Superstatistics and symbolic dynamics of share price returns on different time scales

Dan Xu

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

Queen Mary University of London.

School of Mathematical Sciences

Queen Mary University of London

January 18, 2017

I, Dan Xu, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third partys copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: Dan Xu

Date: 18/01/2017

Details of collaboration and publications:

D. Xu, C. Beck, "Transition from lognormal to χ^2-superstatistics for financial time series," *Physica A*, 453:173-183, 2016

Abstract

Share price returns on different time scales can be well modeled by a superstatistical dynamics. We provide an investigation which type of superstatistics is most suitable to properly describe share price dynamics on various time scales. It is shown that while chi-square-superstatistics works well on a time scale of days, on a much smaller time scale of minutes the price changes are better described by lognormal superstatistics. The system dynamics thus exhibits a transition from lognormal to chi-square-superstatistics as a function of time scale. We discuss a more general model interpolating between both statistics which fits the observed data very well. We also present results on correlation functions of the extracted superstatistical volatility parameter, which exhibits exponential decay for returns on large time scales, whereas for returns on small time scales there are long-range correlations and power-law decays.

We also apply the symbolic dynamics technique from dynamical system theory to analyse the coarse-grained evolution of share price returns. A nontrivial spectrum of Renyi entropies is found. We study how the spectrum depends on the time scale of returns, the sector of stocks considered, as well as the number of symbols used for the symbolic description. Overall our analysis confirms that in the symbol space transition probabilities of observed share price returns depend on the entire history of previous symbols, thus emphasizing the need for a model of share price evolution based on non-Markovian stochastic processes. Our method allows for quantitative comparisons of entirely different complex systems, for example the statistics of coarse-grained share price returns using 4 symbols can be compared with that of other complex systems.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Christian Beck, for his support, invaluable advice and encouragement to me during my PhD study. His dedication and enthusiasm for research was contagious and motivational for me. I would also like to acknowledge Dr. Pau Rabassa who shared many useful computational techniques with me.

Many thanks to my lovely friends for sharing your happiness and giving warmth to me during my study: Ms. Xingya Han, Ms. Xiang Li, Dr. Chao Li, Dr. Miao Lin, Dr. Weizhi Liu, Dr. Shuchen Zhang, Dr. Weiqi Li, Ms. Cheng Chen, and Dr. Xiaotian Yu.

Lastly, I would like to thank my family for all their love and encouragement and my parents, Mr. Kefu Xu and Ms. Qimei Gong for their unconditional love and support. Most of all, great thank goes to my husband, Dr. Xia Chen for being the most solid back up for me all the time.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Previous studies

People might be surprised that the first finding of power-law distributions was published in the social science rather than the natural sciences. It was 100 years ago when Pareto, an Italian social economist, modelled the wealth of individuals by a power-law [5]. Another concept, which is applied in virtually all fields of investigations in the natural science, is the random walk. The formalization of a random walk was also achieved for the first time in a research topic in finance by Bachelier in 1900, and it was until five years later, when Einstein released the first mathematical characterization of a random walk in the natural science [6]. Bachelier stated in his PhD thesis [7] that the distribution of option price changes follows a Gaussian distribution. Note that Bachelier's research was dealing with the behaviour of price changes while mainstream research afterwards was concerned with the differences in the logarithm of price. This includes the well known Black and Scholes option pricing model, which has been the most influential model in option-pricing theory since 1973 when it was published. It was assumed in the Black and Scholes model [8] that stock prices are evolving according to a geometric Brownian motion and the changes in the logarithm of price(log returns) possess a Gaussian distribution. Accordingly, the stock prices are log-normal distributed.

Although the Black and Scholes model has been regarded as a milestone in option-pricing theory since it was published, the assumptions on price evolu-

tion cannot reveal the so called "stylized facts" of financial markets. In 1963, Mandelbrot was the first one who pointed out that the unconditional distribution of empirical returns is non-Gaussian with positive excess kurtosis and fat tail. To reproduce the featured fat tails, Mandelbrot [9] and Fama [10] suggested to use the Levy stable distribution (with index $\alpha < 2$), which decays with a power law($Prob(x_t < x) \sim 1 - x^{-\alpha}$), to describe the price changes. The central limit theorem states that a rescaled sum of n of independent and identically distributed(i.i.d.) random variables will be aggregated to have a normal distribution as n approaches infinity, if those variables have finite variance. The stability property of a Levy distribution is raised by a generalized version of the central limit theory, which suggests that the sum of independent identically distributed random variables possessing the same Levy distribution with power-law tails of index α will tend to be characterized by the same Levy stable distribution of index α as the number of variables tends to infinity. Particularly, when individual variables have $\alpha > 2$, the sum converges to have a Gaussian distribution(a special case of the Levy stable distribution with the index $\alpha = 2$) and if $0 < \alpha \leq 2$, the sum will converge to obey a standard case of the Levy stable distribution with a fat tail property. Due to fact that the variance of Levy stable random variables is infinite for $\alpha < 2$, this method is less handy to implement in reality.

Instead of modelling the entire series of price changes by a specific distribution, one can characterize the tail region on its own by means of statistical extreme value theory [11]. Speaking of fat tails, the largest realizations of price returns(under normalization) are customarily described by a power law, i.e,$Prob(x_t < x) \sim 1 - x^{-\alpha}$. Without a hypothesis on the distribution of the entire return series, one can use the semi-parameter inference to estimate the tail index α. An abundant literature regarding this topic arisen during the nineties has demonstrated that the tail index α lies between 3 and 4 [12][13][14]. At this point of view, the Levy stable hypothesis is rejected.

Now we turn to the second stylized fact of stock returns, which is the known "volatility clustering". It is known that there are very short range correlations in the

logarithm returns as the autocorrelation function decays exponentially. However, it cannot be concluded from pairwise independence that asset returns are independent random variables. In 1969, Mandelbrot recognized that there exists a type of dependency or complex behaviour that affects the returns [15]. This dependency is prominent in absolute returns, squared returns and any other nonlinear functions of returns(referred to as volatility). In these measures, there is a visible long-lasting autocorrelation [16]. Theses findings suggested that volatility itself is an underlying stochastic processes that is characterizing the price returns. Accumulated studies have shown that there is long memory in volatility and the autocorrelation function of the volatility decays with a power law. Evidences of power-law decay in autocorrelations of volatility estimated by squared returns, local standard deviation or local average of the absolute returns are presented in [17][18][14][19][20][21][22].

Studies on the statistical properties of volatility are motivated by several aspects. For a given time, volatility levels reflect the amount of information that is arriving in the market. Volatility is also an indispensable indicator in the measure of investment risk. Moreover, volatility is playing an important role in the modelling of asset prices.

Among the emerging literature of volatility analysis, the GARCH(Generalized autoregressive conditional heteroskedasticity) model is considered as a benchmark model. It was put forward in 1986 by Bollerslev [23] in order to generalize the ARCH model which was introduced by Engle in 1982 [24]. In both models, the return at time t is modelled as a mixture of a normal random variable and a local variance. The variance, representing the squared volatility, varies with time and is deterministic at any time t given prior variances and returns.

Note that the volatilities characterized by conditional variances are not a priori stochastic. To add the randomness to the variance process, Taylor came up with the stochastic volatility(SV) model in 1986 [25]. The instantaneous variance at time t is described by

$$h_t - \alpha = \phi(h_{t-1} - \alpha) + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \tag{1.1}$$

where η_t is the desired random part in the formulation of volatility. α is the mean of h_t. ϕ is the autoregressive parameter which is always restricted to satisfy $|\phi| < 1$ to ensure that the process is stationary.

The requirement for refinement for both the GARCH and the SV models lies in the fact that these two models exhibit exponential decay in the volatility.

Both the GARCH and SV model are required to be refined due to the fact that these two models just exhibit exponential decay of the autocorrelations in the volatility. In order to meet the requirement of long memory in volatility, the initial GARCH model can be extended to the fractionally integrated GARCH(FIGARCH) model, see Baillie et al [26]. The elementary SV model has been extended to a long-memory stochastic volatility model by Breidt **et al** [27]. Besides, there are continuous time stochastic volatility models whose returns and volatilities are driven by stochastic differential equations(SDE), e.g. the Heston model [28].

Recently, with the accessibility of high frequency prices, crowding studies in the financial field are employing asset prices that are recorded more frequently than daily [29]. High frequency data bring advantages in exploring the microstructure of the financial behaviours, which contributes to the estimation of volatility. The stylized facts for high-frequency returns are similar but not quite the same as those for daily returns. For example, although the returns from both daily and small time scale observation exhibit fat tails in their distribution, the kurtosis will increase when observing the price more frequently [29]. The study of high-frequency share prices is still challenging and is worth more effort.

1.2 Superstatistical analysis

In recent years, achievements of researches on the study of financial market are inseparable from similar developments in the field of physics. Since 1990, a voluminous literature dealing with topics in the social sciences has been published by physicists. A new discipline, called **econophysics**, was set up to deal with financial problems by borrowing concepts and methods developed in physical research. Finding the appropriate distribution of log returns is one of the important research

aspects of this discipline [30][31][32][33][34], and is also one of the main objectives of this thesis.

Particularly, documentation of analogies and differences between price dynamics for financial data and similar dynamics in turbulence can be found in [31][35][36]. In financial market, the arriving information will spread from large to smaller scale markets and finally be reacted to by individual investors. On a qualitative level, this is similar to the way by which the injected energy is transferred from the large to smaller scale and finally be dissipated on the smallest scale in turbulence. From the quantitative point of view, the relation between price fluctuations and volatility is parallel to the relation between velocity fluctuations and a generalized inverse temperature parameter.

In 2003, Beck and Cohen introduced the Superstatistics concept, which is a powerful statistical method devoted to study complex systems with spatial-temporal variations in a given environment [37][38]. The core idea of the method is to characterize the given non-equilibrium system by a superpositon of several different statistical models referring to different time scales. For example, the most elementary baseline superstatistics models the velocity of a Brownian particle moving through a turbulent fluid by a superposition of local Gaussians weighted with a probability density $f(\beta)$, where β the slowly varying intensive parameter representing the inverse temperature. General kinds of complex systems have been well described by using this technique of superstatistics, such as hydrodynamics turbulence [39][40][41][42], wind velocity fluctuations [43][44], train delay statistics [3], cancer survival statistics [4] and much more. Most recent progress has been made in superstatistics includes analysis of sealevel fluctuations [45](2015) and statistics of daily rainfall [46](2016).

In the determination of $f(\beta)$ in most complex systems, there are basically three types of distributions to be considered. They are the χ^2 distribution, inverse-χ^2 distribution and log-normal distribution. Thus, the three classes of superstatistics are named according to $f(\beta)$. In the analysis of asset prices, the superstatistical parameter β can be interpreted as the volatility.

In the later chapters of this thesis, I apply superstatistical approach to study and model the stylised facts for financial returns. I investigated the stylized facts (fat tail distribution of returns, absence of autocorrelations of returns and volatility clustering) of share price returns on various time scales, which have been reported in previous studies [47][48][49]. It is verified that different types of superstatistics need to be adopted in describing asset price dynamics on different time scales. Specifically, χ^2-superstatistics works well on a time scale of days. It is equivalent to Tsallis statistics and has also been shown to be connected with the GARCH type model [50]. This finding is consistent with some recent studies [51][52]. On a much smaller time scale of minutes, the correlations of β exhibit power-law behaviour and the price changes are better characterized by lognormal-superstatistics. The asset price dynamics exhibits a transition from lognormal to χ^2-superstatistics as a function of time scale. A more general model interpolating between both statistics, which fits the observed data very well, will be introduced.

1.3 Multifractal analysis

Both a power-law decay of the probability density of unconditional returns and the power-law decline of the autocorrelation in volatilities can be characterized by scaling laws in finance. Fractal (Multifractal) analysis has been proved to be a useful tool to capture the scaling behaviour by means of a singularity spectrum $f(\alpha)$ and generalized dimension function $D(q)$, and has been applied to financial time series [53][54][55].

A fractal has a self-similar structure with a fractal dimension, which is a non-integer lying between the Euclidian and topologic dimension. Fractals cannot be explicitly described by descriptive statistical measures like mean μ because such measures depend on the scale ε of observation in a power law fashion

$$\frac{\mu_2}{\mu_1} = \left(\frac{\varepsilon_2}{\varepsilon_1}\right)^h \tag{1.2}$$

where h is the scaling exponent. However, the fractal dimension, also called the

capacity dimension

$$D = -\lim_{\varepsilon->0} \frac{\log r(\varepsilon)}{\log \varepsilon} \tag{1.3}$$

is defined as a scale-free measure of fractals. ε is scale and $r(\varepsilon)$ is the minimum number of cells with size ε needed to cover the object. The scaling exponent h is referred to as the Hurst exponent H in a fractal time series and it is linked to the capacity dimension by

$$H = 2 - D. \tag{1.4}$$

Note that D is a measure of roughness and higher values correspond to rougher processes, while the Hurst exponent refers to long memory persistence (or anti-persistence) [56]. A Brownian motion (Bm) has $H = 1/2$. A fractional Brownian motion (fBm) has $0 < H < 1/2$ for an anti-persistent process and $1/2 < H < 1$ for a persistent process. The Hurst exponent can be estimated by many methods including rescaled range analysis (R/S) [57], detrended fluctuation analysis (DFA) [58] and detrending moving average (DMA) [55].

While D is a universal measure for a (mono-)fractal time series, it changes along the processes possessing multi-scaling properties. In this case, a local scaling exponent α is defined for multifractal time series as

$$|X_{i+\Delta} - X_i| \propto |\Delta|^{\alpha(i)}. \tag{1.5}$$

α is called the Hölder exponent quantifying the local degree of singularity at a given time i and it varies though time. Calculating the fractal dimension for subsets with the same α yields the singularity spectrum $f(\alpha)$. A mono-scaling process has a unique Hölder exponent α along its path which is identical to the Hurst exponent H.

The singularity spectrum can be obtained via calculating different moments of the proper measure μ

$$\langle \mu^q(\varepsilon) \rangle \propto \varepsilon^{qH(q)}. \tag{1.6}$$

$H(q)$ is the generalized Hurst exponent describing the scaling behaviour of the qth-

moment. The obtained $qH(q)$ vs. q determines different kinds of processes: In the simplest case, $qH(q)$ is linear, which means $H(q)$ is constant and reduces to the Hurst exponent H. Such a process is called mono-scaling. On the other hand, if $qH(q)$ is non-linear, this means $H(q)$ depends on q. Such a process is said to be multi-scaling.

Different measures μ are adopted in different methods. The generalized Hurst exponent method (GHE) studies the qth moments of the increments of the process [59][60]. Multifractal detrended fluctuation analysis(MF-DFA) [61] and multifractal detrended moving average (MF-DMA) [62] both consider the qth moments of a fluctuation measure F, where the F in MF-DFA is determined as the deviation from the local trend and in MF-DMA as the deviation from the moving average.

The box counting method [63] and wavelet transform modulus maxima (WTMM) [64][65] analyse the singularity spectrum by using the partition function

$$Z(\varepsilon,q) = \sum_{i=1}^{r(\varepsilon)} \mu_i^q(\varepsilon), \tag{1.7}$$

$$\tau(q) = \lim_{\varepsilon\to 0} \frac{\log Z(\varepsilon,q)}{\log \varepsilon}, \tag{1.8}$$

where $\tau(q)$ is called the scaling function (or free energy function). It can be expressed by $H(q)$ as

$$\tau(q) = qH(q) - D_T \tag{1.9}$$

where D_T is the topological dimension, which equals 1 for time series.

Finally, $f(\alpha)$ can be obtained from $\tau(q)$ through Legendre transform. The shape of $f(\alpha)$ for a multifractal time series is concave while it reduces to a fixed value for mono-fractals [66].

Each method has both advantages and disadvantages and performs well for particular types of data. Barunik and Kristoufek [67] tested the Hurst exponent estimation of R/S, MF-DFA, DMA and GHE for fat tail distributions. They demonstrated that the R/S method is robust to the presence of heavy tails, however, it is very sensitive to the presence of short-range memory. MF-DFA and DMA ap-

proaches are not suitable for small size data and data with heavier tails. Overall the GHE method gives the best result among all the methods considered. Other comparative studies on multifractal techniques can also be found in the publications of Eke et al. [68] and Salat et al. [69].

The generalized fractal dimension, also called the Rényi dimension, is another scale-free feature of a multifractal time series. It is related to $\tau(q)$ and $H(q)$ by

$$D(q) = \frac{\tau(q)}{q-1} = \frac{qH(q)-1}{q-1}. \tag{1.10}$$

For a time series with non- or mono-fractal property, the shape of $D(q)$ will be constant over different q's. While a multifractal time series will have different Rényi dimensions for different q. There has been a bunch of literatures characterizing the multifractal behaviour for given systems by the spectrum of $D(q)$, for example, [70][71][72][73] explored the non-trivial $D(q)$ spectrum of human genomic sequences.

1.4 Symbolic dynamics

In my thesis, the financial time series is transformed into symbols prior to further analysis. The technique we used for the symbolic representation is called the symbolic dynamics technique, which is powerful in describing trajectories of a dynamical system in a coarse grained way. By cutting the entire symbol sequence into small segments of size N, the probability measure of each possible pattern of size N can be obtained. The Rényi entropy $K(q)$ is derived as

$$K(q) = \lim_{N\to\infty} \frac{1}{1-q}\frac{1}{N} \ln \sum_{j=1}^{\omega} \left(p_i^N\right)^q, \tag{1.11}$$

where ω is the number of allowed sequences of size N. By mapping each allowed symbol sequence of size N to a cell of size ε on the unit interval, $\omega = r(\varepsilon)$, $K(q)$ and $D(q)$ are linked to each other:

$$K(q) \sim -D(q)\ln\varepsilon + C_q, \tag{1.12}$$

where C_q is a constant independent of ε. In this regard, the multifractal feature of a financial symbol sequence can be described in terms of the Rényi entropy $K(q)$. The Rényi entropy is regarded as the dynamical counterpart of the Rényi dimension [74].

There have been many applications of symbolic dynamics in the study of financial time series. Tino [75] *et al.* predicted the direction of daily volatility changes by transforming the volatility changes into a symbolic stream in 1999. Bandt and Pompe [76] developed a technique for estimating the probability distribution based on counting ordinal patterns in 2002. Zanin [77] investigated the financial market efficiency by using the concept of forbidden patterns in 2008.

In the later chapters of this thesis, the symbolic dynamics technique is applied to investigate the share price dynamics for historical data sets of various sectors. The complex behaviour of share prices is quantified in terms of Rényi entropies and Rényi dimensions. It is also demonstrated that the stochastic process of symbol sequences observed for real share price data has significantly non-Markovian character.

1.5 Organisation

The rest of this thesis is organised as follows.

Chapters 2-4 give some general background which is useful for the actual research investigation in chapters 5-7.

Chapter 2 is providing some background knowledge on stochastic volatility. It begins with the GARCH type model in which the volatility is a deterministic variable at any given time. Then we introduce the stochastic volatility model which describes the volatility as a latent stochastic variable. Discrete and continuous stochastic volatility models are both mentioned. Finally, some distributions, which incorporate the stylized facts of asset returns, are introduced.

Chapter 3 gives a review of the superstatistics concept from theoretical approaches to applications. The reason for the occurrence of the often observed Tsallis statistics is explained by fluctuations of an intensive parameter. A generalized

version of Tsallis statistics is constructed based on superstatistics. Three types of distributions used in the description of intensive parameters for most complex systems are given. Finally some applications of different classes of superstatistics are presented.

Chapter 4 discusses fractal phenomena and multifractals. Commonly used multifractal analysis method such as Wavelet transform modulus maxima method and Multifractal detrended fluctuation analysis are introduced. Relevant background knowledge on symbolic dynamics and information measures is provided. Finally, the Rényi dimension and Rényi entropy are introduced as two important measurements of multifractality.

Chapter 5 discusses which type of superstatistics is most suitable to properly describe share price dynamics on various time scales. I firstly look at share price returns on large (daily) time scales and then do a similar analysis on small (minute) time scales. It is shown that while χ^2-superstatistics works well on a time scale of days, on a much smaller time scale of minutes the price changes are better described by lognormal superstatistics. We also present results on correlation functions of the extracted superstatistical volatility parameter, which exhibit exponential decay for returns on large time scales, whereas for returns on small time scales there are long-range correlations and power-law decays. Since the system dynamics exhibits a transition from lognormal to χ^2 superstatistics as a function of time scale, a synthetic model interpolating between both statistics is introduced. I have repeated the same experiment for different types of companies and looked at the similarities and differences among sectors and industries. The work of this chapter has been published in **Physica A** [78].

In chapter 6 the method of symbolic dynamics is applied for the multifractal analysis of financial time series. It offers a coarse-grained way to trace the time evolution of share price trajectories. A nontrivial spectrum of Rényi entropies is found. It is studied how the spectrum depends on the time scale of returns, the sector of stocks considered, as well as the number of symbols used for the symbolic description. It is confirmed that in the symbol space transition probabilities of ob-

served share price returns depend on the entire history of previous symbols. Thus the coarse-grained share price data has significantly non-Markovian character. The content of this chapter is being organized into a new paper.

The thesis is closed with a conclusion in chapter 7.

Chapter 2

Stochastic Volatility

2.1 Stylized facts for financial returns

General properties that all sources of return data manifest are called stylized facts. Through years of investigations in the financial market, daily returns often possess three remarkable properties. The first stylized fact is that the distribution of returns is non-Gaussian, especially it has a fat tail and is almost symmetric. Second, the correlations between returns are almost zero. Third, there are statistically significant positive correlations between absolute or squared returns. All of these properties are related to the fluctuations in volatility.

It is well-known that the volatility of asset prices is changing through time. Many papers have been published devoted to explain the underlying reasons behind such variation in volatility. For example, the volatility of stock market is usually high in the period of economic crises, and afterwards, turns to a normal level [79][80][81]. Also, the variation of volatility is partially affected by the level of the market [82][83]; this phenomenon is the so-called market leverage effect. Moreover, the volatility is also influenced by the information arriving at the market. Based on this postulation, Clark [84] and Epps and Epps [85] proposed a model which shows that volatility varies with the amount of related news reaching the market.

Fluctuations of asset prices and volatilities are described by different models emphasising particular features of asset prices. In this chapter, the ARCH-

type models, stochastic volatility models in both discrete-time and continuous-time frameworks will be introduced in brief, according to Taylor's textbook [86].

2.2 ARCH models

Autoregressive, conditional heteroskedastic(ARCH)-type models describe the dynamics of price and volatility by specifying the conditional variance h_t of the return for period t, conditional on the prior information. The asset return is defined as $r_t = \log \frac{p_t}{p_{t-1}}$ where p_t and p_{t-1} are the asset prices observed at time t and $t-1$.

2.2.1 ARCH(1) model

Assume the return can be described by the equation

$$r_t = \mu + {h_t}^{1/2} z_t \tag{2.1}$$

where z_t are i.i.d. $N(0,1)$ random variables. The distribution of returns over the period t, conditional on all previous returns, is normal with constant mean μ and time-varying conditional variance h_t. Thus r_t can be also represented as

$$r_t | r_{t-1}, r_{t-2}, \ldots \sim N(\mu, h_t) \tag{2.2}$$

The residual at time t is defined as

$$e_t = r_t - \mu = {h_t}^{1/2} z_t. \tag{2.3}$$

The simplest ARCH(1) model by Engle [24] is then represented as

$$h_t = \omega + \alpha e_{t-1}^2, \tag{2.4}$$

where $\omega > 0$ and $\alpha \leq 0$ are the two parameters of the volatility h_t. The volatility in period t is only affected by the return of period $t-1$.

The autocorrelation of the returns is zero

$$cov(r_t, r_{t+\tau}) = cov(e_t, e_{t+\tau}) = \mathbb{E}[{h_t}^{1/2} z_t {h_{t+\tau}}^{1/2} z_{t+\tau}] \tag{2.5}$$

$$= \mathbb{E}[{h_t}^{1/2} z_t {h_{t+\tau}}^{1/2}]\mathbb{E}[z_{t+\tau}] = 0 \tag{2.6}$$

for all $\tau > 0$, where $\mathbb{E}[x]$ denotes the expected value of x.

The forecast error is measured by the deviation of actual squared residuals from predicted squared residuals as

$$v_t = e_t^2 - \mathbb{E}[e_t^2 | r_{t-1}, ...] = e_t^2 - h_t. \tag{2.7}$$

The series of the forecast errors are uncorrelated. Replacing h_t in eq.2.4 by $e_t^2 - v_t$ yields

$$e_t^2 = \omega + \alpha e_{t-1}^2 + v_t \tag{2.8}$$

and hence squared residuals follow an *autoregressive process* of order one, abbreviated as AR(1) process. As α has to satisfy $|\alpha| < 1$ to make sure that the AR(1) process is stationary, recalling $\alpha \leq 0$, we have $\alpha < 1$. The autocorrelation function of squared residuals is equal to $\rho_\tau = \alpha^{|\tau|}$.

The unconditional variance of e_t is

$$Var(e_t) = \mathbb{E}[e_t^2] - (\mathbb{E}[e_t])^2 \tag{2.9}$$

$$= \mathbb{E}[e_t^2] \tag{2.10}$$

$$= \mathbb{E}[h_t z_t^2] \tag{2.11}$$

$$= \mathbb{E}[h_t] \tag{2.12}$$

$$= \omega + \alpha E[e_{t-1}^2] \tag{2.13}$$

and since e_t is stationary, $Var(e_t) = Var(e_{t-1}) = \mathbb{E}[e_{t-1}^2]$, then

$$Var(e_t) = \frac{\omega}{1-\alpha}. \tag{2.14}$$

And since $e_t = h_t^{1/2} z_t$, $\mathbb{E}[h_t^{1/2}] = \mathbb{E}[e_t^2]$, the unconditional variance of the returns is also $\omega/(1-\alpha)$.

The kurtosis of e_t is

$$\mathbf{kur}(r_t) = \frac{\mathbb{E}[e_t^4]}{(\mathbb{E}[e_t^2])^2} \tag{2.15}$$

$$= \frac{\mathbb{E}[h_t^2]\mathbb{E}[z_t^4]}{(\mathbb{E}[h_t])^2(\mathbb{E}[z_t^2])^2} \tag{2.16}$$

$$= \frac{3\mathbb{E}[h_t^2]}{(\mathbb{E}[h_t])^2}. \tag{2.17}$$

By Jensen's inequality, $\mathbb{E}[h_t^2] > (\mathbb{E}[h_t])^2$, we have $\mathbf{kur}(r_t) > 3$. Since if r_t has a normal distribution, $\mathbf{kur}(r_t) = 3$, thus the time-varying volatility h_t results in fat tails in the unconditional distribution of r_t.

2.2.2 GARCH(1,1) model

The ARCH(1) model is generalised to a GARCH model [23] by inserting a lagged variance term to the conditional variance equation as

$$h_t = \omega + \alpha(r_{t-1} - \mu)^2 + \beta h_{t-1}. \tag{2.18}$$

One previous return r_{t-1} and one previous value of the conditional variance h_{t-1} are used to define the conditional variance for period t. The four volatility parameters μ, α, β and ω are required to satisfy $\omega \geq 0, \alpha > 0, \beta > 0$ and $\alpha + \beta < 1$ to determine a non-negative outcome of h_t and to guarantee the stationary of h_t.

Substituting the forecast error $v_t = e_t^2 - h_t$ into Eq.2.18 yields

$$e_t^2 = \omega + (\alpha + \beta)e_{t-1}^2 + v_t - \beta v_{t-1}. \tag{2.19}$$

Hence squared residuals follow an ARMA(1,1) process. The autocorrelation of squared residuals, which are modelled by ARMA(1,1), is $C(\alpha+\beta)^\tau > 0$ for all $\tau > 0$ where C is positive and is determined by α and β as

$$C(\alpha, \beta) = \frac{\alpha(1 - \alpha\beta - \beta^2)}{(\alpha + \beta)(1 - 2\alpha\beta - \beta^2)} \tag{2.20}$$

Other properties of a GARCH(1,1) model can be obtained by using similar calculations as in the ARCH(1) model. The autocorrelation of the returns is zero for all $\tau > 0$; the unconditional variance of r_t is finite and $Var(r_t) = \omega/(1-\alpha-\beta)$; the kurtosis of r_t is always larger than 3.

2.2.3 EGARCH model

If the logarithm of the conditional variance h_t, rather than h_t itself, is characterized by an AR(1) process, with an asymmetric function of the standardized residuals z_{t-1}, the exponential-GARCH(EGARCH) model, proposed by Nelson [87] ,is given as

$$\log(h_t) = \mu_{\log(h)} + \Delta(\log(h_t) - \mu_{\log(h)}) + g(z_{t-1}), \tag{2.21}$$

and

$$g(z_{t-1}) = \vartheta z_{t-1} + \gamma(|z_{t-1}| - E[|z_{t-1}|]), \tag{2.22}$$

where $\mu_{\log(h)}$ is the mean of $\log(h_t)$, $-1 < \Delta < 1$ is the autoregressive parameter of this AR(1) process, ϑ and γ are two parameters of the function $g()$. $g(z_{t-1})$ are i.i.d., and if z_t are normal distributed, $E[|z_{t-1}|] = \sqrt{2/\pi}$. It is obvious that the slope of the function g changes with the sign of z_t, in which way the asymmetry is incorporated into the model.

More details can be found in papers of Bollerslev, Chou, and Kroner [88] and Bollerslev, Engle and Nelson [89] which provide reviews on the ARCH model in terms of theories and empirical applications.

2.3 Stochastic volatility models

Stochastic volatility models provide alternative approaches to describe returns and volatilities. Since the volatility is changing all the time frequently, it is plausible that one can model the volatility by using a random variable. Different stochastic processes, rather than the conditional variance process, for volatility are investigated in this chapter.

2.3.1 Standard stochastic volatility model

The standard SV model introduced by Taylor [25] is defined by

$$r_t = \mu + \sigma_t u_t \tag{2.23}$$

and

$$\log(\sigma_t) - \alpha = \phi(\log(\sigma_{t-1}) - \alpha) + \eta_t, \tag{2.24}$$

for return r_t and volatility σ_t. From Eq.2.23, the excess return $r_t - \mu$ is factorised as a product of σ_t and u_t, where u_t are assumed to be i.i.d. normal random variables: $u_t \sim N(0,1)$. σ_t is the unpredictable part of the volatility whose dynamics is described by Eq.2.24. The logarithm of σ_t is modelled by the AR(1) process. α is the mean of $\log(\sigma_t)$. ϕ is the autoregressive parameter which is always restricted to satisfy $|\phi| < 1$ to ensure that the process is stationary. η_t are i.i.d. white noises with $\mathbb{E}[\eta_t] = 0, \mathbb{E}[\eta_t^2] = \sigma_\eta^2$ and $\mathbb{E}[\eta_t \eta_{t-1}] = 0$. A further assumption is made that η_t are Gaussian distributed, which implies $\log(\sigma_t)$ are also Gaussian distributed. Particularly, $\log(\sigma_t) \sim N(\alpha, \beta^2)$ and $\eta_t \sim N(0, \sigma_\eta^2)$, with $\sigma_\eta^2 = \beta^2(1-\phi^2)$. Thus, Eq.2.24 is also called a Gaussian AR(1) process. The standard stochastic volatility model also assumes that the processes σ_t and u_t are stochastically independent.

The unconditional density function of r_t is given by the integral over the volatility variable σ as

$$f(r_t) = \int_0^\infty \psi(r_t|\mu,\sigma^2)\Lambda(\sigma|\alpha,\beta^2)d\sigma, \tag{2.25}$$

where $\psi(\cdot)$ is the Gaussian probability density function and $\Lambda(\cdot)$ is the lognormal probability density function. The integration uses the lognormal-normal density(LNN), whose shape is symmetric about μ. LNN density integrals can only be numerically evaluated. Note that the density in Eq.2.25 is equal to the density of systems obeying **lognormal superstatistics** for $\mu = 0$, which will be introduced later in Chapter 3.

2.3.2 Asymmetric stochastic volatility model

The standard stochastic volatility model does not specify the volatility process according to the direction of price changes. Asymmetric volatility effects can be captured by postulating that there is some dependence between η_t of the volatility and u_t in the logarithms of prices.

The general SV model that incorporates asymmetric effects is defined as before by

$$r_t = \mu + \sigma_t u_t \tag{2.26}$$

and

$$\log(\sigma_t) - \alpha = \phi(\log(\sigma_{t-1}) - \alpha) + \eta_t \tag{2.27}$$

where now the variables $(u_t, \eta_{t+1})^T$ have bivariate normal distributions with

$$\begin{pmatrix} u_t \\ \eta_{t+1} \end{pmatrix} \sim \text{i.i.d.} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \delta\sigma_\eta \\ \delta\sigma_\eta & \sigma_\eta^2 \end{pmatrix}\right) \tag{2.28}$$

so that δ is the correlation between u_t and η_{t+1} [86].

2.4 Continuous time stochastic processes

The previous models characterize asset price and volatility for discrete moments in time. However, how to characterize the price and volatility dynamics in the continuous time framework is also of particular interest. In this section, univariate and bivariate diffusion processes used in the financial market are introduced.

2.4.1 Geometric Brownian motion

Denote the process of an asset price by $S(t)$, the simplest continuous time model used for describing the dynamics of $S(t)$ is the well-known Geometric Brownian Motion, which is represented by

$$\log S(t) - \log S(0) = (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t), \tag{2.29}$$

where $(\mu - \frac{1}{2}\sigma^2)$ is called the drift function with σ the constant volatility, and $W(t)$ denotes a Wiener process. $S(t)$ is lognormal distributed, which guarantees its positive outcomes.

Another way to represent the Geometric Brownian motion is to use the stochastic differential equation (SDE):

$$dS/S = \mu dt + \sigma dW. \tag{2.30}$$

2.4.2 Bivariate diffusion process

2.4.2.1 The Bivariate Wiener Process

Consider two Wiener processes $W(t)$ and $Y(t)$ which are independent of each other. A third Wiener process $Z(t)$ defined by a linear combination of the previous two processes with a correlation coefficient ρ in the range of -1 to 1 is given as

$$Z(t) = \rho W(t) + \sqrt{1-\rho^2}Y(t). \tag{2.31}$$

The stochastic process $B(t)$, defined by a column vector as [86]

$$B(t) = (W(t), Z(t))^T, \tag{2.32}$$

is called the general bivariate Wiener process. A detailed discussion can be found in [90]. The coefficient ρ then equals the correlation between the increments of the elemental processes:

$$\text{cor}(W(t) - W(s), Z(t) - Z(s)) = \rho. \tag{2.33}$$

This correlation is often described by a compressed expression:

$$dW dZ = \rho dt. \tag{2.34}$$

2.4.2.2 Example

The stochastic volatility (SV) models of the previous section specify a volatility formula and then engage the volatility variable in an equation that defines the evolution of returns. A discrete SV model is an approximation to a bivariate diffusion process for asset price $S(t)$ and its stochastic variance $V(t)$, which is defined by the following equations:

$$dS/S = \mu dt + \sqrt{V} dW \tag{2.35}$$

and

$$dV = a(V)dt + b(V)dZ. \tag{2.36}$$

where dW and dZ are correlated by

$$dW dZ = \rho dt. \tag{2.37}$$

One of the choices is to specify Eq.2.36 as

$$d(\log V) = \kappa(\alpha - \log V)dt + \sigma dZ, \tag{2.38}$$

which generates an *Ornstein-Uhlenbeck*(OU) process [91][92][93] for $\log V$. α is the mean, thus the drift function $a(V) = \kappa(\alpha - \log V)$ is negative or positive when $\log V$ is above or below its mean level. κ is a parameter that governs the rate at which $\log V$ is pulled back towards α. Thus, the OU process is the continuous time extension of the previously mentioned AR(1) process, and Eq.2.38 combined with Eq.2.35 and Eq.2.37 is a continuous stochastic volatility model that can be approximated by the general SV model, which simplifies to the standard SV model when $\rho = 0$.

