Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

THE PHYSICAL PROPERTIES OF AGENT-BASED SYSTEMS

by
Lorenzo Sabatelli

A Thesis submitted to
The University of Dublin
for the degree of

Doctor of Philosophy

DEPARTMENT OF PHYSICS
TRINITY COLLEGE
UNIVERSITY OF DUBLIN

April 2004

DECLARATION

This thesis has not been submitted as an exercise for a degree in any other University.

Except where otherwise stated, the work described has been carried out by the author alone.

I agree that the Library may lend or copy the thesis upon request. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgment.

Signature of the author .....

Lorenzo Sabatelli
April 2004

ACKNOWLEDGMENTS

Firstly, I wish to thank my supervisor, Peter Richmond, for conceiving and setting the research project that I have been working on and for his support and advice during the last three years.

Secondly, I would like to express my gratefulness to Dietrich Stauffer for the extraordinary and completely unselfish help he offered me while working on the consensus models.

I would also like to thank:

Stefano Sanvito for the very interesting conversations we had in the last two years and for his moral support;

Simon Cox, Charles Patterson and Stephan Hutzler for advicing me on many issues concerning my research work and the thesis writing;

Marcel Ausloos, Frank Schweitzer and Janusz Holyst for giving me the possibility of presenting my work in several international meetings;

Roy Asher, Martin Nolan and Hugh Linehan for supporting (financially and morally) my research activities in Hibernian Investment Managers.

I acknowledge support from the European Commission via Marie Curie Industrial Host Fellowship.

Lorenzo Sabatelli

SUMMARY

This thesis deals with the physical properties of agent-based systems. Agent models are powerful tools for understanding 'social' self-organized systems and may provide an explanation of features typically found in many of these systems.

For example, even in quite different societies, the personal wealth is power-law distributed with an exponent close to 1.5. In the Stock Market, the distribution of returns, volatility and market volumes are 'fat tailed'. Volatility and volumes are associated with long memory stochastic processes. The time elapsing between two consecutive transactions itself is a non trivial stochastic process, significantly different from the Poisson process we would observe for independent identically distributed time lags.

The relation between the macroscopic (e.g. the stock market statistical properties) and the microscopic (e.g. the interactions between individuals) features of socio-economic systems is generally non-trivial and the outcome of the system dynamics may often seem counter intuitive. A typical example is provided by the social dilemmas. Here individual decisions that make sense to individual agents can aggregate into outcomes in which everyone suffers.

The simplest way of thinking of an agent, in physical terms, is to draw a comparison with Brownian Motion. However, this approach does not account well for the statistical properties of financial market. A very important extension of the random walk theory is the 'Continuous Time Random Walk' theory, in which the distribution of the waiting times between two consecutive stochastic events is taken into account. We study the waiting time distributions of two data sets taken from the XIX century Irish stock market and from the late XX century world currency market.

Stochastic Dynamical equations based on Generalized Langevin and Generalized Lotka–Volterra equations have also been studied. The probability distributions of wealth obtained from the Generalized Lotka-

Volterra and from an exponential mapping of a simple peer pressure model have been shown to be very close. The results obtained are compatible with the empirical observations of the wealth distribution.

We next consider 'Consensus Models'. In particular, we modify the so-called Sznajd consensus model with synchronous updating introducing an individual agent memory and studying the phase-transition process from no-consensus to consensus as a function of the memory length. The effect of memory has significantly reduced frustration. We find that, in the Sznajd model, the introduction of a small amount of noise may increase the consensus within the system, when in the absence of noise consensus is not achieved. Allowing the noise to depend on the size of fluctuations in the system we obtain a simple model able to account for the statistical properties of financial market volumes.

Motivated by the fact that real social systems display a complex network-like structure, we finally study the so-called Deffuant consensus model on a scale-free network. We propose two different adaptive versions of the model based on the 'relative expectation' of individual agents. The first, leads agents to adjust their expectation to neighbouring environmental states. The second, allows individual agents to adjust their expectations to their past performances.

Unlike the non-adaptive case, the number of surviving opinions as a function of the initial number of the possible opinions presents a non-monotonic behaviour, with a maximum occurring when the number of initial possible opinions approximately equals the number of agents. We suggest that this effect may be due to the existence of a typical average opinion-distance. When the average opinion-distance is above this characteristic value, agents tend to adapt faster, when the average opinion-distance is below this characteristic value, agents have very similar opinions and may easily reach an agreement. When the average distance equals the characteristic value, opinions are too distant for the agents to find an agreement, but also too close to promote quick adaptation.

...to those always looking for...

TABLE OF CONTENTS

LIST OF FIGURES

Figure 5.4: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =301. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

Figure 5.5: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =1000. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

Figure 5.6: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =50. The memory length T=0, 2, 8, 100, 500, 1000. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

Figure 5.7: Variation with L of 1-M_c (one minus the absolute value of the difference in the initial probabilities for +1 and-1) for which in half of the cases a consensus was reached. That may be seen as the phase transition point from the state without consensus to the state with consensus. The estimated slope is --0.39 for T=0, -0.21 for T=2, - 0.11 for T=8. The error bars are of the same size as the point-markers.

Figure 5.8: Variation with T of the difference in the initial probabilities for which in half of the cases a consensus was reached. That may be seen as the phase transition point from the state without consensus to the state with consensus. The estimated slope is 0.46 for L=17 and 0.47 for L=50. The error bars are of the same size as the point-markers.

Figure 6.1: Networks. Above are displayed graphic examples of: a regular lattice, a small-world network and a random network. Small-world networks may be seen as a trade-off between the regular topological structure of a lattice and the random topological structure of a random graph.Below, we can appreciate the pictorial difference existing between Random and Scale-free graphs Moreover, in the random graph, the 5 most connected nodes are connected to 27% of all nodes. In the scale-free graph, the 5 most connected nodes are connected to 60% of all nodes.

Figure 6.2: Number of surviving opinion as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is a fraction (1, 0.2, 0.04 and 0.008) of the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N=1000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers

Figure 6.3: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is 0.008 times the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N=10, 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.4: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is 0.02 times the average local distance. The averages are performed over the direct neighbourhood .The number of agents is equal to N= 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.5: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is equal to the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N= 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.6: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=14. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.7: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=104. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.8: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive

Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=1004. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

Figure 6.9: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=10004. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

LIST OF TABLES

Abbreviations

pdf = probability density function.

ACF = Auto Covariance Function.

CTRW = Continuous Time Random Walk.

STDP = Surviving Time Probability Distribution.

M-L = Mittag-Leffler.

GLV = Generalized Lotka-Volterra.

GLVPP = Generalized Lotka-Volterra with Peer Pressure.

PP = Peer Pressure

REM = Relative Expectation Model.

CHAPTER 1

GENERAL INTRODUCTION

1.1 Complexity

To describe and understand nature it is frequently necessary to appeal to statistical or probabilistic concepts. In physics, that is the case when we deal with large numbers of particles or with systems that are composed of a large number of subsystems. In general one tries to relate the behaviour of the composite system to the properties of the particles or subsystems and their mutual interactions.

There are systems composed of groups of related units (subsystems), for which the degree and nature of the relationships is imperfectly known. Their overall emergent behavior is difficult to predict, even when the subsystem behavior is readily predictable. The time-scales of various subsystems may be very different. Behavior in the long-term and short-term may be markedly different and small changes in inputs or parameters may produce large changes in behavior. Systems with these features are defined as Complex. Examples of Complex Systems are: the atmosphere, turbulent fluids, biological structures and communities, which include human societies.

1.2 Studying communities

The interest of physicists in human-based communities has indeed roots that date back to 1936 when the physicist Majorana wrote a pioneering paper on the essential analogy between statistical laws in physics and in the social sciences [1]. This unorthodox point of view was considered of marginal interest until recently. In the early 1960s, Benoit Mandelbrot, while working at IBM, saw a correlation between the distribution of

incomes in economy and the distribution of cotton prices: both seemed to fluctuate randomly and could not be fitted to the normal distribution curve used as a standard model for plotting variation. Mandelbrot [2] began to analyze a century's worth of cotton price data and, instead of just concentrating on the large long term changes, he included the small scale fluctuations too. This gave Mandelbrot a determination to explore the phenomenon of scaling.

In spite of this, it was only from the 90's that physics community became progressively interested in modeling financial markets, urban development, wealth distributions and the diffusion of opinions and diseases. Dramatic events such as the New York stock market crash and the fast spread of HIV in the 80's brought into questions the existing quantitative models of human interaction and, within the last decade, triggered a series of new interdisciplinary studies. For instance:

- Mantegna and Stanley [3] were the first to publish an extensive study of the scaling properties of data taken from the stock market. They were able to show the existence of significant differences between the statistical distributions and the econometric models (like ARCH, GARCH, etc) used in finance and what is observed for real data.

- Barabasi and Albert [4][5] investigate the scale-free properties of real world networks and proposed a model based on a 'preferential attachment' of nodes mechanism. This was a milestone achievement, since the spreading of opinions and diseases in a society is highly dependent on the topological properties of the system.

- Farmer [6] proposed a market model based on trader ecology and, with Lillo, studied the properties of the impact-price function [7].

- Bouchaud [8] investigated the statistical foundations of existing market models and proposed some new insightful models. He also suggested an explanation based on grand-canonical minority games for the long-time-memory effects observed in market volume and volatility time series [9].

- Solomon proposed an agent-based generalization of Lotka-Volterra models, called Generalized Lotka Volterra (GLV) model, able to account for the power law distributions found in many self-organized social systems (e.g. wealth distribution, city population, etc.). The GLV models offer a 'mesoscopic' description of systems. Their actual connection with microscopic interactions is not obvious, though [10].

- Gopikrishnan and Plerou investigated the statistical properties of market volumes and volatility, identifying some quasi-universal features of these quantities [11] [12].

- Challet studied the Minority Game problem, inspired by the, so called, 'El Farol' bar problem [13].

- Ausloos investigated the statistical properties of exchange rate time series [14], the stock market crash precursors [15] and sought a more scientific approach to the technical analysis methodology used by financial practitioners [16].

- Stauffer studied several consensus models and their critical behaviour [17] [18]. In particular, he developed a consensus model originally introduced by Sznajd [19], showed how it may be used to simulate the vote distribution found for the Brazilian local elections; introduced a synchronously updated version of the same model and studied the phase transitions from a state in which the system does reach total consensus to

a state in which it does not. Nevertheless, quite unrealistically, in these simulations agents are thought to be without memory and random changes of opinion are not allowed. Stauffer also simulated a second consensus model, previously introduced by Deffuant [20], on a scale-free network [21] and proposed a new version of it based on discrete opinion values [22], as opposed to the continuous opinion values (within a certain range) used in the original model. In these first versions of the Deffuant model agents are not allowed to adapt to their local environment and to learn from the past.

- Mainardi and Scalas proposed an approach to the modeling of waiting time distributions in financial markets based on continuous time random walk and fractional Brownian motion [23]. They successfully tested their model using some financial assets (LIFFE bond future), but the general validity of their approach has not been proved yet.

- Richmond studied the Generalized Langevin Equations, Lotka-Volterra and Peer Pressure Models and showed how this models are related and how they may account for some important statistical features found in socio-economic systems [10] [56].

Other significant contributions came from: Galam [24][25], Helbing and Holyst [26], Marsili [27], Schweitzer [28], Iori [29] and many others.

Their work paved the way to a new promising scientific field, today known as socio-econo-physics.

1.3 Agent Modeling

Agent-based models are widely used to describe systems in which many units interact in a way that depends on some internal degrees of freedom. One may naively think of an agent as a particle that may be either attracted or repelled by a given external field. The attraction or repulsion depends on the value taken by an internal variable (which itself obeys a dynamics possibly determined by the external conditions).

A relevant example of agent model is given by the Minority Game model (see appendix), widely studied within the physics community in recent years.

Many (either natural or man-built) systems are best described by networks of adaptive agents. For instance, systems of high technological and intellectual importance such as the Internet, human societies, biological organisms, ecological systems and the electrical power supply network present complex web-like structures. The study of a network is usually performed through the analysis of quantities such as the: average path length connecting two nodes, the clustering coefficient and the connectivity distribution. Traditionally the study of complex networks has been the territory of graph theory. However empirical studies have shown that for many networks, traditional models do not apply. Many real world networks grow continuously and have a power law decaying connectivity distribution. Physicists have now begun to build models of real networks

1.4 The physical properties of agent based systems

1.4.1 The topic of this thesis

This thesis deals with systems made up of agents by using statistical physics tools (such as scaling, correlations, probability distributions, stochastic differential equations and Monte Carlo simulations). We aim at finding a balance between the level of details needed for capturing the internal degree of freedom and describing agent heterogeneity and adaptability, and the simplicity needed for the model to be handled using differential equations, cellular automata and similar tools. Such an approach allows the study of large self-organized social systems that cannot be mastered otherwise.

In Chapter 2 we provide a description of some outstanding quantitative features of socio-economic systems. Among them: the almost 'universal' characters of wealth distributions, the fat tailed distribution and the long memory of volumes and volatility fluctuations in financial markets, the scale-free nature of many real social networks, the brain-twister represented by social dilemmas.

In Chapter 3 we study the waiting time distributions of two data sets taken from the XIX century Irish stock market and from the late XX century world currency market and we test a model proposed by Scalas Raberto and Mainardi [23] and based on Continuous Time Random Walk and fractional Brownian motion results.

In Chapter 4 we show how stochastic dynamical equations such as Generalized Lotka –Volterra equations may be used in agent modelling. We also find a link between the standard Generalized Lotka-Volterra and a simple peer-pressure-based model. This provides also a better

understanding of the underlying microscopic interactions in the Generalized Lotka-Volterra models.

Furthermore in Chapter 5 we deal with the so called 'Consensus dynamics'. Under certain conditions, the 'state' of an agent may spread through the system and become the 'state' of the system. We will refer to models describing this spreading process as 'Consensus Models'. In particular, we consider the so called, Sznajd consensus model with synchronous updating. We show how individual agent memory and random fluctuations may increase the degree of consensus within the system (in a way that, at first glance, may look counter-intuitive). We describe how an interdependence mechanism between the probability of random fluctuations and the global consensus fluctuations may account for the observed market volume dynamics.

Finally, in Chapter 6 we outline the properties and some models of the complex networks observed in nature and we introduce two simple adaptive versions of the Deffaunt consensus model. In the first, individual agents adjust their expectation to their neighbouring environment state. In the second, individual agents adjust their expectations to their past performances (increasing their expectation when they do find someone they agree with, decreasing their expectations when they do not).

1.4.2 Originality of this thesis.

In this thesis I will show that:

1. The waiting time distribution for XIX Century Irish Stock Market can be described by a Mittag-Leffler function, obtained under the assumption of a power-law memory function.

2. Generalized Lotka-Volterra (GLV) and Generalized Lotka-Volterra models with Peer Pressure (GLVPP) lead to power law probability distributions that fit the distribution of personal wealth found in most economies.

3. In the Sznajd Consenus model we observe an apparent paradox: allowing agents to change their state (or opinion) through random fluctuations may result in an increase of the degree of consensus within the system. The Sznajd model dynamics may be used to reproduce the market volume distributions observed in many financial markets.

4. In an adaptive version of the Deffuant consensus model, we observe that the relative number of surviving opinions (number of surviving opinions S over number of agents N), as a function of the relative initial number of opinions (number of initial opinions Q over number of agents N), presents a non-monotonic behavior. A maximum occurs for the initial number of opinions approximately equal to the number of agents N. A higher degree of consensus is reached not only for Q smaller than N, but also for Q much larger than N.

CHAPTER 2

PHENOMENOLOGICAL DESCRIPTION OF SOCIO-ECONOMIC SYSTEMS

2.1 Introduction

Self-organized communities display astonishing features arising from the mutual interaction of individuals. Major progresses have been made in recent years in spotting the statistical regularities of these systems, especially of financial markets. For instance, the presence of power law probability distributions and memory are typical (almost universal) statistical features associated to these systems.

The role of topology and information incompleteness, at a microscopic level, is now better understood by the scientific community, but its full comprehension is still an open challenge.

2.2 Statistical quantities.

Let us now outline very briefly some quantities of interest in the study of Socio-Economic Systems.

- Cumulative Distribution Function (c.d.f.). It is a function giving the probability that the random variable X is less than or equal to x, for every value x.

- Probability Density Function (p.d.f.) The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

$$\Pr ob\left(a < X < b\right) = \int_a^b f(x)dx \tag{2.1}$$

More formally, the probability density function, $f(x)$, of a continuous random variable X is the derivative of the Cumulative Distribution Function $F(x)$.

If $f(x)$ is a probability density function then it must obey two conditions:

1. that the total probability for all possible values of the continuous random variable X is 1:

2. that the probability density function can never be negative: $f(x) > 0$ for all x.

- Central Limit Theorem (CLT) .The Central Limit Theorem states that whenever a random sample of size n is taken from any distribution with mean μ and variance σ, then the sample mean m will be approximately normally distributed with mean μ and variance σ/n. The larger the value of the sample size n, the better the approximation to the normal distribution.

- Auto-covariance (ACVF) and Autocorrelation Functions. The Autocovariance is a measure of the linear dependence between members of a time series of observations, such as weekly share prices or interest rates, and the same values at a fixed time interval later. Autocorrelation is obtained by dividing the Autocovariance by the Variance of the time series. Autocorrelation occurs when

residual error terms from observations of the same variable at different times are correlated (related).

- Leptokurtosis. Leptokurtic is an adjective describing a distribution with high kurtosis. 'High' means the fourth central moment (kurtosis) is more than three times the second central moment; such a distribution has greater kurtosis than a Normal distribution. This term is used in Bollerslev-Hodrick 1992 to characterize stock price returns. Lepto- means 'slim' in Greek and refers to the central part of the distribution.

- Waiting Time Probability Distribution (WTPD). The WTDP is the probability Distribution Function for the time lag between two consecutive events in a time series. For instance, in Chapter 3 we will study the WTPD for the time interval elapsing between two consecutive transactions in Financial Markets.

- Survival Time Probability Distribution (STPD). The STDP gives the probability for a certain event to last longer than a given time interval. In the Financial Market, the STDP tells us how what is the probability that no transaction will occur during a certain time interval.

- Autoregressive (AR) model. It denotes a stochastic process that can be described by a weighted sum of its previous values and a white noise error.

- ARCH models. ARCH stands for Autoregressive Conditional Heteroskedasticity. It is a technique used in finance to model asset price volatility over time. It is observed in much time series data on asset prices that there are periods when variance is high and periods where variance is low. The ARCH econometric model for this (introduced by Engle (1982)) assumes that the variance of the series itself is an AR (autoregressive) time series, often a linear

one. An ARCH model is a discrete time stochastic process of the form:

$$\xi(t) = \eta(t) * s(t) \tag{2.2}$$

where the $\xi(t)$ are independent identically distributed stochastic variables over time, with zero mean and variance equal to 1. $\eta(t)$ is a Standard Normal Variable (i.e. a Gaussian distributed variable with zero mean and variance equal to 1). $s(t)$ is a positive and time-varying quantity, usually obtained as Auto Regressive process from the past values of the time series variance

$$s^2(t) = \alpha_0 + \alpha_1 \sigma^2(t-1) + ... + \alpha_n \sigma^2(t-n). \tag{2.3}$$

Usually $s(t)$ is further modeled to be an autoregressive process. According to Andersen and Bollerslev 1995/6/7, "ARCH models are usually estimated by maximum likelihood techniques." They almost always give a leptokurtic distrbution of asset returns even if one assumes that each period's returns are normal, because the variance is not the same each period. Even ARCH models, however, do not usually generate enough kurtosis in equity returns to match U.S. stock data. Also, their scaling behaviour deviates from what observed in real markets [1]

- Error and Uncertainty

For phenomenological data errors arise from record mistakes and rounding up.

For Montecarlo simulated data, errors have a statistical origin. We perform N simulations (keeping all the deterministic parameters unchanged) and we take the arithmetic average as the most likely value and (a quantity proportional to) the standard deviation, as

error (this way of operating is justified by the Central Limit Theorem).

The Relative Error (or uncertainty) is the estimated amount (or percentage) by which an observed or calculated value may differ from the true value. An estimation of uncertainty is given by the ratio standard deviation/ average value (assuming that the average value is different from zero)

- Statistical Averages.

The vast majority of socio-economic data are 'single realizations'. For instance, if we consider a certain financial index, we may know its time evolution, but we cannot know what that evolution would have been during the same period of time, under different initial conditions. As a consequence, usually we can only perform averages over data taken over a certain time. We will use the symbol $<\ldots>$ to designate time averages.

2.3 Distribution of wealth

Personal wealth is similarly distributed in most human societies. Geographical and political differences seem to play very little role in determining the proportion of people owning a certain fraction of the total wealth. Figure 2.1 shows the wealth and income cumulative distributions found for the United Kingdom. They display a power law asymptotic behaviour with an exponent of 1.85. What is typically found for most of the countries is an exponent fluctuating around 1.5 [30]. This feature is, to an extent, captured by the so called '80/20' rule introduced by the Italian economist Vilfredo Pareto [30]. The '80/20' rule states that the 20% of population owns the 80% of the wealth. This is an empirical evidence, quite true for the United States and other industrialized countries (for others deviations from the rule are found).

Understanding the origin and the ubiquity of these distributions is still an open problem. Significant steps towards shedding light on it have been made in recent years, though.

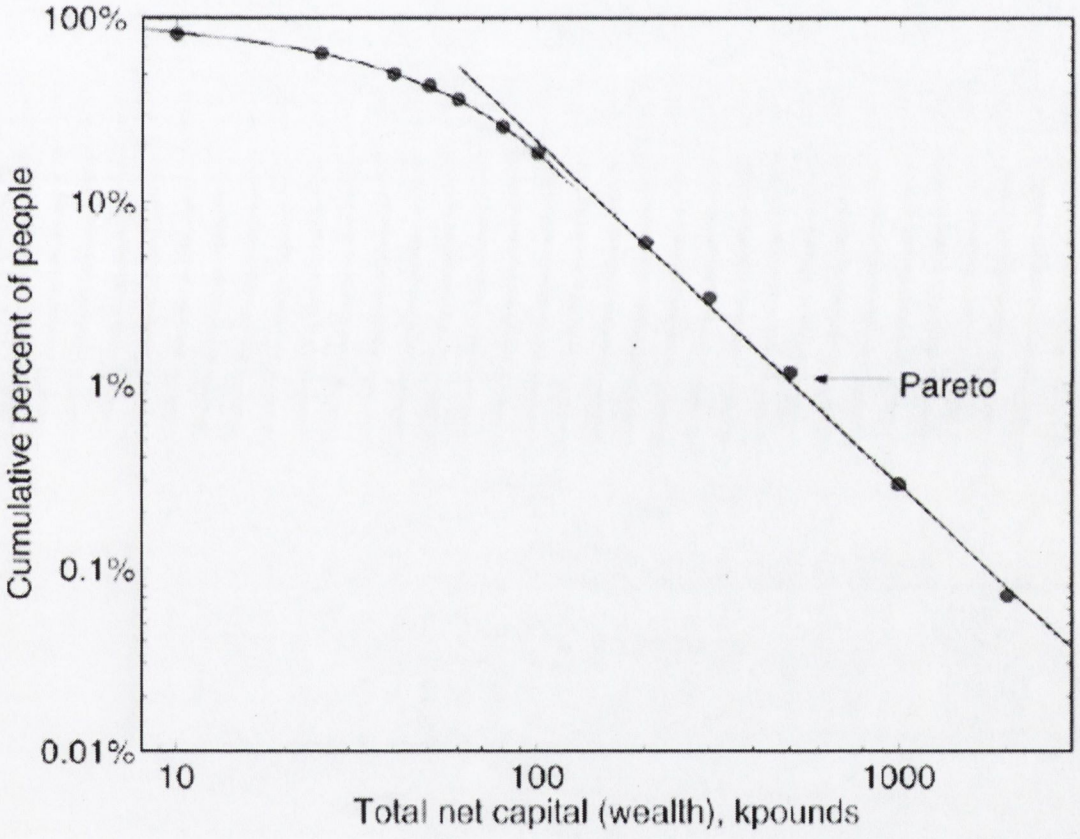

Figure 2.1: Cumulative wealth distributions in the United Kingdom in 1996 (courtesy of Dragulescu & al.; Physica A 299 213-221 (2001)).

2.3 Financial Markets

2.3.1 Introduction

A typical example of a Complex System, consisting of many agents mutually interacting is the financial market.

Tabulated in figure 2.2 is the value of the S&P500 from 1960 to 2000. The first problem one has to face when trying to analyze such a curve is the choice of reference units, both for time and value. A natural choice for the time would be the physical time (what one can measure using a clock) or the trading time (i.e. the physical time elapsing during open market hours. Actually in foreign currency exchange market that coincides with the physical time, i.e. the clock-time). In both cases the time unit (1 year, 1 month, 1 week, 1 day, 1 hour, 20, 10, 5, 1 min...) is to be chosen. Alternatively, one can take tick data (that record transactions when they actually occur), in this case the time lags between two consecutive transactions are not constant and are generated themselves by a (non trivial) stochastic process. The choice depends on the analysis we are carrying on and on its aims (as well as on the quality of the available data set).

Once dealt with the time unit, we have to decide which quantity we want to study. In general we are interested in quantities measuring the change of the price $P(t)$, over a certain time interval Δ. For the purposes of our work we will define the returns as:

$$R_{\Delta}(t) = \frac{P(t+\Delta) - P(t)}{P(t)} \tag{2.4}$$

and logarithmic returns:

$$S_{\Delta}(t) = \ln P(t+\Delta) - \ln P(t) = \ln\left(\frac{P(t+\Delta)}{P(t)}\right) \tag{2.5}$$

These two quantities are basically the same for price differences smaller than 1% (see Figure 2.3 and 2.4) and both do not depend on the price scale.

