

STUDY OF THE SCALING AND TEMPORAL PROPERTIES OF A
SIMPLIFIED EARTHQUAKE MODEL

By

Daniel Groleau

B. Sc. (Physics) Université du Québec à Rimouski

M. Sc. (Physics) Université de Sherbrooke

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
PHYSICS AND ASTRONOMY

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 1997

© Daniel Groleau, 1997

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Physics and Astronomy

The University of British Columbia

6224 Agricultural Road

Vancouver, B.C., Canada

V6T 1Z1

Date:

AUGUST 25, 1997

Abstract

Certain driven systems consisting of a large number of elements evolve towards a critical state with no characteristic time or length scales. This class of phenomena is described as Self-Organized Criticality (SOC). SOC relies on the condition of a slow driving of the systems and the existence of fast burst-like responses of them and includes earthquakes.

We employ a model proposed by Xu et al. for ruptures in an elastic medium, subject to shear stress, and apply it to the study of earthquakes. In the model, the size of an earthquake is defined as the number of ruptures occurring sequentially on the basic units of discretization (squares) of the medium. A histogram of the earthquake sizes shows that the model is not completely scale invariant due to finite-size effects. To take them into account, we implement a finite-size-scaling analysis. The results of this analysis show that the model is scale invariant only when there is stress conservation. So, the model displays SOC in the conservative case only.

We also study the dynamic quantities of the model, in particular the average stress in the system. The sets of average stress values are analyzed using two types of time series analysis. The nonlinear forecasting analysis investigates whether time series exhibit low-dimensional chaotic behavior as opposed to high-dimensional (or stochastic) behavior. We find that the above time series have a nonlinear structure, but with a substantial stochastic component, so SOC is inherently high-dimensional. The appearance of nonlinear structure is due to the fact that the system stops following linearly the external drive when it releases stress through earthquakes. The rescaled range analysis characterizes the time correlations (or memory effects) in the time series. We find strong positive time correlations in the above time series. Their presence is due to the nature of the driving

in the model. These memory effects are destroyed as soon as a large earthquake resets the system.

Table of Contents

Abstract	ii
List of Tables	vii
List of Figures	viii
Acknowledgements	xvii
1 Introduction	1
1.1 Self-Organized Criticality	1
1.2 Identification of Self-Organized Criticality	3
1.3 Geophysics as a Field of Application for Self-Organized Criticality	4
1.4 Interests of the Concept of Self-Organized Criticality for Physicists	5
1.5 Organization of the Thesis	6
2 Lattice Model for the Study of Earthquakes	8
2.1 Approaches to Fracture Modeling	8
2.2 Lattice Models	10
2.2.1 Lattice models considered by engineers	11
2.2.2 Lattice models considered by physicists	12
2.3 A New Lattice Model	15
2.3.1 A new discretization scheme	15
2.3.2 Advantages of that discretization scheme over others	20
2.4 Application to the Study of Earthquakes	21

2.5	Comparison of the Model of Xu et al. to Other Models	27
2.5.1	Model of Xu et al. versus other earthquake models	27
2.5.2	Model of Xu et al. versus models of integrate-and-fire oscillators .	30
3	Study of the Scaling Properties of the Model	32
3.1	Motivation for this Study	32
3.2	Finite-Size Scaling	33
3.3	Results of the Finite-Size-Scaling Analysis	35
3.3.1	Results in the conservative case ($\alpha = 1$)	36
3.3.2	Results in the non-conservative case ($\alpha < 1$)	46
3.4	Stress Ordering in One-Dimensional Lattices	50
3.5	Mean-Field Version of the Model of Xu et al.	54
3.6	Summary of the Important Results and Conclusions	57
4	Time Series Analyses	62
4.1	Time Series of Interest	62
4.2	Nonlinear Forecasting Analysis	69
4.2.1	Introductory discussion	69
4.2.2	State space reconstruction	76
4.2.3	Nonlinear function approximation	78
4.2.4	Nonlinear forecasting algorithm considered	80
4.2.5	Tests of the algorithm with time series having a known structure .	83
4.2.6	Results obtained with the time series of interest	89
4.3	Rescaled Range Analysis	103
4.3.1	Introductory discussion	103
4.3.2	Fractal time series analyses	112
4.3.3	Rescaled range analysis algorithm considered	113

4.3.4	Tests of the algorithm with time series having a known value of H	116
4.3.5	Results obtained with the time series of interest	117
4.4	Summary of the Important Results and Conclusions	126
5	Conclusion	129
	Bibliography	137
	Appendices	142
A	Lattice Green Function Simplification and Properties	143
A.1	Simplification of the Lattice Green Function	143
A.2	Properties of the Lattice Green Function	147
B	Coarse-Graining of the Lattice Green Function	149

List of Tables

3.1	Exponent τ and characteristic earthquake size S_0 for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and three values of α	47
-----	---	----

List of Figures

2.1	One-layer <i>planar</i> region considered in the model of Xu et al. The region is discretized into squares. It is embedded in an infinite medium (providing the necessary external drive) and the effect of any activity originating outside the region is neglected (open boundary conditions). In the model, the interaction between the external drive (represented by the arrows outside the region) and a square of the region is treated in a mean-field way, i.e. it is the same whatever the square location. The arrows at the corners of one of the square inside the region are a schematic representation of a double couple force redistribution following the shear rupture of that square. This figure was modified from [114].	18
3.1	Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. For comparison, we show the curve $C(S) = S^{-0.30}$	37
3.2	Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ and $L \times 2$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. All curves use the fitting parameter $\tau = 0.30$. For the left and middle curves, all ruptures are counted while for the right curves, only ruptures on distinct squares are counted. The left curves were obtained for $L \times 2$ lattices using $\nu = 1.38$, the middle curves for $L \times 1$ lattices using $\nu = 1.38$, and the right curves for $L \times 1$ lattices using $\nu = 1.01$	38

3.3	Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. The segment $S < 50$ is a crossover region and its origin is explained in section 3.4. Beyond $S > 50$, $C(S, L)$ has a linear segment with a slope close to 0.30. This linear segment gets longer as L increases but cuts off at large S . For comparison, we show a curve proportional to $S^{-0.30}$	40
3.4	Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $\theta_l = 0.25$, $\alpha = 1$ and two values of X . Exponents used are $\tau = 0.30$ and $\nu = 1.38$	41
3.5	Log-linear scaling plot of the cumulative size-frequency distributions for a 1000×1 lattice, random σ_{th} , $\theta_l = 0.25$, $\alpha = 1$ and three values of X . Exponents used are $\tau = 0.30$ and $\nu = 1.38$	42
3.6	Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\alpha = 1$ and two values of θ_l . Exponents used are $\tau = 0.30$ and $\nu = 1.38$	44
3.7	Log-linear scaling plot of the cumulative size-frequency distributions for a 1000×1 lattice, random σ_{th} , $X = 0$, $\alpha = 1$ and three values of θ_l . Exponents used are $\tau = 0.30$ and $\nu = 1.38$	45
3.8	Log-log plot of the cumulative size-frequency distributions for a 2000×1 lattice, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and three values of α . The dashed lines are the fits to relation (3.3).	48
3.9	Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 0.94$	49
3.10	Distribution of the stress differences δ_i ($i = 1, 2, \dots, 999$) between the nearest-neighbor elements of a 1000×1 system with constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$	52

- 3.11 Histogram plots of the stress differences δ_i between the nearest-neighbor elements of a 1000×1 system with constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$ at two distinct times. The solid line curve was obtained with data taken at the arbitrary time t_0 (data of figure 3.10) and the curve plotted using a dashed line and lozenges was obtained with data taken at time $t_1 > t_0$. Fifty bins were used to get the histograms. 53
- 3.12 Log-log plot of $1 - \alpha_{eff}$ values versus L for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$ obtained with the original model of Xu et al. The slope of the straight segment was found to be -1.02 58
- 3.13 Log-log plot of the cumulative size-frequency distributions for systems of L elements equally coupled with each other. The values of the parameters are $X = 0$ and the stress thresholds (i.e. the σ_{th}) are taken from a uniform distribution with $\theta_l = 0.25$. From left to right, the distributions refer to systems having $L = 200$, $L = 400$ and $L = 1000$. The values of α differ for all three distributions (see how they were adjusted in the text). 59
- 4.1 Typical time series of type 1: Size of an earthquake sequences (S_i) as a function of the occurrence time (t_i). The whole time series has 5×10^4 data but for clarity, only the first 2000 are plotted. The sequences were generated from a 100×100 system with random σ_{th} . The time at which the first sequence happens is arbitrarily set to $t_1 = 0$ 65
- 4.2 Typical time series of type 1 [variety (b)]: Time intervals $t_i - t_{i-1}$ ($i = 2, 3, \dots, 2000$) between consecutive earthquake sequences. The sequences were generated from a 100×100 system with random σ_{th} (see figure 4.1). 66

4.3	Typical time series of type 2: Average stress ($\bar{\sigma}$) as a function of the long time (t). The data were generated from a 100×100 system with random σ_{th} . The whole time series has 5×10^4 data but for clarity, only the first 500 data (i.e. the end points of the vertical lines) are plotted. The time of the first datum is arbitrarily set to zero.	68
4.4	Autocorrelation function for the time series of type 1 [variety (b)] plotted on figure 4.2. For convenience, the value $r_0 = 1$ is not shown on the graph.	70
4.5	Autocorrelation function for $N = 2000$ data of the time series of type 2 plotted partially on figure 4.3.	71
4.6	Normalized RMS forecasting error $E_m(k)$ as a function of k for time series having $N = 1000$ data generated by the logistic map (with $\mu = 4.0$) and the Hénon map (with $A = 1.4$ and $B = 0.3$). The solid line is for the logistic map time series and the dashed line for the Hénon map time series. The parameters of the nonlinear forecasting algorithm are $N_f = 800$ and $m = 2$	86
4.7	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series having $N = 1000$ data generated by the TAR model (with $\alpha = -0.4$ and $\beta = 0.5$). The parameters of the nonlinear forecasting algorithm are $N_f = 800$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	87
4.8	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series having $N = 1000$ data generated by the AR(1) model (with $\phi = 0.5$). The parameters of the algorithm are $N_f = 800$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	88

4.9	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 1 [variety (c)] for a 100×100 system with random σ_{th} . The sampling time interval Δt is equal to 10 times the average time interval between the earthquake sequences. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	91
4.10	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 1 [variety (a)] for a 1000×1 system with constant σ_{th} . The data in the time series ($N = 331$) are the time intervals between large earthquakes ($S \geq 1000$). The parameters of the nonlinear forecasting algorithm are $N_f = 264$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	92
4.11	Histogram plot of the time intervals between large earthquake sequences ($S \geq 1000$) generated in a 1000×1 system with constant σ_{th} . The data set has 331 data and 20 bins were used to get the histogram. The vertical axis is logarithmic in order to better illustrate that the distribution is exponential.	93
4.12	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with random σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	95
4.13	Normalized RMS forecasting error $E_m(k)$ as a function of k for a first-differenced time series of type 2 for a 100×100 system with random σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	96

4.14	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with constant σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	97
4.15	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with random σ_{th} when the parameter α controlling the level of stress conservation in the system (see chapter 3) is set to 0.9. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	98
4.16	Normalized RMS forecasting error $E_m(k)$ as a function of k for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval Δt is about equal to 1.5 times the average time interval between the earthquake sequences. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	101
4.17	Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series which has the same power spectrum as the time series of figure 4.16, but randomized phases. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).	102
4.18	Fractional Brownian motion (fBm) signals having 8193 data generated by the successive random addition method. From top to bottom, the curves have $H = 0.1$, $H = 0.5$ and $H = 0.9$ and have been shifted for clarity.	107

- 4.19 Running sum of an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} having $N = 16,384$ data. The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences. The trend and the mean have been removed from the time series before computing the running sum. 110
- 4.20 Periodogram estimate of the power spectrum of the running sum of an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval of the time series Δt is equal to 1.5 times the average time interval between the earthquake sequences, where the time series has 16,384 data. We have removed the trend and the mean before calculating the power spectrum. To reduce the fluctuations, we have computed an average power spectrum over 8 non-overlapping subsets (with 2048 data) of the time series resulting from the running sum operation. The dashed lines were obtained by doing linear regressions and have slope $\beta = 2.94 \pm 0.05$ (left segment) and $\beta = 3.46 \pm 0.03$ (right segment). . . . 111
- 4.21 Log-log plot of $(R/S)_M$ versus M for three synthetic time series having $N = 2^{13}$ data generated by the successive random addition method. From top to bottom, the curves correspond to the cases $H = 0.9$, $H = 0.5$ and $H = 0.1$. The dashed line at the right of the figure has a slope of 0.1 and is there to illustrate that the curve for $H = 0.1$ will only slowly reach the slope value $H_{RR} = 0.1$ as M increases. Note that each curve was obtained by averaging $(R/S)_M$ (for a given M) over 5 realizations corresponding to the same value of H . For each curve, we have excluded the point corresponding to $M = 2$ since it is trivial and always has $(R/S)_{M=2} = 1$. 118

- 4.22 Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} having Δt equal to 1.5 times the average time interval between the earthquake sequences ($N = 2^{14}$). The curve corresponding to this time series is plotted using a solid line and lozenges. The dotted line ($H = 0.9$) is a fit of the linear segment $M \in [2^2, 2^7]$ of the latter curve. The curve at the bottom was obtained with 5 random shuffles of the time series: the curve of $(R/S)_M$ versus M was calculated for each realization and an average value of $(R/S)_M$ over the 5 realizations was computed for each M 121
- 4.23 Log-log plot of $(R/S)_M$ versus M for evenly sampled time series of type 2 for $L \times L$ ($L = 25, 50, 100$) systems with random σ_{th} . The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences and the number of data is $N = 2^{14}$ for all three time series. From top to bottom, the curves refer to the 100×100 , 50×50 and 25×25 systems respectively. 122
- 4.24 Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with constant σ_{th} having Δt equal to 1.5 times the average time interval between the earthquake sequences ($N = 2^{14}$). The curve corresponding to this time series is plotted using a solid line and lozenges. The dotted line is a fit ($H = 0.9$) of the linear segment $M \in [2^2, 2^7]$ of the latter curve. The curve at the bottom was obtained with 5 random shuffles of the time series: the curve of $(R/S)_M$ versus M was calculated for each realization and an average value of $(R/S)_M$ over the 5 realizations was computed for each M 123

4.25	Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences and the number of data is $N = 2^{14}$. The curve plotted using a solid line and lozenges was obtained with the original time series whereas the one plotted using a dashed line and pluses was obtained with 5 surrogate time series having the same power spectrum as the original time series but randomized phases. For the latter curve, $(R/S)_M$ was calculated for each realization, for a given M , and an average value of $(R/S)_M$ over the 5 realizations was computed for this value of M	124
------	--	-----

Acknowledgements

I would first like to thank my supervisor, Dr. Birger Bergersen, for his help and financial support over the past five years. His enthusiasm has helped me finding the energy to finish my Ph.D. project, especially in the difficult times. I would also like to thank Dr. Huang-Jian Xu who initiated this work by introducing the earthquake model. I am particularly grateful to him for having helped me in the early stages.

A special thank to Dr. Lawrence Ward of the Department of Psychology at UBC who allowed me to use the code of his nonlinear forecasting algorithm. In addition, I benefited a lot from my discussions with Dr. Ward. I also thank him for having accepted to sit as a member of my Ph.D. committee in the last year.

Thanks also to my other Ph.D. committee members, Dr. Rosemary Knight of the Department of Earth and Ocean Sciences and Drs. Robert Parsons and Nathan Weiss of the Department of Physics and Astronomy, who helped me in different ways.

I gratefully acknowledge the financial support, through graduate fellowships, of *Le Fonds FCAR* of the government of Québec as well as of the National Science and Engineering Research Council of Canada (NSERC).

Merci à mes parents, Louisette et Georges, pour m'avoir toujours supporté moralement durant mes années d'études.

Finalement, sans le support de ma compagne de vie, Renée, je n'aurais probablement jamais terminé ce doctorat. Je la remercie d'avoir toujours été là lorsque j'avais besoin de réconfort et d'encouragement. Sa grande compréhension m'a été indispensable. Je ne pourrais terminer ces remerciements sans mentionner ma fille, Carmen, qui est née durant la dernière année de mon doctorat, plus précisément en pleine rédaction de cette thèse. Son arrivée dans la vie m'a donné l'énergie nécessaire pour terminer mon doctorat.

Chapter 1

Introduction

Once upon a time, science was not divided into disciplines as it is today. There was no distinction between social and natural sciences, and certainly not the fragmentation of the latter into physics, chemistry, biology, geology, etc. However, when one is confronted with a new problem or phenomenon, it is usually impossible to neatly assign the problem to a particular discipline. One needs to recall and fuse together the knowledge one learned before and, if that is insufficient, to consult scientists in other areas [56].

In recent years, under the banner of *complex systems* (or *complex phenomena* to emphasize the fact that the interest is in problems from the real world), some scientists have ventured out of their trained disciplines. Complex systems research is a synthetic approach in the investigation of systems consisting of a large number of interacting, simple elements, irrespective of their origins (see [56] for a general outline of the complex systems research). It has led recently to the introduction of the concept of Self-Organized Criticality (SOC) by Bak, Tang and Wiesenfeld [3]. This concept has been introduced in order to understand the origin of the power laws observed in the study of various phenomena [81, 37, 57, 70, 61, 28].

1.1 Self-Organized Criticality

According to SOC, certain dynamically driven systems consisting of a large number of elements evolve spontaneously towards a critical globally stationary dynamical state with no characteristic time or length scales. This critical state is reached without fine tuning

of an external control parameter, unlike for phase transitions in equilibrium statistical physics. Those systems are made critical by the choice of a threshold (nonlinear) dynamics that forbids them to follow adiabatically the external drive.

To illustrate the basic ideas of SOC, Bak, Tang and Wiesenfeld [3] used a cellular automaton inspired by the creation of avalanches in a pile of sand. In this model, *sand* is added grain by grain at a random location in the lattice (random driving) until somewhere a local slope becomes unstable (i.e. the local slope is larger than a threshold value), and an avalanche is initiated. The pile eventually reaches a stationary critical state, characterized by a critical slope, in which additional grains will fall off the pile via avalanches of all sizes, distributed in size and lifetime according to a power law. It is of interest to note that the relative number of large and small avalanches in the model does not depend on the microscopic mechanisms, i.e. the global behavior of the sandpile can be synthesized from a simple model such as the one above.

Over the years, three facts have been established about SOC [94]. First, some systems qualify as self-organized critical if their large scale evolution obeys a diffusion equation which satisfies a global conservation law [49, 41]. Second, some systems, which exhibit diffusion-like response but do not obey a global conservation law, nevertheless exhibit SOC [77, 26, 1]. In these systems, SOC is believed to be related to the problem of synchronization of coupled threshold oscillators of relaxation [67, 7, 19]. Third, a feedback mechanism must operate which describes the action of the order parameter onto the control parameter [92] and attracts the dynamics to a critical state.

Recently, a unifying formalism has been presented by Gil and Sornette [36] to describe the general conditions under which a system exhibits SOC. They have introduced a continuous formulation in terms of the Landau-Ginzburg theory of phase transitions which takes into account the nonperturbative nature of the threshold dynamics. SOC was identified as the regime where diffusive relaxation is faster than the instability growth

rate whereas synchronization of threshold oscillators was observed in the limit of slow diffusion.

Now that the SOC concept has been described, we can talk about the properties that are necessary for a model to display SOC.

1.2 Identification of Self-Organized Criticality

It is often asserted that the hallmark of SOC is the simultaneous existence of two properties [93]: (1) power law distribution of events, and (2) spatial and temporal correlation functions decaying algebraically. According to [93], property (1) is more characteristic of the class of phenomena described by SOC, since it relies on the condition of a very slow driving of the system and the existence of fast burst-like responses. A number of self-organized critical phenomena can be described as ordinary phase transitions made self-organized [94]. Apart from the absence of tuning and the property of self-organization, the latter are similar to ordinary phase transitions. We will not consider this class of phenomena in this thesis.

For a system to be self-organized critical, it must exhibit a power-law distribution of event (response) sizes. But the opposite is not true: the observation of a power-law distribution of event sizes does not necessary imply that the system is self-organized critical [93]. As an illustration of this, it is of interest to note that a number of numerical and experimental works claiming the observation of SOC in fact rely on a different physical mechanism, called *sweeping of an instability* [93]. This mechanism involves the slow sweeping of a control parameter on the approach to and possibly beyond an instability at the critical value of that control parameter. In other words, the power law arises if a control parameter crosses over from one side of the critical point to the other side. As a

result, the system can not operate persistently since the control parameter does not remain constant. Many examples of how that mechanism operates are provided in [93] and, among them, foreshocks (shocks preceeding a large earthquake) and acoustic emissions during the progressive deformation of materials submitted to increasing stresses.

The SOC concept applies to many different phenomena occurring in nature and in particular, to phenomena in the field of geophysics.

1.3 Geophysics as a Field of Application for Self-Organized Criticality

Geophysics is a field in which many phenomena are well suited to the concept of SOC. In particular, the earthquake phenomenology was argued by several authors [96, 2] to bear the signature of SOC. To illustrate this, we summarize a part of the discussion in the paper of Sornette et al. [94].

Let us consider a model elastic tectonic plate [20] scaled down in the laboratory to perform a mechanical experiment. A shear force (F) is, in the simplest case, imposed on two opposite borders of the plate, the other two borders being free. As F increases, the plate starts to deform. For low F , after some transient during which the plate deforms and adjusts itself to the applied force, the plate becomes static: the strain rate (r) is zero (if any ductile behavior is neglected). As F increases, the transient becomes longer and larger deformations develop within the plate. A plasticity threshold (F_c) will be eventually reached and then the plate will start to flow with a non-zero strain rate ($r > 0$) under fixed F . As F increases above F_c , r increases. In laboratory experiments, such a transition from a brittle to a ductile behavior may be critical. Parameters F and r qualify respectively as the control parameter and the order parameter (r is non-zero only when $F > F_c$).

Now, instead of controlling the shear force exerted on the plate, we imagine that

a constant (very small) shear rate of deformation ($r \rightarrow 0$) is applied at the borders of the plate. Then, the *natural* boundary condition for earthquakes and plate tectonic deformations is recovered. Such a situation has been studied in many works showing the existence of SOC, in particular in the shape of a power-law earthquake size distribution (the well-known Gutenberg-Richter power law [43]). The above boundary condition is nothing but driving the plate at the critical point ($F \rightarrow F_c$) by controlling the order parameter (r) to a very small value ($r \rightarrow 0$), thus ensuring the critical properties of the plate. The fact that r is non-zero even if it is very small proves that the plate is an out-of-equilibrium system.

The above example illustrates the relevance of the concept of SOC to the earthquake phenomenology. We will focus on the earthquake phenomenon throughout the present work.

The SOC concept is interesting not only from a practical point of view (e.g. understanding earthquakes), but also from a theoretical point of view.

1.4 Interests of the Concept of Self-Organized Criticality for Physicists

The interests of the physics community in the study of SOC is twofold. Firstly, SOC, as was shown by Sornette et al. [94], can be mapped onto ordinary criticality in many cases. Then, rather than searching for the underlying mechanism for SOC, physicists can study unstable critical points (i.e. critical points which are not self-organized). This task can be performed by using the toolbox developed in the last decades in this field. However, setting the order parameter of an unstable critical point to an infinitesimal number does not allow for a standard renormalization group procedure, because there is no control parameter and the exponents related to the distance from the critical point do not exist. There have been some attempts [76, 49, 83, 24] at finding a general renormalization group

theory for systems right at their critical point, but much work along this line remains to be done though. Such a theory could provide a new class of tools for critical phenomena as suggested in [94].

Secondly, SOC applies to phenomena actually occurring in nature and so may be useful to their understanding. Among these phenomena are, apart from earthquakes, formation of river networks [100], formation of clouds [72], evolution of populations [1], neuron dynamics [7, 19], etc. The models used to study such phenomena are often idealized in the sense that they capture only the essential features of real systems. If these simple models would turn out to be robust with respect to various modifications, then extrapolation to real situations could be possible. (It is of interest to note that such an approach has been successful in equilibrium statistical mechanics, where universal features in real systems can be understood from the study of simple models). Also, if models used to study a priori not closely related phenomena were to have similar features, then this could prove that a unique universal dynamics is behind those phenomena. Indeed, why could we not learn about the dynamics of neurons of the human body by studying earthquakes?

Our interest in the concept of SOC is oriented towards the understanding of the earthquake phenomenon in the current work.

1.5 Organization of the Thesis

The present work is devoted to the detailed study of a model introduced by Xu et al. [113, 114] which was intended to be applied to earthquakes. This model was shown to be self-organized critical [114]. However, we believe that a more rigorous proof is needed before to state whether or not the model displays SOC. In addition, the temporal aspects of the model have almost entirely been left out in [113, 114] and so, we want to study

them thoroughly in this work.

The organization of the present work goes like this: in chapter 2, we describe the approach used to model fractures in a planar region, the occurrence of earthquakes being closely related to the appearance of fractures in nature. We use a particular lattice model, which has been introduced recently by Xu et al., to do so. In chapter 3, we study the scaling properties of the model by means of a finite-size-scaling analysis, a technique which has been extensively used in the study of equilibrium phase transitions. This will enable us to clearly state whether the model displays SOC. In chapter 4, we implement two types of time series analysis, namely a nonlinear forecasting analysis and a rescaled range analysis. The time series are generated from the model. Finally, in chapter 5, we summarize our results, draw the relevant conclusions and make suggestions for future work.

Chapter 2

Lattice Model for the Study of Earthquakes

In nature, the occurrence of earthquakes is closely related to the appearance of fractures. The modeling of fractures can be done using different approaches (see chapter 5 of [45] for instance). These approaches are discussed in section 2.1. The approach presented in section 2.2 follows an intermediate route and uses lattice models as the tool to study fractures in the presence of disorder. The lattice model that we use in this work is presented in section 2.3 and applied to the study of earthquakes in section 2.4. Finally, it is compared to other earthquake models and to models of integrate-and-fire oscillators in section 2.5.

2.1 Approaches to Fracture Modeling

The description of fracture of samples of materials in the cases where disorder is important is the main goal of the material scientist interested in the fracture phenomenon. One possible approach he (or she) can use is to set up some microscopic rules which determine the mechanical behavior of the medium as well as how rupture will take place. If the medium is homogeneous and the rupture process deterministic, one can very simply know what the macroscopic behavior will be. However, in the presence of disorder, the solution to the problem is not that easy.

Among the various ways of handling disorder directly, three different approaches can be distinguished [45]: mean-field, local models and renormalization methods. Next, we discuss each of them.

The *mean-field* approach treats the surroundings of a volume element as a homogeneous embedding medium. In this way, each element interacts equally with all the others. However, because fracture is sensitive to disorder, this approach gives poor results.

The approach using *local models* tries to reproduce as closely as possible the real microscopic behavior starting from first principles. This is the case, for instance, for molecular dynamics that we now briefly describe. Suppose one chooses an interatomic potential, for example the Lennard-Jones (6-12) potential, to describe the interactions between the particles of the system under study. The particles are placed on a regular lattice. On the outer boundary of the system, an external load is applied by assigning a constant force vector to the particles lying on this boundary. To obtain the motion of the system, time is discretized into slices of duration Δt and one calculates the position that each particle has after the time Δt . In the simplest case, one calculates, using the equations of motion, the velocity of each particle. When all velocities \vec{v} are known, all particles are simultaneously displaced by $\vec{v}\Delta t$. Then, the step is repeated for the next time interval starting from the new positions, and so on. The molecular dynamic framework provides naturally the fracture properties without having to insert by hand any rheological behavior. However, the study of a system with several cracks in a disordered medium is, with the present means of computation, not possible. Much work remains to be done along this line though.

The approach using *renormalization methods* views the fracture process as a critical phenomenon. It expresses the invariance of some physical quantities under a change of length scale. However, the few situations where this technique has been used contain some severe approximations, or use unphysical and simplistic geometries [91, 74].

The above approaches are not adequate to study in an effective way the fracture phenomenon. This pessimistic picture seems to close all doors towards a solution. Nevertheless, the situation is not so hopeless. The approach used in [45] follows an intermediate

route: it describes very simply the mechanical behavior of mesoscopic (not microscopic) elements and considers the collective behavior of these elements in the presence of disorder. We now summarize this approach. The interested reader is referred to [45] for the details.

2.2 Lattice Models

A numerically tractable lattice formulation of fracture, alternative to molecular dynamics, is given by discretizations of continuum equations. In this case, the medium is reduced to a set of points embedded into a grid. Only local laws (e.g. the balance of force) are considered and their implementation involves for each point only a few neighbors. The calculation of collective properties like the equilibrium displacement field is reduced to solving a set of coupled linear equations.

Lattice models have their validity at much larger length scales (than the atomistic length scale) where the medium can be described by continuous fields. The breaking of the lattice is not a natural consequence of the simulation but has to be put into the model by hand as an additional rule of the model. This rule can lean on experimental data or on phenomenological laws, but the fact that it does not follow from first principles shows that lattice models are less fundamental than molecular dynamic models.

One advantage of lattice models is that they allow for the easy introduction of disorder. An even bigger advantage is that they are more *rigid* than a crystal in molecular dynamic models because the number of neighbors is fixed and the breaking of a bond is an irreversible yes-no decision. Lattice models are therefore good candidates to overcome the numerical difficulties encountered in molecular dynamic models.

Lattice models have been studied by engineers and physicists, but the methods used in the two science communities are different. Indeed, engineers use mostly Finite Element

Models (FEM) whereas physicists focus on less formal models. We now discuss in some details lattice models as considered by both engineers and physicists.

2.2.1 Lattice models considered by engineers

In engineering practice, the most commonly used lattice models are FEM. We want here to give a brief outline of the finite element method as applied to fracture problems. The reader is referred to one of the many pedagogical textbooks on the subject for more details (see for instance [105]).

Basically, the FEM are ways of discretizing the general continuum rheological equations on a lattice (usually triangular in two dimensions). The displacement field is approximated by a best fit within a subspace of trial functions. The best fit is found by minimizing the total potential energy over all admissible fields.

In the FEM, the description of a real elastic solid will be accurate only if the displacement field varies slowly over the size of the elements used. In particular, if one wants to introduce a crack in a medium, the displacement field around the crack will be accurate only if the elements are much smaller than the size of the crack itself. This is a serious drawback for the modeling of a dense array of microcracks in a medium. Similar drawbacks, although less dramatic, will appear if one wants to model a heterogeneous solid. In fact, in such a case, the size of the heterogeneities should be large compared to the mesh size. Also, fluctuations of elastic moduli from element to element should be small.

FEM are popular because of their simplicity (although elaborate codes often use all sorts of sophisticated optimizations) and flexibility in dealing with very different problems. These codes are tools and used like black boxes rather than being considered to have a physical interest on their own. The lattice models used by physicists are less sophisticated than FEM. However, they have the advantage of being built in such a way

as to have all the steps directly accessible and easily interpreted physically.

2.2.2 Lattice models considered by physicists

Models considered by physicists tend to have a certain common setting which can be summarized in the following way [45]: the medium is discretized such that all spatial sites are equivalent (contrary to FEM, the grid is not made finer in regions of higher stress) and have the same number of neighbors (regular lattice). The variables that characterize the medium are placed on the sites of the lattice. The equations that describe the medium are discretized so that, for each site, one has one equation per variable which only involves variables on the neighbor sites. Also, a boundary condition on the outer perimeter is implemented. The boundary condition on the internal boundary (crack surface) reflects the response of the bonds that constitute the crack. Then, the set of equations has a unique solution.

As mentioned in [45], the simulation of a rupture process must be done in an *iterative* way: the equations must be solved in order to determine which bonds should be broken, but once the bonds have been broken, the internal boundary condition gives rise to a change in the internal stress. Consequently, the equations must be solved again if one wants to know which bond to break next, and so on. The process is computationally demanding because the set of linear equations must be solved as many times as bonds are broken (unless several bonds are broken simultaneously). The algorithm performed in one iteration can be decomposed into four steps [45]:

- (1) solve the set of equations,
- (2) determine the set of bonds eligible to be broken,
- (3) select from the eligible bonds one which will be broken,

- (4) break the bond, i.e. change its elastic properties.

Each of these steps allows for a large variety of options that describe many possible physical situations. Step (1) describes the nature of the medium and the externally applied constraints, step (2) the connectivity of the crack, step (3) the breaking rule and disorder and step (4) distinguishes for instance between a breakdown and a fuse problem (see [45] for the details). We now briefly discuss each of these steps.

To do step (1), the type of model and the type of external constraints need to be specified. For the elasticity, scalar, vectorial and tensorial models are found in the literature. To our knowledge, there exists only one scalar model dealing with elasticity and it was introduced in [68]. This model simulates antiplane shear deformation along the axis Oz of a thin plate placed in the x - y plane. A constant slow antiplane velocity along Oz is applied at the boundaries of the plate. Vectorial and tensorial models allow for a larger variety of externally imposed constraints than scalar models. Vectorial models of common use are the central-force model [32], the bond-bending model [52] and the beam model [86]. The latter vary in the discretization scheme used to solve the equations of elasticity. The central-force model, the simplest of the three, is a network of Hookean springs which can freely rotate around the sites of the lattice. So, in this model, force can act only in the direction of the bond. The other two vectorial models fall in the same class in the sense that a bond can not rotate around a site without a cost in energy. Tensorial models have been introduced only recently by Chen et al. [15] and Xu et al. [113, 114] and have some advantages over vectorial models, in particular they are computationally less demanding. The discretization scheme used in [113, 114] will be presented in the next section and compared to the one used in [15] in section 2.5.

To study crack propagation [step (2)], one can pursue two philosophies: understand how one crack grows or see how in a stressed medium cracks appear, coalesce and break

the system. The first philosophy is artificial in most experimental situations since the surface of a crack in a lattice is not uniquely defined. The second philosophy corresponds to the majority of experimental situations because all not yet broken bonds are eligible to be broken.

The breaking rule and disorder are two major ingredients of the model [step (3)]. The breaking rule must be put by hand into the model to reflect the phenomenological nature of lattice models. It is usually chosen as simple as possible. Disorder has a major impact on the breaking rule as well as on the breaking process. If disorder is unaffected by the breaking process, then it is *quenched* and otherwise it is *annealed* (interplay of disorder and fracture has to be taken into account). On the length scales that lattice models are supposed to describe, disorder can be reduced to a spatial dependence of the elastic modulus and/or strength, often chosen to be a random distribution.

The bond represents the system on a mesoscopic scale and is characterized by a constitutive law (e.g. force F linearly dependent on the displacement δ). The relation between F and δ can be chosen to remain linear up to a breaking threshold δ_c . The strength of the bond is given by the value of F at $\delta = \delta_c$ and the elastic modulus by the slope of the straight line. Disorder can be implemented by randomly choosing the strengths and/or the elastic moduli of the bonds of the lattice. In quenched systems, the random variables (strength and elastic moduli) are fixed at the beginning of the simulation and the following rupture process is *deterministic*: to break a bond, the equations are solved and the bond for which δ/δ_c is the largest is chosen to be broken. On the other hand, annealed disorder leads to a *probabilistic* rupture process. In this case, the bond to be broken may be chosen according to a probability $p \propto \delta^n$, where n is a given exponent [66].

Finally, regarding step (4), once the bond to be broken has been chosen, its actual removal or replacement has to be implemented. In the case of fracture, the elastic

modulus of that bond can be set to zero or reduced by a certain amount.

2.3 A New Lattice Model

The central-force, bond-bending and beam models (vectorial models) are *physical* in the sense that the basic elements (Hookean springs or elastic beams) have a real physical meaning. However, these models have some drawbacks [45]. One of them, and surely the most severe one, is the computer time involved. It is enormous due to the fact that, at each time step, one has to solve a huge array of linear equations. This is a major obstacle to progress in this area. The (tensorial) lattice model that we use was introduced by Xu et al. [113, 114] and it overcomes some of the difficulties encountered in the above vectorial lattice models. It proposes a new method of discretizing the continuum elastic equations and has some advantages over the latter.

2.3.1 A new discretization scheme

Xu et al. have introduced recently a new scheme to discretize the continuum elastic equations. Because it is not as well known as the discretization schemes presented in [45], we summarize here the most important steps of this new scheme. The interested reader is referred to [114] for the details.

The first step is to construct discrete derivatives. We do this only for a two-dimensional system for simplicity. (Generalization to a three-dimensional system is straightforward in principle). These derivatives are of the form (say for a function $g(\vec{r})$)

$$D_x g(\vec{r}) \equiv \frac{1}{2} \left[g\left(\vec{r} + \frac{\hat{x} + \hat{y}}{2}\right) + g\left(\vec{r} + \frac{\hat{x} - \hat{y}}{2}\right) - g\left(\vec{r} - \frac{\hat{x} + \hat{y}}{2}\right) - g\left(\vec{r} - \frac{\hat{x} - \hat{y}}{2}\right) \right]$$

and

$$D_y g(\vec{r}) \equiv \frac{1}{2} \left[g\left(\vec{r} + \frac{\hat{x} + \hat{y}}{2}\right) - g\left(\vec{r} + \frac{\hat{x} - \hat{y}}{2}\right) - g\left(\vec{r} - \frac{\hat{x} + \hat{y}}{2}\right) + g\left(\vec{r} - \frac{\hat{x} - \hat{y}}{2}\right) \right]$$

where \hat{x} and \hat{y} are unit vectors oriented along the X-axis and Y-axis respectively, and the lattice constant (a) is taken to be unity. Note that if g sits on the nodes of a square lattice, then $D_\alpha g$ ($\alpha = x, y$) sits on the centers of the squares formed by the nodes of the lattice, and vice versa. The above derivatives ensure that the basic unit of discretization is a *square* instead of a bond.

Using these discrete derivatives, the discretization of elastic theory is straightforward. First, a displacement vector \vec{u} is defined on each node (corner of a square). The distortion of a square is characterized by the strain tensor (defined at the center of each square):

$$u_{ij}(\vec{r}) = \frac{1}{2}[D_i u_j(\vec{r}) + D_j u_i(\vec{r})].$$

Initially, all deformations are elastic and the stress tensor is related to the strain tensor through the special Hooke's law:

$$\sigma_{ij}(\vec{r}) = \lambda' \delta_{ij} u_{ll}(\vec{r}) + 2\mu u_{ij}(\vec{r}),$$

where μ is the shear modulus, and λ' a Lamé constant modified to take into account the plane-stress geometry of the problem (see [59] for instance). Summation over repeated indices ($i, j = x, y$) is implied. The discretized elastic equations are given by the force balance condition $F_i = D_j \sigma_{ij} = 0$ which must be satisfied on each square. We further restrict our consideration to the shear mode of rupture (most ruptures in earthquakes are shear mode fractures!). In this case, only two stress components are needed to determine where a rupture will occur: $\sigma_1 = \sigma_{xy}$ (rupture along the X-axis or Y-axis) and $\sigma_2 = (\sigma_{yy} - \sigma_{xx})/2$ (rupture at 45° or 135° to the X-axis). In this thesis, we consider only the stress component $\sigma_1 = \sigma_{xy}$ for simplicity. We do not think that to have two stress components would have a major influence on the results presented in this thesis as well as in [42].

Now, a rupture is assumed to occur at the square centered at \vec{r}_0 in an infinite medium. This happens when $\sigma_{xy}(\vec{r}_0) > \sigma_{th}$, where σ_{th} is the stress threshold of that square. The

additional shear stresses (σ'_{ij}) caused by the rupture are separated into an elastic part (σ^{el}_{ij}) and a non-elastic part (at \mathbf{r}_0 only) [to express the violation of Hooke's law at this location]. The new stress tensor components (σ^{new}_{ij}) can be expressed as a function of the stress tensor components before the rupture (σ^{old}_{ij}) in the following way:

$$\sigma^{new}_{ij}(\vec{r}) = \sigma^{old}_{ij}(\vec{r}) + \sigma'_{ij}(\vec{r}),$$

where

$$\sigma'_{xy}(\vec{r}) = \sigma^{el}_{xy}(\vec{r}) - f\sqrt{2}\delta_{\vec{r},\vec{r}_0} \quad (2.1)$$

$$\sigma'_{xx}(\vec{r}) = \sigma^{el}_{xx}(\vec{r}) \quad (2.2)$$

$$\sigma'_{yy}(\vec{r}) = \sigma^{el}_{yy}(\vec{r}). \quad (2.3)$$

At equilibrium, σ^{new}_{ij} must satisfy the force balance condition on each square and as a result,

$$D_j \sigma'_{ij} = 0 \quad (2.4)$$

since $D_j \sigma^{old}_{ij} = 0$. Using (2.1), (2.2) and (2.3) in (2.4), it can be shown [114] that

$$D_j \sigma^{el}_{ij} + F_i^d = 0,$$

where

$$F_x^d = \frac{f}{\sqrt{2}} \left[\delta_{\vec{r},\vec{r}_0+\frac{\hat{x}+\hat{y}}{2}} - \delta_{\vec{r},\vec{r}_0+\frac{\hat{x}-\hat{y}}{2}} - \delta_{\vec{r},\vec{r}_0-\frac{\hat{x}+\hat{y}}{2}} + \delta_{\vec{r},\vec{r}_0-\frac{\hat{x}-\hat{y}}{2}} \right]$$

and

$$F_y^d = \frac{f}{\sqrt{2}} \left[\delta_{\vec{r},\vec{r}_0+\frac{\hat{x}+\hat{y}}{2}} - \delta_{\vec{r},\vec{r}_0-\frac{\hat{x}+\hat{y}}{2}} - \delta_{\vec{r},\vec{r}_0-\frac{\hat{x}-\hat{y}}{2}} + \delta_{\vec{r},\vec{r}_0+\frac{\hat{x}-\hat{y}}{2}} \right].$$

Therefore, σ^{el}_{ij} can be viewed as generated by the external source $\vec{F}^d = (F_x^d, F_y^d)$, which is only non-zero at the corners of the ruptured square, and is represented on figure 2.1. \vec{F}^d satisfies the conditions that its net force and net torque are zero, and thus is a double couple (a term used in seismology to describe the earthquake source having double couple

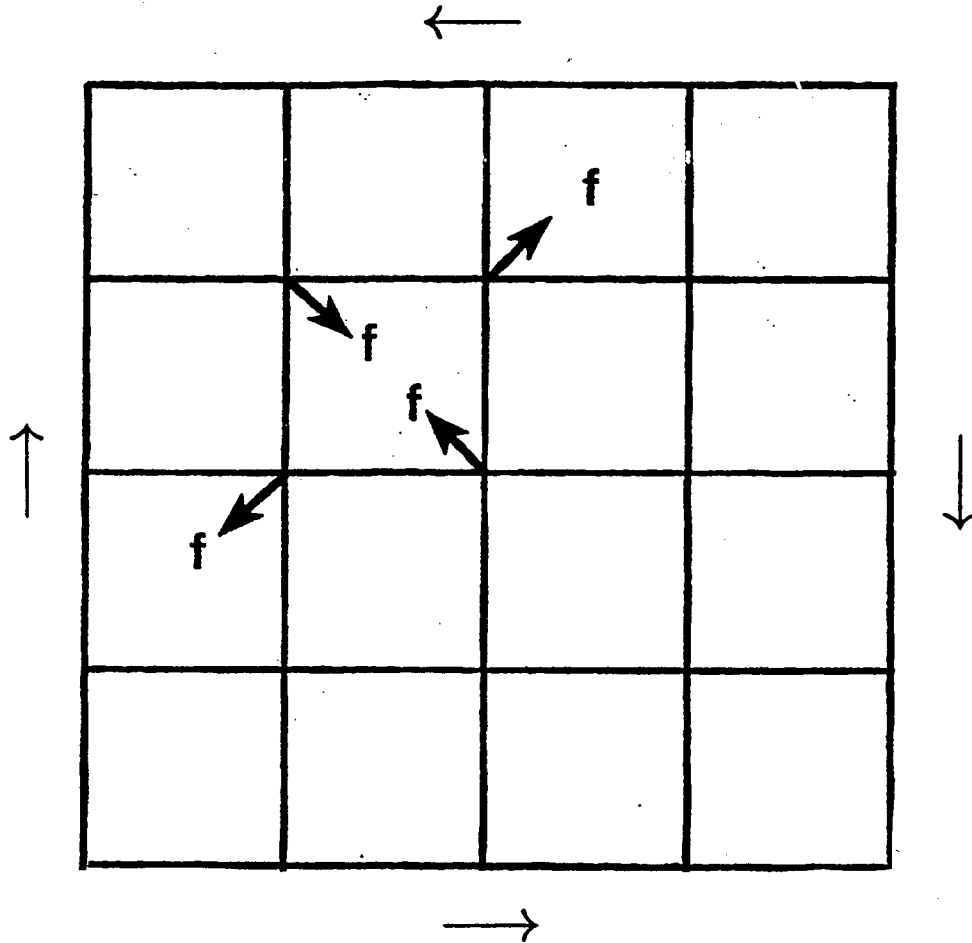


Figure 2.1: One-layer *planar* region considered in the model of Xu et al. The region is discretized into squares. It is embedded in an infinite medium (providing the necessary external drive) and the effect of any activity originating outside the region is neglected (open boundary conditions). In the model, the interaction between the external drive (represented by the arrows outside the region) and a square of the region is treated in a mean-field way, i.e. it is the same whatever the square location. The arrows at the corners of one of the square inside the region are a schematic representation of a double couple force redistribution following the shear rupture of that square. This figure was modified from [114].

force f). The discretization scheme of Xu et al. is in agreement with the fact that a shear rupture can be modeled by a double couple force redistribution [11].

