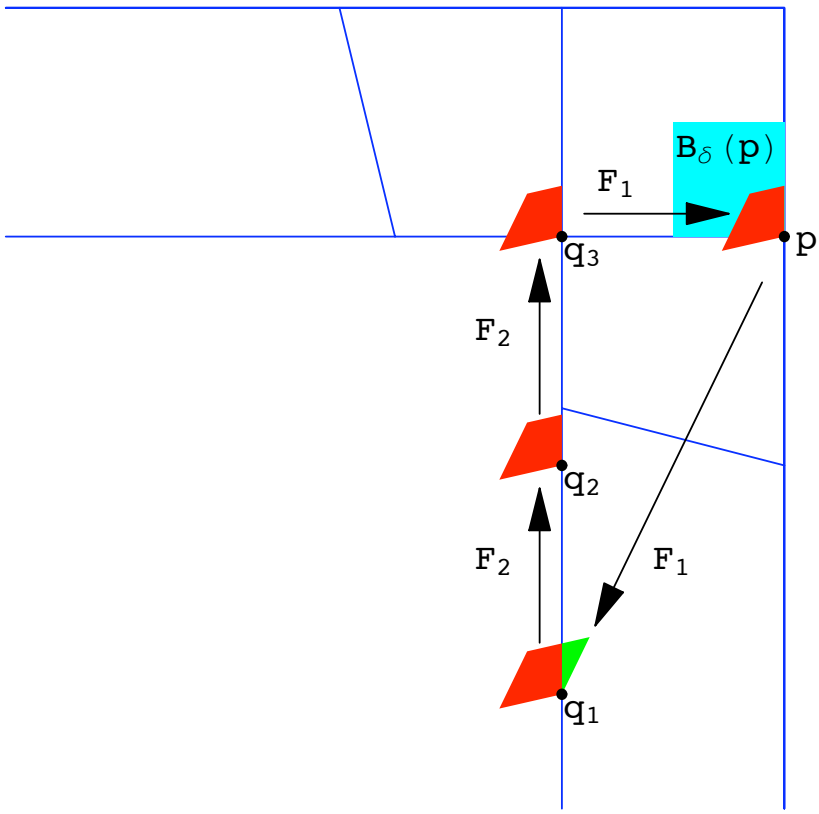


Figure 3.12: Illustration of the fact that singularities are non-removable for $E_c = 7$ and $\epsilon = 1/2$.

scaling we show that the entropy goes to zero and the Lyapunov spectrum collapses completely in the thermodynamic limit.

We do not provide strict mathematical proofs, but still think important to include the discussion of our results from the physical point of view.

3.8.1 Statistics of observables

Let τ be the coordinate measuring the duration of avalanche and let ω correspond to the interval between avalanches (minimal value 1). We will also study the observable s – the avalanche size (defined in §3.2.3). While in the first case we consider only actual avalanches, so that $\tau > 0$, in the second we make distinction between s_0 – all avalanches including the trivial case of under-critical state ($s = 0$) and s_+ – the actual avalanches, so that $s_+ > 0$.

The reason for introducing two different avalanche size observables is the following: s_0 plays a crucial role in mathematical investigation of the model (see §3.8.4), while s_+ is important from physical perspective. In §3.8.3 we will see that only

physical observables allow a thermodynamic limit.

Denote by $\bar{\tau}$, $\bar{\omega}$, $\bar{s}_0$, $\bar{s}_+$ the corresponding mean time-average quantities, each of which is a function on the space-factor M and is defined as follows:

$$\bar{\sigma}(x) = \int_{\Sigma_N^+} \lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} \sigma\big(\hat{f}^i(\mathbf{t}, x)\big)\, d\mu_{\mathrm{Ber}}$$

with μ_{Ber} being the Bernoulli measure on the one-sided shifts (by Birkhoff ergodic theorem the time-average limit exists almost everywhere and is measurable). This function is invariant in the sense: $\bar{\sigma}(x) = \frac{1}{N}\sum_{i=1}^{N} \bar{\sigma}(\hat{f}(i, x))$.

Whenever the system is ergodic with respect to an invariant measure μ the function $\bar{\sigma}$ is constant μ-a.e. and equals the space (ensemble) average

$$\langle\sigma\rangle_\mu = \int\limits_{\Sigma_N^+ \times M} \sigma\, d\mu.$$

If the system has a unique invariant SRB-measure the function $\bar{\sigma}$ is constant $\mu_{\mathrm{Ber}} \times \mu_{\mathrm{Leb}}$-a.e. and equals the space (ensemble) average

$$\langle\sigma\rangle = \int\limits_{\Sigma_N^+ \times M} \sigma\, d\mu_{\mathrm{SRB}}.$$

We will need the maximal values of these observables in a sequel, which we denote by $\tau_{\max}$, $\omega_{\max}$ and $s_{\max} = \max s_0 = \max s_+$ respectively. We shall calculate their asymptotics in L and E_c.

Denote by $\varphi_1 \sim \varphi_2$ the asymptotic equivalence relation meaning that the ratio φ_1/φ_2 has subexponential grow/decay. We denote the equivalence by $\approx$, when the limit of the ratio is 1.

Lemma 3.21 *For $E_c \gg 1$ the maximal avalanche time and size have the asymptotics: $\tau_{max} \sim L^{\gamma_\tau}$, $\omega_{max} = L^d E_c$, $s_{max} \sim L^{d+\gamma_s}$, where $\gamma_\tau = \gamma_s = 1$ for $d = 1$ and $1 < \gamma_\tau, \gamma_s < d$ for $d > 1$.*

Proof. Consider at first the simple case $d = 1$. The maximum avalanche duration and size are achieved when all sites contribute to the avalanche, i.e. their energies are sufficiently big and there is one site with energy greater than $E_c - 1$ to initiate the avalanche. Actually, if some site has small energy, it will serve as a boundary and the avalanche wave reflects from it (to be explained below).

Assuming $E_c > \frac{\epsilon}{1-\epsilon}$ we know from Lemma 3.6 that in the avalanche process each site, whenever overcritical, relaxes until in the next step it receives a sufficient portion of energy to become overcritical and relax etc. In other words, the sites

blink, being under- and over-critical in turn. But in this process they make overcritical their neighbors and the process propagates as a wave, with only difference that its front excites new sites, while in the traversed region there remain blinking overcritical sites.

This wave spreads along the interval $B_L^1 = [1, L] \subset \mathbb{Z}$ towards its boundary and then it reflects from it, bearing now relaxation. In fact, as the front wave reaches the boundary it losses a substantial part of the energy on the boundary sites, which thus cannot be recovered and remain undercritical for the rest of this avalanche. They influence their neighbors to stop being critical and so forth. Thus we obtain the reflected wave that, in contrast with the first one, turns overcritical sites to relaxed. When the wave hits itself, the avalanche process stops.

It is clear that the duration of this avalanche is $\tau_{\max} \approx L$. The number of involved sites corresponds to the area of the triangle with a side $B_L^1 = [1, L]$ and height $L/2$ (recall that each site is overcritical only half time of its blinking period), whence $s_{\max} \approx L^2/2$.

Consider now the case of dimension $d > 1$. Here the scenario of maximal avalanche is more complicated and consists of three stages (the proof is similar to the case $d = 1$, but quite lengthy and will be suppressed). Again the maximum avalanche duration and size are achieved when all sites contribute to the avalanche, though now if some isolated site has smaller energy it serves as a boundary only once but then on the next several waves it receives the required portion of energy and follows the general scheme of motion.

At the first stage a site is excited and it initiates the rhombus-shape wave (a cube in the Manhattan metric, see figure 3.13) that spreads to the boundary of the cube $B_L^d = [1, L]^d \subset \mathbb{Z}^d$ (in time $\approx L/2$ if the center of the rhombus is placed near the center of the cube). The second stage begins as the wave reaches the boundary face and reflects from it (See figure 3.14). The reflected wave is almost momentary overthrown by the coming overcritical wave, which again reflects from the boundary, come now deeper into the interior of the cube, but is overthrown too etc. If one looks along the boundary face, the reflected wave travel along towards a vertex with preserved form (like a soliton) and then disappears into this vertex (there occur strong interactions with other waves in this corner). But if one looks into the perpendicular direction, the collection of reflected waves oscillates (each reflected wave enters deeper and deeper into the cube) contributing to the avalanche duration the sum $1 + 2 + 3 + \ldots$, which stops with the end of the second stage (we do not specify the sum precisely because after some oscillations the wave front becomes more and more eroded by the interactions between overcritical and relaxing waves; this impairs the sum and decreases the exponent, but not too drastically).

The third stage begins as the main body of the overcritical sites becomes disconnected and the avalanches behaves like worms crawling along the high energy fractal-like collection of states. We illustrate this in figure 3.15. From the descrip-

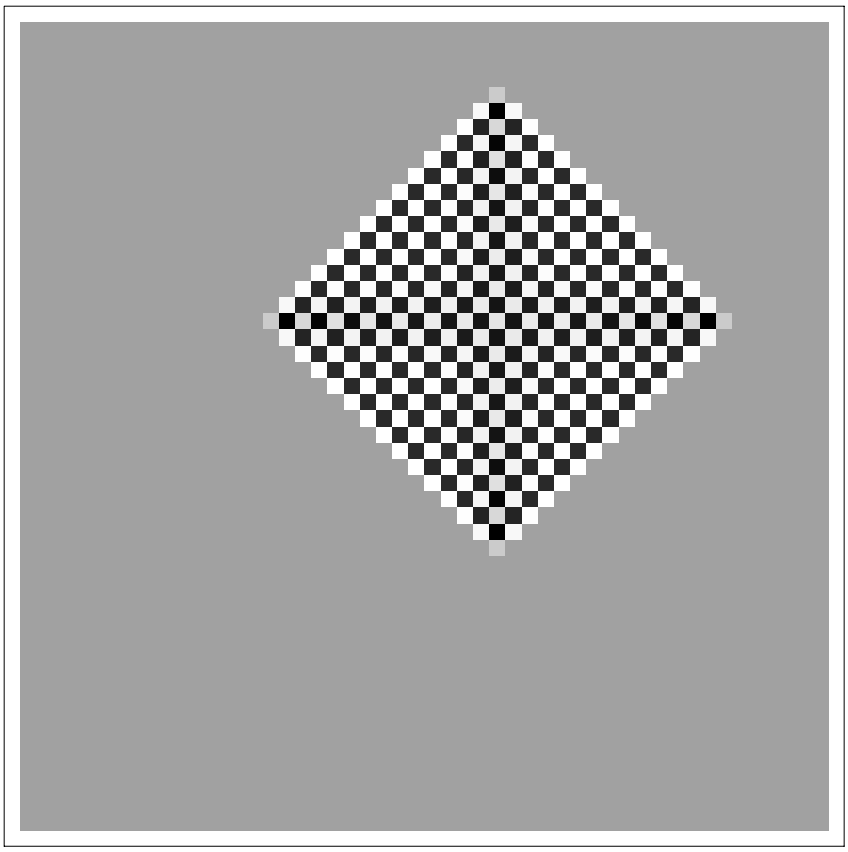


Figure 3.13: The picture is a "snapshot" of the lattice Λ for $d = 2$, $L = 10$, $E_c = 7$ and $\epsilon = 1/2$. A single site in the marginally stable configuration has been exited, and a great avalanche is unfolding in a rhombus shape. This is the first stage of this avalanche. The white squares are overcritical sites, the gray squares are sites with energy just bellow E_c and the black squares have energy approximately equal ϵE_c.

tion of the maximal avalanche process it is clear that the asymptotic exponents γ_τ, γ_s do not depend on the energy $E_c \gg 1$. Let us denote (the sequences are increasing, so the limits exist):

$$\gamma_\tau = \lim_{L\to\infty} \frac{\log \tau_{\max}}{\log L} \quad \text{and} \quad d + \gamma_s = \lim_{L\to\infty} \frac{\log s_{\max}}{\log L}.$$

Duration of the first stage of avalanche is $\sim L$. The above arguments show that the exponent γ'_τ of the second stage is > 1. The last stage can only increase it: $\gamma_\tau \geq \gamma'_\tau$. To see that $\gamma_\tau < d$ we note that in average the number of critical sites on the boundary is about $\kappa_L \sim L^{\gamma_\kappa}$ with $0 < \gamma_\kappa < d$. Thus we have a constant flow of energy out of the system with the average speed $> \frac{1-\epsilon}{2d}\kappa_L E_c$, and the inequality $E_{\text{tot}} \leq L^d E_c$ proves the claim.

To estimate the maximal avalanche size exponent γ_s consider again the second

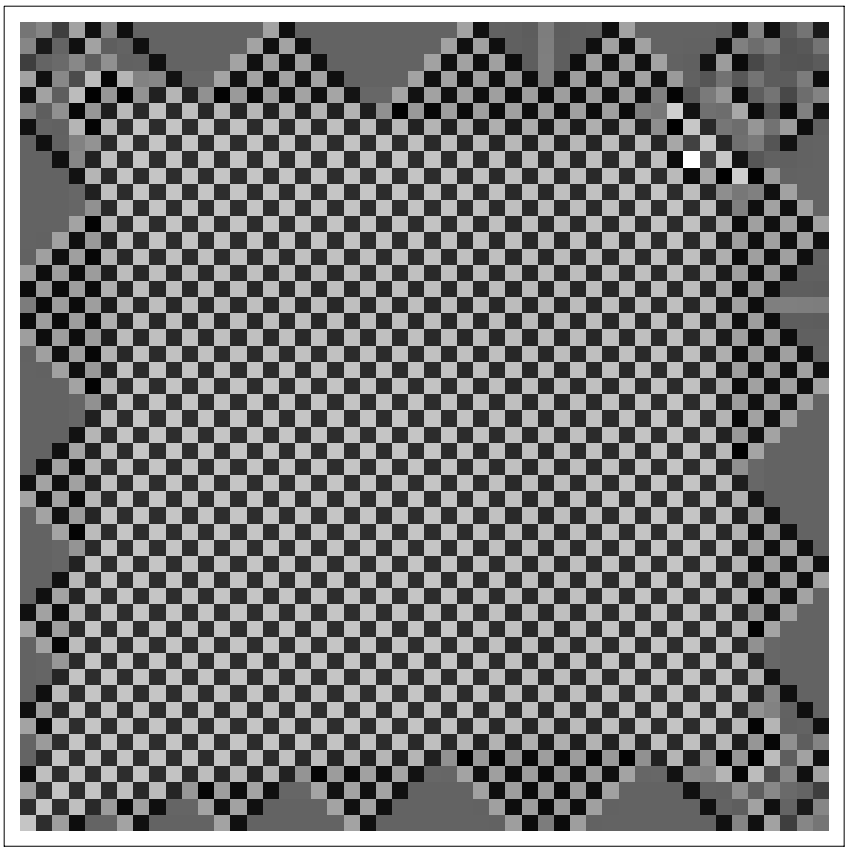


Figure 3.14: The picture shows a "snapshot" of the second stage of the avalanche shown in figure 3.13. We can see the well-shaped (soliton-like) waves of energy near the boundary and observe how their form begins being eroded near the vertices.

stage of the above scenario. The number of involved sites corresponds to the volume of the prism Π_L over the cube B_L^d with height $\ell \approx \frac{1}{2}\tau_{\max} \sim L^{\gamma'_\tau}$, i.e.

$$L^{d+\gamma'_s} \sim \mathrm{Vol}_{d+1}(\Pi_L) = \frac{\ell}{d+1} \cdot \mathrm{Vol}_d(B_L^d) \sim L^{d+\gamma'_\tau}.$$

Thus $\gamma_s \geq \gamma'_s = \gamma'_\tau > 1$. Inequality $\gamma_s < d$ follows from the inequality from above for the duration exponent γ_τ.

The maximal value of the waiting time is obvious. □

3.8.2 Asymptotic of the statistical data

Now we can study statistics of the avalanche data asymptotically (as for the thermodynamic limit).

We will need the following technical statement (informal only for $E_c > 1$):

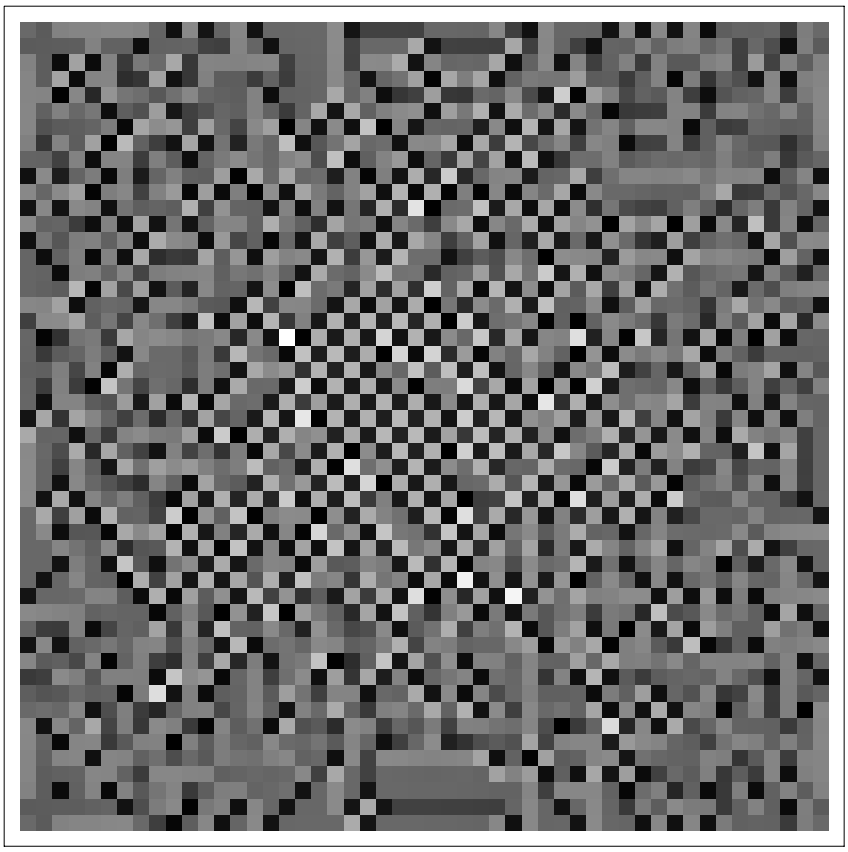


Figure 3.15: The picture shows a "snapshot" of the third stage of the avalanche shown in figure 3.13. The energy configuration has a fractal structure where the avalanche can sustain in regions of high energy.

Lemma 3.22 *Almost every (w.r.t. a random excitation sequence) spatial trajectory returns to the cube* $B_0 = (E_c - 1, E_c]^N$.

Proof. It follows from Proposition 3.4 that $K = \{x \in M \mid \exists i : x_i \in (E_c - 1; E_c]\}$ is a return set, i.e. every trajectory $F^n(\mathbf{t}, x)$ meets $\Sigma_N^+ \times K$. Partition the spatial part $\mathcal{Y}$ of the attractor according to the hyperplanes collection $\mathcal{H} = \cup_{i,m\in\mathbb{N}}\{x_i = E_c - m\}$.

Each such a part can be shifted by an excitation sequence to the cube B_0. The probability of all such sequences (where the avalanche starts from the set B_0 of maximal energy in all sites) is positive. Let $\rho > 0$ denote the minimum of these probabilities over the finite set of all partition elements of Y by $\mathcal{H}$. Then the probability of not entering B_0 in k successive avalanches is less than $(1-\rho)^k$.

Therefore since the number of avalanches tend to infinity as we iterate the dynamics, the measure of trajectories staying away from B_0 is zero. □

Theorem 3.23 *We have the following asymptotic estimates valid as* $E_c \to \infty$ *(and* $N \gg 1$ *fixed) or* $L = \sqrt[d]{N} \to \infty$ *(and* $E_c \gg 1$ *fixed):*

(i) $\bar{\tau} \sim L^{\gamma_\tau}$;

(ii) $\bar{\omega} \sim E_c$;

(iii) $\bar{s}_0 \sim L^{d+\gamma_s}/E_c$;

(iv) $\bar{s}_+ \sim L^{d+\gamma_s}$,

where γ_τ, γ_s are the same exponents as in Lemma 3.21 (thus $\gamma_\omega = \lim_{L\to\infty} \frac{\log \bar{\omega}}{\log L} = 0$).

Proof. The maximal avalanche size is achieved for a certain configuration of states $V \subset M$, which we can bound as follows: $U_1 \subset V \subset U_d$, where

$$U_j = \{x \in M \,|\, \forall i : \big(1 - j\tfrac{1-\epsilon}{2d}\big)E_c < x_i \le E_c \text{ and } \exists i_0 : x_{i_0} > E_c - 1\}.$$

Denote also $\tilde{U}_j = \{x \in M \,|\, \forall i : \big(1 - j\frac{1-\epsilon}{2d}\big)E_c < x_i \le E_c\} \supset U_j$.

To estimate the measure of the sites leading to the maximal avalanche we consider preimages of $\Sigma_N^+ \times \tilde{U}_j$ under the map F. It is clear that one needs $k_\epsilon \in [\frac{2}{1-\epsilon}, \frac{2d}{1-\epsilon}]$ different backwards iterations F^{-i_s}, $s = 1, \dots, k_\epsilon$, to cover the spatial attractor $\mathcal{Y}$. Since the measure μ is F-invariant, we get for its π_s-push-forward: $\mu_s(\tilde{U}_j) \approx \rho_j/k_\epsilon$, which is E_c-independent. Thus we get the same exponent $(d + \gamma_s)$ for $\bar{s}_+$ as for $s_{\max}$.

The same arguments yield the asymptotic of $\bar{\tau}$.

To obtain the asymptotic of $\bar{\omega}$ in L we note that since the amount of lost energy is $< CL^{\gamma_\kappa+\delta}E_c$ (where $\gamma_\kappa < d$ is the quantity from the proof of Lemma 3.21), the average remained energy in a site of configuration obtained from a maximal one after an avalanche is $(L^dE_c - CL^{\gamma_\kappa+\delta}E_c)/L^d \approx E_c$. Thus the waiting time does not grow with L.

The asymptotic of $\bar{\omega}$ in E_c is quite different: If N is fixed but E_c grows, then any state from ∂M becomes at distance $\theta_N \cdot E_c$ after some relatively small number of iterations.

Let us first demonstrate the idea in the simple case $N = 2$. For critical energy $E_c > \epsilon/(1-\epsilon)$ the picture of avalanches is shown on Figure 3.16. We see that in a few steps of the dynamics the configuration becomes far from ∂M, i.e. it strongly contracts in all directions. Actually, it is possible to imagine the situations when the point is mapped to the vertical strip and then is shifted horizontally for a long time by excitations of the first site, but probability of this event exponentially goes to zero as $E_c \to \infty$. Thus in a relatively short time the point from ∂M is mapped into the square $[0, \theta_2 E_c]^2$, where the constant θ_2 is E_c-independent. To achieve the boundary ∂M again it needs $\sim (1-\theta_2)2E_c$ random excitations.

The similar picture happens for $N = 3$, see Figure 3.17. In the general case Theorem 3.5 insures that after some (few) number of steps we get strong contraction, so that the point becomes in the cube $[0, \theta_N E_c]^N$, i.e. far from the boundary ∂M.

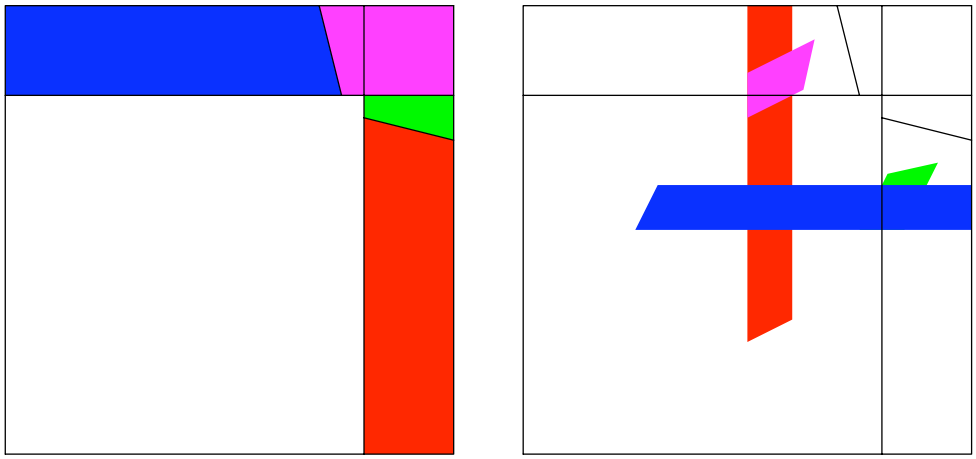


Figure 3.16: The figure illustrates how the continuity domains are mapped under F_1 and F_2 for $N = 2$ and $E_c > \epsilon/(1-\epsilon)$.

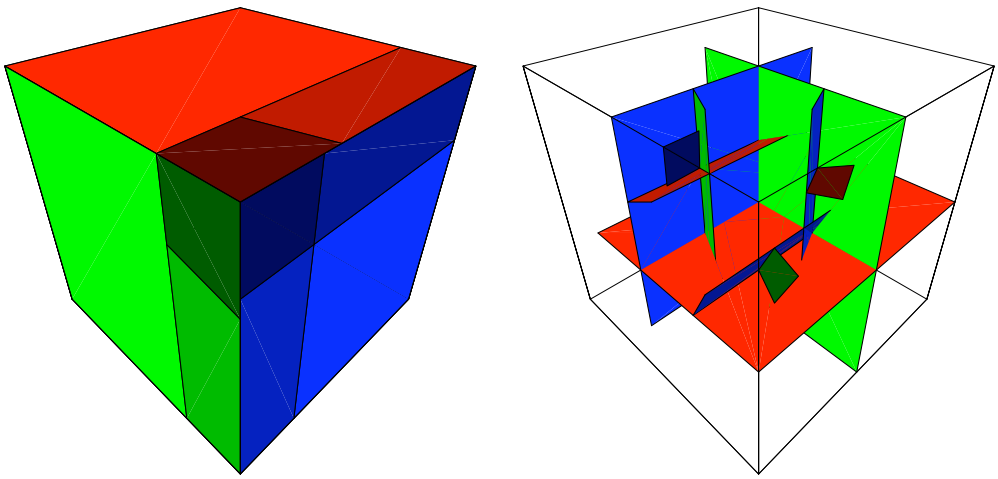


Figure 3.17: The figure illustrates how the faces of the continuity domains are mapped under F_1, F_2 and F_3 for $N = 3$ and $E_c > \epsilon/(1-\epsilon)$. The point of view is at (E_c, E_c, E_c).

Thus we need $\sim (1-\theta_N)NE_c$ excitations to make it overcritical and this implies the claim. Note that the asymptotic for $\bar{\omega}$ is not for the double limit $E_c \to \infty$, $L \to \infty$, but for two partial limits only.

To obtain the estimate for $\bar{s}_0$ we need to estimate the conditional measure $\mu_s(U_j|\tilde{U}_j)$, which coincides with the probability that a randomly chosen configuration $x \in \tilde{U}_j$ and site i satisfy: $x_i > E_c - 1$. This probability is $\approx b_1/E_c$. Thus $\mu(\Sigma_N^+ \times V) \sim \sigma_\epsilon/E_c$ and the mean avalanche size is $\langle s_0 \rangle \sim L^{d+\gamma_s}/E_c$.

This is however the space-average (the arguments below work also for $\bar{s}_+$, $\bar{\tau}$). We would obtain the same for the time-average of Lebesgue a.e.-initial condition if we have an SRB-measure, or for μ-a.e. if we have an ergodic measure μ. But we cannot guarantee existence of an SRB-measure or an ergodic measure. However, for a non-ergodic measure, we can decompose $\mathcal{A}$ into ergodic components, where the Birkhoff theorem works. By Lemma 3.22 each ergodic component (with non-trivial contribution) intersects in its spatial part the top-energy cube B_0 and so for each of it the same asymptotic of space-averages holds with universal exponents but maybe different coefficients. Therefore we obtain the required asymptotic for $\bar{s}_0$.

Another way to get the last asymptotic is via the formula $\bar{s}_0 = \frac{\bar{\omega}\cdot 0 + 1\cdot \bar{s}_+}{\bar{\omega}+1}$. $\square$

For $d=1$, $E_c \in [\frac{1+\epsilon}{1-\epsilon}, \frac{2}{1-\epsilon}]$ we already have $\bar{\tau} \sim N = L^1$ as the theorem states, but $\bar{s}_0 \sim N = L^1$, $N \to \infty$, which shows in the last respect the critical energy $E_c = 2/(1-\epsilon)$ is small.

In [6] the estimate $\gamma_\tau > 1$ was predicted for all d, while this is a feature of the cases $d > 1$. In the latter cases the analytic calculation of exact values of exponents γ_τ, γ_s is a difficult problem.

Remark 3.9 *The difference between cases $d = 1$ and $d > 1$ demonstrated in the theorem is known in the physical literature. The former case is usually considered as the trivial SOC-model.*

3.8.3 Thermodynamic limit

By thermodynamic limit of an observable ϕ we understand the double limit

$$[\phi]_\infty = \lim_{\substack{L\to\infty \\ E_c\to\infty}} \phi$$

if it exists. It is assumed that for physical observables this limit exists (in [6] only limit $L \to \infty$ was considered, though then the value of energy E_c could serve as an essential parameter, which is not desirable in the SOC-paradigm).

As an example of non-physical observable we expose s_0 (in §3.8.1 we called it mathematically relevant): The double limit does not exists because the repeated limits are different:

$$0 = \lim_{L\to\infty} \lim_{E_c\to\infty} \bar{s}_0 \neq \lim_{E_c\to\infty} \lim_{L\to\infty} \bar{s}_0 = +\infty.$$

But s_+ and τ are good physical observables, for the thermodynamic limits exist:

$$[\bar{\tau}]_\infty = \infty, \quad [\bar{s}_+]_\infty = \infty.$$

From §3.4.4 (use iteration arguments in the topological case) and §3.5.2 we obtain:

$$[h_{\text{top}}(\hat{f})]_\infty = [h_{\text{top}}(F)/\bar{\tau}]_\infty = 0 \quad \text{and} \quad [h_{\text{top}}(F)]_\infty = \infty.$$

But with ω the situation is different because the proof (rather than the vague statement of part (ii)) of Theorem 3.23 implies: $\lim_{L\to\infty}\lim_{E_c\to\infty}\bar{\omega} = \infty$, while $\lim_{E_c\to\infty}\lim_{L\to\infty}\bar{\omega}$ is finite. Since ω is definitely physically relevant observable one needs the following reparametrization: $\omega \mapsto \omega/E_c$, which corresponds to contraction of the waiting time via the following ansatz:

We let the energy quantum added at a unit time to the system equal $\delta = \hbar$ (instead of 1 as before), but speed up time in the waiting intervals respectively: $E_c = E_0/\hbar$, $\omega_{\text{new}} = \omega\hbar$. Then the thermodynamic limit (the space part M can be quantized similarly via $L = [l/\hbar]$ with l a finite length) corresponds to the quasi-classical limit $\hbar \to 0$.

The duration of avalanche was suppressed in our definition of dynamics to length one. So if we want to find the entropy of the physical system, where each step of avalanche has time-duration one, we should multiply it by the probability of dropping energy into the system. For every trajectory this equals $\frac{\bar{\omega}}{\bar{\omega}+\bar{\tau}}$. This ratio behaves differently as $L \to \infty$ or $E_c \to \infty$, so we need the reparametrization described above.

In §3.5.2 we calculated the entropy of the "return" Zhang model $h_{\text{top}}(F) = \log N$. But after reparametrization it changes. Denoting by h^{Zhang} the entropy of the reparametrized system we get:

$$h_\mu^{\text{Zhang}} = h_{\mu^u}(\sigma_N^+) \cdot \langle \frac{\bar{\omega}_{\text{new}}}{\bar{\omega}_{\text{new}} + \bar{\tau}} \rangle, \quad h_{\text{top}}^{\text{Zhang}} = d \cdot \log L \cdot \big(\frac{\bar{\omega}_{\text{new}}}{\bar{\omega}_{\text{new}} + \bar{\tau}} \big)_{\max}.$$

This implies:

$$[h_\mu^{\text{Zhang}}]_\infty = [h_{\text{top}}^{\text{Zhang}}]_\infty = 0.$$

Therefore the expanding property is lost in the thermodynamic limit for the original physical system, as was already noticed in [6] for a bit different situation.

Remark 3.10 *Notice that in the reparametrized system $\bar{\omega}_{new} \ll \bar{\tau}$, which is counter-intuitive for certain SOC-examples (sandpile, earthquakes etc, where one expects $\bar{\omega} \gg \bar{\tau}$). This indicates that the Zhang system should be modified by introducing the local contraction of time depending on the avalanche size or speed. We will not consider such gradient-type models here.*

3.8.4 Lyapunov spectrum

In §3.4.3 we showed that the Zhang model is hyperbolic with one positive exponent and the remainder of the spectrum negative. This hyperbolicity is lost in the thermodynamic limit.

Proposition 3.24 *For $E_c \geq \epsilon/(1-\epsilon)$ we have $\sum_{i=1}^{N} \chi_i^- = \bar{s}_0 \log \epsilon$.*

Proof. We cannot ensure the existence of an SRB-measure, so both the Lyapunov exponents and $\overline{s}_0$ should be seen as functions on M. From the general theory of Lyapunov exponents we know that

$$\sum_{i=1}^{N} \chi_i^-(\hat{x}) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n-1} \log\det \mathcal{T}(F^t \hat{x})\,.$$

From Proposition 3.7 we know that the formula $\det(L_{ij}) = \epsilon^{s_{ij}}$ holds for $E_c \geq \epsilon/(1-\epsilon)$. Hence

$$\begin{aligned}
\sum_{i=1}^{N} \chi_i^-(x) &= \sum_{i=1}^{N} \int_{\Sigma_N^+} \chi_i^-(\mathbf{t}, x)\, d\mu_{\text{Ber}} \\
&= \int_{\Sigma_N^+} \sum_{i=1}^{N} \chi_i^-(\mathbf{t}, x)\, d\mu_{\text{Ber}} \\
&= \int_{\Sigma_N^+} \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \log(\det \mathcal{T})(F^t(\mathbf{t}, x))\, d\mu_{\text{Ber}} \\
&= \int_{\Sigma_N^+} \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} s(\mathbf{t}, x)\, d\mu_{\text{Ber}} \log \epsilon \\
&= \overline{s}_0(x) \log \epsilon\,.
\end{aligned}$$

□

Corollary 3.4 $|\overline{\chi}^-| = \frac{\bar{s}_0}{N} \log \frac{1}{\epsilon} \to 0$ *as* $E_c \to \infty$, *but* $|\overline{\chi}^-| \to \infty$ *as* $L \to \infty$.

Thus we should study differently the following cases:

1. $E_c \to \infty$, but L (and d) fixed. Since $\lim_{E_c\to\infty} \bar{s}_0 = 0$, the negative part of the Lyapunov spectrum collapses: $\lim_{E_c\to\infty} \chi_i^- = 0$ for every $1 \leq i \leq N$. The hyperbolicity is lost, but the positive exponent $\chi_0^+ = \log N$ survives. In particular, the entropy does not collapses.

2. $L \to \infty$, but $E_c \gg 1$ fixed. Here only a bounded piece $\chi_1^-, \ldots, \chi_k^-$ of the Lyapunov spectrum collapses, $k = \text{const}$. But the number of elements of this spectrum grows and in average $|\overline{\chi}^-| \to \infty$. In particular, $|\chi^-|_{\max} \to \infty$ as $L \to \infty$. Again, the positive exponent χ_0 and entropy are preserved and though we loose hyperbolicity there are many non-degenerate Oscelledec modes. Moreover they prevail over collapsing modes and so essentially the hyperbolicity is preserved as well.

3. Reparametrized model. This was introduced in §3.8.3 and require the renormalization: multiplication of waiting time by the function $\frac{\bar{\omega}}{\bar{\omega}+\bar{\tau}}$ along the trajectory. In this case the Lyapunov spectrum collapses to zero in any limit $E_c \to \infty$ and $L \to \infty$ and the hyperbolicity is completely lost.

Thus the exponential grow of the statistics is suppressed and we can observe power law statistic as is the basic idea of SOC-phenomenon. The corresponding SOC-exponents are related to the asymptotic of the Lyapunov spectrum (as discussed in [6, 7]), but are difficult to calculate analytically.

3.9 Conclusion

It is of importance to the paradigm of SOC to understand the mechanisms behind the behavior one observes numerically in the Zhang model and other sandpile models, and the aim of this paper is to provide a first step to a rigorous mathematical understanding of the dynamics of Zhang model.

Due to the singularities and non-invertibility of the model there existed very few applicable results, and hence we had to modify known results and develop some new methods in order to describe the dynamical properties of the model.

The result of this work is that the singularities play a modest role in the sense that do not change the main dynamical characterizations such as topological entropy, entropy of SRB-measures and the hyperbolic structure.

We show that any SRB-measure maximizes entropy, and based one this result one could hope to develop a thermodynamic formalism for the Zhang model. However, the presence of essential singularities prevents us from coding the system using Markov chains, except in some special cases where singularities only effect the dynamics for a finite number of time-steps (singularities are removable).

In the physically interesting values of parameters (E_c relatively great compared with $(1-\epsilon)^{-1}$) the attractor has a very complicated topological structure and in general cannot be described as a sub-shift of finite type. The effects of the singularities can also be seen in the rich fractal structure of the spacial attractors.

