TOPICS IN COMPLEX SYSTEMS

A thesis submitted for the degree of Doctor of Philosophy

by

Yee Jiun Yap

Department of Mathematical Sciences, Brunel University

September 2006

Acknowledgements

I want to express my gratitude to my supervisor Prof Geoff Rodgers for introducing me to the field of complex systems. His advice has been very helpful throughout the course of my research. I am also grateful for the discussions which, although highly inspirational, still left much room for me to carry out independent research.

I also want to thank all the members of the Department of Mathematical Sciences for making my stay a wonderful experience. I particularly want to thank Ms Beverley Curr for her help with printing, and Mr Neil Turner for computer support.

Last but not least, I want to thank my parents for their encouragement throughout my research.

Y J Yap

September 2006

Abstract

Brunel University, Uxbridge; Department of Mathematical Sciences; Yee Jiun Yap; Topics in Complex Systems; 2006; Doctor of Philosophy

Fundamental laws of physics, although successful in explaining many phenomena observed in nature and society, cannot account for the behaviour of complex, non-Hamiltonian systems. Much effort has been devoted to better understanding the topological properties of these systems. Neither ordered nor disordered, these systems of high variability are found in many areas of science. Studies on sandpiles, earthquakes and lattice gases have all yielded evidence of complexity in the form of power law distributions. This scale-free characteristic is believed to be the hall-mark of complexity known as self-organised criticality. Systems in the self-organised critical state regulate themselves and are resistant to error and attacks. The aim of this thesis is to further current knowledge of complex systems by proposing and analysing three models of real systems. Statistical mechanics and numerical simulations are used to analyse these models. The first model mimics herd behaviour in social groups and encompasses growth and addition. It has been found that when the growth rate is fast enough, the group size distribution conforms to a power law. When the growth rate is slow, the system runs out of free agents in finite time. The second model aims to capture the basic empirical measurements from hospital waiting lists. This model illustrates how the power law distributions found in empirical studies might arise, but also indicates that these distributions are unlikely to be caused by the preferential behaviour of patients or physicians. The third model is a salary comparison model; the salary distributions of most of its variants are power laws. Both mean field and 1-d versions of the model are analysed, and differences between the two versions are identified by looking at the mean absolute difference between the salaries in each version.

Bibliographical note

Some of this work has been published in

- G. J. Rodgers, Y. J. Yap, Eur. Phys. Journ. B **28**, 129, (2002)
- G. J. Rodgers, Y. J. Yap, T. P. Young, Adv. Compl. Syst. **6**, 215, (2003)
- S. Rawal, G. J. Rodgers, Y. J. Yap, Physica A **357**, 477, (2005)

Contents

Chapter 1

Introduction

1.1 Complex systems in nature and society

For thousands of years, philosophers and thinkers have been trying to understand the world around us. However, modern scientific studies based on hypothesis and supported by experimental evidence only began about three hundred years ago, when Isaac Newton introduced the concept of unbiased interaction between experimentation and theory in *Principia Mathematica.* This led to subsequent scientific studies and discoveries by other thinkers of that time. By the late 19^{th} century, the collective study of nature using mathematics was termed 'physics'.

Physics involves the modelling of the universe using mathematical expressions. All subjects under study are definitive and clear; no ambiguity exists on whether two masses attract each other and the strength of gravity weakens with physical distance. Thus, the behaviour of all matter is clearly accounted for by the laws of physics [1, 2, 3]. In classical mechanics, these

definitive laws allow one to predict the subsequent outcomes if the initial conditions are known. For example, if all the initial conditions of a pendulum bob raised to a certain height are known, it is possible to predict the velocity of the bob at any subsequent time. Although the predictions based on quantum mechanics are in the form of probabilities and eigenvalues, forecasts are still possible. For instance, it is trivial to predict the possible positions of an electron around a nucleus given the wave function and the initial conditions.

However, one might be perplexed if one were to apply the laws of physics to complicated systems such as the weather or a price index in the stock market. Whilst the laws of physics are very effective in predicting the outcomes of linear systems, they tend to fail when applied to more complicated nonlinear systems, as was discovered by Edward Lorenz in 1961. Lorenz was running a simple weather simulation using a set of nonlinear equations when he accidently discovered that two predictions were totaly different despite using the same set of equations and similar values of the initial condition. The cause of the great difference in the outcome turned out to be the tiny difference between the values of the initial state used [4]. This led to subsequent studies of unpredictable systems that are sensitive to initial conditions and the eventual establishment of a new discipline known as chaos theory [5, 6, 7]. Systems that display such unpredictability are known as chaotic .

Ordered systems are deterministic and their future states can be predicted using the laws of physics. As an ordered system becomes more and more disordered, its entropy becomes higher and higher and when a critical point is passed, the system becomes chaotic. However, at precisely the critical point the system is partially predictable, and partially not; it tends to exhibit a

power law behaviour [8, 9, 10], but remains largely unpredictable. Systems at this critical point are known as complex systems [11, 12, 13, 14]. These systems are characterised by high variability and are commonly referred to as being emergent.

Scientists have now realised that many collective systems show emergent characteristics. There have been accounts of many areas of physics where the laws of physics, commonly known as reductionist laws, cannot be used to predict the outcomes given the initial conditions. For example, there have been problems of predicting the outcomes in experiments of solid-state physics, despite the fact that the subjects involved are governed by quantum mechanics. Even theories in particle physics, such as string theory, are emergent in nature. In quantum mechanics, although the Schrödinger equation agrees with experiments involving isolated atoms, it cannot be used to account for the behaviour of more than ten particles. Thus, whilst many laws in physics involve the application of symmetry, many real systems in nature appear to be anti-symmetric.

Truly stable systems only constitute a very small part of the universe, as most systems, be they in nature or society, have symmetry-breaking characteristics. Such systems are often associated with power laws [10, 15]

$$F(x) = x^{\varpi} \tag{1.1}$$

where x is some physical quantity, $F(x)$ is the probability of obtaining x and

ϖ is a constant. Power law distributions are referred to as scale-free since

$$\frac{F(kx)}{F(x)} = k^{\varpi} \tag{1.2}$$

where k is a constant, is independent of x. Many systems in nature and society have been found to exhibit power law behaviour. One such system is the solar flares emitted from the sun [16]. Observations on the solar flares have shown that the energy distribution is a power law with an exponent of 1.6. Size distribution of pulsar glitches [17] and intensity distribution of x-rays from black holes [18] have both been found to be power law with exponents close to unity. Other examples of complex systems in nature include metabolic networks of organisms [19], food webs [20] and protein networks in biological cells [21]. The World Wide Web [22], human sexual contact webs [23] and science collaboration [24] networks are just several of the many systems in society with power law distributions. On the large scale involving the interactions of many elements at critical points, the world can be viewed as a network being regulated by an organizational behaviour. This behaviour, as we shall see in the later sections, organises the system when changes are introduced, causing the system to resist the changes.

To analyse and understand complex systems, statistical mechanics [25, 26, 27] is used to model certain systems in an attempt to better understand their network topologies. Although complex networks have traditionally been the territory of graph theory [28, 29, 30] which focusses on regular graphs, random graphs [13, 31, 32, 33] which have no apparent principles have been proposed as the most fundamental realisation of complex systems. Following

random graphs, other complex networks [13] were also proposed, including small-world [13, 34, 35, 36] and scale-free networks [13, 37, 38, 39]. Each of these complex networks is briefly described in the following subsections.

1.2 Complex networks

1.2.1 Random graphs

A graph [13] is defined as a pair of sets $G = \{P, E\}$ where P is a set of N vertices and E is a set of edges that connect any two vertices together. The number of edges connected to a vertex is known as the degree of the vertex, denoted by k, and the degree of vertex i is denoted by k_i. Vertices and edges are normally represented by dots and straight lines respectively. A typical graph is shown in Fig. 1.1

Random graphs are graphs where any two vertices are connected randomly. This is applicable to complex systems since most of them have an unknown connection algorithm; the ambiguity causes the connections to appear totally random.

In the Erdős-Rényi model [31], a random graph consists of N vertices connected by n edges chosen from $\frac{N(N-1)}{2}$ possible edges. This gives a total of $C_n^{\frac{N(N-1)}{2}}$ possible graphs each with n edges and N vertices, where $C_n^{\frac{N(N-1)}{2}} = \frac{\frac{N(N-1)}{2}!}{n!(\frac{N(N-1)}{2}-n)!}$ is a binomial coefficient. A model equivalent to the random graph is the binomial model [13], defined as a system with N vertices initially with each pair of vertices linked with probability p. Consequently,

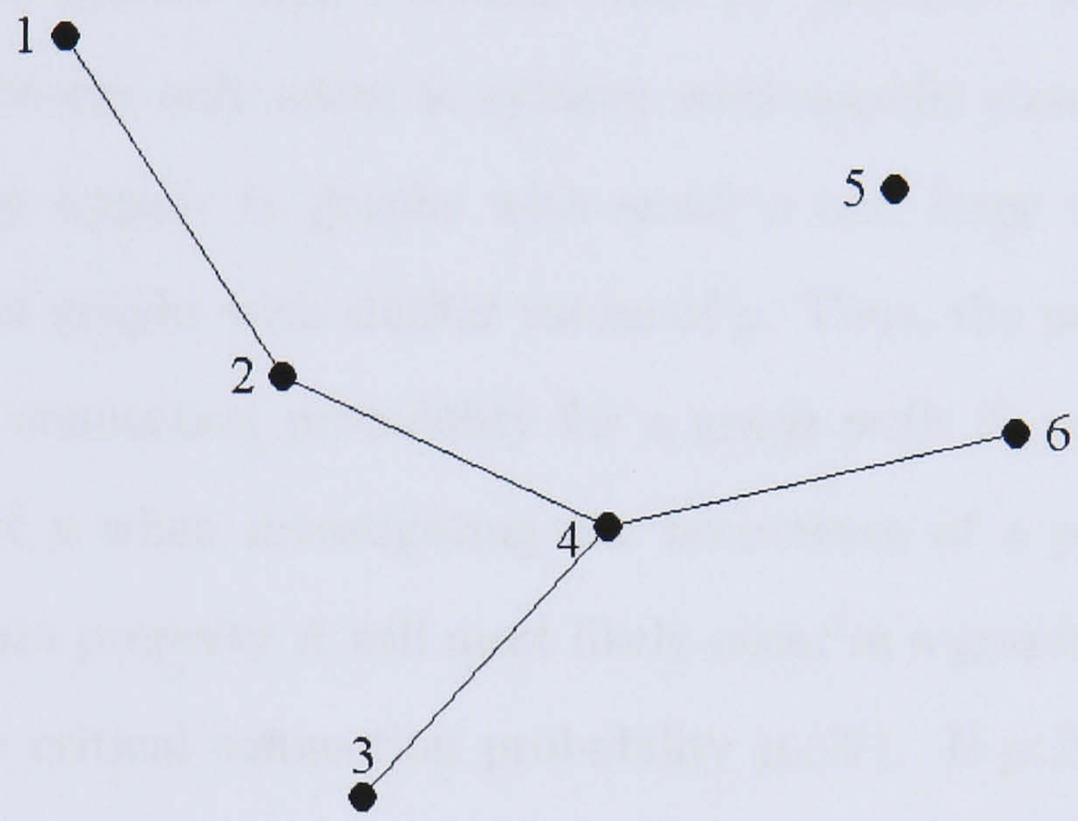

Figure 1.1: An example of a graph with a set of vetices $P = \{1, 2, 3, 4, 5, 6\}$ and edge set $E = \{\{1, 2\}, \{2, 4\}, \{3, 4\}, \{4, 6\}\}$.

the probability of obtaining a graph G_0 with n edges is given by

$$P(G_0) = p^n(1-p)^{\frac{N(N-1)}{2}-n}. \tag{1.3}$$

In real physical systems where the dimensions are mostly finite, the value of p for a system of a fixed size N determines whether a particular property A will occur [13]. Following this line of thought, the property A is likely to occur when p is greater than a critical value p_c. However, many properties in complex networks only occur in systems with specific sizes. For instance, cycles [13] only appear in graphs with small p and large N, and do not occur in smaller graphs with similar values of p. Thus, the parameter $p(N)$, defined as the connection probability for a graph with N vertices, is often used instead of p when investigating the occurrence of a property A. In general, a certain property A will most likely occur in a graph when its $p(N)$ is equal to the critical connection probability $p_c(N)$. If $p(N) \gg p_c(N)$ as $N \to \infty$, almost every graph with $p(N)$ will have property A. However, if $p(N) \ll p_c(N)$, almost all graphs do not possess property A.

An interesting property of the random graph is its diameter [13], defined as the maximal distance between any pair of vertices. If p is not too small, random graphs tend to spread since the number of vertices at distance l from a given vertex is not much smaller than $< k >^l$, where $< k >= pN$ is the mean degree of the graph. This spreading causes the diameters of random graphs to be small.

Another interesting property of a graph to study is its degree distribution

[13]. The degree distribution of a vertex i, denoted by $P_d(k_i = k)$ is given by

$$P_d(k_i = k) = C_k^{N-1} p^k (1-p)^{N-1-k} \tag{1.4}$$

where C_k^{N-1} is the binomial coefficient, and this equation corresponds to a binomial distribution. From Eqn. (1.4) the expectated value of the number of vertices with degree k, denoted by $E(X_k)$, is

$$E(X_k) = NP_d(k_i = k) = NC_k^{N-1} p^k (1-p)^{N-1-k}. \tag{1.5}$$

Consequently, for large N the distribution of X_k is

$$P_n(X_k = r) = \exp(-\lambda_k)\frac{\lambda_k^r}{r!} \tag{1.6}$$

which is a Poisson distribution with mean value $\lambda_k = pN$.

Clustering [13, 35] is another property frequently [40, 41, 42] investigated in graphs, which measures the degree of clustering in the immediate neighbourhood of vertex i. The immediate neighbourhood of vertex i, denoted by M_i, is defined as

$$M_i = \{v_j\} : e_{ij} \in E \tag{1.7}$$

where v_j is a vertex directly connected to vertex i, e_{ij} is the edge connecting them together and E is the set of edges of the graph. For an undirected graph, the clustering coefficient of vertex i, denoted by C_i, is defined as the ratio between the number of edges n_e that exist between the vertices in the immediate neighbourhood of vertex i and the total number of possible edges

$\frac{k_i(k_i-1)}{2}$ that can exist between these vertices. Thus, the C_i of an undirected graph is given by

$$C_i = \frac{2n_e}{k_i(k_i-1)}. \tag{1.8}$$

Similarly, the C_i of a directed graph is

$$C_i = \frac{n_e}{k_i(k_i-1)}. \tag{1.9}$$

For a random graph, the clustering coefficient of the entire graph is given by the mean degree of the graph over all vertices. Following this, the clustering coefficient of random graphs C_{ran} is

$$C_{ran} = p = \frac{< k >}{N}. \tag{1.10}$$

From Eqn. (1.10) it is expected that the plot of $C_{ran}/ < k >$ versus N yields a power law with an exponent of -1. However, plots from many real networks do not conform to this prediction. This is especially true for large ordered lattices where the $C_{ran}/ < k >$ is independent of N, instead of decreasing as N^{-1}. The clustering coefficient of these large ordered lattices is found to depend only on the coordination number of the lattice and not its size [35].

However, the degree distribution of most real networks follows a power law instead of the Poisson distribution Eqn. (1.6) applicable to random graphs [13]. One way to generate scale-free behaviour in a random graph is to force the degree distribution to conform to a power law whilst allowing the vertices to be connected randomly. Although this shows how a power law degree distribution can be produced, it does not help to explain how

the scale-free phenomenon emerges in a complex system as a result of the dynamics of its constituents. This problem is addressed in the following subsection where the scale-free model is described briefly.

1.2.2 The scale-free model

Instead of aiming at producing a topologically correct network, the scale-free model [13] is directed at understanding the dynamics responsible for the scale free features seen in many real networks. Thus, in the construction of this model, emphasis is placed on searching for the mechanism responsible for the power law degree distribution seen in real networks.

The two features that were introduced to produce a power law degree distribution are growth and preferential attachment. In the growth feature, there are m_0 vertices in the system at the initial time step. Then, at each subsequent time step, a new vertex is introduced into the system. This new vertex is attached to m existing vertices where $m \leq m_0$. In the preferential attachment, the new vertex is attached to vertex i with degree k_i with a probability $\Pi(k_i)$, where

$$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j} \tag{1.11}$$

Both of these features are absent in the random graph model and the small-world model introduced in the following subsection.

One way of obtaining the degree distribution is to use the continuum theory [37]. In the continuum theory, k_i is assumed to be a continuous real variable. The change in k_i with time is expected to be proportional to the

probability of a vertex being connected to vertex i. Thus, we have

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\Sigma_{j=1}^{N-1} k_j} \tag{1.12}$$

where m is the number of edges of new vertices introduced at each time step and t is time. The sum in the denominator of Eqn. (1.12) is equal to $2mt - m$ since it is over all vertices in the system minus the newly introduced one. This gives

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t-1} \tag{1.13}$$

for large t. With the initial condition that every new vertex is attached to m existing vertices in the system, the solution of Eqn.(1.13) is

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \tag{1.14}$$

where t_i is the time step at which vertex i is introduced into the system, $k_i(t_i) = m$ and $\beta = \frac{1}{2}$. Thus, the time dependence of the degrees follows the power law.

From Eqn.(1.14) the probability that the degree of vertex i at time t is smaller than k, denoted by $P(k_i(t) < k)$, is given by

$$P(k_i(t) < k) = P\left(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}\right). \tag{1.15}$$

By adding the vertices to the system at equal time intervals, the probability of getting t_i, that is, the probability of vertex i being chosen to be linked to

m new vertices, denoted by $P(t_i)$, is

$$P(t_i) = \frac{1}{m_0 + t} \tag{1.16}$$

where m_0 is the number of vertices at the initial time step as defined earlier.

By substituting Eqn. (1.16) into Eqn.(1.15) we get

$$P\left(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}\right) = 1 - \frac{m^{1/\beta}t}{k^{1/\beta}(t + m_0)}. \tag{1.17}$$

The degree distribution is thus given by

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^{1/\beta}t}{m_0 + t}\frac{1}{k^{1/\beta+1}}. \tag{1.18}$$

For $t \to \infty$

$$P(k) \sim 2m^{1/\beta}k^{-\gamma} \tag{1.19}$$

where $\gamma = \frac{1}{\beta} + 1 = 3$. This agrees with numerical simulation results of the scale-free model in [38].

The average path length of a graph is defined as the average number of vertices between any two vertices. Empirical studies have shown that the average path length of real networks is small. The average path length of the scale-free model is smaller than that of the random graph model, indicating that the former is more efficient in bringing the vertices close. From data plots, the average path length l of the scale-free model is found to follow the logarithmic form

$$l = A\log(N - B) + C \tag{1.20}$$

where A, B and C are constants. However, an analytical explanation for the path length of the scale free model is still lacking.

Besides the path length, there is also no theoretical prediction for the clustering coefficient of the scale free model. From simulations, the clustering coefficient C_{scale} of this model is shown to follow the power law $C_{scale} \sim N^{-0.75}$. Thus, the clustering coefficient of this model decreases rapidly with N, compared to the $C_{ran} \simeq < k > N^{-1}$ in Eqn. (1.10) of the random graph model.

1.2.3 Small-world networks

The small-world model, just like the random graph model, is aimed at constructing a topologically correct network. It was observed that many real networks have small path lengths like those of random graphs. However, the clustering coefficients of these real networks are much larger. The first model which successfully captured these features was the Watts-Strogatz (WS) model [35].

In the WS model, one starts with a regular lattice before initiating a random process. Initially, the system consists of a ring lattice with N vertices, and each vertex is connected to its first $K/2$ neighbours on each side, giving each vertex a total of K neighbours. Then, each edge is randomly reconnected with probability p without any facsimile connections or having any vertices being connected to themselves, giving a total of $pNK/2$ long-range edges. By varying the value of p between zero and unity, one can adjust the degree of randomness of the system, and the small-world model interpolates between

a regular ring lattice and a random graph.

