UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

Fakultät für Physik

DISSERTATION

Dependencies and non-stationarity in financial time series

Von der Fakultät für Physik
der Universität Duisburg-Essen
genehmigte Dissertation
zur Erlangung des Grades
Dr. rer. nat.
von

Dipl.-Phys. Desislava Chetalova

aus Haskovo, Bulgarien

Tag der Disputation: 25. November 2015
Erstgutachter: Prof. Dr. Thomas Guhr
Zweitgutachter: PD Dr. Jens Christian Claussen

Hiermit versichere ich, die vorliegende Dissertation selbstständig, ohne fremde Hilfe und ohne Benutzung anderer als den angegebenen Quellen angefertigt zu haben. Alle aus fremden Werken direkt oder indirekt übernommenen Stellen sind als solche gekennzeichnet.

Die vorliegende Dissertation wurde in keinem anderen Promotionsverfahren eingereicht.

Mit dieser Arbeit strebe ich die Erlangung des akademischen Grades "Doktor der Naturwissenschaften" (Dr. rer. nat.) an.

Ort, Datum	Desislava Chetalova

List of publications

Parts of this thesis are included in the following publications:

[1] Thilo A. Schmitt, Desislava Chetalova, Rudi Schäfer and Thomas Guhr
Non-stationarity in financial time series: Generic features and tail behavior
Europhysics Letters 103, 58003 (2013)

[2] Desislava Chetalova, Thilo A. Schmitt, Rudi Schäfer and Thomas Guhr
Portfolio return distributions: Sample statistics with stochastic correlations
International Journal of Theoretical and Applied Finance 18(2), 1550012 (2015)

[3] Desislava Chetalova, Rudi Schäfer and Thomas Guhr
Zooming into market states
Journal of Statistical Mechanics, P01029 (2015)

[4] Desislava Chetalova, Marcel Wollschläger and Rudi Schäfer
Dependence structure of market states
Journal of Statistical Mechanics, P08012 (2015)

The following publication is not part of this thesis:

[5] Thilo A. Schmitt, Desislava Chetalova, Rudi Schäfer and Thomas Guhr
Credit risk and the instability of the financial system: An ensemble approach
Europhysics Letters 105, 38004 (2014)

Author contributions

Here, I describe my contributions to the publications [1]-[4]:

[1] The letter introduces the correlation averaged multivariate normal distribution, summarizes its analytical derivation and presents an application to financial returns. The project was carried out in collaboration with T. Schmitt and under the supervision of R. Schäfer and T. Guhr. The text was written by T. Guhr with contributions from T. Schmitt, R. Schäfer and me.

[2] The paper provides the complete analytical derivation of the correlation averaged multivariate normal distribution and presents an application to portfolio returns. The project was carried out in collaboration with T. Schmitt and under the supervision of R. Schäfer and T. Guhr. I calculated the average portfolio return distribution and contributed the data analysis. The text was written by me with contributions from T. Schmitt, R. Schäfer and T. Guhr.

[3] I carried out the empirical study on market states. The text was written by me with contributions from R. Schäfer and T. Guhr. The project was supervised by R. Schäfer and T. Guhr.

[4] I carried out the empirical study on copulas. The text was written by me with contributions from M. Wollschläger and R. Schäfer. The project was initiated and supervised by R. Schäfer.

Acknowledgements

First and foremost, I want to express my deepest gratitude to my supervisor Thomas Guhr for giving me the great opportunity to work in his group on such an exciting new field like econophysics. I would like to thank him for his guidance and constant support during the years of my PhD research.

I also would like to thank my second advisor Rudi Schäfer from whom I learned a lot about empirical data and financial markets. I am deeply grateful for his patience and advice throughout these years.

I thank all current and former members of our group, especially Maram Akila, Sabine Lukas, Thilo Schmitt, Yuriy Stepanov, Daniel Wagner, Daniel Waltner, Shanshan Wang, Marcel Wollschläger and of course my office college Tim Wirtz for the nice work atmosphere.

Furthermore, I want to thank Mohammad Assadsolimani, Rudi Schäfer, Yuriy Stepanov and Tim Wirtz who took the time to proofread parts of this thesis and provided valuable feedback.

Finally, I thank my family and my friends for their support throughout these years. Very special thanks go to Mohammad for his constant encouragement and belief in me without which this work would have never been finished.

Abstract

The growing interest of physicists in economic problems has led to the emergence of a new interdisciplinary field called econophysics. It applies methods and concepts of statistical physics to study economic and financial phenomena. This thesis contributes to the research activities in econophysics, focusing mainly on the study of financial markets. Financial markets can be viewed as complex systems which exhibit highly non-stationary behavior. We concentrate in particular on the non-stationarity of correlations between companies.

We begin by introducing an approach to model the non-stationarity of correlations. This approach is based on concepts from random matrix theory and allows us to construct a correlation averaged multivariate normal distribution which takes the non-stationary correlations into account. We perform an empirical study to verify the random matrix approach for financial returns.

We consider two applications of the random matrix approach. First, we study the effect of non-stationary correlations on portfolios, and in particular on the distribution of portfolio returns. We derive an average portfolio return distribution which we compare with the portfolio returns of randomly selected portfolios. Second, we investigate the stability of the correlation structure of market states. The random matrix approach reduces the complexity of the financial market to a single parameter which characterizes the correlation fluctuations due to non-stationarity. This allows us to estimate the fluctuation strength of correlations directly from the empirical return distributions and thus to assess the stability of the correlation structure.

We further extend our market states analysis by studying the empirical dependence structure of market states. We estimate and study empirical pairwise copulas for each market state. In addition, we derive a copula which arises from the correlation averaged multivariate normal distribution and compare it with the empirical pairwise copulas of each state. The results confirm the consistency of our random matrix model with financial data once again.

Finally, we focus on the extreme value statistics of correlated random variables. We derive a distribution for the maximum of correlated normal random variables and extend the result to the non-normal case. We verify our findings in numerical simulations.

Zusammenfassung

Das wachsende Interesse von Physikern an wirtschaftswissenschaftlichen Fragestellungen hat zur Entstehung des neuen interdisziplinären Forschungsgebiets Wirtschaftsphysik geführt. Wirtschaftsphysik nutzt Methoden und Konzepte der statistischen Physik, um wirtschaftswissenschaftliche Phänomene zu studieren. Diese Arbeit trägt zu den Forschungsaktivitäten in Wirtschaftsphysik bei. Der Schwerpunkt liegt dabei auf Finanzmärkten. Diese werden als komplexe Systeme betrachtet, die ein nichtstationäres Verhalten aufweisen. Wir konzentrieren uns insbesondere auf die Nichtstationarität von Korrelationen zwischen Unternehmen.

Zunächst führen wir einen Ansatz zur Modellierung der Nichtstationarität von Korrelationen ein. Dieser beruht auf Konzepte der Zufallsmatrixtheorie und ermöglicht uns eine korrelationsgemittelte multivariate Normalverteilung zu konstruieren, die die nichtstationären Korrelationen mitberücksichtigt. Wir führen eine empirische Studie durch, um den Zufallsmatrixansatz für Aktienrenditen zu verifizieren.

Wir betrachten zwei Anwendungen des Zufallsmatrixansatzes. Als Erstes untersuchen wir den Einfluss der nichtstationären Korrelationen auf Portfolios und insbesondere auf die Verteilung von Portfoliorenditen. Wir leiten eine gemittelte Portfoliorenditeverteilung her, die wir mit den Portfoliorenditen von zufällig zusammengesetzten Portfolios vergleichen. Als Nächstes untersuchen wir die Stabilität der Korrelationsstruktur von Marktzuständen. Der Zufallsmatrixansatz reduziert die Komplexität des Finanzmarktes auf einen einzigen Parameter, der die Fluktuationen der Korrelationen aufgrund der Nichtstationarität charakterisiert. Dies ermöglicht uns die Fluktuationsstärke der Korrelationen direkt aus den empirischen Renditeverteilungen zu schätzen und somit die Stabilität der Korrelationsstruktur zu studieren.

Weiterhin untersuchen wir die Abhängigkeitsstruktur von Marktzuständen. Wir schätzen empirische Copulas für jeden Marktzustand und vergleichen diese mit der Copula, die wir aus der korrelationsgemittelten multivariaten Normalverteilung herleiten. Die Ergebnisse bestätigen erneut die Übereinstimmung unseres Zufallsmatrixmodells mit Finanzdaten.

Zum Schluss befassen wir uns mit der Extremwertstatistik von korrelierten Zufallsvariablen. Wir leiten eine Verteilung für das Maximum von korrelierten normalverteilten Zufallsvariablen her und erweitern das Ergebnis für nicht normalverteilte Zufallsvariablen. Numerische Simulationen bestätigen unsere Ergebnisse.

Contents

Chapter 1

Introduction

1.1 Econophysics – economy and physics

In the last few decades, physicists have developed an increasing interest in economic and financial phenomena. This led to the emergence of a new interdisciplinary research field called econophysics. The term econophysics, a hybrid of the words economy and physics, was first introduced by H. Eugene Stanley at a conference on statistical physics in 1995 to draw attention to the large number of papers written by physicists on the problems of financial markets. Naturally, the question arises – what is actually econophysics? One of the first definitions of econophysics was given by Stanley and Mantegna [6], namely

> "The word econophysics describes the present attempts of a number of physicists to model financial and economic systems using paradigms and tools borrowed from theoretical and statistical physics."

According to another definition [7]

> "Econophysics is an approach to quantitative economy using ideas, models, conceptual and computational methods of statistical physics."

However, econophysics is much more than the application of physical methods to economic problems, it is a new way of thinking about economic and financial systems as complex systems whose internal microscopic interactions generate their macroscopic properties [8].

Although econophysics has established itself only in the recent decades, the involvement of physicists and mathematicians in economic problems has a long history [9], dating back at least to the 16th century when none other than Copernicus studied the behavior of inflation. A century later, another famous astronomer

and mathematician, Edmond Halley, derived the foundations of life insurance. In 1738, Daniel Bernoulli introduced the idea of utility to describe people's preferences. Pierre-Simon Laplace (1812) pointed out that events that seem random and unpredictable can be shown to obey certain laws. This idea was taken up by the astronomer Adolphe Quetelet in 1835 who studied patterns in data sets ranging from economic to social problems and laid the foundations of social physics.

On March 29, 1900 the French mathematician Louis Bachelier, a student of Henri Poincaré, defended his PhD thesis, entitled "Théorie de la Spéculation" [10]. There, he proposed the random walk as a model for price fluctuations while studying the erratic motion of bonds and stock options on the Paris stock exchange. Bachelier developed the mathematical theory of diffusion five years before Albert Einstein's famous work on Brownian motion in 1905 [11]. Furthermore, he constructed a model for French government bonds and provided the first formulation of option pricing in speculative markets. Unfortunately, Bachelier's contributions were not recognized by the scientific community at the time. His work was rediscovered in the late 1950s by the American astrophysicist M. F. Maury Osborne who suggested that stock prices follow a geometric Brownian motion [12]. In the following years, the geometric Brownian motion became an important model for financial markets [13]. In 1973, the physicist Fisher Black together with Myron Scholes, as well as Robert Merton independently of both of them, used the geometric Brownian motion to develop a theory for determining the price of a stock option [14, 15]. Although flawed this theory became an indispensable tool in the capital market practice. One should also mention the work of another French mathematician, Benoit Mandelbrot, mostly known for his studies of fractals, who in the early 1960s pioneered the use of power law distributions for stock price changes [16].

Although physicists had been occasionally involved in economic problems for centuries, it was not until the 1990s that they started seriously turning to economics, and particularly finance. The cause of this turn of events was, on the one hand, the limitations of the traditional approaches in economics and finance, and on the other hand, the evolution of technology and the computerization of trading which led to the storage of huge amounts of data. The availability of the empirical data triggered the interest of theoretical physicists. They treated financial markets as complex systems, applying methods and tools from statistical physics to extract information from the financial data and construct mathematical models based on this information. Along with that, a variety of analogies between financial markets and physical phenomena were identified, e.g., turbulence [17, 18], spin systems [19], universality [20], self-organized criticality [21] and so on. The increasing number of publications marked the emergence of econophysics as a new branch of physics.

Most of the research in econophysics has been focused on the statistical properties of financial time series and their modeling. Several studies have been performed to study the distribution of price changes [22–24], the temporal memory [23, 25], the

cross-correlations among companies [26–28]. Another important area of research concerns the development of theoretical models for financial markets, so-called agent-based models [29–31]. Such models investigate the financial market dynamics from the point of view of single agents and aim at deriving the macroscopic properties of the system from the microscopic interactions. Other areas which undergo intensive investigations deal with the pricing of derivative products [32–34], with aspects of portfolio selection and optimization [35–37]. In addition, there are also studies on broader economic topics, e.g., the distributions of income and size of companies [38, 39], wealth and money distributions of societies [40–42]. These are just a few of the efforts in the field of econophysics. For a more complete overview, the reader is referred to references [43, 44].

Here, we focus our research on the analysis of financial markets, concentrating in particular on the issue of non-stationarity and its consequences.

1.2 Financial markets

As financial markets will be the central topic of this thesis, we begin with a brief review of some basic concepts and terminology. We will focus on stock markets and discuss the mechanism of trading and the formation of the price, whose evolution over time resembles the path of a stochastic process. In statistical analysis, however, price changes, or returns, are the preferred observable. We will summarize some well-established empirical facts about return time series and introduce two important quantities, namely volatility and correlations. In particular, correlations between stocks will play an important role in this thesis. We will learn how to measure correlations and find that they change considerably over time. This fundamental non-stationarity of the financial market will be addressed in the next chapters. Finally, we will briefly discuss risk and the importance of correlations for portfolio optimization. The concepts covered in this section form the foundation of this thesis. For a more detailed introduction into the theory of financial markets, the reader is referred to references [33, 45–47].

1.2.1 Basic concepts

A market is a place where traders meet to exchange products. At each time t, a product has a certain price $S(t)$, which in a free economy is determined by the interplay of supply and demand. If the demand of a product increases (decreases) or the supply falls (rises), the price of the product goes up (down). This interplay between demand and supply generates the price of a product, also called asset.

There are different kinds of markets according to the assets which are traded. A financial market is a market where roughly speaking money is traded in form of different financial instruments. We distinguish between stock, derivative, bond

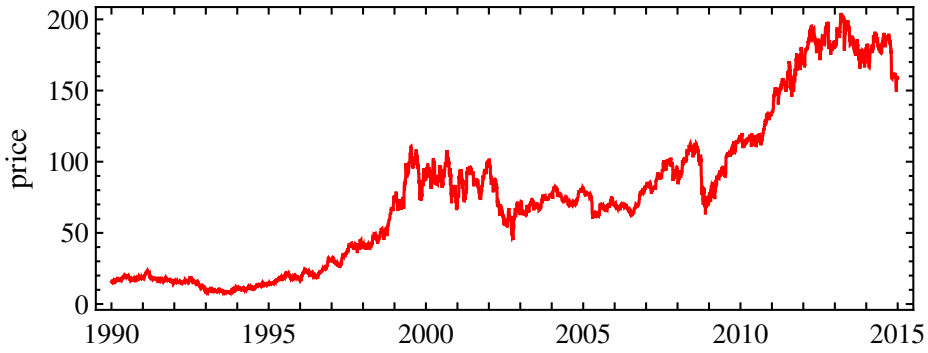


Figure 1.1: Time evolution of the daily price of the IBM stock in the time period 1990 to 2014.

and currency markets named according to the traded financial instruments. In this thesis, we restrict ourselves to stock markets or stock exchanges, where stocks of companies are traded.

Stocks are shares in the ownership of a company. Companies sell stocks to raise investment capital, for instance, to expand their economic activities, without having the obligation to pay the money back. Buying stocks makes the investors co-owners of the company, that is, they have a vote in its strategic decisions and policy. On the other hand, investors buy stocks to gain profit. There are two ways to obtain profit from stocks: First, through dividends, if a company makes a huge profit, it can distribute a part of this profit to the shareholders. Second, if the stock price goes up, the investor can sell it obtaining the difference between the buying and the selling price as a profit. The value of the stock goes up and down according to demand and supply. Typically, the present performance of the company as well as the market expectations of the traders are reflected in the stock price, that means, the better the performance or the higher the market expectations, the higher the price. As supply and demand change, the price of a stock fluctuates over time, see figure 1.1.

Stocks are commonly grouped into stock indices. A stock index is defined as a weighted average over the stocks of selected companies. It is determined by the kind of companies included in the index. For example, one can choose the biggest or the most successful companies in certain countries, like Dow Jones Industrial Average (DJIA), Standard and Poor's 500 (S&P 500) and the Deutscher Aktien Index (DAX), or one can define indices on special markets, for example NASDAQ Composite is an index comprised mainly of technology and growth companies.

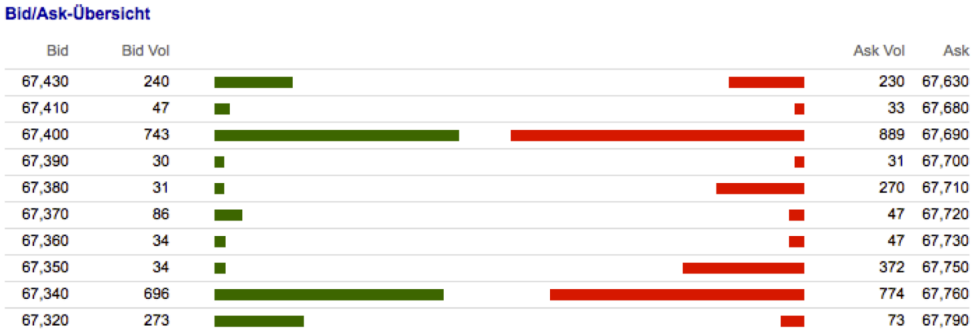


Figure 1.2: An excerpt of the order book of the Xetra stock exchange for the BMW AG stock at some instant of time. On the left hand side are the orders to buy (bids), on the right hand side the orders to sell (asks).

1.2.2 Trading and price formation

Stocks are traded mainly on stock exchanges. Some exchanges are physical locations, where transactions are carried out on a trading floor, e.g., the New York Stock Exchange (NYSE), others are virtual, i.e., the trades are done electronically on the Internet, e.g., the Nasdaq (National Association of Securities Dealers Automated Quotations). Stocks can also be traded over the counter (OTC), that is, a contract is established between two parties without the supervision of an exchange.

Now let's take a closer look on the mechanism of trading at a stock exchange, and in particular on the formation of the price. At a stock exchange, the trading is organized by the clearing house. A clearing house is a financial institution which acts as an intermediary between buyers and sellers providing stability and efficiency. The traders post their orders to buy and sell stocks in an order book. The order book is visible for all traders at the exchange. In this way, it is guaranteed that all traders have the same information of what is offered on the market. If a trader wants to buy or sell a stock, he posts an order. There are two types of orders, limit and market orders. If a trader wants to buy or sell a certain number of stocks at a certain price, he posts a limit order. The word limit refers to the fact that the trader is only willing to trade at the specified price. An excerpt of an order book is shown in figure 1.2. On the left hand side are the orders to buy, the so-called bids, on the right hand side the orders to sell, the so-called asks. The orders are sorted according to the price specified from the trader. Next to the price is the volume, i.e., the number of stocks that is offered to sell or to buy at that price. The best ask $a(t)$ is the lowest price offer to sell and the best bid $b(t)$ the highest offer to buy at time t. A trade takes place when the best bid and the best ask match

$a(t) = b(t)$. Immediately after the trade, the price of the stock is

$$S(t) = a(t) = b(t) \ . \tag{1.1}$$

Often, the volumes do not match and the order is divided. In this case, the price is defined with respect to the last part of the trade. If a trade does not take place within the order's lifetime, the limit order is erased. If a trader wants to buy or sell a stock right away, he can place a market order. In this case, the trader gets the best available price in the order book. Market orders are executed immediately and do not appear in the order book.

Each time a trade happens, the last traded price and the volume are recorded by the exchange. Nowadays, trades happen in time frames of milliseconds, which results in huge amounts of data. In the course of the thesis, we will consider daily closing prices, i.e., the last traded price on the considered trading day.

1.2.3 Random nature of prices

Looking again at figure 1.1, we observe that the time evolution of the price is not smooth but rather irregular with random up and down movements. Based on this observation, price evolution in financial markets is typically modeled in terms of stochastic processes. The idea dates back to 1900, when Louis Bachelier proposed to model stock prices as a random walk, or Brownian motion [10]. He assumed that stock prices move in a random direction (up or down) and have no memory, i.e., future prices depend only on the current price. Based on this assumption, Bachelier derived a differential equation for the evolution of the probability distribution of prices, noting that this equation resembles the diffusion equation, and found as one possible solution the normal distribution. Five years later, Albert Einstein independently discovered the same stochastic process and applied it in thermodynamics [11]. In 1923, Norbert Wiener ultimately proved the existence of Brownian motion and developed the mathematical theory [48]. Therefore, Brownian motion is also known as Wiener process. In terms of stochastic differential equations, the Brownian motion can be written as

$$\mathrm{d}S(t) = \mu \, \mathrm{d}t + \sigma \epsilon \sqrt{\mathrm{d}t} \ . \tag{1.2}$$

It consists of a deterministic part $\mu \, \mathrm{d}t$, where μ is a drift which measures the average growth of the random variable, here the price $S(t)$, and a stochastic part $\sigma \epsilon \sqrt{\mathrm{d}t}$, where ϵ is a random variable which is statistically independent at every infinitesimal step. The parameter σ is called volatility of the price. It is a fundamental quantity in financial markets, which we will discuss in more details in section 1.2.5.

Bachelier's model captured the randomness of the stock market to some extend, but it suffered from the unrealistic property that it allowed negative stock prices.

Prices, however, are always positive. In 1959, Osborne [12], followed by Samuelson [13], proposed a much more realistic description, namely a geometric Brownian motion

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma\epsilon S(t)\sqrt{\mathrm{d}t} \ . \tag{1.3}$$

In a geometric Brownian motion, the prices are log-normally distributed and thus always positive, whereas the differences of the logarithms of the prices are normally distributed.

The geometric Brownian motion is a fundamental model which has countless applications in economic modeling. The probably most prominent one is the Black and Scholes model for the calculation of option prices [14]. However, the geometric Brownian motion is known to provide only a first approximation of what is observed in real data. More realistic descriptions are necessary in order to capture the empirical properties of financial time series.

1.2.4 Returns

Although prices are what we observe in financial markets, most financial studies involve financial returns instead. The return of a stock k at time t is defined as the relative price change during a certain time interval Δt

$$r_k(t) = \frac{S_k(t+\Delta t) - S_k(t)}{S_k(t)} \quad , \tag{1.4}$$

where $S_k(t)$ is the price of the stock k at time t and Δt is referred to as return interval. The return interval can be any amount of time from a couple of seconds to years. In figure 1.3 the daily closing prices of the Citigroup stock are shown together with the daily returns, $\Delta t = 1$ day, in the time period 1992 to 2014.

There are two main reasons for using returns instead of prices. On the one hand, it is advantageous for investors since the return directly indicates the possible profit from the trade, if positive, but also the potential loss, if negative. On the other hand, the benefit of using returns instead of prices is normalization: it enables us to compare the time series of two or more different stocks despite the unequal absolute prices $S_k(t)$ and the local trends. This is an important requirement for multivariate statistical analysis.

In recent years, considerable attention has been given to the distribution of returns. The most important empirical finding is that the unconditional distribution of returns is non-Gaussian and has heavy tails. This is illustrated in figure 1.4, which shows the distribution of daily returns for the Citigroup stock in the time period 1992 to 2014. Compared to a normal distribution, the return distribution exhibits a higher peak around zero and heavier tails. A distribution of this kind is called leptokurtic. It implies that the probability for extreme events is higher than in the

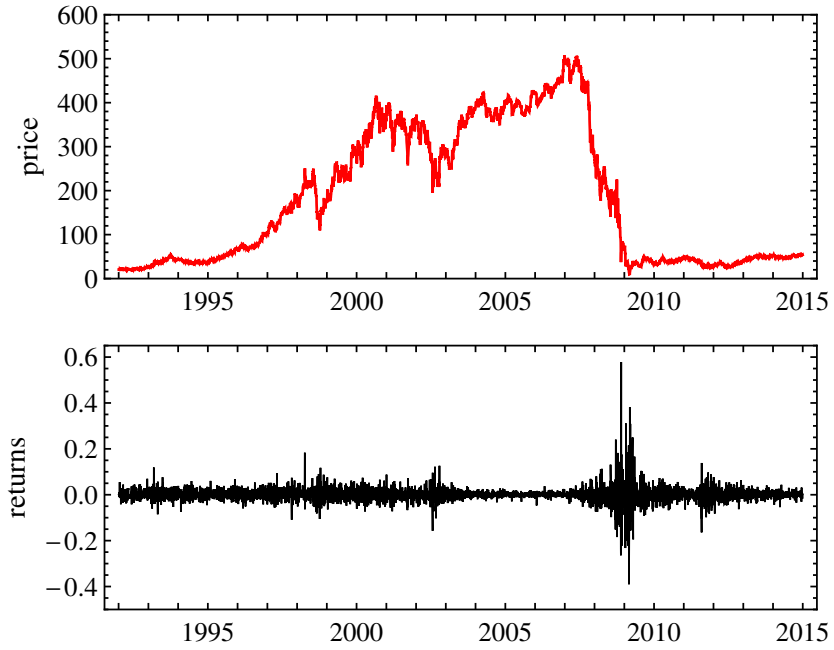


Figure 1.3: (top) Daily closing prices and (bottom) daily returns, $\Delta t = 1$ day, of the Citigroup stock in the time period 1992 to 2014.

case of the normal distribution. This is important for financial risk since it means that large price fluctuations are much more common than one would expect. In particular, empirical studies generally agree that the smaller the return interval Δt, the heavier the tails. On the other hand, as the return interval increases, the distribution of returns tends to approximate normality [43].

Although the leptokurtic nature of the return distribution had been occasionally observed since the 1920s [49–51], the first systematic account for this phenomenon was provided in 1963 by Benoit Mandelbrot. He rejected normality as a distributional model for asset returns and proposed an alpha stable, or stable Paretian, distribution [16] instead. This was supported later on by the work of Eugene Fama [52, 53], among others, and established the stable Paretian distribution as a model for asset returns. A major consequence of Mandelbrot's assumption was the infinite variance of the Paretian distribution. This posed a serious problem when dealing with real data and was unpopular among many economists used to work with models based on finite second moments. Many alternatives were proposed in the literature, including a Student's t-distribution [54, 55], a mixture of normal distributions [56], a hyperbolic [57], a normal inverse Gaussian distribution [58],

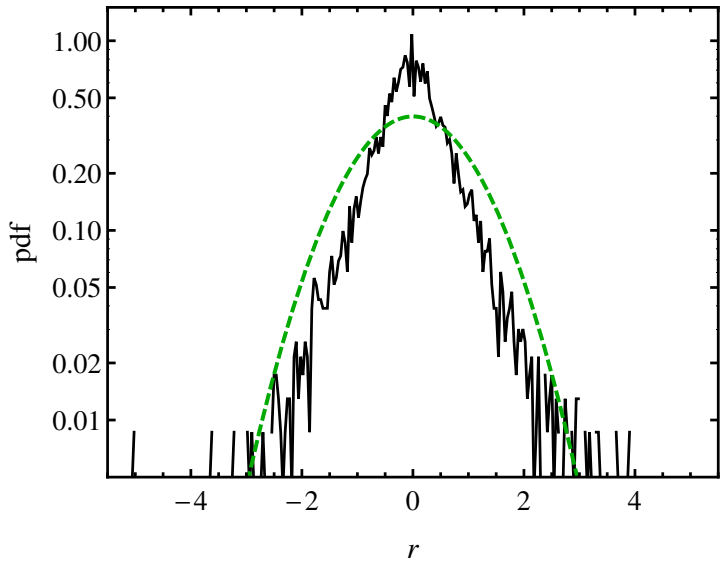


Figure 1.4: Probability density function (pdf) of the normalized (to mean zero and standard deviation one) daily returns of the Citigroup stock in the time period 1992 to 2014. For comparison, a standard normal distribution $\mathcal{N}(0,1)$ is plotted as a green dashed curve.

and others. Still, no general consensus exists on the exact form of the tails.