Another specification of a *square-root* process (Cox, Ingersoll, and Ross [94]) for $V(t)$ is used in the Heston model [28], which is of the form

$$dV = \kappa(\alpha - V)dt + \sigma\sqrt{V} dZ. \tag{2.39}$$

The drift function is attained as in the OU process. A further enhancement in the volatility function $b(V) = \sigma\sqrt{V}$ ensures that the process only produces positive values. Closed-form option prices can be calculated from a stochastic volatility model that specifies the $V(t)$ as a square-root process with any value of ρ.

More theory of continuous-time stochastic process can be found in the books by Cont and Tankov [95] and Etheridge [96]. Shephared provided a review of the stochastic volatility models in the text [97] and he also published a detailed survey of SV and ARCH models in [98].

2.5 Probability Distributions for Returns

Probability distributions are only plausible to describe returns if they are consistent with the mentioned stylized facts. They are expected to have fat tails and must be proximately symmetric. We are going to review some of the distributions that have these features.

Clark [84] and some others have suggested that the asset returns can be modelled by a mixture of normal distributions. Then, the conditional normal distribution of returns:

$$r_t|\omega_t \sim N(\mu, f(\omega_t)) \tag{2.40}$$

are defined by a constant mean μ and a function f of mixing variables ω_t which represents the conditional variance $\sigma_t^2 = f(\omega_t)$. There should be some assumptions that make the unconditional distribution of returns have excess kurtosis.

The mixing variable ω_t is possibly related to several observable quantities such as number of transactions [99], trading volume [84] or number of new informations that arrive at the market on day t [100]. The unconditional distribution of returns can be constructed by adopting a particular distribution for the conditional variance $\sigma_t^2 = f(\omega_t)$.

Praetz [101] suggested that ω_t^2 is distributed as an inverse gamma distribution and unconditional returns are then distributed with a generalized Student-t distribution with degrees of freedom $v > 2$. Then central moments of order $N < v$ are finite and the unconditional returns have a finite kurtosis $3 + 6/(v-4)$ when $v > 4$. The

Student-t distribution converges to a Gaussian distribution as $v \to \infty$. A gamma distribution has also been suggested for σ_t^2 by Madan dan Seneta [102], which implies that $\frac{1}{\sigma_t^2}$, instead of σ_t^2, is inverse gamma distributed.

A lognormal distribution for σ^2 is advocated by Clark [84] so that the resulting unconditional distribution for returns is the so-called lognormal-normal distribution. This result is supportive of the standard stochastic volatility model and is still popular in the literature dealing with stochastic volatility. The unconditional kurtosis is $3\exp(V)$, where V is the variance of $\log(\sigma_t^2)$. Besides, all the other moments are also finite. However the density function can only be represented by an integral and evaluated numerically.

Barndorff-Nielsen and Shephard [103] also showed that gamma and inverse gamma distribution are appropriate choices for the conditional variance in the modelling of returns recorded over various time scales.

Returns measured over time lags which are less than or equal to 5 minutes are referred to as high-frequency returns. There exist similarities in the stylized facts of daily and high-frequency returns. First, fat tail behaviour is also captured in high-frequency return. Additionally, the kurtosis of the distribution is observed to increase as the frequency of return measurement increase. Second, there are almost no correlations between high-frequency returns, however, the first-lag autocorrelation is usually negative. Third, there is nonnegligible dependence among high-frequency absolute returns. The magnitude of dependence has periodic U-shape corresponding to the intraday variation in volatility. Finally, volatility can have short but high bursts due to the fact that the intraday price is sensitive to the news absorbed by the market and the return becomes volatile after such a news release.

High-frequency returns can also be reasonably modelled by other mixtures of normal distributions. Among those literatures, Tauchen and Pitts [100], Clark [84] and Taylor [25] showed strong evidences that a lognormal distribution for the time-varying σ_t^2 was most appropriate.

Note that the the idea of constructing the unconditional returns by a mix-

ture of normal distributions, weighted by some appropriate densities for the conditional variance, is consistent with the concept of Superstatistics [37][38]. This technique was originally developed in the context of nonequilibrium statistical mechanics(Chapter 3) and can be applied to financial systems as well(Chapter5). The mentioned distributions, which are commonly chosen for characterizing σ_t^2, are also favoured in Superstatistics when it comes to fitting the variance related parameter β. Barndorff-Nielsen and Shephard [103] also stated that the class of distribution used to describe ω_t^2 must be selfdecomposable. It will be explained in detail in chapter 3 that gamma, inverse gamma and lognormal distribution are all selfdecomposable.

Chapter 3

Superstatistics

3.1 Motivation for Superstatistics

In this chapter we introduce a technique that is useful in the study of many kinds of nonequilibrium systems, called superstatistics [37][38][104]. The name of this method can be briefly understood as a superposition of different statistics on different time scales. It generalises Boltzmann statistics and explaines why the Tsallis distribution is quite often observed in many driven nonequilibrium systems.

3.1.1 Boltzmann statistics

Consider an ideal gas that is composed of many molecules that do not interact with each other except during collisions. Each particle in this system possess as a certain level of energy and the entire system contains an amount of E energy. For an isolated system under consideration, the total number of particles N, the volume of the gas V as well as the energy E are all fixed. Then questions of how is E assigned among the particles and how are the particles distributed among the system are of interest to Boltzmann's theory.

If one looks at such a system from the classical mechanics perspective, the N particles are each in a six-dimension phase space which can be divided into $W \ll N$ cells with local volume ω_i. Each cell consists of n_i particles and holds $\varepsilon_i = \frac{1}{2}mv_i^2$ kinetic energy at its center. Then the **macrostate** of the system is represented by the number of particles, the volume and the energy of $i = 1, 2, ..., W$ cells, i.e. $\{n_i\}$, $\{\omega_i\}$ and $\{\varepsilon_i\}$. The disorder number, $W(\{n_i\})$, represents the number of possible

microstates for the given macrostate.

It is easy to calculate that the number of ways to allocate n_1 particles to cell $i = 1$ is $\frac{N!}{n_i!(N-n_1)!}$, and the number of ways to allocate n_2 particles to cell $i = 2$ is $\frac{N!}{n_i!(N-n_1-n_2)!}$ and so on. Then the ways to distribute $\{n_i\}$ among $i = 1,2,...,W$ cells is

$$\begin{aligned} &\frac{N!}{n_1!(N-n_1)!} \times \frac{N!}{n_2!(N-n_1-n_2)!} \times \cdots \\ &= \frac{N!}{n_1!n_2!\cdots n_W!} \\ &= \frac{N!}{\prod_{i=1}^{W} n_j!}. \end{aligned} \tag{3.1}$$

Since each cell has a volume ω_i, any of the n_i particles can be assigned to anywhere in ω_i. The number of ways to allocate particles locally in their own cells for W cells is $\prod_{i=1}^{W} \omega_i^{n_i}$. Then the disorder number is enumerated as

$$W(\{n_i\}) = \frac{N!}{\prod_{i=1}^{W} j_i!} \prod_{i=1}^{W} \omega_i^{n_i}. \tag{3.2}$$

Under the A Priori assumption which states that all microstates have equal probabilities, the macrostate at equilibrium must include a maximum $W(\{n_i\})$. Thus $W(\{n_i\})$ also denotes a thermodynamic probability. Maximising $W(\{n_i\})$ while keeping the number of particles N and energy E fixed as

$$\sum_{i=1}^{W} n_i = N, \tag{3.3}$$

$$E = \sum_{i=1}^{W} n_i \varepsilon_i, \tag{3.4}$$

yields the Maxwell-Boltzmann distribution

$$\bar{n}_i = \frac{1}{Z} e^{-\beta \varepsilon_i}, \quad i = 1,2,...,W \tag{3.5}$$

where Z is the partition function and $\beta = \frac{1}{k_B T}$, where k_B is the Boltzmann constant and T is the temperature. Note that the probability to find a particle in cell i is

proportional to n_i under the a priori assumption, $\bar{p}_i \sim \bar{n}_i$. The equilibrium entropy S_{eq} in terms of $\bar{p}_i$ is

$$S_{eq}(\{\bar{p}_i\}) = -k_B \sum_{i=1}^{W} \bar{p}_i \log \bar{p}_i + constant. \tag{3.6}$$

Moreover, the Boltzmann's formula

$$S(E) = k_B \log W(E) \tag{3.7}$$

describes the relation between entropy and the number of probable microstates $W(E)$, corresponding to a certain system of energy E. Eq.3.7 is obtained from Eq.3.6 for $\bar{p}_i = 1/W$.

3.1.2 Tsallis statistics

For more complex systems where W cannot be postulated, the Boltzmann statistics is not adequate to reflect the information provided by the system. The probability of finding the occupancy in certain regions in the phase space should be determined by the system's dynamics. Tsallis [105] generalised the Boltzmann statistics by defining the entropy as

$$S_q = k_B \frac{1 - \sum_{i=1}^{W} p_i^q}{q-1} \tag{3.8}$$

where q[1] is a real parameter that should be adapted to the system under consideration. The Boltzmann-Gibbs-Shannon entropy is a special case of Tsallis entropy as when $q = 1$, Eq.3.8 leads to

$$S_1 = -k_B \sum_{i=1}^{W} p_i \log p_i \tag{3.9}$$

Maximising S_q under two constraints

$$\sum_i P_i(q) = 1 \tag{3.10}$$

[1]The q parameter related to the Tsallis statistics is different from the q parameter in the context of multifractal analysis in this thesis.

and

$$\sum_i P_i(q)E_i = U_q \tag{3.11}$$

where

$$P_i(q) = \frac{p_i^q}{\sum_{i=1}^W p_i^q} \tag{3.12}$$

are the so called escort probabilities [66], yields

$$\bar{p}_i = \frac{1}{Z_q}[1-(1-q)\tilde{\beta}E_i]^{1/(1-q)} \tag{3.13}$$

where Z_q^{-1} is a normalisation factor in order that $\sum_i \bar{p}_i = 1$, E_i is the energy level of the system and $\tilde{\beta}$ is related to the temperature. In the case of $q = 1$, $\bar{p}_i \sim e^{-\beta E_i}$ is of the exponential form of the Maxwell-Boltzmann distribution. While for $q \neq 1$, $\bar{p}_i$ is a power law. In addition, Eq.3.8 indicates a non-extensivity of the physical entropy for $q \neq 1$ as

$$S_q(A+B) = S_q(A) + S_q(B) + (1-q)S_q(A)S_q(B) \tag{3.14}$$

where A and B are two independent systems. In fact, many non-equilibrium systems may have this property.

In the following, we will discuss why it is the Tsallis statistics that is adapted by to many non-equilibrium physical systems, adopting the **superstatistical** point of view.

3.1.3 Superstatistics and derivation of Tsallis statistics

Let us first give a brief introduction of what is superstatistics [37][104][38] and then provide a simple example, a Brownian particle moving in a turbulent fluid.

Consider a nonequilibrium system which is spatio-temporally inhomogeneous. Such a system can be partitioned into many spatial cells with different values of an intensive parameter β. This parameter may stand for the inverse temperature, the energy dissipation or a chemical potential and so on, depending on the system under consideration. Locally, each cell, with a constant β, is assumed to reach equilibrium within a very short relaxation time. The local behaviour can be described by an ordinary statistics, i.e.Boltzmann factor $e^{-\beta E}$, where E is an effective energy of

the cell. In the long term, the intensive parameter β is fluctuating on a time scale T which is much larger than the relaxation time that a cell needs to reach local equilibrium. Then the entire inhomogeneous system is described by a superposition of ordinary Boltzmann factors weighted with a probability density distribution $f(\beta)$. In this case, the long term stationary distribution of this supertatistical system is defined as

$$p(E) = \int_0^\infty f(\beta)\frac{1}{Z(\beta)}e^{-\beta E}d\beta \tag{3.15}$$

where $e^{-\beta E}$ is normalised by a constant factor $Z(\beta)$.

Later in this section we will use an example to show that Eq.3.15 becomes equivalent to the Tsallis distribution if $f(\beta)$ is a χ^2 distribution. But generally, the intensive parameter β does not stick to χ^2 distribution in different systems. Three often observed distributions will be discussed later.

Consider a Brownian particle of mass 1 moving in a fluid with its velocity u modelled by a Langevin equation

$$\dot{u} = -\gamma u + \sigma L(t) \tag{3.16}$$

where $\gamma > 0$ represents the friction constant, $L(t)$ is the white noise term with a Gaussian probability distribution and σ denotes the strength of the noise. If γ and σ are constant-in other word, the inverse temperature denoted by $\beta := \gamma/\sigma^2$ is constant-then the stationary probability density of u is Gaussian distributed with variance β^{-1}. However, the environment through which the particle is moving can be spatio-temporally inhomogeneous due to the external dynamics of environmental changes, like the situation sketched in Fig.3.1.

As shown in Fig.3.1, the particle is moving through an environment that consist of many spatial cells. The relaxation time in each cell, which the Brownian particle needs to reach local equilibrium, is γ^{-1}. Each cell is assigned a local inverse temperature β_i, however, while the particle is moving through the environment, the inverse temperature varies from cell to cell on a large time scale $T \gg \gamma^{-1}$ and it has

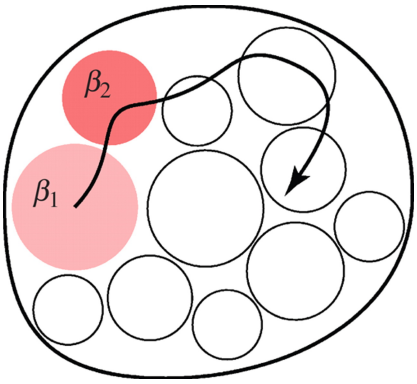


Figure 3.1: A Brownian particle is moving through a spatially inhomogeneous environment with fluctuating inverse temperatures β_i dominate different regions. (Picture taken from [1].)

the probability density distribution $f(\beta)$. Then one gets the conditional probability of the velocity u conditional on an observed intensive parameter β,

$$p(u|\beta) = \sqrt{\frac{\beta}{2\pi}} \exp\left(-\frac{1}{2}\beta u^2\right), \tag{3.17}$$

which is Gaussian. The joint probability of observing a certain u together with a certain β is

$$p(u,\beta) = p(u|\beta) f(\beta), \tag{3.18}$$

and the marginal probability of u, regardless of the observation of β, is

$$p(u) = \int_0^\infty p(u|\beta) f(\beta) d\beta, \tag{3.19}$$

which is a generalisation of the canonical distribution. As a result, Eq.3.19 is said to be superstatistical due to the fact that it is a "statistics ($f(\beta)$) of a statistics (Gaussian)".

To be specific, if γ and σ are fluctuating in a way that β is χ^2 distributed, then

$$f(\beta) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{n}{2\beta_0}\right)^{n/2} \beta^{n/2-1} \exp\left(\frac{-n\beta}{2\beta_0}\right), \tag{3.20}$$

where n represents the degrees of freedom. β_0 denotes the average of the varying β

which is calculated as

$$\langle\beta\rangle = \int_0^\infty \beta f(\beta) d\beta = \beta_0. \tag{3.21}$$

and the variance is given by

$$\langle\beta^2\rangle - \beta_0^2 = \frac{2}{n}\beta_0^2. \tag{3.22}$$

Substituting Eq.3.17 and Eq.3.20 into Eq.3.19 yields

$$p(u) = \int_0^\infty p(u|\beta) f(\beta) d\beta = \frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{\frac{\beta_0}{\pi n}} \frac{1}{\left(1+\frac{\beta_0}{n}x^2\right)^{\frac{n}{2}+\frac{1}{2}}} \tag{3.23}$$

which can be rewritten in the form of a Tsallis distribution

$$p(u) \sim \frac{1}{\left(1+\frac{1}{2}(q-1)\tilde{\beta}u^2\right)^{\frac{1}{q-1}}} \tag{3.24}$$

with the following identifications:

$$\frac{1}{q-1} = \frac{n}{2}+\frac{1}{2} \iff q = 1+\frac{2}{n+1} \tag{3.25}$$

$$\frac{1}{2}(q-1)\tilde{\beta} = \frac{\beta_0}{n} \iff \tilde{\beta} = \frac{2}{3-q}\beta_0 \tag{3.26}$$

Note that we have obtained a power law distribution 3.24 from the fluctuating β with distribution 3.20.

3.2 Observed universality classes

Note that β is a positive fluctuating parameter determined by the spatio-temporal dynamics, so a Gaussian distribution is not proper for $f(\beta)$. In the following some possible kinds of distributions are considered [39].

3.2.1 χ^2-superstatistics

Suppose there are microscopic random variables ξ_j, $j = 1,\ldots,J$ independently affecting β in an additive way. According to the Central Limit Theorem, the rescaled

sum of the random variables $\frac{1}{\sqrt{J}}\sum_{j=1}^{J}\xi_j$ will be close to a Gaussian random variable X_1 when J is large enough. Then there can be a collection of n independent Gaussian random variables $X_1,\ldots,X_n$ and each of them are contributing by those microscopic random variables. If $\beta=\sum_{i=1}^{n}X_i^2$, then β is positive as expected and distributed according to the χ^2 distribution, i.e.

$$f(\beta)=\frac{1}{\Gamma(\frac{n}{2})}\left(\frac{n}{2\beta_0}\right)^{n/2}\beta^{n/2-1}e^{-\frac{n\beta}{2\beta_0}}. \tag{3.27}$$

$f(\beta)$ has two parameters, where β_0 is the average of β and n is the degree of freedom. We have shown that a χ^2 type of $f(\beta)$ leads to χ^2-superstatistics according to Eq.3.19, which is equivalent to Tsallis statistics.

3.2.2 Inverse χ^2-superstatistics

On the other hand, if, instead of the inverse temperature β, β^{-1} is the sum of n squared Gaussian random variables X_i ,$\beta^{-1}=\sum_{i=1}^{n}X_i^2$, where X_i are generated by the same way in 3.2.1. The resulting β has a inverse χ^2 distribution:

$$f(\beta)=\frac{\beta_0}{\Gamma(\frac{n}{2})}\left(\frac{n\beta_0}{2}\right)^{n/2}\beta^{-n/2-2}e^{-\frac{n\beta_0}{2\beta}}. \tag{3.28}$$

3.2.3 Lognormal-superstatistics

β may be generated by multiplicative random processes. Start from a local cascade random variable $X_1=\prod_{j=1}^{J}\xi_j$, where J is the number of cascade steps and the ξ_j are positive microscopic random variables. The random variable $\frac{1}{\sqrt{J}}\log X_1=\frac{1}{\sqrt{J}}\sum_{j=1}^{J}\log\xi_j$ becomes Gaussian as $J\to\infty$ according to the Central Limit Theorem. Hence X_1 has log-normal distribution. If there are n such product contributions to β, that is, $\beta=\prod_{i=1}^{n}X_i$, then $\log\beta=\sum_{i=1}^{n}\log X_i$ is a sum of Gaussian random variables, and is Gaussian as well. Thus β has a log-normal distribution, i.e.,

$$f(\beta)=\frac{1}{\sqrt{2\pi}s\beta}\exp\left\{\frac{-(\ln\frac{\beta}{\mu})^2}{2s^2}\right\}, \tag{3.29}$$

where μ and s are mean and variance parameters.

3.3 Application to turbulent flow

The first example of application of supersatistics is the statistics of Taylor-Couette turbulent flow. The data used in this example is a velocity time series $v(t)$ measured in turbulent Taylor-Couette flow, which is taken from the experiments made by The velocity differences are defined by $u(t) := v(t+\delta) - v(t)$, where $v(t)$ represents the velocity and δ denotes the time lag by which the velocity series is extracted. The stationary probability distribution of the velocity difference $P(u)$ is known to be non-Gaussian. The value of Reynolds number and time lag used in the paper of der Straeten and Beck [2] are $Re = 540000$ and $\delta = 16$, and the number of data points is $n \approx 2 \times 10^7$.

Let β represent the time-varying inverse variance in u. Comparing the histogram of β with the χ^2, inverse-χ^2 and log-normal distributions yields that the log-normal distribution

$$f(\beta) = \frac{1}{\alpha\sqrt{2\pi}} \frac{1}{\beta} e^{-(\ln\beta - \theta)^2/2\alpha^2} \tag{3.30}$$

with $\alpha = 0.5222$ and $\theta = -1.51$ is the best fit to the experimental distribution of β, see Fig.3.2.

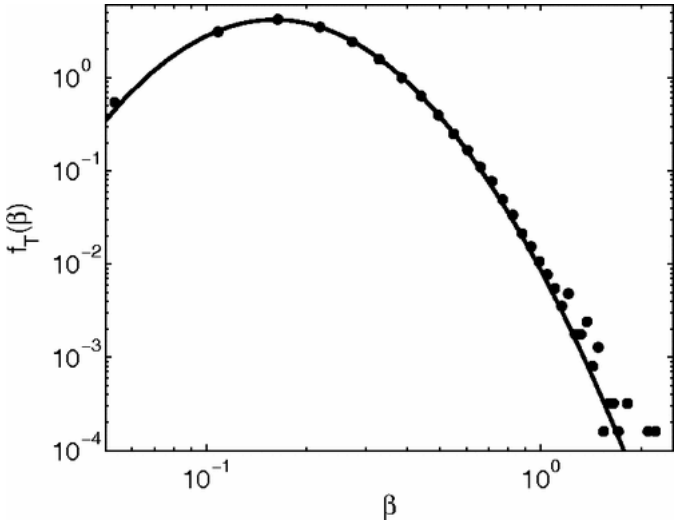


Figure 3.2: The probability distribution(dots) of the inverse variance parameter β calculated from the experimental turbulent data is well fitted by the log-normal distribution$f(\beta)$(solid line) with parameters $\alpha = 0.5222$ and $\theta = -1.51$. This figure is taken from [2]

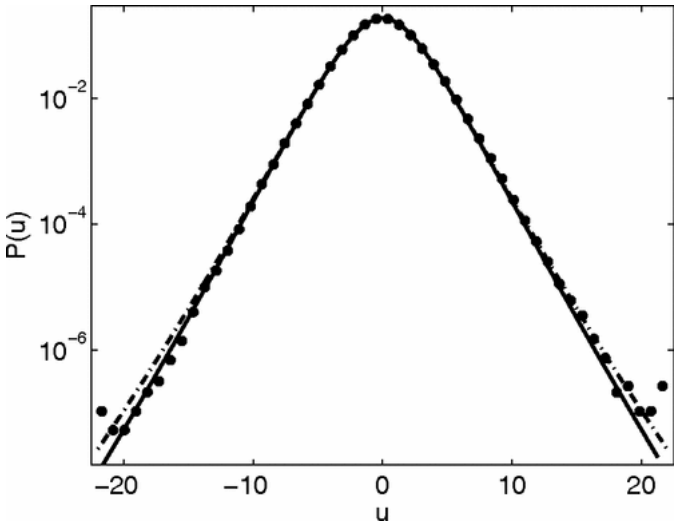


Figure 3.3: The empirical distribution of u(dots) extracted from the time series of velocity differences in turbulent Taylor-Couette flow is fitted by $p(u)$(dashed line). This figure is taken from [2]

Then the unconditional distribution of u can be expressed as

$$P(u) := \int_0^\infty d\beta f_T(\beta) \sqrt{\frac{\beta}{2\pi}} e^{-\frac{1}{2}\beta u^2} \tag{3.31}$$

which can be understood as a superposition of local Gaussians weighted with a distribution of the slowly varying parameter β. Fig.3.3 shows that the empirical distribution of u can be well modelled by $p(u)$, by which we can conclude that the distribution of velocity differences in turbulent Taylor-Couette flows can be well described by log-normal superstatistics.

For quantum turbulent flows, χ^2-superstatistics yields a better fit [106][107].

3.4 Other applications of Superstatistics

The superstatistical technique has been widely applied to many different kinds of complex systems. Here we will describe some further applications, corresponding to the other two classes of superstatistics: χ^2 superstatistics and inverse-χ^2 superstatistics.

3.4.1 Application to train delays

The probability distribution of train delays in England is found to be a q-exponential functions with asymptotic power law tails, and the underlying reason for the occurrence of q-exponentials can be explained by means of superstatistics [3].

Train delays are often observed at various stations in the UK. Suppose that the waiting time t until the actual departure is described by a Poisson process [93] , the probability distribution of the waiting time t is given by

$$P(t|\beta) = \beta e^{-\beta t}, \tag{3.32}$$

which represents the probability distribution of the delay time t conditional on a given positive rate parameter β. The above Poisson distribution is already normalised since

$$\int_0^\infty P(t|\beta)dt = \beta \int_0^\infty e^{-\beta t}dt = 1. \tag{3.33}$$

Under the assumption that the delay time is no less than 1 minute, a large value of β indicates that most trains depart with a short period of delay, whereas a small β means there is a higher probability for rather long delays. In reality, on a large time scale, β is fluctuating in terms of the railway network environment which makes the above model superstatistical. Temporal variations in β may happen due to a surge in passenger numbers during the begin of the holiday season, bad weather conditions or many other extreme events like derailments, strikes, etc. There can also be spatial variations in β, in addition to the temporal ones.

The observed long term distribution of the delay time becomes a superposition of local Poisson distributions with a fluctuating parameter β. If the slowly changing parameter β is distributed in accordance with a probability density $f(\beta)$, the marginal distribution of train delays is given by

$$p(t) = \int_0^\infty f(\beta)p(t|\beta)d\beta = \int_0^\infty f(\beta)\beta e^{-\beta t} \tag{3.34}$$

which is actually to be compared with the experimental data. Assuming the fact that

the n causes of fluctuations in β can be represented by n Gaussian random variables and they are affecting β in an additive way, β is obtained to be χ^2 distributed with n degrees of freedom.

For such $f(\beta)$, Eq.3.34 can be evaluated to a q-exponential form

$$p(t) \sim (1+b(q-1)t)^{\frac{1}{1-q}} \tag{3.35}$$

with $q = 1+2/(n+2)$ and $b = 2\beta_0/(2-q)$ where β_0 is the average β. The q-exponential is actually a power law for large t.

The fitting result shown in Fig.3.4 demonstrates the above model is in good accordance with the recorded train delays with fitted parameters b and q.

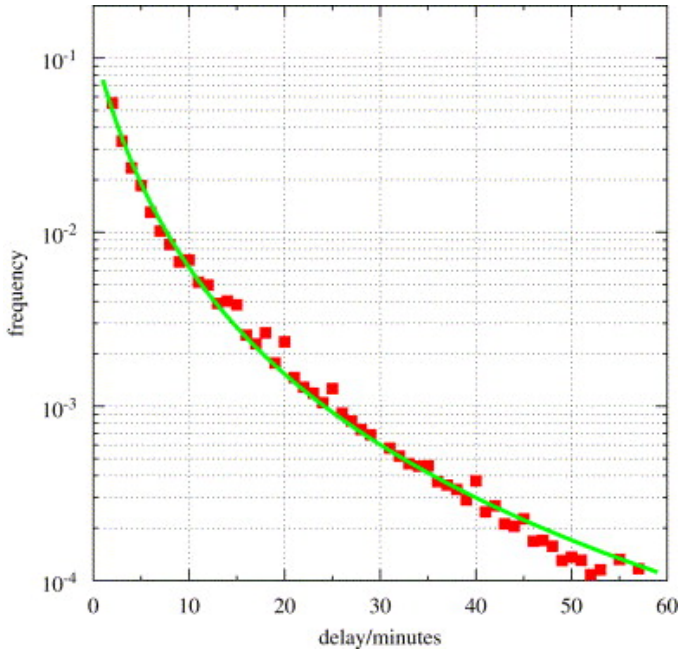


Figure 3.4: All-route data of train delay times with the best fitted χ^2 superstatistics. This picture is taken from [3].

In summary, the British railway delays are well described by a χ^2 superstatistics.

The χ^2 type of superstatistics is quite often observed to be applicable in a variety of complex systems. We will give three further applications of this type which are applications to sea-leval fluctuations(2015) [45], daily rainfall(2016) [46] and traffic flow(2016) [108].

3.4.2 Application to sea-level fluctuations

P. Rabassa and C. Beck present an analysis of the probability densities of measured sea-levels at five coastal locations in the UK [45]. They investigated the time series of the difference of the recorded non-tidal residuals. The non-tidal residual is the stochastic part of the three components of sea-level measurement. Concretely, each observed sea-level $Z(t)$ is the sum of an annual mean $M(t)$, an astronomical tide $X(t)$ and a non-tidal residual $Y(t)$. As the finding of slight skewness in the histogram of non-tidal residuals conflicts with the stipulation of local Gaussian behaviour, the statistics of non-tidal residual differences $Y(t_{i+1}) - Y(t_i)$ rather than the residual themselves are of interest. By virtue of the Kolmogorov-Smirnov(K-S) test, it is clearly verified that the experimental distribution of the sea-level differences can be well modelled by χ^2 superstatistics.

3.4.3 Application to daily rainfall

G.C. Yalcin, P. Rabassa and C. Beck investigated the probability distribution of daily rainfall amounts of different locations around the world [46]. It is demonstrated that the daily rainfall amounts possess a q-exponential distribution with $q \approx 1.3$. On the other hand, though the waiting time between the rainy days is tentatively observed to obey an exponential distribution, it is manifested that a q-exponential distribution with exponent q close to 1 as $q \approx 1.05$ allows a better fit to the empirical data. Observed q-exponentials are interpreted by χ^2-superstatistics, for which there exist two well separated time scales in the system. For the daily rainfall dataset, the probability distribution is modelled by a superposition of local Gaussian distributions weighted by the density of a variance parameter β, where β is fluctuating slowly on a large time scale and follows a χ^2-distribution. Differently, the waiting time is described as a superstatistical Poisson process where the data follows Poisson process locally with a positive rate parameter β which is χ^2-distributed on the large time scale.

3.4.4 Application to traffic flow

C. Kosun and S. Ozdemir applied the technique of superstatistics to the modelling of one-directional traffic flow monitored at Izmir highway [108]. They demonstrated that the vehicle speeds, the differences of vehicle speed and the shuffled vehicle speeds are all observed to follow q-exponentials, which implies χ^2-superstatistics. However, for data sets of speeds and speed differences, the exponent q obtained directly from the fitted q-exponential distribution differs from the value of q calculated from the parameter of the χ^2-distribution characterizing the variance parameter. They claim that the reason of this is due to the fact that the index q of the q-exponential distribution is a time-independent parameter while the q, which is obtained from superstatistics(3,26) and associated with the variance for given data slices, is time-dependent. Such phenomenon is also noticed in the analysis of financial data which will be presented in Chapter 5.

3.4.5 Application to cancer survival

The second type of superstatistics, inverse χ^2-superstatistics is observed to be applicable in the study of metastasis and cancer survival [4]. To have insights into the mechanisms of metastasis, cancer progression is analysed in the context of a complex nonequilibrium system. It is been illustrated in [4] that the superstatistical model can fit the probability distributions of the survival time of breast cancer patients perfectly.

Based on the fact that the proliferation of distant metastases is the immediate cause of death for over 90% of the cancer patients [109], death from cancer can be regarded as a consequence of multiple occurrences of cancer cell metastasis. Fig.3.5 describes the two pathways through which distant metastases may arise: the lymphatic system and the vascular system. For example, cancer cells of the primary tumor can invade into the local lymph nodes, proliferating while spreading to distant sites. On the other hand, cancer cells could also transmit to distant organs by intervening the blood circulation.

Let T be the survival time of a cancer patient. Then T is a continuous random variable that denotes the waiting time until n events of metastases have accumu-

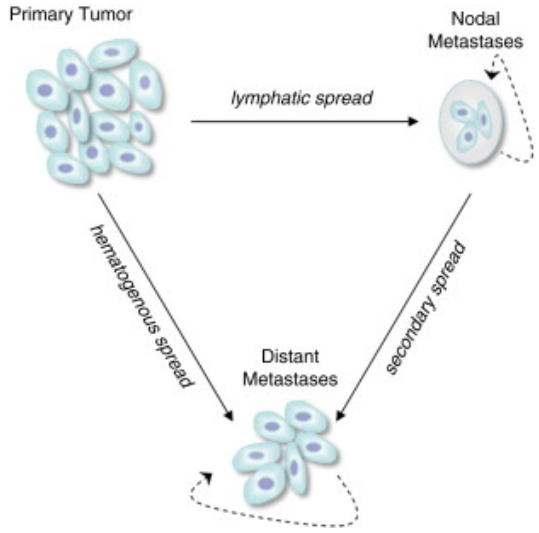


Figure 3.5: Macroscopic pathways of the metastatic cascade. This picture is taken from [4].

lated to induce death. Each event obeys a Poisson process and is exponentially distributed. Thus, the density of T is a gamma distribution $\Gamma(n, \frac{1}{\lambda})$ with parameter λ [110],

$$f(t|\lambda) = \frac{t^{n-1}\lambda^n \exp -\lambda t}{\Gamma(n)} \tag{3.36}$$

where $n \geq 1$ controls the shape and $\lambda \geq 0$ adapts the scale. However the above model becomes superstatistical due to the fact that the parameter λ fluctuates, indicating various levels of aggressiveness of the cancer propagation. It is explained in [4] that the probability distribution of λ is an inverse-χ^2 distribution with n degrees of freedom

$$f(\lambda) = \frac{\lambda_0(n\lambda_0/2)^{n/2}}{\Gamma(n/2)}\lambda^{-n/2-2}\exp\left(\frac{-n\lambda_0}{2\lambda}\right). \tag{3.37}$$

The probability density function of survival time is obtained by multiplying Eq.3.36 and Eq.3.37, and integrating over all possible λ:

$$p(t) = \int_0^\infty \frac{t^{n-1}\lambda^n e^{-\lambda t}}{\Gamma(n)} \frac{\lambda_0(n\lambda_0/2)^{n/2}}{\Gamma(n/2)}\lambda^{-n/2-2}e^{\frac{-n\lambda_0}{2\lambda}}d\lambda, \tag{3.38}$$

or

$$p(t) = \frac{(n\lambda_0)^{3n/4}}{\Gamma(n)\Gamma(n/2)}\left(\frac{t}{2}\right)^{3n/4-1}\left[\frac{\sqrt{2n\lambda_0 t}}{n}K_{n/2+1}\left(\sqrt{2n\lambda_0 t}\right) - K_{n/2}\left(\sqrt{2n\lambda_0 t}\right)\right], \tag{3.39}$$

where $K_\nu(z)$ is the modified Bessel function. This density function with estimated parameters is fitted to the observed distribution of survival time, see Fig.3.6.