2.3.2Correlation Functions

A number of authors have now published results for the linear time correlation (see also appendix 2) function:

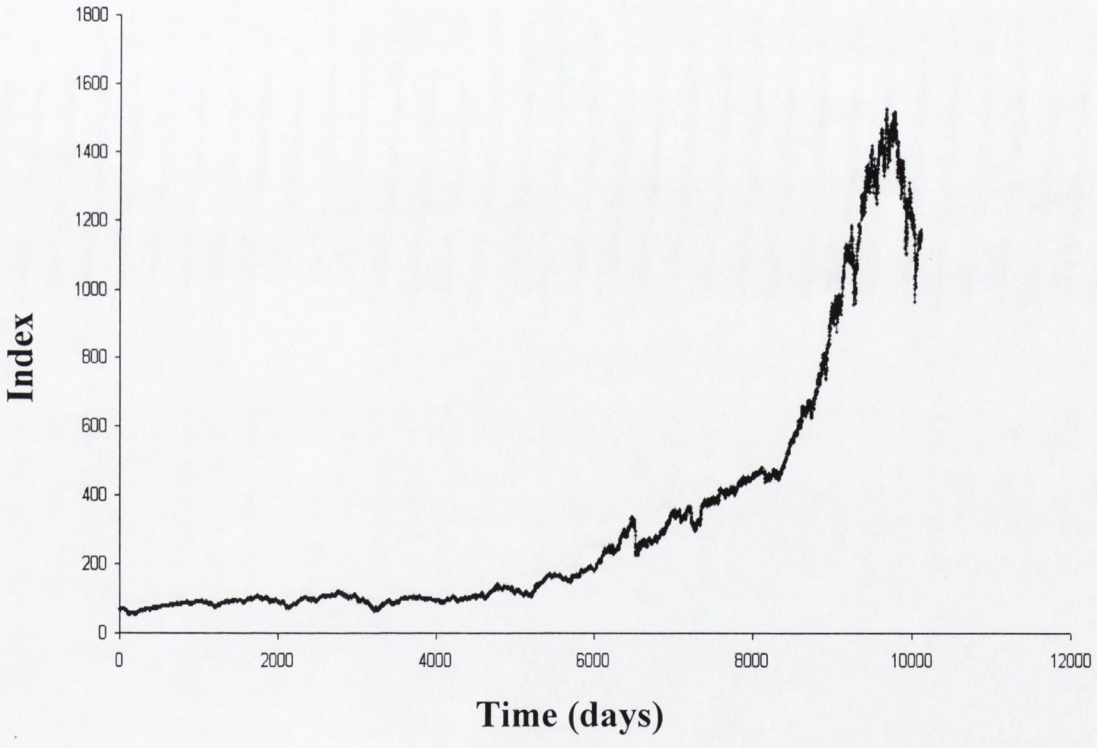

Figure 2.2: Index Standard and Poor 500, daily closing value data taken from 1960 to 2001.

$$C(\tau) = \frac{\langle S_\Delta(t+\Delta) S_\Delta(t) \rangle - \langle S_\Delta(t) \rangle^2}{\langle S_\Delta(t)^2 \rangle - \langle S_\Delta(t) \rangle^2} \tag{2.6}$$

The result for daily data of S&P500, taken over the period 1960-2000, is tabulated by Figure 2.5.

There is almost no auto-correlation between returns. This feature is typical of daily data and is a necessary (but not necessarily sufficient) condition for assuming that stock fluctuations are independent distributed variable. However with the availability of minute by minute data, new features of this function can be seen.

Things are different with high frequency data. In fact, one finds that returns separated by a time lag smaller than 10-30 minutes are significantly positively correlated and the autocorrelation function decays exponentially. After that time the correlation is very small and oscillates around zero (as shown by Gopikrishnan & al.) [3].

2.3.3 Probability Distribution Functions

The Probability Distribution Function (p.d.f.) of returns does not belong to the "gaussian world", at least when returns are taken over time intervals smaller than months. As shown by Mantegna and Stanley [3], the scaling behaviour of the p.d.f. of returns is remarkably different from a gaussian distribution up to ten thousand minutes. The frequency distribution of daily log returns is shown in Figure 2.6.

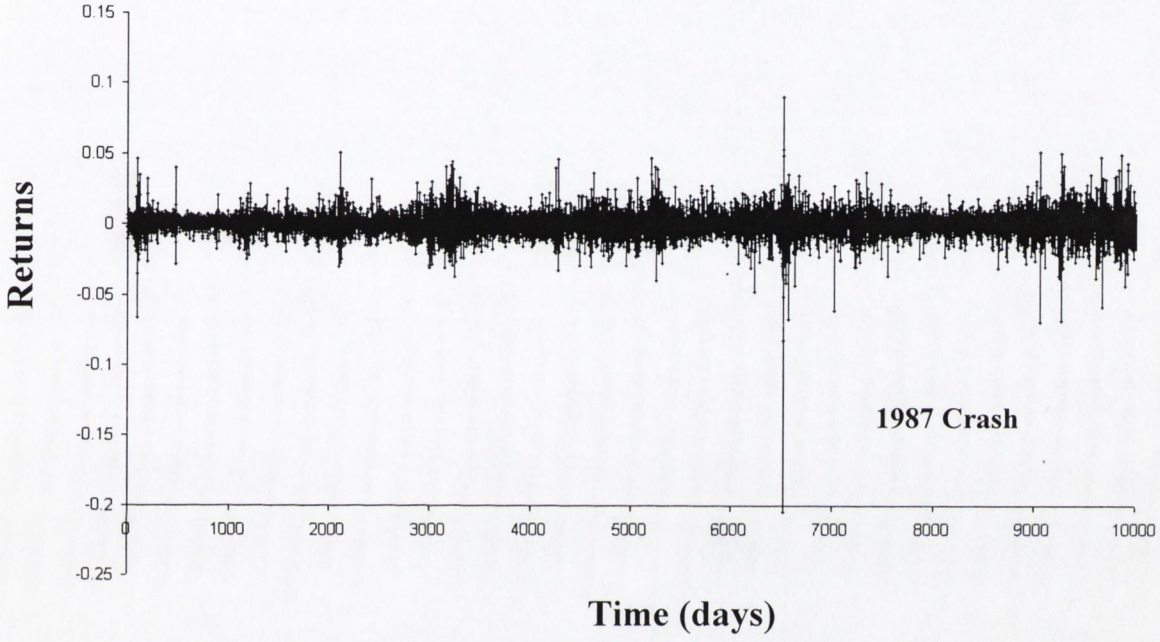

Figure2.3: Standard and Poor 500, daily returns

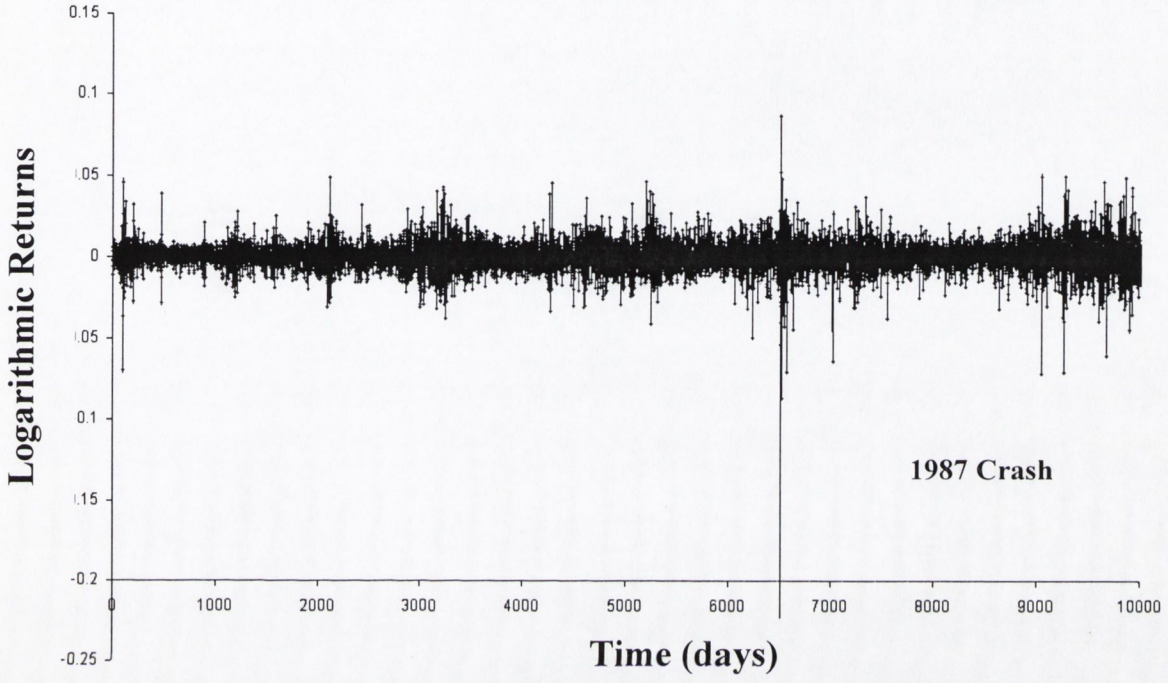

Figure 2.4: Standard and Poor 500, daily logarithmic returns.

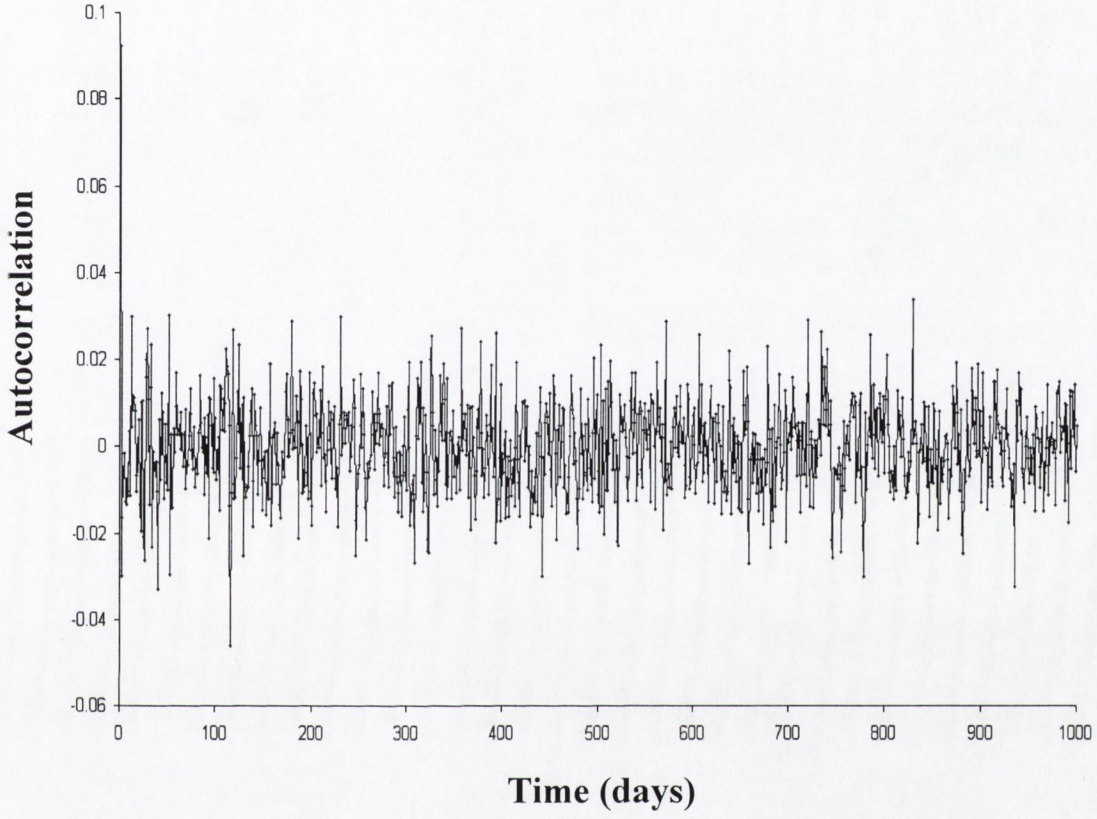

Figure 2.5: Standard and Poor 500, autocorrelation function of the logarithmic returns. The data show the absence of any significant correlation.

A typical feature of return distributions is what econometricians call leptokurtosis (long neck, thin shoulders and fat tails). Analysis based on high frequency data illustrates the presence of two tail regions. The far-tail decays as a power law with an exponent equal to 4; the near-tail is better approximated by a power law with exponent close to 1.5.

Such power law or fat tail regions have been known since the early work of Mandelbrot who suggested that the probability distribution function decayed as a power law with an index of 1.7. Following the work of Mandelbrot, stable Levy distributions were proposed as the form for fat tails. Mantegna and Stanley found that S&P500 can be fit by a truncated Levy distribution [3].Tang and Huang [31] suggest the following for the probability distribution function:

$$P(x)=\frac{C}{\left(x^2+a^2\right)^{1/2}}\exp\left[-\alpha\left(x^2+a^2\right)^{1/2}\right] \tag{2.7}$$

where α, a and C are parameters.

Bouchaud states [8] that a reasonable fit to most markets can be obtained using a symmetrical truncated Levy distribution defined in Fourier space as

$$\ln\widetilde{P}(z)=\frac{A}{\cos\pi\mu/2}\left[\lambda^{\mu}-\left(\lambda^2+z^2\right)^{\mu/2}f(z)\right] \tag{2.8}$$

where

$$f(z)=\cos\left(\mu \arctan\frac{|z|}{\lambda}\right) \tag{2.9}$$

A is the scale factor, μ is the power law exponent for the 'near tails' that for most markets is approximately 1.5. The parameter λ characterizes

the exponential decay of the far tails. Plerou and collaborators have also examined the distribution of profits for 1000 of the largest US companies [32]. The data clearly exhibits a power law but the exponent for the cumulative distribution function is essentially 3. The probability distribution function decays thus with an index of 4.

2.3.4 Volatility

Another very important variable, in the study of financial markets, is volatility. Volatility is a quantity expressing the magnitude of price fluctuations, often identified with the sample variance of p.d.f. of logarithmic returns, calculated over a certain time window (see Appendix 3).

Alternatively, as in [11] and here, one can take the absolute value or of the logarithmic returns defined in (2.2) (or of the returns, defined in (2.1)) as a measure for the volatility.

The time series of volatility present some typical, almost ubiquitous (independently of the asset or index related) features:

- Clusters (high volatility periods come with high volatility periods and the other way around)

- Long memory.

The autocorrelation function of volatility decays slowly following a power law with an exponent estimated between 0.3 and 1 [11].

A process called volatility cascade has also been observed. The time series of volatility calculated over time windows of days is positively correlated to the time series of volatility calculated over shorter time-series, say minutes, shifted forward. In other words, the time series of volatility calculate taking a larger time-window (coarse volatility) can be used to forecast the behavior of volatility calculated, taking a smaller

time-window (fine volatility) [33]. This feature can be explained by the heterogeneous market hypothesis [33], which postulates the co-existence, in the market, of investors with very different time horizons.

The pdf of high frequency volatility can be fit by a log gaussian pdf around his pick, and by a power law in the (positive!) tail. With an exponent estimated around 4 for S&P500 Index.

Volatility is strongly correlated [33] to volumes of assets exchanged in the market and to the number of agents operating in the market.

This also explains the seasonality effects in the time series of volatility of Foreign currency markets. In fact, there are three main currency markets in the World: New York, Tokio, London. Their opening and closing times, due to the different time zones, are not the same, though the activity periods are overlapping in part. As a result, the Foreign Exchange Market is open 24 hours a day, but the opening or closing of the three major markets produce periodical changes in the number of agents trading (and in the volumes traded). This effect is evident in the volatility time series and produces an intra-daily seasonal pattern.

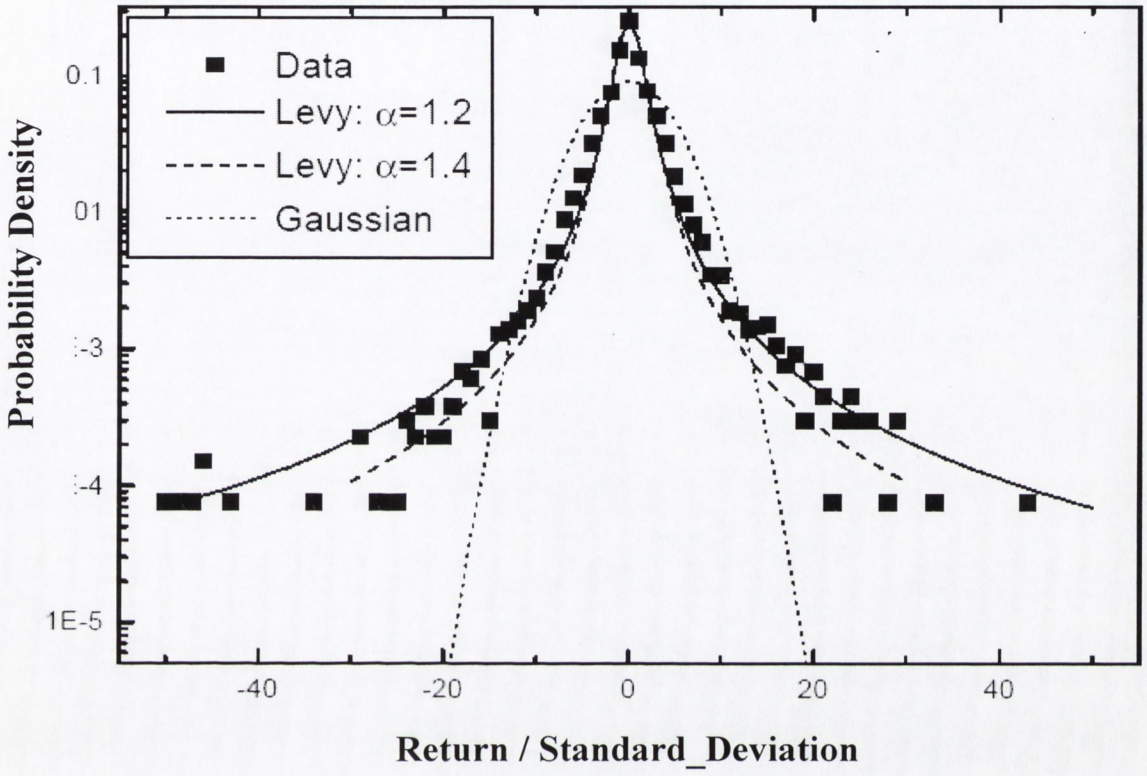

Figure 2.6: Standard and Poor 500 Returns. The curve displays the probability density for the normalized logarithmic returns (returns divided by the sample standard deviation). The figure is taken from http://soma.uchicago.edu/~cowan/finance/Papers/Matacz.pdf

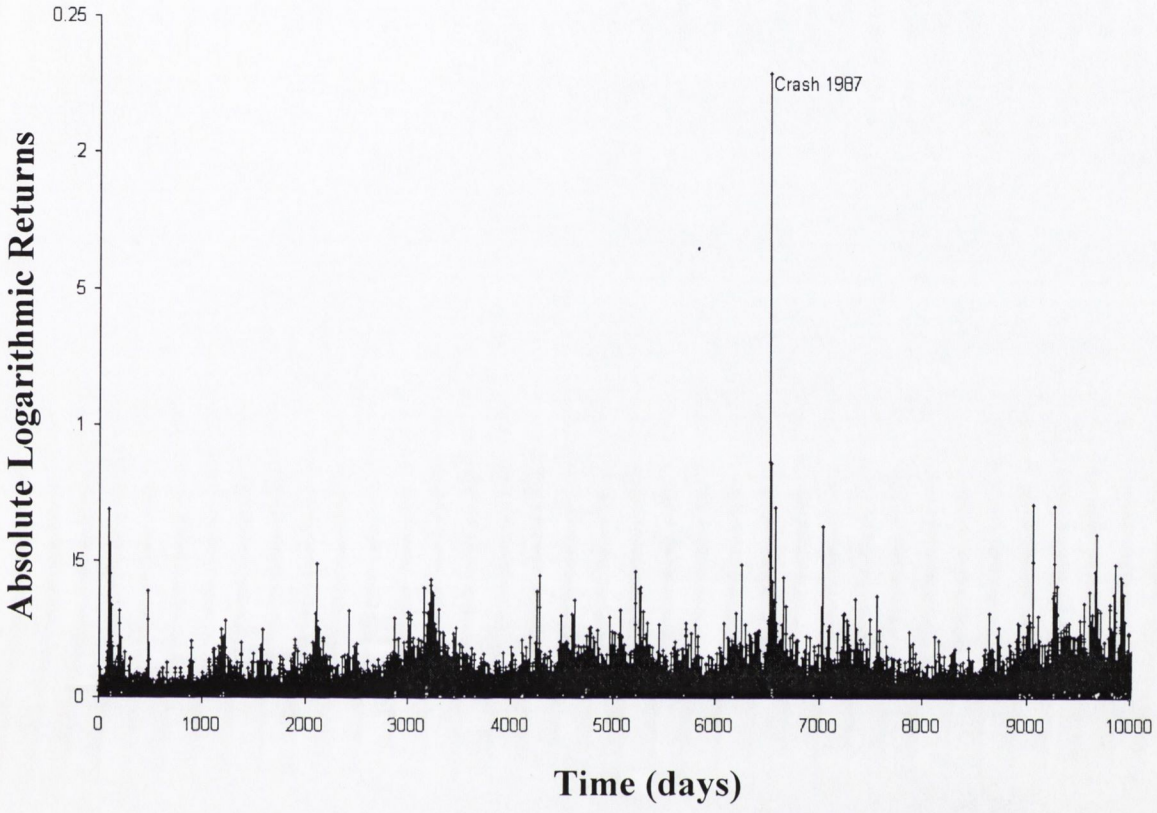

Figure 2.7: Standard and Poor 500 Volatility (absolute value of the logarithmic (daily) returns).

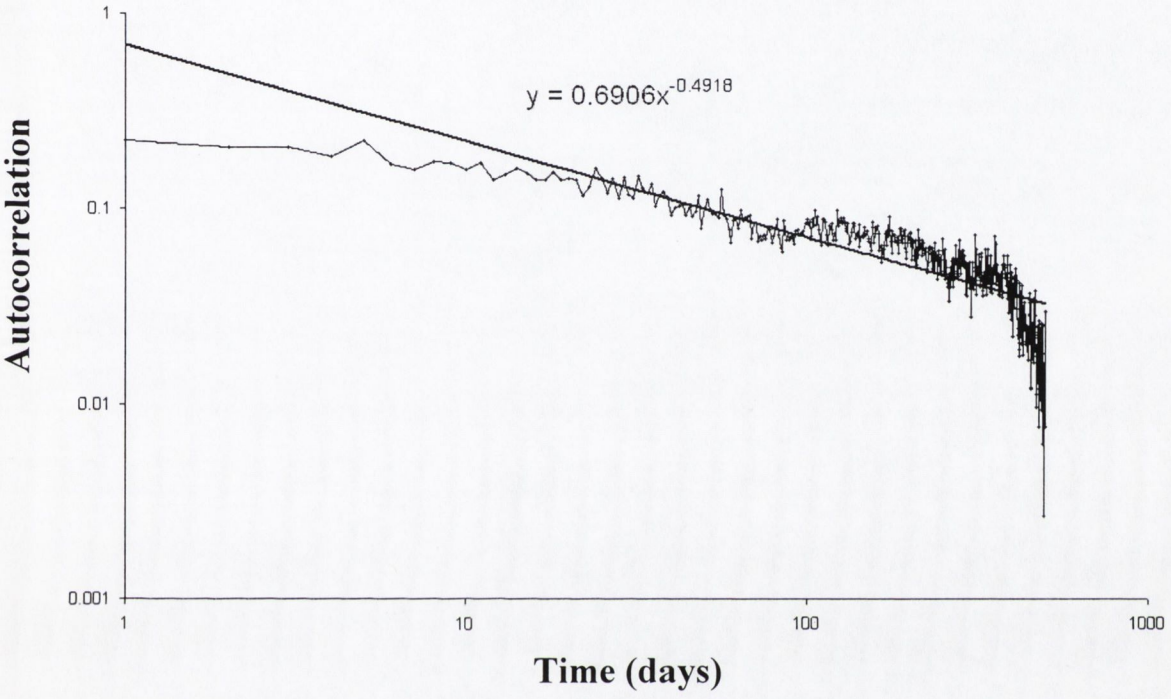

Figure 2.8: Standard and Poor 500. Autocorrelation function of the absolute value of the logarithmic returns. The curve displays a power law asymptotic behaviour, with an exponent close to 0.5.

2.3.5. Market Volumes

The volume is the traded amount of a certain asset.

The probability density distribution of volumes has been studied by Gopikrishnan and co-workers [12]. For many (highly capitalized) stocks traded at New York Stock Exchange it has been found a power law behaviour for the tail with exponent 2.5 (i.e. 1.5 for the cumulative distribution).

More recent studies (Lillo, Farmer and Mantegna) seem to suggest a different behaviour (close to an exponential decay) for the volume distribution of stocks traded at London Stock Exchange [7], but it is not clear yet whether this discrepancy between the two data-sets is due to an actual difference between the two markets or to a different methodological approach followed by Farmer and Lillo.

2.3.6. Cross correlation between stocks and data clustering

Assets in a stock market and, in particular stocks, are not mutually independent. Working out the synchronous cross-correlations

$$\rho_{ij} = \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{\langle S_i^2 \rangle - \langle S_i \rangle^2} \sqrt{\langle S_j^2 \rangle - \langle S_j \rangle^2}} \tag{2.10}$$

between pairs of stocks, we can measure this mutual dependence. In order to classify stocks on the base of their dependence, a natural choice is to introduce a distance, defined by the following properties:

$$\begin{cases} d_{ij} = 0 \Leftrightarrow i = j \\ d_{ij} = d_{ji} \\ d_{ij} \leq \max\{d_{ik}, d_{kj}\} \end{cases} \tag{2.11}$$

In practice, one works out the cross correlation coefficient ρ_{ij} defines a distance $\hat{d}_{ij} = \sqrt{1-\rho_{ij}^2}$ which using a proper algorithm [3][27] builds a space characterized by the three properties above, finding eventually an hierarchy.