1.2.5 Volatility

Volatility is an important concept in financial markets. It is a measure for the variation of the price over time. Relatively high volatility means that the price can change dramatically over a short time interval in either direction, whereas lower volatility means smaller fluctuations and changes at a steady pace over a period of time.

Unlike financial returns, volatilities are not directly observable on the market. Hence, they have to be estimated. A common approach in the literature is to use the standard deviation of the returns over a certain time window

$$\sigma_k = \sqrt{\langle (r_k(t) - \langle r_k(t) \rangle)^2 \rangle} = \sqrt{\langle r_k(t)^2 \rangle - \langle r_k(t) \rangle^2} \ , \tag{1.5}$$

where $\langle \cdot \rangle$ denotes the average over the considered time window T given by

$$\langle r_k(t) \rangle = \frac{1}{T} \sum_{t=1}^{T} r_k(t) \ . \tag{1.6}$$

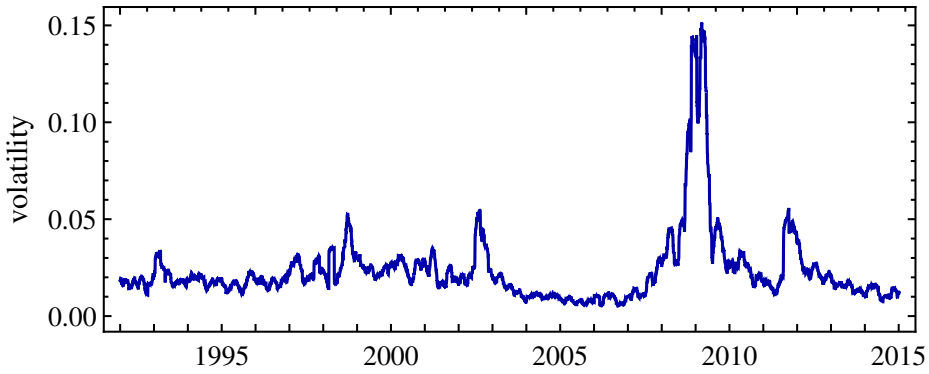


Figure 1.5: Volatility for the Citigroup stock in the time period 1992 to 2014, estimated on a 40-days time window.

Another definition of volatility uses the absolute or squared returns. Both definitions are common in the literature and lead to the same generic features. Throughout this thesis, we will mainly use the first definition (1.5).

In the geometric Brownian motion, volatility is a constant parameter. However, empirical studies show that the volatility is a fluctuating quantity, see e.g., references [59, 60]. Figure 1.5 shows the volatility for the Citigroup stock estimated on a 40-days time window. We observe that the volatility fluctuates over time. Furthermore, we note that the volatility of returns is high for extended periods and then low for subsequent extended periods. This phenomenon is called volatility clustering. It was first observed by Mandelbrot and formulated as "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes" [16]. Volatility clustering is strictly correlated with two more properties of returns: the absence of linear autocorrelation and the presence of non-linear autocorrelation. The linear autocorrelation of raw returns is often insignificant, except for very small intraday time scales ($\simeq$ 20 minutes). On the other hand, the autocorrelation function of absolute returns remains positive over lags of several weeks and decays slowly to zero. Similar behavior is observed for the autocorrelation of squared returns. These observations are regarded as a quantitative manifestation of the volatility clustering itself and indicate that returns are not independent across time. For more details on the volatility clustering, the reader is referred to reference [61].

Modeling volatility is a crucial issue in risk management. Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models [62] are commonly used to capture the volatility fluctuations. The simplest GARCH model used in practice

is the GARCH(1,1) model given by

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= a_0 + a_1 r_{t-1}^2 + b_1 \sigma_{t-1}^2 \ , \end{aligned} \tag{1.7}$$

where σ_t is the volatility at time t and ϵ_t is a random process with zero expectation value and unit variance. The parameters a_0, a_1 and b_1 are free parameters which have to be estimated by a fit to the historical time series. In this model, the volatility at time t depends on the volatility and the squared returns at the previous time step $t-1$. Another set of models used in volatility modeling are the so-called stochastic volatility models [63, 64]. Whereas in GARCH-type models volatility is completely determined by past information, stochastic volatility-type models assume that volatility evolves as a stochastic process over time. For further information on volatility models, the interested reader is referred to reference [65].

1.2.6 Correlations between stocks

Correlations between stocks are an important feature of financial markets, which will play a central role in this thesis. They provide information about the relationship between stocks of different companies. In order to quantify correlations, we consider the return time series (1.4). Since different stocks have different mean values and volatilities, we first have to normalize the return time series of each stock in the following way

$$M_k(t) = \frac{r_k(t) - \langle r_k(t) \rangle}{\sqrt{\langle r_k^2(t) \rangle - \langle r_k(t) \rangle^2}} \ , \tag{1.8}$$

where $\langle \cdot \rangle$ is the average over the considered time window T. The normalization allows to treat all stocks on equal footing. The correlation coefficient for two stocks k and l is then defined as

$$C_{kl} = \langle M_k(t) M_l(t) \rangle = \frac{1}{T} \sum_{t=1}^{T} M_k(t) M_l(t) \ . \tag{1.9}$$

With equation (1.8), we can write the correlation coefficient in the form

$$C_{kl} = \frac{\langle r_k(t) r_l(t) \rangle - \langle r_k(t) \rangle \langle r_l(t) \rangle}{\sqrt{\langle r_k^2(t) \rangle - \langle r_k(t) \rangle^2} \sqrt{\langle r_l^2(t) \rangle - \langle r_l(t) \rangle^2}} \ . \tag{1.10}$$

This is the Pearson correlation coefficient. Its values lie in the range between -1 and $+1$. If the correlation coefficient between the stocks k and l is positive, the stocks tend to move in the same direction. On the other hand, if the correlation coefficient

is negative, they tend to move in opposite directions. For -1 we have completely anti-correlated time series, for $+1$ completely correlated, i.e., identical time series. If the correlation coefficient is equal to zero, the time series are uncorrelated.

In case of K stocks, we can order the normalized time series $M_k(t)$, $k = 1, \ldots, K$ in a rectangular $K \times T$ data matrix M. Thus, the time average (1.9) can be viewed as a matrix product of the data matrix M with its transpose $M^\dagger$. This leads to the $K \times K$ correlation matrix

$$C = \frac{1}{T} M M^\dagger \ . \tag{1.11}$$

It is a positive definite, real symmetric matrix, which contains the correlation coefficients of all pairs of stocks.

The correlation matrix is a central object in the study of nature and dynamics of financial markets and plays a crucial role in the practice of modern finance. Empirical studies show that correlations change significantly over time, see e.g., references [66–69]. Figure 1.6 shows the correlation matrix for 452 stocks of the S&P 500 index for the last two quarters of 2008 and the first two quarters of 2009. The stocks are ordered according to their industry sectors. The industry sectors correspond to the blocks on the diagonal. On the other hand, the correlations between the sectors are given by the off-diagonal blocks. We observe that the correlation matrix changes significantly over time. This is not surprising since the market conditions but also the business relations between companies change over time. Hence, if we want to capture the current structure of the market, we have to determine the correlation coefficients from most recent data, i.e., on short time windows. This turns out to be rather difficult since the finite length of the time series introduces measurement noise. If one considers K stocks, the correlation matrix contains $K(K-1)/2$ entries, which must be determined from K time series of length T. If T is not large compared to K, the correlations are dressed with noise. The shorter the time series, the noisier the correlation matrix. In other words, there is a trade-off between choosing the time series long enough to keep the noise dressing at a reasonable level and choosing them short enough to provide a good estimate of current correlations. In practice, the noise dressing cannot be avoided. There are various techniques to reduce the noise and uncover the true correlations, see e.g., references [70, 71].

The Pearson correlation coefficient (1.10) is the most common measure of dependence between financial time series. Sometimes, however, its use can be problematic. First, the correlation coefficient averages over time-varying trends and volatilities, which leads to estimation errors due to this non-stationarity. In chapter 4, we will deal with this problem introducing the method of local normalization [72], which removes the non-stationarity on a local scale. Second, the Pearson correlation coef-

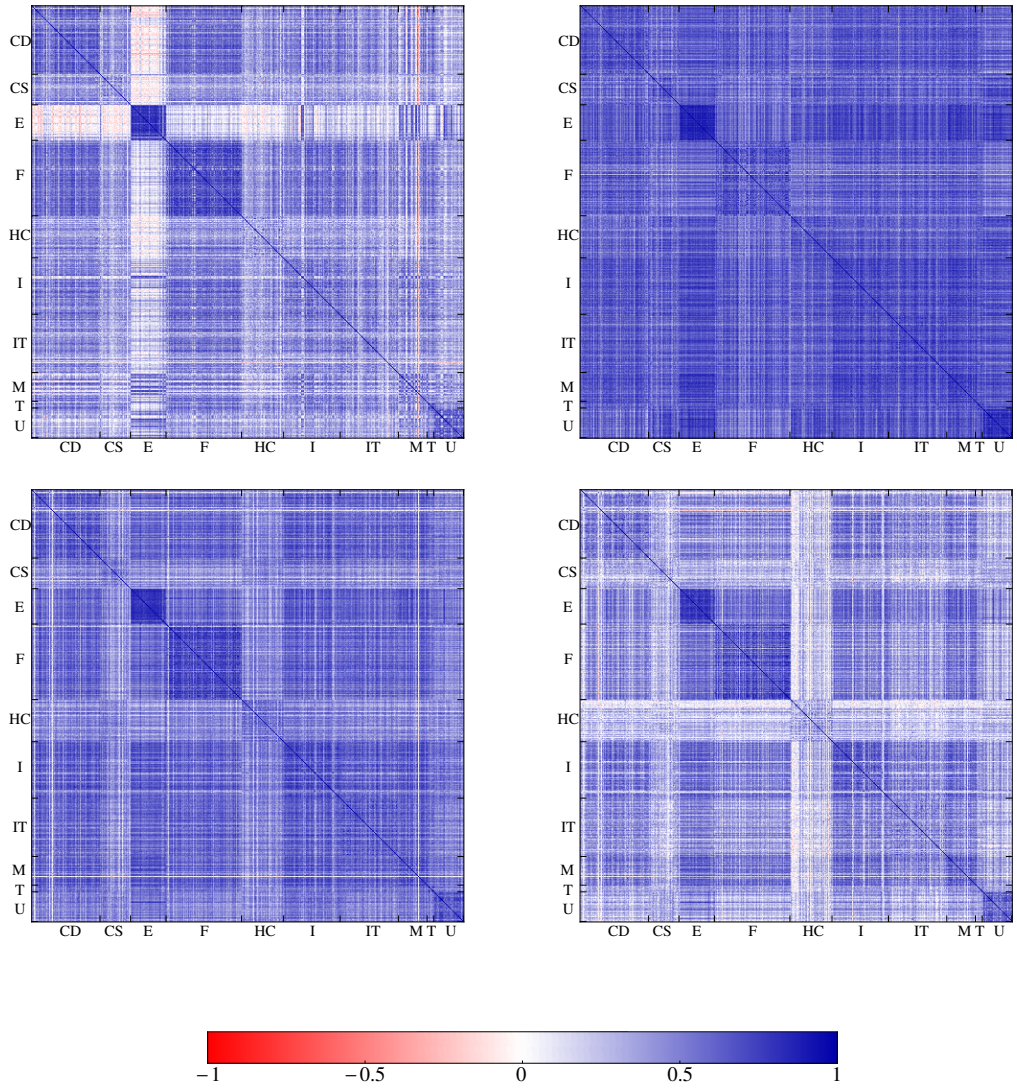


Figure 1.6: Correlation matrices for 452 stocks from the S&P 500 index for (top) the third and fourth quarter of 2008 and (bottom) the first and second quarter of 2009. Industry branches legend: CD, Consumer Discretionary; CS, Consumer Staples; E, Energy; F, Financials; HC, Health Care; I, Industrials; IT, Information Technology; M, Materials; T, Telecommunication Services; U, Utilities.

ficient only accounts for linear dependencies. It is a natural measure of dependence only for elliptical distributions, which include for instance the multivariate normal and Student's t-distributions. Furthermore, it is a good dependence measure only for distributions with finite variance. In section 5, we will thus introduce the concept of copulas to study the statistical dependence of financial returns.

Instead of the correlation matrix one often uses the covariance matrix. On the diagonal the covariance matrix contains the variances of each time series, and the off-diagonal elements give the covariances

$$\Sigma_{kl} = \langle r_k(t) r_l(t) \rangle - \langle r_k(t) \rangle \langle r_l(t) \rangle \ . \tag{1.12}$$

We note the relation

$$\Sigma = \sigma C \sigma \tag{1.13}$$

where $\sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_K)$ is a diagonal matrix whose entries are the standard deviations of each time series.

1.2.7 Risk and portfolio management

Trading with stocks not always leads to profit, but it also involves a certain amount of risk. In general, risk is defined as deviation of the actual outcome of an investment from its expected outcome. In finance, risk is a term used to imply the negative deviation, meaning the potential for financial losses.

Roughly, one distinguishes four types of risk [47]: Market risk refers to the risk of losing money due to the unpredictable fluctuations of the asset prices. It includes the risk from investments in stocks, bonds, currencies, commodities, traded derivatives, etc. Credit risk is the risk associated with a default of a borrower, when the borrower cannot pay or can only partially pay his obligations to the creditor. Operational risk is defined as the "risk of losses resulting from inadequate or failed internal processes, people and systems or from external events". Examples include rogue traders, limit violations, insufficient controlling, fraud, IT-failures and attacks, system unavailability, catastrophes such as fire, earthquakes, floods, etc. Liquidity risk is associated with the risk that a bank is unable to satisfy all claims of payment against it, i.e., becomes illiquid. The bank thus would default on some payments. Liquidity risk in essence appears very similar to credit risk.

Risk is an important issue in portfolio management and optimization. Usually, investors do not invest just in one single stock but into several stocks and financial instruments simultaneously. The ensemble of these investments is called a portfolio. More precisely, a portfolio is a linear combination of financial assets. Here, we consider purely stock portfolios. To measure the gains and the losses for a given portfolio, one is interested in the portfolio return. The return of a portfolio

consisting of K stocks is defined as the weighted sum of the individual returns

$$R = \sum_{k=1}^{K} w_k r_k = w^\dagger r \ , \tag{1.14}$$

where r_k is the return of the stock k and the coefficients w_k represent the fraction of money invested in stock k, also referred to as fractions of wealth. The fractions of wealth are dimensionless and obey the normalization condition

$$\sum_{k=1}^{K} w_k = 1. \tag{1.15}$$

This is the so-called budget constraint, implying that the total amount of money invested in the portfolio is fixed. We note that the weights can be positive or negative. Negative weights imply the possibility of short selling, that is, trading with stocks which one does not own.

What is the risk of a given portfolio? According to Markowitz portfolio theory [73], the risk of a portfolio is simply the variance of the portfolio return

$$\Omega^2 = \langle (R - \langle R \rangle)^2 \rangle \ , \tag{1.16}$$

where the average $\langle \cdot \rangle$ runs over the entire evaluation period of the portfolio. Note that this is a good risk measure only for normally distributed returns. For heavy-tailed return distributions there are more appropriate risk measures, e.g., Value at Risk. Nevertheless, here we confine ourselves to the variance. Inserting equation (1.14) into equation (1.16) allows us to express the portfolio variance in terms of the correlation matrix C and the volatilities of the single stocks σ_k

$$\Omega^2 = \sum_{k=1}^{K} \sum_{l=1}^{K} w_k w_l \sigma_k \sigma_l C_{kl} = w^\dagger \sigma C \sigma w = w^\dagger \Sigma w \ . \tag{1.17}$$

The relation implies that the risk of a portfolio depends on the correlations between the stocks. Thus, one could reduce the risk of the portfolio by including uncorrelated or even better anti-correlated stocks. Intuitively, by including several different assets, even if one of them has a large loss due to its variability, chances are the others will not. This is called diversification. In the end though, even with a large number of stocks, risk cannot be avoided altogether. Diversification only reduces the risk due to the presence of correlations, referred to as unsystematic risk. The risk that always remains is called systematic, or market risk. It results from the general trends which affect the whole market not just particular stocks or industry. Thus, the systematic risk is inherent to the whole market and cannot be reduced

by diversification.

Portfolio optimization is a procedure to minimize the risk of a portfolio for a desired portfolio return by choosing appropriate set of weights. Mathematically, this is an optimization problem of the form

$$\begin{aligned} \text{minimize} \quad & \Omega^2 = w^\dagger \Sigma w \\ \text{subject to} \quad & \widehat{R} = \sum_{k=1}^{K} w_k \mu_k = w^\dagger \mu \\ & \sum_{k=1}^{K} w_k = w^\dagger \mathbf{1} = 1 \ , \end{aligned} \tag{1.18}$$

where the desired portfolio return $\widehat{R}$ is the weighted sum of the expected returns μ_k for each stock k. Here, we consider only the budget constraint (1.15) assuming that short selling is allowed. Forbidding short selling would impose an additional constraint, namely $w_k \geq 0$. The optimization problem (1.18) can be solved by means of Lagrange multipliers. The corresponding Lagrangian reads

$$\mathcal{L} = \frac{1}{2} w^\dagger \Sigma w - \alpha(w^\dagger \mu - \widehat{R}) - \beta(w^\dagger \mathbf{1} - 1) \ . \tag{1.19}$$

The first term is the portfolio risk to be minimized while the second and the third term account for the two constraints with the Lagrange multipliers α and β. At the optimal value, the total differential vanishes

$$\mathrm{d}\mathcal{L} = \sum_{k=1}^{K} \frac{\partial \mathcal{L}}{\partial w_k} \mathrm{d}w_k + \frac{\partial \mathcal{L}}{\partial \alpha} \mathrm{d}\alpha + \frac{\partial \mathcal{L}}{\partial \beta} \mathrm{d}\beta \, \Big|_{\text{opt}} = 0 \ . \tag{1.20}$$

This leads to a system of $K + 2$ equations whose solution is a set of optimal values. For the minimum variance portfolio, i.e., the portfolio with the lowest risk, the weights are given by

$$w_{\text{opt}} = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}^\dagger \Sigma \mathbf{1}} \ . \tag{1.21}$$

We will come back to portfolio optimization in chapter 3, where we will construct optimal portfolios with minimum variance.

1.3 Random matrix theory

In this section, we present random matrix theory as an example for the application of physical models in finance. After a short historical introduction, we will discuss the common application of random matrix theory to the study of correlation

matrices. In chapter 2, we will put forward a new application of random matrix theory aiming at modeling the non-stationarity of correlations.

1.3.1 Background

Random matrix theory (RMT) is concerned with the study of large matrices, whose entries are random variables, and in particular with their eigenvalues and eigenvectors.

Random matrices were first introduced by John Wishart in 1928 in the context of multivariate statistics [74]. Wishart studied sample covariance matrices of observations from a multivariate normal distribution and formulated their distribution. However, the concept of random matrices did not attract much attention at the time. Its major breakthrough came in the 1950s when Eugene Wigner introduced random matrices in nuclear physics in order to understand the energy levels of complex nuclei, which the existing models failed to explain [75]. His idea was to replace the unknown Hamiltonian of a heavy nucleus by an ensemble of random matrices which share the same general properties with the Hamiltonian in question, e.g., symmetries, invariances, etc. Studying the random matrix ensemble, Wigner was able to make profound statements about the spectral statistics of heavy nuclei in agreement with the experimental data. Later on, the mathematical foundations of RMT were established in a series of papers by Freeman Dyson [76–79]. Dyson viewed RMT as "a new kind of statistical mechanics". In statistical mechanics, one considers an ensemble of all possible states of a system, the microcanonical, canonical or macrocanonical one. Observables are averaged over this ensemble. In RMT, the averaging is done over an ensemble of random matrices, which share similar properties with the system. For a complex quantum system, RMT predictions represent an average over all possible interactions. Deviation from the universal predictions of RMT identify system-specific, non-random properties of the system, providing clues about the underlying interactions [80].

Today, RMT finds applications not only in nuclear physics but also in other fields like quantum chaos, quantum chromodynamics, quantum gravity, etc. For an extensive review describing many of the applications in physics see reference [81]. RMT is also successful beyond physics, e.g., in number theory [82], wireless communication [83], neural network theory [84]. Here, we concentrate on the application of RMT in finance.

1.3.2 Application to financial markets

In finance, RMT is usually applied to study the statistical properties of empirical correlation matrices [20, 26, 80, 85–87]. The idea behind this application is the observation that due to the finiteness of the return time series, the correlation

coefficients are noise dressed, and therefore the correlation matrix is to a large extent random.

Motivated by this observation, Laloux *et al.* [26] compared the properties of an empirical correlation matrix to a null hypothesis – purely random correlation matrix constructed from finite time series of uncorrelated assets. Let

$$X = \frac{1}{T} A A^{\dagger} \tag{1.22}$$

be a random correlation matrix constructed from independent Gaussian elements A of size $K \times T$. By construction, X belongs to the type of matrices referred to as Wishart matrices in multivariate statistics. The statistical properties of matrices such as X are known [88]. In particular, the eigenvalue density in the limit $K \to \infty$, $T \to \infty$ and $Q = T/K \geq 1$ fixed, reads

$$\rho(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda} \ , \quad \lambda \in [\lambda_{\min}, \lambda_{\max}] \ , \tag{1.23}$$

where σ^2 is the variance of the elements of A, and $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of X respectively, given by

$$\lambda_{\max/\min} = \sigma^2(1 + 1/Q \pm 2\sqrt{1/Q}) \ . \tag{1.24}$$

These eigenvalues determine the bounds of the theoretical eigenvalue distribution (1.23). The central result of the study of Laloux *et al.* was the remarkable agreement between the theoretical prediction (1.23) and the distribution of the eigenvalues of an empirical correlation matrix C, see figure 1.7. They found that the majority of the eigenvalues lie within the RMT bounds with exception of few largest eigenvalues. Similar results were observed in further studies, see e.g., references [80, 85–87]. In particular, Plerou *et al.* [20] analyzed the eigenvalues of the correlation matrices within the RMT bound for universal properties of random matrices and found a good agreement with the results for the Gaussian orthogonal ensemble (GOE), implying a large degree of randomness in the measured correlation coefficients. Furthermore, they examined the eigenvectors corresponding to the eigenvalues outside the RMT bounds and found that the distribution of their components displayed systematic deviations from the RMT prediction and that these deviating eigenvectors were stable in time. They analyzed the components of the deviating eigenvectors and found that the largest eigenvalue corresponded to an influence common to all stocks, namely the market itself. The analysis of the remaining deviating eigenvectors showed distinct groups, whose identities corresponded to conventionally-identified business sectors.

These findings have been further applied to separate the noise from the true

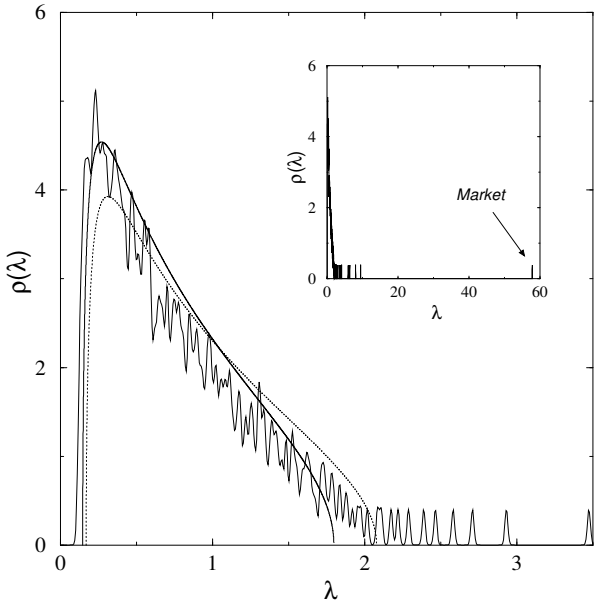


Figure 1.7: Empirical eigenvalue density of a correlation matrix C, extracted from 406 assets of the S&P 500 index during the time period 1991–1996. For comparison, the RMT prediction (1.23) for $Q = 3.22$ and $\sigma^2 = 0.85$ (solid line) and $\sigma^2 = 0.74$ (dotted line) is plotted. Inset: Same plot including the largest eigenvalue, which corresponds to the market. Taken from [26].

information, i.e., the deviations from the RMT predictions, contained in empirical correlation matrices. In particular, it has been shown that noise filtering techniques based on RMT are beneficial for portfolio optimization [37, 70, 89–91].

We point out that all these studies consider an individual large correlation matrix estimated over a long time period and study its statistics. The ensemble is fictitious, it comes into play only via the ergodicity argument, i.e., the average over one long spectrum equals an ensemble average over random matrices. In this thesis, we will put forward a new application of RMT. We will consider an ensemble of empirical correlation matrices created by the fluctuating correlations and model this ensemble by an ensemble of random matrices.

1.4 Outline of the thesis

In the previous sections, we have established some basic facts about financial markets, which are prominent examples of complex systems. Non-stationarity is an

important feature of financial markets. It manifests itself, on the one hand, in the fluctuating volatilities of single stocks, and on the other hand, in the time-varying correlations between different stocks. Here, we focus on the non-stationarity of correlations.

Financial correlations are of crucial importance to assess financial markets as a whole. They fluctuate considerably in time, because the market conditions but also the business relations between companies change over time. Here, we address this non-stationarity employing random matrices. More precisely, the fluctuating correlations create an ensemble of correlation matrices which we model by a random matrix ensemble. We point out that our approach is conceptually different from previous applications of random matrices in finance, which address the statistical properties of correlation matrices and model the estimation errors due to the finiteness of the financial time series, see section 1.3.2.

The thesis is organized as follows: In chapter 2, we introduce an approach to model non-stationary correlations by an ensemble of Wishart random matrices. Averaging the multivariate normal distribution over the random matrix ensemble leads to a correlation averaged distribution, which yields a realistic quantitative description of heavy-tailed multivariate return distributions. The model demonstrates that the non-stationarity of correlations between stocks leads to heavy tails in the multivariate return distributions. Furthermore, it reduces the complexity of a financial market to a single parameter which characterizes the correlation fluctuations due to non-stationarity.

The non-stationarity of correlations has an impact on return distributions of portfolios consisting of correlated financial instruments. In chapter 3, we apply our random matrix approach to derive a distribution for the return of a stock portfolio, which takes the non-stationary correlations into account. The resulting distribution is tested in a comparison with empirical portfolio return distributions of randomly selected portfolios.

Correlation matrices can be used to identify states of a financial market based on similarities in the correlation structure at different times. Each market state has a characteristic correlation structure and time evolution. In chapter 4, we identify market states and investigate the stability of the correlation structure of these states. In particular, we address the question: Are the correlations of a given state stationary or do they fluctuate around the average correlation matrix of the state? We study the stability of the correlation structure by means of our random matrix model and find indications for correlation fluctuations within market states.

So far, we have only considered the Pearson correlation coefficient as a measure of statistical dependence. However, it only measures the linear dependence between time series. In chapter 5, we extend our market states analysis by studying the dependence structure of market states using pairwise copulas of daily returns. We derive a bivariate K-copula, which arises from our random matrix model, and

compare it with the empirical pairwise copulas of each market state. We arrive at a consistent picture within our random matrix model.