Consequently, this model is suitable to represent a social network where everyone is connected to some people far away as well as someone close-by. Examples of this type of social network includes scientific collaboration, familial relationship and worldwide branches of a business group.

According to the work by Watts and Strogatz [35], when $p = 0$, the average path length and clustering coefficient at p, denoted by $l(p)$ and $C_{small}(p)$ respectively, are such that $l(0) \simeq N/2K$ and $C_{small}(0) \simeq 3/4$. For $p \to 1$ the model tends towards a random graph with $l(1) \sim \ln(N)/\ln(K)$ and $C_{small}(1) \sim K/N$. By looking at the two extremes, it seems that a small l corresponds to a small C_{small}, and a large l corresponds to a large C_{small}. However, the work of Watts and Strogatz [35] has shown that there is a region within which the $l(p)$ drops rapidly with p but the $C_{small}(p)$ stays almost constant. This results in a small l and large C_{small}, two features commonly seen in real networks.

In Ref. [34] it was noticed that l starts to decrease only when $p \geq 2/NK$, thereby indicating the presence of at least one shortcut. This suggests that a p-dependent crossover length N^* is present and it has been found that $l \sim N$ for $N < N^*$ and $l \sim \ln(N)$ for $N > N^*$. Barthélémy and Amaral [43] conjectured that the characteristic path length is given by

$$l(N, p) \sim N^* F\left(\frac{N}{N^*}\right) \tag{1.21}$$

where

$$F(u) = \begin{cases} u & \text{if } u \ll 1 \\ \ln(u) & \text{if } u \gg 1 \end{cases} . \tag{1.22}$$

The expression for $C_{small}(p)$ can be derived by using a slightly different definition of C_{small}, denoted as $C'_{small}(p)$ [44]. $C'_{small}(p)$ is defined as the mean number of edges between the neighbours of a vertex over the mean number of possible edges between those neighbours. To obtain the $C'_{small}(p)$, we first look at a regular lattice with $C_{small}(0)$. The probability of two neighbours which are connected to each other and to a vertex i at $p = 0$ remaining connected at $p > 0$ is $(1-p)^3$, so that $C'_{small}(p) \simeq C_{small}(0)(1-p)^3$. The deviation of $C'_{small}(p)$ from $C_{small}(p)$ has been numerically shown to be zero as $N \to \infty$ [44].

1.3 Error and attack tolerance of complex networks

Complex networks tend to display a high degree of tolerance against errors and attacks made on them [13]. For example, the large global phone-call networks are not disrupted by small malfunctions and technical errors in individual systems, and biological cells continue to evolve and multiply despite some defects in the DNA structures. The tendencies to resist changes in these systems are largely due to the presence of redundant connections. However, the ability of a network to maintain stability based on topological properties would be an interesting topic to pursue.

An error occurring in a network or an attack made on it can be represented

by the removal of one or more vertices from a graph representing the network [13]. To investigate the robustness of a system in response to an error or attack, some of the vertices of a complex system are removed and the effects of these removals are observed. Initially, the network is connected, and at each subsequent time step, a vertex is removed from the system. The removal of a vertex results in the loss of all edges connected to it. An illustration of the vertex removal is shown in Fig. 1.2.

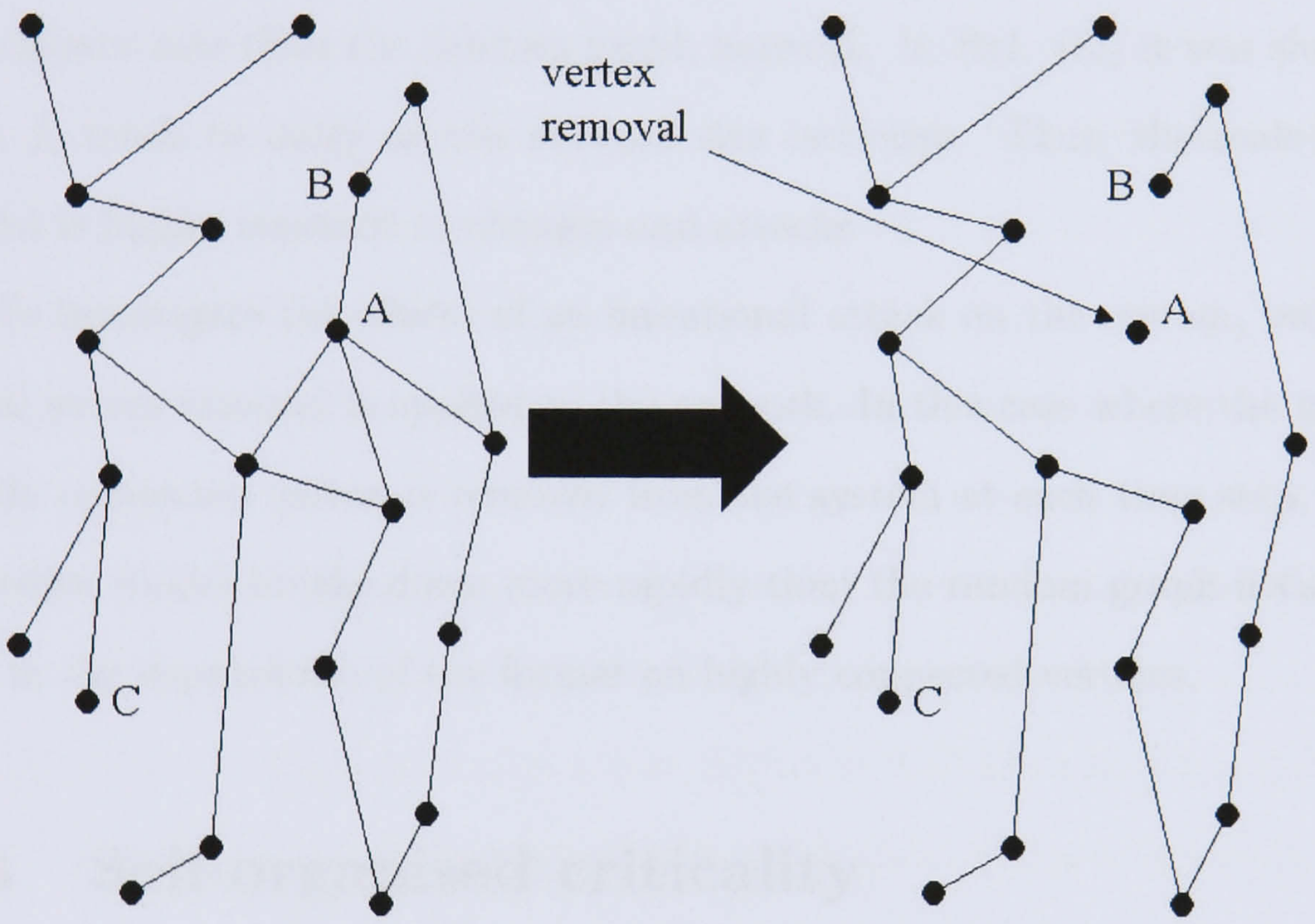

Figure 1.2: An illustration of the vertex removal. After the removal of vertex A, the network breaks down into two isolated clusters. The path length between B and C is five in the undisturbed state, and this increases to eleven after the vertex removal.

One way to measure the effects of vertex removal on the system is to look at the relationship between the fraction f of the vertices removed and C_L, the fraction of vertices in the largest cluster. Also, the effects can be

measured by looking at how f changes with l, the average path length of the largest cluster.

For a random graph network the C_L decreases as the f increases, and for a critical f, denoted by f_c, the C_L drops to zero. The l increases with f and reaches a maximum value at f_c. The l decreases as the f increases beyond f_c.

For the scale-free model, although the C_L decreases as f increases, the former reaches zero at a higher f. Also, the l of the scale-free model increases at a slower rate than the random graph network. In Ref. [13] it was shown that f_c tends to unity as the network size increases. Thus, the scale-free model is highly resistant to changes and attacks.

To investigate the effects of an intentional attack on the system, preferential vertex removal is applied to the network. In this case where the most highly connected vertex is removed from the system at each time step, the scale-free model breaks down more rapidly than the random graph network due to the dependence of the former on highly connected vertices.

1.4 Self-organised criticality

In the previous section, the tendency of complex systems to overcome any effects brought on by errors and attacks has been discussed. However, it is interesting to know that not only can an unstable complex system resist changes, it can do so on its own without any external influences. Although at a disordered state, it regulates into a phase which has the same properties as those found in the ordered systems at the critical point. Examples include

the size distribution of avalanches in the sandpile model [45, 46, 47] and scaling behavior in the lattice gas model [48, 49, 50]. This inclination of disordered systems to form patterns solely through the process of self-assembly is commonly referred to as self-organised criticality (SOC) [16, 51, 52] in complexity theory.

Despite tremendous efforts by scientists to pin down SOC mathematically, it still lacks a mathematical formalism. Although much is still unknown about complex systems, empirical studies have shown that the distribution of frequency with which events occur in a system at SOC state tends to follow a power law. Examples of such systems that display this SOC behaviour include the financial markets and hospital waiting lists, both of which will be discussed in detail in later chapters.

An interesting question to ask is what features of a system are required in order for it to display SOC. The answer lies in the time scales of the external and internal dynamical processes [52]. For SOC to occur, the internal relaxation process must be much faster than the process of external influencing. This separation of time scales is due to the presence of threshold and metastability in many complex systems. A good example would be the pushing of a heavy object lying on the floor. A force is applied to the object which slowly builds up over time. When the applied force is smaller than the friction between the object and the floor, the object remains stationary. As the force on the object builds up, the stress between the bottom of the object and the floor increases. Thus, although stationary, the object goes through a series of metastable states each corresponding to certain energy levels. The friction is the threshold for motion; once the applied force is larger than the

friction, the object jerks off along the surface for a short while before stopping and remaining stationary again. These metastable states are barely stable as a slight application of force can result in the object jerking forward or remaining stationary. The relaxation process releases the stress which sends vibrations along the interacting surfaces in the same way [53] as tectonic plates sending off vibrations during earthquakes, and the distribution of energy released during earthquakes has been shown to be power law [10, 15] which has an exponent of 1, commonly referred to as the Gutenberg-Richter law [54, 55]. In short, SOC behaviour occurs in slowly driven threshold systems dominated by interactions. Thus, a constructive definition for complex systems would be Slowly Driven Interaction-Dominant (SDIDT) systems [52] in contrast to SOC, which is a phenomenological definition.

In disordered networks, just like the two-body system mentioned above, threshold and metastability also play a vital role in producing SOC [52]. Consider a network of vertices connected by edges. External forces are applied randomly to the vertices, and each time a force is applied to a vertex, the energy in it starts to build up. A signal can only travel from vertex A to vertex B if A and B are connected and the energies of all vertices connected between A and B, inclusive of B, are above a certain threshold. Due to the error tolerance characteristic of SDIDT systems, the internal dynamics of the system will attempt to conserve its original state of relaxation. An analogy to this is inertia described in Newton's first law of motion. The internal dynamics will stop every time the relaxation process is over. The slow driver eventually brings the energies of a portion of the vertices to rise above the threshold and a new signal starts to move through the system.

Since the driver applies a force to a randomly selected vertex, the path taken by the signal is a random walk [56, 57]. It is known that in higher dimensions, random walks are associated with fractals [58, 59, 60] which display self-similarity. Since the activated areas are composed of fractals of greatly varying sizes, the distribution of relaxation durations are expected to be a power law.

Although power law distributions occur in many complex systems, care must be taken when using a power law distribution as the indicative sign that a system is at SOC state [52]. To establish that a system is truly at SOC, both spatial and temporal quantities must display scale-free characteristics. This is because power laws can arise in many systems without any underlying critical state. So, both spatial fractals and fractal time series must be identified before one can claim that a system is critical. However, spatial fractals are hard to detect. In most investigations of systems with spatial structures, the size distribution of an event is measured after triggering the system into an avalanche or a self-sustaining chain of events.

The fact that power law distributions stem from so many seemingly different systems indicates that the occurrence of SOC is independent of minor details. With this principle of universality, one can safely discard all but the qualitatively important features when constructing a model of a complex system without losing the SOC feature. Besides making it easier to analyse the dynamics of the interacting agents, a simplified model also allows one to understand the real world intuitively without all the complicated features. Thus, all models introduced in this thesis are simplified models of the real dynamical systems.

1.5 Criticality in a cellular automaton: the sandpile model

The concept of self-organized criticality was illustrated in 1988 by Bak, Tang and Weisenfeld [45] using a sandpile model. In this paradigm, grains of sand are added onto the pile, causing avalanches to occur. Although the model is oversimplified and is far from adequate to model the dynamics of a real physical sandpile, it possesses the necessary ingredients to capture the essence of SOC so commonly observed in real systems.

Consider a system with a two dimensional grid with coordinates (x, y). The number of grains of sand at (x, y) is denoted by $Z(x, y)$. The grains in this model are assumed to be cubes of equal sizes which can be stacked on top of one another. A grain is added to a randomly chosen site, so that

$$Z(x, y) \rightarrow Z(x, y) + 1 \tag{1.23}$$

and this process is repeated over a certain number of time steps. The authors of this model proposed that grains of sand at a site will topple once a certain arbitrarily set threshold Z_{cr} is reached. For instance, if the threshold is set to 8, and if the Z of a site (x, y) has reached 9, one grain will move into each of the surrounding sites. Consequently,

$$Z(x, y) \rightarrow Z(x, y) - 4 \tag{1.24}$$

and

$$Z(x \pm 1, y) \rightarrow Z(x \pm 1, y) + 1 \tag{1.25}$$

$$Z(x, y \pm 1) \rightarrow Z(x, y \pm 1) + 1 \tag{1.26}$$

for $Z(x, y) > Z_{cr}$. This is illustrated in Fig. 1.3. When the Z of a site at the boundary of the pile exceeds the threshold, the grains topple out of the system and are no longer considered.

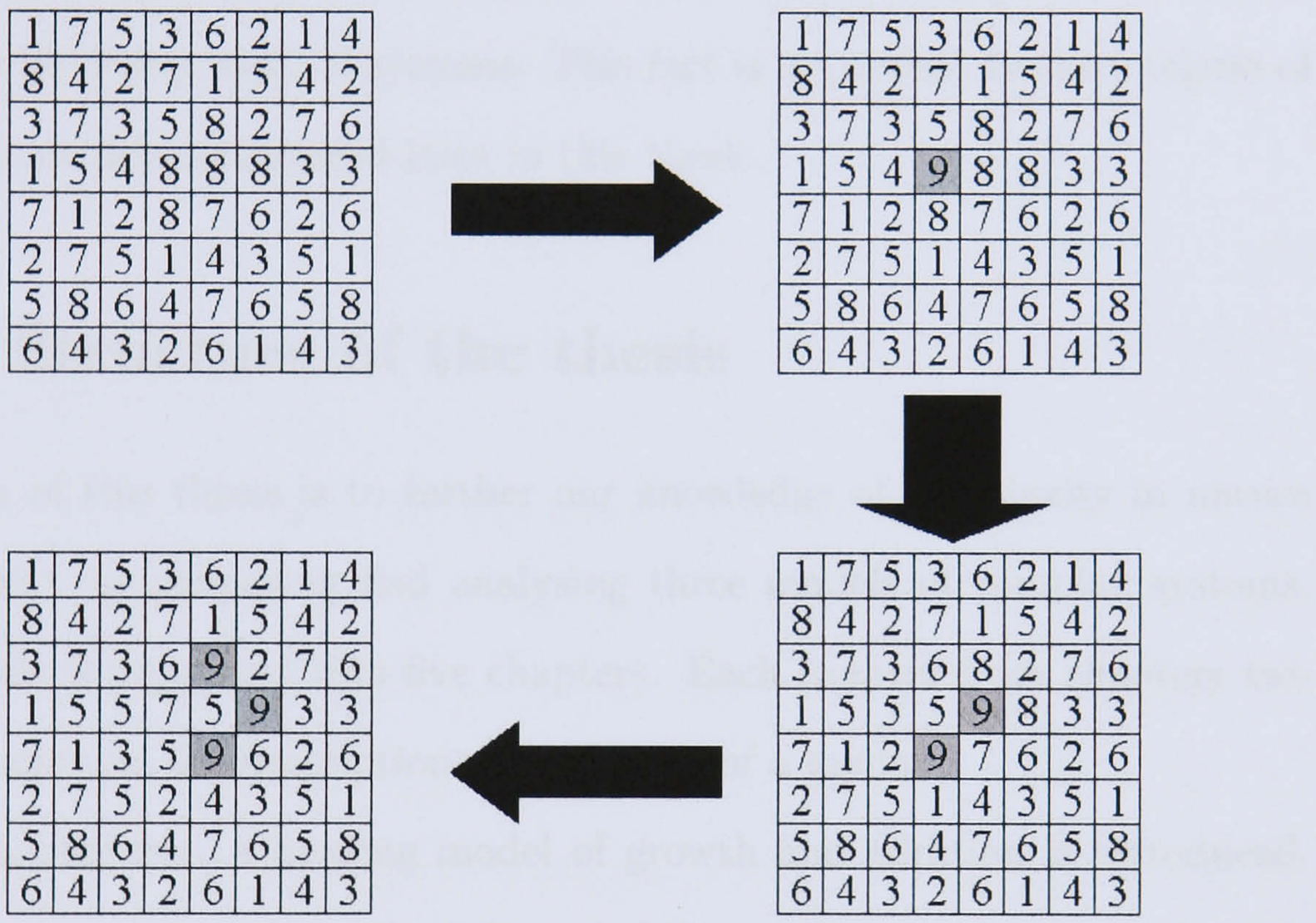

Figure 1.3: An illustration of the sandpile model algorithm. The addition of a single grain to a site in the pile causes a series of toppling events. The grey boxes correspond to the unstable sites.

Initially, all sites are stable and no toppling events occur. As time progresses, a single toppling event occurs which does not induce further events since the neighbours are unlikely to gain a height exceeding Z_{cr}. As more

and more grains are added to the system, the toppling at a site induces other toppling events, causing an avalanche. It has been shown that the size distribution of the avalanches follows a power law with an exponent of approximately 1.1 [16], indicating that the system is at SOC state. This should come as no surprise if one recalls from the previous section that the ingredients essential for SOC are a threshold and metastability.

The randomness in the choosing of sites in this model does not affect the power law behaviour, and this is an indication that randomness does not play a role in the complexity of systems. This fact is important to the analysis of all of the models introduced later in this thesis.

1.6 Structure of the thesis

The aim of this thesis is to further our knowledge of complexity in nature and society by proposing and analysing three models of complex systems. The thesis is organized into five chapters. Each chapter from chapters two to four contains the description and analysis of a model.

In chapter two, a herding model of growth and addition is introduced. At each time step either (1) with probability p the system grows through the introduction of a new agent or (2) with probability $1-p$ a free agent already in the system is added at random to a group of size k with rate A_k. Two versions of the model, $A_k = k$ and $A_k = 1$ are solved and in both versions different types of behaviour have been found. When $p > 1/2$, all of the moments of the distribution of group sizes are linear in time for large time and the group size distribution follows a power law. When $p = 1/2$, the

time taken by the system to run out of free agents scales as the square of the initial number of free agents in the system. When the $p < 1/2$, the number of free agents runs out in a finite time, and this time is proportional to the initial number of free agents in the system.

In chapter three, a model of hospital waiting lists is introduced. Patients entering the system must choose a waiting list to join, based on its length. At the same time patients leave the system as they get served. This model illustrates how the power law distributions found in empirical studies might arise, but indicates that the mechanism causing the power law is unlikely to be the preferential attachment of the patients or their physicians.

In chapter four, a model of salary distributions is introduced. In this model, a higher paid individual does nothing but the lower paid individual leaves the organisation and is replaced by another, whose salary is picked from a power law distribution. The resulting distribution of the salaries has been found to be power law with a different exponent. Variants of this model have also been used to show that the resulting distribution is dependent on the distribution from which the new individual's salary is chosen. It has also been found that the exponent of the resulting distribution is dependent on the total number of individuals comparing salaries. A mean field version of the model is also studied by carrying out numerical simulations. A comparison of the two versions is made by looking at the mean absolute difference between salaries in each version.

Conclusions and discussions are presented in the final chapter and possible future work proposed.