Next, $\sigma_{ij}^{el}(\vec{r})$ can be determined by using the above earthquake source and the method of Fourier transformation. Since σ_{xy} is the *only relevant* component of the stress tensor, we drop the subscript (from now on, stress simply refers to the shear stress component σ_{xy}). It can be shown [114] that the redistributed field is given by

$$\sigma'(\vec{r}) \equiv \sigma'_{xy}(\vec{r}) = -\tilde{f}G(\vec{r} - \vec{r}_0), \quad (2.5)$$

where $\tilde{f} = f\sqrt{2}(\lambda' + \mu)/(\lambda' + 2\mu)$ and G is a lattice Green function expressed as

$$G(\vec{r}) = \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \int_{-\pi}^{\pi} \frac{dk_y}{2\pi} \frac{\sin^2 k_x \sin^2 k_y}{(1 - \cos k_x \cos k_y)^2} e^{i\vec{k} \cdot \vec{r}}. \quad (2.6)$$

In appendix A, we illustrate how to simplify the calculation of G by reducing the number of integrals in (2.6) from two to one. This has a positive impact on the accuracy of the computation of $G(\vec{r} - \vec{r}_0)$ especially at large value of $\vec{r} - \vec{r}_0$. Also, in appendix B, we do a coarse-graining procedure to overcome an unphysical property of G . Our coarse-graining procedure is different from the one introduced by Chen et al. [16] in a study of the model of Xu et al. in the sense that it is not implemented in the derivation of the equations of the model, but instead applied directly to the Green function (2.6) [see appendix B].

In the model, the unit of discretization (square) may be sufficiently large that the whole square will not break in an individual rupture event. Instead, it will retain some shear stress. Assuming that

$$\sigma^{new}(\vec{r}_0) = X\sigma^{old}(\vec{r}_0), \quad (2.7)$$

where X is the fraction of the original stress which is not released, (2.5) can be rewritten as

$$\sigma'(\vec{r}) = -(1 - X)\sigma^{old}(\vec{r}_0)G(\vec{r} - \vec{r}_0)/G(\vec{0}). \quad (2.8)$$

After the original rupture at \vec{r}_0 , which is followed by stress redistribution, it can happen that the stress exceeds the stress threshold on several squares. We will then let them all break independently and add the respective stress redistribution contributions. This procedure was used as well in [113, 114]. Other rules for the order in which the squares break could have also been considered. For instance, we could have picked only the square whose stress exceeds its stress threshold most. We could, as well, have considered a self-consistent scheme in which the redistributed stresses from different broken squares are coupled as in [6]. We employ the independent rupture procedure because it is the simplest.

2.3.2 Advantages of that discretization scheme over others

The discretization scheme introduced by Xu et al. is rather simple and easy to implement. Before applying it to the study of earthquakes, we mention the advantages it has over discretization schemes discussed in [45]:

- (i) Each square is characterized by a strength. The length scale is set by the size of the square. This makes it possible to study engineering type problems (e.g. crack growth in ceramics) as well as geological scale problems (e.g. plate tectonics and earthquakes).
- (ii) The stress tensor is dealt with directly, so the simulation process is explicit in its physical interpretation.
- (iii) When open boundary conditions are used, one focuses on a finite region embedded in an infinite medium and the effect of any activity originating outside the region is neglected. In this case, the method of Xu et al. becomes straightforward since the calculation of the lattice Green function can be done as for an infinite region.

2.4 Application to the Study of Earthquakes

The phenomenology of earthquake occurrence is complex (for a review, see [55]). Many statistical relations have been deduced over the last century from the observation of earthquakes, which can be viewed as sequences of ruptures releasing the elastic stress built up slowly by the movement of tectonic plates. Among them are the well-known Gutenberg-Richter [43] and Omiri [78] power laws. The former relation expresses the frequency of occurrence of earthquakes of a given magnitude whereas the latter expresses the aftershock occurrence decay with time from the mainshock.

The global dynamics of earthquakes is rich and complex. There have been many attempts at earthquake modeling, yet there are still many unsolved problems due to the fact that one does not know how to model the physics well [55]. Knopoff suggested to introduce at least five constraints when modeling the complex phenomenology of earthquake occurrence [55]:

- (i) plate tectonics to restore the energy dissipated in earthquakes,
- (ii) a wide range of earthquake sizes,
- (iii) a brittle-ductile rheology to provide for the time delays between clustering events (aftershocks in particular) and the abrupt occurrence of ruptures,
- (iv) stress redistribution upon rupture to provide the mechanical coupling between earthquake events,
- (v) a complex geometry, to take into account the distribution of earthquakes as well as the finite size of the fracture surfaces.

The model of Xu et al. is in principle able to incorporate those five constraints as well as the tensor character of the stresses (the latter is also believed to be important for

earthquake modeling [55]). However, in practice, aftershocks will appear in this model only if a static fatigue law is introduced by hand into the model [113, 114] (static fatigue is believed to be the mechanism responsible for aftershocks [89]). Also, the model is two-dimensional and only a one-layer system is considered (see figure 2.1), so the complex geometry of fracture surfaces is approximately dealt with. In the version of the model of Xu et al. used in this work and in [42], aftershocks are absent since no static fatigue law is considered. However, even with this simplification, we will see that this version of the model still has many complex features.

To apply the model of Xu et al. to the study of earthquakes, we implement the following algorithm, divided into twelve steps:

1. Consider a finite (planar) region embedded in an infinite medium, neglecting the effect of any activity originating outside the region (open boundary conditions).
2. Divide the region into $L_1 \times L_2$ squares each having a size equal to unity.
3. Set the stresses on all squares to zero at the beginning of the simulation.
4. At the beginning of the simulation, introduce disorder by assigning random numbers between θ_l and θ_u to the stress thresholds (σ_{th}) of the squares. The random distribution of stress thresholds is taken as uniform in this work.
5. Drive the system, i.e. increase the stresses uniformly until on a square the stress exceeds the stress threshold (this is called a rupture). The amount of stress added to every square is $\Delta\sigma$. An earthquake sequence then begins and the driving process stops.
6. Increment the long time scale (defined by variable t throughout this work) by the amount $\Delta\sigma$.

7. Reset the stress threshold of the ruptured square to a random number between θ_l and θ_u (rule I) or to the constant value 1.0 (rule II).
8. Calculate the stress redistribution [using (2.8)].
9. Check if other ruptures occur in the region afterwards. If so, then step 7 is repeated for all the ruptured squares.
10. Repeat step 8 for all the ruptured squares (independently) and do step 9 again. The earthquake sequence stops when the stress is lower than the stress threshold on all squares.
11. Count the number of units (S) that have ruptured during the sequence. S is a measure of the magnitude of the earthquake. Note that if a square ruptures more than once during the sequence, every occurrence is counted (unless otherwise specified).
12. Go to step 5 to generate the next earthquake sequence.

We describe next most of the above steps in some details.

Step 1 describes the region and the boundary conditions considered. Physically, the planar (two-dimensional) region represents a thin piece of the earth's crust, or at least a region for which the largest horizontal dimension is much larger than its thickness (vertical dimension) [see figure 2.1]. Therefore, it means that as soon as earthquakes occur in this one-layer region, the faults generated break the region over its entire thickness. The activity happening outside this region is neglected. However, the medium surrounding the region of interest (see figure 2.1) is essential since it provides the necessary external drive (see step 5 for the details). The boundary conditions that represent best such a situation are the *open boundary conditions*. The latter greatly facilitate the calculation of the lattice Green function (2.6) since it can be done as for an infinite region.

In most of the models studying SOC, the continuous space is discretized into elements (step 2). It is the existence of such distinct elements that is the basis of the complex response of the model to an external drive. This is due to the fact that each element acts as an oscillator of relaxation. The coupling between many such oscillators with different characteristics is known to generate complicated behaviors. It could be argued that the results presented in this work are specific to the discrete nature of the model of Xu et al. and are thus not relevant to real geology. However, as suggested by Sornette et al. [95], it is clear that the presence of heterogeneities in the earth's crust generates a distribution of characteristic scales. Somehow, models such as the one of Xu et al. provide simple ways to represent the heterogeneities which are rescaled from the infinitesimal scale up to the smallest scale considered in the models, i.e. the lattice constant (a).

In step 4, we take into account the fact that the region modeled is heterogeneous. However, it is of interest to note that disorder is incorporated *solely* in the stress thresholds and not in the elastic constants. The latter form of disorder would destroy the translational invariance of the lattice Green function as mentioned in [114]. This represents one limitation of the model.

The driving of the region (step 5) is done *slowly* and *uniformly*. The slow driving (this defines the long time scale) is not an artificial constraint imposed but, as we have seen in chapter 1, is fundamental to get SOC in a model. This constraint makes sense for plate tectonic problems. The uniform driving condition is a *mean-field* approximation of what happens (or happened) in nature. Indeed, in the case of the Asian continent deformed by the indentation of India for instance, the stress increase was distributed non-uniformly over large distances. In such a case, the model of Xu et al. would treat the interaction between the external drive (India in the above example) and the region (the Asian continent in the above example) as being independent of the position in the region. This is another limitation of the model.

In step 6, it is implied that the rate of stress increase (p) is set to the value 1. The choice of such a value is motivated by practical considerations: because no aftershocks or foreshocks can happen in the current version of the model of Xu et al., it is not necessary to use a small value of p (e.g. $p = 10^{-9}$) since no rupture can occur between two earthquake sequences. As a result, the amount $\Delta\sigma$ is added *as a whole* rather than being added little by little to all the squares of the lattice, and this makes the computation much faster. It is of interest to note that the value of p determines the order of magnitude of the long time scale (defined by variable t): $t \sim p^{-1} = 1$. Also, once a sequence has begun, t is *frozen*. This reflects the fact that $t/t_{seq} \gg 1$, where t_{seq} is the short time scale (or time scale for the duration of an earthquake sequence). As a result, the two time scales in the model are well separated, a fact which is important to obtain SOC in a model as discussed in chapter 1.

The resetting rules of the stress threshold of a ruptured square are defined in step 7. Rules I and II do not have a physical basis contrary to the rules used by Xu et al. in [113, 114], for which slow annealing of the fracture surfaces was allowed. In particular, rules I and II avoid the problem of having to deal with the whole range of time scales implicit in the annealing process of a fracture surface. The big advantage of using rules I and II is that they allow for the easy investigation of the effect of disorder on the scaling properties of the model (see chapter 3).

Step 10 defines an independent redistribution procedure when more than one square break simultaneously. When such a procedure is used, the model of Xu et al. is said to be a *series* model (according to the terminology in [35]). This means that the time for the stress on a ruptured square to be released is shorter than the time needed to redistribute (transmit) that stress to the other squares. Therefore, we have a series of stress releases, followed by a series of stress transmissions, followed by a series of stress releases and so on. We will see in chapter 4 that this independent redistribution procedure gives rise to

earthquake sequences that are temporally not correlated. The model of Xu et al. is also said to be *quasistatic* because it does not describe the details of the dynamical process of rupture, but only the stress distribution before and after an earthquake sequence (treated as an instantaneous event on the long time scale). In other words, the inertial forces *on the long term*, which are very small compared to other forces, are neglected and thus, the region is always in mechanical equilibrium. Therefore, the model of Xu et al. describes the avalanche analogue of a fracture, or a moving dislocation [97]. According to Sornette et al. [95], a series (or dislocation) model is more adapted to the study of large earthquakes, i.e. the ones which break the entire thickness of the earth's crust. The latter situation is precisely the one which we consider in this work (see the above paragraph related to step 1).

In this work, we estimate the magnitude of an earthquake in a simplified way by counting the number of squares that have broken during an earthquake sequence (step 11). The use of such an estimator of the earthquake magnitude (i.e. S) is convenient since most of the work that has been done on earthquake models considered a similar estimator. As a result, the comparison of our results to the ones obtained by other people will be easier. Obviously, other estimators of the earthquake magnitude could have been used, such as $\bar{\sigma}_{bef} - \bar{\sigma}_{aft}$, where $\bar{\sigma}_{bef}$ and $\bar{\sigma}_{aft}$ are the average stress in the system before the earthquake and the average stress in the system after the earthquake respectively. However, we do not believe that the scaling properties of the model of Xu et al. (see chapter 3) will be *qualitatively* dependent on the estimator of the earthquake magnitude used, and so, we consider only the above definition in the present work.

The algorithm outlined above describes how to apply the model of Xu et al. to the study of earthquakes. In its current version (and also in [42]), this model conserves the stress if the small fraction of it lost at the boundaries of the region due to the long-range stress redistribution is ignored. (This fraction goes to zero as the size of the region goes

to infinity). In addition, stress dissipation can be incorporated into the bulk of the region as well by multiplying $\sigma'(\vec{r})$ at all \vec{r} except \vec{r}_0 (rupture location) by a factor α between 0.0 and 1.0 [see (2.8)]. The reason for introducing parameter α is to investigate the effect of the degree of stress nonconservation on the scaling properties of the model, a task which will be performed in chapter 3.

In all cases, the system organizes itself into a stationary state where *on average* the stress added to the region balances the stress removed from it. The stationary state of the model can be obtained by initially discarding a large number of earthquake sequences. Afterwards, by generating a sufficient number of subsequent sequences, it can be checked whether the cumulative size-frequency distribution obeys the Gutenberg-Richter power law [43]

$$C(S) \propto S^{-\tau}, \quad (2.9)$$

where $C(S)$ is the fraction of sequences during which S or more squares have ruptured, and τ is the exponent of the power law.

2.5 Comparison of the Model of Xu et al. to Other Models

Before we study the scaling properties of the model of Xu et al. (see chapter 3), we compare the latter to other earthquake models and to models of integrate-and-fire oscillators.

2.5.1 Model of Xu et al. versus other earthquake models

Friction and fracture are two equally important phenomena occurring during earthquakes. To our knowledge, only the model of Mora and Place [71] is able to incorporate both friction and fracture in a unique model. Most of the earthquake models deal with either friction or fracture. The models dealing with friction only are *stick-slip models* and the ones dealing with fracture only are *lattice models* (see section 2.2). The model of Xu et

al. (presented in section 2.3) is a lattice model. We now describe in some details the model of Mora and Place, stick-slip models and other lattice models applied to the study of earthquakes.

The model of Mora and Place is based on molecular dynamics. While it is infeasible to directly simulate the number of particles required for geophysical scales, this model can be used to simulate larger units such as pieces of rock. The interactions between the particles are specified through potentials which are compatible with geophysical observations. Up to now, the number of particles considered in the model of Mora and Place has not been sufficient to simulate in a realistic way the behavior of a two-dimensional tectonic plate (see [25]).

Stick-slip models focus on regions with a pre-existing fault. One of the first model of this type was the one-dimensional model of Burridge and Knopoff [12]. It consists in a chain of blocks lying on a rough stationary surface and connected by springs. Also, the blocks are all connected by flat springs to a top surface moving with a certain speed relatively to the bottom surface (system driving). The dynamics of the model is defined as follows: as long as all the F_i 's (F_i is the total force on block i) are smaller than a constant threshold value (F_{th}), then all the blocks are stuck. As the system is being driven, eventually one block i will have $F_i > F_{th}$ and the excess force will be relaxed. The excess force will be transferred equally to the two neighboring blocks, which can in turn slip if they overcome the static friction force with the bottom surface. A chain reaction (sequence) might thus happen, which can involve an arbitrary number of blocks (N). The frequency distribution of sequences involving N slipping blocks was shown not to be a power law by Burridge and Knopoff [12].

The model of Burridge and Knopoff has inspired a great deal of work in the last decade. For instance, Carlson and Langer [13] added a nonlinear friction law to the original model and got the power law that Burridge and Knopoff were unable to obtain.

Nakanishi [73] also got a power law by studying a coupled map lattice version of the original model of Burridge and Knopoff. Despite these findings, the model of Burridge and Knopoff as well as its many versions were shown to NOT be self-organized critical [93]. On the other hand, the model of Olami et al. [77] was shown to display SOC [17]. The latter model was also inspired from the model of Burridge and Knopoff for its derivation but, apart from this, it can be regarded as a two-dimensional generic representation of a nonconservative model, since the final equations of the model of Olami et al. do not contain the physical constants of the original model. In this sense, we do not consider the model of Olami et al. as a stick-slip model. *Stick-slip* is a label which will be reserved for the models displaying explicitly the physical constants of the model of Burridge and Knopoff.

The stick-slip models are in principle able to solve the fully dynamic problem where, during an earthquake, a rupture propagates with a speed close to that of shear waves. However, these models are too simplified from a geometrical point of view: the region modeled is that of a pre-existing fault with many asperities coupled only to their nearest neighbors and so, nothing is said about how the fault was generated and about the effect of the surrounding medium.

The creation of faults is related to the fracture phenomenon and can be modeled using lattice models such as the ones discussed in sections 2.2 and 2.3. These models use a continuum (macroscopic) description of matter as opposed to the discontinuous (microscopic) description used in the model of Mora and Place [71]. The continuum description of matter allows for the easy prediction of where the rupture will occur but poorly describes the evolution of discontinuities such as faults [27, 106]. Apart from the models discussed in [45], other models using a continuum description of matter include the ones of Chen et al. [15] (see also a modified version of it in [6]) and Xu et al. [113, 114] (see section 2.3). The latter have been applied to the study of earthquakes.

They differ in the discretization scheme used to solve the continuum elastic equations. In [15] (and [6]), the basic unit of discretization is a bond whereas in [113, 114], it is a square. As a result, the earthquake source is a dipole in [15] (and [6]) and a quadrupole (double couple) in [113, 114]. A rupture is followed by long-range stress redistribution in both models however. This should be contrasted with stick-slip models and the model of Olami et al. [77] in which the stress redistribution is to the nearest neighbors only.

2.5.2 Model of Xu et al. versus models of integrate-and-fire oscillators

A class of models, which are mainly applied to the study of biological systems, has appeared in the last decade. These models focus on large assemblies of oscillators which can spontaneously evolve to a state of large scale organization. Collective synchronization is the best known phenomenon of this kind. This effect has attracted much interest in the study of fireflies [10], pacemaker cells of the heart [82], cells of the pancreas [90], etc. Most of the works on synchronization have used models in which the interactions between the oscillators are smooth and continuous in time. Comparatively few works have been done with models in which the interactions are episodic and pulselike, although these models are relevant to the study of certain species of fireflies for instance. Models of this type are often called models of *Integrate-and-Fire Oscillators* (IFO).

Models of IFO are closely related to the models of Xu et al. [113, 114] and Olami et al. [77]. This can be seen by identifying the state variable of earthquake models (usually a shear component of the stress tensor) with the voltage magnitude of models of IFO. It is of interest to note that the models of IFO study the level of mutual entrainment (i.e. the synchronization) between the elements whereas the earthquake models focus on the size distribution of the sequences generated (if it is a power law, then the model might display SOC). Considering the fact that both of these classes of models contain the same basic ingredients then, from this point of view, SOC and synchronization can

be considered as two sides of the same coin. For completeness, we mention that, contrary to earthquake models, the models of IFO could allow for more general driving processes than slow uniform driving (e.g. random driving of individual elements).

Models of IFO exist in at least two varieties. In some models [69, 7], the oscillators are equally coupled with each other, so the firing of one oscillator leads to an equal redistribution of voltage to all the other oscillators in the system. In other models (see [19] for example), the oscillators are coupled only to their nearest neighbors, so the firing of one oscillator leads to a voltage redistribution to the nearest neighbors only. Therefore, as in earthquake models, the firing of one element leads to the redistribution of the state variable either to all the other elements or to the nearest neighbors of that element only.

Chapter 3

Study of the Scaling Properties of the Model

The scaling properties of the model of Xu et al. [113, 114] are studied by means of a finite-size-scaling analysis. In section 3.1, we give the motivation for such a study and in section 3.2, we present how to implement finite-size scaling in the context of SOC in an earthquake model. The results of the analysis, obtained by considering one-dimensional lattices, are presented in section 3.3 in both the conservative and non-conservative cases. The stress ordering in one-dimensional lattices is investigated in section 3.4. This investigation is useful in explaining some of the results obtained in section 3.3. We introduce in section 3.5 a mean-field version of the model. Finally, in section 3.6, we summarize the important results obtained in this chapter and draw our conclusions.

3.1 Motivation for this Study

The Gutenberg-Richter power law (2.9) is obtained only in the limit of an infinite system. Obviously, the power law may also appear for a finite system, but it will prevail only over a *limited range*. The system is never completely scale invariant. It is then important to sort out effects due to the finite size of the system and those due to the finite size of the basic unit of discretization. Such a problem arises because systems of interest are heterogeneous on scales much smaller than the discretization unit, so that the continuum limit is not meaningful. Quantities relating to the stress threshold (σ_{th}), the degree of stress retained by a ruptured square (X) [see section 2.3] and the degree of stress conservation (α) [see section 2.4] may therefore depend upon the size of the discretization unit and the size

of the system in a non-trivial way. This type of dependence was realized by Weibull [110], but has since often been forgotten. To illustrate more specifically these ideas, we discuss the effect of parameter X introduced in section 2.3 [see (2.7)]. The case $X = 0$ corresponds to a rupture opening the whole element, as in [6]. The interpretation of a rupture event and the scaling properties are somewhat different if $X \neq 0$. We expect parameter X (as well as the stress threshold distribution) to depend on the size of the discretization unit. In fact, two discretizations of the same system can be considered, one with a coarser mesh than the other. Then, a coarser mesh element will rupture more often than a finer mesh element, but the total number of rupture events in the whole system will be the same. These ideas will be illustrated quantitatively in section 3.3.

Disentangling the effects due to the finite size of the system from those due to the finite size of the basic unit of discretization is the motivation of the study in this chapter. To achieve this, we need to vary the various parameters of the model of Xu et al. From what we said above, it is clear that varying the parameters of the model is not merely a mathematical procedure: it is a physically motivated one. The results obtained from the study in this chapter will ultimately serve to confirm whether the model of Xu et al. really displays SOC. Before we present our results, we briefly describe finite-size scaling.

3.2 Finite-Size Scaling

It is now well known that a finite system can not have a true singularity in its thermodynamic properties at a non-zero temperature. For instance, this is easily seen in the case of the Ising model in any dimension. Experimental systems, even if very large, are always finite. For all practical purposes, true phase transitions do occur in these finite systems and this fact must be reconciled with the rigorous proof that no phase transition can occur in any finite system. This is one of the roles of finite-size scaling. As an illustration

of this, we consider a $L \times L \times L$ system. Such a system falls into the category mentioned above of a system without a true thermodynamic equilibrium phase transition. In this situation, the susceptibility and specific heat remain finite at all temperatures and the power law divergences of the infinite system are replaced by rounded peaks. Finite-size scaling describes the dependence of the peak height and the range in temperature over which the power-law behavior is observed on the system size L .

Finite-size scaling has been extensively used in the study of equilibrium phase transitions (see for instance [84]). Its use in the present work is justified by the fact that, as argued by Sornette et al. [94], SOC can often be mapped onto ordinary criticality. The way to implement mathematically finite-size scaling, in the context of SOC in an earthquake model, is to replace the Gutenberg-Richter power law (2.9) (valid in the limit of an infinite system) by

$$C(S, L) = S^{-\tau} g[(S - 1)L^{-\nu}], \quad (3.1)$$

where L is the largest dimension in a $L_1 \times L_2$ lattice, g a scaling function and ν an exponent that expresses how the finite-size effects scale with the size of the system. When comparing systems with different parameter values, we also attempt

$$C(S, L) = S^{-\tau} g[(S - 1)L^{-\nu}/f(X, \theta_l, \alpha)], \quad (3.2)$$

where we assume that parameter θ_u , the upper bound for the stress threshold distribution (see step 4 of section 2.4), is set to a fixed value. The function f is included to test our hypothesis that, if the mesh size for a given system is changed, then some of the parameter values must also be changed. The term $(S - 1)$ in the argument of the scaling function ensures that $g(0) = 1$, i.e. when scaling, we map the size interval $[1, \infty)$ on $[0, \infty)$. It is of interest to note that the function g acts as a correction factor for the finite size of the system and the finite size of the discretization unit.

We are now ready to do our finite-size-scaling analysis. Our hope is that finite-size scaling will turn out to be as useful for SOC as for ordinary criticality.

3.3 Results of the Finite-Size-Scaling Analysis

We consider lattices with $L \times L_2$ units, where $L_2 = 1$ (unless specified otherwise). The lattice configuration considered is therefore a *strip* (i.e. a one-dimensional lattice). We believe that the scaling properties of the model of Xu et al. will be sensitive to this particular lattice configuration though, in the sense that with two-dimensional lattices, we would probably find different values for the exponents τ and ν [see (3.1) and (3.2)]. However, the use of one-dimensional lattices allows one to go to large lattice sizes, and check if the scaling laws are valid over the whole range of lattice sizes.

Now, regarding the parameters of the model, θ_u is fixed to the value 1 while θ_l , X (see subsection 3.3.1) and α (see subsection 3.3.2) are allowed to vary. In particular, parameter α can be set to 1 (conservative case) and to values smaller than 1 (non-conservative case). Also, the degree of disorder in the distribution of stress thresholds ($\sigma_{th} \in [\theta_l, \theta_u \equiv 1.0]$) can be changed by using one or the other resetting rules (see step 7 in section 2.4). The use of rule I maintains the degree of disorder of the initial distribution of stress thresholds, whereas the use of rule II makes disorder disappear after a small number of iterations. Below, the latter two situations will be respectively identified as *random* σ_{th} and *constant* σ_{th} .

We let the model reach a stationary state by discarding the first 10^5 earthquake sequences and include 10^6 sequences to do the finite-size-scaling analysis.

3.3.1 Results in the conservative case ($\alpha = 1$)

We show in figure 3.1 the cumulative size-frequency distribution $C(S, L)$ for random σ_{th} and various L values. The main features are the power-law behavior for small events, and the L -dependent fall-off of the frequency of large events. The power-law behavior extends over two or three decades depending on L . In figure 3.2, we note that the data sets of figure 3.1 obey quite well the scaling behavior predicted by (3.1) [see the middle curves]. The exponent $\tau = 0.30$ was found and cannot be changed by more than 0.01 before there is a noticeable deterioration in the data collapse. As can be seen from figure 3.1, this value of τ is smaller than the one obtained from a linear fit to the straight segment of the log-log plot (especially for $L = 200$ and $L = 400$). However, the exponent τ is the same if we count the number of ruptures on distinct units or if we count all ruptures. The exponent ν , on the other hand, is different in these two cases, as is the scaling function g . When all ruptures are counted, the same exponents are found for the $L \times 1$ and $L \times 2$ lattices, but the scaling functions are different. We attempted a data collapse of the data sets for the $L \times 1$ and $L \times 2$ lattices (when all ruptures are counted) by rescaling the horizontal axis for one of the sets, but found it to be significantly worse than that of figure 3.2.

We now turn to the effect of having constant σ_{th} . From figure 3.3, we see that large events become less common. We were unable to find a scaling function which fits both small and large events for $L \times 1$ lattices, no matter if all ruptures or ruptures on distinct units (results not shown) are counted. However, from figure 3.3, it is clear that the *shoulder* broadens as L increases, so $C(S, L)$ might scale with L if the small events (i.e. those with $S < 50$) are not considered in the data collapse. We found that, for $\tau = 0.30$ and $\nu = 1.38$ (values obtained above with $L \times 1$ systems having random σ_{th}), the scaling function (3.1) works well with systems having constant σ_{th} when only the events with

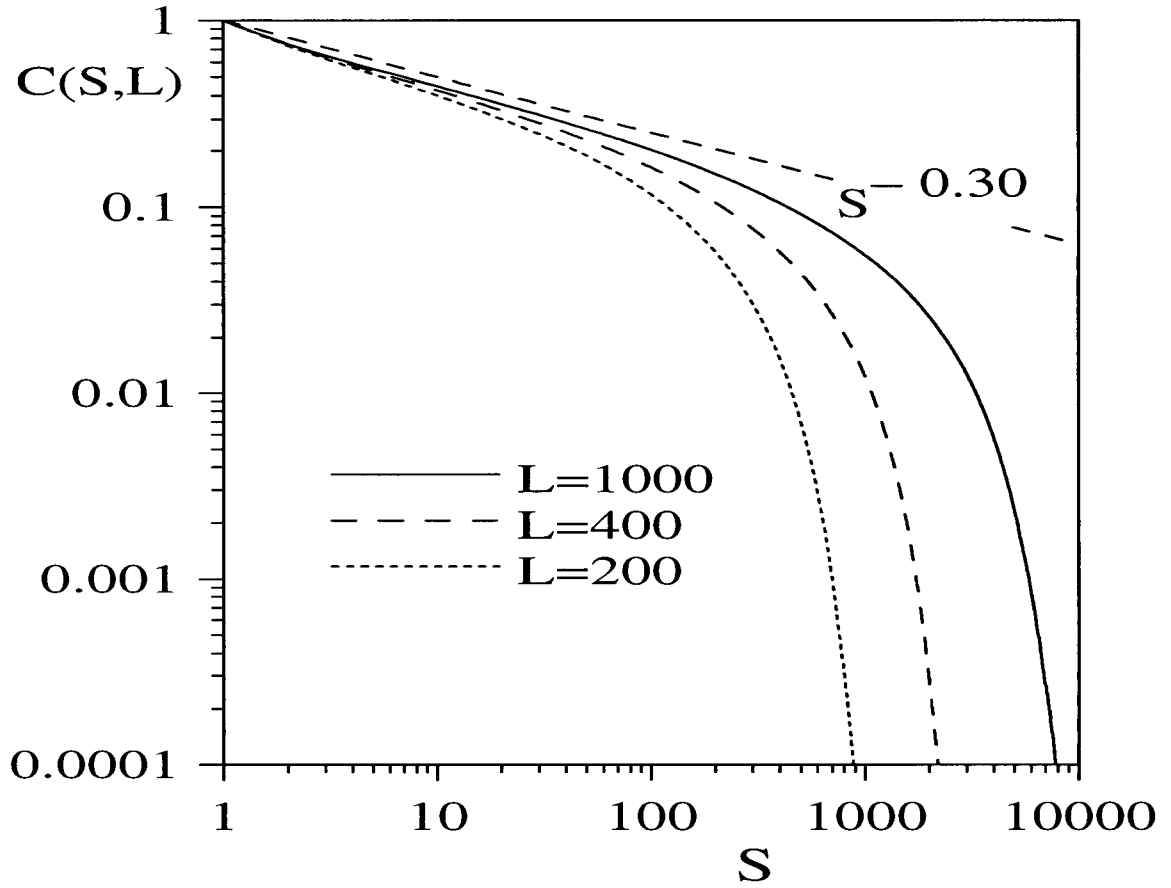


Figure 3.1: Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. For comparison, we show the curve $C(S) = S^{-0.30}$.

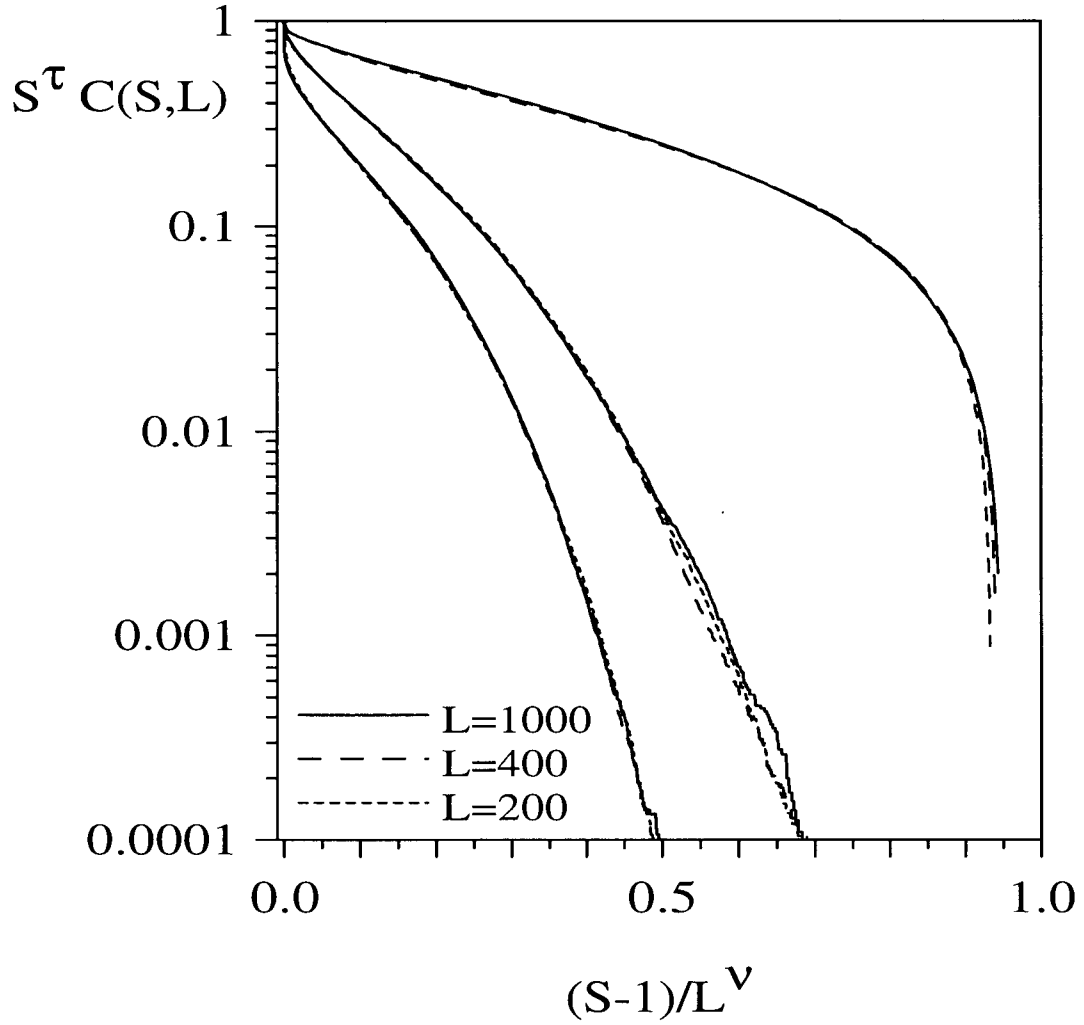


Figure 3.2: Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ and $L \times 2$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. All curves use the fitting parameter $\tau = 0.30$. For the left and middle curves, all ruptures are counted while for the right curves, only ruptures on distinct squares are counted. The left curves were obtained for $L \times 2$ lattices using $\nu = 1.38$, the middle curves for $L \times 1$ lattices using $\nu = 1.38$, and the right curves for $L \times 1$ lattices using $\nu = 1.01$.

$S > 50$ are included in the fit (results not shown). This shows in particular that exponent $\tau = 0.30$ is universal in the sense that it is independent of the degree of disorder in the model. On the other hand, the *crossover region* ($S < 50$) on figure 3.3 is not a universal feature since it appears only for systems with constant σ_{th} . We have checked that the crossover region is a robust feature which is not sensitive to the initial configuration of the model. Therefore, even if an ensemble average of initial configurations is carried out (see [45]), the crossover region remains. The presence of the crossover region will be explained in the next section by studying the stress ordering in one-strip systems with constant σ_{th} .

We next investigate the effect of the parameter X controlling the fraction of stress retained by a ruptured square. The data collapses on figure 3.4 are seen to be good for systems with random σ_{th} . Also, for a fixed L , we note that an increase in X extends the power law to larger S . This makes sense because a larger X means that a ruptured element retains more stress, so it will be easier to have it involved in the near future in an earthquake sequence (because stress remains close to stress threshold). As a result, the largest size for an earthquake sequence increases. In figure 3.5, we show, for a 1000×1 lattice and random σ_{th} , a fit to the scaling form

$$C(S, L) = S^{-\tau} g[(S - 1)L^{-\nu} / f(X)].$$

The fit, although not perfect, implies that a system which has been discretized by a fine mesh and a small value of X (i.e. a ruptured square almost opens) can be *approximated* by a coarser mesh and a larger value of X (i.e a larger fraction of the stress is retained by the ruptured square).

The effect of changing the lower bound θ_l for the uniform stress threshold distribution is shown on figure 3.6. The data collapses are very good for systems with random σ_{th} . For a fixed L , we note that a decrease in θ_l extends the power law to larger S . Comparing

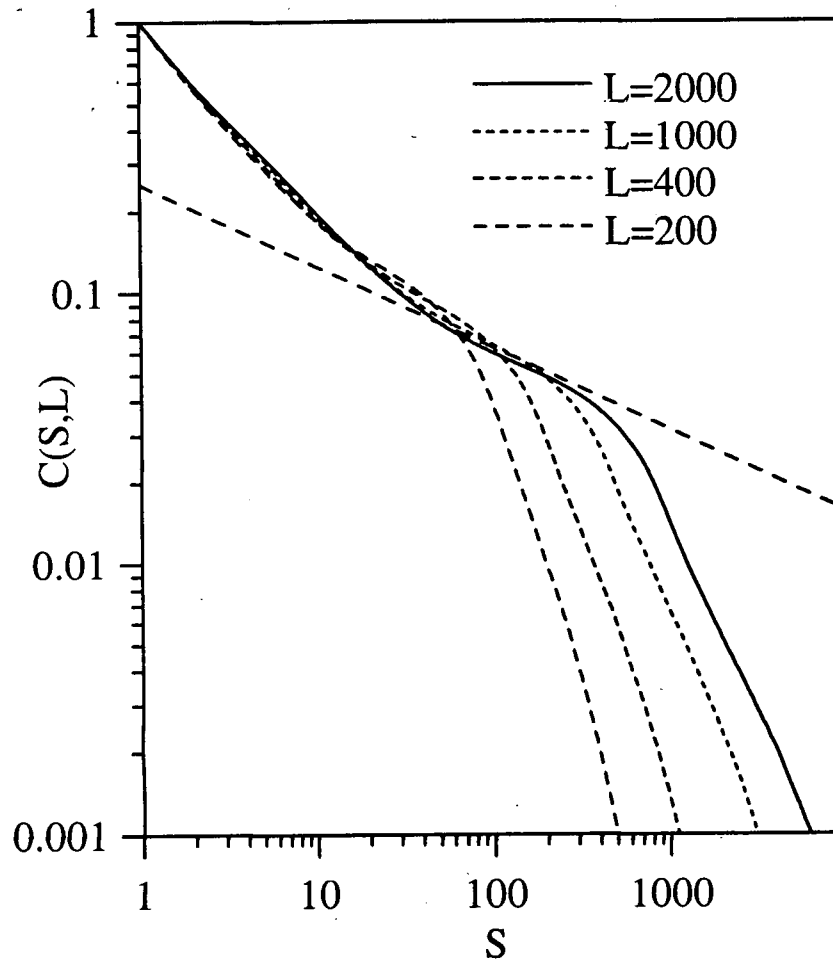


Figure 3.3: Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. The segment $S < 50$ is a crossover region and its origin is explained in section 3.4. Beyond $S > 50$, $C(S, L)$ has a linear segment with a slope close to 0.30. This linear segment gets longer as L increases but cuts off at large S . For comparison, we show a curve proportional to $S^{-0.30}$.

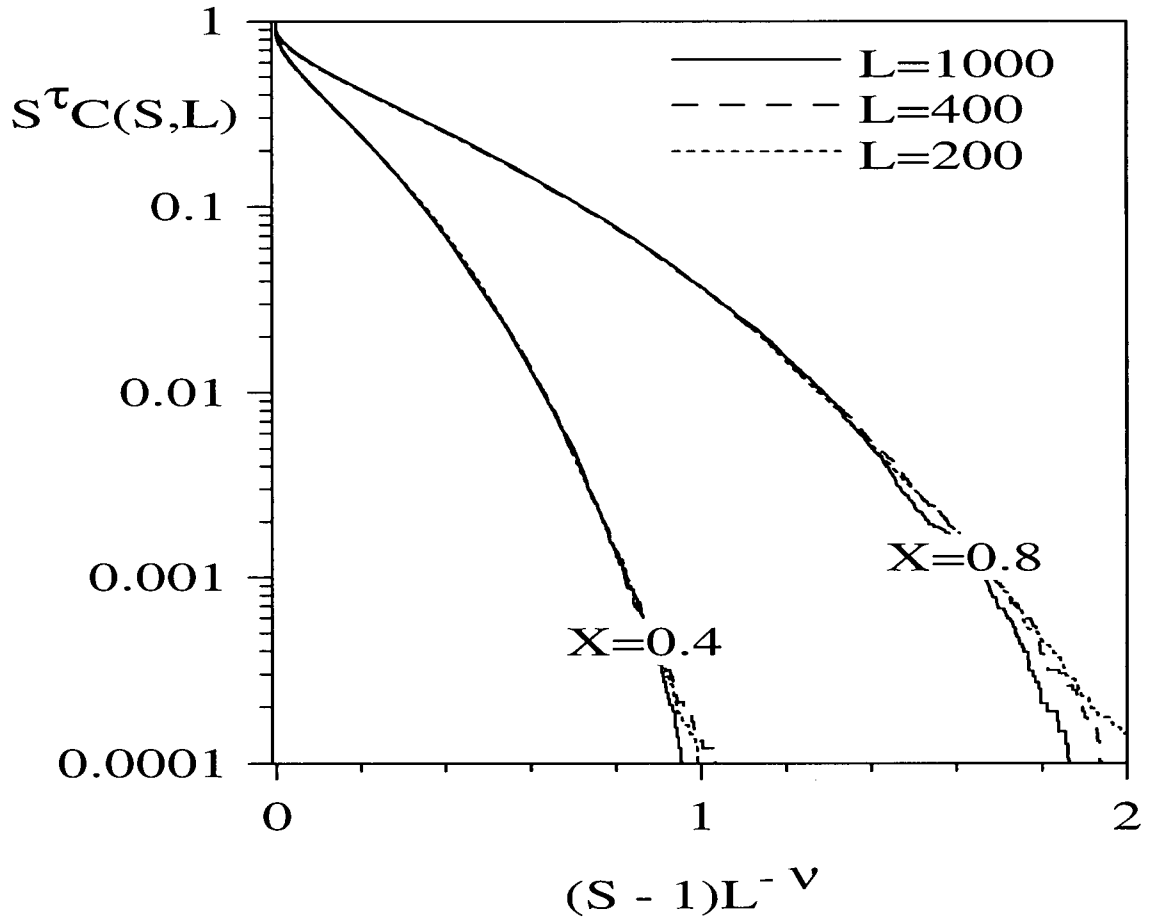


Figure 3.4: Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $\theta_l = 0.25$, $\alpha = 1$ and two values of X . Exponents used are $\tau = 0.30$ and $\nu = 1.38$.