Correlations are an important issue also in the context of extreme values. Extreme value statistics deals with the statistics of the maximum or minimum of a set of random variables. In chapter 6, we review some basic aspects of the extreme value statistics of independent and identically distributed random variables and perform a simulation study on the convergence rate of extremes. In many applications, however, data are correlated. Thus, we proceed with the correlated case. We derive a maximum distribution for a sample of equally correlated normal random variables and extend it to the non-normal case introducing a deformation function. Our results are verified in numerical simulations.

To conclude this thesis, we summarize and discuss our findings in chapter 7.

Chapter 2

Random matrix approach to non-stationarity of correlations

2.1 Introduction

Financial markets are highly non-stationary complex systems. In particular, correlations between companies change significantly over time. This fundamental non-stationarity of the financial market was discussed in section 1.2.6. Here, we address the questions: How can we model the time-varying correlations, and what are the consequences of this non-stationarity?

To take the non-stationarity into account, we put forward a new approach based on random matrices. As discussed in section 1.3.2, there are numerous applications of random matrix theory in finance. Usually, they address the statistical properties of an individual large empirical correlation matrix. Due to measurement noise, its eigenvalue density is to a great extent consistent with the eigenvalue density of random matrices. In contrast to such applications, we consider an ensemble of empirical correlation matrices created by the fluctuating correlations and model this ensemble by an ensemble of random matrices. To this end, we choose a Wishart random matrix ensemble constructed such that it has on average the same correlation structure as the ensemble of empirical correlation matrices.

This new application of random matrices in finance was first proposed by Münnix *et al.* [92] in the context of credit risk. It has been used to derive estimates for the price and loss distribution of credit portfolios with an average correlation level of zero. Here, we extend this result by allowing non-zero average correlations. In particular, we study applications of the random matrix approach to stock markets. An application to credit risk is discussed in reference [5].

The random matrix approach allows us to study the impact of the non-stationarity

on statistical observables which depend on the correlations, e.g., multivariate return distributions. In particular, we demonstrate that non-stationary correlations between stocks lead to heavy tails in the multivariate return distribution of a stock market.

We point out that our approach yields a model for the unconditional return distributions, taking into account the non-stationarity of the correlations by an ensemble of random matrices. This is different from other models, see e.g., references [93–97], where correlations are modeled by stochastic processes.

The chapter is organized as follows: We begin by setting up the random matrix model in section 2.2. Although introduced in the context of financial markets, this model has relevance to any complex system with non-stationary correlations. Thus, we present our model in a more general fashion, deriving the sample statistics for a random sample with multivariate normally distributed realizations but a randomly drawn correlation matrix. We transfer these results to stock returns in section 2.3 and perform an extensive empirical study to justify our approach in section 2.4. We conclude our results in section 2.5. The contents of this chapter are published in references [1–3].

2.2 Correlation averaged normal distribution

Consider a sample of K dimensional random vectors x, each drawn from a multivariate normal distribution with the probability density function (pdf)

$$g(x|\Sigma_{\mathrm{s}}) = \frac{1}{\sqrt{\det 2\pi\Sigma_{\mathrm{s}}}} \exp\left(-\frac{1}{2}x^{\dagger}\Sigma_{\mathrm{s}}^{-1}x\right) \ . \tag{2.1}$$

Here, Σ_{s} is the covariance matrix of the realization x and † denotes the transpose. We recall the relation between the covariance and the correlation matrix $\Sigma_{\mathrm{s}} = \sigma C_{\mathrm{s}} \sigma$, where C_{s} contains the pairwise correlations of the elements of each random vector x and $\sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_K)$ is the diagonal matrix of the standard deviations of each random variable. We further assume that σ is fixed, i.e., it is the same for each realization.

In the following, we address the question: What is the distribution of the sample if the correlation matrix C_{s} is drawn randomly for each realization? To this end, we propose an approach based on random matrices. Our ansatz is to replace the covariance matrix of each realization by a random matrix

$$\Sigma_{\mathrm{s}} \quad \longrightarrow \quad \sigma W W^{\dagger} \sigma \ . \tag{2.2}$$

The model matrix W is a rectangular $K \times N$ real random matrix, where the parameter N formally represents the length of the K model time series. This

parameter is an important ingredient of our model. Its meaning will become apparent in the following, but for now it is just an arbitrary parameter. The elements of the random matrix W are drawn from a Gaussian distribution

$$w(W|C,N) = \sqrt{\frac{N}{2\pi}}^{KN} \frac{1}{\sqrt{\det C}^N} \exp\left(-\frac{N}{2}\operatorname{tr} W^\dagger C^{-1} W\right) , \tag{2.3}$$

where C is the average correlation matrix of the whole sample. Thus, we construct an ensemble of random correlation matrices $WW^\dagger$ which follow a Wishart distribution [74] of the form

$$\tilde{w}(WW^\dagger|C,N) = \frac{\sqrt{N}^{KN}\sqrt{\det WW^\dagger}^{N-K-1}}{\sqrt{2}^{KN}\Gamma_K(N/2)\sqrt{\det C}^N} \exp\left(-\frac{N}{2}\operatorname{tr} C^{-1} WW^\dagger\right) . \tag{2.4}$$

Here, $\Gamma_K(\cdot)$ denotes the multivariate Gamma function given by

$$\Gamma_K(a) = \pi^{K(K-1)/4} \prod_{k=1}^{K} \Gamma\left(a + \frac{1-k}{2}\right) . \tag{2.5}$$

The Wishart correlation matrix ensemble fluctuates around the average correlation matrix C. By construction, the ensemble average of the model correlation matrix $WW^\dagger$,

$$\left\langle WW^\dagger \right\rangle = \int \mathrm{d}[W]\ w(W|C,N)\ WW^\dagger = C , \tag{2.6}$$

equals the average correlation matrix C. Note that the measure $\mathrm{d}[W]$ is the product of the differentials of all matrix elements. Furthermore, we point out that the variance of the elements of $WW^\dagger$ is given by

$$\operatorname{var}([WW^\dagger]_{kl}) = \frac{C_{kl}^2 + 1}{N} , \tag{2.7}$$

where C_{kl} is the kl-th element of C. It is determined by the average correlation matrix C scaled with the parameter N. Thus, N characterizes the fluctuations around C. The larger N, the narrower the distribution of the elements of $WW^\dagger$ becomes. In the limit $N \to \infty$, the random correlation matrix $WW^\dagger$ is fixed without fluctuations.

We note that the random correlation matrix C_s is invertible in the case $N \geq K$. For $N < K$, as we will find in the data later on, the resulting matrix is not invertible. Nevertheless, the pdf (2.1) is well defined in terms of proper δ functions, as we show in appendix A.1.

The key idea of our approach is to average the multivariate normal distribution (2.1) with the random covariance matrix (2.2) over the Gaussian distribution (2.3) of W

$$\langle g\rangle(x|C,N) = \int \mathrm{d}[W]\ w(W|C,N)\ g(x|\sigma WW^\dagger\sigma)\ . \tag{2.8}$$

For the calculation of this integral, it is advantageous to use the Fourier transform of the multivariate normal distribution (2.1) given by

$$g(x|\sigma WW^\dagger\sigma) = \frac{1}{(2\pi)^K}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\exp\left(-\frac{1}{2}\omega^\dagger\sigma WW^\dagger\sigma\omega\right)\ , \tag{2.9}$$

where ω is a K component real vector and the measure $\mathrm{d}[\omega]$ is the product of the differentials of the individual elements. The integration runs over the entire real axis for each component of ω. Inserting equations (2.3) and (2.9) into equation (2.8) leads to

$$\begin{aligned}\langle g\rangle(x|C,N) = \int \mathrm{d}[W]\ \sqrt{\frac{N}{2\pi}}^{KN}\ \exp\left(-\frac{N}{2}\,\mathrm{tr}\,W^\dagger C^{-1}W\right)\\ \times\ \frac{\sqrt{\det C}^{-N}}{(2\pi)^K}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\exp\left(-\frac{1}{2}\omega^\dagger\sigma WW^\dagger\sigma\omega\right).\end{aligned} \tag{2.10}$$

We notice that the term $\omega^\dagger\sigma WW^\dagger\sigma\omega$ is a scalar bilinear form, which can be written as a trace. Since the trace is invariant under cyclic permutations, we reorder the terms and express the scalar bilinear form in the exponent as

$$\omega^\dagger\sigma WW^\dagger\sigma\omega = \mathrm{tr}(W^\dagger\sigma\omega\omega^\dagger\sigma W)\ . \tag{2.11}$$

Merging the two traces and rearranging the terms, we arrive at

$$\begin{aligned}\langle g\rangle(x|C,N) = \sqrt{\frac{N}{2\pi}}^{KN}\ \frac{\sqrt{\det C}^{-N}}{(2\pi)^K}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\\ \times\int \mathrm{d}[W]\ \exp\left(-\frac{1}{2}\,\mathrm{tr}\left(WW^\dagger\left(NC^{-1}+\sigma\omega\omega^\dagger\sigma\right)\right)\right)\ .\end{aligned} \tag{2.12}$$

Here, and in similar cases later on, we may exchange the order of integration as the Fourier representation (2.9) is robust in a distributional sense while the Gaussian distribution does not inflict any convergence problems. Since the integral over W

is simply Gaussian, we have

$$\langle g\rangle(x|C,N) = \frac{\sqrt{N}^{KN}\sqrt{\det C}^{-N}}{(2\pi)^K}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\frac{1}{\sqrt{\det(NC^{-1}+\sigma\omega\omega^\dagger\sigma)}^N}. \quad (2.13)$$

To rewrite the determinant, we use a corollary of the Sylvester's determinant theorem [98]

$$\det\left(X+uv^\dagger\right) = \det\left(X\right)\left(1+v^\dagger X^{-1}u\right) \quad (2.14)$$

for any invertible square matrix X and vectors u and v, for which $uv^\dagger$ is a matrix with rank one. As C^{-1} is an invertible $K\times K$ matrix and $\sigma\omega\omega^\dagger\sigma$ is a dyadic matrix with rank one, we can apply equation (2.14) to write the determinant as

$$\begin{aligned}\det\left(NC^{-1}+\sigma\omega\omega^\dagger\sigma\right) &= N^K\det\left(C^{-1}+\frac{1}{N}\sigma\omega\omega^\dagger\sigma\right)\\ &= N^K\det C^{-1}\left(1+\frac{1}{N}\omega^\dagger\sigma C\sigma\omega\right)\ . \end{aligned}\quad (2.15)$$

Replacing the determinant in equation (2.13) with equation (2.15) leads to

$$\langle g\rangle(x|C,N) = \frac{1}{(2\pi)^K}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\frac{1}{\sqrt{1+\omega^\dagger\sigma C\sigma\omega/N}^N}\ . \quad (2.16)$$

We rewrite the remaining integral using the representation of the Gamma function

$$\frac{\Gamma(\eta)}{a^\eta} = \int_0^\infty \mathrm{d}z\ z^{\eta-1}\mathrm{e}^{-az} \quad (2.17)$$

for real and positive variables a and η. We identify a with the radicand of the square root and η with $N/2$ and cast equation (2.16) into the form

$$\langle g\rangle(x|C,N) = \frac{1}{(2\pi)^K\Gamma(N/2)}\int_0^\infty \mathrm{d}z\ z^{N/2-1}\mathrm{e}^{-z}\int \mathrm{d}[\omega]\ \mathrm{e}^{-\mathrm{i}\omega^\dagger x}\exp\left(-\frac{z}{N}\omega^\dagger\sigma C\sigma\omega\right). \quad (2.18)$$

The ω integral yields a multivariate Gaussian. Thus, equation (2.18) can be expressed as a one-dimensional average over a Gaussian involving a χ^2 distribution with N degrees of freedom, as demonstrated explicitly in appendix A.2. This is reminiscent of a compounding [99] or mixture [100] approach in statistics, where a new distribution is obtained by averaging over a parameter of a given distribution.

To proceed with the calculation, we introduce the new fixed matrix

$$\Sigma = \sigma C \sigma \quad , \tag{2.19}$$

which represents the average covariance matrix in the sense that C is the average correlation matrix and σ the diagonal matrix of the standard deviations. Performing the ω integral, we arrive at

$$\begin{aligned}\langle g \rangle(x|\Sigma, N) &= \frac{1}{(2\pi)^K \Gamma(N/2)\sqrt{\det \Sigma}} \int_0^\infty \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{\pi N}{z}}^K \exp\left(-\frac{N}{4z} x^\dagger \Sigma^{-1} x\right) \\ &= \frac{\sqrt{\pi N}^K}{(2\pi)^K \Gamma(N/2)\sqrt{\det \Sigma}} \int_0^\infty \mathrm{d}z \; z^{N/2-K/2-1} \exp\left(-z - \frac{N}{4z} x^\dagger \Sigma^{-1} x\right) \; .\end{aligned} \tag{2.20}$$

Finally, we use the representation of the modified Bessel function of the second kind of order ν [101]

$$\mathcal{K}_\nu(a) = \frac{a^\nu}{2^{\nu+1}} \int_0^\infty \mathrm{d}t \; t^{-\nu-1} \exp\left(-t - \frac{a^2}{4t}\right) \; . \tag{2.21}$$

Identifying a with $\sqrt{N x^\dagger \Sigma^{-1} x}$ and ν with $(K-N)/2$, we cast equation (2.20) into the form

$$\langle g \rangle(x|\Sigma, N) = \sqrt{\frac{N}{4\pi}}^K \frac{\sqrt{2}^{K-N+2}}{\Gamma(N/2)} \frac{\sqrt{N x^\dagger \Sigma^{-1} x}^{\frac{N-K}{2}}}{\sqrt{\det \Sigma}} \mathcal{K}_{\frac{K-N}{2}}\left(\sqrt{N x^\dagger \Sigma^{-1} x}\right) \; . \tag{2.22}$$

The distribution depends only on the average covariance matrix Σ and the free parameter N which characterizes the fluctuations around Σ. Furthermore, due to the invariance of the Wishart distribution, the vector x enters the result only via the bilinear form $x^\dagger \Sigma^{-1} x$. The pdf exhibits heavy tails, the smaller N, the heavier the tails. In the limit $N \to \infty$, it converges towards the multivariate normal distribution.

We note that the multivariate distribution (2.22) belongs to the broad family of elliptical distributions [102, 103], which generalize the multivariate normal distribution while inheriting many of its useful properties. Moreover, it includes several heavy-tailed distributions, which makes it very attractive for modeling of financial data. In particular, the multivariate Student's t-distribution has received much attention in the context of modeling multivariate financial returns [104, 105].

We now visualize the distribution (2.22) for the bivariate case, $K = 2$, with

$x = (x_1, x_2)$ and the covariance matrix

$$\Sigma = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} . \tag{2.23}$$

In this case, the pdf depends only on two parameters, the correlation coefficient c and the free parameter N. Figure 2.1 shows the bivariate pdf for different correlation coefficients $c = -0.5, 0, 0.5$ and values of $N = 4, 20$.

2.3 Application to stock returns

We now transfer our findings to financial data, aiming at deriving a multivariate distribution for stock returns which takes the time-varying correlations into account.

Consider a market consisting of K stocks. For each stock k, $k = 1, \ldots, K$, we measure the return time series $r_k(t)$ in a given observation period T_{obs}. At each time t, $t = 1, \ldots, T_{\text{obs}}$, we assume that the empirical return vector

$$r(t) = (r_1(t), \ldots, r_K(t))$$

is multivariate normally distributed with a different covariance matrix $\Sigma_t = \sigma C_t \sigma$

$$g(r|\Sigma_t) = \frac{1}{\sqrt{\det 2\pi\Sigma_t}} \exp\left(-\frac{1}{2} r^\dagger \Sigma_t^{-1} r\right) . \tag{2.24}$$

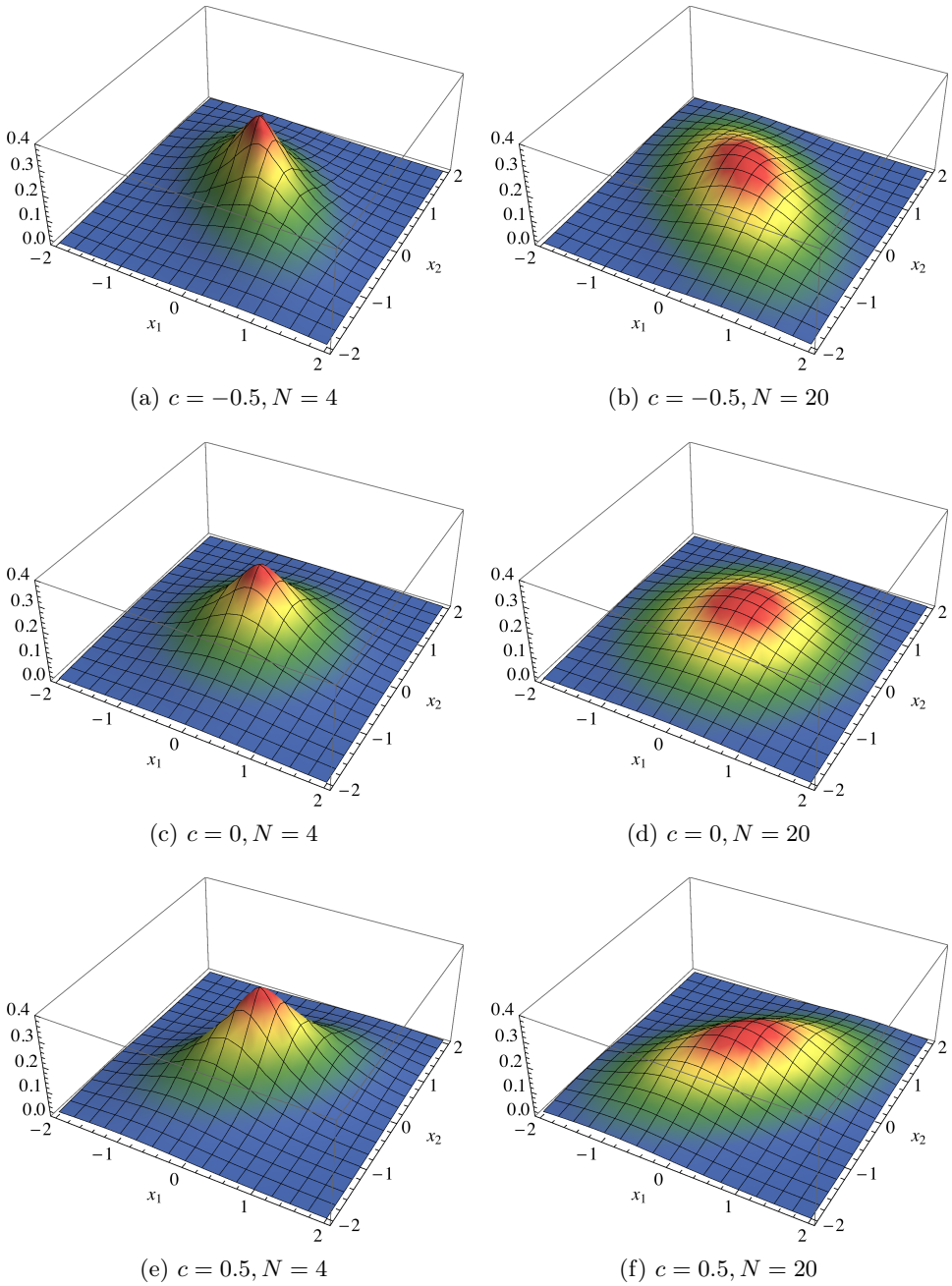
Here, we suppress the argument t of r to simplify the notation. As in the previous section, we now model the time-dependent correlation matrices C_t by an ensemble of Wishart random matrices $WW^\dagger$, which fluctuate around the average empirical correlation matrix C evaluated over the whole observation period. Averaging the multivariate normal distribution (2.24) with the random correlation matrix $WW^\dagger$ over the Wishart ensemble leads to a heavy-tailed distribution for the multivariate returns

$$\langle g \rangle (r|\Sigma, N) = \int \mathrm{d}[W]\ w(W|C, N)\ g(r|\sigma W W^\dagger \sigma) \tag{2.25}$$

$$= \frac{1}{(2\pi)^K \Gamma(N/2) \sqrt{\det \Sigma}} \int_0^\infty \mathrm{d}z\ z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{\pi N}{z}}^K \exp\left(-\frac{N}{4z} r^\dagger \Sigma^{-1} r\right) \tag{2.26}$$

$$= \sqrt{\frac{N}{4\pi}}^K \frac{\sqrt{2}^{K-N+2}}{\Gamma(N/2)} \frac{\sqrt{N x^\dagger \Sigma^{-1} x}^{\frac{N-K}{2}}}{\sqrt{\det \Sigma}} \mathcal{K}_{\frac{K-N}{2}} \left(\sqrt{N x^\dagger \Sigma^{-1} x}\right) , \tag{2.27}$$

where $\mathcal{K}_\nu$ is the modified Bessel function of the second kind of order $\nu = (K-N)/2$.



(a) $c = -0.5, N = 4$ (b) $c = -0.5, N = 20$

(c) $c = 0, N = 4$ (d) $c = 0, N = 20$

(e) $c = 0.5, N = 4$ (f) $c = 0.5, N = 20$

Figure 2.1: Bivariate correlation averaged normal distribution for different correlation coefficients c and N values.

The distribution (2.27) depends on the average empirical covariance matrix $\Sigma = \sigma C \sigma$, evaluated over the whole observation period, and a free parameter N, which governs the variance of the Wishart ensemble (2.7). Thus, N characterizes the fluctuation strength of correlations around the average correlation matrix C. The larger N, the smaller the fluctuations around C, eventually vanishing in the limit $N \to \infty$. The latter corresponds to a stationary case with no fluctuations around C.

We note that in our model we assumed fixed standard deviations σ_k for each random variable. However, empirical return time series exhibit non-stationary volatilities, as discussed in section 1.2.5. In this case, not only the correlations but also the volatilities change over time. Instead of the correlation matrix we can express the full covariance matrix as a random matrix

$$\Sigma_t \longrightarrow AA^\dagger \tag{2.28}$$

with $AA^\dagger$ being a Wishart random matrix. This leads to the same result (2.27), as explicitly shown in reference [1]. Mathematically, it does not make a difference whether we perform the calculation with a random covariance or a random correlation matrix, see appendix A.3. Thus, our approach does not contradict the empirical observation of fluctuating volatilities.

2.4 Empirical verification of the random matrix approach

In the following, we perform an empirical study to verify the random matrix approach for stock returns. We have three goals. First, we verify the assumption of multivariate normally distributed return vectors. We show that this assumption is justified on short time horizons, on which the covariances can be viewed as fixed. Second, we confirm that the correlation averaged normal distribution of returns describes the empirical multivariate returns. To this end, we look at the aggregated distribution of returns, obtained by rotating the return vectors into the eigenbasis of the covariance matrix. Hence, the aggregated distribution captures the properties of the multivariate return distribution as a whole. Third, we study the fine structure of the aggregated return distribution related to the principal components.

The data used in this and all empirical studies in the sequel is obtained from Yahoo Finance [106]. For more details and a full list of the stocks used in the empirical studies see appendix B. We consider adjusted daily closing prices, i.e., prices adjusted for splits and dividends, and take into account only stocks for which the price time series are complete and cover the whole observation period. From

the price time series, we calculate the return time series (1.4), which are the object of study in our analysis.

2.4.1 Multivariate normal distribution for returns

In section 2.3, we assumed that the distribution of the empirical return vectors at each time t can be described by a multivariate normal distribution (2.24) with a covariance matrix $\Sigma_t = \sigma C_t \sigma$. We now verify this assumption empirically on short time windows, on which the covariance matrix can be viewed as fixed.

We test the assumption for a dataset consisting of $K = 306$ stocks of the S&P 500 index continuously traded in the time period between 1992 and 2012. We compute the daily return time series, $\Delta t = 1$ trading day, and divide these time series into windows of length T, so short that the sampled covariances can be viewed as constant within these windows. Here, we choose a window length of $T = 25$ trading days. In this case, the corresponding covariance matrices are not invertible since the length of the time series $T = 25$ is much smaller than the number of stocks $K = 306$. To carry out the data analysis, we take all pairs r_k, r_l of returns which, according to our assumption, should follow a bivariate normal distribution with a 2×2 covariance matrix $\Sigma^{(k,l)}$. The bivariate covariance matrix is always invertible. We then rotate each two component vector (r_k, r_l) into the eigenbasis of the covariance matrix $\Sigma^{(k,l)}$ and normalize the elements of the resulting vector with the square root of the corresponding eigenvalues. In this way, the components become comparable and can be aggregated into a single univariate distribution, shown in figure 2.2. We find a good agreement with a normal distribution.

This observation confirms our assumption that the multivariate distribution of returns can be described by a multivariate normal distribution on short time horizons, here 25 trading days, where the covariances are sufficiently constant. We emphasize that the agreement is required on short time horizons only. On longer time horizons, the covariance matrix changes, which, as we show later on, lifts the tails of the multivariate return distribution.

2.4.2 Aggregated distribution of returns

Next, we compare the correlation averaged return distribution (2.27) with empirical returns. To visualize the comparison, we reduce the dimension by calculating an univariate average return distribution. We start with the integral (2.26) and rotate the return vector r into the eigenbasis of the covariance matrix Σ, normalizing each component of the rotated vector with the eigenvalues λ_i

$$\tilde{r}_i = \frac{[Ur]_i}{\sqrt{\lambda_i}}, \quad i = 1, \ldots, K \ . \tag{2.29}$$

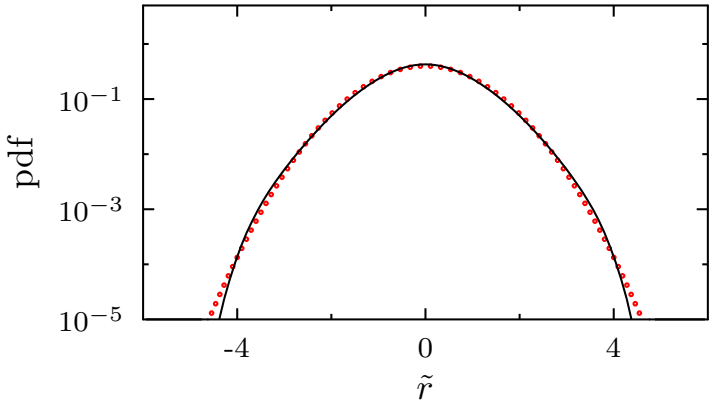


Figure 2.2: Aggregated distribution of returns, here denoted by $\tilde{r}$, for fixed covariances from the S&P 500 dataset, $\Delta t = 1$ trading day and window length $T = 25$ trading days. The red circles show a normal distribution. Taken from [1].