The final result is a taxonomic description of a stock market, in which stocks are gathered in clusters, according to the correspondence of their price movements. This is extremely important when trying to build an optimal portfolio. Since a strong cross correlation between a set of stocks entails that they will tend to move in the same direction, therefore if you have in your portfolio only positively and highly correlated stocks, a pitfall in your market sector would not be compensated by the performances of other sectors. The consequences can be very destabilizing for individual, corporate or public finances.

2.4 Vote distributions.

Other interesting features of social systems are related to the statistics of elections. As we will see in Chapter 5, this topic relates to the dynamics of opinions within social systems. The distribution of votes among the candidates in the Brazilian [34] and Indian elections [35] display a power law behaviour (overlooking the finite size effects observed at the beginning and at the end of these distributions) with exponents 1 and 1.3.

2.5 Microscopic features of socio-economic systems

2.5.1 Introduction

Understanding the microscopic structure of self-organized social systems means understanding the way information is received, processed and transmitted by individual agents.

A key role is played by the topological structure of the system, that determines who interacts with who.

A second fundamental factor is that agents often operate without knowing what other agents are doing at that time or will do next. This leads to 'social dilemmas': situations in which decisions that make sense to each individual can aggregate into outcomes in which everyone suffers.

2.5.2 Networks

In a typical social system [36]:

- Nobody interacts with everybody else and not everybody interacts with the same number of people (degree distribution).
- Cliques form, for instance, exclusive circles of friends where everyone knows everybody else (clustering).
- Despite an often large size, there is a relatively short path (i.e. common acquaintances, business partners, etc) between two people (small world).

Interestingly enough, 'real world' network features deviate quite a lot from what one would expect from purely randomly connected networks.

In fact, the degree distributions found decay much slower than Poissonian distributions and the clustering is much more pronounced.

Empirical studies on networks (performed by several researchers) have found that the degree distribution to display often a power law decaying behaviour. The degree distribution of the collaboration between movie actors has, a power law asymptotic tail with an exponent of about 2.3 [4], for the science collaboration an exponent equal to 1.2 was found [5]. For the web of human sexual contacts the exponent found is 3.5 [37] and, finally, for the phone call network the exponent found is 2.1 [38]. Similar studies have been performed for the clustering coefficients of some social systems. A popular manifestation of the 'small world' is the 'six degree of separation' concept, uncovered by the social psychologist Stanley Milgram (1967), who concluded that there was a path of acquaintances with typical length about six between most pairs of people in the United States[39].

2.5.3 Social dilemmas

A paradigmatic example of a 'social dilemma' is given by the, so called, Prisoner's dilemma, developed by Merrill Flood and Melvin Dresher at the Rand Corporation, a US Government think-tank, in the early 1950's [40].

Two suspects are taken into custody and separated. The district attorney is certain that they are guilty of a specific crime, but he does not have adequate evidence to convict them at a trial. He points out to each prisoner that each has two alternatives: to confess the crime, the police are sure they have done, or not to confess. If they both do not confess, the district attorney states that he will book them on some very minor trumped-up charge such as petty larceny and illegal possession of weapon, and they will both receive minor punishment; if they both confess, they will be prosecuted, but he will recommend less than the most severe sentence; but if one confesses and the other does not, then

the confessor will receive lenient treatment for turning state's evidence whereas the latter will get "the book" slapped at him. They are both better off if they do not squeal and get 1 year each for the minor misdemeanor. But as they are tempted to get away with the smallest sentence possible, they end up with an outcome that is worse for both of them they both get 15 years in prison.

The numbers in the cells represent the utility that every player is getting from each outcome. The higher the number, the more utility he/she gets. The number before the coma stands for the payoff of Player 1. The number after the coma stands for the payoff of the second player. The payoff of each player depends on the action of the other person (that is why it is a strategic interaction). To figure out the payoffs of the players, one has to do the following:

If Player 1 Cooperates (Do not Squeal), Player 2 has two options: to Cooperate himself, or to Defect (Squeal). For these alternatives he can get payoffs of 4 or 5. Naturally he will choose 5.

If Player 1 Defects (Squeals) Player 2 can choose between the two possible outcomes and payoffs in the bottom left and bottom right cell. Between 0 and 1 naturally he will choose 1 (Defect).

		Player 2	
		Cooperate (Do not Squeal)	Defect (Squeal)
Player 1	Cooperate (Do Not Squeal)	4,4	0,5
	Defect (Squeal)	5,0	1,1

Table1. Payoff of different strategies in the Prisoner's dilemma. The outcome for a single player depends on the actions of both players. If they both cooperate they both get 4 points each; if one cooperate and the other squeals, the first get 0 points and the second 5. If they both squeal they only get 1 point each.

Now if one does the same for player 1 you will discover that if Player 2 cooperates (Does not Squeal), Player 1 is better of Defecting (again 5 is greater than 4). If Player 2 Defects, Player 1 is again better off defecting (choice between 1 and 0).

In the Prisoner's Dilemma the rationality seems to frustrate the mutual good. The extension of this result to real life is not straightforward, since other elements have to be taken into account, first of all, the records of past actions and the experience.

Nevertheless, the dilemma still holds when there is no such a thing as memory (or any other element that may guide our actions) and we are asked to make a decision whose outcome depends heavily on what other people are deciding, without us knowing.

2.5.4 Playing with rationality: a guessing game

Consider now a group of people taking part in a simple game. Nobody knows what the others' answers are. The aim of the game is to pick a (real) number between 0-100 ([0,100]). The winner will be the respondent who chooses the number closest to 2/3rds of the average number chosen.

In a world of equally rational players, the only answer that may ensure everybody wins is 'zero'.

In the real world that is not the case. One single player may get it wrong and still win, preventing the others ('rationally altruistic') from winning. In other words there is not a right answer: the winning answer ultimately depends on the answers given by the others.

Table 2 provides some examples from this game.

Player	Guess1	Guess2	Guess 3	Guess 4
A	0	0.023	24	10.5
B	0	0.001	0	0.5
C	0	0	0	0
D	0	0	0	0
E	0	0	0	1
F	0	0	0	0
G	0	0	0	0
H	0	0	0	0
Winning answer	0	0.002	2	1
Winner(s)	A,B,C,D,E,F,G,H	B	B,C,D,E,F,G,H	E

Table 2: Eight players are asked to pick a (real) number between 0-100 ([0,100]). The winner will be the respondent who chooses the number closest to 2/3rds of the average number chosen. Guess 1: everybody choose zero and, as a consequence, everybody win. Guess 2: A and B choose numbers larger than zero, B wins since 0.001 is the answer closer to 2/3rds of the average (0.002). Guess 3: Everybody win except from A. Guess 4: E wins since 1 is equal to 2/3rds of the average.

CHAPTER 3

DIFFUSIVE AGENTS

3.1 Introduction

The simplest way of modeling fluctuations in a social system is to draw a comparison with the Brownian Motion and to think of an agent as a Brownian particle. One may then assume that changes of a system macroscopic quantity are determined by the sum of many agent contributions. However, this approach does not account for some relevant statistical properties of financial market and of other socio-economic systems and requires generalization. A very important extension of the random walk theory is the 'Continuous Time Random Walk' theory, in which the distribution of the waiting times between consecutive stochastic events is taken into account. In this chapter we study the waiting time distributions of two data sets taken from the XIX century Irish stock market and from the late XX century world currency market.

3.2 Langevin approach

The theory of Brownian motion is traditionally associated to the name of A. Einstein (1905), but in reality a first version of that theory was introduced before by Bachelier (1900) to describe price movements [41].

The basic idea consists in assuming the change of coordinate driven by a deterministic function of the coordinate plus a stochastic term (basically white noise), according to the equation

$$\frac{dx}{dt} = a(x) + \hat{a}_0(t) \tag{3.1}$$

Where $a(x)$ is a deterministic function of position and $\hat{a}_0(t)$ is a stochastic component defined by the following properties:

$$\langle \hat{a}_0(t) \rangle = 0 \qquad and \qquad \langle \hat{a}_0(t)\hat{a}_0(t') \rangle = 2D_0\delta(t-t') \tag{3.2}$$

The Fokker-Plank equation associated to equation (3.1) is:

$$\frac{\partial P(x,t)}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(P(x,t)\right)\right) - \frac{\partial}{\partial x}\left(a(x)P(x,t)\right) \tag{3.3}$$

Where $P(x,t)$ is the probability density function of the system.

When thinking about molecules, Einstein [41] argued that anyone molecule moving through the system is buffeted by other molecules and the total effect may be interpreted as the effect of uncorrelated random noise. By analogy, we may loosely think of agents as Brownian particles and assume that the change of the x coordinate is determined by the sum of agent contributions. Generally speaking each individual agent action is made of two components: the first is deterministic and dependent only on the average coordinate x, the second is purely stochastic and uncorrelated to any other agent behaviour.

Unfortunately the basic Langevin approach may not be sufficient to describe complex dynamics (for instance is not able to explain the emergence of power-law tails in the distribution $P(x)$) and requires generalization. We can write

$$\dot{x} = F(x) + G(x)\eta(t) \tag{3.4}$$

where $F(x)$ is the drift-term and $G(x)$ is the 'local' variance.

$$\begin{cases} \langle \eta(t) \rangle = 0 \\ \langle \eta(t)\eta(t') \rangle = 2D\delta(t-t') \end{cases} \tag{3.5}$$

A simple example of a generalized Langevin model is the Black-Scholes model [3][8]

$$\dot{P} = \mu P + P\eta(t) \tag{3.6}$$

where $\eta(t)$ fulfils the conditions in equations (3.5).

It leads to a log normal distribution of prices. This model, though based upon some relatively unrealistic assumptions (as continuity of price change variable and also completely uncorrelated price movements) is currently used in finance for options evaluation and pricing.

Anyway, only generalizing the basic Langevin approach we may obtain power law tails in the probability distribution functions [44].

3.3 Continuous Time Random Walk

3.3.1 Theory

The "classic" model of random walk is based on the idea of instantaneous jumps, each one occurring after a certain well-defined and constant time. A more realistic model should regard the time as a continuous variable and should assume that also the time lags between two consecutive jumps can be a stochastic process. This is the basic idea

of Continuous Time Random Walk introduced by Montroll and Weiss in 1965 [42].

On a microscopic scale, this may be explained by the fact that each fluctuation is determined by a single agent and by the fact that agents do not act at regular time intervals. For instance, in a market, orders are not submitted at regular times and the time-lag between two transactions may depend on several factors such as the number of traders operating in the market at a certain time and the recent past performances of the market.

To the sake of analytical tractability, we assume that the jump pdf $\lambda(\xi)$ is independent of the waiting-time p.d.f. $\psi(t)$ so that the jumps ξ_i (at instant t_i; i = 1; 2; 3; : : :) are i.i.d. (independent identically distributed) random variables, all having the same probability density $\lambda(\xi)$. We can therefore write $\varphi(\xi,t) = \lambda(\xi)\psi(t)$.

The jump p.d.f. $\lambda(\xi)$represents the p.d.f. for transition of the walker from a point $x$to a point $x+\xi$. The waiting-time p.d.f. represents the p.d.f. that a step is taken at the instant $t_1 + t$after the previous one (that happened at the instant t_1) so it is also called the pausing-time p.d.f. The probability that a given inter-step time interval is greater or equal to τ will be denoted by $\Psi(t)$ which is defined in terms of $\psi(t)$by

$$\Psi(\tau) = \int_{\tau}^{\infty} \psi(t')\,dt' = 1 - \int_{0}^{\tau} \psi(t')\,dt' \tag{3.7}$$

$\Psi(\tau)$ is the probability that the diffusing quantity x does not change value during the time interval of duration τ after a jump. We also note, taking t_0=0 that $\Psi(t)$ is the survival probability until time t at the initial position x_0 = 0.

Let us now denote by $p(x; t)$ the p.d.f. of finding the random walker at the position x at time instant t .We assume the initial condition $p(x,t) = \delta(x)$, meaning that the walker is initially at the origin

$x = 0$. Montroll and Weiss have shown that the Fourier-Laplace transform of $p(x,t)$ satisfies a characteristic equation:

$$\int_0^t \Phi(t-t')\frac{\partial}{\partial t'}p(x,t')dt' = -p(x,t) + \int_{-\infty}^{\infty}\lambda(x-x')p(x',t)dx' \tag{3.8}$$

$\Phi(t-t')$ plays here the role of "memory function", when $\Phi(t) \propto \delta(t)$ we find the well known results for a non memory process.

If we consider a process with a long memory function [23], given by

$$\Phi(t-t') = \frac{t^{-\beta}}{\Gamma(1-\beta)}, \qquad t \geq 0, \qquad 0 < \beta < 1 \tag{3.9}$$

we find the so called Caputo fractional time derivative of order β that implies for the Surviving Time Distribution a particular form

$$\Psi(t) = \begin{cases} = \sum_{n=0}^{\infty}(-1)^n \frac{t^{\beta n}}{\Gamma(\beta n+1)}, & t \geq 0 \\ \sim \frac{\sin(\beta\pi)}{\pi}\frac{\Gamma(\beta)}{t^{\beta}} & t \to \infty \end{cases} \tag{3.10}$$

known as Mittag-Leffler generalized function.

3.3.2 Testing the continuous-time random walk model

We have tested the model presented in equations 3.26-3.27-3.28, using two non-homogeneous (the time between two consecutive transactions is not constant) market data sets [54].

We analyse 10 stocks taken from Irish Stock Market[1] in 19th century (5 railways and 5 banks). The data were recorded between the 1st of January

[1] During the 19th century, Irish stock activity was recorded on a daily basis and characterized by waiting times that varied from a day to some months.

1850 and the 31st of December 1854. In order to minimize spurious effects related to single stocks and to capture the typical behaviour we work out the "average" Survival Time Probability Distribution (STPD) function.

The distribution in figure 3.1 presents a sharp cut off for survival times bigger than 150 days. Below that time we can clearly identify two regions: the first one, ranging between 1 and 10 days presents an exponential behaviour (characteristic time 5 days), the second is well fitted by a power law function having an exponent $\beta = 0.4$. These two regions are very well fitted by the Mittag-Leffler function:

$$\Psi(\tau) = E_{\beta}\left[-(\gamma\tau)^{\beta}\right] \tag{3.11}$$

Deals were done on a 'matched bargain basis' with members of the exchange (i.e., stockbrokers) bringing buyers and sellers together in essentially the same way as is done today via electronic trading, say. The only difference is that there are many more buyers and sellers. In the 19th century one informed one's stockbroker what one wished to trade (say sell) a certain quantity and requested he get the best quote or do the deal subject to a price limit (here a minimum price). He maintained a ledger of such pending deals and sought to meet, on the stock exchange floor, other member of the exchange trying to do the opposite deal. Recent studies [5] of different 19th century markets find that they were well integrated. Dublin traded international shares - it was not solely a regional market. The largest shares - Banks and key railways - represented quality investments for UK investors and were also traded in London. Smaller shares would primarily have been of only local interest. We see a mix of local small investment and large ones integrated with London and hence the world stock markets. World trends are thus reflected in the Irish market for which, at the time, there were no exchange controls. (From 1801 to 1922 Ireland was part of UK).

where $\gamma = 0.025$. We estimate an uncertainty for the fit parameters value better than 10% and a $\tilde{\chi}^2 > 0.06$ (overlooking the cut off region).

Waiting times range typically from a day to some months. In each time series we have 1826 days and a number of waiting times (i.e. of transactions) ranging between 140 and 160.

We have performed the same analysis for Japanese Yen Currency. The data are taken between the 1989 and the 1998, the waiting times range from a minute to some hours. The STPD function (figure 3.2) shows in this case a very different behavior. Above 100 minutes, the distribution does not drop off sharply and presents two small humps or shoulders. The times associated with these shoulders may be associated with closing and opening times of the main currency markets in the World (New York, London, Tokyo).

For shorter times we no longer have a region that can be fitted easily with an exponential function – or at least any exponential behavior has a relaxation time of less than a minute. This region appears to exhibit power law essentially over the complete range of data. But note that the power law now is greater than unity and the origin therefore must involve a mechanism outside that considered by Mainardi.

The shoulders do exhibit a very small region showing a power law on the boundary of the values admitted for β in (3.9). In fact, β ranges here between 0.9 and 1.1. However, the distribution appears to be dissimilar to the Mittag-Leffler function.

In conclusion, according to our studies, continuous time random walk models, with a power law memory function in the waiting time dynamic, seem to reproduce the Survival time probability distributions for Irish Stock Market of 19^{th} century and it would be interesting to see whether this holds for other early markets. However in their simple form, they do not appear to replicate the more complex features of modern currency markets. Again it would be interesting to explore this theory in the context of other contemporary assets.

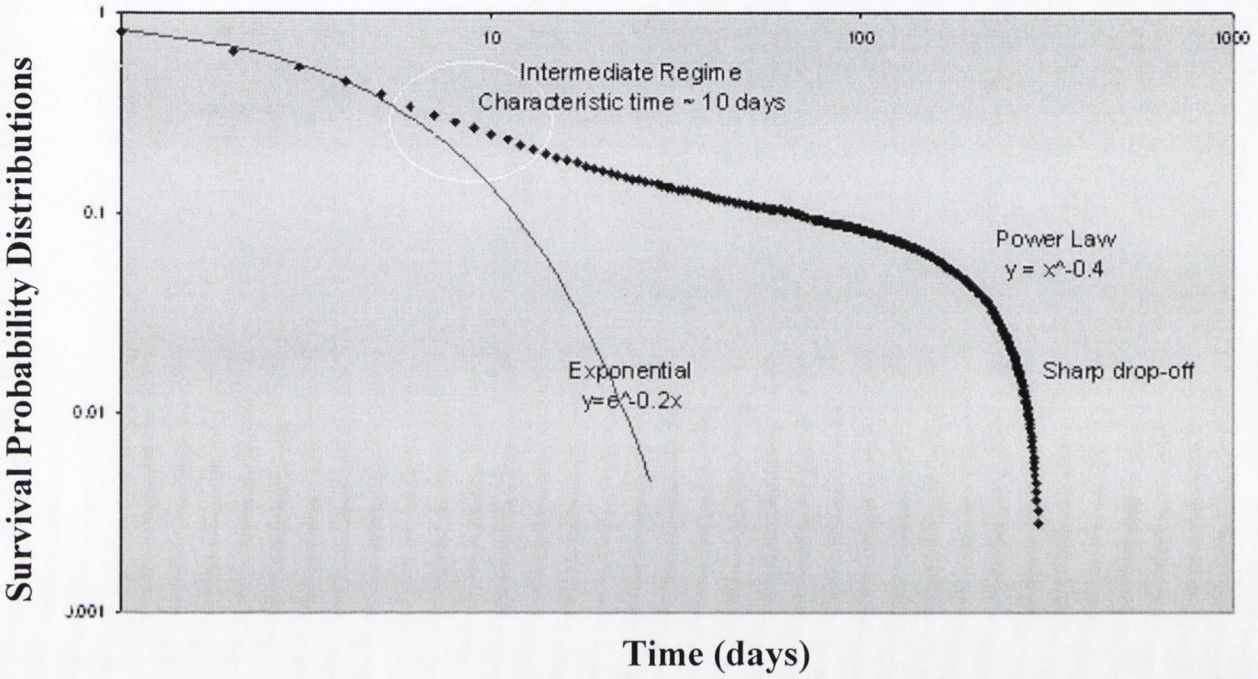

Figure 3.1: Average Survival Time Probability Distribution function for Irish Stock Market data between 1850 and 1854. Fit parameters for Mittag-Leffler function: $\gamma = 0.025$, $\beta = 0.4$. The relative error for the exponent is about 0.1.

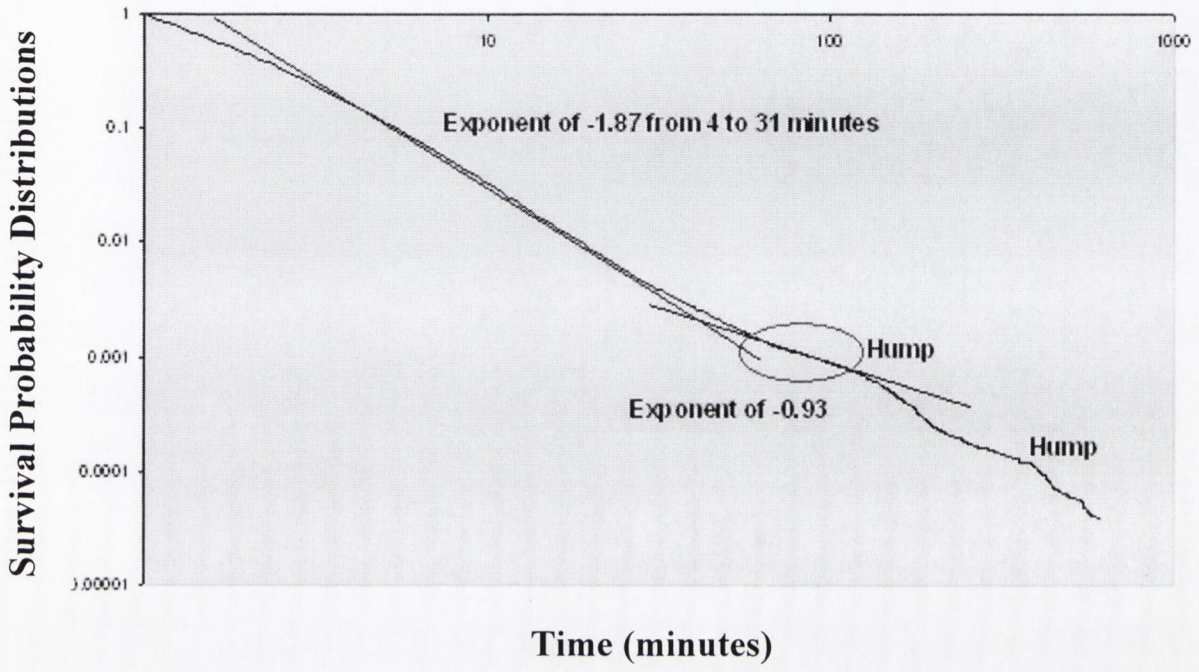

Figure 3.2: Survival Time Probability Distribution function for Yen Currency Market data between 1989 and 1998. The relative error for the exponents is about 0.1.

CHAPTER 4

STOCASTIC DYNAMICAL AGENTS

4.1 Introduction

Some statistical models, such as ARCH and GARCH can generate leptokurtic (i.e. having long neck, thin shoulder and fat tails) probability distributions [3], but they use several parameters that lack of a clear connection with the underling microscopic structure. In this chapter we will deal with 'mesoscopic' agents, each of which is subject to specific forces and random fluctuations. This allows us to generate fat tailed distributions depending on few parameters having a more intuitive meaning at a 'mesoscopic' level.

4.2 Generalized Lotka-Volterra (GLV) Models

In the 90s, Sorin Solomon and co-workers introduced the Generalized Lotka-Volterra (GLV) models. These generalize the Logistic equation

$$\omega(t+1) = p\omega(t)(1-\omega(t)) \qquad p,\omega \in \Re_+ \tag{4.1}$$

Where p denotes a constant parameter and ω is a function of time.

They assume that the quantity ω above (that from this time forward we will call wealth) is given by the contribution of many agents, whereby individual wealth changes are due to chance, cooperation and competition with other agents[10][44].

The most general expression of GLV is given by

$$\omega_i(t+1) = [1+\lambda(t)]\omega_i(t) + \sum_{j=1}^{N} a_{i,j}\omega_j(t) - \sum_{j=1}^{N} a_{j,i}\omega_i(t) - \sum_{j=1}^{N} c_{i,j}\omega_i(t)\omega_j(t) \tag{4.2}$$

$$\omega_j(t+1) = \omega_j(t) \qquad j = 1,....N; j \neq i \tag{4.3}$$

where i is an integer chosen randomly in the range $1 \leq i \leq N$, at each time step t, where $\lambda(t)$ are identical independent variables drawn from a given distribution $\Pi(\lambda)$, which does not depend on i and t; $a_{i,j}$ and $c_{i,j}$ are constant.

The first term on the right hand side of Eq. (4.2), describes the effect of stochastic factors at the individual level. In an ecological system this term represents variations in the population of a given species (births and deaths affected by external conditions but not affected by the interaction with other species). In a stock market, it represents the increase (or decrease) by a random factor $\lambda(t)$ of the capital of the investor i, between the time t and $t+1$.

The second and the third terms are called auto-catalytic terms. In an ecological system, the second term represents the dependence of population i on the availability of food, in the form of population j. The third term represents the fact that population i itself may be the food of some other species. In an economic system the second and third terms represent trade between investors or firms i and j, such as buying and selling, respectively.

The fourth term in Eq. (4.2), describes saturation effects due to the competition for limited resources. In an ecological model, that term implies that large populations tend to exhaust the available resources they depend on. In an economic system, this term is related to the saturation, due to the final size of the economy.