Our analysis of the effect of the singularities allows us to take the thermodynamic limit of the main dynamical quantities, showing that the entropy goes to zero and that the Lyapunov-spectrum completely collapses to zero (for the re-scaled physical

model). From a physical point of view this is interesting. Typically chaotic dynamics (positive entropy and positive Lyapunov exponents) is an indication of exponential speeds of mixing (exponential decay of correlations) [5], which is not compatible with the power-law statistics of the SOC-hypothesis. The loss of hyperbolicity in the thermodynamic limit hence supports the critical behavior observed numerically [30], [57], [25].

To conclude: We have shown that the Zhang model is chaotic hyperbolic dynamical system, where all the entropy is produced by the random driving of the system, and due to the singularities the orbit structure is richer than for a topological Markov chain. The hyperbolicity implies that under weak conditions there exists a SRB-measure (self-organization), and in the thermodynamic limit the hyperbolicity is lost and we may expect power-law statistics (criticality). In practice the systems that are studied have finite size and finite critical energy. Hence they are chaotic, but with small entropy. The SOC-hypothesis is that these weakly chaotic systems have SRB-measures, but that the rates of convergence to these measures are exponential but slow compared to a unit step in an avalanche, causing what appears to be power-law statistics.

Appendix

A Topological entropy of piecewise affine maps

A.1 The Buzzi theorem and its generalization

The statement as it is done in [16] does not apply to our situation. In [17] Buzzi noted that it extends to isometries and contractions. In fact, the assertion holds always, but since we cannot make a simple reference we write an adapted proof for the convenience of the reader.

Theorem 3.25 *Let $X \subset \mathbb{R}^d$ be a bounded polytope and $f : X \to X$ a piecewise affine map. Then*

$$H_{\text{sing}}(f) \leq \lambda^+(f) + H_{\text{mult}}(f)\,,$$

where

$$\lambda^+(f) = \overline{\lim_{n\to\infty}} \sup_{x\in X} \frac{1}{n} \max_k \log \|\Lambda^k d_x f^n\|\,.$$

Proof. Let $\mathcal{P} = \{P_i\}$ be the continuity partition and $f_i := f|_{P_i}$ be affine maps. Fix $\epsilon > 0$ and $T = T(\epsilon) \geq \frac{d}{\epsilon}\log(\sqrt{d}+1)$ such that for $n \geq T$ we have:

$$\text{mult}(\mathcal{P}^n) \leq \exp((H_{\text{mult}}(f)+\epsilon)n),\ \|\Lambda df^n\| \leq \exp((\lambda^+(f)+\epsilon)n).$$

Take $r = r(\epsilon)$ to be compatible with the partition $\mathcal{P}^T$ (for f^T), i.e. any r-ball intersects maximally $\text{mult}(\mathcal{P}^T)$ partition elements.

We will prove that each non-empty cylinder $C(\mathbf{a}) = [P_{a_0}\dots P_{a_{lT-1}}]$ of length $|\mathbf{a}| = lT$ can be partitioned into a collection $Q(\mathbf{a}) = \{W\}$ satisfying the following properties:

1. $\sum_{|\mathbf{a}|=lT} \text{card}(Q(\mathbf{a})) \leq C_0 \exp((\lambda^+(f) + H_{\text{mult}}(f) + 3\epsilon)lT)$

2. $\text{diam}(f_{a_{lT-1}} \circ \dots \circ f_{a_0}(W)) \leq r\,.$

Let us prove this claim by induction assuming it holds for some $l \geq 0$. The base of induction is obvious and C_0 is the minimal cardinality of an $r/2$-ball cover.

Take a partition element $W \in Q(\mathbf{a})$ that is used to cover the cylinder $[P_{a_0} \dots P_{a_{lT-1}}]$. By the induction hypothesis it has diameter less than r, so it can be continued to cover a non-empty cylinder of length $(l+1)T$ in at most $\mathrm{mult}(\mathcal{P}^T)$ ways. So to cover the cylinders $[P_{a_0} \dots P_{a_{lT-1}} P_{b_0} \dots P_{b_{T-1}}]$ we make a division of W:

$$W = \bigcup_{i=1}^{\gamma} W_i' \,, \gamma \leq \mathrm{mult}(\mathcal{P}^T)\,.$$

Let

$$W_i'' = f_{a_{Tl-1}} \circ \dots \circ f_{a_0}(W_i')$$

and

$$W_i''' = f_{b_{T-1}} \circ \dots \circ f_{b_0}(W_i'')\,.$$

By the assumption $\mathrm{diam}(W_i'') < r$ for all $i = 1, \dots, \gamma$, but the sets W_i''' may have greater diameter. We need to divide the sets W_i''' so that they have diameter less than r, and then pull this refinement back to the partition of sets W_i'.

Let L be the differential of $f_{b_{T-1}} \circ \dots \circ f_{b_0}$ on W_i''. We can assume that L is symmetric and take $\{e_k\}$ to be a basis of eigenvectors corresponding to eigenvalues $\lambda_1, \dots, \lambda_d$. Let $\{v_k\}$ be a basis in the vector subspace corresponding to W_i'''. We can choose this basis to be orthonormal and triangular with respect to $\{e_k\}$. Divide W_i''' by the hyperplanes

$$\psi_j(x) \stackrel{\mathrm{def}}{=} \langle v_j, x \rangle = p\frac{r}{\sqrt{d}}, \quad p \in \mathbb{Z}, \quad j = 1, \dots, d\,.$$

This defines cells $\tilde{W}$ of diameter less than r. Since $\psi_j(W_i''') = \psi_j(L(W_i''))$ has diam $\leq |\lambda_i| r$, the number of cells $\tilde{W}$ needed to cover W_i''' is less than or equal to

$$(\sqrt{d}+1)^d |\lambda_1|^+ \dots |\lambda_d|^+ \leq (\sqrt{d}+1)^d \|\Lambda df^n\| \leq \exp\big((\lambda^+(f) + 2\epsilon)T\big)\,,$$

where $|\lambda_i|^+ = \max\{|\lambda_i|, 1\}$.

Therefore the total cardinality of the new partition is less than or equal to

$$\begin{aligned}\mathrm{mult}(\mathcal{P}^T) \exp\big((\lambda^+(f) + 2\epsilon)T\big) \exp\big((\lambda^+(f) + H_{\mathrm{mult}}(f) + 3\epsilon)lT\big) \\ \leq \exp\big((\lambda^+(f) + H_{\mathrm{mult}}(f) + 3\epsilon)(l+1)T\big)\,.\end{aligned}$$

This proves the statement. □

The theorem holds as well for most degenerate piece-wise affine systems, but there can be problems with $H_{\mathrm{mult}}(f)$. Namely the latter is not defined if the image

of continuity domain contains a boundary face of a continuity domain. But if we assume the image and the faces always meet transversally, no problems occur and the above theorem applies literally.

In degenerate cases of the Zhang model, the above requirement holds for most parameters. For instance, if $N = 2$ and $\epsilon = 0$, then all $E_c \notin \mathbb{Z}$ satisfy the request. For integer E_c the theory fails, but one can look just to the whole image set, which has dimension $< N$: its Poincaré return map is piece-wise affine and it satisfies the requirements. In the above example $N = 2$, $\epsilon = 0$ the dynamics is confined to two one dimensional lines, whence the multiplicity entropy is zero (by dimensional reasons) and $h_{\text{top}}(F) = \log 2$ in this case.

A.2 Entropy of conformal piecewise affine skew-products

We say that an affine map is conformal modulo degenerations if the image on some subspace transversal to the kernel is mapped conformally to its image. A piecewise affine map is said to be conformal modulo degenerations if all its affine components are conformal modulo degenerations. We will assume that degenerations satisfy the transversality requirement of A.1.

In [36] we noticed that $H_{\text{mult}} = 0$ for piece-wise affine conformal maps. This easily extends to allow degenerations. Now we consider a more general situation of skew-product systems of Zhang's type.

Theorem 3.26 *Let $f_i : X \to X$ be piece-wise affine non-strictly contracting and conformal modulo degenerations, $i = 0, \dots, N-1$. Define $F : \Sigma_N^+ \times X \to \Sigma_N^+ \times X$ by the formula $F(\mathbf{t}, x) = (\sigma_N^+(\mathbf{t}), f_{t_0}(x))$, $\mathbf{t} = t_0 t_1 \cdots \in \Sigma_N^+$, $x \in X \subset \mathbb{R}^d$. Then we have: $h_{\text{top}}(F) = \log N$ (so that a-posteriori the variational principle holds).*

Remark 3.11 *For $N = 2$ and $\epsilon = 0$ the affine components of have rank 1, and hence the above Theorem 3.26 shows that the topological entropy is $\log N$. The same holds for $E_c = \epsilon = 1/3$.*

Proof. It follows from Theorem 3.25 that it suffice to prove that $H_{\text{mult}}(F) = 0$.

To achieve the desired equality note that preimages of the time-like singularity planes $\{t = n/N\}$ never intersect under inverse iterations of $F(t, x) = (Nt \bmod 1, f_{[Nt]}(x))$. Thus $H_{\text{mult}}(F) = \sup_{\mathbf{t}} H_{\text{mult}}(f_{\mathbf{t}})$. We will prove that $H_{\text{mult}}(f_{\mathbf{t}}) = 0$ for all $\mathbf{t} \in \Sigma_N^+$.

A piecewise affine map can be considered as an ordered triple $(X, \mathcal{P}, f)$, where X is a polytope in $\mathbb{R}^d$, $\mathcal{P} = \{P_i\}$ is a partition of X made up of pairwise disjoined polytopes (with certain faces of boundary included, so that the whole boundary is distributed between polytopes) and $f_i := f|_{P_i} : P_i \to X$ are affine maps. We let $X' = \cup \operatorname{Int}(P_i)$ and $\operatorname{Sing}(f) = X \setminus X'$.

For a piecewise affine map $(X, \mathcal{P}, f)$ and a point $x \in X$ construct a piecewise affine map $(X_x, \mathcal{P}_x, f_x)$, called the differential of f at x, by letting

1. $X_x = \{y \in \mathbb{R}^d \,|\, \exists \epsilon_0 > 0 \,\text{s.t.} \forall \epsilon \in (0, \epsilon_0) : \, x + \epsilon y \in X\} \subset \mathbb{R}^d$ is the tangent cone to X.

2. $\mathcal{P}_x$ is the partition of X_x consisting of non-empty sets

$$P_x = \{y \in \mathbb{R}^d \,|\, \exists \epsilon_0 > 0 \,\text{s.t.} \forall \epsilon \in (0, \epsilon_0) : \, x + \epsilon y \in P\}\,,$$

   where $P \in \mathcal{P}$.

3. $f_x : P_x \to X_{f(x)}$ is the collection of maps

$$f_x(y) = \lim_{\epsilon \to 0^+} \frac{f(x + \epsilon y) - \lim_{\delta \to 0^+} f(x + \delta y)}{\epsilon}\,, \quad P_x \in \mathcal{P}_x\,.$$

Consider iterated differentials and denote $f_{x_1,\ldots,x_n} := (\ldots (f_{x_1})_{x_2} \ldots)_{x_n}$. Note that $(f^n_{\mathbf{t}})_x = (f_{t_{n-1}})_{f^n_{\mathbf{t}}(x)} \circ \ldots (f_{t_1})_{f^1_{\mathbf{t}}(x)} \circ (f_{t_0})_x$.

Proposition 3.27 *There exists a constant $C \in \mathbb{R}_+$ such that for any subspace $W \subset \mathbb{R}^d$ and any $(x_1, \ldots, x_r) \in X \times W^{r-1}$ we have:*

$$\operatorname{mult}((f^n_{\mathbf{t}}|_W)_{x_1,\ldots,x_r}) \le C \Big(\sup_{V \subset \mathbb{R}^d} \sup_{y_1,\ldots,y_{r+1}} \operatorname{mult}(f_{\mathbf{t}}|_V)_{y_1,\ldots,y_{r+1}} \Big)^n, \qquad \forall n \ge 0,$$

where the collection of points $(y_1, \ldots, y_{r+1})$ runs over $X \times V^r$ with the condition $\operatorname{rank}(y_2, \ldots, y_{r+1}) + \operatorname{codim} V = \operatorname{rank}(x_2, \ldots, x_r) + \operatorname{codim} W + 1$.

The theorem follows from this, because for

$$\mu(r) = \overline{\lim_{n \to \infty}} \frac{1}{n} \log \sup_{\substack{y_1 \in X, y_2 \ldots, y_r \in V \subset \mathbb{R}^d \\ \operatorname{rank}(y_2,\ldots,y_r) + \operatorname{codim} V = r-1}} \operatorname{mult}(f^n_{\mathbf{t}}|_V)_{y_1,\ldots,y_r}$$

we have: $H_{\text{sing}}(f_{\mathbf{t}}) = \mu(1) \le \mu(2) \le \cdots \le \mu(d+1) = 0$.

To prove the proposition note that when $r = 0$ we have (in this case we do not need W, V):

$$\operatorname{mult}(f^n_{\mathbf{t}}) \le \Big(\sup_{x \in X} \operatorname{mult}(f_{\mathbf{t}})_x \Big)^n = \Big(\max_{0 \le j < N} \sup_{x \in X} \operatorname{mult}(f_j)_x \Big)^n.$$

Let $r \ge 1$. Consider the continuity partition $\mathcal{P}^{\mathbf{t}}_{x_1,\ldots,x_r}$ for $(f_{\mathbf{t}})_{x_1,\ldots,x_r}$, which is just the continuity partition $\mathcal{P}^{(t_0)}_{x_1,\ldots,x_r}$ of $(f_{t_0})_{x_1,\ldots,x_r}$ (one iteration), and let $\mathcal{P}^{n,\mathbf{t}}_{x_1,\ldots,x_r}$ for $(f^n_{\mathbf{t}})_{x_1,\ldots,x_r}$ be the iterated partition. Note that the latter is the collection of all non-empty intersections

$$P^{t_0} \cap (f_{t_0})^{-1}_{P^{t_0}}(P^{t_1}) \cap (f_{t_0})^{-1}_{P^{t_0}}(f_{t_1})^{-1}_{P^{t_1}}(P^{t_2}) \cap \cdots \cap (f_{t_0})^{-1}_{P^{t_0}} \ldots (f_{t_{n-1}})^{-1}_{P^{t_{n-1}}}(P^{t_n}),$$

where P^{t_i} are elements of $\mathcal{P}^{(t_i)}_{x_1,\dots,x_r}$ and f_P denotes the restriction of the (differential of the) map to the corresponding continuity domain. Every element of these partitions is invariant under the shift by vectors from $\mathrm{span}(x_2,\dots,x_r)$. Therefore it intersects the unit sphere $S_1(x_2,\dots,x_r)^\perp$ in the orthogonal complement. Consider the induced partition on the sphere and refine it so that every element has diameter no greater than ε. Denote by $n(\mathcal{P}^{n,\mathbf{t}}_{x_1,\dots,x_r},\varepsilon)$ the minimal cardinality of such a refinement. Let also $m(\mathcal{P}^{(t_i)}_{x_1,\dots,x_r},\varepsilon)$ be the maximal number of elements of $\mathcal{P}^{(t_i)}_{x_1,\dots,x_r}$ that an ε-ball $B(y,\varepsilon)\cap S_1^\perp$ of $S_1(x_2,\dots,x_r)^\perp$ can meet.

Denote by $n(\mathcal{P}\cap W,\varepsilon)$, $m(\mathcal{P}\cap W,\varepsilon)$ the corresponding quantities in the subspace W. Then from the above formula for the iterated partition:

$$n(\mathcal{P}^{n+1,\mathbf{t}}_{x_1,\dots,x_r}\cap W,\varepsilon)\le n(\mathcal{P}^{n,\mathbf{t}}_{x_1,\dots,x_r}\cap W,\varepsilon)\cdot m(\mathcal{P}^{(t_{n+1})}_{y_1,\dots,y_r}\cap V,\varepsilon),$$

where $y_1=f^n_{\mathbf{t}}(x_1)$, $y_2=f^{n\prime}_{\mathbf{t}}(x_2),\dots,$ $y_r=f^{n\prime}_{\mathbf{t}}(x_r)$ and $V=f^{n\prime}_{\mathbf{t}}(W)$ with $f^{n\prime}_{\mathbf{t}}=(f^n_{\mathbf{t}})_{x_1,\dots,x_r}$. Therefore

$$n(\mathcal{P}^{n+1,\mathbf{t}}_{x_1,\dots,x_r}\cap W,\varepsilon)\le n(\mathcal{P}^{n,\mathbf{t}}_{x_1,\dots,x_r}\cap W,\varepsilon)\cdot\sup_V\sup_{y_1,\dots,y_r} m(\mathcal{P}^{(t_{n+1})}_{y_1,\dots,y_r}\cap V,\varepsilon),$$

where the supremum is taken over all $V\subset\mathbb{R}^d$ and $(y_1,\dots,y_r)\in X\times V^{r-1}$ such that codimension of $\langle y_2\dots,y_r\rangle$ in V equals codimension of $\langle x_2\dots,x_r\rangle$ in W.

Since for a fixed ε the number $n(\mathcal{P}^{t_0}_{x_1,\dots,x_r},\varepsilon)$ is finite and

$$m(\mathcal{P}^{(t_i)}_{y_1,\dots,y_r}\cap V,\varepsilon)\le\sup_{y\in S_1(y_1,\dots,y_r)^\perp\cap V}|\mathcal{P}^{(t_i)}_{y_1,\dots,y_r}\cap B(y,\varepsilon)\cap V|,$$

the claim follows from the following statement. Fix $i\in[0,N)$.

Lemma 3.28 *There exists $\varepsilon>0$ (depending only on i) such that for all $V\subset\mathbb{R}^d$ and all $(y_1,\dots,y_r)\in X\times V^{r-1}$, with $\mathrm{rank}(y_2,\dots,y_r)<\dim V$, and $y\in S_1(y_2,\dots,y_r)^\perp\cap V$ there exists $y'\in(y_2,\dots,y_r)^\perp\cap V$ satisfying:*

$$|\mathcal{P}^{(i)}_{y_1,\dots,y_r}\cap B(y,\varepsilon)\cap V|\le\mathrm{mult}((f_i|_V)_{y_1,\dots,y_r,y'}).$$

This statement, modulo our notations and restrictions to V, is proved in [17]. The proposition and hence the theorem follow. □

A.3 Estimates on entropy by angular expansion rates

It is possible to estimate the effect of angular expansion on the topological entropy of a piecewise affine map $f:X\to X$ by its spherizations. Define the piecewise smooth map $d^{(s)}_x f:ST_xX\to ST_{f(x)}X$ given at $x\in X'$ by the formula

$$d^{(s)}_x f(v)=\frac{d_xf(v)}{\|d_xf(v)\|}\,.$$

For $x \in \operatorname{Sing}(f)$ and $v \notin T_x \operatorname{Sing}(f)$ (the tangent cone) we let $d_x^{(s)} f(v) = \lim_{\epsilon \to +0} d_{x+\epsilon v}^{(s)} f(v)$. For other $(x, v) \in STX$ the map is not defined. The angular expansion of f is exactly the expansion in the fibers of its spherization.

If $d_x f$ is degenerate we restrict to the orthogonal component of its kernel, and consider the map

$$d_x f|_{\operatorname{Ker}(d_x f)^\perp} : \operatorname{Ker}(d_x f)^\perp \to \operatorname{Im}(d_x f)\,.$$

Then the map $S_x(f) = d_x^{(s)} f|_{\operatorname{Ker}(d_x f)^\perp} : S \operatorname{Ker}(d_x f)^\perp \to S \operatorname{Im}(d_x f)$ between $(\operatorname{rank}(d_x f) - 1)$-dimensional spheres is given by the formula

$$v \mapsto \frac{d_x f|_{\operatorname{Ker}(d_x f)^\perp}(v)}{\|d_x f|_{\operatorname{Ker}(d_x f)^\perp}(v)\|}.$$

For $i < d$ we define

$$\rho_i(f) = \overline{\lim_{n \to \infty}} \frac{1}{n} \sup_{(x,v)} \max_{0 \le k \le i} \log \|\Lambda^k d_v S_x(f^n)\|.$$

Let $m_* = \min_x \dim \operatorname{Ker}(d_x f)$ and $d_* = d - m_* = \max_x \operatorname{rank}(d_x f)$, where d is the dimension of X. The numbers $\rho_i(f)$ can be non-zero only for $i < d_*$.

We have: $\rho_0(f) = 0$. The number $\rho_1(f)$ measures the maximal exponential rate with which angles can increase under the map f. The numbers $\rho_i(f)$ for $i < d$ measures the maximal rate of expansion of the restrictions to i-dimensional spheres. If f is conformal, then $\rho_i(f) = 0$ for all i.

Theorem 3.29 ([36]). *For piece-wise affine maps $H_{\mathrm{mult}}(f) \le \sum_{i=1}^{d_*-1} \rho_i(f)$.*

We define the maximal expansion rate

$$\lambda_{\max}(f) = \overline{\lim_{n \to \infty}} \sup_x \frac{1}{n} \log \|d_x f^n\|\,,$$

and the minimal finite expansion rate

$$\lambda_{\min}(f) = -\overline{\lim_{n \to \infty}} \sup_x \frac{1}{n} \log \|(d_x f^n|_{\operatorname{Ker}(d_x f^n)})^{-1}\|\,.$$

In [36] we show that

$$\rho_i(f) \le i\Big(\lambda_{\max}(f) - \lambda_{\min}(f)\Big)\,.$$

This gives the following result (the same bound holds for $h_{\mathrm{top}}(f)$):

Theorem 3.30 *For a piecewise affine map f it holds:*

$$H_{\mathrm{sing}}(f) \le \lambda^+(f) + \frac{d_*(d_* - 1)}{2}\Big(\lambda_{\max}(f) - \lambda_{\min}(f)\Big)\,.$$

B Generalization of Moran's formula

Consider an IFS $(M^d, f_1, \dots, f_N)$ on a Riemannian manifold M, where the maps f_i can possess singularities, but we assume that they are mild in a sense that the number of continuity domains is finite, any of them has piece-wise smooth boundary and the map, restricted to any of the domains, smoothly extends to the adjacent singularities, so that the number of continuity domains for f_i is finite, each of them is a smooth manifold with boundary and f_i has a smooth extension to the closure of every such domain (this is the case the case of the Zhang model).

Remark that the IFS can be interpreted as the dynamical system $(\Sigma_N^+ \times M, F)$, $\hat{f}(\mathbf{t}, x) = (\sigma_N^+ \mathbf{t}, f_{t_0} x)$. Attractor $\mathcal{Y}$ of the IFS can be defined via the attractor of the extended system F, which has the form $\mathcal{A} = \Sigma_N^+ \times \mathcal{Y}$.

Let $\|\cdot\|$ be the norm on TM generated by the metric on M. Denote

$$s_i^+ = \max_{x \in M} \|d_x f_i\|, \qquad s_i^- = \big(\max_{x \in M} \|d_x f_i^{-1}\|\big)^{-1}.$$

We assume that the maps are non-degenerate (this is just for simplicity of arguments) and strictly contracting, so that $0 < s_i^- \le s_i^+ < 1$.

Let $\eta = \max_{x \in M} \#\{i \,|\, x = f_i(y_i) \text{ for some } y_i \in \mathcal{Y}\}$ be the maximal multiplicity of overlaps and $\varkappa_i = \max_{x \in \mathcal{Y}} \#\{y \in \mathcal{Y} \,|\, x = f_i(y)\}$ be the maximal multiplicity of self-overlaps (we assume it is finite) on the attractor. Denote also by ϑ_i the multiplicity of the continuity partition for $f_i|_{\mathcal{Y}}$, i.e. the maximal number of continuity domains intersecting the attractor and meeting at one point of it.

Theorem 3.31 *Let $\underline{D} = \alpha$, $\overline{D} = \beta$ be the solutions of the equations*

$$\sum_{i=1}^N \tfrac{1}{\varkappa_i} |s_i^-|^\alpha = \eta, \qquad \sum_{i=1}^N \vartheta_i |s_i^+|^\beta = 1.$$

Then the Hausdorff dimension of the attractor satisfies:

$$\underline{D} \le \dim_H(\mathcal{Y}) \le \overline{D}.$$

In addition to Hausdorff dimension we will need some other dimensional characteristics (see [44] for details). Denote by $\mathfrak{N}(X, \delta)$ the minimal cardinality of covers of X by balls of radius δ. Then the lower and upper box dimensions are defined by the formula:

$$\underline{\dim}_B(X) = \varliminf_{\delta \to +0} \log \mathfrak{N}(X, \delta) / \log \tfrac{1}{\delta}, \quad \overline{\dim}_B(X) = \varlimsup_{\delta \to +0} \log \mathfrak{N}(X, \delta) / \log \tfrac{1}{\delta}.$$

When these quantities are equal, their value is also called fractal dimension.

Consider a Borel probability measure $\mu \in \mathcal{M}(X)$ (an SRB-measure on the attractor can be taken in the SOC-context if it exists). The upper and lower pointwise dimensions are defined then as

$$\underline{d}_\mu(x) = \varliminf_{\delta\to+0} \log \mu(B(x,\delta))/\log\delta, \qquad \overline{d}_\mu(x) = \varlimsup_{\delta\to+0} \log \mu(B(x,\delta))/\log\delta.$$

When they are equal and constant a.e. the measure μ is called exact-dimensional. This is precisely the case, when $\operatorname{supp}\mu = X$ and we have equality in the general chain of inequalities (together with (3.13) below):

$$\operatorname{ess.inf} \underline{d}_\mu(x) \le \dim_{\mathrm{H}}(X) \le \underline{\dim}_B(X) \le \overline{\dim}_B(X), \tag{3.12}$$

where by essential infimum we mean its upper bound taken over all subsets $U \subset X$ of measure 1 (and similar for $\operatorname{ess.sup} \overline{d}_\mu(x)$). The last two inequalities are known and the first one follows from the inequality $\operatorname{ess.inf} \underline{d}_\mu(x) \le \dim_{\mathrm{H}}(\mu)$ ([44]), where $\dim_{\mathrm{H}}(\mu) = \lim_{\delta\to+0} \inf\{\dim_{\mathrm{H}}(Z) \,|\, \mu(Z) > 1-\delta\}$.

Other dimensional characteristics of the measure are defined similarly and satisfy:

$$\underline{\dim}_B(\mu) \le \overline{\dim}_B(\mu) \le \operatorname{ess.sup} \overline{d}_\mu(x). \tag{3.13}$$

Note that $\overline{\dim}_B(\mu) \le \overline{\dim}_B(X)$, while the quantities $\operatorname{ess.sup} \overline{d}_\mu(x)$ and $\overline{\dim}_B(X)$ are in general incomparable.

The known formulas for the Hausdorff and other dimensions are generalizations of Moran's result ([42, 28]) and are based on the Bowen's equation (using the idea of coding); in this case one usually obtains exact-dimensionality [44]. In the SOC-context coding becomes problematic in the presence of singularities (unless the properties of the SRB-measure are clarified) and thus we cannot easily establish exact-dimensionality or formula for the dimension.

We prove instead the inequality of the theorem for all the various dimensions from (3.12) and (3.13), which we denote just by $\dim(\mathcal{Y})$:

$$\underline{D} \le \dim(\mathcal{Y}) \le \overline{D}.$$

Proof. Let us consider at first the upper box dimension $\overline{\dim}_B(\mathcal{Y})$. The function $\mathfrak{N}$ satisfies the inequalities:

$$\tfrac{1}{\varkappa_i}\mathfrak{N}(\mathcal{Y}, \delta/s_i^-) \le \mathfrak{N}(f_i(\mathcal{Y}), \delta) \le \vartheta_i \mathfrak{N}(\mathcal{Y}, \delta/s_i^+).$$

The inequality from above is obtained as follows. Let $\mathcal{S} = \{x_j\}$ be a δ-spanning set, i.e. a collection of points from X with $U_\delta(\mathcal{S}) = X$. Then $f_i(\mathcal{S}) = \{f_i(x_j)\}$ may fail to be a $\delta \cdot s_i^+$-spanning set thanks to singularities. Whenever $\delta \ll 1$, every δ-ball intersects maximally ϑ_i domains of continuity for f_i meeting $\mathcal{Y}$. Then we need to add maximally ϑ_i points for each ball $U_\delta(x_j)$ intersecting singularities. The inequality from below is proved similarly.

Now we have:

$$\mathcal{Y} = f_1(\mathcal{Y}) \cup \cdots \cup f_N(\mathcal{Y})$$

and the same for U_δ-neighborhoods. This implies:

$$\mathfrak{N}(\mathcal{Y}, \delta) \leq \sum_{i=1}^{N} \mathfrak{N}(f_i(\mathcal{Y}), \delta) \leq \sum_{i=1}^{N} \vartheta_i \mathfrak{N}(\mathcal{Y}, \delta/s_i^+). \tag{3.14}$$

Denote $\sigma(\delta) = \mathfrak{N}(\mathcal{Y}, \delta)\delta^{\overline{\dim}_B(\mathcal{Y})}$. This functions grows sub-polynomially:

$$\overline{\lim_{\delta \to +0}} \frac{\log \sigma(\delta)}{\log 1/\delta} = 0. \tag{3.15}$$

Lemma 3.32 *Let $\lambda_i > 1$ be some numbers and $p_i > 0$ be some probabilities, $\sum_{i=1}^{N} p_i = 1$. Then (3.15) implies:*

$$\underline{\lim_{\delta \to +0}} \frac{\sum_{i=1}^{N} p_i \sigma(\lambda_i \delta)}{\sigma(\delta)} \leq 1.$$

Proof. Suppose the lower limit is $> \kappa > 1$. Then for every sufficiently small δ there exists $i \in [1, N]$ such that $\sigma(\lambda_i \delta) \geq \kappa \sigma(\delta)$.

Denote $\bar{\lambda} = \max_{1 \leq i \leq N} \lambda_i$. Let $C = \max_{\delta \in [1/\bar{\lambda}, 1]} \sigma(\delta)$. Then:

$$\sigma(\delta) \leq \frac{1}{\kappa}\sigma(\lambda_{i_1}\delta) \leq \frac{1}{\kappa^2}\sigma(\lambda_{i_1}\lambda_{i_2}\delta) \leq \cdots \leq \frac{1}{\kappa^{s(\delta)}}\sigma(\lambda_{i_1} \ldots \lambda_{i_{s(\delta)}}\delta),$$

where $s(\delta)$ is the first number such that $\lambda_{i_1} \ldots \lambda_{i_{s(\delta)}}\delta \in [1/\bar{\lambda}, 1]$. This number can be estimated as follows: $s(\delta) \geq -\log \delta / \log \bar{\lambda} - 1$, whence:

$$\frac{\log \sigma(\delta)}{\log 1/\delta} \leq \frac{\log C - s(\delta) \log \kappa}{\log 1/\delta} \leq \frac{\log(C\kappa)}{\log 1/\delta} - \frac{\log \kappa}{\log \bar{\lambda}}.$$

Therefore $\overline{\lim}_{\delta \to +0} \frac{\log \sigma(\delta)}{\log 1/\delta} \leq -\frac{\log \kappa}{\log \bar{\lambda}} < 0$ and we get a contradiction. This proves the lemma. $\square$

Now to obtain the inequality from above for $\overline{\dim}_B(\mathcal{Y})$ divide (3.14) by $\mathfrak{N}(\mathcal{Y}, \delta)$. Denoting $\varpi = \sum_{i=1}^{N} \vartheta_i |s_i^+|^{\overline{\dim}_B(\mathcal{Y})}$, $\lambda_i = 1/s_i^+$ and $p_i = \vartheta_i |s_i^+|^{\overline{\dim}_B(\mathcal{Y})}/\varpi$ we get:

$$\frac{1}{\varpi} \leq \sum_{i=1}^{N} p_i \frac{\sigma(\lambda_i \delta)}{\sigma(\delta)}.$$

Thus Lemma 3.32 implies that $1/\varpi \leq 1$ or

$$1 \leq \sum_{i=1}^{N} \vartheta_i |s_i^+|^{\dim \mathcal{Y}}$$

and the first claim $\overline{\dim}_B(\mathcal{Y}) \le \overline{D}$ follows from the contraction $|s_i^+| < 1$. The same arguments show another statement that $\overline{d}_\mu(x) \le \overline{D}$.

The inequality from below follows from

$$\mathfrak{N}(\mathcal{Y}, \delta) \ge \frac{1}{\eta} \sum_{i=1}^{N} \mathfrak{N}(f_i(\mathcal{Y}), \delta) \ge \frac{1}{\eta} \sum_{i=1}^{N} \frac{1}{\varkappa_i} \mathfrak{N}(\mathcal{Y}, \delta / s_i^-),$$

which implies $\sum_{i=1}^{N} \frac{1}{\varkappa_i} |s_i^-|^{\underline{\dim}_B(\mathcal{Y})} \le \eta$ and the same for the lower pointwise dimension: $\underline{d}_\mu(x) \ge \underline{D}$ a.e.

In this case we should define

$$\sigma(\delta) = \mathfrak{N}(\mathcal{Y}, \delta) \delta^{\underline{\dim}_B(\mathcal{Y})} \text{ or } \sigma(\delta) = \big(\operatorname{ess.inf} - \log \mu(B(x, \delta)) \big) \delta^{\underline{\dim}_B(\mu)}$$

respectively and use

Lemma 3.33 *Let $\lambda_i > 1$ be some numbers and $p_i > 0$ be some probabilities, $\sum_{i=1}^{N} p_i = 1$. Then:*

$$\varliminf_{\delta \to +0} \frac{\log \sigma(\delta)}{\log 1/\delta} = 0 \quad \Longrightarrow \quad \varlimsup_{\delta \to +0} \frac{\sum_{i=1}^{N} p_i \sigma(\lambda_i \delta)}{\sigma(\delta)} \ge 1.$$

This is proved similarly to Lemma 3.32. The inequalities for the Hausdorff dimension follows now from (3.12). □

Chapter 4

Entropy via multiplicity

Joint with B. Kruglikov

Abstract: The topological entropy of piecewise affine maps is studied. It is shown that singularities may contribute to the entropy only if there is angular expansion and we bound the entropy via the expansion rates of the map. As a corollary we deduce that non-expanding conformal piecewise affine maps have zero topological entropy. We estimate the entropy of piecewise affine skew-products. Examples of abnormal entropy growth are provided.

Refrence for this paper: B. Kruglikov, M. Rypdal, *Entropy via multiplicity*, Discrete and Continuous Dynamical Systems **16**, 395 (2006)

1 Introduction

For a smooth map f of a compact manifold the Ruelle-Margulis inequality together with the variational principle [15, 32] tells us that the topological entropy of f is bounded by the maximal sum over positive Lyapunov exponents. For maps with singularities this result is no longer true and there are examples of piecewise smooth maps, where the topological entropy exceeds what can be predicted from the rate of expansion.

In this paper we study the class of piecewise affine maps. It follows from [16] that for piecewise affine maps the entropy is bounded by the rate of expansion and the growth in the multiplicity of singularities. The latter was shown by J. Buzzi to be zero for piecewise isometries [17], but his proof does not generalize to non-expanding piecewise affine maps. In fact, we exhibit an example of a piecewise affine contracting map with positive topological entropy.

We show that the growth of multiplicity is an effect caused by angular expansion that can be estimated by the expansion rates of the map f. As a corollary we obtain that the topological entropy of a piecewise conformal map can be estimated by the expansion rate of the map as in the smooth compact case. It follows that piecewise affine non-expanding conformal maps have zero topological entropy.

In the second part of the paper we study the topological entropy of piecewise affine maps of skew-product type and obtain a formula which bounds the entropy of the skew products in terms of the entropy and multiplicity growth of the factors. The estimate includes a term which indicates that the entropy of a skew-product system may be greater than the sum of the maximal entropy of its factors and we give an example where this is realized.

Our main results have several corollaries for which one can calculate the topological entropy of various classes of piecewise affine maps.

2 Definitions and main results

2.1 Piecewise affine maps and topological entropy

Definition 4.1 *We say that $(X, \mathcal{Z}, f)$ is a piecewise affine map if*

1. $X \subset \mathbb{R}^n$
2. $\mathcal{Z} = \{Z\}$ *is a finite collection of open, pairwise disjoint polytopes such that* $X' := \cup_{Z \in \mathcal{Z}} Z$ *is dense in* X.
3. $f_Z := f|_Z : Z \to X$ *is affine for each* $Z \in \mathcal{Z}$

The maps f_Z are called the affine components of the map f. The linear part of f_Z is denoted f'_Z, and if $x \in Z$ we denote $d_x f = f'_Z$. Let $\mathrm{PAff}(X; X)$ be the set of

piecewise affine maps on X and let $U_f = X' \cap f^{-1}(X') \cap f^{-2}(X') \cap \dots$ be the set of points in X with well-defined infinite orbits. Let $\mathcal{Z}^n$ be the continuity partition of the piecewise affine map f^n. We always assume X to be compact.

Since the maps we consider in this paper have singularities (consult [33, 53]) we must explain what we mean by topological entropy.