The matrix U is an orthogonal matrix which diagonalizes the covariance matrix as

$$\Sigma = U^\dagger \Lambda U \ , \tag{2.30}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$ is the diagonal matrix of the eigenvalues. The procedure leads to a factorization of the integral (2.26)

$$\langle g \rangle(\tilde{r}|\Sigma, N) = \frac{1}{(2\pi)^K \Gamma(N/2)} \int_0^\infty \mathrm{d}z \ z^{N/2-1} \mathrm{e}^{-z} \prod_{i=1}^K \sqrt{\frac{\pi N}{z}} \exp\left(-\frac{N}{4z}\tilde{r}_i^2\right) \ . \tag{2.31}$$

Integrating out all but one of the components of the rotated and scaled vector $\tilde{r}$, which we call $\tilde{r}_k$, leads to

$$\langle g \rangle(\tilde{r}_k|N) = \int_{\mathbb{R}^{K-1}} \mathrm{d}\tilde{r}_1 \ldots \mathrm{d}\tilde{r}_{k-1} \mathrm{d}\tilde{r}_{k+1} \ldots \mathrm{d}\tilde{r}_K \ \langle g \rangle(\tilde{r}|\Sigma, N) \tag{2.32}$$

$$= \frac{1}{(2\pi)^K \Gamma(N/2)} \int_0^\infty \mathrm{d}z \ z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{\pi N}{z}}^K \exp\left(-\frac{N}{4z}\tilde{r}_k^2\right)$$

$$\times \int_{\mathbb{R}^{K-1}} \mathrm{d}\tilde{r}_1 \ldots \mathrm{d}\tilde{r}_{k-1} \mathrm{d}\tilde{r}_{k+1} \ldots \mathrm{d}\tilde{r}_K \prod_{i=1}^{k-1} \exp\left(-\frac{N}{4z}\tilde{r}_i^2\right) \prod_{j=k+1}^{K} \exp\left(-\frac{N}{4z}\tilde{r}_j^2\right) \ . \tag{2.33}$$

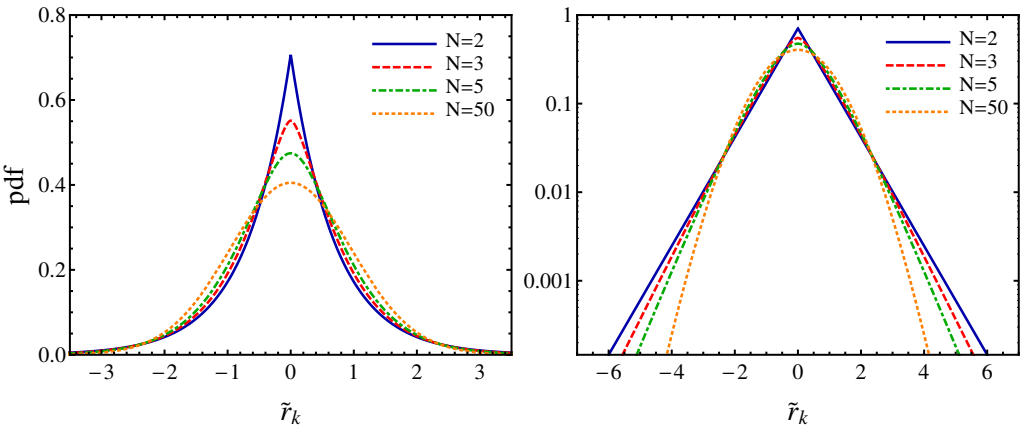


Figure 2.3: Illustration of the average return distribution $\langle g\rangle(\tilde{r}_k|N)$ for different values of N, plotted linearly (left) and logarithmically (right). Solid, dashed, dashed-dotted and dotted lines correspond to $N = 2, 3, 5, 50$, respectively.

Performing the $K-1$ Gaussian integrals, we arrive at

$$\langle g\rangle(\tilde{r}_k|N) = \frac{\sqrt{\pi N}}{(2\pi)\Gamma(N/2)} \int_0^\infty \mathrm{d}z\; z^{(N-1)/2-1} \exp\left(-z - \frac{N}{4z}\tilde{r}_k^2\right) \tag{2.34}$$

$$= \frac{\sqrt{2}^{1-N}\sqrt{N}}{\sqrt{\pi}\;\Gamma(N/2)} \sqrt{N\tilde{r}_k^2}^{\frac{N-1}{2}}\; \mathcal{K}_{\frac{N-1}{2}}\left(\sqrt{N\tilde{r}_k^2}\right) . \tag{2.35}$$

Again, we find a modified Bessel function of the second kind, this time of the order $\nu = (N-1)/2$. We note that all components of $\tilde{r}$ are equally distributed. Figure 2.3 shows the pdf (2.35) for different values of N. The distribution has exponential tails which become more and more dominant the smaller N. For large N, it approaches the normal distribution. This is shown in figure 2.4, where the kurtosis excess γ_2 of the average return distribution (2.35) is depicted versus the parameter N. We observe that as N increases, the kurtosis excess $\gamma_2 = 6/N$ slowly decreases to zero, which corresponds to a normal distribution.

We now compare the average return distribution (2.35) with the aggregated distribution of the empirical returns. Here, we consider a dataset consisting of $K = 258$ stocks of the NASDAQ Composite index traded in the time period from January 1992 to December 2013. The aggregated distribution is obtained by rotating the returns into the eigenbasis of the empirical covariance matrix evaluated over the whole 22-year observation period. Normalizing the rotated vectors by the empirical eigenvalues allows to view all of them on equal footing and to aggregate

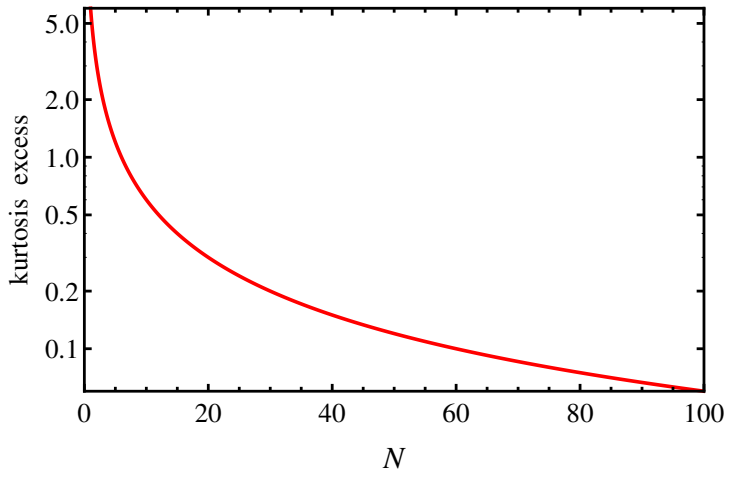


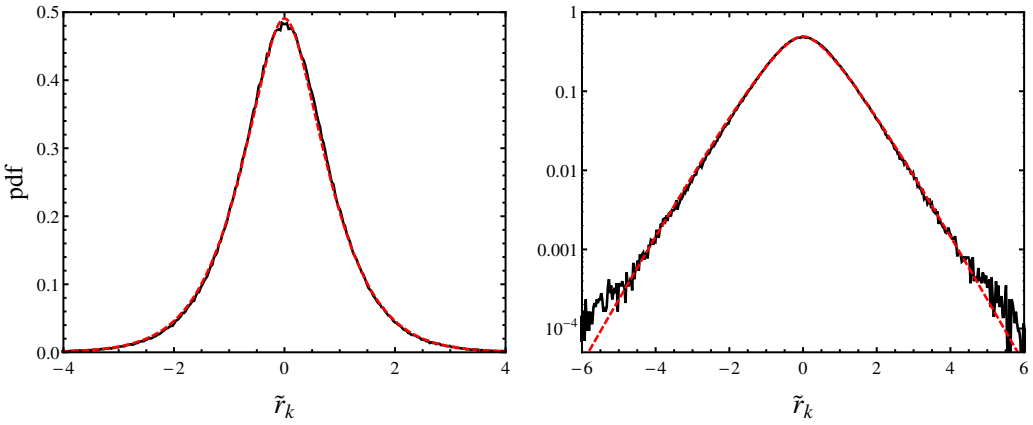
Figure 2.4: Kurtosis excess versus the parameter N, plotted logarithmically. As N increases, the average return distribution converges towards the normal distribution.

them into one histogram, which then captures the properties of the multivariate distribution as a whole. Figures 2.5 shows the aggregated distribution for daily, $\Delta t = 1$ trading day, and monthly returns, $\Delta t = 21$ trading days, compared with the average return distribution (2.35). We observe a good agreement between theory and data, deviations appear beyond the third decade. The free parameter N is determined with a minimum distance estimation method based on a Cramer-von Mises statistics [107]. We calculate the Cramer-von Mises distance, i.e., the integral of the squared difference between empirical and model distribution function, for different values of N and choose the N value which yields the smallest distance. For daily returns, we find the best agreement for $N = 4.2$. For monthly returns, higher values are needed, here $N = 13.5$. We point out that although our analytical result was derived for integer values of N, we can easily extend this result to real values. The good agreement between model and data justifies our ansatz to model non-stationary correlations by an ensemble of Wishart random matrices.

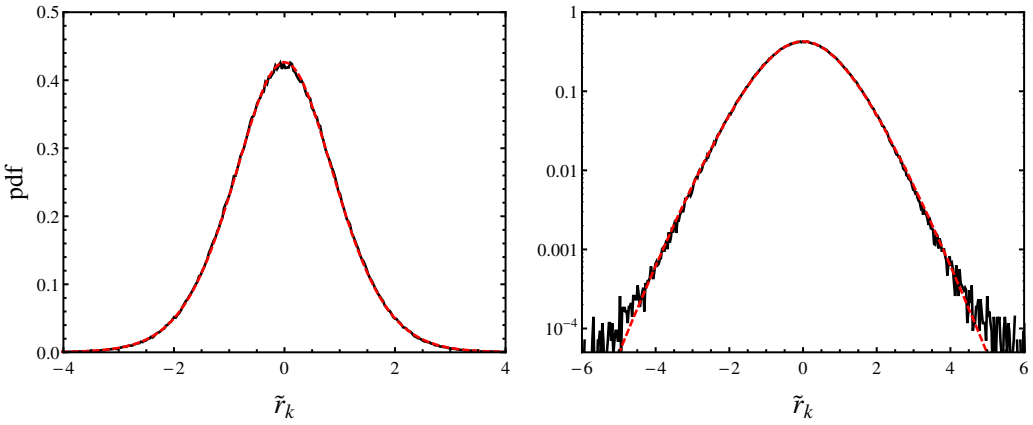
Importantly, the aggregated distribution exhibits heavy tails which result from the non-stationarity of the correlations, the smaller N, the stronger the correlation fluctuations and the heavier the tails.

2.4.3 Fine structure of the aggregated return distribution

So far, we have looked at the aggregated distribution of all returns, which represents the distribution for the entire market.



(a) $\Delta t = 1$ trading day, $N = 4.2$



(b) $\Delta t = 21$ trading days, $N = 13.5$

Figure 2.5: Aggregated distribution of the rotated and scaled returns for (a) $\Delta t = 1$ and (b) $\Delta t = 21$ trading days in the observation period 1992–2013, plotted linearly (left) and logarithmically (right). For comparison, an average return distribution (2.35) is plotted as a red dashed line.

In the following, we study the distributions of the single components $\tilde{r}_i$ (2.29), which represent the principal components of the original data. We consider the NASDAQ dataset from the previous section. While in our model all components have the same distribution (2.35), in the data we find significant deviations for the largest 20 and the smallest 4 eigenvalues. The remaining components have similar statistics and are consistent with the aggregated distribution of all returns. Figure 2.6 shows the distributions of the normalized components $\tilde{r}_i$ averaged over 43 components. We observe that the average distribution of the first 43 components, which belong to the largest eigenvalues, has different statistics compared to the rest. The associated eigenvectors have different structures and interpretations. The first component $\tilde{r}_1$ belongs to the largest eigenvalue λ_1, which describes the whole market. The other large eigenvalues correspond to eigenvectors that have only a subset of components different from zero. These eigenvectors roughly correspond to market sectors. Furthermore, the smallest eigenvalues are most sensitive to measurement noise. A histogram of the eigenvalues of the empirical covariance matrix evaluated over the whole observation period is depicted in figure 2.7.

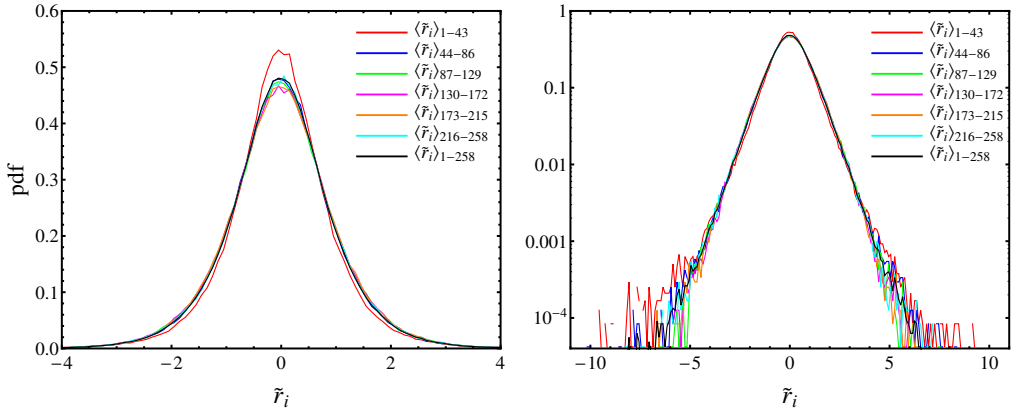


Figure 2.6: Distribution of the normalized components $\tilde{r}_i$ averaged over 43 components compared with the aggregated distribution of all returns (black), plotted linearly (left) and logarithmically (right).

The different statistics of some of the principal components is not an obstacle to aggregate all components together when looking at the multivariate distribution as a whole. No model can capture all aspects of reality. Although we find an overall good agreement with the aggregated distribution of all returns, the fine structure cannot be fully captured by our model.

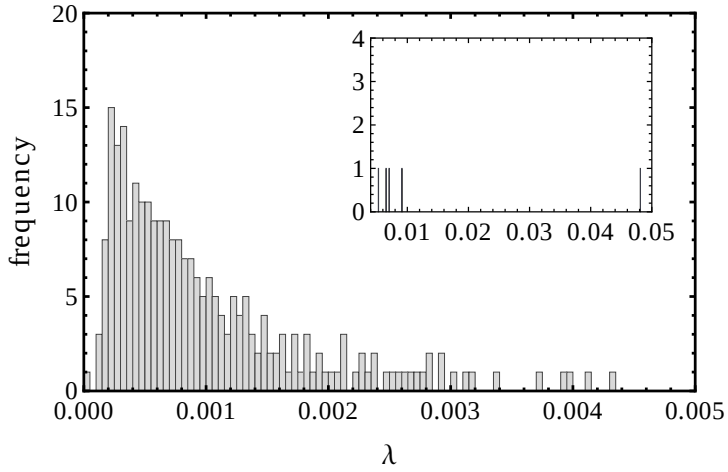


Figure 2.7: Histogram of the eigenvalues of the average empirical covariance matrix evaluated over the 22-year observation period 1992–2013. The inset shows the largest eigenvalues.

2.5 Summary

Non-stationarity is a common feature of complex systems. Here, we introduced an approach to model fluctuating correlations in a sample of multivariate normally distributed realizations by an ensemble of Wishart random matrices. Averaging the multivariate normal distribution over this ensemble yields an elliptical distribution with heavy tails, which can be expressed in terms of a modified Bessel function of the second kind. It depends only on the average covariance matrix of the sample and a parameter which characterizes the fluctuations around the average covariance matrix.

We applied our results to stock returns and verified the random matrix approach empirically. To this end, we assumed multivariate normally distributed return vectors at each fixed time step and showed that this assumption is justified on short time horizons, where the covariances or correlations are sufficiently constant. On longer time horizons, the fluctuating correlations lift the tails. The correlation averaged return distribution takes these non-stationary correlations into account. It yields a realistic quantitative description of the aggregated distribution of all returns, which represents the return distribution for the entire market. In addition, we also studied the fine structure of the return distribution related to the principal components. Although we find an overall good agreement with the aggregated distribution of all returns, the fine structure cannot be fully captured by our model.

Our approach demonstrated that the non-stationarity of the correlations between

stocks leads to heavy tails in the multivariate distribution of returns. In addition, it reduces the complexity of a correlated market to a single parameter which characterizes the correlation fluctuations. Thus, it enables us to estimate the fluctuation strength of correlations in a given time interval directly from the empirical return distributions. We will use this feature in chapter 4 to study the correlation structure of market states.

Chapter 3

Portfolio return distributions and non-stationary correlations

3.1 Introduction

Correlations are an important issue in modern portfolio theory. In particular, portfolio optimization relies heavily on the correlation matrix between stocks. Here, we study the implications of time-varying correlations on portfolios, and in particular on the distribution of portfolio returns.

In chapter 2, we introduced an approach to model non-stationary correlations by an ensemble of Wishart random matrices. This approach allowed us to derive a multivariate distribution, which yields a good quantitative description of empirical stock returns. Here, we apply our results to derive a distribution for the return of a portfolio, which takes the non-stationary correlations into account. We further show that this average portfolio return distribution yields a good description of empirical portfolio returns, in particular in the central part of the distribution. The very fat tails, however, cannot be fully captured by our model.

The chapter is organized as follows: In section 3.2, we use the correlation averaged return distribution to derive an average distribution for a portfolio return. In section 3.3, we compare our result with the portfolio return distributions of randomly selected portfolios considering different portfolio weights. We conclude our findings in section 3.4. The contents of this chapter are published in reference [2].

3.2 Average portfolio return distribution

We begin with the derivation of the average portfolio return distribution. Consider a purely stock portfolio consisting of K stocks. We recall that the portfolio return

is given by the weighted sum of the individual stock returns

$$R = \sum_{k=1}^{K} w_k r_k = w^\dagger r \quad , \tag{3.1}$$

where w_k is the weight of the k-th stock. The weights obey the normalization condition $\sum_{k=1}^{K} w_k = 1$.

To calculate the average portfolio return distribution, we have to integrate over the correlation averaged multivariate return distribution $\langle g \rangle(r|\Sigma, N)$ (2.27) and filter for those returns that lead to a given portfolio return R. Thus, we have to compute the following filter integral

$$\langle f \rangle(R|\Sigma, N) = \int \mathrm{d}[r] \; \langle g \rangle(r|\Sigma, N) \; \delta\left(R - w^\dagger r\right) \; , \tag{3.2}$$

where the measure $\mathrm{d}[r]$ is the product of the differentials for each stock k. Using a Fourier transformation, the delta function can be expressed in the following way

$$\delta\left(R - w^\dagger r\right) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}\nu \; \exp\left(-\mathrm{i}\nu R + \mathrm{i}\nu w^\dagger r\right) \; . \tag{3.3}$$

Inserting the Fourier transform (3.3) and the correlation averaged return distribution (2.26) into equation (3.2) and reordering the terms leads to

$$\langle f \rangle(R|\Sigma, N) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}\nu \; \mathrm{e}^{-\mathrm{i}\nu R} \int \mathrm{d}[r] \; \langle g \rangle(r|\Sigma, N) \; \mathrm{e}^{\mathrm{i}\nu w^\dagger r} \tag{3.4}$$

$$= \frac{\sqrt{N\pi}^K}{(2\pi)^{K+1} \; \Gamma(N/2)\sqrt{\det \Sigma}} \times \int_0^\infty \mathrm{d}z \; z^{(N-K)/2-1} \mathrm{e}^{-z} \int_{-\infty}^{+\infty} \mathrm{d}\nu \; \mathrm{e}^{-\mathrm{i}\nu R} \int \mathrm{d}[r] \exp\left(\mathrm{i}\nu w^\dagger r - \frac{N}{4z} r^\dagger \Sigma^{-1} r\right) . \tag{3.5}$$

The integral over r is Gaussian. Thus, we have

$$\langle f \rangle(R|\Sigma, N) = \frac{1}{(2\pi) \; \Gamma(N/2)} \int_0^\infty \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \int_{-\infty}^{+\infty} \mathrm{d}\nu \; \mathrm{e}^{-\mathrm{i}\nu R} \; \exp\left(-\frac{z}{N} \nu^2 w^\dagger \Sigma w\right) . \tag{3.6}$$

The integral over ν is Gaussian as well and we have

$$\langle f\rangle(R|\Sigma,N) = \frac{1}{\Gamma(N/2)}\sqrt{\frac{N}{4\pi w^\dagger \Sigma w}}\int_0^\infty \mathrm{d}z\; z^{(N-1)/2-1}\exp\left(-z-\frac{NR^2}{4zw^\dagger\Sigma w}\right)\;. \tag{3.7}$$

Once more, the last integral is a representation of the modified Bessel function of the second kind, this time of the order $\nu = (N-1)/2$. Thus, we obtain

$$\langle f\rangle(R|\Sigma,N) = \frac{\sqrt{2}^{1-N}}{\sqrt{\pi}\;\Gamma(N/2)}\sqrt{\frac{N}{w^\dagger\Sigma w}}^{\frac{N+1}{2}}\;|R|^{\frac{N-1}{2}}\;\mathcal{K}_{\frac{N-1}{2}}\left(|R|\sqrt{\frac{N}{w^\dagger\Sigma w}}\right) \tag{3.8}$$

for the average distribution of a portfolio return. Again, this is a heavy-tailed distribution. It depends only on the free parameter N, which characterizes the correlation fluctuations in the considered observation period, and on the scale variable

$$\alpha = w^\dagger \Sigma w \quad , \tag{3.9}$$

which can be computed from the portfolio weights and the empirical covariance matrix Σ evaluated over the whole observation period. We notice that α is the bilinear form of the weights w with the covariance matrix Σ and has a very direct economic relevance. It represents the variance of the entire portfolio and thus the portfolio risk according to Markowitz portfolio theory, see section 1.2.7. Normalizing the portfolio return R in the following way

$$\widehat{R} = \frac{R}{\sqrt{\alpha}} \tag{3.10}$$

leads to the density function

$$\langle f\rangle(\widehat{R}|N) = \frac{\sqrt{2}^{1-N}}{\sqrt{\pi}\;\Gamma(N/2)}\sqrt{N}^{\frac{N+1}{2}}\,|\widehat{R}|^{\frac{N-1}{2}}\;\mathcal{K}_{\frac{N-1}{2}}\left(|\widehat{R}|\sqrt{N}\right), \tag{3.11}$$

in which N is the only free parameter. We note the resemblance with the average return distribution (2.35).

3.3 Comparison with empirical portfolio returns

In the following, we compare our analytical results for the portfolio return R (3.8) and the rescaled portfolio return $\widehat{R}$ (3.11) with empirical portfolio returns. We

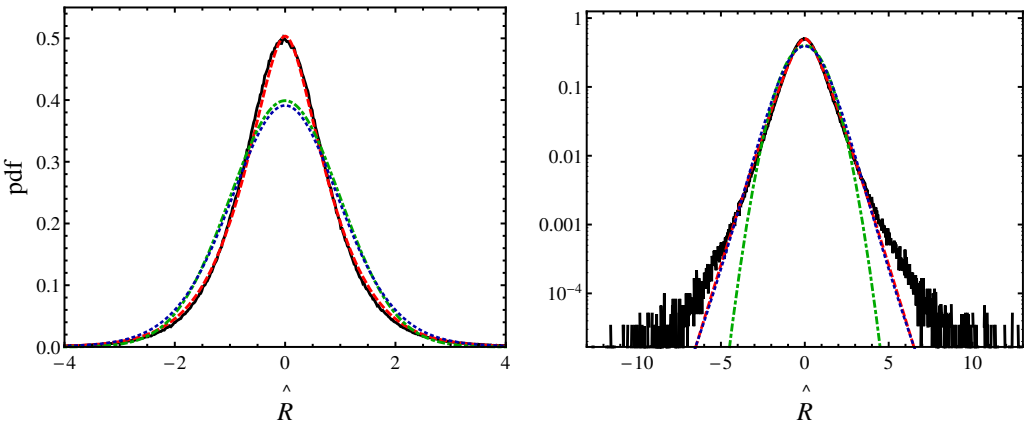


Figure 3.1: Histogram of the rescaled empirical portfolio return $\widehat{R}$ (solid black) for daily returns and weights $w_k \sim \mathcal{U}(-0.5, 0.5)$ compared with the average portfolio return pdf $\langle f \rangle(\widehat{R}|N)$ (dashed red) with $N = 3.9$, plotted linearly (left) and logarithmically (right). The green dashed-dotted line shows a normal distribution $\mathcal{N}(0, 1)$ and the blue dotted line a Student's t-distribution with degrees of freedom $\nu = 12.73$.

consider a dataset of 241 stocks of the NASDAQ Composite index in the time period from January 1992 to March 2012. We construct empirical portfolios each consisting of K stocks, which are randomly chosen from all available stocks. For each stock, we calculate the stock returns (1.4) on a given return interval Δt, for each portfolio, the corresponding portfolio return. Thereby, we will consider different kinds of weights: random, equal and optimal weights.

First, we study the rescaled portfolio return $\widehat{R}$. Figure 3.1 shows the histogram of $\widehat{R}$ taking into account 600 random portfolios of 20 stocks each. We use daily returns, $\Delta t = 1$ trading day, and consider positive and negative weights drawn from a symmetric uniform distribution $\mathcal{U}(-a, a)$ with $a = 0.5$ in such a way that the normalization condition is satisfied. Compared with the normal distribution with $\mu = 0$ and $\sigma = 1$, the histogram has a higher peak around zero and fatter tails. The Student's t-distribution with $\nu = 12.73$ describes the tails much better than the normal distribution, but it fails to describe the center of the histogram. The parameters of the normal and the Student's t-distributions are estimated with the maximum likelihood method. Compared with both distributions, the average portfolio return distribution (3.11) resembles the data much better in the center of the histogram. In the tails, it matches the Student's t-distribution. We obtain the best agreement for $N = 3.9$. Still, there are some deviations in the far tails. We determine the free parameter N with a minimum distance estimation method using

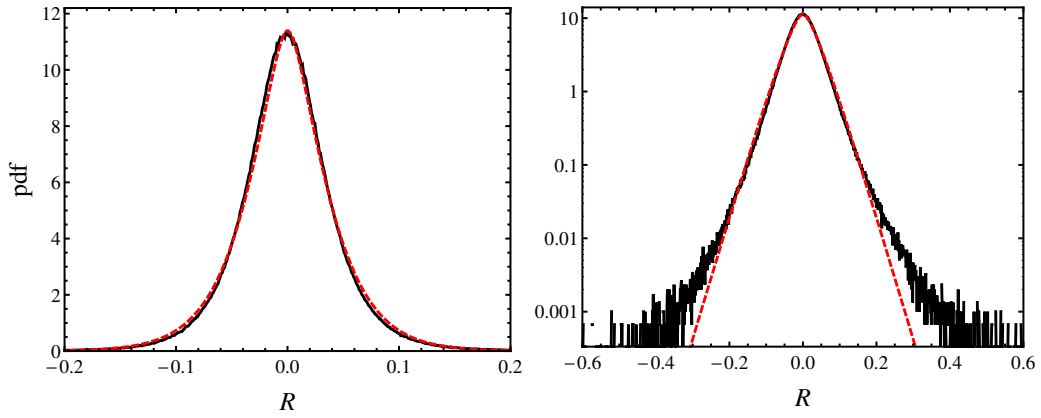


Figure 3.2: Histogram of the empirical portfolio return R (solid black) for daily returns and weights $w_k \sim \mathcal{U}(-0.5, 0.5)$ compared with the average portfolio return pdf $\langle f \rangle(R|\Sigma, N)$ (dashed red) with $\alpha = 2.09 \times 10^{-3}$ and $N = 3.5$, plotted linearly (left) and logarithmically (right).

a weighted Cramer-von Mises statistics. Estimation methods based on weighted Cramer-von Mises statistics are often applied to accentuate the discrepancies between model and empirical distribution in those parts of the distribution where more sensitivity is desired [107, 108]. A prominent example of weighted Cramer-von Mises statistics is the Anderson-Darling statistics [109, 110], which uses an appropriate weighting function in order to give more weight to the tails of the distribution. Here, we aim at fitting particularly the center of the distribution where the best match between model and data is observed. To give more weight to the center of the distribution, we use a Gaussian function of the form $\exp\left(-y^2/2c^2\right)$ with $c = 0.07$ as a weighting function. The maximum likelihood method yields about 10% smaller values for the parameter N since it also takes the tails into account.