A simpler version of this model is based on the equations:

$$\omega_i(t+1) = [1+\lambda(t)]\omega_i(t) + a\,\bar{\omega}\,(t) - C(\omega_1,...\omega_n,t)\omega_i(t) \tag{4.4}$$

$$\omega_j(t+1) = \omega_j(t) \qquad j = 1,....N; j \neq i \tag{4.5}$$

where $C(\omega_1,...\omega_n,t)$ is a general function of the $\omega_j(t)$'s that includes an explicit time dependence:

$$\overline{\omega}\ (t) = \frac{1}{N}\sum_{i=1}^{N}\omega_i(t) \tag{4.6}$$

Introducing a set of normalized variables

$$x_i = \frac{\omega_i(t)}{\overline{\omega}\ (t)} \tag{4.7}$$

and, using a mean-field approximation, Eq. (4.2) can be couched in terms of a generalized Langevin Equation. This allowed Peter Richmond (2000) to work out analytically a very important feature of GLV models: they asymptotically lead to power law tailed probability density functions. [44]

$$P(x) \propto x^{-1-\alpha}\exp\left[\frac{1-\alpha}{x}\right] \tag{4.8}$$

where $\alpha = 1+2a/\left(D+a^2\right)$ and D is the variance of $\lambda(t)$. This is consistent with the distribution of the data obtained through simulating the system described by the (discrete) equation (4.4) and (4.5).

4.3 Peer pressure and diffusion

We consider now a one-dimensional diffusion model characterized by a drift term originating from a kind of "peer pressure" that each particle feels depending on its position and functionally dependent on the probability density distribution.

There is an interaction between the walkers, which leads to a drift or a bias term which opposes the diffusion spreading and promotes aggregation. The basic idea is that a walker perceives the populations of other walkers, over a certain range, both right and left of her own location and has a drift in the more crowded direction.

Examples in which aggregation processes compete with diffusion process are:

- Insects (chemotactic signals)
- Charity
- Stock price movements
- Investors in a market

The idea can be encapsulated in a non linear equation, proposed by Marsili [45]

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\left(V(x,t)P\right) + D\frac{\partial^2}{\partial x^2}P \tag{4.9}$$

where $P(x,t)$ is the distribution of the locations of particles at time t with a drift velocity:

$$V(x,t) = \frac{\lambda}{2}\frac{\Phi_+(x,t) - \Phi_-(x,t)}{\Phi_+(x,t) + \Phi_-(x,t)} \tag{4.10}$$

where

$$\Phi_+(x,t) = \int_x^\infty dy e^{(x-y)/\xi} P(y,t) \tag{4.11}$$

$$\Phi_-(x,t) = \int_{-\infty}^x dy e^{(y-x)/\xi} P(y,t) \tag{4.12}$$

are the measure of population within a range ξ.

If $\xi \to 0$ or $\lambda \to 0$ we find the usual diffusion equation of brownian motion.

Taking:

$$\begin{aligned} r &= \frac{\lambda \xi}{2D} \\ Q &= \Phi_+ + \Phi_- \\ \Pi &= \Phi_+ - \Phi_- \end{aligned} \tag{4.13}$$

Q and Π can be seen as conjugate Hamiltonian variables of a system where:

$$H(\Pi, Q) = \frac{\Pi^2}{2\xi} + V(Q) \tag{4.14}$$

is the Hamiltonian

$$V(Q) = -\frac{Q^2}{2\xi} + \frac{2C}{1+r} Q^{r+1} \tag{4.15}$$

is the potential and where the initial spatial coordinate x plays the role of time.

The steady solution is strongly dependent on the parameter r.

If r <1 the only solution (physically acceptable) is $P_s(x) = 0$. Therefore the probability density spreads out and vanishes as in standard diffusion.

For r>1 the steady solution has the form

$$p_s(x) = A(r) \cosh\left[\frac{r-1}{2\xi}(x - x_0)\right]^{-2r/(r-1)} \tag{4.16}$$

where $A(r)$ is a quite complicated function of r.

At r=1 there is a non-trivial change of behaviour, corresponding to a proper phase transition (accompanied by a spontaneous breaking of the translation symmetry).

The centre of the steady distribution depends (in a complicated manner) on the initial conditions

One can also define a kind of "localization length" ℓ which diverges as $\ell = |1-r|^{-1}$ and measures the spread of the probability density distribution.

4.4 Peer pressure effects in Generalized Lotka-Volterra Models.

Consider a simple type of agent model consisting of a set {i} of agents, each of which can have an intrinsic attribute, x_i. This could be proportional to the financial health or wealth of the agent. Now assume that this gives rise to some kind of peer pressure or a force that tends to make 'people keep up with the Jones' as they say in the UK. Such a force of interaction $f(x_i\text{-}x_j)$ between the different agents is similar to that between molecules interacting via an inter-particle potential. The strength of this kind of 'social' force, as has been pointed out by for example, Krawiecki, Holyst and Helbing [26], is not dependent on spatial neighbourhood; rather its origin lies in communication between agents. The spatial topology of agents may be considered to be unimportant. The simplest type of peer pressure is introduced by assuming

$$f(X_i - X_j) = -f(X_j - X_i) \tag{4.17}$$

The peer pressure force has not only a magnitude but also a direction depending on whether the argument, x, is positive or negative.

Now assume the agent dynamics can be captured by the following set of Langevin equations:

$$\frac{dX_i}{dt} = \xi_i + \sum_j f(X_i - X_j) \tag{4.18}$$

The stochastic terms, $\{\xi_i\}$ are defined as:

$$\left\langle \xi_i(t)\xi_j(t') \right\rangle = 2D\delta_{ij}\delta(t-t'); \quad \left\langle \xi_i(t) \right\rangle = 0 \tag{4.19}$$

We now compute the many agent probability distribution function

$$P_N(\mathbf{x},t) = \left\langle \prod_{i=1}^{N} \delta[X_i(t) - x_i] \right\rangle \tag{4.20}$$

A standard calculation gives:

$$\frac{\partial P_N(\mathbf{x}|t)}{\partial t} = \frac{D}{2}\sum_i \frac{\partial^2 P_N(\mathbf{x}|t)}{\partial x_i^2} - \sum_{i,j} \frac{\partial}{\partial x_i}\Big[f(x_i - x_j) P_N(\mathbf{x}|t\} \Big] \tag{4.21}$$

According to Richmond [56], we may reduce the N coordinate representation. This yields a hierarchy of equations of agent probability distribution functions:

$$\frac{\partial P_1(x|t)}{\partial t} = \frac{D}{2}\frac{\partial^2 P_1(x|t\}}{\partial x^2} - \frac{\partial}{\partial x}\int dx' f(x-x') P_2(x,x'|t) \tag{4.22}$$

Clearly a solution is only possible if we have a route to truncating this infinite hierarchy of coupled equations. The simplest way forward is to assume a mean field approximation:

$$P_2(x,x'|t) \to P_1(x|t)P_1(x'|t) \tag{4.23}$$

Dropping the suffix when N=1 we obtain

$$\frac{\partial P(x|t)}{\partial t} = \frac{D}{2}\frac{\partial^2 P(x|t)}{\partial x^2} - \frac{\partial}{\partial x}\phi(x|t)P(x|t) \tag{4.24}$$

The mean field force is given by:

$$\phi(x|t) = \int dx' f(x-x')P(x'|t) \tag{4.25}$$

Now consider the agent pair force $f(x)$, having a directional character. Clearly a number of functions fit this form. A particularly simple choice is [55][56]:

$$\begin{aligned} f(x) &= ae^{-x/\zeta} \quad x > 0 \\ f(x) &= -ae^{x/\zeta_1} \quad x < 0 \end{aligned} \tag{4.26}$$

If $a > 0$, we see that agents with attribute $x_j < x_i$ are attracted to agent i. Equally agents with attribute $x_j > x_i$ choose to avoid agent i. This mimics a kind of 'keeping up with Jones' social force.

This now leads to the expression:

$$\phi(x|t) = a\int_x^\infty dx' e^{-(x'-x)/\zeta} P(x'|t) - a\int_{-\infty}^x dx' e^{-(x-x')/\zeta_1} P(x'|t) \tag{4.27}$$

This form proves useful in illustrating the link to the Lotka-Volterra approach developed by Solomon and co-workers [47].

Using our particular choice of force defined in equation (4.26), we introduce the change of variable:

$$w_i = e^{X_i/\zeta} \tag{4.28}$$

Using the theorem of Ito the stochastic equation for our new variable w_i defined in (4.28) is:

$$\zeta \frac{dw_i}{dt} = w_i \varepsilon(t) + a \sum_{w_j < w_i} w_j - a \sum_{w_j > w_i} \frac{w_i^{1+\zeta/\zeta_1}}{w_j^{\zeta/\zeta_1}} + \frac{D}{2} w_i \qquad (4.29)$$

Note that now the stochastic function $\varepsilon_i(t)$ takes only positive values and the additional term D may be absorbed within the mean value of this stochastic function.

We refer to (4.29) as GLVPP (Generalised Lotka-Volterra with Peer Pressure). Note that they are exact consequences of introducing the specific force (4.26) into the initial Langevin dynamics defined by equations (4.18). No mean field approximation has been used to reach this point. For large values of w_i, the third term on the RHS of equation (4.19) may be neglected. Equally the sum over $w_j < w_i$ may be extended over all w_j. As a result equation (4.19) reduces essentially to the generalised Lotka-Volterra equations proposed by Solomon for wealth dynamics. These latter equations have been solved using both a mean field approach and exact numerical methods. Both approaches show that the distribution of wealth, w, displays a power law or fat tail [24].

The comparison between our simple peer pressure model and the GLV model of Solomon is shown in Figure 4.1. The points are computed from the Langevin equation not the mean field theory so is an exact solution. Clearly the difference for the parameters used is minimal [25]. For both models we find a power-law tail with an exponent between 1.3 and 1.6, compatible with the exponent of the wealth distributions of many industrialized countries.

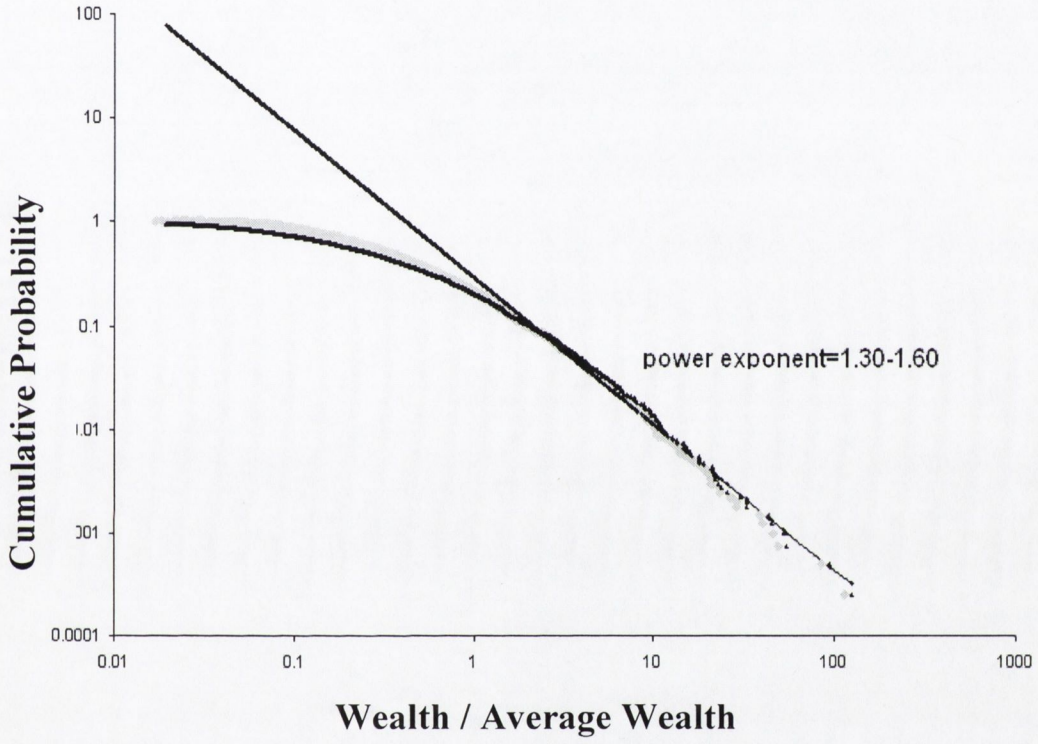

Figure 4.1: GLVPP vs. GLV. The curve displays the cumulative probability density (the probability of finding a ratio Wealth/Average Wealth larger than the value reported on the x-axes)of the data obtained both GLVPP (black) and GLV (grey) models, with the following parameter choice: c=0, aN=0.00023, D=0.00083, $\zeta 1=\zeta 2=1$. 10 samples each made of 1000 agents have been studied. The straight lines (black-grey) fitting the tails in the log-log plot show that they exhibit a power-law behaviour with exponent between 1.3 and 1.6. The relative error for each data point is about 5%. The error bars are of the same size as the point-markers in the graph.

CHAPTER 5

CELLULAR AUTOMATA, PERCOLATION AND THE SZNAJD MODEL

5.1 Introduction

5.1.1 Cellular Automata

Cellular automata are idealizations of physical systems in which all the relevant 'physical' quantities (including space and time) are discrete.

A typical cellular automaton consists of a regular uniform lattice (or 'array') in which, at each time step, each site is characterised by one set of (one or more) discrete variables.

The time evolution happens through updating the variables at each site. The variables at one site are affected by the values of variables at sites in its ``neighborhood'' on the previous time step. The neighborhood of a site is typically taken to be the site itself and all immediately adjacent sites. The variables at each site are updated simultaneously (``synchronously''), based on the values of the variables in their neighborhood at the preceding time step, and according to a definite set of 'local rules.'

Cellular automata [43] were originally introduced by von Neumann and Ulam (under the name of ``cellular spaces'') in 1963-6 as a possible idealization of biological systems, with the particular purpose of modelling biological self-reproduction. They have been applied and reintroduced for a wide variety of purposes, and referred to by a variety of names, including ``tessellation automata,'' ``homogeneous structures,'' ``cellular structures,'' ``tessellation structures,'' and ``iterative arrays.''

In physics many systems may be easily represented in terms of cellular automata. A paradigmatic case is given by the Ising Model [53] (an elementary form of cellular automata).

5.1.2 Consensus Models

Simple cellular automaton based models, such as the Ising or Potts models, are useful not only to understand physical phenomena such as ferromagnetism [53], but also to study the effect of interactions between human beings in a society. This may cast light on the way opinions or behaviours spread throughout society.

In many social systems, agent interaction is local or short ranged. In spite of this, under certain conditions, the 'state' of an agent may spread through the system and become the 'state' of the system. The word 'state' is very generic and may mean: opinion, attitude, behaviour, action, infection, etc, depending on the specific case. We will refer to models describing this spreading process as 'Consensus Models', since they may describe the process through which a society in which initially different opinions are present may end up with upholding a single opinion (or just few of them).

5.1.3 Percolation theory

The opinion spreading process within a society may be described as a percolation process [36] [57].

Consider a regular d-dimensional lattice whose edges are present with probability p and absent with probability 1-p or, alternatively, a lattice in which all bonds are present and a node may be occupied with probability p.

Percolation theory studies the emergence of paths that percolate the lattice (starting at one side and ending at the opposite side).

A typical feature is the existence of a critical value of p, p_c called the percolation threshold. For p values below p_c, only few connections are present, thus only small clusters of nodes connected by edges can form. For p values above p_c, a percolating cluster of nodes connected by edges appears. This cluster is called infinite cluster, since its size diverges as the size of the lattice increases.

Quantities of interest in percolation are:

- The percolation probability, P, denoting the probability that a given nodes belongs to the infinite cluster:

$$P = P_p(|C| = \infty) = 1 - \sum_{s<\infty} P_p(|C| = s) \tag{5.1}$$

where $P_p(|C| = s)$ denotes the probability that the cluster at the origin has size s.

- The average cluster size $\langle s \rangle$, defined as

$$\langle s \rangle = E_p(|C|) = \sum_{s=1}^{\infty} sP_p(|C| = s) \tag{5.2}$$

and gives the expectation value of cluster sizes.

For $P>0$, $\langle s \rangle$ is infinite.

In this case is useful to consider the average size of finite clusters by taking away from the system the infinite cluster:

$$\langle s \rangle^f = E_p(|C|, |C| < \infty) = \sum_{s<\infty} sP_p(|C| = s) \tag{5.3}$$

- The cluster size distribution, n_s, defined as the probability of a node being the left hand end of a cluster of size s,

$$n_s = \frac{1}{s} P_p(|C| = s) \tag{5.4}$$

n_s does not coincide with the probability that a node is part of a cluster of size s . By fixing the position of the node in the cluster (at the left hand end of the cluster), we are choosing one of the s

possible nodes of the cluster, reflected in the division of $P_p(|C| = s)$ by s, and counting every cluster only once (loosely speaking it's like to take all the clusters of size s containing a given point, divide them into s classes defined by the position of that point (from the left hand end to the right hand end) and take only the first. This gives a quantity proportional to the number of clusters of size s).

General results from the percolation theory are:

In the *sub-critical phase* ($p < p_c$)

The probability for a cluster of having size s is given by:

$$P_p(|C| = s) \propto e^{-\alpha(p)s} \qquad s \to \infty \tag{5.5}$$

for

$$\begin{cases} \alpha(p_c) = 0 \\ \alpha(p) \to \infty \quad as \quad p \to \infty \end{cases} \tag{5.6}$$

In the *super-critical phase* ($p > p_c$)

According to the principle ansatz of percolation theory (and to the results so far outlined), the most general expression for the cluster size distribution can be written as:

$$n_s(p) \propto \begin{cases} s^{-\tau} f_-\left(|p - p_c|^{1/\sigma} s\right) & as \quad p \le p_c \\ s^{-\tau} f_+\left(|p - p_c|^{1/\sigma} s\right) & as \quad p \ge p_c \end{cases} \tag{5.7}$$

where τ and σ are critical exponents; $f_{-}(x) \propto e^{-Ax}$ and $f_{+}(x) \propto e^{-Bx^{d-1/d}}$; $x >> 1$.

A second ansatz is that the correlation length diverges near the percolation threshold following a power law:

$$\xi(p) \propto |p - p_c|^{-\nu} \tag{5.8}$$

for $p \to p_c$

The correlation length ξ and the characteristic size s_ξ for the exponential cut-off (i.e. the maximum size that contribute significantly to cluster averages) are related by a power law:

$$s_\xi \propto \xi^{1/\sigma\nu} \tag{5.9}$$

Follows that the percolation probability is given by:

$$P(p) \propto |p - p_c|^{\beta} \; ; \; \beta = \frac{\tau - 2}{\sigma} \tag{5.10}$$

and the average size of finite clusters obeys:

$$\langle s \rangle^{f} \propto |p - p_c|^{-\gamma} \tag{5.11}$$

with $\gamma = \frac{3 - \tau}{\sigma}$

Also, below the critical threshold, clusters are fractals since their size does not scale as their radius to the d-th power, but as

$s(r) \propto r^{d_f}$ $(d_f = 1/\sigma\nu)$. Above the critical threshold clusters become normal d-dimensional objects.

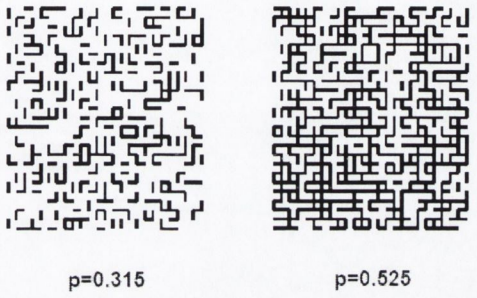

Figure 5.1. An example of bond percolation.

5.2 The Sznajd Model

In this chapter we will deal with the, so called, Sznajd Model. It tries to capture some imitative features of human social behaviour. In fact, it is well understood, within human societies, that it is generally easier to change someone's opinion by acting within groups than by acting alone. For example, a single person stopping on the street and staring at the sky is usually ignored (or perhaps considered eccentric). However, if several people stare into the sky, they readily induce others to do the same.

Equally it is not unusual to come across door-to-door sale agents working in couples rather than on an individual basis. In democratic electoral campaigns, small groups (typically couples) of political party activists may visit, door-to-door, potential electors and seek to gain their votes. Trade union movements often try to coordinate their actions in order to strengthen their position against management than if everybody tries to negotiate it alone. The underlying principle, in all of the above-mentioned examples may be captured by the famous Abraham Lincoln's injunction: "United we stand, divided we fall". This principle was developed into a computational model by Sznajd [19].

A simple version of the Sznajd model may be implemented on a two-dimensional lattice (the Sznajd model was originally implemented on a one-dimensional lattice, but for our purposes, the two-dimensional is the simplest non trivial version). Each site carries a spin, S, that may be either up or down. This represents either positive or negative opinions on any question. Two neighbouring parallel spins, representing for instance two neighbouring people sharing the same opinion in the model, are able to convince their neighbours of this opinion. If these two neighbours are not parallel or in step, then they have no influence on their neighbours. Allowing such a system to evolve from one time step to another via a random sequential updating mechanism always leads to a complete orientation of spins after a sufficiently long time which for the social systems is analogous to complete consensus, provided that the initial net orientation of spins is greater than zero [17].

5.3 Sznajd Model and Vote Distributions

Interestingly, the distribution of votes obtained from the simulation of a multi-dimensional Sznajd Model (in which each candidate is associated to one dimension, i.e. a possible opinion) on a cubic lattice show a good agreement with that obtained from the results of the local Brazilian elections. Bernardes & al. [34] used a modified version of the Sznajd model in which a pair of neighbours in agreement convinces its ten nearest neighbours to the same opinion. A cubic lattice of size $L \times L \times L$ represents the set of voters. A number N_{tot} of candidates, $N_{tot} = L^3$, is fixed in the beginning of the simulation. The value $n = 1, 2, \ldots, N_{tot}$ of a site S on the lattice means that this voter prefers the candidate n. The implementation of the model goes through two stages. First, the initial conditions are produced and, after that, the simulation of the electoral campaign (only voters can influence other voters) is performed.

As in real elections, the candidates have different initial chances of being voted for (representing more money for electoral campaigns, more initial visibility etc.). This is modeled by a probability P_c of convincing, calculated from the label n of the candidate

$$P_c = (n/N_{tot})^2 \tag{5.12}$$

It means that the higher is the label n of a candidate, the higher is the probability of convincing a voter.

In the first stage, all the sites have value zero, meaning that there are no committed voters. Then, all the sites are visited exactly once, in random order. For each visit, the voter is 'invited' to adopt a candidate, chosen at random. A random number r is generated and compared with P_c.

If $r \leq P_c$ the candidate is accepted by that voter. If the candidate convinces the voter, this voter tries to convince the neighboring sites. Once again, a new random number is extracted and compared with P_c. If successful, i.e. $r \leq P_c$, the voter tries to convince the neighborhood as follows: all the six neighboring sites are checked; for each that has the same value of the candidate chosen before, all the ten neighbouring sites of this bond of two sites assume the same value (as in the usual Sznajd prescription). If nobody has chosen the same candidate, only the originally selected voter is committed to this candidate.

In the second stage, the usual Sznajd process is performed: going to random sites on the lattice, a neighboring site is chosen at random and if the two sites have the same value (they prefer the same candidate), all the ten neighbors change in favor of that candidate.

The simulation results are compatible with the hyperbolic law

$$N(v) \propto 1/v \tag{5.13}$$

observed in Brazilian and Indian elections, in which a fraction $N(v)$ of the total number of candidates have v votes each. Deviations from a simple power law for both very large numbers of votes and very small numbers are observed.

In the same paper a similar result is obtained simulating the model on a scale free network.

5.4 Synchronous updating mechanism

If the random sequential updating is replaced by a synchronous updating mechanism the possibility of reaching a consensus is reduced quite dramatically [18].

The introduction of a synchronous updating rule is justified when contradictory information, that has to be processed simultaneously, is

present and when the effects of a single interaction may last for a certain time and conflict with the effects of a subsequent interaction.

The updating is performed by going systematically through the lattice to find the first member of the pair, then choosing randomly the second member of the pair within the neighborhood of the first. Having in this way completed the assembly of couples, each agent then orients her/himself according to her/his neighbours at time step t . Like-minded couples will induce their neighbours to turn to the same state (opinion). However a single agent may often belong, simultaneously, to the neighbourhood of more than one couple (of likeminded agents). In this case, if the couples have different opinions, she/he doesn't know what to do (frustration) and ends up doing nothing, i.e. sticks with her/his previous opinion. Frustration may prevent the system from reaching total consensus.

5.5 Memory

A feature of the agents in the models discussed above is the complete absence of memory. The past plays no role. We now assume that agents are endowed with memory [58]. For the sake of simplicity, they are all thought to have the same memory span T and updating mechanism is synchronous. Agents are keen to change opinion when in the neighbourhood of a like-minded couple, as in the models above. An agent resorts to her/his individual history when frustration occurs. In that case, the new state turns out to be the most frequent of her/his own $T+1$ most recent (T accounts for the past, one for the present). In an n-dimensional case (n opinions), this rule proves to be more efficient when T is an integer multiple of n, because that guaranties the existence of one opinion more frequent than the others.

5.6 Phase transitions

Let's now draw an analogy with the ferromagnetic systems. We may then associate a spin (either pointing up or down) to each agent. In a two-dimensional Sznajd model with asynchronous updating, the system always ends up at a fixed point: either all spins point up or they all point down. When synchronous updating is chosen things are different. The system converges to a fixed point (all spins up or all spins down), only if the asymmetry (absolute difference) in the initial distribution of opinions is above a critical value M_c that depends not only on the size of the lattice, but also on the memory length T [fig.5.2-5.6].

Fig.5.7 shows M_c as a function of L (the lattice side size). The slope decreases as T increases.

For small values of L the three curves displayed seem to converge to the same point. This is easily understood because for small lattices one expects the impact of frustration on total consensus to be small. As a consequence, the role of the memory T is not as important as it would be for large values of L.

For large L the curves monotonically decay to zero, this suggests that asymptotically (for infinitely large lattice) no consensus is possible.

Fig.5.8 shows M_c as a function of T. The slope of the two curves appears to be independent of L - an aspect that requires further study.