Let $U_n = \cap_{k=0}^{n-1} f^{-k}(X')$, then $U_f = \cap_{n\geq 0} U_n$. Take a metric d defining the standard topology on X. Let $d_n^f = \max_{0\leq k<n}(f^k)^* d$, and define $S(d_n^f, \epsilon)$ to be the minimal number of (d_n^f, ϵ)-balls needed to cover U_n. Define

$$h_{\text{top}}(f) = \lim_{\epsilon\to 0} \overline{\lim_{n\to\infty}} \frac{1}{n} \log S(d_n^f, \epsilon)\,.$$

This number is independent of the choice of metric on X and is finite because it is bounded by $d \cdot \sup_x \log(\|d_x f\| + |\mathcal{Z}|)$. If we have a finite cover of U_f by (d_n^f, ϵ)-balls, then concentric $(d_n^f, 2\epsilon)$-balls cover U_n. Therefore $h_{\text{top}}(f)$ equals the (n, ϵ)-entropy $h_{\text{top}}(f|_{U_f})$, which coincides with the (upper=lower by f-invariance of the set U_f) capacity entropy $Ch_{U_f}(f)$ [44]. This bounds from above the topological entropy $h_{U_f}(f)$ of non-compact subsets by Pesin and Pitskel' [45], so that we have $h_{U_f}(f) \leq h_{\text{top}}(f)$. Since we estimate $h_{\text{top}}(f)$ from above, this bounds the other entropy too.

Remark 4.1 *We observe that even in the presence of singularities the property $h_{\text{top}}(f^T) = Th_{\text{top}}(f^T)$ holds for $T \in \mathbb{N}$. The proof uses the fact that for all $\epsilon > 0$ there is a number $\delta(\epsilon) > 0$ such that $B_d(x, \delta(\epsilon)) \cap Z \subset B_{d_T^f}(x, \epsilon) \cap Z$ $\forall x \in Z \subset X'$, where $Z \in \mathcal{Z}$ is a continuity partition (cf. [32]).*

2.2 Singularity entropy

One can also measure the orbit growth of a piecewise affine map through the growth of continuity domains. The singularity entropy of f is

$$H_{\text{sing}}(f) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log |\mathcal{Z}^n|\,.$$

For non-expanding piecewise affine maps $h_{\text{top}}(f) \leq H_{\text{sing}}(f)$ (indeed every (d, ϵ)-ball in an element of $\mathcal{Z}^n$ is a (d_n^f, ϵ)-ball). More generally:

Proposition 4.1 *For $f \in \text{PAff}(X, X)$: $h_{\text{top}}(f) \leq H_{\text{sing}}(f)$.*

Proof. Consider the continuity partition $\mathcal{Z}$ of f. Subdivide it by hyperplanes into the pieces of diameter $\leq \epsilon$. The number of hyperplanes does not exceed $K \cdot d/\epsilon$, $K = \operatorname{diam} X$. Denote the new partition by $\tilde{\mathcal{Z}}$ and its f-iteration by $\tilde{\mathcal{Z}}^n$. Elements of this latter have d_n^f-diameter $\leq \epsilon$.

Let Q be a partition element of $\mathcal{Z}^n$. The first n iterations of all $x \in Q$ experience the same sequence of affine maps $A_1, \dots, A_{n-1}$. Preimages of the added hyperplanes by $\mathrm{Id}, A_1, A_2 \cdot A_1, \dots, A_{n-1} \cdots A_1$ form a subdivision of Q by no more than $N = nKd/\epsilon$ hyperplanes. N hyperplanes subdivide $\mathbb{R}^d$ into no more than $C_d N^d$ pieces. Therefore

$$S(d_n^f, \epsilon) \leq |\tilde{\mathcal{Z}}^n| \leq C_d \Big(\frac{nKd}{\epsilon}\Big)^d \cdot |\mathcal{Z}^n|$$

and the claim follows. □

The reverse inequality is not true in general, but holds if all singular points are unstable ([49] or [35] §4.1):

Proposition 4.2 *Let $f \in \mathrm{PAff}(X, X)$. If $f_Z(x) \neq f_{Z'}(x)$ for all $x \in \partial Z \cap \partial Z'$ with $Z \neq Z'$, then $H_{\mathrm{sing}}(f) \leq h_{\mathrm{top}}(f)$.*

Proof. We claim that there exists $\delta_0 > 0$ such that $d(f(x), f(y)) \geq \delta_0$ whenever $d(x, y) \leq \delta_0$ and x and y are elements of different domains of continuity. Otherwise there exist sequences $\{x_m\} \subset Z$ and $\{y_m\} \subset Z'$ such that $d(x_m, y_m) \to 0$ and $d(f_Z(x_m), f_{Z'}(y_m)) \to 0$. Since X is compact, we can extract a convergent subsequence $x_{m_n} \to z$. The point z lies in $\partial Z \cap \partial Z'$ and $y_{m_n} \to z$. By the continuity of the maps f_Z, $f_{Z'}$ and the metric d we get: $f_Z(z) = f_{Z'}(z)$, which contradicts our assumption.

Our claim implies that a small (d_n^f, ϵ)-ball cannot contain elements from two different domains of continuity for $n > 1$. Thus $S(d_n^f, \epsilon) \geq |\mathcal{Z}^n|$ for $n > 1$ and $\epsilon < \delta_0/2$. □

If we enumerate $\mathcal{Z} = \{Z_1, \dots, Z_N\}$ and let $g : U_f \to \Sigma_N^+$ be the map associating to x the code $(a_0 a_1 \dots)$ of its orbit, i.e. $f^k(x) \in Z_{a_k}$, then $H_{\mathrm{sing}}(f)$ is the topological entropy of the symbolic system $(g(U_f), \sigma_N^+)$.

It is also possible to show that the singularity entropy $H_{\mathrm{sing}}(f)$ coincides with the topological entropy $h_{\mathrm{top}}(\hat{f})$ for generalized polygon exchanges from [27], when f is a piece-wise affine exchange (this follows from Proposition 2.8 of [27] and our Proposition 4.1).

2.3 Expansion rates and multiplicity entropy

Definition 4.2 *The multiplicity of the partition $\mathcal{Z}^n$ at a point $a \in X$ is $\mathrm{mult}(\mathcal{Z}^n, a) = |\{Z \in \mathcal{Z}^n \mid \bar{Z} \ni a\}|$, and the multiplicity of $\mathcal{Z}^n$ is $\mathrm{mult}(\mathcal{Z}^n) = \sup_{a \in X} \mathrm{mult}(\mathcal{Z}^n, a)$. The multiplicity entropy [16] of f is defined as*

$$H_{\mathrm{mult}}(f) = \overline{\lim_{n \to \infty}} \frac{1}{n} \log \mathrm{mult}(\mathcal{Z}^n)\,.$$

Definition 4.3 *For $f \in \mathrm{PAff}(X,X)$ we define*

$$\lambda^+(f) = \overline{\lim_{n\to\infty}} \sup_{x\in U_n} \frac{1}{n} \max_{0\le k\le d} \log \|\Lambda^k d_x f^n\| \,.$$

We also let (the second quantity can equal $-\infty$ for non-invertible maps)

$$\lambda_{\max}(f) = \overline{\lim_{n\to\infty}} \sup_{x\in U_n} \frac{1}{n} \log \|d_x f^n\| \text{ and } \lambda_{\min}(f) = -\overline{\lim_{n\to\infty}} \sup_{x\in U_n} \frac{1}{n} \log \|(d_x f^n)^{-1}\| \,.$$

Theorem 4.3 *For any $f \in \mathrm{PAff}(X;X)$ it holds:*

$$h_{\mathrm{top}}(f) \le \lambda^+(f) + H_{\mathrm{mult}}(f) \,.$$

This result is basically due to Buzzi [16]. However, he only proves it for a special class of strictly expanding maps, and he considers the entropy of coding $H_{\mathrm{sing}}(f)$ instead of $h_{\mathrm{top}}(f)$. Hence we modify his proof. In fact, the proof shows that $H_{\mathrm{sing}}(f) \le \lambda^+(f) + H_{\mathrm{mult}}(f)$ for any piece-wise affine f.

2.4 The spherization and angular expansions

For any k-dimensional submanifold $N^k \subset X$ and $x \in N$ we define the spherical bundle $STN = \{v \in TN : \|v\| = 1\}$, where $\|\cdot\|$ is the Euclidian norm on every $T_xN \subset \mathbb{R}^d$. Let f be non-degenerate, i.e. each affine component is non-degenerate. The spherization of f is defined to be the piecewise smooth map $d_x^{(s)}f : ST_xX \to ST_{f(x)}X$ given at $x \in X'$ by the formula

$$d_x^{(s)}f(v) = \frac{d_x f(v)}{\|d_x f(v)\|} \,.$$

For $x \in \mathrm{Sing}(X) \stackrel{\mathrm{def}}{=} X \setminus X'$ and $v \notin T_x \mathrm{Sing}(X)$ (the tangent cone) we define $d_x^{(s)}f(v) = \lim_{\epsilon\to+0} d_{x+\epsilon v}^{(s)} f(v)$ (this is related to the first pseudo-differential of Tsujii, see §3.2). For other $(x,v) \in STX$ the map is not defined.

The angular expansion of f is exactly the maximal expansion of its spherization.

Definition 4.4 *For a non-degenerate map $f \in \mathrm{PAff}(X,X)$ and $i < d$ we define*

$$\rho_i(f) = \overline{\lim_{n\to\infty}} \frac{1}{n} \sup_{x\in U_n} \max_{0\le k\le i} \sup_{v\in S^{d-1}} \log \|\Lambda^k d_v d_x^{(s)} f^n\|.$$

Note that

$$\rho_i(f) = \overline{\lim_{n\to\infty}} \frac{1}{n} \sup_{N^{i+1}\subset X} \sup_{x\in U_n\cap N^{i+1}} \sup_{v\in ST_xN^{i+1}} \log \left\| \Lambda\, d_v \left[d_x^{(s)} f^n|_{ST_xN^{i+1}} \right] \right\| .$$

(N^{i+1} runs over all local $(i+1)$-dimensional submanifolds of X and we choose the norm $\|\Lambda T\| = \max \|\Lambda^k T\|$).

Informally speaking, the number $\rho_1(f)$ measures the maximal exponential rate with which angles can increase under the map f. Similarly, the number $\rho_i(f)$ measures the maximal rate of expansion of the i-dimensional spherical volume under f. Clearly $\rho_0(f) = 0$ for any $f \in \mathrm{PAff}(X, X)$, and if f is conformal, i.e. all the affine components of f are conformal, then $\rho_i(f) = 0$ for all i.

Theorem 4.4 *$H_{\mathrm{mult}}(f) \leq \sum_{i=1}^{d-1} \rho_i(f)$ for any non-degenerate $f \in \mathrm{PAff}(X; X)$.*

The following corollaries are direct consequences of Theorem 4.4.

Corollary 4.1 *If $f \in \mathrm{PAff}(X; X)$ is conformal, then $h_{\mathrm{top}}(f) \leq \lambda^+(f)$.*

We say that a piecewise affine map is non-expanding if all its affine components are non-expanding, i.e. the eigenvalues of the linear part of each affine component have absolute values not exceeding 1.

Corollary 4.2 *If $f \in \mathrm{PAff}(X; X)$ is conformal non-expanding, then $h_{\mathrm{top}}(f) = 0$.*

It is shown in §4.1 that

$$\rho_i(f) \leq \overline{\lim_{n\to\infty}} \sup_{x\in U_f} \max_{0\leq k\leq i} \frac{1}{n} \log \|\Lambda^k d_x f^n\| - i\lambda_{\min}(f) \leq i\Big(\lambda_{\max}(f) - \lambda_{\min}(f)\Big),$$

This gives the following estimate:

Corollary 4.3 *For a non-degenerate map $f \in \mathrm{PAff}(X; X)$ it holds:*

$$H_{\mathrm{mult}}(f) \leq \frac{d(d-1)}{2}\big(\lambda_{\max}(f) - \lambda_{\min}(f)\big).$$

Hence we see that the topological entropy of a non-degenerate piecewise affine map f can be bounded using only its expansion rates. In fact, we have

$$h_{\mathrm{top}}(f) \leq \lambda^+(f) + \frac{d(d-1)}{2}\big(\lambda_{\max}(f) - \lambda_{\min}(f)\big).$$

It is not clear if the above estimates are optimal, but they show the nature of increase of the topological entropy.

Let us call $f \in \mathrm{PAff}(X; X)$ asymptotically conformal if $\lambda_{\max}(f) = \lambda_{\min}(f)$.

Corollary 4.4 *For asymptotically conformal $f \in \mathrm{PAff}(X; X)$: $h_{\mathrm{top}}(f) \leq \lambda^+(f)$.*

Remark 4.2 *It is essential that in our definition of piece-wise affine maps we consider a finite number of continuity domains. With countable number of domains (this is related to countable Markov chains) the above theorems become wrong. In fact, according to [2] every aperiodic measure preserving transformation can be represented as an interval exchange with countable number of intervals. In particular, there exist piece-wise isometries with infinite number of continuity domains, which have positive topological and metric entropies.*

2.5 Piecewise affine skew products

Definition 4.5 *We say that $S\tilde{\times}T \in \mathrm{PAff}(X \times Y, X \times Y)$ is a piecewise affine skew product if it has the form $f(x, y) = (S(x), T_x(y))$ for some $T_x \in \mathrm{PAff}(Y, Y)$.*

The following results hold for piecewise affine skew products:

Theorem 4.5 *If $S\tilde{\times}T$ is a piecewise affine skew product, then*

$$h_{\mathrm{top}}(S) \leq h_{\mathrm{top}}(S\tilde{\times}T) \leq h_{\mathrm{top}}(S) + H_{\mathrm{mult}}(S) + \sup_{\mathbf{x}} \left(\lambda^+(T_{\mathbf{x}}) + H_{\mathrm{mult}}(T_{\mathbf{x}})\right).$$

where $\mathbf{x} = (x_0, x_1, \dots)$ is an orbit of S in X and $T_{\mathbf{x}}$ is the dynamics along this orbit, i.e. $T^n_{\mathbf{x}} = T_{x_{n-1}} \circ \cdots \circ T_{x_0}$ (see §3.3 for details).

From Theorem 4.5 we can deduce several simple corollaries:

Corollary 4.5 *If $\dim(X) = 1$ and $T_x \in \mathrm{PAff}(Y; Y)$ are non-expanding and conformal for all $x \in X'$, then $h_{\mathrm{top}}(S\tilde{\times}T) = h_{\mathrm{top}}(S)$.*

Corollary 4.6 *Let $X = [0, 1]^d$ and $A \in \mathrm{PAff}(X; X)$ be defined by $x \mapsto Ax \mod \mathbb{Z}^d$ for some $A \in \mathrm{GL}_d(\mathbb{R})$. If $T_x \in \mathrm{PAff}(Y; Y)$ are non-expanding and conformal for all $x \in X'$, then*

$$h_{\mathrm{top}}(A\tilde{\times}T) = h_{\mathrm{top}}(A) = \log \mathrm{Jac}^+ A\,,$$

where

$$\mathrm{Jac}^+ A = \prod_{\lambda \in \mathrm{Sp}(A)} \max\{|\lambda|, 1\}\,.$$

(eigenvalues $\lambda \in \mathrm{Sp}(A)$ are repeated according to their multiplicities).

Corollary 4.7 *Let $\Sigma^+_N = \{1, \dots, N\}^{\mathbb{Z}_{\geq 0}}$ and σ^+_N be the left shift on Σ^+_N. Take $T_1, \dots, T_N \in \mathrm{PAff}(Y; Y)$ to be non-expanding and conformal. Then for the map $f : \Sigma^+_N \times Y \to \Sigma^+_N \times Y$, $(\mathbf{t}, y) \mapsto (\sigma^+_N \mathbf{t}, T_{t_0}(y))$, we have: $h_{\mathrm{top}}(f) = \log N$.*

We can consider more general random maps $\sigma^+_N \tilde{\times} T$ with T depending on more than one (but finite) entries of $\mathbf{t}$. The claim of the theorem still holds.

Remark 4.3 *The class of piecewise affine skew-products of the form*

$$\sigma_N^+ \tilde{\times} T : \Sigma_N^+ \times Y \to \Sigma_N^+ \times Y$$

is physically relevant and appears in the Zhang sandpile model of Self-Organized Criticality [6, 7, 8]. The maps T_i correspond to the avalanches and are contracting [35], though not conformal. So we can get an estimate for the entropy.

In general, we cannot ensure existence of Borel probability invariant measures for piecewise affine maps. In fact, there are examples with no invariant measure at all. In this case the variational principle $h_{\text{top}}(f) = \sup h_\mu(f)$ fails (it can therefore fail even in the presence of invariant measures — take a disjoint union of such a system and an identity). However we can give an estimate for the metric entropy $h_\mu(f)$ whenever an invariant measure exists (we should suppose that it is not supported on singularities, because f is not defined there).

Theorem 4.6 *Let $S\tilde{\times}T$ be a piecewise affine skew product and $T_x \in \text{PAff}(Y;Y)$ be non-expanding for all $x \in X'$. If μ is a $S\tilde{\times}T$-invariant Borel probability measure on $X \times Y$, then*

$$h_{\pi_*\mu}(S) \leq h_\mu(S\tilde{\times}T) \leq h_{\pi_*\mu}(S) + H_{\text{mult}}(S)\,,$$

where $\pi : X \times Y \to X$ is the projection to X.

Corollary 4.8 *Let $A\tilde{\times}T$ be as in Corollary 4.6 with A expanding. If μ is a measure of maximal entropy for $A\tilde{\times}T$ on $X \times Y$, then $\pi_*\mu$ is absolutely continuous with respect to the Lebesgue measure on X.*

3 Proof of the theorems

3.1 Proof of Theorem 4.3

Fix $\epsilon > 0$ and let $T = T(\epsilon) \in \mathbb{N}$ be such that $\text{mult}(\mathcal{Z}^n) \leq \exp((H_{\text{mult}}(f) + \epsilon)n)$ for all $n \geq T$ and

$$\forall x \in U, n \geq T, 0 \leq k \leq d : \ \|\Lambda^k d_x f^n\| \leq \exp((\lambda^+(f) + \epsilon)n)$$

We suppose that $\sqrt{d} + 1 \leq \exp(\epsilon T/d)$. Take $r = r(\epsilon)$ to be compatible with the partition $\mathcal{Z}^T$, i.e. any r-ball intersects maximally $\text{mult}(\mathcal{Z}^T)$ partition elements.

We will prove inductively on l that each $Z \in \mathcal{Z}^{lT}$ can be covered by a family $Q_Z = \{W\}$ satisfying the following properties:

1. $\sum_{Z \in \mathcal{Z}^{lT}} \text{card } Q_Z \leq C_0 \exp((\lambda^+(f) + H_{\text{mult}}(f) + 3\epsilon)lT)$

2. $\operatorname{diam}(f^{lT}(W)) \le r$.

3. $\forall x, y \in W: \ d_l^{f^T}(x,y) < \epsilon$ and $d_{lT}^f(x,y) < \delta(\epsilon)$ with $\delta(\epsilon)$ from Remark 4.1.

The base of induction $l = 0$ is obvious and $C_0 \le |\mathcal{Z}|(\operatorname{diam} X/\min\{r,\epsilon\})^d$.

Take a partition element $W \in Q_Z$ that is used to cover the set $Z \in \mathcal{Z}^{lT}$. By the induction hypothesis it can be continued to cover an element of length $\mathcal{Z}^{(l+1)T}$ in at most $\operatorname{mult}(\mathcal{Z}^T)$ ways. So to cover the cylinders $Z \in \mathcal{Z}^{(l+1)T}$ we make a division of W:

$$W = \bigcup_{i=1}^{\gamma} W_i' \ , \gamma \le \operatorname{mult}(\mathcal{Z}^T)\,.$$

Let $W_i'' = f^{Tl-1}(W_i')$ and $W_i''' = f^T(W_i'')$. By the assumption $\operatorname{diam}(W'') \le r$, but the set W''' may have greater diameter than r. Thus we need to divide the sets W_i''' and pull this refinement back to the partition of sets W_i'.

The image $f^T(W_i'')$ is the image of one affine component of f^T. Let L^T denote the linear part of this affine component. We can assume that L^T is symmetric and take $\{e_k\}$ to be a basis of eigenvectors corresponding to eigenvalues $\lambda_1^T, \dots, \lambda_d^T$. Let $\{v_k\}$ be a basis in the vector subspace corresponding to W_i'''. We can choose this basis to be orthonormal and triangular with respect to $\{e_k\}$. Divide W_i''' by the hyperplanes

$$\psi_j(x) \stackrel{\text{def}}{=} \langle v_j, x\rangle = p\frac{\min\{r,\epsilon\}}{\sqrt{d}}, \quad p \in \mathbb{Z}, \quad j = 1, \dots, d\,.$$

This defines cells $\tilde{W}$ of diameter less than $\min\{r,\epsilon\}$. Since $\psi_j(W''') = \psi_j(L^T(W''))$ has diam $\le |\lambda_i^T| \min\{r,\epsilon\}$, the number of cells $\tilde{W}$ needed to cover W''' is less than or equal to

$$\begin{aligned}(\sqrt{d}+1)^d|\lambda_1^T|^+ \dots |\lambda_d^T|^+ &\le (\sqrt{d}+1)^d \sup_{x\in U_T} \max_{1\le k\le d} \|\Lambda^k d_x f^T\| \\ &\le \exp\big((\lambda^+(f) + 2\epsilon)T\big)\,,\end{aligned}$$

where $|\lambda|^+ = \max\{|\lambda|, 1\}$.

Therefore the total cardinality of the new partition is less than or equal to

$$\begin{aligned}\operatorname{mult}(\mathcal{Z}^T)\exp\big((\lambda^+(f)+2\epsilon)T\big)\exp\big((\lambda^+(f)+H_{\text{mult}}(f)+3\epsilon)lT\big) \\ \le \exp\big((\lambda^+(f)+H_{\text{mult}}(f)+3\epsilon)(l+1)T\big)\,.\end{aligned}$$

The elements of the partition Q_Z have diameter less than ϵ in the metric $d_l^{f^T}$. By Remark 4.1 each partition element has diameter less than some number $\delta(\epsilon) > 0$ in the metric d_{lT}^f. This proves the statement.

3.2 Proof of Theorem 4.4

Define the bundles

$$S^{(k)}TX = \{(x, v_1, \dots, v_k) \,|\, x \in X, v_1 \in ST_xX, v_i \in ST_xX \cap \langle v_1, \dots, v_{i-1}\rangle^{\perp}\}.$$

They form the spherical towers:

$$S^{(d)}TX \xrightarrow{\pi_d} S^{(d-1)}TX \xrightarrow{\pi_{d-1}} \cdots \xrightarrow{\pi_2} S^{(1)}TX = STX \xrightarrow{\pi_1} X$$

with fibers $S^0, S^1, \dots, S^{d-1}$ respectively.

The spherization $d^{(s)}f : STX \to STX$ induces the maps $S^{(k)}f : S^{(k)}TX \to S^{(k)}TX$. Although defined on $S^{(k)}TX'$ they extend over the strata of $\mathrm{Sing}(X)$ as in §2.4 (modulo spherization this corresponds to the differential $f_{x,v_1,\dots,v_{k-1}}(v_k)$ of Tsujii and Buzzi [55, 17]):

$$S^{(k)}f(x, v_1, \dots, v_k) = \lim_{\epsilon_1 \to +0} \cdots \lim_{\epsilon_k \to +0} d^{(s)}_{x+\epsilon_1 v_1 + \cdots + \epsilon_k v_k} f(x, v_1, \dots, v_k) =$$
$$= \left(\lim_{\epsilon_1 \to +0} \cdots \lim_{\epsilon_k \to +0} f\Big(x + \sum\nolimits_1^k \epsilon_i v_i\Big), \lim_{\epsilon_1 \to +0} \cdots \lim_{\epsilon_k \to +0} d^{(s)}_{x+\sum_1^k \epsilon_i v_i} f(v_1), \ \dots \right).$$

In particular, $S^{(d)}f$ is defined everywhere on $S^{(d)}TX$ (but is not continuous).

Let $\mathcal{Z}^n_{(x,v_1,\dots,v_{k-1})}(S^{(k)}f|S^{(k-1)}f)$ be the continuity partition of the piecewise smooth map $(S^{(k)}f)^n$ restricted to the fiber $\pi_k^{-1}(x, v_1, \dots, v_{k-1})$. Define

$$H_{\mathrm{sing}}(S^{(k)}f|S^{(k-1)}f) = \overline{\lim_{n\to\infty}} \sup_{(x,v_1,\dots,v_{k-1})} \frac{1}{n} \log |\mathcal{Z}^n_{(x,v_1,\dots,v_{k-1})}(S^{(k)}f|S^{(k-1)}f)| \,.$$

Define $\mathrm{mult}(\mathcal{Z}^n_{(x,v_1,\dots,v_{k-1})}(S^{(k)}f|S^{(k-1)}f), v_k)$ as the multiplicity of the partition $\mathcal{Z}^n_{(x,v_1,\dots,v_{k-1})}(S^{(k)}f|S^{(k-1)}f)$ at the point $(x, v_1, \dots, v_k) \in S^{(k)}TX$. Clearly $\mathrm{mult}(\mathcal{Z}^n_{(x,v_1,\dots,v_{k-1})}(S^{(k)}f|S^{(k-1)}f), v_k) = |\mathcal{Z}^n_{(x,v_1,\dots,v_k)}(S^{(k+1)}f|S^{(k)}f)|$. So if we define the conditional multiplicity entropy $H_{\mathrm{mult}}(S^{(k)}f|S^{(k-1)}f)$ as the growth rate of the above multiplicity (as in definition 4.2) we get:

$$H_{\mathrm{mult}}(S^{(k)}f|S^{(k-1)}f) = H_{\mathrm{sing}}(S^{(k+1)}f|S^{(k)}f).$$

In particular, we have: $H_{\mathrm{mult}}(f) = H_{\mathrm{sing}}(S^{(1)}f|f)$. In Theorem 4.3 we dealt with piece-wise affine maps. The spherizations $S^{(k)}f$ are not piece-wise affine, but they are spherically piece-wise affine along fibers, meaning that they map spherical polygons into spherical polygons. In addition, the differentials of $S^{(k)}f$ are uniformly continuous on continuity domains. Thus the arguments from the proof of Theorem 4.3 apply and we can estimate the conditional entropy via the fiber expansion rate and the conditional multiplicity entropy:

$$H_{\mathrm{sing}}(S^{(1)}f|f) \le \rho_{d-1}(f) + H_{\mathrm{mult}}(S^{(1)}f|f)\,.$$

Doing its once more for $H_{\mathrm{mult}}(S^{(1)}f|f) = H_{\mathrm{sing}}(S^{(2)}f|S^{(1)}f)$ we get:

$$H_{\mathrm{sing}}(S^{(2)}f|f) \leq \rho_{d-2}(f) + H_{\mathrm{mult}}(S^{(2)}f|f)\,.$$

Applying the same argument $d-1$ times yields: $H_{\mathrm{mult}}(f) \leq \rho_{d-1}(f) + \cdots + \rho_1(f)$.

3.3 Proof of Theorem 4.5

The inequality $h_{\mathrm{top}}(S) \leq h_{\mathrm{top}}(S\tilde{\times}T)$ is obvious since S is a quotient of $S\tilde{\times}T$. We will now prove the upper bound for $h_{\mathrm{top}}(S\tilde{\times}T)$. Denote $d_X = \dim X$.

Let Γ be the set of all sequences $((x_0,z_0),(x_1,z_1),(x_2,z_2),\dots) \in (X \times \mathbb{R}^{d_X})^{\mathbb{Z}_{\geq 0}}$ satisfying $S(x_i+\epsilon z_i) \to x_{i+1}$, $d_{x_i+\epsilon z_i}S(z_i) \to z_{i+1}$, $x_i + \epsilon z_i \in X'$ as $\epsilon \to +0$. Define $T_{(x,z)} = \lim_{\epsilon\to+0} T_{x+\epsilon z}$ and let $\mathcal{Y}^{n,\mathbf{x}}$ be the collection of non-empty sets

$$Y_{(x_0,z_0)} \cap T^{-1}_{(x_0,z_0)}(Y_{x_1,z_1}) \cap \cdots \cap (T_{(x_{n-1},z_{n-1})} \circ \dots T_{(x_0,z_0)})^{-1}(Y_{(x_n,z_n)})\,,$$

for $\mathbf{x} = ((x_0,z_0),(x_1,z_1),(x_2,z_2),\dots) \in \Gamma$, where the sets $Y_{(x_i,z_i)}$ are the continuity domains of the maps $T_{(x_i,z_i)}$. Let

$$P^n(x) = \{Y \in \mathcal{Y}^{n,\mathbf{x}} \,|\, \mathbf{x} = ((x_0,z_0),(x_1,z_1),\dots) \in \Gamma, x_0 = x\}$$

be the continuity partition of T on the fiber $\pi^{-1}(x)$ iterated for n steps along all possible S-orbits starting from $x \in X$. Clearly

$$|P^n(x)| \leq \mathrm{mult}(S^n,x) \sup_{\mathbf{x}\in\Gamma} |\mathcal{Y}^{n,\mathbf{x}}|\,.$$

To simplify notations we will write elements of Γ as $\mathbf{x} = (x_0,x_1,x_2,\dots)$, where x_i consists of a point in X and a vector in $\mathbb{R}^{d_X}$. If $x_i \in X'$ the vector z_i is not essential. Denote $T^n_{\mathbf{x}} = T_{x_{n-1}} \circ \cdots \circ T_{x_0}$ for $\mathbf{x} = (x_0,x_1,x_2,\dots) \in \Gamma$ and let

$$H_{\mathrm{mult}}(T_{\mathbf{x}}) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathrm{mult}(T^n_{\mathbf{x}})\,.$$

Lemma 4.7 *For a piecewise affine skew product $S\tilde{\times}T$ it holds:*

$$h_{\mathrm{top}}(S\tilde{\times}T) \leq h_{\mathrm{top}}(S) + H^{\mathrm{fiber}}_{\mathrm{sing}}(T|S)\,,$$

where

$$H^{\mathrm{fiber}}_{\mathrm{sing}}(T|S) = \sup_{x\in X} \overline{\lim_{n\to\infty}} \frac{1}{n} \log |P^n(x)|\,.$$

Remark 4.4 *The statement of the lemma is similar to Bowen's Theorem 17 [12], but the direct generalization fails, see Example 2.*

Proof. Assume at first that the maps T_x are non-expanding for all $x \in X'$. This assumption is not crucial, but it simplifies the proof.

Let $\epsilon > 0$ be arbitrary. Denote $a = H^{\text{fiber}}_{\text{sing}}(T|S)$, and fix $\alpha > 0$ and $m_\alpha = [1/\alpha] \in \mathbb{N}$. For all $x \in X$ we let

$$n_\alpha(x) = \min\{n \geq m_\alpha \,|\, \frac{1}{n} \log |P^n(x)| \leq a + \alpha\}\,.$$

Let $\mathcal{Z}$ be the continuity partition for the map S and $\mathcal{Z}^n$ correspond to S^n. The multiplicity functions $\text{mult}(\mathcal{Z}, x)$ and $|P^n(x)|$ are upper semi-continuous. Consequently, the same is true for the function $n_\alpha(x)$. So $n_\alpha := \sup_{x \in X} n_\alpha(x)$ is finite. Then we have $0 < m_\alpha \leq n_\alpha(x) \leq n_\alpha < +\infty$ for all $x \in X$.

Let $r_\alpha > 0$ be compatible with all the partitions $\mathcal{Z}^n$, $m_\alpha \leq n \leq n_\alpha$, i.e. any ball of radius r_α can intersect at most $\text{mult}(\mathcal{Z}^n)$ different elements of the partition $\mathcal{Z}^n$ for $m_\alpha \leq n \leq n_\alpha$. We can assume that $r_\alpha < \epsilon$.

Let E_n denote a (n, r_α)-spanning set of minimal cardinality for S in X. For each $x \in X$ consider the singularity partition $\{Z \cap (\{x\} \times Y) \,|\, Z \in P^{n_\alpha(x)}(x)\}$. Each element of this partition is an open polytope in the fiber $\{x\} \times Y$, and hence we can subdivide the partition so that each element has diameter no greater than ϵ. Denote the resulting partition of $\{x\} \times Y$ by F_x. The refinement can be done in such a way that $|F_x| \leq C_0/\epsilon^d \cdot |P^{n_\alpha}(x)|$ for some constant $C_0 \in \mathbb{R}_+$.

For $x \in X$ we let $t_0(x) = 0$ and define recursively

$$t_{k+1}(x) = t_k(x) + |F_{S^{t_k(x)}(x)}|\,.$$

Let $q(x) = \min\{k > 0 \,|\, t_{k+1}(x) \geq n\}$. We will denote $q = q(x)$. For $x \in E_n$, $z_0 \in F_x$, $z_1 \in F_{S^{t_1(x)}(x)}, \ldots, z_q \in F_{S^{t_q(x)}(x)}$ denote

$$V(x; z_0, \ldots, z_q) = \big\{w \in X \times Y \,\big|\, d((S\tilde{\times}T)^{t+t_k(x)}(w), (S\tilde{\times}T)^t(z_k)) < 2\epsilon \\ \forall\, 0 \leq t \leq |F_{S^{t_{k-1}(x)}(x)}|,\, 0 \leq k \leq q(x)\big\}\,.$$

Then $\cup_{x, z_0, \ldots, z_q} V(x; z_0, \ldots z_q) = U_n \times Y$ and for any $(n, 4\epsilon)$-separating set $K \subset X \times Y$ for $S\tilde{\times}T$ we have $|K \cap V(x; z_0, \ldots z_q)| \leq 1$. Thus if K is a maximal $(n, 4\epsilon)$-set, then the cardinality of K is bounded by the number of ways we can choose $x, z_0, \ldots, z_q$ modulo the partitions specified above. For fixed $x \in X$ the number Π_x of such admissible combinations satisfies

$$\Pi_x \leq \prod_{k=0}^{q(x)} |F_{S^{t_k(x)}(x)}|\,.$$

Since $q(x) \leq n/m_\alpha$ we have:

$$\begin{aligned}
\log \Pi_x &\leq (q(x)+1)\log\frac{C_0}{\epsilon^d} + \sum_{k=0}^{q(x)} \log|P^{n(S^{t_k(x)}(x))}(S^{t_k(x)}(x))| \\
&\leq \frac{n+m_\alpha}{m_\alpha}\log\frac{C_0}{\epsilon^d} + (a+\alpha)\sum_{k=0}^{q(x)} n(S^{t_k(x)}(x))) \\
&\leq \frac{n+m_\alpha}{m_\alpha}\log\frac{C_0}{\epsilon^d} + (a+\alpha)(n+n_\alpha)
\end{aligned}$$

Let $Q(S\tilde{\times}T, n, 4\epsilon)$ denote the cardinality of a maximal $(n, 4\epsilon)$-separating set for $S\tilde{\times}T$ in $X \times Y$. We have that

$$\frac{1}{n}\log Q(S\tilde{\times}T, n, 4\epsilon) \leq \frac{1}{n}\log|E_n| + \Big(\frac{1}{m_\alpha} + \frac{1}{n}\Big)\log\frac{C_0}{\epsilon^d} + (a+\alpha)(1+\frac{n_\alpha}{n})\,.$$

This yields

$$\overline{\lim_{n\to\infty}}\,\frac{1}{n}\log Q(S\tilde{\times}T, n, 4\epsilon) \leq h_{\text{top}}(S) + \frac{1}{m_\alpha}\log\frac{C_0}{\epsilon^d} + a + \alpha\,.$$

Since the left hand side does not depend on α we can let $\alpha \to 0$. Then $m_\alpha \to \infty$ and we get

$$\overline{\lim_{n\to\infty}}\,\frac{1}{n}\log Q(S\tilde{\times}T, n, 4\epsilon) \leq h_{\text{top}}(S) + a\,.$$

Finally let $\epsilon \to 0$.

If the maps T_x have expansions, we may need to make a better refinement of the partitions F_x. But this gives a sub-exponential increase of Π_x (see proof of Proposition 4.1) and so does not change the inequality. □

Lemma 4.8 $\forall \mathbf{x} \in \Gamma: \ \overline{\lim_{n\to\infty}}\,\frac{1}{n}\log|\mathcal{Y}^{n,\mathbf{x}}| \leq \lambda^+(T_\mathbf{x}) + H_{\text{mult}}(T_\mathbf{x})\,.$

The proof of Lemma 4.8 is similar to the proof of Theorem 4.3 and will be omitted.