In the following, we will consider the portfolio return R directly, as it is the economically relevant quantity, and will investigate the impact of the portfolio weights. Figure 3.2 shows the histogram of the empirical portfolio return R compared with the average portfolio return distribution (3.8). Again, we take into account 600 random portfolios of size $K = 20$ and consider daily returns and weights drawn from a symmetric uniform distribution $\mathcal{U}(-a, a)$ with $a = 0.5$. The minimum distance estimation method yields $N = 3.5$. The portfolio variance α is computed for each portfolio from the corresponding weights and covariance matrix and then averaged over all available portfolios. The theoretical curve agrees well with the histogram in the central part, there are some deviations in the tails.

Figure 3.3 shows the range of α and N values induced by the different portfolios. Choosing different distribution parameters for the portfolio weights affects the portfolio variance α, see figure 3.4. It increases monotonically with the distribution width $2a$. The parameter N, on the other hand, remains nearly constant.

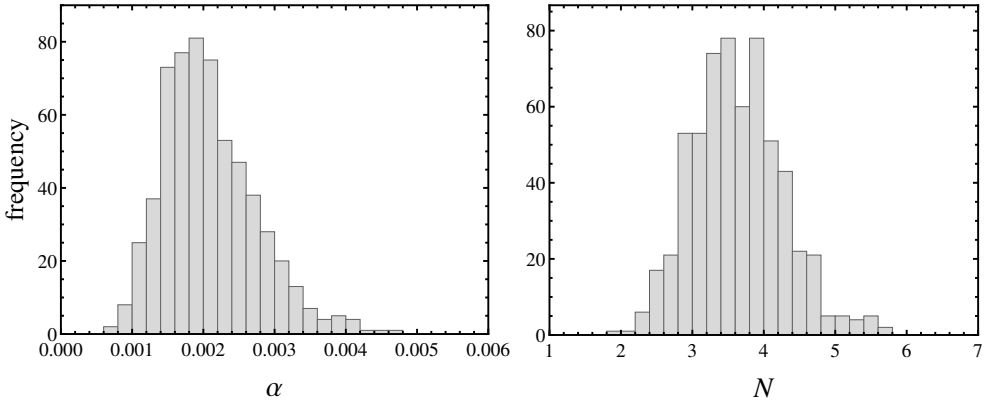


Figure 3.3: Histograms of the portfolio variance α (left) and the N value (right) of all 600 portfolios in the case $w_k \sim \mathcal{U}(-0.5, 0.5)$.

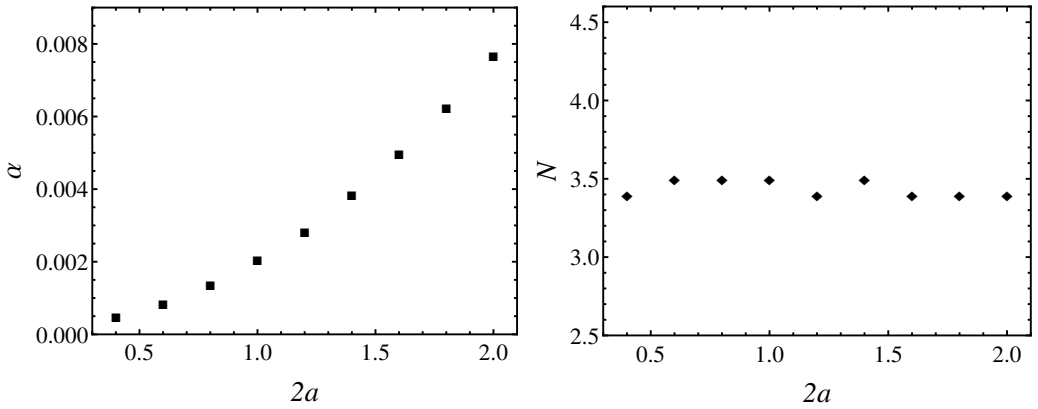


Figure 3.4: The parameters α (left) and N (right) averaged over all portfolios versus the uniform distribution width $2a$.

We now examine the case when all weights are equal to $1/K$. Since all weights are positive, this case represents a situation in which short selling is not allowed. Figure 3.5 shows the histogram of the empirical portfolio return R using daily returns and equal weights $w_k = 1/K$ with $K = 20$. The histogram is asymmetric with a heavier tail on the right hand side. Although the deviations in the tails are slightly more pronounced, the average return distribution with $N = 3.2$ still agrees

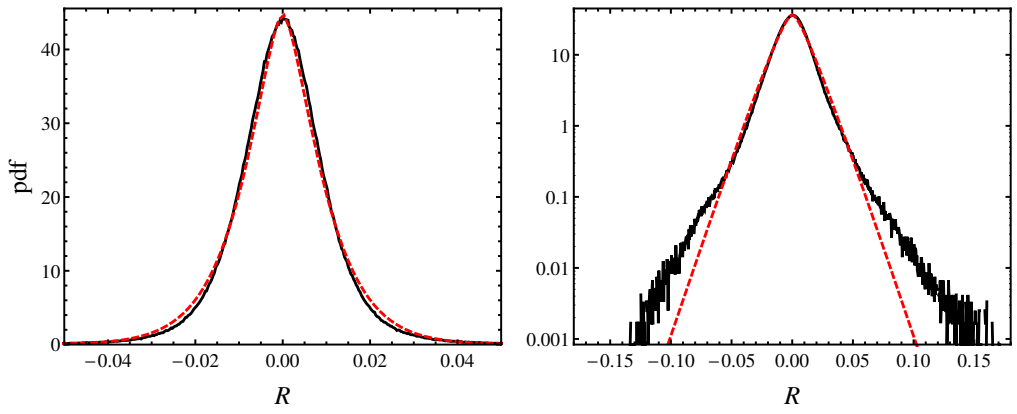


Figure 3.5: Histogram of the empirical portfolio return R (solid black) for daily returns and equal weights $w_k = 1/20$ compared with the average portfolio return pdf $\langle f \rangle(R|\Sigma, N)$ (dashed red) with $\alpha = 2.17 \times 10^{-4}$ and $N = 3.2$, plotted linearly (left) and logarithmically (right).

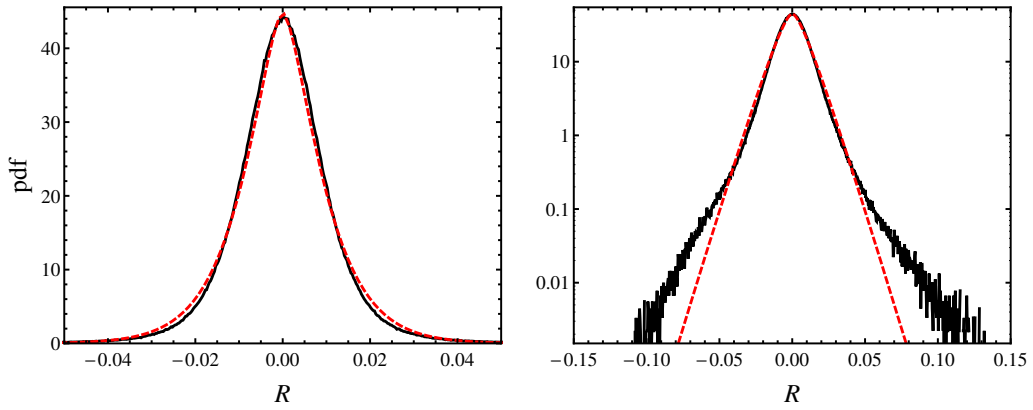


Figure 3.6: Histogram of the empirical portfolio return R (solid black) for daily returns and optimal weights compared with the average portfolio return pdf $\langle f \rangle(R|\Sigma, N)$ (dashed red) with $\alpha = 1.38 \times 10^{-4}$ and $N = 3.4$, plotted linearly (left) and logarithmically (right).

well with the histogram in the central part. The portfolio variance, on the other hand, is much smaller than in the case of uniform portfolio weights, which implies a lower risk for equally-weighted portfolios.

Lastly, we consider a set of optimal portfolio weights. We recall that the optimal weights are the solution of the optimization problem (1.18). Here, we consider

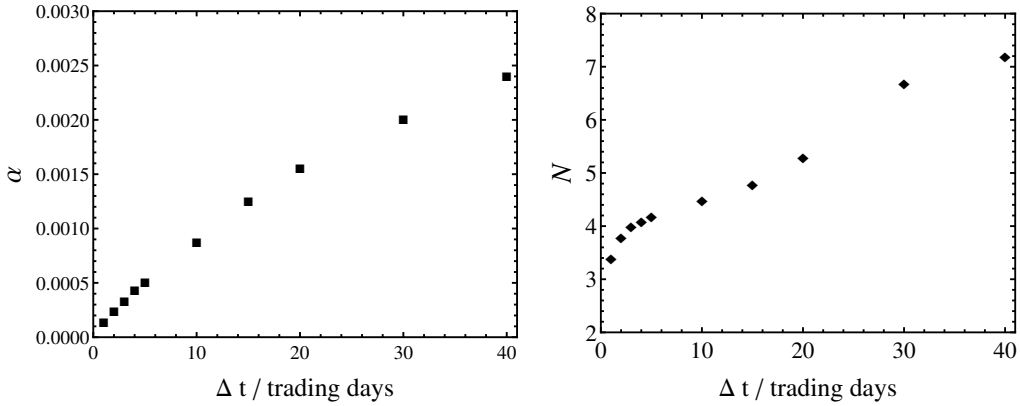


Figure 3.7: The parameters α (left) and N (right) averaged over all portfolios versus the return interval Δt using optimal weights.

minimum variance portfolios with weights

$$w_{\text{opt}} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^{\dagger}\Sigma\mathbf{1}} \ , \tag{3.12}$$

where Σ is the covariance matrix computed for each portoflio in the whole observation period and $\mathbf{1} = (1, \ldots, 1)$ a K dimensional vector. The histogram of the empirical portfolio return R compared with the average portfolio return distribution with $N = 3.4$ is shown in figure 3.6. The histogram is asymmetric with a heavier tail on the right hand side. Again, we observe a good agreement in the central part and deviations in the tails. Indeed, we find the smallest portfolio variance α, about a factor 1.5 smaller than the second best α for $w_k = 1/20$.

So far, we have only considered daily returns, $\Delta t = 1$ trading day. Let us now take a look at other return intervals. Figure 3.7 shows the parameters α and N for different return intervals between one day and two months. We observe that N increases with Δt, which leads to a more Gaussian-like distribution. The portfolio variance increases too.

Till now, we have used 600 random portfolios of size $K = 20$. What if we vary the number of portfolios or/and the number of stocks? Increasing the number of portfolios reveals more of the tails of the histogram. Figure 3.8 shows the case of varying the number of stocks K. As the number of stocks K increases, the value of N increases too. The variance on the other hand decreases. In other words, we still see the benefit of diversification.

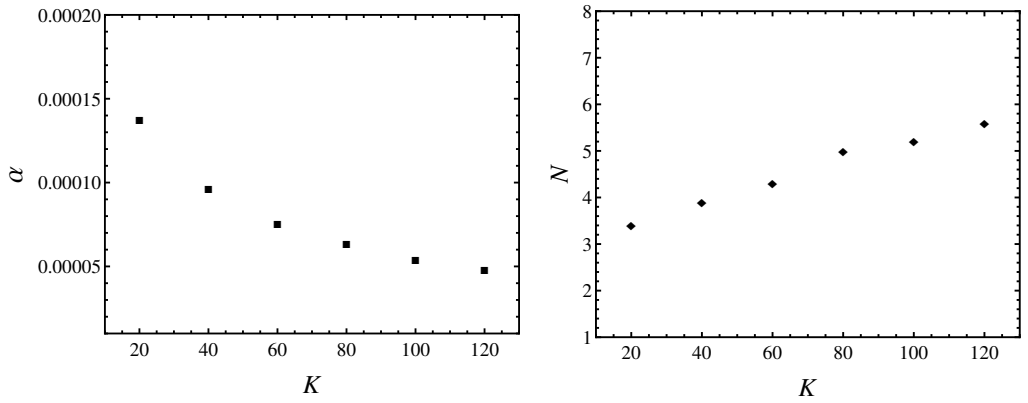


Figure 3.8: The parameters α (left) and N (right) averaged over all portfolios versus the portfolio size K for daily returns, $\Delta t = 1$ day, and optimal weights.

3.4 Summary

We presented an application of the random matrix approach introduced in chapter 2. In particular, we studied the implications of time-varying correlations on portfolio return distributions. The random matrix approach allowed us to derive a distribution for the return of a portfolio, which takes the non-stationary correlations into account. It is a heavy-tailed distribution which depends only on two parameters: the bilinear form of the portfolio weights with the average empirical covariance matrix, which represents a portfolio variance, and a free parameter which characterizes the fluctuation strength of correlations.

We found that the average portfolio return distribution describes the empirical data well, in particular in the central part of the distribution. This behavior is independent of the choice of the portfolio weights. Still, there are deviations in the far tails. This can be traced back to the fact that although the choice of the Wishart distribution for the random correlation matrix ensemble is indeed a reasonable assumption, it obviously cannot capture all empirical details. Nevertheless, our results have a remarkable agreement with the data up to the third decade.

It is important to note that the average portfolio return distribution yields a good description of heavy-tailed portfolio return distributions with only one free parameter. It provides a better fit to the empirical data than the Student's t-distribution, which is one of the standard heavy-tailed distributions used in financial economics, see e.g., references [54, 55].

Chapter 4

Market states: A correlation structure analysis

4.1 Introduction

The correlation matrix is a central object in the study of financial markets. In the previous chapters, we addressed the non-stationarity of correlations. We introduced a random matrix approach to model this non-stationarity and discussed its impact on return distributions. Here, we use correlation matrices to identify states of a financial market and study their correlation structures by means of our random matrix approach.

The concept of different market states or regimes in which the market operates is not entirely new to the economics literature, see e.g., references [111, 112]. Our study is based on the definition of market states as first suggested by Münnix *et al.* [69]. In their empirical study, Münnix *et al.* analyzed the correlation structure of a financial market over a long time period and observed significant structural changes. They identified market states as clusters of correlation matrices with similar correlation structures and found that each market state has a characteristic correlation structure and temporal behavior. Over time, the market switches back and forth between these states.

Here, we take a closer look at the statistics of market states. In particular, we investigate the stability of the corresponding correlation structures by estimating the fluctuations of true correlations due to their non-stationarity. As fluctuations due to measurement noise compete with actual fluctuations, we cannot study the stability of the correlation structure directly from the empirical correlation matrices. To circumvent this problem, we use the random matrix approach introduced in chapter 2. It reduces the effect of fluctuating correlations to a single parameter which

measures the fluctuation strength. Our approach provides a method to estimate the correlation fluctuations due to non-stationarity directly from the empirical data, and thus to assess the stability of the correlation structure. In addition, we look at the correlation structure dynamics and investigate the relationship between average correlation and correlation fluctuations.

The chapter is organized as follows: In section 4.2, we identify market states for the NASDAQ Composite stock market in the time period $1992 - 2013$ and study their dynamics. We apply the random matrix approach to study the stability of the correlation structure of each market state in section 4.3 and the correlation structure dynamics in the whole observation period in section 4.4. We conclude our findings in section 4.5. The contents of this chapter are published in reference [3].

4.2 Market states: Identification and dynamics

We begin with the identification of market states as clusters of correlation matrices with similar correlation structures. We consider the same dataset used in section 2.4.2. It consists of $K = 258$ stocks of the NASDAQ Composite index traded in the 22-year period from January 1992 to December 2013, i.e., 5542 trading days. For each stock k we calculate the return time series

$$r_k(t) = \frac{S_k(t + \Delta t) - S_k(t)}{S_k(t)} \ , \quad k = 1, \ldots, K \ , \tag{4.1}$$

where $S_k(t)$ is the price of the k-th stock at time t and Δt is the return interval. We choose Δt to be one trading day and calculate the daily returns for each stock.

The main object of interest in the following is the $K \times K$ correlation matrix C which contains the correlation coefficients between all pairs of return time series. We recall that the correlations between time series are commonly measured via the Pearson correlation coefficient

$$C_{kl} = \frac{\langle r_k(t) r_l(t) \rangle - \langle r_k(t) \rangle \langle r_l(t) \rangle}{\sqrt{\langle r_k^2(t) \rangle - \langle r_k(t) \rangle^2} \sqrt{\langle r_l^2(t) \rangle - \langle r_l(t) \rangle^2}} \ , \tag{4.2}$$

where $\langle \cdot \rangle$ denotes the average over a time window yet to be specified. As discussed in section 1.2.6, the Pearson correlation coefficient is the most common measure of dependence. Sometimes, however, it can be problematic, in particular for non-linear dependencies or for non-stationary data. The latter is extremely relevant for financial data since drift and volatilities fluctuate considerably in time. Thus, the correlation coefficient averages over time-varying trends and volatilities, which results in an estimation error of the correlations.

In order to eliminate the estimation error due to non-stationary trends and

volatilities, we employ the method of local normalization introduced by Schäfer *et al.* [72]. For each return time series k we subtract the local mean $\mu_k(t)$ and divide by the local standard deviation $\sigma_k(t)$

$$\hat{r}_k(t) = \frac{r_k(t) - \mu_k(t)}{\sigma_k(t)} = \frac{r_k(t) - \langle r_k(t) \rangle_n}{\sqrt{\langle r_k^2(t) \rangle_n - \langle r_k(t) \rangle_n^2}} \quad , \tag{4.3}$$

where the local average

$$\langle r(t) \rangle_n = \frac{1}{n} \sum_{j=0}^{n-1} r(t - j\Delta t) \tag{4.4}$$

runs over the most recent n sampling points. For daily data we use $n = 13$, which yields nearly normally distributed time series, as discussed in reference [72]. The local normalization removes the local trends and variable volatilities while preserving the correlations between the time series. Thus, the locally normalized time series are better suited for correlation analysis than the original time series [72]. An alternative approach would be to use the residuals of a GARCH fit [62], which also yields stationary time series. We choose the local normalization, as it does not require any model assumptions.

We now divide the locally normalized time series $\hat{r}(t)$ into disjoint two-month time intervals and estimate the correlation matrix for each interval. The choice of two months, i.e., 42 trading days, for the estimation interval of the correlation matrices is a trade-off between reducing the estimation noise and still being able to resolve changes in the correlation structure. In previous empirical studies, a characteristic time scale of three months has been found for the correlation dynamics [113, 114]. Thus, we obtain a set of 131 correlation matrices for the whole 22-year observation period.

To identify the market states, we perform a clustering analysis based on the Partitioning Around Medoids (PAM) algorithm [115], where the number of clusters is estimated via the gap statistic [116]. The clustering analysis separates the set of 131 correlation matrices into six groups based on the similarity of their correlation structures. Each group is associated with a market state. Figure 4.1 (top) shows the time evolution of the market states. The market switches back and forth between states: Sometimes it remains in a state for a long time, sometimes it jumps briefly to another state and returns back or evolves further. On longer time scales, the market evolves towards new states, whereas previous states die out. How frequently does the market switch between states? Figure 4.1 (bottom) shows the number of jumps from one state to another calculated on a one-year sliding window. After a stable five-year period, we observe that the market begins to switch between

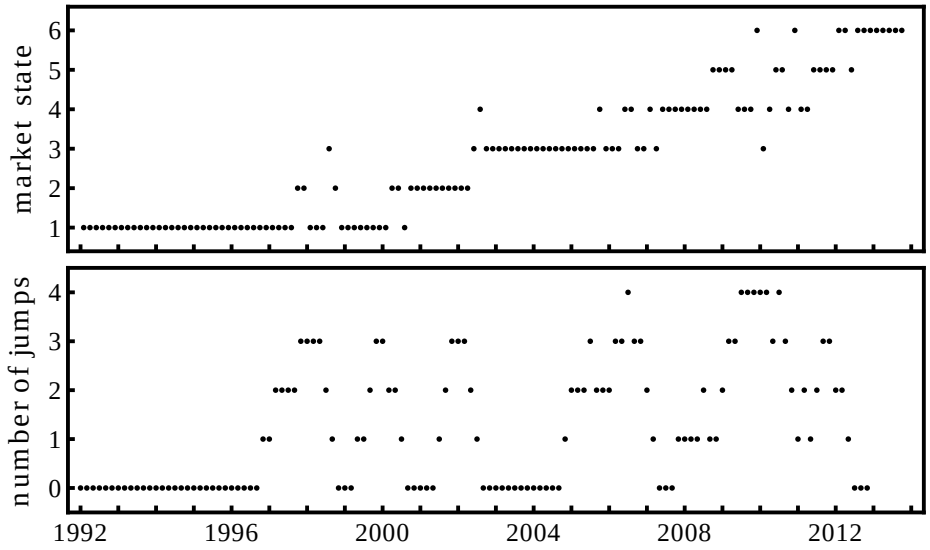


Figure 4.1: Top: Time evolution of the market in the observation period $1992-2013$. Each point represents a correlation matrix measured over a two-month time window. Bottom: Number of jumps between states calculated on a one-year sliding window. The first point represents the number of jumps in the period $1/92 - 12/92$, the second point–in the period $3/92 - 2/93$ and so on.

states. The highest number of jumps per year can be found in the period around 2010. In figure 4.2 we compare the number of jumps frequency in both halves of the observation period. In the second half of the observation period the number of jumps per year increases. At the same time, the lifetime, i.e., the time the market stays in a certain state before it jumps to another one, decreases. Figure 4.3 shows the histograms of the lifetime in both halves of the observation period. The first half contains mostly long-lived states. In the second half the frequency of the short-lived states increases considerably, while the frequency of the long-lived states decreases.

To illustrate the different correlation structures of each state, we sort the stocks according to their industry sector and calculate the corresponding average correlation matrices, see figure 4.4. We indeed recognize different characteristic correlation structures. State 1 shows an overall weak correlation. In state 2 we have the strongest correlation within the technology sector and between technology and capital goods, whereas in states 3 and 4 the correlation within the finance sector is the strongest. We observe that the average correlation level increases from state

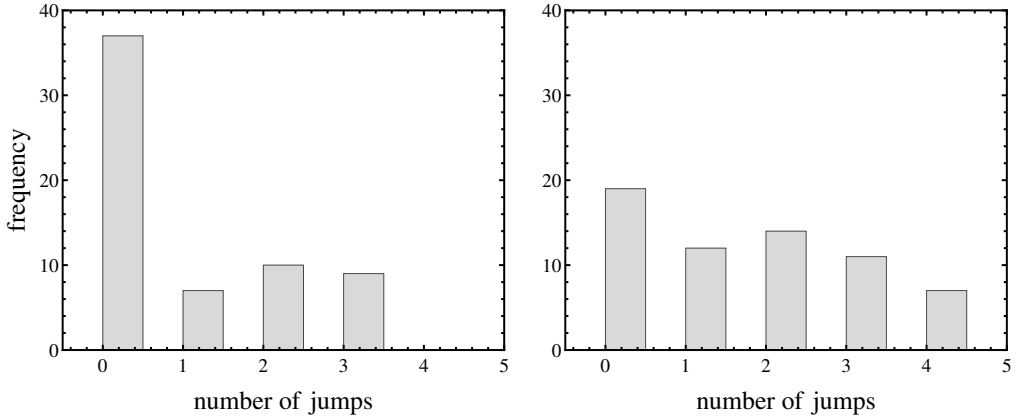


Figure 4.2: Histograms of the number of jumps between states in the first half $1992 - 2002$ (left) and in the second half $2003 - 2013$ (right) of the observation period.

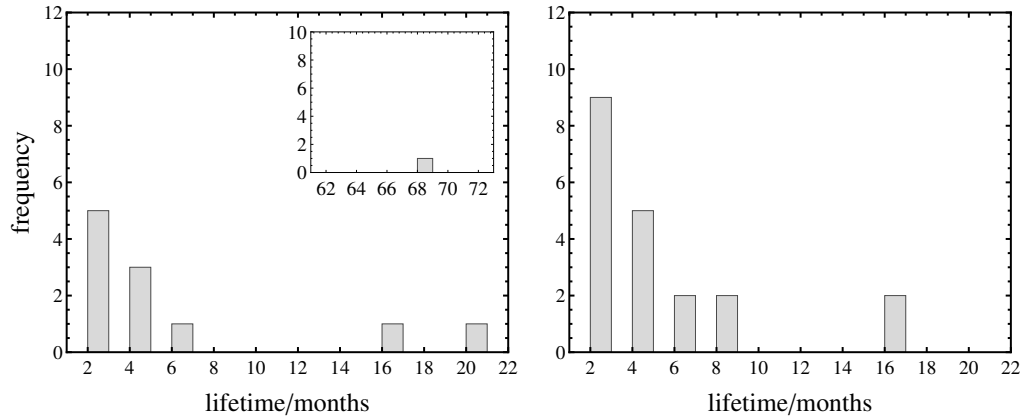
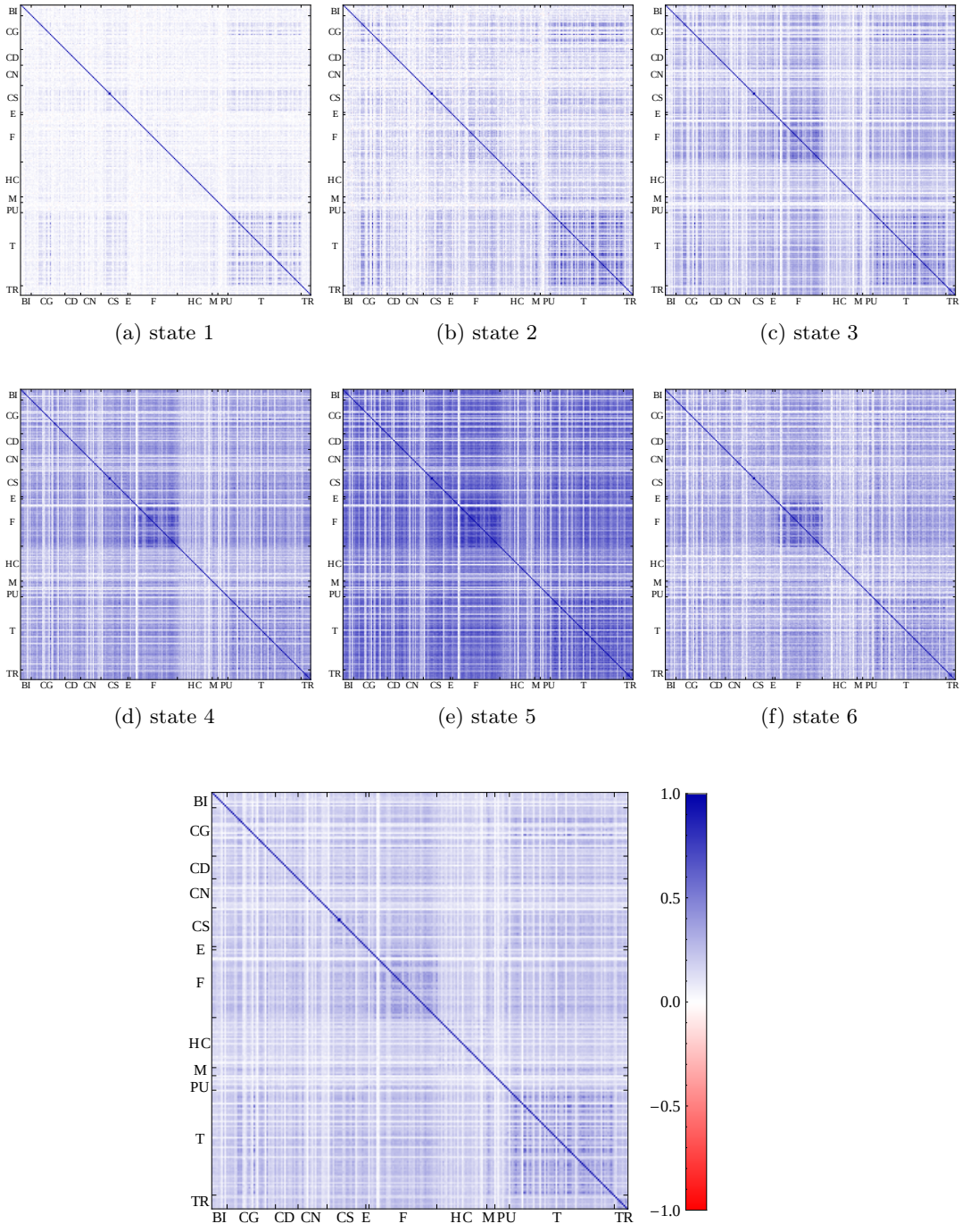


Figure 4.3: Histograms of the lifetime in months in the first half $1992 - 2002$ (left) and in the second half $2003 - 2013$ (right) of the observation period.

to state, reaching its highest value in state 5. In state 6 the average correlation level decreases. The finance sector, however, is still strongly correlated. Further, we note that the health care sector is weakly correlated to the rest of the market in almost all states.