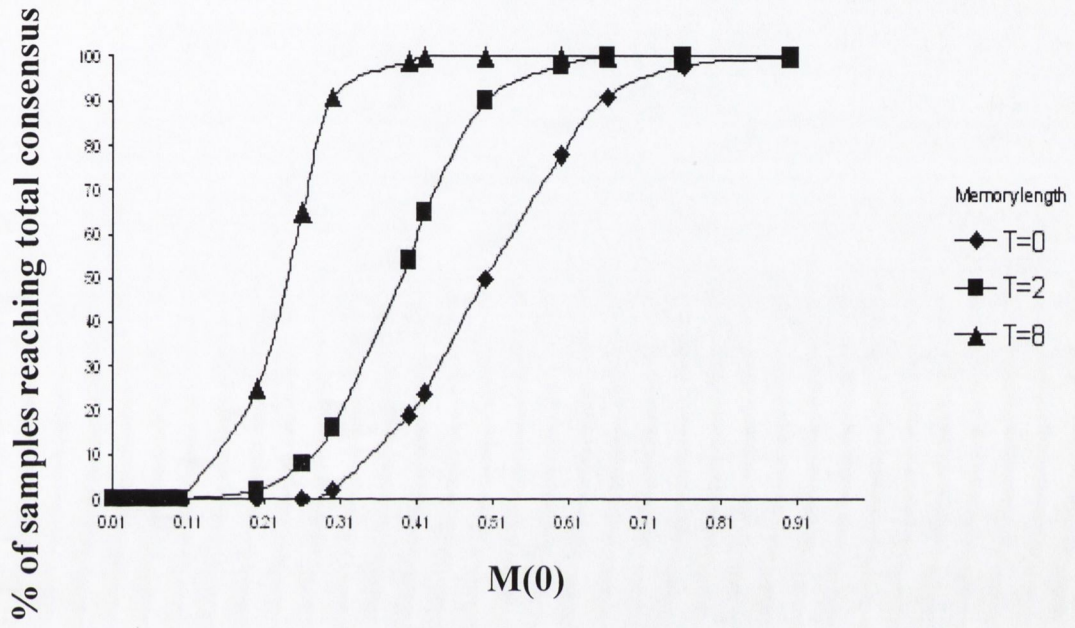

Figure 5.2: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =17. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

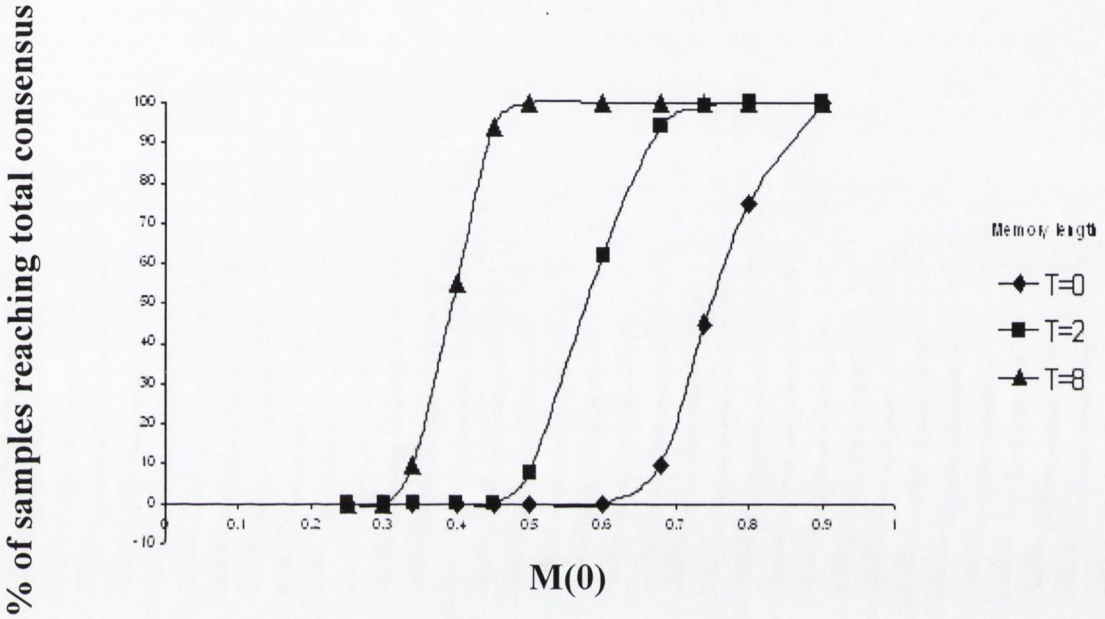

Figure 5.3: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =101. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

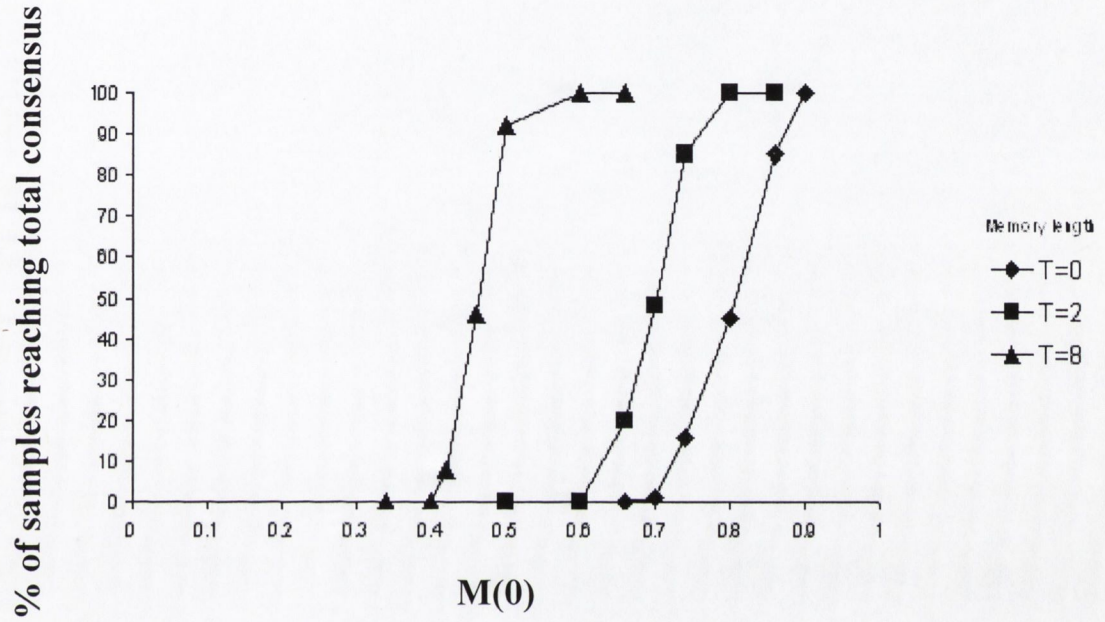

Figure 5.4: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =301. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

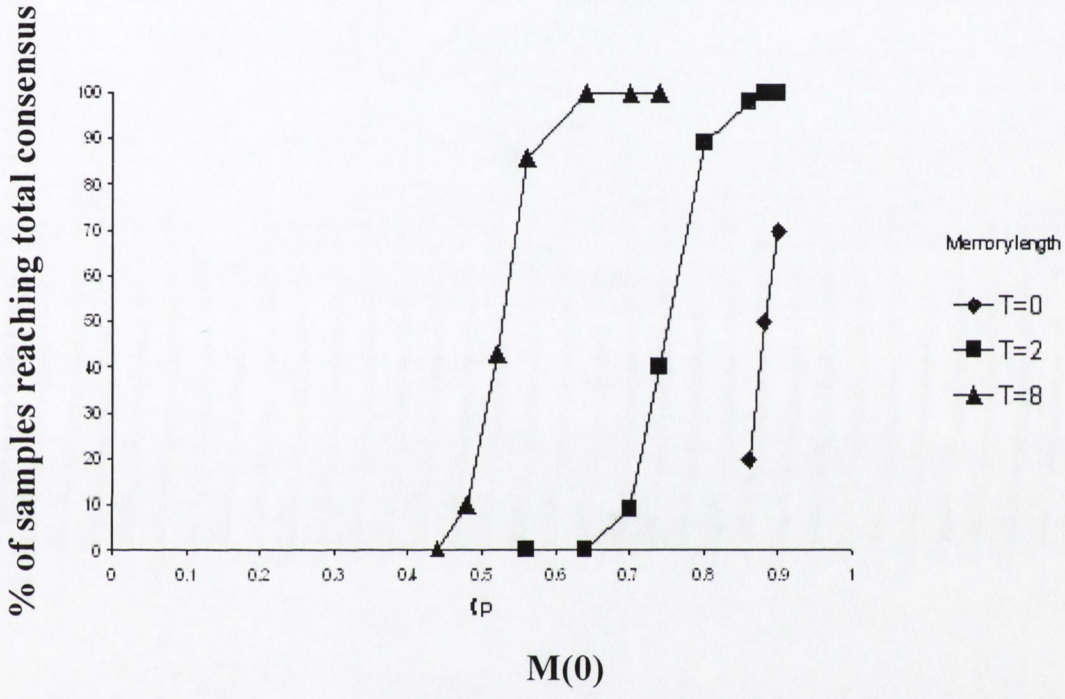

Figure 5.5: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =1000. The memory length T=0, 2, 8. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

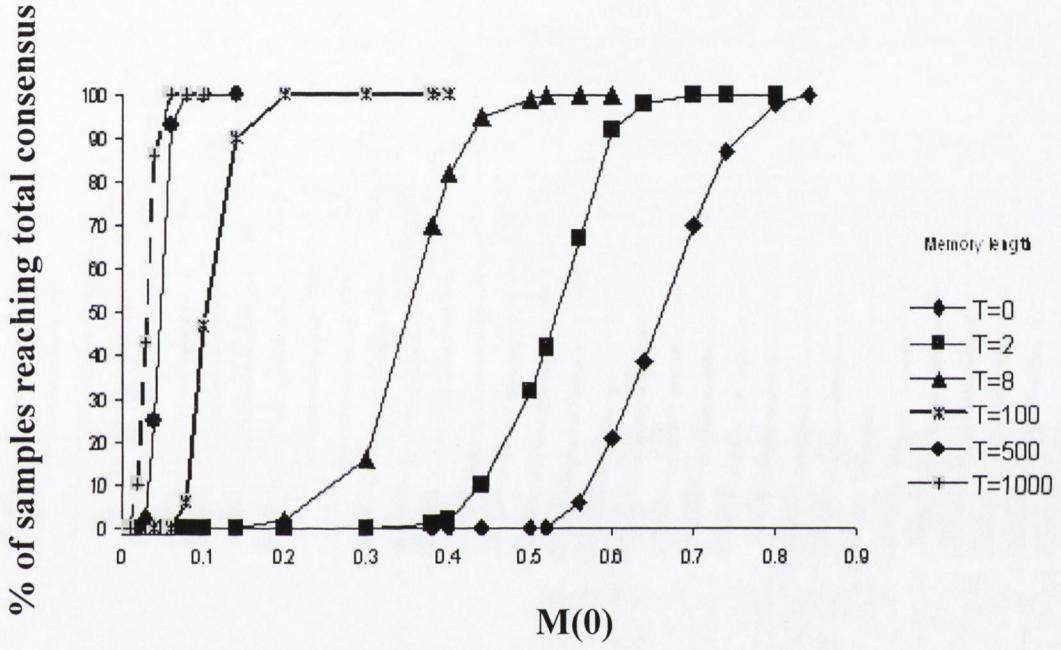

Figure 5.6: The phase transition from a no-consensus state to a total consensus state is driven by the value $M(0)$ of the magnetization at time zero. The lattice linear dimension is L =50. The memory length T=0, 2, 8, 100, 500, 1000. The transition point is shifted towards zero as the agent memory length T increases. The error bars are of the same size as the point-markers.

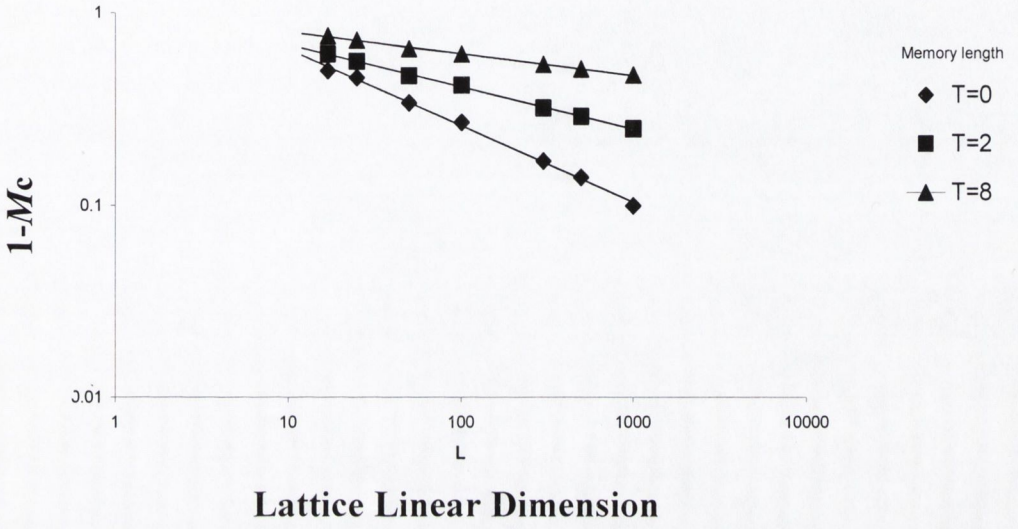

Figure 5.7: Variation with L of $1-M_c$ (one minus the absolute value of the difference in the initial probabilities for +1 and-1) for which in half of the cases a consensus was reached. That may be seen as the phase transition point from the state without consensus to the state with consensus. The estimated slope is --0.39 for T=0, - 0.21 for T=2, - 0.11 for T=8. The error bars are of the same size as the point-markers.

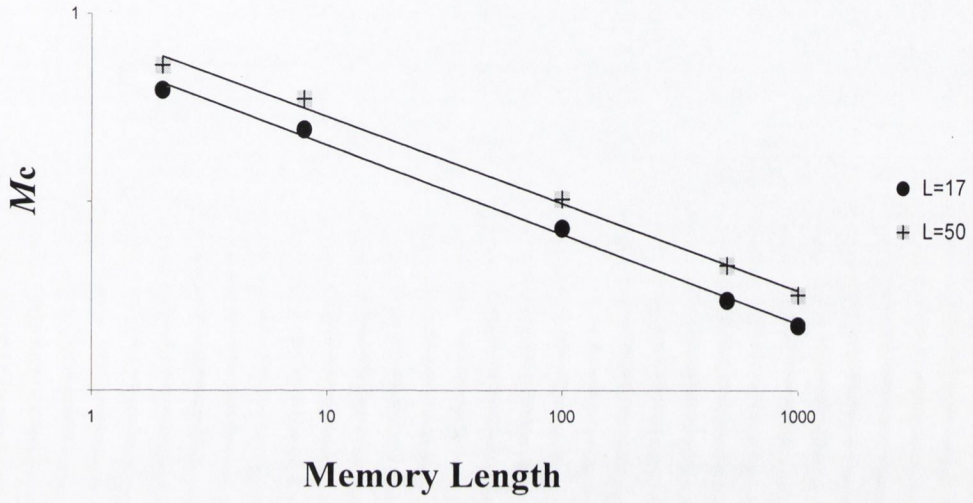

Figure 5.8: Variation with T of the difference in the initial probabilities for which in half of the cases a consensus was reached. That may be seen as the phase transition point from the state without consensus to the state with consensus. The estimated slope is 0.46 for L=17 and 0.47 for L=50. The error bars are of the same size as the point-markers.

5.7 Synchronous updating in the Sznajd Model with added noise.

We examine now the influence of noise in the two dimensional Sznajd model where updating is synchronous [59].

The noise here accounts for:

- External factors that may influence a specific agent.
- Differences in the way agents react to their environment and deviations from the general rule of interaction.
- A looser connection of some agents with their environment (i.e. a higher degree of independence)

We define the Magnetization in the following way

$$M = \frac{|N_+ - N_-|}{N_+ + N_-} \tag{5.14}$$

where N_+ is the number of up spins and N_- is the number of down spins and neglect memory effects. During the evolution of the system from the initial random state into a stable configuration, we allow each spin to flip randomly with probability q where $0<q<0.5$.

This is implemented applying first the noise-free rule to the whole lattice, then going systematically through the lattice and allowing each site a further flip (with probability q). This two-step updating completes one time step.

Figures 5.9 shows the effect of temperature or noise on the final magnetization, $M(\infty)$ for lattice size of 50. Figures 5.10 and 5.11 are similar results for lattice sizes of 100 and 250 respectively.

When $M(0)> M_c$ the final orientation, $M(\infty)$ decreases monotonically as q increases. When $M(0)< M_c$ the final orientation $M(\infty)$ displays non-monotonic behaviour as a function of q. For very small values of q (typically below 0.005), M increases and may reach values close (but still below) 1, for intermediate values of q (roughly between 0.006 and 0.06)

$M(\infty)$ may decrease slowly or even display oscillations depending on the value of the initial magnetization. For larger values of q, $M(\infty)$ quickly decays to zero. These results have been obtained using Monte Carlo simulations and averaging over ensembles of 500 realizations for each set of the three parameters $\{L, M(0), q\}$. We will call this effect 'different-is-uniform effect', in fact this results suggest that allowing the agents to randomly change their state (and be in that sense 'different') may result in a higher degree of uniformity in the whole system.

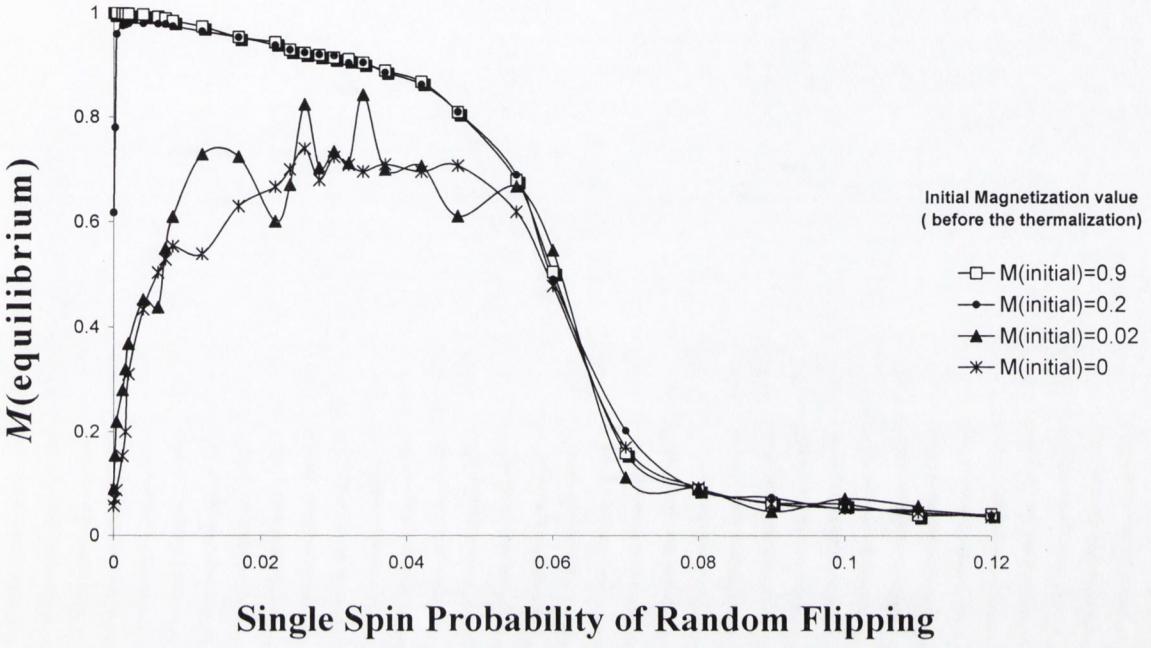

Figure 5.9: Magnetization at equilibrium as a function of the probability q (for a spin to flip randomly at each time step of each run of the simulation), for L=50 and initial magnetization values 0.9, 0.2, 0.02 and zero. The error bars are of the same size as the point-markers.

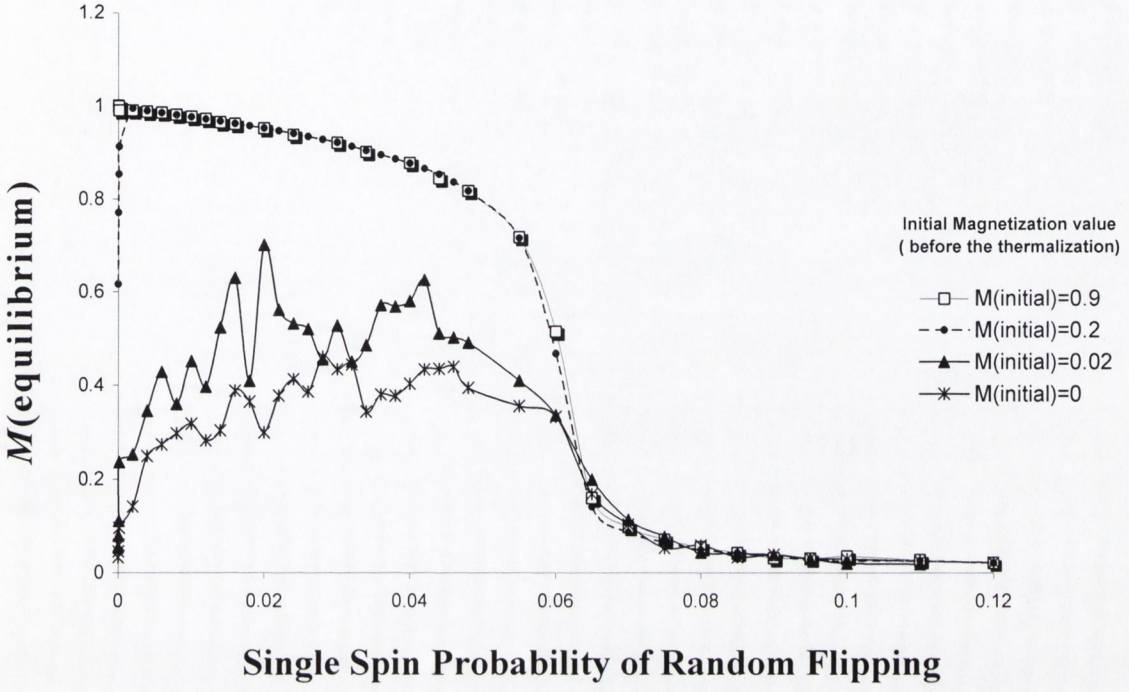

Figure 5.10: Magnetization at equilibrium as a function of the probability q (for a spin to flip randomly at each time step of each run of the simulation), for L=100 and initial magnetization values 0.9, 0.2, 0.02 and zero. The error bars are of the same size as the point-markers.

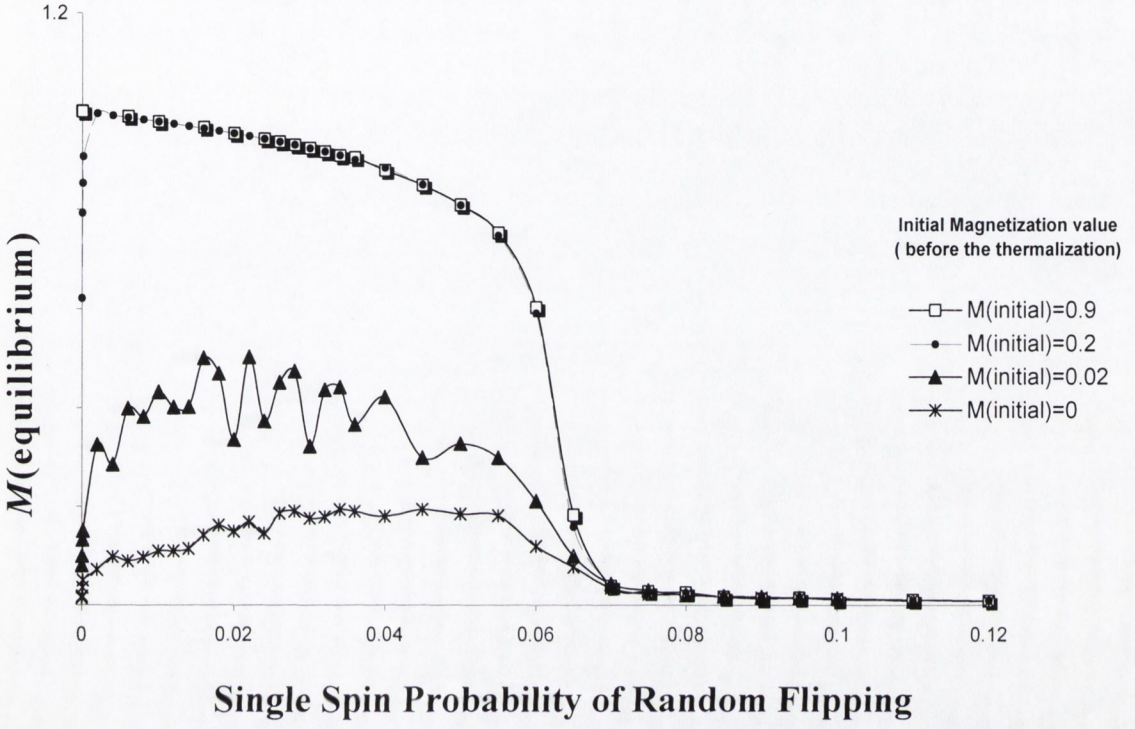

Figure 5.11: Magnetization at equilibrium as a function of the probability q (for a spin to flip randomly at each time step of each run of the simulation) , for L=250 and initial magnetization values 0.9, 0.2, 0.02 and zero. The error bars are of the same size as the point-markers.