Combining Lemmata 4.7 and 4.8 we obtain:

$$\begin{aligned}
h_{\text{top}}(S\tilde{\times}T) &\leq h_{\text{top}}(S) + \sup_{\mathbf{x}\in\Gamma}\overline{\lim_{n\to\infty}}\,\frac{1}{n}\log\Big[\text{mult}(S^n, x_0)|\mathcal{Y}^{n,\mathbf{x}}|\Big] \\
&\leq h_{\text{top}}(S) + H_{\text{mult}}(S) + \sup_{\mathbf{x}\in\Gamma}\overline{\lim_{n\to\infty}}\,\frac{1}{n}\log|\mathcal{Y}^{n,\mathbf{x}}| \\
&\leq h_{\text{top}}(S) + H_{\text{mult}}(S) + \sup_{\mathbf{x}\in\Gamma}\Big(\lambda^+(T_\mathbf{x}) + H_{\text{mult}}(T_\mathbf{x})\Big)\,.
\end{aligned}$$

3.4 Proof of Theorem 4.6

Let μ be an f-invariant Borel probability measure on $X \times Y$. Denote the projection of μ to X by $\mu_X = \pi_* \mu$ and let $\{\nu_x\}$ be the canonical family of conditional measures on the fibers $\pi^{-1}(x)$. By the generalized Abramov-Rokhlin formula [11] (Bogenschütz and Crauel removed restrictions on the the maps S and T_x in the original formula [1]) we have:

$$h_\mu(S\tilde{\times}T) = h_{\mu_X}(S) + h_\mu(T|S)\,,$$

where

$$h_\mu(T|S,\xi) = \lim_{n\to\infty} \frac{1}{n} \int H_{\nu_x}\Big(\bigvee_{k=0}^{n-1} (T_{S^{k-1}(x)} \circ \cdots \circ T_x)^{-1}(\xi)\Big)\, d\mu_X(x)\,,$$

for a measurable partition ξ of Y and $h_\mu(T|S) = \sup_\xi h_\mu(T|S,\xi)$, where the supremum is taken over all finite measurable partitions ξ with finite entropy and the refinement $\bigvee_{k=0}^{n-1}(T_{S^{k-1}(x)} \circ \cdots \circ T_x)^{-1}(\xi)$ is understood with respect to all orbits $(x, Sx, \ldots, S^{k-1}x)$ starting at x (as in §3.3).

For $\epsilon > 0$ we choose ξ of $\mathrm{diam}(\xi) \le \epsilon$ such that $h_\mu(T|S) \le h_\mu(T|S,\xi) + \epsilon$. Clearly

$$\frac{1}{n} H_{\nu_x}\Big(\bigvee_{k=0}^{n-1} (T_{S^{k-1}(x)} \circ \cdots \circ T_x)^{-1}(\xi)\Big) \le \frac{1}{n} \log \mathrm{mult}(S^n, x)$$
$$+ \sup_{\mathbf{x}\in\Gamma_x} \frac{1}{n} H_{\nu_x}\Big(\bigvee_{k=0}^{n-1} (T_{(x_{n-1},z_{n-1})} \circ \cdots \circ T_{(x_0,z_0)})^{-1}(\xi)\Big)\,.$$

where $\Gamma_x \subset \Gamma$ is the set of sequences $((x_0,z_0),(x_1,z_1)\ldots)$ with $x_0 = x$. If the maps T_x are non-expanding, then the last term μ_X-uniformly tends to zero, since by Ruelle-Margulis inequality [32] $h_{\nu_x}(T_{\mathbf{x}}) = \inf_\xi h_{\nu_x}(T_{\mathbf{x}},\xi)$ vanishes, where

$$h_{\nu_x}(T_{\mathbf{x}},\xi) = \lim_{n\to\infty} \frac{1}{n} H_{\nu_x}\Big(\bigvee_{k=0}^{n-1} (T_{(x_{n-1},z_{n-1})} \circ \cdots \circ T_{(x_0,z_0)})^{-1}(\xi)\Big).$$

In fact, our framework is more general, but the standard proof of Ruelle inequality as e.g. in [39] works here. We should only indicate why singularities do not contribute to the entropy. Since the maps T_x are non-expanding, the image $T^n_{\mathbf{x}}(\xi_\alpha)$ of a partition element ξ_α meets $< C_0$ (independent of ϵ) elements of the partition ξ unless the iteration enters ϵ-neighborhood of the singularities, in which case we should subdivide the partition element into no more than $K = \max_x |P(x)|$ pieces, $P(x)$ being the continuity partition of T_x, and iterate them. Therefore there are $< C_0 K^{s(n)}$ elements of $(T^n_{\mathbf{x}})^{-1}(\xi)$ meeting ξ_α, where $s(n)$ is the number of entrances of ξ_α into the ϵ-neighborhood of singularities.

By Kac's theorem this happens with frequency $\lim \frac{s(n)}{n}$ equal to the measure $\varrho(\epsilon)$ of this ϵ-neighborhood, which tends to zero as $\epsilon \to 0$. Thus the contribution to the entropy $h_{\nu_x}(T_{\mathbf{x}}, \xi)$ does not exceed $\varrho(\epsilon) \log K$.

Therefore $h_\mu(T|S) \le H_{\text{mult}}(S) + \epsilon + \varrho(\epsilon) \log K$. Let $\epsilon \to 0$.

4 Proof of the corollaries

Corollaries 4.1, 4.2 and 4.7 are obvious. Corollaries 4.3, 4.4 follow from the following section. Corollary 4.5 is implied by the fact that $H_{\text{mult}}(S) = 0$ if $\dim(X) = 1$ and Corollary 4.8 follows from our Theorem 4.6 and Theorem 3 of Buzzi [16].

4.1 Estimate for the angles expansion rates

Define

$$\lambda^+_{[i]}(f) = \overline{\lim_{n\to\infty}} \sup_{x\in U_n} \max_{0\le k\le i} \frac{1}{n} \log \|\Lambda^k d_x f^n\|.$$

Obviously $\lambda^+_{[i]}(f) \le i\lambda^+_{[1]}(f) = i \cdot \lambda_{\max}(f)$.

Proposition 4.9 *The following estimate holds:* $\rho_i(f) \le \lambda^+_{[i]}(f) - i\lambda_{\min}(f)$.

Proof. Let us first calculate the differential of the spherical transformation ${}^{(s)}A : S^{d-1} \to S^{d-1}$, ${}^{(s)}A(x) = \dfrac{Ax}{\|Ax\|}$, corresponding to linear $A : \mathbb{R}^n \to \mathbb{R}^n$.

Lemma 4.10 *If* $A \in \mathrm{GL}(\mathbb{R}^d)$, *then* $d\big({}^{(s)}Ax\big)(v) = P^\perp_w\left(\dfrac{Av}{\|Ax\|}\right)$, *where* $P^\perp_w$ *is the orthogonal projection along* $w = {}^{(s)}A(x)$.

In fact, $d\|Ax\| = \langle {}^{(s)}Ax, d(Ax)\rangle$ and so

$$d\big({}^{(s)}Ax\big)(v) = \frac{Av}{\|Ax\|} - {}^{(s)}A(x) \cdot \left\langle {}^{(s)}A(x), \frac{Av}{\|Ax\|}\right\rangle.$$

From this lemma we get: $\|d\big({}^{(s)}Ax\big)(v)\| \le \dfrac{\|Av\|}{\|Ax\|} \le C \cdot \dfrac{\max|\operatorname{Sp}(A)|}{\min|\operatorname{Sp}(A)|}$ for $\|x\| = \|v\| = 1$, where the constant C depends on the eigenbasis of A (in non semi-simple case – normal basis) only. Since this eigenbasis is the same for all iterates of A and we have a finite number of pieces in f, we get for all v: $\|d_v d^{(s)}_x f^n\| \le C_n \cdot \dfrac{\max|\operatorname{Sp}(d_x f^n)|}{\min|\operatorname{Sp}(d_x f^n)|}$ (with sub-exponentially growing C_n).

Therefore the maximal vertical (spherical) Lyapunov exponent for the map $d^{(s)}f$ at the point $(x, v) \in STX$, i.e. $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \|d_v d^{(s)}_x f^n\|$, does not exceed the difference

$\overline{\chi}_{\max}(x) - \underline{\chi}_{\min}(x)$ ($\overline{\chi}_i(x), \underline{\chi}_i(x)$ are usual upper resp. lower Lyapunov exponents of f). This latter quantity does not exceed the difference of uniform Lyapunov exponents $\lambda_{\max}(f) - \lambda_{\min}(f)$.

Similarly, we have: $\|\Lambda^k d\big({}^{(s)}Ax\big)\| \leq \dfrac{\|\Lambda^k A\|}{\|Ax\|^k}$ and

$$\|\Lambda^k d_v d_x^{(s)} f^n\| \leq C_n \cdot \frac{\max\{|\lambda_1 \cdots \lambda_k| \,:\, \lambda_j \in \mathrm{Sp}(d_x f^n)\}}{\min|\mathrm{Sp}(d_x f^n)|^k}$$

with $\overline{\lim\limits_{n\to\infty}} \frac{1}{n} \log C_n = 0$ (we take repeated eigenvalues, so the numerator equals $|\lambda_1 \cdots \lambda_k|$ if $|\lambda_1| \geq \ldots |\lambda_k| \geq \ldots |\lambda_d|\}$), whence the claim. □

Note that $\rho_i(f)$ is conformally invariant in the cocycle sense: The cocycle $\mathcal{A} : X \to GL(\mathbb{R}^n)$, $x \mapsto d_x f$, can be changed by any cocycle $\alpha : X \to \mathbb{R}$, $\mathcal{A} \mapsto \alpha \cdot \mathcal{A}$. Then the Lyapunov-type characteristics $\rho_i(f)$ do not change.

However if the cocycle α has different upper and lower Lyapunov exponents, then the quantity $\lambda_{\max}(f) - \lambda_{\min}(f)$ used in the bound is not invariant.

4.2 Proof of Corollary 4.6

It was shown by Buzzi [16] that $H_{\mathrm{mult}}(A) = 0$. This was proven for strictly expanding maps, but the proof extends literally for any non-degenerate A. Thus the bound from above follows from Theorem 4.3.

Let's prove that $h_{\mathrm{top}}(A) \geq \mathrm{Jac}^+ A$. Assume A is semi-simple. Let λ be an eigenvalue with $|\lambda| > 1$ and v the corresponding unit eigenvector. We divide $[0,1]^d$ into domains $k\epsilon/|\lambda|^n \leq \langle x, v\rangle \leq (k+1)\epsilon/|\lambda|^n$. A (d_n^f, ϵ)-ball intersects no more than two such domains, the total number of which is less than $\sqrt{d} \cdot |\lambda|^n/2\epsilon$.

The same holds for other $\lambda \in \mathrm{Sp}(A)$, so the number of (d_n^f, ϵ)-balls to cover $[0,1]^d$ is at least $C_0(\sqrt{d}/2\epsilon)^m(\mathrm{Jac}^+ A)^n$, where m is the number of eigenvalues with absolute value greater than 1 and C_0 some n-independent constant.

If $A \in \mathrm{GL}_d(\mathbb{R})$ is not semi-simple, the estimates change sub-exponentially, implying the same result. Note that the formula of the corollary holds true even in the case, when A is degenerate, though arguments should be modified.

5 Examples

Let us demonstrate some of the angular expansion effects.

Example 1: Let X be a triangle with vertices in $a = (-1, 0)$, $c = (1, 0)$ and $f = (0, 1)$. Divide this triangle in two by taking X_1 to be the left triangle with vertices $a = (-1, 0)$, $b = (0, 0)$ and $f = (0, 1)$, see Figure 4.1.

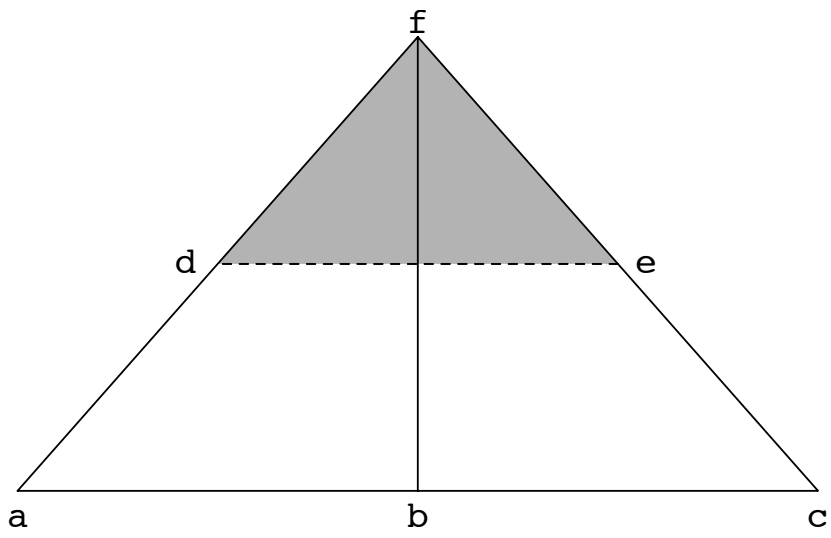


Figure 4.1: Shows the domains of continuity and their image in Example 1.

Let X_2 be the right triangle with vertices $b = (0,0)$, $c = (1,0)$ and $f = (0,1)$. Let X be compact, i.e. the sides are contained in X, and let X_1 and X_2 be open. Then $\mathcal{Z} = \{X_1, X_2\}$ is a finite collection of open disjoined polytopes in X, and $X_1 \cup X_2$ is dense in X.

We define a map S on $X' = X_1 \cup X_2$ by the formula

$$S(x) = \begin{cases} A_1 x + B_1 \text{ if } x \in X_1 \\ A_2 x + B_2 \text{ if } x \in X_2 \end{cases} \quad \text{where } A_1 = \frac{1}{2}\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}, \ A_2 = \frac{1}{2}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

and $B_1 = (1/2, 1/2)$, $B_2 = (-1/2, 1/2)$. This maps both X_1 and X_2 to the triangle with vertices d, e and f. Observe that $S^n(x)$ tends to the point $(0,1)$ with exponential speed for all $x \in U_f$. This implies that $h_{\text{top}}(S) = 0$. However the multiplicity of $\mathcal{Z}^n$ at the point $f = (0,1)$ is 2^n, whence $H_{\text{mult}}(S) = \log 2$.

It is also easy to see that $H_{\text{sing}}(S) = \log 2$. The eigenvalues of A are $\frac{1}{2}, 1$, so the map is non-expanding and $\lambda^+(f) = 0$. Changing $S(x) \mapsto \frac{1}{2}(S(x) + f)$ we obtain a strictly contracting piecewise affine map with positive H_{sing} and H_{mult}.

The growth of multiplicity is produced by angular expansion: on $ST_f X \simeq S^1$ the spherization is conjugated to the map $\theta \mapsto 2\theta$, whence $\rho(S) = \log 2$.

Example 2: The map in Example 1 can be modified to obtain positive topological entropy. We give here an example of a piecewise affine non-expanding map with positive topological entropy in dimension 3, but it is also possible to construct such an example in dimension 2, see [37].

Let $Y = [0,1]$ and let $S : X_1 \cup X_2 \to X$ be as in Example 1. For each $x \in X_1 \cup X_2$ we take a piecewise affine map $T_x \in \mathrm{PAff}(Y,Y)$. For $x \in X_1$ we let $T_x = Id_Y$ and for $x \in X_2$ we let T_x be the interval exchange

$$T_x(y) = \begin{cases} y + 1/2 & \text{if } y \in (0, 1/2), \\ y - 1/2 & \text{if } y \in (1/2, 1). \end{cases}$$

The map $f = S\tilde{\times}T \in \mathrm{PAff}(X \times Y, X \times Y)$ has three domains of continuity Z_1, Z_2 and Z_3. These domains and the images are shown in Figure 2.

As in Example 1 the cardinality of the continuity partition of S grows like 2^n, but in this example we see that if $x_1, x_2 \in X$ are elements of different continuity domains for S, then $d(f(x_1, y), f(x_2, y)) \geq 1/2$ for all $y \in Y$. Hence the number of (d_n^f, ϵ)-balls needed to cover $X \times Y$ is at least $\mathcal{Z}^n$, where $\mathcal{Z}^n$ is the continuity partition of f^n. This implies $h_{\mathrm{top}}(f) \geq \log 2$. The opposite inequality follows from Theorem 4.4, whence $h_{\mathrm{top}}(f) = \log 2$.

This example has several interesting aspects. First of all it is an example of a non-expanding piecewise affine map with positive entropy. We can easily modify f to make it strictly contracting without changing its topological entropy.

A second important point is that $h_{\mathrm{top}}(S) = h_{\mathrm{top}}(T_x) = 0$ for all $x \in X'$, so the the entropy of the skew product $S\tilde{\times}T$ exceeds the combined entropy of its factors. This shows that the term $H_{\mathrm{mult}}(S)$ on the right hand side of the inequality in Theorem 4.5 cannot be removed. Also this justifies Remark 4.4.

Example 3: Let us consider a system with continuity domains as in Figure 4.3. The dynamics f is piece-wise affine and is given by the rules:

$$ABC \longrightarrow AEF, \quad ADC \longrightarrow AFE, \quad BDFE \longrightarrow BDFE,$$

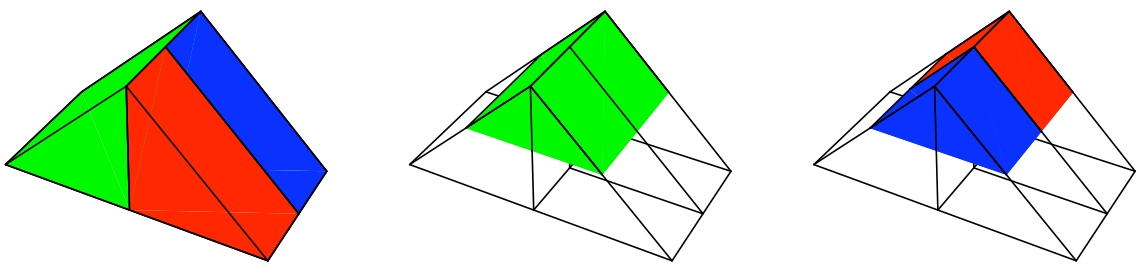


Figure 4.2: The left figure shows the three domains of continuity for the skew-product $S\tilde{\times}T$ in Example 2. Two right figures show the images of the continuity domains.

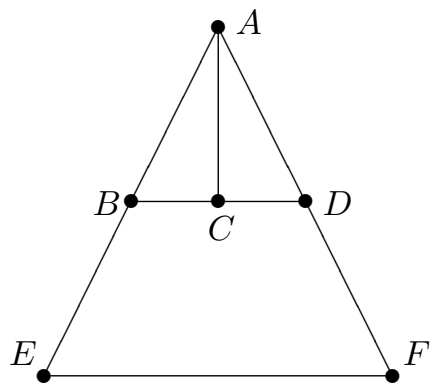


Figure 4.3: Example of a system with zero Lyapunov exponents and positive entropy.

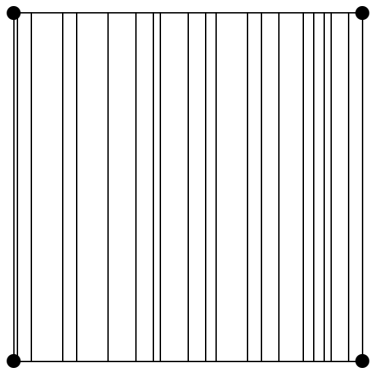


Figure 4.4: A countable piece-wise isometry consisting of random vertical shifts.

the last map being the identity. Since every point eventually comes into the domain $BDFE$, all the Lyapunov exponents vanish. But the multiplicity of the point A is $\log 2$ and this easily yields $h_{\text{top}}(f) = \log 2$.

This shows that in the estimates of the main theorems we cannot change the difference $\lambda_{\max}(f) - \lambda_{\min}(f)$ to the maximal difference of upper and lower Lyapunov exponents $\sup_{x \in U_f} \max_{i,j}(\overline{\chi}_i(x) - \underline{\chi}_j(x))$. However we suggest that the estimate in Theorem 4.3 can be refined by changing $\lambda^+(f)$ to the maximal sum of positive upper Lyapunov exponents $\sup_{x \in U_f} \sum \overline{\chi}_i^+(x) \leq \lambda^+(f)$.

Example 4: Consider the following map f of T^2 to itself. We represent the torus as a glued square, which is partitioned into countable number of rectangles $\Pi_i = I_i \times [0,1]$. We define $f(x,y) = (x + \alpha, y + \beta_i)$ if $x \in I_i$. Thus the map is a piece-wise isometry with countable number of continuity domains Π_i, see figure 4.4.

The number α is chosen irrational. The intervals I_i (with their lengths $l_i = |I_i|$) and the shift lengths β_i are supposed to be sufficiently generic. The value of $h_{\text{top}}(f)$

(note that singularity entropy $H_{\text{sing}}(f)$ does not have sense here) depends on the speed of convergence of the series $\sum_{i=1}^{\infty} l_i$.

Consider, for instance, the case of rapid convergence, when l_i decrease exponentially or at least polynomially, namely $l_i \leq Ci^{-r}$ for some $r > 1$ and $C \in \mathbb{R}_+$. Then the frequency with which an interval of length ϵ meets a singularity of the base $\cup \partial I_i \subset S^1$ under rotation by the angle α is at most $p_\epsilon \sim \epsilon^{\frac{r-1}{r}}$ (for exponential convergence $p_\epsilon \sim \epsilon \log \frac{1}{\epsilon}$). The number of (d_n^f, ϵ)-balls to cover the torus satisfies: $S(d_n^f, \epsilon) \leq c \cdot (\frac{1}{\epsilon})^{p_\epsilon \cdot (n+2)}$. Consequently $h_{\text{top}}(f) = 0$.

On the other hand, if the series $\sum_{i=1}^{\infty} l_i$ converges slowly, then the entropy may become positive and even infinite. For example, if $l_i \sim 1/i \log i (\log \log i)^2$, then the frequency with which an ϵ-interval meets $[1/\epsilon]$ different intervals I_i under rotation by the angle α has asymptotic $\sigma_\epsilon \sim 1/\log \log \frac{1}{\epsilon}$. Thus choosing the shifts β_i and geometry of the decomposition $I = \cup I_i$ appropriately we may arrive to $S(d_n^f, \epsilon) \sim (\frac{1}{\epsilon})^{\sigma_\epsilon \cdot (n+2)}$, which yields $h_{\text{top}}(f) = \infty$ (note that f is a skew-product with vanishing entropies of the base and the fibers).

Note that $h_{\text{top}}(f) = 0$ for a piecewise isometry f with finite number of continuity domains [17] and the same holds for conformal non-expanding maps [35] (and Corollary 4.2). For infinite number of domains this fails.

Example 5: In Corollary 4.4 we stated that if a piecewise affine map f is asymptotically conformal, i.e. $\lambda_{\max}(f) = \lambda_{\min}(f)$, then $H_{\text{mult}}(f) = 0$. This result does not generalize to piecewise smooth maps. This is shown by the following example due to Buzzi [17]:

Let $X = [-1, 1]^2$ and define $f : X \to X$ by

$$(x, y) \mapsto (x/2, y/2 - \operatorname{sgn}(y) x^2)\,.$$

We observe that

$$\operatorname{Jac}(f) = \begin{bmatrix} 1/2 & 0 \\ -2x \operatorname{sgn}(y) & 1/2 \end{bmatrix} \qquad \text{for } y \neq 0\,,$$

so $\lambda_{\max}(f) = \lambda_{\min}(f) = 1/2$. However it is easy to verify that multiplicity of the origin grows like 2^n, whence $H_{\text{mult}}(f) = \log 2$.

Chapter 5

A piecewise affine contracting map with positive entropy

Joint with B. Kruglikov

Abstract: We construct the simplest chaotic system with a two-point attractor on the plane.

Reference for this paper: B. Kruglikov, M. Rypdal, *A piece-wise affine contracting map with positive entropy*, Discrete and Continuous Dynamical Systems **16**, 393 (2006)

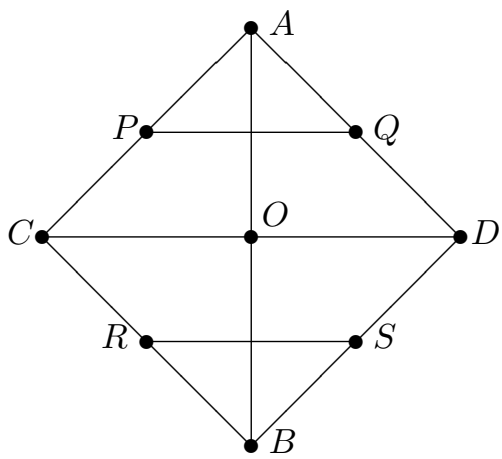
If $f : X \to X$ is an isometry of the metric space (X, d), then the topological entropy vanishes: $h_{\text{top}}(f) = 0$ (for definitions and notations, consult e.g. [32]).

This follows from the fact that the iterated distance $d_n^f = \max\limits_{0 \le i < n} (f^i)^*(d)$ equals d. If f is distance non-increasing, the same equality holds, and again $h_{\text{top}}(f) = 0$.

Whenever f can have discontinuities of some tame nature so that f is piece-wise continuous, even the isometry result becomes difficult. In dimension two for invertible maps, it was proven by Gutkin and Haydn [27]. In arbitrary dimension, Buzzi proved that piece-wise affine isometries have zero topological entropy [17].

In the same paper after the theorem (remark 4), it is claimed that the result holds for arbitrary piece-wise (non-strictly) contracting maps. This latter is wrong, however, and the goal of this note is to present a counter-example.

Example: Let X be a rhombus $ADBC$ with vertices $(\pm 1, 0), (0, \pm 1)$; see the figure below. Let O be its center and P, Q, R, S be on the sides as is shown.



Let f be partially defined on the rhombus, namely let it be defined on the interior of four big triangles forming the rhombus. These triangles are bijectively mapped by f as follows:

$$ACO \longrightarrow APQ, \quad ADO \longrightarrow BRS, \quad BCO \longrightarrow AQP, \quad BDO \longrightarrow BSR.$$

Thus the piece-wise affine map is defined.

If P, Q and R, S are middle-points of the intervals AC, AD and BC, BD, then the map is not strictly contracting. But if they are closer to the vertices A and B respectively than to C and D, then f is strictly contracting. In any case, the attractor of the system is the two-point set $\{A, B\}$. Notice that the points belong to the singularity set, where the map f is not (uniquely) defined.

Taking $\varepsilon = \frac{1}{2}$, we observe that the cardinality of minimal (n, ε)-spanning set satisfies: $2^{n+2} \le N(f, n, \varepsilon) \le 2^{n+3}$. In fact, if we partition CD into 2^n equal

intervals Z_iZ_{i+1}, then every d_n^f ε-ball is contained in some triangle AZ_iZ_{i+1} or BZ_iZ_{i+1}, and every such a triangle is covered by two d_n^f ε-balls.

Therefore the topological entropy $h_{\mathrm{top}}(f) = \log 2$ is positive. In addition, the Lyapunov spectrum is strictly negative at each point (for strict contractions), and no invariant measure exists, so the variational principle breaks.

The result of Buzzi [17] generalizes, however, in the following fashion:

Theorem. *Let f be a piece-wise affine map with restriction to each continuity component being conformal (non-strict) contraction. Then $h_{\mathrm{top}}(f) = 0$.*

Now we can repeat Buzzi's remark 4 [17]: the proof of his theorem 3 applies almost literally to the above case of piece-wise affine conformal contracting maps. Therefore we omit the proof.

Remark 5.1 *It is obvious that if the attractor consists of one point only, then $h_{\mathrm{top}}(f) = 0$. If the phase space $X \subset \mathbb{R}^1$ is one-dimensional and the map is (non-strictly) contracting, then again $h_{\mathrm{top}}(f) = 0$. We don't even need to require piece-wise affine property. This follows from the Buzzi proposition 4 [16], yielding $h_{\mathrm{top}}(f) \leq H_{\mathrm{mult}}(f)$, where $H_{\mathrm{mult}}(f)$ is the multiplicity entropy because the latter always vanishes in dimension one.*

Thus, our example with two-points attractor and two-dimensional phase-space X is the simplest possible example of a contracting chaotic system.

Part 2:

Stochastic modeling of temporal fluctuations

Chapter 6

Introduction to the stochastic model

Abstract: This chapter is an introduction to the second part of the thesis. We explain how to derive an approximate stochastic model of the toppling activity in sandpile models and discuss some fundamental concepts related to self-similar stochastic processes and stochastic differential equations. Numerical verification and applications of the stochastic model are provided in chapters 7 and 8.

1 Introduction

In the first part of the thesis we presented descriptions of the BTW and Zhang models within the framework of dynamical systems theory, and the results and conclusions where written in terms of mathematical quantities specifically designed to study chaotic maps acting in low dimensions. These results are useful for the general understanding of SOC, but it is difficult to compare them directly with experimental/observational data, and they do not answer any of the classical questions regarding SOC.

Traditionally, investigations of SOC focus on the scaling properties of the statistical distributions of certain avalanche observables. For instance, it is commonly believed that the probability distributions[1] of avalanche size and avalanche duration are power laws:

$$P_{\text{size}}(s) \sim s^{-\nu} \text{ and } P_{\text{dur}}(\tau) \sim \tau^{-\alpha}\,.$$

The exponents ν and α vary from model to model, and this has lead to the notion of universality classes, which simply are collections of SOC models with the same set of critical exponents. However, as computer simulations of sandpiles have become more sophisticated it has become evident that the determination of critical exponents is more difficult than what was first believed.

One of the problems is the lack of finite-size scaling for avalanche observables. In the BTW model for instance, the shape of the probability distributions depend on the system size a non-trivial way, making it difficult to extrapolate numerical results to the thermodynamic/continum limit.

Remark 6.1 *Consider a sandpile on a lattice with $N \times N$ sites and let $P_{size}(s;N) = s^{-\nu}G_N(s)$ denote the distribution of avalanche sizes. Simple finite size scaling means that we can write $G_N(s) = g(s/N^D)$, where $D > 0$ is some exponent and $g(r)$ is some (N-independent) function which decays rapidly to zero for $r > 1$. This can be tested numerically by plotting $s^\nu P_{size}(s;N)$ versus s/N^D for various N. For systems with simple finite-size scaling these curves should coincide.*

The second problem is that the probability distribution of some avalanche observables (examples are durations in the BTW and Zhang models) are not scaling unless their definitions are modified. The classical definition of an avalanche is a set of subsequent time steps for which the configuration is non-stable. This definition can be expressed in terms of the toppling activity signal as follows:

If $x(k) \in \{0, 1, \dots\}$ is the number of overcritical sites in the lattice at time step k, then an avalanche is a sequence of subsequent time steps for which $x(k) > 0$. In other words, the duration of an avalanche is the return time of $x(k)$ to the line

[1] We define these as the probability of having an avalanche of size (or duration) larger than s (or τ), i.e. $P_{\text{size}} = \text{Prob}[\text{size of avalanche} > s]$

$x = 0$, and the size of an avalanche is the area under the graph of $x(k)$ between subsequent zeros.

The problem with this definition is that it cannot be applied directly to experimental/observational data. Such data always includes noise, and it is impossible to determine if the value of x is exactly equal to zero or not. In this case we are forced to introduce a small threshold $x_{\text{th}} > 0$, and change the definition so that avalanches are subsequent times with $x(k) > x_{\text{th}}$. The durations of avalanches then become return times to the set $x < x_{\text{th}}$, and the definition of sizes is modified accordingly.

Definition of avalanches with respect to a positive threshold on the toppling activity was first introduced in [48], where it was demonstrated that the waiting-time statistics for the BTW sandpile undergoes a significant change with this modification. Another important feature is that the distributions for durations in the BTW and Zhang models now become good power laws. (See chapters 7 and 8.)

A third advantage of using a threshold is that, when combined with a stochastic differential equation for the toppling activity, it allows us to describe the essential characteristics of (height-type) sandpile models by a single parameter $H \in (0, 1)$. In the next two chapters we will show how the toppling process (in the the continum limit) can be described by the equation

$$dX(t) = f(X)\, dt + \sigma \sqrt{X(t)}\, dB_H(t)\,, \tag{6.1}$$

where $B_H(t)$ is fractional Brownian motion with Hurst exponent H. For a threshold $X_{\text{th}} > 0$ we can change variable to $Y(t) = X(t) - X_{\text{th}}$ to get

$$\begin{aligned} dY(t) &= f(Y + X_{\text{th}})\, dt + \sigma \sqrt{X_{\text{th}} + Y(t)}\, dB_H(t) \\ &= \Big(f(X_{\text{th}}) + \mathcal{O}(Y)\Big)\, dt + \sigma \Big(\sqrt{X_{\text{th}}} + \mathcal{O}(Y)\Big)\, dB_H(t)\,. \end{aligned}$$

If the threshold is chosen such that $f(X_{\text{th}}) = 0$, then $Y(t)$ is essentially proportional to $B_H(t)$ on short time scales, and the corresponding duration statistics is simply the return time (to zero) of fractional Brownian motion.

In what follows we give an introduction to derivation of equation 6.1.

2 From a dynamical to a stochastic point of view

In the first part of the thesis our basic structure was a map $\Omega \to \Omega$ (where $\Omega = \Sigma_N^+ \times M$), and we made general characterizations of its orbits $\omega_0, \omega_1, \omega_2, \cdots \in \Omega$. We will now change our approach by fixing an observable $\phi : \Omega \to \mathbb{R}$. The sequences $x(k) = \phi(\omega_k)$ depend only on the initial condition ω_0 and we can write $x(k) : \Omega \to \mathbb{R}$, where

$$x(k)(\omega) = \phi(\omega_k) \text{ given } \omega_0 = \omega\,.$$

If we equip Ω with the probability measure $m = \mu_{\text{Ber}} \times \mu_{\text{Leb}}$ we obtain a family of random variables $\{x(k) \,|\, k \in \mathbb{N}\}$. By definition this is a discrete-time stochastic process.

Remark 6.2 *In the following we will drop the brackets and denote the process by $x(k)$. For each $\omega \in \Omega$, the sequence $k \mapsto x(k)(\omega)$ is called a realization of the process.*

It is often convenient to describe the process in terms of the family $\{P_{k_1 \cdots k_n} | k_i \in \mathbb{N}, n > 0\}$ of finite-dimensional marginals. These can be constructed from the process by defining random vectors $\mathbf{x}_{k_1 \cdots k_n} = (x(k_1), \ldots, x(k_n))$, and letting $P_{k_1 \cdots k_n}$ be the distribution of the probability measure $(\mathbf{x}_{k_1 \cdots k_n})_* m$ on $\mathbb{R}^n$, i.e.

$$\begin{aligned} P_{k_1 \cdots k_n}(y_1, \ldots, y_n) &= m\Big(\mathbf{x}_{k_1 \cdots k_n}^{-1}\big((-\infty, y_1] \cap \cdots \cap (-\infty, y_n]\big)\Big) \\ &= \text{Prob}[x(k_1) \leq y_1, \ldots, x(k_n) \leq y_n]\,. \end{aligned}$$

Two stochastic processes are said to be equal in distribution (denoted by $\equiv$) if the families of finite dimensional marginals coincide. Moreover, there is a theorem of Kolmogorov that ensures that for any consistent[2] family of probability distributions $\mathcal{P} = \{P_{k_1 \cdots k_n} | k_i \in \mathbb{N}, n \geq 0\}$ there is a real valued stochastic process, whose family of finite dimensional marginal distributions equals $\mathcal{P}$.

Remark 6.3 *If the random variables $x(k)$ and $y(k)$ have the same distribution (for fixed k) we write $x(k) \sim y(k)$. Clearly, $x(k) \equiv y(k)$ implies that $x(k) \sim y(k)$ for all k. The inverse statement is in general false.*

3 The branching model

Let $\phi(\omega)$ be the number of overcritical sites in the configuration corresponding to $\omega \in \Omega$. Then the corresponding process $x(k)$ is the toppling activity of the sandpile model. We will derive an approximation of $x(k)$ using stochastic differential equations. The model will be valid for the BTW and Zhang models, but has a simpler derivation for the so-called random neighbor model [20]. This a modified version of the two-dimensional BTW model, where the toppling rule is such that when site i topples, then the grains are distributed to $2d$ random sites $j_1(i), \ldots, j_{2d}(i)$. Sites on the boundary of the lattice distribute grains to less than $2d$ random sites, but the details of this implementation is not essential for the following.

[2]This means that if $\{k_{s_1}, \ldots, k_{s_m}\} \subset \{k_1, \ldots, k_n\}$, then $P_{k_{s_1} \cdots k_{s_n}}$ is the marginal of $P_{k_1 \cdots k_n}$ corresponding to the index set $\{s_1, \ldots, s_m\}$.

Let $OC(k)$ denote the set of overcritical sites at time step k and let $x(k)$ be the toppling activity. If $OC(k) = \{i_1(k), \ldots, i_{x(k)}(k)\}$ we define

$$\xi_{n,k} := \xi(i_n(k)) := \sum_{m=1}^{2d} \Theta\Big(\big(z(k+1)\big)_{j_m(i_n(k))} - 2d \Big) ,$$

where $z(k+1)$ is the configuration at the time step $k+1$. This defines a family of random variables $\{\xi_{n,k} \,|\, k \in \mathbb{N}, n = 1, \ldots, x(k)\}$.

Note that $\xi_{n,k}$ can be interpreted as the number of overcritical sites at time $k+1$ which are produced by the n-th overcritical site at time k. In the random neighbor model, site i distributes grains to sites which typically are far away from i and from each other. Thus it is reasonable to make the following mean-field approximations:

1. All random variables $\xi_{n,k}$ are independent and identically distributed.

2. $E[\xi_{n,k}] = 1$ for all $n \in \mathbb{N}$ and all $k = 1, \ldots, x(k)$.

Let $\sigma^2 = \mathrm{Var}[\xi_{n,k}]$ and notice that since

$$x(k+1) - x(k) = \sum_{n=1}^{x(k)} (\xi_{n,k} - 1)$$

we have

$$x(k+1) = x(k) + \sigma \sum_{i=n}^{x(k)} w_i ,$$

where w_i are independent random variables on $\{-1, \ldots, 2d\}$ with unit variance and zero mean.