(a) state 1 (b) state 2 (c) state 3

(d) state 4 (e) state 5 (f) state 6

(g) Average correlation matrix

Figure 4.4: (a)–(f) Average correlation matrices for each market state. (g) Overall average correlation matrix. Industry sectors legend: BI: Basic Industries, CG: Capital Goods, CD: Consumer Durables, CN: Consumer Non-Durables, CS: Consumer Services, E: Energy, F: Finance, HC: Health Care, M: Miscellaneous, PU: Public Utilities, T: Technology, TR: Transportation.

4.3 Stability of the correlation structure for each market state

In the prevoius section, we used a clustering analysis to group our original set of 131 correlation matrices, calculated on two-month time intervals of the 22-year observation period, into six distinct groups based on the similarity of the correlation structure. Each of these six groups, which we identify as different market states, is characterized by its average correlation matrix $C^{(i)}$. In this section, we study the stability of the correlation structure for each state. In particular, we address the question: Are the correlations of a given state stationary or do they fluctuate around the respective average state correlation matrix $C^{(i)}$, and if so, how strongly? This question, however, cannot be answered by looking at the empirical correlation matrices. Since they are calculated on very short time intervals of two months, they contain a considerable amount of noise. The clustering is not so sensitive to the noise level, since it is based on a distance measure which averages over the noise. This noise, however, competes with actual fluctuations of the correlations. Thus, it prevents us from studying the stability of the correlation structure directly. In order to assess the actual fluctuations of correlations due to non-stationarity, we use the random matrix model introduced in chapter 2. Instead of assuming a constant correlation matrix for each market state, we assume a Wishart ensemble of correlation matrices which allows for fluctuations around the average state correlation matrix $C^{(i)}$. Further, we assume conditional normality, i.e., return vectors following a multivariate normal distribution conditioned on a fixed correlation or covariance matrix. Not only is this assumption common in the literature (e.g. in GARCH models and stochastic volatility models), but it is also justified by empirical data, as discussed in chapter 2. The two extreme assumptions of stationary versus non-stationary correlations lead to different sample statistics for the multivariate returns observed within a given market state. In case of stationary correlations, we would expect a normal distribution. In case of non-stationary correlations, we expect a distribution of the form (2.35) instead. Once the average state correlation matrix is fixed, we may use the parameter $N^{(i)}$ to fully characterize the fluctuations. By fitting to the empirical return distributions we obtain a measure $N^{(i)}$ for these fluctuations. In a further step, we compress the information contained in each $C^{(i)}$ into a single number, namely an average correlation coefficient $c^{(i)}$. This allows us to study the relationship between fluctuations and average market correlation.

We obtain the return time series for each market state in the following way: We take the original daily return time series $r(t)$ and divide it into a sequence of disjoint two-month intervals. We merge all intervals belonging to a given state according to the cluster analysis described in section 4.2. We note that the return

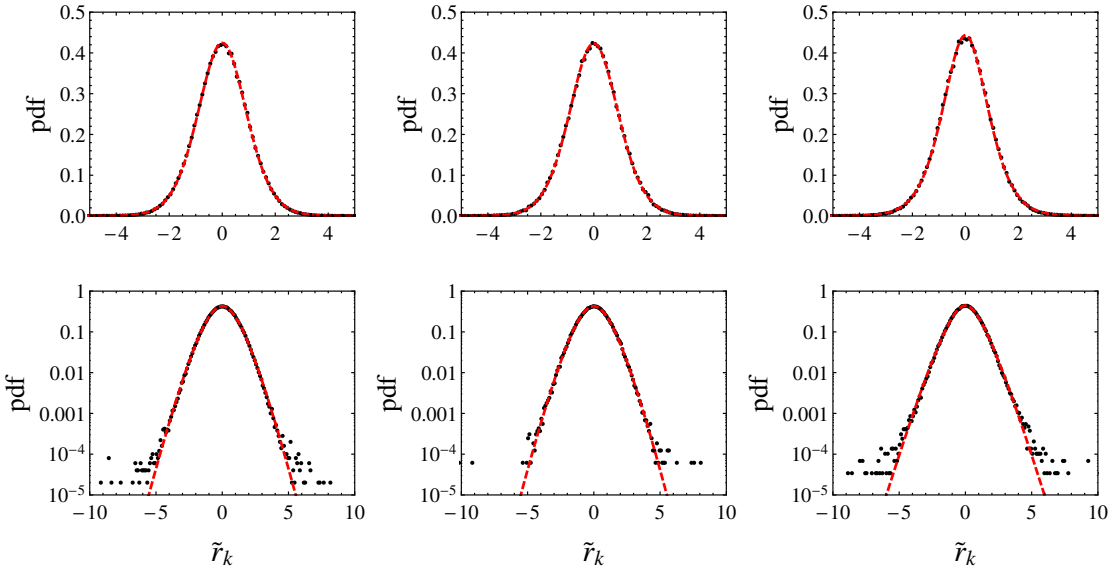
time series for the six market states differ in length.

For the comparison with the model, we rotate the empirical return vectors for each state into the eigenbasis of the state covariance matrix $\Sigma^{(i)}$ and normalize the components of the rotated vectors with the corresponding empirical eigenvalues. We then aggregate all components into a single histogram and compare it with the average return distribution (2.35). Figure 4.5 shows the results for each market state. It is already obvious from the heavy-tailed empirical return distributions that the assumption of stationary correlations within a market state has to be dismissed. Instead, we observe a clear indication for fluctuations around each average state correlation matrix $C^{(i)}$. For the whole observation period we found a much smaller N, see figure 2.6(a). In this case, the fluctuations are stronger than for the single states. The parameter $N^{(i)}$ is estimated by the maximum likelihood method and depicted for each state together with the average correlation $c^{(i)}$ in figure 4.6. We obtain $c^{(i)}$ by averaging over the off-diagonal correlation coefficients $C_{kl}^{(i)},\ k \neq l$ of the average correlation matrix for a given state. We observe that the states 1 and 2, which cover the period 1992 to roughly 2002, are rather stable. We find low average correlation with high N values, i.e., weak fluctuations. In state 3 and 4 the fluctuations increase. While the N values for both states are equal, the average correlation is rising. In state 5, first appearing during the crisis in 2008, the fluctuations increase further. It is the most unstable state with the smallest N value and the highest average correlation. In state 6 the fluctuations and the average correlation decrease, the market stabilizes. To examine the relationship between average correlation and fluctuations we look at the scatter plot between $c^{(i)}$ and $N^{(i)}$, see figure 4.7. We observe a clear decreasing trend, i.e., a negative correlation.

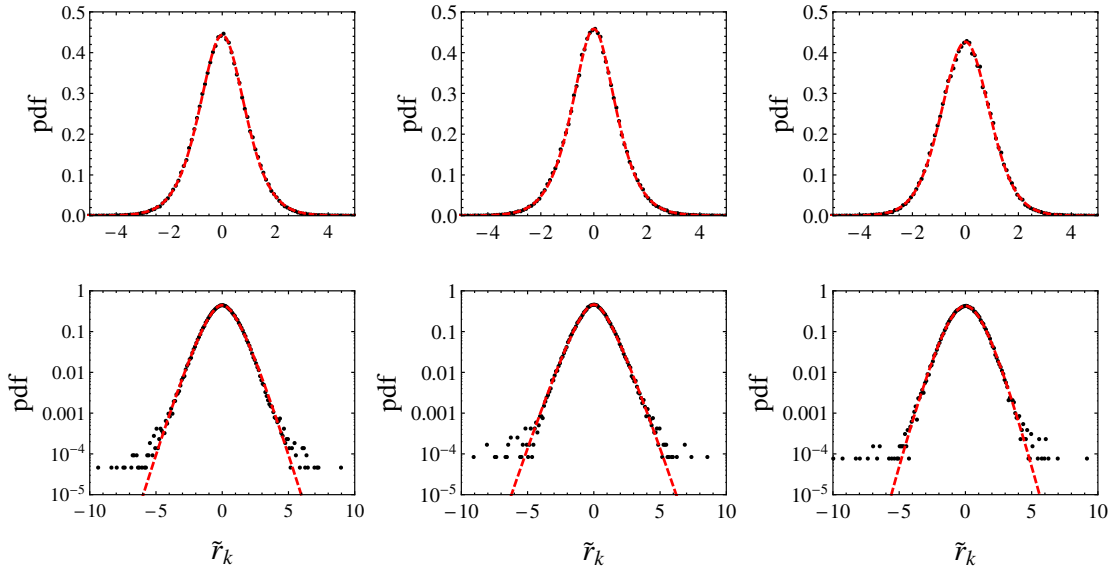
4.4 Dynamics of the correlation structure

To further investigate the relationship between fluctuations and average correlation, we now take a closer look at the variation of the correlation structure over time. To this end, we examine the parameter N and the average correlation c computed on a sliding window of 500 trading days shifted by 21 trading days, see figure 4.8. As in the previous section, the parameter N for each time window is estimated by fitting to the aggregated distribution of the rotated and scaled returns, where for the rotation we use the average covariance matrix in the given time window. The parameter c is obtained by averaging over the off-diagonal elements of the average correlation matrix for the corresponding time window. We recognize four distinct regimes:

- The first regime covers the period 1/1992 to 9/1996, which mainly corresponds to the stable market state 1, see figure 4.1. Here, we find the lowest average



(a) state 1: $N^{(1)} = 12.8$ (b) state 2: $N^{(2)} = 13.4$ (c) state 3: $N^{(3)} = 7.8$



(d) state 4: $N^{(4)} = 7.8$ (e) state 5: $N^{(5)} = 6.1$ (f) state 6: $N^{(6)} = 12.0$

Figure 4.5: (a)–(f) Histograms of the rotated and rescaled returns for each market state compared with the average return distribution (2.35).

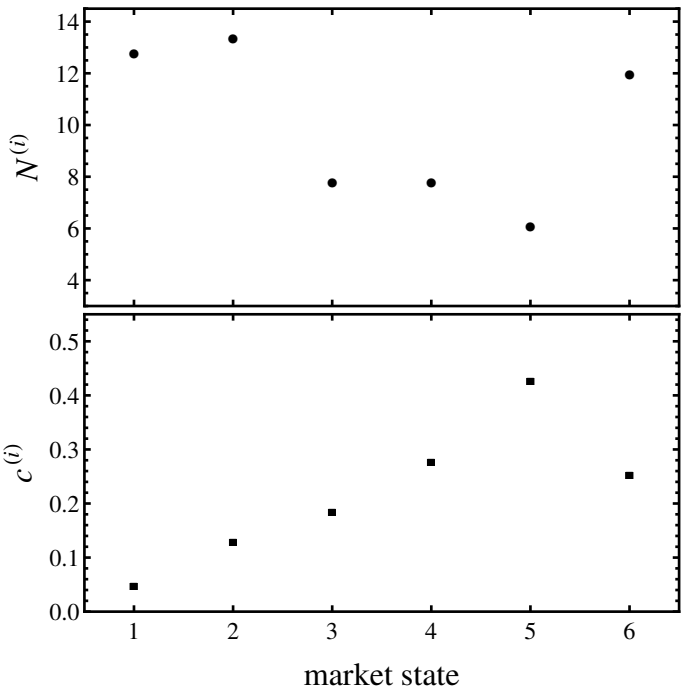


Figure 4.6: The parameter $N^{(i)}$ (top) and the average correlation $c^{(i)}$ (bottom) for each market state.

correlation and the highest N value. While the average correlation is relatively stable in this period, the N value shows a clear decreasing trend indicating increasing fluctuations.

- The second regime covers the period 10/1996 to 9/2006, which corresponds to the stable states 1 and 2 and the more unstable states 3 and 4. While the average correlation in this period is steadily growing, the N value is mostly stable. Compared to the first regime, we find smaller N values because of the transitions between the different market states.

- The third regime, beginning 10/2006 to 5/2009, covers mostly the period before and during the financial crisis in 2008 and corresponds to the unstable states 4 and 5. We observe a sharp increase in the average correlation, which is over two times larger compared with the first regime. Indeed, in times of market instabilities collective behavior is induced which results in larger correlations. Here, we find the smallest N values, i.e., the strongest fluctuations.

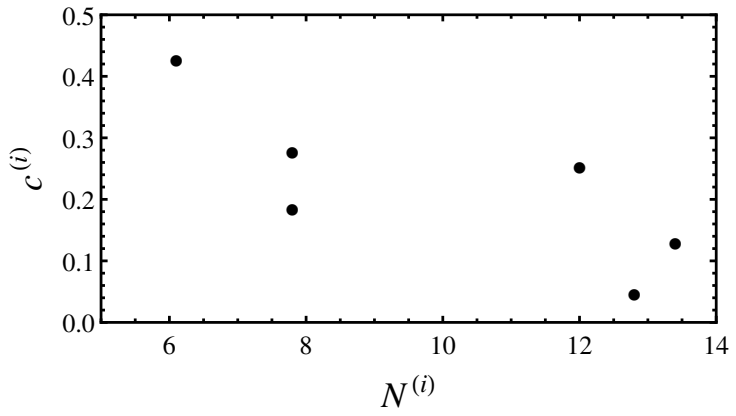


Figure 4.7: Scatter plot $c^{(i)}$ versus $N^{(i)}$.

- The last regime, beginning 6/2009, covers the rest of the observation period and corresponds to the unstable states 4 and 5 and the stable state 6. The fluctuations decrease slightly. The average correlation increases at first even further compared to the previous regime, but decreases again after 2010. The market stabilizes after the crisis.

The relationship between average correlation and fluctuations for the two-year time window is depicted in figure 4.9(a). We observe an overall negative correlation between c and N. Moreover, the data corresponding to the four regimes cluster into different regions: a stable region (regime I) characterized by low average correlation and weak fluctuations, which are typical for calmer periods; an unstable region (regime III) characterized by high average correlation and strong fluctuations, typical for crisis periods; and an intermediate region (regime II and IV) characterized by varying average correlation and more moderate fluctuations.

Finally, we examine the dependence between average volatility σ and fluctuations for the two-year time window, shown in figure 4.9(b). Again, we find clustering into regions as observed before: a stable region with weak fluctuations and nearly constant volatility $\sigma \approx 0.03$; an unstable region with strong fluctuations and high volatility; and an intermediate region. In this case, we do not recognize a clear trend, σ and N show no clear dependence. This justifies our interpretation of N as correlation rather than covariance fluctuations.

4.5 Summary

To achieve a better understanding of the financial market, the concept of market states as clusters of correlation matrices was introduced in reference [69]. Here,

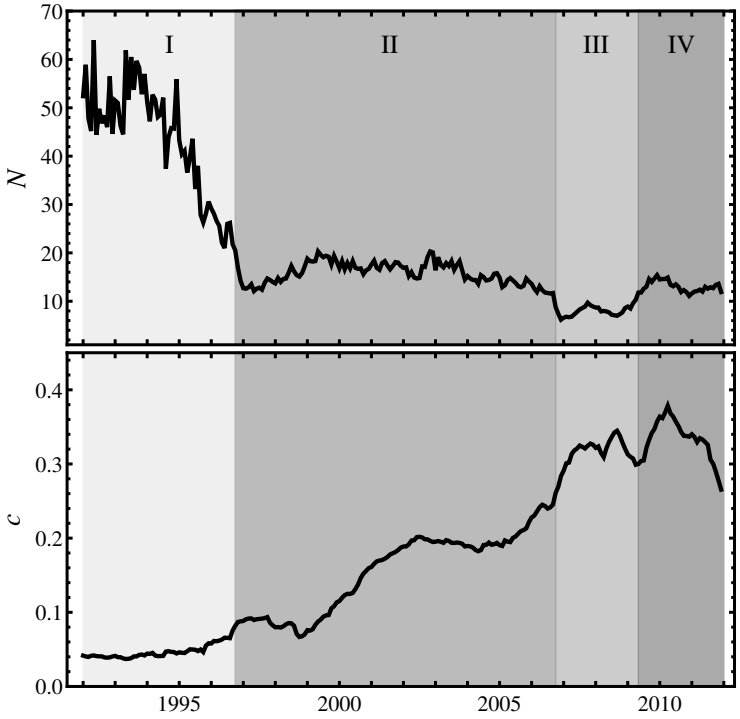


Figure 4.8: Time evolution of the parameter N (top) and the average correlation c (bottom) calculated on a sliding window of 500 trading days shifted by 21 trading days. The plots are divided into four regimes indicated by different gray scales.

we took a closer look at the statistics of market states studying the NASDAQ Composite market over a period of 22 years. To this end, we used our random matrix approach, which models the non-stationarity of true correlations by a random matrix ensemble. Alongside with a heavy-tailed distribution for the stock returns, this approach provides a method to study the correlation structure by estimating the fluctuation strength of correlations directly from the empirical return distributions.

Our study provides a better understanding of the market state dynamics as well as of the stability of the corresponding correlation structure. Despite the non-stationarity of the market, we found a set of quite stable states in which the market operates. We discussed their statistical properties and studied the dynamics of the correlation structure in the whole observation period using a sliding window analysis. We found four distinct regimes with different statistical behavior. The analysis

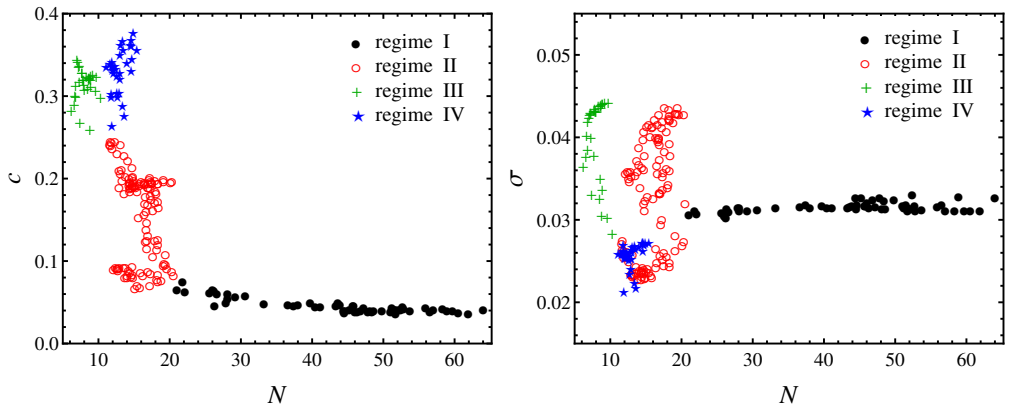


Figure 4.9: Scatter plots c versus N (left) and σ versus N (right), for the time window of 500 trading days.

revealed a remarkable relationship between average correlation and fluctuations. Strong fluctuations most likely occur during periods of high average correlation. Unstable periods are thus characterized not only by larger correlations and high volatilities but also by strong correlation fluctuations. Furthermore, we studied the relationship between fluctuations and average volatility. In this case, we did not find a clear trend, volatility and fluctuations are mostly independent of each other.

Another, conceptual aspect of the present study should be mentioned. At first sight, the following two results might appear contradictory: when studying the entire time interval of 22 years from 1992 to 2013, we identify, on the one hand, a small number of distinct states in which the market operates, but on the other hand, we claim that the return distribution for this entire time interval can be modeled by a random matrix ansatz. The simultaneous existence of few distinct states and of an ensemble of homogenously distributed correlation matrices in the random matrix ensemble might seem incompatible. Importantly, these two features can coexist. The present study may be viewed as a refined resolution of the really existing (random) matrix ensemble in terms of a superposition of sub-ensembles around the distinct states. Certainly, one might come up with statistical observables that can make this fine structure of the ensemble visible – as we do in the present study. However, the plain return distribution itself is a highly relevant quantity, see e.g., reference [5]. To study it, the data are aggregated, i.e., represented in the eigenbasis of the mean covariance matrix for the entire time interval. This involves a rotation of the return vector with an orthogonal matrix and thus further randomization which is just an additional averaging over the sub-ensembles. This is why our random matrix ensemble works despite of the existence of distinct market

states.

Such effects are quite common in random matrix models and one of the reasons for their remarkable robustness. Wigner's original random matrix ansatz [117] based on a rotation invariant and homogenous ensemble was heavily criticized by many nuclear physicists. They argued that no realistic Hamilton matrix of a nucleus, calculated in some basis, will look like a random matrix, because it will contain many strict zeros due to selection rules. Thus, there was a blatant disagreement with the idea of a rotation invariant ensemble. This then led to the embedded random matrix ensembles [118] which correctly incorporate the selection rules. Nevertheless, the statistical observables of particular interest are indistinguishable for Wigner's original ansatz and for the embedded random matrix ensembles.

Chapter 5

Dependence structure of market states: A copula approach

5.1 Introduction

So far, we have considered the correlation as a measure of dependence. However, the correlation coefficient only captures the linear dependence between two random variables. Non-linear dependencies are not captured appropriately. Here, we choose a copula approach to study the statistical dependence of stock returns.

The concept of copulas was introduced by Abe Sklar in 1959 [119, 120] to study the linkage between multivariate distribution functions and their univariate marginals. Since then, copulas have gained growing importance as a tool for modeling statistical dependence of random variables in many fields. In finance, the usage of copulas is relatively new, but it has already found application in risk management, see e.g., references [121–127], derivative pricing, see e.g., references [128–133], and portfolio optimization, see e.g., references [134–137]. For an overview of the literature on applications of copulas in finance, the reader is referred to references [138, 139]. Copulas allow to separate the dependence structure of random variables from their marginal distributions. This is sometimes useful in statistical applications as the dependence structure and the marginal distributions can be modeled separately and joined together resulting in new multivariate distributions with different behavior. For a discussion on difficulties in the application of copulas, the reader is referred to references [140–142]. Here, we simply view the copulas as providing a standardized way for the study of statistical dependences. The marginal distributions are mapped to the uniform distribution; the statistical dependence is considered in terms of the marginal cumulative distribution functions.

In chapter 4, we identified market states as clusters of similar correlation matrices

and studied their corresponding correlation structures. Here, we choose a copula approach to study the dependence structure of these market states. To this end, we estimate empirical copulas for many stock return pairs and average over all of them to obtain an empirical pairwise copula for each market state. To estimate the empirical copulas we use both original returns, which exhibit time-varying trends and volatilities, as well as locally normalized ones, which show stationary behavior. The resulting empirical copulas provide different information. The copulas for the original returns describe the dependence structure for the full time horizon, i.e., on a global scale, whereas the copulas for the locally normalized returns describe the dependence structure on a local scale.

The empirical pairwise copulas for each market state are compared with a bivariate K-copula, which arises from the random matrix approach introduced in chapter 2. The approach yields a multivariate return distribution in terms of a modified Bessel function of the second kind, a so-called K-distribution. In chapter 4, the K-distribution was found to provide a good description of the heavy-tailed empirical return distributions for each market state. Here, we aim to arrive at a consistent description within the random matrix model studying the agreement between K-copula and empirical dependence structure for each market state. In addition, our study provides further evidence for asymmetric dependencies between financial returns [143–145]. We find an asymmetry in the tail dependence of empirical pairwise copulas, which we study in more detail.

The chapter is organized as follows: In section 5.2, we review the basic concepts of copulas, stating the main result in the copula theory, the Sklar's theorem, which we use to derive the K-copula. In section 5.3, we study the empirical copula densities for each of the market states identified in chapter 4 and compare them with the K-copula density. We conclude our findings in section 5.5. The contents of this chapter are published in reference [4].

5.2 Copula

We begin with a short introduction to the concept of copulas. For more details with an emphasis on the statistical and mathematical foundations of copulas the reader is referred to the textbooks of Joe [146] and Nelsen [147].

5.2.1 Basic concepts

Consider two random variables X and Y. We confine ourselves to the bivariate case since we will only study pairwise copulas. The joint distribution of X and Y contains all the statistical information about them. It can be expressed either in terms of the joint probability density function (pdf) $f_{X,Y}(x, y)$ or in terms of the

joint cumulative distribution function (cdf) $F_{X,Y}(x, y)$, where

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \mathrm{d}x' \int_{-\infty}^{y} \mathrm{d}y' \; f_{X,Y}(x', y') \; . \tag{5.1}$$

From the joint pdf $f_{X,Y}(x, y)$ one can extract the individual distributions of X and Y as follows

$$f_X(x) = \int_{-\infty}^{\infty} \mathrm{d}y \; f_{X,Y}(x, y) \; , \tag{5.2}$$

and analogously for Y. The densities $f_X(x)$ and $f_Y(y)$, called marginal probability density functions, and the corresponding marginal cumulative distribution functions $F_X(x)$ and $F_Y(y)$ describe the individual statistical behavior of the random variables.

When dealing with correlated random variables, one is interested in their statistical dependence. The Pearson correlation coefficient is commonly used as a measure of dependence. It is defined as

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} \; , \tag{5.3}$$

where $\mathrm{Cov}(X, Y)$ is the covariance of both random variables and σ_X and σ_Y are the respective standard deviations. However, the correlation coefficient only measures the linear dependence between the random variables. Non-linear depnendencies are not captured appropriately.

Copulas provide a natural way to study the statistical dependence of random variables. What is the copula of X and Y? The probability integral transformation

$$U_i = F_i(i) \qquad i = X, Y \tag{5.4}$$

leads to new random variables called the ranks of X and Y, respectively. The distribution of the ranks is uniform on the unit interval $[0, 1]$, regardless of the original distribution F_i. The copula of X and Y is defined as the joint distribution of their ranks (U_X, U_Y). Thus, a copula is a multivariate distribution function with uniform marginals on the unit interval. It completely describes the dependence structure between the random variables. Any measure of dependence which is scale invariant, i.e., invariant under strictly increasing transformations of the underlying variables, can be expressed in terms of the copula alone. Such measures of dependence are the rank correlation and the tail dependence coefficients. Importantly, the Pearson correlation coefficient cannot be expressed in terms of the copula [148].

A central result in the copula theory is the Sklar's theorem. It states that if

$F_{X,Y}$ is a bivariate distribution function with marginal distributions F_X and F_Y, then there exists a copula such that

$$F_{X,Y}(x,y) = \mathrm{Cop}_{X,Y}(F_X(x), F_Y(y)) \ . \tag{5.5}$$

If F_X and F_Y are continuous, then the copula is unique, else the copula is defined only on the range of F_X and F_Y. Conversely, for any marginal distributions F_X and F_Y and copula $\mathrm{Cop}_{X,Y}$, the function $\mathrm{Cop}_{X,Y}(F_X(x), F_Y(y))$ defines a bivariate distribution function with marginals F_X and F_Y. The Sklar's theorem provides the theoretical foundation of the widespread use of copulas in generating multivariate from univariate distributions. Furthermore, it enables us to extract the dependence structure directly from the joint distribution function

$$\mathrm{Cop}_{X,Y}(u,v) = F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)) \ , \tag{5.6}$$

where F_X^{-1} and F_Y^{-1} represent the inverse cumulative distribution functions, the so-called quantile functions. From the copula (5.6), one can compute the copula density by taking the derivative

$$\mathrm{cop}_{X,Y}(u,v) = \frac{\partial^2 \mathrm{Cop}_{X,Y}(u,v)}{\partial u \partial v} = \frac{f_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v))}{f_X(F_X^{-1}(u)) f_Y(F_Y^{-1}(v))} \ , \tag{5.7}$$

where $f_{X,Y}$ denotes the joint pdf and f_X and f_Y are the marginal pdfs.