5.8 Microstructure

To gain further insight into this non-monotonic behaviour, we began with a randomly selected configuration of spins on the lattice, having a fraction p up and $1-p$ down (say $p=0.4$, i.e. initial magnetization $M(0)=0.2$), and no added random noise ($q=0$). After only a few hundred iterations, the system evolved into an archipelago-like structure. Spins could be observed aligned in the manner of islands located in a sea of opposite orientation. For values of p close to 0.5, the 'landscape' is irregular and jagged. The geometrical structure of these small islands is now crucial. Spins surrounded by four others of the same sign (so forming cross-shaped islands) form a stable state since, at any time, they are prevented, at zero temperature, from flipping by frustration. Islands made up of one or more copies (even overlapping) of such a cross-shaped structure are destined to last and no total orientation can be reached in the system.

However now the addition of random noise can flip a fraction q of spins, at each time step. The occasional flipping of one spin within a cross-shaped island now removes the frustration and breaks the previous stability of the island structure. Equally some new cross-shaped islands may also be randomly created. M thus is now dependent on the balance between creation and destruction of cross-shaped islands. The effect is illustrated by figures 5.12 and 5.13 which represent the microstructure at two consecutive time steps. The area located by the arrow show how an island has become destabilized by the noise.

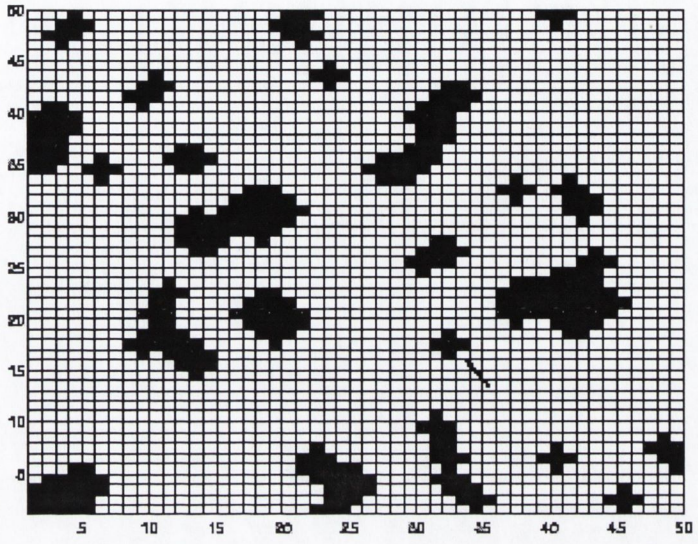

Figure 5.12: This shows the microscopic detail of a calculation for a lattice of linear dimension L=50, a probability of random flipping q=0.001 and an initial magnetization M(0)=0.2.The arrow points to the 'cross-shaped' island (positive spins, black) centred in the site {33,17} and surrounded by the sea (negative spins, white) that has formed after 396 time-steps.

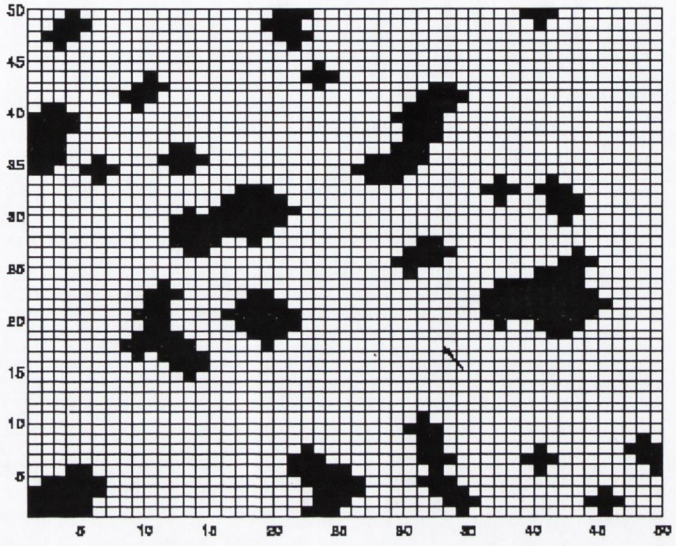

Figure 5.13: This shows the microscopic detail for the calculation shown in figure 11 at the subsequent time step. The random flipping of spin {33, 16} (see the arrow) has now eased the frustration, it would otherwise exhibit at zero temperature, destabilising the 'cross shaped' island. In this particular case it took just one time-step for the other four positive spins forming the island to be flipped via interaction with the surrounding negative spins. In general that may take up to tens of steps.

5.9 A toy model for market volumes

It is well understood that collective behavior may play a key role in markets [6][9][13] [15]. Trends, bubbles and crashes may be seen as an expression of that on a macroscopic level. If we assume that market movements are mainly determined by few big players, such as the main Mutual Funds, able to move very large asset volumes, a key problem is to understand the very nature of volume fluctuations.

We focus on the process that makes the owners of fund quotes decide whether to buy or sell, trying to model it in terms of opinion dynamics. We then assume the balance between the buyers and sellers of fund quotes to determine the volume of the buy (or sell) order that the fund manager will place in the market.

Think of a market dominated by a single big investor, say an investment fund.

After a transaction initially each of N fund share traders is as likely to be a potential buyer as to be a potential seller. These N agents interact over a certain time before they make a new deal. Here we assume that to happen through the simultaneous Sznajd-like interaction described above. The average strength of the mutual influence (inversely measured by the probability of random fluctuations for a spin) may change from a transaction to the next, driven by random factors and by the risk perception.

A simple choice is to assume, for each time step:

$$\Delta q(t) = \sigma \frac{\left| |D(t)| - B(t-1) \right|}{B(t-1)} \eta(t) \tag{5.15}$$

$$\eta(t) \in N(0,1)$$
$$D(t) \sim \quad M_t(\infty)$$
$$B(t) = \alpha B(t-1) + (1-\alpha)|D(t)|$$

Where the demand $D(t)$ at time t is given by the asymptotic value of the Magnetization $M_t(\infty)$ at the t-run of the simulation; B is an exponentially weighed 'benchmark' for the demand and $1/(1-\alpha)$ is a measure of the average agent memory length for the demand [60].

This is justified by recent studies [7] that suggest the impact function (price changes as a function of traded share volumes) to have power law behaviour with an exponent value typically ranging from 0.2 to 0.5.

The assumption (made at the beginning of this paragraph) of a single big market player makes B also a measure for the market volatility (since volumes grow as some power (between 2 and 5) of the absolute returns [11]).

Fig. 5.14 and 5.15 display the results of a simulation of this model. The exponents found for the Cumulative Distribution Function (between 1.4 and 1.6) and for the Auto-Correlation Function (between 0.3 and 0.4) are compatible with the exponents typically found for the time series of financial market volumes. For instance, for the New York Stock Exchange [6] the exponents found range in between 1.4 and 1.7. for the c.d.f. and between 0.3 and 0.38 for the ACF.

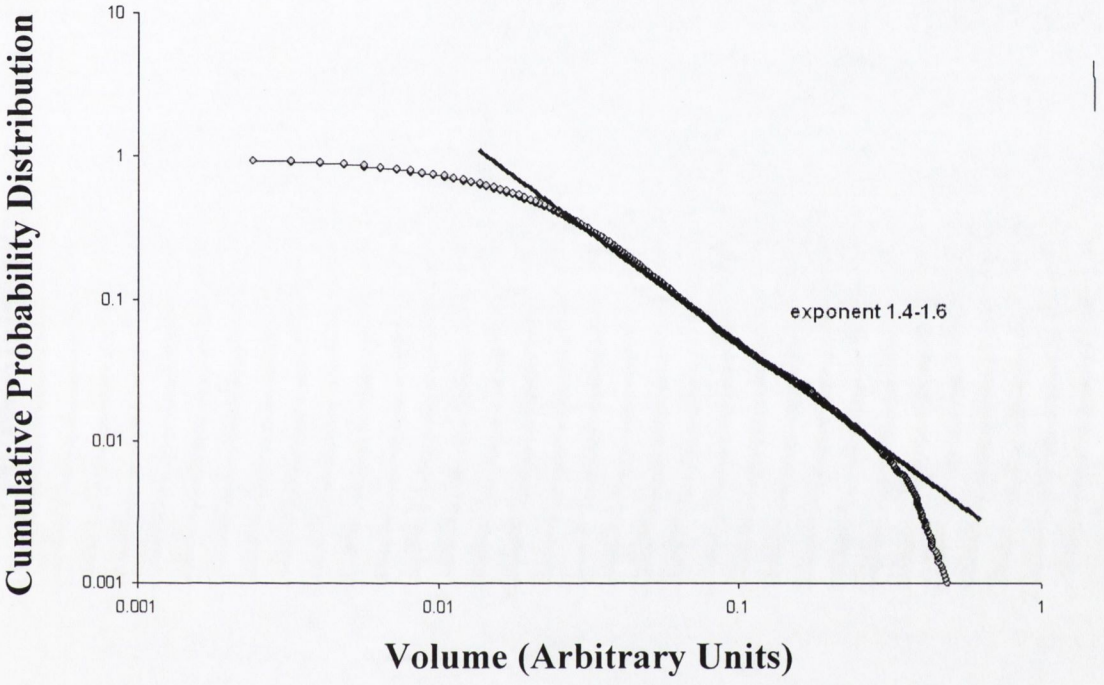

Figure 5.14: Cumulative probability distribution for the simulated demand (volume of shares). The decay follows a power-law behaviour with an exponent (between 1.4 and 1.6) close to the value of 1.5 found in NYSE volume distributions. Number of transactions (data points) =20000, $\{\alpha=0.98, \sigma=0.005$, N=2500$\}$. The error bars are of the same size as the point-markers.

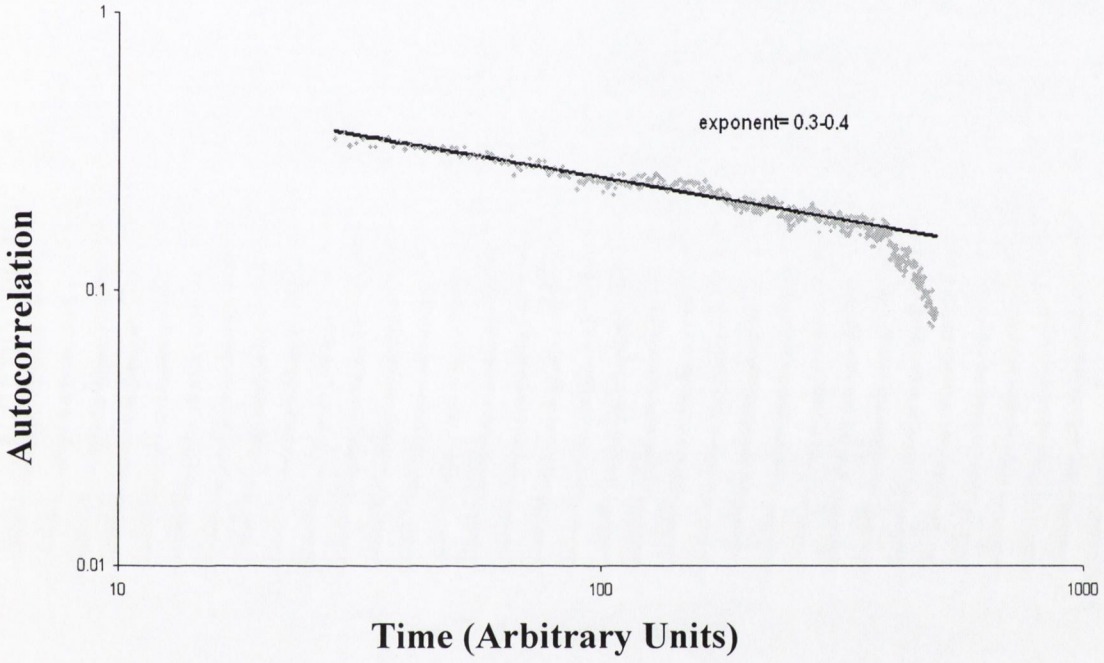

Figure 5.15: Autocorrelation function for the simulated demand (volume of shares). The chart displays a power law decay with an exponent smaller than 1. Number of transactions (data points) = 20000 $\{\alpha=0.98,\ \sigma=0.005,\ N=2500\}$. The error bars are of the same size as the point-markers.

CHAPTER 6

COMPLEX NETWORKS AND THE DEFFUANT MODEL

6.1 Introduction

Lattice models have the gift of (a relative) topological simplicity, but fail to capture the complex topology of socially organized systems.

Systems of high technological and intellectual importance such as the Internet, human societies, biological organisms, ecological systems and the electrical power supply network present complex web-like structures.

Here we outline the basic descriptive properties of networks and some models able to capture the most relevant properties of real social networks (namely the scale-free character of the degree distribution and the small-world aspect). We describe the Deffuant Consensus Model implemented on a scale free network and we modify it introducing two simple adaptive versions.

6.2 Networks

6.2.1 Definitions

Following a recent review paper from Barabasi and Albert [36] in this paragraph we outline very briefly the properties of Networks. In mathematical terms, a network is represented by a graph. A graph is a pair of sets $G = \{P, E\}$, where P is a set of N nodes $P_1, P_2 ....., P_N$ and E is a set of edges that connect two elements of P.

Given a node i connected to k_i other nodes, the clustering coefficient of node i is:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{6.1}$$

where E_i is the number of edges between the k_i neighboring nodes and $\frac{1}{2} k_i(k_i - 1)$ is the number of edges that would exist between them if they were part of a clique (each one of k nodes were connected to all the other k nodes).

The clustering coefficient of a network is given by the average clustering coefficient (the average is performed over all nodes).

The distribution $P(k)$ of the number of neighbors is called degree distribution and happens to be a power law distribution for many real networks.

The average number edges connecting two randomly chosen nodes in a given network defines the average path length of that network.

6.2.2 Scale free networks

Many real networks are (approximately) scale free (SF).

In fact, their degree distribution follows a power-law for large k (up to a certain cut-off value). Furthermore, even for those networks for which $P(k)$ has an exponential tail, the degree distribution significantly deviates from a Poisson.

The problem of the origin of the power-law degree distribution in networks was first addressed by Barabasi and Albert (1999) [4].

They introduced the Scale-free Model (SF), based on two ingredients: growth and preferential attachment.

1. Growth: Starting with a small number (m_0) of nodes, at every time step we add a new node with m ($\leq m_0$) edges that link the new node to m different nodes already present in the system.

2. Preferential attachment: When choosing the nodes to which the new node connects, we assume that the probability Π that a new node will be connected to node i depends on the degree k_i of node i, such that $\Pi(k_i)=\frac{k_i}{\sum_j k_j}$. After t time steps the outcome of this algorithm is a network with $N=t+m_0$ nodes and mt edges, evolving into a scale-invariant state. The probability for a node to have k edges follows a power-law behaviour with an exponent (independent of m) $\gamma_{SF}=3$.

Some of the properties of SF networks are still to be fully unveiled. According to simulations, the average path length in a SF model is smaller than in a random graph and grows approximately logarithmically with the network size N. This means that the scale free topology is more efficient in bringing the nodes close than the random network topology. Non trivial correlations are found between the degree of connection of nodes connected by an edge.

The clustering coefficient is many times higher than that of a random network. Differently from the Small world model (in which the clustering coefficient is independent of N) and from the random graphs (in which the clustering coefficient decays as N^{-1}), the clustering coefficients of the Free Scale model decreases with the network size following approximately a power law $C \propto N^{-0.75}$.

The spectral density $\rho(\lambda)$ of the scale-free model is continuous, but it is significantly different form the semicircular shape found for random graph spectral density.

Numerical simulations indicate that the bulk of $\rho(\lambda)$ has a triangle-like shape with top lying well above the semi-circle and edges decaying as a power-law. This power-law decay is due to eigenvectors 'localized' on

the highest degree nodes. As for random graph (and unlike small world network models), the principal eigenvalue, λ_1, is clearly separated from the bulk of the spectrum. A lower bound for λ_1 can be given as the square-root of the network's largest degree k_1. Since the node degrees in the scale free model increase as $N^{1/2}$, it results that λ_1 increases approximately as $N^{1/4}$. Numerical results indicate that λ_1 deviates from the expected behaviour for small network sizes, reaching it asymptotically for $N \to \infty$. This crossover would suggest the presence of correlations between the longest row vectors, offering additional evidence for correlations in the scale-free model.

The principal eigenvalue plays an important role in the moments of $\rho(\lambda)$, determining the loop structure of the scale-free network. In contrast with the sub-critical random graphs ($p < 1/N$), where the fraction of loops becomes negligible, in a scale-free network the fraction of loops with more than four edges increases with N and their growth rate increases with the size of the loop (Note that the fraction of triangles decreases with N).

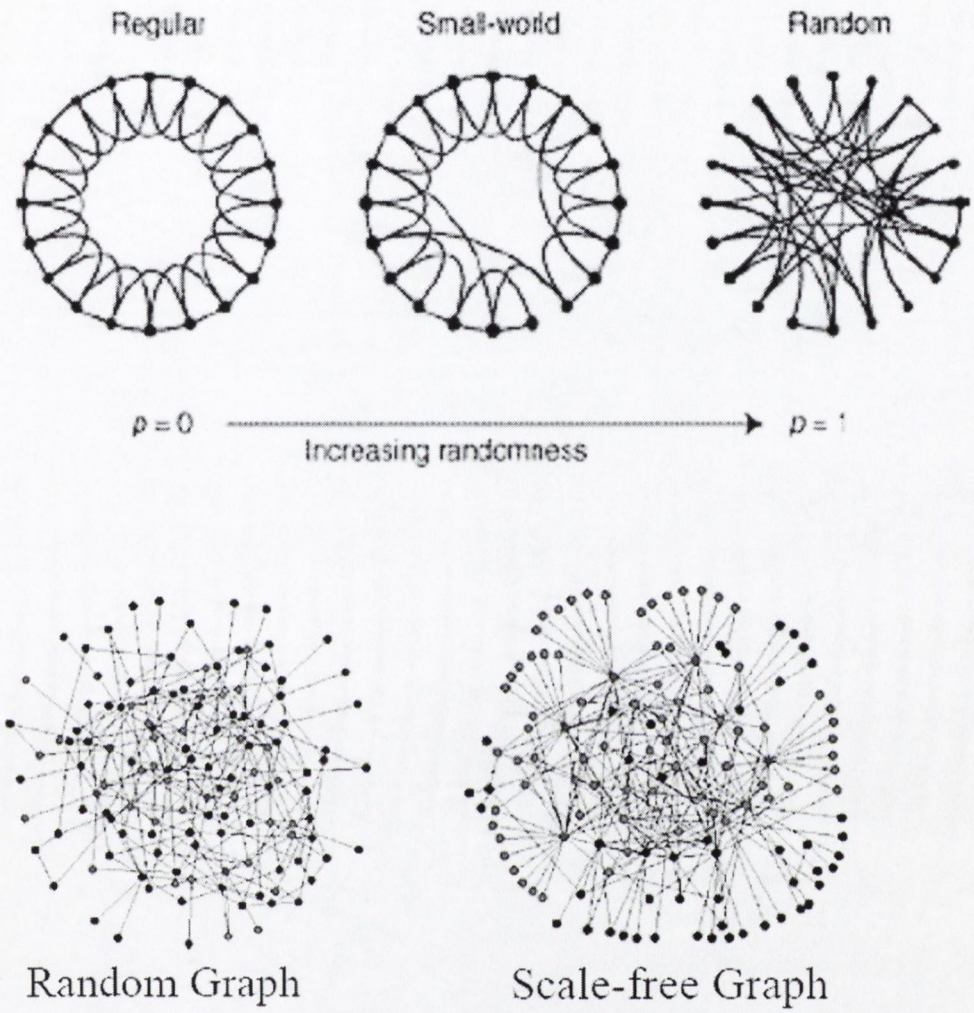

Figure 6.1: Networks. Above are displayed graphic examples of: a regular lattice, a small-world network and a random network. Small-world networks may be seen as a trade-off between the regular topological structure of a lattice and the random topological structure of a random graph.Below, we can appreciate the pictorial difference existing between Random and Scale-free graphs Moreover, in the random graph, the 5 most connected nodes are connected to 27% of all nodes. In the scale-free graph, the 5 most connected nodes are connected to 60% of all nodes.

6.3 The Deffuant Model on scale free networks

6.3.1 General overview of the Deffuant Model

Deffuant et co-workers [20] introduced a consensus model that was later implemented on a Barabasi Network by Stauffer and Meyer-Ortmanns [51]. The model is based upon the assumption that two agents are likely to agree if their initial opinion are not too different.

As a first step, Stauffer and Meyer-Ortmanns constructed a scale-free network (through the preferential attachment mechanism) and assigned to each site a random number S between zero and one as initial opinion. Then for each iteration, every site A was updated once by selecting randomly one site B from the sites connected with A (in the directed case the selection is made only from the m sites which A had selected as friends). If then the opinions S_A and S_B differ by more than a constant confidence bound ε between zero and one, A and B refuse to discuss and do not change their opinion. Otherwise both move closer to the position of the other by an amount:

$$\delta = \mu(S_A - S_B) \tag{6.2}$$

with in their simulations, i.e. A takes the opinion $S_A - \delta$ and B the opinion $S_B + \delta$. The parameter μ characterizes the flexibility in changing opinion.

Therefore ε may be interpreted as a measure for the tolerance of people to other opinions.

They find that for ε larger than about 0.4 a full consensus is reached; only one opinion survives. For smaller ε, no consensus is reached and the number F of fixed opinions increases with decreasing ε. For small ε, $F \propto N$ for large N.

A discrete-opinion version of the previous model was also implemented by Stauffer and Sousa [22].

Instead of allowing for the opinions any real number between 0 and 1, they took discrete numbers $q = 1,2,....Q$, as in the Sznajd model [17]. Only people differing by ±1 in their opinion can convince each other thus $1/Q$ corresponds to the confidence interval of the previous models. They find that for large N the number S of surviving final opinions roughly equals Q for not too small Q; for $Q = 2$, on the other hand, nearly always a complete consensus was found.

If, however, Q grows to values closer to N, then the finite size of the network is felt: S is lower than Q and approaches $N + m$, that means everybody keeps its own opinion and the simulation stops soon. A finite-size scaling formula

$$S = (Q-1) f(Q/N); \quad f(x \to 0) = 1 \qquad f(x \to \infty) = 1/x \tag{6.3}$$

fits reasonably the same simulated data, except for small Q.

6.3.2 Deffuant model on Barabasi Network with local relative confidence bound and time adaptive agents.

In it's original version, the Deffuant Model does not take into account the time evolution and the local dependence of the confidence bound.

In a second paper [52] Deffuant and al. propose a new model of interaction called relative agreement model, which is a variant of the previously discussed bounded confidence. In this model, uncertainty (quantified by the confidence bound) as well as opinion can be modified by interactions. They assume the uncertainty of an agent B to diminish after interacting with a more confident (smaller uncertainty) agent A and to increase after interacting with a less confident (larger uncertainty) agent C.

We propose two further models [61] in which agents take discrete opinions (integer number from 1 to Q) and can change their confidence

bound. The confidence bound measures the individual degree of tolerance. Agents tend to optimize their degree of tolerance (confidence bound). Loosely speaking, they try to find an agreement with their neighbours, without conceding more then what is needed. This assumption seems to apply to many social situations ranging from the choice of trading partners and political allies to the choice of sexual partners. In social systems the keenness of individuals to agree with someone else is often conditioned by the environment. For instance, people with quite different political view are more likely to sink their differences when are surrounded by extreme minded people. The opposite is also true. In many 'moderate' political systems parties having very close policies may often engage in fierce battles.

We can call this relative expectation models (REMs).In the first REM, we assume that, at each time step, each agent works out an individual confidence bound based on the individual experience and neighborhood. The larger is the bound, the smaller is the expectation.

A very simple assumption, to implement this idea, is to take as a threshold, for a agent i at a time t, a fraction η of the average distance between agent i and it is direct neighbors, over the past time step.

Figures 6.2-6.5 display the results found.

The number of surviving opinions decreases as η increases. The ratio number of surviving opinions over number of agents decreases as the network size increases.