A realization of this process will exist as long as long as $x(k) \geq 0$, but this problem can be avoided by defining

$$x(k+1) = \max\Big\{ x(k) + \sigma \sum_{i=n}^{x(k)} w_i, 0 \Big\} .$$

We now have a stochastic model for the toppling activity $x(k)$. Notice that it does not include the random dropping of sand grains onto the lattice, and new avalanches are not initiated after the first avalanche has terminated. Nevertheless, if we assume that the toppling activity in one avalanche is independent of the previous avalanches, then the stochastic model describes the statistical distribution of the avalanche size and duration. For instance, if we choose $x(0) = 1$, then the duration of an avalanche is a random variable defined by $\tau = \min\{k \,|\, x(k) = 0\}$.

It is not difficult to derive an analytic expression for the distribution function of avalanche duration based on this stochastic model, but it is easier if we first proceed to a continuous time limit. Before we do this we will use the next section to describe some of the fundamental properties of Brownian motion and Itô stochastic differential equations.

4 Brownian motion

A stochastic process can be defined for continuous times $t \in \mathbb{R}$ or $t \in [0, \infty)$ in the same way as we defined discrete processes. This means that we can think of a process as a family $\{X(t)|t \in \mathbb{R}\}$ of random variables of the same probability space, or as a family $\{P_{t_1,\ldots,t_n}|t_1, \ldots, t_n \in \mathbb{R}, n > 0\}$ of finite-dimensional marginals. As before, two processes are said to be equal in distribution if their families of finite-dimensional marginals coincide.

4.1 Self-similar processes

A process $X(t)$ is self-similar if there for all $a > 0$ exist $b > 0$ such that

$$X(at) \equiv bX(t)\,.$$

The simplest kind of self-similar processes is when $X(t)$ is independent of t almost surely. We call such a process trivial. For any non-trivial process there exists a real number $H > 0$ such that $b = a^H$. We call H the Hurst exponent of the process, and we say that $X(t)$ is H-self-similar. If $X(t)$ is a non-trivial self-similar process, then $X(0) \sim a^H X(0)$, and letting $a \to 0$ we see that $X(0) = 0$ almost surely.

In general we say that a process $X(t)$ has stationary increments if all the joint distributions of $X(t + \Delta t) - X(t)$ are independent of t. Increments are said to be independent if the random variables $X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent for any $n \in \mathbb{N}$ and $t_1, \ldots, t_n \in \mathbb{R}_{\geq 0}$. A Gaussian self-similar process with stationary and independent increments is called Brownian motion. Equivalently we can define Brownian motion as a process $B(t)$ with the following properties:

1. $B(0) = 0$ almost surely
2. Realizations are almost surely continuous curves
3. Independent increments
4. $X(t + \Delta t) - X(t) \sim \mathcal{N}(0, \Delta t)$ for all $\Delta t > 0$

It is easy to see that if $B(t)$ satisfies proporties 1-4, then so does $a^{-1/2}B(at)$ for all $a > 0$. Hence Brownian motion is self-similar with $H = 1/2$.

4.2 Construction of Brownian motion

It is instructive to see how we can construct Brownian motion as a limit of discrete random walks: Fix a parameter $\delta x > 0$, let $x(k) = 0$ and

$$x(k+1) = x(k) + \delta x\, w(k)\,,$$

where $w(k)$ are independently and identically distributed (i.i.d.) random variables with finite variance and zero mean. Assume with no loss of generality that $\mathrm{Var}[w(k)] = 1$. Denote $\Delta x(k) = x(k+1) - x(k)$. Then

$$x(k) = \sum_{i=1}^{k} \Delta x(k) = \delta x \sum_{i=1}^{k} w(k)\,.$$

This is a sum of i.i.d. random variables, so by the central limit

$$\mathrm{Prob}\left[a < \frac{x(k)}{\delta x\sqrt{k}} < b\right] \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-y^2/2}\, dy \;\text{ as } k \to \infty\,.$$

Re-scale the time variable by $t = k \cdot \delta t$ and let $B(t) = x(k)$. Then

$$\mathrm{Prob}\left[a < \frac{B(t)}{\delta x \delta t^{-1/2}\sqrt{t}\sqrt{k}} < b\right] \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-y^2/2}\, dy \;\text{ as } k \to \infty\,.$$

Choose δx depending on δt such that $\delta x \sim \delta t^{1/2}$ as $\delta t \to 0$, and let $\sigma = \lim_{\delta t \to 0} \delta x \delta t^{-1/2}$. Holding t fixed and letting $\delta t, \delta x \to 0$ we get

$$\mathrm{Prob}[a < B(t) < b] = \frac{1}{\sqrt{2\pi\,\sigma^2\, t}} \int_a^b e^{-\frac{y^2}{2\sigma^2 t}}\, dy\,.$$

The process $B(t)$ in the limit $\delta x, \delta t \to 0$ is the Brownian motion.

Denote $\Delta X B(t) = B(t+\Delta t) - B(t)$ and let $P(B,t)$ and $\rho(\Delta B, t)$ denote the probability density functions of $B(t)$ and $\Delta B(t)$ respectively. From the construction it is clear that $\rho(B,t) = \rho(B)$ is independent of t (as it should be when the increment process is stationary). Notice also that ρ depends on Δt. In fact, $\rho(\Delta B) = P(\Delta B, \Delta t)$.

4.3 The diffusion equation

Let P and ρ the probability density functions for Brownian motion as defined above. Then

$$P(X, t+\Delta t) = \int_{\mathbb{R}} P(X - \Delta X, t)\rho(\Delta X) d\Delta X\,.$$

Expanding in Δt and ΔX we get

$$P(X,t) + \frac{\partial P}{\partial t}(X,t)\Delta t + \mathcal{O}(\Delta t^2) =$$
$$\int_{\mathbb{R}} \Big(P(X,t) - \frac{\partial P}{\partial X}(X,t)\Delta X + \frac{1}{2}\frac{\partial^2 P}{\partial X^2}(X,t)\Delta X^2 + \mathcal{O}(\Delta X^3)\Big)\rho(\Delta X)\, d\Delta X\,.$$

Since

$$\int \Delta X \rho(\Delta X) d\Delta X = 0\,, \int \Delta X^2 \rho(\Delta X) d\Delta X = \Delta t\,,$$

and higher moments of ΔX grow faster than linearly in Δt, we get the diffusion equation

$$\frac{\partial P}{\partial t} = \frac{1}{2}\frac{\partial^2 P}{\partial X^2}$$

in the limit $\Delta t \to 0$. This is an example of a (Kolmogorov) Fokker-Planck equation.

5 Stochastic differential equations

The Brownian motion is a special case of a solution to the Itô stochastic differential equation (SDE)

$$dX(t) = \mu(X,t)\, dt + \sigma(X,t)\, dB(t)\,, \tag{6.2}$$

which is a notation representing the integral equation

$$X(t) = X(0) + \int_0^t \mu(X(t'),t')dt' + \int_0^t \sigma(X(t'),t')\, dB(t')\,.$$

The second integral on the right hand side of the equation is called the Itô integral. For a real valued stochastic process $H(t)$ this integral is defined as the limit (convergence is in probability)

$$\int_0^t H(s)\, dB(s) = \lim_{\epsilon\to 0} \sum_i H(s_{i-1})\Big(B(s_i) - B(s_{i-1})\Big)\,,$$

where $\{[s_{i-1}, s_i]\}$ is a partition of $[0,t]$ with diameter $\leq \epsilon$.

Itô's lemma tells us that Itô's equation transforms in a non-trivial way under a change of variable $Y(t) = f(X(t),t)$:

Theorem 6.1 (Itô's lemma) *Assume that the stochastic process $X(t)$ is a solution of the stochastic differential equation $dX(t) = \mu(X,t)\, dt + \sigma(X,t)\, dB(t)$ and that $Y(t) = f(X,t)$, where $f(X,t)$ is a smooth function of two real variables. Then $Y(t)$ is a solution of the stochastic differential equation*

$$dY(t) = \Big(\mu\frac{\partial f}{\partial X} + \frac{\partial f}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial X^2}\Big)\, dt + \sigma\,\frac{\partial f}{\partial X}\, dB(t)\,.$$

Remark 6.4 *The classical example of use of the Itô lemma is in finance. The famous Black-Scholes equation describes a price by the equation*

$$dX(t) = \mu X\, dt + \sigma X dB(t)\,.$$

Preforming the substitution $Y(t) = \log X(t)$ *and using Itô's lemma we get*

$$dY(t) = (\mu - \sigma^2/2)\, dt + \sigma dB(t)\,.$$

In other words, the logarithmic price is a simple Brownian motion with a drift term.

By using the Itô lemma one can derive the Fokker-Planck equation for processes satisfying equation 6.2: If $P(X,t)$ is the probability density function of $X(t)$, then

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial X}(\mu P) + \frac{1}{2}\frac{\partial^2}{\partial X^2}(\sigma^2 P)\,. \tag{6.3}$$

6 SDE formulation of the branching process

We now return to the discrete branching model:

$$x(k+1) - x(k) = \sigma \sum_{i=1}^{x(k)} w_i\,.$$

By the central limit theorem, the sum on the right hand side of the equation can be approximated by a Gaussian random variable with zero mean and variance $\sigma^2 x(k)$. Since $B(t+\delta t) - B(t) \sim \mathcal{N}(0, \delta t)$ we can write

$$\sigma \sum_{i=1}^{x(k)} w_i \approx \frac{\sigma\sqrt{x(k)}}{\sqrt{\delta t}} \Big(B(t+\delta t) - B(t)\Big)\,.$$

Similar to the construction of Brownian motion from a random walk we want to take a continuous limit of this discrete process. Let $\delta x, \delta t > 0$, define $t = \delta t \cdot k$ and let $X(t) = x(k)\,\delta x$. Then

$$\begin{aligned}\frac{X(t+\delta t) - X(t)}{\delta x} &= x(k+1) - x(k) \approx \sigma\,\frac{1}{\sqrt{\delta t}}\,\sqrt{x(k)}\Big(B(t+\delta t) - B(t)\Big) \\ &= \sigma\sqrt{\frac{X(t)}{\delta x \cdot \delta t}}\Big(B(t+\delta t) - B(t)\Big)\,.\end{aligned}$$

This gives the equation

$$X(t+\delta t) - X(t) = \sqrt{\frac{\delta x}{\delta t}}\,\sigma\,\sqrt{X(t)}\,\Big(B(t+\delta t) - B(t)\Big)\,.$$

Let δt depend on δx in such a way that $\lim_{\delta x \to 0} \delta x / \delta t = 1$. When δx and δt go to zero we get the Itô stochastic differential equation

$$dX(t) = \sigma \sqrt{X(t)}\, dB(t)\,.$$

Remark 6.5 *For a standard random walk we can obtain a Brownian motion by letting $\delta x \sim \delta t^{1/2}$, whereas here we must choose $\delta x \sim \delta t$. Although the process $X(t)$ is not self-similar in a strict sense, the relation between δx and δt indicates self-similar properties with $H = 1$.*

Notice that use of the central limit theorem requires $x(k) \gg 1$. This means that the model is not valid for small activities. In particular, the initial condition $x(k) = 1$ corresponds to $X(0) = \delta x \to 0$, and then $X(t)$ is trivial. Hence we have to consider initial conditions $X(0) = Y$, where $Y > 0$ is chosen in the rescaled coordinates.

Another important property is that the process only is defined for positive X, and even for a positive initial condition $X(0) = Y > 0$, the solution $X(t)$ will cease to exist after finite time (because $X(t)$ crosses the line $X = 0$). We can see this by considering fixed $\delta t, \delta x > 0$. Then $X(t + \delta t) = X(t) + \sigma \sqrt{X(t)}\, w(t; \delta t)$, where $w(t; \delta t)$ are independent Gaussian random variables with zero mean and variance δt. Since conditional probability density function of $X(t + \delta t) - X(t)$ given $X(t) = X$ is

$$\rho(\Delta X; X, \delta t) = \frac{1}{\sqrt{2\pi\sigma^2 \delta t\, X}}\, e^{-\frac{\Delta X^2}{2\sigma^2 \delta t\, X}}\,,$$

the probability of "jumping" from a positive value $X = X(t) > 0$ to a negative value $X(t + \delta t) < 0$ is equal to

$$\int_{-\infty}^{X} \rho(\Delta X, X, \delta t)\, d\Delta X = \frac{1}{2} - \frac{1}{2}\,\mathrm{Erf}\left(\frac{\sqrt{X}}{\sqrt{2}\,\sigma\sqrt{\delta t}}\right) = \frac{1}{2} - \frac{1}{2}\,\mathrm{Erf}\left(\frac{\sqrt{x(k)}}{\sqrt{2}\,\sigma}\sqrt{\frac{\delta x}{\delta t}}\right).$$

This quantity is bounded away from zero in the limit $\delta x, \delta t \to 0$.

When a solution $X(t)$ ceases to exist, we can interpret this as the termination of an avalanche. This allows us to calculate the distribution function for avalanche durations analytically. Let $P(X, t)$ be defined by

$$\mathrm{Prob}[a < X(t) < b \text{ and } X(t') > 0\ \forall t' \in [0, t] \,|\, X(0) = Y] = \int_a^b P(X, t)\, dX\,.$$

Then the probability that an avalanche is still running after a time t is

$$\mathrm{Prob}[\tau > t] = \int_0^\infty P(X, t)\, dX\,.$$

In particular, the probability density function for avalanche duration is

$$p_{\rm dur}(t) = -\frac{d}{dt}\,{\rm Prob}[\tau > t] = -\int_0^\infty \frac{\partial P}{\partial t}(X,t)\,dX\,. \tag{6.4}$$

We now apply the Fokker-Planck equation (see equation 6.3)

$$\frac{\partial P}{\partial t} = \frac{\sigma^2}{2}\,\frac{\partial^2}{\partial X^2}(X\,P)\,. \tag{6.5}$$

This gives

$$\begin{aligned} p_{\rm dur}(t) &= -\frac{\sigma^2}{2}\int_0^\infty \frac{\partial^2}{\partial X^2}(XP)\,dX = \frac{\sigma^2}{2}\,\lim_{X\to 0}\frac{\partial}{\partial X}(XP) \\ &= \frac{\sigma^2}{2}\,\lim_{X\to 0} P(X,t)\,. \end{aligned}$$

The termination of realizations when they reach the line $X = 0$ corresponds to the Fokker-Planck equation (Eq.6.5) with an absorbing boundary in $X = 0$ and initial condition $P(X,0) = \delta_Y(X)$. Using separation of variables and superposition via Hankel transforms it is not difficult to establish the solution

$$P(X,t) = \frac{1}{2}\sqrt{\frac{Y}{X}}\int_0^\infty J_1(s\sqrt{Y})\,J_1(s\sqrt{X})\,\exp\left(-\frac{\sigma^2\,s^2}{8}\,t\right)s\,ds\,,$$

where J_1 denotes the Bessel function of first kind. Then

$$\begin{aligned} \lim_{X\to 0} P(X,t) &= \frac{\sqrt{Y}}{4}\int_0^\infty s^2\,J_1(s\sqrt{Y})\,\exp\left(-\frac{\sigma^2\,s^2}{8}\,t\right)ds \\ &= \frac{4\,X_0}{\sigma^4 t^2}\,\exp\left(-\frac{2\,Y}{\sigma^2\,t}\right), \end{aligned}$$

and

$$p_{\rm dur}(\tau) = \frac{2\,Y}{\sigma^2\tau^2}\,\exp\left(-\frac{2\,Y}{\sigma^2\,\tau}\right).$$

For $\tau \gg 2Y/\sigma^2$ we have $p_{\rm dur}(\tau) \sim \tau^{-\alpha}$, with $\alpha = 2$.

7 Extension to the BTW and Zhang models

The result $p_{\rm dur}(\tau) \sim \tau^{-2}$ is known to agree well with numerical simulation of the random neighbor sandpile model. However, the exponent $\alpha = 2$ does not agree with simulations of the BTW and Zhang models. We will now generalize the model in order to obtain good descriptions of the toppling activities in the BTW and Zhang models as well. The central element is this generalization is introduction of fractional noise.

7.1 Fractional noise terms

Recall that the main assumption leading to the stochastic differential equation $dX(t) = \sqrt{X(t)}dB(t)$ is that the random variables $\xi_{n,k}$ are identically and independently distributed. Our first generalization involves relaxing this requirement by only assuming that we for a fixed $k \in \mathbb{N}$ have independence between $\xi_{1,k}, \ldots, \xi_{x(k),k}$. This is all we need to apply the central limit theorem and obtain the stochastic difference equation

$$x(k+1) - x(k) = \sigma \sqrt{x(k)}\, w(k)\,,$$

where $w(k)$ are Gaussian.

Previously we assumed that $w(k)$ where independent for different k, and then we could write $w(k) = B(k+1) - B(k)$. A generalization of this would be to replace $B(t)$ with a Gaussian process with dependent increments. The obvious candidate for such a process is a fractional Brownian motion $B_H(t)$, which can be defined as an H-self-similar Gaussian process with stationary increments.

We can also define $B_H(t)$ explicitly: For $H \in (0,1)$ a fractional Brownian motion can be defined from a Brownian motion by the following formula [41]:

$$B_H(t) = C \int_0^t K(t,t')\, dB(t')$$

where

$$K(t,t') = \left(\frac{t}{t'}\right)^{H-\frac{1}{2}} (t-t')^{H-\frac{1}{2}} - \left(H - \frac{1}{2}\right) t'^{\frac{1}{2}-H} \int_{t'}^t s^{H-\frac{3}{2}} (s-t')^{H-\frac{1}{2}} ds$$

An important property of $B_H(t)$ is that

$$\text{Prob}[a < B_H(t+\Delta t) - B_H(t) < b] = \frac{1}{\sqrt{2\pi}\Delta t^H} \int_a^b e^{-\frac{\Delta X^2}{2\Delta t^{2H}}}\, d\Delta X\,.$$

Moreover $B_{1/2}(t)$ is Brownian motion, but for $H \neq 1/2$ increments are not independent. Inversely, any H-self-similar Gaussian process with stationary increments is equal to $B_H(t)$ in distribution.

A generalization of the branching process can be obtained by writing

$$x(k+1) - x(k) = \sigma \sqrt{x(k)} \Big(B_H(k+1) - B_H(k) \Big)\,.$$

As before we obtain a stochastic differential equation by letting $t = k\,\delta t$ and $X(t) = x(k)\,\delta x$. If we let relate δx and δt such that $\lim_{\delta t \to 0} \delta x / \delta t^{2H} = 1$, then we get the stochastic differential equation

$$dX(t) = \sigma \sqrt{X(t)}\, dB_H(t)$$

in the limit $\delta x, \delta t \to 0$:

Remark 6.6 *There is no Fokker-Planck formulation for stochastic differential equations driven by fractional noise, so the method we used to calculate the distribution of avalanche durations in the case $H = 1/2$, no longer applies for $H \neq 1/2$.*

Once we have generalized the model to include the possibility of fractional Brownian noise terms we can start to compare it with the numerical simulations of the BTW and Zhang models. The fist step in this analysis is to verify that the conditional probability distribution of $x(k+1) - x(k)$ are Gaussians with variance proportional to $x(k)$. The results of this analysis is presented in the next two chapters.

The second step of the analysis is to verify that the noise term is produced by a self-similar process with stationary increments. From numerical simulation of the sandpiles we obtain time series representing the signal $x(k)$. From this signal we can construct a signal representing $w(k)$ by taking

$$w(k) = \frac{x(k+1) - x(k)}{\sigma\sqrt{x(k)}} \, .$$

We have modeled $w(k)$ as a difference $B_H(k+1) - B_H(k)$, so we can obtain a discrete approximation to $B_H(t)$ by taking the sums

$$\sum_{k'=0}^{k} \frac{x(k'+1) - x(k')}{\sigma\sqrt{x(k')}} \, .$$

If our model is correct, then this process should have the characteristics of a fractional Brownian motion. This means that we should check that it is Gaussian, that the increments are stationary and that the process is H-self-similar.

It is easy to check that the process is Gaussian, and weak stationarity of increments can be tested by constructing various realizations of starting from zero, and verifying that the moments of the increment process are time independent. A more difficult problem is to verify self-similarity. From a strict mathematical point of view we are forced to calculate all finite dimensional marginals of the process to verify self-similarity. Obviously, this is not a practical method of analysis for numerical data, and instead we resort to the analysis of structure functions.

For a fractional Brownian motion $B_H(t)$ we have

$$\rho(\Delta B) = \frac{1}{\sqrt{2\pi\Delta t^{2H}}} \, e^{-\frac{\Delta B}{2\Delta t^{2H}}} \, .$$

This implies that the q-th moment satisfies

$$\begin{aligned} S_q(\Delta t) &= E\big[\big|B(t+\Delta t) - B(t)\big|^q\big] = \int_{\mathbb{R}} |\Delta B|^q \rho(\Delta B)\, d\Delta B \\ &= \frac{1}{\sqrt{2\pi\Delta t^{2H}}} \int_{\mathbb{R}} |\Delta B|^q e^{-\frac{|\Delta B|}{2\Delta t^{2H}}}\, d\Delta B \propto \Delta t^{Hq} \, . \end{aligned}$$

The functions $S_q(\Delta t)$ are called the structure functions of the process. In particular the double-logarithmic plot of $S_2(\Delta t)$ versus Δt is called the variogram of the process. In the physical literature (if this curve is a straight line) the slope divided by two is called the Hurst exponent of the process. Of course, this coincides with our definition of the Hurst exponent in the case of fractional Brownian motion, since we have $S_q(\Delta t) \propto \Delta t^{2H}$. However, there are processes for which $S_q(\Delta t) \propto \Delta t^{2H}$ but where, for instance, $S_4(\Delta t) \sim \Delta t^r$ for $r < 4H$. If increments of such a process are stationary, then we know that the process is not a self-similar process. Actually there exist a generic class of processes, which are omnipresent in nature, where the exponents in the different structure functions depend on q non-lienarly. These processes are called multi-fractals, and they can be characterized by the zeta-function $\zeta(q)$ which is defined through the relation $S_q(\Delta t) \sim \Delta t^{\zeta(q)}$. We know that for fractional Brownian motion the zeta-funciton is of the form $\zeta(q) = Hq$, but for multi-fractal processes the zeta-functions are concave.

Since multi-fractal processes are as generic as self-similar processes, it is very important to test for multi-fractality before modeling a signal as a self-similar process. The results from this test are presented in chapter 7.

Remark 6.7 *The stochastic model $dX(t) = \sigma\sqrt{X(t)}\,dB_H(t)$ has two essential weaknesses when compared with the toppling activity in real sandpiles. First of all, the model does not reflect the toppling process correctly for very small values of $x(k)$. Secondly, there is no information about the system size in the model. However, both of these problems can be solved by heuristically introducing a drift term $f(X)\,dt$ to the model. The shape of this term, and its interpretation is discussed in chapters 7 and 8.*

With the drift term our stochastic model has the final form

$$dX(t) = f(X)\,dt + \sigma\sqrt{X(t)}\,dB_H(t)\,.$$

Appendix

A Stochastic Integrals

A stochastic differential equation $dX(t) = \mu(X(t), t)dt + \sigma(X(t), t)dB(t)$ is a notation for the integral equation

$$X(t) = X(0) + \int_0^t \mu(X(t'), t')\, dt' + \int_0^t \sigma(X(t'), t')\, dB(t')\,.$$

The second integral on the right hand side is a so-called Itô integral, and it can be defined, not only for Brownian motion, but for all semi-martingales. However, fractional Brownian motions are not semi-martingales, and thus it is not completely trivial to extend this integral to include fractional Brownian motion $B_H(t)$. In fact, there exist several definitions of the integral

$$\int_a^b X(t)\, dB_H(t)$$

for $H \neq 1/2$, and today it is not generally agreed which of these definition is "correct". In this section we present an approach given by Øksendal [58].

Remark 6.8 *Brownian motion $B(t)$ is a martingale. This means that the conditional expectation of $B(t)$ given the trajectory $\{B(t') \,|\, t' \leq s\}$ up to a time $s \leq t$, is equal to $B(s)$. That is,*

$$E\Big[B(t)\Big|\{B(t')|t' \leq s\}\Big] = B(s)\ \forall s \leq t\,.$$

There is also a notion of local martingales, which loosely defined, are processes $X(t)$ for which the stopped processes

$$X(\min\{t, t_k\})\Theta(t_k - t)$$

are martingales for all k in an infinite almost surely increasing and divergent sequence $\{t_k\}$ of stopping times. A semi-martingale is a process $X(t)$ which can be

decomposed into a sum $M(t) + V(t)$, where $M(t)$ is a local martingale and $V(t)$ is a càdlàg (right continuous with left limits) process of bounded variation.[3]

A.1 Brownian motion and white noise

We have so far defined Brownian motion $B(t)$ in terms of its family of finite-dimensional marginals and we have constructed a discrete random walk process which approximates $B(t)$ as $\delta t, \delta x \to 0$. However, it is also possible to define the process $B(t)$ explicitly by specifying a probability space $(\Omega, \mathcal{F}, \mu)$ and a family $\{B(t) : \Omega \to \mathbb{R} | t \in \mathbb{R}\}$ of random variables. The standard way of doing this is to choose Ω to be the set of tempered distributions on $\mathbb{R}$, i.e. $\Omega = \mathcal{S}(\mathbb{R})^*$, where $\mathcal{S}(\mathbb{R}) \subset C^\infty(\mathbb{R})$ is the Schwartz space of rapidly decreasing functions on $\mathbb{R}$. In other words, the elements of $\omega \in \Omega$ are functionals taking functions $f \in S(\mathbb{R})$ to real numbers $\langle \omega, f \rangle := \omega(f) \in \mathbb{R}$. Let $\mathcal{F}$ be the Borel σ-algebra on Ω. It can be shown that there exists a unique Borel probability measure μ on Ω satisfying the equation

$$\int_\Omega e^{i\langle \omega, f \rangle} d\mu(\omega) = e^{-\frac{1}{2}||f||^2_{L^2(\mathbb{R})}} . \tag{6.6}$$

This defines our probability space $(\Omega, \mathcal{F}, \mu)$.

With this set-up every function $f \in \mathcal{S}(\mathbb{R})$ defines a random variable by $\omega \mapsto \langle \omega, f \rangle$. Hence a family $\{f_t | t \in \mathbb{R}\}$ of functions will define a stochastic process $\omega \to \langle \omega, f_t \rangle$. Brownian motion can be defined simply by specifying

$$f_t(s) = \begin{cases} 1 & \text{for } s \in [0, t] \\ 0 & \text{otherwise} \end{cases}$$

when $t > 0$ and

$$f_t(s) = \begin{cases} -1 & \text{for } s \in [t, 0] \\ 0 & \text{otherwise} \end{cases}$$

for $t < 0$.

Two important properties of the measure defined by equation 6.6 is that for all $f \in \mathcal{S}(\mathbb{R})$:

$$E[\langle \omega, f \rangle] = \int \langle \omega, f \rangle \, d\mu(\omega) = 0$$

and

$$E[\langle \omega, f \rangle^2] = ||f||^2_{L^2(\mathbb{R})} .$$

[3] Bounded variation on an interval means that $\sup_{\mathcal{P}} \sum_n |V(t_i) - V(t_{i-1})| < +\infty$, where the supremum is taken over all finite partitions $\{[t_i, t_{i+1}]\}$ of the interval. Intuitively, this means that $V(t)$ has only jump singularities.

Using these properties it is not difficult to verify that the family of random variable defined by $\{f_t\}$ satisfies the defining properties of Brownian motion.

A more elegant description of Brownian motion is possible if we choose an orthonormal basis of Hermite functions $\{\xi_n\}$ in $L^2(\mathbb{R})$ and expand the functions f_t:

$$f_t = \sum_n (\xi_n, f_t)_{L^2(\mathbb{R})}\, \xi_n = \sum_n \Big(\int_{\mathbb{R}} \xi_n(s)\, f_t(s)\, ds \Big)\, \xi_n = \sum_n \Big(\int_0^t \xi_n(s)\, ds \Big)\, \xi_n \,.$$

Then Brownian motion can be written as

$$B(t)(\omega) = \sum_n \Big(\int_0^t \xi_n(s)\, ds \Big)\, \langle \omega, \xi_n \rangle \,.$$

This shows that Brownian motion can be written as an expansion over random variables $\langle \cdot, \xi_n \rangle$ with time-dependent coefficients. The result suggests that one should try to find a basis of in which all random variables on the form $\langle \cdot, f \rangle$ (with $f \in L^2(\mathbb{R})$) can be expanded. Such a basic set exists, and its elements can be given explicitly by the formula

$$\mathcal{H}_{\mathbf{k}}(\omega) = h_{k_1}(\langle \omega, \xi_1 \rangle) h_{k_2}(\langle \omega, \xi_2 \rangle) \cdots h_{k_n}(\langle \omega, \xi_n \rangle) \,,$$

where h_i are the Hermite polynomials and $\mathbf{k} = (k_1, \dots, k_n)$ is a multi-index. It is then natural to define the space of random variables as the set $(\mathcal{S})^*$ of all formal expressions on the form

$$\sum_{\mathbf{k} \in I} c_{\mathbf{k}} \mathcal{H}_{\mathbf{k}} \,,$$

with some conditions on $c_{\mathbf{k}}$ being satisfied. Here I is the set of all finite-length multi-indices $\mathbf{k} = (k_1, \dots, k_n)$ with $k_i \in \{0, 1, 2, \dots\}$.

Note that if $\mathbf{e}_n = (0, 0, \dots, 0, 1)$ (where the 1 is in the n-th entry), then we can write

$$B(t) = \sum_n \Big(\int_0^t \xi_n(s) ds \Big)\, \mathcal{H}_{\mathbf{e}_n} \,.$$

We observe that the map $B(\cdot) : \mathbb{R} \to (\mathcal{S})^*$ is differentiable, and we can define the process

$$W(t) = \frac{dB(t)}{dt} = \sum_n \xi_n(t)\, \mathcal{H}_{\mathbf{e}_k} \,.$$

This process is called white noise.

The Wick product of two elements $F_1, F_2 \in (\mathcal{S})^*$ is defined by the formula

$$F_1 \diamond F_2 = \sum_{\mathbf{k}, \mathbf{l} \in I} c_{\mathbf{k}}^{(1)} c_{\mathbf{l}}^{(2)} \mathcal{H}_{\mathbf{k}+\mathbf{l}} \,,$$

where

$$F_1 = \sum_{\mathbf{k}} c_{\mathbf{k}}^{(1)} \mathcal{H}_{\mathbf{k}} \text{ and } F_2 = \sum_{\mathbf{k}} c_{\mathbf{k}}^{(2)} \mathcal{H}_{\mathbf{k}} \,.$$

Using this product we can formulate Itô integration with respect to Brownian motion in terms of white noise. In fact, if $X(t)(\omega)$ is an Itô integrable process, then its integral with respect to $B(t)$ can be written as

$$\int_0^t X(t')\, dB(t') = \int_0^t X(t) \diamond W(t)\, dt \,. \tag{6.7}$$

A.2 Fractional Brownian noise

With the white-noise formalism presented above, the generalization to fractional Brownian noise is straight forward. It is not difficult to see that fractional Brownian motion can be defined through the family of functions $\{\mathcal{M}_H f_t\}$, where the operator $\mathcal{M}_H : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is given by

$$(\mathcal{M}_H \phi)(s) = \begin{cases} C_H \int_{\mathbb{R}} \frac{\phi(s-s')-\phi(s)}{|s'|^{3/2_H}}\, ds' & H \in (0, 1/2) \\ \phi(s) & H = 1/2 \\ C_H \int_{\mathbb{R}} \frac{f(s')}{|s'-s|^{3/2-H}}\, ds' & H \in (1/2, 1) \end{cases} .$$

Here C_H is a normalization constant only depending on H.

We define an inner product $(f, g)_{L^2(\mathbb{R})} = (\mathcal{M}_H f, \mathcal{M}_H g)_{L^2(\mathbb{R})}$ and let $L^2_H(\mathbb{R})$ be the space of functions f with $||f||_{L^2_H(\mathbb{R})} < \infty$. In this space we have an orthonormal basis $\eta_n = \mathcal{M}_H^{-1} \xi_n$. It is easy to verify that $(\mathcal{M}_H f, g)_{L^2(\mathbb{R})} = (f, \mathcal{M}_H g)_{L^2(\mathbb{R})}$ for $f, g \in L^2(\mathbb{R}) \cap L^2_H(\mathbb{R})$, and this gives

$$\begin{aligned} B_H(t)(\omega) &= \langle \omega, \mathcal{M}_H f_t \rangle = \langle \omega, \sum_n (f_t, \eta_n)_{L^2_H(\mathbb{R})} \mathcal{M}_H \eta_n \rangle \\ &= \langle \omega, \sum_n (\mathcal{M}_H f_t, \xi_n)_{L^2(\mathbb{R})} \xi_n \rangle = \sum_n (f_t, \mathcal{M}_H \xi_n)_{L^2(\mathbb{R})} \langle \omega, \xi_n \rangle \\ &= \sum_n \Big(\int_0^t \mathcal{M}_H \xi_n(s)\, ds \Big) \mathcal{H}_{\mathbf{e}_n}(\omega) \,. \end{aligned}$$

This shows that $B_H(\cdot) : \mathbb{R} \to (\mathcal{S})^*$ is a differential map and the differential is

$$W_H(t) = \sum_n \mathcal{M}_H \xi_n(t)\, \mathcal{H}_{\mathbf{e}_n} \,.$$

The process $W_H(t)$ is called fractional Brownian noise with Hurst exponent H. By analogy with equation 6.7 we define

$$\int_0^t X(t') dB_H(t') = \int_0^t X(t') \diamond W_H(t')\, dt' \,.$$

Also note that the stochastic differential equation

$$dX(t) = \mu(X(t), t)\, dt + \sigma(X(t), t)\, dB(t)$$

can be written as

$$\frac{dX}{dt} = \mu(X(t), t) + \sigma(X(t), t) \diamond W_H(t)\,,$$

where the differentiation of $X(t)$ is with respect to the map $X(\cdot) : \mathbb{R} \to (\mathcal{S})^*$. In particular, the equation $dX(t) = \sigma\sqrt{X(t)}\, dB(t)$ can be written as

$$\frac{dX}{dt} = \sqrt{X(t)} \diamond W_H(t)\,. \tag{6.8}$$

Remark 6.9 *Linear stochastic differential equations on the form*

$$\frac{dX}{dt} = \mu X + \sigma \diamond W_H(t)\,,$$

can be solved in terms of the Wick exponential which is formally defined as

$$\exp^{\diamond}(F) = \sum_{n=0}^{\infty} \frac{1}{n!} F^{\diamond n}\,.$$

This can be used to derive a generalized Itô-lemma which works for these linear equations. However, this is approach does not work for equation 6.8. The main difficulty here is that the square root is not a Wick-type square root, and we have a combination of two different products. This makes it difficult to write a formal solution.

Chapter 7

A stochastic theory for the toppling activity in SOC

Joint with K. Rypdal

Abstract: A stochastic theory for the toppling activity in sandpile models is developed, based on a simple mean-field assumption about the toppling process. The theory describes the process as an anti-persistent Gaussian walk, where the diffusion coefficient is proportional to the activity. It is formulated as a generalization of the Itô stochastic differential equation with an anti-persistent fractional Gaussian noise source and a deterministic drift term. An essential element of the theory is re-scaling to obtain a proper thermodynamic limit. When subjected to the most relevant statistical tests, the signal generated by the stochastic equation is indistinguishable from the temporal features of the toppling process obtained by numerical simulation of the Bak-Tang-Wiesenfeld sandpile.

Reference for this paper: M. Rypdal, K. Rypdal, *A stochastic model of the toppling activity in self-organized critical systems*, Preprint: arXiv:0710.4010v2 (2008), (Accepted by the New Journal of Physics).

1 Introduction

The existence of self-organized critical dynamics in complex systems has traditionally been demonstrated through numerical simulation of certain classes of cellular automata referred to as sandpile models [30, 19]. Non-linear, spatio-temporal dynamics is always essential for the emergence of SOC behavior, but the details of this dynamics for a specific natural system is often poorly understood and/or not accessible to observation. In many cases the information available is in the form of time-series of spatially averaged data like stock-price indices, geomagnetic indices, or global temperature data. For scientists who deal with such data a natural question to ask is: are there specific signatures of SOC dynamics that can be detected from such data?

In this letter we shall report some results which provide a partial answer to such a question. Some important statistical features of the toppling activity are common to most weakly driven sandpile models described in the literature, and these are used to formulate a stochastic model for the toppling activity signal. A benchmark case against which our results are tested, is a numerical study of the Bak-Tang-Wiesenfeld (BTW) sandpile [3]. A crucial step in our work is a re-scaling of the dynamical variables which allows a natural passage to the thermodynamic (continuum) limit. We demonstrate that this leads to new results concerning SOC scaling laws. We find that the probability density function (pdf) for the toppling activity is a stretched exponential, close to the Bramwell-Holdsworth-Pinton distribution [13], or close to a gaussian, depending on whether the sandpile is so slowly driven that avalanches are well separated, or it is driven so hard that several avalanches run simultaneously. The pdf for avalanche durations is unique in the thermodynamic limit, but is not a power law, unless we redefine the meaning of an avalanche to be the activity burst between successive times for which the activity rises above a positive threshold. Implementing such a threshold yields an exponent for the avalanche duration pdf of 1.63, in agreement with [48], but in contradiction to [40]. It also gives power-law quiet-time statistics as in [48] and thus refuting the claim in [9] that SOC implies power-law distributed avalanche durations, but Poisson-distributed quiet times.