5.2.2 Examples of copulas

Here, we provide two examples of copulas, namely the Gaussian and the Student's t-copula, which find numerous applications in financial modeling, see e.g., references [104, 105].

- **Gaussian copula**: The Gaussian copula is the dependence structure associated with the multivariate normal distribution. In the two-dimensional case with the covariance matrix

$$\Sigma = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} \ , \tag{5.8}$$

the bivariate Gaussian copula is given by

$$\mathrm{Cop}_c(u,v) = \Phi_c(\Phi^{-1}(u), \Phi^{-1}(v)) \ . \tag{5.9}$$

Here, Φ_c describes the cumulative distribution function of a bivariate standard normal distribution with correlation c

$$\Phi_c(x,y) = \int_{-\infty}^{x} \mathrm{d}x' \int_{-\infty}^{y} \mathrm{d}y' \; \frac{1}{2\pi\sqrt{1-c^2}} \exp\left(-\frac{(x')^2+(y')^2-2x'y'}{2(1-c^2)}\right) \quad (5.10)$$

and Φ^{-1} is the inverse distribution function of a univariate standard normal distribution

$$\Phi(x) = \int_{-\infty}^{x} \mathrm{d}x' \; \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x')^2}{2}\right) \; . \quad (5.11)$$

- **Student's t-copula**: The Student's t-copula is the dependence structure associated with the multivariate Student's t-distribution. The bivariate Student's t-copula reads

$$\mathrm{Cop}_{c,\nu}(u,v) = t_{c,\nu}(t_\nu^{-1}(u), t_\nu^{-1}(v)) \; . \quad (5.12)$$

Here, $t_{c,\nu}$ describes the cumulative distribution function of a bivariate Student's t-distribution with correlation c and degrees of freedom ν

$$t_{c,\nu}(x,y) = \int_{-\infty}^{x} \mathrm{d}x' \int_{-\infty}^{y} \mathrm{d}y' \; \frac{1}{2\pi\sqrt{1-c^2}} \left(1+\frac{(x')^2+(y')^2-2x'y'}{\nu(1-c^2)}\right)^{-(\nu+2)/2} \quad (5.13)$$

and t_ν^{-1} the inverse distribution function of a univariate Student's t-distribution

$$t_\nu(x) = \int_{-\infty}^{x} \mathrm{d}x' \; \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\nu\pi}} \left(1+\frac{(x')^2}{\nu}\right)^{-(\nu+1)/2} \; . \quad (5.14)$$

The Gaussian and Student's t-copula belong to the class of elliptical copulas. The densities of both copulas are shown in figure 5.1 for the same correlation coefficient $c = 0.5$. We observe that the Student's t-copula assigns more probability to the tail events than the Gaussian copula. Increasing the degrees of freedom ν decreases the probability for extreme co-movements. In the limit $\nu \to \infty$, the Student's t-copula converges towards the Gaussian copula.

5.3 K-copula

The random matrix approach introduced in chapter 2 allowed us to derive a multivariate distribution for stock returns, which can be expressed in terms of a

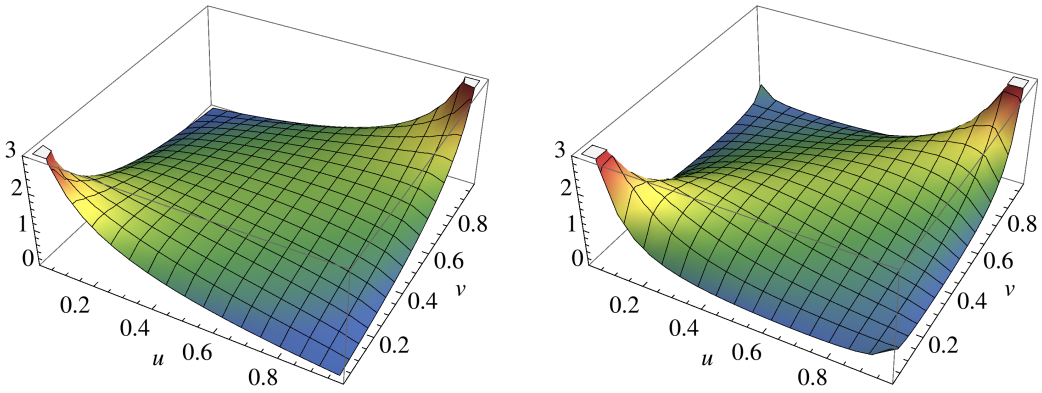


Figure 5.1: Left: Gaussian copula density with $c = 0.5$. Right: Student's t-copula density with $c = 0.5$ and $\nu = 3$.

modified Bessel function of the second kind. Distributions of this kind are often called K-distributions. The K-copula is the dependence structure which arises for the K-distribution (2.27). It was first used in reference [149], where it was found to describe the empirical dependencies in financial data much better than a Gaussian copula.

We now apply the Sklar's theorem to derive a bivariate K-copula from the K-distribution. For the bivariate case $K = 2$, the pdf of the vector $r = (r_1, r_2)$ reads

$$f_{c,N}(r_1, r_2) = \langle g \rangle(r|\Sigma, N) = \frac{1}{\Gamma(N/2)} \int_0^\infty \mathrm{d}z \, \frac{z^{N/2-1} \mathrm{e}^{-z}}{\sqrt{1-c^2}} \times \frac{N}{4\pi z} \exp\left(-\frac{N}{4z} \frac{r_1^2 - 2cr_1r_2 + r_2^2}{1-c^2} \right). \tag{5.15}$$

Here, we used the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 c \\ \sigma_1\sigma_2 c & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}, \tag{5.16}$$

where c denotes the average correlation coefficient estimated over the considered sample of returns. We chose the standard deviations one, $\sigma_1 = \sigma_2 = 1$, since the copula is independent of the marginal distributions. Then, the marginal density

functions are identical, $f_1(r_1) = f_2(r_2)$, where

$$f_1(r_1) = \int_{-\infty}^{\infty} \mathrm{d}r_2 \; f_{c,N}(r_1, r_2) = \frac{1}{\Gamma(N/2)} \int_0^{\infty} \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{N}{4\pi z}} \exp\left(-\frac{N}{4z} r_1^2\right) \; . \tag{5.17}$$

According to equation (5.6) the bivariate K-copula is given by

$$\mathrm{Cop}_{c,N}(u, v) = F_{c,N}(F_1^{-1}(u), F_2^{-1}(v)) \; , \tag{5.18}$$

where c and N are the parameters of the copula, F^{-1} denotes the inverse of the marginal cdf

$$\begin{aligned} F_1(r_1) &= \int_{-\infty}^{r_1} \mathrm{d}\xi \; f_1(\xi) = \int_{-\infty}^{r_1} \mathrm{d}\xi \frac{1}{\Gamma(N/2)} \int_0^{\infty} \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{N}{4\pi z}} \exp\left(-\frac{N}{4z}\xi^2\right) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \; \frac{\sqrt{N}\sqrt{N\xi^2}^{\frac{N-1}{2}}}{\sqrt{\pi}\; \Gamma(N/2)\sqrt{2}^{N-1}} \; \mathcal{K}_{\frac{1-N}{2}}\left(\sqrt{N\xi^2}\right) \; . \end{aligned} \tag{5.19}$$

and $F_{c,N}$ is the cumulative distribution function of the bivariate distribution (5.15)

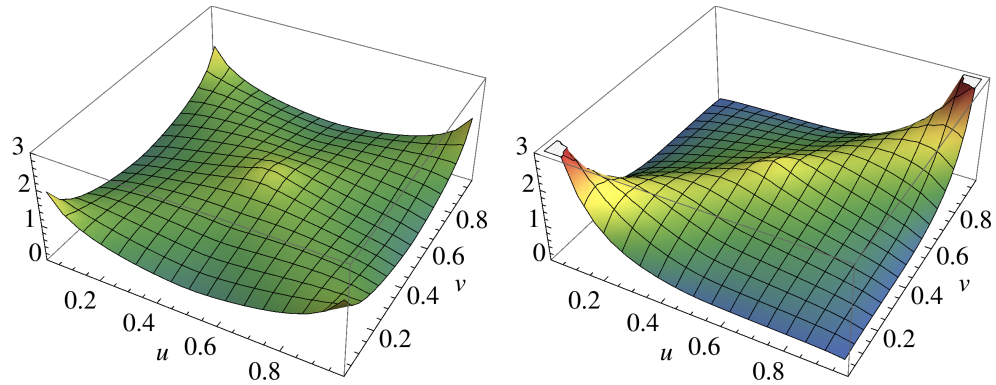
$$\begin{aligned} F_{c,N}(r_1, r_2) &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \; f_{c,N}(\xi, \zeta) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \int_0^{\infty} \frac{\mathrm{d}z}{\Gamma(N/2)} \frac{z^{N/2-1}\mathrm{e}^{-z}}{\sqrt{1-c^2}} \frac{N}{4\pi z} \exp\left(-\frac{N}{4z}\frac{\xi^2 - 2c\xi\zeta + \zeta^2}{1-c^2}\right) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \; \frac{N\sqrt{\frac{N(\xi^2-2c\xi\zeta+\zeta^2)}{1-c^2}}^{\frac{N-2}{2}}}{\pi\; \Gamma(N/2)\sqrt{2}^N\sqrt{1-c^2}} \; \mathcal{K}_{\frac{2-N}{2}}\left(\sqrt{\frac{N(\xi^2 - 2c\xi\zeta + \zeta^2)}{1-c^2}}\right) \; . \end{aligned} \tag{5.20}$$

The K-copula density can be obtained from the K-copula (5.18) by differentiation

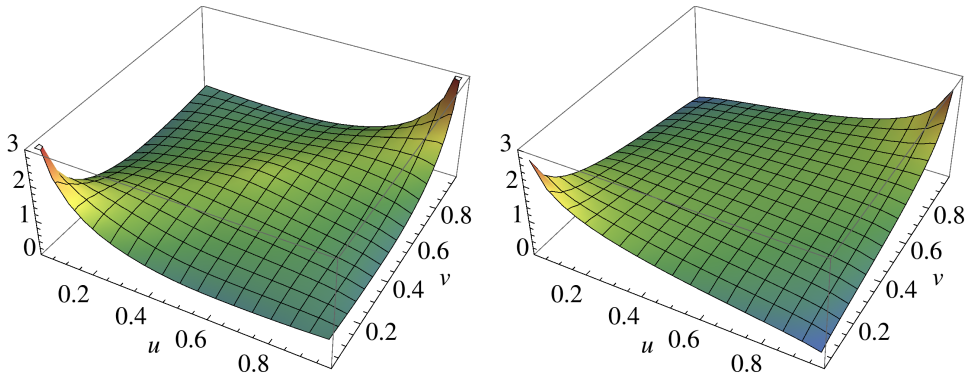
$$\mathrm{cop}_{c,N}(u, v) = \frac{\partial^2 \mathrm{Cop}_{c,N}(u, v)}{\partial u \partial v} = \frac{f_{c,N}(F_1^{-1}(u), F_2^{-1}(v))}{f_1(F_1^{-1}(u)) f_2(F_2^{-1}(v))} \; . \tag{5.21}$$

It depends only on the average correlation c and the free parameter N, which characterizes the fluctuations around Σ. Figure 5.2 shows the K-copula density for different parameter values. The stronger the average correlation c and the

lower the parameter N, the higher is the probability for extreme co-movements. As N increases, the probability for extreme co-movements decreases. In the limit $N \to \infty$, the K-copula converges towards the Gaussian copula.



(a) $c = 0, N = 3$ (b) $c = 0.5, N = 3$



(c) $c = 0.3, N = 4$ (d) $c = 0.3, N = 40$

Figure 5.2: K-copula densities $\text{cop}_{c,N}(u, v)$ for different parameter values.

Furthermore, we note that the K-copula is a symmetric copula. It is based on the elliptical distribution (2.27) and thus it also belongs to the class of elliptical copulas. As the Student's t-copula, the K-copula has an additional parameter to the average correlation c. For small N and ν values the K- and the Student's t-copula exhibit different behavior, in particular in the center and the corners of the copula density. As the parameters increase, both copulas become more and more similar, converging towards a Gaussian copula in the limit $\nu, N \to \infty$.

5.4 Empirical results

In this section, we will study the statistical dependence for each of the six market states identified in chapter 4. We present the empirical pairwise copula densities for each market state in section 5.4.1 and compare them with the K-copula densities in section 5.4.2. In section 5.4.3, we study the asymmetry of the tail dependence of the empirical copula densities in more detail.

5.4.1 Empirical pairwise copulas for each market state

To estimate the empirical pairwise copula of two return time series $r_k(t)$ and $r_l(t)$, we first have to transform them into uniformly distributed time series. To achieve this, we employ the empirical distribution function

$$u_k(t) = F_k(r_k(t)) = \frac{1}{T}\sum_{\tau=1}^{T} \mathbf{1}\{r_k(\tau) \leq r_k(t)\} - \frac{1}{2T} \ , \tag{5.22}$$

where $\mathbf{1}$ is the indicator function, T denotes the length of the time series and the factor $1/2$ ensures that the values of the transformed time series $u_k(t)$ lie in the interval $(0, 1)$. The empirical copula density of the time series $r_k(t)$ and $r_l(t)$ is then the two-dimensional histogram of the transformed time series $u = u_k(t)$ and $v = u_l(t)$.

An accurate estimation of the copula density requires a large amount of data. Thus, for each state we compute the copula densities of all $K(K-1)/2$ stock pairs as two-dimensional histograms of the transformed time series and then average over all pairs

$$\mathrm{cop}^{(i)}(u, v) = \frac{2}{K(K-1)} \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} \mathrm{cop}_{k,l}^{(i)}(u, v) \ , \qquad i = 1, \ldots, 6 \ , \tag{5.23}$$

where the upper index i denotes the state number. For the bin size of the histograms we choose $\Delta u = \Delta v = 0.05$. We note that as the length of the time series for each market state is rather short, we are not able to study the full copulas for each stock pair k, l separately. Hence, we cannot make any direct statements about similarity and dispersion regarding the full dependence structure. Our results only yield statements about the empirical dependence structure on average.

We estimate the empirical copula densities for both original (4.1) and locally normalized returns (4.3). As discussed in chapter 4, the original returns exhibit time-varying trends and volatilities. In contrast, the locally normalized returns show stationary behavior. It is important to note that the copulas for the original and the locally normalized returns contain different information. The copulas for

the original returns describe the dependence structure for the full time horizon, i.e., on a global scale. On the other hand, the copulas for the locally normalized returns provide information about the statistical dependence on a local scale.

Figure 5.3 shows the empirical pairwise copula densities for the original returns for each of the six market states. We observe a variation of the dependence structure from state to state, particularly visible in the tails. In state 1, which covers the period from 1992 to roughly 2000, we find a rather flat copula density, indicating low dependence between return pairs. In state 2 we observe deviations from the flat copula density particularly in the tails, which become more and more pronounced in state 3 and 4. State 5, first appearing during the financial crisis in 2008, exhibits the strongest dependence. The dependence decreases again in state 6.

Figure 5.4 shows the empirical pairwise copula densities for the locally normalized returns. The dependence structures of the six states are mostly preserved after performing local normalization. Deviations are observed in the lower-left and the upper-right corners where the copula densities for the original returns exhibit higher peaks. This can be explained as follows: Events in the corners reflect periods with high volatility. In periods with high volatility large returns are more common than in calmer periods, see figures 1.3 and 1.5. Large negative returns obtain ranks closer to zero and contribute to the lower corners of the copula densities, whereas large positive returns obtain ranks closer to one and thus contribute to the upper corners. In addition, in periods with high volatility the correlations between stocks increase. This leads to a stronger dependence and thus to higher peaks in the corners of the copula densities for the original returns.

Furthermore, we observe that the empirical copula densities are asymmetric with respect to opposite corners. We find a stronger dependence in the lower tail than in the upper one, that is, the dependence between large negative returns is stronger than the dependence between large positive ones. This asymmetry is an important feature of empirical copula densities and thus we will discuss it in more details in section 5.4.3.

5.4.2 Comparison with the K-copula

In the following, we compare the empirical pairwise copula densities for each market state with the K-copula density introduced in section 5.3. The K-copula density is obtained in the following way: We calculate the K-copula according to equation (5.18), where the integrals are performed numerically. The K-copula density for each bin of size $\Delta u = \Delta v = 0.05$ is estimated as follows

$$\begin{aligned}\mathrm{cop}_{\bar{c},N}(u,v) = \mathrm{Cop}_{\bar{c},N}(u,v) - \mathrm{Cop}_{\bar{c},N}(u,v-\Delta v) - \mathrm{Cop}_{\bar{c},N}(u-\Delta u,v) \\ + \mathrm{Cop}_{\bar{c},N}(u-\Delta u,v-\Delta v)\ .\end{aligned} \tag{5.24}$$

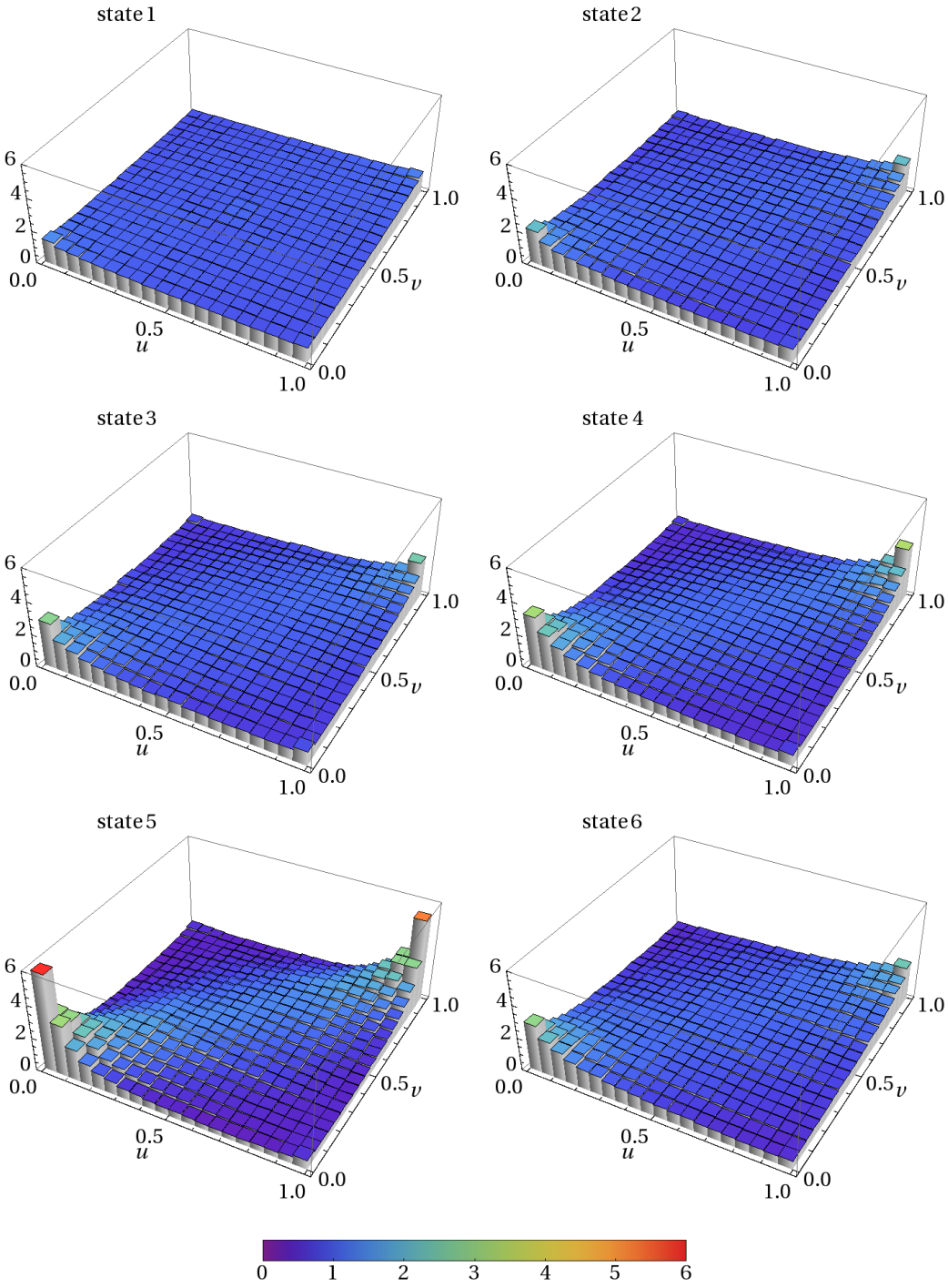


Figure 5.3: Empirical pairwise copula density $\text{cop}^{(i)}(u, v)$ for the original returns.

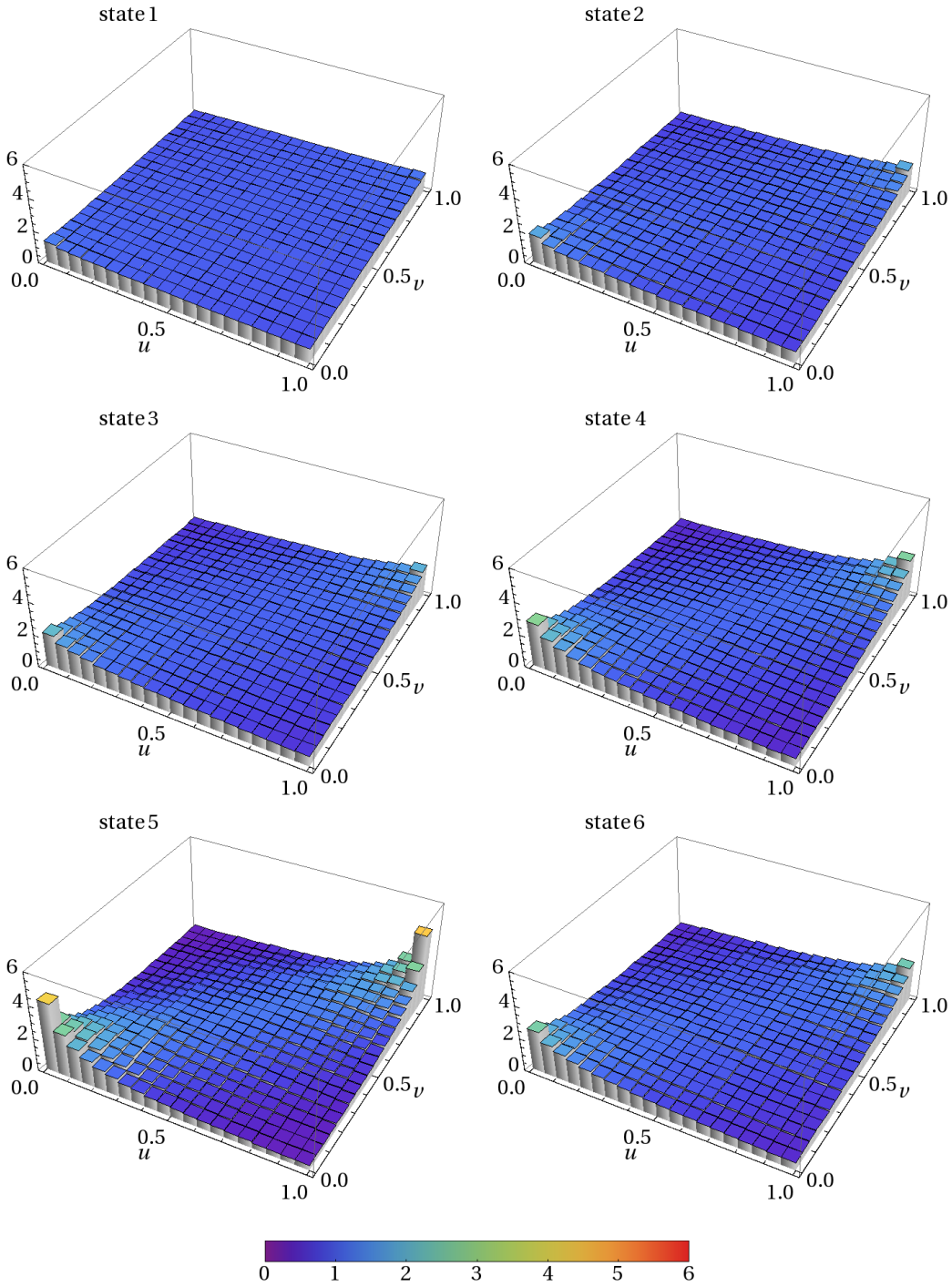


Figure 5.4: Empirical pairwise copula density $\mathrm{cop}^{(i)}(u, v)$ for the locally normalized returns.

returns		state 1	state 2	state 3	state 4	state 5	state 6
original	$\bar{c}$	0.046	0.13	0.17	0.25	0.42	0.22
	N	41.7	11.7	8.4	5.6	2.8	10.0
loc. normalized	$\bar{c}$	0.048	0.13	0.19	0.28	0.43	0.25
	N	70.7	28.6	29.8	15.8	7.4	20.4

Table 5.1: Parameters of the K-copula density $\mathrm{cop}^{(i)}_{\bar{c},N}(u,v)$ for the original and the locally normalized returns.

returns	state 1	state 2	state 3	state 4	state 5	state 6
original	0.11	0.41	1.22	2.47	5.53	2.84
loc. normalized	0.047	0.38	0.70	1.43	2.94	1.50

Table 5.2: Least sum of squared differences between empirical and K-copula densities.

For the comparison, we compute the difference between empirical and analytical copula density for each state

$$\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\bar{c},N}(u,v) \ , \qquad i = 1,\ldots,6 \ , \tag{5.25}$$

where $\bar{c}$ is the average correlation coefficient of all $K(K-1)/2$ stock pairs for the considered state. The free parameter N is determined by a fit which minimizes the sum of squared differences between empirical and analytical copula density. The parameter values for each market state are summarized in table 5.1. The differences between the empirical copula density and the K-copula density for each state are presented in figure 5.5 for the original and in figure 5.6 for the locally normalized returns. Overall, we find a good agreement between empirical and analytical copula densities for the original returns. The K-copula seems to capture the dependence structure of the first three states very well. Small but statistically significant deviations from the K-copula density are observed for state 4 and 6. Only the dependence structure of state 5 cannot be captured by the K-copula. For the locally normalized returns we find a better agreement, which is reflected in the smaller least squares, see table 5.2. Deviations are observed mainly in the corners of the copula densities.