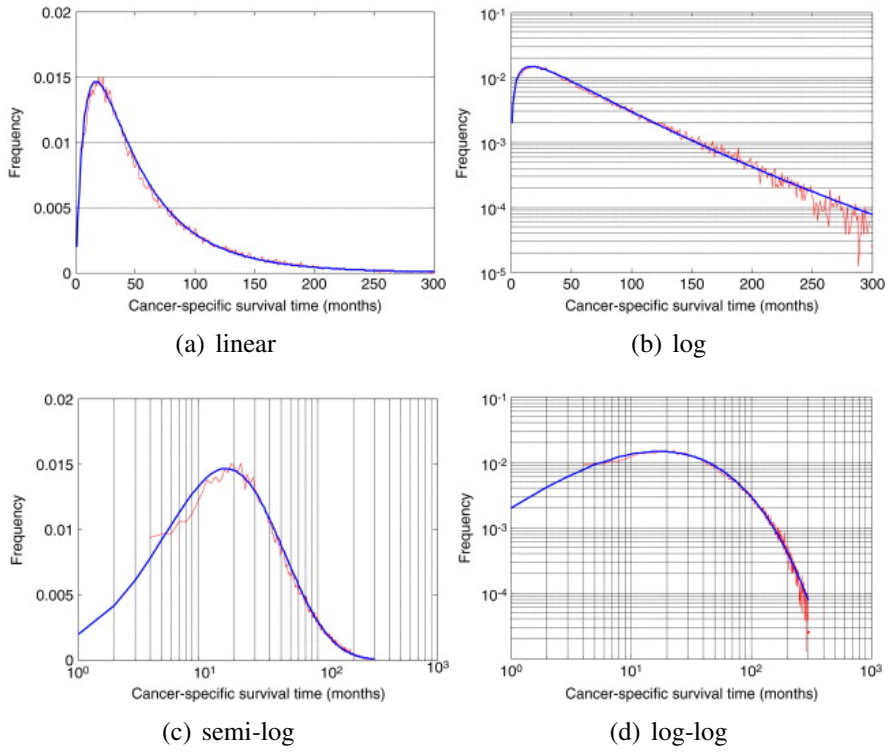


Figure 3.6: Data of the survival statistics together with the fitted curve as given by Eq.3.38 ($n = 3.00004, \lambda_0 = 0.08712$). This picture is taken from [4].

Clearly, the inverse-χ^2 superstatistical model is in a good agreement with the real data, keeping in mind that the data shown are survival distributions conditional on death from cancer.

A more recent paper [111] published in 2015 makes a careful investigation of migration processes of cancer cells in a heterogenous environment and comes to the conclusion that a superstatistical model is in very good agreement with the data.

3.5 Application to finance

Many well established concepts in mathematical finance (such as the Black-Scholes model) are based on the assumption that an index or a stock price follows a geometric Brownian motion, and as a consequence the log returns of these processes are Gaussian distributed. But nowadays it is well known that the log returns of realistic stock prices are typically non-Gaussian with fat tails [112][50][113][2][114][115][116][117][118]. Such behaviour revealed in the financial market can also be well captured by superstatistical models.

In finance, early applications of the superstatistics concept were worked out by Duarte Queiros et al. [112, 50] and Ausloos et al. [113]. Van der Straeten and Beck [2] analysed daily closing prices of the Dow Jones Industrial Average index (DJI) and the S&P 500 index. They verified that both log-normal superstatistcs and χ^2 superstatistics result in good approximations. Biro and Rosenfeld [114] also studied the data sets of the Dow Jones index and verified that the distribution of log returns is well fitted by a Tsallis distribution. Katz and Li Tian [115] showed that the probability distributions of daily leverage returns of 520 North American industrial companies during the 2006-2012 financial crisis comply with the q-Gaussian distribution which can be generated by χ^2 superstatistics. They also verified in [116] that the Tsallis entropic parameter q obtained by direct fitting to q-Gaussians coincides with the q obtained from the shape parameters of the χ^2 distribution fitted to the histogram of the volatility of the returns. Gerig, Vicente and Fuentes [117] consider a similar model that indicates that the volatility of intra day returns is well described by the χ^2 distribution, see also [118] for related work in this direction.

In a later chapter of this thesis, various datasets of historical share prices will be carefully analysed and it will be demonstrated which type of superstatistics is best suited to model the dynamics.

Chapter 4

Multifractal Analysis

4.1 Fractals

A fractal is a complicated set that exhibits selfsimilar structures on arbitrary scales. Fractals can be found all over in nature. Moreover, fractals can be artificially formed by iterative processes. In this section we will first look at how fractal structures are built by studying some simple examples of fractals and we will then explain a dimension parameter which is called the 'fractal dimension'.

4.1.1 Examples of fractals

4.1.1.1 Cantor sets

One example of a fractal is the so called classical Cantor set. We start with a unit interval, equally divide the interval into three subintervals of length $\varepsilon_1 = \frac{1}{3}$. Removing the central subinterval, there are two left-over subintervals which lie on the left and right ends of the original unit interval. These two subintervals are again segmented into three subintervals of size $\varepsilon_1 = \frac{1}{9}$ respectively and the central subintervals are taken out again. If we repeat such a procedure infinitely many times, the classical Cantor set is obtained (see Fig.4.1).

Such subset of the claimed Cantor set contains an infinite amount of real numbers and the elements of this set are uncountable. The real numbers in this set can

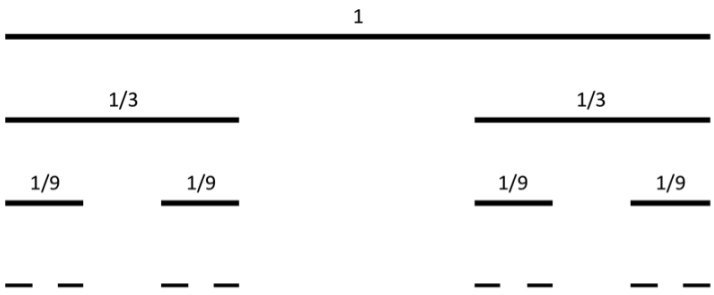


Figure 4.1: Construction of the classical Cantor set.

be represented by the ternary notation as

$$x = \sum_{k=1}^{\infty} s_k 3^{-k} \tag{4.1}$$

where s_k is equal to 0 or 2.

4.1.1.2 Sierpinski carpet

A 2-dimensional version of the Cantor set is the so called Sierpinski carpet. We start with a unit square and divide this square into $3^2 = 9$ identical subsquares of side length $1/3$ of the original square. We delete the central subsquare and repeat this 'divide then delete' procedure for the other eight subsquares which were surrounding the central one that has been removed. The 'remainder' after infinitely repetition of such iteration is called the 'Sierpinski carpet' (See Fig.4.2). Note that the area of this remainder is zero. As a 2D version of the Cantor set, the Sierpinski carpet contains the Cantor sets as subsets. These particular subsets are lying on the two central lines through the Sierpinski carpet, one of which is horizontal and the other one is vertical.

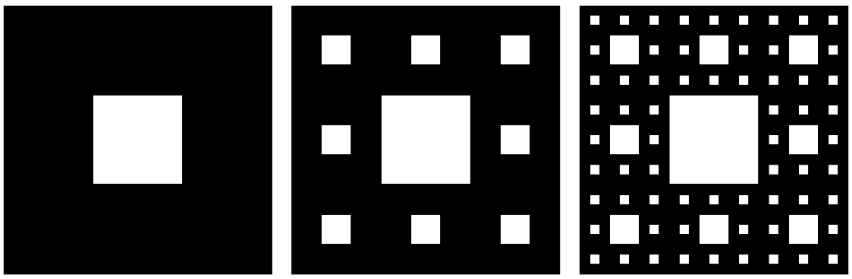


Figure 4.2: Construction of the Sierpinski carpet.

4.1.1.3 Sierpinski sponge

Again, the 3-dimensional analog of the Sierpinski carpet is called the Sierpinski sponge (also named as Monger sponge). Cutting a unit cube into $3^3 = 27$ identical subcubes of side length $1/3$ of the original cube, and removing the 6 subcubes in the middle of each face and 1 subcube in the very centre of the original cube, we will have $27 - 6 - 1 = 20$ subcubes remaining. Repeating this iterative process for each of the remaining subcubes, we obtain the so called Sierpinski sponge (see Fig.4.3). Similar to the Sierpinski carpet, the volume of the sponge is zero when the above procedure is infinitely operated.

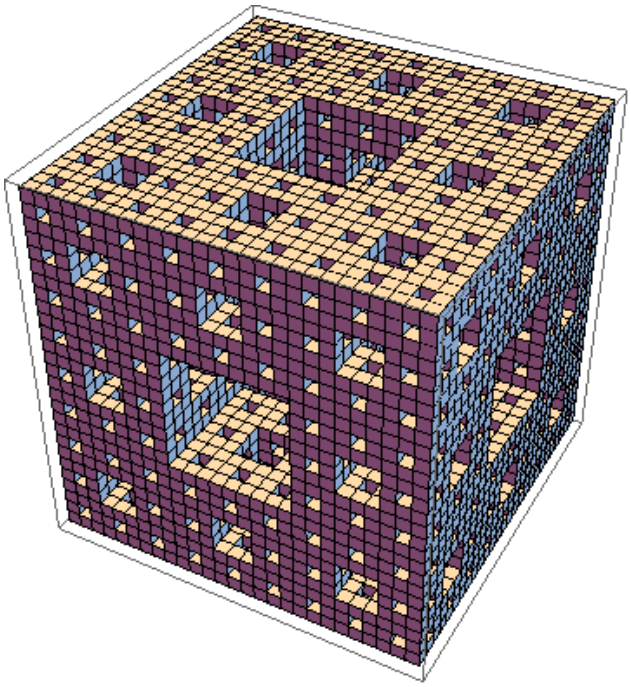


Figure 4.3: Third step of the construction of the Sierpinski sponge.

4.1.2 The fractal dimension

There is no doubt that traditional geometrical objects have an integer dimension that we denote as $D(0)$ here. For example, a fibre has dimension 1, a piece of paper has dimension 2, a book has dimension 3, moreover, a grain of dust has dimension 0. But how is the dimension $D(0)$ defined for a fractal is our interest in this section. Let define $r(\varepsilon)$ the number of d-dimension cells of volume ε^d that are needed to fully cover a typical object. Here d is called the embedding dimension, which is an

arbitrary integer so that $D(0) \leq d$. As we mentioned, $D(0)$ is the dimension of the object under discussion. For example, small cells of volume ε^3 ($d = 3$) can be used to cover not only a book, but also a piece of paper or a fibre. In order to cover these objects thoroughly, we need

$$r(\varepsilon) \sim \varepsilon^{D(0)} \tag{4.2}$$

small cells, where $D(0)$ adopts 3, 2, 1 accordingly. The dimension is determined by the growth rate of $r(\varepsilon)$ with respect to $\varepsilon \to 0$ as

$$D(0) = -\lim_{\varepsilon \to 0} \frac{\ln r(\varepsilon)}{\ln \varepsilon}. \tag{4.3}$$

This equation can be applied among traditional geometrical objects as well as fractals. However, the fractal dimension , also named the box dimension or capacity, will give a noninteger value in Eq.4.3.

For simple fractals with rigorous self-similarity and straightforward construction prescription, an easier method can be put forward to determine the fractal dimension. The fractal dimensions of the examples we discussed in 4.1.1 are all achievable by this method. In the case of the classical Cantor set, we can use intervals of length ε as the boxes. For the Sierpinski carpet, the boxes can be selected as squares of side length ε and area ε^2. For the Sierpinski sponge, the boxes can be selected to be cubes of side length ε and volume ε^3. Since the classical Cantor set, the Sierpinski carpet and the Sierpinski sponge are initialized as a unit interval, a unit square and a unit cube, we set an initial side length $\varepsilon_0 = 1$ and cut down its size by a factor a at each step of the construction. For all the three examples we mentioned above, we have $a = 1/3$. After the Nth iteration, the box size is diminished to

$$\varepsilon = a^N \varepsilon_0. \tag{4.4}$$

In each step, the increment factor of boxes that is required to fully cover the next stage of the fractal is G. So the total number of the boxes used to cover the fractal after the Nth step is

$$r(\varepsilon) = G^N r(\varepsilon_0). \tag{4.5}$$

Removing N from Eq.4.5 by Eq.4.4, we obtain for $N \to \infty$

$$N \sim \ln \varepsilon / \ln a, \tag{4.6}$$

$$r(\varepsilon) \sim \exp\left[(\ln G)(\ln \varepsilon)/\ln a\right] \tag{4.7}$$

Now we can reassemble Eq.4.3 and obtain the fractal dimension from the two parameters G and a as

$$D(0) = \lim_{\varepsilon \to 0} \frac{\ln r(\varepsilon)}{\ln \varepsilon} = -\frac{\ln G}{\ln a}. \tag{4.8}$$

Hence we can simply calculate the fractal dimensions for the classical Cantor set, the Sierpinski carpet and the Sierpinski sponge as

$$G = 2, D(0) = \frac{\ln 2}{\ln 3} = 0.6309..., \tag{4.9}$$

$$G = 8, D(0) = \frac{\ln 8}{\ln 3} = 1.8928..., \tag{4.10}$$

$$G = 20, D(0) = \frac{\ln 20}{\ln 3} = 2.7268..., \tag{4.11}$$

respectively.

4.2 Multifractals spectrum analysis

Multifractals are a generalization of fractals where a single fractal dimension is not enough to fully describe the set or system under consideration. The measure of the ith box of size ε is represented by $\mu_i(\varepsilon)$. Then one could obtain the singularity exponent of each box as

$$\alpha_i = \frac{\log \mu_i(\varepsilon)}{\log \varepsilon} \quad \text{or} \quad \mu_i(\varepsilon) \sim \varepsilon^{\alpha_i} \tag{4.12}$$

where α_i is also called Holder index. The number $r_\alpha(\varepsilon)$of boxes with the same Holder index α scales with the fractal dimension $f(\alpha)$ of the subsets with the same α as

$$r_\alpha(\varepsilon) \sim \varepsilon^{f(\alpha)}. \tag{4.13}$$

The function of $f(\alpha)$ is called multifractal spectrum. The shape of $f(\alpha)$ converges to a point for non- and mono-fractals while the spectra of multifractal datasets are humped over a relatively large range of α. A bunch of papers in the field of multifractal analysis evolve around finding the multifractal spectrum $f(\alpha)$ from a given data set. Here we will introduce three most common methods.

4.2.1 Box-counting method

In the original box-counting method [63], one divides the d-dimensional shpase space into boxes of equal size ε and compute a normalized measure $\mu_i(\varepsilon)$ for each box. The number of boxes with non-zero measure is $r(\varepsilon)$. One can then introduce a partition function

$$Z(\varepsilon,q)=\sum_{i=1}^{r(\varepsilon)}\mu_i^q(\varepsilon)\sim\varepsilon^{\tau(q)} \tag{4.14}$$

which scales with the scaling function $\tau(q)$ for $-\infty\leq q\leq\infty$. When $q>0$ the partition function emphasizes the effect of the regions with higher $\mu_i(\varepsilon)$. On the contrary, the partition function emphasizes the effect of the region with lower $\mu_i(\varepsilon)$ when $q<0$.

The relation between the multifractal spectrum $f(\alpha)$ and the scaling function $\tau(q)$ can be obtained by the substitution of Eq.4.12 and Eq.4.13 into the expression for the partition function in Eq.4.14:

$$Z(\varepsilon,q)=\sum_{i=1}^{r(\varepsilon)}\varepsilon^{-f(\alpha)}\varepsilon^{\alpha q}=\sum_{i=1}^{r(\varepsilon)}\varepsilon^{\alpha q-f(\alpha)}=\varepsilon^{\tau(q)}; \tag{4.15}$$

equivalently,

$$\sum_{i=1}^{N(\varepsilon)}\varepsilon^{\alpha q-f(\alpha)-\tau(q)}=1 \tag{4.16}$$

By using the stationary phase method it can be found that

$$\tau(q)=q\alpha(q)-f(\alpha(q)), \tag{4.17}$$

where $\alpha(q)$ is value for which the expression $q\alpha-f(\alpha)$ is maximized for a given q. This condition is equivalently interpreted by $f'(\alpha)=q$ under the assumption that

$f(\alpha)$ is differentiable. Substituting $f'(\alpha) = q$ back into Eq.4.17 gives $\tau'(q) = \alpha$. Hence Eq.4.17 is the Legendre transform of $f(\alpha)$, for this reason it contains the same information. The generalized dimension, which is also known as the Rényi dimension of order q, is defined as

$$D(q) = \lim_{s\to 0} \frac{1}{q-1} \frac{\ln Z(\varepsilon, q)}{\ln \varepsilon} = \frac{\tau(q)}{q-1}. \tag{4.18}$$

which will be discussed in detail later.

4.2.2 Wavelet transform modulus maxima method

Another method to calculate $f(\alpha)$ is the wavelet transform modulus maxima method [64][65], which comes from the original wavelet transform approach [119] in signal theory. The wavelet transform of a discrete time series $x(i), i = 1, ..., n$ is defined as

$$T_\psi(t,s) = \frac{1}{s} \sum_{i=1}^{n} \psi\left(\frac{i-t}{s}\right) x(i) \tag{4.19}$$

where ψ is called the mother wavelet which is shifted by a position parameter t and stretched by a scale parameter s [120]. The wavelet coefficients depend on both the position and the scale and the transform serves as a tool for the decomposition of the time series into the time-scale plane. The mother wavelet ψ, which should have zero mean, are most frequently chosen to be the mth derivative of a Gaussian

$$\psi^{(m)}(x) = \frac{d^m}{dx^m}\left(e^{-x^2/2}\right). \tag{4.20}$$

After picking out the local maxima of T_ψ, one can calculates the partition function for the moduli of the maxima

$$Z(s,q) = \sum_{j=1}^{j_{\max}} |T_\psi(t_j, s)|^q \tag{4.21}$$

so that $|T_\psi(t_j - 1, s)| < |T_\psi(t_j, s)|$ and $|T_\psi(t_j + 1, s)| \leq |T_\psi(t_j, s)|$ for $j = 1, ..., j_{\max}$ where t_j denotes the positions of the local maxima of $|T_\psi(t_j, s)|$. For any given scale s, if a maximum at a certain time position t_j is identified to be smaller than a

maximum at a near position $t_j^{'} \approx t_j$ for a lower scale $s' < s$, then $T_\psi(t_j, s)$ is replaced by $T_\psi(t_j^{'}, s')$.

If the time series is supposed to have fractal properties, the partition function is expected to scale with the scaling function τ as in Eq.4.14

$$Z(s,q) \sim s^{\tau(q)}. \tag{4.22}$$

The singularity spectrum $f(\alpha)$ is related to $\tau(q)$ by the same formulas as in the box-counting method [120]:

$$\alpha = \tau'(q) \quad \text{and} \quad f(\alpha) = q\alpha - \tau(q) \tag{4.23}$$

4.2.3 Multifractal detrended fluctuation analysis

The multifractal detrended fluctuation analysis [61] is an alternative method that has become a commonly used technique in determination of the multifractal spectrum. It is a generalized version of the detrended fluctuation analysis [58] that serves as a tool for identifying the scaling properties of mono-fractal time series. The procedure for determining the multifractality of a given time series $x(i), i = 1, ..., n$ is outlined as follows [120].

We first calculate the cumulative profile $Y(j)$ by

$$Y(j) = \sum_{i=1}^{j} (x(i) - \langle x \rangle), \quad j = 1, ..., n \tag{4.24}$$

One first divides the n elements of Y into N_s non-overlapping segments of length s starting from the beginning of the time series and then one divides another N_s non-overlapping segments of length s starting from the end. Since n/s may not be an integer, such $2N_s$ segments can make sure that none of the data at the end or at the beginning of $Y(j)$ is ignored. Each segment, labelled with $v (v = 1, ..., 2N_s)$, has a local tendency $P_v^{(l)}$, which is obtained by fitting an lth order polynomial. The square average of the deviation from the local tendency for the vth segments and for the

segment lengths s is calculated by

$$F^2(v,s) = \frac{1}{s}\sum_{j=1}^{s}\{Y[(v-1)s+j] - P_v^{(l)}(j)\}^2 \tag{4.25}$$

This average is calculated for all the segments and for all the possible scale s. The qth-order fluctuation function over all the segments is obtained by:

$$F_q(s) = \left(\frac{1}{2N_s}\sum_{j=1}^{2N_s}[F^2(v,s)]^{q/2}\right)^{1/q} \tag{4.26}$$

For a time series with fractal properties, the fluctuation function $F_q(s)$ is expected to scale like

$$F_q(s) \approx s^{H(q)}, \tag{4.27}$$

where $H(q)$ is called the generalized Hurst exponent. For a mono-fractal series, $H(q)$=const and for a series with multifractal structure, $H(q)$ is a decreasing function of q. The singularity spectrum of the Holder exponents $f(\alpha)$ which is our interest here is connected to the generalized Hurst exponent by the following formulas [120]:

$$\alpha = H(q) + qH'(q), \quad f(\alpha) = q[\alpha - H(q)] + 1. \tag{4.28}$$

4.3 Symbolic dynamics method and Rényi entropy

The generalized dimension, which is the so-called the Rényi dimension, is an important index of the multifractal properties for a given time series. The key point in the implementation of this generalized dimension is to identify a normalized measure(probability measure) for a given support. In this thesis we use the symbolic dynamics technique to analyse the multifractal property of the share price dynamics in terms of the Rényi dimension and the Rényi entropy.

4.3.1 Symbolic dynamics

The method of symbolic dynamics offers a coarse-grained way to trace the time evolution of a signal such as the share price trajectories [66]. The first step is to construct the symbol sequence from the dynamical system under consideration.

Symbol sequence can only be generated by a proper partition of the phase space, i.e. the set of possible values that are relevant for the given times series. For share price, we will often just use 2 symbols u and d for up or down movements(more details in chapter 6). The easiest way is to choose d-dimensional cubes of equal size, the so called *boxes*, to partition a general d-dimensional phase space X. The *size* of the cubes is determined by their length ε of the side. The entire phase space X is covered by a total number of R boxes. The boxes do not overlap and they are labeled by an index i that runs from 1 to R.

However, a more flexible way is to use *cells* of variable size rather than *boxes* to partition the phase space. The cells are also disjoint, meaning that they do not have a point in common. Unlike the boxes, the cells could have different sizes and shapes. The cells A_i are also labeled by an index i. In general, if we have the following two conditions:

$$A_i \cap A_j \tag{4.29}$$

and

$$\bigcup_{i=1}^{R} A_i = X \tag{4.30}$$

we call it a partition of X into cells. If the cells are particularly of equal size, we call them boxes.

Suppose we have a sequence of data points, say a time series, the first data point x_0 is in the cell labeled by i_0, the second data point is in the cell i_1 and so on. In this way, the data time series is attributed to a symbol sequence

$$i_0, i_1, i_2, ..., i_n, ... \tag{4.31}$$

Such a sequence of symbols i_n is called a 'symbol sequence', and the word 'symbolic dynamics' describes the mapping from the phase space to the symbol space.

The trajectory is recorded in a coarse-grained way by this technique. Then if the size of the boxes, or cells, is relatively large, the description of the trajectory will be rather rough, for the fact that many trajectories are represented by one symbol sequence. In order to get a detailed description, the size of the boxes has to be taken relatively small. We will look at the case when ε is going to 0, meaning that the total number of the boxes R goes to infinity. Moreover, to obtain a rather complete description by the symbol sequence, we are interested in the length of the sequence N going to infinity. We will investigate the situation of $\varepsilon \to 0$ and $N \to \infty$ in the following chapters.

4.3.2 Information measures

4.3.2.1 Bit-numbers

Consider a bit-storage unit of a computer as a trigger switch with two possible positions, a number A of such switches can generate

$$N = 2^A \tag{4.32}$$

different states, also named N different bit patterns. Each of the bit patterns corresponds to a particular integer in the range of 0 to $N-1$. This can be explained by defining such an integer m as the decimal representation of a binary number of length A:

$$m = \sum_{k=0}^{A-1} s_k 2^k \tag{4.33}$$

where s_k is 0 or 1, representing one of the two possible positions of the kth switch. From Eq.4.32 we know that, to select one of N events, we need

$$A = \frac{\ln N}{\ln 2} \tag{4.34}$$

bits. If we use $\ln 2$ as a unit, then the 'bit-number' is defined as

$$b = \ln N. \tag{4.35}$$

Now, suppose there is a sample set of N 'elementary' events α of equal probability, where N is relatively large. We divide this set into R disjoint subsets. Each subset i is regarded as a compound event, containing N_i 'elementary' events. That is,

$$N = \sum_{i=1}^{R} N_i. \tag{4.36}$$

Then, the probability of the compound event i, namely, the probability of finding an elementary event α lying in the subset i, is given by

$$p_i = N_i/N. \tag{4.37}$$

We denote the set of all p_i as the probability distribution p. In the following, we essentially follow C.Beck and F.Schögl [66].

To select an elementary event from the large sample set, it always requires a minimum bit-number of $\ln N$, no matter whether we first select the subset i which α is in and then select α out of this subset, or we select α directly from the large sample set. Also, the minimum bit-number required for selecting an elementary event α out of the subset i it lies in is $\ln N_i$. Now, let denote the bit-number needed to select the subset i by b_i. Then we can construct the formula

$$b_i + \ln N_i = \ln N, \tag{4.38}$$

meaning that the sum of the bit-number needed for selecting the subset i and the bit-number needed for selecting α out of the subset i, is equivalent to the bit-number required for selecting α directly from the large sample set. As a result, conforming Eq.4.37 and Eq.4.38, we obtain that

$$b_i = -\ln p_i \tag{4.39}$$

is the bit-number lacked for an observer to identify whether the event i will definitely occur, under the situation that he only knows the probability p_i of this event.

4.3.2.2 The Shannon information

Since b_i measures a lack of knowledge, $-b_i$ measures a knowledge. The mean value of $-b_i$ we can get, after a long series of observations where each event i happens with probability p_i, is

$$I(b) = \sum_{i=1}^{R} p_i \ln p_i. \tag{4.40}$$

This function of the distribution p is known as the 'Shannon information' [121][122]. It is regarded as a measure for the knowledge of the observer about the question of which event of the sample set is to be expected, if he knows only the distribution p.Note that as $0 \leq p_i \leq 1$, $I(p) \leq 0$, and it reaches its maximum value 0 with an optimum knowledge. That means $I(p) = 0$ when only one of the R microstates occurs. On the other hand, $I(p)$ takes on its minimum value $-\ln N$ when all events $i = 1, ..., R$ are distributed with a uniform distribution, namely,

$$p_i = 1/R. \tag{4.41}$$

Here the knowledge reaches minimum since no event is distinguished.

The negative Shannon information measure

$$S(p) = -I(p) \tag{4.42}$$

is called the Shannon entropy, which is always positive. It measures the lack of knowledge regarding the above circumstances.

4.3.2.3 The Rényi information

Now we introduce another information measure , the so called Rényi information [123][124],

$$I_\beta(p) = \frac{1}{\beta - 1} \ln \sum_{i=1}^{r} (p_i)^\beta, \tag{4.43}$$

where the parameter β is an arbitrary real number; r is the number of all nonempty states i, in other word, i runs through all the nonzero probabilities and all the events with probability zero are excluded.

One important feature we can capture in Eq.4.43 is that, when $\beta = 0$,

$$I_0(p) = -\ln \sum_{i=1}^{r} 1 = -\ln r. \tag{4.44}$$

That means $I_0(p)$ decreases in a logarithmical way with the number of all the realised events.

A power expansion with respect to $\varepsilon = \beta - 1$ shows that for $\beta \to 1$ one has in first order of ε

$$\begin{aligned} \sum_{i=1}^{r} p_i^{1+\varepsilon} &= \sum_{i=1}^{r} p_i \exp(\varepsilon \ln p_i) \approx \sum_{i=1}^{r} p_i (1 + \varepsilon \ln p_i) \\ &= 1 + \varepsilon \sum_{i=1}^{r} p_i \ln p_i. \end{aligned} \tag{4.45}$$

Hence,

$$\begin{aligned} \lim_{\varepsilon \to} I_{1+\varepsilon}(p) &= \lim_{\varepsilon \to} \frac{1}{\varepsilon} \ln(1 + \varepsilon \sum_{i=1}^{r} p_i \ln p_i) \\ &= \sum_{i=1}^{r} p_i \ln p_i \end{aligned} \tag{4.46}$$

or

$$\lim_{\beta \to 1} I_\beta(p) = I(p). \tag{4.47}$$

Thus I_1 is identical with the Shannon information. We therefore regard the Rényi information as a generalisation of the Shannon information. The Shannon information is contained as a special case $\beta = 1$.

4.3.3 Multifractals and Rényi dimensions

A fractal with a probability measure associated with its support is called a multifractal. The information about the probability distribution can be statistically described by the Rényi dimensions. We divide the d-dimensional phase space into boxes of equal size. The d-dimensional boxes are cubes with side length ε. We exclude the boxes with zero probability and label those boxed with nonzero probability by $i = 1, 2, ..., r$. The number of the boxes with nonzero probability r, is distinguished

from the total number of boxes $R \sim \varepsilon^{-d}$. Other than keeping the number R of boxes finite, we are interested in the limit behaviour of box size $\varepsilon \to 0$, which implies $R \sim \varepsilon^{-d} \to \infty$. When ε approaches 0, the Rényi information defined as Eq.4.43 diverges. However, we can re-introduce the Rényi dimension [124] $D(\beta)$ as the changing of the Rényi information with respect to $\varepsilon \to 0$:

$$D(q) = \lim_{\varepsilon \to 0} \frac{I_q(p)}{\ln \varepsilon} = \lim_{\varepsilon \to 0} \frac{1}{\ln \varepsilon} \frac{1}{q-1} \ln \sum_{i=1}^{r} p_i^q. \tag{4.48}$$

The Rényi dimension is a decreasing function of q and will stays finite for $\varepsilon \to 0$. Depending on the values of q, the Rényi dimension reveals different features of the system. For $q = 0$, the negative Rényi information becomes equal to the logarithm of the number $r(\varepsilon)$ of nonempty boxes. This is the smallest number of boxes necessary to cover the entire fractal. Hence,

$$D(0) = -\lim_{\varepsilon \to 0} \frac{\ln r(\varepsilon)}{\ln \varepsilon} \tag{4.49}$$

is the fractal dimension which has already been discussed in the previous section. For $q \to 1$, according to Eq.4.47, the Rényi information becomes the Shannon information. Therefore

$$D(1) = \lim_{\varepsilon \to 0} \frac{1}{\ln \varepsilon} \ln \sum_{i=1}^{r} p_i \ln p_i, \tag{4.50}$$

which describes how the Shannon information flows with the refinement of the box size, is called the *information dimension*. In general, q determines how much weight is assigned to the probability measure p, When $q > 0$, $D(q)$ is mostly influenced by the boxes with high probability while when $q < 0$, the boxes with lower probability have stronger effects on $D(q)$.

Similar as the $f(\alpha)$ spectrum, the spectrum of $D(q)$ vs. q is also regarded as an indicator of multifractality. For a signal with non- or mono-fractal property, the shape of $D(q)$ will be constant over different q's. While a multifractal signal will have different Rényi dimensions for different q.

In [125], C. Beck put forward a rigorous restriction on the Rényi dimension,

where a general upper and lower bound of the Rényi dimension was proved:

$$\frac{q'-1}{q'}D(q') \geq \frac{q-1}{q}D(q) \qquad \text{for } q' > q, q'q > 0. \tag{4.51}$$

This bound is valid for arbitrary probability distributions. If we substitute $+\infty$ and $-\infty$ for q in Eq.6.13, we can obtain an upper bound

$$D(q) \leq \frac{q}{q-1}D(+\infty) \qquad \text{for } q > 1 \tag{4.52}$$

and a lower bound

$$D(q) \geq \frac{q}{q-1}D(-\infty) \qquad \text{for } q < 0. \tag{4.53}$$

4.3.4 Statistics of dynamical symbol sequence and Rényi entropies

Whereas in the above sections we have discussed the static properties of dynamical systems, now we look at the dynamical features of the symbol sequences. First, we partition the entire phase space A into subsets A_i,

$$\{A\} = \{A_1, A_2, ..., A_R\} \tag{4.54}$$

where A_i are either boxes of equal size ε, or cells of variable size and R is the total number of the cells.

Recall from subsection 4.3.1, each trajectory $x_0, x_1, ...$ is attributed to a symbol sequence $i_0, i_1, i_2, ...$ by the following condition:

$$x_n \in A_{i_n}. \tag{4.55}$$

Consider a finite symbol sequence $i_0, i_1, ..., i_{N-1}$ of length N, the number of all allowed sequences is denoted by $\omega(N)$. The probability of the jth symbol sequence is represented as

$$p_j^{(N)} = p(i_0, i_1, ..., i_{N-1}), \qquad j = 1, 2, ..., \omega(N). \tag{4.56}$$

Compared to the Rényi dimension we have studied above, where we partitioned the phase space using the boxes of equal size ε and looked at the scaling behaviour of this static partition function in the limit $\varepsilon \to 0$, here we use a finite partition $\{A\}$ whose cells may have different finite size. The quantity named the 'Rényi entropy of order q',

$$K(q) = \lim_{N\to\infty} \frac{-I_q^{(N)}}{N} = \lim_{N\to\infty} \frac{1}{1-q}\frac{1}{N} \ln \sum_{j=1}^{\omega} (p_j^{(N)})^q, \tag{4.57}$$

describes the scaling behaviour of the dynamical partition function with respect to the time N rather than with respect to the box size ε. It is a measure of the changing of the negative Rényi information for time $N \to \infty$.

The Rényi entropy is equal to the 'Kolmogorov-Sinai entropy' [126][127] when $\beta \to 1$. It describes the average information loss with respect to the time N.

Now suppose that each box of scale ε that has non-zero probability in the phase space represents a symbol sequence, and the probability of this symbol sequence is equal to the probability measure of that box, then the Rényi dimension and the Rényi entropy are clearly connected. Jizba and Arimitsu discussed how these two measures are related to each other in detail in the paper "The world according to Rényi: thermodynamics of multifractal systems" [74]. In the later experimental chapter on multifractal analysis of the share prices, we will discuss the multifractal properties of the considered data by looking at the spectra of both measures, which are connected with the help of the symbolic dynamics technique.

Chapter 5

Superstatistics in Finance

We now come to the central results off this thesis. We will demonstrate that share price returns on different time scales can be well modeled by a superstatistical dynamics. Here we provide an investigation which type of superstatistics is most suitable to properly describe share price dynamics on various time scales. It is shown that while χ^2-superstatistics works well on a time scale of days, on a much smaller time scale of minutes the price changes are better described by lognormal superstatistics. The system dynamics thus exhibits a transition from lognormal to χ^2 superstatistics as a function of time scale. We discuss a more general model interpolating between both statistics which fits the observed data very well. We also present results on correlation functions of the extracted superstatistical volatility parameter, which exhibits exponential decay for returns on large time scales, whereas for returns on small time scales there are long-range correlations and power-law decay.