The sandpile models considered in this short paper deal with a $d \geq 2$-dimensional lattice of N^d sites each of which are occupied by a certain integer number of quanta which we conveniently can think of as sand grains. The dynamics on the lattice is given by a toppling rule which implies that if the number of grains on a site exceeds a prescribed threshold, the grains on that site are distributed to its nearest neighbors. If the occupation number of some of these neighbors exceed the toppling threshold these sites will topple in the next time step, and the dynamics continues as an avalanche until all sites are stable. The details of this toppling rule can vary, but a useful theory for a broad class of natural phenomena should not be very

sensitive to such detail.

In natural systems the SOC dynamics is usually driven by some weak random external forcing. In sandpile models this can be modeled by dropping of sand grains at randomly selected sites at widely separated times. In numerical algorithms this is often done by dropping sand grains only at those times when no avalanche is running. This ensures that the drive does not interfere with the avalanching process. Usually it will then only take a few time steps from one avalanche has stopped until a new starts, so for a large system the quiet times between avalanches will appear insignificant compared to their durations.

A more physical drive would be to drop sand also during avalanches. If the dropping rate is slower than the typical duration of a system-size avalanche the drive would still not interfere with the avalanche dynamics, but the quiet times would depend on the statistical distribution of dropping times, which is typically a Poisson distribution. In many natural systems, however, avalanching occurs all the time, corresponding to a higher driving rate. In such cases, and also because there will always be noise in time-series data, we cannot identify the start and termination of an avalanche from a zero condition of our observable. In practice we have to define avalanches as *bursts* in the time series identified by a threshold on the signal [48]. In a sandpile simulation such bursts are correlated and therefore the quiet times between the bursts are power-law distributed even if the dropping of sand grains is chosen to be a Poisson process. Hence if focus is on modeling features that can be detected in observational data we shall think of avalanches as activity bursts starting and terminating at a non-zero threshold value. Moreover, one of the main results of this work is that power-law shape of the pdf for avalanche duration is true *only* if one defines avalanches in this way.

2 The stochastic model

The BTW model was first developed in the seminal paper [3], and is described in many monographs like [30, 19]. If z_i is the number of sand grains occupying the i'th site, the toppling rule is $z_i \to 0$ and $z_j \to z_i + z_i/4$ if z_i is overcritical and j is a nearest neighbor of i. Whenever the configuration has no overcritical sites a grain is added to a random site with uniform probability on the lattice. In a continuously driven system, grains are added at random times at a preset average rate, even when avalanches are running.

We shall assume that the lattice has linear extent $L = 1$ with N^d sites, so the thermodynamic limit $N \to \infty$ can be thought of as a continuum limit. The sandpile evolves in discrete time steps labeled by $k = 1, 2, 3 \ldots$, and the number of sites whose occupation number exceeds the toppling threshold at time k is called the toppling activity $x_N(k)$. The toppling increment is $\Delta x_N(k) \stackrel{\text{def}}{=} x_N(k+1) - x_N(k)$. Let us define two active sites as dynamically connected if they have at least one

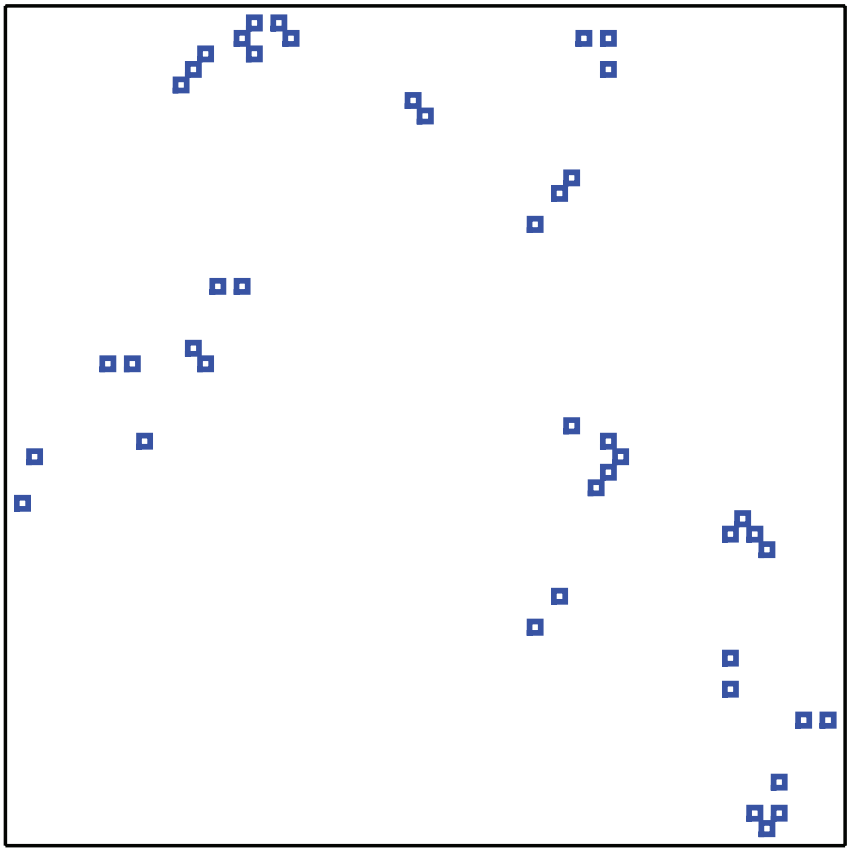


Figure 7.1: The blue squares mark the active (toppling) sites at a randomly chosen instant during an avalanche in a BTW lattice. Note that all the marked sites belong to one avalanche only, i.e. they all ere the result of one site going unstable some time in the past due to the feeding of a grain at a marginally stable site. The figure illustrates that a large avalanche in a large sandpile at a given time will consist of a large number of small disconnected clusters.

common nearest neighbor, and define a connected cluster as a collection of active sites which are linked trough such connections. From numerical simulations of sandpiles we observe that such clusters never consist of more than a few elements and that the instantaneous number of clusters n_N increases in proportion to x_N (see figure 7.1). This implies that at each time k we can label the clusters by $i = 1, \ldots, cx_N(k)$, where $c < 1$ is a constant depending on the specific toppling rule and the dimension d of the sandpile. We can then decompose the increment $\Delta x_N(k)$ into a sum of local increment contributions $\xi_{N,i}(k)$ produced by each of the clusters, i.e. $\Delta x_N(k) = \sum_{i=1}^{cx_N(k)} \xi_{N,i}(k)$. We think of the local increment contributions as random variables which take values in a finite sample space. Indeed, if each cluster i only consists of a single overcritical site, then $\xi_{i,N}$ takes values in the set

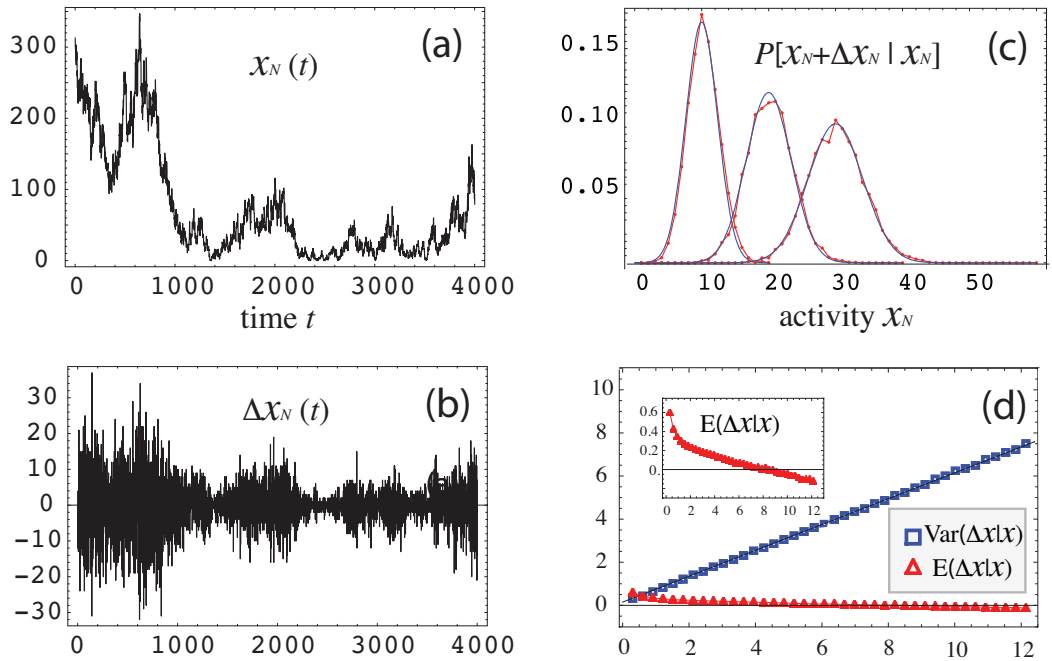


Figure 7.2: a): A realization of the toppling activity $x_N(t)$ in the BTW sandpile. b): The increments $\Delta x_N(t) = x_N(t+1) - x_N(t)$ of the trace in (a), showing that $\Delta x_N(t)$ is large when $x_N(t)$ is large. c): Conditional pdfs of $x_N + \Delta x_N$ for $x_N = 10$, 20, 30 respectively. d) The conditional mean and variance of Δx_N versus x_N.

$\{-1, 0, \ldots, 2d-1\}$.

As a first step to a stochastic model we make a mean-field assumption [54, 29], which impiles that $\xi_{N,i}(k)$ and $\xi_{N,j}(k)$ are statistically independent for $i \neq j$. Then the central limit theorem states that in the limit $N \to \infty$, $x_N(k) \to \infty$ the conditional probability density $P[\Delta x_N(k)|x_N(k)]$ of an increment $\Delta x_N(k)$, given $x_N(k)$, is Gaussian with variance $\sigma^2 x_N(k)$, where $\sigma^2 = c^2(E[\xi_{N,i}^2|x_N] - (E[\xi_{N,i}|x_N])^2)$. This has been verified numerically in the two-dimensional BTW-model as shown in figure 7.1. The figure demonstrates the need to introduce a conditional probability: The conditional variance of the increments is proportional to x_N and the conditional mean is not zero.

In fact, numerical simulations show that the the conditional mean increment, $E[\Delta x_N|x_N]$, is positive for small x_N, reflecting the natural tendency for the activity to grow when it is small. On the other hand the mean increment decays exponentially to zero for moderate x_N, and becomes negative when x_N is comparable to the activity of a system-size avalanche, reflecting the limiting influence of the finite system size. These effects will be incorporated as a drift-term correction to the

model, but for now we consider for simplicity of argument a Gaussian process with non-stationary increments and no drift term:

$$\Delta x_N(k) = \sigma \sqrt{x_N(k)}\, w(k)\,, \tag{7.1}$$

where $w(k)$ is a stationary Gaussian stochastic process with unit variance. In Sec. 3 we demonstrate from the numerical sandpile data that the *normalized* toppling process

$$W(k) \stackrel{\text{def}}{=} \sum_{k'=0}^{k} w(k') = \sum_{k'=0}^{k} \frac{\Delta x_N(k')}{\sigma\sqrt{x_N(k')}}$$

has the characteristics of a fractional Brownian walk with Hurst exponent $H \approx 0.37$ on time scales shorter than the characteristic growth time for a system-size avalanche, consistent with a power spectrum which scales like $f^{-1.74}$. Thus we model the normalized increment process as $w(k) = W_H(k+1) - W_H(k)$, where $W_H(k)$ is a fractional Brownian walk with Hurst exponent H. For the transition to the thermodynamic limit, where time will become a continuous variable, we can think about $W_H(k)$ as the result of a discrete sampling of the (continuous-time) fractional Brownian motion (fBm) $W_H(t)$. This process has the property $\langle |W_H(t+\tau) - W_H(t)|^2\rangle = \tau^{2H}$. We now have a stochastic difference equation

$$\Delta x_N(k) = \sigma \sqrt{x_N(k)}\, (W_H(k+1) - W_H(k)). \tag{7.2}$$

Numerical simulations show that $x_N \sim N^{D_1}$ [1], where $0 < D_1 \leq d$ can be interpreted as a fractal dimension of the set of active sites imbedded in the d-dimensional lattice space. This property is used to re-scale $x_N(k)$ such that it has a well-defined limit as $N \to \infty$. We also have to re-scale the time variable by letting $t = k\Delta t$, where $\Delta t = N^{-D_2}$. The value of D_2 will become apparent if we define the normalized activity variable $X_N(t) = N^{-D_1} x_N(t/\Delta t)$, such that the corresponding increment becomes

$$\Delta X_N(t) = N^{HD_2 - D_1}\, \sigma \sqrt{X_N(t)}\, \Delta W_H(t)\,, \tag{7.3}$$

where $\Delta W_H(t) = W_H(t+\Delta t) - W_H(t)$. A well-defined thermodynamic limit $N \to \infty$ requires $D_2 = D_1/2H$, for which Eq. (7.3), by introduction of the limit function $X(t) = \lim_{N\to\infty} X_N(t)$, reduces to the stochastic differential equation

$$dX(t) = f(X)\, dt + \sigma \sqrt{X(t)}\, dW_H(t), \tag{7.4}$$

[1] The BTW model does not exhibit perfect finite-size scaling [19] and hence the scaling $x_N \sim N^{D_1}$ is not valid for very large activity. The effect of imperfect scaling with increasing N can be built into Eq. 7.4 through an N-dependent drift term. However, the distributions of duration and size of sub-system size avalanches (defined by a threshold $X_c > 0$) is not sensitive to this feature of the BWT model. We have given a detailed treatment of this problem in [51].

where we have heuristically added a drift term $f(X)\,dt$ to account for the non-zero mean of the conditional increment. We take $f(X)$ to be an exponentially decaying function based on the numerical results from the sandpile. In the 2-dimensional BTW model we find that $D_1 \approx 0.86$ and hence $D_2 = 1.16$. This defines re-scaled coordinates $X_N = x_N/N^{0.86}$ and $t_N = k/N^{1.16}$.

3 The normalized toppling process is a fractional Brownian walk

A fractional Brownian walk is a self-affine stochastic process with gaussian increments and self-affinity (Hurst) exponent H. In strict mathematical terms a stochastic process $W(k)$ is self-affine only if the rescaled process $c^{-H}W(ck)$ is equal *in distribution* to $W(k)$ for any positive stretching factor c. This means that for any sequence of time points $k_1, \ldots, k_n$ and any positive c, the random variables $c^{-H}W(ck_1, \ldots, ck_n)$ have the same joint distribution as the random variables $W(k_1, \ldots, k_n)$. This definition can be used for theoretical purposes, but not as a practical tool to verify the self-affinity of actual time-series. This can be done, however, by means of multifractal analysis. The simplest method is to compute structure functions,

$$S(l, q) = E[|W(k + l) - W(k)|^q]. \tag{7.5}$$

For a multifractal process $S(l, q) \sim l^{\zeta(q)}$, where $\zeta(q)$ may be a non-linear function of q, while for a monofractal (self-affine) process $\zeta(q) = qH$, where H is the self-affinity exponent (see for instance the review [10]. This means that for a self-affine process we have linear relations between $\log(S(l, q))$ and $\log(l)$ and between $\zeta(q)$ and q. These linear relationships are shown for the normalized toppling activity of numerical sandpile simulations in figure 7.3 , and demonstrates that this process is a fractional Brownian walk with Hurst exponent $H = 0.37$.

4 Analysis of avalanches

A time series $X(t) \geq 0$, representing a succession of avalanches with zero quiet times, can be constructed numerically from the discrete-time version of Eq. (7.4) by integrating the equation using realizations of the fractional Gaussian noise process $\Delta W_H(t)$. At those times when $X(t)$ drops below zero we consider the avalanche as terminated, and a new, independent realization of $\Delta W_H(t)$ is generated and used to produce the next avalanche. From long, stationary time-series generated from the stochastic model and from the sandpile model this way, we can construct pdfs $\mathcal{P}(X)$ which turn out to give almost identical results for the two models (see figure 7.4). The shape of this pdf is universal in the thermodynamic limit: a stretched

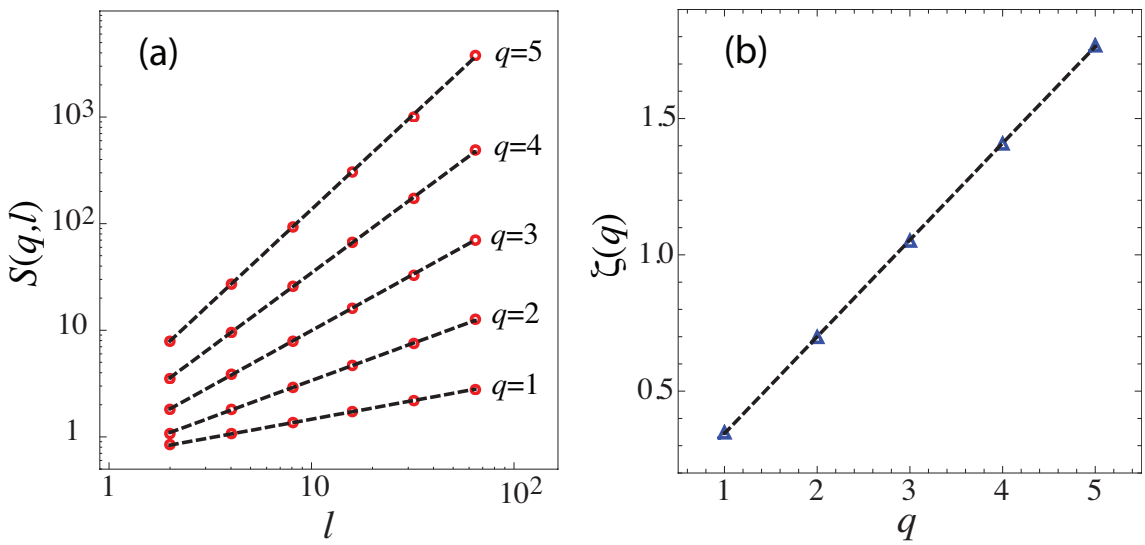


Figure 7.3: a) Structure functions $S(q,l)$ plotted versus l in a log-log plot for $q = 1,\ldots,5$. b) The scaling exponent for the structure functions (the slopes of the lines in (a)) plotted against q.

exponential $\mathcal{P}(X) \sim \exp(-aX^{\mu})$ with $\mu \approx 0.5$. A different pdf appears if the time-series are constructed by launching the avalanches at random times (Poisson-distributed) with characteristic time between launches shorter than the growth time of a system size avalanche. In this case several avalanches may run simultaneously, and for moderate strength of the drive $\mathcal{P}(X)$ from both models are close to the Bramwell-Holdsworth-Pinton distribution, which was claimed to be valid for the toppling-activity in the BTW-model in [13]. However, for stronger drive the pdf tends more towards a gaussian. These results are merely empirical, based on the numerical sandpile simulations. So far, there are no hard analytical results on the shape of the activity pdf.

Consider a solution of Eq. (7.4) with initial condition $X(0) = Y > 0$, and let $P(X,t)$ be the evolution of the density distribution in X-space of an ensemble of realizations of the stochastic process $X(t)$ all launched at activity $X = Y$ at time $t = 0$. Every realization $X(t)$ will sooner or later terminate at a finite time $t = \tau$ for which $X(\tau-1) > 0$ and $X(\tau) \leq 0$, and then we remove it from the ensemble. $P(X,t)$ contains information about all commonly considered avalanche characteristics. For example, it is easily found from from Eq. (7.4) that, on time scales shorter than the growth time of a system-size avalanche, $X(t)$ is a self-similar process with non-stationary increments and self-similarity exponent $h = 2H$ [51]. Hence the variance of $X(t)$ with respect to $P(X,t)$ will scale as $\sim t^{2h}$. That this relation holds for the 2-dimensional BTW model can easily be verified through numerical simulation (figure 7.5(a)).

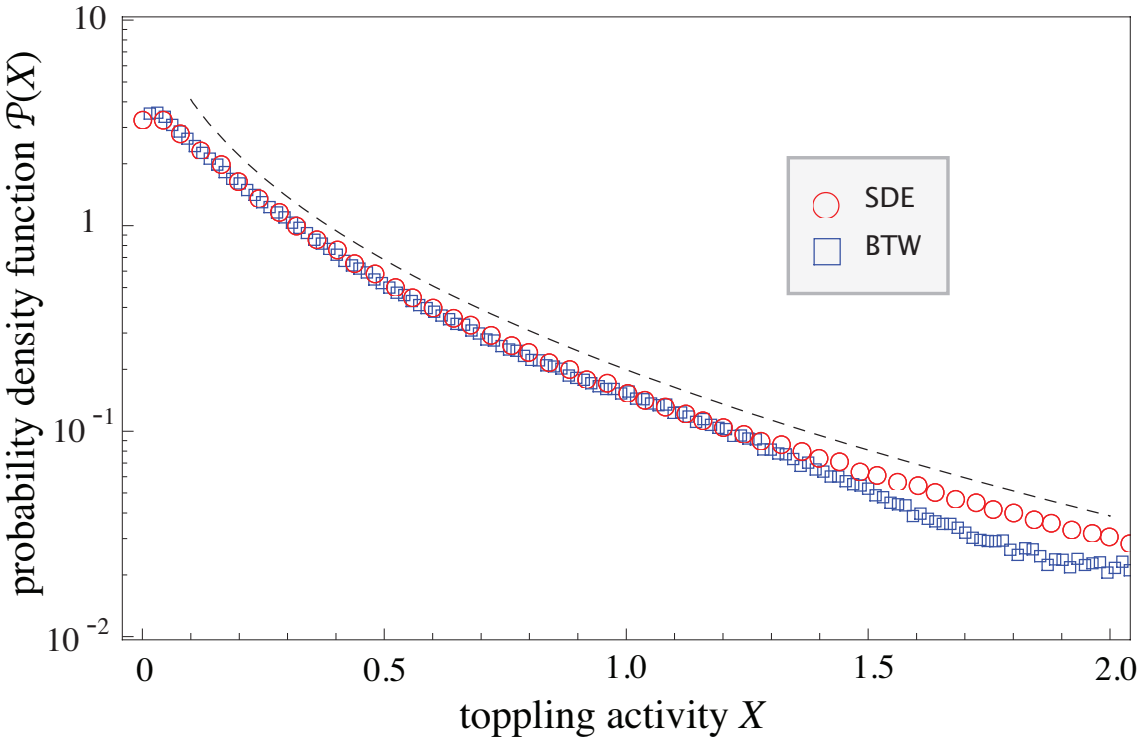


Figure 7.4: Logarithmic plots of $\mathcal{P}(X)$ from simulations of the 2-dimensional BTW sandpile for $N = 1024$. Also shown is $\mathcal{P}(X)$ found from simulations of Eq. (7.4), and a stretched exponential fit (dashed curve, vertically shifted for visibility). All pdfs are scaled to unit variance.

We can also compute the survival probability $\rho(\tau) = \int_0^\infty P(X,\tau)\,dX$, which is the probability that a realization of an avalanche has not terminated at the time τ. This function is related to the pdf for avalanche durations by $p_{\rm dur}(\tau) = -\rho'(\tau)$, so that $p_{\rm dur}(\tau)$ is a power law if and only if $\rho(\tau)$ is a power law. figure 7.5(b) shows the function $\rho(\tau)$ for numerical simulations of the BTW sandpile in the re-scaled coordinates X_N and t_N, demonstrating that the pdf for avalanche durations does *not* represent a power law. The power-law form $\rho(\tau) \sim \tau^{0.5}$ proposed in [40] can only be obtained as a tangent to the log-log plot of $\rho(\tau)$ at a given duration time τ, and the slope of this tangent depends crucially on the duration time τ for which this tangent is drawn.

The situation changes if we let all avalanches terminate when X drops below a small threshold $X_c > 0$ as proposed in [48]. In this case avalanche durations are the return times to the line $X = X_c$, and by changing coordinates to $Y = X - X_c$ we see that this corresponds to the return times to the time axis of the process given by the stochastic differential equation $dY(t) = \sigma\sqrt{X_c + Y(t)}\,dW_H(t)$. For small avalanches where $X(t) - X_c \ll X_c$ we can approximate this expression with

$dY(t) = \sigma\sqrt{X_c}\, dW_H(t)$, i.e. can approximate $Y(t)$ by a fractional Brownian motion with Hurst exponent H. Using the result of Ding and Yang [24] on the return times of a fractional Brownian motion we get $p_{\text{dur}}(\tau) \sim \tau^{2-H} = \tau^{-1.63}$.

Numerical simulations of the BTW model verify this result: The survival function $\rho(\tau)$ becomes a power law on time scales shorter than a system-size avalanche (see figure 7.6(a)), and the slope of the graph in a log-log plot is approximately -0.63, which corresponds to a scaling of the pdf for duration times on the form $p_{\text{dur}}(\tau) \sim \tau^{-1.63}$. The result is also reproduced by simulations of Eq. (7.4) with an exponentially decaying drift term. figure 7.6(b) shows the log-log plot of the pdf for duration times in the stochastic differential equation and a line with slope -1.63, demonstrating that the avalanche statistics in the BTW sandpiles is captured by the stochastic differential equation. From the scaling $\rho(\tau) \sim \tau^{-\alpha}$ we can deduce an exponent for the pdf of avalanche size as well. On the time scales where the toppling activity can be approximated by a fractional Brownian motion $W_H(t)$, the signal disperses with time as $X \sim t^H$, the size of an avalanche of duration τ scales like $S(\tau) \sim \int_0^\tau t^H\, dt \sim \tau^{H+1}$. Assuming that the pdf for avalanche size is on the form $p_{\text{size}}(S) \sim S^{-\nu}$, the relation $p_{\text{size}}(S)\, dS = p_{\text{dur}}(\tau)\, d\tau$ yields $\tau^{-\nu(H+1)+H} \sim \tau^{-\alpha-1}$, so

$$\nu = \frac{H+\alpha+1}{H+1} = \frac{2}{H+1}\,. \tag{7.6}$$

With $H = 0.37$ we obtain $\nu = 1.46$. The dependence of α and ν on H is the same as obtained in [56] and [18].

We also remark that if we omit the drift term and let $H = 1/2$ and $X_c = 0$ we obtain the so-called mean-field theory of sandpiles. In this case the stochastic differential equation has a corresponding Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2}{\partial X^2}(XP)\,.$$

If we solve this equation on the interval $[0, \infty)$ with absorbing boundary conditions in $X = 0$ we can obtain an analytical expression for $P(X, t)$, and from some straightforward algebra we find for large τ that $p_{\text{dur}}(\tau) \sim \tau^{-2}$ [51]. Since $X_c = 0$ we can not approximate the toppling activity by a Brownian motion on any scale and thus $X(t)$ diverges like $\sim t^h$, where $h = 2H$. By replacing H with $h = 2H$ in Eq. (7.6) we get $p_{\text{size}}(S) \sim S^{-3/2}$, in agreement with previous mean field approaches [54, 29].

5 Concluding remarks

We point out that the validity of Eq. 7.4 is not restricted to the BTW model. For instance, the equation has been verified for the Zhang model [51, 57], though

with a different Hurst exponent H. Time series of global quantities derived from numerical simulation of different sandpile and turbulent fluid systems can be shown to be adequately described by generalizations of Eq. 4, where H, the specific form of the "diffusion coefficient" in the stochastic term, and the drift term, all depend on the system at hand [51].

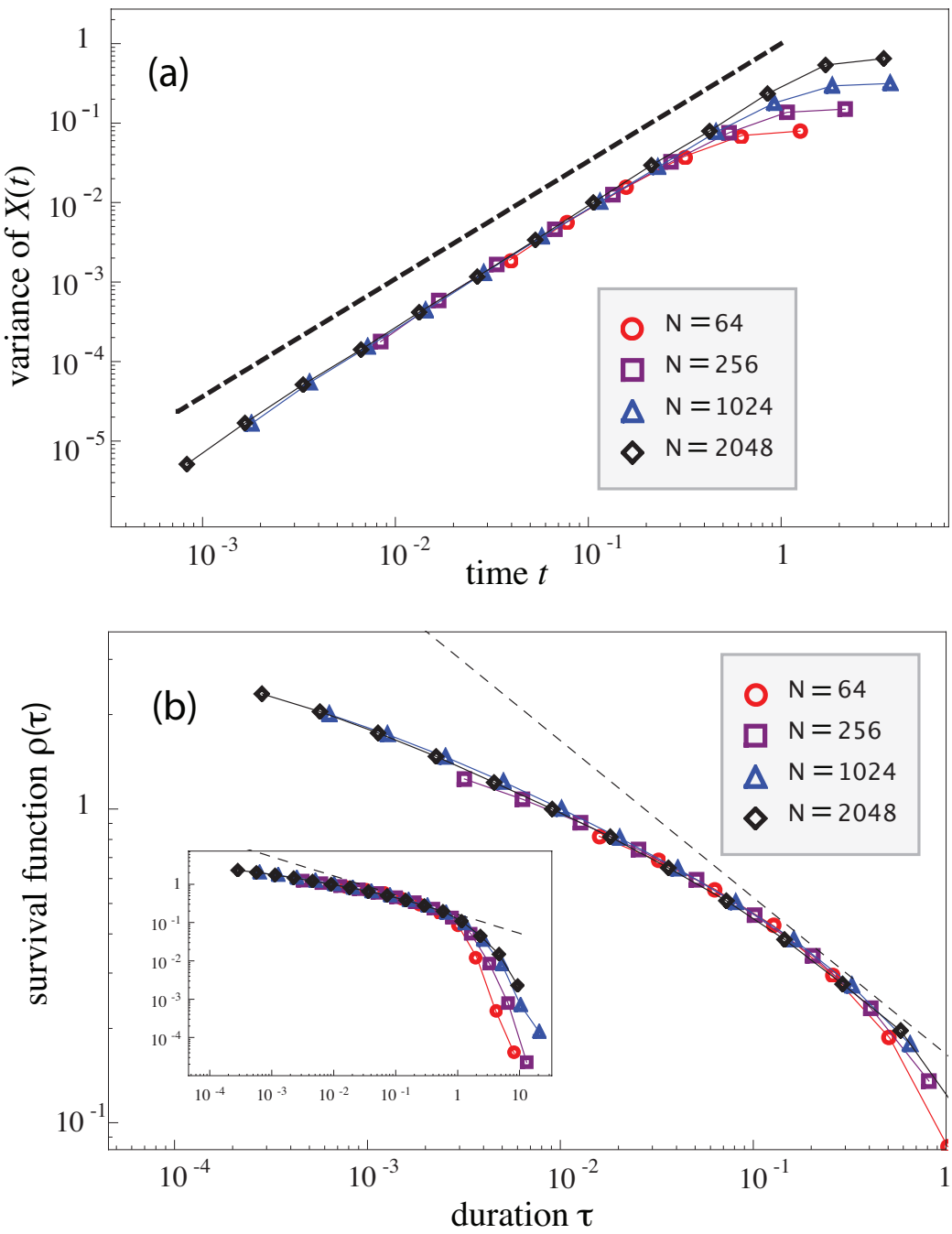


Figure 7.5: a) Double-logarithmic plots of the variance of $X(t)$ with respect to the pdf $P(X,t)$. The variance grows like t^{2h}, with $h = 2H = 0.74$ for times less than the duration of a system size avalanche. b) Double-logarithmic plots of the survival function $\rho(\tau)$ in the re-scaled coordinates X_N and t_N, demonstrating that the pdf of avalanche durations is not a power-law. The dotted line has slope -0.5.

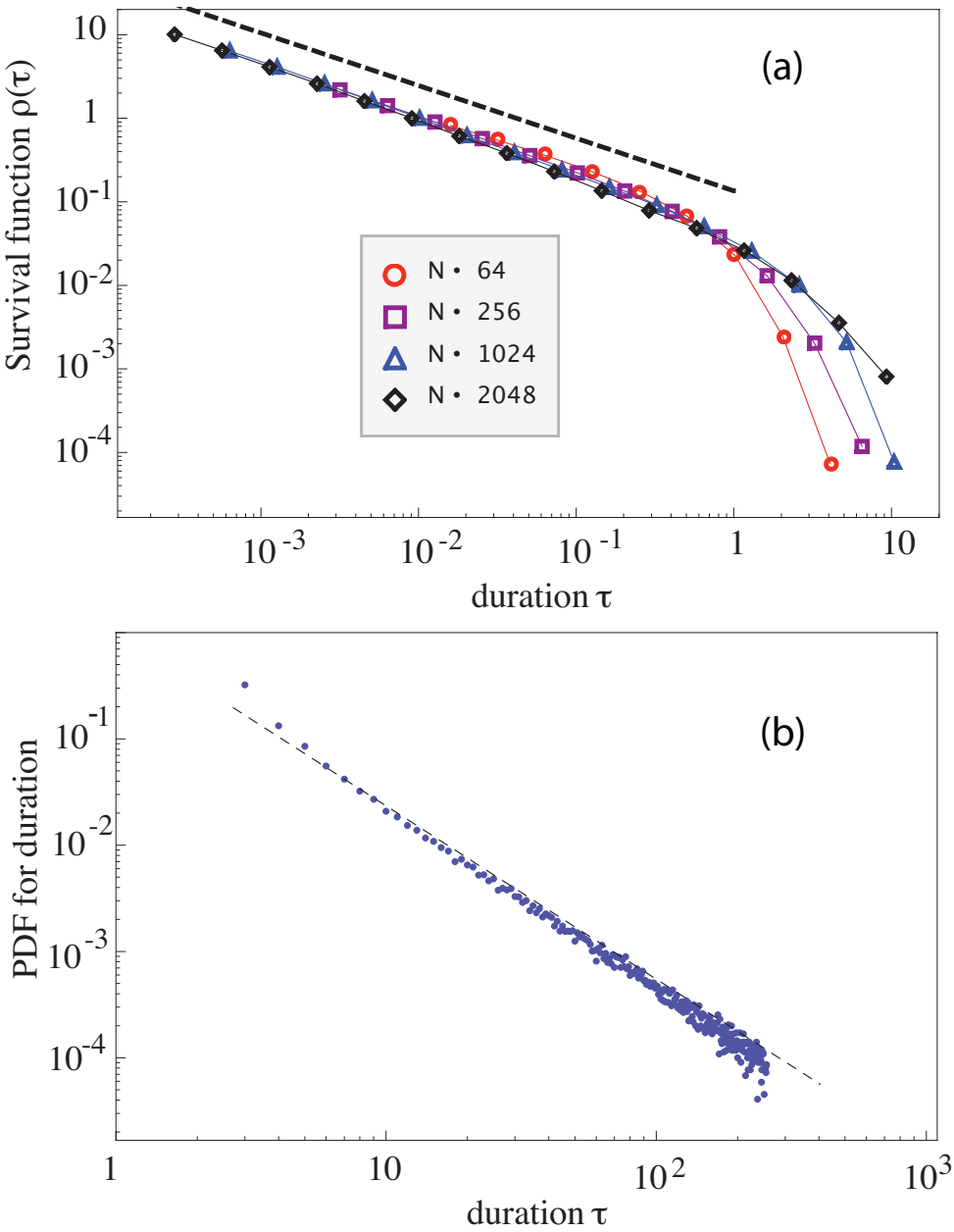


Figure 7.6: a) The survival function for the BTW sandpile in re-scaled coordinates X_N and t_N for $N = 64, 256, 1024, 2048$ when the durations are defined by putting a small threshold X_c on the toppling activity. The function shows power-law behavior with exponent -0.63 for avalanches smaller than system size. b) The pdf for duration times from simulations of Eq. (7.4) when avalanches are defined in the same way as for the sandpiles. The dotted line has slope -1.63.

Chapter 8

Modeling temporal fluctuations in sandpiles

Joint with K. Rypdal

Abstract: We demonstrate how to model the toppling activity in avalanching systems by stochastic differential equations (SDEs). The theory is developed as a generalization of the classical mean field approach to sandpile dynamics by formulating it as a generalization of Itoh's SDE. This equation contains a fractional Gaussian noise term representing the branching of an avalanche into small active clusters, and a drift term reflecting the tendency for small avalanches to grow and large avalanches to be constricted by the finite system size. If one defines avalanching to take place when the toppling activity exceeds a certain threshold the stochastic model allows us to compute the avalanche exponents in the continum limit as functions of the Hurst exponent of the noise. The results are found to agree well with numerical simulations in the Bak-Tang-Wiesenfeld and Zhang sandpile models. The stochastic model also provides a method for computing the probability density functions of the fluctuations in the toppling activity itself. We show that the sandpiles do not belong to the class of phenomena giving rise to universal non-Gaussian probability density functions for the global activity. Moreover, we demonstrate essential differences between the fluctuations of total kinetic energy in a two-dimensional turbulence simulation and the toppling activity in sandpiles.

Reference for this paper: M. Rypdal, K. Rypdal, *Modeling of temporal fluctuations in sandpiles*, Preprint: arXiv:08073416v1 (2008), (Accepted by Physical Review E).