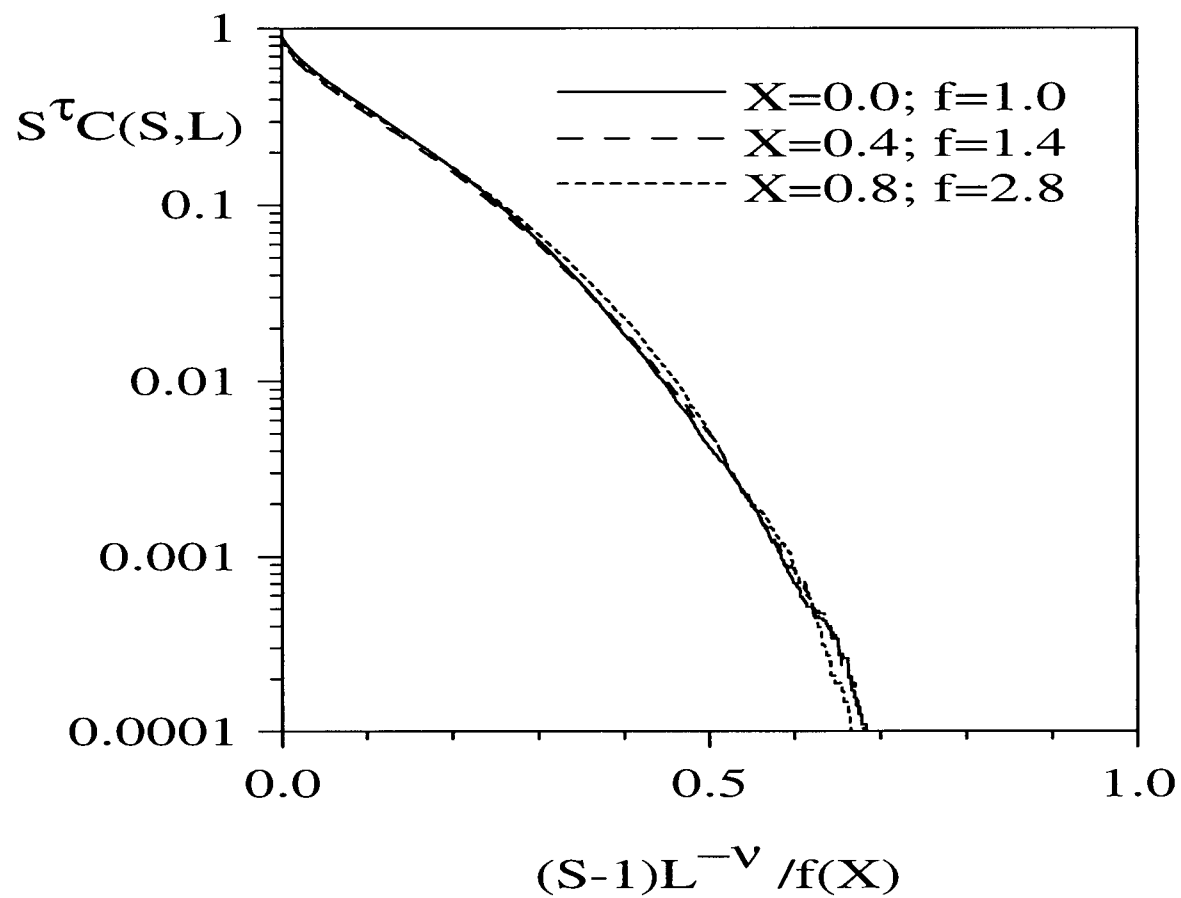


Figure 3.5: Log-linear scaling plot of the cumulative size-frequency distributions for a 1000×1 lattice, random σ_{th} , $\theta_l = 0.25$, $\alpha = 1$ and three values of X . Exponents used are $\tau = 0.30$ and $\nu = 1.38$.

figures 3.4 and 3.6, we see that, for a given L , an increase in X has much the same effect on the power law as a decrease in θ_l . We have also attempted a fit to the scaling form

$$C(S, L) = S^{-\tau} g[(S - 1)L^{-\nu}/f(\theta_l)].$$

The quality of the fit (see figure 3.7) is comparable to the one on figure 3.5. This result implies that a system which has been discretized by a fine mesh and a large value of θ_l , can be *approximated* by a coarser mesh using a smaller value of θ_l . We expect that this near equivalence can be made even better if we simultaneously adjust both X and θ_l .

We have also investigated the effects of changing parameters X and θ_l in systems with constant σ_{th} . Again, we fail to find a satisfactory scaling form which is valid for both small and large events since a crossover region at small S is still present. As we did above, we attempted data collapses for various X and θ_l by focusing only on the events outside the crossover region and found that with $\tau = 0.30$ and $\nu = 1.38$, the scaling function (3.1) works well. This shows that $\tau \approx 0.30$ is a universal value.

Before we consider the case of non-conservative systems ($\alpha < 1$), we comment on the value obtained for exponent τ . We have obtained through a finite-size-scaling analysis $\tau = 0.30 \pm 0.01$, which seems to be robust. This value can be compared with the one obtained in [68] (0.31 ± 0.04), which was obtained with a long-range scalar model with quenched disorder (stress thresholds are kept equal to their initial random values throughout a simulation). Our value of τ is in agreement with the latter, but is not in agreement with the value (0.4) obtained by Xu et al. [114] and Chen et al. [15] for a two-dimensional system. As we have checked for a $100 \times L_2$ system (results not shown), the discrepancy between our value of τ and theirs is significant since as L_2 increases, the cumulative size-frequency distribution settles onto a curve with a linear segment having a slope close to 0.4. This was observed for systems with both random σ_{th} and constant σ_{th} . (In some way, this discrepancy shows that if we would have studied the scaling properties

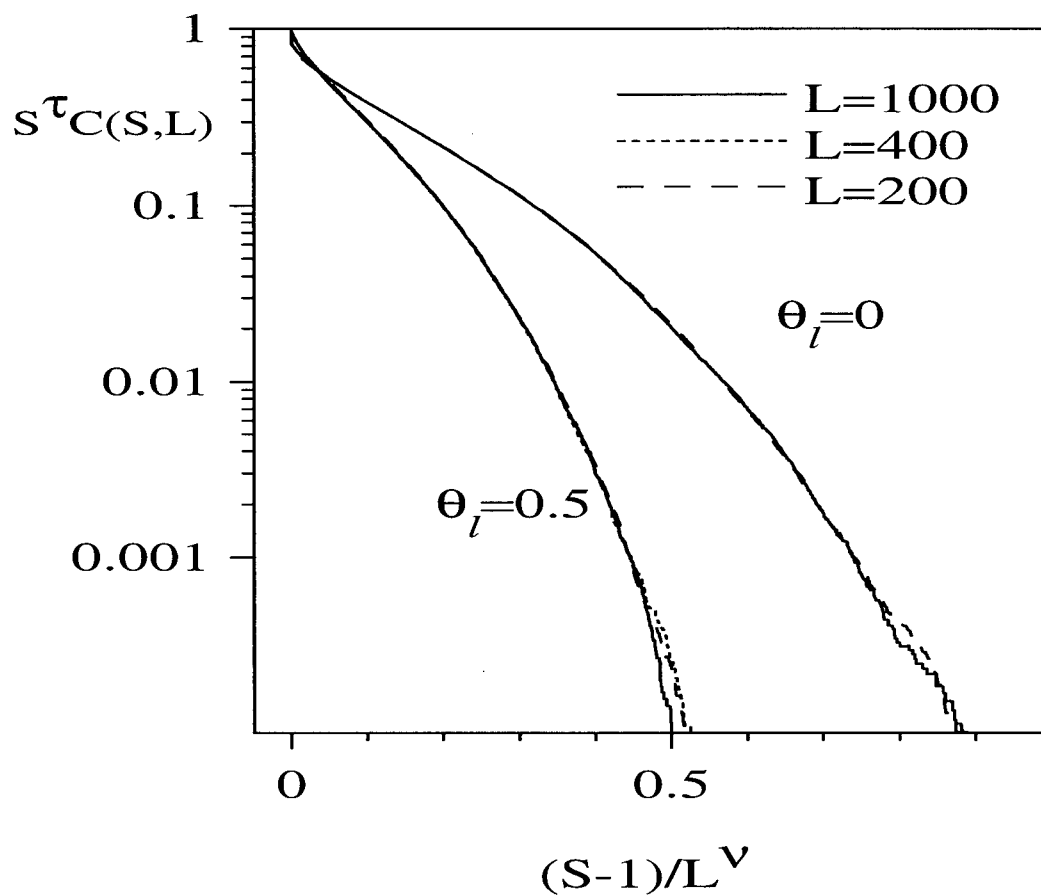


Figure 3.6: Log-linear scaling plots of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\alpha = 1$ and two values of θ_l . Exponents used are $\tau = 0.30$ and $\nu = 1.38$.

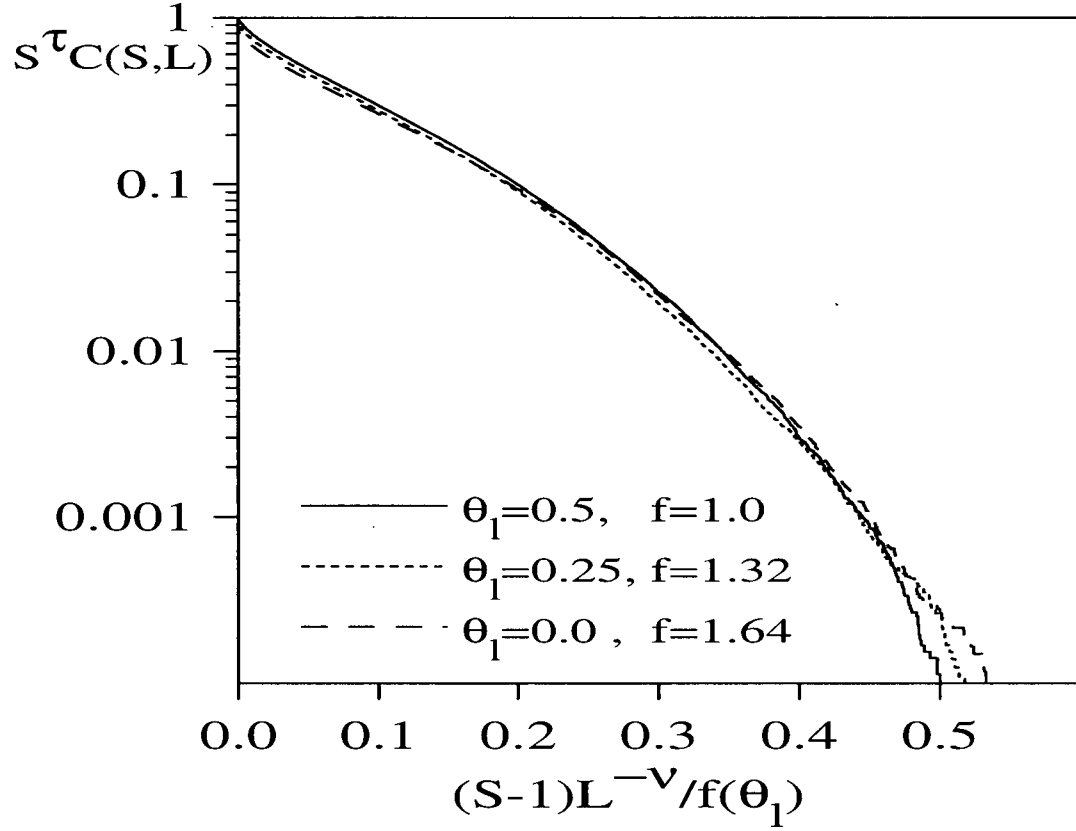


Figure 3.7: Log-linear scaling plot of the cumulative size-frequency distributions for a 1000×1 lattice, random σ_{th} , $X = 0$, $\alpha = 1$ and three values of θ_l . Exponents used are $\tau = 0.30$ and $\nu = 1.38$.

of isotropic lattices instead of one-strip lattices, then we would have obtained a different value of τ , but apart from that, qualitatively similar results). Finally, Christensen and Olami [17] obtained a different value (0.22 ± 0.05) through a finite-size-scaling analysis of the short-range, two-dimensional model with constant σ_{th} in [77]. Therefore, the models in [77] and [114] seem to belong to different universality classes. Some light is shed on this question in section 3.5.

3.3.2 Results in the non-conservative case ($\alpha < 1$)

In this subsection, we examine two effects for $L \times 1$ systems with random σ_{th} and constant σ_{th} when X is fixed to 0 and θ_l to 0.25. The first one is the effect of parameter α for a fixed L , and the second is the effect of L for a fixed $\alpha < 1$.

In figure 3.8, we present the cumulative size-frequency distributions for different values of the parameter α controlling the level of stress conservation for a 2000×1 system with random σ_{th} . A decrease in α makes the power law cut-off happen at increasingly smaller S . Also, as α decreases, the magnitude of the slope of the straight segment of the log-log plot increases, an effect which was also noticed in [17]. In figure 3.9, we show $C(S, L)$ for $\alpha = 0.94$, random σ_{th} and different lattice sizes (L). As expected, as L increases, the straight segment of the log-log plot extends to larger S . However, a saturation is noticed for $L > 2000$. We have fitted the cumulative size-frequency distribution for fixed L and α (see the dashed lines on figure 3.8) to

$$C(S) = S^{-\tau} \exp[(1 - S)/S_0 + bS^2 + cS^3 + \dots]. \quad (3.3)$$

For large L , we found that only the first term in the argument of the exponential is significant, where S_0 is a characteristic earthquake size. In this case, (3.3) reduces to a form similar to the one fitted in [58]. In table 3.1, we give the values of τ and S_0 for various values of α and L . For a fixed α , $S_0(L)$ increases with L and appears to saturate. It is

α	L	τ	S_0
0.94	4000	0.482	52.2
0.94	2000	0.477	50.0
0.94	1000	0.490	48.2
0.94	400	0.419	37.2
0.94	200	0.354	27.4
0.97	2000	0.428	115
0.97	1000	0.405	98.5
0.97	400	0.361	70.6
0.99	2000	0.386	407
0.99	1000	0.367	307
0.99	400	0.277	149

Table 3.1: Exponent τ and characteristic earthquake size S_0 for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and three values of α .

possible with our data that $S_0(L)$ diverges, but slower than L ($S_0(L) \sim L^\beta$ with $\beta \ll 1$) as suggested by János and Kertész [50]. The discussion in section 3.5 suggests, however, that this is unlikely. From table 3.1, we also note that, for a fixed L , S_0 increases with α and τ increases as α decreases. We therefore corroborate quantitatively conclusions obtained qualitatively from the curves on figure 3.8.

We have also studied the above effects for systems with constant σ_{th} (results not shown). Again, as in the conservative case, a crossover region is observed. Then, we found that a straight line can be fitted, for fixed L and α , to the log-log plot of $C(S, L)$ versus S for values of S outside the crossover region up to a cut-off value much smaller than in the conservative case. We noted that the slope of the straight segment increases and the cut-off value decreases as α decreases. Also, for a fixed α , we found that the cut-off value saturates as L increases. Because the cut-off value (for a fixed L) is smaller for $\alpha < 1$ than for $\alpha = 1$, we can again associate a characteristic earthquake size S_0 to

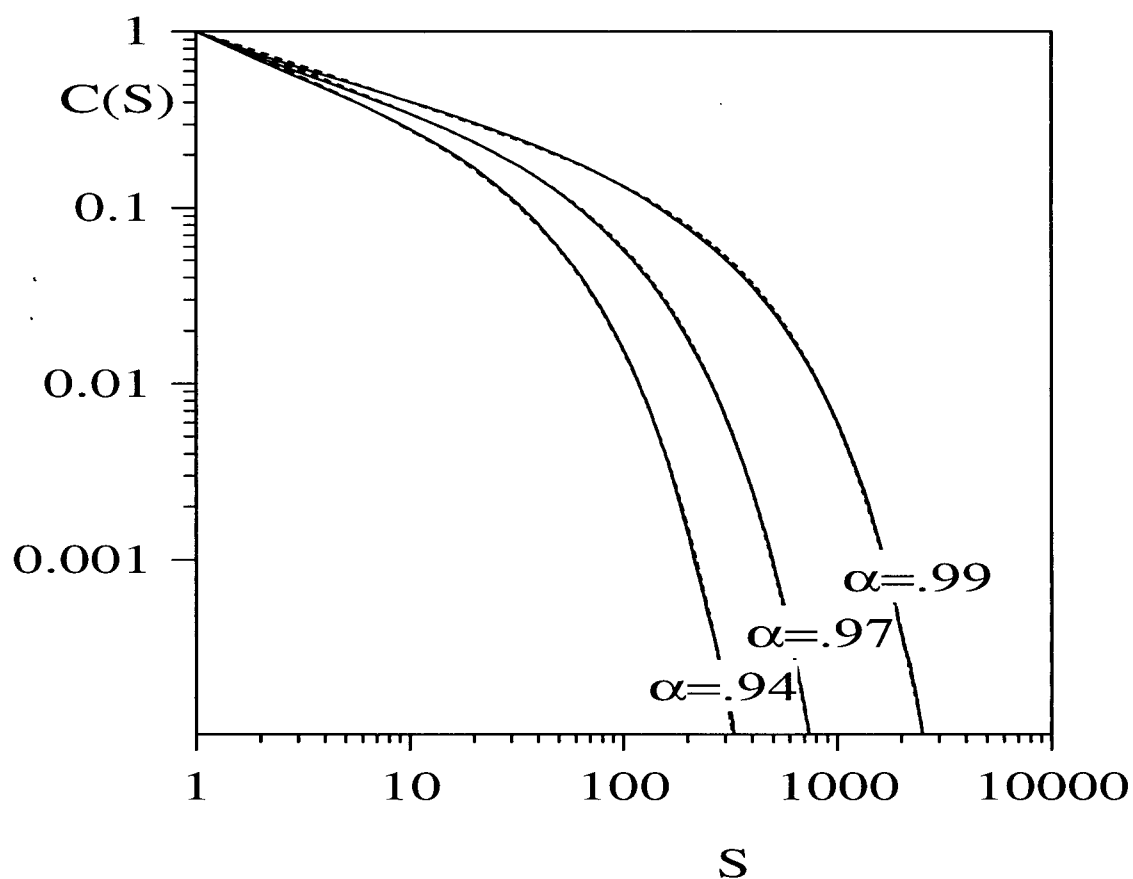


Figure 3.8: Log-log plot of the cumulative size-frequency distributions for a 2000×1 lattice, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and three values of α . The dashed lines are the fits to relation (3.3).

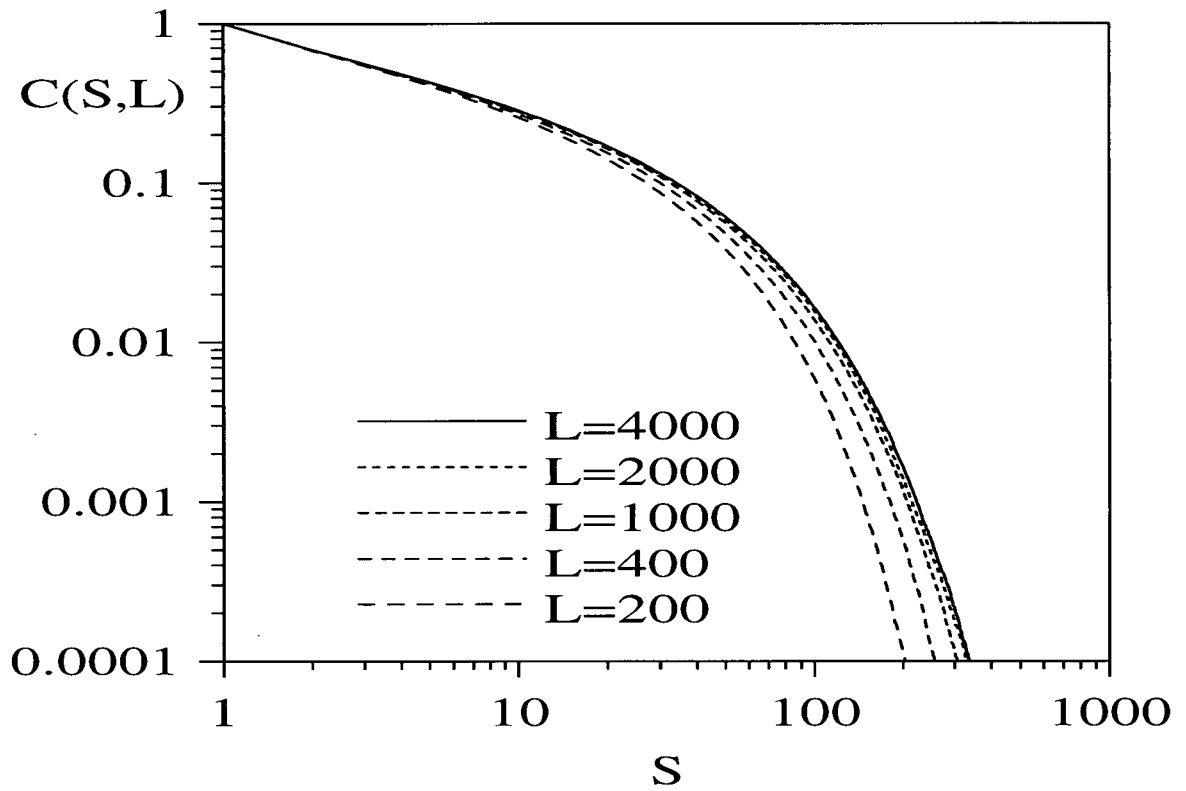


Figure 3.9: Log-log plot of the cumulative size-frequency distributions for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 0.94$.

$C(S, L)$.

Finally, it was observed that simulations running for longer times are providing data which are not altering our results. This is in contrast to a study of the short-range model in [77], which found that the time for total invasion of the interior of large lattices by SOC can be very large, especially for values of α close to 0.0 [67]. We believe that because the model of Xu et al. is long-range, the invasion is much faster than in the model in [77].

3.4 Stress Ordering in One-Dimensional Lattices

In the previous section, we found that for one-strip systems with constant σ_{th} , the cumulative size-frequency distribution $C(S, L)$ exhibits a crossover region for $S < 50$, which seems to be robust. We also discovered that for $S > 50$, up to a cut-off value scaling with L , the log-log plot of $C(S, L)$ versus S displays a straight segment (see figure 3.3). The goal of this section is to explain the presence of that crossover region. To achieve this, we study the stress ordering in one-strip systems, and in particular in one-strip systems with constant σ_{th} . To our knowledge, a study of the stress ordering in a system has only been performed by Grassberger [38] in the case of the short-range model in [77]. We begin by describing the stress ordering study and afterwards, attempt an explanation for the presence of the crossover region for systems with constant σ_{th} .

The stress ordering in a system can be either *local* or *global* [38]. To explain the presence of the crossover region, it is sufficient to study the local stress ordering. In fact, in the case of a one-strip system with constant σ_{th} , the study of both kinds of stress ordering is redundant since if some local stress ordering is found in a system, then automatically global stress ordering also prevails in the system. Local stress ordering is studied by considering the stress differences between nearest-neighbor elements in a system, i.e. $\delta_i \equiv |\sigma_{i+1} - \sigma_i|$ ($i = 1, 2, \dots, L - 1$ for a $L \times 1$ system). In comparison,

global stress ordering can be studied by focusing on the fitness, defined as $f_i \equiv \sigma_{th}^i - \sigma_i$, where $i = 1, 2, \dots, L$ for a $L \times 1$ system, σ_{th}^i being the stress threshold of the i th element. A histogram plot of the δ_i 's can be obtained and used to conclude whether local stress ordering exists in a system at any given time. The presence of peak(s) in the histogram ultimately proves that local stress ordering prevails in the system.

We begin our exploration by considering a system with constant σ_{th} . On figure 3.10, we plotted the δ_i 's for a 1000×1 system. A series of patches of various sizes are observed at this particular time (arbitrarily defined as t_0 , where here *time* refers to the long time scale) for small δ_i . In particular, the two biggest patches on figure 3.10 have sizes 61 (at $i \approx 800$) and 46 (at $i \approx 600$), which means that the number of elements at about the same stress level is 62 and 47 in these two cases respectively. Assemblies of elements next to one another, which are at about the same stress level, will be referred to as *GROUPS* of elements below. As long as one element in a group has its stress over σ_{th} , then the other elements in the group will *likely* rupture in the same sequence. We have checked that the groups of elements are not static features, i.e. they do not involve the same elements at later times. A histogram of the data of figure 3.10 was obtained and the result is plotted on figure 3.11 (solid line). The other curve on figure 3.11 was obtained with data taken at time $t_1 > t_0$. The histograms at both times display a peak at $\delta_i \approx 0$, showing that local stress ordering prevails in the one-strip system with constant σ_{th} . We have also checked that local stress ordering is maintained at any time by varying parameters X and θ_l (results not shown). The quantitative difference a larger value of X (or a smaller value of θ_l) makes, as compared to $X = 0$ (or $\theta_l = 0.25$), is to restrict the δ_i 's to be smaller numbers. For instance, we have checked that for $X = 0.8$, the δ_i 's cut off approximately at 0.2 (as compared to about 1.0 for $X = 0$ on figure 3.11).

At this point, we attempt an explanation for the presence of the crossover region observed for systems with constant σ_{th} (see figure 3.3): during an earthquake sequence,

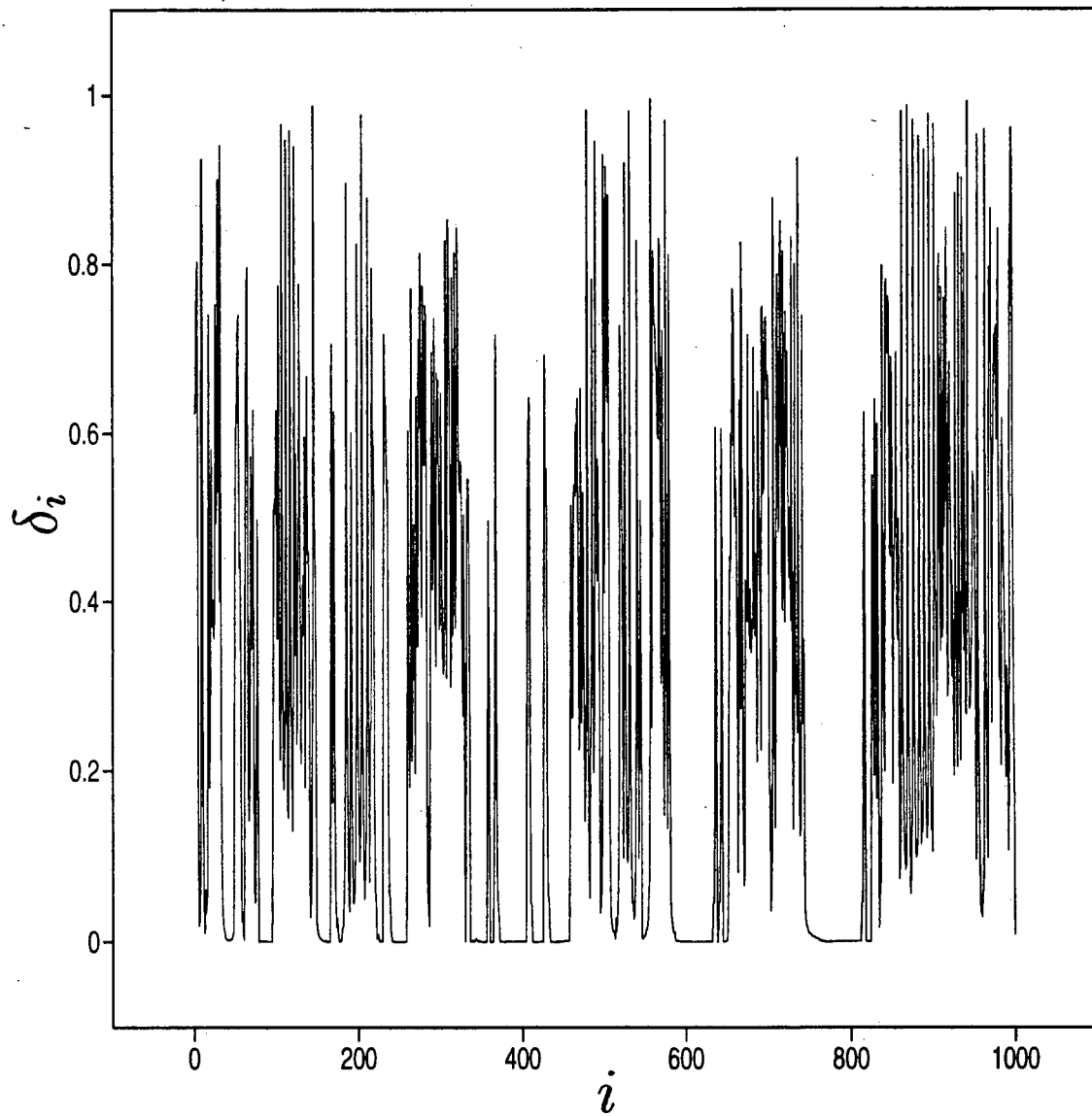


Figure 3.10: Distribution of the stress differences δ_i ($i = 1, 2, \dots, 999$) between the nearest-neighbor elements of a 1000×1 system with constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$.

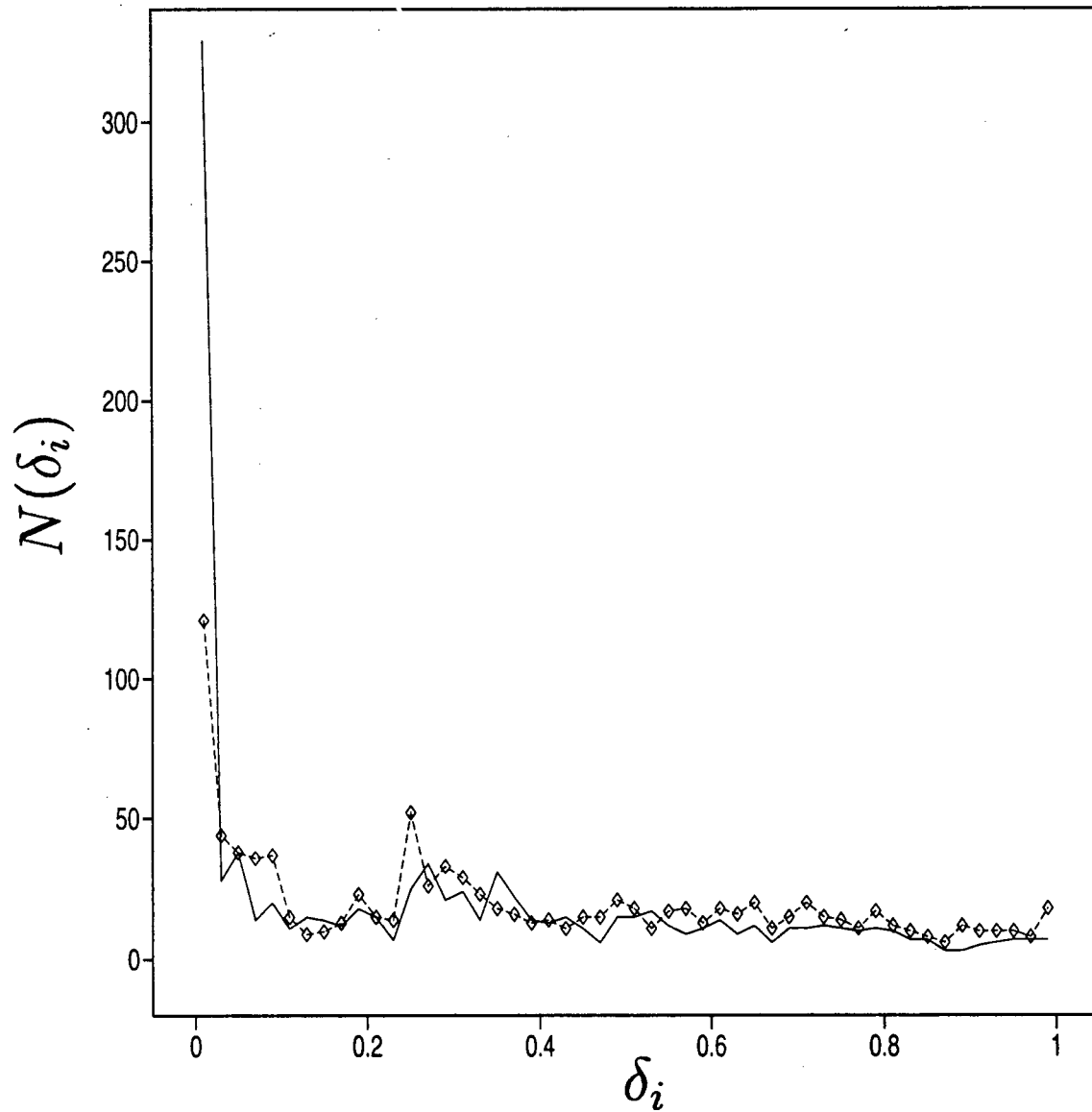


Figure 3.11: Histogram plots of the stress differences δ_i between the nearest-neighbor elements of a 1000×1 system with constant σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$ at two distinct times. The solid line curve was obtained with data taken at the arbitrary time t_0 (data of figure 3.10) and the curve plotted using a dashed line and lozenges was obtained with data taken at time $t_1 > t_0$. Fifty bins were used to get the histograms.

when the ruptured elements are found mostly in a given group of elements or in a couple of small groups of elements away from other groups in the system, then the size of the sequence (S) is limited by the number of elements in the group(s), so S tends to be small (i.e. in the crossover region). On the contrary, if the ruptured elements are found *mostly* in groups of appreciable sizes close to one another, then S tends to be large. In other words, the localized (versus delocalized) nature of the activity in a sequence determines whether S is small or large. We believe that this explanation is plausible considering the fact that for a 1000×2 system with constant σ_{th} , no crossover region is observed (results not shown) since the presence of a second strip of elements allows the delocalization of the activity to take place. Along the same line, we have noticed that systems with random σ_{th} never display patches (and in a more general sense, local stress ordering) such as the ones on figure 3.10 (results not shown), so the localization of the activity never happens in these systems. This explains why systems with random σ_{th} always display simple cumulative size-frequency distributions (as on figure 3.1 for instance).

We now turn to a simplified version of the original model of Xu et al. by considering a mean-field version of it.

3.5 Mean-Field Version of the Model of Xu et al.

We present in this section a mean-field version of the model of Xu et al. in which all the elements are *coupled equally* with each other. As a result, it is not necessary to use a Green function and to refer to a lattice structure anymore. The additional stresses (σ') caused by the rupture of a given square are now all expressed as

$$\sigma' = \frac{(1 - X)\sigma^{old}}{L - 1}, \quad (3.4)$$

where L is the number of squares in the system and the other symbols have the same meaning as in the case of relation (2.8). Therefore, the rupture of a square is followed by

a stress *decrease* $[(1 - X)\sigma^{old}]$ on this element and a stress *increase* $[(1 - X)\sigma^{old}/(L - 1)]$ on all the other $L - 1$ elements, which leads to the same total stress value before and after the rupture. Such a situation leads to a total stress growing without limit in time. To prevent this from happening, we can multiply σ' in (3.4) by a constant α between 0.0 and 1.0. We use the same symbol as in subsection 3.3.2, but it is important to keep in mind that in the mean-field model, nonconservation of stress is *essential* for the model to describe earthquakes. In the mean-field model, α should be adjusted to values close to 1 in order to not have earthquakes releasing *artificially* too much stress from the system. We will say more about this below.

In section 3.3, we presented results obtained with a version of the original model of Xu et al. Systems with random σ_{th} and constant σ_{th} were both considered. In the case of the mean-field model, systems with constant σ_{th} lead, in the stationary state, to trivial size-frequency distributions, since groups of elements which ruptured simultaneously in the past, will rupture simultaneously in the future as well. Indeed, for a given such system, depending on the number of elements (N_e) in each group, the resulting size-frequency distribution consists of a series of peaks, one for each distinct value of N_e . Therefore, the mean-field model is NOT self-organized critical when disorder in the stress threshold distribution is absent, since no power law can be obtained for the size-frequency distribution. In systems like this, it remains the possibility of studying the synchronization in the system, but we will not do such a study here (see however [69, 7] for a discussion of mean-field models of integrate-and-fire oscillators). The study of systems with random σ_{th} with the mean-field model is non-trivial.

The study of systems with random σ_{th} with the mean-field model can be done by implementing the algorithm of section 2.4, except for steps 1 and 2. Only resetting rule I (see step 7) of the algorithm is meaningful and so is the only one we use. To do step 8 of the algorithm of section 2.4, i.e. to calculate the stress redistribution after the rupture

of an element, relation (3.4) is used rather than (2.8). Again, we let the model reach a stationary state by discarding the first 10^5 earthquake sequences generated and include 10^6 sequences to obtain the cumulative size-frequency distribution $C(S, L)$ for a given number of elements (L) in the system. Now, if we would set α to a given value and plot $C(S, L)$ for a few values of L , then we would obtain a graph similar to the one on figure 3.9, with no data collapse of the distributions for different values of L . Because we want to attempt a data collapse of the distributions for different L , we use an alternate route: for a given L , α will be set equal to α_{eff} , the *effective value* of α for a $L \times 1$ system (when $\alpha = 1$) studied with the original model of Xu et al. We plotted on figure 3.12 $\log(1 - \alpha_{eff})$ versus $\log(L)$ for $L \in [100, 1000]$, where α_{eff} was obtained by averaging over 2000 earthquake sequences generated by the original model of Xu et al. with $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$. From figure 3.12, we see that $\log(1 - \alpha_{eff})$ decreases about linearly with $\log(L)$. The slope of the straight line was found to be -1.02 . The fact that $\alpha_{eff} \neq \alpha$ even if these two quantities are close together results from the long-range nature of the model of Xu et al.: there is always a fraction of the redistributed stress from a ruptured element which is lost at the boundaries of the system. This fraction goes to zero as the size of the system goes to infinity (i.e. $L \rightarrow \infty$). Now, using for α in the mean-field model the value of α_{eff} on figure 3.12 for the appropriate L , we did simulations of the mean-field model for $L = 200$, $L = 400$ and $L = 1000$. A fit of the cumulative size-frequency distributions to the finite-size-scaling function (3.1) was attempted with the values $\tau = 0.3$ and $\nu = 1.38$, but the data collapse was found to be bad (results not shown). For this reason, we simply plotted on figure 3.13 $C(S, L)$ versus S without finite-size scaling. It is obvious from the figure that even if $C(S, L)$ extends to larger S as L increases, it is not possible to get a satisfying data collapse for both small and large S . As a result, the exponent τ seems to be sensitive to the details of the model, i.e. in this case the fact that the elements are equally coupled with each other and not

arranged in a lattice as in the original model of Xu et al. It is also possible that the trick we have used above to fix α for a given L in the mean-field model does not allow to get optimal results, which could explain why no satisfying data collapse was obtained. It is relevant to mention that α_{eff} obtained from the original model of Xu et al. is not unique for a given L in the mean-field model since it depends on how the squares are joined together in the original model. For instance, α_{eff} obtained for $L = 400$ is not the same if the lattice configuration is a strip (400×1) rather than a two-dimensional 20×20 arrangement.

This study of the mean-field model was just a first attempt at trying to investigate the effect of changing the rules of the model. Much work along this line remains to be done.

3.6 Summary of the Important Results and Conclusions

In this chapter, we have studied the scaling properties of the original model of Xu et al. as well as a mean-field version of it by means of a finite-size-scaling analysis. In the mean-field model, the elements are equally coupled with each other. In the original model, one-dimensional ($L \times 1$) lattices were considered for convenience. We summarize the results obtained in this chapter and draw our conclusions.

For the original model, the finite-size-scaling function (3.1) works well over the whole range of S for systems with random σ_{th} and for $S > 50$ for systems with constant σ_{th} . For systems with constant σ_{th} , there is a crossover region for $S < 50$ and it appears whatever the parameter values in the model. The existence of such a crossover region is related to the localized nature of the activity on a group of elements or on a couple of small groups of elements at about the same stress level (and therefore likely to rupture together as long as one of them will do so) isolated from any other group in the system.

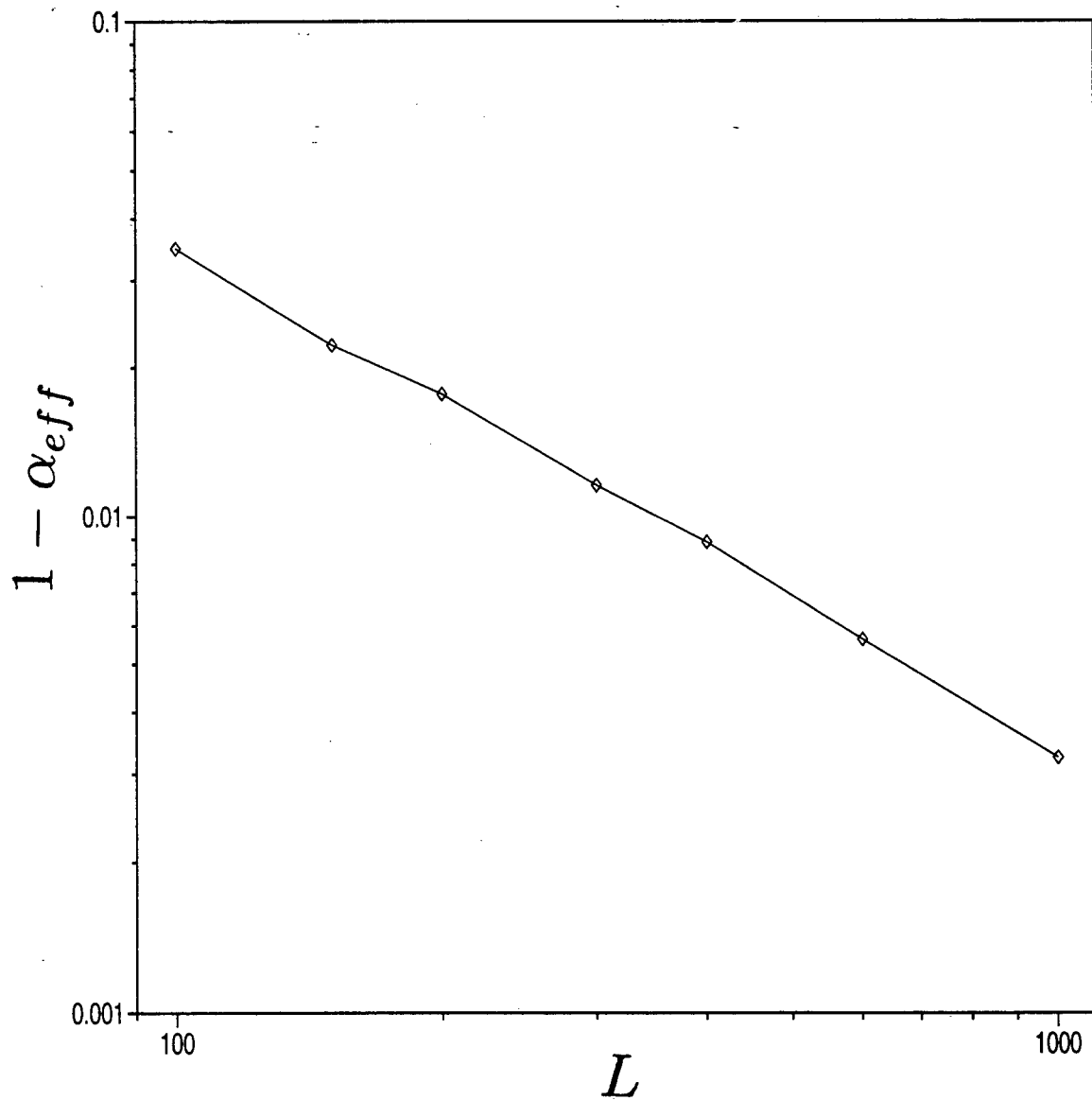


Figure 3.12: Log-log plot of $1 - \alpha_{eff}$ values versus L for $L \times 1$ lattices, random σ_{th} , $X = 0$, $\theta_l = 0.25$ and $\alpha = 1$ obtained with the original model of Xu et al. The slope of the straight segment was found to be -1.02 .