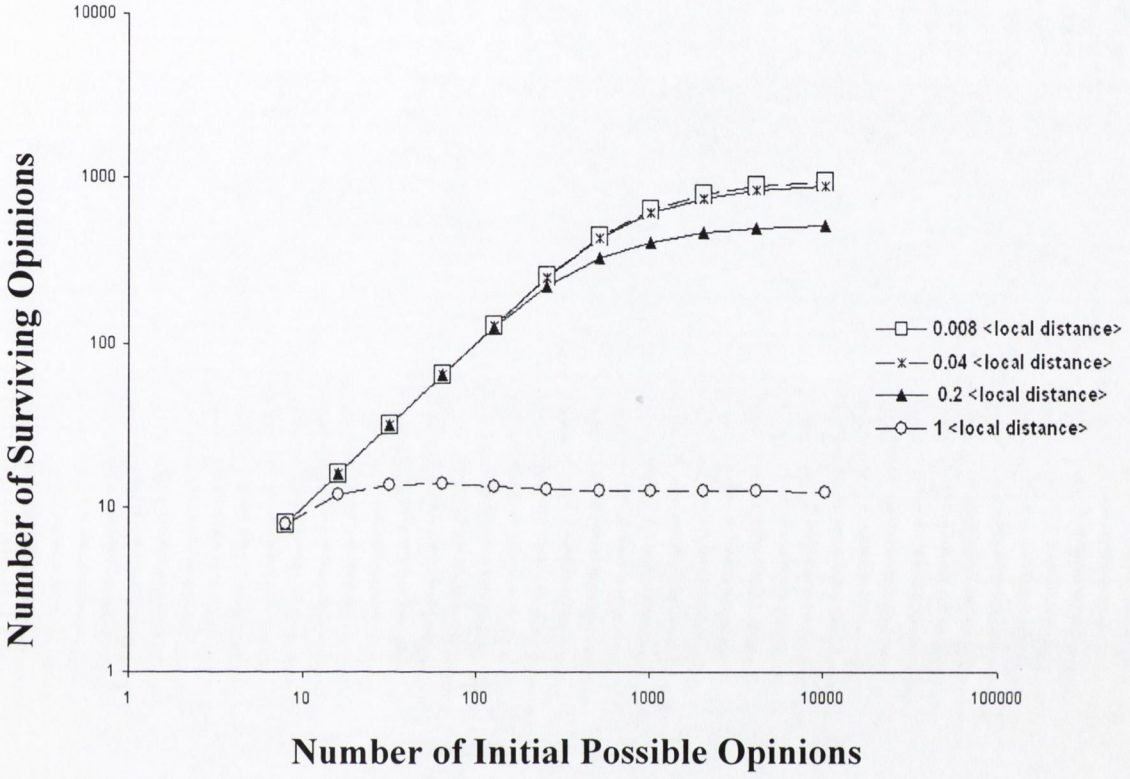

Figure 6.2: Number of surviving opinion as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is a fraction (1, 0.2, 0.04 and 0.008) of the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N=1000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

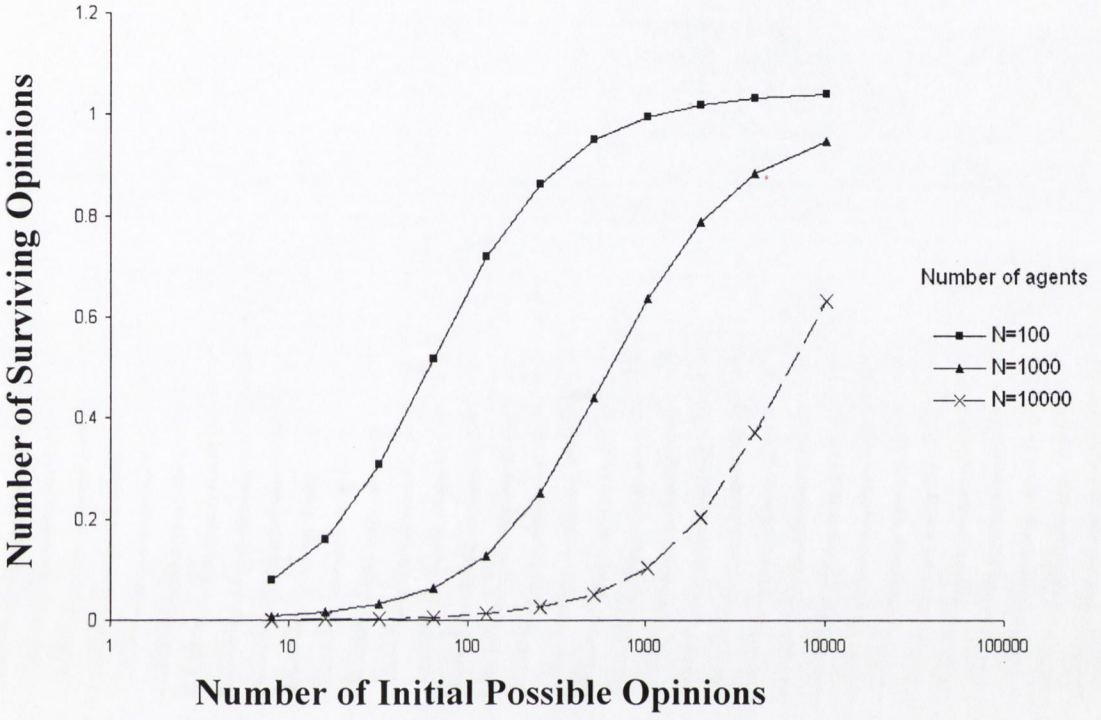

Figure 6.3: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is 0.008 times the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N=10, 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

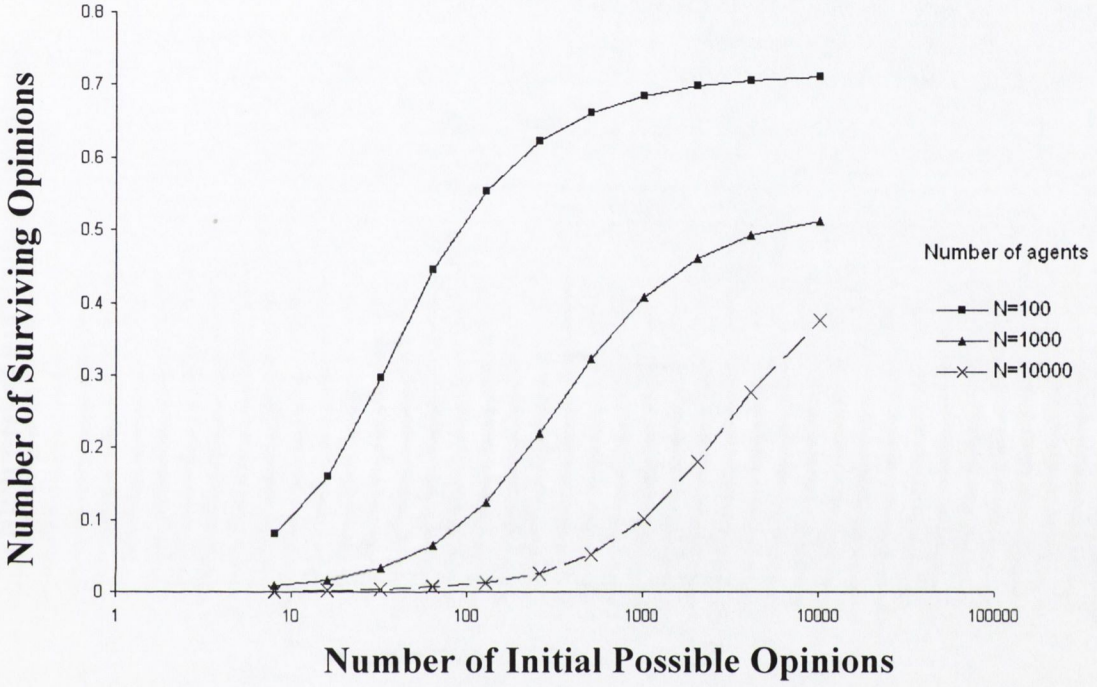

Figure 6.4: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is 0.02 times the average local distance. The averages are performed over the direct neighbourhood .The number of agents is equal to N= 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

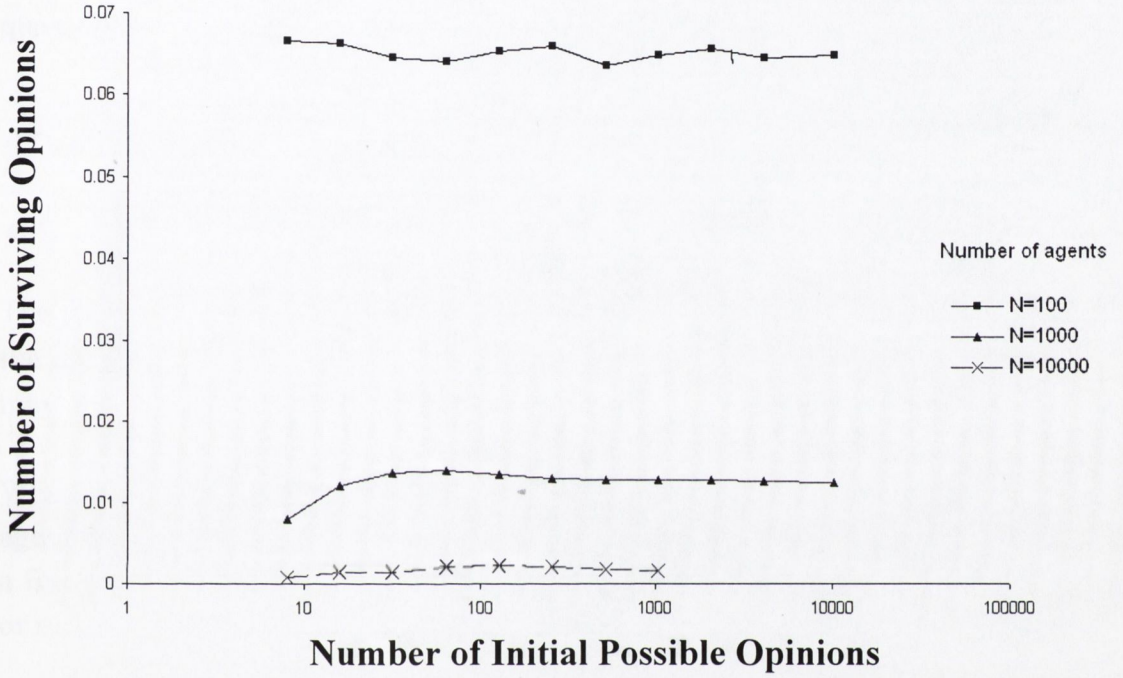

Figure 6.5: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in an adaptive Deffuant-like model. The consensus threshold is equal to the average local distance. The averages are performed over the direct neighbourhood. The number of agents is equal to N= 100, 1000, 10000. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

In the second REM, we assume, each individual confidence bound ε to increase each time there is no agreement, and to decrease each time there is an agreement of a quantity:

$$\Delta\varepsilon_A = \pm\frac{1}{\tau}\varepsilon_A$$
$$\varepsilon_A(t+1) = \max\left(1,\ \varepsilon_A(t) + \Delta\varepsilon_A(t)\right) \tag{6.4}$$

The second equation simply means that ε cannot be smaller than 1. At time zero all the agents have the same confidence bound, equal to 1. During the time evolution ε may get higher values.

When an agent A meet a neighbour B whose opinion is within the threshold value, agent A approaches B opinion of a quantity equal to the previous opinion value plus a fraction (0.5 in this case) of the relative opinion distance between A and B. B may or may not do the same, depending on its confidence bound.

We obtain the results displayed in figures 6.6-6.9, for N= 14, 104, 1004, 10004 and τ= 10, 50, 333, 2500. The relative number of surviving opinions, as a function of the initial number of opinions, presents a non-monotonic behavior. With a maximum occurring for the *initial number of opinions* Q approximately equal to the number of agents N.

When the number of initial opinions Q is much larger then N we find :

$$\frac{S}{N}\left(\frac{Q}{N} \to \infty, \tau, N\right) = d(\tau, N) \tag{6.5}$$

Where S is the Number of Surviving Opinions.

For Q much larger than N the number of surviving opinions is much smaller then for Q=N. This effect may appear counter-intuitive and remarkably different from what observed in the non-adaptive case, but it may be explained by the existence of a characteristic opinion-distance.

$$d_{opinion} = \left\langle \left| q_i - q_j \right| \right\rangle_{Q=N} = \frac{1}{3}(N+1) \tag{6.6}$$

When the average opinion-distance d is:

$< d_{opinion}$

($Q < N$), agents are more likely to have similar opinions and may easily reach an agreement;

$> d_{opinion}$

($Q > N$), agents tend to enlarge more their confidence bound to avoid isolation;

$\approx d_{opinion}$

($Q = N$), opinions are too distant for the agents to find an agreement, but also too close to promote an efficient adaptation process.

What observed may be called 'far-is-close effect', since agents start from a wider range and end up with a narrower range of opinions than in.

.

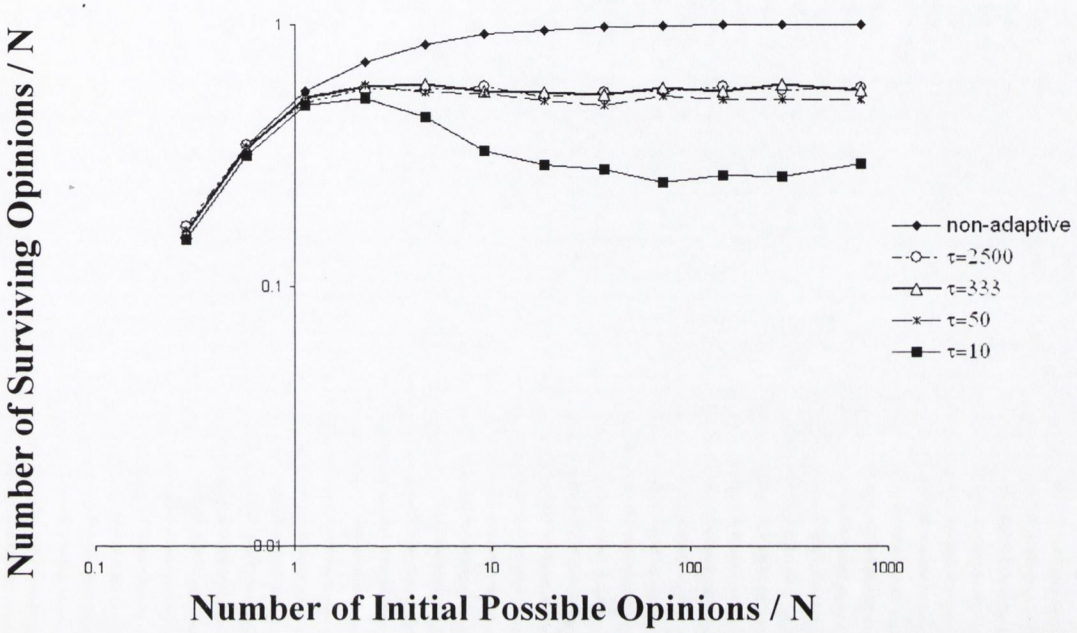

Figure 6.6: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=14. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

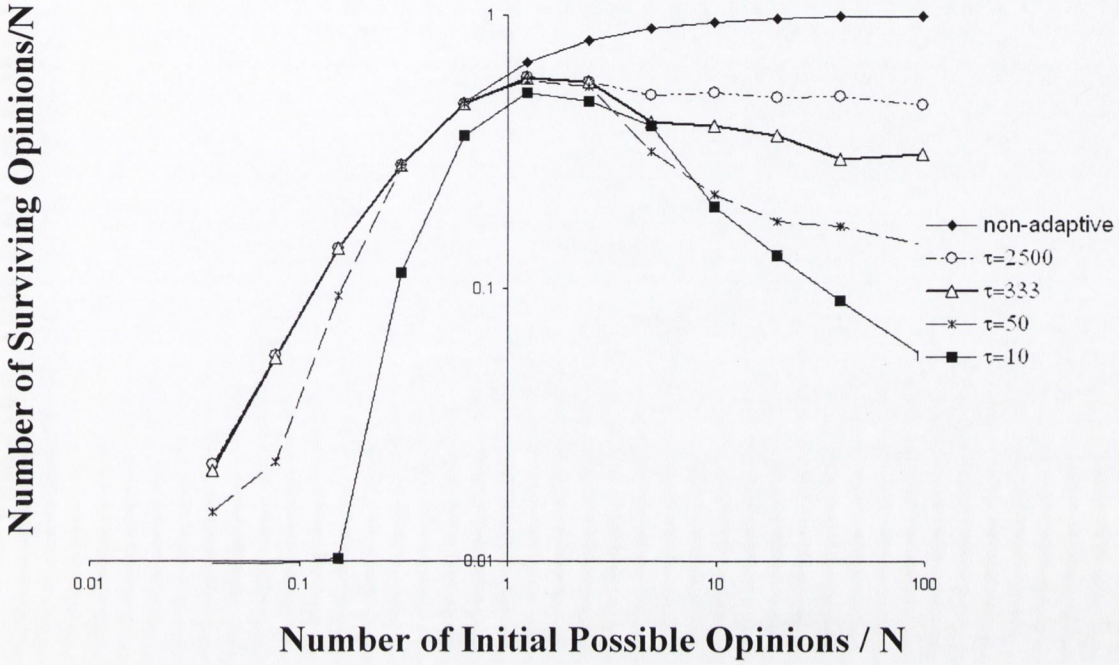

Figure 6.7: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=104. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

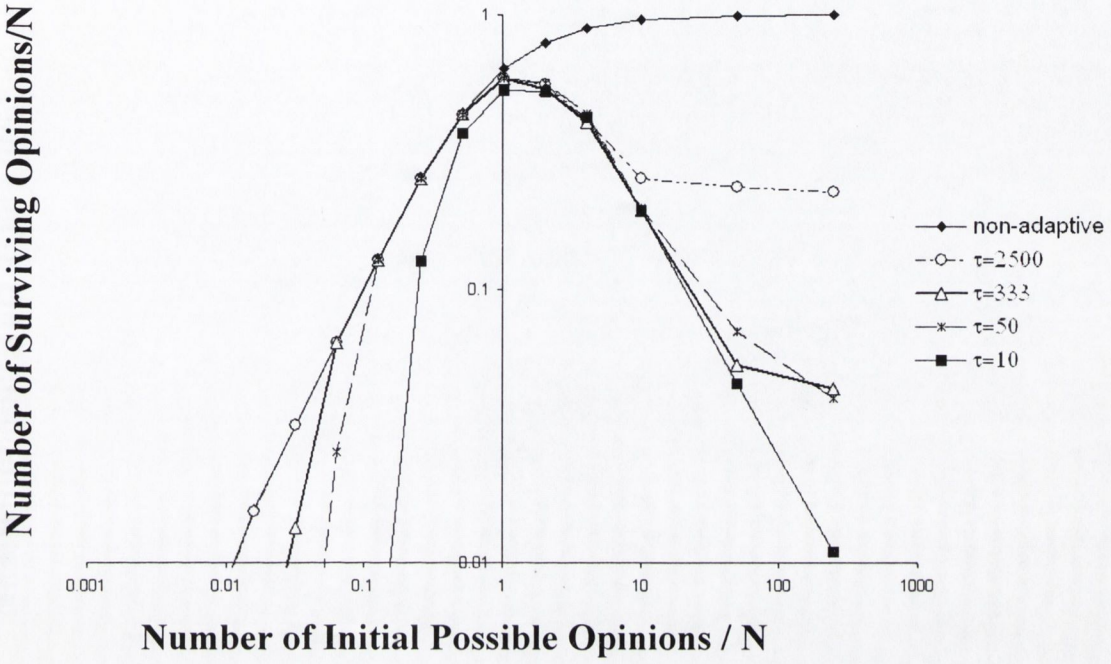

Figure 6.8: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=1004. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

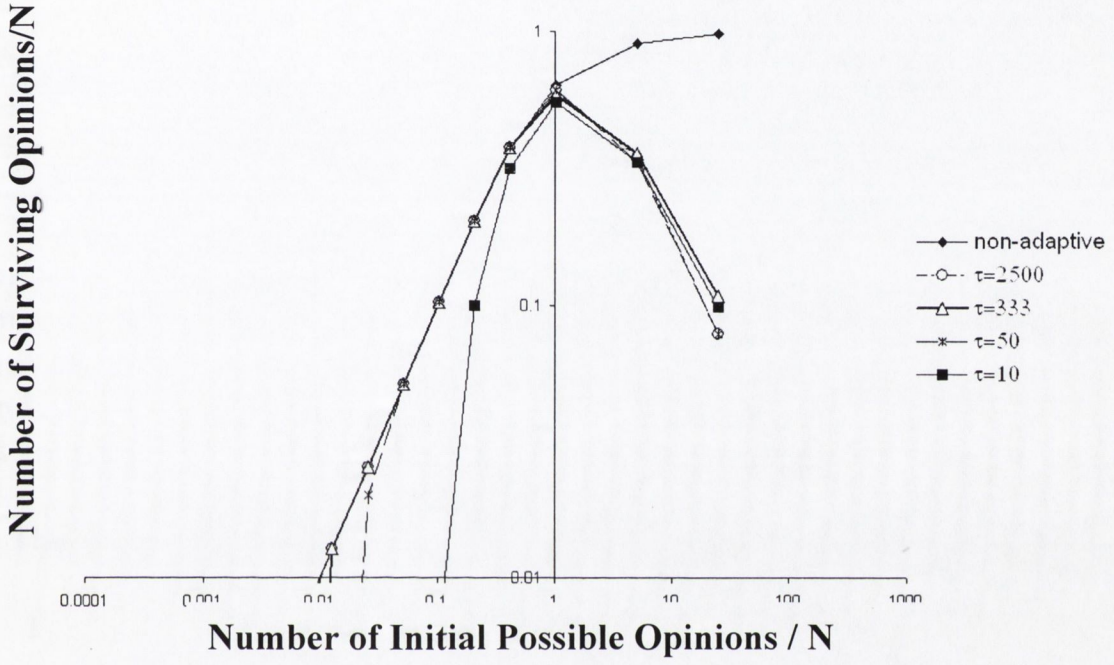

Figure 6.9: Number of surviving opinions divided by the number N of agents, as a function of the initial number of opinions in a time adaptive Deffuant-like model. The consensus threshold changes over time depending on the agent experience: it increases (decreases) of a factor τ depending on whether the agent is successful (or unsuccessful) in finding a like-minded neighbour. The results obtained for τ=10, 50, 333, 2500 are compared to those obtained for the non adaptive case. The averages are performed over the direct neighbourhood. The number of agents is equal to N=10004. Each agent has m=4 direct connections. The simulations have been carried out over 1000 samples. The error bars are of the same size as the point-markers.

CHAPTER 7

CONCLUSIONS

7.1 Summarizing

Understanding the relation between microscopic and macroscopic features of social self-organized systems is still an open challenge. In this thesis we have started from a simple Langevin-like approach and moved to agent-based modeling as it allows a better understanding of the underlying microscopic structure of socio-economic systems and enable us to reproduce the microscopic statistical features of those systems.

The results of this thesis may be summarized as it follows:

1. The waiting time distribution found for XIX Century Irish Stock Market can be described by a Mittag-Leffler function, obtained under the assumption of a power-law memory function (for the whole system). For some other markets (e.g. the XX century Japanese currency market) this model does not apply.

2. Generalized Lotka-Volterra (GLV) and Generalized Lotka-Volterra models with Peer Pressure (GLVPP) lead to power law probability distributions that (for an appropriate choice of the parameters) fit the distribution of personal wealth found in most economies.

3. The Sznajd Consensus model dynamics may be used to reproduce the market volume distributions observed in many financial

markets (e.g. New York Stock Exchange). A very interesting feature of the Sznajd model is represented by an apparent paradox (different-is-equal effect): allowing agents to change their state (or opinion) through random fluctuations may result in an increase of the degree of consensus within the system. In other words, in social systems in which this model applied, allowing individuals to be different form their neighbors may result in a higher uniformity.

4. Two simple adaptive versions of the Deffuant model, based on the 'relative expectation' of individual agents, have been studied on scale-free networks. The first, leads agents to adjust their expectation to neighbouring environmental states. The second, allows individual agents to adjust their expectations to their past performances (increasing their expectation when they do find someone they agree with, decreasing their expectations when they do not). Adaptive mechanisms play a key role in increasing the possibility for the system to converge onto few final opinions, starting from a wide range of initial opinions.
Particularly interesting is the second of these two models. The relative number of surviving opinions (number of surviving opinions over number of agents), as a function of the relative initial number of opinions (number of initial opinions over number of agents), presents a non-monotonic behavior. A maximum occurs for the initial number of opinions approximately equal to the number of agents N. A higher degree of consensus is reached not only for Q smaller than N, but also for Q much larger than N (far-is-close effect). This may appear counter-intuitive and remarkably different from what observed in the non-adaptive case, but it may be explained by the existence of a typical average distance, so that when the average opinion distance is above such a value, agents have to change their opinions faster (to avoid isolation), when the average opinion-distance is below the value, agents have very similar opinions and may easily reach an agreement, while around

the typical value, opinions are too distant for the agents to find an agreement, but also too close to promote a quick adaptation process.

The common denominator of the Sznajd and Deffuant-REM (Deffuant with relative expectation model) consensus models and Marsili-Richmond peer pressure models is that in both social environment significantly influence agent behaviour.

GLVPP, Peer Pressure, Sznajd and Deffuant models try to capture different aspects of socio interaction (loosely speaking: 'keep up with the Jones' (GLVPP), going with the crowed (PP), united-we-stand-divided-we-fall (Sznajd model)), bargaining (Deffuant model)).

One of the most important differences between GLVPP and PP models on the one hand and Sznajd and Deffuant models on the other has to do with the topology. In the first everybody does interact with everybody else. In the second this does not happen.

There are not only differences related to the interaction and the topology, but also to the equilibrium solution. In the two continuous models agent start having the same wealth and end up with their wealth being power- law or 'tent' distributed. In the discrete models, at the beginning there are different opinions, at the end only few (or just one) opinions survive. In the first case the evolution reduces the 'order' within the system, in the second increases it.

7.2 Outlook

In order to understand the emergence and evolution of self-organized social systems, we need to push forward agent-based modeling and to achieve a more realistic representation of the systems we are trying to analyse.

Key steps, in doing that, would be:

1. Building new models enabling each agent to both select a specific social environment and adapt to it. For instance, one may think of a Deffuant-REM model in which the network connections may change according to a dynamical mechanism.

2. Assessing the impact of agent behaviour on the topological and dynamical properties of a system. For instance, understanding what makes agents build a previously inexistent network (along the lines of some recent studies on self-assembling networks within ant communities [62]) and how the properties of their local interaction affects topological quantities such as the average clustering coefficient, the average path length and the degree distribution.

3. Studying the impact of global conditions and network topology on the individual agent behaviour. For instance, studying how, given a certain interaction, an agent performance depends on the local clustering coefficient and on the connectivity degree distribution.

Applications of possible findings would certainly be topical for a variety of decision-making processes ranging from market regulation to health and safety.

Bibliography

[1] Majorana, E.; *Il valore delle leggi statistiche nella fisica e nelle scienze sociali.* Scientia **36**, 58-66 (1942).

[2] Mandelbrot, B.B.; *The fractal geometry of nature.* W.H. Freeman, San Francisco (1982).