In this chapter, we will carefully analyse for various data sets of historical share prices which type of superstatistics is best suited to model the dynamics. While Tsallis statistics ($= q$-statistics) is known to be equivalent to χ^2 superstatistics [37][128], there are other types of superstatistics, such as lognormal superstatistics and inverse χ^2 superstatistics [39], which are known to be different from q-statistics (though all these different statistics generate similar distributions if the variance of the fluctuations in β is small [37]). We show that in our analysis χ^2-superstatistics appears best suitable to describe the daily price changes, whereas

on much smaller time scales of minutes lognormal superstatistics seems preferable. We analyse the relevant time scale of the changes in the superstatistical parameter β and present results for the decay of correlations in β. For small return time scales, correlation functions exhibit power law decay and there are long memory effects. In the final section, we develop a synthetic stochastic model that fits the data well. This is kind of a hybrid model interpolating between lognormal and χ^2-superstatistics.

5.1 Superstatistics of log-returns of share prices on a large time scale

Non-equilibrium system dynamics can often be regarded as as superposition of a local equilibrium dynamics and a slowly fluctuating process of some variance variable β [37]. These types of 'superstatistical' nonequilibrium models are also useful for financial time series [112][50]. In this chapter, the empirical data we use as an example is the historical stock prices of Alcoa Inc(AA), which is an American company that engages in the production and management of primary aluminium, fabricated aluminium and alumina. We have looked at shares of many other companies as well (see Tab.5.1) in section 5.4, with similar results. Our data set covers the period January 1998 to May 2013. The source data are minutely recorded share prices. The last minute price quoted each day is used as the daily closing price. We study the log return R_i denoted by

$$R_i = \log\left(\frac{S_{i+1}}{S_i}\right) \tag{5.1}$$

where $i = 0,1,2,...,N$; S_i and S_{i+1} are two successive daily closing prices. We consider the normalised log returns

$$u_i = \frac{R_i - \langle R\rangle}{\sqrt{\langle R^2\rangle - \langle R\rangle^2}} \tag{5.2}$$

which have been rescaled to have variance 1. The symbol $\langle\cdots\rangle$ denotes the long-time average.

From the simplest superstatistics model point of view, the entire time series of stock prices can be divided into N smaller time slices T. We will later determine the optimal window size T. Within each T, the financial volatility β is approximately temporarily constant and the log return of the stock price is assumed to be Gaussian distributed. β has some probability distribution $f(\beta)$ to take a particular value in a given slice. The conditional probability $p(u|\beta)$ is

$$p(u|\beta) = \sqrt{\frac{\beta}{2\pi}} \exp\left(-\frac{1}{2}\beta u^2\right) \tag{5.3}$$

and the marginal probability distribution of u under long time observation is the average over local Gaussians weighted with the probability density $f(\beta)$

$$p(u) = \int p(u|\beta) f(\beta) d\beta. \tag{5.4}$$

The integration over β yields non-Gaussian behaviour with fat tails.

We now describe our technique to obtain the optimal window size T for a given time series. Firstly we split the time series into

$$n = \lfloor \frac{N}{\Delta t} \rfloor \tag{5.5}$$

equal intervals, where Δt is the window size taken and N the total number of data points. Generally the kurtosis of a random variable u is defined as

$$\kappa = \frac{\langle u^4 \rangle}{\langle u^2 \rangle^2} \tag{5.6}$$

and it is equal to 3 for a Gaussian distribution of arbitrary variance. For a given window size Δt, the kurtosis in the jth window is given by

$$\kappa_{\Delta t}(j) = \frac{\frac{1}{\Delta t}\sum_{i=(j-1)\Delta t+1}^{j\Delta t} u_i^4}{\left(\frac{1}{\Delta t}\sum_{i=(j-1)\Delta t+1}^{j\Delta t} u_i^2\right)^2}, \tag{5.7}$$

where $j = 1, 2, ..., n$. When we have all the values of kurtosis for all windows, we

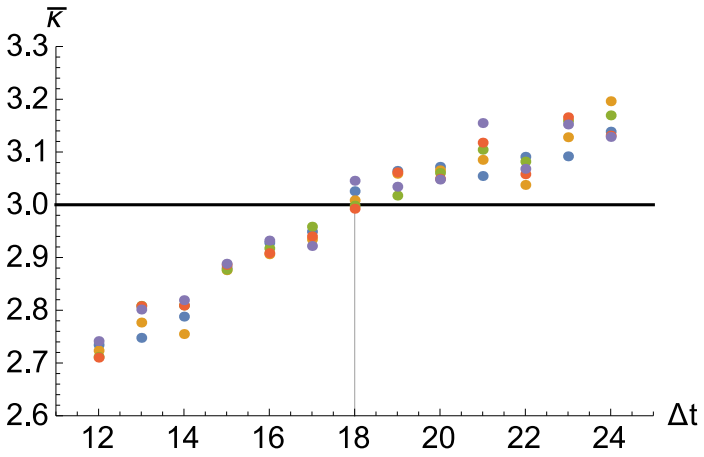


Figure 5.1: Determination of the optimal window size for the Alcoa share price data. The intersection with the line kurtosis $\bar{\kappa} = 3$ yields $T = 18 \pm 0.5$. The various values of $\bar{\kappa}$ for a given Δt (indicated by different colors in the online version) are obtained for different translational shifts of the sliding windows. The scattering of the data can be used to estimate the standard deviation as $\delta\bar{\kappa} \sim 0.03$.

can calculate an average kurtosis of the n windows as

$$\bar{\kappa}_{\Delta t} = \frac{1}{n}\sum_{j=1}^{n} \kappa_{\Delta t}(j). \tag{5.8}$$

The aim is to achieve an optimum window size such that for a given data set the distribution in each window is as close as possible to a Gaussian, but with varying variance. For this purpose the optimal window size T should satisfy the condition

$$\bar{\kappa}_{\Delta t} = 3. \tag{5.9}$$

Fig. 5.1 shows how the average kurtosis changes with the window size. We obtain from condition (5.9) the optimal window size 18 ± 0.5 for this example. The result makes financial sense. 18 trading days correspond to a time scale of about 3-4 weeks. It is a typical time scale where market volatility changes, due to events such as changes in the confidence in the future economic development, anticipated interest changes, and so on. See also [129] for related work.

With the given optimal window size, we can now calculate the local volatility

parameter β in each time interval as

$$\beta_k = \frac{1}{\frac{1}{T-1}\sum_{i=(k-1)T+1}^{kT}(u_i - \bar{u_i})^2} \tag{5.10}$$

where $k = 1,2,...,n$. Note that the variance of u in each window is β^{-1}. One can then plot a histogram of the β_k and fit it with some suitable model distribution.

Here we will consider three distributions to be compared with our experimental distribution of β, which were previously advocated in [39]. The first one is the χ^2 -distribution for which $f(\beta)$ is given by

$$f_1(\beta) = \frac{1}{\Gamma\left(\frac{d_1}{2}\right)}\left(\frac{d_1}{2\beta_0}\right)^{d_1/2}\beta^{d_1/2-1}e^{-d_1\beta/2\beta_0}. \tag{5.11}$$

The second one is the inverse χ^2-distribution where

$$f_2(\beta) = \frac{\beta_0}{\Gamma\left(\frac{d_2}{2}\right)}\left(\frac{d_2\beta_0}{2}\right)^{d_2/2}\beta^{-d_2/2-2}e^{-d_2\beta_0/2\beta}. \tag{5.12}$$

The third distribution that will be tested is the log-normal distribution for which the probability density function is given by

$$f_3(\beta) = \frac{1}{\sqrt{2\pi}s\beta}\exp\left(\frac{-(\ln\beta-\mu)^2}{2s^2}\right) \tag{5.13}$$

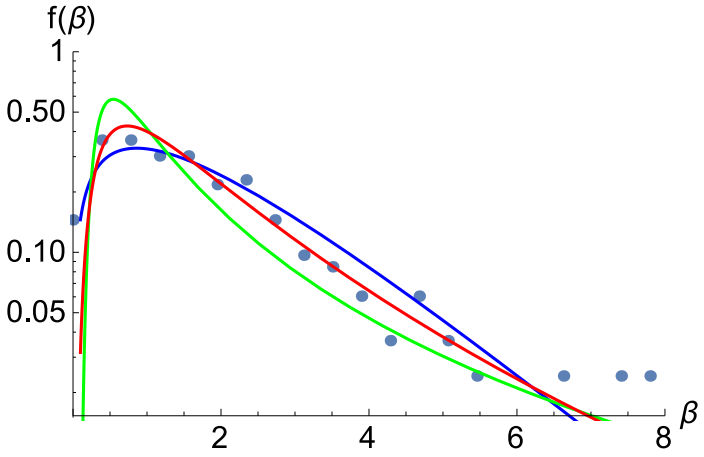
where

$$\mu = \ln\beta_0 - \frac{s^2}{2}. \tag{5.14}$$

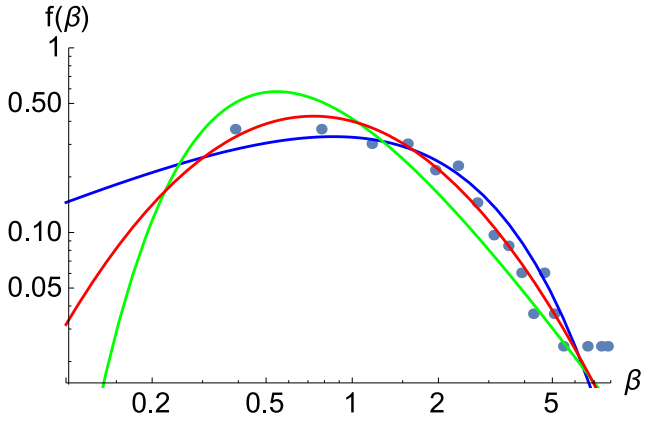
The β_0 in Eq. (5.14), (5.11), (5.12) is the mean value of β, given by

$$\beta_0 = \langle\beta\rangle = \frac{1}{n}\sum_{k=1}^{n}\beta_k, \tag{5.15}$$

and d_1, d_2, s are parameters. Lognormal superstatistics often occurs for complex systems described by a cascading dynamics [130], whereas χ^2 and inverse χ^2 superstatistics are more common for additive degrees of freedom contributing to a



(a)



(b)

Figure 5.2: Best possible fits that can be achieved for the distribution of the volatility β of Alcoa shares (plotted by dots), in a log-linear (top) and double logarithmic plot (bottom). Blue: χ^2 distribution $f_1(\beta)$ with $d_1 = 1.51, \beta_0 = 2.19$, Green: inverse χ^2 distribution $f_2(\beta)$ with $d_2 = 0.45, \beta_0 = 2.19$, Red: lognormal distribution $f_3(\beta)$ with $s = 0.87, \mu = 0.45$,.

fluctuating temperature or inverse temperature [39].

We have fitted our experimental histograms $f(\beta)$ with the above distributions. Given β_0, we used the "method of moments" method to find the optimum d_1, d_2 and s of Eq. (5.11), (5.12), (5.13) in order to obtain the optimum fit to our observed $f(\beta)$.

It can be seen in Fig. 5.2 that lognormal, χ^2- and inverse χ^2 superstatistics all yield a more or less decent fit, though inverse χ^2-superstatistics seems less favorable.

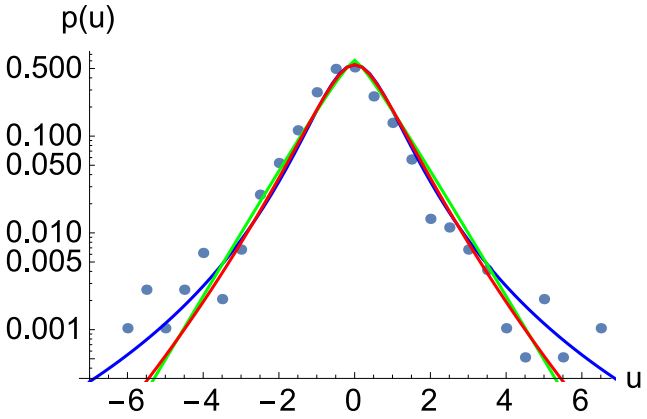


Figure 5.3: Comparison of the histogram of u (plotted by dots) with the 3 types of superstatistics, integrated with the same parameters as in Fig. 5.2. Blue: χ^2 Superstatistics $p_1(u)$, Green: inverse χ^2 Superstatistics $p_2(u)$, Red: lognormal Superstatistics $p_3(u)$.

Still for consistency we also need to check the validity of Eq. (5.4). We thus also compare the original histogram of returns u with the following integrals where the parameters take the same values as in Fig.5.2:

$$p_i(u) = \int \sqrt{\frac{\beta}{2\pi}} \exp\left(-\frac{1}{2}\beta u^2\right) f_i(\beta) d\beta \quad i = 1,2,3 \tag{5.16}$$

As shown in Fig. 5.3, for the integrated densities χ^2 superstatistics seems to fit better to the probability density of u compared with lognormal superstatistics and inverse χ^2 superstatistics.

Thus, if *independent* variation of the volatility parameter in each interval is assumed, then the data clearly point to χ^2 superstatistics, equivalent to Tsallis statistics [128]. On the other hand, independence of β_k may not always be a good approximation. There can be strong correlations of the volatility parameter β_k, and variations of the time scales where it is approximately constant. In that case more complicated dynamics arise, and one could then possibly get a better fit for the integrated distributions $p(u)$ if other effective parameters are used. For this reason, we also allowed the fitting parameters for $p_1(u), p_2(u), p_3(u)$ to take on other possible values. The result of this 'amended superstatistics' is shown in Fig. 5.4.

After the adjustment, we find in Fig.5.4 that in fact *all three* superstatistics can

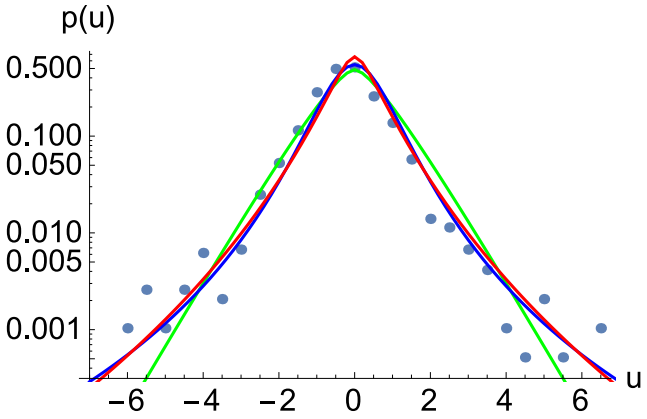


Figure 5.4: Amended Superstatistics Blue: χ^2 Superstatistics $p_1(u)$ with $d_1 = 1.51, \beta_0 = 2.19$, Green: inverse χ^2 Superstatistics $p_2(u)$ with $d_2 = 1.2, \beta_0 = 2.19$, Red: lognormal Superstatistics $p_3(u)$ with $s = 1.2, \mu = 0.65$.

describe $p(u)$ quite well. To distinguish between them, one would need much more data so that the tail behaviour would be clearer. In practice, more data are available if one considers price changes on much smaller time scales than days. This will be done in the next section.

5.2 Short time scales

Let us extend our analysis to returns on much smaller time scales. Let s_i be the stock price for every recorded minute, in our example chosen as that of Alcoa Inc(AA). The total number of data points is about 150 million. We look at the returns

$$r_i = \log\left(\frac{s_{i+\tau}}{s_i}\right) \tag{5.17}$$

where τ is an integer in units of minutes. The log returns are again normalized to variance 1:

$$u_i = \frac{r_i - \langle r \rangle}{\sqrt{\langle r^2 \rangle - \langle r \rangle^2}} \tag{5.18}$$

There is one small technical problem for these types of data, as the returns are not given overnight but only during normal working hours. This can lead to big overnight jumps and affect the analysis. For this reason, if $s_{i+\tau}$ and s_i are from two successive trading days, we removed the corresponding $\log\left(\frac{s_{i+\tau}}{s_i}\right)$. $\tau = 1$ means

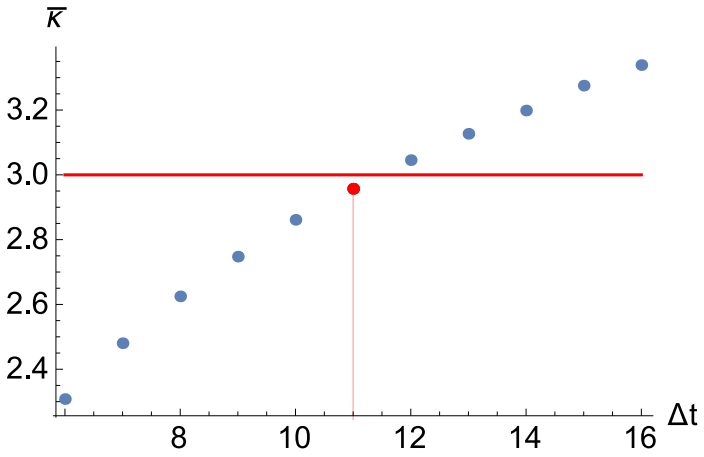


Figure 5.5: Determination of the optimum window size for the 1-minute data of Alcoa. The intersection with the line $kurtosis = 3$ yields $T \approx 11$.

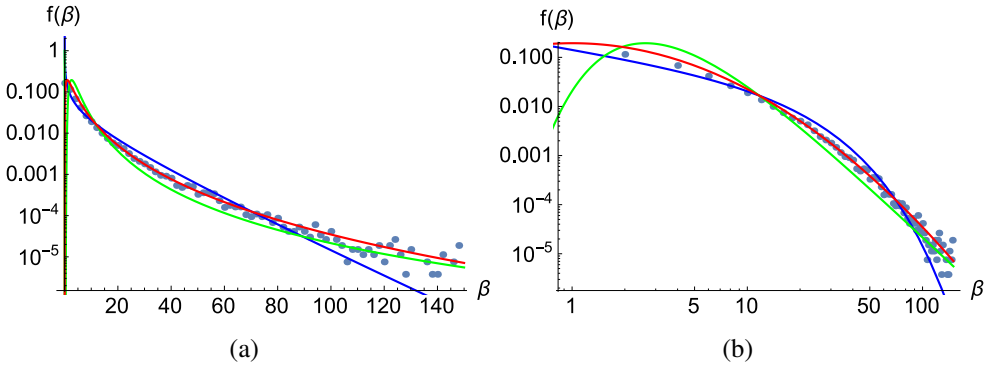


Figure 5.6: Best fits that can be achieved for the distribution of the short-scale volatility parameter β (time scale of returns: 1 minute). Blue: χ^2 distribution $f_1(\beta)$ with $d_1 = 0.13, \beta_0 = 6.33$, Green: inverse χ^2 distribution $f_2(\beta)$ with $d_2 = 2.83, \beta_0 = 6.33$, Red: lognormal distribution $f_3(\beta)$ with $s = 1.11, \mu = 1.23$, top: log-linear plot, bottom: double logarithmic plot.

the log return is extracted every minute. Again we determined the optimal window size, using the same technique as in the previous section. We obtain $T \approx 11$ (see Fig. 5.5).

Again this time scale of about 11 minutes makes sense. It is a typical time scale on which new relevant information becomes availbale to the traders, leading to changes in the small-scale volatility. It also coincides with typical time scales on which observed correlations in short-term returns start to decay [131]. Our results of fitting the three types of superstatistics are shown in Fig. 5.6-5.8.

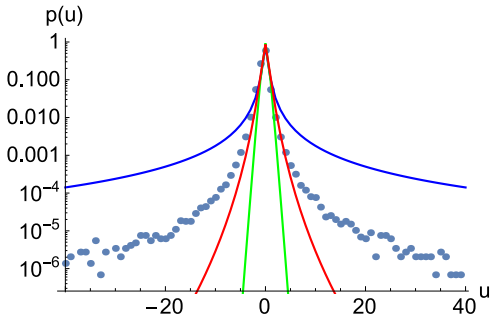


Figure 5.7: Comparison of histogram of u (plotted by dots) with the integrated superstatistics distributions, using the same parameters as in Fig.5.6. Blue: χ^2 Superstatistics $p_1(u)$, Green: inverse χ^2 Superstatistics $p_2(u)$, Red: lognormal Superstatistics $p_3(u)$. None of the curves is a good fit, indicating the presence of strong correlations for the volatility parameter β_k.

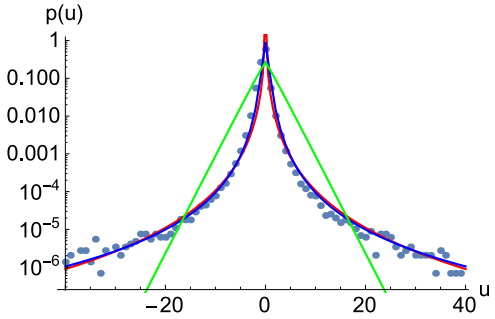


Figure 5.8: Amended Superstatistics Blue: χ^2 Superstatistics $p_1(u)$ with $d_1 = 0.36, \beta_0 = 6.33$, Red: lognormal Superstatistics $p_3(u)$ with $s = 2.7, \mu = 3.9$, Green: inverse χ^2-superstatistics ($d_2 = 0.2, \beta_0 = 1.8$).

As can be seen in Fig. 5.6, the lognormal distribution is by far best fit of $f(\beta)$ if the time scale is 1 minute.

Fig. 5.7 shows a clear difference as compared to the daily data in Fig. 5.3: The integrated formula now does *not* give good fits to $p(u)$. The reason is that the β_k on a time scale of minutes are not anymore statistically independent, hence random sampling of Gaussians with different variance is not appropriate anymore.

After the free adjustment in the parameters of $p_1(u), p_2(u), p_3(u)$, again both χ^2 and lognormal superstatistics can provide good fits of $p(u)$. See Fig. 5.8.

If one does not allow for parameter amendments, then we can conclude that there is a transition from χ^2 to lognormal superstatistics when the time scale changes from 1 day to 1 minute. Also, a more general conclusion seems to be

that the assumption of a sequence of independent volatility parameters β_k is not valid, as we are getting in general differences between the optimum fit of $f(\beta)$ and the corresponding fit of $p(u)$ written as an integral over Gaussians with the same corresponding parameters.

T.-L. Xu also discussed another possible reason why $p(u)$ could deviate from the histogram of the experimental data in her Msc thesis[132] where she applied the superstatistical technique and repeated the method of our thesis. She explained this phenomenon in terms of the problem of discreteness[103]. Share prices are discrete random variables and their values are recorded at discrete time points. The average of a large amount of such independent random variables will converge to a Gaussian variable based on the central limit theorem. This is the basis of the assumption that in each window the normalized log return is Gaussian distributed with a constant inverse parameter β. However when the time lag τ in Eq.5.17 is below 5 minutes[103], there may not be enough data points during the period τ to implement the central limit theorem. In this situation, the assumption that the normalized log return is Gaussian in each window is not valid.

5.3 Correlation functions

For the development of a suitable dynamical model, it is very important to look not only at probability densities but also on correlation functions and memory effects [133][134][135][136][137][138]. In our case there are two types of correlation functions: the one of the original data u_i,

$$C_u(t) = \frac{1}{N-t}\sum_{i=1}^{N-t} u_i u_{i+t} - \langle u_i \rangle^2 \tag{5.19}$$

and those of the volatility parameter β_k,

$$C_\beta(t) = \frac{1}{n-t}\sum_{k=1}^{n-t} \beta_k \beta_{k+t} - \langle \beta_k \rangle^2. \tag{5.20}$$

Figs. 5.9-5.12 show $C_u(t)/C_u(0)$ and $C_\beta(t)/C_\beta(0)$, both for the daily returns as well as for the 1-minute returns. As is illustrated in Fig. 5.9 and 5.10, $C_u(t)$ decays

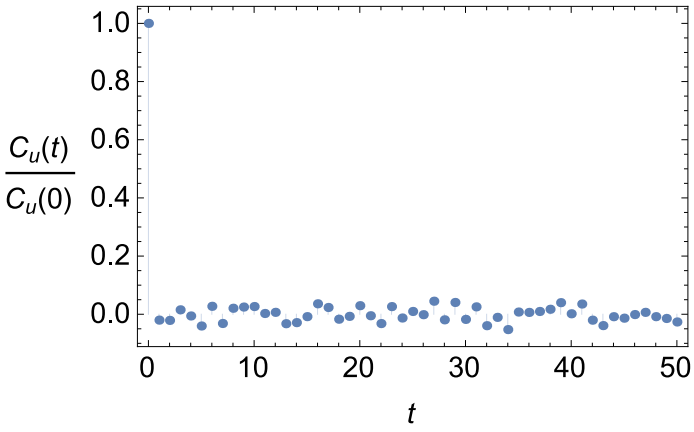


Figure 5.9: Correlation function of log-returns u on a daily time scale for AA shares. The time unit of t is days.

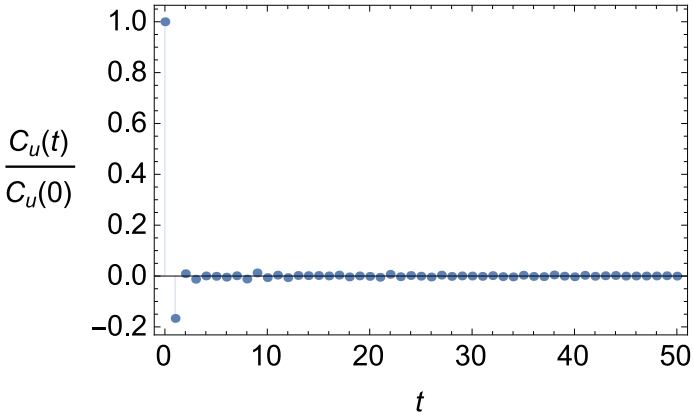


Figure 5.10: Correlation function of log-returns u on a time scale of minutes. The time unit of t is minutes.

almost immediately to zero, both for the daily and minute data.

More interesting is the correlation function $C_\beta(t)$. We did an analysis of the decay rates of correlation functions of the volatility for many different shares from different sectors. The data used for all those companies are in the same period from January 1998 to May 2013. The results are summarized in Tab.5.1. We observe that the correlation functions of volatility decay in an exponential way for daily returns, $C_\beta(t) \sim e^{-\gamma t}$, whereas for minute return there is a power law decay $C_\beta(t) \sim t^{-\alpha}$ with a periodic modulation, see Figs. 5.11-5.12 for the example of AA shares.

The period of oscillations that we observe in figures such as Fig. 5.12 corresponds (roughly) to one trading day and is consistent with periodic oscillations of

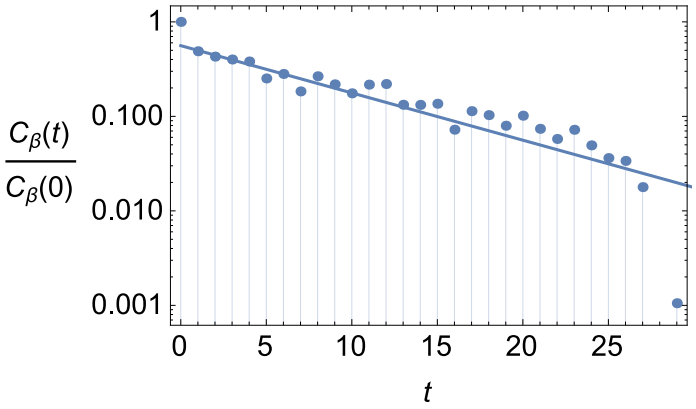


Figure 5.11: Correlation function of volatility β for returns on a daily time scale for AA shares.

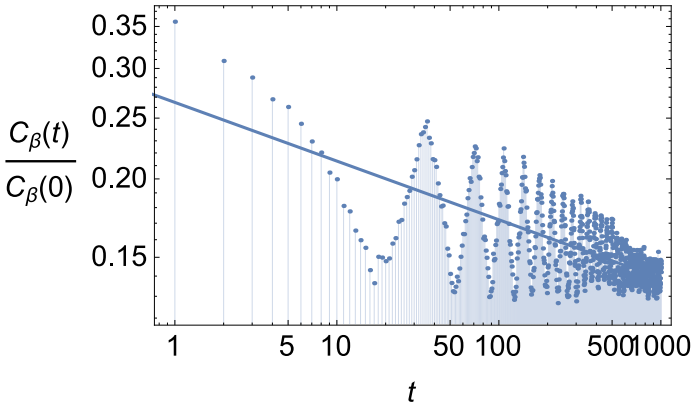


Figure 5.12: Correlation function of volatility β for returns on a time scale of minutes.

intraday volatility reported previously in [136].

The data in Tab.5.1 is summarized as plots of (γ, α) in Fig.5.13. It can be observed that the decay parameters γ and α depend slightly on the sectors. For example, shares from basic materials (except for **DD.**) have the strongest correlation decay (largest γ) on the daily scale and shares from industrial goods (except for **GE.**) have relatively strong correlation decay (large α) on the minute scale. Note that a strong decay of the volatility correlation function in a sense measures a 'volatility of a volatility' and is an interesting quantity to study.

Table 5.1: Decay rates(with standard errors) of correlation functions of volatility for shares of different sectors.

Company	Sector	$\gamma\pm$SE	T[Days]	$\alpha\pm$SE	T[Mins]
Alcoa Inc. (AA)	basic materials	0.115±0.0165	18	0.094±0.0029	11
Chevron Corporation (CVX)	basic materials	0.152±0.0157	20	0.164±0.0035	13
E. I. du Pont de Nemours and Company (DD)	basic materials	0.060±0.0035	17	0.125±0.0037	13
Exxon Mobil Corporation (XOM)	basic materials	0.138±0.0158	19	0.139±0.0030	13
General Motors Company (GM)	consumer goods	0.084±0.0094	16	0.112±0.0028	11
The Coca-Cola Company (KO)	consumer goods	0.056±0.0040	15	0.101±0.0032	13
Mondelez International, Inc. (MDLZ)	consumer goods	0.098±0.0236	15	0.171±0.0055	11
Altria Group Inc. (MO)	consumer goods	0.040±0.0036	15	0.104±0.0035	13
The Procter & Gamble Company (PG)	consumer goods	0.056±0.0056	15	0.117±0.0037	12
American International Group, Inc. (AIG)	financial	0.079±0.0046	15	0.108±0.0024	11
American Express Company (AXP)	financial	0.079±0.0089	19	0.090±0.0025	13
Bank of America Corporation (BAC)	financial	0.056±0.0040	20	0.073±0.0033	13
Citigroup Inc. (C)	financial	0.054±0.0043	17	0.059±0.0020	13
JPMorgan Chase & Co. (JPM)	financial	0.072±0.0076	18	0.089±0.0024	13
The Travelers Companies, Inc. (TRV)	financial	0.081±0.0062	15	0.117±0.0045	11
Johnson & Johnson (JNJ)	healthcare	0.037±0.0034	16	0.102±0.0036	13
Merck & Co. Inc. (MRK)	healthcare	0.071±0.0075	16	0.113±0.0046	12
Pfizer Inc. (PFE)	healthcare	0.064±0.0063	17	0.088±0.0029	14
UnitedHealth Group Incorporated (UNH)	healthcare	0.084±0.0102	15	0.178±0.0043	11
The Boeing Company (BA)	industrial goods	0.064±0.0063	15	0.131±0.0039	12
Caterpillar Inc. (CAT)	industrial goods	0.061±0.0057	15	0.143±0.0033	12
General Electric Company (GE)	industrial goods	0.066±0.0072	18	0.062±0.0018	14
Honeywell International Inc. (HON)	industrial goods	0.057±0.0029	15	0.126±0.0041	12
3M Company (MMM)	industrial goods	0.056±0.0075	15	0.148±0.0049	12
United Technologies Corp. (UTX)	industrial goods	0.077±0.0073	18	0.166±0.0049	12
The Walt Disney Company (DIS)	services	0.064±0.0042	18	0.088±0.0032	13
The Home Depot, Inc. (HD)	services	0.058±0.0046	16	0.109±0.0038	13
McDonald's Corp. (MCD)	services	0.039±0.0039	15	0.088±0.0031	13
Wal-Mart Stores Inc. (WMT)	services	0.036±0.0029	16	0.092±0.0035	13
Cisco Systems, Inc. (CSCO)	technology	0.038±0.0046	15	0.096±0.0020	13
Hewlett-Packard Company (HPQ)	technology	0.078±0.0082	15	0.131±0.0026	12
International Business Machines Corporation (IBM)	technology	0.047±0.0035	15	0.145±0.0038	13
Intel Corporation (INTC)	technology	0.050±0.0037	17	0.096±0.0022	13
Microsoft Corporation (MSFT)	technology	0.059±0.0066	16	0.104±0.0026	13
AT&T, Inc. (T)	technology	0.055±0.0036	16	0.087±0.0022	13
Verizon Communications Inc. (VZ)	technology	0.063±0.0054	17	0.105±0.0033	12

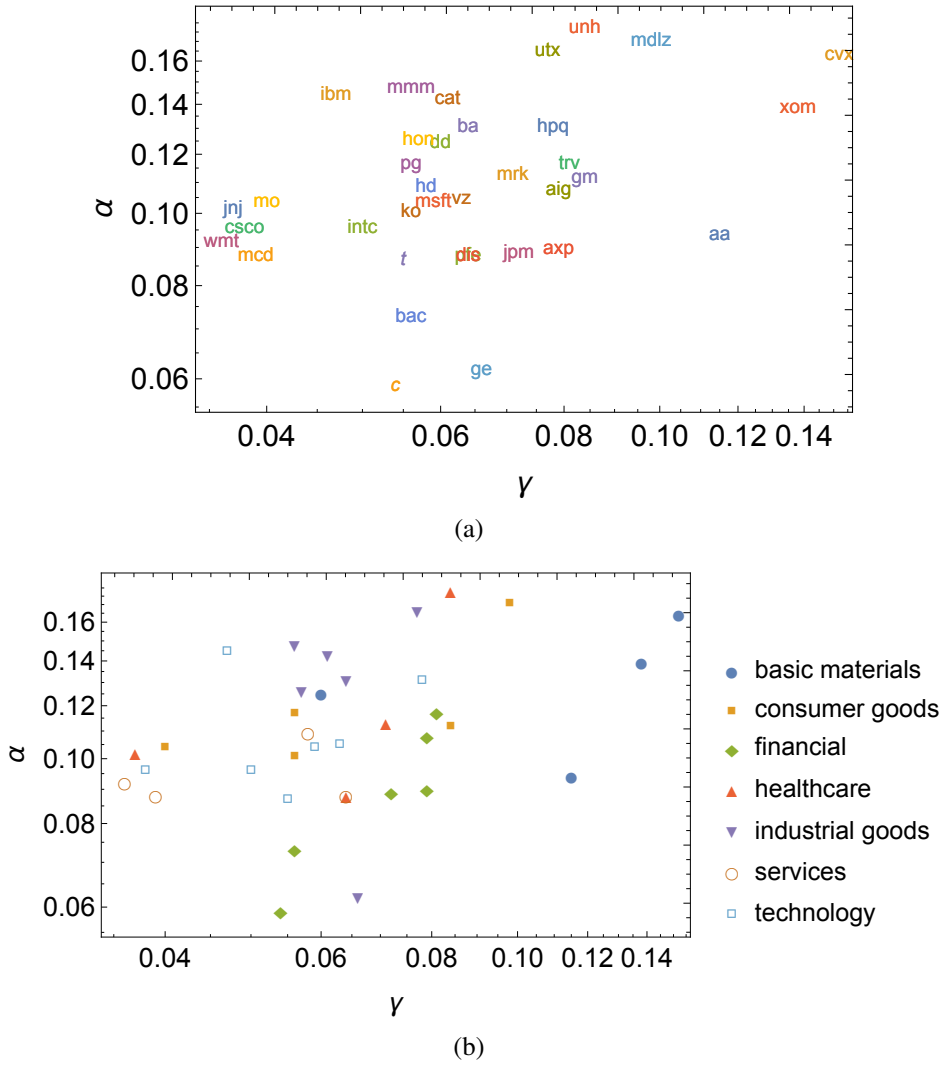
5.4 Synthetic model

Based on the results of the previous sections, it is desirable to construct a simple superstatistical dynamical model that incorporates the possibility of both lognormal and χ^2 superstatistics on different scales, and allows for different decay patterns of correlation functions.