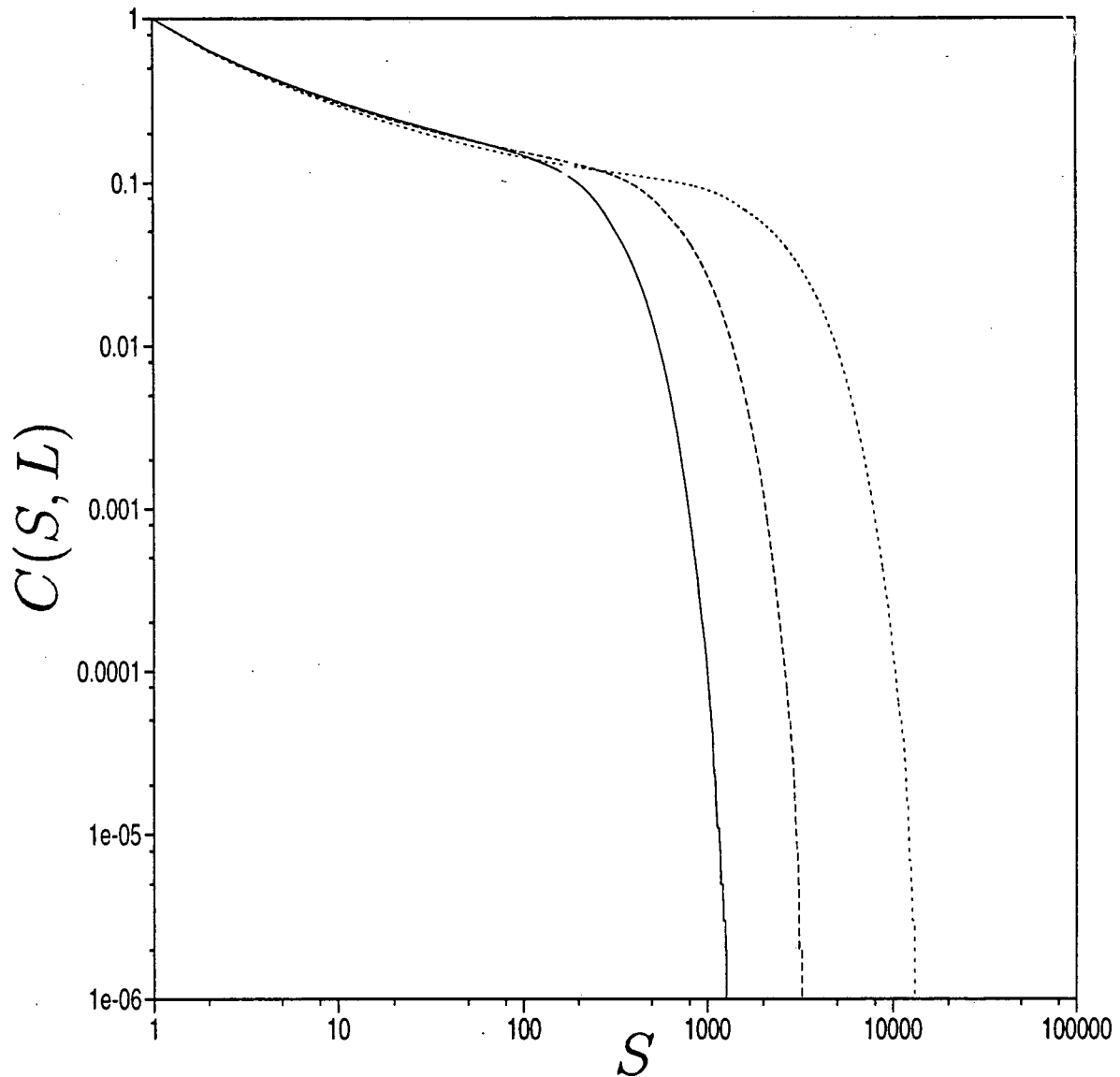


Figure 3.13: Log-log plot of the cumulative size-frequency distributions for systems of L elements equally coupled with each other. The values of the parameters are $X = 0$ and the stress thresholds (i.e. the σ_{th}) are taken from a uniform distribution with $\theta_l = 0.25$. From left to right, the distributions refer to systems having $L = 200$, $L = 400$ and $L = 1000$. The values of α differ for all three distributions (see how they were adjusted in the text).

The crossover region disappears for two-dimensional systems with constant σ_{th} as well as for systems with random σ_{th} .

The effects of parameters X (controlling the degree of stress retained by a ruptured square) and θ_l (controlling the width of the uniform distribution of stress thresholds) were also investigated and we found a fair agreement with the finite-size-scaling function (3.2) over the whole range of S for systems with random σ_{th} and for S larger than about 50 (outside the crossover region) for systems with constant σ_{th} . This result implies that a system which has been discretized by a fine mesh and a small value of X (or a large value of θ_l) can be *approximated* by a coarser mesh using a larger value of X (or a smaller value of θ_l).

The exponents of the finite-size-scaling functions (3.1) and (3.2) were found to be $\tau = 0.30 \pm 0.01$ and $\nu \approx 1.38$. This value of τ is robust and in particular, it is independent of the degree of disorder in the system (i.e. it is the same for systems with random σ_{th} and constant σ_{th}). However, we have checked that for systems with more than one strip, there is a crossover of the exponent τ to $\tau \approx 0.4$. Therefore, the results in this chapter are quantitatively specific to one-strip systems.

Introduction of stress nonconservation (controlled by parameter α) results in the appearance of a characteristic earthquake size (S_0), which is smaller than L for large L . For a fixed L , S_0 increases with α and the exponent τ increases as α decreases. The latter observation is in agreement with the findings of Christensen and Olami [17]. These results were found to be independent of the degree of disorder in the system.

For the mean-field model, only the study of systems with random σ_{th} is non-trivial. In addition, this model necessitates stress non-conservation in order to apply to earthquakes. We have used a trick to fix parameter α for a given number of elements (L) in the mean-field model. When such a trick is used, we get a rather poor fit to the finite-size-scaling function (3.1) when $\tau = 0.30$ and $\nu = 1.38$ are used. We conjectured that this could be

due to the sensitivity of the exponent τ to the details of the model. However, it is also possible that the trick that we have used to fix α does not work well.

The general conclusions that we can draw from our study of the scaling properties are that the model of Xu et al. displays SOC in the conservative case but not in the non-conservative case. In the conservative case, this was proven by the fact that $C(S, L)$ is a power law which cuts off due to finite-size effects. On the other hand, in the non-conservative case, a characteristic earthquake size smaller than the system size appears. In such a case, $C(S, L)$ is best fitted to a function of the type (3.3) and the model is NOT self-organized critical.

Chapter 4

Time Series Analyses

In this chapter, we implement two types of analysis of time series generated by the original model of Xu et al. These time series are presented in section 4.1. In section 4.2, we perform a nonlinear forecasting analysis. This analysis investigates whether the data in a time series generated from a dynamical system exhibit low-dimensional chaotic behavior, as opposed to high-dimensional (or stochastic) behavior. The second time series analysis consists in doing a rescaled range analysis, which allows computing the Hurst exponent (see section 4.3). This exponent enables one to learn about the time correlations (or memory effects) in a system. Finally, in section 4.4, we summarize the results obtained in this chapter and draw our conclusions.

4.1 Time Series of Interest

At least two types of time series can be generated by the model of Xu et al. The first type involves time series consisting in the size of the earthquake sequences (S) at different values of the long time (t): $\{S_i\}$ ($i = 1, 2, \dots, N$). A typical time series of this type is plotted on figure 4.1 for a 100×100 system with random σ_{th} . A time series looking qualitatively similar is obtained for a two-dimensional system with constant σ_{th} as well as for one-dimensional systems (see chapter 3). (Most of the work in this chapter will be performed on time series for two-dimensional systems, but for completeness, we will sometimes analyze time series for one-dimensional systems with constant σ_{th} since we saw in chapter 3 that stress ordering prevails in these systems. Also, unless otherwise

specified, the values of the parameters of the model of Xu et al. will always be assumed to be $\alpha = 1$, $X = 0$ and $\theta_l = 0.25$). From figure 4.1, we note that S_i fluctuates greatly over time and that the sequences are instantaneous events on the long time scale. Also, even if it is not completely clear from the figure, the time intervals between consecutive earthquake sequences vary over several orders of magnitude. Therefore, a time series such as the one plotted on figure 4.1 is highly variable in both its dependent variable (S_i) and independent variable (t_i). At this point, it is desirable to fix one of these variables and allow the other one to vary. This can be done according to one of the three following procedures:

- (a) focus only on the earthquake sequences having $S = S_0$ or $S \geq S_0$. Then, the time intervals between these consecutive sequences constitute a time series of interest,
- (b) focus on all the earthquake sequences whatever S . Then, the time intervals between these consecutive sequences constitute another time series of interest,
- (c) sample the original time series every Δt time units. Now, because there is in general not a sequence happening at a particular sampling time, the sequence which is closest in time to the latter will be treated as having happened at that sampling time. This sequence will have its size retained as a datum for a newly generated time series consisting in the size of earthquake sequences *about* equally sampled in time. (For completeness, we mention that another related procedure that takes advantage of the fact that the earthquake sequences are point events could be used: sum the size of all the earthquakes occurring in a time window of a given size. The time corresponding to the midpoint of the window, for instance, can then have this sum of earthquake sizes assigned to it. The resulting time series is also evenly sampled. This procedure is appealing but we will not consider it in this thesis).

Consequently, the time series of type 1 exist in three varieties, defined as varieties (a), (b) and (c). An example of a time series of variety (b) is plotted on figure 4.2. It can be obtained from figure 4.1 by calculating the time intervals between the consecutive earthquake sequences. As mentioned above, it is obvious from figure 4.2 that these time intervals vary over several orders of magnitude. Now, regarding the time series of variety (a), their structure depends on the value of the parameter S_0 . When S_0 is small, the time intervals between the consecutive sequences are rather small whereas for S_0 large, the time intervals are larger (results not shown). To this effect, we will try to confirm by a different technique the results obtained by Chen et al. [16] with the model of Xu et al.: the large earthquakes are temporally correlated but the small earthquakes are not. Finally, for the time series of variety (c), we noted that parameter Δt has little effect on their structure (time series not shown).

The second type of time series (hereafter called type 2) which can be obtained from the model of Xu et al. involves the average stress in the system ($\bar{\sigma}$) as a function of the long time (t), where $\bar{\sigma}$ is defined as

$$\bar{\sigma} = \left(\frac{1}{N_s} \right) \sum_{i=1}^{N_s} \sigma_i,$$

N_s being the number of elements in the system and σ_i the stress on element i . Here, we have chosen not to take the absolute value of σ_i in calculating $\bar{\sigma}$. As we have checked, this choice has no effect on the time series for one-dimensional systems since the σ_i 's are all positive. On the other hand, for two-dimensional systems, σ_i is sometimes negative, but we have checked that this occurs for about 1% of the elements only and so, the effect on $\bar{\sigma}$ is minor. A time series of this type for a 100×100 system with random σ_{th} is plotted on figure 4.3. A time series looking qualitatively similar is obtained for a two-dimensional system with constant σ_{th} as well as for one-dimensional systems. From figure 4.3, we note that a well-defined pattern appears: diagonal lines are connected to

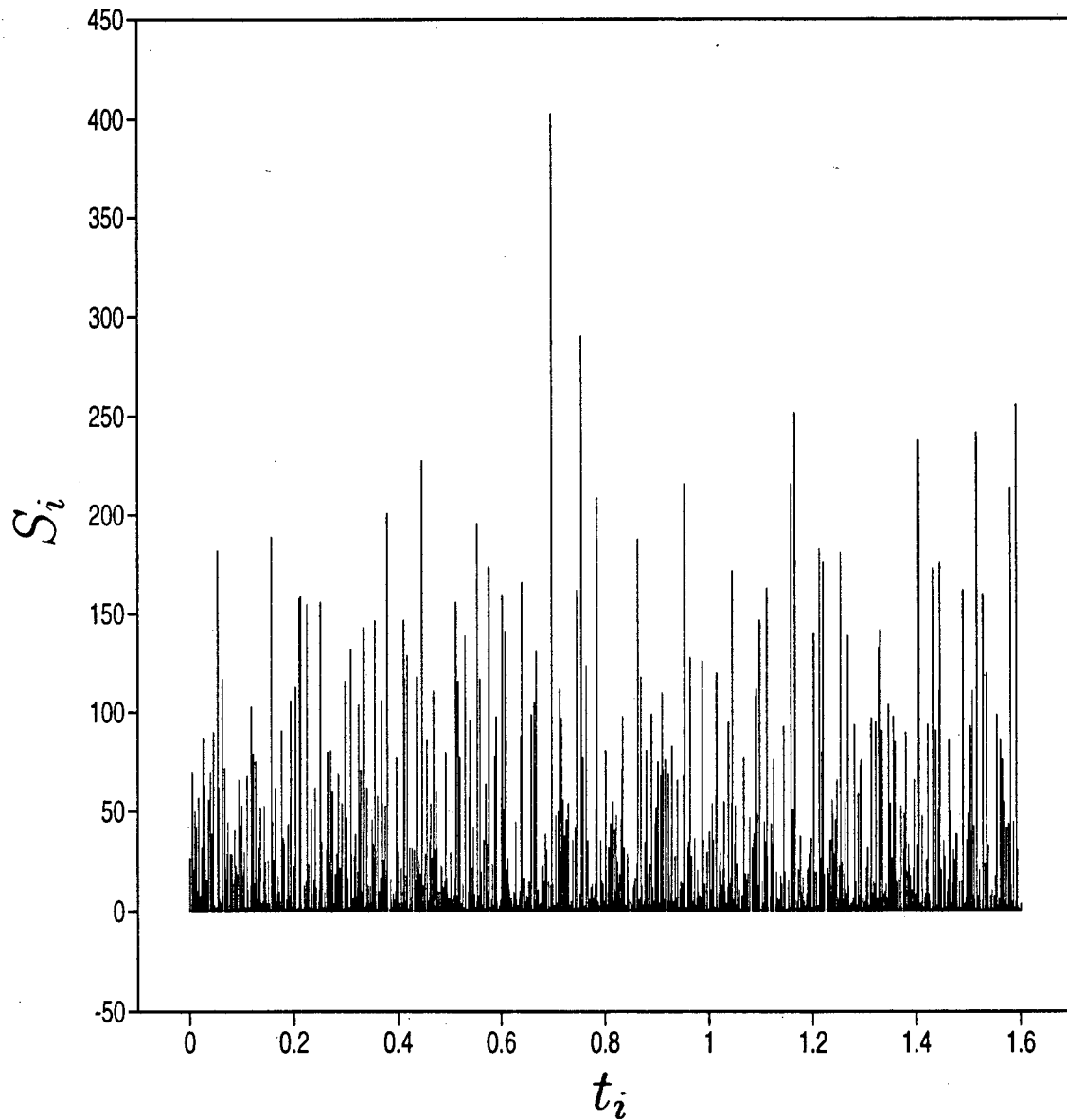


Figure 4.1: Typical time series of type 1: Size of an earthquake sequences (S_i) as a function of the occurrence time (t_i). The whole time series has 5×10^4 data but for clarity, only the first 2000 are plotted. The sequences were generated from a 100×100 system with random σ_{th} . The time at which the first sequence happens is arbitrarily set to $t_1 = 0$.

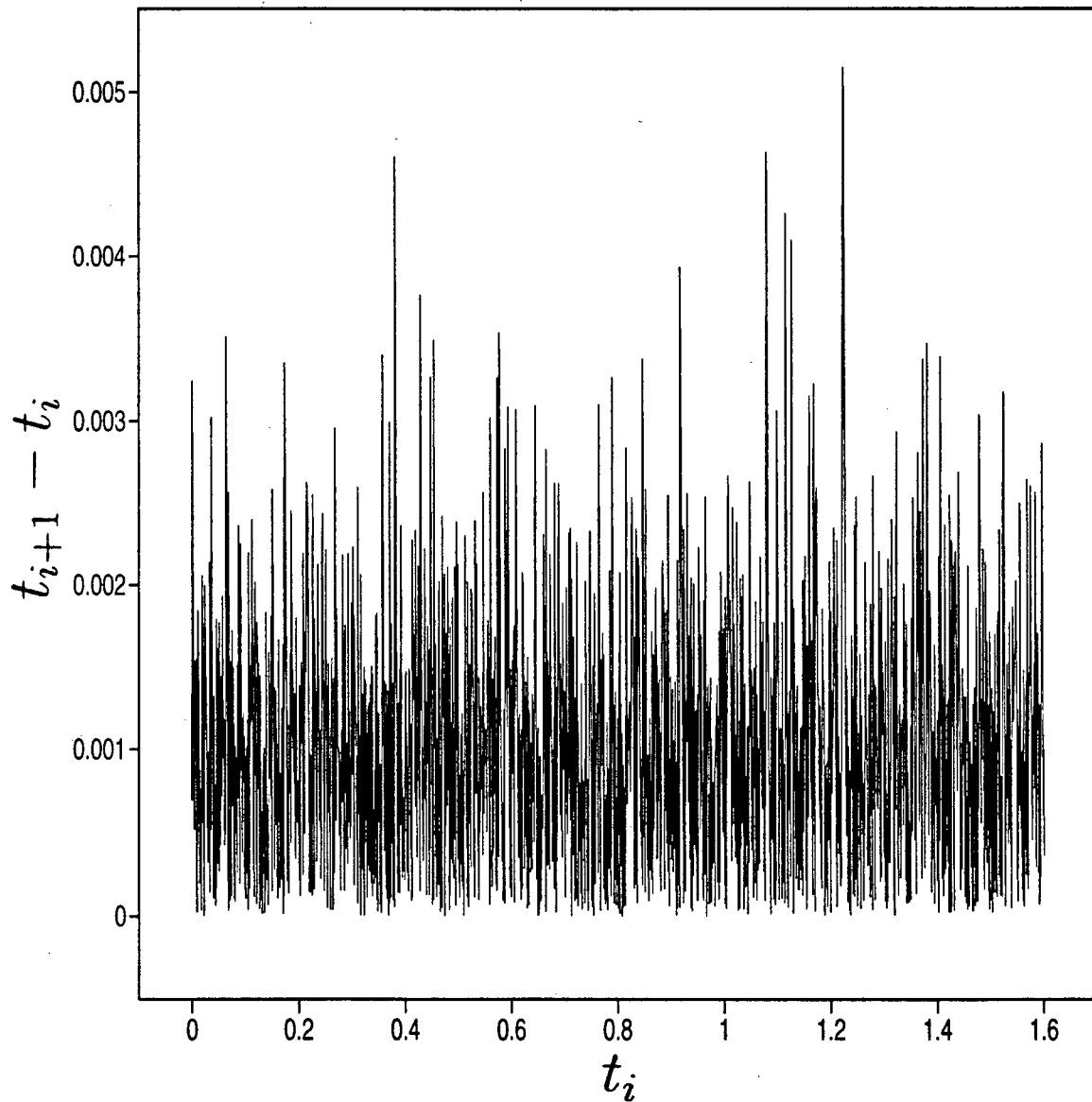


Figure 4.2: Typical time series of type 1 [variety (b)]: Time intervals $t_i - t_{i-1}$ ($i = 2, 3, \dots, 2000$) between consecutive earthquake sequences. The sequences were generated from a 100×100 system with random σ_{th} (see figure 4.1).

vertical lines, which are themselves connected to diagonal lines, and so on. The diagonal lines symbolize average stress increases due to system driving. The slope of the diagonal lines (i.e. the rate of stress increase) is symbolized by p and was set to 1 for convenience (see section 2.4). The vertical lines represent average stress decreases due to earthquake occurrences and give a hint about the magnitude of the earthquakes. A time series such as the one on figure 4.3 is unevenly sampled. We will also consider below evenly sampled versions of the time series of type 2 by sampling every Δt time units the original time series of type 2. This is possible since, as explained in section 2.4, we assume that $\bar{\sigma}$ increases linearly with time between two earthquake sequences.

As a preliminary study, we can analyze the time series plotted partially on figures 4.2 and 4.3 by considering the linear correlation coefficient [85]

$$r_l = \frac{\sum_i (x_i - \bar{x})(x_{i+l} - \bar{x}')}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (x_{i+l} - \bar{x}')^2}},$$

where l is called the lag and \bar{x} and \bar{x}' are the mean of the time series $\{x_i\}$ and $\{x_{i+l}\}$ ($i = 1, 2, \dots, N$) respectively. As suggested in [54], r_l is reliable only for $l \leq N/4$ and thus, we do not calculate r_l for $l > N/4$. The set of values r_l ($l = 1, 2, \dots, N/4$) and the plot of r_l against l is referred to as the AutoCorrelation Function (ACF). The ACF for the time series plotted on figure 4.2 and the time series plotted partially on figure 4.3 are displayed on figures 4.4 and 4.5 respectively. The ACF for the time series of type 1 ($N = 1999$) exhibits oscillations around 0, and r_l remains small for all l except $l = 0$ ($r_0 = 1$ is not shown on figure 4.4). This seems to reveal that the data in this time series are not correlated. On the other hand, the ACF for the time series of type 2 ($N = 2000$) decreases with l (for $l < 100$) and fluctuates around 0 at larger l . It is of interest to note that the bump of the ACF between $l \approx 160$ and $l \approx 220$ (see figure 4.5) is a feature reflecting the presence of non-trivial correlations in the time series of type 2. However,

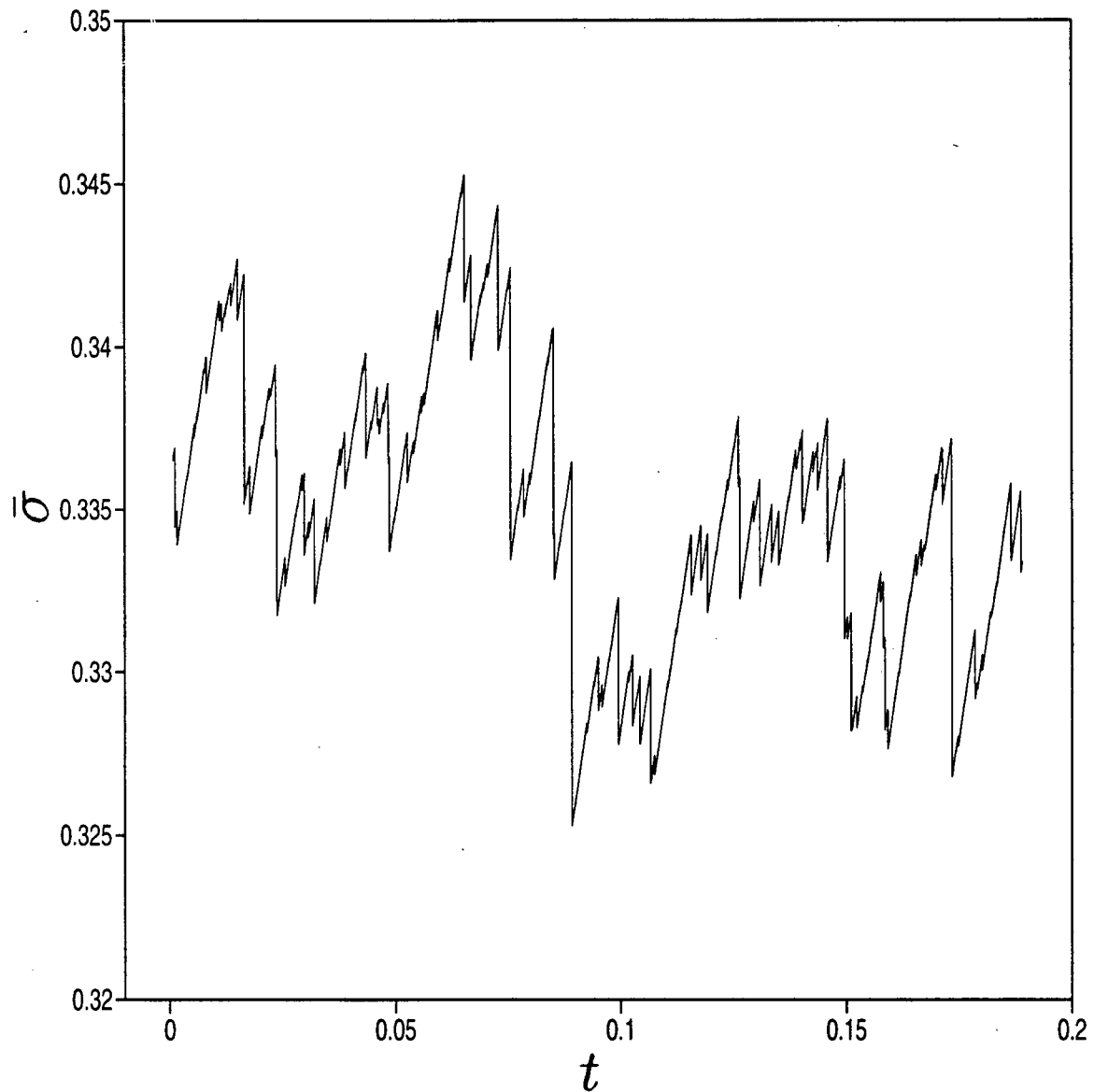


Figure 4.3: Typical time series of type 2: Average stress ($\bar{\sigma}$) as a function of the long time (t). The data were generated from a 100×100 system with random σ_{th} . The whole time series has 5×10^4 data but for clarity, only the first 500 data (i.e. the end points of the vertical lines) are plotted. The time of the first datum is arbitrarily set to zero.

the ACF is not powerful enough to reveal exactly what type of correlations are present in the latter time series. More sophisticated analyses are required to do so. Two of them are considered in this chapter and are outlined below.

In summary, we consider in this chapter two types of time series. Note that only the time series of type 2 are explicitly linked to the state variable (i.e. the stress) of the model. Time series consisting in the value of the stress at a given location in the system as a function of the long time could have also been considered but, as we have checked, they exhibit large fluctuations and depend on the location chosen to *measure* the stress, so they are not adequate for a time series analysis.

4.2 Nonlinear Forecasting Analysis

In this section, we implement the first of two types of time series analysis, namely a nonlinear forecasting analysis. We begin with a general introductory discussion in subsection 4.2.1. Afterwards, we describe the two important ingredients entering into a nonlinear forecasting algorithm, namely the state space reconstruction (subsection 4.2.2) and the nonlinear function approximation (subsection 4.2.3). The forecasting algorithm that we use is presented in subsection 4.2.4 and tested by considering time series having a known structure in subsection 4.2.5. In subsection 4.2.6, we apply this algorithm to the analysis of the two types of time series generated by the model of Xu et al. and discuss the results obtained.

4.2.1 Introductory discussion

Until recently, it was assumed that randomness was caused by extreme complication, i.e. the presence of many irreducible Degrees Of Freedom (DOF). This led to Kolmogorov's theory of random processes, which he defined in terms of the joint probability distribution

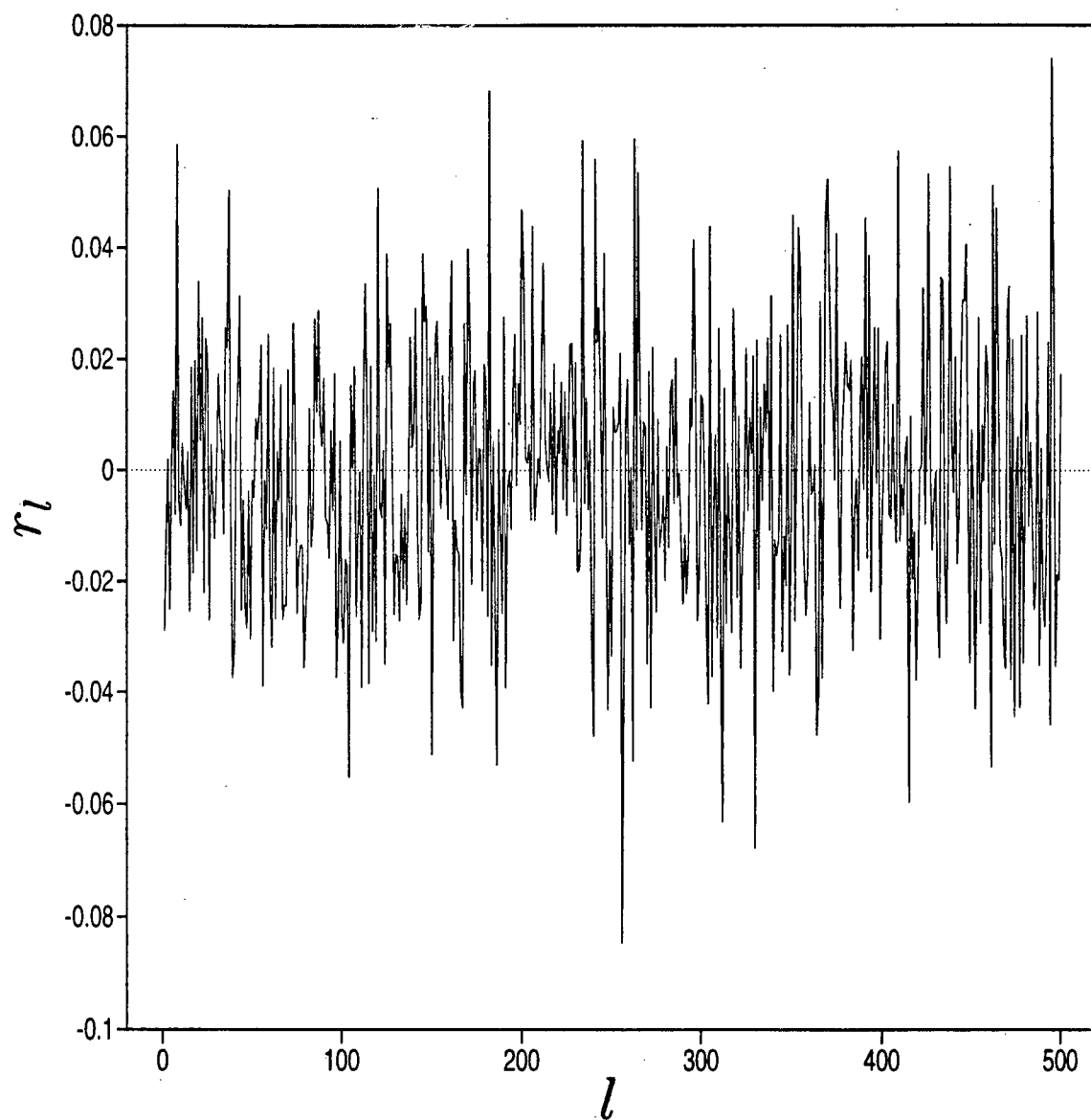


Figure 4.4: Autocorrelation function for the time series of type 1 [variety (b)] plotted on figure 4.2. For convenience, the value $r_0 = 1$ is not shown on the graph.

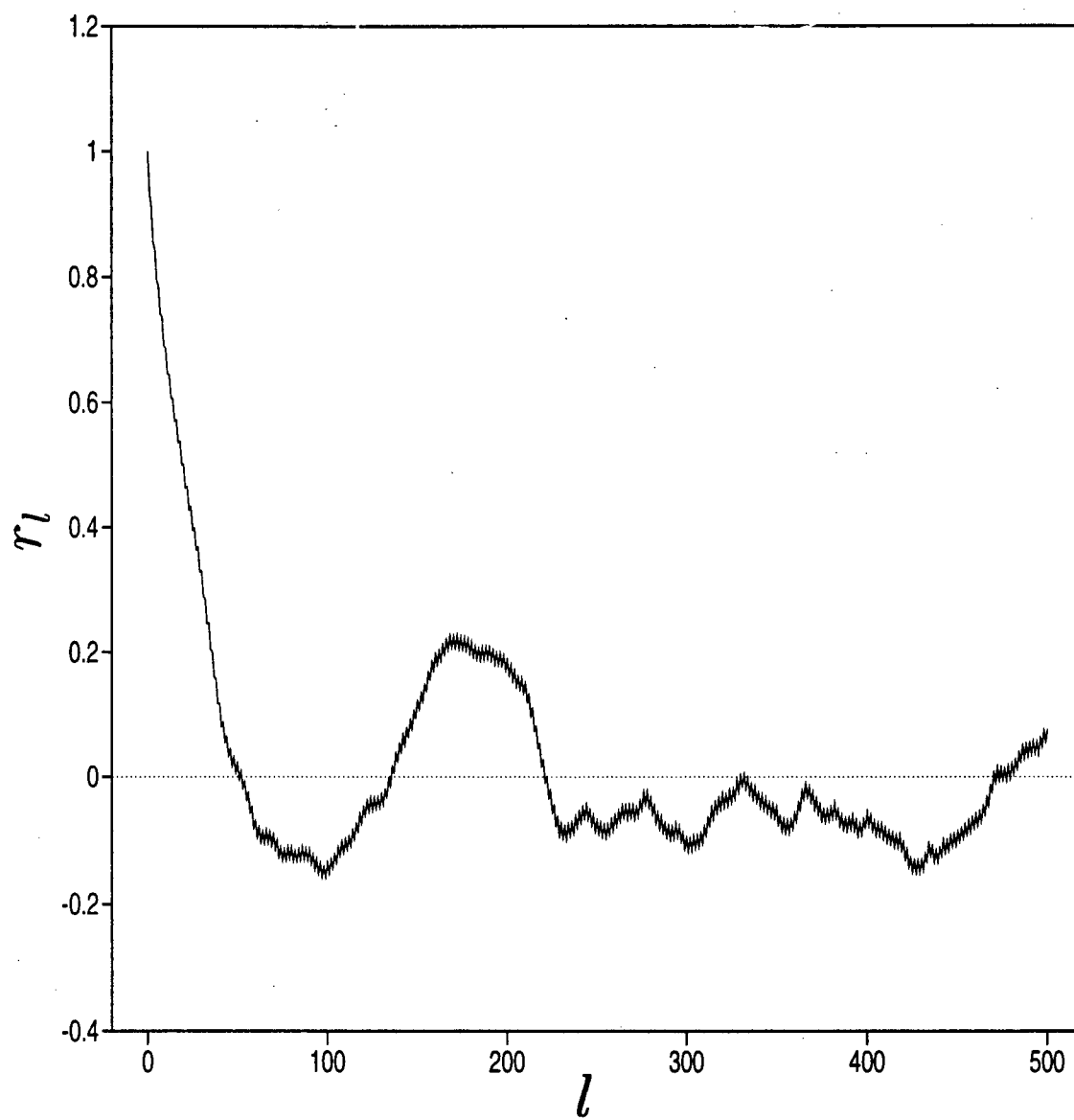


Figure 4.5: Autocorrelation function for $N = 2000$ data of the time series of type 2 plotted partially on figure 4.3.

P . For a time series $\{x_i\}$ ($i = 1, 2, \dots, N$), the d^{th} order probability distribution is

$$P(\xi_1, \dots, \xi_d) = \text{Probability}(x_1 < \xi_1, \dots, x_d < \xi_d).$$

The theory of random processes copes with situations in which inadequate information is available and makes no statements about the causes of the randomness. As a result, with more data or more accurate observations, a phenomenon that previously seemed random might become more predictable, and hence less random. Many of the classic examples of randomness are not complicated. This is the case for a flipping coin for example. The dynamics of a flipping coin involves only a few DOF. Its randomness comes from a sensitive dependence on the initial conditions: a small perturbation causes a much larger effect at a later time, making prediction difficult. When sensitive dependence on the initial conditions occurs in a sustained way, it is called *chaos* (see [21] for an outline at the introductory level). Since chaos is defined in the context of deterministic dynamics, in some strict sense chaos is not random. But it is relevant to note that chaos amplifies noise exponentially, so any uncertainty present initially is amplified to macroscopic proportions in a finite time. As a result, short-term determinism becomes long-term randomness.

The flipping coin being a chaotic system, it is inherently nonlinear. A nonlinear system is not necessarily chaotic though! The flipping coin is an autonomous system since it receives no time dependent input from the outside. Non-autonomous (or driven) systems are rather common: they are found in many fields of science, for example in engineering, chemistry, biology, physical geography and economics. An external agent provides the necessary drive, which can take on many different forms: periodic and random are the most common forms. When the driving process is periodic, a Poincaré section can reduce the system to an autonomous mapping problem. On the other hand, when the driving process is random, it is not possible to simplify the problem in such a way. This is the case in vibration tests performed on mechanical systems. A generic representation of

the kind of mechanical systems that are studied in engineering is provided by the driven Duffing oscillator. It is described by the equation (see [47] for instance)

$$x'' - x_0'' + 2\zeta\omega_n(x' - x_0') + \omega_n^2(x - x_0) + \alpha\omega_n^2(x - x_0)^2 + \beta\omega_n^2(x - x_0) |x - x_0| = 0, \quad (4.1)$$

where ω_n is the natural frequency in the absence of nonlinearity, ζ the damping coefficient and β (respectively α) the coefficient of the antisymmetric (respectively symmetric) nonlinear restoring force. The oscillator is driven through the acceleration x_0'' , from which x_0' and x_0 can be computed. Substantial high-order harmonic responses alert to the presence of a nonlinearity in the system. For given values of α and β , the behavior of equation (4.1) is chaotic for appreciable driving levels and nearly linear for low driving levels [47]. The driven Duffing oscillator is useful to illustrate what is going on in the model of chapter 2, with the important difference that the former has only a few DOF. As seen in section 2.4, the systems studied with the model of Xu et al. are slowly driven and thus the amount of stress added to the system is just enough to cause one rupture in it. As a result, the systems studied with the model of Xu et al. will not be treated as input-output systems: we will consider instead that only a single time series is available for the forecasting analysis and use the techniques appropriate for such situations (see below). Note that the important role played by the external drive in the model of Xu et al. should not be minimized though.

Both chaotic and random time series have broadband spectra, so that distinguishing them requires some other statistics than the autocorrelation function and the Fourier power spectrum. There are a number of statistics which are sensitive to nonlinear behavior, namely the short-term prediction error, the Lyapunov exponents [112] and the correlation dimension [39]. The short-term prediction error is small for a chaotic system and large for a random process. The Lyapunov exponents quantify the rate at which

nearby trajectories in the state space of a system diverge. For a system to be considered as chaotic, at least one of its Lyapunov exponents must be positive (necessary but not sufficient condition). The correlation dimension (ν) *estimates* the dimension of the subset of the state space of a system (often called the attractor of the system) that the trajectories approach. For example, fluid flows have an infinite-dimensional state space, but can have low-dimensional attractors [60]. The correlation dimension can be obtained by applying the algorithm of Grassberger and Procaccia [39]:

$$\nu_n = \lim_{\epsilon \rightarrow 0} \log C_n(\epsilon) / \log(\epsilon),$$

where C_n is the correlation integral for the n -histories of the time series

$$C_n(\epsilon) = \lim_{N \rightarrow \infty} \left(\frac{2}{N(N-1)} \right) \sum_{i>j} \theta(\epsilon - \|\vec{x}_i^n - \vec{x}_j^n\|),$$

$\theta(x)$ being the Heaviside function ($\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ otherwise) and the norm used is the maximum norm. The n -histories of the time series are constructed by using the vectors

$$\vec{x}_i^n = (x_{i-n+1}, x_{i-n+2}, \dots, x_i).$$

If the time series comes from a chaotic system, then ν_n does not increase with n and converges to the value ν . An alternative way of characterizing the attractor of a system is through the fractal (or Hausdorff) dimension (D). This number can be obtained by using the box-counting algorithm for instance (see [40]). In most cases, if $D < 2$, then ν and D are close together, but it is important to keep in mind that D has only to do with the geometrical structure of the attractor whereas ν is sensitive to the dynamical process of coverage of the attractor [39]. As a result, ν tends to be a more relevant measure of the attractor than D . The correlation dimension (ν) is useful to distinguish between chaos and randomness. However, the use of the correlation dimension as a statistic to distinguish between chaos and randomness has its limitations (see for instance [79]).

Along the same line, three economists (Brock, Dechert and Scheinkman) have recently introduced a test based on the correlation dimension: the BDS test [9]. It is a test against IID (Independent Identically Distributed) noise (null hypothesis). The test can be modified to single out the nonlinear behavior by subtracting out the linear component of the data and applying the BDS test to the residuals. A procedure advocated in [102] is instead to combine the BDS test with a technique called *bootstrapping with surrogate data sets*. The result is a powerful tool to detect any nonlinearity in a time series.

In this work, we consider only the short-term prediction error. This statistic is very powerful in distinguishing between low-dimensional chaos and randomness, but does not have the limitations observed with the correlation dimension. An example of the use of the short-term prediction error for such a purpose is provided by the nonlinear forecasting algorithm of Farmer and Sidorowich [29]. In this algorithm, the error of a local linear predictor is plotted against a variable smoothing parameter. If maximum smoothing allows to get the best forecasts, then the structure in the time series is described by a linear stochastic model. If the smallest forecasting errors are obtained for minimum smoothing, then the time series is generated from a low-dimensional chaotic system. If any amount of smoothing less than the maximum but more than the minimum is needed to get the best forecasts, then a nonlinear stochastic model describes the structure in the time series. We will outline in subsection 4.2.4 an algorithm which can be used to construct these different types of models. For now, we mention that linear stochastic models have a long history, going back at least to Yule [116] and more recently Box and Jenkins [8]. On the other hand, nonlinear deterministic models have been constructed by the dynamical systems community since about 1987 (see [23] for instance) and independently, the statistics community has constructed nonlinear stochastic models since about 1980 (for a review see [103]).

In this work, we assume that a good model from first-principles can not be found,

and so we must find it directly from the time series obtained from the observation of the dynamical system. This situation is what happens for a dynamical system studied with the model of Xu et al. since the collective behavior of the large number of elements in the system does not arise from a small set of equations. (Generally, self-organized critical systems are *believed* to have many DOF [14]). The time series generated in this thesis are all scalar (see section 4.1), i.e. they each involve only one variable, which is less than the number of variables necessary to fully describe the dynamical system. These time series are not contaminated by observational noise. However, dynamical noise affects the time series because the dynamics is in general not deterministic. To this effect, it is relevant to note that because time series for systems with both random σ_{th} and constant σ_{th} are considered, then it is possible to study the effect of the dynamical noise on the system dynamics. In any case, if neither observational noise nor dynamical noise are too large and the randomness in the time series is caused by low-dimensional chaotic behavior, then building a dynamical model directly from the time series involves two steps: *state space reconstruction* and *nonlinear function approximation*. These are the two important ingredients entering into a nonlinear forecasting algorithm. Before we present the version of the nonlinear forecasting algorithm that we use, we describe in some details these two ingredients.

4.2.2 State space reconstruction

In most physical situations, only a single scalar time series $\{x_i\}$ ($i = 1, \dots, N$) obtained by making measurements on a dynamical system is available. The time series has in general fewer variables than the number of variables necessary to fully describe the system. An illustration of this situation is provided by a fluid flow experiment, which in principle can be modeled by the Navier-Stokes equations. In the case in which only a single measure of a given velocity component at a given point in space is performed, then the resulting

time series is inadequate to provide the initial conditions for the Navier-Stokes equations. To build a model directly from the time series, the state space of the dynamical system must be reconstructed from this unique time series.