[3] Mantegna, R.N. & al.; *An Introduction to Econophysics: Correlations and Complexity in Finance.* Cambridge University Press, Cambridge (2000).

[4] Barabasi, A.-L. & al.; *Emergence of Scaling in Random Networks.* Science **286**, 509 (1999).

[5] Barabasi, A.-L & al.; *Evolution of the social network of scientific collaborations.* Physica A **311**, 590-614 (2002).

[6] Farmer, J.D.; *Market Force, Ecology and Evolution.* Industrial and Corporate Change **11**, (5) 895-953 (2002).

[7] Lillo, F. & al.; *Master curve for price-impact function.* Nature, **421**, n. 6919, p. 129-130 (2003).

[8] Bouchaud, J.P. & al.; *Theory of Financial risks and Derivative Pricing: from Statistical Physics to Risk Management.* Cambridge University Press; Cambridge (2003).

[9] Bouchaud, J.P. & al.; *On a Universal Mechanism for Long Ranged Volatility Correlations.* (pre-print cond-mat/0012156).

[10] Malcai, O. & al..; *Theoretical Analysis and Simulations of the Generalized Lotka-Volterra Model*. Phys Rev E **66,** 31-102 (2002).

[11] Liu, Y. & al.; *The statistical properties of the volatility of price fluctuation*. Phys. Rev. E **60**, 1390 (1999).

[12] Gopikrishnan, P. et al.; *Statistical Properties of Share Volume Traded in Financial Markets*. Phys. Rev. E., **62,** R4493 (2000).

[13] Challet, D. & al.; *On the Minority Game: Analytical and Numerical Studies*. Physica A **256**, 514 (1998).

[14] Ausloos, M. & al.; *Patterns, Trends and Predictions in stock market indices and foreign currency exchange rates*. New Vistas in Statistical Physics-- Applications in Econophysics, Bioinformatics, and Pattern Recognition, Delray Beach, Florida, USA, May 9-12, 2001, Proceedings, L.T. Wille. Ed. (Springer Verlag, Berlin, 2002).

[15] Ausloos, M & al.; *Crashes: symptoms, diagnoses and remedies*. Empirical Science of Financial Fluctuations, Tokyo, Japan, Nov.15-17, 2000 Proceedings, Berlin, Springer Verlag (2001).

[16] Ausloos, M.& al.; *Mechanistic approach to generalized technical analysis of share prices and stock market indices*. Third EPS Conference on Application of Physics in Financial Analysis, London, UK, Dec. 3-7, 2001.

[17] Stauffer, D.; *Monte Carlo Simulations of the Sznajd model*. Journal of Artificial Societies and Social Simulation **5**, paper 1 (2002) (jasss.soc.surrey.ac.uk).

[18] Stauffer, D.; *Frustration from Simultaneous Updating in Sznajd Consensus Model*. J.Mathematical Sociology **28**, 25-33 (2004).

[19] Sznajd-Weron, K. & al.; *Opinion Evolution in Closed Communities*. Int. J. Mod. Phys. C **11**, 1157-1166 (2000).

[20] Deffuant, G. & al.; *Mixing beliefs among interacting agents.* Adv. Complex Syst. **3**: 87-98 (2000).

[21] Stauffer, D. & al.; *Simulation of Consensus Model of Deffuant et al on a Barabasi-Albert Network.* Int.J.Mod.Phys.C **15**, issue 2 (2003).

[22] Stauffer, D. & al.; *Discretized opinion dynamics of Deffuant on scale-free networks.* To appear in the Journal of Artificial Societies and Social Simulation in June 2004; (jasss.soc.surrey.ac.uk).

[23] Mainardi, F & al.; *Fractional calculus and continuous-time finance II: the waiting-time distribution*, Physica A, **287**, 468-481 (2000).

[24] Galam, S.; *Sociophysics: a personal testimony.* Physica A **336,** 49-55 (2004).

[25] Galam, S.; *Fragmentation versus Stability in Bimodal Coalitions.* Physica A **230**, 174 (1996).

[26] Krawiecki, A. & al.; *Volatility Clustering and Scaling for Financial Time Series due to Attractor Bubbling.* Phys. Rev. Lett., **89**, 15 (2002).

[27] Marsili, M.; *Dissecting financial markets: Sectors and states.* Quantitative Finance, **2**, 297-302 (2002).

[28] Behera, L. & al.; *On Spatial Consensus Formation: Is the Sznajd Model Different from a Voter Model?* Int. J. Mod. Phys C, **14**, 10, 1331-1354 (2003)

[29] Iori, G. & al; *Patterns of consumption in a discrete choice model with asymmetric interactions.* presented at "Complex Behaviour in Economics" Aix-en-Provence 3-7 May, 2000.

[30] Pareto,V.;*Course d' Economie Politique*. Lausanne and Paris (1897).

[31] Tang, L.-H. & al.; *Modeling High-frequency Economic Time Series*. Proceedings of the Dynamics Days Asia Pacific Conference, 13-16 July, 1999, Hong Kong (Physica A, 2000).

[32] Plerou, V. & al.; *Scaling of the distribution of price fluctuations of individual companies*. Phys. Rev. E **62**, R3023 (2000).

[33] Daccorogna, M.; *Introduction to High-Frequency finance*. Academic Press (Elsevier), 1st edition (2001).

[34] Bernardes, A. T. & al. ; *Election results and the Sznajd model on Barabasi network* ;.Eur. Phys. J. B. (submitted).

[35] Gonzales, M. C. & al; *Opinion formation on a Deterministic Pseudo-fractal Network*. Int. J. Mod. Phys C **15**,1 (2004).

[36] Albert, R. & al.; *Statistical mechanics of complex networks*. Reviews of Modern Physics **74**, 47 (2002).

[37] Liljeros, F. & al; *The Web of Human Sexual Contacts*. Nature **411**, 907-908 (2001).

[38] Aiello, W. & al; in Proc. *The network of phone calls*. 32nd ACM Symp. Theor. Comp. (2000)

[39] Milgram, S.; *The small world problem*. Psych. Today. **2**, 60-67 (1967).

[40] Ball, P.; *The self made tapestry: pattern formation in nature*. Oxford University Press, New York (2001).

[41] Eintein, A.; *On the Theory of Brownian Motion* (Published in 1906).-Bachelier, L.; *The Theory of Speculation.* The Random Character of Stock Prices, Paul H. Cootner (editor), Cambridge: MIT. Translated from the 1900 doctoral thesis.

[42] Montroll, E. W. and Weiss, G. H.; *Random walks on lattices II*, J. Math. Phys. 6, 167-181 (1965).

[43] Wolfram, S.; A new kind of science; Wolfram Media, Champaign, Illinois (2002).

[44] Richmond, P. & al; *Power Laws are Boltzmann Laws in Disguise.* International Journal of Modern Physics C, **12**, 3, 333-343 (2001).

[45] Cecconi, F. & al; *Diffusion, peer pressure and tailed distributions.* Phys. Rev. Lett., **89**, 8 (2002).

[46] Solomon, S. & al.; *Power Laws of Wealth, market order and market returns.* Physica A **299**, 188-197 (2001).

[47] Malcai, O. & al; *Theoretical Analysis and Simulations of the Generalized Lotka-Volterra Model.* Phys Rev E **66**, 31-102 (2002).

[48] Louzoun, Y. & al..; *Volatility driven market in a generalised Lotka-Volterra formalism.* Physica A **302**, 220-233 (2001).

[49] Erdos, P. and Renyi, A.; *On the evolution of random graphs.* Bull. Inst. Int. Stat. **38**, 343 (1961).

[50] Watts, D. J. and Strogatz, S. H.; *Collective dynamics of small-world networks.* Nature **393**, 440 (1998)

[51] Stauffer, D. and Meyer-Ortmanns, H.; *Simulation of consensus model of Deffuant et al. on a Barabási-Albert network.* Int. J. Mod. Phys.C **15**, 2 (2004).

[52] Deffuant, G.; Amblard, F.; Weisbuch, G. and Faure, T.; *How can extremism prevail? A study based on the relative agreement interaction model.* Journal of Artificial Societies and Social Simulation 5, issue 4, paper 1 (jasss.soc.surrey.ac.uk) (2002).

[53] Huang, K.; Statistical Mechanics; John Wiley & Sons (New York,1987).

[54] Sabatelli, L. & al; *Waiting time distributions in financial markets.* Eur Phys J B **27,** 2, 273–275 (2002).

[55] Richmond, P. & al.; *Peer pressure effects in Generalized Lotka-Volterra models.* To be published in Physica A.

[56] Richmond, P. & al.; *Langevin processes, agent models and socio-economic systems.* To be published in Physica A.

[57] Stauffer, D. & al.; *Introduction to Percolation Theory.* 2nd ed.; Taylor and Francis, London 1994 (second printing)

[58] Sabatelli, L. & al.; *Phase transitions, memory and frustration in a Sznajd-like model with synchronous updating.* Int. J. Mod.,Phys C **14,** 9 (2003).

[59] Sabatelli, L. & al.; *Non-monotonic spontaneous magnetization in a Sznajd-like Consensus Model.* Physica A **334**,1-2, 274-280, (2004).

[60] Sabatelli, L. & al.; *A consensus based dynamics for market volumes.* To appear in Physica A.

[61] Sabatelli, L. & al.; *Deffuant model on a scale-free Network with adaptive local consensus thresholds.* Submitted.

[62] Schweitzer, F. & al.; *Self-Assembling of Networks in an Agent-Based Model.* Phys. Rev. E **66** (2002).

Appendix

In Appendix 1 we outline the definition and properties of Auto-Co-Variance Functions. Autocovariance and Autocorrelation (a normalized version of the first), are important to detect memory effect within time series. In Appendix 2 we review the basic properties of Financial Market Volatility. In Appendix 3, we describe the principle behind the random number generation algorithm used in the simulations within this thesis.

Appendix 1: Auto-Co-Variance

Introduction to Auto-Co-Variance

The auto-covariance function (ACVF) measures for a given time series, $x(t)$, the correlation between fluctuations around the average value, μ at time t_1 with similar fluctuations at another time t_2. It is formally expressed as follows [1][2]:

$$C(t_1,t_2) \equiv E\{x(t_1),x(t_2)\} - \mu(t_1)\mu(t_2) \qquad (1).$$

$$E\{x(t_1),x(t_2)\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1 x_2 f(x_1,x_2;t_1,t_2)dx_1 dx_2 \qquad (2)$$

and

$$\mu(t) \equiv E\{x(t)\} \equiv \int_{-\infty}^{+\infty} xf(x,t)dx \qquad (3)$$

are the *expectation values* of $x(t)$ and $x(t_1)x(t_2)$. $f(x,t)$ is the probability density of observing the random value x at time t. $f(x_1,x_2;t_1,t_2)$ is the joint probability density of observing x_1 at time t_1 and x_2 at time t_2.

For $t_1 = t_2$ the ACVF is equal to the variance associated with the stochastic process $x(t)$.

If $f(x,t)$ does not depend explicitly upon the time t, then clearly $\mu(t_1) = \mu(t_2) = \mu$ and both $E\{x(t_1), x(t_2)\}$ and $C(t_1, t_2) = C(\tau)$ depend only on $\tau = t_2 - t_1$. Such a process is said to be " wide-sense" stationary. The ACVF may now be thought of as a "memory function", i.e. a function that quantifies how much memory of past fluctuations is retained by the system after a time τ.

In general, the ACVF is non-zero over a range $|\tau| < \tau^*$ and fluctuations within this range are statistically dependent. In particular, if the ACVF is positive over this range, positive fluctuations are more likely to follow positive fluctuations and negative fluctuations are more likely to follow negative fluctuations after a time τ. If the ACVF is negative, positive fluctuations are more likely to follow negative fluctuations and vice versa.

The Properties of the Auto-Covariance Function.

There is a very important relation between the Fourier transformed $\tilde{x}'(\omega) = \int_{-\infty}^{+\infty} x'(t)\exp(i\omega t)dt$ of the stochastic process $x'(t) = x(t) - \mu$ and the Fourier transformed $\tilde{C}(\omega) = \int_{-\infty}^{+\infty} C(\tau)\exp(i\omega\tau)dt$ of the ACVF $C(\tau)$ of $x(t)$:

$$\tilde{C}(\omega) = E\{\tilde{x}'(\omega)\tilde{x}'^{*}(\omega)\} \qquad (4)$$

This result is also known as Wiener-Khintchine theorem and. $\tilde{x}'(\omega)\tilde{x}'^{*}(\omega)$ is also called Power Spectrum (PS) of $x'(t)$[1][2].

This result allows us to study the PS and establish more detail about the ACVF. If the ACVF of $x(t)$ is zero for $\tau \neq 0$ the power spectrum is flat, in other words all the frequency components contribute with the same weight to PS. If $x(t)$ is a short range correlated process and $C(\tau) = \sigma^2 \exp(-|\tau|/\lambda)$, the P.S. becomes:

$$\widetilde{C}(\omega) = \frac{2\sigma^2 \lambda}{1+(2\pi\lambda\omega)^2} \qquad (5)$$

If $x(t)$ is characterized by a power law: $C(\tau) = A|\tau|^{\eta-1}$, where A is a constant coefficient and η is a number bigger than zero and smaller than 1, the PS is

$$\widetilde{C}(\omega) = B|\omega|^{-\eta} \qquad (6)$$

The range of frequencies, for which we can rely on analyses of the power spectrum is given approximately by

$$\omega_{\min} = 1/T \quad \text{and} \quad \omega_{Max} = 1/\Delta t \qquad (7)$$

where T is the time length of the time series and Δt is the time lag between two data.

An important point to make at this stage is that most workers, (especially physicists) assume that expectation values may be equated with time ave$E\{x(t)\} \equiv < x(t) > = \lim_{T\to\infty} \frac{1}{T} \int x(t)dt$ (8)

$$E\{x(t)\} \equiv < x(t) > = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)dt \quad (8)$$

and

$$E\{x(t_1)x(t_2)\} \equiv < x(t_1)x(t_2) > = \lim_{T\to\infty} \int_0^T x(t_1)x(t_2)dt \qquad (9)$$

When that is possible the process $x(t)$ is said to be ergodic.

To discriminate whether or not a process is ergodic can be a very hard task.

General requirements for a process $x(t)$ to be ergodic are: to be stationary and to have a time averages in (10) and (11) independent by the sample (sample function) chosen of $x(t)$.

The behaviour of the ACVF is very important in order to assess to what extent the Central Limit Theorem (CLT) may fully describe the asymptotic behaviour of a process $x(t)$ (see Central Limit Theorem).

CLT assumes variables are statistically independent. A stochastic process of identical distributed random variables, having a zero ACVF value for τ different from zero, meet the CLT requirements, hence the process obtained summing N of those iid variables is expected to converge to a stable distribution (i.e. to a Gaussian distribution if the variance associate to is finite, to a Levy distribution if the variance is infinite or not defined, according to the generalized CLT) as N increases.

If $x(t_i)$ is a stationary process with variance $\sigma^2 = C(0)$ and $C(\infty) = 0$, the variance of the variable

$$S_N(t_j) = \sum_{i=1}^{N} x(t_{j+i}) \qquad with \qquad j = 0, N, 2N.....kN \quad (10)$$

can be expressed in terms of $C(\tau)$ as follows:

$$\sigma_{S_N}{}^2 = E\{S_N{}^2\} - E\{S_N\}^2 = N\sigma^2 + 2N\sum_{\tau=1}^{N}(1-\frac{\tau}{N})C(\tau) \qquad (11)$$

If $C(\tau)$ decays faster than $1/\tau$ for large τ, the sum over τ tends to a constant for large N, and thus the variance of the sum still grows as N, as for the usual Central Limit Theorem.

If $C(\tau)$ decays for large τ as a power law ("fractional Brownian motion") $\tau^{-\nu}$ $with \nu \prec 1$ then the variance grows faster than N, as $N^{2-\nu}$. In this case the standard CLT needs to be amended, and there is no general solution for the problem of the limit distribution in these cases [2].

Studying the standard deviation σ_{S_N} of S_N as function of N one can work out some properties of $C(\tau)$.

In order to do that one takes a time series of length T and divides it in $n = \frac{T}{N}$ intervals, works out for each interval $S_N(i)$ $i = 1,...n$ and estimates $\sigma_{S_N}{}^2$ as

$$\overline{\sigma}_{S_N}{}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(S_N(i) - \overline{S}_N^n(i))^2 \qquad (12)$$

where

$$\bar{S}_N^n = < S_N > = \frac{1}{n}\sum_{i=1}^{n} S_N(i) \qquad (13)$$

if the process $x(t)$ is stationary.

If the process is not stationary because of a trend, we can still use the relation (14) to work out the asymptotic properties of auto-covariance, but we must take $\bar{S}_N^n$ as a non constant function of i representing the value of the function best fitting the points $\{i, S_N(i)\}$. In this case we talk of *Detrended Fluctuation Analysis (DFA).*[3]

From such an analysis one can infer some asymptotic properties of ACVF, according to the following criteria:

- If σ_{S_N} is proportional to $N^{0.5}$, this means that there is no long range correlation, the stochastic process has fast decaying autocorrelation or no autocorrelation.
- If σ_{S_N} is proportional to N^{α}, with $0.5 < \alpha < 1$, $C(\tau)$is positive and has a power law behaviour with exponent $\nu = 2 - 2\alpha$.
- If σ_{S_N} is proportional to N^{α}, with $0 < \alpha < 0.5$, $C(\tau)$is negative (anti-correlation) and has a power law behaviour with exponent $\nu = 2 - 2\alpha$.
- If σ_{S_N} is proportional to N^{α}, with $\alpha > 1$, $C(\tau)$ is non zero, but does not have anymore a power law structure.

There are studies suggesting how, in some cases, just one of the methods outlined might be not enough to work out the properties of ACVF. In general, it is advisable to use not only ACVF direct expression, but also PS and DFA, relying only on results supported by at least two of those methods[5].

Also, a certain care in the interpretation of results is advisable when dealing with distributions having "heavy tails".

In fact, in that case, the estimated value of the ACVF (sample ACVF) can be affected by bigger statistical errors than in the case of a gaussian

probability density function, due to a slower rate of convergence to the actual ACVF [5]

Bibliography

[1] Mantegna, R. N. & al; *An Introduction to Econophysics: Correlations and Complexity in Finance.* Cambridge University Press, Cambridge (2000).

[2] Bouchaud, J.P. & al.; *Theory of Financial risks and Derivative Pricing: from Statistical Physics to Risk Management.* Cambridge University Press; Cambridge (2003).

[3] Liu, Y. & al.; *The statistical properties of the volatility of price fluctuation.* Phys. Rev. E **60**, 1390 (1999).

[4] Ramgarajan, G. & al.; *Integrated approach in the assessment of long range correlation in time series data.* Phys. Rev. E, **61**, .5, (2000)

[5] Davis, R. A. & al.; *The sample autocorrelations of financial time series models.* In: Fitzgerald, W.J. & al. (Eds.) *Nonlinear and Nonstationary Signal Processing*, Cambridge University Press, Cambridge (2001).

Appendix 2: Volatility

Volatility is commonly associated with risk and opportunities for making profits. It is a measure of the magnitude of asset price fluctuations and is commonly formulated using one of two different routes. Historic volatility is a statistical measure based on past price movements. Implied volatility is calculated from quoted traded option premiums or prices and for this purpose, the use of a pricing model such as the Black Scholes theory is required. This latter route is used to estimate whether option premiums are relatively cheap or expensive. A measure of volatility in the future would clearly enable us to predict future price movements more precisely. Since this is not possible, historic and implied volatilities are widely used as the best guesses for future volatility. These are not always equivalent.

Generally, when evaluating volatility, we consider different time periods. We may look at the volatility over the past week, the past month, the past three months, the past six months and even over longer periods. The longer time period will clearly yield more of an average volatility.

Suppose we have the time series for the price, $\{p_0, p_1, p_2 \ldots\ldots\ldots p_N\}$, of an asset. The prices are taken at regular intervals, τ that may be annual, monthly, weekly, daily or even finer steps taken over a time period T where $T = N\tau$.

We now compute the log price return time series:

$$r_n = \ln\left(p_n / p_{n-1}\right) \qquad (1)$$

Clearly the average return is

$$r = \frac{1}{N}\sum_{n=1}^{N} r_n \qquad (2)$$

The historic volatility is then defined as

$$\sigma = \sqrt{\left[\left(\frac{1}{N}\sum_{n=1}^{N} r_n^2\right) - r^2\right]} \qquad (3)$$

Another definition sometimes used is:

$$\sigma = \left(\frac{1}{N}\sum_{n=1}^{N} |r_n|\right) - r \qquad (4)$$

With the emergence of tic data that is not homogeneous in time and for which the time lag between data points is no longer constant, care must be take with the computation of volatility. A common approach is to use Exponential Moving Average (EMA) operators to generalize the usual statistical tools devised for homogeneous time series [1].

The probability distribution function of the volatility can be described by a log normal distribution for small values; for large values it exhibits a fat tail that can be characterised by a power law.

The Autocorrelation function of the volatility also exhibits a long tail with characteristic power law behaviour. Generally the exponent of the power law is small lying between 0.2 and 0.6[2].

There is a strong positive correlation between volatility and market volume, suggesting that volatility is intimately related to the scale of market activity.

When evaluating the purchase of an option, not only the historical volatility but also the so-called implied volatility of the underlying security is evaluated. There are many different models for pricing options and most require a value for the volatility in order to estimate the price. Frequently using the historic volatility with most models will yield a

price relatively close to the true price. The estimate is however affected by the time interval or lag between two consecutive prices used for the determination. According to an empirical rule, the best estimate is obtained taking a time interval as long as the time to maturity, T, of the option [3].

However, there are times when the calculated price is quite different to that at which the option is trading. This is generally ascribed to the fact that the historical volatility being used is incorrect since all other inputs such as the time until expiration, the strike price, dividends to be paid by the stock, the current risk free interest rate are known. The marketplace in effect is assuming a different volatility from that predicted using the historical volatility. The way out of this dilemma is to use the option-pricing model in reverse. We know the option price and all other variables except the volatility that the market price implies. Therefore, instead of using the equation to solve for the option price, we use the model to solve for the volatility assuming all other variable are known. Using quoted option prices together with other variables such as time to expiry, strike price, dividends to be paid, underlying stock price and current risk free interest rate we may compute the so-called implied volatility.

Many traders refer to implied volatility as the premium. To be precise, the word premium refers to the option price relative to the underlying security. Nevertheless, traders will say things like, "Premium levels are high." or "Premium levels are low." What the trader is really referring to is the implied volatility. The implied volatility is high or the implied volatility is low.

Thus using the Black-Scholes pricing formula for European options we may estimate the market expectation of volatility, σ using market price data for a “benchmark” option with a time to maturity, T, using the Black Scholes formula for the price, p:

$$p = c(x_0, t, T, r, \sigma, x_s) . \qquad (5)$$

The result for the implied volatility, $\sigma(x_s,T)$ as a function of strike price, x_s and x_0, exhibits a characteristic behaviour as a function of x_s-x_0. The curve may frequently be likened to the smile of Cheshire cat. This is the so-called volatility smile. The implied volatility reaches a minimum value for $x_s = x_0$ (at-the-money options) that is smaller than the actual volatility σ of the underlying asset.

The volatility smile effect has been shown to be intimately related to the presence of a non-zero and generally positive kurtosis k(T) for the distribution of returns of the underlying asset. Bouchaud [2] has derived the following approximate, analytic expression for implied volatility:

$$\sigma(x_s,T) = \sigma\left[1+\frac{k(T)}{24}\left(\frac{(x_s-x)^2}{\sigma^2 x_0^2 T}-1\right)\right] \qquad (6)$$

ARCH, GARCH and related stochastic processes have been widely used in an attempt to develop time dependent models of volatility [4]. HARCH models succeed in modelling long ranged memory of volatility[1]. Other stochastic volatility models that invoke a dependence on past volatility values rather than simply returns have also been explored. All these models generally fail to describe correctly the scaling behaviour that characterises return and volatility distribution functions and can be cumbersome to compute. This has led to a new approach based on the use of agent methods rooted in statistical physics for the characterisation of these stochastic processes. Recently, however, LeBaron [4] has shown how simple stochastic volatility models that include driving processes with different time scales can display power laws and scale invariance similar to actual financial data

Bibliography

[1] M. Dacorogna et al.; *An Introduction to high frequency finance*. Academic Press; San Diego (2001)

[2] Bouchaud, J.P. & al.; *Theory of Financial risks and Derivative Pricing: from Statistical Physics to Risk Management*. Cambridge University Press; Cambridge (2003).

[3] Mantegna, R. N. & al; *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge (2000).

[4] LeBaron, B.; *Stochastic Volatility as a Simple Generator of Financial Power-Laws and Long Memory*. Preprint available from blebaron@brandeis.edu.

Appendix 3: Random number generation

Random numbers needed in Computer simulations have been generated through Linear Congruential Random Number Generator algorithm.

Starting with an integer, called ISEED, we produce an odd integer IBM = 2 *ISEED -1, then we perform the multiplication IBM=IBM*16807.

Implemented on any 32-bit machine this procedure gives integers IBM distributed randomly between -2^{31} and 2^{31}.

In fact, when two numbers are multiplied a 32-bit computer throws away the leading bits exceeding 32 bits. Since the first bit refers to the number sign (say 0 for positive and 1 for negative), the multiplication of two positive numbers may still produce a negative number. More important the last (and least significant) 32 digits, that are retained, look completely unpredictable and would pass most non-trivial tests for randomness.

Empirically the best results are obtained multiplying IBM times 16807 (in the past other numbers, such as 65539 and 65549, have been used). In general it is required for the IBM numbers to be odd integers.

Bibliography

Stauffer, D.; Chap1. Computational Physics, Hoffmann-Schreiber Editors, Springer-Verlag Berlin Heidelberg (1996).