Here we propose the following model. We start from a linear superstatistical Langevin equation

$$\dot{u} = -\gamma u + \sigma L(t) \tag{5.21}$$

where $L(t)$ is Gaussian white noise and the 'inverse temperature' β, in accordance



(a)

(b)

Figure 5.13: Plots of (γ, α) for shares(a) of different sectors(b).

with Einstein's theory of Brownian motion, is defined as

$$\beta = \frac{\gamma}{2\sigma^2}. \tag{5.22}$$

This equation is –by construction– superstatistical as we do not keep the parameter β constant but regard it as a random variable that fluctuates on a large time scale. Let us now consider $n+1$ Gaussian random variables X_i, $i = 0, 1, 2, \ldots, n$ which are statistically independent and have the same variance and mean 0 (except for X_0 which may have potentially a different variance and different mean). We then write

β as

$$\beta = \kappa e^{X_0} + (1-\kappa)(X_1^2 + X_2^2 + \ldots + X_n^2), \tag{5.23}$$

where $\kappa \in [0,1]$ is a parameter. We now see that if $\kappa = 1$, this system generates lognormal superstatistics, as $\log \beta = X_0$ is a Gaussian random variable. On the other hand, if $\kappa = 0$ this system generates χ^2-superstatistics with n degrees of freedom, as in this case $\beta = \sum_{i=1}^{n} X_i^2$ is χ^2 distributed. Choosing any value of $\kappa \in [0,1]$ one can interpolate between lognormal and χ^2 superstatistics, getting a mixed type of behaviour.

The Gaussian random variables X_i can again be simulated by ordinary linear Langevin equations of the form

$$\dot{X}_i = -\Gamma X_i + \Sigma L_i(t), \quad i = 0, \ldots, n \tag{5.24}$$

For constant Γ and Σ these equations generate the Ornstein Uhlenbeck process, i.e. a Gaussian Markov process with exponential decay of correlation functions. More complicated dynamics, leading e.g. to power law decay of correlation functions, can be constructed if the driving forces in these linear stochastic differential equations are not Gaussian white noise but more complicated correlated processes, or critical maps with a near-vanishing Liapunov exponent [139].

Fig. 5.14 and Fig. 5.15 show that indeed the observed distributions of $f(\beta)$ for Alcoa shares are best fitted by intermediate distributions (a superposition of a log-normal and χ^2 distribution with appropriate weights). The parameter κ increases if one goes from larger to smaller time scales of returns. The mixed synthetic model is able to reproduce the transition scenario of observed densities from χ^2 superstatistics to lognormal superstatistics in a quantitatively correct way, giving good fits on any time scale.

We did this analysis for a variety of time scales τ of returns, taking again the example of Alcoa shares. In Fig. 5.16 we show how the parameter κ depends on the time scale of returns. As expected, the parameter κ that best fits the observations decreases as a function of time scale. In fact we observe a logarithmic dependence

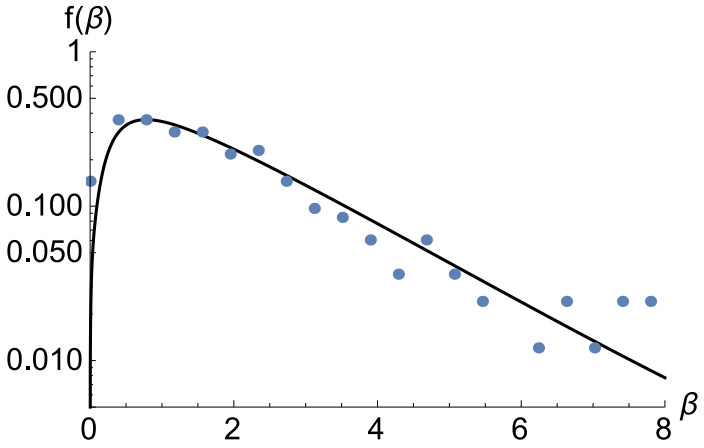


Figure 5.14: Mixed distribution fit to $f(\beta)$ with $\kappa = 0.36$ on the daily time scale.

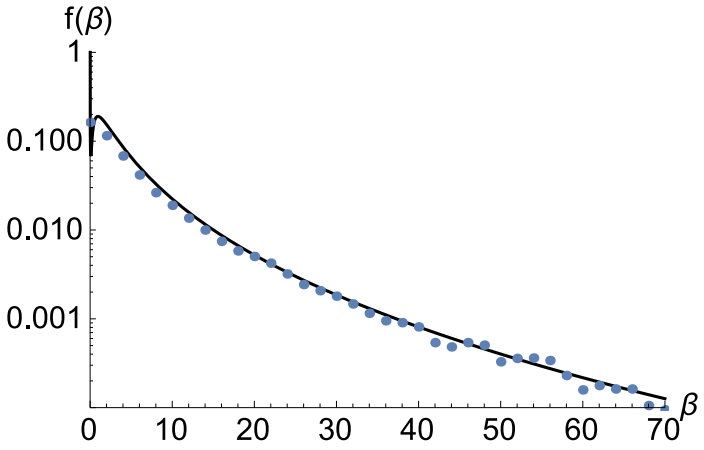


Figure 5.15: Mixed distribution fit to $f(\beta)$ with $\kappa = 0.92$ on the time scale of minutes.

if the time scale is not too big, see the straight line fit in Fig. 5.16.

We carefully repeated our analysis for various companies from different sectors. The summary is shown in Tab.5.2 and Fig.5.17. It can be observed that κ has larger values for small times scales than that for 2 to 3 hours time scales τ or daily time scales τ. Our results for different sectors are visualized in Fig.5.17. For all sectors κ changes most rapidly near $\tau \approx 100$.

5.5 Conclusion

Many investigations of complex systems in the past have focused on the application of a particular statistics, for example q-statistics [128], and then studying the effect

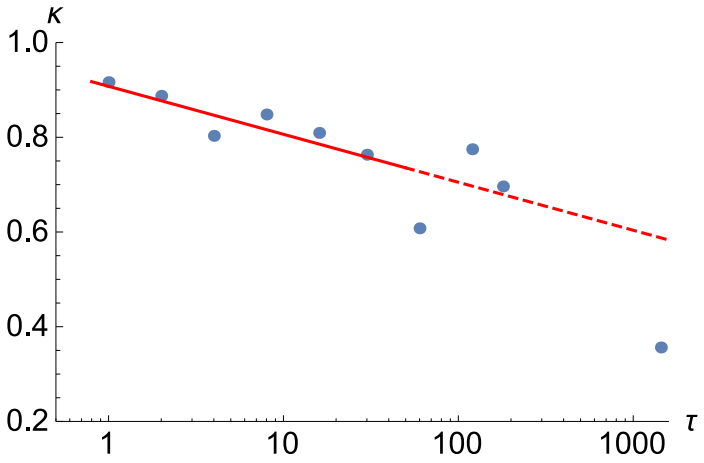
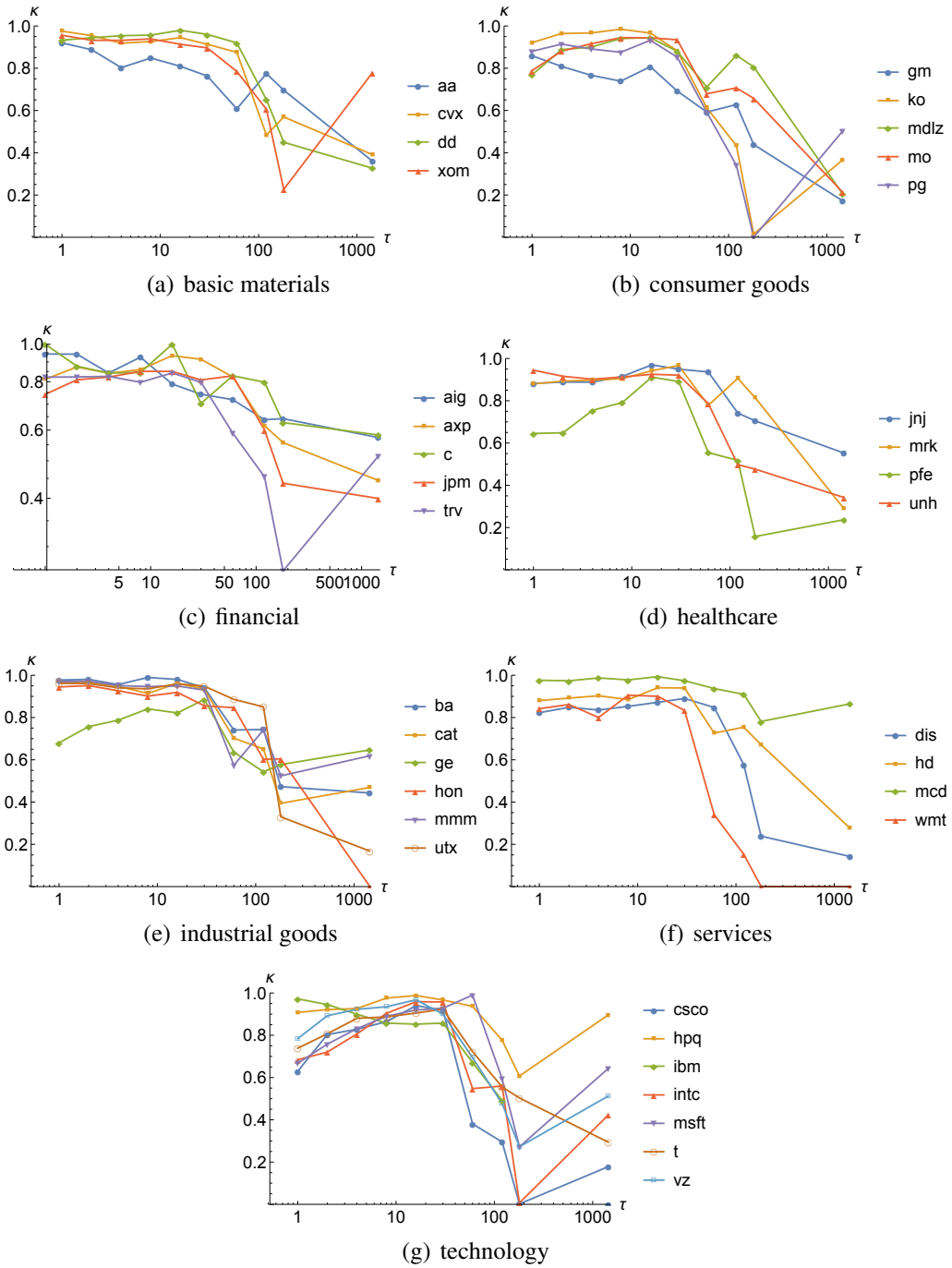


Figure 5.16: Parameter κ describing the relative weight of lognormal and χ^2 superstatistics in the mixed model as a function of the time scale τ of return. κ decreases if the time scale τ is increased. For not too big time scales τ a logarithmic dependence is observed: The straight line corresponds to a fit of the first six data points of the form $\kappa = 0.907 - 0.044 \log \tau$.

of varying system parameters, which may change the entropic index q. Here we have shown that for financial time series it is sometimes useful to consider broader classes of statistics and even proceed from one class of superstatistics to another when the scale or other system parameters under consideration are changed. The example we considered in detail in this chapter were share price returns of various companies. We provided evidence that there is a transition scenario from lognormal superstatistics to χ^2 superstatistics, with lognormal superstatistics giving a better fit to the data on small time scales and χ^2 superstatistics ($= q$-statistics) on larger time scales. We constructed a hybrid superstatistical model that allows to implement both types of superstatistics, with a weighting parameter κ that describes how far away we are from one of the two cases. Correlation functions of the extracted superstatistical volatility parameter β_k were shown to exhibit different qualitative behavior as a function of the time scale of returns, with exponential decay on large time scales and power law decay on small time scales, modulated by intraday periodicity. The decay parameters of the exponential or power law decay were extracted from the data and were shown to depend slightly on the sector of shares considered. The general transition scenario from lognormal to χ^2 superstatistics as a function of



(a) basic materials (b) consumer goods

(c) financial (d) healthcare

(e) industrial goods (f) services

(g) technology

Figure 5.17: Plots of τ versus κ for 36 companies from seven sectors.

Table 5.2: The parameter κ as a function of τ for 36 companies from seven sectors.

		τ[Mins]									
Company	Sector	1	2	4	8	16	30	60	120	180	1440
aa	basic materials	0.920	0.888	0.803	0.848	0.809	0.763	0.608	0.775	0.696	0.360
cvx	basic materials	0.976	0.954	0.918	0.926	0.945	0.913	0.875	0.484	0.569	0.391
dd	basic materials	0.933	0.944	0.954	0.957	0.979	0.960	0.920	0.651	0.450	0.327
xom	basic materials	0.956	0.932	0.932	0.939	0.913	0.896	0.785	0.606	0.227	0.780
gm	consumer goods	0.859	0.809	0.767	0.739	0.807	0.692	0.593	0.628	0.439	0.173
ko	consumer goods	0.921	0.964	0.967	0.985	0.966	0.883	0.613	0.434	0.015	0.365
mdlz	consumer goods	0.768	0.888	0.900	0.940	0.945	0.880	0.711	0.861	0.806	0.206
mo	consumer goods	0.789	0.881	0.916	0.945	0.944	0.935	0.679	0.707	0.657	0.214
pg	consumer goods	0.880	0.914	0.891	0.876	0.932	0.852	0.592	0.340	0.000	0.502
aig	financial	0.944	0.943	0.843	0.926	0.789	0.743	0.719	0.638	0.642	0.574
axp	financial	0.811	0.873	0.842	0.860	0.934	0.914	0.823	0.616	0.557	0.445
c	financial	1.000	0.877	0.842	0.845	1.000	0.703	0.829	0.798	0.627	0.583
jpm	financial	0.742	0.809	0.822	0.851	0.851	0.809	0.829	0.601	0.437	0.399
trv	financial	0.822	0.823	0.826	0.797	0.843	0.798	0.589	0.455	0.260	0.514
jnj	healthcare	0.882	0.888	0.888	0.915	0.970	0.949	0.936	0.742	0.705	0.552
mrk	healthcare	0.880	0.893	0.898	0.905	0.942	0.967	0.778	0.906	0.814	0.289
pfe	healthcare	0.645	0.648	0.755	0.791	0.912	0.891	0.555	0.518	0.158	0.237
unh	healthcare	0.945	0.916	0.902	0.912	0.925	0.921	0.786	0.497	0.476	0.342
ba	industrial goods	0.978	0.981	0.956	0.990	0.980	0.941	0.741	0.743	0.473	0.443
cat	industrial goods	0.971	0.965	0.948	0.914	0.962	0.934	0.702	0.651	0.394	0.469
ge	industrial goods	0.681	0.757	0.788	0.841	0.822	0.886	0.636	0.543	0.577	0.647
hon	industrial goods	0.945	0.951	0.927	0.901	0.919	0.855	0.847	0.602	0.604	0.003
mmm	industrial goods	0.970	0.972	0.951	0.947	0.949	0.930	0.574	0.742	0.524	0.619
utx	industrial goods	0.962	0.961	0.940	0.936	0.960	0.948	0.886	0.851	0.330	0.168
dis	services	0.823	0.848	0.836	0.853	0.872	0.890	0.848	0.577	0.239	0.143
hd	services	0.880	0.892	0.903	0.885	0.941	0.940	0.726	0.754	0.671	0.277
mcd	services	0.976	0.974	0.987	0.978	0.993	0.974	0.937	0.910	0.782	0.865
wmt	services	0.842	0.862	0.800	0.906	0.901	0.833	0.340	0.153	0.000	0.000
csco	technology	0.630	0.803	0.828	0.864	0.940	0.920	0.381	0.295	0.002	0.178
hpq	technology	0.908	0.922	0.926	0.976	0.988	0.967	0.937	0.776	0.607	0.896
ibm	technology	0.972	0.945	0.900	0.857	0.853	0.857	0.671	0.491	0.000	0.000
intc	technology	0.684	0.720	0.806	0.904	0.959	0.958	0.548	0.560	0.007	0.422
msft	technology	0.668	0.758	0.831	0.890	0.920	0.928	0.990	0.594	0.272	0.643
t	technology	0.739	0.807	0.880	0.887	0.905	0.924	0.723	0.556	0.503	0.294
vz	technology	0.784	0.893	0.923	0.935	0.968	0.900	0.694	0.476	0.273	0.512

the time scale of returns, however, is a robust phenomenon and occurs for all sectors in a similar way. Our detailed investigations of 3 different companies, as displayed in Fig.5.17, shows that the parameter κ most rapidly changes close to the limit scale $\tau \approx 100$min, at which point the system makes a pronounced change from lognormal to χ^2 superstatistics.

Chapter 6

Symbolic dynamics and multifractal analysis in finance

Fat-tailed probability distributions and long-range correlations in financial data were investigated in the previous chapter by the superstatistical method. These two features are known as being closely related to multifractality in a time series [53][61][54]. We will use an approach borrowed from dynamical systems theory to understand share prices in a coarse grained way. In this chapter we draw the method of symbolic dynamics [70][71][72][73] into our study on financial data which offers a coarse-grained way to trace the time evolution of share price trajectories. We can then study the attributed probability of a given coarse-grained orbit of the so-called symbol sequence and introduce the Rényi dimension and Rényi entropy to measure stochastic properties in a quantitative way. Different datasets will be compared.

6.1 Symbolic dynamics of share prices

Symbol sequence can be generated by a proper partition of the phase space. Since the log returns of a share price $\log \frac{S_{i+1}}{S_i}$ can theoretically be moving anywhere along the real number axis and the average of which is assumed to be 0 (Fig.6.1) , the easiest way is to partition the phase space interval $X = (-\infty, \infty)$ into two disjoint subsets $A_1 = (-\infty, 0)$ and $A_2 = (0, \infty)$. Note here A_1 and A_2 are all open and the zero point is not under consideration. This is because for very small time scale price data, say minutely recorded prices, due to the finite precision, a rounded-off price

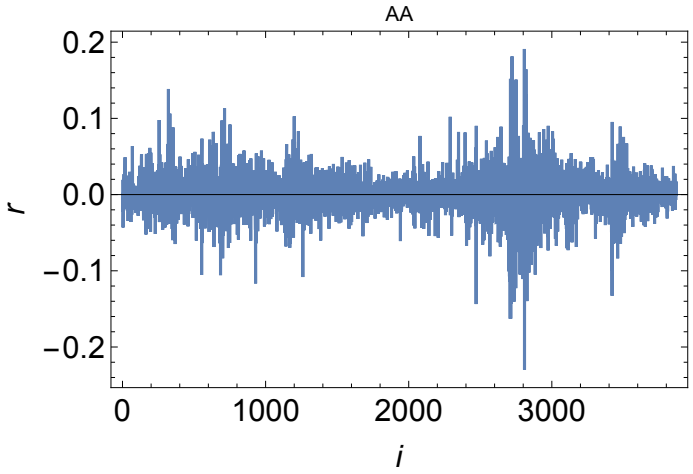


Figure 6.1: The log returns of a share(Alcoa Inc., over the period 1998 to 2013) fluctuate around 0.

will somehow stay unchanged until the same value is caught again at the next point in time, hence, we eliminate the effect of zero changes on the probability measure for both subsets.

We start with a dataset of daily stock prices of AA Inc. that covers the period January 1998 to May 2013. For any given value (except 0) of a log return at a time, it has roughly equal chances to fall into either of the two sets A_1 and A_2. If the log return is an element of A_1, which is equivalent to say $S_{i+1} < S_i$, then we denote such a price decrease by d. Otherwise a price in A_2 stands for $S_{i+1} > S_i$ and is denoted by u which means a price increasing. By this method we can allocate to the time series of share prices an attributed *symbol sequence*

$$i_0, i_1, i_2, ..., i_n, ... \text{where } i_n \in \{d, u\}$$

As we only deal with the trajectory over a limited period of time, we divide up the whole sequence into R segments of equal length N. We use a sliding window of certain length N that moves from the beginning of then entire symbol sequence to the right by one step at each time and extract all the segments enveloped in that window. By using this method, we can get more segments than that if we cut the whole sequence into non-overlapping segments.

Because each symbol has only two choices to settle, which are either u or d,

for any given N we get

$$\omega(N) = 2^N$$

allowed sequences. Since the partition of the phase space is rather simple and our dataset is big enough to satisfy $R >> \omega(N)$, there would be many segments that correspond to the same symbolic pattern. Then we can easily acquire the probabilities of each allowed symbol sequences by computing the frequencies of different sequences that occur. Namely

$$\begin{aligned} p_j^{(N)} &= p(i_0,...,i_{N-1}) \\ &= \frac{r(i_0,...,i_{N-1})}{\sum_{i_0,...,i_{N-1}} r(i_0,...,i_{N-1})} \\ &= \frac{r(i_0,...,i_{N-1})}{R} \qquad j = 1,...,\omega(N), \end{aligned} \tag{6.1}$$

where j runs over all the possible sequences and $r(i_0,...,i_{N-1})$ is the number of times that a certain sequence $i_0,...,i_{N-1}$ appears. Now by encoding each allowed symbol sequence of length N into a decimal in a range between 0 and 1, we can produce a scatter plot to visualise our results. Any given sequence of symbols $i_0,...,i_{N-1}$ can be represented by a sequence of bits, particularly we assign

$$x_{i_n} = \begin{cases} 1 & \text{if } i_n = u \\ 0 & \text{if } i_n = d \end{cases} \tag{6.2}$$

where $n = 0,1,...,N-1$. We regard the encoded binary version $\mathbf{x}^{(N)} = x_{i_0},...,x_{i_n},...,x_{i_{N-1}}$ of a N-step share price movement as a binary number, which can be eventually converted into a decimal fraction. One can implement this by defining the coordinate assigned to a given symbol sequence

$$\alpha(\mathbf{x}^{(N)}) = \sum_{n=1}^{N} x_{i_n} 2^{-n}, \tag{6.3}$$

note that $\alpha(\mathbf{x}^{(N)}) \in [0,1)$. By now, we allocate each symbol sequence a position on the real number axis so that we can easily illustrate our results on the frequency of

a given symbol dynamics, as shown in Fig.6.2. The probabilities of the empirical

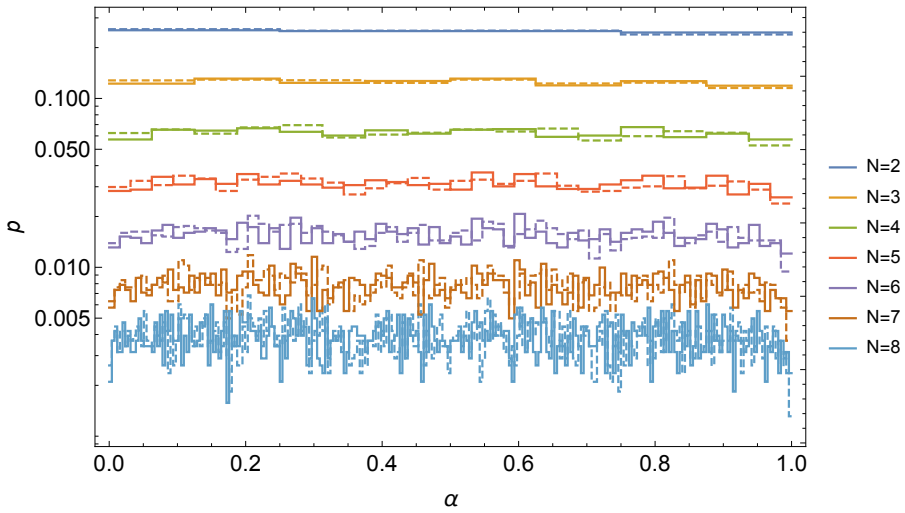


Figure 6.2: Joint probabilities of the daily share prices movement dynamics for symbol sequence of length 2 to 8 for Alcoa Inc.(solid lines) and those of purely random symbols sequences(dashed lines). The amount of symbols is the same.

symbol sequences are compared with random sequences with a same amount of symbols. The difference is not clearly visible. Note that N is kept as $N \leq 8$ because of the fact that we only have limited data points of AA daily price values. Similar bounds on N were also used in [70][71][72][73]. Using higher N is beyond the computing power and will result in unreliable numerical statistics since there are many patterns which do not occur.

We are also interested in the conditional probabilities which are defined by

$$\begin{aligned} &\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_0 = x_0) \\ =&\frac{\mathbb{P}(X_n = x_n, Xn-1 = x_{n-1}, Xn-2 = x_{n-2}, ..., X_0 = x_0)}{\mathbb{P}(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_0 = x_0)}. \end{aligned} \tag{6.4}$$

Conditional probabilities need to satisfy

$$\sum_{\omega \in \Omega} P(\omega | B) = 1. \tag{6.5}$$

Table.6.1 shows all the possible conditional probabilities of for symbol sequences of $N = 3$ in 4 columns labelled as $p(i_3 = \mathrm{d} | i_2 = \mathrm{d}, i_1)$, $p(i_3 = \mathrm{u} | i_2 = \mathrm{d}, i_1)$,

Table 6.1: Conditional probabilities of large scale sequences

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
d	0.483063	0.517199	0.493853	0.506409
u	0.523150	0.477112	0.515060	0.484135

$p(i_3 = \mathrm{d}|i_2 = \mathrm{u}, i_1)$, $p(i_3 = \mathrm{u}|i_2 = \mathrm{u}, i_1)$, where the observations of i_1 are listed on the left which can be either d or u.

It can be seen that the condition described by Eq.6.5 is satisfied, e.g.$p(d|d,d) + p(u|d,d) = 0.483063 + 0.517199 \approx 1.00$(also satisfied for other combinations).

6.2 Non-Markovian property

Conditional probability a is crucial indicator in the discussion of the Markovian property. From the definition of the discrete-time Markov chains, one has

$$\begin{aligned} &\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_0 = x_0) \\ =&\mathbb{P}(X_n = x_n | Xn - 1 = x_{n-1}). \end{aligned} \tag{6.6}$$

where $\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_0 = x_0)$ is defined by Eq.6.4.

If the Markovian property is satisfied, one must have

$$\begin{aligned} p(i_3 = \mathrm{d}|i_2 = \mathrm{d}, i_1 = d) &= p(i_3 = \mathrm{d}|i_2 = \mathrm{d}, i_1 = u) \\ p(i_3 = \mathrm{u}|i_2 = \mathrm{d}, i_1 = d) &= p(i_3 = \mathrm{u}|i_2 = \mathrm{d}, i_1 = u) \\ p(i_3 = \mathrm{d}|i_2 = \mathrm{u}, i_1 = d) &= p(i_3 = \mathrm{d}|i_2 = \mathrm{u}, i_1 = u) \\ p(i_3 = \mathrm{u}|i_2 = \mathrm{u}, i_1 = d) &= p(i_3 = \mathrm{u}|i_2 = \mathrm{u}, i_1 = u) \end{aligned}$$

which means each column in Table.6.1 must has same values. From the empirical results displayed in Table.6.1, one can conclude the daily share price data has slightly non-Markovian character and the conditional probabilities of observed share price returns depend on the history of previous symbols.

6.3 Renyi dimensions

For given probability distributions, it's practically meaningful to find a proper way to measure multifractal features. In this part we consider the Rényi Dimension from which some useful features will be revealed for the symbolic share price dynamics. Before the investigation of the Rényi Dimension we first look at the *Rényi information* defined as

$$I_q(p^{(N)}) = \frac{1}{q-1} \ln \sum_{j=1}^{\omega(N)} (p_j^{(N)})^q \tag{6.7}$$

where q is an arbitrary real number and $\omega(N)$ is the number of allowed symbol sequences for a certain N, so that j labelled all the probabilities. We regard the Rényi information as a generalisation of the Shannon information as for $q \to 1$,

$$\lim_{q\to 1} I_q(p^{(N)}) = \sum_{j=1}^{\omega(N)} p_j^{(N)} \ln p_j^{(N)} = I(p^{(N)}) \tag{6.8}$$

where $I(p^{(N)})$ denotes the Shannon information.

Another important feature we can capture in Eq.6.7 is that when $q = 0$,

$$I_0(p^{(N)}) = -\ln \omega(N), \tag{6.9}$$

which means $I_0(p^{(N)})$ decreases in a logarithmical way with the number of allowed symbol sequences. Note that when we have a limited number of symbol sequences, Eq.6.7 yields a finite value as $\omega(N)$ is finite.

Recall Eq.6.3 and from that shown in Fig.6.2, for given N, any symbol share price movements are mapped onto the phase space of $(0,1]$ and are located with equal distances. We call the distance of two neighboring coordinates *box size* and denote it by ε. In this way ε is simply determined by N as

$$\varepsilon = \frac{1}{2^N}. \tag{6.10}$$

That means ε is getting smaller as N grows and when $N \to \infty$, ε is approaching 0, in which case the Rényi information defined by Eq.6.7 diverges. Now we look at

the *Rényi dimension* which is important as it stays finite for $\varepsilon \to 0$.

$$D(q) = \lim_{\varepsilon \to 0} \frac{I_q(p^{(N)})}{\ln \varepsilon} = \lim_{\varepsilon \to 0} \frac{1}{\ln \varepsilon} \frac{1}{q-1} \ln \sum_{j=1}^{\omega(N)} (p_j^{(N)})^q. \tag{6.11}$$

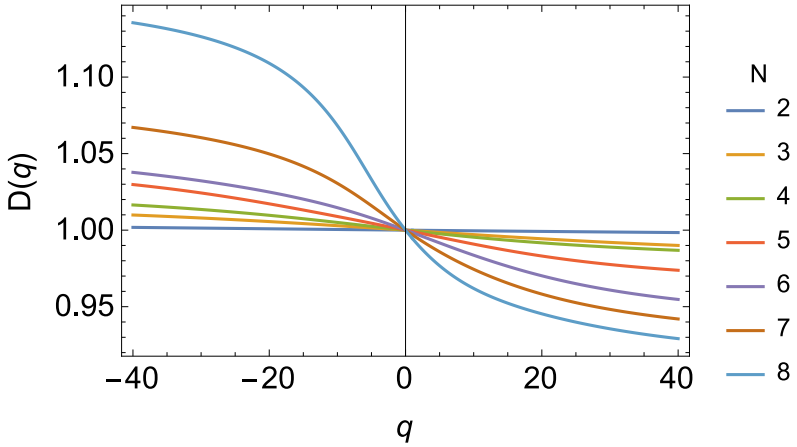


Figure 6.3: Finite-N Rényi dimensions for the daily share prices movement dynamics for symbol sequence of length 2 to 8 for Alcoa Inc.

The plots of Rényi dimensions for our daily share prices is shown in Fig.6.3 for various finite N. However, as the size of our dataset is limited, the smallest ε can be achieved is $\frac{1}{2^8}$, when N takes 8. The Rényi dimensions are monotonically decreasing and one can see the smaller the ε is the stronger dependence on q. In fact, $D(q)$ maintain positive for all q. Also, all lines intersect at 1 when $q = 0$. This is because

$$D(0) = \lim_{\varepsilon \to 0} -\frac{1}{\ln \varepsilon} \ln \sum_{j=1}^{\omega(N)} 1 = \lim_{\varepsilon} -\frac{\ln 2^N}{\ln \frac{1}{2^N}} = 1. \tag{6.12}$$

In [125], C.Beck(1990) put forward a rigorous general bound on Rényi dimensions where a general upper and lower bound of the Rényi dimension was proved:

$$\frac{q'-1}{q'} D(q') \geq \frac{q-1}{q} D(q) \qquad \text{for } q' > q, q'q > 0. \tag{6.13}$$

If we substitute $+\infty$ and $-\infty$ for q in Eq.6.13, we can obtain an upper bound

$$D(q) \leq \frac{q}{q-1} D(+\infty) \qquad \text{for } q > 1 \tag{6.14}$$

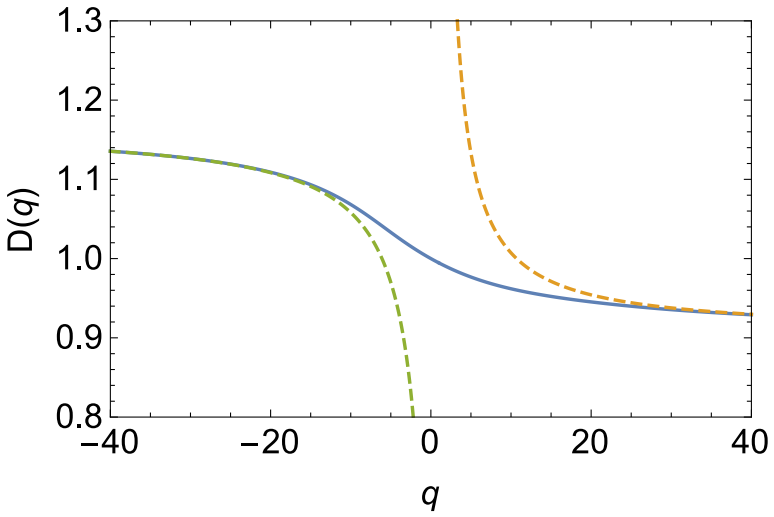


Figure 6.4: Rényi dimensions with the upper(orange dash line) and lower(green dash line) bounds for the daily share prices movement dynamics for symbol sequence of 8 for Alcoa Inc .

and a lower bound

$$D(q) \geq \frac{q}{q-1} D(-\infty) \qquad \text{for } q < 0. \tag{6.15}$$

Here we check these conditions for our dataset with $N = 8$ in Fig.6.4. It is shown that Eq.6.14 and Eq.6.15 are indeed followed by our experimental data .

6.4 Renyi entropies

Unlike the Rényi dimension, which describes how the Rényi information grows with the refining scale ε, the *Rényi entropy* $K(q)$ measures the changing of the negative Rényi information over time $N \to \infty$.

$$K(q) = \lim_{N\to\infty} \frac{-I_q^{(N)}}{N} = \lim_{N\to\infty} \frac{1}{1-q} \frac{1}{N} \ln \sum_{j=1}^{\omega(N)} (P_j^{(N)})^q \tag{6.16}$$

We have plots of Rényi entropies where N grows from 2 to 8 in Fig.6.5.