A popular method for reconstructing the state space of a dynamical system was introduced by Packard et al. [80] and put on a firm mathematical foundation by Takens [101]. In this method, the past behavior of a single scalar time series is used to get information about the present state of the system. This information can be represented as a delay vector of embedding dimension m and delay time τ :

$$\vec{X}_i = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})^\dagger, \quad (4.2)$$

where the components of a delay vector are called the *delay coordinates*. The \dagger symbol insures that the delay vectors are column vectors, which is the standard way of representing them in the literature. The optimal choice of parameters m and τ has been the subject of many theoretical studies (see [34] for instance). In practice, τ is arbitrary but it should not be too small or too large. Also, m is often chosen by trial and error, starting with a small value and increasing it, searching for optimal results.

Other methods of state space reconstruction of common use include the principal component technique and reconstruction using derivative coordinates (see [30] for a brief outline of these methods). Also, filtering is sometimes used in combination with any reconstruction method. It is of interest to note that principal component, derivative and delay coordinates, and filtered versions thereof, are all related to each other by linear transformations whereas the transformation from delay coordinates to the original coordinates (i.e. the data of the time series) is nonlinear. As demonstrated by Fraser [33], nonlinear coordinate transformations can be greatly superior. It is not clear which method of state space reconstruction works best and what causes one method to be better than another though. In the forecasting algorithm presented in subsection 4.2.4, delay

coordinates are used to reconstruct the state space of the dynamical system.

4.2.3 Nonlinear function approximation

Once the state space of the dynamical system has been reconstructed, the next task is to predict the future behavior of the scalar time series obtained by making measurements on this system. This is called *forecasting* and it involves an extrapolation in time, in the sense that one uses data from one domain (the past) to extrapolate to the behavior of data in a disjoint domain (the future). State space reconstruction makes it possible to convert the extrapolation of a time series into a problem of interpolation in the state space of the system. This problem can be expressed mathematically in a simple way. Indeed, using the fact that the delay vectors [see (4.2)] contain information about the states of the system at different times, we have

$$x_{i+T} = f^T \vec{X}_i = f^T(x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})^\dagger, \quad (4.3)$$

where, for simplicity, we assume that only the first coordinate of \vec{X}_{i+T} is forecasted, T being called the *forecasting time*. The function $f^T : \mathcal{R}^m \rightarrow \mathcal{R}$ holds for all i . The function f^T can, in general, only be determined approximately. The most commonly used approach to time series forecasting assumes that f^T is a linear function [8]. However, almost any real system contains nonlinearities, which must be taken care of by a nonlinear f^T . A particular and important kind of nonlinearity is one due to chaos, for which the scalar time series is generated by motion on a D -dimensional attractor. According to Sauer et al. [87], it is sufficient to take $m > 2D$ for a smooth f^T to hold for all i and if $D < m < 2D$, then (4.3) holds for *almost* all i . In contrast, if the time series is generated by a stochastic process then, for all m , a noise term is expected to appear on the right-hand side of (4.3) and the delay vectors are expected to fill out a set of dimension m [87]. This situation also prevails if the time series is generated by motion

on an attractor with dimension D large. Indeed, in such a case, there will be insufficient data to approximate f^T with a deterministic model with $m > D$ and in this sense, a high-dimensional deterministic system is equivalent to a low-dimensional nonlinear stochastic system [14]. On the other hand, when D is small, then an accurate approximation of f^T can always be obtained with a modest number of data ($N \sim 10^D$).

The approximation of function f^T requires that a representation as well as a method of selecting the parameters of the representation be chosen. When f^T is nonlinear, there is an infinite variety of possible representations, and in the absence of any prior knowledge, no reason to prefer one representation over another. In many cases, a given representation is chosen because it allows a fast algorithm for fitting the parameters. It is beyond the scope of this thesis to review the most popular nonlinear function representations (see however [30]), but it is relevant to mention that representations can be both *global* and *local*. We describe next one way of obtaining a local approximation of f^T since this method is used in the nonlinear forecasting algorithm considered below. A local approximation usually enables one to get a good fit since the variations are less extreme than when a global representation is used.

The basic idea of a local approximation of f^T is to break up the domain of f^T into local neighborhoods and fit different parameters in each one. There are several ways to assign the neighborhoods (see [30] for a review). One way of doing this was introduced by Farmer and Sidorowich [29] and is now commonly used: they impose a metric on the state space, denoted by $\| \cdot \|$, and find the k nearest neighbors of \vec{X}_i [see (4.3)], i.e. the k delay vectors \vec{X}_j with $j < i$ that minimize $\| \vec{X}_i - \vec{X}_j \|$. Afterwards, in each neighborhood, a local predictor for x_{i+T} is constructed. The simplest approach to do so is an approximation by the nearest neighbor ($k = 1$), i.e. $x_{i+T}^{pred} = x_{j+T}$. For example, to predict tomorrow's weather, one would search the historical record and find the weather pattern the most similar to that of the last few days, and predict that

tomorrow's weather pattern will be the same as the neighboring pattern one day later. A superior approach is the *local linear approximation* ($k > 1$). This approach consists in fitting a linear polynomial to the pairs (\vec{X}_j, x_{j+T}) . The fit can be made in any of several ways. Often, least squares by singular-value decomposition (see [85] for instance) is realized. In principle, approximations using higher-order polynomials can be performed but in practice, this may or may not give an improvement depending on the time series analyzed. The method of Farmer and Sidorowich has the advantage of being quick and accurate, but the disadvantage that the approximations are discontinuous. Continuity may be guaranteed by weighted local fitting (see an illustration of this in a particular case in [98]) but for simplicity, simple local fitting is often considered. In the nonlinear forecasting algorithm presented next, simple local linear fitting is implemented.

4.2.4 Nonlinear forecasting algorithm considered

The nonlinear forecasting algorithm presented below investigates whether the data in a scalar time series generated from a dynamical system exhibit a low-dimensional chaotic behavior, as opposed to a high-dimensional or stochastic behavior. The algorithm constructs models with a variable smoothing parameter (k) which at one extreme defines a nonlinear deterministic model, and at the other extreme, a linear stochastic model. The accuracy of the short-term forecasting error as a function of the smoothing parameter reveals much about the underlying dynamics generating the aperiodic time series. In particular, if any amount of smoothing less than the maximum allows one to get the best forecasts, then the structure in the time series is described by a nonlinear stochastic model. The smoothing parameter (k) represents the number of nearest-neighbor delay vectors included in the local linear fitting of the function f^T (see the discussion in the previous subsection).

In addition to its use for detecting low-dimensional chaos, the nonlinear forecasting

algorithm considered below can be used to investigate whether time series from high-dimensional deterministic systems can be approximately modeled with low-dimensional nonlinear stochastic models. This problem has recently attracted the attention of some physicists who study collective phenomena such as spatiotemporal chaos [22], robust intermittency [53] and Self-Organized Criticality (SOC) [3].

A version of a nonlinear forecasting algorithm is now outlined. It was mainly inspired from the algorithm introduced by Casdagli [14] and described by Ward [109]. The implementation of this algorithm involves the following steps:

- (a) Divide the analyzed time series into a fitting set x_1, \dots, x_{N_f} and a testing set x_{N_f+1}, \dots, x_N . As suggested in [102], the fitting set should be appreciably larger than the testing set.
- (b) Choose an embedding dimension m and set the delay time τ as well as the forecasting time T to 1.
- (c) Choose a delay vector \vec{X}_i with $N_f + m \leq i \leq N - 1$ for a 1-step ahead forecasting test.
- (d) Compute the distances d_{ij} of test vector \vec{X}_i from the delay vectors \vec{X}_j ($m \leq j \leq N_f - 1$) in the fitting set using the ordinary Euclidean norm.
- (e) Order the distances d_{ij} , find the k nearest neighbors $\vec{X}_{j(1)}, \vec{X}_{j(2)}, \dots, \vec{X}_{j(k)}$ of \vec{X}_i and fit a model of the following form:

$$x_{j(l)+1} \approx \alpha_0 + \sum_{n=1}^m \alpha_n x_{j(l)-n+1} \quad (l = 1, \dots, k). \quad (4.4)$$

The heapsort algorithm described in Press et al. [85] is used to order the distances d_{ij} . The parameters $\alpha_0, \dots, \alpha_m$ are obtained by solving the normal equations for the linear system of equations (4.4) using singular-value decomposition (see [85]).

- (f) Use the fitted model (4.4) to estimate a 1-step ahead forecast $\hat{x}_{i+1}(k)$ for test vector \vec{X}_i and compute its error

$$e_i(k) = | \hat{x}_{i+1}(k) - x_{i+1} | .$$

- (g) Repeat steps (d)-(f) for all allowed i in the testing set and compute the normalized Root-Mean-Square (RMS) forecasting error

$$E_m(k) = \frac{1}{\sqrt{N_t} \times \sigma} \left[\sum_i e_i^2(k) \right]^{\frac{1}{2}}$$

where σ is the standard deviation of the time series and $N_t = N - N_f - m$ is the number of allowed vectors in the testing set. The time series is forecastable only if $E_m(k) < 1$.

- (h) Repeat step (g) for several representative values of k between $2(m+1)$ and $N_f - m$.
- (i) Vary the embedding dimension m .

This algorithm can readily be applied to the analysis of time series by studying the curves $E_m(k)$ as a function of k for a few values of m . We will do that first for time series having a known structure in order to illustrate the different types of curves that result from the analysis. Afterwards, we will apply this algorithm to the analysis of the time series generated by the model of Xu et al. (see section 4.1).

In the literature, it is not always clear whether a nonlinear forecasting algorithm should be applied to the raw time series or the first-differenced time series. On the one hand, Sugihara and May [99] have analyzed first-differenced time series, where the first-difference operation was performed, they say, to clarify the suspected nonlinearities by reducing the effects of any short-term linear autocorrelations. On the other hand, Casdagli [14] has analyzed raw time series that he assumed stationary, this assumption

being important since only stationary time series should be analyzed with the algorithm. According to Kendall and Ord [54], a time series is stationary if its mean and variance are constant over time and if the structure of the time series depends only upon the relative position in time of the observations. These conditions must be satisfied at least to a reasonable degree of approximation for a time series to qualify as stationary. In the case of the time series generated by the model of Xu et al., we have checked that they are reasonably stationary. This means that it is not necessary to first-difference the time series since they are stationary in the mean (i.e. they do not exhibit a linear drift in the mean). However, the first-difference operation performed in [99] was shown to improve the results. For this reason, we will take no chance and consider both raw time series and their first-differenced counterparts. This obviously applies only to the time series generated by the model of Xu et al. since the ones considered in the next subsection have a known structure and we therefore know what results to expect from the algorithm.

4.2.5 Tests of the algorithm with time series having a known structure

In this subsection, we test the forecasting algorithm outlined in subsection 4.2.4 by analyzing four time series having $N = 1000$ data and a known structure. Before we do this, we describe how we generate these time series.

The first two time series are generated by low-dimensional chaotic maps, namely the *logistic map* and the *Hénon map*. The logistic map is univariate and is defined as [65]

$$x_i = \mu x_{i-1}(1 - x_{i-1}),$$

where μ is the parameter of the map. For $\mu \geq 3.58$, the map exhibits chaos for *almost* all values of μ . The time series was obtained for the value $\mu = 4.0$, a value in the chaotic regime. The Hénon map is bivariate and is described by the pair of equations [44]

$$x_i = 1 - Ax_{i-1}^2 + y_{i-1}$$

$$y_i = Bx_{i-1},$$

which can be combined to get the single equation

$$x_i = 1 - Ax_{i-1}^2 + Bx_{i-2}.$$

For the special values $A = 1.4$ and $B = 0.3$, the map is known to have a chaotic behavior.

The last two time series have a substantial stochastic component. One of them is generated by a nonlinear stochastic model, the Threshold AutoRegressive (TAR) model. The TAR model is expressed mathematically as [104]

$$x_i = \alpha x_{i-1} + u_i \quad (x_{i-1} < 1)$$

$$x_i = \beta x_{i-1} + u_i \quad (x_{i-1} \geq 1),$$

where u_i is distributed according to a Gaussian with zero-mean and standard deviation equal to 1 (see [85] for the description of an algorithm to produce Gaussian deviates). To obtain the time series, we used the values $\alpha = -0.4$ and $\beta = 0.5$. The last time series is generated by a linear stochastic model, a AR(1) (first-order autoregressive) model (see [54] for instance), i.e.

$$x_i = \phi x_{i-1} + u_i,$$

where ϕ is a constant. To obtain the time series, we used the value $\phi = 0.5$.

For the logistic map, the TAR model and the AR(1) model, it is necessary to choose one starting point for the iterations, whereas for the Hénon map, it is necessary to choose two. The starting point(s) has (have) some effect on the resulting time series. However, if, say, the first 500 data generated in an iteration process are discarded, then the resulting time series will not be sensitive to the starting point(s). The four time series can then be readily analyzed using the nonlinear forecasting algorithm presented in the previous subsection.

The two time series generated by low-dimensional chaotic maps were first analyzed with the forecasting algorithm and the results are displayed on figure 4.6 for an embedding dimension (m) equal to 2. Comparable results are obtained for $m > 2$ and are not shown. As expected, the best forecasts (i.e. the smallest $E_m(k)$) are obtained, for both time series, when the smoothing parameter (k) is small. It is relevant to note also from the figure that excellent forecasts (i.e. $E_m(k) \ll 1$) are obtained at small k and that they are much better than the ones obtained at large k . This reflects the fact that no stochastic component is present in the time series. It is also possible from figure 4.6 to compare the forecasting errors for the logistic map and the Hénon map. Indeed, at small k , $E_m(k)$ is *always* smaller for the logistic map than for the Hénon map. This feature reflects the fact that the attractor for the Hénon map has a more complex structure than the one for the logistic map.

If we now turn to the time series having a stochastic component, the results of the analysis are plotted on figure 4.7 (time series for the TAR model) and figure 4.8 (time series for the AR(1) model) for three values of m in both cases. On figure 4.7, we see that the best forecasts are obtained at intermediate values of k ($k \approx 400 - 500$) and that the improvement in accuracy over a linear stochastic hypothesis (corresponding to the largest values of k) is only of about 2 – 3%. This is typical of a nonlinear time series having a substantial stochastic component. On figure 4.8, we notice that the forecasts get better as k increases, but no upturn in $E_m(k)$ at large k is observed. These features are typical of a linear time series having a substantial stochastic component. Finally, it is interesting to compare figure 4.6 to figures 4.7 and 4.8. We note that $E_m(k)$ is always much smaller for the time series generated by the low-dimensional chaotic maps than for the stochastic time series, as it should.

Now that we have obtained typical curves of $E_m(k)$ versus k for time series having a known structure, we turn to the analysis of the time series introduced in section 4.1 and

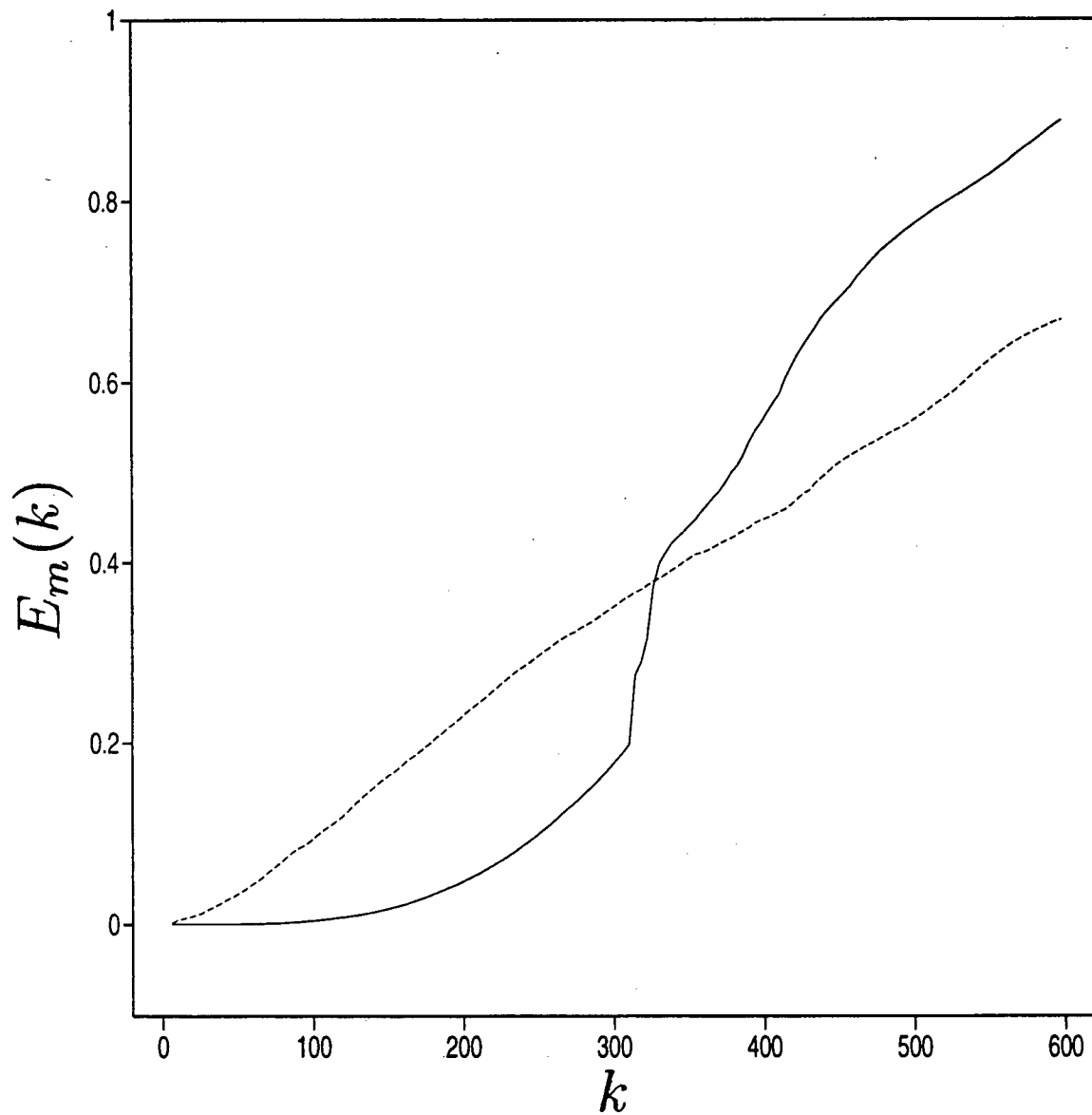


Figure 4.6: Normalized RMS forecasting error $E_m(k)$ as a function of k for time series having $N = 1000$ data generated by the logistic map (with $\mu = 4.0$) and the Hénon map (with $A = 1.4$ and $B = 0.3$). The solid line is for the logistic map time series and the dashed line for the Hénon map time series. The parameters of the nonlinear forecasting algorithm are $N_f = 800$ and $m = 2$.

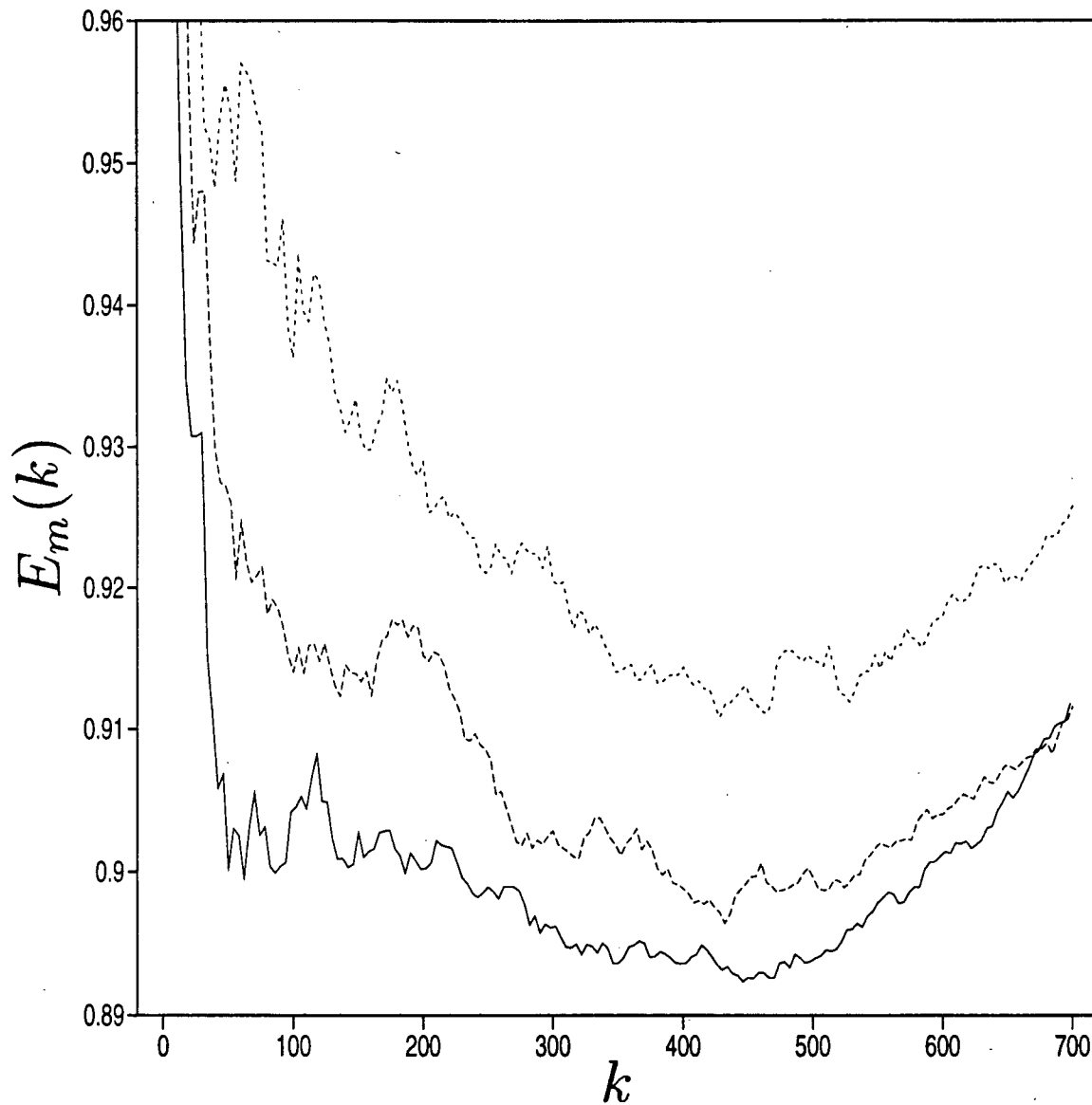


Figure 4.7: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series having $N = 1000$ data generated by the TAR model (with $\alpha = -0.4$ and $\beta = 0.5$). The parameters of the nonlinear forecasting algorithm are $N_f = 800$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

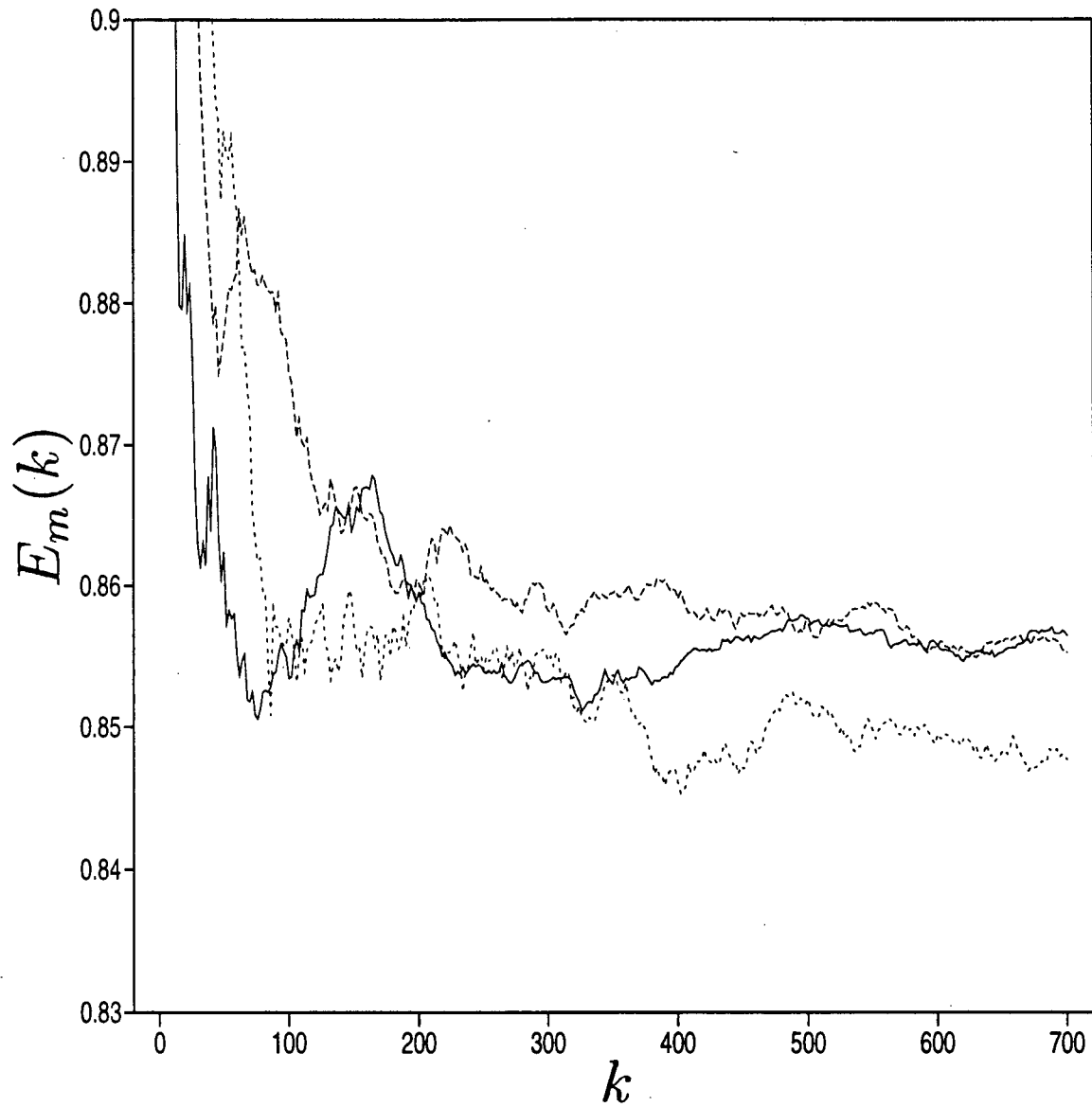


Figure 4.8: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series having $N = 1000$ data generated by the AR(1) model (with $\phi = 0.5$). The parameters of the algorithm are $N_f = 800$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

generated by the earthquake model of Xu et al.

4.2.6 Results obtained with the time series of interest

Unless otherwise specified, all the time series considered in the analysis have $N = 2000$ data. As discussed in subsection 4.2.4, we consider both raw time series and first-differenced versions of them. The two types of time series outlined in section 4.1 are analyzed. Also, for a given type of time series, time series for systems with both random σ_{th} and constant σ_{th} are considered.

We first analyze the time series of type 1. As discussed in section 4.1, the time series of type 1 exist in three varieties, namely (a), (b) and (c). We have analyzed the three varieties of time series of type 1 and found, for a 100×100 system, that only the time series of variety (c) are forecastable (i.e. have $E_m(k) < 1$). The results for a time series of variety (c) for a system with random σ_{th} are plotted on figure 4.9. Similar curves are obtained with a time series for a 100×100 system with constant σ_{th} (results not shown). On figure 4.9, we observe a monotonic decrease of $E_m(k)$ with k . Also, we notice that the smallest forecasting errors are at most 5% better than the standard deviation of the corresponding time series. This suggests that the time series of variety (c) are weakly forecastable, probably because they have a large stochastic component. In addition, no nonlinear structure exists in the time series of variety (c). This was also found to be true for the time series of varieties (a) and (b) and whether or not the time series are first-differenced. Now, we turn to the time series of type 1 for a one-dimensional system. Contrary to the case of the time series for a 100×100 system, all three varieties of time series are forecastable, but they still have a large stochastic component and no nonlinear structure (results not shown). A time series which is of particular interest is the one consisting in the time intervals between large earthquakes. As we saw in chapter 3, the one-dimensional systems with constant σ_{th} are particularly interesting since they display

stress ordering. For this reason, we included only the results for a time series of variety (a) ($S > 1000$) for a 1000×1 system with constant σ_{th} (see figure 4.10). The curve for $m = 2$ in particular seems to show that the best forecasts are obtained at intermediate values of k , but it should not be concluded that the time series has a nonlinear stochastic structure since it has not enough data. Instead, we prefer to safely state that the time series has probably no nonlinear structure but a large stochastic component. The best forecasting errors are 25% better than the standard deviation of the data. This may not be sufficient to allow the prediction of a future large earthquake in the model of Xu et al. To further confirm this possibility, we have plotted on figure 4.11 the histogram $N(\delta_t)$ of the time intervals δ_t between large earthquakes. The plot on figure 4.11 is a log-linear one in order to better illustrate that the distribution is roughly exponential, i.e. $N(\delta_t) \sim e^{-\delta_t/\delta_0}$. Obviously, because of the small number of data used to get the histogram, then the straight line on figure 4.11 does not cover the whole range of δ_t . We can then conclude from figure 4.11 that the large events generated in the model seem to be Poisson events. (As a result, we have confirmed explicitly that it was correct above to conclude from the curves on figure 4.10 that no nonlinear structure was present in the data). We got similar results with the time intervals between large events generated in a 1000×1 system with random σ_{th} as well as in 100×100 systems. This result is surprising and seems to contradict a conclusion of Chen et al. [16] obtained in a study of the model of Xu et al. that the large events are temporally correlated. However, we believe that our conclusion is not in contradiction with theirs simply because we studied a series version of the model and they studied a parallel version of it (see the discussion in section 2.4 and [35]). Consequently, it seems that only the more complicated to study parallel models can give rise to temporally correlated large events.

We next turn to the analysis of the time series of type 2, which are unevenly sampled. The $E_m(k)$ versus k curves obtained for a 100×100 system with random σ_{th} are plotted

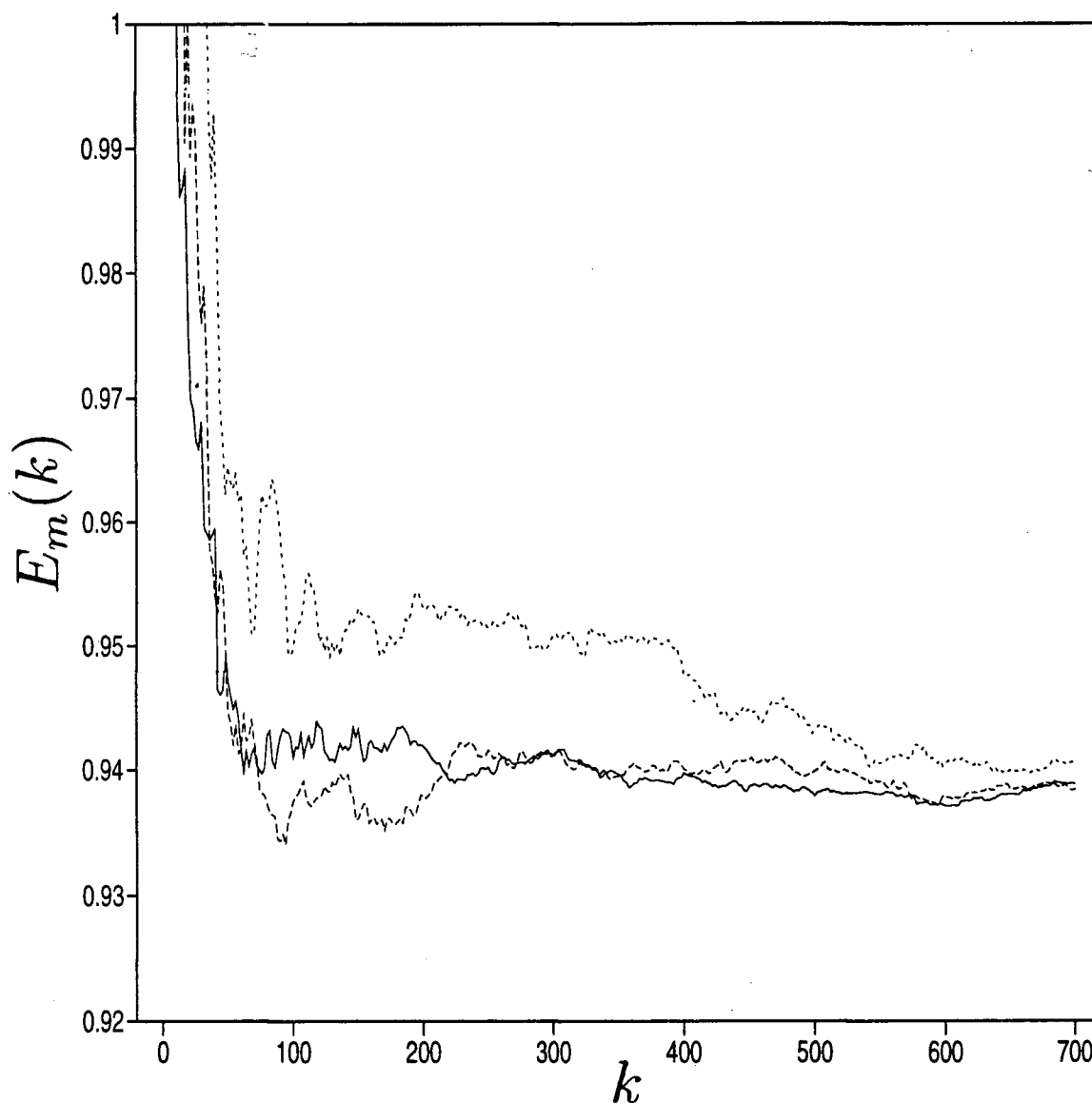


Figure 4.9: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 1 [variety (c)] for a 100×100 system with random σ_{th} . The sampling time interval Δt is equal to 10 times the average time interval between the earthquake sequences. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

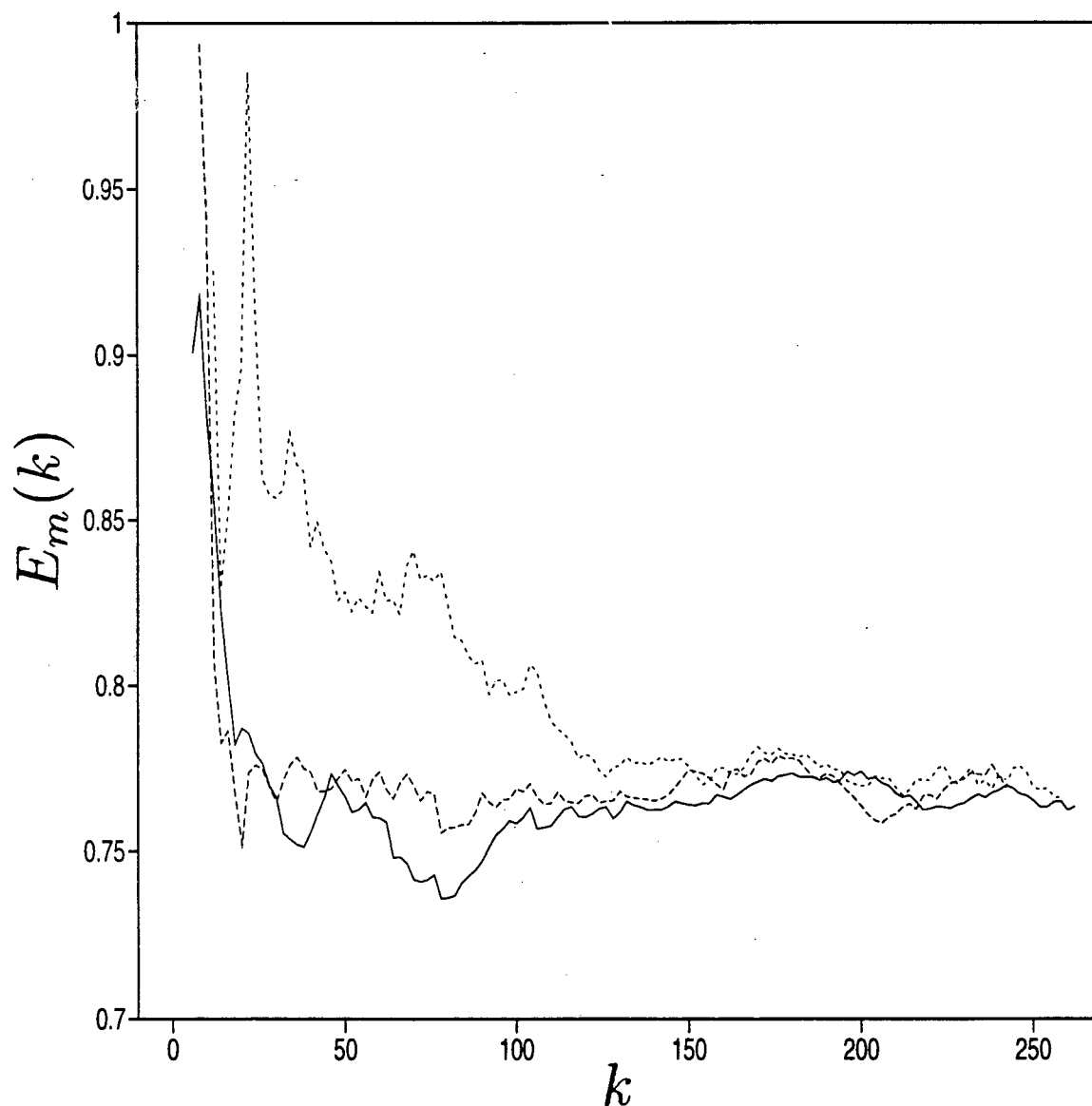


Figure 4.10: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 1 [variety (a)] for a 1000×1 system with constant σ_{th} . The data in the time series ($N = 331$) are the time intervals between large earthquakes ($S \geq 1000$). The parameters of the nonlinear forecasting algorithm are $N_f = 264$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

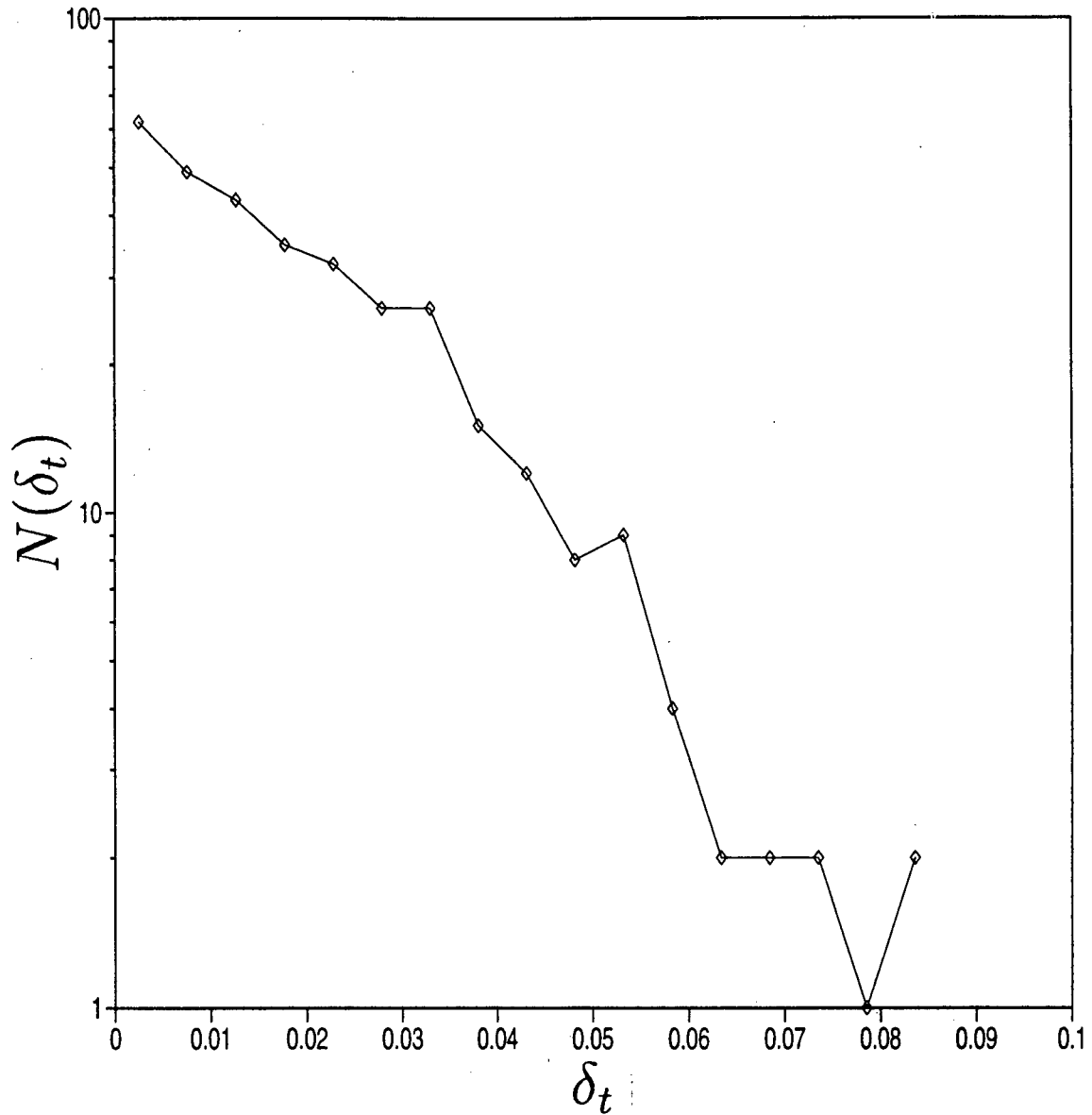


Figure 4.11: Histogram plot of the time intervals between large earthquake sequences ($S \geq 1000$) generated in a 1000×1 system with constant σ_{th} . The data set has 331 data and 20 bins were used to get the histogram. The vertical axis is logarithmic in order to better illustrate that the distribution is exponential.

on figure 4.12 (raw time series) and figure 4.13 (first-differenced time series). The curves obtained on both figures reveal a weak nonlinear stochastic structure (see figure 4.7 for comparison). However, $E_m(k)$ is appreciably smaller and the curves are less noisy for the raw time series, showing that the first-difference operation has blown up the stochastic component. It is therefore disadvantageous to perform first-differencing on the time series of type 2. On figure 4.12, the improvement in predictive accuracy of the nonlinear stochastic models (intermediate k) over the linear stochastic models (large k) is at most of 4% and it is at most of 3% on figure 4.13. As a matter of comparison, we show on figure 4.14 the $E_m(k)$ versus k curves for a time series for a 100×100 system with constant σ_{th} (without first-differencing). Again, we see that the time series has a weak nonlinear stochastic structure, where the improvement in predictive accuracy of the nonlinear stochastic models (intermediate k) over the linear stochastic models (large k) is about 8 – 9%. We believe that this 8 – 9% improvement for the time series for a system with constant σ_{th} as compared to at most 4% for the time series for a system with random σ_{th} reflects the fact that the latter system has a more noisy dynamics than the former. We have also considered a time series for a system with random σ_{th} when the parameter α controlling the level of stress conservation (see chapter 3) in the system is set to 0.9 (rather than 1.0 for the time series of figure 4.12). The results obtained by analyzing this time series are plotted on figure 4.15. A nonlinear stochastic structure is still observed though weaker than in the case $\alpha = 1.0$ (see figure 4.12). As compared to the case $\alpha = 1.0$, $E_m(k)$ is now about 30% smaller on figure 4.15, showing that the introduction of stress nonconservation in the system allows for better forecasts. We believe that this is related to the fact that, as seen in chapter 3, stress nonconservation destroys the scale invariance and introduces a characteristic earthquake size in the system. For completeness, we finally mention that the analysis of time series for one-dimensional systems leads to similar features as the ones observed with two-dimensional systems.