Chapter 2

Growth and addition in a herding model

2.1 Coherent group reactions in a society

A 'herd instinct' exists in a large portion of the animal kingdom, from cattle in the Mongolian deserts to traders in the stock market of Wall Street. This "herd instinct" is an inhibitive characteristic in a lot of higher order organisms, a force that drives an individual to follow a group either to a place or in making some crucial decisions. This behaviour may be due to the intention of an individual to identify itself to a group in a certain style, but in most cases, this is a survival tactic and is a product of evolution and natural selection. For example, a cow is better at defending itself against a predator when in a group [61], and a trader follows the decisions of a group of traders [62] when he is unsure of his own judgements.

Since individuals pick groups to follow, there must exist an algorithm

for choosing which group to follow. In certain situations, this may be a totally random decision, such as in the case of fashion where a person chooses which types of clothing to wear. The randomness in this situation is largely attributed to the mood of the individual at a particular time, as is nausea in the case of mass hysteria. However, in most cases it is a result of carefully made decisions based on various factors, especially when it concerns the survival of the individual. For example, a cow may choose to follow a larger herd as it offers better protection than a smaller one. In the internet, a web site is more likely to be attached to a larger web ring as it provides a greater pool of web surfers, thereby ensuring its utilisation. Also, an investor chooses to follow a certain equity fund with a famous fund manager.

In recent years, the algorithm of the herding behavior has provided a basis for physicists to model various social and natural systems. A successful attempt was made by Cont and Bouchaud [62] where a herding model of traders in the stock market was introduced. Each trader in this model chooses a group to follow and remains in the same group throughout the duration of the simulation. This is followed by the Equíluz-Zimmermann model [63] which is a kinetic version of the original Cont-Bouchaud model. In this model, groups of agents can either coagulate or fragment at each time step. Both of these models will be discussed in greater detail in the following subsections.

Motivated by this work, a herding model with a growth and addition algorithm is introduced in this chapter. In this model, instead of just having a system with a static size, growth is included to better represent population growth in certain social networks, such as the herds of nomadic travelers. The growth feature depicts the increase in size of the system in the form of

births, making it more realistic. The newborn child may choose to follow the existing group or any other groups in the system. Each individual is attached to a group with a certain probability, and three cases are introduced corresponding to three different attachment probabilities. This model is described in detail in section two.

2.2 Herding models

2.2.1 Previous attempts to explain heavy-tails in empirical studies

Transactions in stock markets have generated much wealth for many traders and is an instrumental part of the economy. It is thus beneficial to study it in detail and research has been conducted to try and understand more about the dynamics of various aggregate market variables. However, attempts at fitting the distributions of stock prices and returns with Gaussian distributions [64, 65, 66] have failed, even after the consideration of heteroskedasticity in the data. Applications of fundamental economic variables, including the arrival of information, have failed to explain the 'bursts' of volatility. The distinctive feature of the plots is a heavy-tail [67, 68, 69] with large excess kurtosis κ, defined as

$$\kappa = \frac{\mu_4}{\sigma^4} - 3 \tag{2.1}$$

where μ_4 and σ are the fourth central moment [64, 70] and the standard deviation of the quantity measured, respectively. The κ of a Gaussian distribution is zero, but the κ of daily returns in the stock market is found to

range from 2 to 50 [71, 72]. This excess kurtosis indicates a slow decay of the daily return distribution and is a result of the large fluctuations of the prices. In some cases, the tail takes the form of an exponential decay [73], and in others it conforms to a power law [72]. In Ref. [74] two statistics TP and TE corresponding to power law and exponential decay were introduced, and these statistics were used to test whether a sample tail is power law or exponential.

Various attempts have been made to explain these heavy-tails. Among these are some statistical mechanisms proposed to account for the large fluctuations of the aggregate market variables [75, 76, 77]. The model in [77] made use of the fact that conditional heteroskedasticity leads to an unconditional distribution of returns with heavy-tails. The drawback of this model is the assumption that the return is conditionally normal. A conditional non-normal distribution has been shown to produce better heavy-tails [78].

In [75] random variables are drawn from Lévy [79, 80, 81] distributions. The problem with this model is that the infinite sample variance is not in agreement with empirical data; the latter showed that the sample variance remains constant beyond a certain value of sample size.

Brownian motion was used as a subordinated process in [76] and [82] in an attempt to mimic the large fluctuations in the stock prices. The subordinator was chosen to be the trading volume and number of trades in [76] and [82] respectively. Although the randomness associated with Brownian motion does resemble the random fluctuations of the stock prices, none of the choices for the subordinator gives a normal distribution of increments, suggesting that Brownian motion does not provide a complete explanation for the heavy-

tails.

Following the failure of introducing statistical mechanisms to account for the heavy-tails, the existence of a more fundamental market mechanism was proposed. In [83] and [84] a model based on the interaction between fundamental traders and noise traders was introduced. Although the computer simulation results do yield heavy-tail distributions in the asset return, the model is too complicated to be analysed properly, making it difficult to determine the specific feature of the model that causes the heavy-tail phenomenon. The complexity of this model also prevents any predictions being made that can be compared to empirical results.

2.2.2 Early theoretical studies of herding behavior

As the search for an explanation of the heavy-tails continues, attention was turned to the herding behaviour among traders. It has long been known among stock traders that the large price fluctuations is a result of the 'crowd' effect of the traders. The empirical evidence of herd behavior in speculative markets was reported in a number of articles [85, 86, 87]. However, this has only recently appeared in the economics literature [88, 89, 90].

In [88] and [89], the individuals take turns to make decisions on the transactions, and the decision of each individual is based on that of the preceding trader. This ordering of traders is highly unrealistic since the unordered actions of the traders is essential for the fluctuations in the aggregate market variables.

A model based on unordered actions of traders was introduced in [90]

where random pairs of agents in the system make the same decision on the transaction. However, the plots of aggregate variables do not conform to the heavy-tails seen in the empirical studies. This is due largely to the unrealistic imitation algorithm where any two agents have the same tendency to make the same decision.

A herding model successful in explaining the heavy-tails seen in the distributions of aggregate market variables was introduced by Cont and Bouchaud in [62]. In this article, it was shown that the return distribution based on the model matches the heavy-tail seen in empirical studies, thereby indicating a link between herd behaviour and large return fluctuations. This herding model is described in the following subsection.

2.2.3 The Cont-Bouchaud model

Cont and Bouchaud [62] introduced a system of vertices denoted by i where $1 \leq i \leq N$, with each vertex representing a trader in a stock market. The vertices are grouped into clusters denoted by α where $1 \leq \alpha \leq n_c$. At each time step, the traders can carry out one of three possible actions on the single asset of the system: buying, selling or no trading. To imitate the herd behavior of the stock market traders, all traders in the same cluster make the same decision. The action of cluster α is represented by ϕ_α which takes values of $+1$, -1 and 0 corresponding to buying, selling and no trading respectively. The price of the asset at time t is denoted by $x(t)$. In that

model, the return or price change Δx is assumed to be

$$\Delta x = \frac{1}{\lambda} \sum_{\alpha=1}^{k} W_\alpha \phi_\alpha \tag{2.2}$$

where W_α is the size of cluster α and λ is the amount of excess demand required to move the price of the asset by one unit, and k is the total number of clusters. The distribution of Δx has been shown to follow a power law [62] which agrees with empirical results [91].

2.2.4 The Eguíluz-Zimmermann model

In the Cont-Bouchaud Model described in the previous subsection, all members of a cluster remain in the same group throughout the entire simulation process. The topology of the model is thus static and although this may resemble the loyalty of investors to their respective fund managers, it is rarely observed in actual financial markets. Investors, unlike loyal subjects in herds or tribal clans, are only driven by profits and personal financial successes. Thus, a more dynamic model is needed to represent this feature.

In response to this, a kinetic version of the Cont-Bouchaud Model was introduced by Eguíluz and Zimmermann in [63]. This was solved exactly by D'Hulst and Rodgers [94] and since then a number of generalised versions of this model have been proposed [95, 96, 97]. This Eguíluz-Zimmermann model, retains most of the features in the original herding model whilst allowing the topology of the network to evolve instead of letting it stay static. In general, the model is characterised by random information dispersion, coherent group decisions and adjustments of the network in response to the

actions of a cluster.

Consider a system of N vertices representing traders with each vertex labeled i where $1 \leq i \leq N$. The three possible actions ϕ_i of a trader i are buying, selling and no trading, corresponding to $\phi_i = +1$, $\phi_i = -1$ and $\phi_i = 0$ respectively. At the initial time step, all vertices are not linked i.e. they are all isolated, and $\phi_i = 0 \; \forall \; i$. At each time step t_l, where $1 \leq l \leq \infty$, one of the vertices is randomly selected. With probability a, the vertex makes the same random decision as all the vertices linked to it on whether to buy or sell the single asset of the system. At the end of each time step, all links between the vertices in the cluster are removed. With probability $1 - a$, the vertex remains inactive, i.e., no trading, and it is randomly connected to any other vertices in the network. This process is repeated for L time steps.

Thus, the clustering coefficient c_{t_l}, defined as the average number of vertices a vertex is connected to at time t_l, will increase so long as a randomly selected trader at each time step remains inactive. In fact, no trading activity occurs in any part of the network until one cluster decides to carry out actual transactions, after which all coalitions in the network are removed. Therefore, the averaged clustering coefficient C is controlled by the parameter a. For $a = 1$, trading is carried out by one agent at each time step, leaving no chance for the formation of clusters. All of the vertices will remain isolated throughout the entire simulation process, and no herding behaviour will be seen. For $a \ll 1$, trading rarely happens in the system, and this prolonged state of inactivity between tradings allows for the build up of large clusters. With larger clusters, the order is also significantly bigger when a trading decision finally arrives, thereby maximising the impact of herding on the

market. This extreme scenario occurs when $a \ll O(1/N)$. Since the parameter a determines the degree of herding behaviour in the system, a 'herding parameter' h in terms of a

$$h = \frac{1}{a} - 1 \tag{2.3}$$

can be defined. When $a = 1$, $h = 0$ and no herding occurs. For $a < 1$, $h > 0$ and herding is observed. Besides indicating the presence of herding behaviour, h also allows one to work out the number of edges between any two vertices. The latter is important as it concerns the rate of dispersion of the information.

When a selected vertex is inactive, the random linkage formed corresponds to random information dispersion in actual financial markets. The linkage may bring about the merging of two clusters, and the information is dispersed throughout the newly merged cluster. An order is placed to an external centralised market maker after a cluster initiates trading . The removal of links between vertices after a trading activity is realistic since the information which binds the members of a cluster together is useless after a transaction.

The price index dynamics in this model adopts the update rule in [102]. The price index P is such that at time step t_{l+1},

$$P(t_{l+1}) = P(t_l) \exp\left[\frac{s_l}{\lambda}\right] \tag{2.4}$$

where s_l is the order size of the active cluster at that time step, and λ is a parameter that controls the size of the update and can be used as a measure

of the liquidity of the market. The price return R is given by

$$R(t_l) = \ln P(t_l) - \ln P(t_{l-1}) \tag{2.5}$$

and from Eqn. (2.4) the price return is proportional to the order size.

Numerical simulations in [63] have shown that the C increases with h, displaying the impact of cluster sizes on the herding activity. The simulations have also shown power law decays in the distributions of R for various values of h. The exponent of the decay α was found to be 1.5. For values of h smaller than a critical value h^*, there exists an exponential cutoff, and the C for these values of h are also far from C^*, the critical C. For $h > h^*$, the entire range of R can be fitted with power law, and the C is close to C^*. In addition to that, the probability of getting extremely high returns also increases for this range of h.

Since the return R is dependent on the order size s, a relationship exists between the distributions of return and cluster size. Numerical simulations have shown that the distribution of cluster size follows a power law with an exponent $\beta = 2.5$ [63]. The distribution of R is equal to the product of the distribution of cluster and the probability of getting a cluster which is proportional to the cluster size. The probability density of R is therefore given by

$$P(R) \approx ss^{-\beta} \approx R^{-\alpha} \tag{2.6}$$

Thus, there is a unit difference between the exponents of the return and cluster size distributions.

2.2.5 Other related models

In Refs. [98, 99] a model based on the kinetics of the order book was proposed in which either market or limit bids to buy or sell were made at random. It was shown that this model leads to a power law distribution of returns under simple assumptions about the kinetics of this process.

In Ref. [100] a herding model in which, at each time step, either an incoming agent joins an existing group or a group is fragmented into individual agents, and the probability of each of these events is fixed. This is a simpler version of a model proposed in [101] in which the above steps occur at each time step, but with rates determined by the number of individual agents.

2.3 Introducing growth and addition to a herding model

In an attempt to further the work in [62] and [63], a herding model is introduced in this chapter which is characterised by two features: growth and addition. At each time step, one of two events can happen. With probability p a growth event takes place in which an agent is created but remains free i.e. remains unconnected. With probability $q = 1 - p$ a free agent already in the system joins a group of size k with a rate proportional to k, where $k = 1, 2, ...$ The number of groups of size $k > 1$ at time t, denoted by $n_k(t)$, evolves like

$$\frac{dn_k(t)}{dt} = \frac{q}{M(t)}[(k-1)n_{k-1} - kn_k] \tag{2.7}$$

where $M(t)$ is the number of agents in the system given by

$$M(t) = \sum_{k=1}^{\infty} k n_k(t). \tag{2.8}$$

The first and second terms on the right hand side of Eqn. (2.7) correspond to the addition of a new agent to a group of $k-1$ and k respectively. Since the probability of an addition event at each time step is q, and the probability that an existing agent joining a group of size $k-1$ is proportional to $k-1$, the product of q, $k-1$ and n_{k-1} in the first term on the right hand side of Eqn. (2.7) corresponds to the increase in the number of groups of size k formed due to the destruction of the groups of size $k-1$. Similarly, in the second term, the product of q, k and n_k corresponds to the decrease in the number of groups of size k due to the formation of a group of size $k+1$. The number of free agents behaves like

$$\frac{dn_1(t)}{dt} = -q\left[\frac{n_1}{M(t)} + 1\right] + p. \tag{2.9}$$

The first term of Eqn. (2.9) represents the destruction of free agents by addition and the second term describes the arrival of free agents. By summing Eqn. (2.7) over all $k > 1$ we have

$$\begin{aligned}
\frac{d}{dt}\left(\sum_{k=2}^{\infty} n_k(t)\right) &= \frac{q\sum_{k=2}^{\infty}[(k-1)n_{k-1} - kn_k]}{M(t)} \\
&= \frac{q\left(\sum_{k=2}^{\infty} kn_k - n_k\right)}{M(t)} - \frac{q\sum_{k=2}^{\infty} kn_k}{M(t)}
\end{aligned}$$

$$= \frac{qn_1}{M(t)}. \tag{2.10}$$

By defining

$$N(t) = \sum_{k=1}^{\infty} n_k(t) \tag{2.11}$$

and adding Eqn. (2.9) to Eqn. (2.10) gives

$$\frac{dN(t)}{dt} = 2p - 1. \tag{2.12}$$

Similarly, multiplying Eqn. (2.7) by k and summing over all $k > 1$ gives

$$\begin{aligned}
\frac{d}{dt}\left(\sum_{k=2}^{\infty} kn_k(t)\right) &= \frac{q}{M(t)} \sum_{k=2}^{\infty} \left[k(k-1)n_{k-1} - k^2 n_k\right] \\
&= \frac{q}{M(t)} \left[\sum_{k=2}^{\infty} k^2 n_k + \sum_{k=1}^{\infty} kn_k + n_1 - \sum_{k=2}^{\infty} k^2 n_k\right] \\
&= q\left[1 + \frac{n_1}{M(t)}\right],
\end{aligned} \tag{2.13}$$

and adding it to Eqn. (2.9) we have

$$\frac{dM(t)}{dt} = p. \tag{2.14}$$

Eqn. (2.12) indicates that on average, the number of groups increases by one with probability $2p - 1$. Eqn. (2.14) represents the fact that the number of agents increases by 1 with probability p. It is obvious from Eqn. (2.12) that on average, when $p > 1/2$ the system is growing, when $p = 1/2$ the system is static and when $p < 1/2$ the system decreases in size and runs out of free monomers in a finite time. Each of these cases is analysed separately in the

following subsections.

2.3.1 Case 1: $p > \frac{1}{2}$

For large t, the form of Eqns. (2.7, 2.9, 2.12, 2.14) suggests that for $p > 1/2$ the solution for n_k is linear in time. In this limit, Eqns. (2.12, 2.14) can be solved to yield

$$N(t) = (2p - 1)t \tag{2.15}$$

and

$$M(t) = pt \tag{2.16}$$

for large time. By writing

$$n_k = tc_k \tag{2.17}$$

we find that for $k > 1$

$$c_k = \frac{1-p}{p}[(k-1)c_{k-1} - kc_k]. \tag{2.18}$$

To solve Eqn. (2.18) we first work out an initial condition from Eqn. (2.9). From Eqn. (2.9) and using Eqns. (2.16) and (2.17) we have

$$\begin{aligned} c_1 &= -q\left[\frac{c_1}{p} + 1\right] + p \\ &= p(2p-1). \end{aligned} \tag{2.19}$$

An iteration of Eqn. (2.18) shows that

$$c_k = \frac{c_1(k-1)!}{(2+\frac{p}{1-p})(3+\frac{p}{1-p})(4+\frac{p}{1-p})\ldots(k+\frac{p}{1-p})}$$

$$
\begin{aligned}
&= \frac{p(2p-1)(k-1)!}{(2+\frac{p}{1-p})(3+\frac{p}{1-p})(4+\frac{p}{1-p})\ldots(k+\frac{p}{1-p})} \\
&= \frac{p(2p-1)(k-1)!}{(2+\frac{1}{1-p}-1)(3+\frac{1}{1-p}-1)(4+\frac{1}{1-p}-1)\ldots(k+\frac{1}{1-p}-1)} \\
&= \frac{p(2p-1)\Gamma(k)}{\frac{\Gamma(k+\frac{1}{1-p})}{\Gamma(1+\frac{1}{1-p})}} \\
&= p(2p-1)\Gamma\left(1+\frac{1}{1-p}\right)\frac{\Gamma(k)}{\Gamma(k+\frac{1}{1-p})} \qquad (2.20)
\end{aligned}
$$

where

$$\Gamma(n+1) = n! \qquad (2.21)$$

for integer values of n. As $k \to \infty$

$$c_k \sim k^{-\frac{1}{1-p}}. \qquad (2.22)$$

This result can be used to model the financial markets. Similar to the Cont-Bouchaud Model, one can imagine that, at each time step, an agent is randomly chosen, independent of the group kinetic process. The entire group to which the selected agent is linked will carry out the actions of buying or selling an asset with equal probability. In this way, a group of size k trades with rate kc_k. It is assumed that the traded amount is proportional to the size of the group. Thus, the distribution of returns $R(k)$ for a commodity, which is equivalent to the distribution of the difference between the number of buyers and sellers, behaves like

$$R(k) \sim kc_k = p(2p-1)\Gamma\left(1+\frac{1}{1-p}\right)\frac{\Gamma(k+1)}{\Gamma(k+\frac{1}{1-p})}. \qquad (2.23)$$

In the limit $k \to \infty$ we have

$$R(k) \sim k^{-\beta} \tag{2.24}$$

with

$$\beta = \frac{p}{1-p}. \tag{2.25}$$

The distribution of returns is thus a power law for large k. It is obvious that β can take any value greater than unity for $1/2 < p < 1$, with $\beta \to 1$ as $p \to 1/2$ and $\beta \to \infty$ as $p \to 1$.