The Renyi entropies have the same dependence on q as $D(q)$. This is because in the context of the current study, each box with size ε represents an observation of a symbol sequence with length N. According to Eq.6.11, Eq.6.16 and Eq.6.10,

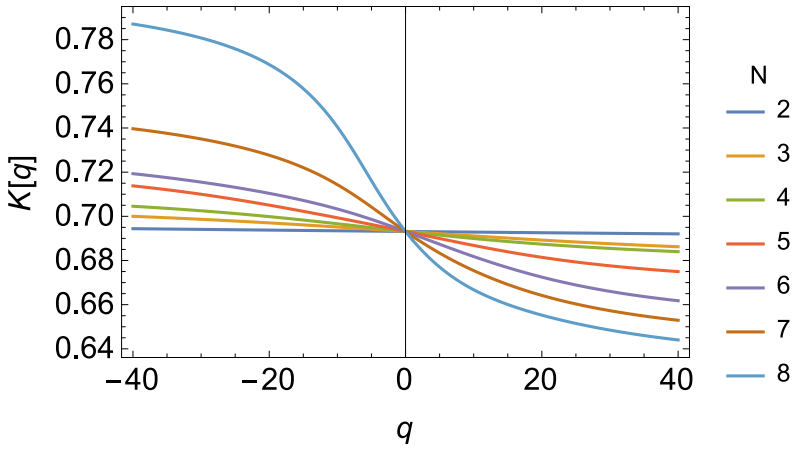


Figure 6.5: Rényi entropies for the daily share prices movement dynamics for $N = 2, \ldots, 8$ for Alcoa Inc.

the Rényi dimension $D(q)$ is related to the Rényi entropy by

$$D(q) = K(q) \ln 2. \tag{6.17}$$

the multifractal properties of symbol sequences can be described in terms of Rényi entropy, which is regarded as the dynamical counterpart of Rényi dimension. In our case the two contain the same information, since we map the probabilities of observed symbol sequence onto a multifractal probability measure on the unit interval.

Again the intersection when $q = 0$ is worth mentioning, as

$$K(0) = \lim_{N\to\infty} \frac{1}{N} \ln 2^N = \ln 2. \tag{6.18}$$

is the topological entropy, no matter what value N takes. Moreover for $q \to 1$, based on Eq.6.8,

$$K(1) = \lim_{N\to\infty} \frac{-I(p^{(N)})}{N} \tag{6.19}$$

where the negative Shannon information $-I(p^{(N)})$ is the *Shannon entropy* and $K(1)$ is also called the *Kolmogorov-Sinai entropy*.

From the non-trivial spectra shown in Fig.6.5, we can say that the symbol sequences are not independent and identically distributed random variables and there

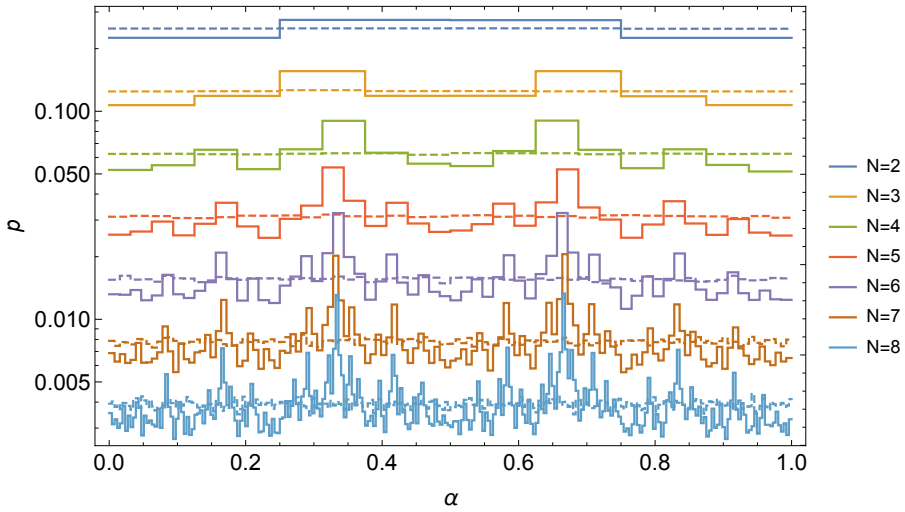


Figure 6.6: Joint probabilities of the minute scale share prices movement dynamics for symbol sequence of length 2 to 8 for Alcoa Inc.(solid lines) and those of purely random symbols sequences(dashed lines). The amount of symbols is the same.

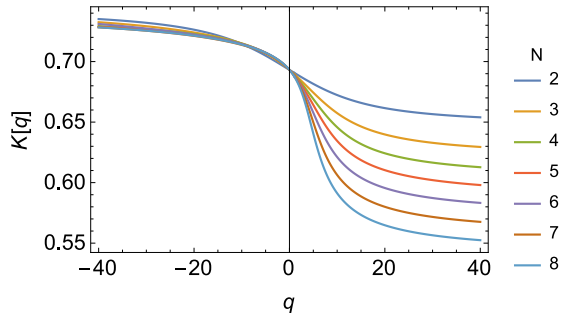
are non-trivial correlations in the price movement dynamics.

6.5 Small time scales

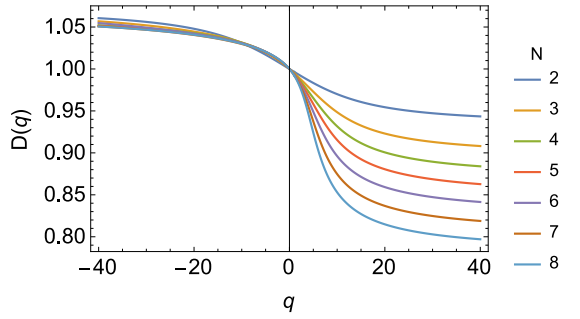
We have quantified the joint probabilities of the daily share price movement by Rényi dimension and Rényi entropy, and have numerically provided evidence that the share price movement has a slightly non-Markovian property. We are now interested in how these measurements work for the small time scale symbolic dynamics for the same company. Instead of using the daily share prices, the new dataset includes the minutely recorded share prices for the same period from 1998 to 2013 that covers about 1.5 million data points. Since there is a substantial increase in the data size, dividing the whole sequence into non-overlapping segments will not cause big errors in the result of our experiment. By using the same partition method of the previous section the probability distributions of symbol sequence with different length are shown in Fig.6.6.

Compared with Fig.6.2, the probability distributions on small time scales are significantly different from those of daily time scale. Random symbol sequences tend to have probabilities that are evenly distributed over all the possible patterns. All the sub-figures in Fig.6.6 have symmetry around $\frac{1}{2}$ and it seems there is some

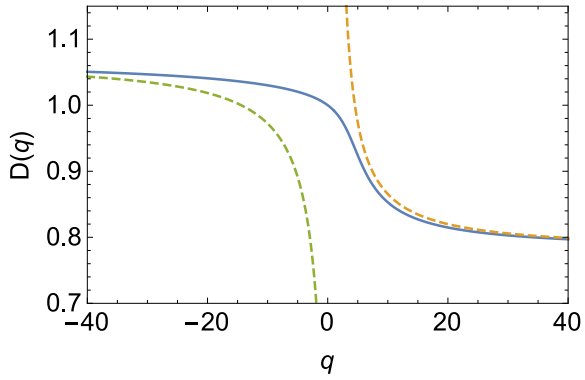
self-similarity. We show the Rényi entropies and the Rényi dimensions together with bounds for small scale data in Fig.6.7.



(a)



(b)



(c)

Figure 6.7: (a)The Rényi entropy of minute scale symbolic dynamics for $N = 2, ..., 8$. (b) the Rényi dimensions of minute scale symbolic dynamics for $N = 2, ..., 8$. (c)The Rényi dimension for $N = 8$ with the upper and lower bounds.

We see, similar to the daily time scale, both the Rényi entropy and the Rényi dimension are monotonically decreasing with respect to q, with larger N having stronger dependence than smaller one for $q > 0$. However, the scaling behaviour for negative q can be hardly distinguished for different N.

Comparison of the multifractal spectra for large and small time scale symbol sequences of different length N is shown in Fig.6.8. We can see when $q > 0$, the

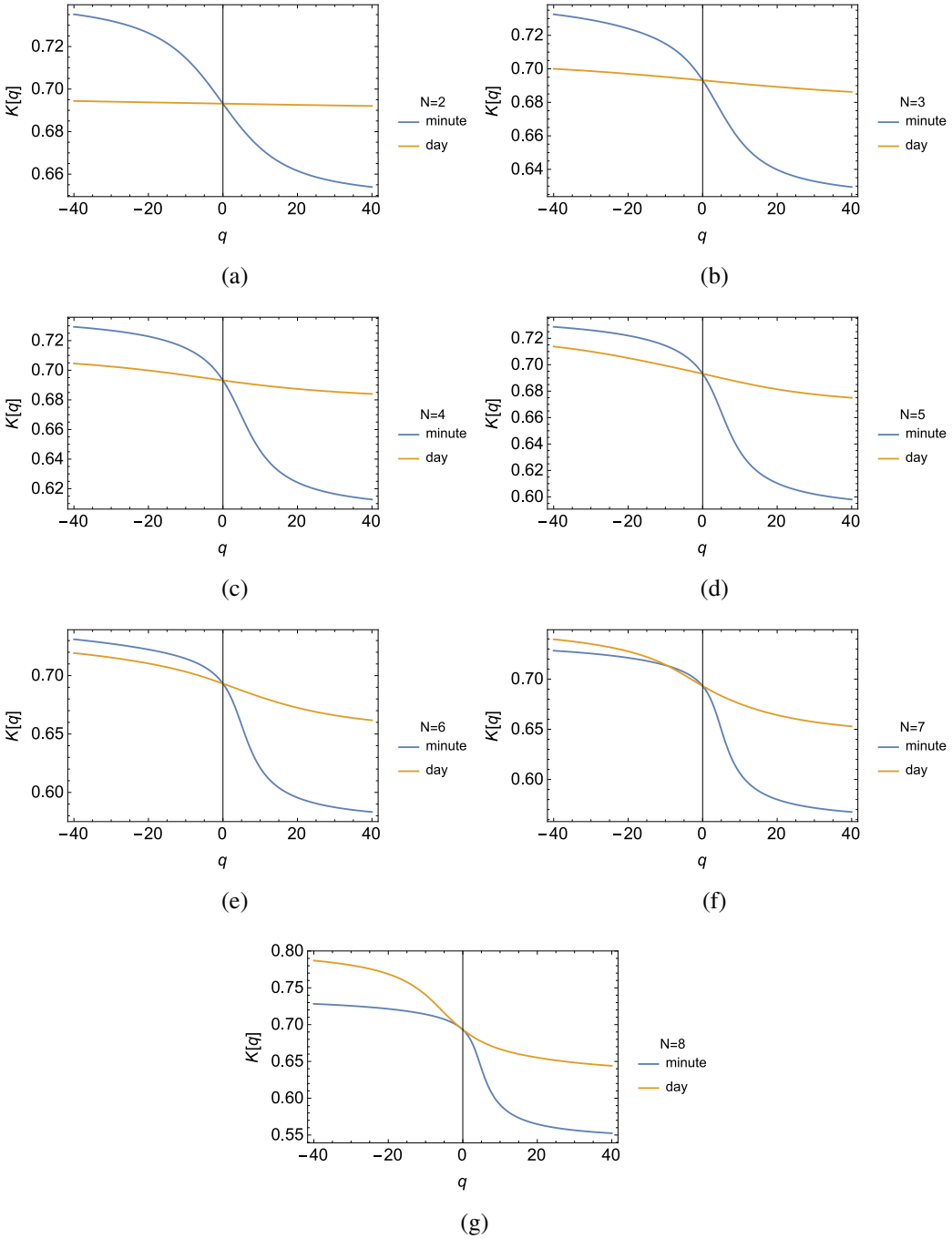


Figure 6.8: Comparison of the Rényi entropy for large and small time scale symbol sequences of different length N

Rényi entropies of the minute scale data always have stronger dependence on q, no

Table 6.2: Conditional probabilities of small scale sequences

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
d	0.474592	0.525938	0.567106	0.431542
u	0.432986	0.568584	0.52401	0.475199

matter how long the symbol sequence is. While for $q < 0$, daily data produce more flat spectra of Rényi entropy than minute data for small N(2-6) and have stronger q-dependence than minute data for larger N(7-8). We also test the non-Markovian character for the small time scale data by constructing a similar table as in the previous section. The conditional probabilities of small time scale share price changes for $N = 3$ are listed in Table.6.2. We can say the non-Markorvian character is also visible for the small scale share prices.

6.6 4-symbol partition

By now we have used the simplest method to study the symbolic dynamics of the share price movements. To detect further details, we can generate a refined version of the phase space partition $\mathbf{A} = \{A_1, A_2\}$ where $A_1 = (-\infty, 0)$ and $A_2 = (0, \infty)$. Assume there exists a real number c where the log returns have equal probabilities to fall into each element of the partition $\mathbf{B} = \{B_1, B_2, B_3, B_4\}$ where

$$B_1 = (-\infty, -c), B_2 = (-c, 0), B_3 = (0, c), B_4 = (c, +\infty). \tag{6.20}$$

In other words, the probabilities $p(B_i), i = 1,2,3,4$ of the log returns lying in each set B_i are identical to $1/4$. For Alcoa Inc. share prices on daily and minutely time scales, we find c is equal to 0.014 and 0.00088, respectively. Instead of denoting the time evolution by u and d, we redefine our symbol sequence by

$$i_0, i_1, ..., i_n, ... \qquad i_n \in \{b_1, b_2, b_3, b_4\} \tag{6.21}$$

where b_i corresponds to a log return in B_i, resulting in the number of allowed sequences, for a given length N, $\omega(N) = 4^N$. Also we update the bit sequence to a 4-level sequence by

$$x_{i_n} = \begin{cases} 0 & \text{if } i_n = b_1 \\ 1 & \text{if } i_n = b_2 \\ 2 & \text{if } i_n = b_3 \\ 3 & \text{if } i_n = b_4, \end{cases} \tag{6.22}$$

and any such sequence of length N can be encoded to a coordinate based on the representation

$$\alpha(\mathbf{x}^{(N)}) = \sum_{n=1}^{N} x_{i_n} 4^{-n}, \quad \alpha(\mathbf{x}^{(N)}) \in [0,1) \tag{6.23}$$

We show the plots for the joint probabilities of allowed symbol sequences of length N for both large and small time scales in Fig.6.9 and Fig.6.10. Again we use overlapping segments for daily data and non-overlapping segments for minutely data. The multifractal spectra are shown in Fig.6.11 and Fig.6.12.

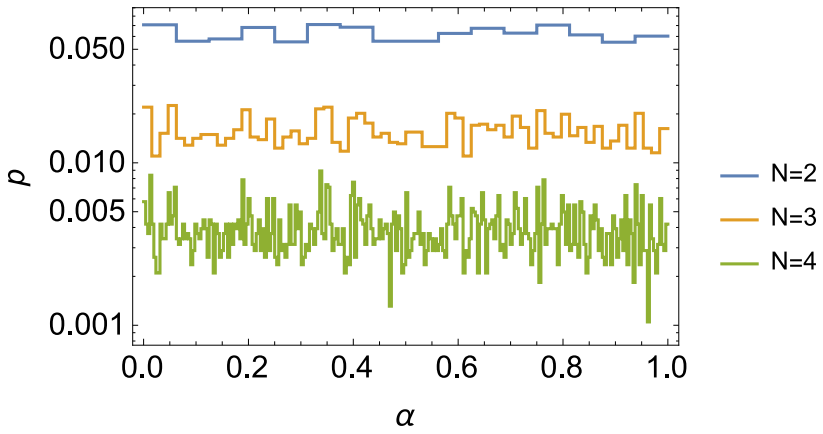


Figure 6.9: Probabilities for the 4-level sequences on large time scales for Alcoa Inc.

The overall shapes of the multifractal spectra produced from the 4-level symbol sequences are similar to that of the 2-level sequences. We will compare them by picking up a certain sequence length N and plot the curves from 2- and 4-level symbolic description in one figure for both small and large time scales as in Fig.6.13. We use $N = 4$ for the large time scale and $N = 5$ for the small time scale in the comparison since those are the longest sequence lengths that we can achieve for 4-level sequences. It can be clearly seen that the Rényi entropy of the 4-level sym-

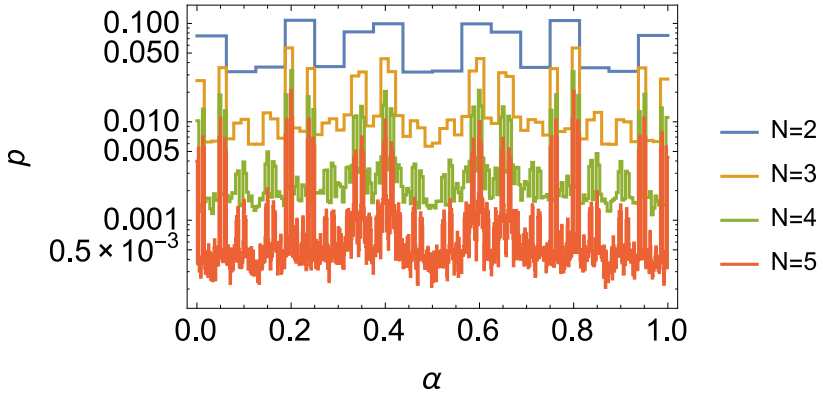
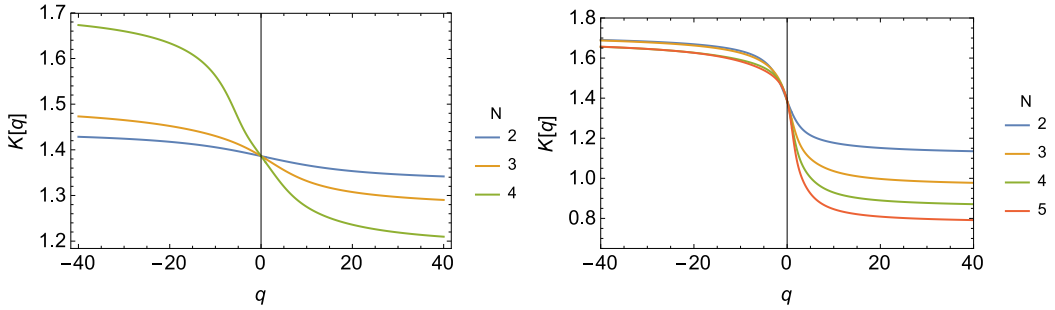


Figure 6.10: Probabilities for the 4-level symbol sequence on small time scales for Alcoa Inc.



(a) large time scale (b) small time scale

Figure 6.11: Rényi entropies for the 4-level symbol sequence of Alcoa Inc. All curves intersect at $(0, \ln 4)$.

bol sequence has much stronger dependence on q than that of the 2-level symbol sequence, for both large and small time scales.

We have also checked if the 4-level symbol sequences also have the non-Markovian property. If the 4-level symbol sequences are Markovian, then one would expect in Table.6.3 and 6.4 that each row has same values no matter what the columns is. This is not the case in both tables, thus we confirm that 4-level symbol sequences also display the non-Markovian character.

6.7 Differences in companies

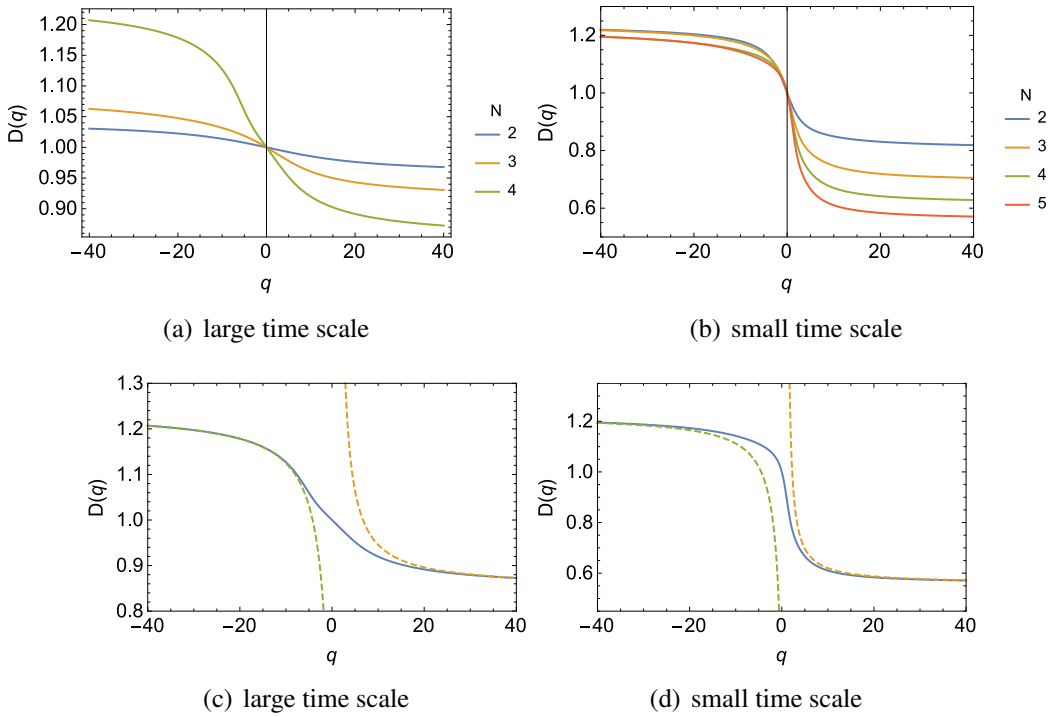
So far we have mainly taken one company as an example, to introduce our study based on symbolic dynamics. We now repeat the same experiment for share prices

Table 6.3: Conditional probabilities for 1-day time scale symbol sequences using an alphabet of 4 symbols

	0	1	2	3
$p(0\|0,\bullet)$	0.311193	0.221756	0.275773	0.297476
$p(1\|0,\bullet)$	0.155596	0.259502	0.275773	0.21567
$p(2\|0,\bullet)$	0.214871	0.283093	0.224358	0.204515
$p(3\|0,\bullet)$	0.318602	0.235911	0.224358	0.282602
$p(0\|1,\bullet)$	0.252403	0.199314	0.200889	0.239379
$p(1\|1,\bullet)$	0.229032	0.302662	0.32226	0.269301
$p(2\|1,\bullet)$	0.252403	0.310044	0.301334	0.218006
$p(3\|1,\bullet)$	0.266425	0.188241	0.175778	0.273576
$p(0\|2,\bullet)$	0.257986	0.172459	0.252985	0.222807
$p(1\|2,\bullet)$	0.221778	0.275934	0.256877	0.24651
$p(2\|2,\bullet)$	0.244408	0.295096	0.237416	0.308138
$p(3\|2,\bullet)$	0.27609	0.256772	0.252985	0.222807
$p(0\|3,\bullet)$	0.31162	0.257077	0.229227	0.33487
$p(1\|3,\bullet)$	0.211594	0.271099	0.308414	0.204401
$p(2\|3,\bullet)$	0.2039	0.23838	0.262569	0.191354
$p(3\|3,\bullet)$	0.273148	0.233706	0.195885	0.269636

Table 6.4: Conditional probabilities for 1-minute time scale symbol sequences using an alphabet of 4 symbols

	0	1	2	3
$p(0\|0,\bullet)$	0.351143	0.203533	0.185173	0.324927
$p(1\|0,\bullet)$	0.0835436	0.267288	0.260761	0.0745199
$p(2\|0,\bullet)$	0.0851713	0.294743	0.318656	0.0795201
$p(3\|0,\bullet)$	0.475625	0.226481	0.224449	0.525986
$p(0\|1,\bullet)$	0.207999	0.143277	0.114662	0.189087
$p(1\|1,\bullet)$	0.300433	0.35597	0.331887	0.295299
$p(2\|1,\bullet)$	0.327608	0.391438	0.444153	0.352179
$p(3\|1,\bullet)$	0.183645	0.11066	0.110332	0.167896
$p(0\|2,\bullet)$	0.164538	0.115243	0.117974	0.181273
$p(1\|2,\bullet)$	0.342409	0.441557	0.38673	0.322428
$p(2\|2,\bullet)$	0.297351	0.326313	0.354183	0.307034
$p(3\|2,\bullet)$	0.189177	0.11653	0.137152	0.192841
$p(0\|3,\bullet)$	0.523817	0.221966	0.229088	0.468126
$p(1\|3,\bullet)$	0.0804287	0.319912	0.294505	0.0857177
$p(2\|3,\bullet)$	0.0744439	0.270986	0.277406	0.0833825
$p(3\|3,\bullet)$	0.323635	0.176169	0.205183	0.360224



(a) large time scale (b) small time scale

(c) large time scale (d) small time scale

Figure 6.12: (a)(b)Rényi dimensions for the 4-level symbol sequence of Alcoa Inc. All curves intersect at (0, 1). (c)(d)Rényi dimensions with the upper and lower bounds(dashed curves).

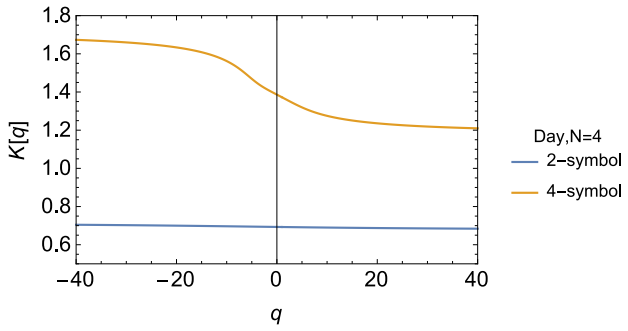
Table 6.5: companies belonging to different sectors

Company	**Sector**	**Industry**
Alcoa Inc. (AA)	basic materials	Aluminum
Bank of America Corporation (BAC)	financial	Money Center Banks
General Electric Company (GE)	industrial goods	Diversified Machinery
Intel Corporation (INTC)	technology	Semiconductor - Broad Line
Johnson & Johnson (JNJ)	healthcare	Drug Manufacturers - Major
The Coca-Cola Company (KO)	consumer goods	Beverages - Soft Drinks
Wal-Mart Stores Inc. (WMT)	services	Discount, Variety Stores

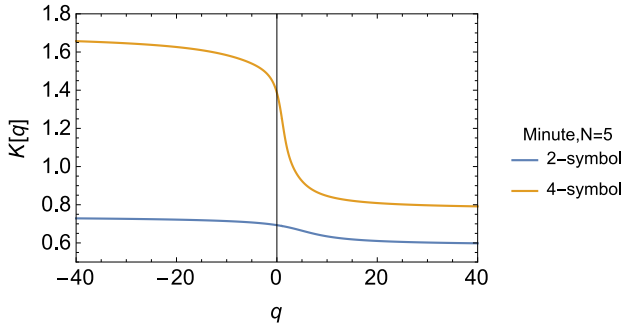
of 7 different companies belonging to different industries and sectors (Table.6.5). The plots of joint probabilities are shown in Fig.6.14-6.19.

To quantitatively analyse these joint probabilities, the corresponding Rényi entropies are shown in Fig.6.20

It can be observed that the spectrum of Rényi entropies is different for different companies. For 2-level sequences, Bank of America Corporation seems to have the lowest Rényi entropy for $q > 0$, on the other hand, the healthcare company



(a) large time scale



(b) small time scale

Figure 6.13: Comparisons of the 2-level and 4-level symbol sequence for given N in term of Rényi entropy (Alcoa Inc.)

Johnson & Johnson seems to have the highest Rényi entropy for $q < 0$. When using the more-refined 4-level partition function, Bank of America Corporation has the strongest dependence on q for both negative and positive vales of q. Overall the q-dependence is most pronounced for financial stocks. This could have to do with the fact that financial stocks have relatively strong fluctuations and exhibit non-trivial correlations, described by a non-trivial spectrum of Rényi entropies. They are more volatile than, say, healthcare sector stocks.

6.8 Memory length and predictability revealed by the conditional probabilities

The investigated non-Markovian behaviour indicates that the memory length of the symbolic share price is longer than 1. Let us provide evidence that the price movements depend on the entire history of previous symbols by investigating the condi-

Table 6.6: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1 minute scale share price

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.451	0.550	0.548	0.451
d	0.475	0.526	0.567	0.432
u	0.433	0.569	0.524	0.475
d, d	0.488	0.514	0.551	0.444
d, u	0.421	0.577	0.533	0.472
u, d	0.459	0.543	0.579	0.419
u, u	0.451	0.556	0.515	0.479
d, d, d	0.488	0.506	0.536	0.460
d, d, u	0.441	0.556	0.531	0.468
d, u, d	0.463	0.536	0.597	0.413
d, u, u	0.445	0.575	0.516	0.470
u, d, d	0.490	0.526	0.558	0.437
u, d, u	0.404	0.583	0.528	0.462
u, u, d	0.463	0.536	0.563	0.440
u, u, u	0.461	0.549	0.508	0.492
d, d, d, d	0.517	0.514	0.523	0.467
d, d, d, u	0.441	0.555	0.521	0.500
d, d, u, d	0.476	0.524	0.576	0.430
d, d, u, u	0.440	0.561	0.502	0.486
d, u, d, d	0.478	0.528	0.572	0.434
d, u, d, u	0.387	0.603	0.537	0.446
d, u, u, d	0.442	0.541	0.564	0.413
d, u, u, u	0.455	0.564	0.515	0.491
u, d, d, d	0.462	0.508	0.543	0.448
u, d, d, u	0.443	0.565	0.545	0.447
u, d, u, d	0.452	0.553	0.617	0.398
u, d, u, u	0.454	0.593	0.509	0.472
u, u, d, d	0.455	0.526	0.563	0.452
u, u, d, u	0.423	0.561	0.529	0.481
u, u, u, d	0.487	0.524	0.555	0.440
u, u, u, u	0.475	0.527	0.494	0.491

tional probabilities for longer symbol sequences. The conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for both 1 minute and 1 day share prices are listed in Table.6.6 and Table.6.7.

For instance, in both tables, the value in column $p(u|d,\bullet)$, row (d, d, u) represent the conditional probability $p(i_5 = u|i_4 = d, i_3 = u, i_2 = d, i_1 = d)$, which can be written in a compact form as $p(u|d,u,d,d)$. Assume the price has a memory length

Table 6.7: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1 day scale share price

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.503	0.497	0.505	0.495
d	0.483	0.517	0.494	0.506
u	0.523	0.477	0.515	0.484
d, d	0.467	0.533	0.492	0.508
d, u	0.513	0.487	0.510	0.488
u, d	0.498	0.502	0.496	0.505
u, u	0.534	0.466	0.520	0.480
d, d, d	0.496	0.505	0.526	0.474
d, d, u	0.516	0.484	0.536	0.465
d, u, d	0.513	0.488	0.491	0.509
d, u, u	0.506	0.494	0.496	0.504
u, d, d	0.442	0.558	0.458	0.542
u, d, u	0.509	0.491	0.483	0.513
u, u, d	0.485	0.516	0.498	0.502
u, u, u	0.564	0.437	0.546	0.454
d, d, d, d	0.463	0.537	0.518	0.482
d, d, d, u	0.481	0.519	0.483	0.517
d, d, u, d	0.488	0.512	0.530	0.471
d, d, u, u	0.515	0.485	0.449	0.551
d, u, d, d	0.395	0.605	0.449	0.551
d, u, d, u	0.469	0.531	0.462	0.530
d, u, u, d	0.448	0.552	0.508	0.492
d, u, u, u	0.556	0.445	0.555	0.445
u, d, d, d	0.527	0.473	0.533	0.468
u, d, d, u	0.557	0.444	0.581	0.419
u, d, u, d	0.539	0.461	0.451	0.550
u, d, u, u	0.496	0.505	0.543	0.458
u, u, d, d	0.488	0.512	0.466	0.534
u, u, d, u	0.545	0.455	0.505	0.496
u, u, u, d	0.519	0.481	0.486	0.515
u, u, u, u	0.572	0.429	0.535	0.465

of 3 only, then one should observe the following identity:

$$\begin{aligned} p(u|d,u,d,d) =& p(u|d,u,d,u) \\ =& p(u|d,u,d,d,d) \\ =& p(u|d,u,d,d,u) \\ =& p(u|d,u,d,u,d) \\ =& p(u|d,u,d,u,u), \end{aligned} \tag{6.24}$$

which means only three movements in the past will affect the next movement while what happened previously to the latest three movements is not supposed to have

influence on the next movement. But this is not what we see in both of the tables. Thus we have provided again numerical evidence that the share price movement is depending on the entire history of previous symbols.

Moreover, in Table.6.6, it can be observed that $p(u|d,\bullet)$ and $p(d|u,\bullet)$, irrelevant of the length of the symbol sequence, always have probabilities that are slightly higher than 0.5, while $p(d|d,\bullet)$ and $p(u|u,\bullet)$ are always slightly lower than 0.5. The data in this table has been summarized in a bar chart with error bars in Fig.6.21. Each symbol length N corresponds to 4 bars representing the means of the conditional probabilities of the 4 types($p(d|d...), p(u|d...), p(d|u...), p(u|u...)$), which are separated by columns in Table6.6. Except for the case of $N = 2$, where the means are exactly the conditional probabilities of $p(d|d), p(u|d), p(d|u), p(u|u)$, the error bars(mean$\pm$ standard deviation) are shown for other N's. It is clearly seen in Fig.6.21 that no matter how long the symbol sequence is, the average probability of observing a $u(d)$ conditional on a $d(u)$ in the previous step is always higher than the average probability of observing a $u(d)$ conditional on a $u(d)$ in the previous step. This means the occurrence of two successive movements in alternate directions has higher probability than the occurrence of two successive movements in one single direction. We can say that on the small time scale, a price increasing(decreasing) is more likely to be followed by a price decreasing(increasing). This result is also observable in the joint probabilities shown in Fig.6.6 where the highest probabilities are always hold by the symbol sequences with alternate d and u, e.g. 1,0,1,... or 0,1,0... However, this pattern is not valid for daily time scale share prices. Similar to the 2-level minute data, the 4-level sequences also get the highest joint probabilities for the patterns with alternate 3 and 0, e.g. 3,0,3,... or 0,3,0...

Indeed, this fact is observed for all companies on the minute time scale. Details of conditional probabilities and corresponding bar charts are listed in Appendix A and Appendix B.

6.9 Stationary test

In this section, we will investigate how the multifractal features of the financial time series are affected by an unstable period of financial crisis. This analysis is focused on a financial company, Bank of America (BAC), which was seriously influenced by the 2008 financial crisis.

Fig.6.22 shows the daily log returns r of BAC from January 1998 to May 2013. It can be clearly observed that the log returns with index in the range 2400-3000 exhibit large fluctuations; this the period corresponds to the 2008 financial crisis

We consider two new return series by splitting the original time series into two parts of equal amount of data, so that only the second part includes the period of financial crisis. To have enough data for this analysis, we use the minute scale log returns. Distributions of the normalized minute scale log returns are plotted in Fig.6.23 for both the time span without financial crises (January 1998 - August 2005) and that with the crisis (September 2005 - May 2013).

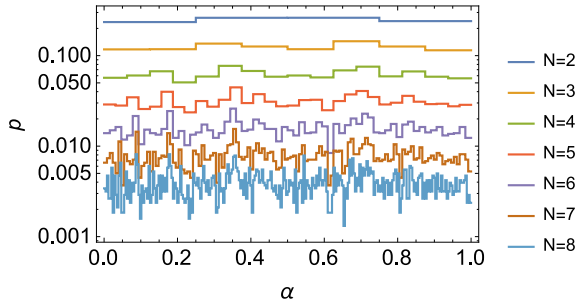
It is easy to see that the distribution of the data including the crisis period has fatter tails than the data excluding the crisis period. It means that the strongly fluctuating log-returns in the duration of financial crisis increase the probability of extreme values.