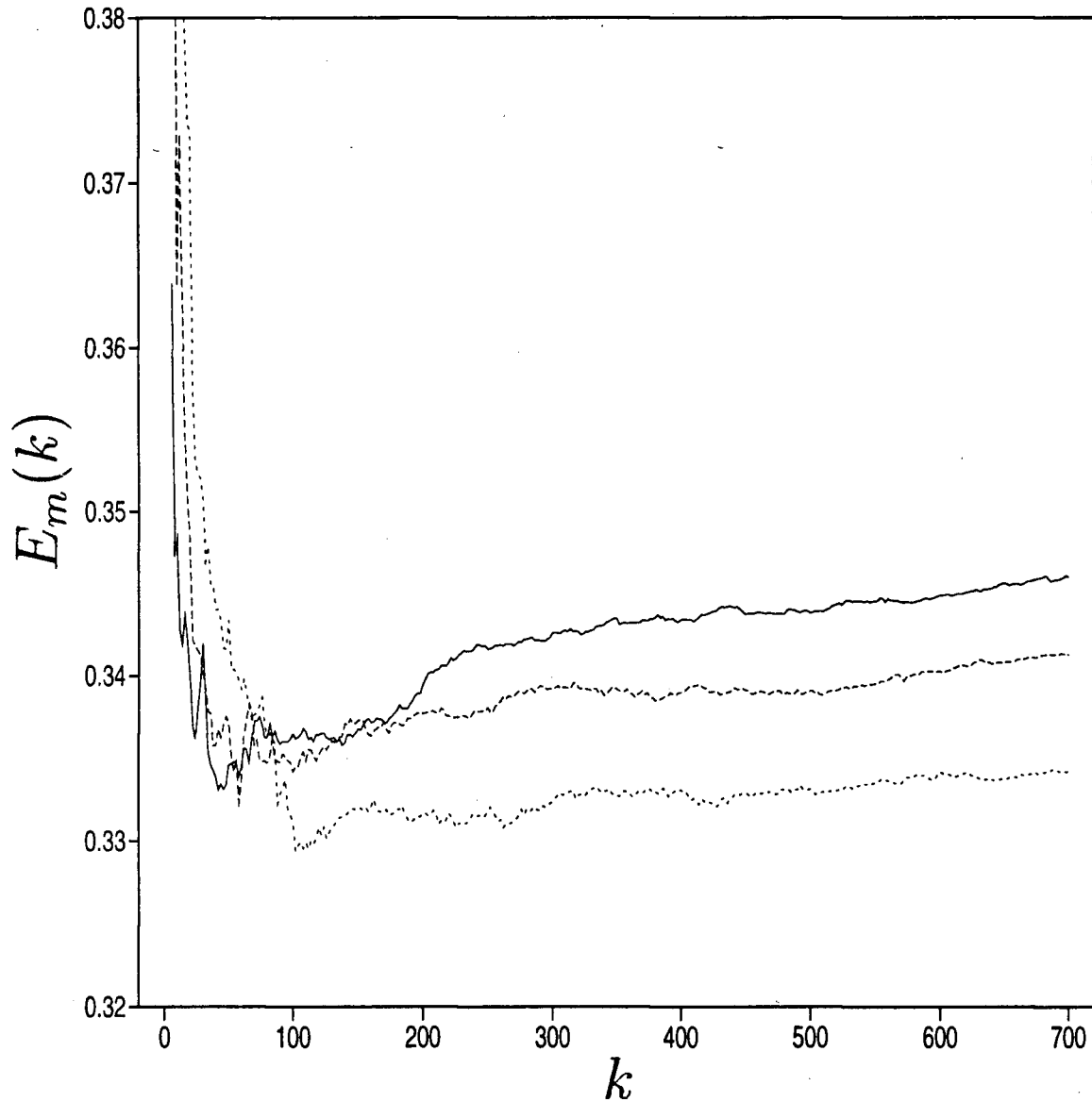


Figure 4.12: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with random σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

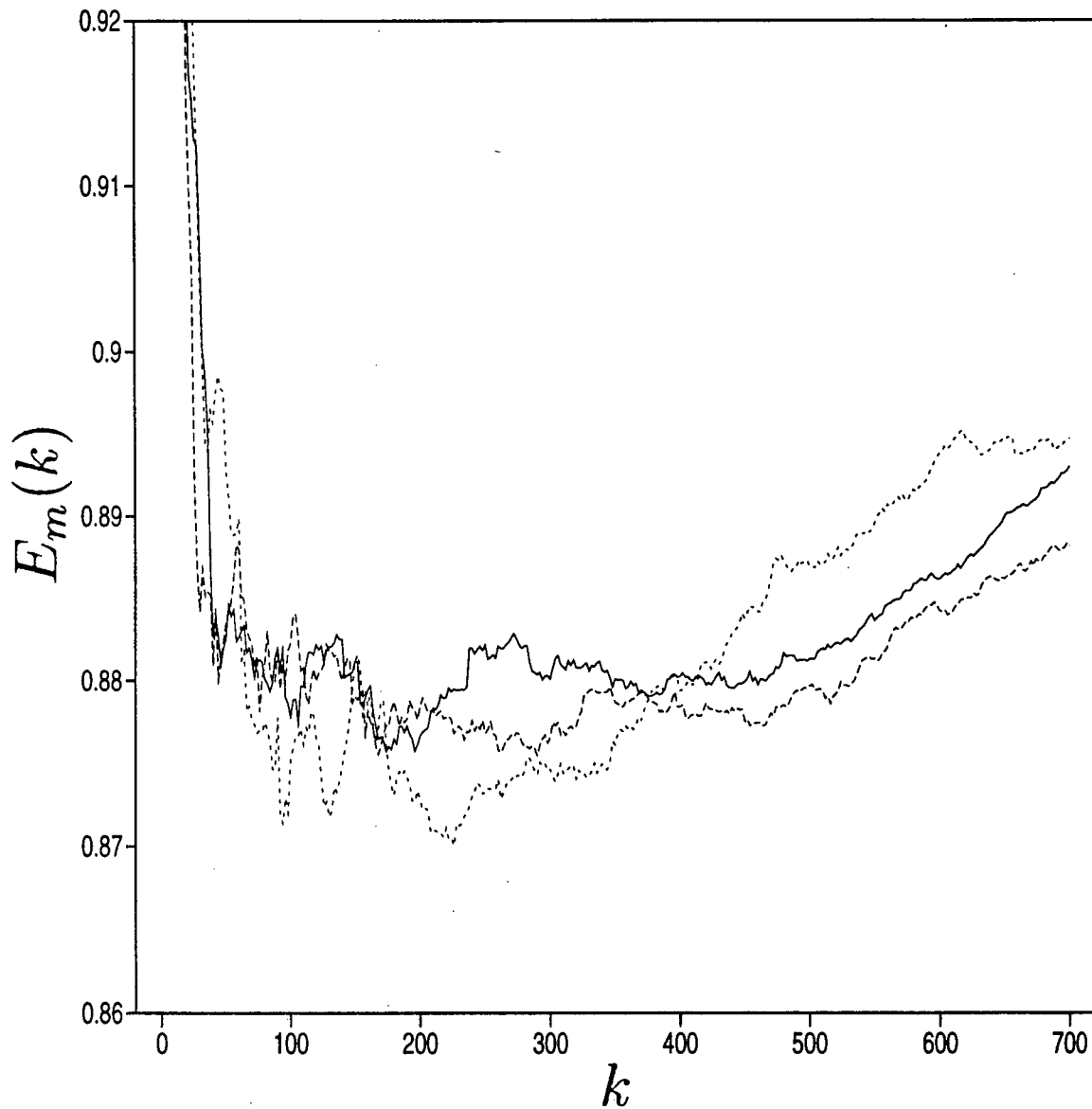


Figure 4.13: Normalized RMS forecasting error $E_m(k)$ as a function of k for a first-differenced time series of type 2 for a 100×100 system with random σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

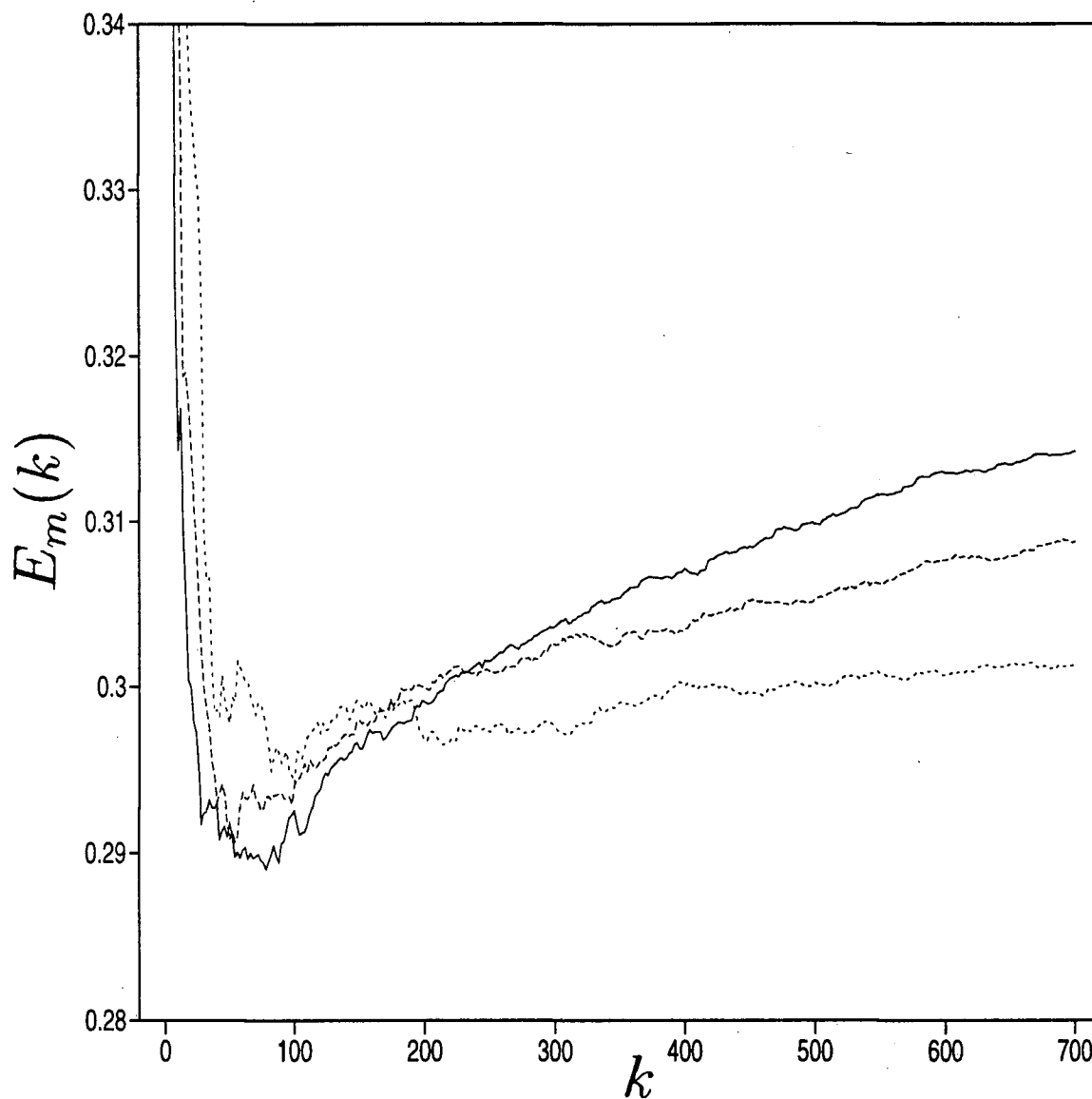


Figure 4.14: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with constant σ_{th} . The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

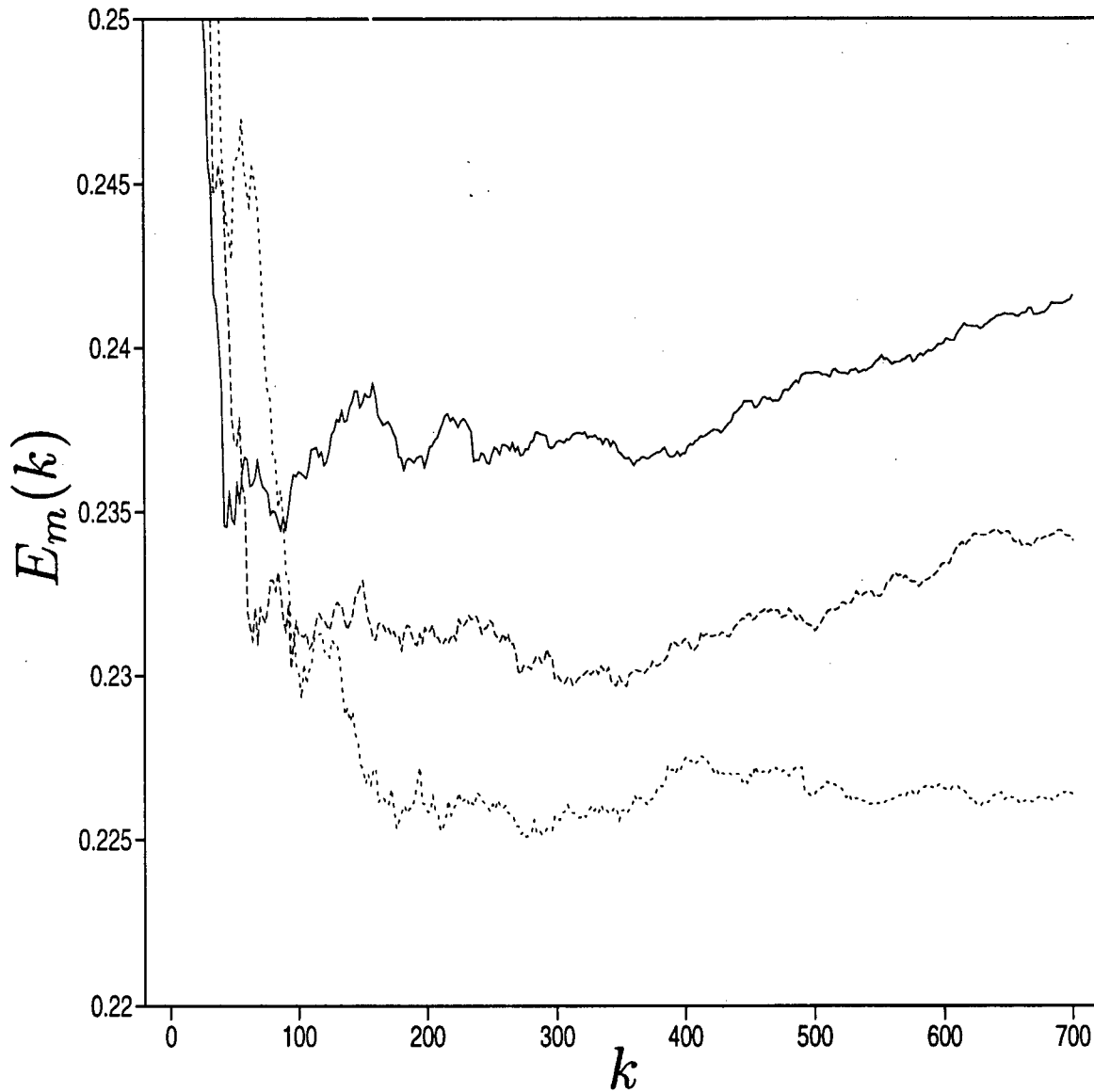


Figure 4.15: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series of type 2 for a 100×100 system with random σ_{th} when the parameter α controlling the level of stress conservation in the system (see chapter 3) is set to 0.9. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

The appearance of nonlinear structure on figures 4.12, 4.14 and 4.15 is connected to the fact that, when the stress releases due to earthquakes happen, then the system *stops following linearly* the external drive (threshold dynamics!). In addition, the presence of a stochastic component in the time series simply reflects the fact that the response of the various elements in the system varies from one element to the other: not all elements are triggered simultaneously as one element ruptures. The main consequence of these findings is that the phenomenon of SOC that the model of Xu et al. was shown to display in chapter 3 is inherently high-dimensional. This conclusion might be specific to the model studied in this thesis though. It could have been guessed from the start since the systems we study are made of a large number of elements and rarely is low-dimensional chaos found in systems with many DOF, away from the transition to chaos. This was illustrated by Casdagli [14] through a nonlinear forecasting analysis of a time series consisting in a component of the velocity at a point in a fluid in a highly turbulent regime (i.e. the Reynolds number is high).

We have checked for the robustness of the nonlinear stochastic structure found above by analyzing evenly sampled time series of type 2. The results we show are for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} , but similar results are obtained for a system with constant σ_{th} . Figure 4.16 displays the $E_m(k)$ versus k curves for such a time series. The comparison of the magnitude of $E_m(k)$ on figures 4.12 and 4.16 leads us to conclude that the sampling procedure has weakened even more the nonlinear stochastic structure: the improvement in predictive accuracy of the nonlinear stochastic models over the linear stochastic models is about 2% on figure 4.16 as compared to 4% on figure 4.12. Now, to show that the observation of a nonlinear stochastic structure in the time series of type 2 is not an artifact of the algorithm, we have randomized the phases of the frequency components of the evenly sampled time series of type 2 whose results are plotted on figure 4.16, but without altering its power

spectrum. The resulting time series is called a surrogate data set of the original data set. We generated the latter using the method of Theiler et al. [102]. In this method, the frequency components are multiplied by a factor $e^{i\phi}$, where ϕ is uniformly distributed between 0 and 2π . Afterwards, the phases are symmetrized to take into account the fact that the original time series is real. The $E_m(k)$ versus k curves for one such surrogate data set are plotted on figure 4.17. We note from this figure that the weak nonlinear stochastic structure apparent on figure 4.16 has disappeared and that the resulting curves look like the ones for a linear stochastic time series (see figure 4.8). As a result, we have proved that the nonlinear stochastic structure observed on figure 4.16, even if it is weak, is real and some weakly coherent phase relationships exist between the frequency components of the system.

In summary, we have found that the time series of type 1 are weakly forecastable or not at all in most of the cases considered. Only the time series of the time intervals between large earthquakes generated in a one-dimensional system was found to have some degree of forecastability. For all the time series of type 1, no nonlinear structure was found. On the other hand, the time series of type 2 were found to be weakly nonlinear but with a substantial stochastic component. These results suggest that SOC is inherently a high-dimensional phenomenon. Note however that it is possible that, with a parallel version of the model (see chapter 2), this conclusion does not hold true, but this remains to be investigated. Also, we have checked that there are weakly coherent phase relationships between the frequency components. We will see in the next section that the absence of strong phase relationships between the Fourier components in the evenly sampled time series of type 2 enables us to implement a rescaled range analysis. The information that we will get from this second type of analysis is complementary to the one extracted in this section. In the next section, we will restrict our attention to the time series of type 2 since, as we saw in the present section, the time series of type 1, whatever the variety,

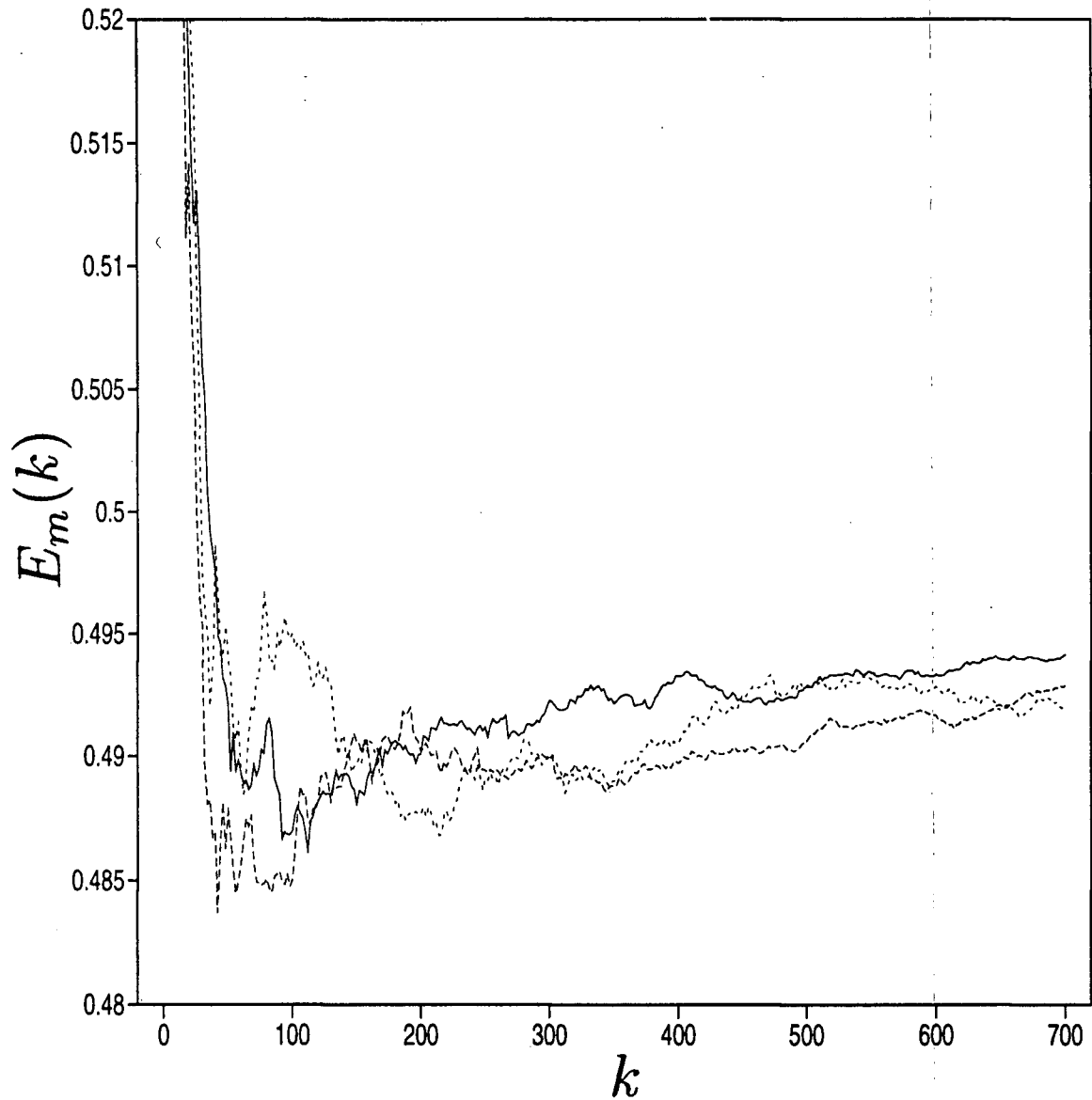


Figure 4.16: Normalized RMS forecasting error $E_m(k)$ as a function of k for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval Δt is about equal to 1.5 times the average time interval between the earthquake sequences. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

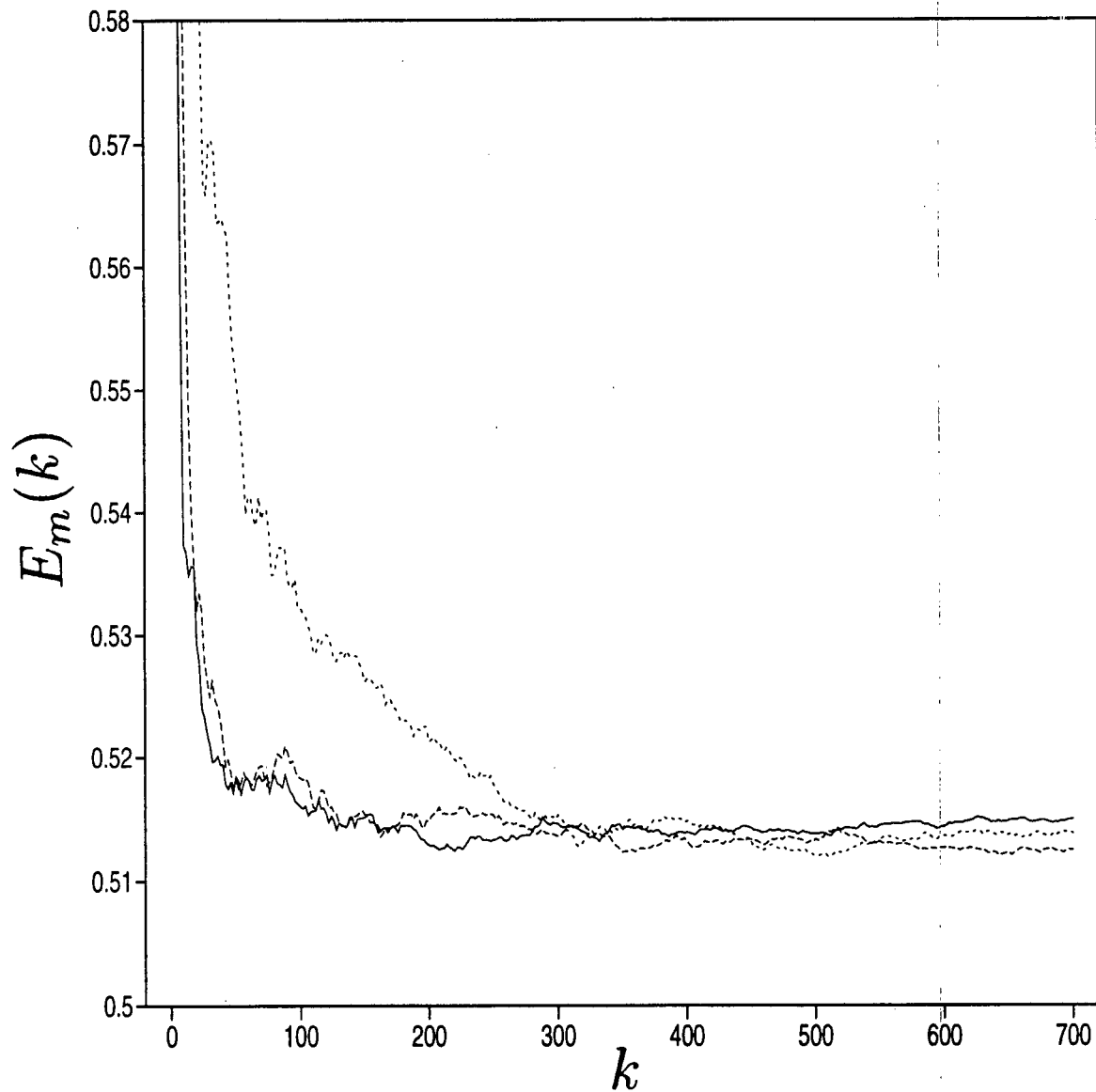


Figure 4.17: Normalized RMS forecasting error $E_m(k)$ as a function of k for a time series which has the same power spectrum as the time series of figure 4.16, but randomized phases. The parameters of the nonlinear forecasting algorithm are $N_f = 1600$ and $m = 2$ (solid line), $m = 3$ (dashed line), $m = 5$ (dotted line).

very much look like white noise and so, are not of interest anymore.

4.3 Rescaled Range Analysis

In this section, we perform a second type of time series analysis, namely a rescaled range analysis. We begin with a general introductory discussion in subsection 4.3.1, followed by a brief discussion of the fractal analyses in subsection 4.3.2. We present the algorithm used to do the rescaled range analysis in subsection 4.3.3. This algorithm is tested on synthetic time series generated by the method of successive random addition in subsection 4.3.4 and applied to the evenly sampled time series of type 2 in subsection 4.3.5. In the latter subsection, we also discuss the results obtained.

4.3.1 Introductory discussion

The power spectrum analysis is often used to examine irregular time series. It is rather common in nature to observe power spectra which have no eminent peaks and exhibit instead appreciably more power at low frequencies than at high frequencies. In such cases, the power spectrum density $P(f)$ is in most cases approximated as a power-law function of the frequency (f): $P(f) \propto f^{-\beta}$, where β is called the spectral exponent and is positive. In this approximation, β is one of the quantities describing the irregularity of the time series. Time series which exhibit power-law behavior in their power spectrum are often termed *fractal time series* [115], but as shown by Higuchi [46], this is only a sufficient condition. Indeed, it is necessary for a time series to qualify as fractal that the phase distribution of the frequency components be random. To illustrate this more clearly, we follow the ideas of Higuchi [46]. When the total number of data (N) in the evenly sampled time series $\{x(j)\}$ ($j = 1, 2, \dots, N$) is even, $x(j)$ can be expressed in terms

of the Fourier series representation as follows:

$$x(j) = \sum_{k=0}^{N/2} A_k \cos(2\pi jk/N) + \sum_{k=1}^{N/2-1} B_k \sin(2\pi jk/N) = \sum_{k=0}^{N/2} C_k \cos(2\pi jk/N - \theta_k),$$

where $C_k = \sqrt{A_k^2 + B_k^2}$ and $\theta_k = \tan^{-1}(B_k/A_k)$ ($k \neq 0, N/2$). For $k = 0$ and $k = N/2$, $C_k = A_k$. Both θ_0 and $\theta_{N/2}$ are set to zero. The variables C_k and θ_k represent the amplitude and phase for this given value of k respectively. The power spectrum density for a given k is

$$P(k) = C_k^2/N, \quad (4.5)$$

where the phase θ_k is absent from the relation. If the time series has a power-law behavior (with exponent β) in the power spectrum, i.e. if $P(k) \propto k^{-\beta}$, then (4.5) can be used to obtain C_k for all k except $k = 0$. Higuchi has fixed C_0 to zero. Now, to construct simulated data having a power-law spectrum, the phase θ_k must be determined for a given k . For simplicity, Higuchi assumed that θ_k is given by a random variable with a uniform distribution in the range $[0, \theta_{max}]$, where θ_{max} can vary within the range $[0, 360^\circ]$. The correlation between the phases is set to be $\langle \theta_k \theta_{k'} \rangle = \delta_{kk'}$, where δ_{ij} denotes the Kronecker delta. Higuchi has considered the case $\beta = 2$ and varied the value of θ_{max} in the range $[0, 360^\circ]$. His results have shown that the resulting time series $\{x(j)\}$ becomes more and more irregular (fractal) as θ_{max} increases, but definitely that for small θ_{max} , the time series is not fractal. This shows that a quantitative analysis using only the power-law exponent β is inappropriate for describing the irregularity of a fractal time series: the phase distribution of the frequency components must also be random.

When β is in the interval $1 < \beta < 3$ and $\theta_{max} = 360^\circ$, the time series $\{x(j)\}$ obtained by the above procedure is a fractional Brownian motion (fBm) time series. Fractional Brownian motion has been introduced by Mandelbrot and Van Ness [62] as a generalization of the ordinary Brownian motion (corresponding to $\beta = 2$). When $\beta = 2$, x represents the positions of a particle undergoing a one-dimensional random

walk. Fractional Brownian motion (fBm) time series satisfy the relationship formalized by Mandelbrot and Van Ness

$$x(hj + j_0) - x(j_0) \stackrel{d}{=} h^H [x(j + j_0) - x(j_0)]$$

for any integer $h > 0$ and j_0 , where $\stackrel{d}{=}$ implies that both sides of the equation have the same distribution function. The exponent H is called the Hurst exponent ($0 < H < 1$) and it is related to the spectral exponent β through the relation $\beta = 2H + 1$. It is relevant to mention that self-affinity and not self-similarity applies to the fBm time series since in general the stochastic variable (x) does not have the dimension of time. However, since in the literature *fractal time series* is often used to define both fractal and self-affine time series, we will stick to this expression in this thesis. Now, a property which is particular to fBm is that it has long-run correlations, in the sense that past increments are correlated with future increments. This property can be illustrated more easily by shifting the data index (j), so that j now runs from $-N/2 + 1$ to $N/2$ (the data with $j = 0$ had previously $j = N/2$). Also, for convenience, the new $x(0)$ is set to zero and all the other $x(j)$ are shifted accordingly. Then, the correlation function of future increments $x(j) - x(0)$ with past increments $x(0) - x(-j)$ normalized by the variance of $x(j)$ can be expressed as [31]

$$C(j) = \frac{\langle -x(-j)x(j) \rangle}{\langle x(j)^2 \rangle} = 2^{2H-1} - 1.$$

As can be easily checked, $C(j)$ vanishes for all j in the case $H = 0.5$. On the other hand, $C(j)$ is non-zero independent of j when $H \neq 0.5$. This remarkable feature of fBm leads to persistence for $H > 0.5$ and to antipersistence for $H < 0.5$. Persistence means that if we have in the past a positive (or negative) increment, then we also have *on average* a positive (or negative) increment in the future. On the other hand, antipersistence signifies that an increasing (or decreasing) trend in the past implies a decreasing (or increasing) trend in the future.

The generation of fBm curves requires special methods. The two most popular methods are successive random addition [107] and spectral synthesis [88]. It is beyond the scope of this thesis to outline one or the other of these methods. We have generated fBm curves on figure 4.18 for three values of H by the successive random addition method (see [31] for a quick algorithm to achieve this task). We see from the figure that as H gets larger, then the fBm curves get smoother and vice versa. Indeed, the curve for $H = 0.1$ exhibits local noise of the same order of magnitude as the total excursions of the curve whereas the one for $H = 0.9$ displays rather clear trends with relatively little noise. The curve for $H = 0.5$ represents a crossover between these two behaviors. As a result, we see that exponent H characterizes the degree of roughness of a fBm curve.

At this point, we would like to show the relevance of the concepts introduced above for fractal times series for the evenly sampled time series of type 2. The stochastic variable of the time series of type 2, i.e. the average stress ($\bar{\sigma}$), is like the water level in hydrology. The accumulated differences from the mean water level is called the *volume accumulation* in that field. By analogy with hydrology, we can introduce the volume accumulation of the average stress in the model of Xu et al. Such a quantity is plotted on figure 4.19. The curve on this figure was obtained by a running sum of an evenly sampled time series of type 2 from which we originally removed the trend and the mean. A similar curve would be obtained with an unevenly sampled time series of type 2. (Note that, for convenience, we restrict the discussion to evenly sampled time series of type 2, but since from section 4.2 we know that the original time series of type 2 and their evenly sampled counterparts have a similar structure, then this restriction does not have serious implications). The curve on figure 4.19 is reminiscent of a fBm curve (for comparison, see figure 4.18), but this does not prove that it is actually a fBm curve! We have also obtained the power spectrum for the time series plotted on figure 4.19. We now outline the procedure that we have used to achieve this task.

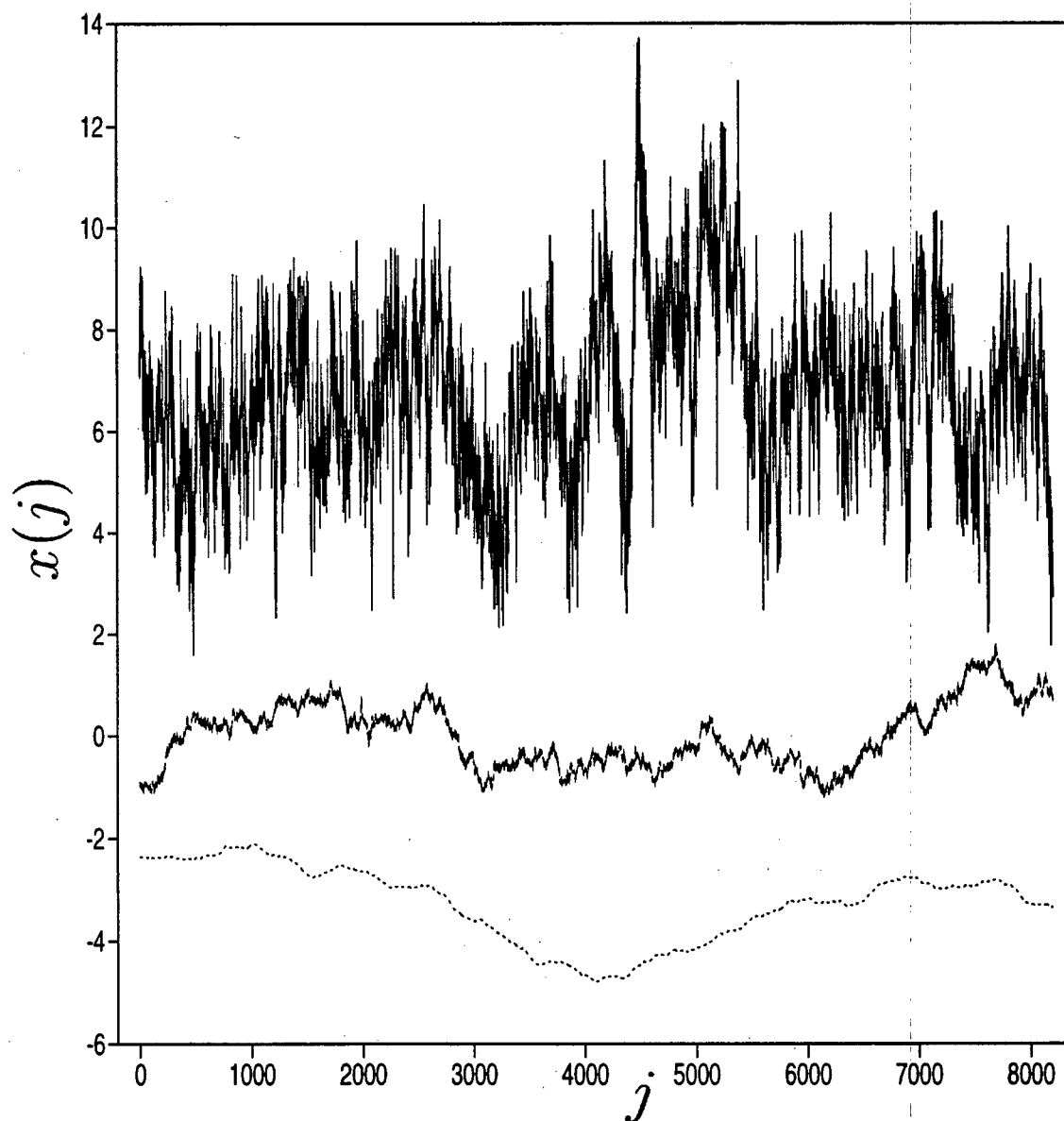


Figure 4.18: Fractional Brownian motion (fBm) signals having 8193 data generated by the successive random addition method. From top to bottom, the curves have $H = 0.1$, $H = 0.5$ and $H = 0.9$ and have been shifted for clarity.

The discrete Fourier transform of an evenly sampled time series (with sampling time interval Δt) $\{x(j)\}$ ($j = 0, 1, \dots, N - 1$) is defined as

$$X(k) = \sum_{j=0}^{N-1} x(j) e^{2\pi i j k / N}, \quad (4.6)$$

where $k = 0, 1, 2, \dots, N - 1$. The discrete Fourier transform can be easily obtained using the Fast-Fourier Transform (FFT) technique [85]. Afterwards, the periodogram estimate of the power spectrum can be obtained at $N/2 + 1$ frequencies using (see [85] for instance)

$$\begin{aligned} P(f_0) &= \frac{1}{N^2} |X(0)|^2 \\ P(f_k) &= \frac{1}{N^2} (|X(k)|^2 + |X(N - k)|^2) = \frac{2}{N^2} |X(k)|^2 \\ P(f_{N/2}) &= \frac{1}{N^2} |X(N/2)|^2, \end{aligned}$$

where the frequencies f_k ($k = 0, 1, 2, \dots, N/2$) are defined as

$$f_k = \frac{k}{N\Delta t}$$

and the time series is assumed real (i.e. not complex). To prevent the leakage of certain frequencies to adjacent frequencies, it is necessary to use a technique called *data windowing*. The window function that we use is of the form

$$w(j) = \frac{1}{2} \{1 - \cos[(j + 1/2)10\pi/N]\}$$

$$w(N - j - 1) = w(j)$$

where the values of j considered are the ones for which the argument of the cosine function is smaller than π (otherwise $w(j)$ is set to 1). In other words, the window function influences only the first 10% of the data at both ends of the time series. The way the window function enters into the calculation of the power spectrum is by replacing in (4.6) $x(j)$ by $w(j)x(j)$. We have obtained the power spectrum for the data set plotted

on figure 4.19 by the above calculation (see figure 4.20). We have not performed such a calculation for the running sum of an original time series of type 2 because it is not evenly sampled. The Lomb normalized periodogram (see [85] for instance) could be used to estimate the power spectrum in the latter case, but this computation has not been accomplished here. A power-law decrease of the power spectrum is observed on figure 4.20. However, there seems to have different β for different frequency intervals: we fitted, for $f_k \Delta t < 0.04$ and $f_k \Delta t \in [0.04, 0.3]$, $\beta = 2.94 \pm 0.05$ and $\beta = 3.46 \pm 0.03$ respectively. As a result, the running sum of the time series of type 2 has a more complicated power spectrum than a fBm time series. The change in slope on figure 4.20 is probably related to finite-size effects in time. We will return to these considerations below.

Despite the fact that figures 4.19 and 4.20 seem to suggest that the evenly sampled time series of type 2 are fractal time series, we know from section 4.2 that they are not pure fractal time series since they have a nonlinear component. However, they surely have an important fractal component because the power spectrum on figure 4.20 can be fitted to two power laws and from section 4.2, we know that they have only weakly coherent phase relationships between the frequency components. The presence of more than one component in a time series is commonly observed in nature. For example, while the human heartbeat intervals have been hypothesized to be fractal in nature (this was based solely on the observation of a power-law dependence in the power spectrum!), there are also well-defined harmonic oscillations due to the feed-back regulation of blood pressure and the external forcing by respiration [111]. Now, despite the fact that the phase relationships between the frequency components are not fully random in the evenly sampled time series of type 2, it is possible to use fractal time series analyses, even if in principle the latter are applicable only to pure fractal time series. This is what we want to do in this section by focusing on one fractal time series analysis, the rescaled range analysis. The goal of this analysis is to calculate the Hurst exponent H , which

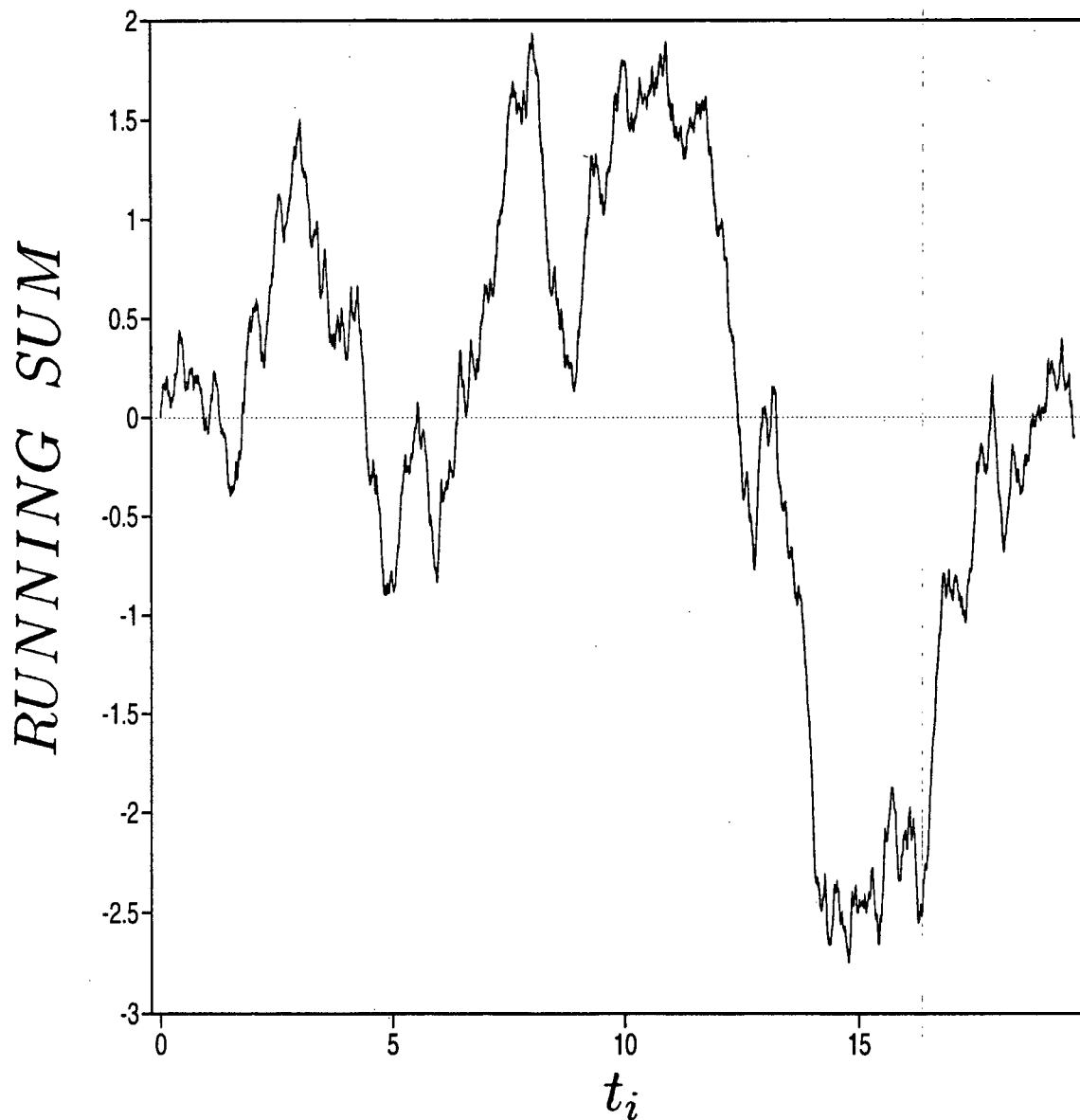


Figure 4.19: Running sum of an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} having $N = 16,384$ data. The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences. The trend and the mean have been removed from the time series before computing the running sum.