2.3.2 Case 2: $p = \frac{1}{2}$

When $p = 1/2$ the size of the system remains static, on average. By having an initial condition of N free agents we have

$$N(0) = M(0) = n_1(0) \equiv N. \tag{2.26}$$

From Eqn. (2.7) and using Eqn. (2.17) we have

$$n_k(t) = \frac{(k-1)n_{k-1}t}{2N + (1+k)t}. \tag{2.27}$$

For small k where $k > 1$, Eqn. (2.27) becomes

$$n_k(t) \simeq \frac{n_{k-1}t}{2N + t}. \tag{2.28}$$

From Eqn. (2.26), $n_1(0) = N$, so

$$n_1(t) = \frac{2N^2}{2N + t}. \tag{2.29}$$

From Eqn. (2.28) it can be seen that

$$n_k(t) = \frac{n_1 t^{k-1}}{(2N + t)^{k-1}}. \tag{2.30}$$

Substituting Eqn. (2.29) into Eqn. (2.30) gives the full time dependent solution

$$n_k(t) = 2N^2 \frac{t^{k-1}}{[t + 2N]^k}. \tag{2.31}$$

From Eqn. (2.31) it can be seen that n_k grows initially before decaying to zero. On average, the system runs out of free agents when $n_1(t) \sim O(1)$ which occurs on the time scales $t \sim O(N^2)$. This observation is in agreement with a simple random walk argument; at time $t = 0$ there are N free agents which increases by 1 with probability 1/2 and decreases, by either 1 or 2, with probability 1/2. The time taken to reach zero free agents is thus expected to scale as $t \sim O(N^2)$.

From Eqn. (2.31) we have

$$N(t) = N \tag{2.32}$$

and

$$M(t) = N + \frac{t}{2}. \tag{2.33}$$

For $t \ll N$, the number of groups of size k grows like

$$n_k(t) \sim N \left[\frac{t}{2N}\right]^{k-1}. \tag{2.34}$$

In the limit $N \ll t \ll N^2$, n_k decays as

$$n_k \sim \frac{1}{t}. \tag{2.35}$$

In this limit, the r^{th} moment $M_r(t)$, defined by

$$M_r(t) = \sum_{k=1}^{\infty} k^r n_k(t), \tag{2.36}$$

decays as

$$M_r(t) \sim \left(\frac{t}{2}\right)^r. \tag{2.37}$$

2.3.3 Case 3: $p < \frac{1}{2}$

For $p < \frac{1}{2}$ the number of agents in the system increases but the number of groups of free agents falls, and the system runs out of free agents in a finite time. To solve Eqn. (2.9) , it is first rewritten as

$$\begin{aligned} \frac{dn_1}{dt} &= -q\left[\frac{n_1}{N\left(\frac{N+pt}{N}\right)} + 1\right] + p \\ &= -q\left[\frac{n_1}{NT} + 1\right] + p \end{aligned} \tag{2.38}$$

where $T \equiv \left(\frac{N+pt}{N}\right)$. Eqn. (2.38) is then rewritten as

$$\frac{dn_1}{dt} + q\frac{n_1}{NT} = p - q. \tag{2.39}$$

Since $\exp[\ln T^{\frac{1-p}{p}}] = T^{\frac{1-p}{p}}$, by multiplying LHS of Eqn. (2.39) by $\exp[\ln T^{\frac{1-p}{p}}]$ and RHS by $T^{\frac{1-p}{p}}$ and using the fact that

$$\frac{d\left[\exp[\ln T^{\frac{1-p}{p}}]\right]}{dt} = \exp[\ln T^{\frac{1-p}{p}}]\left[\frac{dn_1}{dt} + \frac{q}{NT}n_1\right] \tag{2.40}$$

Eqn. (2.39) is solved to give

$$n_1(t) = N\left[\frac{2(1-p)}{\left[1+\frac{pt}{N}\right]^{\frac{1}{p}-1}} - (1-2p)\left(1+\frac{pt}{N}\right)\right]. \tag{2.41}$$

At $t = \tau$,

$$n_1(\tau) = 0 \tag{2.42}$$

and τ is

$$\tau = \frac{N}{p}\left[\left[\frac{2(1-p)}{1-2p}\right]^p - 1\right]. \tag{2.43}$$

Therefore, τ/N is the average length of time for the system to run out of free agents, and this is shown as a function of p in Figure 2.1.

In the limit when no new agents are added, $p \to 0$ and we have

$$\tau = N\log 2. \tag{2.44}$$

This time is shorter than N since some of the free agents form a dimer

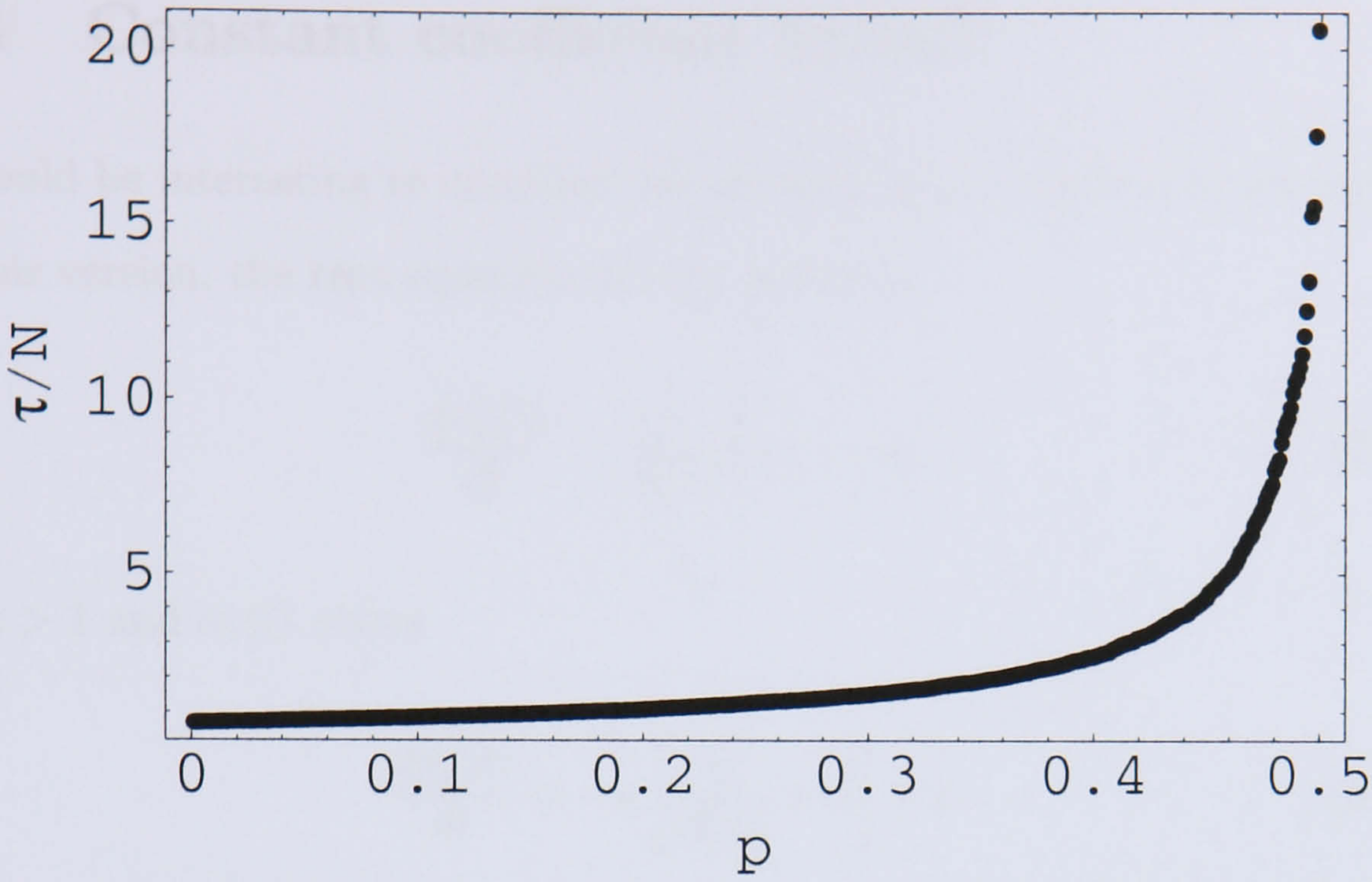

Figure 2.1: Average of the time τ/N it takes the system to run out of monomers, as a function of p. This simulation was done with $N = 10^4$ over 10^7 time steps.

with another monomer, eliminating two free agents in one time step. When $p \to 1/2$ from below, the average time to run out of free agents diverges as

$$\tau \sim 2N(1-2p)^{-\frac{1}{2}}. \tag{2.45}$$

The number of dimers, trimers etc... as well as the total number of agents all become zero in finite time, but the time taken to run out of free agents is shorter and is thus the most important timescale in the system.

2.4 Constant coefficient kernel

It would be interesting to consider the constant kernel version of the model. In this version, the rate equation for the system is

$$\frac{dn_k(t)}{dt} = \frac{q}{N(t)}[n_{k-1} - n_k] \tag{2.46}$$

for $k > 1$ and $n_1(t)$ obeys

$$\frac{dn_1(t)}{dt} = -q\left[\frac{n_1}{N(t)} + 1\right] + p. \tag{2.47}$$

Eqn. (2.46) can be solved by using a method similar to that used in the previous section. For $p > 1/2$ we have

$$n_k(t) = (2p-1)^2 \frac{(1-p)^{k-1}}{p^k} t \tag{2.48}$$

and for $p = 1/2$, the solution is

$$n_k(t) = N \frac{t^{k-1}}{(2N)^{k-1}(k-1)!} \exp\left(-\frac{t}{2N}\right). \tag{2.49}$$

When $p < 1/2$ the system runs out of free agents in time τ which is given by

$$\tau = \frac{N}{1-2p}\left[1 - \left[\frac{1-2p}{1-p}\right]^{\frac{1}{p}-2}\right]. \tag{2.50}$$

As $p \longrightarrow 0$ we have

$$\tau = N(1 - \exp[-1]) \tag{2.51}$$

which, as in the previous version of the model, is less than N because sometimes two free agents are destroyed in one time step. When $p \to 1/2$ from below, τ diverges logarithmically as

$$\tau \sim -2N \log(1 - 2p) \tag{2.52}$$

which is slower than the algebraic divergence seen in the linear kernel model in the previous section.

2.5 Comparison with other herding models

In the the Cont-Bouchaud model, the topology of the network is fixed for all time steps. However, our model allows for the population to grow in size over time.

In the Eguíluz-Zimmermann model, the size of the system is fixed for all times, a feature shared by the Cont-Bouchaud model. However, unlike the former, the agents in our model remain loyal to their respective groups and no links are broken.

Despite these differences, all three herding models produce the heavy tail distributions seen in empirical studies. This implies that the herding property of these models plays a vital role in generating the heavy tail distributions. For a model without any spatial structures, the occurrence of power laws in temporal quantities is the sole indicator of SOC behavior. Since all three herding models do not have any spatial structures, the power law distributions seen in these models imply that these systems, and probably the actual

financial markets, are self-similar.

2.6 Conclusions and discussion

A kinetic version of the herding model has been introduced which encompasses growth and addition. These two mechanisms allow for the creation of new agents as well as the formation of groups of agents. When the growth is fast enough the group size distribution conforms to a power law with a parameter dependent exponent. When the number of groups of agents does not change on average, the time it takes the system to run out of free agents scales as the square of the initial number of free agents in the system. When the rate of growth is sufficiently slow, the system runs out of free agents in a finite amount of time, and this time is shown to be proportional to the initial number of free agents in the system. The kinetics for this case do not allow the power law group size distribution to develop. This behaviour is seen in both the linear and constant kernel versions of the model.

Comparisons with other herding models have also been made. The heavy tails seen in all three herding models, despite their differences, imply that the herding property shared by them is the cause of heavy tails seen in empirical studies.

Models in Refs. [63, 94, 100] which are mean field in character appear to be well suited to model simple processes in social networks. These include processes in which an asset is exchanged between people or in which groups of people are formed who all share the same asset. The suitability of these models is in part due to the non-local nature of social interactions in

a modern, highly connected world. Processes for which this approach would appear relevant range from information or rumour spreading, particularly in financial markets, through the take up of the latest craze to epidemiological studies of disease and epidemic spread.

This model is especially relevant to epidemiological studies of disease spread due to its growth mechanism. When this model is applied to population studies, the growth mechanism corresponds to the increase of the population size over time, and the addition mechanism corresponds to the forming of social groups within which individuals are in contact with each other. Probabilistic infection mechanisms, such as those used in standard epidemiological models, can be included to mimic disease spreading. In addition, this model can also be used as an extension to previous herding models discussed earlier.

An interesting point to note in this work is that this model only displays SOC behaviour when $p > \frac{1}{2}$. Unlike previous herding models mentioned earlier in this thesis, this model possesses a growth feature which can be seen as a driver similar to the dropping of sand in the sandpile model. Grains of sand are dropped at a rate of one per unit time step in the sandpile model, and this slow rate is a prerequisite for SOC to occur as mentioned in Chapter 1. Thus, an interesting feature to investigate is the critical rate beyond which no SOC will occur. The herding model introduced in this work is ideal for this investigation since the probability of attachment p at each time step is a continuous real number. The cutoff $p > \frac{1}{2}$ in this work shows that SOC only occurs when on average at least one new agent is introduced after two time steps, thereby extending below the single grain addition per time step in the

sandpile model. Although this has only been found to be true in this model, future research on complexity models may include a search for the general critical driving rate for SOC to occur.

Chapter 3

Simple models of waiting lists

3.1 Motivation for modelling hospital waiting lists

Besides having an adequate supply of physicians and medical equipment, the efficiency of the healthcare system is also instrumental to the health of patients. In particular, the length of waiting lists in hospitals affects the treatment rate of patients, and shorter waiting lists reduce the risk of medical complications due to prolonged delay for treatment.

The hospital waiting lists in Northern Ireland are the longest in the country despite the region receiving the second largest funding for acute health services [103]. The number of inpatients waiting for treatment per 1000 population in 2004 was 30.04, compared to the 25.87 in Wales and 22.23 in England [103]. This figure is lowest in Scotland where the acute health services are the best funded in the country [103]. It was stated in a report [103]

by the House of Commons that the under utilisation of operating theatres is a cause for the lengthy waiting times. Another cause stated in the report is the lack of nurses and beds in the region.

This problem of long waiting lists causes anxiety and suffering among patients waiting for their operations. In response to this crisis, the House of Commons called for better theatre management and control. According to the report [103], the efficiency of operating theatres can be increased by establishing theatre policy and guidelines, and the use of computerised data collection systems.

In attempts to better understand the system of hospital waiting lists, a number of studies have been carried out on their length distribution. This was shown to be power law in Refs. [104, 105]. In Ref. [104], the waiting lists for four dermatologists over a six year period were analysed and the double log plot of the magnitude of monthly variations of waiting time and the frequency of occurrence of this magnitude showed a straight line. In Ref. [105] chaos and complexity theories were used to analyse hospital waiting list data between 1998 and 2001 and the plots of frequency versus quarter to quarter change of waiting times show a power law, independent of surgical specialty and hospital location. To support a view that patients or physicians prefer shorter waiting lists, it has also been shown that there is a correlation between the referral rate to lists and the length of the waiting lists [106]. This correlation had a time-lag which is of the order of the time it takes information about the length of a list to spread through the system.

Power law distributions usually signal that a system is self-organising, and thus resistant to changes. This implies that, for instance, small changes

in the rate at which patients are treated, or the number of consultants, will have little effect on the distribution of waiting lists, or on the length of the longest list in the system. This question is clearly important to many systems, including those in healthcare; can greater efficiency be addressed by small changes to the current system, or is a radical overhaul required?

Similarly, if the distribution is power law, then the average queue length and the length of the longest queue are dependent on the value of the exponent in the power law. It would thus seem to be important to build microscopic models to examine this system and determine under what conditions power laws occur and the dependence of the exponent of the power law on other parameters in the system.

A model of waiting lists which captures the essence of the problem is introduced in this chapter. In this model the lists lengthen or shorten as patients join the lists to be treated or leave the list after treatment. The model illustrates how power-law distributions might arise, but indicates that the mechanism is unlikely to be the preferential behaviour of patients or their physicians. The difference between this model and those proposed by queueing theorists is that in our model there is an infinite number of service channels and queue-length dependent rates. This model is based on Queueing Theory [107, 108, 109, 110] which is described briefly in the next section.

3.2 Queueing Theory and its applications to queue modelling

The purpose of Queueing Theory is to capture the essence of systems of waiting lines. The three basic elements present in a typical queueing model are the arrival paradigm, the service process and the queue-discipline. In a typical model, new agents are created at every time step and select a slot to attach to. Agents at one end of the lines are served and leave the system as departures.

The arrival paradigm can be ordered or completely random, depending on the nature of the system. In ordered arrivals, agents enter the system at a fixed rate $\lambda = \frac{1}{T}$, where T is the total number of time steps considered in the model. An example of such systems is the queueing of printing jobs in a printer. However, when the arrival paradigm is totally random, the probability that no agent enters the system in the time interval $(t, t + \Delta t)$ is $1 - \lambda\Delta t + o(\Delta t)$, that one agent enters the system is $\lambda\Delta t + o(\Delta t)$, and that more than one agent enters the system is $o(\Delta t)$, where $o(\Delta t)$ denotes quantities that become negligible when compared with Δt as $\Delta t \to 0$

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0. \tag{3.1}$$

The events occurring in the interval $(t, t + \Delta t)$ are independent of any other arrival or non-arrival events not overlapping this time interval. If we consider the finite interval $(t, t + T)$, the probability of N arrivals in this interval is a

Poisson Distribution

$$Pr(N = n) = \frac{\exp(-\lambda T)(\lambda T)^n}{n!} \tag{3.2}$$

where $n = 0, 1, 2, ...$

Of particular importance is the service process in the system which affects the rate at which agents leave the system. The two important features of the service process are the rate at which agents are being served and the times when the service is accessible. The former is influenced by the time taken to serve an agent, and is taken to be a random variable X in many models where the distribution of X is known.

The queue-discipline determines how an agent is selected from the queue for service. For single capacity systems, an agent is selected for service after a service has been completed. The selection process is based either on the order of arrival or on a specific classification of the agents. For multi-capacity systems, more than one agent can be served when the service is available. In some cases, the specific needs of the agents are taken into account. The agents in such systems are served by specialised servers and this significantly reduces the throughput. For systems without specialisation, three types of queue-discipline are frequently used in models. The first is one in which agents select the servers in strict rotation. The second and third involve agents choosing which queues to join upon arrival and agents forming a single queue respectively. The first is a mathematically simple system. However, the second and third are more difficult to be analysed using mathematics, and the new model introduced in this chapter includes the second type of

queue-discipline.

This concept of arrival, service and queue-discipline of Queuing Theory is applied to the new model of hospital waiting lists which is described in detail in the following section. Analysis is also carried out to obtain the length distribution of the waiting lists.

3.3 A model of waiting lists with infinite number of channels

In this section, a model with the minimal ingredients to capture the basic empirical measurements from hospital waiting lists is introduced. Imagine a system in which at time t there are n_k waiting lists of length k. Thus, at time t there are

$$\sum_{k=1}^{\infty} k n_k(t) \tag{3.3}$$

patients waiting in the lists. At each time step, a patient in a list of length k is treated with rate A_k and a new patient arrives at the system and joins a list of length k with rate B_k. These treatment and arrival events correspond, respectively, to the service and arrival elements of the Queueing Theory described in the previous section. A new waiting list of length zero is created with rate r. The list distribution $n_k(t)$ satisfies the equation

$$\frac{dn_k(t)}{dt} = A_{k+1}n_{k+1} - A_k n_k + B_{k-1}n_{k-1} - B_k n_k + r\delta_{k,0} \tag{3.4}$$

for $k = 0, 1, 2, ...$ where

$$A_k = \frac{pa_k}{\Sigma_{k=0}^{\infty} a_k n_k(t)} \tag{3.5}$$

and

$$B_k = \frac{qb_k}{\Sigma_{k=0}^{\infty} b_k n_k(t)}. \tag{3.6}$$

The rate parameters p, q and r can be used to adjust the overall relative rates of patient treatment, arrival and of list creation. The time-independent functions a_k and b_k model the rates of leaving and joining a list of length k, respectively. The function a_k is the service rate in a list of length k, and b_k controls the propensity of patients to join a list of length k, and models the patients' preference to join shorter lists.