1 Introduction

The aim of this paper is to present a consistent framework for the modelling of temporal fluctuations, including definition and computation of avalanche exponents, in sandpile (height) models such as the Bak-Tang-Wiesenfeld (BTW) and Zhang models [3, 57]. One of the defining properties of self-organized criticality is that avalanche duration and size are quantities subject to scaling, i.e. $p_{\mathrm{dur}}(\tau) \sim \tau^{-\alpha}$ and $p_{\mathrm{size}}(s) \sim s^{-\nu}$ [30, 19]. The calculation of the exponents α and ν in the thermodynamic limit $N \to \infty$ (N^d is the number of sites in the d-dimensional lattice) has proven to be a difficult task in dimensions $d = 2$ and $d = 3$. This is partly due to the lack of simple finite-size scaling in models such as the BTW model [19]. Moreover, it has recently been pointed out [50] that the difficulty may originate from the fact that the τ and s do not scale when defined in the traditional sense, which is to consider the duration of an avalanche as the time interval where toppling takes place between two successive zeroes in the toppling activity. In this paper we shall denote avalanches defined this way as *type-I avalanches*.

The scaling property can be restored, however, if one defines the duration of an avalanche as the time interval when the toppling activity $x(t)$ exceeds a prescribed threshold x_{th}. We shall use the term *type-II avalanches* when the start and end of an avalanche is determined via such a threshold criterion. The idea of using a threshold on the toppling activity in the definition of avalanches was first introduced in [48], where it was argued that any avalanche analysis of real-world activity time series must define avalanches from a threshold, since there is no way to uniquely determine whether a non-negative continuous-valued experimental quantity is actually zero, or just small. For type-II avalanches it can be shown by numerical simulation that the quiet times between avalanches in the BTW model are power-law distributed. For type-I avalanches the quiet times only depend on the statistics of the driver, which is usually assumed to be Poisson distributed.

The present work represents the first systematic investigation of type-II avalanche statistics for the BTW and Zhang sandpiles. We are particularly interested in the continuum limit where the system size $L = 1$ is considered fixed and the spatial resolution increases as $N \to \infty$. The power-law statistics of avalanche observables have cutoffs for large avalanches due to the finiteness of the system, but as we increase the resolution we see an increasing range of scaling for smaller avalanches. As $N \to \infty$ we keep the threshold fixed relative to the time-averaged activity $\langle x \rangle$, which scales as $\langle x \rangle \sim N^{D_1}$, where $0 < D_1 < 2$. For the BTW sandpile, numerical simulations yields $D_1 \approx 0.86$ [50]. This means that the number of overcritical sites corresponding to the threshold value diverges like $x_c \sim N^{D_1}$ in the limit $N \to \infty$.

We model the toppling activity in the continum limit by a stochastic differential equation [58] for a normalized toppling activity $X(t)$. In its simplest form this

equation is on the form

$$dX(t) = \sigma \sqrt{X(t)}\, dW(t)\,, \tag{8.1}$$

where $W(t)$ is the Wiener process. Without the factor $\sqrt{X(t)}$ on the right hand side we would simply have that $X(t) = \sigma W(t) + X_0$ is a Brownian motion with diffusion coefficient $D = \sigma^2/2$. This factor, however, gives rise to a non-uniform (X-dependent) diffusion coefficient $D = \sigma^2 X(t)/2$, and the stochastic process $X(t)$ will have non-stationary increments. This model can be perceived as a continuous version of the classical mean field theory of sandpiles [29]. We use its corresponding Fokker-Planck equation to derive that $\alpha = 2$ and $\nu = 3/2$ when avalanches are defined in the type-I sense. This is the same results obtained by mean field theory [29, 19, 30]. For type-II avalanches the effect of the non-uniform diffusion coefficient vanishes for avalanches of durations short compared to that of a system-size avalanche, and for the purposes of calculating α and ν we can assume that the the toppling activity is a standard Brownian motion. Solving the Fokker-Planck equation for Brownian motion (or equivalently using the known distribution of first return times in Brownian motion) we obtain $\alpha = 3/2$ and $\nu = 4/3$.

For the modeling of non-trivial sandpile models (BTW and Zhang) the stochastic differential equation takes the form

$$dX(t) = f(X)\, dt + \sigma \sqrt{X(t)}\, dW_H(t)\,. \tag{8.2}$$

Two important generalizations of the stochastic model are included here. First, from numerical simulations of sandpiles we find that for small activities $X(t)$ there is an effective positive drift term. This term is dominant for very small activity, since the diffusion term is negligible for very small $X(t)$ due to the X-factor in the diffusion coefficient. The positive drift term is X-dependent and quickly decreases as X increases, but strongly influences the avalanche statistics because it contributes to prevent avalanches from terminating when $X(t)$ approaches zero. We believe that this effect is responsible for destroying the scaling of avalanche duration and size (when these are defined in the type-I sense). However, if the drift term is small compared to the diffusion term for $X > X_{th}$, the drift term will not affect the avalanche statistics if one employs a threshold X_{th} to define type-II avalanches. This explains why scaling of size and duration is restored when when type-II avalanches are introduced. Another generalization, which is essential for the avalanche statistics, is the Hurst exponent H of the noise term. The mean-field approach to sandpiles implicitly assumes that $H = 1/2$. However, this is not the case for the BTW and Zhang models. Actually, analysis of numerical simulations of the sandpiles show that $H = 0.37$ for the BTW model and $H = 0.75$ for the Zhang model.

As in the mean-field model, the effect of the non-uniform diffusion coefficient vanishes as the threshold increases, and hence keeping the threshold X_{th} fixed and increasing N we can, for the purposes of computing α and ν for avalanches where

X never grows much greater than X_{th}, consider the toppling activity as a fractional Browian motion. Using the result of Ding and Yang [24], that the first return time in fractional Brownian motion scales like $\sim \tau^{H-2}$, we obtain the general results $\alpha = 2 - H$ and $\nu = 2/(1 + H)$.

Although the drift term and the non-uniformity of the diffusion coefficient are not important to calculate the avalanche exponents for type-II avalanches whose duration are short enough not to be limited by the finite system size, they are important on the time scales where the toppling activity is a stationary process. These are scales sufficiently long that the toppling activity of avalanches is limited by the boundaries. The stochastic equation (8.2) is fully equipped to handle these time scales. A good example of the applicability of these aspects of the stochastic model is the computation of the probability density function (PDF) of the temporal fluctuations in the activity signal itself.

For weakly driven sandpiles the PDFs of the fluctuations in the toppling activity are stretched exponentials. This result is reproduced by simulation of the stochastic model [50]. For stronger driving the activity exhibits fluctuations which are more confined around a mean value where the drive and dissipation balance each other. It has been claimed that the PDFs of the toppling activity in sandpiles are examples of universal Bramwell-Holdsworth-Pinton (BHP) distributions [13, 14], a certain class of asymmetric PDFs commonly seen in complex systems. Our sandpile simulations show that the BHP distributions can only be seen if one fine tunes the driving rate to a certain value, and for other driving rates the PDFs belong to a much wider class of distributions. For sufficiently strong drive Gaussian PDFs are observed. These can also be obtained from the stochastic model if one correctly models the drift term in this parameter range.

The rest of the paper is structured as follows: In Sec. 2 we explain and derive the stochastic model for the toppling activity. In Sec. 3 we compute the avalanche exponents for type-I and type-II avalanches for the mean field case ($H = 1/2$) by solving a Fokker-Planck equation, and for $H \neq 1/2$ we compute the avalanche exponents for type-II avalanches as a function of H. The results allow us to predict the avalanche exponents for sandpile models by computation of the Hurst exponents. These results are then tested against numerical simulations of the BTW and Zhang models and are shown to agree well. In Sec. 4 we present results which indicate that type-II avalanches exhibit so-called finite-size scaling, even though type-I avalanches do not.

In Sec. 5 we use the stochastic theory to calculate the PDFs of the toppling activity signal, both for strongly driven sandpiles and in the weak driving limit. The method is finally applied to the fluctuations in kinetic energy in a two-dimensional turbulence simulation. In this case the process is given by a different kind of stochastic differential equation:

$$dX(t) = b\,dt + c\,e^{a\,X}\,dW(t)\,. \tag{8.3}$$

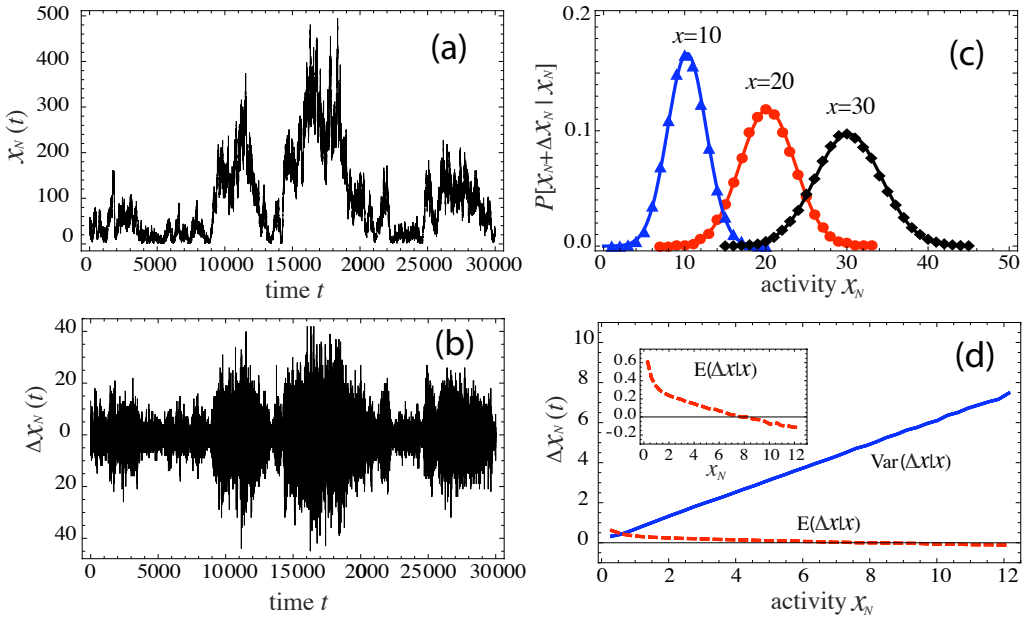


Figure 8.1: a) A realization of the toppling activity $x(k)$ in the BTW sandpile. b) The increments $\Delta x(k) = x(k+1) - x(k)$ of the trace in (a), showing that $\Delta x(k)$ is large when $x(k)$ is large. c) Conditional PDFs of $x + \Delta x$ for $x = 10, 20, 30$ respectively. d) The conditional mean and variance of Δx versus x.

This equation gives rise to a Fischer-Tippet-Gumbel (FTG) distribution [26, 21], which is very close to the BHP. The differences between the stochastic models for sandpiles activity and kinetic energy fluctuations in 2D turbulence may represent an essential distinguishing feature between 2D turbulent dynamics and the kind of avalanching dynamics which are observed in the classical sandpile models.

In Sec. 6 we summarize and conclude the work.

2 The stochastic model

Let $x(k)$ denote the number of overcritical sites at time step k in a sandpile. A common feature of (height-type) sandpile models is that the typical size of increments Δx is proportional to the square root of the toppling activity x. To be more precise, the conditional probability of an increment $\Delta x(k) = x(k+1) - x(k)$, given $x = x(k)$, is

$$P(\Delta x|x) = \frac{1}{\sqrt{2\pi\sigma^2 x}}\, e^{-\frac{\Delta x^2}{2\sigma^2 x}} \,. \tag{8.4}$$

In figure 8.1 this property is verified for the BTW model (it holds in the Zhang model as well). This result can be explained as follows: At a given time k there

are $x = x(k)$ overcritical sites, which we can enumerate $i = 1, 2, \ldots, x$. In the next time step $k \to k+1$ each site i distributes energy to its neighbors, and will usually (always in the BTW model) become subcritical. If none of the neighbors receive sufficient energy to become overcritical, the contribution to Δx from site i is $\xi_i = -1$. If exactly one of the neighbors become overcritical, then $\xi_i = 0$, and so on. For the two-dimensional models the maximal value of ξ_i is 4 (3 for the BTW model) since a site maximally can excite 4 neighboring sites. Hence, for this case, we consider ξ_i to be random variables with realizations in $\{-1, 0, 1, 2, 3, 4\}$. The randomness originates from the local configuration in the vicinity of the overcritical site, which for these purposes is considered to be random. In other words, we think of the configuration on the lattice as a random background.

As an approximation we consider the different realizations of ξ_i (at a fixed time k) as independent of each other. We also consider the distribution of ξ_i to be identical for all overcritical sites and for all times. In this approximation the total increment Δx can be written as a sum of independent, identically distributed, random variables $\Delta x = \xi_1 + \cdots + \xi_x$ and by the central limit theorem we have (8.4) provided that the local means $\langle \xi \rangle$ are zero. Then $\sigma^2 = \langle \xi^2 \rangle = (-1)^2 p_{-1} + \cdots + 4^2 p_4$, where $\mathbf{p} = (p_{-1}, \ldots, p_4)$ is the probability vector for the local increment processes.

If the local processes at different times k are independent of each other, then $x(k)$ is a Markov process which satisfies a stochastic difference equation

$$\Delta x(k) = \sigma \sqrt{x(k)}\, w(k)\,,$$

where $w(k)$ is a stationary, normalized, and uncorrelated Gaussian process, i.e. $w(k) = W(k+1) - W(k)$, where $W(t)$ is the Wiener process. Under a rescaling of time $t = k\,\delta t$ and $X(t) = x(k)\,\delta x$ we have

$$\begin{aligned}
\Delta X(t) &= X(t+\delta t) - X(t) = \delta x \Big(x(k+1) - x(k) \Big) \\
&= \delta x\, \sigma \sqrt{x(k)}\, w(k) = \delta x^{1/2}\, \sigma \sqrt{X(t)} \Big(W(k+1) - W(k) \Big) \\
&= \left(\frac{\delta x}{\delta t} \right)^{1/2} \sigma \sqrt{X(t)} \Big(W(t+\delta t) - W(t) \Big)\,.
\end{aligned}$$

In the last step we used the self-affinity of the Wiener process, $W(t/\delta t) \stackrel{d}{=} \delta t^{-1/2} W(t)$. For $\delta x = \delta t$ we have a well defined model in the limit $\delta t, \delta x \to 0$, namely the Itô stochastic differential equation Eq. (8.1).

The first generalization of this model is obtained if we relax the requirement that the local increment processes $\xi_i(k)$ at time k are independent of the local increment processes $\xi_j(k')$, $j = 1, \ldots, x(k')$ at previous times $k' < k$. In this case we need to model memory effects in the stationary Gaussian process

$$w(k) = \frac{\Delta x(k)}{\sigma \sqrt{x(k)}}\,.$$

From the power spectrum or the variogram of the activity signal from numerical simulation of the BTW and Zhang models we find that $w(k)$ can be accurately modelled as a colored noise characterized by a Hurst exponent H. That is, $w(k) = W_H(k+1) - W_H(k)$, with W_H being a normalized (diffusion coefficient $= 1$) fractional Brownian motion (fBm). If we perform the rescaling $t = k\,\delta t$ and $X(t) = x(k)\,\delta x$ in the case $H \neq 1/2$ we obtain

$$\Delta X(t) = \left(\frac{\delta x}{\delta t^{2H}}\right)^{1/2} \sigma \sqrt{X(t)} \left(W_H(t+\delta t) - W_H(t)\right),$$

and by requiring that $\delta x = \delta t^h$, with $h = 2H$, we obtain the stochastic differential equation

$$dX(t) = \sigma \sqrt{X(t)}\, dW_H(t)\,. \tag{8.5}$$

If we assume that the the stochastic process $X(t)$ is self-affine with self-affinity exponent h, i.e. $X(st) \stackrel{d}{=} s^h X(t)$, it is easy to verify that Eq. (8.5) is invariant with respect to the transformation $t \to st$ if $h = 2H$. Thus, the exponent $h = 2H$ is the self-affinity exponent of the process $X(t)$, where H is the Hurst exponent determining the color of the noise process driving the stochastic differential equation. Observe that the reason why $h \neq H$ is the non-stationarity of the increment process due to the the factor $\sqrt{X(t)}$ in Eq. (8.5). Note also that the case $H = 1/2$ corresponds to $h = 1$.

The self-affinity of $X(t)$ described by Eq. (8.5) implies that there is no upper bound on the fluctuations on increasing time scales, i.e. Eq. (8.5) describes the activity of an infinite sandpile where the activity is never influenced by the system boundaries. From a physical viewpoint, however, it is more interesting to consider the dynamics of a finite sandpile in the continuum (thermodynamic) limit, and for this purpose it is natural to let the scaling factor δx depend on N such that $X_N = x_N \delta x(N)$ is bounded in the limit $N \to \infty$, for instance such that $\lim_{N\to\infty} \max(X_N) = 1$. Such a bound on $X(t)$ can be obtained by the introduction of a drift term $f(X)\,dt$ leaving the stochastic equation in the form of Eq. (8.2), where $f(X)$ is negative for large X. The form of $f(X)$ can be found from sandpile simulations by computing the conditional mean $E(\Delta x|x)$ of the increments, and is shown in figure 8.1d. It appears that $f(X)$ is a decreasing function, positive for small X and negative for large X, and $f(X) = 0$ for a characteristic activity $X_c \sim 1$. Without the stochastic term the drift term establishes X_c as a stable fixed point for the dynamics.

Since $f(X = 0) > 0$ solutions of Eq. (8.2) with initial condition $X(0) > 0$ exist and are positive for all $t > 0$. This means that while Eq. (8.5) has solutions for which $X(t) = 0$ after a finite time (avalanches terminate), avalanches described by Eq. (8.2) will never terminate in the meaning $X(t) = 0$ (type-I avalanches). This signifies that type-I termination never occurs in the the continuum limit. If sandpiles

of increasing N are simulated, and the type-I avalanche durations are computed in the rescaled coordinates, the durations generally grow without bounds for increasing N. The reason is that the effective threshold for type-I termination in a discrete sandpile is $x_N = 1$, but in rescaled coordinates this threshold $X_N = x_N\,\delta x(N)$ goes to zero as $N \to \infty$. As this rescaled threshold vanishes the duration in rescaled time goes to infinity. Since the type-I termination is a discreteness effect, the resulting PDFs of avalanche durations depends on N (system discreteness) and are not power-laws. As we shall demonstrate later, the introduction of activity thresholds which are defined in the rescaled coordinates, and hence remain finite in the continuum limit, will give rise to PDFs of durations of type-II avalanches which converge to a specific power-law in this limit.

Eq. (8.2) remains valid also for sandpiles which are driven by continuous feeding of sand during avalanches. The drive adds a positive contribution to the drift function $f(X)$ for $X < X_c$, but a negative contribution for $X > X_c$, because for large activites $f(X)$ mainly accounts for the increased boundary losses. The result is a steeper $f(X)$, which tends to confine the activity closer to the fixed point X_c. On the other hand the increased drive also increases the diffusion coefficient (by increasing σ) due to a larger number of new active clusters initiated per unit time. The net effect is a positive shift of X_c and that the fluctuations in X are confined to a smaller region around X_c. For sufficiently strong drive the range of variation in $X(t)$ becomes so small that the diffusion coefficient does not vary much. It is nevertheless important to model it correctly in order to calculate the PDFs of the fluctuations in toppling activity.

3 Calculation of avalanche exponents

We denote the stochastic model Eq. (8.1) (which is Eq. (8.2) with $H = 1/2$ and $f(X) = 0$) the *mean field model* of sandpiles. This is because its underlying assumptions and the results derived from it coincide with what is known as the mean field solution of sandpiles in the literature [29]. Eq. (8.1), together with its corresponding Fokker-Plack formulation can be used to calculate the avalanche exponents α and ν. The idea is that each avalanche corresponds to a realization $X(t)$ with some initial condition $X(0) = X_0$ ($X_0 \ll X_{max}$). The avalanche propagates until the realization $X(t)$ terminates at $t = t_1$ in the meaning that $X(t) > 0$ for $t < t_1$ and $X(t_1) = 0$. Calculating the ratio of surviving realizations at different times t in an ensemble will provide information about the distribution of avalanche durations. The avalanche size distribution can then be obtained by using a general relationship between the self-affinity exponent $h = 2H = 1$ and the duration statistics.

The scenario outlined above can be mathematically formulated as follows: Let

$P(X,t)$ be the probability density function such that

$$\text{Prob}[X(t) \text{ exists and } b < X(t) < a] = \int_a^b P(X,t)\, dX\,.$$

Then the the probability $\rho(t)$ that an avalanche still runs after a time t (we call it the survival function) is given by

$$\rho(t) = \int_0^\infty P(X,t)\, dX\,, \tag{8.6}$$

and the probability density function for durations is $p_{\text{dur}}(\tau) = -\rho'(\tau)$. The density $P(X,t)$ can be calculated by solving the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{\sigma^2}{2}\,\frac{\partial^2}{\partial X^2}(X\,P) \tag{8.7}$$

on $X \in [0,\infty)$, $t \in [0,\infty)$, subject to an absorbing boundary condition $\lim_{X\to 0} XP(X,t) = 0$ and an initial condition $P(X,0) = \delta(X - X_0)$.

To correctly incorporate the absorbing boundary condition we let $U = XP$, and solve the corresponding Fokker-Planck equation for U with boundary condition $U(0) = 0$ to get

$$P(X,t) = \int_0^\infty G(X,Y,t)\, P(Y,0)\, dY\,,$$

where

$$G(X,Y,t) = \frac{1}{2}\sqrt{\frac{Y}{X}}\int_0^\infty J_1(s\sqrt{Y})\, J_1(s\sqrt{X})\, \exp\left(-\frac{\sigma^2\, s^2}{8}\, t\right) s\, ds\,.$$

Remark 8.1 *The most elegant way to obtain the solution is to use the integral transform pair*

$$\begin{aligned} \hat{F}(s) &= \frac{1}{2}\int_0^\infty F(X)\, J_1(s\sqrt{X})\,\sqrt{X}\, dX \\ F(X) &= \frac{1}{\sqrt{X}}\int_0^\infty \hat{F}(s)\, J_1(s\sqrt{X})\, s\, ds\,. \end{aligned}$$

Taking the transform of the Fokker-Planck equation we obtain the ODEs

$$\frac{\partial \hat{P}}{\partial t}(s,t) = -\frac{\sigma^2\, s^2}{4}\,\hat{P}(s,t)\,,$$

which we can solve and take the inverse transform. It is also easy to solve the Fokker-Plack equation for $U = XP$ by separation of variables.

If $P(X,0) = \delta(X - X_0)$ the solution of the absorbing boundary problem is $P(X,t) = G(X, X_0, t)$. Moreover

$$\lim_{X\to 0} P(X,t) = \frac{\sqrt{X_0}}{4} \int_0^\infty s^2\, J_1(s\sqrt{X_0})\, \exp\left(-\frac{\sigma^2\, s^2}{8}\, t\right) ds = \frac{4\, X_0}{\sigma^4 t^2}\, \exp\left(-\frac{2\, X_0}{\sigma^2\, t}\right). \tag{8.8}$$

From Eqs. (8.6), (8.7), (8.8), and the absorbing boundary condition at $X = 0$ we find that

$$\frac{d\rho}{dt} = -\frac{\sigma^2}{2} \lim_{X\to 0} P(X,t) = -\frac{2\, X_0}{\sigma^2 t^2}\, \exp\left(-\frac{2\, X_0}{\sigma^2\, t}\right).$$

Hence the PDF for avalanche durations τ is

$$p_{\rm dur}(\tau) = \frac{2\, X_0}{\sigma^2 \tau^2}\, \exp\left(-\frac{2\, X_0}{\sigma^2\, \tau}\right),$$

and for $\tau >> 2X_0/\sigma^2$ we have $p_{\rm dur}(\tau) \sim \tau^{-\alpha}$, with $\alpha = 2$.

This result crucially depends on the correct formulation of the Fokker-Planck equation. If the stochastic process $X(t)$ were a classical Brownan motion, the Fokker-Planck equation would have the form of the standard heat equation, and by performing the analogous calculations for this equation we get $p_{\rm dur}(\tau) \sim \tau^{-3/2}$. The same scaling relation ($\sim \tau^{3/2}$) is obtained if one (incorrectly) employs the Stratonovich formulation of the Fokker-Planck equation rather than the Itoh form [58].

From the derivation of the stochastic model we have seen that the process $X(t)$ has a self-affinity exponent $h = 2H$ (for $H = 1/2$ we have $h = 1$). This means that $X(t)$ disperses like $\sim t^h$. For long avalanches this implies that the size of the avalanche scales as

$$s = \int_0^\tau X(t)\, dt \sim \int^\tau t^h\, dt \sim \tau^{h+1}\,. \tag{8.9}$$

This property is easy to check directly by studying the relation between duration and sizes of avalanches in sandpiles. We find that the relation holds for large avalanches, whereas for very small avalanches $s \sim \tau^1$. If the survival function scales as $\rho(\tau) \sim \tau^{-\delta} = \tau^{-\alpha+1}$, then using (8.9) together with $p_{\rm size}(s)\, ds = p_{\rm dur}(\tau)\, d\tau$ yields that if $p_{\rm size}(s) \sim s^{-\nu}$, then

$$\nu = \frac{1 + h + \delta}{1 + h} = \frac{\alpha + h}{1 + h}\,. \tag{8.10}$$

In the case $H = 1/2$ ($h = 1$) and $\alpha = 2$ this gives $\nu = 3/2$.

This mean-field solution is known to be quite correct for the random neighbor sandpile model [20], but numerical simulation shows that is fails for type-I as well as type-II avalanches in the BTW and Zhang models. The computation of these exponents are shown for these models in figures 8.2, 8.3, and 8.4. The computation

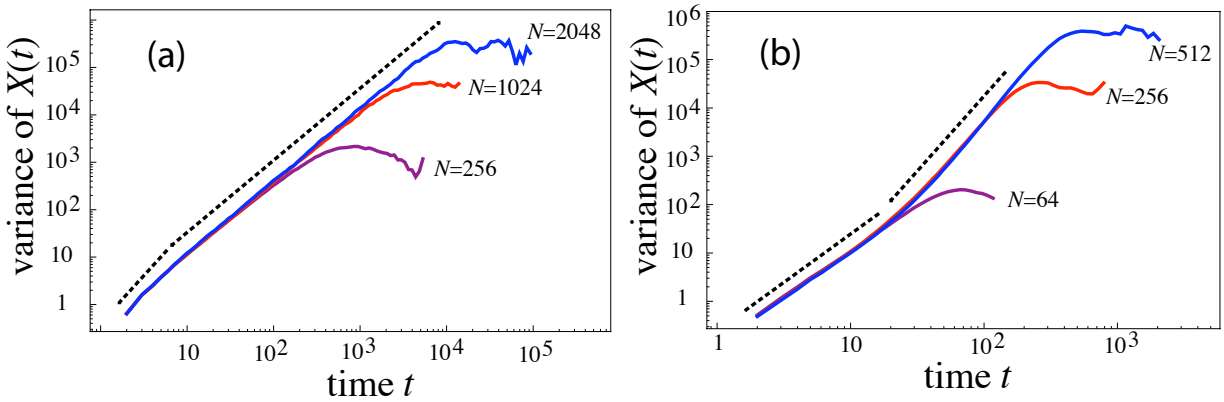


Figure 8.2: a) Determining $h = 2H$ in the Zhang model through the calculation of the variance of the activity $X(k)$ in avalanches still running at time t. On shorter time scales there is a Hurst exponent $H = 0.5$ (the first dashed line has slope $2h = 4H = 2$) on longer time scales the Hurst exponent is $H = 0.75$ (the second dashed line has slope $2h = 4H = 3$). b) Determining $h = 2H$ in the BTW model by the same method as in (a). In this case we have $H = 0.37$.

of h presented in figure 8.2 for type-I avalanches yield $H = h/2 = 0.37$ for the BTW model, and $H = 0.75$ for the Zhang model. This invalidates the Fokker-Planck formulation, which can be strictly justified only for a white noise source term ($H = 1/2$). This is one reason for the failure of the mean-field approach, but there are also others.

For avalanche duration and size type-I avalanches do not yield good power-law PDFs. The reason for this was discussed in the previous section: letting avalanches terminate when $x_N(t) = 0$ in a sandpile with N^d sites corresponds to using an effective threshold for termination in the rescaled coordinates $X_N(t)$ which goes to zero as $N \to \infty$. The termination process depends on the discreteness of the system and one cannot expect convergence to scale-invariant behavior in the continuum limit.

For type-II avalanches Figs. 8.3 and 8.4 show good scaling for duration and size in BTW and Zhang models, but the scaling exponents differ from those of the mean-field approach. The discrepancy is partly due to the fact that the sandpile models have $H \neq 1/2$, but is also related to the observation that the introduction of a finite termination threshold X_{th} modifies the exponent h as it appears in Eqs. (8.9) and (8.10). In the following we shall demonstrate that it is possible to obtain analytical results for type-II avalanches from the stochastic model which are in agreement with the corresponding sandpile simulations.

First we observe that omission of the drift term will only have effect on avalanches

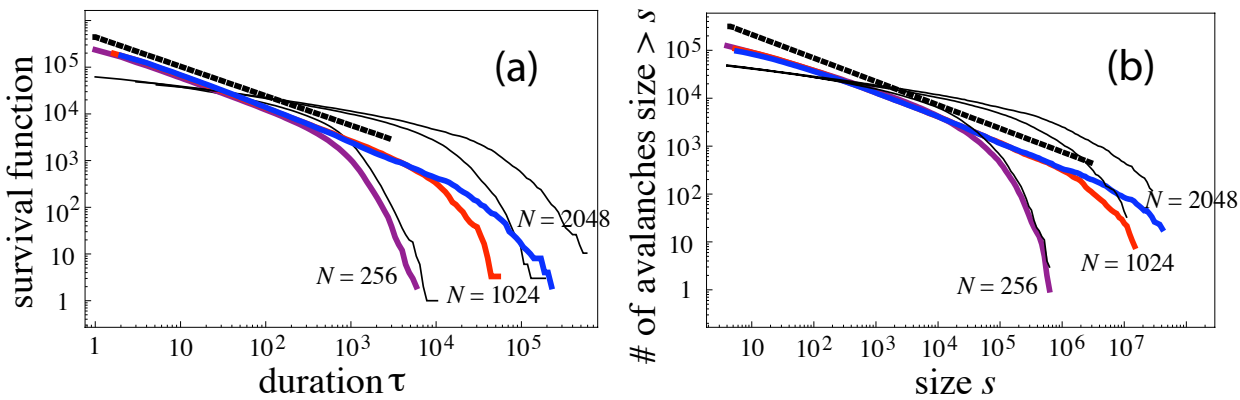


Figure 8.3: a) The survival function of avalanches in the BTW model computed with thresholds $\langle X \rangle/3$ (type-II) and without thresholds (type-I, the dotted curves). The dashed line corresponds to $\alpha = 2 - H = 1.63$. b) The probability density function of avalanches with size $> s$ in the BTW model computed with thresholds $\langle X \rangle/3$ and without thresholds (the dotted curves). The dashed line corresponds to $\nu = 2/(1 + H) = 1.49$.

which are so large that they are strongly limited by boundary dissipation. Next, we notice that the effect of the positive drift term for small activities is eliminated for type-II avalanches if $X_{th} > X_c$. In other words, it is only the cut-offs of the power-law PDFs due to finite system size which are lost by this omission. Thus, by considering a threshold X_{th} which is not much smaller than the mean activity $\langle X \rangle$, and considering only avalanches which are sufficiently small not to be influenced by the boundary dissipation, we have $X(t) \sim X_{th}$ and thus can justify the substitution $X(t) \to X_{th}(t)$ on the right hand side in Eq. (8.5). This equation then reduces to the equation for an fBm with Hurst exponent H. For an fBm the duration of avalanches is given by the return time statistics for $W_H(t)$, which is known to scale like τ^{H-2} [24], i.e. we have

$$\alpha = 2 - H \,. \tag{8.11}$$

Now we need to address a slightly subtle point: the exponent h, as defined in Eq. (8.9), is $h = 2H$ for times τ so large that $X(\tau) \gg X(0)$. Only on such time scales will the effect of the factor $\sqrt{X(t)}$ in the stochastic term show up in the scaling. However, this makes sense only for type-I avalanches where we can choose $X(0) \ll \langle X \rangle$. For type-II avalanches we have $X(0) \approx X_{th}$, and it is more natural to consider the opposite limit where τ is so small that $\hat{X}(\tau) \equiv X(\tau) - X_{th} \ll X(0) \approx X_{th}$. On these time scales the activity measured relative to the threshold level scales like an fBm with Hurst exponent H, i.e. $\hat{X}(t) \sim t^H$, because $X(t) \approx X_{th}$. The avalanche size of type-II avalanches defined as $\hat{s}(\tau) \equiv \int_0^\tau \hat{X}(t)\, dt$ then scale as $\sim \tau^{1+H}$. Thus, for type-II avalanches we have $h = H$, and hence Eq. (8.10) for this

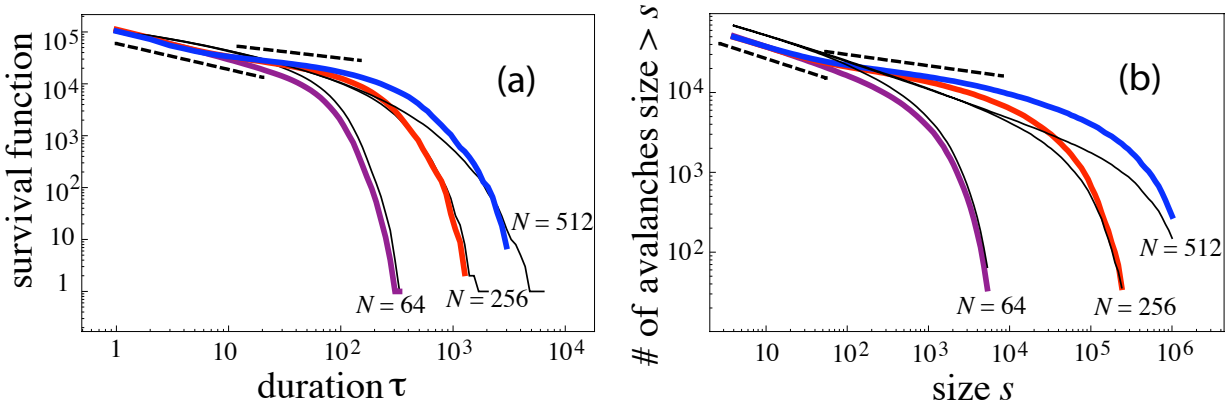


Figure 8.4: a) The probability of having avalanches of duration $> \tau$ in the Zhang model computed with thresholds $\langle X \rangle/3$ and without thresholds (the dotted curves). The first dashed line corresponds to $\alpha = 2 - H = 1.50$ obtained with $H = 0.50$, and the second dashed line corresponds to $\alpha = 2 - H = 1.25$ obtained with $H = 0.75$. b) The probability of having avalanches with size $> s$ in the Zhang model computed with thresholds $\langle X \rangle/3$ and without thresholds (the dotted curves). The first dashed line corresponds to $\nu = 2/(1 + H) = 1.33$ obtained with $H = 0.50$, and the second dashed line corresponds to $\nu = 2/(1 + H) = 1.14$ obtained with $H = 0.75$.

case reduces to

$$\nu = \frac{2}{1 + H} \,. \tag{8.12}$$

In the mean field limit $H = 1/2$ these exponents reduce to those given in the right hand column in Table 8.1. The results obtained analytically by approximating the coefficient on right hand side in the stochastic equation by its threshold value has been verified by numerical solutions of the full equation. Similar results for fractional Brownian motion have been obtained in [56] and [18].

Table 8.1: Exponents in mean-field solutions ($H = 1/2$) of the stochastic equation with zero threshold (type-I) and large threshold (type-II).

	type-I avalanches	type-II avalanches
H	1/2	1/2
h	1	1/2
α	2	3/2
ν	3/2	5/4

Eqs. (8.11) and (8.12) show that the calculations of α and ν reduces to determining the Hurst exponent of the normalized increment process

$$w(k) = \frac{\Delta x(k)}{\sqrt{x(k)}} ,$$

which can easily be constructed from the toppling activity signal. The corresponding Hurst exponent of the motion

$$W(k) = \sum_{i=0}^{k} w(k)$$

can be calculated using standard techniques such as taking the power spectrum or constructing variograms. Alternatively we can find H by computing the variance of $X(t)$ in an ensemble of realizations starting with small initial values of X. This variance scales like $\langle X^2(t) \rangle \sim t^{2h} = t^{4H}$. figure 8.2 shows this computation for the BTW and Zhang models. Since we find $H = 0.37$ for the BTW model the type-II scaling exponents are $\alpha = 1.63$ and $\nu = 1.49$. This is verified by numerical calculation of α and ν using a threshold $\langle x \rangle/3$, as shown in figure 8.3.

In the Zhang model the process $w(k)$ is more complicated. figure 8.2 shows that $W(k)$ has a Hurst exponent $H = 0.5$ on short time scales and a different Hurst exponent $H = 0.75$ on longer timescales. This means that short avalanches should satisfy the mean-field solution $\alpha = 1.5$ and $\nu = 1.33$, whereas longer avalanches should have exponents $\alpha = 1.25$ and $\nu = 1.14$. All of these predictions are verified by the direct calculation of α and ν as shown in figure 8.3.

For increasing N the long-avalanche scaling dominates an increasing portion of the graph, so in the continuum limit the mean-field solution only prevails at infinitely small scales in rescaled coordinates. It is however a nice verification of our method to see that the relation between the avalanche exponents and the Hurst exponent correctly predicts the avalanche exponents on short time scales as well.

These results on the BTW and Zhang models are summarized in table 8.2.