Now we repeat the analysis in the previous sections to compare the multifractal features of the two return series by means of Rényi entropies. The result is shown in Fig.6.24, where it is obvious that $K(q)$ of data influenced by the financial crisis has much stronger dependence on q than $K(q)$ of the data which does not include the crisis period. We can draw the conclusion that an unstable period with large price fluctuations increases the multifractality of the time series.

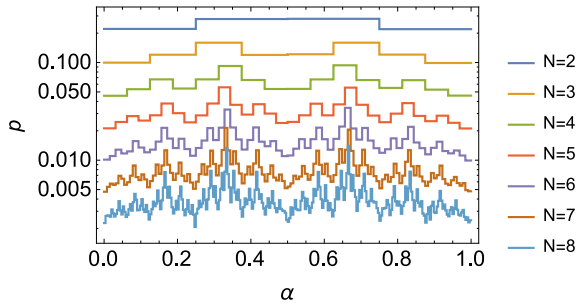
This result is in agreement with Morales et al.'s paper [140] where they confirmed that multifractality increased when the crisis occurred by using the Generalized Hurst exponent method.

6.10 Conclusion

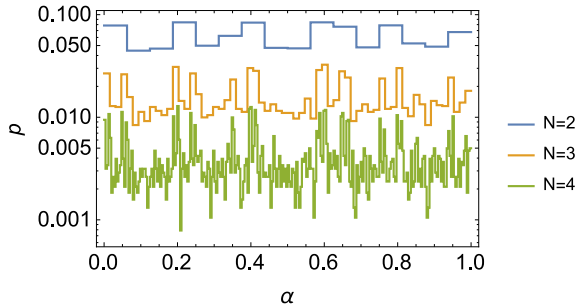
In this chapter, we applied the symbolic dynamics technique to analyse the evolution of share price returns in a coarse-grained way. Non-trivial spectra of Rényi entropies and Rényi dimensions were found which manifest nontrivial correlations in the symbolic share price dynamical systems. Such spectra were found to depend on the time scale of returns, the number of symbols used for the symbolic description and the sector of stocks considered. Also, The-Markovian property is investigated for the price movement dynamics and we provided evidence that the price movements depend on the entire history of previous symbols. It was also seen that alternating symbolic sequences occur with slightly higher than 0.5 conditional probability for the minute time scale data and this is observed for all the 7 companies taken into consideration in this chapter. This can possibly be interpreted as a small-sale feature related to the fact that a smaller price in the previous time interval motivates to buy as a kink of 'bargain' effect, and vice versa. In this way alternating sequences may get a slightly higher probability than random sequences, though the effect observed is only tiny.



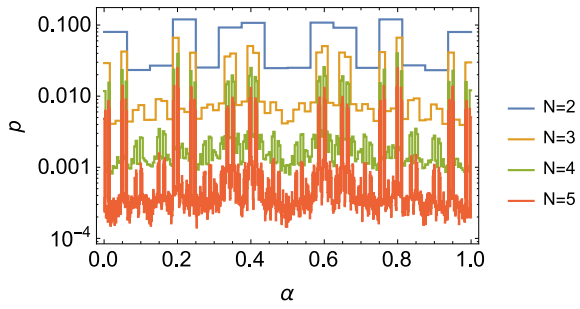
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

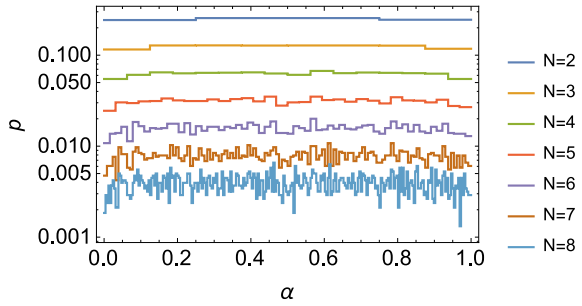


(c) large time scale, 4-level symbol sequence

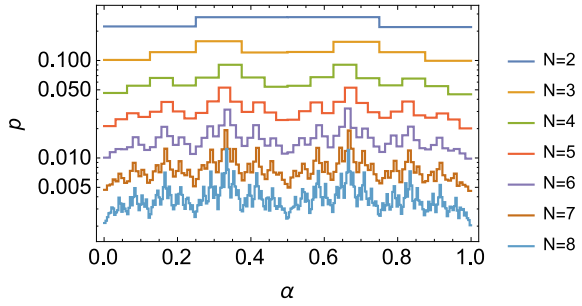


(d) small time scale, 4-level symbol sequence

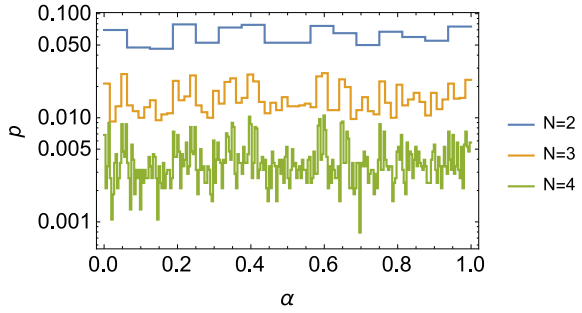
Figure 6.14: Joint probabilities of share price dynamics for Bank of America Corporation (BAC)



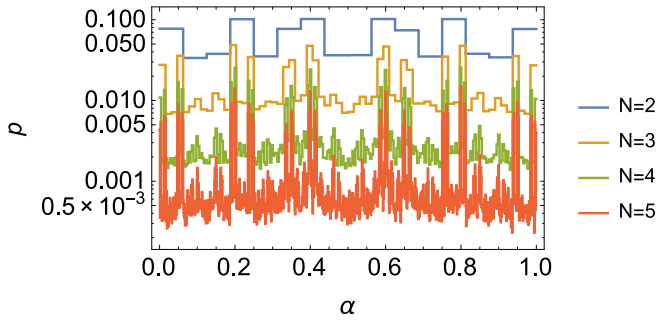
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

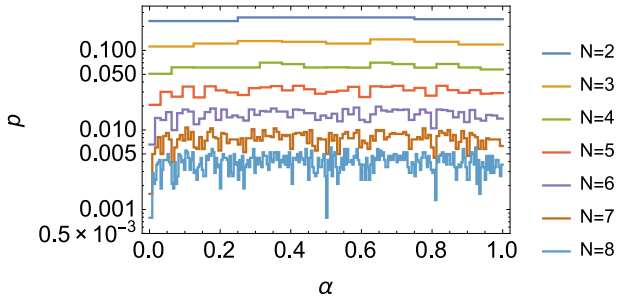


(c) large time scale, 4-level symbol sequence

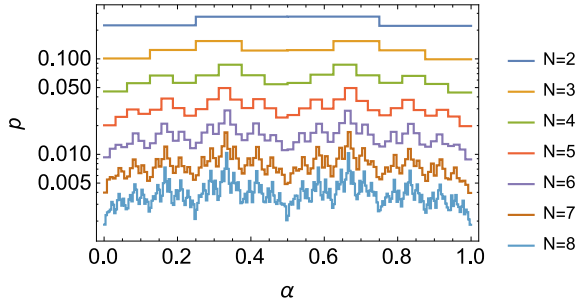


(d) small time scale, 4-level symbol sequence

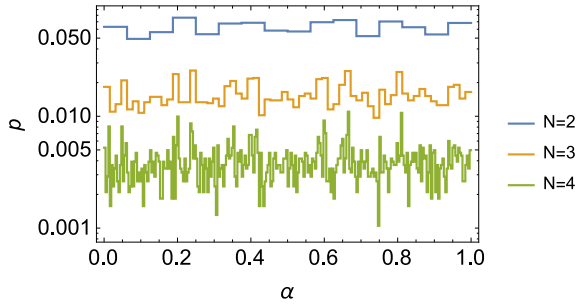
Figure 6.15: Joint probabilities of share price dynamics for General Electric Company (GE)



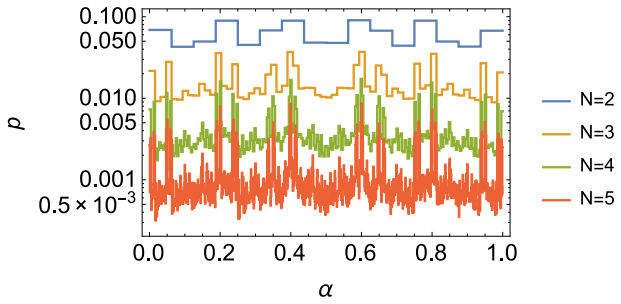
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

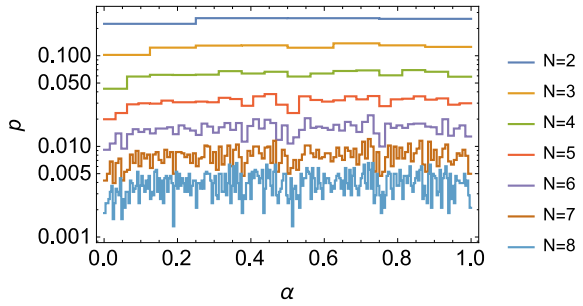


(c) large time scale, 4-level symbol sequence

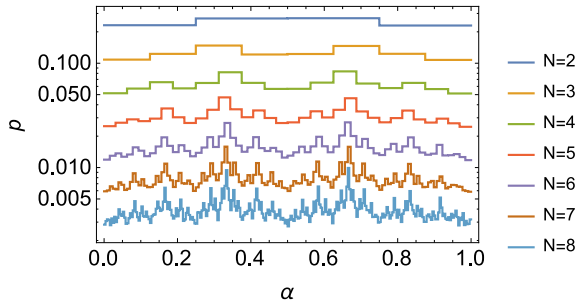


(d) small time scale, 4-level symbol sequence

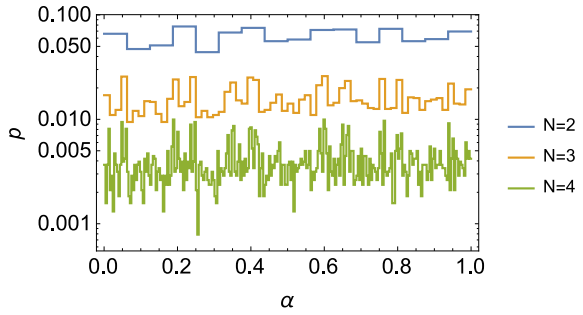
Figure 6.16: Joint probabilities of share price dynamics for Intel Corporation (INTC)



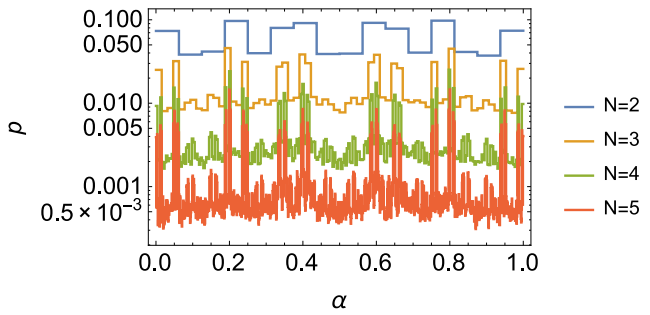
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

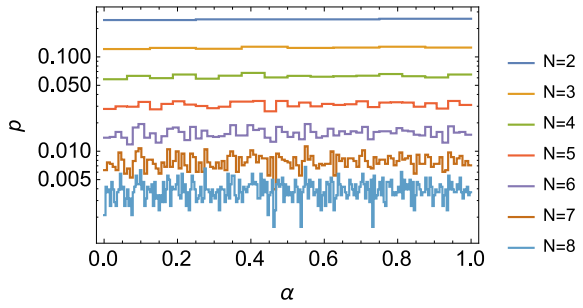


(c) large time scale, 4-level symbol sequence

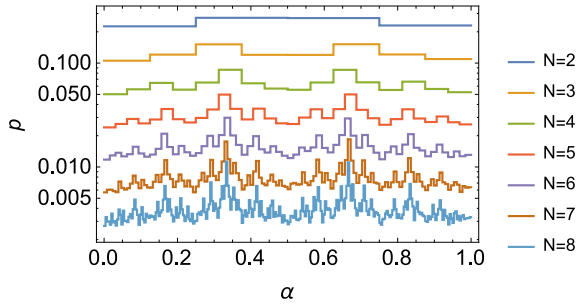


(d) small time scale, 4-level symbol sequence

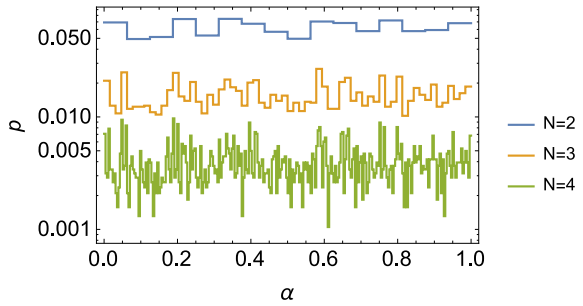
Figure 6.17: Joint probabilities of share price dynamics for Johnson & Johnson (JNJ)



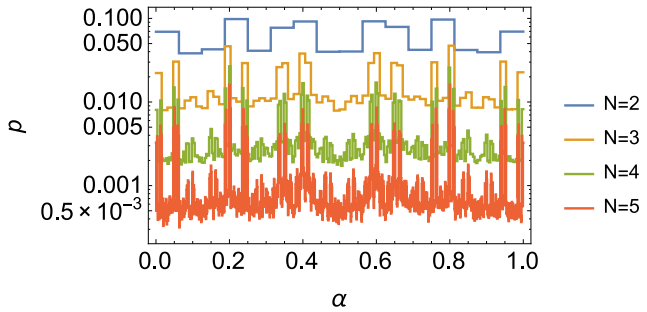
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

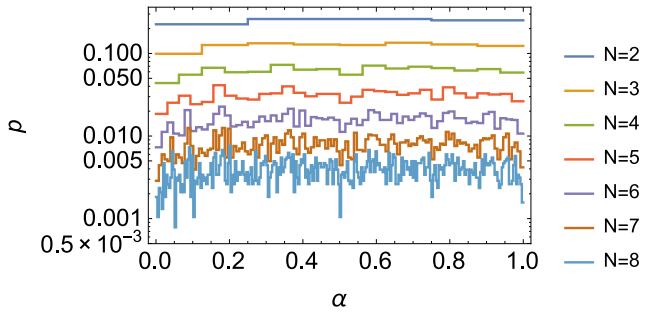


(c) large time scale, 4-level symbol sequence

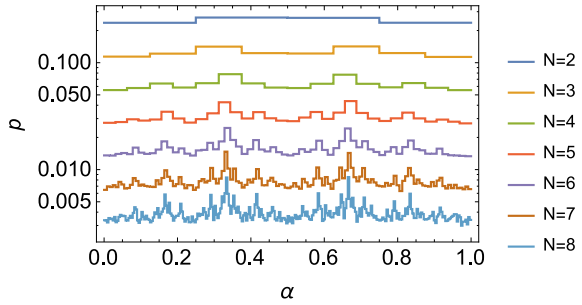


(d) small time scale, 4-level symbol sequence

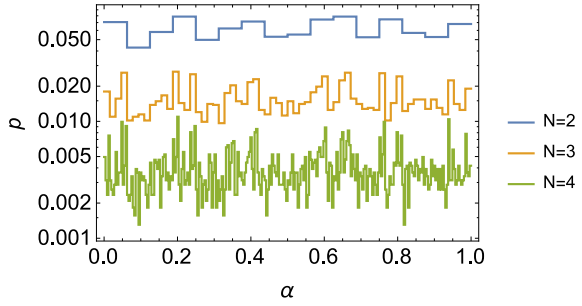
Figure 6.18: Joint probabilities of share price dynamics for The Coca-Cola Company (KO)



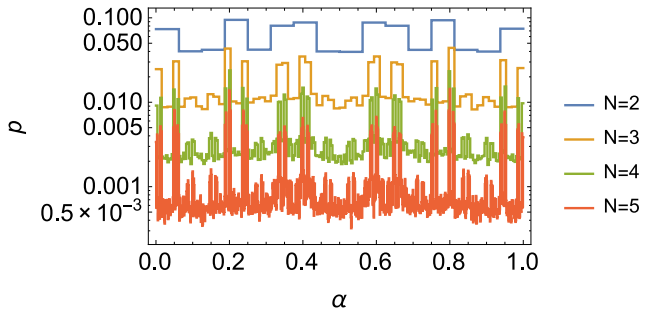
(a) large time scale, 2-level symbol sequence



(b) small time scale, 2-level symbol sequence

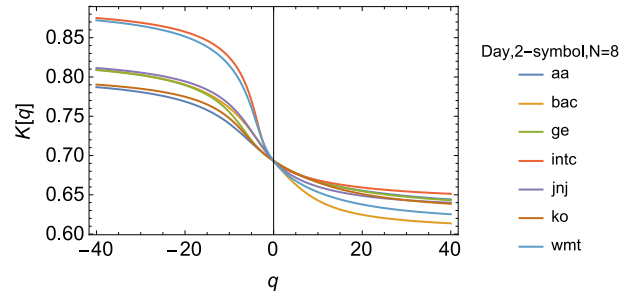


(c) large time scale, 4-level symbol sequence

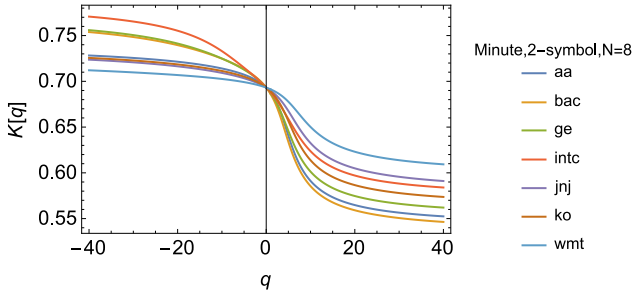


(d) small time scale, 4-level symbol sequence

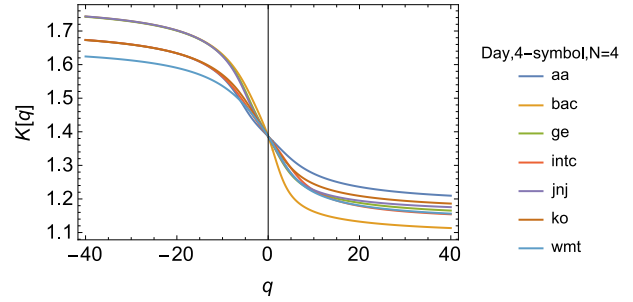
Figure 6.19: Joint probabilities of share price dynamics for Wal-Mart Stores Inc. (WMT)



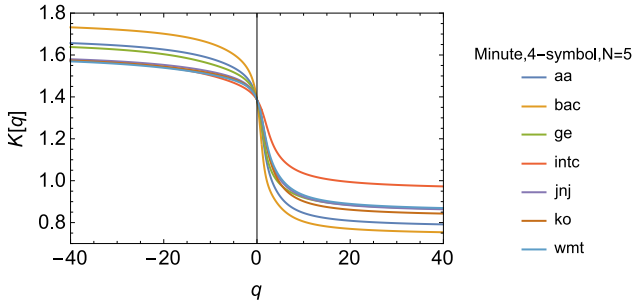
(a) large time scale



(b) small time scale



(c) large time scale



(d) small time scale

Figure 6.20: Rényi entropies for 7 companies. (a)(b)2-level sequence. (c)(d) 4-level sequence.

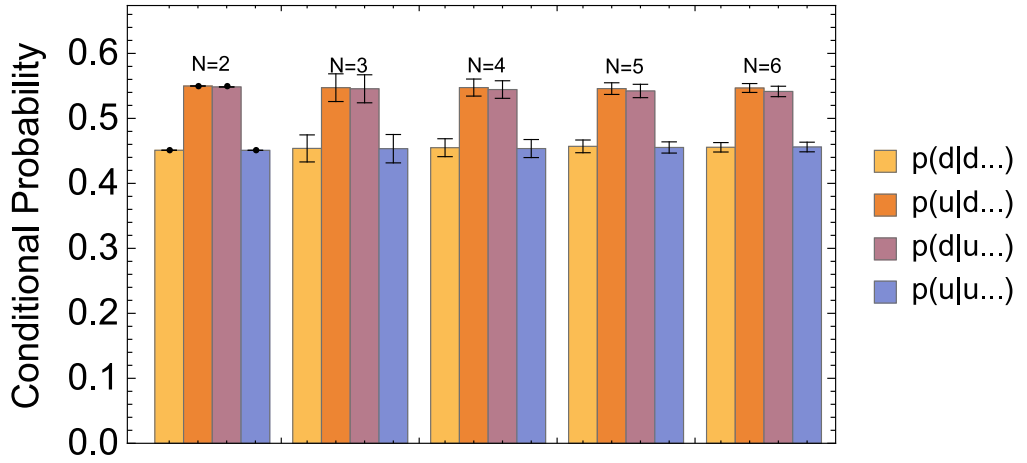


Figure 6.21: For $N = 2,...,6$, the mean values of the conditional probabilities of the 4 types($p(d|d...), p(u|d...), p(d|u...), p(u|u...)$) are represented by the height of the 4 bars(with error bars). No matter how long the symbol sequence is, the average probability of $p(u|d...)$ and $p(d|u...)$ is always higher than the average probability of $p(d|d...)$ and $p(u|u...)$

.

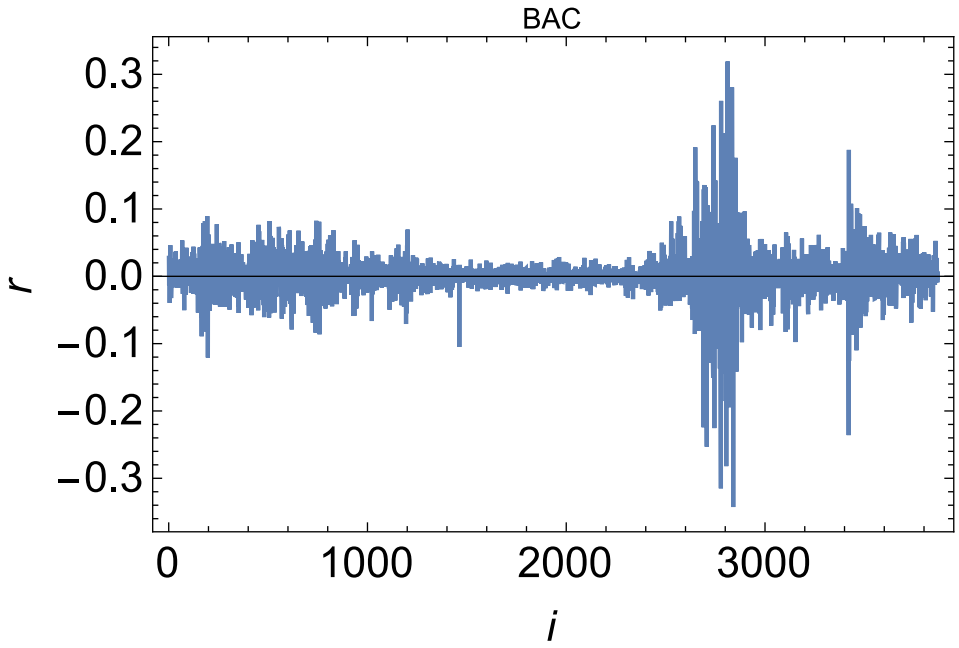


Figure 6.22: The daily log returns of Band of America (BAC) from January 1998 to 2013.

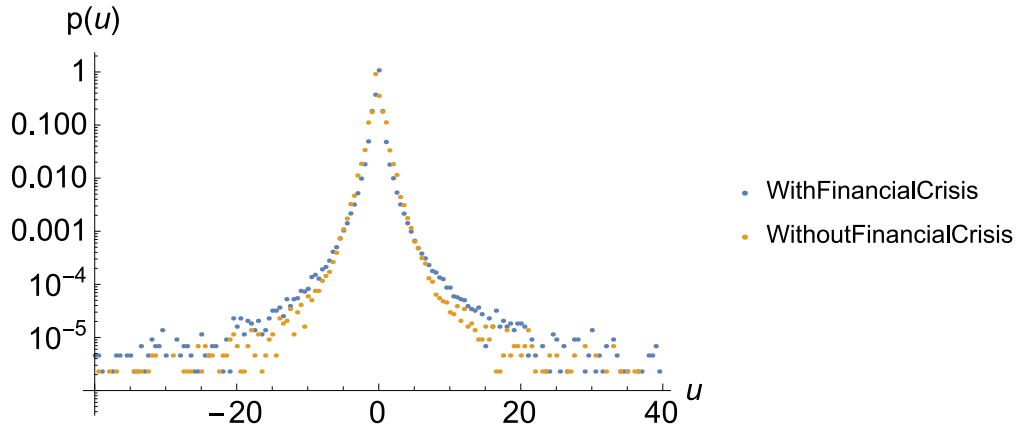


Figure 6.23: Distributions of log returns for time series including and not including the financial crisis 2008.

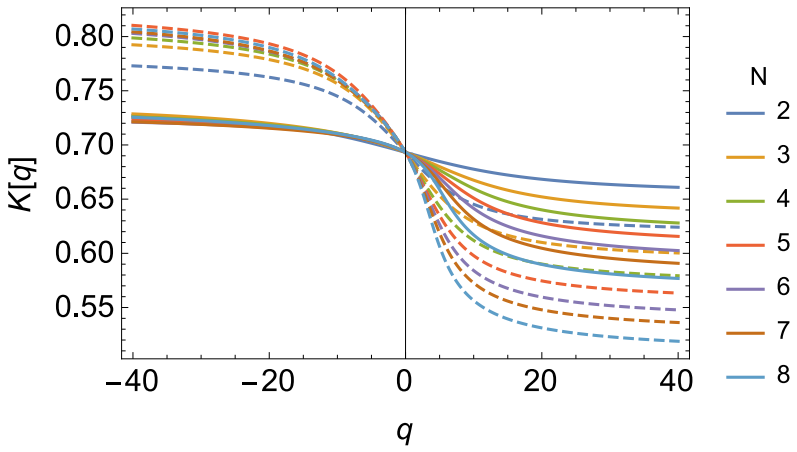


Figure 6.24: Rényi entropies of the above time series. The dashed lines include the financial crisis.

Chapter 7

Conclusion

This PhD thesis was devoted to the investigation of share price dynamics in terms of the statistics of log return distributions, correlation functions of share prices and their volatilities, and changes in the behaviour of volatility on various time scales. The study also explored the coarse-grained evolution of share price returns on various time scales, the probabilities of the generated symbol sequences and the corresponding Rényi entropies. All the analyses took into account share price datasets from different industries and sectors. Many of the studies in the field of finance have targeted at revealing the stylized facts of asset prices and seeking models for describing the price dynamics properly. Our study was engaged in answering the following three questions.

- What statistics is applicable in the modelling of log returns?
- Is there a difference between the characterization of log returns on small and large time scales?
- What properties can be disclosed from the coarse-grained evolution of share price returns?

The background of the applied methods and previous work in the literature were introduced in chapter 2 Stochastic volatility, chapter 3 Superstatistics and chapter 4 Symbolic dynamics and multifractal analysis. The concrete data analyses were performed in chapter 5 Superstatistics for share price returns and chapter

6 Symbolic dynamics and multifractal analysis for share price returns. The theoretical methods used, along with the main results of our thesis regarding the three questions, are summarized here.

- Share price returns, which on different time scales exhibit non-Gaussian distributions with fat tails, can be well modelled by superstatistical methods. The time series of returns is regarded as a superposition of local Gaussian processes weighted with a process of a slowly changing variance parameter β. The method of superstatistics is borrowed from nonequilibrium statistical mechanics, and put here into finance context.

- It was shown that χ^2-superstatistics works well for daily price returns, while on a much smaller time scale of minutes the price returns are better modelled by lognormal superstatistics. The system dynamics performs a transition from lognormal to χ^2-superstatistics as a function of time scale.

- We introduced a new superstatistical dynamical model that incorporates both lognormal and χ^2-superstatistics, with a weighting parameter $\kappa \in [0,1]$ that describes how far away a given model of share price dynamics is from one of the two cases. $\kappa = 1$ and $\kappa = 0$ respectively imply a lognormal and a χ^2 superstatistics and a value in between 0 and 1 generates a mixed superstatistics.

- Correlation functions of the extracted superstatistical volatility parameter β_k are shown to exhibit different qualitative behaviour as a function of the time scale of returns, with exponential decay on large time scales, modulated by intraday periodicity.

- Similar phenomena were found for all the companies considered in thesis thesis. However, the relevant parameters(such as decay rates of correlation functions) differ slightly among different stocks.

- The share price returns were analysed in a coarse-grained way using the symbolic dynamics technique from dynamical system theory. Symbol sequences were analysed for share price changes on time scales of 1 day and 1 minute,

from different stock sectors. The corresponding probabilities of the allowed symbol sequences were calculated. A nontrivial spectrum of Rényi entropies was found.

- The stylized fact of multifractality and the nontrivial correlations in the share prices are quantitatively measured by a nontrivial spectrum of Rényi entropies. It was numerically confirmed that the stochastic process of symbol sequences observed for real share price data has significantly non-Markovian character and the transition probabilities of observed share price returns depend on the entire history of previous symbols.

- The Rényi entropies have more dependency on the scale parameter q for 4-level symbol sequences compared to 2-level symbol sequence. It was also found that both the time scales of returns and the sector of stocks considered appear to affect the detailed form of the spectrum of Rényi entropies.

- For the small time scale share prices, we prenatally observed that a price increasing(decreasing) is slightly more likely to be followed by a price decreasing(increasing). The occurrence of two successive movements in alternate directions has higher probability than the occurrence of two successive movements in one single direction.

Some previous studies have already applied the technique of superstatistics in the area of finance [112][50][113][2][114][115][116][117][118]. χ^2-superstatistics, or equivalently Tsallis statistics are often observed to work well to describe the daily returns [112][50][113][2][114][115][116][117][118]. In this thesis, we present for the first time a systematic analysis where the time scale changes from minutes to daily scales, choosing the optimum superstatistical model properly according to the given time scale. In a previous paper by E. Van der Straeten and C. Beck [2], it was found that it is sometimes hard to distinguish which of lognormal and χ^2-superstatistics is best suitable to model the price changes. In my study here, such observed distributions were shown to be better modelled by a mixed case of superstatistics.

Not surprisingly, using a mixture of different distributions in the characterization of financial time series is prevailing recently. Non-Gaussian process like Lély processes are successful in characterizing the heavy tails of financial data but they fail to pose finite moments. Alexander [141] and Tan [142] have pointed out that a mixture of statistics will make promising improvement in modelling the excess kurtosis while ensuring finite moments. Superstatistics, which depicts the considered system as a mixed statistics of a Gaussian and a non-Gaussian distribution, has made nonnegligible contributions in the field nonequilibrium statistical mechanics, and can be applied in finance as well.

Previous studies on volatility extractions[143][144] from a time series of returns used different techniques from ours presented in this thesis. They usually cut the N data points into $n = N/\Delta$ time slices with equal length Δ and calculate an average of the absolute returns or the squared returns for each time slice. The histogram of the averages is fitted with some known distributions. The type of the fitted distribution has been observed not to change with the varying Δ, while different Δ result in different sets of parameters for that distribution. Recently, with the accessibility of high-frequency intraday data, Δ is often set to one day in previous studies [145][117]. Note that in the concept of superstatistics, Δ is fixed to a large time scale T. In addition, for every time slice with length T, the volatility β is defined as the local inverse variance. Because the T was determined as the optimal window size, in which the return is Gaussian distributed, the distribution of the return can be represented by a mixture of Gaussian distributions weighted by the density of β, $f(\beta)$.

As to the symbolic dynamics method, it can be applied to symbol sequences generated by time series of other complex systems as well. In this way the Renyi entropies associated with such a symbolic description allow for a quantitative comparison of the dynamical properties in symbol space, making it easy to compare different complex systems. In fact, one can in this way compare entirely different complex systems, for example the Rényi entropies associated with share price changes (using an alphabet of 4 symbols) can be compared with those of genomic

sequences [72]. The most important dynamical information is then encoded in form of the shape of the function $K(q)$, allowing the application of thermodynamic tools [66]. It is an advantage that a quantitative comparison of different systems is possible, solely based on the multifractal properties of the coarse-grained symbolic description.

Meanwhile, this study presented here encountered some limitations which need need to be considered. For example, in the original superstatistical framework, the integral over the volatility parameter β is performed under the assumption that those volatility parameters are changing independently on a given time scale. However, in fact there are nonnegligible correlations of the volatility, especially for data on the small time scales. In this case, a better fit for the integrated distribution $p(u)$ is obtained by adjusting the parameters in $p(u)$, rather than adopting the estimated parameter from $f(\beta)$ directly. This has, to some extent, added complexity to the operation of superstatistics. A similar issue is also recognized by a recent study on the superstatistics of vehicular traffic flow by C. Kosuna and S. Ozdemir(2016) [108]. Thus, a more general superstatistical model that releases the assumption of independent variation of the variance parameter might be desired.

Xu [132] explained the reason for the daily log-returns being modelled better than the minute data by superstatistics in terms of the problem of discreteness [103]. When the timespan τ, over which each data point is extracted, is below 5 minutes, there may not be enough data points during the period τ to implement the central limit theorem. In this situation, the assumption that the normalized log return being Gaussian in each window is not valid.

In the application of the symbolic dynamics technique for share prices, the length N of the symbol sequence is restricted to $N \leq 8$ for daily prices and small time scale prices due to the limited amount of data and computational resources that are limited. Larger N would induce large stochastic errors because of the fact that the statistics is not high enough to estimate the frequency of a given symbol sequence in a reliable way. To eliminate the data size limitation, a proper model, which emulates the evolution of share price based on the observed non-Markovian

character, is required. The analysis of an artificial time series will in turn help to build up the awareness of additional features of share prices.

A more comprehensive and profound understanding of the characters of asset prices contributes to constructing share price evolution models with improved accuracy and validity. The establishment of such models can further provide the foundations for plenty of further financial analysis such as option valuation and investment risk management and facilitates the construction of more efficient trading strategies. Biondo, Pluchino, Rapisarda and Helbing explored in an recent paper [146] that the performance of the most used trading strategies, which are based on the historical time series, is not better than a completely random strategy. This finding is remarkable in the research area of finance engineering and modern finance theory where the construction of a profitable investment strategy is still an open question. My study confirms the stylized facts of the historical share prices in terms of superstatistical and coarse-grained multifractal features on various time scales and developed a new model that fulfils the desired features.