The first term on the right hand side of Eqn. (3.4) is the gain in the number of lists of length k after a patient in a list of length $k+1$ has been treated. Similarly, the second term is the loss in the number of lists of length k after a patient in a list of length k has been treated. The third term represents the gain in the number of lists of length k when a new patient joins a list of length $k-1$. The fourth term is the loss term. The final term corresponds to the creation of a new list of zero length.

It is a trivial matter to normalise Eqns.(3.5) and (3.6). For example, consider the third term on the right hand side of Eqn. (3.4). The probability per time step of a new patient joining a particular list of length $k-1$ is

$$P = \frac{b_{k-1}}{\Sigma_{k=0}^{\infty} b_k n_k}. \tag{3.7}$$

Thus the rate of creating a list of length k by patient arrival is equal to the

rate of patient arrival q multiplied by the number of lists of length $k-1$, n_{k-1}, multiplied by the probability of a patient joining a list of length $k-1$, Eqn. (3.7) above. This product gives $B_{k-1}n_{k-1}$, which is the third term on the right hand side of Eqn. (3.4). The other terms can be derived in a similar way. Equations similar to Eqn. (3.4) have been used within the field of econophysics [111] and in random growing networks [13].

In the following subsections, different values of the parameters p, q, r, a_k and b_k are used in order to model the various scenarios relevant to the modelling of healthcare systems. In principle, the functions a_k and b_k can take any value , although it is required that $A_0 = a_0 = 0$ because a patient cannot be removed from a waiting list of zero length.

3.3.1 Case 1: $a_k = k$ and $b_k = 1$

This case corresponds to a situation in which patients join a list at a constant rate, independent of its length. At each time step every patient is equally likely to be treated and thus removed from a list, irrespective of the length of the list the patient is in. This choice of joining and removal rates simplifies Eqns. (3.4 - 3.6) considerably. These equations then become

$$\frac{dn_k(t)}{dt} = \frac{p}{M(t)}[(k+1)n_{k+1} - kn_k] + \frac{q}{N(t)}[n_{k-1} - n_k] + r\delta_{k,0} \tag{3.8}$$

where

$$N(t) = \sum_{k=0}^{\infty} n_k(t) \tag{3.9}$$

and

$$M(t) = \sum_{k=0}^{\infty} k n_k(t). \tag{3.10}$$

Thus, the $N(t)$ evolves like

$$\frac{dN(t)}{dt} = r \tag{3.11}$$

and the $M(t)$ like

$$\frac{dM(t)}{dt} = q - p. \tag{3.12}$$

In this scenario, in the large time limit and when $q > p$ and $r > 0$, the n_k for all k and the moments $N(t)$ and $M(t)$ are all linear in time. We write

$$n_k(t) = t s_k \tag{3.13}$$

where s_k obeys

$$s_k = \frac{p}{q-p}[(k+1)s_{k+1} - k s_k] + \frac{q}{r}[s_{k-1} - s_k] + r\delta_{k,0}. \tag{3.14}$$

This equation is solved using the generating function $g(\omega)$ which is defined by

$$g(\omega) = \sum_{k=0}^{\infty} \omega^k s_k. \tag{3.15}$$

First, Eqn. (3.14) is multiplied through by ω^k and summed over all k, giving

$$g(\omega) = \sum_{k=0}^{\infty} \left[\frac{p}{q-p}[(k+1)s_{k+1}\omega^k - k s_k \omega^k] + \frac{q}{r}[\omega^k s_{k-1} - \omega^k s_k] + r\delta_{k,0}\omega^k \right]. \tag{3.16}$$

By using the fact that

$$\frac{dg(\omega)}{d\omega} = \sum_{k=0}^{\infty} k\omega^{k-1} s_k \tag{3.17}$$

Eqn. (3.16) is reduced to

$$g(\omega) = (1-\omega)[\frac{p}{q-p}\frac{dg}{d\omega} - \frac{q}{r}g(\omega)] + r. \tag{3.18}$$

Eqn. (3.18) will be solved for $\omega \to 0$ and $\omega \to 1$. For $\omega \to 0$, and by placing

$$\frac{dg(\omega)}{d\omega} = \sum_{k=1}^{\infty} k\omega^{k-1} s_k \tag{3.19}$$

and Eqn. (3.15) into Eqn. (3.18) we have

$$\sum_{k=0}^{\infty} \omega^k s_k = \frac{p}{q-p} \sum_{k=1}^{\infty} k\omega^{k-1} s_k - \frac{q}{r} \sum_{k=0}^{\infty} \omega^k s_k. \tag{3.20}$$

Since

$$\sum_{k=1}^{\infty} k\omega^{k-1} s_k = \sum_{k=0}^{\infty} (k+1)\omega^k s_{k+1} \tag{3.21}$$

Eqn. (3.20) can be rearranged to give

$$\sum_{k=0}^{\infty} \omega^k \left(1 + \frac{q}{r}\right) s_k = \sum_{k=0}^{\infty} \omega^k \frac{p}{q-p}(k+1)s_{k+1}. \tag{3.22}$$

By the theorem of identity, if

$$\sum_{i=0}^{\infty} a^i c_i = \sum_{i=0}^{\infty} b^i d_i \tag{3.23}$$

then

$$c_i = d_i \ \forall \ i \geq 0. \tag{3.24}$$

So, by comparing the coefficients of Eqn. (3.22) we have

$$s_k = \frac{\left[\frac{(r+q)(q-p)}{rp}\right]^k s_0}{k!}. \tag{3.25}$$

Thus,

$$s_k = s_0 \exp\left[\frac{(r+q)(q-p)}{rp}\right] \tag{3.26}$$

for $\omega \to 0$. It can be seen from Eqn. (3.26) that for $q > p$ the number of lists of size k increases exponentially, and for $q < p$ this list size distribution follows an exponential decay. This agrees with the intuition that if the arrival rate is greater than the treatment rate, the number of patients waiting for treatment is large, and if the patients get treated more quickly than the arrival of new patients, the size of the waiting population dwindles.

For $\omega \to 1$, an integrating factor $(1-\omega)^{\frac{q-p}{p}} \exp \frac{q(p-q)}{pr}\omega$ is used to solve Eqn. (3.18). Multiplying Eqn. (3.18) by this factor gives

$$\begin{aligned} &\exp\left[\frac{q(p-q)}{pr}\omega\right]\left[-\frac{q-p}{p}(1-\omega)^{\frac{q-2p}{p}} g(\omega) + (1-\omega)^{\frac{q-p}{p}} \frac{dg}{d\omega}\right] \\ = \ &\frac{q-p}{p} \exp\left[\frac{q(p-q)}{pr}\omega\right](1-\omega)^{\frac{q-p}{p}}\left[\frac{q}{r}g(\omega) + \frac{r}{1-\omega}\right]. \end{aligned} \tag{3.27}$$

Since

$$\frac{d\left[\exp\left[\frac{q(p-q)}{pr}\omega\right](1-\omega)^{\frac{q-p}{p}} g(\omega)\right]}{d\omega}$$

$$
\begin{aligned}
&= \exp\left[\frac{q(p-q)}{pr}\omega\right]\left[\frac{q(p-q)}{pr}(1-\omega)^{\frac{p-q}{p}} - \frac{q-p}{p}(1-\omega)^{\frac{q-2p}{p}}g(\omega) + (1-\omega)^{\frac{q-p}{p}}\frac{dg}{d\omega}\right] \\
&\approx \exp\left[\frac{q(p-q)}{pr}\omega\right]\left[-\frac{q-p}{p}(1-\omega)^{\frac{q-2p}{p}}g(\omega) + (1-\omega)^{\frac{q-p}{p}}\frac{dg}{d\omega}\right] \qquad (3.28)
\end{aligned}
$$

Eqn. (3.27) becomes

$$
\frac{d\left[\exp\left[\frac{q(p-q)}{pr}\omega\right](1-\omega)^{\frac{q-p}{p}}g(\omega)\right]}{d\omega} = \frac{q-p}{p}\exp\left[\frac{q(p-q)}{pr}\omega\right](1-\omega)^{\frac{q-p}{p}}\left[\frac{q}{r}g(\omega) + \frac{r}{1-\omega}\right] \qquad (3.29)
$$

and integrating gives

$$
\begin{aligned}
g(\omega) &= \frac{r(q-p)}{p}(1-\omega)^{\frac{p-q}{p}}\exp\left[\frac{q\omega}{pr}(q-p)\right] \\
&\quad \times \int_\omega^1 \left[\frac{q(1-x)g(x)}{r^2} + 1\right](1-x)^{\frac{q}{p}-2}\exp\left[\frac{qx}{pr}(p-q)\right]dx \\
&\approx \frac{r(q-p)}{p}(1-\omega)^{\frac{p-q}{p}}\exp\left[\frac{q\omega}{pr}(q-p)\right]\int_\omega^1 (1-x)^{\frac{q}{p}-2}\exp\left[\frac{q(p-q)}{pr}x\right]dx. \qquad (3.30)
\end{aligned}
$$

In the limit $\omega \to 1$ we have

$$
g(1) - g(\omega) \sim (q-p)(1-\omega) \qquad (3.31)
$$

so that

$$
s_k \sim k^{-2} \qquad (3.32)
$$

as $k \to \infty$. Therefore the distribution of the waiting lists behaves as $n_k \sim k^{-\beta}$ for large k, with $\beta = 2$ independent of p, q and r. Thus this scenario gives rise to a distribution of waiting lists that is asymptotically power law.

3.3.2 Case 2: $b_k = 1$ and $p = 0$

In this situation the patients enter at a constant rate, and new waiting lists are created at a constant rate, but no patients are ever treated. In this situation Eqn. (3.14) becomes

$$s_k = \frac{q}{r}[s_{k-1} - s_k] + r\delta_{k,0}. \tag{3.33}$$

From induction, it can be seen that

$$\begin{aligned} s_k &= \left(\frac{q}{q+r}\right)^k s_0 + \left(\frac{q}{q+r}\right)^{k-1} \frac{r^2}{q+r}\delta_{k,0} + \left(\frac{q}{q+r}\right)^{k-2} \frac{r^2}{q+r}\delta_{k,0} + \\ &\ldots + \frac{r^2}{q+r}\delta_{k,0} \end{aligned} \tag{3.34}$$

giving an exponential distribution of waiting lists

$$s_k = \frac{r^2}{q+r}\left[\frac{q}{q+r}\right]^k. \tag{3.35}$$

The number of patients waiting in the lists diverges as $q \to \infty$. Although at first glance this model may seem unrealistic, it represents a situation that occurs in many healthcare systems when some sort of crisis occurs and in order to deal with this, no elective referrals take place. For instance, this can occur when there is a local or national disaster or when there is a winter flu epidemic.

3.3.3 Case 3: $a_k = k$ and $b_k = 1$ when $p = q = 1$ and $r = 0$

Here the number of patients in the system is fixed on average, and does not grow as in the previous case. Thus a power law is not expected. The rate equation for this case is given by

$$\frac{dn_k(t)}{dt} = \frac{1}{M(t)}[(k+1)n_{k+1} - kn_k] + \frac{1}{N(t)}[n_{k-1} - n_k] \tag{3.36}$$

where $N(t)$ and $M(t)$ are given in Eqns. (3.9) and (3.10). In this case the $n_k(t)$, $N(t)$ and $M(t)$ are all independent of time. In the limit $t \to \infty$ there is a stationary solution

$$n_k = N(0)\frac{\alpha^k}{k!}\exp[-\alpha] \tag{3.37}$$

where $\alpha = \frac{M(0)}{N(0)}$. Thus in this case the length of the waiting lists obeys a Poisson distribution which is a feature of queueing systems with events occurring at fixed average rates as described in the previous section. This case has shown that once the system stops growing continuously and becomes effectively closed, the power laws disappear.

3.3.4 Case 4: $b_k = \frac{1}{k+1}$ and $r = 0$

This situation corresponds to the scenario where patients join a waiting list with a rate inversely proportional to its length, which is the first case considered that directly models the preferential behaviour of patients. Two variants of this model are considered. Firstly, a patient is chosen at random and re-

moved from the list, so that $a_k = k$. Secondly, a list is chosen at random and, if possible, a patient is removed from the list, so that $a_k = 1$ for $k > 0$ and $a_0 = 0$. A simulation of both these models has been performed with 1000 waiting lists, $p = q = 1$ and initially 10 patients on each list. The simulation results of both models show that the distribution of the length of the waiting lists is very different from power law, and is closer to a Gaussian distribution. The mean value of k for $a_k = 1$ remains fixed at 10 patients per list. For $a_k = k$, the mean grows linearly with time. Figure 3.1 shows the results for the average value of n_k against k for 100 samples for the case $a_k = k$.

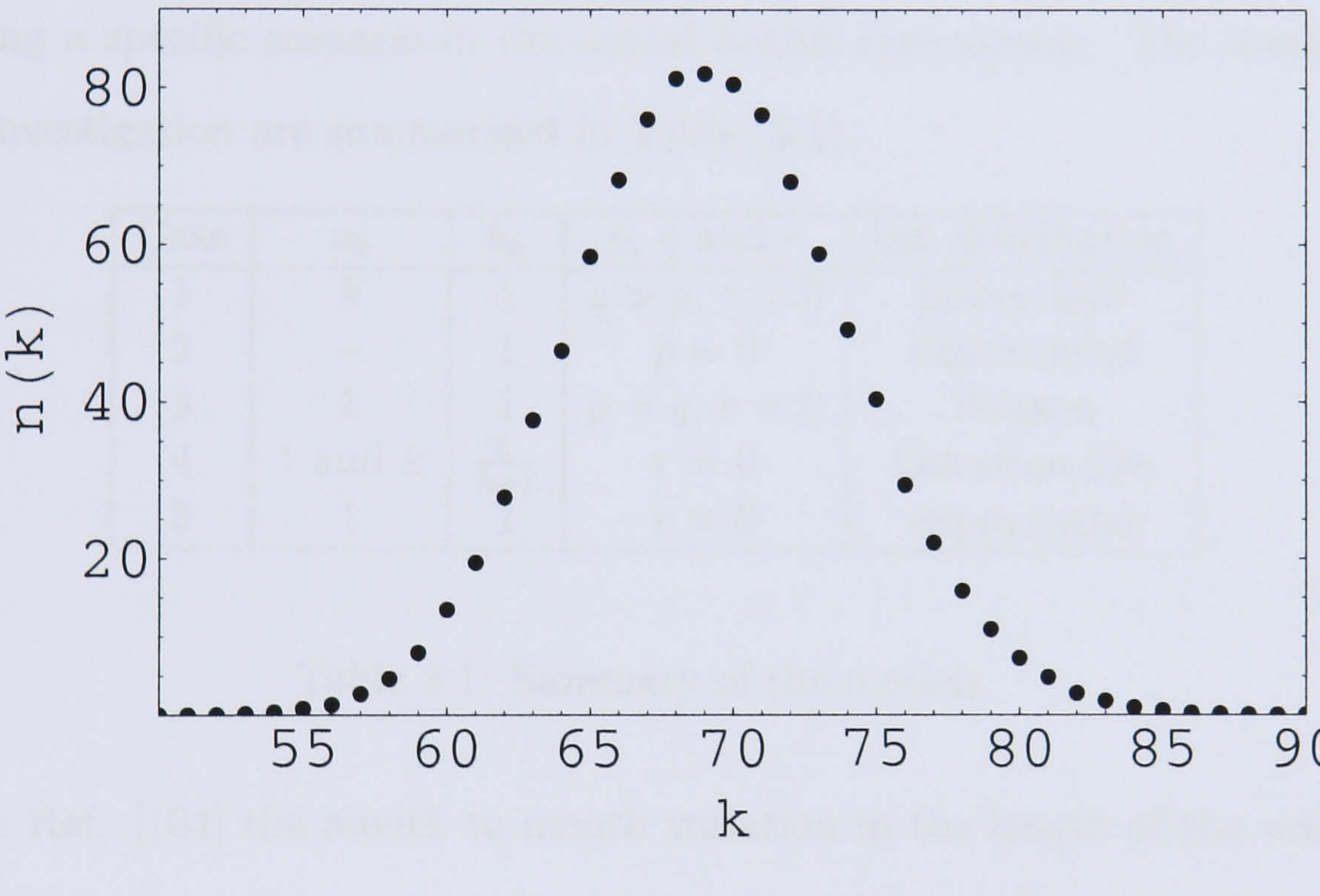

Figure 3.1: A graph of n_k against k for Case 4 when $a_k = k$.

3.3.5 Case 5: $a_k = b_k = 1$ for $k > 0$ with $b_0 = 1$ and $r = 0$

This case represents a biased random walk along a semi-infinite chain in which $n_k(t)$ is proportional to the probability of finding the walker on site k

at time t. Due to this reason, when $p > q$ the system evolves to a stationary distribution with

$$n_k(\infty) \sim \left(\frac{q}{p}\right)^k \tag{3.38}$$

which is exponential. For $p < q$ the number of patients in the system grows with t and the mean length of the lists is $(q-p)t$.

3.4 Conclusions and discussion

A set of models of hospital waiting lists have been introduced, each representing a specific scenario in the actual health care system. The results of the investigation are summarised in Table (3.1).

Case	a_k	b_k	p, q and r	list distribution
1	k	1	$q > p$, $r > 0$	power Law
2	–	1	$p = 0$	exponential
3	k	1	$p = q$, $r = 0$	Poisson
4	1 and k	$\frac{1}{k+1}$	$r = 0$	Gaussian-like
5	1	1	$r = 0$	exponential

Table 3.1: Summary of the models.

In Ref. [104] the month to month variation in the length of the waiting lists of four dermatology specialists was measured over a six year period. The distribution of the fractional change λ in the length of the waiting lists, $b(\lambda)$, was shown to be power law. In the previous section, the number of waiting lists of length k at time t was considered, and these quantities are relevant

via the formula

$$b(\lambda) = \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{\infty}\sum_{k'=1}^{\infty} n_k(t)n_{k'}(t+1)\delta\left(\lambda - \frac{k}{k'}\right). \tag{3.39}$$

When the measurement period is very much longer than the time between successive measurements, $T \gg 1$, and the distribution of waiting list lengths is only weakly dependent on time, we can write

$$b(\lambda) = \sum_{k=1}^{\infty}\sum_{k'=1}^{\infty} n_k n_{k'}\delta(\lambda - \frac{k}{k'}). \tag{3.40}$$

Therefore if we have a power law distribution of waiting lists, $n_k \sim k^{-\beta}$, then it follows that the distribution of the fractional change is also power law, $b(\lambda) \sim \lambda^{-\beta}$, with the same exponent.

The length of the longest waiting lists can be estimated by assuming that these power law distributions are generally valid for hospital waiting lists. In particular, if the number of lists of length k behaves like $n_k \sim k^{-\beta}$, then the number of lists m is related to the length of the maximum list, N_{max} by

$$N_{max} \sim m^{\frac{1}{\beta-1}}. \tag{3.41}$$

This formula can be used to estimate N_{max} and β from real data. There are 4349 surgical consultants in England and Wales, and adding the non-consultant surgeons, gives a total of, say, $m = 5000$ surgical waiting lists [112]. This suggests that the longest waiting list in England and Wales is 5000 patients if the exponent β is equal to 2 [104, 113], and 292 patients if $\beta = 2.5$ [105]. The longest reported waiting list was found to be 1800 patients

[114], which would imply that the exponent in the power law is $\beta = 2.14$.

However, a number of variants to the above models are possible. For instance, the possibility of patients hopping from list to list or the possibility of the removal of waiting lists with low average length can be introduced. The former would create a network of waiting lists, allowing for a more complicated system. Similarly, each list can be given a different rate to allow patients to be treated at different rates in different lists. These additions would make the models more realistic and without affecting the overall conclusions. Namely, that power laws can arise in some models of waiting lists, but when some type of preferential choice is introduced, there are no power laws. This observation seems to be in agreement without any basic understanding of when power laws arise in complex systems. A power law distribution normally indicates that extremal events are much more likely to occur than would be expected with an exponential or Gaussian distribution. Therefore it seems unlikely that a power law distribution of waiting lists will arise as a result of patients' reluctance to join long waiting lists. If power laws do exist in hospital waiting lists, then the the mechanisms creating them are probably indistinguishable from the random selection of lists by the patients. Since the models do not have any spatial structures, these power law distributions indicate that these systems of waiting lists are self-organised, a property that allows the lists to regulate themselves so as to resist any attempts at changing them. This explains why the measures taken to reduce the length of waiting lists are always ineffective in the long term.