Remark 8.2 *The numerical simulations of the Zhang model are run using the standard toppling rule $z_i \to 0$ and $z_j \to z_i + z_i/4$ if z_i is overcritical and j is a nearest neighbor of i. Whenever the configuration has no overcritical sites a random site i is chosen with respect to uniform probability and a mass ϵ is added to this site: $z_i = z_i + \epsilon$. In the simulations presented in this paper we use $\epsilon = 0.1$. In the strongly driven Zhang models presented in Sec. 5 the feeding times (which can now be during avalanches) are Poisson distributed and we have used $\epsilon = 0.25$.*

For the BTW model we have used the standard toppling rule $z_i \to z_i - 4$ and $z_j = z_j + 1$ through out the paper. As usual a mass $\epsilon = 1$ is added to a random site whenever the configuration is stable.

Table 8.2: Exponents for type-II avalanches computed from $H = 0.37$ in the BTW model, and $H = 0.50$ (short time scales) and $H = 0.75$ (long time scales) in the Zhang model.

	BTW	Zhang (short times)	Zhang (long times)
H	0.37	0.50	0.75
α	1.63	1.50	1.25
ν	1.49	1.33	1.14

4 Finite-size scaling for type-II avalanches

A systematic technique for the determination of the avalanche exponents α and ν is the so called moment analysis [19]. We illustrate how this works for the size distribution of the BTW model. The analysis confirms our result $\nu = 0.49$ for the BTW model for type-II avalanches.

Let $p_{\text{size}}(s; N)$ be the PDF of avalanche size s in the two-dimensional BTW model with N^2 sites and a threshold $\langle x \rangle/3$. Let us assume finite-size scaling (FSS). This means that for $N, s \gg 1$ we have

$$p_{\text{size}}(s; N) \propto s^{-\nu}\, G(s/N^D)\,, \tag{8.13}$$

for some exponent $D > 0$, where $G(r)$ falls off quickly for $r > 1$. From Eq. (8.13) we observe that

$$\langle s^q \rangle \propto N^{D(1+q-\nu)} \int_{1/N^D}^{\infty} r^{q-\nu}\, G(r)\, dr\,.$$

The integral tends to a constant as $N \to \infty$ so we have $\langle s^q \rangle \sim N^{D(1+q-\nu)}$. Thus, if we plot $\langle s^q \rangle$ versus N in a log-log plot the slope of a fitted straight line will give us an estimate of the exponent $\zeta(q) = D(1+q-\nu)$ for $q = 1, 2, 3, \ldots$. figure 8.5a shows the computation of moments of s as a function of N for type-II avalanches obtained from simulation of the BTW model, and figure 8.5b shows the exponent $\zeta(q)$ versus q. We find that $\zeta(q) \sim q^{2.81}$, hence that $D = 2.81$. Moreover, when plotted in a log-log plot, the intersection of $\zeta(q)$ with the first axis corresponds to the value $\nu - 1$. Hence, based on our predictions this intersection should be in the point 0.49. This value is plotted as a dotted horizontal line in figure 8.5b, confirming our result with good accuracy.

Strictly, this method requires that we have data collapse when re-scaling the avalanche sizes by s/N^D. The shape of the scaling function $G(r)$ can be seen by plotting $s^\nu p_{\text{size}}(s; N)$ versus s/N^D. This is shown in figure 8.6. The data collapse for the available system sizes is not perfect, but it is better than for type-I

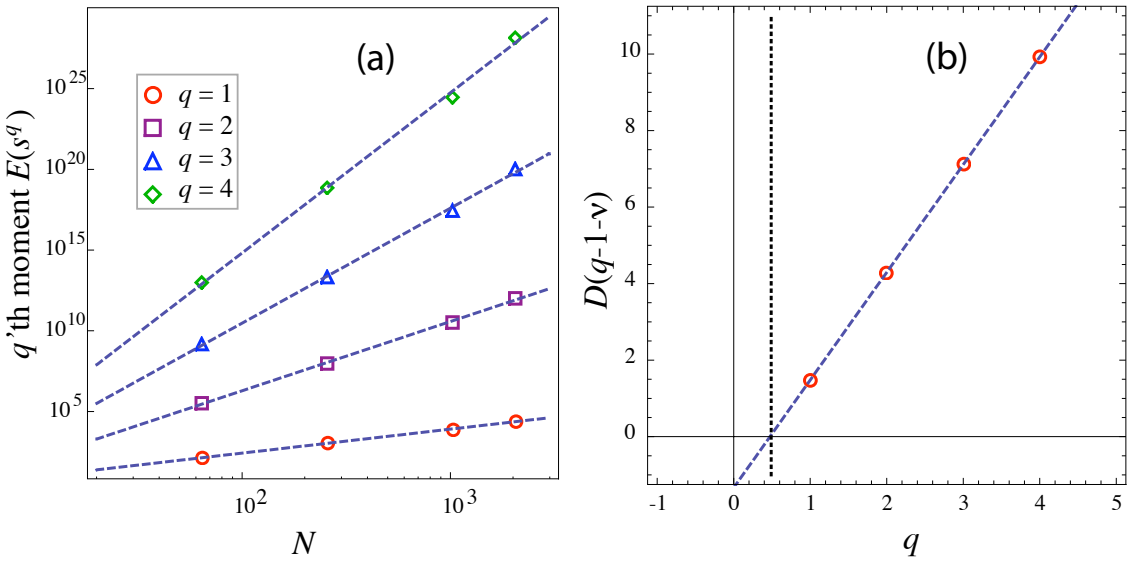


Figure 8.5: a) The moments $\langle s^q \rangle$ plotted as functions of N for the size distribution of the BTW model ith respect to a threshold $\langle X \rangle/3$. b) The shape of the structure function $\zeta(q)$. The slope of the line is $D = 2.81$ and the intersection with the first axis is $\nu - 1 = 0.49$. This value is indicated by the dotted vertical line.

avalanches. In fact, it seems that we might have convergence to a N-independent scaling function as $N \to \infty$, and that the lack of data collapse observed here is simply due to the fact that we are not able to simulate sufficiently large sandpiles. Thus the general lack of scaling for for type-I avalanches, including the finite-size scaling, seems to be restored for type-II avalanches.

5 The PDFs of the toppling activity

When calculating the PDFs of the toppling activity signal we have to distinguish between the weakly and strongly driven sandpiles. For the weakly driven sandpiles the toppling activity covers a large range and the high-activity tail of the PDF decays like a stretched exponential. This property can be reproduced by simulations of Eq. (8.2) where $f(X)$ decays exponentially to zero as X increases. figure 8.7 compares the PDF obtained from such a simulation of the stochastic differential equation (run with $H = 0.37$) with the PDF of the weakly driven BTW model. The stochastic model is run by initiating new avalanches (realizations) whenever the previous avalanche terminates, thus representing the classical slow drive of the sandpile model. Since the Hurst exponents of the Zhang and BTW models are different from 1/2, the application of the Fokker-Planck equation can not be used to

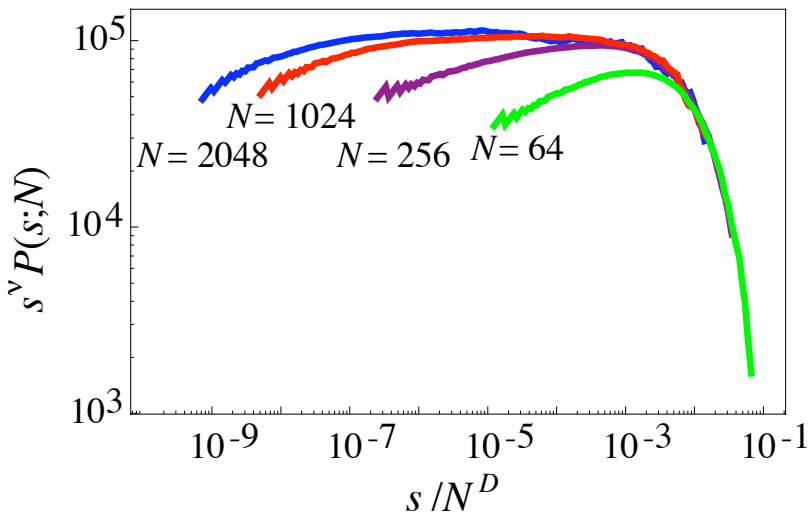


Figure 8.6: Attempted data collapse for the avalanches in the BTW model with respect to a threshold $\langle X \rangle/3$. We have plotted $s^{\nu-1}\int_s^\infty p_{\text{size}}(s')\,ds'$ as a function of s/L^D with $D = 2.81$ for $N = 64, 256, 1024, 2048$.

calculate the shape of the PDFs, and thus we have to rely on numerical simulations.

It is also interesting to study the PDFs of toppling activity in strongly driven sandpiles. In particular in the light of the recent claims that this toppling activity belongs to a class of fluctuating global quantities with a universal non-gaussian shape [13, 14]. These PDFs are reminiscent of distributions derived from the extreme value theory of statistics [26, 21], which deals with a sequence of n identically distributed, independent random variables $y_1, \ldots, y_n$. If the tail of the PDF of these variables decays faster than any power-law, the PDF of the m'th largest value drawn from each of M realizations of this sequence converge to the Gumbel class of stable distributions in the limit $M \to \infty$:

$$G_k(y) = K(e^{-\xi(y-s)}e^{-e^{-\xi(y-s)}})^m. \tag{8.14}$$

The distribution for the largest value in each realization ($m = 1$) is often called the Fischer-Tippet-Gumbel distribution (FTG). The FTG with zero mean and unit variance requires $K = \xi = \pi/\sqrt{6}$ and $s = \gamma\sqrt{6}/\pi$, where $\gamma \approx 0.58$ is the so-called Euler constant.

A distribution obtained from the spin-wave approximation to the 2D XY model for equilibrium crititical fluctuations in a finite-size magnetized system is the so-called Bramwell-Holdsworth-Pinton (BHP) distribution [14], which corresponds to Eq. (8.14) with k having the non-integer value $m = \pi/2 \approx 1.57$. With $K = 2.16$, $\alpha = 1.58$, $\xi = 0.93$ and $s = 0.37$ this distribution has zero mean and unit variance.

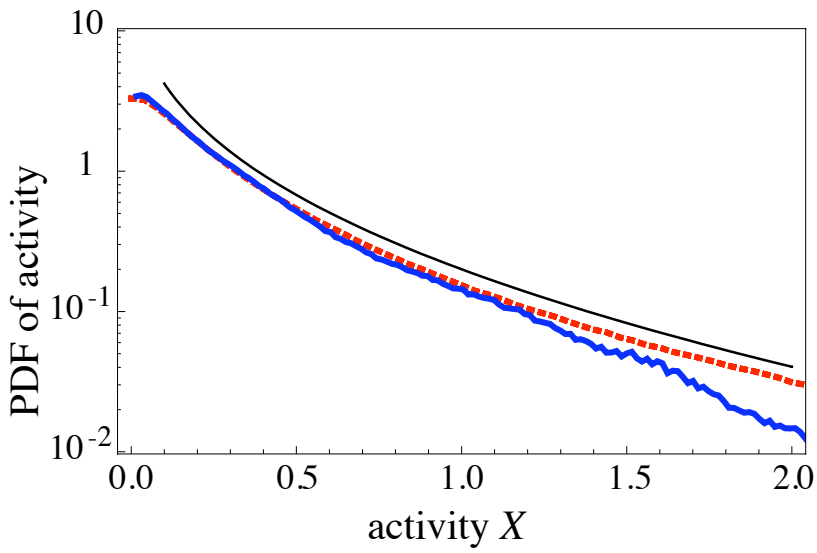


Figure 8.7: The PDF of the fluctuations in the toppling activity of the slowly driven BTW sandpile together with the corresponding PDF produced from the stochastic model. The dashed line is a fitted stretched exponential.

The difference between the normalized FTG and BHP distributions is rather small, and it is difficult to distinguish between the two based in experimental and numerical data.

Analysis of our simulations show the claim that the toppling activity in strongly driven sandpiles has a PDF similar to the BHP or FTG distributions is wrong, unless the driving rate of the system is fine tuned to some particular value. Actually, the normalized PDFs of the toppling activity is only insensitive to variation of the driving rate in the limits of weak and strong drive. In the limit of weak drive the PDFs are stretched exponentials as shown in figure 8.7 and in the limit of strong drive the PDFs are close to Gaussian.

figure 8.8 shows the PDFs of the toppling activity in the strongly driven Zhang model where the feeding rate is a Poisson process with characteristic time scales $\lambda = 0$ (feeding one unit ϵ of mass in every time step), $\lambda = 2.3$, and $\lambda = 3.0$. For $\lambda = 0$ and $\lambda = 2.3$ the sandpile is running, in the meaning that avalanches never terminate. For $\lambda = 3$, the sandpile is no longer running, and we see a deviation from the Gaussian shape in the left tail of the PDF.

The Gaussian PDFs for strongly driven sandpiles can actually be explained in terms of the stochastic model in Eq. (8.2). The idea is that the range of fluctuations is confined by the drift term $f(X)$, which now has a different shape than in the slowly driven sandpile. figure 8.9 shows the shape of both the diffusion coefficient $D(X) = \sigma^2 X/2$ and the drift term for the strongly driven Zhang model. The

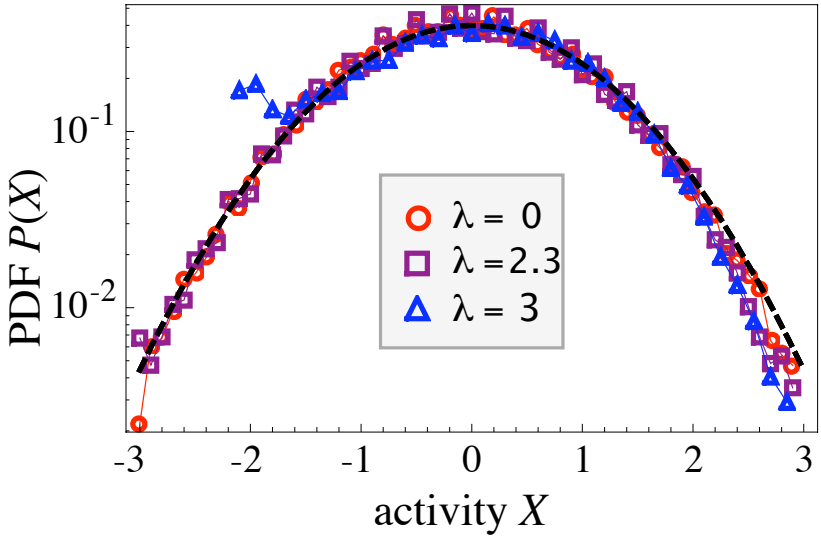


Figure 8.8: The PDF of the fluctuations in the toppling activity of the strongly driven Zhang sandpile driven BTW sandpile for different values of λ (see text). The dashed line is a fitted Gaussian.

diffusion coefficient has the same form as for the slowly driven sandpile, whereas the drift term is well approximated by a parabola: $f(X) = -aX^2 + bX + c$. As explained in Sec. 1 this drift term confines the fluctuations in toppling activity to a bounded region around the positive root X_c of $f(X)$.

Due to the higher rate of random feeding, the memory effects described by the Hurst exponent $H \neq 1/2$ becomes inessential in the strongly driven sandpile, allowing us to give an approximate description of the time dependent PDF through the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial X}\Big(f(X)\,P\Big) + \frac{\sigma^2}{2}\,\frac{\partial^2}{\partial X^2}(X\,P)\,. \tag{8.15}$$

Stationary solutions of this equation must satisfy

$$-\frac{d}{dX}\Big(f(X)\,P\Big) + \frac{\sigma^2}{2}\,\frac{d^2}{dX^2}(X\,P) = 0\,,$$

and substituting $-aX^2 + bX + c$ for $f(X)$, we have the Gaussian solution

$$P(X) = \frac{1}{\sqrt{2\pi}\,\Sigma}\,\exp\left(-\frac{(X-\mu)^2}{2\,\Sigma^2}\right),$$

with

$$\mu = \frac{b}{a} \;\text{ and }\; \Sigma = \frac{1}{\sqrt{2}}\,\frac{\sigma}{a}\,.$$

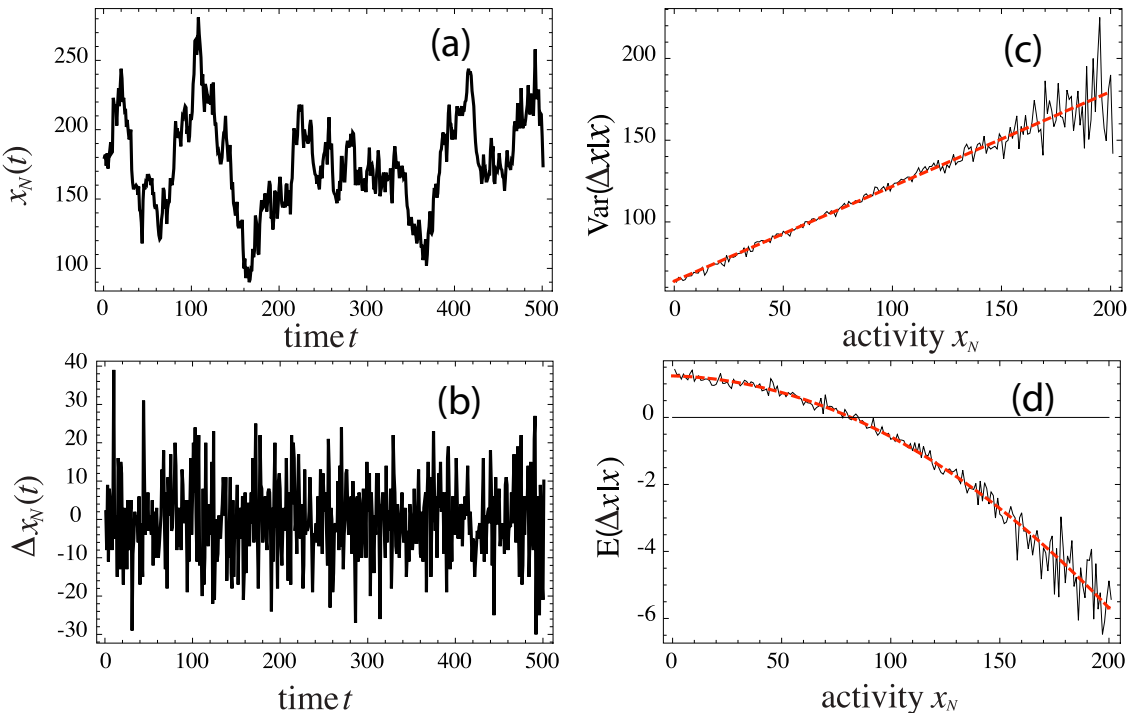


Figure 8.9: a) Part of a time series for the toppling activity in a strongly driven Zhang sandpile. b) The increments $\Delta x(t) = x(t + \Delta t) - x(t)$ of the signal in (a). As for the slowly driven case. c) The conditional variance of the increments Δx given the value of x. d) The conditional mean of increments. This curve is fitted to a parabola.

To further substantiate that the toppling dynamics of sandpiles is fundamentally different from the fluctuating quantities that give rise to BHP and FTG distributions we apply the above analysis to the fluctuations in total kinetic energy in a simulation of two-dimensional Navier-Stokes turbulence. The two-dimensional geometry is chosen because of the inverse energy cascade in k-space caused by merging of small vortices and formation of large structures, reminiscent of formation of large avalanches from emerging from localized random perturbation in sandpile models. In the simulation energy is injected via a source term on a characteristic, small spatial scale throughout the simulation area, and is dissipated as loss through the open boundary.

We compute the conditional mean and conditional variance of the increment process. These results are presented in figure 8.10. We observe that contrary to the sandpile models, the diffusion coefficient for this process grows exponentially with X. Moreover, the drift term can be approximated by a small positive constant except when the kinetic energy is very large. This leads us to the stochastic

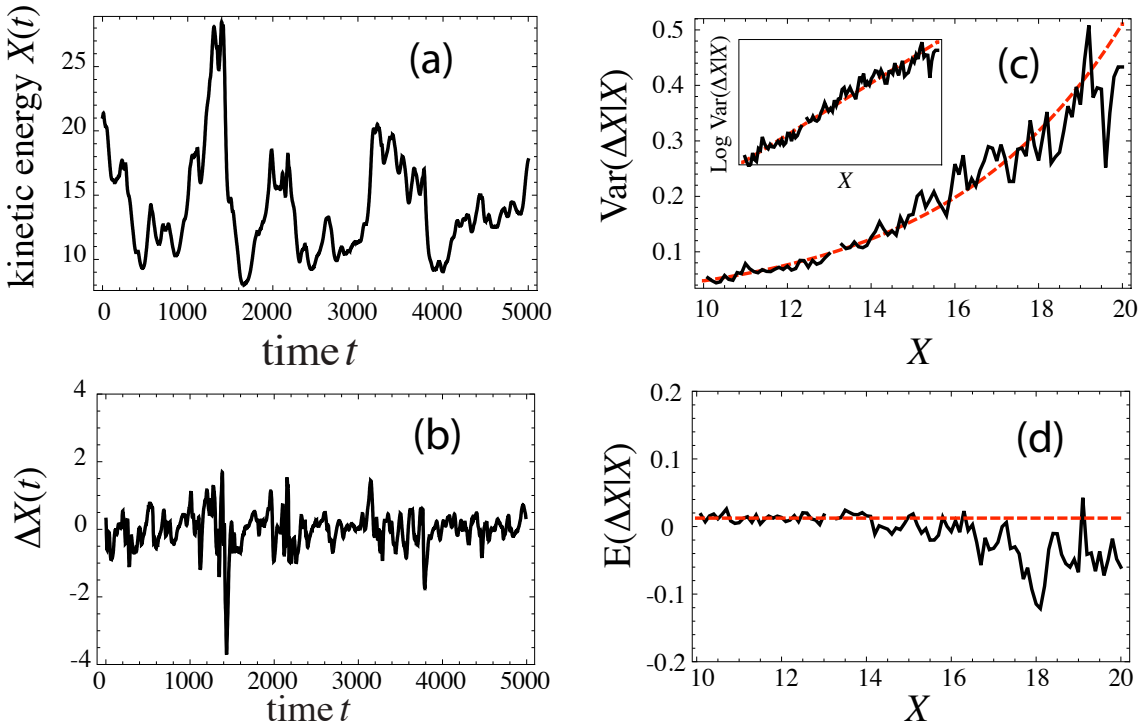


Figure 8.10: a) Part of a time series for the kinetic energy in the 2D turbulence simulation. b) The increments $\Delta X(t) = X(t + \Delta t) - X(t)$ of the signal in (a). As for the toppling activity in sandpiles we see that the increments are large when the kinetic energy itself is large. c) The conditional variance of the increments ΔX given the value of X. The inset shows the logarithm of this variance versus X. The dashed curves corresponds to a fitted exponential function. d) The conditional mean of increments. We see that the mean is approximately constant for a large range of X.

differential equation Eq. (8.3), and the corresponding Fokker-Planck equation is

$$\frac{\partial P}{\partial t} = -b\,\frac{\partial P}{\partial X} + \frac{c^2}{2}\,\frac{\partial^2}{\partial X^2}\left(e^{2aX}P\right). \tag{8.16}$$

Stationary solutions of this equation must satisfy

$$-b\,\frac{dP}{dX} + \frac{c^2}{2}\,\frac{d^2}{dX^2}\left(e^{2aX}P\right) = 0\,.$$

We can put $c = 1$ without loss of generality, and obtain the solution

$$P(X) = \frac{1}{\beta}e^{-(X-\mu)/\beta}e^{-e^{-(X-\mu)/\beta}}\,,$$

where $\beta = 1/2a$ and $\mu = (1/2a)\log(1/2a)$. This is the standard FTG distribution. figure 8.11 shows the normalized PDF for the fluctuations in total kinetic energy

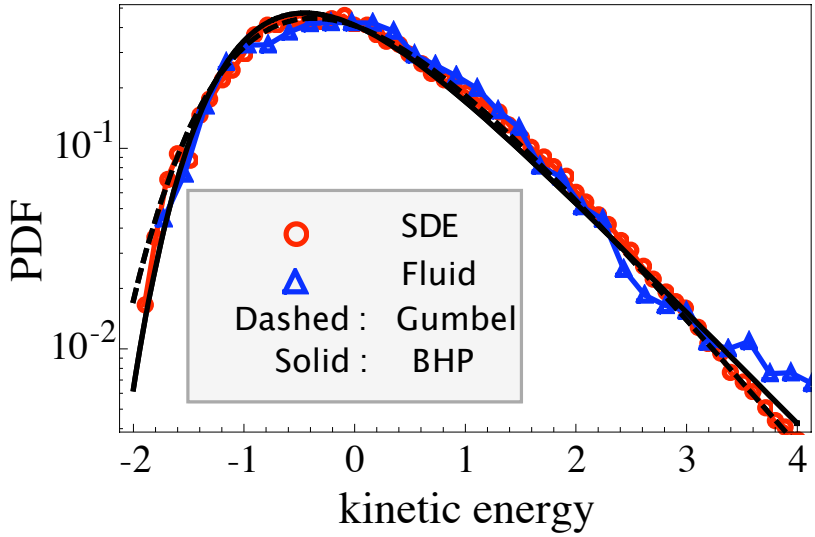


Figure 8.11: The normalized PDF for the fluctuation of kinetic energy in the 2D turbulence simulation together with a normalized FTG PDF (smooth solid curve), the normalized BHP PDF (smooth dashed curve) and a PDF obtained from the simulation of Eq. 8.3 (dotted curve).

obtained from the fluid simulation along with a normalized FTG distribution and the PDF obtained from the simulation of Eq. (8.3).

6 Conclusions

When output from simple models for avalanching systems are compared to observational data of real-world systems one has to deal with the problem of establishing a correspondence between the model variables and the observables of the natural system. In general, this is usually not an obvious task, since the model is usually not derived from first physical principles. The observation may be spatiotemporal, or just temporal. Likewise we may choose to analyze the spatiotemporal output from an avalanche model, or just some spatially integrated quantity like the total activity variable in a sandpile, presented as a time series. In this paper we have focused on the latter, and in particular on reproducing the statistical properties of such time-series by modeling the stochastic process by means of a stochastic differential equation.

Since our interest is avalanching dynamics we focus on avalanche statistics, and we have to face the problem of how to define what an avalanche is when our observational data is in the form of a time series. In general we cannot expect that such a definition is necessarily equivalent with a definition based on spatiotemporal

data, at least not in those cases where feeding of new "sand grains" occurs while avalanches are running. In those cases several spatially separated avalanches may run simultaneously, but these cannot be separated in a pure temporal analysis. For such a continuously driven model system the temporal signal may never be zero, and in an observational time series noise also makes it impossible to use a zero signal condition to separate an avalanching state from a quiet state. Thus, the natural way out is to define the avalanching state by means of a threshold level on the activity signal, giving rise to the concept of type-II avalanches.

In this paper we have shown that the toppling activity in sandpiles, and also global kinetic energy in a 2D fluid simulation, can be modeled by stochastic differential equations. The modeling clarifies that the main discrepancy between the mean-field approach and the actual BTW and Zhang models is related to the Hurst exponent of the activity process, which is different for the two sandpile models. It also clarifies the origin of the differences between the scaling exponents for type-I and type-II avalanches, and why type-II avalanches exhibit clearer scaling characteristics than type-I avalanches.

It follows from the theory how to rescale coordinates to approach the thermodynamic limit, and the results obtained for finite-size scaling in the BTW model in Sec. 4 gives a strong indication that this limit actually exists.

For continuously driven sandpiles the stochastic equation can be cast into a Fokker-Planck equation due to the loss of memory caused by the random feeding. This allows an analytic solution for the activity PDF, which is a Gaussian for the sandpile models, but gives rise to the FTG distribution for the 2D fluid simulation. These results show that the non-Gaussian universal PDF described in [13, 14] is not relevant for strongly driven sandpiles, but may be so for certain turbulence models. The stochastic theory relates the difference between Gaussian and FTG distributed activity signal to a the difference between a linear and exponential X dependence of the diffusion coefficient in the stochastic equation.

References

[1] L. Abramov, V. Rokhlin, *The entropy of a skew product of measure-preserving transformations*, Amer. Math. Soc. Transl. Ser. 2 **48**, 255 (1966).

[2] P. Arnoux, D. Ornstein, B. Weiss, *Cutting and stacking, interval exchanges and geometric models*, Israel J. Math. **50**, 160 (1985).

[3] P. Bak, C. Tang, K. Wiesenfeld, *Self-organized criticality: An explanation of the 1/f noise*, Phys. Rev. Lett. **59**, 381 (1987).

[4] P. Bak, C. Tang, K. Wiesenfeld, *Self-organized criticality*, Phys. Rev. A **38**, 364 (1988).

[5] V. Baladi, *Spectrum and Statistical Properties of Chaotic Dynamics*, Proc. 3rd European Congress of Mathematics, Birkhauser, 203 (2001).

[6] Ph. Blanchard, B. Cessac, T. Krüger, *What can we learn about SOC from Dynamical System Theory*, J. Stat. Phys. **98**, 375 (2000).

[7] B. Cessac, P. Blanchard, T. Krüger, *Lyapunov exponents and transport in the Zhang model of self-organized criticality*, Phys. Rev. E **64**, 016133 (2001).

[8] B. Cessac, P. Blanchard, T. Krüger J. Meunier, *Self-Organized Criticality and Thermodynamic Formalism*, J. Stat. Phys. **115**, 1283 (2004).

[9] G. Boffetta *et al.*, *Power Laws in Solar Flares: Self-Organized Criticality or Turbulence?*, Phys. Rev. Lett. **83**, 4662 (1999).

[10] G. Boffetta, A. Mazzino, A. Vulpiani, *Twenty-five years of multifractals in fully developed turbulence: a tribute to Giovanni Paladin*, J. Phys. A: Math. Theor. **41**, 363001 (2008).

[11] T. Bogenschütz, H. Crauel, *The Abramov-Rokhlin Formula*, Ergodic theory and related topics, III (Gustrow, 1990), 32–35, Lecture Notes in Math. **1514**, Springer, Berlin, 1992.

[12] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces*, Trans. A.M.S. **153**, 401 (1971).

[13] S. Bramwell, P. Holdsworth, J. Pinton, *Universality of rare fluctuations in turbulence and critical phenomena*, Nature **396**, 552 (1998).

[14] S. Bramwell *et al.*, *Universal Fluctuations in Correlated Systems*, Phys. Rev. Lett. **84**, 3744 (2000).

[15] J. Brown, *Ergodic theory and topological dynamics*, Pure and Applied Mathematics **70**, Academic Press [Harcourt Brace Jovanovich Publ.], New York-London (1976).

[16] J. Buzzi, *Intrinsic ergodicity of affine maps in* $[0,1]^d$, Mh. Math. **124**, 97 (1997).

[17] J. Buzzi, *Piecewise isometries have zero topological entropy*, Ergod. Th. & Dynam. Sys. **21**, 1371 (2001).

[18] A. Carbone *et al.*, *Analysis of clusters formed by the moving average of a long-range correlated time series*, Phys. Rev. E **69**, 026105 (2004)

[19] K. Christensen, N. Moloney, *Complexity and Criticality*, Imperial College Press (2005).

[20] K. Christensen, Z. Olami, *Sandpile models with and without an underlying spatial structure*, Phys. Rev. E **48**, 3361 (1993)

[21] S. Coles, *An Introduction to Statistical modeling of Extreme Values*, Springer, Bristol (2001).

[22] M. Denker, C. Grillenberger, K. Sigmund, *Ergodic theory on compact spaces*, Lecture Notes in Math. **527.** Springer-Verlag, Berlin-New York (1976).

[23] D. Dhar, R. Ramasway, *Exactly solved Model of Self-Organized Critical Phenomena*, Phys. Rew. Let. **63**, 1659 (1089).

[24] M. Ding, W. Yang, *Transition through onoff intermittency in Nd: YAG laser systems pumped by laser diodes*, Phys. Rev. E **52**, 207 (1995).

[25] A. Giacommetti, A. Dias-Guilera, *Dynamical properties of the Zhang model of self-organized criticality*, Phys. Rev. E **58**, 247 (1998).

[26] E. Gumbel, *Statistics of Extremes*, Diver Publication, Mineola, New York (1958).

[27] E. Gutkin, N. Haydn, *Topological entropy of polygon exchange transformations and polygonal billiards*, Ergod. Th. & Dynam. Sys. **17**, 849 (1997).

[28] J. Hutchinson, *Fractals and self-similarity*, Indiana University Math. Journal **30**, 713 (1981).

[29] E. Ivashkevich, V. Priezzhev, *Introduction to the sandpile model*, Physica A **97**, 254 (1998).

[30] H. Jensen, *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological systems*, Cambridge Lecture Notes in Physics 10, Cambridge University Press (1998).

[31] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Etudes Sci. Publ. Math. **51**, 137 (1980).

[32] A. Katok, B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press (1995).

[33] A. Katok, J. Strelcyn, *Invariant manifolds, Entropy and Billiards; Smooth Maps with Singularities*, Lecture Notes in Math. **1222**, Springer, Berlin, 1986.

[34] B. Kozelov, T. Kozelova, *Cellular automata model of magnetospheric-ionospheric coupling*, Annales Geophysicae **21**, 1931 (2003).

[35] B. Kruglikov, M. Rypdal, *Dynamics and entropy in the Zhang model of SOC*, J. Stat. Phys. **122**, 975 (2006). This paper is included as chapter 3 in this thesis.

[36] B. Kruglikov, M. Rypdal, *Entropy via multiplicity*, Discrete and Continuous Dynamical Systems **16**, 395 (2006). This paper is included as chapter 4 in this thesis.

[37] B. Kruglikov, M. Rypdal, *A piece-wise affine contracting map with positive entropy*, Discrete and Continuous Dynamical Systems **16**, 393 (2006). This paper is included as chapter 5 in this thesis.

[38] F. Ledrappier, L. Young, *The metric entropy of diffeomorphisms*, I-II, Ann. of Math. (2) **122**, 509 (1985).

[39] F. Ledrappier, *Quelques proprietes des exposants caracteristiques*, Lecture Notes in Math. **1097**, 305–396, Springer, Berlin, 1984.

[40] S. Lübeck, K. Usadel, *Numerical determination of the avalanche exponents of the Bak-Tang-Wiesenfeld model*, Phys. Rev. E **55**, 4095 (1997).

[41] M. Maejima, *Selfsimilar processes*, Princeton University Press (2002).

[42] P. Moran, *Additive functions of intervals and Hausdorff measure*, Proc. Camb. Phil. Soc. **42**, 15 (1946).

[43] Y. Pesin, *Dynamical systems with generalized hyperbolic attractors: hyperbolic, ergodic and topological properties*, Ergod. Th. & Dynam. Sys. **12**, 123 (1992).

[44] Y. Pesin, *Dimension theory in dynamical systems*, Chicago Lect. in Math. Ser., The University of Chicago Press (1997).

[45] Y. Pesin, B. Pitskel, *Topological pressure and the variational principle for non-compact sets*, Funct. Anal. Appl. **18**, 50 (1984).

[46] K. Petersen, *Chains, entropy, coding*, Ergod. Th. & Dynam. Sys. **6**, 415 (1986).

[47] M. Pollicott, M. Yuri, *Dynamical systems and ergodic theory*, London Mathematical Society Student Texts **40**, Cambridge University Press (1998).

[48] M. Paczuski, S. Boettcher, M. Baiesi, *Interoccurrence Times in the Bak-Tang-Wiesenfeld Sandpile Model: A Comparison with the Observed Statistics of Solar Flares*, Phys. Rev. Lett. **95**, 181102 (2005).

[49] M. Rypdal, *Dynamics of the Zhang model of Self-Organized Criticality*, Master Thesis in mathematics, University of Tromsø (2004), e-printed: http://www.math.uit.no/seminar/preprints.html.

[50] M. Rypdal, K. Rypdal, *A stochastic model of the toppling activity in self-organized critical systems*, Preprint: arXiv:0710.4010v2 (2008), (Accepted by the New Journal of Physics). This paper is included as chapter 7 in this thesis.

[51] M. Rypdal, K. Rypdal, *Modeling of temporal uctuations in sandpiles*, Preprint: arXiv:08073416v1 (2008), (Accepted by Physical Review E). This paper is included as chapter 8 in this thesis.

[52] D. Ruelle, *Chaotic evolution and strange attractors*, Lezioni Lincee, Cambridge University Press (1989).

[53] J. Schmeling, S. Troubetzkoy, *Dimension and invertibility of hyperbolic endomorphisms with singularities*, Ergod. Th. & Dynam. Sys. **18**, 1257 (1998).

[54] C. Tang, P. Bak, *Mean Field Theory of Self-Organized Critical Phenomena*, J. Stat. Phys. **51**, 797 (1988).

[55] M. Tsujii, *Absolutely continuous invariant measures for expanding piecewise linear maps*, Invent. Math. **143**, 349 (2001).

[56] N. Watkins *et al.*, *A kinetic equation for linear fractional stable motion with applications to space plasma physics*, Preprint: arXiv:0803.2833v1 (2008)

[57] H. Zhang, *Scaling theory of Self-Organized Criticality*, Phys. Rev. Lett. **63**, 470 (1988).

[58] B. Øksendal, *Stochastic Differential Equations. An Introduction with Applications*, Springer (2000).