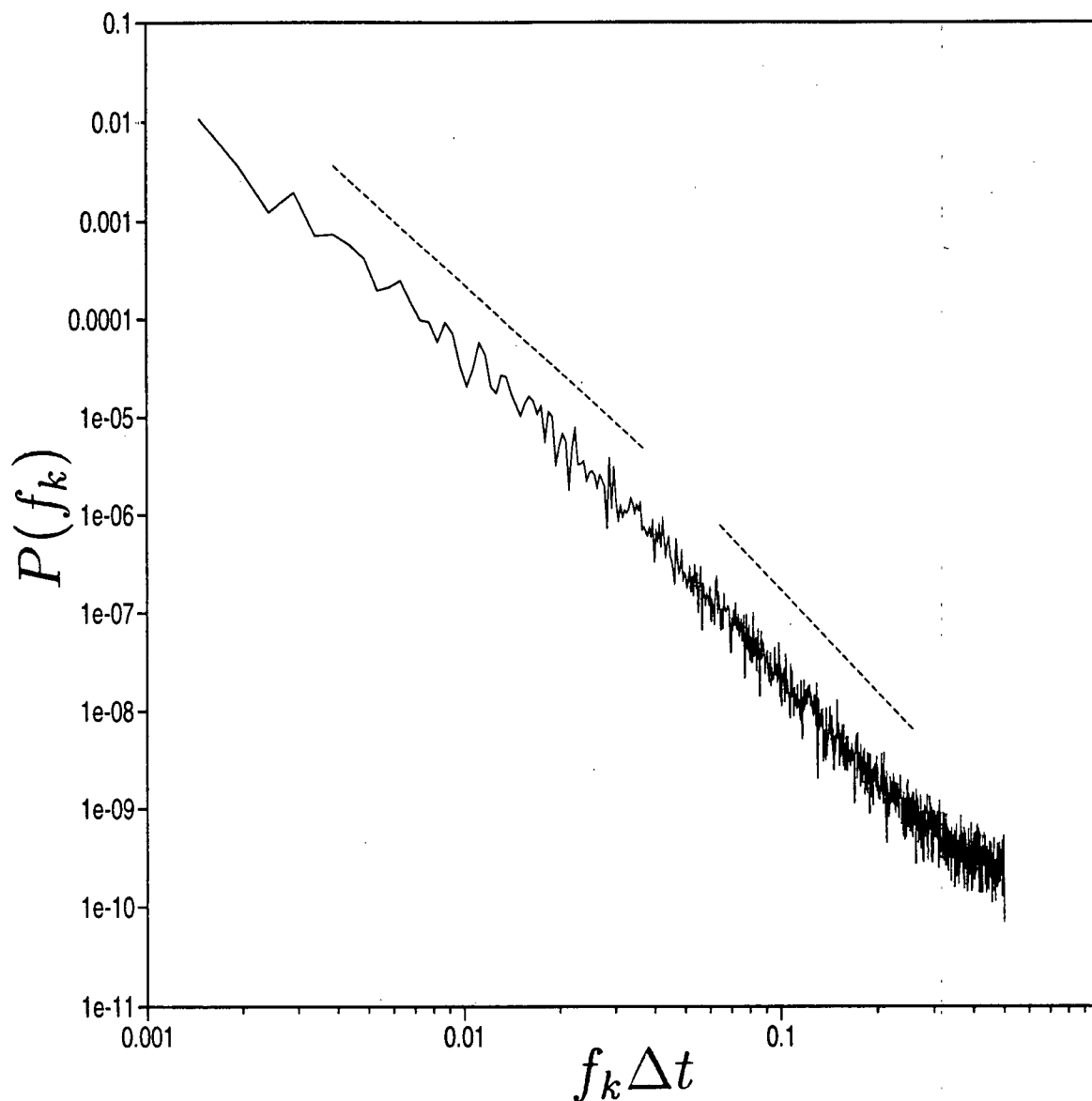


Figure 4.20: Periodogram estimate of the power spectrum of the running sum of an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval of the time series Δt is equal to 1.5 times the average time interval between the earthquake sequences, where the time series has 16,384 data. We have removed the trend and the mean before calculating the power spectrum. To reduce the fluctuations, we have computed an average power spectrum over 8 non-overlapping subsets (with 2048 data) of the time series resulting from the running sum operation. The dashed lines were obtained by doing linear regressions and have slope $\beta = 2.94 \pm 0.05$ (left segment) and $\beta = 3.46 \pm 0.03$ (right segment).

characterizes the time correlations in a time series. Note that the use of the rescaled range analysis in cases where the time series has a weak fractal component is unreliable. (This shows the danger of applying blindly the rescaled range analysis to any time series because the latter will always return a value of H whatever the importance of the fractal component). In cases like that, the only technique that we think could be used is the coarse-grained spectral analysis of Yamamoto and Hughson [115]. This technique extracts the fractal component from a time series and estimates what percentage of the power spectrum is due to the fractal component (this percentage is 100 for a fBm time series and zero for a pure harmonic signal).

4.3.2 Fractal time series analyses

Fractal time series analyses assume that the moments (such as the mean and variance) of the processes studied depend on the number of pieces included at a given resolution. Because the number of pieces depends on the resolution, these moments do not have fixed, limiting values. For example, when studying the fluctuations of a fractal (or self-affine) time series, the difference between the maximum and minimum values of the time series (i.e. the range) increases with time. Therefore, as more data are analyzed, the variance increases continually. A few fractal techniques are commonly used. Among them are the dispersional analysis and the rescaled range analysis.

The dispersional analysis involves the measurement of the variance of a signal at a succession of different levels of resolution [5]. This technique has been widely used in experimental physics (e.g. in molecular beam epitaxy) but also in studies related to medicine and physiology. It seems that the general physics community as well as the medical physics community have used different names to describe the same analysis. The dispersional analysis is considered as a relatively recent technique. On the other hand, the rescaled range analysis was introduced by Hurst [48] in the 1950's. Its use is still

mostly concentrated in the earth sciences [75], but it spreads slowly to other disciplines [18, 31]. The rescaled range analysis has benefited over the years from the works of Mandelbrot and Wallis [63, 64] and Bassingthwaite and Raymond [4] among others. Basically, this analysis studies how the range of cumulative deviations of the data from the mean (R) divided by the standard deviation of the data (S) (or the ratio R/S) varies with the size of the time window (or the number of data in the window M for an evenly sampled time series). For self-affine (or fractal) time series, R/S is a power-law function of M [$R/S \propto M^H$], where H is the Hurst exponent. The rescaled range analysis is the only fractal analysis that we consider in this thesis.

Before we describe the algorithm to implement the rescaled range analysis, we summarize the main ideas brought about by the pioneering work of Hurst [48]. For the Nile river, Hurst observed that records of levels at a gauge did not vary randomly, but showed series of low flow years and high flow years (i.e. *memory effects*!). Because the flow in a large river system such as the Nile depends on the water content in a large drainage area, the amount of water stored in the drainage area will increase in prolonged periods of higher than average precipitation and the excess amount of water stored will contribute to the discharge in drier years, and vice versa. Similar ideas were expressed later by Mandelbrot and Wallis [63]. The fBm concept is capable to take these memory effects into account. However, some authors [75] suggested that it not necessary to resort to a complicated concept such as fBm: autocorrelation (or a lack of serial independence) is generally accepted as a feasible explanation for the existence of memory effects.

4.3.3 Rescaled range analysis algorithm considered

The rescaled range algorithms that are usually found in the literature proceeds by dividing the analyzed time series into subseries of length M : for each subseries, $R(M)$ and $S(M)$ are computed. The manner in which the time series is divided into subseries, and the

range of values of M over which the calculation is performed, control the determination of the Hurst exponent H . The procedures advocated in [108] are flawed since the subseries overlap for any M , and as a result, the values of $R(M)/S(M)$ are not independent. The only way to achieve the independence of the calculated ratio values is to choose, for a given M , nonoverlapping subseries. This procedure was advocated in [64, 4] and used in [18]. It will be used in this work. For a given M , it is recommended in [4] to obtain an average value of R/S over the various subseries. These are the values that we will use in the fit to obtain H . Now, the fit of $R/S \propto M^H$ can be done either using an optimization procedure for parameterization of the nonlinear regression or a linear regression between $\log(R/S)$ and $\log(M)$. The former is theoretically more correct but it was shown in [4] that it provides values of H indistinguishable from the ones obtained by the latter for data sets not too small. For this reason, we will use the linear regression below to determine H . The algorithm that we implement is similar to the one in [18]. It does not include trend correction [4] for simplicity, but this choice should not be seen as limiting since trend correction does not always have a positive effect on the results as we have checked. Hurst did not consider trend correction in his original method [48].

The times series $\{x(j)\}$ ($j = 1, 2, \dots, N$) analyzed will be such that they have a number of data (N) which is an integral power of two. They are divided into non-overlapping subseries of M data, where M is an integral power of two ranging between 2 and $N/2$. The number of subseries of length M is $N(M) = N/M$ and is an integer whatever M . We compute the quantity $(R/S)_M$ for each M , where the latter is an average value of R/S over the non-overlapping subseries. We now outline the algorithm that we will use.

We start by calculating the mean, $\langle x \rangle_{n,M}$, of the n -th subseries of length M , which is

$$\langle x \rangle_{n,M} = \frac{1}{M} \sum_{j=(n-1)M+1}^{nM} x(j).$$

We also calculate the standard deviation $S_{n,M}$ of the n -th subseries of length M , which

is

$$S_{n,M} = \left[\left(\frac{1}{M} \right) \sum_{j=(n-1)M+1}^{nM} (x(j) - \langle x \rangle_{n,M})^2 \right]^{1/2},$$

where M and not $M - 1$ is used since the *population* standard deviation rather than the sample standard deviation is considered. This choice has only little effect on the results as we have checked. Our use of the population standard deviation is guided by the fact that in the literature, it is this quantity that is usually considered. In addition, we compute for all the allowed values of j the quantity

$$Y_{n,M}(j) = \sum_{k=(n-1)M+1}^j (x(k) - \langle x \rangle_{n,M}),$$

where $(n-1)M+1 \leq j \leq nM$. The range $R_{n,M}$ in the n -th subseries is obtained by subtracting the least value of $Y_{n,M}(j)$ from the greatest value of $Y_{n,M}(j)$. The rescaled range of the n -th subseries of length M is determined by the relation

$$(R/S)_{n,M} = (R_{n,M})/(S_{n,M}).$$

An average rescaled range, denoted by $(R/S)_M$, is then defined by

$$(R/S)_M = \left(\frac{1}{N(M)} \right) \sum_{n=1}^{N(M)} (R/S)_{n,M}$$

and computed for each allowed value of M . Afterwards, a straight line is fitted to the plot of $\log[(R/S)_M]$ versus $\log(M)$, the slope of which is precisely the Hurst exponent H .

We are now ready to apply this algorithm to the analysis of times series. We will do this first for time series having a known value of H . This preliminary study will enable us to test whether the above algorithm gives reliable results. In a second time, we will apply the algorithm to the analysis of the evenly sampled time series of type 2 introduced in section 4.1.

4.3.4 Tests of the algorithm with time series having a known value of H

The time series that we consider in this subsection were generated by the successive random addition method [107] using the algorithm presented by Feder [31]. As discussed by Churilla et al. [18], it is clear that the time series to be analyzed with the rescaled range algorithm are the time series of the increments, not the time series of the accumulated increments, where the former is obtained from the latter by a first-difference operation. This is because the rescaled range analysis computes the range of *cumulative deviations* of the data from the mean (R).

We have analyzed with the algorithm of subsection 4.3.3 synthetic time series with a known value of H having $N = 2^{13}$ data. The results are plotted on figure 4.21 for $H = 0.1$, $H = 0.5$ and $H = 0.9$. The curves are almost perfect straight lines, except the one for $H = 0.1$. Also, note that even at large M , the slope of the latter is still larger than 0.1, showing that the slope value 0.1 is reached only asymptotically. The method of generation of the time series is not responsible for the absence of a linear segment for the case $H = 0.1$ on figure 4.21 since we have obtained rather straight lines for the other two cases. It is now well documented (see [4] for instance) that the rescaled range analysis does not work well on time series having negative correlations (i.e. $H < 0.5$) and our results are simply in agreement with this fact. It is fortunate that the rescaled range analysis works fine when applied to signals having $H \geq 0.5$ because, in the real world, the latter are common while the signals with $H < 0.5$ are relatively rare. Now, we have fitted straight lines to the curves having $H \geq 0.5$ on figure 4.21 and found the values $H_{RR} \approx 0.55$ and $H_{RR} \approx 0.83$ for the curves for $H = 0.5$ and $H = 0.9$ respectively. The discrepancies between H_{RR} and H are of about +10% for $H = 0.5$ and -8% for $H = 0.9$. As a result, the rescaled range analysis is a biased method which overestimates H when $H \sim 0.5$, underestimates H when $H \sim 0.9$ and provides wrong estimates of H

when H is small. These findings are in agreement with the ones in [4] in particular. Note that these limitations of the rescaled range analysis are not so severe if one is interested only in the order of magnitude of H for an unknown signal though. The reliability of the estimate of H provided by the analysis can be checked for instance by shuffling the order of the data and reapplying the algorithm to the resulting time series. If largely different estimates of H are obtained for the original time series and the shuffled time series, then one can safely conclude that the value of H obtained for the former is representative of a real physical effect.

We can now safely apply the algorithm of subsection 4.3.3 to the analysis of the evenly sampled time series of type 2 since it works rather well on synthetic time series having a known value of H , at least if $H \geq 0.5$. Our hope is that the time correlations in the evenly sampled time series of type 2 are positive.

4.3.5 Results obtained with the time series of interest

Evenly sampled time series of type 2 for systems with both random σ_{th} and constant σ_{th} are analyzed in this subsection. As mentioned in subsection 4.3.1, a running sum of an evenly sampled time series of type 2 is reminiscent of a fBm time series, even if in fact we have found that it is not a pure fBm time series. The evenly sampled time series of type 2 are therefore suitable to be analyzed with the algorithm of subsection 4.3.3. However, since the phase relationships between the frequency components in these time series are not fully random (due to the nonlinear component), we will investigate if this has an effect on the results of the rescaled range analysis.

The results of the analysis for an evenly sampled time series of type 2 for a system with random σ_{th} having $N = 2^{14}$ data are plotted on figure 4.22. The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences. We got similar results with larger values of Δt but in these cases, the $(R/S)_M$ versus

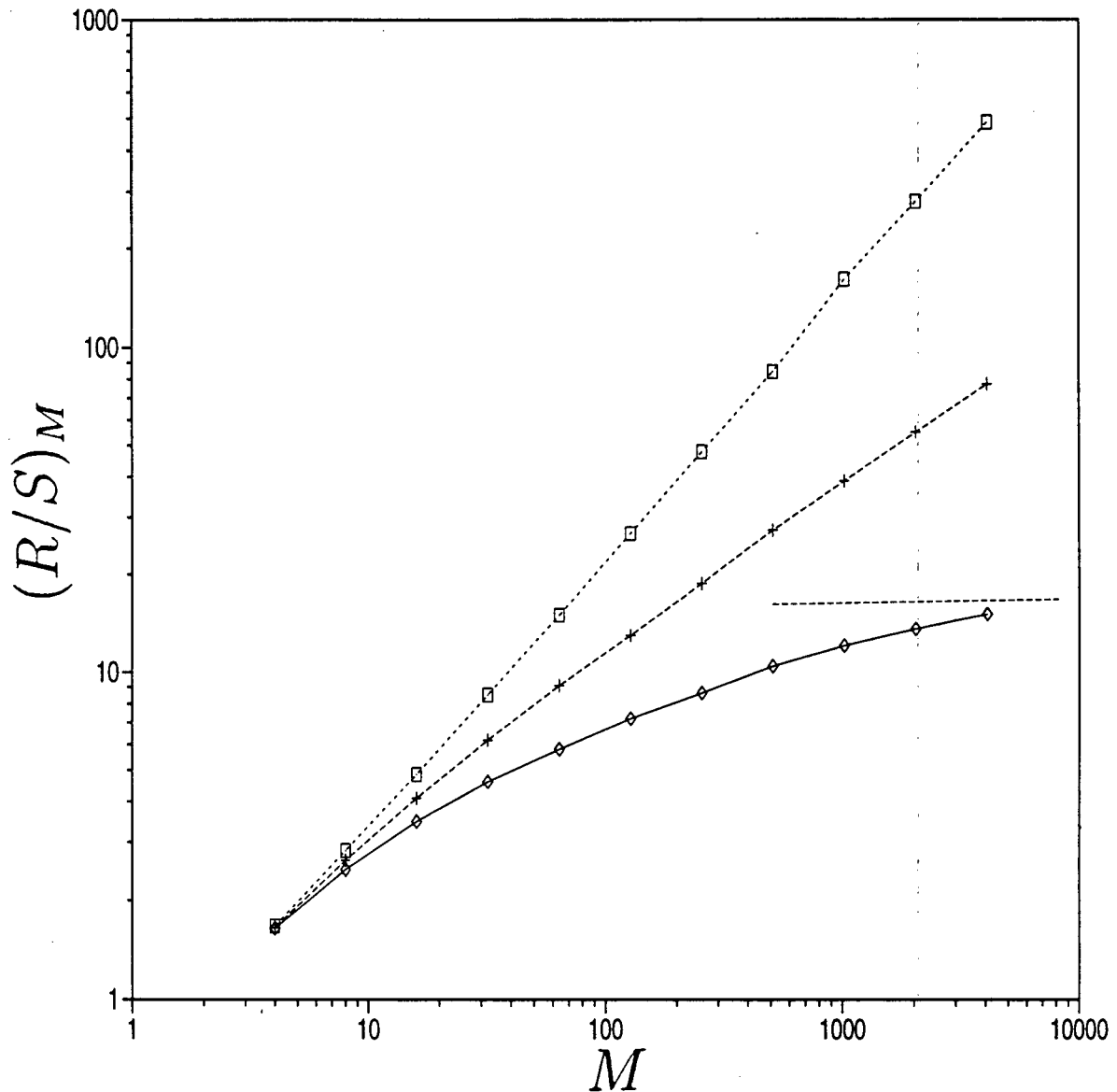


Figure 4.21: Log-log plot of $(R/S)_M$ versus M for three synthetic time series having $N = 2^{13}$ data generated by the successive random addition method. From top to bottom, the curves correspond to the cases $H = 0.9$, $H = 0.5$ and $H = 0.1$. The dashed line at the right of the figure has a slope of 0.1 and is there to illustrate that the curve for $H = 0.1$ will only slowly reach the slope value $H_{RR} = 0.1$ as M increases. Note that each curve was obtained by averaging $(R/S)_M$ (for a given M) over 5 realizations corresponding to the same value of H . For each curve, we have excluded the point corresponding to $M = 2$ since it is trivial and always has $(R/S)_{M=2} = 1$.

M curves cut-off at increasingly smaller M (results not shown) since there are less data. For $M = 2$, the ratio is always equal to 1, so we did not include this trivial point on figure 4.22. What we observe on the figure is a segment of a given slope for the smallest values of M and another one with a lower slope for the largest values of M . The slope for the segment $M \in [2^2, 2^7]$ is $H \approx 0.90$ (see figure 4.22). To check if the decrease of the slope at large M is not an artifact of the time series under consideration, we have also analyzed evenly sampled time series of type 2 for 50×50 and 25×25 systems with random σ_{th} . The $(R/S)_M$ versus M curves for these two time series as well as for the 100×100 time series (see figure 4.22) are plotted on figure 4.23, where the sampling time interval Δt is the same for all three time series. From this figure, we observe that, as the system size gets smaller, then the change in slope tends to occur at smaller M . As a result, the change in slope on figure 4.22 is a real effect, and represents some kind of finite-size effect in time. Note that since all three curves on figure 4.23 span the same M interval, then the change in slope can not be due to the finite number of data in the time series: the system size solely determines at what value of M it will take place. Now, the presence of positive correlations in the 100×100 system with random σ_{th} is real as we have checked by randomizing the order of the data. The curve obtained by averaging over 5 random shuffles is plotted on figure 4.22 (bottom curve). This curve is rather straight, with a slope equal to $H \approx 0.55$. Therefore, randomizing the order of the data has caused the destruction of the positive correlations that exist in the system. We have also found similar results for an evenly sampled time series of type 2 for a 100×100 system with constant σ_{th} (see figure 4.24). Again, we notice that the linear segment $M \in [2^2, 2^7]$ of the curve has a slope $H \approx 0.90$. There is still a change in slope due to finite-size effects. In addition, we noted that randomizing the order of the data has destroyed the time correlations (see the curve at the bottom of figure 4.24). Thus, we can say that the evenly sampled time series of type 2 for systems with both random σ_{th} and constant σ_{th}

have positive correlations that are affected by finite-size effects.

We have also analyzed evenly sampled time series of type 2 for one-dimensional systems and found results (results not shown) which are qualitatively similar to the ones obtained with two-dimensional systems. In addition, time series for which parameter α controlling the degree of stress conservation in the system (see chapter 3) is set to values smaller than 1 provide results again qualitatively similar to the ones obtained with time series for which $\alpha = 1$.

In section 4.2, we have conjectured that the evenly sampled time series of type 2 have weakly coherent phase relationships between their frequency components. The latter are responsible for the observation of a weak nonlinear stochastic structure. Now, we would like to investigate if these weakly coherent phase relationships have any effect on the $(R/S)_M$ versus M curves. This can be done by obtaining the $(R/S)_M$ versus M curves for a given time series and for a surrogate time series having the same power spectrum as the original time series but randomized phases (see the algorithm of Theiler et al. [102] for how to generate such surrogate data sets). We have performed such a calculation for a time series for a 100×100 system with random σ_{th} as well as for 5 surrogate time series of the above type. The results are plotted on figure 4.25 and clearly show that the weakly coherent phase relationships between the frequency components have a minor effect on the $(R/S)_M$ versus M curve. Therefore, this confirms that it was correct in the first place to analyze the evenly sampled time series of type 2 with the rescaled range analysis because the nonlinear component is sufficiently weak, so the fractal component of the time series is clearly dominant.

Finally, to conclude this section, we give our arguments to justify the results obtained with the evenly sampled time series of type 2. We believe that the existence of strong positive correlations is due to the nature of the driving in the model of Xu et al. because the amount of stress added to the system between the earthquake sequences depends

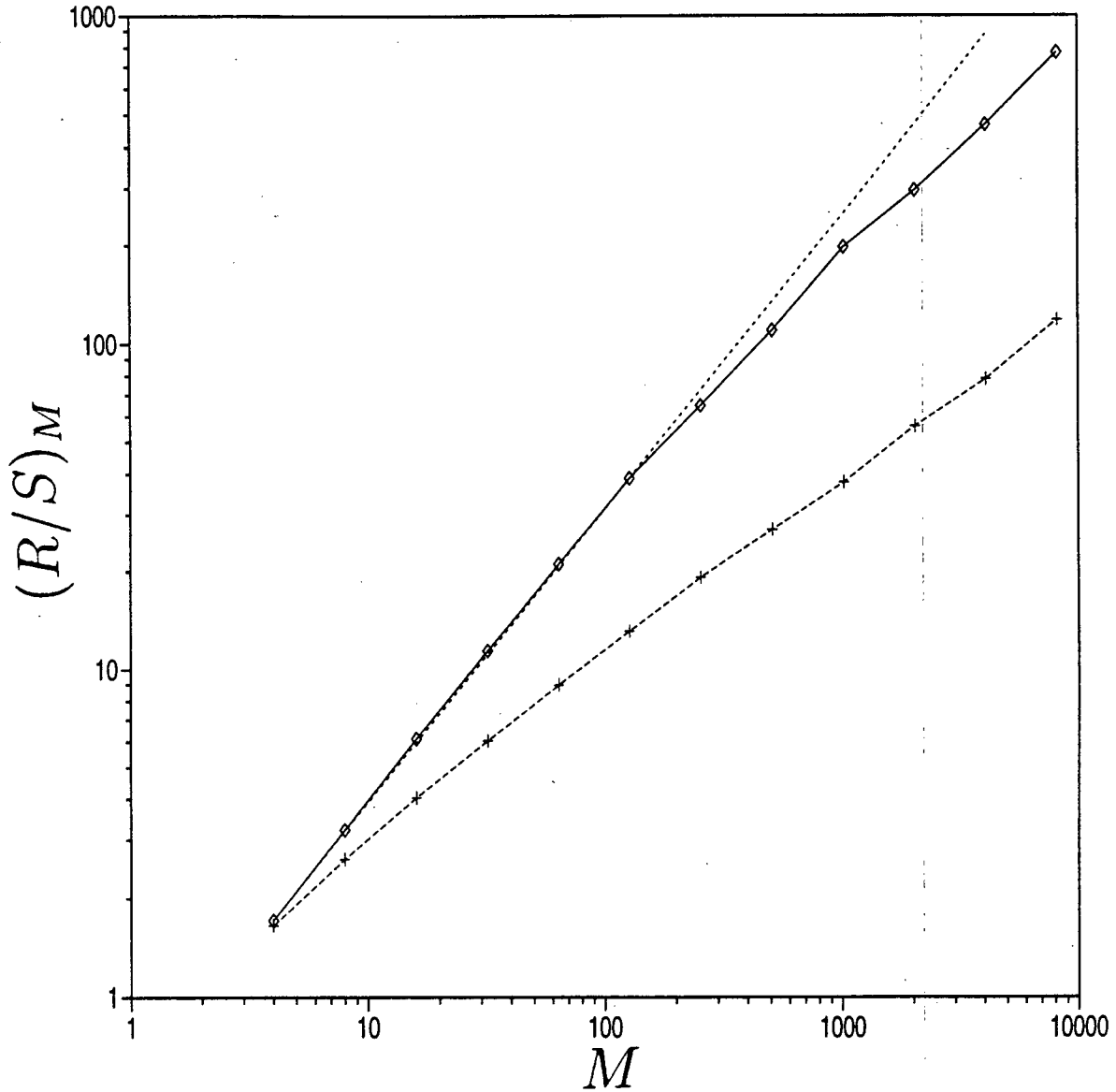


Figure 4.22: Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} having Δt equal to 1.5 times the average time interval between the earthquake sequences ($N = 2^{14}$). The curve corresponding to this time series is plotted using a solid line and lozenges. The dotted line ($H = 0.9$) is a fit of the linear segment $M \in [2^2, 2^7]$ of the latter curve. The curve at the bottom was obtained with 5 random shuffles of the time series: the curve of $(R/S)_M$ versus M was calculated for each realization and an average value of $(R/S)_M$ over the 5 realizations was computed for each M .

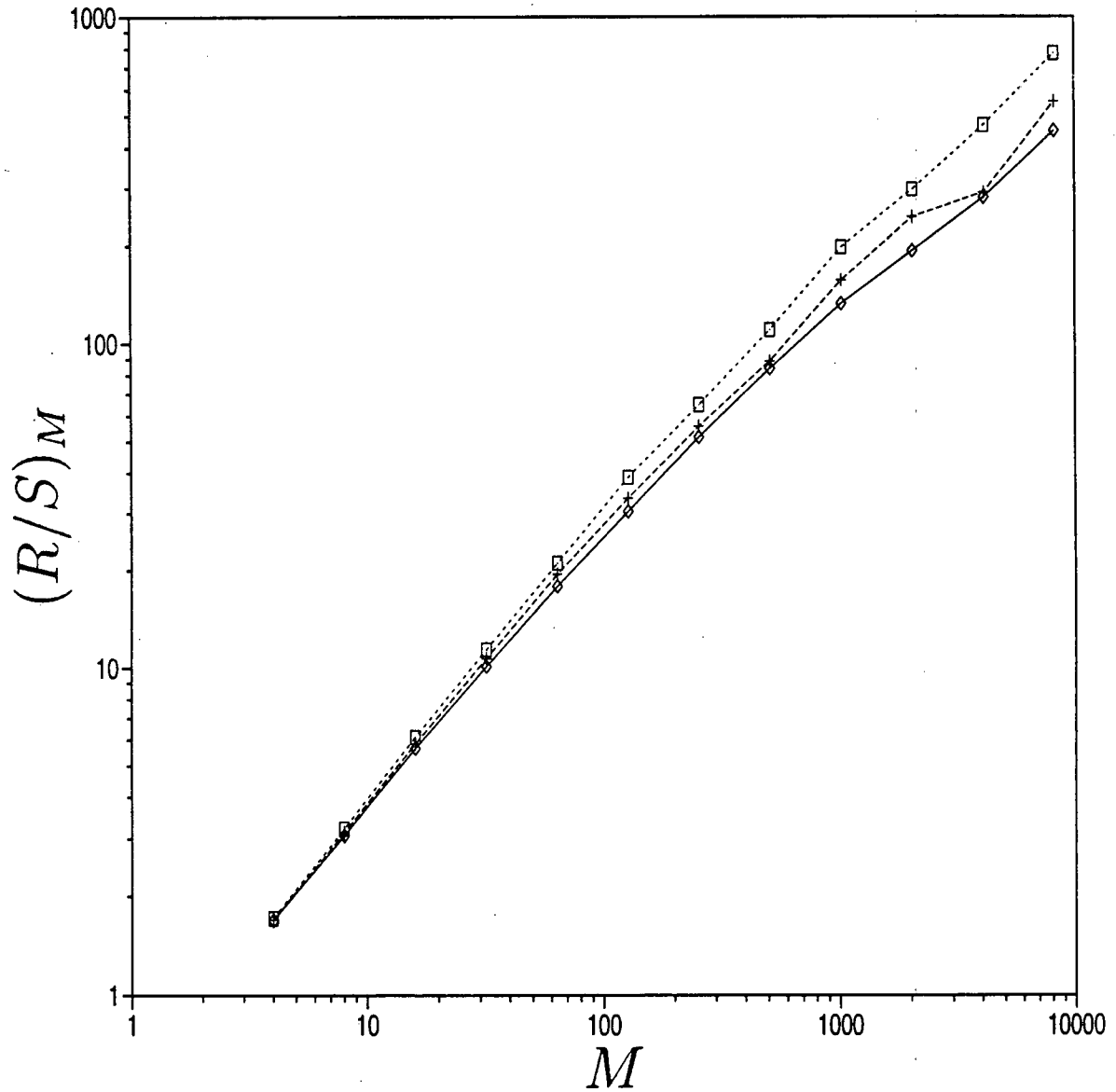


Figure 4.23: Log-log plot of $(R/S)_M$ versus M for evenly sampled time series of type 2 for $L \times L$ ($L = 25, 50, 100$) systems with random σ_{th} . The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences and the number of data is $N = 2^{14}$ for all three time series. From top to bottom, the curves refer to the 100×100 , 50×50 and 25×25 systems respectively.

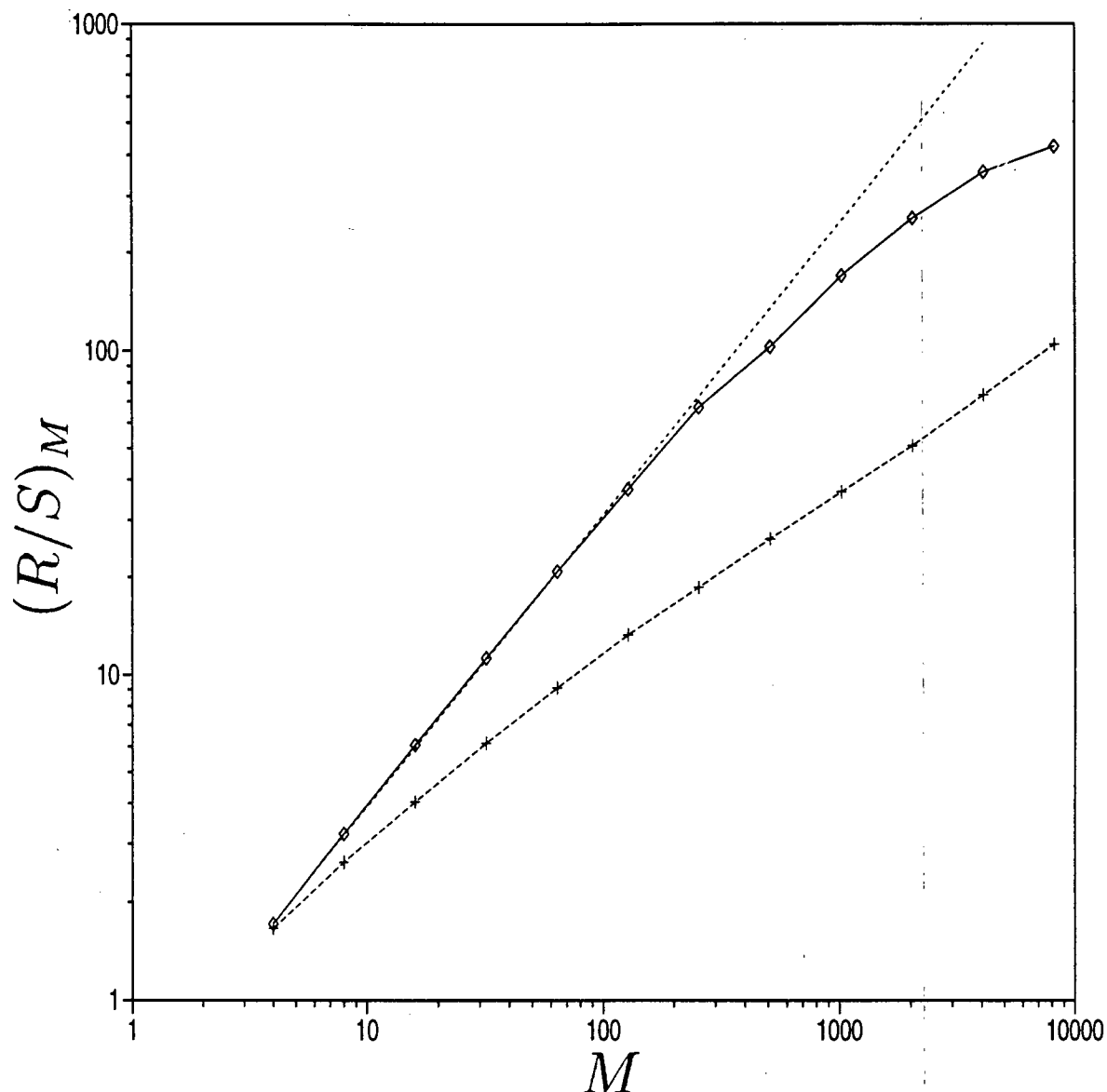


Figure 4.24: Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with constant σ_{th} having Δt equal to 1.5 times the average time interval between the earthquake sequences ($N = 2^{14}$). The curve corresponding to this time series is plotted using a solid line and lozenges. The dotted line is a fit ($H = 0.9$) of the linear segment $M \in [2^2, 2^7]$ of the latter curve. The curve at the bottom was obtained with 5 random shuffles of the time series: the curve of $(R/S)_M$ versus M was calculated for each realization and an average value of $(R/S)_M$ over the 5 realizations was computed for each M .

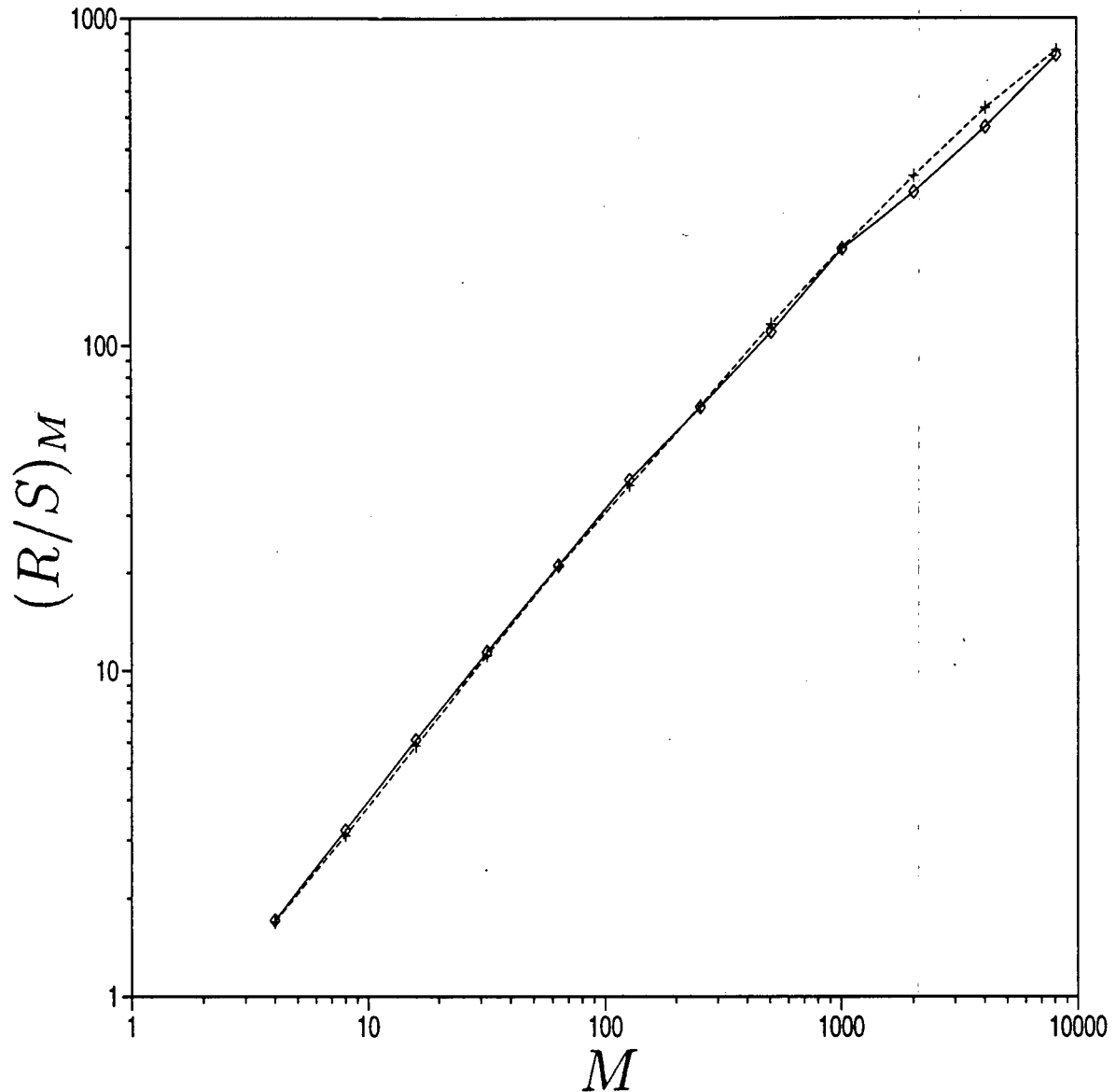


Figure 4.25: Log-log plot of $(R/S)_M$ versus M for an evenly sampled time series of type 2 for a 100×100 system with random σ_{th} . The sampling time interval Δt is equal to 1.5 times the average time interval between the earthquake sequences and the number of data is $N = 2^{14}$. The curve plotted using a solid line and lozenges was obtained with the original time series whereas the one plotted using a dashed line and pluses was obtained with 5 surrogate time series having the same power spectrum as the original time series but randomized phases. For the latter curve, $(R/S)_M$ was calculated for each realization, for a given M , and an average value of $(R/S)_M$ over the 5 realizations was computed for this value of M .

on the *past history* (see chapter 2). In addition, the correlations are positive ($H > 0.5$) and not negative ($H < 0.5$) because the system tends to accumulate stress until a large earthquake sequence occurs and releases a sizable fraction of that stress. This can be seen on figure 4.3 for instance where sometimes the average stress added to the system is much less than the average stress released from it, but this large stress release is due to the past accumulation of stress in the system (stress reservoir is full!). Now, we have also noticed that these positive correlations are affected by finite-size effects. We attempt here an explanation for the latter. We have seen in chapter 3 that the cumulative size-frequency distribution is affected by the finite size of the system in the sense that the power law for small earthquake sizes does not extend to large earthquake sizes (see figure 3.1 for instance). On the other hand, we have checked in section 4.2 that the histogram ($N(\delta_t)$) of the time intervals (δ_t) between large earthquakes is roughly exponential, i.e. proportional to $e^{-\delta_t/\delta_0}$, where δ_0 is a characteristic time interval and can be determined by a linear regression of $\log[N(\delta_t)]$ versus δ_t . We have performed such a linear regression for the large earthquakes ($S \geq 200$) generated in a 100×100 system with random σ_{th} and found the value $\delta_0 \approx 0.154$. We conjecture that this characteristic time interval between large earthquakes is responsible for the decrease of the slope at large M on figure 4.22. Indeed, the cut-off happens at $M \approx 128$ on this figure, which corresponds to a window of size $M\Delta t$. Therefore, using the value of Δt for the corresponding time series, we get $M\Delta t = 0.167 \sim \delta_0$. As a result, we have demonstrated that the finite-size effects affecting the $(R/S)_M$ versus M curves are due to a characteristic time interval between large earthquakes. These large earthquakes have the effect of resetting the system by destroying the memory effects in it. Since, as we have checked, δ_0 gets smaller as the system size decreases, then this explains why on figure 4.23 the change in slope occurs at increasingly smaller M as the system size gets smaller. Note in addition that if we would go to arbitrarily large M , then we would see that the slope past the cut-off point

approaches 0.5, the value in the absence of any memory effects.

4.4 Summary of the Important Results and Conclusions

In this chapter, we have performed two types of analysis of time series generated by the model of Xu et al. These time series are of types 1 and 2. The time series of type 1 involve the size of the earthquake sequences (S) as a function of the occurrence time. They exist in three varieties: varieties (a), (b) and (c) consist in the time intervals between consecutive earthquake sequences with $S = S_0$ (or $S \geq S_0$), the time intervals between consecutive earthquake sequences whatever S and the sizes of earthquake sequences about evenly sampled in time respectively. The time series of type 2 involve the average stress in the system as a function of the long time.