Chapter 4

Modelling the distribution of salaries

4.1 Comparative models

In chapter two we have seen the evolutionary trait of herding in higher order organisms where individuals join up in groups to increase their chances of survival. Yet another interesting product of evolution is the act of comparison. Many higher order organisms have the tendency of comparing one individual's quantity and quality of resources with those of other individuals. A relatively poorer individual normally tries to gain more wealth to catch up with the others. Thus, just like the herd behaviour in chapter two, constructive comparison is beneficial to the survival of a species.

In the following subsections, various models of comparison are described, including sorting and mixing, auctions and card comparison in card games. These are followed by the introduction of three models of salary comparison

where employees in companies compare their salaries and the less well paid employees take actions in an attempt to improve their salaries.

4.1.1 Sorting and mixing, segregating and integrating

In Ref. [115], models of sorting and mixing or segregating and integrating are discussed. These models can represent a population which consists of people who are responding to each other. The response of one person can trigger off another's response which in turn can cause further responses. These individuals may be responding to specific characteristics of the population that include discrete variables, such as nationality, sex and race, or continuous variables such as age, income, height and IQ. In Ref. [115], each of the individuals is assigned a certain preference about the population characteristics and their response is specified. An individual's response can be taken as the departure, joining or rejoining of a population. Ref. [115] discusses some situations where this type of behaviour may occur. An example is to consider a nursing home where individuals are likely to leave if people are older than themselves on average. This results in the departure of younger people which increases the average age and this in turn encourages other younger people to leave. Eventually only the oldest individuals will remain in the nursing home. This is just one example of many. These models can also be related to colleges where students may join or leave due to the rank of the college; or even sports teams where the joining or departure will depend on the quality of the team e.g. number of games won in total.

Such models are important as they can mimic real life processes or even

artificial situations [115], and the resulting distributions of the latter can be useful for two reasons. Firstly, these models can be built on to represent a real life situation and secondly, they give a starting point to the type of analysis that is required, some of the phenomena to be anticipated and the important questions that need to be asked.

4.1.2 Distribution characteristic comparison in a model of auctions

Recently, a simple model of auctions was proposed by D'Hulst and Rodgers [116] which takes the concept of comparing a certain characteristic of a distribution and carrying out a specific response. In Ref. [116], an auction is considered as a competition between sellers who compete to attract buyers, where the lower price is most appealing. A model is introduced in which two players have an infinite set of numbers described by a probability distribution. At each time step, these two players bid against each other by drawing a number at random from their corresponding distributions. Both players then compare numbers and the player with the smaller number wins and does nothing, while the player with the larger number replaces his losing number with a number chosen at random from a uniform distribution, and this process is repeated at each time step. If both players have identical initial distributions, the system reaches a steady state and the probability distribution for each player evolves to a power law with an exponent dependent on the number of players.

4.1.3 Comparison in Card Games

An approach similar to the auction model is taken in Ref. [117] where an elementary two player card game is studied. In this model, the players compare their card at each time step and the holder of the smaller card wins and gets both cards. An exact solution is obtained for arbitrary initial conditions, and similar to the results found in Ref. [116], the system approaches a steady state where the card densities are proportional to each other. However, both of the players end up with a different number of cards overall and it is possible for one player to gain all cards and win.

4.2 Models of salaries: mean field version

Motivated by previous models of comparison, new models of salary distributions are introduced in this chapter. In the first model, two individuals (employees) who both work for the same organisation compare salaries. The higher paid individual does nothing and remains in the organisation but the lower paid individual leaves the organisation and is replaced by another whose salary is picked from a power law distribution. This seems realistic as an employee who is unhappy with his salary tends to leave and join a new organisation. In addition to that, the new individual's salary is chosen from a distribution which is power law as the personal income distribution is proven to obey a power law distribution [118, 119, 120]. The mean-field version of this model, where any individual can compare their salary to anyone else in the organisation, is solved analytically. This model is also studied numerically in a lattice form where any individual may only compare salaries

with their nearest neighbour. This is more likely as employees in an organisation tend to associate only with their close colleagues. Several variations of this model are introduced and they are again solved analytically and studied using numerical simulations.

In the following subsections several variants of the mean field version of the model are introduced. These are followed by a 1-d lattice version of Model A. Finally, conclusions are drawn and results are discussed in the final section of this chapter.

4.2.1 Model A

In this model, two individuals, who both work for the same organisation, compare salaries at each time step. The individual with the higher salary does nothing but carries on as normal, while the individual with the lower salary leaves and is replaced by another whose salary is chosen from a power law distribution, $(\gamma - 1)x^{-\gamma}$, where $\gamma > 1$. Thus, the number of individuals $A(x,t)$ with salary $x > 1$ at time t evolves like

$$\frac{dA(x,t)}{dt} = -A(x,t)\int_x^\infty A(y,t)dy + \frac{\gamma-1}{x^\gamma}\int_1^\infty A(y,t)\int_y^\infty A(z,t)dzdy. \quad (4.1)$$

The first term on the right hand side of Eqn. (4.1) corresponds to the destruction of numbers in $A(x,t)$ when one of the two competing individuals has the lower salary x, which results in him leaving. In this term, the product of the number of individuals with salary x and the total number of individuals with salaries greater than x corresponds to the probability of the loss of an individual with salary x. The second term on the right hand side cor-

responds to the creation of salary x after one of the individuals has been replaced. The $\int_1^\infty A(y,t)\int_y^\infty A(z,t)dzdy$ in this term is a constant, and the $\frac{\gamma-1}{x^\gamma}$ corresponds to the probability that the newly arriving agent has a salary x.

Introducing the cumulative probability distribution

$$F(x,t) = \int_x^\infty A(y,t)dy \tag{4.2}$$

the salary distribution $A(x,\infty)$ can be obtained when the system reaches a stationary state. This can be done by rewriting Eqn. (4.1) in terms of $F(x,\infty)$ to give

$$0 = \frac{dF(x)}{dx}F(x) - \frac{\gamma-1}{x^\gamma}\int_1^\infty \frac{dF(y)}{dy}F(y)dy. \tag{4.3}$$

Eqn. (4.3) can be reduced to

$$0 = \frac{dF(x)}{dx}F(x) + \frac{1}{2}\frac{\gamma-1}{x^\gamma} \tag{4.4}$$

and by integrating Eqn. (4.4) we have

$$F(x,\infty) = x^{\frac{1-\gamma}{2}}. \tag{4.5}$$

Thus, we obtain

$$A(x,\infty) = \frac{\gamma-1}{2}x^{\frac{-(1+\gamma)}{2}}. \tag{4.6}$$

Hence the resulting distribution approaches a power law. This power law distribution is obtained as the new individual, who joins the organisation, has

his salary picked from a power law distribution. The final salary distribution is thus dependent on the distribution from which the new individuals's salary is chosen; however the resulting distribution has a different exponent.

If the new individuals are chosen from an exponential distribution $\gamma \exp[-\gamma(x-1)]$ then by using the same steps as above, it is found that the resulting salary distribution is $\frac{\gamma}{2}\exp[\frac{-\gamma(x-1)}{2}]$.

4.2.2 Model B

This is a variant of the first model and at each time step, after both individuals have compared salaries, the individual with the lower salary either with probability p leaves and is replaced, or with probability $1-p$ has his salary matched to that of the other individual. Consequently, $A(x,t)$ with a salary $x > 1$ at time t evolves like

$$\frac{dA(x,t)}{dt} = -A(x,t)\int_x^\infty A(y,t)dy + p\frac{\gamma-1}{x^\gamma}\int_1^\infty A(y,t)\int_y^\infty A(z,t)dzdy + (1-p)A(x,t)\int_1^x A(y,t)dy. \quad (4.7)$$

The first term on the right hand side of Eqn. (4.7) corresponds to the destruction of numbers in $A(x,t)$ when one of the two individuals has the lower salary, which results in them leaving. The second term on the right hand side corresponds to the creation in new numbers in $A(x,t)$ after one of the individuals has been replaced with probability p. The third term on the right hand side represents the creation of new numbers in $A(x,t)$ when the salary of one of the individuals is modified with probability $1-p$.

Using the same method as that of the previous model,

$$F(x,\infty) = \frac{1-p}{2-p}\left[1-\sqrt{1+p\frac{(2-p)}{(p-1)^2}x^{1-\gamma}}\right] \tag{4.8}$$

(see appendix C)hence

$$A(x,\infty) = \frac{(\gamma-1)px^{-\gamma}}{2\sqrt{(p-1)^2+p(2-p)x^{1-\gamma}}}. \tag{4.9}$$

Eqn. (4.6) is obtained when $p = 1$.

4.2.3 Model C

This is another variation of Model A, however instead of having two individuals compare salaries, this model has α individuals compare salaries, where $\alpha > 1$. The individual with the lower salary leaves and is replaced by another individual whose salary is chosen from the same power law distribution as the previous models.

The rate equation for this variant of the first model is

$$\frac{dA(x,t)}{dt} = -A(x,t)\left(\int_x^\infty A(y,t)dy\right)^\alpha + \frac{\gamma-1}{x^\gamma}\int_1^\infty A(y,t)\left(\int_y^\infty A(z,t)dz\right)^\alpha dy. \tag{4.10}$$

Using the same method as used in the previous models,

$$F(x,\infty) = x^{\frac{1-\gamma}{\alpha+1}} \tag{4.11}$$

(see appendix C)hence

$$A(x,\infty) = \frac{\gamma - 1}{\alpha + 1} x^{\frac{-(\alpha+\gamma)}{\alpha+1}}. \tag{4.12}$$

The resulting distribution is power law and has an exponent which is on the number of individuals comparing salaries, and this exponent is $\alpha + 1$. Hence, for all initial salary distributions, the resulting distribution approaches a power law with an exponent which is dependent on the number of individuals who are comparing salaries. This is a generalised version of the first model.

4.3 Differences between mean field and lattice models

A number of simulations have been performed on Model A in which two individuals compare salaries. These were compared with a 1-d version of Model A in which at each time step, an individual is chosen at random, and that individual randomly selects his or her neighbours. For this pair of individuals, as in the mean field model, the individual with the highest salary does nothing whereas the one with the lower salary leaves and is replaced with a new individual whose new salary is selected at random.

Figure 4.1 shows the salary distribution for Model A in which the new individuals have a salary selected from a power law with exponent $\gamma = 2$. The simulation was carried out over 10^8 time steps. From Eqn.(4.6), it can be seen that the resultant distribution should have exponent $\frac{(\gamma+1)}{2} = \frac{3}{2}$, and

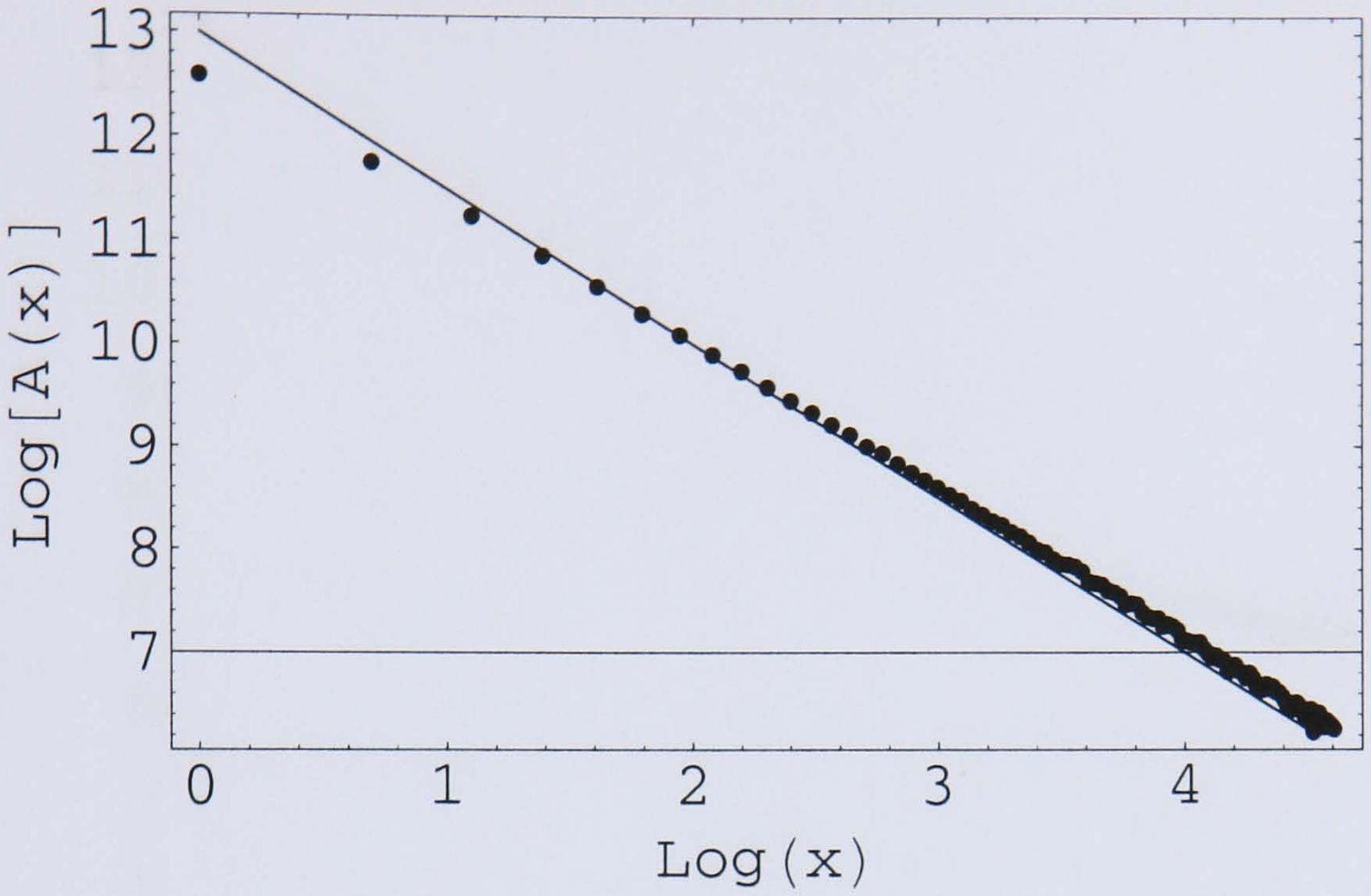

Figure 4.1: Salary distribution for Model A in which the new individuals have a salary selected from a power-law distribution with exponent $\gamma = 2$.

this is indeed the case, and the line with slope 1.5 fits the data very well. In Figure 4.2 the same distribution for the 1-d system is shown, and it is a power law with an exponent close to that of the mean field version.

A comparison between the the mean field and the 1-d model can be made by looking at the mean absolute difference between the salaries of traders, denoted by $< z >$. To do this, we first consider $F(z)$, the distribution of the absolute difference between the salaries of two neighbouring individuals. $F(x)$ is given by

$$F(z) = \int_{-\infty}^{\infty} P(x)P(x-z)dx + \int_{-\infty}^{\infty} P(x)P(x+z)dx. \qquad (4.13)$$

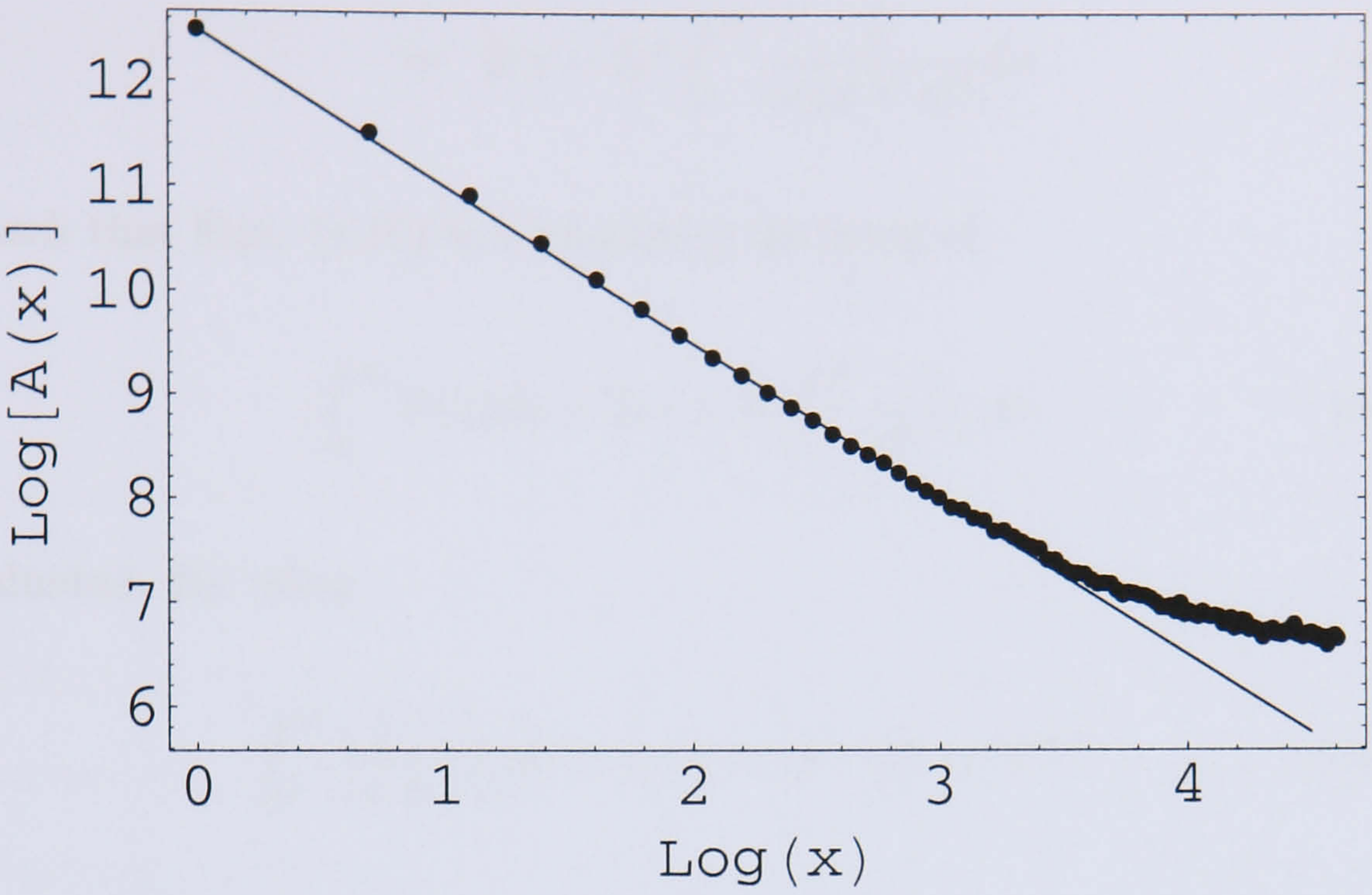

Figure 4.2: Salary distribution for Model A on a 1-d lattice in which the new individuals have a salary selected from a power-law distribution with exponent $\gamma = 2$.