Appendix A

6 companies' conditional probabilities

Table A.1: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Bac

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.441	0.557	0.561	0.440
d	0.451	0.544	0.572	0.429
u	0.434	0.570	0.549	0.450
d, d	0.458	0.535	0.559	0.451
d, u	0.421	0.577	0.554	0.448
u, d	0.442	0.553	0.587	0.414
u, u	0.456	0.547	0.533	0.464
d, d, d	0.464	0.539	0.530	0.477
d, d, u	0.429	0.566	0.559	0.450
d, u, d	0.428	0.566	0.605	0.400
d, u, u	0.442	0.565	0.560	0.449
u, d, d	0.459	0.537	0.564	0.429
u, d, u	0.398	0.591	0.558	0.433
u, u, d	0.446	0.545	0.582	0.431
u, u, u	0.475	0.541	0.526	0.461
d, d, d, d	0.478	0.520	0.521	0.476
d, d, d, u	0.435	0.555	0.530	0.473
d, d, u, d	0.445	0.538	0.564	0.438
d, d, u, u	0.439	0.547	0.533	0.455
d, u, d, d	0.463	0.552	0.560	0.422
d, u, d, u	0.386	0.592	0.586	0.442
d, u, u, d	0.438	0.549	0.579	0.413
d, u, u, u	0.463	0.523	0.522	0.458
u, d, d, d	0.453	0.532	0.570	0.462
u, d, d, u	0.433	0.569	0.570	0.444
u, d, u, d	0.444	0.593	0.621	0.382
u, d, u, u	0.440	0.579	0.555	0.463
u, u, d, d	0.486	0.528	0.558	0.444
u, u, d, u	0.429	0.572	0.534	0.467
u, u, u, d	0.457	0.533	0.541	0.434
u, u, u, u	0.486	0.521	0.529	0.470

Table A.2: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Ge

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.445	0.554	0.559	0.442
d	0.453	0.545	0.567	0.435
u	0.439	0.558	0.553	0.451
d, d	0.458	0.544	0.544	0.455
d, u	0.424	0.574	0.554	0.445
u, d	0.450	0.551	0.582	0.424
u, u	0.457	0.536	0.550	0.453
d, d, d	0.457	0.538	0.522	0.478
d, d, u	0.450	0.567	0.542	0.459
d, u, d	0.434	0.570	0.583	0.418
d, u, u	0.443	0.555	0.544	0.456
u, d, d	0.449	0.549	0.551	0.428
u, d, u	0.416	0.582	0.557	0.448
u, u, d	0.465	0.538	0.581	0.419
u, u, u	0.475	0.525	0.552	0.447
d, d, d, d	0.475	0.540	0.495	0.492
d, d, d, u	0.448	0.568	0.518	0.458
d, d, u, d	0.434	0.547	0.556	0.446
d, d, u, u	0.454	0.542	0.546	0.432
d, u, d, d	0.453	0.555	0.555	0.438
d, u, d, u	0.409	0.595	0.574	0.429
d, u, u, d	0.447	0.543	0.584	0.421
d, u, u, u	0.457	0.543	0.547	0.452
u, d, d, d	0.463	0.549	0.529	0.453
u, d, d, u	0.430	0.584	0.578	0.446
u, d, u, d	0.448	0.554	0.612	0.387
u, d, u, u	0.444	0.546	0.547	0.445
u, u, d, d	0.446	0.542	0.569	0.440
u, u, d, u	0.451	0.552	0.557	0.461
u, u, u, d	0.473	0.525	0.563	0.439
u, u, u, u	0.479	0.498	0.554	0.489

Table A.3: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Intc

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.448	0.551	0.556	0.445
d	0.450	0.552	0.556	0.442
u	0.446	0.554	0.555	0.446
d, d	0.451	0.551	0.540	0.452
d, u	0.437	0.565	0.551	0.444
u, d	0.453	0.553	0.566	0.438
u, u	0.456	0.540	0.552	0.449
d, d, d	0.442	0.541	0.530	0.477
d, d, u	0.438	0.574	0.544	0.453
d, u, d	0.443	0.561	0.570	0.431
d, u, u	0.453	0.562	0.560	0.448
u, d, d	0.461	0.542	0.548	0.440
u, d, u	0.433	0.569	0.539	0.431
u, u, d	0.457	0.543	0.561	0.439
u, u, u	0.465	0.536	0.560	0.444
d, d, d, d	0.462	0.568	0.510	0.482
d, d, d, u	0.438	0.562	0.521	0.439
d, d, u, d	0.446	0.561	0.545	0.447
d, d, u, u	0.455	0.565	0.559	0.457
d, u, d, d	0.444	0.568	0.543	0.434
d, u, d, u	0.426	0.583	0.546	0.433
d, u, u, d	0.461	0.540	0.557	0.425
d, u, u, u	0.460	0.523	0.566	0.455
u, d, d, d	0.436	0.519	0.550	0.450
u, d, d, u	0.448	0.558	0.564	0.449
u, d, u, d	0.443	0.564	0.578	0.425
u, d, u, u	0.456	0.565	0.555	0.441
u, u, d, d	0.471	0.534	0.565	0.430
u, u, d, u	0.443	0.560	0.572	0.435
u, u, u, d	0.473	0.539	0.553	0.439
u, u, u, u	0.491	0.507	0.569	0.448

Table A.4: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Jnj

	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.462	0.536	0.541	0.460
d	0.468	0.531	0.550	0.451
u	0.454	0.542	0.536	0.468
d, d	0.476	0.528	0.535	0.467
d, u	0.438	0.555	0.534	0.468
u, d	0.464	0.532	0.571	0.436
u, u	0.469	0.526	0.525	0.477
d, d, d	0.483	0.522	0.507	0.487
d, d, u	0.447	0.560	0.528	0.466
d, u, d	0.458	0.554	0.574	0.442
d, u, u	0.469	0.544	0.526	0.471
u, d, d	0.477	0.528	0.531	0.463
u, d, u	0.428	0.553	0.540	0.470
u, u, d	0.470	0.516	0.543	0.447
u, u, u	0.478	0.521	0.518	0.479
d, d, d, d	0.478	0.521	0.509	0.472
d, d, d, u	0.461	0.539	0.521	0.461
d, d, u, d	0.466	0.541	0.529	0.438
d, d, u, u	0.464	0.540	0.539	0.479
d, u, d, d	0.468	0.526	0.546	0.440
d, u, d, u	0.423	0.569	0.533	0.448
d, u, u, d	0.467	0.514	0.555	0.435
d, u, u, u	0.483	0.517	0.533	0.466
u, d, d, d	0.480	0.516	0.529	0.467
u, d, d, u	0.461	0.570	0.529	0.464
u, d, u, d	0.444	0.568	0.588	0.434
u, d, u, u	0.458	0.545	0.513	0.456
u, u, d, d	0.486	0.551	0.566	0.476
u, u, d, u	0.454	0.553	0.539	0.481
u, u, u, d	0.492	0.503	0.561	0.444
u, u, u, u	0.495	0.507	0.513	0.479

Table A.5: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Ko

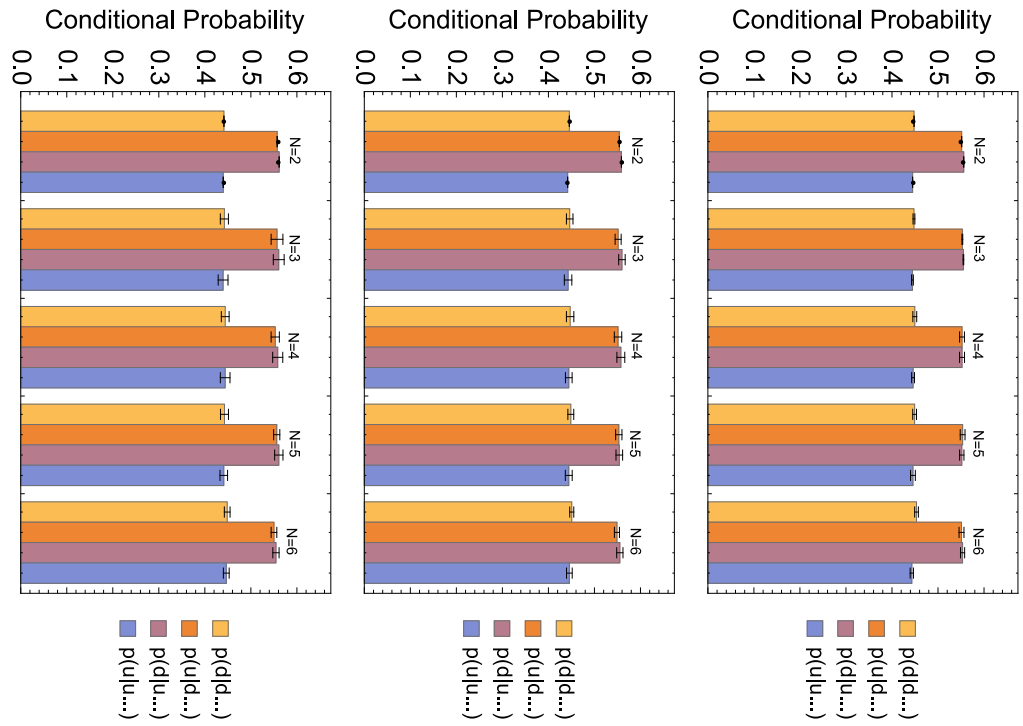
	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.454	0.548	0.540	0.458
d	0.467	0.534	0.556	0.440
u	0.442	0.560	0.526	0.475
d, d	0.476	0.531	0.534	0.460
d, u	0.429	0.569	0.532	0.473
u, d	0.462	0.542	0.567	0.430
u, u	0.455	0.547	0.515	0.480
d, d, d	0.479	0.517	0.518	0.470
d, d, u	0.445	0.561	0.520	0.483
d, u, d	0.456	0.550	0.577	0.420
d, u, u	0.449	0.570	0.515	0.463
u, d, d	0.468	0.540	0.552	0.454
u, d, u	0.419	0.580	0.544	0.451
u, u, d	0.462	0.524	0.542	0.442
u, u, u	0.481	0.544	0.514	0.490
d, d, d, d	0.490	0.537	0.528	0.488
d, d, d, u	0.458	0.540	0.514	0.478
d, d, u, d	0.459	0.522	0.579	0.431
d, d, u, u	0.460	0.553	0.517	0.478
d, u, d, d	0.459	0.535	0.558	0.441
d, u, d, u	0.408	0.599	0.555	0.450
d, u, u, d	0.457	0.554	0.543	0.406
d, u, u, u	0.470	0.540	0.531	0.488
u, d, d, d	0.470	0.508	0.515	0.476
u, d, d, u	0.430	0.561	0.517	0.468
u, d, u, d	0.453	0.563	0.587	0.401
u, d, u, u	0.432	0.572	0.540	0.466
u, u, d, d	0.466	0.544	0.546	0.460
u, u, d, u	0.457	0.577	0.523	0.493
u, u, u, d	0.468	0.515	0.536	0.451
u, u, u, u	0.463	0.524	0.494	0.509

Table A.6: Conditional probabilities of the 2-level symbol sequences with length up to $N = 6$ for 1-minute scale share price of Wmt

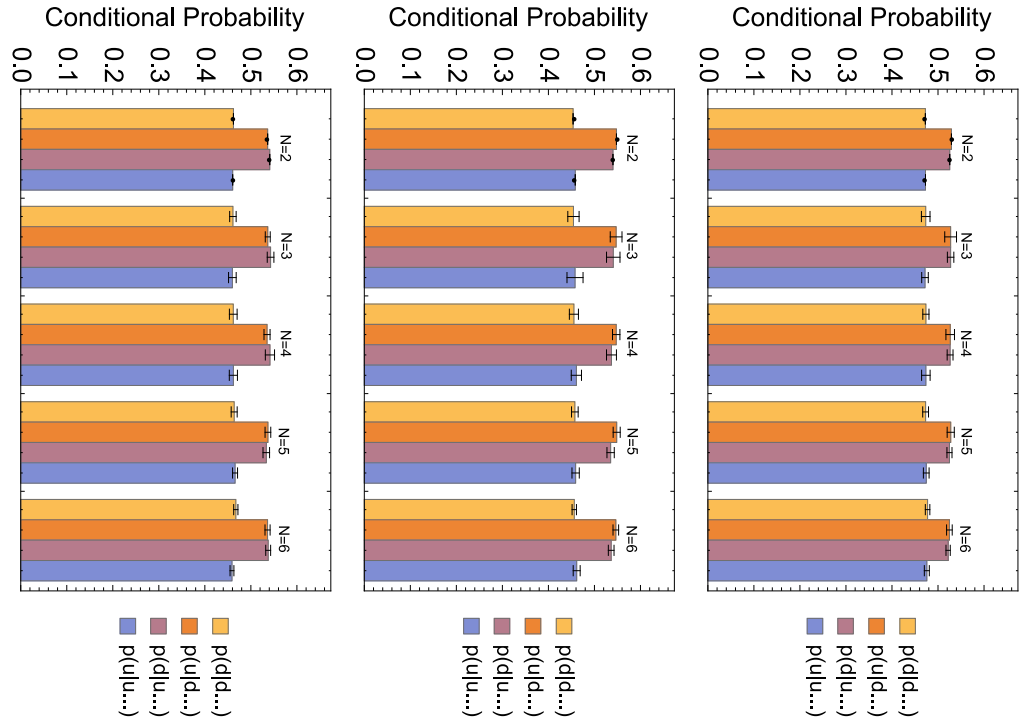
	$p(d\|d,\bullet)$	$p(u\|d,\bullet)$	$p(d\|u,\bullet)$	$p(u\|u,\bullet)$
	0.473	0.529	0.525	0.473
d	0.483	0.515	0.534	0.465
u	0.464	0.540	0.520	0.479
d, d	0.486	0.508	0.525	0.482
d, u	0.455	0.553	0.521	0.476
u, d	0.475	0.521	0.544	0.447
u, u	0.478	0.524	0.514	0.489
d, d, d	0.494	0.502	0.510	0.496
d, d, u	0.461	0.544	0.513	0.470
d, u, d	0.461	0.523	0.546	0.441
d, u, u	0.470	0.538	0.516	0.490
u, d, d	0.490	0.530	0.544	0.469
u, d, u	0.442	0.566	0.534	0.473
u, u, d	0.484	0.513	0.529	0.467
u, u, u	0.483	0.504	0.507	0.489
d, d, d, d	0.496	0.494	0.503	0.513
d, d, d, u	0.472	0.543	0.501	0.488
d, d, u, d	0.484	0.523	0.529	0.464
d, d, u, u	0.482	0.525	0.512	0.492
d, u, d, d	0.510	0.516	0.530	0.461
d, u, d, u	0.430	0.575	0.550	0.478
d, u, u, d	0.466	0.516	0.548	0.447
d, u, u, u	0.490	0.536	0.507	0.482
u, d, d, d	0.478	0.493	0.506	0.477
u, d, d, u	0.445	0.539	0.512	0.461
u, d, u, d	0.465	0.542	0.555	0.426
u, d, u, u	0.478	0.545	0.515	0.487
u, u, d, d	0.480	0.517	0.538	0.466
u, u, d, u	0.476	0.549	0.516	0.503
u, u, u, d	0.500	0.500	0.535	0.475
u, u, u, u	0.485	0.483	0.496	0.493

Appendix B

Bar charts of 6 companies' conditional probabilities



(a) Bank of America Corporation(bac) (b) General Electric Company(ge) (c) Intel Corporation(intc)



(d) Johnson&Johnson(jnj) (e) The Coca-Cola company(ko) (f) Wal-Mart Stores Inc(wmt)

Figure B.1: For $N = 2,...,6$, the mean values of the conditional probabilities of the 4 types($p(d|d...), p(u|d...), p(d|u...), p(u|u...)$) are represented by the height of the 4 bars(with error bars). No matter how long the symbol sequence is, the average probability of $p(u|d...)$ and $p(d|u...)$ is always higher than the average probability of $p(d|d...)$ and $p(u|u...)$. This phenomenon occurs for all the six companies.

Bibliography

[1] C Beck. Generalized statistical mechanics for superstatistical systems. *Phil. Trans. Royal Soc.*, 369(1935):453, 2011.

[2] E. V. der Straeten and C Beck. Superstatistical fluctuations in time series: Applications to share-price dynamics and turbulence. *Phys. Rev. E*, 80:036108, 2009.

[3] K. Briggs and C Beck. Modelling train delays with q-exponential functions. *Physica A*, 378:498, 2007.

[4] L. L. Chen and C Beck. A superstatistical model of metastasis and cancer survival. *Physica A*, 387:3162, 2008.

[5] V. Pareto. *Cours d'Economie Politique*. Lausanne and Paris, 1897.

[6] A. Einstein. On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-Kinetic Theory of Heat. *Ann. Physik*, 17:549–560, 1905.

[7] L. Bachelier. Theorie de la speculation. *Annales Scientifiques de I'Ecole Normale Superieure*, III-17:21–86, 1900.

[8] F Black and M. Scholes. The valuation of options and Corporate Liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[9] B. B. Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 26:394–419, 1963.

[10] E. F. Fama. Mandelbrot and the stable Paretian hypothesis. *Journal of Business*, 36:420–429, 1963.

[11] R. Reiss and M. Thomas. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhuser Basel, 1997.

[12] D. Jansen and C. de Vries. On the frequency of large stock market returns: putting booms and busts into perspective. *Review of Economics and Statistics*, 23:18–24, 1991.

[13] K. G. Koedijk, M. Schafgans, and C. de Vries. The tail index of exchange rate returns. *Journal of International Economics*, 29:93–108, 1990.

[14] T. Lux. The stable Paretian hypothesis and the frequency of large returns: An examination of major German stocks. *Applied Economics Letters*, 6:463–475, 1996.

[15] B. B. Mandelbrot. Long-run linearity, locally Gaussian processes, H-spectra and infinite variance. *International Economic Review*, 10:82–111, 1969.

[16] Z. Ding, C. Granger, and R. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106, 1993.

[17] M. M. Dacorogna, U. A. Müller, R. J. Nagler, R. B. Olsen, and O. V Pictet. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance*, 12:413–438, 1993.

[18] N. Crato and P. J. F. de Lima. Long-range dependence in the conditional variance of stock returns. *Economics Letters*, 45:281–285, 1994.

[19] T. Mills. Stylized facts of the temporal and distributional properties of daily FTSE returns. *Applied Financial Economics*, 7:599–604, 1997.

[20] P. Cizeau, Y. Liu, M. Meyer, C. K. Peng, and H. E. Stanley. Volatility Distribution in the S&P 500 Stock Index. *Physica A*, 245:441–445, 1997.

[21] Y. Liu, P. Cizeau, M. Meyer, C. K. Peng, and H. E. Stanley. Quantification of Correlations in Economic Time Series. *Physica A*, 245:437–440, 1997.

[22] M. Raberto, E. Scalas, G. Cuniberti, and M. Riani. Volatility in the Italian Stock Market: An Empirical Study. *Physica A*, 269:148–155, 1999.

[23] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[24] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.

[25] S. J. Taylor. *Modelling Financial Time Series*. Wiley, 1986.

[26] R. T. Baillie, T. Bollerslev, and H. O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74:3–30, 1996.

[27] F. J. Breidt, N. Crato, and P. de Lima. On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83:325–348, 1998.

[28] S. L. Heston. A Closed Solution For Options With Stochastic Volatility, With Application to Bond and Currency Options. *Review of Financial Studies*, 6(2):327–343, 1993.

[29] M.M. Dacorogna, R. Gençay, U. A. Müller, R.B. Olsen, and O.V. Pictet. *An Introduction to High-Frequency Finance*. CA: Academic Press, San Diego, 2001.

[30] J. P. Bouchaud and M. Potters. *Theories des Risques Financiers*. Eyrolles, Alea-Saclay, 1997.

[31] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge. Turbulent Cascades in Foreign Exchange Markets. *Nature*, 381:767–770, 1996.

[32] P. Gopikrishnan, M. Meyer, L. A. N. Amaral, and H. E. Stanley. Inverse Cubic Law for the Distribution of Stock Price Variations. *Eur. Phys. J. B*, 3:139–140, 1998.

[33] R. N. Mantegna. Levy Walks and Enhanced Diffusion in Milan Stock Exchange. *Physica A*, 179:232–242, 1991.

[34] R. N. Mantegna and H. E. Stanley. Scaling Behavior in the Dynamics of an Economic Index. *Nature*, 376:46–49, 1995.

[35] R. N. Mantegna and H. E. Stanley. Turbulence and Financial Markets. *Nature*, 383:587–588, 1996.

[36] R. N. Mantegna and H. E. Stanley. Stock Market Dynamics and Turbulence: Parallel Analysis of Fluctuation Phenomena. *Physica A*, 239:255–266, 1997.

[37] C Beck and E. G. D. Cohen. Superstatistics. *Physica A*, 322:267–275, 2003.

[38] E. G. D. Cohen. Superstatistics. *Physica D*, 139(1):35–52, 2004.

[39] C Beck, E. G. D. Cohen, and H. L. Swinney. From time series to superstatistics. *Phys. Rev. E*, 72:056133, 2005.

[40] A. Reynolds. Superstatistical mechanics of tracer-particle motions in turbulence. *Phys. Rev. Lett.*, 91:084503, 2003.

[41] C Beck. Statistics of 3-dimensional Lagrangian turbulence. *Phys. Rev. Lett.*, 98:064502, 2007.

[42] C Beck. Lagrangian acceleration statistics in turbulent flows. *Europhys. Lett.*, 64(2):151–157, 2003.

[43] S. Rizzo and A. Rapisarda. Environmental atmospheric turbulence at Florence airport. *AIP Conf. Proc.*, 742:176–181, 2004.

[44] T. Laubrich, F. Ghasemi, J. Peinke, and H. Kantz. Statistical analysis of wind speed fluctuation and increments of non-stationary atmospheric boundary layer turbulence. *eprint arXiv: 0811.3337*, 2008.

[45] P. Rabassa and C Beck. Superstatistical analysis of sealevel fluctuations. *Physica A*, 417:18, 2015.

[46] C. Yalcin, P Rabassa, and C Beck. Extreme event statistics of daily rainfall: Dynamical systems approach. *J. Phys. A Math. Theor.*, 49:154001, 2016.

[47] A. Chakraborti, I.M. Toke, M. Patriarca, and F. Abergel. Econophysics review: I. empirical facts. *Quantitative Finance*, 11:991, 2010.

[48] P. Suárez-García and D.Gómez-Ullate. Scaling, stability and distribution of the high-frequency returns of the ibex35 index. *Physica A*, 392:1409, 2013.

[49] M.Denys, T.Gubiec, R.Kutner, M. Jagielski, and H.E. Stanley. Universality of market superstatistics. *Phys. Rev. E*, 94:042305, 2016.

[50] S. M. D. Queiros and C. Tsallis. On the connection between financial processes with stochastic volatility and nonextensive statistical mechanics. *Eur. Phys. J. B*, 48:139, 2005.

[51] C. Tsallis. Inter-occurrence times and universal laws in finance, earthquakes and genomes. *Chaos Solitions & Fractas*, 88, 2016.

[52] S. Devi. Financial market dynamics: Superdiffusive or not? *arXiv:1608.07752*, 2016.

[53] J. Barunik, T. Aste, T. Di Matteo, and R. Liu. Understanding the source of multifractality in financial markets. *Physica A*, 391:4234, 2012.

[54] P. Suárez-García and D. Gómez-Ullate. Multifractality and long memory of a financial index. *Physica A*, 394:226, 2014.

[55] A. Carbone, G. Castelli, and H. Stanley. Time-dependent Hurst exponent in financial time series. *Physica A*, 344:267, 2004.

[56] T. Di Matteo. Multi-scaling in finance. *Quantitative Finance*, 7:21, 2007.

[57] B. B. Mandelbrot and J. R. Wallis. Noah, Joseph, and operational hydrology. *Water Resources Research*, 4, 1968.

[58] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. *Phys. Rev. E*, 49:1685–1689, 1994.

[59] N. Nava, T. Di Matteo, and T. Aste. Anomalous volatility scaling in high frequency financial data. *Physica A*, 447:434–445, 2016.

[60] R.J. Buonocore, T. Aste, and T. Di Matteo. Measuring multiscaling in financial time-series. *Chaos, Solitons and Fractals*, 88:38, 2016.

[61] J. W. Kantelhardt, S. A. Zschiegner, E. K-Bunde, S. Havlin, A. Bunde, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316:87–114, 2002.

[62] G.-F. Gu and W.-X. Zhou. Detrending moving average algorithm for multifractals. *Phys.Rev.E*, 82:1, 2010.

[63] J. Feder. *Fractals*. Plenum Press, New York, 1988.

[64] J. F. Muzy, E. Bacry, and A. Arneodo. Wavelets and multifractal formalism for singular sigulars: application to turbulence data. *Physical Review Letters*, 67:3515, 1991.

[65] J. F. Muzy, E. Bacry, and A. Arneodo. The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos*, 4:245, 1994.

[66] C. Beck and F. Schögl. *Thermodynamics of Chaotic Systems An Introduction*. Number 4 in Cambridge Nonlinear Science Series. Cambridge University Press, 1993.

[67] J. Barunik and L. Kristoufek. On Hurst exponent estimation under heavy-tailed distributions. *Physica A*, 389:3844, 2010.

[68] A. Eke, P. Herman, B. G. Sanganahalli, F. Hyder, P. Mukli, and Z. Nagy. Pitfalls in fractal time series analysis: fMRI BOLD as an exemplary case. *Front. Physiol.*, 3:417, 2012.

[69] H. Salat, R. Murcio, and E. Arcaute. Multifractal methodology. *arXiv:1606.02957*, Submitted on 8 Jun 2016.

[70] Z. G. Yu, V. Anh, and K. S. Lau. Measure representation and multifractal analysis of complete genomes. *Phys. Rev. E*, 64:031903, 2001.

[71] Z. G. Yu, V. Anh, and K. S. Lau. Multifractal and correlation analyses of protein sequences from complete genomes. *Phys. Rev. E*, 68:021913, 2003.

[72] C. Beck and A. Provata. Multifractal information production of the human genome. *Europhys. Lett.*, 95:58002, 2011.

[73] C. Beck and A. Provata. Multifractal analysis of nonhyperbolic coupled map lattices: Application to genomic sequences. *Phys. Rev. E*, 83:066210, 2011.

[74] P. Jizba and T. Arimitsu. The world according to Rényi: thermodynamics of multifractal systems. *Ann. Phys.*, 312(1):17–59, 2004.

[75] P. Tino, C. Schittenkopf, G. Dorffner, and E. J. Dockner. A Symbolic Dynamics Approach to Volatility Prediction. In *IN COMPUTATIONAL FINANCE, (PROCEEDINGS OF THE SIXTH INTERNATIONAL CONFERENCE ON COMPUTATIONAL FINANCE), LEONARD N. STERN SCHOOL OF BUSINESS*, pages 6–8. MIT Press, 1999.

[76] C. Bandt and B. Pompe. Permutation entropy:A natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102, 2002.

[77] M. Zanin. Forbidden patterns in financial time series. *Chaos*, 18:013119, 2008.

[78] D. Xu and C. Beck. Transition from lognormal to χ^2-superstatistics for financial time series. *Physica A*, 453:173–183, 2016.

[79] R. R. Officer. The variability of the market factor of the New York Stock Exchange. *Journal of Business*, 46:434–453, 1973.

[80] G. W. Schwert. Stock volatility and the crash of 87. *Review of Financial Studies*, 3:77–102, 1990.

[81] H. J Voth. Political risk and stock price volatility: evidence from 10 countries during the interwar period. 2003.

[82] A. A. Christie. The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics*, 10:407–432, 1982.

[83] G. R. Duffee. Stock returns and volatility: a firm-level analysis. *Journal of Financial Economics*, 37:399–420, 1995.

[84] K. Clark. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41:135–155, 1973.

[85] T. W. Epps and M. L. Epps. The stochastic dependence of security price changes and transaction volumes: implications for the mixture-of-distributions hypothesis. *Econometrica*, 44:305–321, 1976.

[86] S. J. Taylor. *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press, Princeton and Oxford, 2007.

[87] D. B. Nelson. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59:347–370, 1991.

[88] T. Bollerslev, R. Y. Chou, and K. F. Kroner. ARCH modeling in finance: a review of the theory and empirical evidence. *Journal of Econometrics*, 52:5–59, 1992.

[89] T. Bollerslev, R. F. Engle, and D. B. Nelson. ARCH models. In *Handbook of Econometrics*, volume 4, pages 2959–3038. North Holland, Amsterdam, 1994.

[90] S. E. Shreve. *Stochastic Calculus for Finance II Continuous-Time Models*. Springer-Verlag, New York, 2004.

[91] L. O. Scott. Option pricing when the variance changes randomly: theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22:419–438, 1987.

[92] J. B. Wiggins. Option values under stochastic volatility: theory and empirical estimates. *Journal of Financial Economics*, 19:351–372, 1987.

[93] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North Holland, Amsterdam, 1981.

[94] J. C. Cox, J. E. Ingersoll, and S. A. Ross. The relationship between forward prices and futures prices. *Journal of Financial Economics*, 9:321–346, 1981.

[95] R. Cont and P. Tankov. *Financial modelling with Jump Processes*. Chapman & Hall / CRC, 2003.

[96] A. Etheridge. *A course in financial calculus*. Cambridge University Press, 2002.

[97] N. Shephard, editor. *Stochastic Volatility: Selected Readings*. Oxford University Press, 2005.

[98] N. Shephard. Statistical aspects of ARCH and stochastic volatility. In *Time Series Models: In econometrics, finance and other fields*, pages 1–67. Chapman & Hall, London, 1996.

[99] T. Ané and H. Geman. Order flow, transaction clock and normality of asset returns. *Journal of Finance*, 55:2259–2284, 2000.

[100] G.E. Tauchen and M. Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, 51:485–505, 1983.

[101] P.D. Praetz. The distribution of share price changes. *Journal of Business*, 45:49–55, 1972.

[102] D.B. Madan and E. Seneta. The variance gamma model for share market returns. *Journal of Business*, 63:511–524, 1990.

[103] O.E. Barndorff-Nielson and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(2):167–241, 2001.

[104] C Beck. Superstatistics: theory and applications. *Continuum Mech. Thermodyn.*, 16:293, 2004.

[105] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, 52:479, 1988.

[106] S. Miah and C. Beck. Lagrangian quantum turbulence model based on alternating superfluid/normal fluid stochastic dynamics. *Europhys. Lett.*, 108:40004, 2014.

[107] C. Beck and S. Miah. Statistics of Lagrangian quantum turbulence. *Phys. Rev. E*, 87:031002, 2013.

[108] C. Kosun and S. Ozdemir. A superstatistical model of vehicular traffic flow. *Physica A*, 444:466–475, 2016.

[109] P. Mehlen and A. Puisieux. Metastasis: a question of life or death. *Nature Rev. Cancer*, 6(6):449–458, 2006.

[110] S. M. Ross. *Stochastic Processes*. John & Wiley Sons, 2nd edition, 1996.

[111] C. Metzner, C. Mark, J. Steinwachs, L. Lautscham, F. Stadler, and B. Fabry. Superstatistical analysis and modelling of heterogeneous random walks. *Nat. Commun.*, 6:7516, 2015.

[112] S. M. D. Queiros. On the emergence of a generalized Gamma distribution. Application to traded volume in financial markets. *Europhys. Lett.*, 71:339, 2005.

[113] M. Ausloos and K. Ivanova. Dynamical model and nonextensive statistical mechanics of a market index on large time windows. *Phys. Rev. E*, 68:046122, 2003.

[114] T. S. Biro and R. Rosenfeld. Microscopic Origin of Non-Gaussian Distributions of Financial Returns. *Physica A*, 387:1603–1612, 2008.

[115] Y. A. Katz and L. Tian. q-Gaussian distributions of leverage returns, first stopping times, and default risk. *Physica A*, 392:4989, 2013.

[116] Y. A. Katz and L. Tian. Superstatistical fluctuations in time series of leverage returns. *Physica A*, 405:326, 2014.

[117] A. Gerig, J. Vicente, and M. A. Fuentes. Model for Non-Gaussian Intraday Stock Returns. *Phys. Rev. E*, 80:065102, 2009.

[118] C. Vamos and M. Craciun. Separation of components from a scale mixture of Gaussian white noises. *Phys. Rev. E*, 81:051125, 2010.

[119] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi. *Wavelets and their Applications*. ISTE Ltd, London and Newport Beach,CA, 2007.

[120] P. Oświęecimka, J. Kwapień, and S. Drożdż. Wavelet versus detrended fluctuation analysis of multifractal structures. *Phys. Rev. E*, 74:016103, 2006.

[121] C. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[122] N. Wiener. *Cybernetics or Control and Communication in the Animal and the Machine*. Wiley, New York, 1949.

[123] A. Rényi. On Measures of Entropy and Information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 547–561. University of California Press, Berkeley, Calif., 1961.

[124] A. Rényi. *Probability Theory*. North-Holland, Amsterdam, 1979.

[125] C. Beck. Upper and lower bounds on the Renyi dimensions and the uniformity of multifractals. *Physica D*, 41:67–68, 1990.

[126] A. N. Kolmogorov. A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Dokl. Akad. Sci. USSR*, 119:861, 1958.

[127] Ya. G. Sinai. On the concept of entropy for a dynamic system. *Dokl. Akad. Sci. USSR*, 124:768, 1959.

[128] C. Tsallis. *Introduction to Nonextensive Statistical Mechanis: Approaching a Complex World*. Springer, 2009.

[129] S. J. Camilleri. Month-related seasonality of stock price volatility: Evidence from the Malta stock exchange. *Bank of Valletta Review*, 37:49, 2008.

[130] C. Beck. Chaotic cascade model for turbulent velocity distributions. *Phys. Rev. E*, 49:3641, 1994.

[131] R. Vicente, C. M. de Toledo, V. B. P. Leite, and N. Caticha. Underlying dynamics of typical fluctuations of an emerging market price index: The Heston model from minutes to months. *Physica A*, 361:272, 2006.

[132] T.-L. Xu. Modelling High-frequency Non-Gaussian Share Price Returns Using Superstatistics. Msc thesis, QMUL. July 2016.

[133] J Cotter. Uncovering long memory in high frequency UK futures. 2004.

[134] T. G. Andersen and T. Bollerslev. Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns. *Journal of Finance*, 52(3):975–1005, 1997.

[135] T. G. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. *J. Empirical Finance*, 4:115, 1997.

[136] T. Bollerslev, J. Cai, and F. M. Song. Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the US Treasury bond market. *J. Empirical Finance*, 7:37, 2000.

[137] S. H. Kang and S. M Yoon. Long memory features in the high frequency data of the Korean stock market. *Physica A*, 387:5189, 2008.

[138] K. P. Evans and A. E. H. Speight. Intraday periodicity, calendar and announcement effects in Euro exchange rate volatility. *Res. Int. Business Finance*, 24:82, 2010.

[139] U. Tirnakli, C. Tsallis, and C. Beck. A closer look at time averages of the logistic map at the edge of chaos. *Phys. Rev. E*, 79:056209, 2009.

[140] R. Morales, T. Di Matteo, R. Gramatica, and T. Aste. Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series. *Physica A*, 391:3180, 2012.

[141] C. Alexander. Normal mixture diffusion with uncertain volatility: Modelling short-and long-term smile effects. *J. Bank. Financ.*, 28(12):2957, 2004.

[142] K. Tan and Shozo Tokinaga. Identifying Returns Distribution by Using Mixture Distribution Optimized by Genetic Algorithm. pages 119–122, 2006.

[143] R. Reno and R. Rizza. Is volatility lognormal? Evidence from Italian futures. *Physica A*, 322:620, 2003.

[144] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C. K. Peng, and H. E. Stanley. Statistical properties of the volatility of price fluctuations. *Phys. Rev. E*, 60:1390, 1999.

[145] T. Takaishi. Analysis of Realized Volatility in Superstatistics. *Evol. Inst. Econ. Rev.*, 7(1):89–99, 2010.

[146] A.E. Biondo, A. Pluchino, A. Rapisarda, and D. Helbing. Are random trading strategies more successful than technical ones. *PloS one*, 8(7):e68344, 2013.