A nonlinear forecasting analysis was first implemented. It investigates whether the data in a scalar time series generated from a dynamical system exhibit low-dimensional chaotic behavior as opposed to high-dimensional (or stochastic) behavior. The analysis was performed by using an algorithm which constructs models for the structure in the time series with a variable smoothing parameter (k). The accuracy of the short-term forecasting error as a function of k reveals much about the underlying dynamics generating the time series. We have applied a nonlinear forecasting algorithm to the analysis of the time series of types 1 and 2. The time series of type 1, whatever the variety, were found to be weakly or not at all forecastable. The only time series of type 1 that have some degree of forecastability are the ones involving the time intervals between large earthquakes generated in one-dimensional systems. For these time series as well as those involving the time intervals between large earthquakes generated in two-dimensional systems, we have obtained histograms of the data and found that they have an exponential distribution. This is in contradiction with the finding that large earthquakes are temporally correlated

in the model of Xu et al. [16]. We argued that this is because the model of Xu et al. in the form used in this thesis is a series model. For all the time series of type 1, no nonlinear structure was found, but they all have a large stochastic component. For this reason, they are of little interest since they seem to be indistinguishable from white noise. The time series of type 2 were found to be forecastable, with a nonlinear but high-dimensional structure. For the evenly sampled time series of type 2, we have checked that the weakly coherent phase relationships between the Fourier components are responsible for the observation of this structure. We suggested that the appearance of nonlinear structure was due to the fact that the system stops following linearly the external drive (threshold dynamics!) when it releases stress through earthquakes. A direct consequence of these results is that self-organized criticality is inherently a high-dimensional phenomenon, as previously suspected by other authors. Note that this conclusion might be particular to the model studied in this thesis.

Fractal time series have power-law behavior ($f^{-\beta}$) in their power spectrum. This is however only a sufficient condition since, as shown by Higuchi [46], fractal time series must have random phase relationships between their frequency components. A common example of a fractal time series is provided by the record of the positions of a Brownian particle moving in a one-dimensional Euclidean space. Fractional Brownian motion (fBm) is a generalization of the ordinary Brownian motion exhibiting long-run time correlations. The latter are positive if $H > 0.5$, negative if $H < 0.5$ and zero if $H = 0.5$ (ordinary Brownian motion), where H is called the Hurst exponent and it characterizes the degree of roughness of the record of the motion. From the results of the nonlinear forecasting analysis, we know that the evenly sampled time series of type 2 are not pure fractal time series since they have weakly coherent (and not truly random) phase relationships between the frequency components, but at least they have an important fractal component. The determination of the Hurst exponent H characterizing the time

correlations in time series with an important fractal component can be accomplished by implementing a rescaled range analysis. This analysis consists in studying how the range of cumulative deviations of the data from the mean (R) divided by the standard deviation of the data (S) (or the ratio R/S) varies with the size of the time window (or the number of data M in the window for evenly sampled time series). The plot of $\log(R/S)$ versus $\log(M)$ is linear with a slope H for a pure fractal time series. The quantity we calculated for each M was $(R/S)_M$, which is an average value of R/S over non-overlapping subseries of M consecutive data. The values of M were chosen to be equally spaced on a logarithmic scale. The analysis of the evenly sampled time series of type 2 has led to the following results: the time series of type 2 have strong positive correlations ($H \sim 0.9$) which are affected by finite-size effects since H decreases for large time windows. The weakly coherent phase relationships between the frequency components were shown to have little impact on our results. The presence of strong time correlations in the time series of type 2 is due to the nature of the driving in the model because the amount of stress added to the system between the earthquake sequences depends on the past history. The correlations are positive (and not negative) since the system tends to accumulate stress until a large earthquake occurs and releases a sizable fraction of that stress. A characteristic time interval between large earthquakes is responsible for the appearance of finite-size effects in time. The large earthquakes cause the destruction of the memory effects in the system by resetting the latter.

Chapter 5

Conclusion

The research on complex phenomena has led recently to the introduction of the concept of Self-Organized Criticality (SOC) by Bak, Tang and Wiesenfeld [3]. This concept has been presented in order to understand the origin of the power laws observed in nature. According to SOC, certain driven systems consisting of a large number of elements evolve towards a critical state with no characteristic time or length scales. The critical state is reached without fine tuning of a control parameter, unlike for phase transitions in equilibrium statistical physics. A simplified sandpile model illustrates the ideas of SOC. In this model, sand is added grain by grain at a random site on the pile. This random driving process lasts until one site becomes unstable, i.e. when the local slope becomes larger than a threshold value. Then, grains are redistributed to the nearest neighbors of the unstable site and might themselves become unstable as well, and so on, leading to the creation of an avalanche. After a while, the pile eventually reaches a critical state, characterized by a critical slope, in which additional grains fall off the pile via avalanches of all sizes, distributed in size and lifetime according to a power law. According to Sornette [93], the class of phenomena described by SOC relies on the condition of a very slow driving of the systems and the existence of fast burst-like responses of them. Earthquakes are one of the first example of phenomena that comes to mind for which such a time scale separation is apparent. Our interest in the SOC concept has been oriented towards the understanding of earthquakes in this thesis.

The earthquake occurrence is related to the appearance of fractures as well as to

friction. In principle, both fracture and friction should be incorporated in an earthquake model. However, in practice, this is a formidable task and realistic situations have yet to be correctly modeled. The model that we have used throughout this thesis studies the creation of faults in a planar elastic region and was introduced by Xu et al. [113, 114]. This model has its validity at much larger length scales than the atomistic one, where the medium can be described by continuous fields. The model of Xu et al. involves the following:

- it focuses on a one-layer planar region which is discretized into $L_1 \times L_2$ squares,
- it assigns the stress tensor directly to each square,
- it solves the elastic equations for all the squares before and after the rupture of a given square.

In the model, a square ruptures when $\sigma > \sigma_{th}$, where σ is the stress on that square ($\sigma = \sigma_{xy}$ if there is only one shear mode of rupture, σ_{xy} being one component of the stress tensor) and σ_{th} the stress threshold, and this is followed by a double couple stress redistribution. We have applied the model of Xu et al. to the study of earthquakes by implementing the following:

- embed the region of interest in an infinite medium and neglect the effect of any activity originating outside the region (i.e. we use open boundary conditions),
- uniformly drive the region by adding stress amount $\Delta\sigma$ to all the squares (this is enough to cause the rupture of the square which has σ the closest to σ_{th}), where the surrounding medium provides the necessary external drive,
- increment the long time scale (defined by variable t) by the amount $\Delta\sigma$,

- once an earthquake sequence has begun, freeze t since earthquakes are instantaneous events on the long time scale,
- reset σ_{th} of a ruptured square to a new value (for a system with random σ_{th}) or set it to a constant value (for a system with constant σ_{th}),
- when more than one square break simultaneously, use the independent redistribution procedure,
- count the total number of broken squares (S) during the sequence to estimate the size of the earthquake.

The parameters of the model are:

- θ_l , which controls the width of the uniform (random) distribution of stress thresholds ($\sigma_{th} \in [\theta_l, 1.0]$),
- X , which fixes the fraction of stress retained by a ruptured square,
- α , which introduces stress nonconservation in the redistribution of stress ($\alpha = 1$ corresponds to the conservative case).

Once the model has reached the stationary state, we have checked that the Gutenberg-Richter power law $C(S) \propto S^{-\tau}$ is satisfied, where $C(S)$ is the cumulative fraction of earthquakes with S or more broken squares. However, this power law is affected by finite-size effects: there are the effects of the finite size of the system and of the finite size of the squares. To take these effects into account, we have performed a finite-size-scaling analysis. For simplicity, we have considered systems with $L \times 1$ squares (i.e. lattices arranged in a strip). The finite-size-scaling function that we have considered includes a factor which corrects for the finite-size effects affecting the power law:

$$C(S, L) = S^{-\tau} g[(S - 1)L^{-\nu} / f(X, \theta_l, \alpha)].$$

The results that we have obtained are, for $\alpha = 1$:

- for fixed parameter values, when L is allowed to vary, then the finite-size-scaling function works well over the whole range of S for systems with random σ_{th} and for $S > 50$ for systems with constant σ_{th} . For systems with constant σ_{th} , there is a crossover region for $S < 50$,
- for fixed L , when θ_l or X is allowed to vary, then the above function works fine for all S for systems with random σ_{th} and for $S > 50$ for systems with constant σ_{th} . These results imply that a system which has been discretized by a fine mesh and a small X (or a large θ_l) can be approximated by a coarser mesh using a larger X (or a smaller θ_l).

In all the cases, $\tau \approx 0.30$ gives the best results and is a robust value for $L \times 1$ systems. However, for $L \times L$ systems, a preliminary check has convinced us that $\tau \approx 0.4$ provides optimal results. This shows that, at least quantitatively, the results obtained with $L \times 1$ systems are particular to this lattice configuration. We have also considered a mean-field version of the model of Xu et al. in which all the elements are equally coupled with each other and discovered that the finite-size-scaling function with $\tau \approx 0.30$ does not work well. This shows that τ is also sensitive to the details of the model. Despite the sensitivity of τ to the lattice configuration and the details of the model, we can state that the model of Xu et al. displays SOC since the finite-size-scaling hypothesis works fine. This conclusion holds only in the conservative case ($\alpha = 1$). In the non-conservative case ($\alpha < 1$), the model is not self-organized critical since, as we have checked, a characteristic earthquake size $S_0 < L$ appears, S_0 being a function of α and L .

The quantity that we have considered in the finite-size-scaling analysis is the power law $C(S) \propto S^{-\tau}$, which is an average over time as well as over space of the activity in

the system. In the stationary state, this is a static distribution. We have also studied in this thesis the dynamic quantities of the model, such as, for instance, the size of the earthquake sequences (S) and the average stress in the system ($\bar{\sigma}$), $\bar{\sigma}$ referring to a spatial average over the squares of the system. The sets of values of S and $\bar{\sigma}$ over the long time (t) make up time series which we have called time series of types 1 and 2 respectively. The time series of type 1 exist in three varieties: varieties (a), (b) and (c) consist in the time intervals between consecutive earthquake sequences with $S = S_0$ (or $S \geq S_0$), the time intervals between consecutive earthquake sequences whatever S and the sizes of earthquake sequences about evenly sampled in time respectively. These time series have been analyzed using two different types of analysis.

A nonlinear forecasting analysis was first implemented. It investigates whether the data in a scalar time series generated from a dynamical system exhibit low-dimensional chaotic behavior as opposed to high-dimensional (or stochastic) behavior. That analysis was performed by using an algorithm which constructs models for the structure in the time series with a variable smoothing parameter (k). The accuracy of the short-term forecasting error as a function of k reveals much about the underlying dynamics generating the time series. We have applied a nonlinear forecasting algorithm to the analysis of the time series of both types. The time series of type 1, whatever the variety, were found to be weakly or not at all forecastable. The only time series of type 1 that have some degree of forecastability are the ones involving the time intervals between large earthquakes generated in $L \times 1$ systems. For all the time series of type 1, no nonlinear structure was found, but they have a large stochastic component. For this reason, they are of little interest since they seem to be indistinguishable from white noise. On the other hand, the time series of type 2 were found to be forecastable, with a nonlinear but high-dimensional structure. We suggested that the appearance of nonlinear structure in

these time series was due to the fact that the system stops following linearly the external drive (threshold dynamics!) when it releases stress through earthquakes. A direct consequence of these results is that SOC is inherently high-dimensional, at least in the model studied in this thesis. It might be possible to observe a low-dimensional behavior in other versions of the model, but this remains to be proven.

Since it is usually the magnitude of the earthquakes which is known in practical situations, the results obtained with the time series of type 1 seem to close all doors towards the prediction of earthquakes. However, it is relevant to note that the definition of S used in this thesis is not tied to the state variable (i.e. the stress) in the model. If we would have considered as a definition of S the difference between the average stress in the system before and after the earthquakes (i.e. $\bar{\sigma}_{bef} - \bar{\sigma}_{aft}$), then the results obtained with the time series of type 2 suggest that it might be possible to make some predictions on the latter quantity. On a related note, we mention that $\bar{\sigma}_{bef} - \bar{\sigma}_{aft}$ is of the same order of magnitude as the quantity called *stress drop* in the literature on seismology, the latter being related to the magnitude of the earthquakes (see [51] for instance).

The second type of time series analysis that we have performed is a rescaled range analysis. In principle, it is applicable only to fractal time series, which have random phase relationships between their frequency components as well as a power-law behavior in their power spectrum [46]. However, in practice, it is usually sufficient for the time series to have an important fractal component for the rescaled range analysis to apply. This is the case for the evenly sampled times series of type 2 which, in addition to a weak nonlinear component, have an important fractal component as we have checked. The rescaled range analysis determines an exponent, the Hurst exponent H ($0 < H < 1$), characterizing the time correlations in time series with an important fractal component (and in general in fractal time series). This analysis consists in studying how the range of cumulative deviations of the data from the mean (R) divided by the standard deviation of the data

(S) (or the ratio R/S) varies with the size of the time window (or the number of data M in the window for evenly sampled time series). The plot of $\log(R/S)$ versus $\log(M)$ is a straight line with a slope H for fractal time series. We have used an algorithm to achieve that task and found the following results for the evenly sampled time series of type 2: they have strong positive correlations ($H \approx 0.9 > 0.5$) which are affected by finite-size effects since H decreases for large time windows. The presence of strong time correlations in the time series of type 2 is due to the nature of the driving in the model because the amount of stress added to the system between the earthquake sequences depends on the past history. The correlations are positive and not negative (when $H < 0.5$) since the system tends to accumulate stress until a large earthquake occurs and releases a sizable fraction of that stress. We have checked that a characteristic time interval between large earthquakes causes the appearance of the finite-size effects in time. The large earthquakes are responsible for the destruction of the memory effects in the system because they reset the latter.

The work performed in this thesis has enabled us to better understand the earthquake phenomenon but is in no way exhaustive. In the future, we suggest considering more realistic versions of the model of Xu et al., for instance a parallel version of it. A parallel model treats the situations where more than one square rupture simultaneously in a self-consistent manner, by assuming that the rupture at one site is affected by the ruptures at the other sites. Chen et al. [16] have initiated that task, but much work remains to be done though. For instance, it would be interesting to see if improved forecastability is obtained with the time series of type 2 and if the memory effects are modified in this version of the model. An extra suggestion for the future is to focus on a definition of the size of the earthquakes (S) which is more closely tied to the dynamics than the one used in this thesis: $S = \bar{\sigma}_{bef} - \bar{\sigma}_{aft}$. In particular, it would be interesting to investigate, by a finite-size-scaling analysis, what is the effect on the power-law exponent τ of this choice

for S . This should be done for both the series and the parallel versions of the model.

Bibliography

- [1] P. Bak and K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).
- [2] P. Bak and C. Tang, J. Geophys. Res. B **94**, 15,635 (1989).
- [3] P. Bak, C. Tang and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
- [4] J. M. Bassingthwaite and G. R. Raymond, Ann. J. Biomed. Eng. **22**, 432 (1994).
- [5] J. M. Bassingthwaite and G. R. Raymond, Ann. J. Biomed. Eng. **23**, 491 (1995).
- [6] R. Bhagavatula, K. Chen and C. Jayaprakash, J. Phys. A:Math. Gen. **27**, L155 (1994).
- [7] S. Bottani, Phys. Rev. Lett. **74**, 4189 (1995).
- [8] G. E. P. Box and G. M. Jenkins, *Time Series Analysis Forecasting and Control* (Holden-Day, San Francisco, 1970).
- [9] W. A. Brock, W. D. Dechert and J. Scheinkman, *A Test for Independence Based on the Correlation Dimension* [Working Paper, Department of Economics, University of Wisconsin (Madison), University of Houston, University of Chicago, 1986].
- [10] J. Buck, Quart. Rev. Biol. **63**, 265 (1988).
- [11] R. Burridge and L. Knopoff, Bull. Seism. Soc. Am. **54**, 1875 (1964).
- [12] R. Burridge and L. Knopoff, Bull. Seism. Soc. Am. **57**, 341 (1967).
- [13] J. M. Carlson and J. S. Langer, Phys. Rev. Lett. **62**, 2632 (1989).
- [14] M. Casdagli, J. Roy. Statist. Soc. B **54**, 303 (1991).
- [15] K. Chen, P. Bak and S. P. Obukhov, Phys. Rev. A **43**, 625 (1991).
- [16] K. Chen, R. Bhagavatula and C. Jayaprakash, J. Phys. A:Math. Gen. **30**, 2297 (1997).
- [17] K. Christensen and Z. Olami, Phys. Rev. A **46**, 1829 (1992).
- [18] A. M. Churilla, W. A. Gottschalke, L. S. Liebovitch, L. Y. Selector, A. T. Todorov and S. Yeandle, Ann. J. Biomed. Eng. **24**, 99 (1996).

- [19] Á. Corral, C. J. Pérez, A. Díaz-Guilera and A. Arenas, Phys. Rev. Lett. **75**, 3697 (1995); Á. Corral, C. J. Pérez, A. Díaz-Guilera and A. Arenas, Phys. Rev. Lett. **74**, 118 (1995).
- [20] P. Cowie, C. Vanneste and D. Sornette, J. Geophys. Res. B **98**, 21,809 (1993).
- [21] J. P. Crutchfield, J. D. Farmer, N. H. Packard and R. S. Shaw, Sci. Am. **255** (No 6), 46 (1986).
- [22] J. P. Crutchfield and K. Kaneko in *Directions in Chaos*, volume I, ed H. Bai-Lin (World Scientific, Singapore, 1988).
- [23] J. P. Crutchfield and B. S. McNamara, Complex Syst. **1**, 417 (1987).
- [24] A. Díaz-Guilera, Europhys. Lett. **26**, 177 (1994).
- [25] F. Donzé, P. Mora and S.-A. Magnier, Geophys. J. Int. **116**, 46 (1994).
- [26] B. Drossel and F. Schwabl, Phys. Rev. Lett. **69**, 1629 (1992).
- [27] P. England and D. McKenzie, Geophys. J. Roy. Astr. Soc. **70**, 295 (1982).
- [28] E. Fama, Management Science **11**, 404 (1965).
- [29] J. D. Farmer and J. J. Sidorowich, Phys. Rev. Lett. **59**, 845 (1987).
- [30] J. D. Farmer and J. J. Sidorowich in *Evolution, Learning and Cognition*, ed Y. C. Lee (World Scientific, Singapore, 1988).
- [31] J. Feder, *Fractals* (Plenum Press, New York, 1988).
- [32] S. Feng and P. N. Sen, Phys. Rev. Lett. **52**, 216 (1984).
- [33] A. M. Fraser, Physica D **34**, 391 (1989).
- [34] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
- [35] A. Gabrielov, W. I. Newman and L. Knopoff, Phys. Rev. E **50**, 188 (1994).
- [36] L. Gil and D. Sornette, Phys. Rev. Lett. **76**, 3991 (1996).
- [37] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sum of Independent Random Variables* (Addison-Wesley, Reading, 1954).
- [38] P. Grassberger, Phys. Rev. E **49**, 2436 (1994).
- [39] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).

- [40] H. S. Greenside, A. Wolf, J. Swift and T. Pignataro, *Phys. Rev. A* **25**, 3453 (1982).
- [41] G. Grinstein, D.-H. Lee and S. Sachdev, *Phys. Rev. Lett.* **64**, 1927 (1990).
- [42] D. Groleau, B. Bergersen and H.-J. Xu, *J. Phys A: Math. Gen.* **30**, 3407 (1997).
- [43] B. Gutenberg and C. F. Richter, *Bull. Seism. Soc. Am.* **34**, 185 (1944).
- [44] M. Hénon, *Commun. Math. Phys.* **50**, 69 (1976).
- [45] H. J. Herrmann and S. Roux (eds) *Statistical Models for the Fracture of Disordered Media* (North-Holland, Amsterdam, 1990).
- [46] T. Higuchi, *Physica D* **46**, 254 (1990).
- [47] N. F. Hunter Jr. in *Nonlinear Modeling and Forecasting*, eds M. Casdagli and S. Eubank (Addison-Wesley, Redwood City, 1991).
- [48] H. E. Hurst, *Trans. Amer. Soc. Civ. Engrs.* **116**, 770 (1951).
- [49] T. Hwa and M. Kardar, *Phys. Rev. Lett.* **62**, 1813 (1989).
- [50] I. M. Jánosi and J. Kertész, *Physica A* **200**, 179 (1993).
- [51] H. Kanamori and D. L. Anderson, *Bull. Seism. Soc. Am.* **65**, 1073 (1975).
- [52] Y. Kantor and I. Webman, *Phys. Rev. Lett.* **52**, 1891 (1984).
- [53] J. D. Keeler and J. D. Farmer, *Physica D* **23**, 413 (1986).
- [54] M. Kendall and J. K. Ord, *Time Series* (Edward Arnold, Kent, 1990).
- [55] L. Knopoff in *Disorder and Fracture*, eds J. C. Charmet, S. Roux and E. Guyon (Plenum Press, New York, 1990).
- [56] L. Lam and V. Naroditsky (eds) in *Modeling Complex Phenomena*, Proceedings of the Third Woodward Conference, San Jose State University, April 12-13 1991 (Springer-Verlag, New York, 1992).
- [57] P. Lévy, *Théorie de l'Addition des Variables Aléatoires* (Gauthier-Villars, Paris, 1937-1954).
- [58] L. Lise and H. Jedtoft Jensen, *Phys. Rev. Lett.* **76**, 2326 (1996).
- [59] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944).

- [60] B. Malraison, P. Atten, P. Bergé and M. Dubois, *J. Phys. Lettre (France)* **44**, L897 (1983).
- [61] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).
- [62] B. B. Mandelbrot and J. W. Van Ness, *SIAM Rev.* **10**, 422 (1968).
- [63] B. B. Mandelbrot and J. R. Wallis, *Water Resours. Res.* **4**, 909 (1968).
- [64] B. B. Mandelbrot and J. R. Wallis, *Water Resours. Res.* **5**, 228-241; 242-259; 260-267; 321-340; 967-988 (1969).
- [65] R. M. May, *Nature* **261**, 459 (1976).
- [66] P. Meakin, G. Li, L. M. Sander, E. Louis and F. Guinea, *J. Phys. A: Math. Gen.* **22**, 1393 (1989).
- [67] A. A. Middleton and C. Tang, *Phys. Rev. Lett.* **74**, 742 (1995).
- [68] P. Miltenberger, D. Sornette and C. Vanneste, *Phys. Rev. Lett.* **71**, 3604 (1993).
- [69] R. E. Mirollo and S. H. Strogatz, *SIAM J. Appl. Math.* **50**, 1645 (1990).
- [70] E. W. Montroll and M. Shlesinger in *Nonequilibrium Phenomena II, From Stochastic to Hydrodynamics*, Studies in statistical mechanics XI, eds J. L. Lebowitz and E. W. Montroll (North-Holland, Amsterdam, 1984).
- [71] P. Mora and D. Place, *Int. J. Mod. Phys. C* **4**, 1059 (1993).
- [72] K. Nagel and E. Raschke, *Physica A* **182**, 519 (1992).
- [73] H. Nakanishi, *Phys. Rev. A* **43**, 6613 (1991).
- [74] W. I. Newman and A. M. Gabrielov, *Int. J. Fract.* **50**, 1 (1991).
- [75] C. P. North and D. I. Halliwell, *Math. Geol.* **26**, 531 (1994).
- [76] S. Obukhov in *Random Fluctuations and Pattern Growth*, eds H. E. Stanley and N. Ostrowsky (Kluwer, Dordrecht, 1988).
- [77] Z. Olami, H. J. S. Feder and K. Christensen, *Phys. Rev. Lett.* **68**, 177 (1991).
- [78] F. Omiri, *J. College Sci. Tokyo Imper. Univ.* **7**, 111 (1895).
- [79] A. R. Osborne and A. Provenzale, *Physica D* **35**, 357 (1989).
- [80] N. H. Packard, J. P. Crutchfield, J. D. Farmer and R. S. Shaw, *Phys. Rev. Lett.* **45**, 712 (1980).

- [81] V. Pareto, *Cours d'économie politique*. Reprinted as a volume of *Oeuvres Complètes* (Droz, Geneva, 1896-1965).
- [82] C. S. Peskin in *Mathematical Aspects of Heart Physiology* (Courant Institute of Mathematical Sciences, New York University, New York, 1975).
- [83] L. Pietronero, A. Vespignani and S. Zapperi, *Phys. Rev. Lett.* **72**, 1690 (1994).
- [84] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, Singapore, 1994).
- [85] W. H. Press, S. A. Teukolski, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1992).
- [86] S. Roux and E. Guyon, *J. Phys. Lettre (France)* **46**, L999 (1985).
- [87] T. Sauer, J. A. Yorke and M. Casdagli, *Embedology* (Technical Report, University of Maryland, 1990).
- [88] D. Saupe in *The Science of Fractal Images*, eds H. O. Peitgen and D. Saupe (Springer-Verlag, New York, 1988).
- [89] C. H. Scholz, *The Mechanics of Earthquakes and Faulting* (Cambridge University Press, Cambridge, 1990).
- [90] A. Sherman, J. Rinzel and J. Keizer, *Biophys. J.* **54**, 411 (1988).
- [91] R. F. Smalley, D. L. Turcotte and S. Solla, *J. Geophys. Res. B* **90**, 1894 (1985).
- [92] D. Sornette, *J. Phys. I (France)* **2**, 2065 (1992).
- [93] D. Sornette, *J. Phys. I (France)* **4**, 209 (1994).
- [94] D. Sornette, A. Johansen and I. Dornic, *J. Phys. I (France)* **5**, 325 (1995).
- [95] D. Sornette, P. Miltenberger and C. Vanneste, *Pageoph.* **142**, 491 (1994).
- [96] A. Sornette and D. Sornette, *Europhys. Lett.* **9**, 197 (1989).
- [97] J. A. Steketee, *Can. J. Phys.* **36**, 192 (1958); J. A. Steketee, *Can. J. Phys.* **36**, 1168 (1958).
- [98] K. Stokbro and D. K. Umberger in *Nonlinear Modeling and Forecasting*, eds M. Casdagli and S. Eubank (Addison-Wesley, Redwood City, 1991).
- [99] G. Sugihara and R. M. May, *Nature* **344**, 734 (1990).

- [100] H. Takayasu and H. Inaoka, *Phys. Rev. Lett.* **68**, 966 (1992).
- [101] F. Takens in *Dynamical Systems and Turbulence*, eds D. Rand and L.-S. Young (Springer-Verlag, Berlin, 1981).
- [102] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank and J. D. Farmer in *Non-linear Modeling and Forecasting*, eds M. Casdagli and S. Eubank (Addison-Wesley, Redwood City, 1991).
- [103] H. Tong, *Non-linear Time Series: A Dynamical System Approach* (Clarendon Press, Oxford, 1990).
- [104] H. Tong and K. Lim, *J. Roy. Statist. Soc. B* **42**, 245 (1980).
- [105] P. Tong and J. Rossetos, *Finite Element Method: Basic Technique and Implementation* (MIT Press, Cambridge, 1977).
- [106] J. P. Vilotte, M. Daignières and R. Madariaga, *J. Geophys. Res. B* **87**, 10,709 (1982).
- [107] R. F. Voss in *The Science of Fractal Images*, eds H. O. Peitgen and D. Saupe (Springer-Verlag, New York, 1988).
- [108] J. R. Wallis and N. C. Matalas, *Water Resours. Res.* **6**, 1583 (1971).
- [109] L. M. Ward in press in *Nonlinear Dynamics in Human Behavior*, ed W. Sulis (World Scientific, Singapore).
- [110] W. Weibull, *J. Appl. Mech.* **8**, 293 (1951).
- [111] B. J. West and A. L. Goldberger, *Am. Scientist* **75**, 354 (1987).
- [112] A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano, *Physica D* **16**, 285 (1985).
- [113] H.-J. Xu, B. Bergersen and K. Chen, *J. Phys. A: Math. Gen.* **25**, L1251 (1992).
- [114] H.-J. Xu, B. Bergersen and K. Chen, *J. Phys. I (France)* **3**, 2029 (1993).
- [115] Y. Yamamoto and R. L. Hughson, *Physica D* **68**, 250 (1993).
- [116] G. U. Yule, *Philos. Trans. Roy. Soc. London A* **226**, 267 (1927).

Appendix A

Lattice Green Function Simplification and Properties

Firstly, we want to illustrate how to simplify the calculation of the lattice Green function (G) introduced in the discretization scheme defined by Xu et al. and presented in chapter 2. This will be done by reducing the number of integrals in (2.6) from two to one. This will allow a much more accurate evaluation of the value of $G(\vec{r} - \vec{r}_0)$, especially at large $\vec{r} - \vec{r}_0$.

Secondly, we want to give some useful properties of the lattice Green function which are easy to prove.

A.1 Simplification of the Lattice Green Function

Referring to (2.6) and (2.8), the lattice Green function is expressed as

$$G(\vec{r} - \vec{r}_0) = \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \int_{-\pi}^{\pi} \frac{dk_y}{2\pi} \frac{\sin^2 k_x \sin^2 k_y}{(1 - \cos k_x \cos k_y)^2} e^{i\vec{k} \cdot (\vec{r} - \vec{r}_0)}, \quad (\text{A.1})$$

where $\vec{r} - \vec{r}_0 \equiv (X, Y)$, X and Y being two integer numbers proportional to the lattice constant a (defined as 1 in this work for simplicity) and which can be negative, zero or positive. If the lattice has $L_1 \times L_2$ squares, then $|X| = 0, 1, 2, \dots, L_1 - 1$ and $|Y| = 0, 1, 2, \dots, L_2 - 1$. We will see below that $G(X, Y) = G(-X, Y) = G(X, -Y) = G(-X, -Y)$ so it is sufficient to focus on $X \geq 0$ and $Y \geq 0$ throughout this appendix. Then, $\vec{k} \cdot (\vec{r} - \vec{r}_0) = k_x X + k_y Y$ and the exponential factor in (A.1) can be written as

$$e^{i\vec{k} \cdot (\vec{r} - \vec{r}_0)} = \cos(k_x X + k_y Y) + i \sin(k_x X + k_y Y). \quad (\text{A.2})$$

The right-hand side of (A.2) can be further developed using obvious trigonometric identities to give four terms. Out of these four terms, only one provides a non-zero contribution to $G(\vec{r} - \vec{r}_0)$ whatever the values of X and Y [$\cos(k_x X) \cos(k_y Y)$]. The other three terms are such that they render at least one of the integrals in (A.1) odd and as a result, their contributions are zero since the integration is done over a complete cycle. Keeping only the relevant term of (A.2), (A.1) becomes

$$G(X, Y) = \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \sin^2 k_x \cos(k_x X) \int_{-\pi}^{\pi} \frac{dk_y}{2\pi} \frac{\sin^2 k_y \cos(k_y Y)}{(1 - \cos k_x \cos k_y)^2}. \quad (\text{A.3})$$

Let us rewrite (A.3) as

$$G(X, Y) = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \sin^2 k_x \cos(k_x X) \left[\int_0^{\pi} dk_y \frac{\sin^2 k_y \cos(k_y Y)}{(1 - c \times \cos k_y)^2} \right], \quad (\text{A.4})$$

where $c \equiv \cos k_x$ is treated as a constant when the integration over k_y is done. Defining the integral over k_y as I , where

$$I = \int_0^{\pi} dk_y \frac{\sin^2 k_y \cos(k_y Y)}{(1 - c \times \cos k_y)^2}, \quad (\text{A.5})$$

(A.4) can now be written as

$$G(X, Y) = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \sin^2 k_x \cos(k_x X) \times I. \quad (\text{A.6})$$

The next step is to obtain an analytical expression for I for various values of Y . The simplest case is $Y = 0$. Setting $Y = 0$ in (A.5), we have:

$$I_{Y=0} = \int_0^{\pi} dk_y \left[\frac{\sin k_y}{1 - c \times \cos k_y} \right]^2.$$

Using a software such as Mathematica, an analytical expression for $I_{Y=0}$ can be easily derived:

$$I_{Y=0} = \frac{\pi(1 - \sqrt{1 - c^2})}{c^2 \sqrt{1 - c^2}}. \quad (\text{A.7})$$

For values of Y different from 0, I can be simplified further by using software Mathematica. We thus obtain, starting from (A.5),

$$I = \frac{1}{2c^2} \int_0^\pi dk_y \left\{ \frac{-2 \cos(k_y Y) + cY \cos[(1-Y)k_y] - cY \cos[(1+Y)k_y]}{-1 + c \times \cos k_y} \right\}. \quad (\text{A.8})$$

It is not possible to simplify I further without fixing $Y > 0$. For simplicity, we consider only the smallest possible values of $Y > 0$. Then, setting $Y = 1$ in (A.8), and using software Mathematica, we obtain, after some simplification of the resulting expression,

$$I_{Y=1} = \frac{\pi}{c^3} \left[-2 + \frac{2 - c^2}{\sqrt{1 - c^2}} \right]. \quad (\text{A.9})$$

Repeating this procedure for $Y = 2$, $Y = 3$ and $Y = 4$, we obtain respectively

$$I_{Y=2} = \frac{\pi}{c^2} \left[2 - \frac{6}{c^2} + \frac{6 - 5c^2}{c^2 \sqrt{1 - c^2}} \right] \quad (\text{A.10})$$

$$I_{Y=3} = \frac{\pi}{c^5} \left[2(-8 + 5c^2) + \frac{16 - 18c^2 + 3c^4}{\sqrt{1 - c^2}} \right] \quad (\text{A.11})$$

$$I_{Y=4} = \frac{\pi}{c^6} \left[4(-10 + 9c^2 - c^4) + \frac{40 - 56c^2 + 17c^4}{\sqrt{1 - c^2}} \right]. \quad (\text{A.12})$$

We could pursue this procedure for all the necessary values of Y but we stop here for brevity.

Now, we can substitute back into (A.6) the analytical expression of I for the given value of Y and try to cast the expression in a convenient form. Let us illustrate this in the case $Y = 0$. Using $c \equiv \cos k_x$ in (A.7) and substituting in (A.6), we have

$$G(X, 0) = \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \sin^2 k_x \cos(k_x X) \frac{1 - \sqrt{1 - \cos^2 k_x}}{\cos^2 k_x \sqrt{1 - \cos^2 k_x}}.$$

Using obvious trigonometric identities, we can simplify the integrand in the latter expression and cast the integral in the form

$$G(X, 0) = \int_{-\pi}^{\pi} \frac{dk_x}{2\pi} \frac{\sin^2 k_x \cos(k_x X) (1 - |\sin k_x|)}{\cos^2 k_x |\sin k_x|}. \quad (\text{A.13})$$

Now, the integrand of (A.13) is even in k_x and so, using the fact that $\sin k_x \geq 0$ for $k_x \in [0, \pi]$, we can get rid of the absolute values and simplify (A.13) further:

$$G(X, 0) = \int_0^\pi \frac{dk_x}{\pi} \frac{\sin k_x \cos(k_x X)}{1 + \sin k_x}. \quad (\text{A.14})$$

We can make the change of variables $k_x \rightarrow k_x' + \pi/2$ which leads to $\sin(k_x' + \pi/2) = \cos k_x'$ and $\cos[(k_x' + \pi/2)X] = \cos(k_x'X) \cos(\frac{\pi}{2}X) - \sin(k_x'X) \sin(\frac{\pi}{2}X)$. Out of the two terms in the latter identity, only the first one leads to an integrand even in k_x' (the bounds of the integral over k_x' are $-\pi/2$ and $\pi/2$) and so, to a non-zero contribution to $G(X, 0)$. As a result, (A.14) becomes

$$G(X, 0) = \pm \left(\frac{2}{\pi}\right) \int_0^{\pi/2} dk_x' \frac{\cos k_x' \cos(k_x'X)}{1 + \cos k_x'}, \quad (\text{A.15})$$

where we have used the fact that for an integer X , $\cos(\frac{\pi}{2}X) = 1$ for $X = 0, 4, 8, \dots$ [(+) sign in (A.15)], $\cos(\frac{\pi}{2}X) = -1$ for $X = 2, 6, 10, \dots$ [(-) sign in (A.15)], and $\cos(\frac{\pi}{2}X) = 0$ for X odd.

We can repeat a similar procedure for $G(X, 1)$, $G(X, 2)$, $G(X, 3)$ and $G(X, 4)$ using (A.9), (A.10), (A.11) and (A.12) respectively. For brevity, we just give the results:

$$\begin{aligned} G(X, 1) &= \pm \left(\frac{2}{\pi}\right) \int_0^{\pi/2} dk_x' \frac{(1 - \cos k_x') \sin(k_x'X)}{(1 + \cos k_x') \tan k_x'} \\ G(X, 2) &= \pm \left(\frac{2}{\pi}\right) \int_0^{\pi/2} dk_x' \frac{(1 - 2 \cos k_x') \cos k_x' \cos(k_x'X)}{(1 + \cos k_x')^2} \\ G(X, 3) &= \pm \left(\frac{2}{\pi}\right) \int_0^{\pi/2} dk_x' \frac{(1 - 3 \cos k_x')(1 - \cos k_x') \sin(k_x'X)}{(1 + \cos k_x')^2 \tan k_x'} \\ G(X, 4) &= \pm \left(\frac{2}{\pi}\right) \int_0^{\pi/2} dk_x' \frac{(1 - 4 \cos k_x')(1 - \cos k_x') \cos k_x' \cos(k_x'X)}{(1 + \cos k_x')^3} \end{aligned}$$

When Y is an odd integer, then the (+) sign applies to $X = 1, 5, 9, \dots$ and the (-) sign to $X = 3, 7, 11, \dots$ whereas when Y is an even integer, then the (+) sign applies to $X = 0, 4, 8, \dots$ and the (-) sign to $X = 2, 6, 10, \dots$

To summarize, let us give the expressions for $G(X, Y)$ when Y is fixed to a given value (odd or even integer) different from 0 (the case $Y = 0$ is special and then $G(X, 0)$ is given by (A.15)). It can be easily checked that the above results for $G(X, 1)$, $G(X, 2)$, $G(X, 3)$ and $G(X, 4)$ are special cases of the following general expressions:

- Y is an even integer different from 0

$$G(X, Y) = \pm \left(\frac{2}{\pi} \right) \int_0^{\pi/2} dk_x' \frac{(1 - Y \cos k_x')(1 - \cos k_x')^{\frac{Y}{2}-1} \cos k_x' \cos(k_x' X)}{(1 + \cos k_x')^{\frac{Y}{2}+1}}, \quad (\text{A.16})$$

where the (+) sign is for $X = 0, 4, 8, \dots$ and the (−) sign for $X = 2, 6, 10, \dots$

- Y is an odd integer

$$G(X, Y) = \pm \left(\frac{2}{\pi} \right) \int_0^{\pi/2} dk_x' \frac{(1 - Y \cos k_x')(1 - \cos k_x')^{\frac{Y-1}{2}} \sin(k_x' X)}{(1 + \cos k_x')^{\frac{Y+1}{2}} \tan k_x'} \quad (\text{A.17})$$

where the (+) sign is for $X = 1, 5, 9, \dots$ and the (−) sign for $X = 3, 7, 11, \dots$

These expressions clearly show that if $X + Y$ is an even integer, then $G(X, Y) \neq 0$ whereas if $X + Y$ is an odd integer, then $G(X, Y) = 0$. We will see in appendix B how to overcome this unphysical property of $G(X, Y)$. For now, let us describe some useful properties of $G(X, Y)$.

A.2 Properties of the Lattice Green Function

Using (A.3), it can be checked that the following property applies to $G(X, Y)$:

$$G(X, Y) = G(-X, Y).$$

The proof of such a property can be done by replacing in the expression for $G(-X, Y)$ k_x by $-k_x$ and reorganizing the integrand, which will turn out to be identical to the one of $G(X, Y)$. In a similar way, the following properties also apply:

$$G(X, Y) = G(X, -Y)$$

$$G(X, Y) = G(-X, -Y).$$

Another useful property of $G(X, Y)$ is

$$G(X, Y) = G(Y, X)$$

which can be proved by replacing in the expression for $G(Y, X)$ [using (A.3)] k_x by k_y and k_y by k_x . These changes of variables can be done since k_x and k_y are dummy variables. That property greatly simplifies the task of calculating $G(X, Y)$ for all X and Y . Indeed, we can then limit ourselves to the computation of $G(X, Y)$ for $X \geq Y$.

We have also checked that $G(X, Y)$ falls off as $1/r^2$ at large $r = \sqrt{X^2 + Y^2}$, where r is the distance from the rupture location to a given square of the lattice. The latter property can be useful when dealing with very large lattice sizes because the calculation of $G(X, Y)$ using (A.16) and (A.17) is sometimes not accurate at large r .

Finally, it is of interest to note that $G(X, Y)$ can be multiplied by a global constant without changing the redistributed stresses. This can be seen by referring to (2.8).

Appendix B

Coarse-Graining of the Lattice Green Function

As we have seen in appendix A, $G(X, Y) = 0$ for $X + Y$ equal to an odd integer and $G(X, Y) \neq 0$ for $X + Y$ equal to an even integer, where X and Y are assumed to be positive or zero but not negative. (The latter restriction is not severe since from the properties of $G(X, Y)$ outlined in appendix A, the signs of X and Y are irrelevant). This unphysical property of $G(X, Y)$ can be overcome by a *coarse-graining* procedure.

First, we illustrate this procedure on a one-dimensional strip having $2L_1$ squares. Such a lattice configuration is considered in chapter 3 for the study of the scaling properties of the model of Xu et al. We imagine that one of the elementary squares (called the source square) breaks. During the stress redistribution which follows, we can mentally combine the squares pairwise, so that there are now L_1 pairs of squares in the lattice. The lattice Green function associated with the pairs can be expressed as

$$\bar{G}^1(X, 0) = G(2X, 0),$$

where $2X$ is the horizontal distance between the left (right) source square to a given left (right) square (in unit of a), and $X = 0, 1, 2, \dots, L_1 - 1$ is the distance between the pairs in unit of $2a$ (a was fixed to 1 in chapter 2).

For completeness, it can be mentioned that in the case of a two-dimensional $(2L_1) \times (2L_2)$ lattice of squares having $a = 1$, we can, during the stress redistribution which follows the rupture of an elementary square (again called the source square), generalize the above procedure by mentally combining four of them such as to form a bigger

square with lattice constant $a = 2$. The lattice now has $L_1 \times L_2$ squares with $a = 2$. The most natural coarse-grained lattice Green function can be expressed in this case as

$$\begin{aligned}\bar{G}^2(X, Y) = & G(2X, 2Y) + \frac{1}{4}[G(2X + 1, 2Y + 1) + G(2X + 1, 2Y - 1) + \\ & G(2X - 1, 2Y + 1) + G(2X - 1, 2Y - 1)]\end{aligned}$$

where $X = 0, 1, 2, \dots, L_1 - 1$ and $Y = 0, 1, 2, \dots, L_2 - 1$ refer to distances between the squares with $a = 2$.

\bar{G}^1 and \bar{G}^2 will be used in (2.8) to replace G when the lattice of squares is one-dimensional and two-dimensional respectively.

Now, using the properties of the elementary (not coarse-grained) lattice Green function (G) outlined in appendix A, we can easily prove the following properties of the coarse-grained lattice Green function \bar{G} (it refers to both \bar{G}^1 and \bar{G}^2):

$$\bar{G}(X, Y) = \bar{G}(-X, Y) = \bar{G}(X, -Y) = \bar{G}(-X, -Y)$$

$$\bar{G}(X, Y) = \bar{G}(Y, X).$$

We have also checked that $\bar{G}(X, Y)$ falls off as $1/r^2$ at large $r = \sqrt{X^2 + Y^2}$. It is remarkable that the elementary lattice Green function (G) and the coarse-grained lattice Green function (\bar{G}) have similar properties. This gives us confidence that the coarse-graining procedure that we use does not alter G that much.