Due to symmetry, Eqn. (4.13) can be written as

$$F(z) = 2\int_{-\infty}^{\infty} P(x)P(z+x)dx. \tag{4.14}$$

This can be used to compute the average value of the absolute difference $< z >$ between the salaries of the two neigbouring individuals. For the case with the power law distribution,

$$P(x) = \frac{\gamma - 1}{x^{\gamma}} \tag{4.15}$$

so that

$$F(z) = 2\int_{1}^{\infty} \left(\frac{\gamma - 1}{x^{\gamma}}\right)\left(\frac{\gamma - 1}{(x+z)^{\gamma}}\right) dx$$

$$= 2(\gamma-1)^2 \int_1^\infty \frac{1}{x^\gamma (x+z)^\gamma} dx. \tag{4.16}$$

To check that Eqn. (4.16) is normalised, the integral

$$\int_0^\infty F(z)dz = 2(\gamma-1)\int_1^\infty \frac{1}{x^{2\gamma-1}}dx \tag{4.17}$$

is evaluated. By using

$$\int_0^\infty \frac{\gamma-1}{(x+z)^\gamma}dz = [-(x+z)^{1-\gamma}]_0^\infty = x^{1-\gamma}, \tag{4.18}$$

Eqn. (4.17) becomes

$$\begin{aligned}
\int_0^\infty F(z)dz &= \int_0^\infty 2(\gamma-1)^2 \int_1^\infty \frac{1}{x^\gamma (x+z)^\gamma} dxdz \\
&= 2(\gamma-1)\int_0^\infty \int_1^\infty \frac{1}{x^\gamma}\frac{\gamma-1}{(x+z)^\gamma} dxdz \\
&= 2(\gamma-1)\int_1^\infty x^{1-2\gamma} dx \\
&= 1,
\end{aligned} \tag{4.19}$$

thereby showing that Eqn. (4.16) is normalised. $< z >$ is given by

$$\begin{aligned}
< z > &= \int_0^\infty zF(z)dz \\
&= 2(\gamma-1)^2 \int_0^\infty \int_1^\infty \frac{z}{[x(x+z)]^\gamma} dxdz.
\end{aligned} \tag{4.20}$$

By using

$$\int_0^\infty \frac{z(\gamma-1)}{(x+z)^\gamma}dz = \frac{x^{2-\gamma}}{\gamma-2} \tag{4.21}$$

Eqn. (4.20) is solved to give

$$< z >= \frac{2(\gamma-1)}{(\gamma-2)(2\gamma-3)} \tag{4.22}$$

where $\gamma > \frac{3}{2}$. From the numerical simulation for $\gamma = 11$, the $< z >$ is found to be 0.259. However, substituting $\gamma = 11$ into Eqn. (4.22) gives $< z >= 0.526$. This is due to the earlier assumption that the salary distribution of the individuals is approximately the same as the distribution from which new salaries are chosen $P(x) = \frac{\gamma-1}{x^\gamma}$. However, as shown in the previous section, the resultant salary distribution is $A(x,\infty) = \frac{\gamma-1}{2} x^{\frac{-(1+\gamma)}{2}}$. Thus, the γ in Eqns. (4.16) to (4.22) should be replaced by $\frac{\gamma+1}{2} = 6$, giving $< z >= 0.278$. The difference between the theoretical and simulation results is due to the range of integration assumed in the theory; since the steady state distribution is power law, all moments are divergent, so that there must exist an upper and a lower cutoff, denoted by x_1 and x_2 respectively, such that the power law solution is only applicable for the range $x_1 < x < x_2$. Thus, $< z >$ is rewritten as

$$\begin{aligned} < z > &= \frac{2(\gamma-1)}{\gamma-2} \int_{x_1}^{x_2} x^{2-2\gamma} dx \\ &= \frac{2(\gamma-1)}{\gamma-2} \left[\frac{x_2^{3-2\gamma}}{3-2\gamma} - \frac{x_1^{3-2\gamma}}{3-2\gamma} \right] \end{aligned} \tag{4.23}$$

From Figure 4.2 it can be seen that $x_2 \sim 20$, so that the first term of Eqn. (4.23) is negligible, and the resultant value of $< z >$ is only dependent on the second term. Although it is difficult to estimate the value of x_1 from the same figure, one would expect the 'true' value of x_1 to be close to unity and

slightly greater than it. So, an estimated value $x = 1.010$ can be used, giving a theoretical $< z >= 0.297$ which is close to the simulation value 0.259. Note that the closer the chosen value of x_1 to unity, the closer the theoretical value to the simulation result.

For the exponential distribution,

$$\begin{aligned} F(z) &= \int_{z+1}^{\infty} P(x)P(x-z)dx + \int_{1}^{\infty} P(x)P(x+z)dx \\ &= 2\int_{1}^{\infty} P(x)P(x+z)dx \end{aligned} \tag{4.24}$$

and

$$P(x) = \frac{\gamma}{2}\exp[-\frac{\gamma(x-1)}{2}]. \tag{4.25}$$

This gives

$$\begin{aligned} F(z) &= \frac{\gamma^2}{2}\exp^{-\frac{\gamma z}{2}}\int_{1}^{\infty}\exp^{-\gamma(x-1)}\,dx \\ &= \frac{\gamma^2}{2}\exp^{-\frac{\gamma z}{2}}\,\frac{1}{\gamma} \\ &= \frac{\gamma}{2}\exp^{-\frac{\gamma z}{2}}\,. \end{aligned} \tag{4.26}$$

Thus,

$$< z >= \frac{\gamma}{2}\int_{0}^{\infty} z\exp^{-\frac{\gamma z}{2}}\,dz = \frac{2}{\gamma}. \tag{4.27}$$

The numerical simulation for $\gamma = \frac{3}{2}$ yields $< z >= 1.251 \pm 0.149$. From Eqn. (4.27), we have $< z >= \frac{4}{3} = 1.333$. Thus, the theoretical prediction agrees with the simulation result.

For the uniform distribution,

$$P(x) = \frac{1}{2\sqrt{1-x}} \tag{4.28}$$

and

$$F(z) = \frac{1}{4}\int_0^1 \frac{1}{(1-x)^{\frac{1}{2}}(1-(x-z))^{\frac{1}{2}}}dx + \frac{1}{4}\int_0^L \frac{1}{(1-x)^{\frac{1}{2}}(1-(x+z))^{\frac{1}{2}}}dx, \tag{4.29}$$

so that

$$\begin{aligned} < z > &= \int_0^1 zF(z)dz \\ &= \int_0^1 z\left(\frac{1}{4}\int_0^1 \frac{1}{(1-x)^{\frac{1}{2}}(1-(x-z))^{\frac{1}{2}}}dx\right)dz \\ &\quad + \int_0^L z\left(\frac{1}{4}\int_0^L \frac{1}{(1-x)^{\frac{1}{2}}(1-(x+z))^{\frac{1}{2}}}dx\right)dz \end{aligned} \tag{4.30}$$

where L is the upper limit for which Eqns. (4.29, 4.30) are valid.

To solve Eqn. (4.30), the Composite Simpson's rule [121, 122]

$$\int_a^b f(x)dx = \frac{b-a}{3}\left[f(a) + 2\sum_{i=1}^{m-1} f(a+2hi) + 4\sum_{i=1}^{m} f(a+2h(i-1)) + f(b)\right] \tag{4.31}$$

is applied, where a and b are the limits of the integral chosen, $h = \frac{b-a}{2m}$, m is the number of subintervals used, i is an integer and $f(x)$ is the function to be integrated. The software *Mathematica* was used to carry out the iterations, and details of these routines are included in the appendix of this thesis. It was apparent from iterations that a suitable upper limit is $L < 0.5$. The value of $< z >$ was obtained from iterations using 10000 subintervals and

the results show that $< z >= 0.303$. The numerical simulation of the model yield $< z >= 0.297 \pm 0.025$. Thus, the theoretical value falls within the error range of the simulation value.

Table (4.1) summarizes the differences between the average values of z obtained analytically and a simulation of both the mean-field and 1-d systems.

	lattice (simulation)	mean field (simulation)	mean field (theory)
power law	0.566 ± 0.043	0.259 ± 0.030	0.297
exponential	2.609 ± 0.181	1.251 ± 0.149	1.333
uniform	0.433 ± 0.021	0.297 ± 0.025	0.303

Table 4.1: $< z >$ for various distributions.

4.4 Conclusions and discussion

A mean-field model of salary distributions has been introduced, along with several variants of the model. Numerical simulations of the Model A in lattice form have also been studied.

Solutions for the mean field models at steady state have been found. Model A is the simplest model where any two individuals, both working for the same organisation, compare salaries. The individual with the lower pay leaves and is replaced by another whose salary is chosen from a power law distribution $(\gamma - 1)x^{-\gamma}$. Solving this model analytically reveals that the resulting salary distribution of the organisation is power law with a different exponent.

The resulting distributions of all other mean field variants of this model are all power law, apart from Model B, which is power law only when $p = 1$. This implies that all resulting salary distributions are dependent on the distribution from which the new individual's salary is picked. Nevertheless, the exponent of the resulting distribution is strongly dependent on the number of individuals comparing salaries. This is in agreement with the results from Model C, where the resulting distribution has an exponent $-(\alpha+\gamma)/(\alpha+1)$, where α is the number of individuals which another is comparing salaries with. Thus, the exponent of the resulting distribution decreases because of the α dependence and as α increases, it causes a further decrease in the size of the exponent.

The fact that power laws occur in Models A and C indicate that these models are at SOC state, and this can be understood in terms of threshold and metastability described earlier in chapter one. In order for SOC to occur, there must be a threshold in the system, and in this model, the salary of the richer individual serves as a threshold for the poorer individual. However, when $p = 1$ in Model B corresponding to no action taken by the poorer individual, no criticality is observed, and this is due to the absence of actions taken when the threshold is reached.

This work shows that if a total population has a power law salary distribution then this mechanism, which is basically a comparison and exchange process, produces a power law in a sub-population. Consequently this mechanism is one, possibly one of many, which might explain the ubiquity of power laws in social and economic systems.

Since these salary models lack any spatial structure, the occurrence of

power laws in these models indicates the reluctance of the systems to change even when subjected to external driving forces. This is probably a sign of the will of nature to preserve the trait of comparison, which should not come as a surprise, given the benefits of comparison to survival mentioned in the earlier parts of this chapter.

Chapter 5

Summary and outlook

Motivated by recent advances in complexity research, three models that mimic real systems in the financial market, the health care system and salary comparison in companies have been proposed in this thesis. The aim of these investigations is to further our understanding of complexity so commonly found in real systems.

Recent advances in computer technologies have greatly facilitated the work of the research community. As the performance of the microprocessor becomes better, many scientific theories of complicated, many-body systems can now be tested using numerical simulations. In many cases, researchers can even use numerical simulations as a substitute for lengthy calculations. This ability to study a system without having to describe it completely using mathematical expressions is even more important in the case of complex systems where the upshots of a simple algorithm, such as the algorithm of the sandpile model [45], are extremely complicated.

Numerical simulations play a major role in the research done in this thesis.

Each study has been carried out using an algorithm that captures the essence of the real system. Statistical mechanics has been used to solve the rate equations and to compute various quantities. Theoretical results have been shown to compare well with numerical simulations.

In chapter two, a model of herding is proposed which, besides mimicking the coherent group reactions in social communities, also encompasses the elements of growth and addition. These two mechanisms allow for the generation of new agents and the grouping of existing agents in the system. In this model, at each time step either [1] with probability p the system grows through the introduction of a new agent or [2] with probability $1-p$ a free agent already in the system randomly joins a group of size k with rate A_k. Three separate cases have been investigated. When the growth of the system is sufficiently fast corresponding to $p > 1/2$, the group size distribution is found to follow a power law with a parameter dependent exponent. When the number of groups stays constant corresponding to $p = 1/2$, the time taken by the system to run out of free agents scales as the square of the initial number of free agents in the system. When the rate of growth is slow enough corresponding to $p < 1/2$, the system runs out of free agents in a finite amount of time, and this time has been shown to be proportional to the initial number of free agents in the system.

In chapter three, a model of hospital waiting lists is proposed by using ideas from Queueing Theory. Patients enter the system and join a list based on the length of the latter. At the same time, patients get served and leave the system. In this model, a_k is the service rate in a list of length k and b_k corresponds to the tendency of a patient to join a list of length k. The

parameters p, q and r are used to adjust the overall relative rates of patient departure, arrival and list creation. Five different cases were considered in this model. For the case when $a_k = k$, $b_k = 1$, $q > p$ and $r = 0$, the list distribution has been found to be a power law. For the case when $b_k = 1$ and $p = 0$, the list distribution is exponential. For the case when $a_k = k$, $b_k = 1$, $p = q$ and $r = 0$, the list distribution has been found to be a Poisson distribution. For the fourth case, the configurations $b_k = 1/(k+1)$ and $r = 0$ are used. Two variants of this case are considered. It has been found that for both $a_k = 1$ and $a_k = k$, the list distribution is Gaussian-like. The mean number of patients for the former variant remains stationary over all time, but that of the latter grows linearly with time. For the final case, $a_k = 1$, $b_k = 1$ and $r = 0$, and the list distribution has been found to be exponential. This model of hospital waiting lists illustrates how the power law distributions found in empirical studies might arise, but indicates that the preferential attachment of the patients or their physicians is unlikely to be responsible for the occurrence of the power law distribution.

In chapter four, a model of salary comparison is introduced. An individual with higher salary does nothing but the lower paid individual leaves the organisation and is replaced by another, whose salary is picked from a power law distribution. The solution for the mean field version of the model has been found, which is a power law with an exponent different from that of the distribution from which the salary is picked. Two variants of the model were considered. In the first variant, two individuals compare salaries at each time step and the individual with the lower salary either with probability p leaves and is replaced, or with probability $1 - p$ has his salary matched to

that of the other individual. In the second variant, α individuals compare salaries instead of fixing at two, where $\alpha > 1$. The individual with the lower salary leaves and is replaced by another individual whose salary is chosen from the same power law distribution as the previous variants. The resulting salary distributions of the mean field version of the second variant is power law, but the first variant is power law only when $p = 1$. The implication is that all resulting salary distributions are dependent on the distribution from which the new individual's salary is picked. However, the exponent of the resulting salary distribution is strongly dependent on the number of individuals comparing salaries. This agrees well with the results from the second variant, where the resulting distribution has an exponent $-(\alpha + \gamma)/(\alpha + 1)$. Therefore, the exponent of the resulting distribution decreases because of the α dependence and as the latter increases, it causes a further decrease in the size of the exponent. Both mean field and 1-d versions of the model have been studied. The two versions are compared by looking at the mean absolute difference between salaries in each version.

The analysis of the three models proposed in this thesis has yielded some insights on how certain systems might have worked in the real world. For instance, the group size distribution of the herding model in chapter two has been shown to depend on the growth rate of the system, and if the growth is fast, the group size distribution will conform to a power law. If this is applied to the real world, a financial market that constantly takes in a large number of new traders will be expected to end up with most traders not consulting with anyone before making the decision on whether to buy or sell his market shares. Even for those who do consult with each other, they mostly do so

in small groups, making knowledge sharing or group loyalty minimal. An interesting point to note is that by having most traders consulting in small groups, their volumes of transactions are small, and so the decisions of most groups do not have a strong impact on the price of the shares. By having few groups with strong impacts, the system reduces the chance of sudden big changes in the stock price, and this feature of resistance to changes is a hall-mark of SOC.

The analysis of the proposed models of hospital waiting list has also provided an insight on the healthcare system. In particular, it has shown the mechanisms required to keep the length of waiting lists at the minimum. From the results in Table (3.1), it seems that the most desirable list distribution is the power law corresponding to Case 1. Thus, one way to to keep the waiting lists in a hospital short is probably to make sure that a significant portion of existing patients are treated before admitting more patients into the hospital. Patients on longer lists are treated first, and new patients join lists randomly.

In the model of salary comparison, it had been shown that when lower paid patients demand their salaries to match those of their colleagues, most of the employees in the entire population end up with very low salaries. The number of resultant lowly-paid employees in this case is much higher than those of the cases when the lower paid employees leave the company. Thus, to ensure a higher-paid population on average, it is probably a good idea for the employees to leave the company rather than demanding a matching salary.

Power laws are observed in all of the three models proposed in this the-

sis, a further implication that fractals are ubiquitous in nature and society. Future work on complex systems might suitably involve the identification of more self-similar characteristics in the form of power law distributions. The search for other properties of complex systems should also continue. However, the ultimate goal is probably to develop a mathematical framework that establishes a formal definition of complex systems.

All numerical simulation codes given in the following appendixes are in Fortran 77. The algorithms in these programs can be adapted to be used in any other programming languages.

Appendix A

Source Code for Chapter 2

The code given in this appendix is for the herding model in chapter 2. This program can be used to simulate case 3 of the herding model. The time needed for the simulation is approximately 4 hours.

```
integer N,A(10000),Nsteps,sum,k,kk,sumN
integer x
real y,z,sum1,p,jd,xd
c The above are the declarations used in this program. The A(10000) is a
```

matrix that records the number of groups of a certain size. In this program, a matrix size of 10000 is used. However, the reader is free to choose any size he or she pleases. Nsteps corresponds to the number of time steps used to simulate the model. sum records the size of the system at any time step. sumN records the total size of the system at the end a particular time step. sum1 is used to determine which group size a free agent chooses to attach to. p is the probability p used in the thesis. jd and xd are the values of $\frac{\tau}{N}$ and p respectively which will be printed onto an output file.

```
do 110 x=1,499
N=10000
c N is the total size of the system.
mran=1001
c mran is the seed for the random number generator.
p=0.001*x
c All values of A(i) are now allocated a zero value.
do 10 i=1,10000
A(i)=0
10 continue
    c And the value of A(1) is given a value of N so that all agents are free
initially.
    A(1)=N
    sum=N
    Nsteps=10000000
    sumN=0
    c The simulation is now started.
    do 20 j=1,Nsteps
    z=ran(mran)
    If(z.lt.p)then
    A(1)=A(1)+1
    sum=sum+1
    c The above statements correspond to the growth feature of the herding
model.
    else
```

c If a growth feature fails to occur, the addition feature will now be carried out.

```
A(1)=A(1)-1
y=ran(mran)
sum1=0
do 40 k=1,sum
sum1=sum1+A(k)/real(sum-1)
if(sum1.gt.y)then
kk=k
goto 50
endif
40 continue
50 A(kk)=A(kk)-1
A(kk+1)=A(kk+1)+1
if(A(1).eq.0)then
goto 70
endif
endif
do 7 i=1,10000
sumN=sumN+A(i)
7 continue
if(sumN.lt.0)then
goto 70
endif
20 continue
```

```
70 continue
jd=real(j)/real(10000)
xd=0.001*x
c Values of xd and jd are now printed onto an output file.
write(2,*)xd,jd
110 continue
stop
end
```

Appendix B

Source Code for Chapter 3

This is a program for the case 4 simulation mentioned in chapter 3. The time need for the simulation is approximately one hour and a half.

```
      integer A(1000),B(100),t,A1,r3,r4,sum2,B1,s3,s4,i,j,s5
      real sum1,A2,r1,r2,r6,s1,s2,F
      integer consa,consb,k,r5,consc,nsamples,g,C(100)
      real d1,BB,A3,D(100)
      c A(i) is the length of list i. B(i) is the number of lists of
      c length i.
      c C(i) is used to calculate the total number of lists with length i
      c averaged over the 100 samples used in this program.
      c D(i) is the value of the total number of lists with length i
      c averaged over the 100 samples used in this program.
      mran=1001
      nsamples=100
      Do 13 i=1,100
```

```
C(i)=0

13 continue

Do 10 g=1,nsamples

c All lists start off with 10 patients.

Do 1 i=1,1000

A(i)=10

1 continue

Do 14 i=1,100

B(i)=0

14 continue

B(10)=1000

c The list joining and service mechanisms are now started.

Do 2 t=1,10000000

sum1=0.000

Do 6 k=1,99

A1=k+1

A2=real(1)/real(A1)

A3=B(k)*A2

sum1=sum1+A2

6 continue

r1=ran(mran)

r2=r1*1000

r3=int(r2)

r4=A(r3)

r5=r4+1
```

```
r6=real(1)/real(r5)

BB=real(r6)/real(sum1)

d1=ran(mran)

If(d3.lt.BB)then

consa=B(r4)

If(consa.gt.0)then

A(r3)=A(r3)+1

B(r4)=B(r4)-1

B(r5)=B(r5)+1

endif

endif

sum2=0

Do 7 k=1,100

B1=B(k)*k

sum2=sum2+B1

7 continue

s1=ran(mran)

s2=s1*1000

s3=int(s2)

s4=A(s3)

s5=s4-1

F=real(s3)/real(sum2)

If(s1.lt.F)then

consb=A(s3)

consc=B(s4)
```

```
If(consb.gt.0)then
If(consc.gt.0)then
A(s3)=A(s3)-1
B(s4)=B(s4)-1
B(s5)=B(s5)+1
endif
endif
endif
2 continue
Do 8 i=1,100
C(i)=C(i)+B(i)
8 continue
10 continue
Do 11 i=1,100
D(i)=real(C(i))/real(nsamples)
11 continue
Do 5 j=1,100
write(1,*)D(j),j
5 continue
stop
end
```

Appendix C

Source Code for Chapter 4, *Mathematica* routines and derivations of equations

These are the programs for the Model A simulations mentioned in chapter 4. The time needed for the simulation is approximately 4 minutes.

```
      c This is the program for the mean field version of Model A with power
     law distribution. Note that a similar code can be used for other distributions.
          integer t,num(100)
          real trader,money1,money2,money(1000000)
          integer selectedtrader,passivetrader,uu
          real st,pt,ms,mp,trader1
          real dif,sum,average
          real mn,mnn,mnnn
      c The money(1000000) is the matrix for the amount of money for each
```

person.

```
mran=1085

a=0.000000001

Do 1 i=1,1000000

money1=ran(mran)

money2=(money1+aaaa)**(-1)

money(i)=money2

1 continue

cc Now start the simulation

Do 2 t=1,100000000

cc the traders are first selected

trader=ran(mran)

selectedtrader=int(trader*1000000)

trader1=ran(mran)

passivetrader=int(trader1*1000000)

st=selectedtrader+1

pt=passivetrader+1

ms=money(st)

mp=money(pt)

c Now perform the salary comparison

If(ms.lt.mp)then

money1=ran(mran)

money2=(money1+a)**(-1)

money(st)=money2

endif
```

```
If(ms.gt.mp)then
money1=ran(mran)
money2=(money1+a)**(-1)
money(pt)=money2
endif
2 continue
sum=0.0
c Now measure the mean average difference between the salaries of traders.
```

Note that the same code is applicable to all other distributions.

```
Do 4 i=1,999999
uu=i+1
dif=money(i)-money(uu)
If(dif.lt.0)then
dif=-dif
endif
sum=sum+dif
4 continue
average=real(sum)/real(1000000)
print*,average
Do 811 i=1,100
num(i)=0
811 continue
Do 81 i=1,100
Do 82 j=1,1000000
mn=money(j)
```

```
mnn=i*1

mnnn=(i+1)*1

If(mn.lt.mnnn.and.mn.gt.mnn)then

num(i)=num(i)+1

endif

82 continue

write(889,*)mnn,num(i)

81 continue

stop

end
```

This is the program for the lattice version of Model A with power law distribution.

```
real trader,money1,money(1000000)

real money2,a

integer selectedtrader,passivetrader,uu

real st,pt,ms,mp,trader1

real dif,sum,average

real mn,mnn,mnnn

integer t,num(100)

c Similar to the previous program, except that this is for the power law
```

distribution of the 1-d version. all other distributions can be adjusted accordingly. The time needed for the simulation is approximately 4 minutes.

```
mran=1085

a=0.0000001

Do i=1,1000000
```

```
money1=ran(mran)

money2=(money1+a)**(-1)

money(i)=money2

1 continue

Do 2 t=1,100000000

trader=ran(mran)

selectedtrader=int(trader*1000000)

trader1=ran(mran)

If(trader1.gt.0.5)then

passivetrader=selectedtrader+1

endif

If(trader1.lt.0.5)then

passivetrader=selectedtrader-1

endif

st=selectedtrader+1

pt=passivetrader+1

ms=money(st)

mp=money(pt)

If(ms.lt.mp)then

money1=ran(mran)

money2=(money1+aaaa)**(-1)

endif

If(ms.gt.mp)then

money1=ran(mran)

money2=(money1+aaaa)**(-1)
```

```
money(pt)=money2

endif

2 continue

sum=0.0

Do 4 i=1,999999

uu=i+1

dif=money(i)-money(uu)

If(dif.lt.0.0)then

dif=-dif

endif

write(888,*)dif

sum=sum+dif

4 continue

average=real(sum)/real(1000000)

print*,average

Do 811 i=1,100

num(i)=0

811 continue

Do 81 i=1,100

Do 82 j=1,1000000

mn=money(j)

mnn=i*1.000

mnnn=(i+1)*1.000

If(mn.lt.mnnn.and.mn.gt.mnn)then

num(i)=num(i)+1
```

```
endif

82 continue

write(889,*)num(i)

81 continue

stop

end
```

Mathematica routines

These are the routines used in the iterations for the uniform distribution in Chapter 4. The aim of these routines is to obtain a theoretical value of $< z >$.

The following corresponds to the first term of the LHS of Eqn. (4.29).

```
a = 0; b = 0.999999; m = 10000; h = (b - a) / (2 m)
f[x_] := (0.25) * (((1 - x) ^ (-0.5)) * ((1 - x + z) ^ (-0.5)));
SumOdd = 0; For[k = 1, k <= m, k++, SumOdd = SumOdd + f[a + h (2 k - 1)];];
SumEven = 0; For[k = 1, k <= m - 1, k++, SumEven = SumEven + f[a + h * 2 * k];];
NIntegrate[z * (h * (f[a] + 4 * SumOdd + 2 * SumEven + f[b]) / 3), {z, 0, 0.999999}]
```

The 'For' loops above sum up the odd and even terms in Eqn. (4.31). The following corresponds to the second term of the LHS of Eqn. (4.29).

```
a = 0; b = 0.499999; m = 10000; h = (b - a) / (2 m)
f[x_] := (0.25) * (((1 - x) ^ (-0.5)) * ((1 - x - z) ^ (-0.5)));
SumOdd = 0; For[k = 1, k <= m, k++, SumOdd = SumOdd + f[a + h (2 k - 1)];];
SumEven = 0; For[k = 1, k <= m - 1, k++, SumEven = SumEven + f[a + h * 2 * k];];
NIntegrate[z * (h * (f[a] + 4 * SumOdd + 2 * SumEven + f[b]) / 3), {z, 0, 0.499999}]
```

Derivations of equations

At steady state, Eqn. (4.7) becomes

$$\begin{aligned} 0 &= -A(x,t)\int_x^\infty A(y,t)dy + p\frac{\gamma-1}{x^\gamma}\int_1^\infty A(y,t)\int_y^\infty A(z,t)dzdy \\ &\quad +(1-p)A(x,t)\int_1^x A(y,t)dy. \end{aligned} \tag{C.1}$$

Substituting Eqn. (4.2) into Eqn. (C.1) gives

$$\frac{2-p}{2}(F(x))^2 - (1-p)F(x) - \frac{p}{2}x^{1-\gamma} = 0. \tag{C.2}$$

Solving Eqn. (C.2) leads to Eqn. (4.8).

At steady state, Eqn. (4.10) becomes

$$0 = -A(x,t)\left(\int_x^\infty A(y,t)dy\right)^\alpha + \frac{\gamma-1}{x^\gamma}\int_1^\infty A(y,t)\left(\int_y^\infty A(z,t)dz\right)^\alpha dy. \tag{C.3}$$

Substituting Eqn. (4.2) into Eqn. (C.3) gives

$$\frac{dF(x)}{dx}(F(x))^\alpha = \frac{1-\gamma}{x^\gamma(\alpha+1)}. \tag{C.4}$$

Integrating Eqn. (C.4) leads to Eqn. (4.11).

Bibliography

[1] H. Young, R. Freedman, University Physics with Modern Physics. 11^{th} ed. Addison-Wesley, Reading, (2003)

[2] W. Rees, Physics by Example: 200 Problems and Solutions, Cambridge University Press, Cambridge, (1994)

[3] I. Grant, W. Phillips, Elements of Physics, Oxford University Press, Oxford, (2001)

[4] E. N. Lorentz, Journ. Atmosph. Sciences **20**, 130, (1963)

[5] P. Smith, Explaining Chaos, Cambridge University Press, Cambridge, (1998)

[6] J. Gleick, Chaos: Making a New Science, Penguin, New York, (1988)

[7] B. Hao, Chaos, World Scientific, Singapore, (1984)

[8] L. A. Adamic, R. M. Lukose, A. R. Puniyani, B. A. Huberman, Phys. Rev. E **64**, 046135-1, (2001)

[9] M. Faloutsos, P. Faloutsos, C. Faloutsos, Proc. ACM SIGCOMM, Comput. Commun. Rev. **29**, 251, (1999)

[10] S. Abe, N. Suzuki, Phyica. A **350**, 588, (2005)

[11] Y. Bar-Yam, Dynamics of Complex Systems, Westview Press, Boulder, (2003)

[12] M. Waldrop, Complexity: The Emerging Science at the Edge of Order and Chaos, Simon & Schuster, New York, (1992)

[13] R. Albert, A-L. Barabási, Rev. Mod. Phys. **74**, 47, (2002)

[14] http://134.184.131.111/COMPLEXI.html.

[15] J. M. Carlson, J. S. Langer, B. E. Shaw, C. Tang, Phys. Rev. A **44**, 884, (1991)

[16] P. Bak, How Nature Works: the science of self-organized criticality, Springer-Verlag, New York, (1996)

[17] P. Morley, R. Garcia-Pelayo, Europhys. Lett. **23**, 185, (1993)

[18] S. Mineshige, M. Takeuchi, H. Nishimori, Astrophys. Journ. **435**, L125, (1994)

[19] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A-L. Barabási, Nature **407**, 651, (2000)

[20] R. J. Williams, N. D. Martinez, Nature **404**, 180, (2000)

[21] H. Jeong, S. P. Mason, A-L. Barabási, Z. N. Oltvai, Nature **411**, 41, (2001)

[22] R. Albert, H. Jeong, A-L. Barabási, Nature **401**, 130, (1999)

[23] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, Y. Åberg, Nature **411**, 907, (2001)

[24] A-L. Barabási, H. Jeong, E. Ravasz, Z. Néda, A. Schubert, T. Vicsek, cond-mat/0104161

[25] D. Chandler, Introduction to Modern Statistical Mechanics, Oxford University Press, New York, (1987)

[26] D. McQuarrie, Statistical Mechanics, Harper & Row, New York, (1976)

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

[28] B. Bollobás, Graph Theory: An Introductory Course, Springer-Verlag, New York, (1979)

[29] W. Tutte, Graph Theory as I Have Known It, Oxford University Press, Oxford, (1994)

[30] S. Skiena, Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica, Addison-Wesley, Redwood City, (1988)

[31] P. Erdős, A. Rényi, Publ. Mathem. Debrecen **6**, 290, (1959)

[32] P. Erdős, A. Rényi, Publ. Mathem. Inst. Hung. Acad. Sci. **5**, 17, (1960)

[33] P. Erdős, A. Rényi, Bull. Inst. Int. Stat. **38**, 343, (1961)

[34] D. Watts, Small Worlds: The Dynamics of Networks Between Order and Randomness, Princeton University Press, New Jersey, (1999)

[35] D. J. Watts, S. H. Strogatz, Nature **393** 440, (1998)

[36] M. E. J. Newman, C. Moore, D. J. Watts, Phys. Rev. Lett. **84**, 3201, (2000)

[37] A-L. Barabási, R. Albert, Science **286**, 509, (1999)

[38] A-L. Barabási, R. Albert, H. Jeong, Physica A **272**, 173, (1999)

[39] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **89**, 218701, (2002)

[40] E. Ravasz, A-L. Barabási, Phys. Rev. E **67**, 026112, (2003)

[41] P. Holme, B. J. Kim, Phys. Rev. E **65**, 026107, (2002)

[42] J. C. Nacher, N. Ueda, T. Yamada, M. Kanehisa, T. Akutsu, BMC Bioinformatics **5**, 207, (2004)

[43] M. Barthélémy, L. A. Nunes Amaral, Phys. Rev. Lett. **82**, 3180, (1999)

[44] A. Barrat, M. Weigt, Eur. Phys. Journ. B **13**, 547, (2000)

[45] P. Bak, C. Tang, K. Wiesenfeld, Phys. Rev. A **38**, 364, (1988)

[46] D. Dhar, Phy. Rev. Lett. **64**, 1631, (1990)

[47] S. R. Nagel, Rev. Mod. Phys. **64**, 321, (1992)

[48] H. J. Jensen, Phys. Rev. Lett. **64**, 3103, (1990)

[49] T. Fiig, H. Jensen, Journ. Stat. Phys. **71**, 653, (1993)

[50] A. Hader, K. Saadouni, Y. Boughaleb, Physica A **342**, 561, (2004)

[51] P. Bak, C. Tang, K. Wiesenfeld, Phys. Rev. Lett. **59**, 381, (1987)

[52] H. Jensen, Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems, Cambridge University Press, Cambridge, (1998)

[53] P. Bak, C. Tang, Journ. Geophys. Res. **94**, 15635, (1989)

[54] K. Christensen, L. Danon, T. Scanlon, P. Bak, Proc. Nat. Acad. Sci. USA **99**, 2509, (2002)

[55] P. Bak, K. Christensen, L. Danon, T. Scanlon, Phys. Rev. Lett. **88**, 178501, (2002)

[56] M. N. Barber, B. W. Ninham, Random and Restricted Walks: Theory and Applications, Gordon and Breach, New York, (1970)

[57] E. B. Dykin, V. A. Uspenskii, Random Walks, Heath, New York, (1963)

[58] J. Feder, Fractals, Plenum, New York, (1988)

[59] B. Mandelbrot, The Fractal Geometry of Nature, Freeman, New York, (1983)

[60] M. Matsuzaki, Phil. Trans. Soc. Lond. A **348**, 449, (1994)

[61] W. D. Hamilton, Journ. Theor. Biol. **31**, 295, (1971)

[62] R. Cont, J-P. Bouchaud, Macroeconom. Dyn. **4**, 170, (2000)

[63] V. M. Eguíluz, M. G. Zimmermann, Phys. Rev. Lett.

[64] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 2^{nd} ed. McGraw-Hill, New York, (1984)

[65] J. Havil, Gamma: Exploring Euler's Constant, Princeton University Press, Princeton, (2003)

[66] J. Kenney, E. Keeping, Mathematics of Statistics, Pt. 2, 2^{nd} ed. Van Nostrand, Princeton, (1951)

[67] R. Mantegna, H. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, Cambridge, (2000)

[68] P. Gopikrishnan, M. Meyer, L. A. N. Amaral, H. E. Stanley, Eur. Phys. Journ. B **3**, 139, (1998)

[69] M. M. Dacorogna, U. A. Muller, R. J. Nagler, R. B. Olsen, O. V. Pictet, Journ. Int. Money Finance **12**, 413, (1993)

[70] W. Press, B. Flannery, S. Teukolsky, W. Vetterling, Numerical Recipes in Fortran: The Art of Scientific Computing, 2^{nd} ed. Cambridge University Press, Cambridge, (1992)

[71] J. Campbell, A. Lo, C. McKinlay, The Econometrics of Financial Markets, Princeton University Press, Princeton, (1997)

[72] A. Pagan, Journ. Emp. Finance **3**, 15, (1996)

[73] R. Cont, M. Potters, J-P. Bouchaud, Scaling in Stock Market Data: Stable Laws and Beyond, Proc. CNRS Workshop on Scale Invariance, (1997)

[74] V. F. Pisarenko, D. Sornette, physics/0403075

[75] B. B. Mandelbrot, Journ. Business **36**, 392 (1963)

[76] P. Clark, Econometrica **41**, 135, (1973)

[77] R. Engle, ARCH: Selected Readings, Oxford University Press, Oxford, (1995)

[78] T. Bollerslev, R. N. Chou, K. F. Kroner, Journ. Econometrics **52**, 5, (1992)

[79] P. Lévy, Calcul des Probabilités, Gauthier-Villars, Paris, (1925)

[80] J. Nolan, Stable Distributions: Models for Heavy Tailed Data, Birkhäuser, Boston, (2005)

[81] R. H. Rimmer, J. P. Nolan, Mathematica Journ. **9**, 776, (2005)

[82] H. Geman, T. Ané, Stochast Subordination, RISK, **9**, 146, (1996) (September)

[83] P. Bak, M. Paczuski, M. Shubik, Physica A **246**, 430, (1997)

[84] T. Lux, Journ. Econom. Behav. Org. **33**, 143, (1998)

[85] D. S. Scharfstein, J. C. Stein, Amer. Econom. Rev. **80**, 465, (1990)

[86] M. Grinblatt, S. Titman, R. Wermers, Amer. Econom. Rev. **85**, 1088, (1995)

[87] B. Trueman, Rev. Financ. Stud. **7**, 97, (1994)

[88] A. Bannerjee, Rev. Econom. Stud. **60**, 309, (1993)

[89] S. Bikhchandani, D. Hirshleifer, I. Welch, Journ. Politc. Econ. **100**, 992, (1992)

[90] A. Orléan, Journ. Econom. Behav. Org. **28**, 257, (1995)

[91] R. Cont, Scaling and Correlation in Financial Time Series. Science & Finance working paper.

[92] P. Dirac, Quantum Mechanics, 4^{th} ed. Oxford University Press, London, (1958)

[93] S. Gasiorowicz, Quantum Physics, Wiley, New York, (1974) **85**, 5659, (2000)

[94] R. D'Hulst, G. J. Rodgers, Int. Journ. Theor. Appl. Finance **3**, 609, (2000)

[95] R. D'Hulst, G. J. Rodgers, Physica A **280**, 554, (2000)

[96] R. D'Hulst, G. J. Rodgers, Eur. Phys. Journ. B **20**, 619, (2001)

[97] D. Zheng, G. J. Rodgers, P. M. Hui, R. D'Hulst, Physica A, **303**, 77, (2001)

[98] S. Maslov, Physica A **278**, 571, (2000)

[99] S. Maslov, M. Mills, Physica A **299**, 234, (2001)

[100] G. J. Rodgers, D. Zheng, Physica A **308**, 375, (2002)

[101] R. D'Hulst, G. J. Rodgers, Eur. Phys. Journ. B **21**, 447, (2001)

[102] J. D. Farmer, adap-org/9812005

[103] House of Commons, The Use of Operating Theatres in the Northern Ireland Health and Personal Social Services, **HC 414**, 7^{th} report of Session 2003-04, The Stationery Office, London, (2005)

[104] D. P. Smethurst, H. C. Williams, Nature **410**, 652, (2001)

[105] M. C. Papadopolous, M. Hadjitheodossiou, C. Chrysostomou, C. Hardwidge, B. A. Bell, Journ. Roy. Soc. Med. **94**, 613, (2001)

[106] D. P. Smethurst, H. C. Williams, Journ. Roy. Soc. Med. **95**, 287, (2002)

[107] B. Bunday, Basic Queueing Theory, Edward Arnold. Basic Queueing Theory, Edward Arnold, London. (1986)

[108] D. Cox, W. Smith, Queues, Methuen & Co, London, (1961)

[109] D. Gross, C. Harris, Fundamentals of Queueing Theory, 3^{rd} ed. Wiley, New York, (1998)

[110] A. Allen, Probability, Statistics, and Queueing Theory with Computer Science Applications, 2^{nd} ed. Academic Press, Orlando, (1990)

[111] L. Kullmann, J. Kertesz, Journ. Physica A **299**, 121, (2001)

[112] Private Communication, R. Brackenbury, Royal College of Surgeons.

[113] R. P. Freckleton, W. J. Sutherland, Nature **413**, 382 (2001).

[114] http://society.guradian.co.uk/NHSstaff/story/0,7991.438272.00.html.

[115] T. Schelling, Micromotives and Macrobehavior, W. W. Norton and Company, New York, (1978)

[116] R. D'Hulst, G. J. Rodgers, Physica A **294**, 447, (2001)

[117] E. Ben-Naim, P. L. Krapivsky, Eur. Phys. Journ. B **25**, 239, (2002)

[118] K. Okuyama, M. Takayasu, H. Takayasu, Physica A **269**, 125, (1999)

[119] V. Pareto, Le Cours d'Économie Politique, Macmillan, London, (1897)

[120] B. B. Mandelbrot, Int. Econom. Rev. **1**, 79, (1960)

[121] H. Jeffreys, B. Jeffreys, Methods of Mathematical Physics, 3^{rd} ed. Cambridge University Press, Cambridge, (1988)

[122] M. Trott, The Mathematica GuideBook for Programming, Springer-Verlag, New York, (2004)

List of Figures

List of Tables