It is important to note that the K-copula captures the overall empirical dependence structure rather well with only one free parameter. In reference [149] it was shown that the K-copula describes the empirical dependencies in financial data

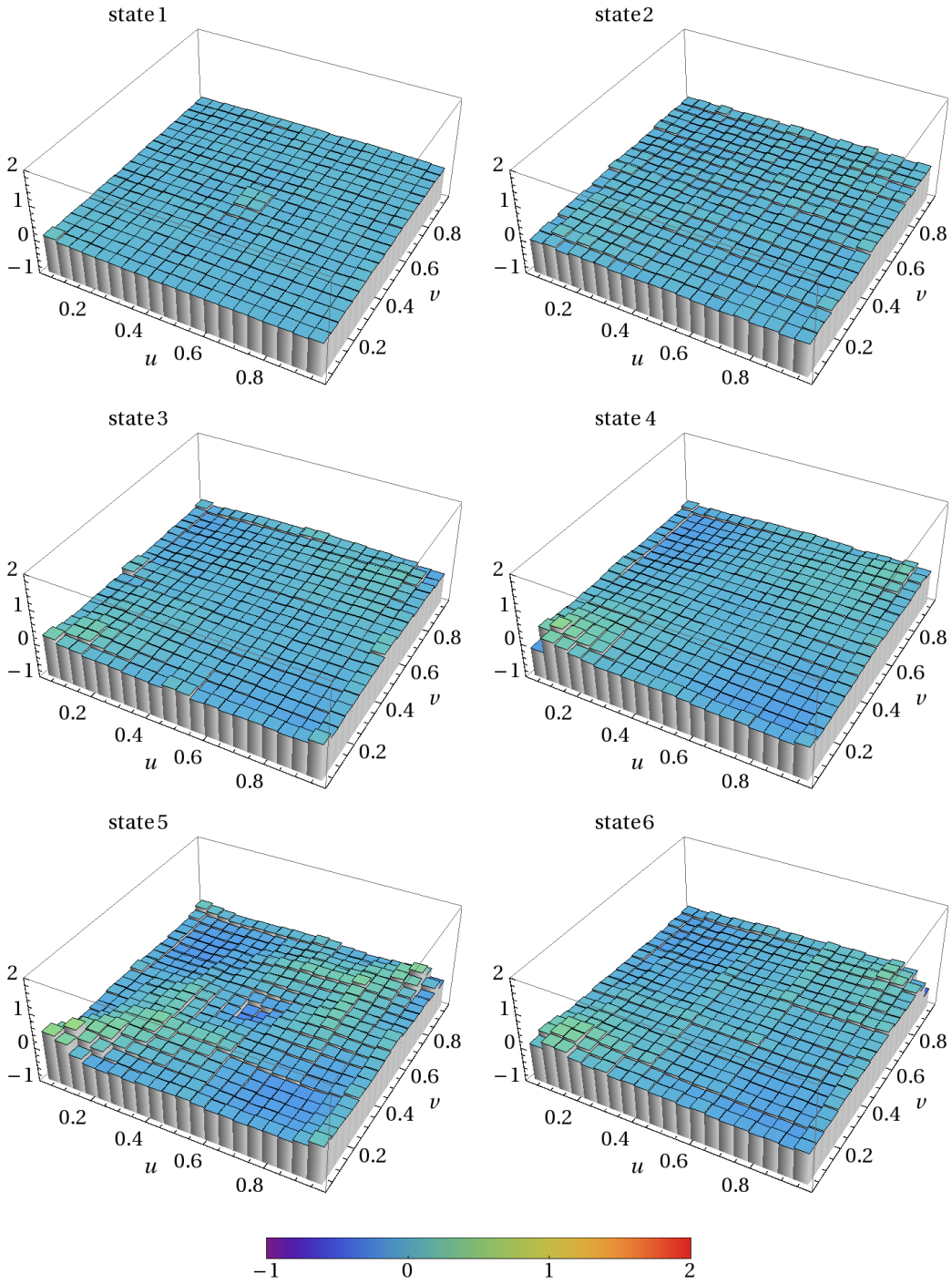


Figure 5.5: Difference between the empirical copula density and the K-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\bar{c},N}(u,v)$ for the original returns.

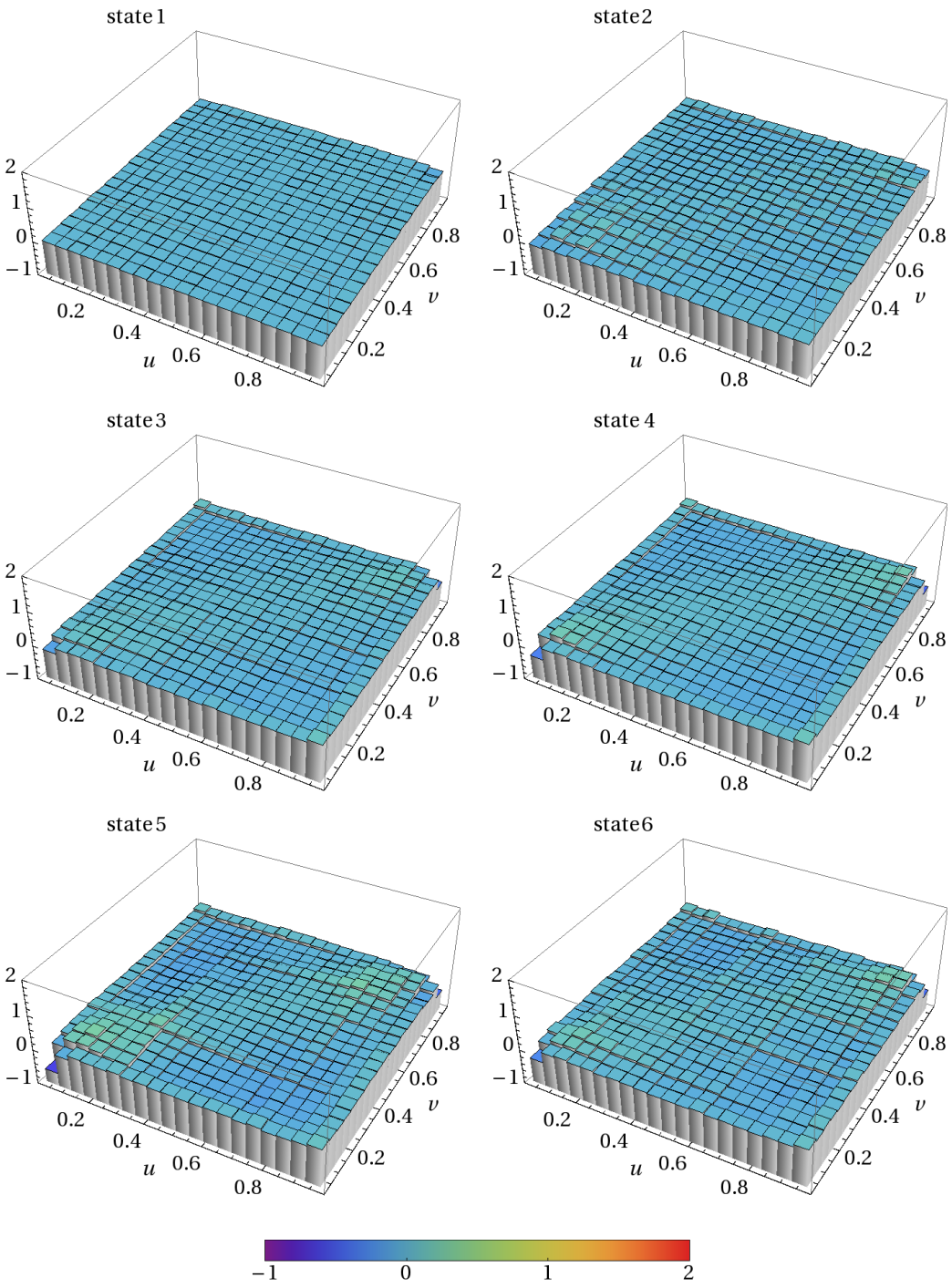


Figure 5.6: Difference between the empirical copula density and the K-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\tilde{c},N}(u,v)$ for the locally normalized returns.

much better than the Gaussian copula. It also provides a slightly better fit to empirical returns than the Student's t-copula as shown in appendix C.

Nevertheless, neither the K-copula nor the Student's t-copula can account for the asymmetry observed in the empirical data due to their symmetric nature. The skewed Student's t-copula is an alternative proposed by Demarta *et al.* [150], which was found to account for asymmetric dependencies in financial data [151, 152]. It captures the empirical dependence structure of the original returns better than the K-copula due to the presence of an additional parameter which accounts for the asymmetry [149]. We discuss the asymmetry in more detail in the following.

5.4.3 Asymmetry of the tail dependence

Asymmetric dependence between returns has been reported by several authors, see e.g., references [143–145]. Our study provides further evidence revealing a stronger lower tail dependence in the empirical copula densities of market states.

We now take a closer look at this asymmetry. To this end, we estimate the tail dependence in the four corners for all $K(K-1)/2$ empirical pairwise copulas according to

$$\begin{aligned}
\mathrm{LL}^{(i)}_{k,l} &= \int_0^{0.2} \mathrm{d}u \int_0^{0.2} \mathrm{d}v \; \mathrm{cop}^{(i)}_{k,l}(u,v) \;, \\
\mathrm{UL}^{(i)}_{k,l} &= \int_{0.8}^{1} \mathrm{d}u \int_0^{0.2} \mathrm{d}v \; \mathrm{cop}^{(i)}_{k,l}(u,v) \;, \\
\mathrm{UU}^{(i)}_{k,l} &= \int_{0.8}^{1} \mathrm{d}u \int_{0.8}^{1} \mathrm{d}v \; \mathrm{cop}^{(i)}_{k,l}(u,v) \;, \\
\mathrm{LU}^{(i)}_{k,l} &= \int_0^{0.2} \mathrm{d}u \int_{0.8}^{1} \mathrm{d}v \; \mathrm{cop}^{(i)}_{k,l}(u,v) \;, \qquad i = 1, \ldots, 6 \;,
\end{aligned} \tag{5.26}$$

where the upper index $i = 1, \ldots, 6$ denotes the state number and the lower indices represent a stock pair k, l. Here, LL and UU refer to the lower-lower and upper-upper corners, respectively, and represent the positive tail dependence, whereas UL and LU refer to the upper-lower and lower-upper corners, respectively, and represent the negative tail dependence. The asymmetry in the tail dependence can

now be quantified by the differences

$$\begin{aligned} \alpha_{k,l}^{(i)} &= \mathrm{UU}_{k,l}^{(i)} - \mathrm{LL}_{k,l}^{(i)} \ , \\ \beta_{k,l}^{(i)} &= \mathrm{LU}_{k,l}^{(i)} - \mathrm{UL}_{k,l}^{(i)} \ , \qquad i = 1, \ldots, 6 \ , \end{aligned} \tag{5.27}$$

where $\alpha_{k,l}^{(i)}$ captures the asymmetry of the positive and $\beta_{k,l}^{(i)}$ the asymmetry of the negative tail dependence for each stock pair k,l. Figure 5.7 shows the histograms of the asymmetry values exemplarily for state 5. For the returns we find a negative offset for the values of $\alpha_{k,l}^{(5)}$, whereas the values of $\beta_{k,l}^{(5)}$ are centered around zero. This indicates, on average, an asymmetry in the positive tail dependence, i.e., simultaneous large negative returns are more likely to occur than simultaneous large positive returns. On the other hand, we do not find such an asymmetry in the negative tail dependence. For the locally normalized returns we find a much weaker asymmetry in the positive tail dependence than for the original returns and once again no asymmetry in the negative tail dependence. Whereas for the original returns the tail dependence reflects periods with high volatility, for the stationary locally normalized returns all periods contribute equally to the tail dependence.

We note that the means of the asymmetry values are the relevant quantities, the asymmetry values for each pair are distributed around the mean due to statistical fluctuations. The standard deviation for $\alpha_{k,l}^{(5)}$ is 0.01 and for $\beta_{k,l}^{(5)}$ 0.007. Indeed, the asymmetry effect is very small. Still, it is clearly visible, see figure 5.3.

In the following, we study the asymmetry values for each market state. Figure 5.8 shows the mean asymmetry values $\bar{\alpha}^{(i)}$ and $\bar{\beta}^{(i)}$ for each market state, obtained by averaging over all $\alpha_{k,l}^{(i)}$ and $\beta_{k,l}^{(i)}$ for the given state. On a local scale the asymmetry in the positive tail dependence is much weaker. Only state 5 still exhibits a certain amount of asymmetry. On the other hand, the asymmetry in the negative tail dependence is negligibly small for both original and locally normalized returns.

5.5 Summary

We studied the dependence structure of market states by means of a copula approach. To this end, we estimated empirical pairwise copulas for each state. We used both original and locally normalized returns. The local normalization removes the time-varying trends and volatilities of the original time series while preserving the dependence structure. The copula of the locally normalized returns describes the dependence structure on a local scale, whereas the copula of the original returns provides information about the dependence structure on a global scale, i.e., for the full time horizon.

We compared the empirical pairwise copulas for each state with a bivariate K-

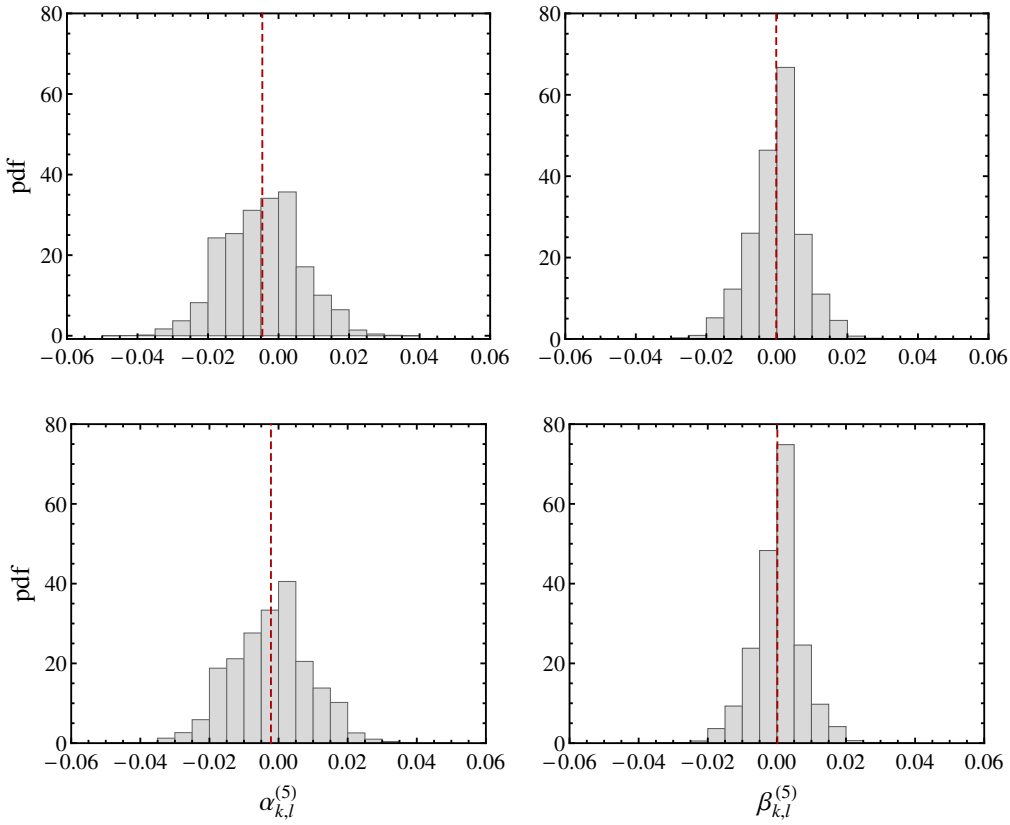


Figure 5.7: Histograms of the asymmetry values for all stock pairs k,l, exemplarily for state 5. Left: asymmetry for the positive tail dependence $\alpha^{(5)}_{k,l}$, right: asymmetry for the negative tail dependence $\beta^{(5)}_{k,l}$. Top: for returns, bottom: for locally normalized returns. The dashed red lines represent the corresponding mean values.

copula. The bivariate K-copula results from the random matrix model introduced in chapter 2 to model non-stationary correlations. It is a symmetric, elliptical copula, which depends on two parameters: the average correlation coefficient, estimated over the considered sample of returns, and a free parameter which characterizes the fluctuations around the average correlation in this sample. Overall, the K-copula captures the empirical dependence structure of market states. We found a good agreement, in particular for the empirical copulas estimated on a local scale. Thus, we obtain a consistent description within the random matrix model: The K-distribution describes the empirical multivariate return distributions and the

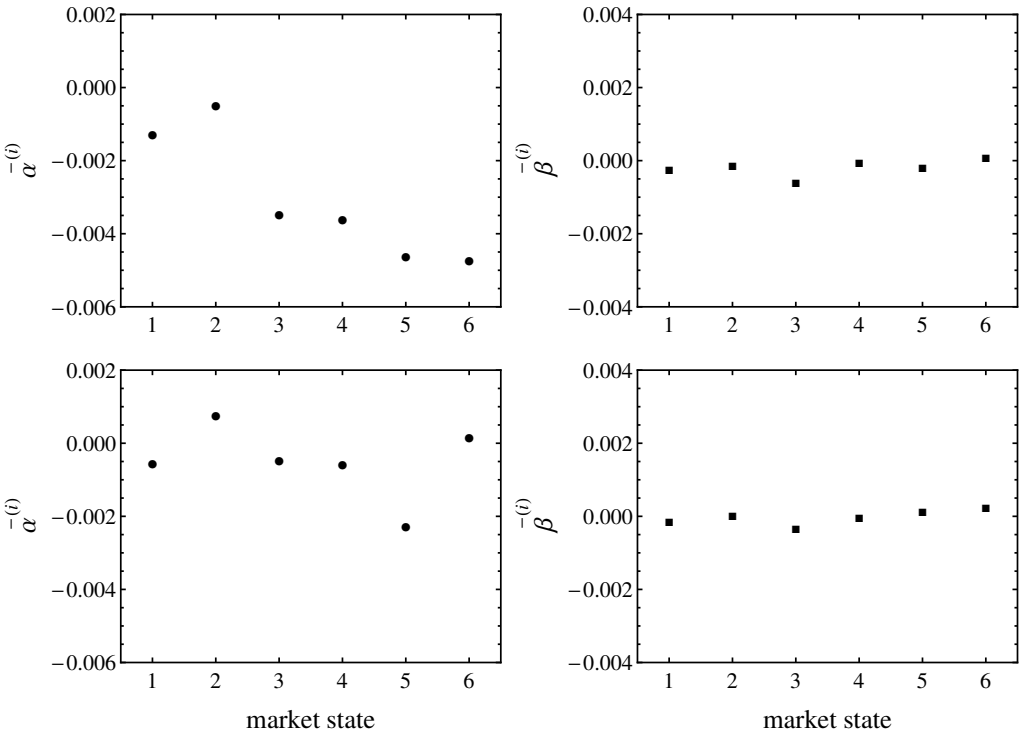


Figure 5.8: The mean asymmetry values for each market state. Left: asymmetry for the positive tail dependence $\bar{\alpha}^{(i)}$, right: asymmetry for the negative tail dependence $\bar{\beta}^{(i)}$. Top: for returns, bottom: for locally normalized returns.

K-copula captures the empirical dependence structure.

Furthermore, we found an asymmetry in the positive tail dependence, i.e., a stronger lower tail dependence, indicating a larger probability for simultaneous extreme negative returns. This asymmetry cannot be captured by our model. It is more pronounced on a global scale, as the tail dependence reflects periods with high volatility. On a local scale we find a much weaker asymmetry. However, in times of crisis the asymmetry is still clearly present.

Chapter 6

Extreme value statistics

6.1 Introduction

Extreme value statistics deals with the statistical theory of extreme events, i.e., events which occur with very low probability. Although quite rare, such events are of great importance since they may have catastrophic consequences. Examples range from natural disasters like floods, hurricanes, earthquakes to stock market crashes or large insurance losses in finance.

Classical extreme value statistics is concerned with the maximum or the minimum of a set of independent and identically distributed random variables. Its origin goes back to the 1920s, when a number of scientists recognized the need for a general statistical theory of extreme values. In 1927, Maurice Fréchet published the first results on the asymptotic distributions of largest values [153]. A year later the same problem was studied independently by Ronald Fisher and Leonard Tippett. While Fréchet found only one possible limit distribution, Fisher and Tippett showed that the limit distributions can only be one of three types [154]. In 1943, Boris Gnedenko presented a rigorous foundation for the theory of extreme values by providing necessary and sufficient conditions for the Fisher and Tippett theorem [155]. The theoretical development was followed by a number of papers dealing with practical applications. In particular, the German mathematician Emil Gumbel made many significant contributions focusing mainly on meteorological phenomena [156]. During the late 1990s the focus was shifted towards insurance and finance as applications in financial risk started to occur [123, 157].

In many applications, however, data are correlated. Dependence causes features which were not encountered in the classical extreme value statistics. In recent years, many efforts have thus been made to develop a theory for correlated random variables. For weakly correlated random variables, it has been shown that the

problem basically reduces to the case of uncorrelated random variables [158]. On the other hand, there still does not exist a general theory for strongly correlated random variables. Most of the theoretical efforts are focused in finding exactly solvable cases which may shed some light on the issue of the universality classes of the extreme value theory for strongly correlated random variables [158–160]. Identifying such universality classes, if they do exist, is a challenging but still open problem [161].

Here, we present some preliminary work on a new project on extreme value statistics, aiming at a contribution to the theory of correlated random variables. In particular, we address the question: What is the distribution of the maximum of a correlated random sample? To answer this question, we begin with a sample drawn from a multivariate normal distribution and derive the distribution of the maximum assuming equally correlated random variables. We then use this result to obtain a maximum distribution for the non-normal case, and in particular for a sample of correlated t-distributed random variables. The validity of these results is confirmed in numerical simulations.

The chapter is organized as follows: In section 6.2, we begin with a short review of the extreme value statistics of independent and identically distributed random variables and study the convergence rate of the sample maximum towards the limit distributions for a finite sample size. To this end, we perform a simulation study considering some of the most common distributions which arise in statistical applications. In section 6.3, we derive the maximum distributions for correlated normal and non-normal random samples and compare our results with numerical simulations and financial data. We conclude our findings in section 6.4. The contents of this chapter are based on unpublished notes [162] and private communication with Thomas Guhr.

6.2 Extreme value statistics of uncorrelated random variables

We begin with a short introduction into the theory of extreme values following mostly the textbook of Coles [163]. For a more detailed introduction the reader is referred to the textbooks [164–166].

6.2.1 Basic concepts

Let $x_1, \ldots, x_N$ be a sample of independent and identically distributed (iid) random variables with common distribution function

$$F(x) = \Pr[x_n \leq x] \ . \tag{6.1}$$

The maximum of the sample is defined as the largest value

$$m_N = \max\{x_1, \ldots, x_N\} \ . \tag{6.2}$$

For instance, x_n may represent daily losses or returns, so that m_N represents the maximum loss or return of the considered sample. Here, we focus on the theory of the sample maximum. Corresponding results for the minimum can easily be obtained from those for the maximum by using the identity

$$\min\{x_1, \ldots, x_N\} = -\max\{-x_1, \ldots, -x_N\} \ . \tag{6.3}$$

The distribution of m_N can be derived exactly for all values of N as follows

$$\begin{aligned} Q_N(x) = \Pr[m_N \leq x] &= \Pr[x_1 \leq x, \ldots, x_N \leq x] \\ &= \prod_{n=1}^{N} \Pr[x_n \leq x] = F^N(x) \ . \end{aligned} \tag{6.4}$$

In the last steps, we used the fact that the random variables are independent and identically distributed, which allows us to write the distribution of the maximum as a product of the initial distribution F. The corresponding density function of the maximum is obtained by taking the derivative

$$q_N(x) = \frac{\mathrm{d}}{\mathrm{d}x} Q_N(x) = \frac{\mathrm{d}}{\mathrm{d}x} F^N(x) = N F^{N-1}(x) f(x) \ , \tag{6.5}$$

where $f(x) = dF(x)/dx$ is the initial probability density function.

In practice, the initial distribution F is often unknown, making the calculation of Q_N impossible. Of course, one could estimate F from measurements and use this estimate to calculate the distribution of m_N according to equation (6.4). The disadvantage of this approach is that even small discrepancies in F could result in large discrepancies for Q_N.

An alternative approach is to look at the behavior of Q_N as N goes to infinity. In this case, the distribution of m_N degenerates to a point mass at x_0, i.e., $Q_N(x) \to 0$ if $x < x_0$ and $Q_N(x) \to 1$ if $x \geq x_0$. This problem can be avoided by a renormalization of the maximum

$$m_N^* = \frac{m_N - b_N}{a_N} \tag{6.6}$$

for sequences of constants $a_N > 0$ and $b_N \in \mathbb{R}$. Appropriate choices of a_N and b_N stabilize the scale and location of m_N^* as N increases, avoiding the difficulties that arise with the variable m_N. Therefore, we seek limit distributions for the normalized maximum m_N^* rather than m_N. The determination of the range of all possible limit distributions for m_N^* is the central result in extreme value theory,

derived by Fisher and Tippett [154] in 1928 and rigorously proved by Gnedenko [155] in 1943.

Theorem 6.1 (Fisher-Tippett-Gnedenko Theorem) *If there exist normalizing constants $a_N > 0$ and $b_N \in \mathbb{R}$ and some non-degenerate distribution function G such that*

$$Pr\left(\frac{m_N - b_N}{a_N} \leq x\right) = F^N(b_N + a_N x) \xrightarrow[N\to\infty]{} G(x) \ , \tag{6.7}$$

then the distribution function G belongs to one of the following three distribution types:

- *Type I: Gumbel distribution*

$$\Lambda(x) = \exp\left(-\mathrm{e}^{-x}\right), \ x \in \mathbb{R} \ , \tag{6.8}$$

- *Type II: Fréchet distribution*

$$\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp\left(-x^{-\alpha}\right), & x > 0 \end{cases}, \ \alpha > 0 \ , \tag{6.9}$$

- *Type III: Weibull distribution*

$$\Psi_\alpha(x) = \begin{cases} \exp\left(-(-x)^{\alpha}\right), & x < 0 \\ 1, & x \geq 0 \end{cases}, \ \alpha > 0 \ . \tag{6.10}$$

If the theorem holds for suitable choices of a_N and b_N, then we say that G is an extreme value distribution and F is in the domain of attraction of G.

The theorem implies that if the maximum m_N can be stabilized with appropriate normalizing constants a_N and b_N, the normalized maximum m_N^* has a limiting distribution which must be one of the three distribution types. The remarkable feature of this result is that the three types of extreme value distributions are the only possible limits for the distributions of the normalized maximum regardless of the initial distribution of the random variables.

Motivated by statistical applications, von Mises and Jenkinson introduced an unifying representation of the three extreme value distribution types [167, 168]

$$G_\xi(x) = \begin{cases} \exp\left(-(1+\xi x)^{-1/\xi}\right) & \text{if } \ \xi \neq 0 \\ \exp\left(-\mathrm{e}^{-x}\right) & \text{if } \ \xi = 0 \end{cases} \tag{6.11}$$

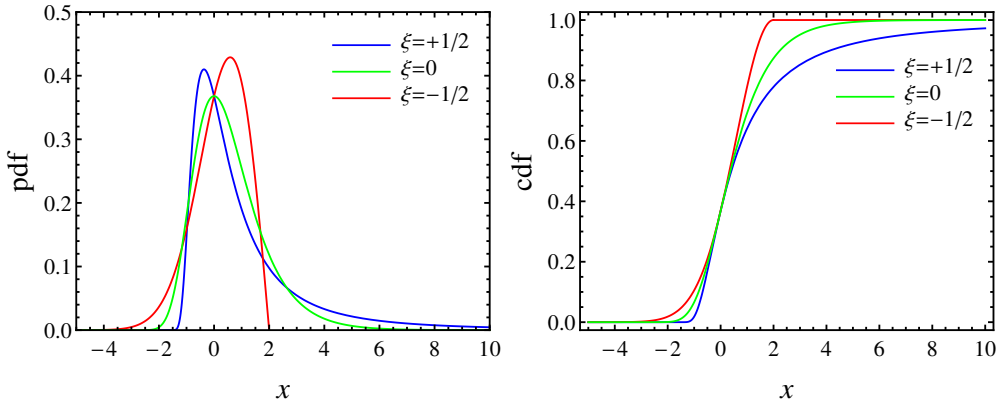


Figure 6.1: Probability density function (left) and cumulative distribution function (right) of the GEV distribution for different ξ values.

for $1 + \xi x > 0$. This is the generalized extreme value (GEV) distribution. The parameter ξ is a shape parameter, also called tail parameter. It governs the tail behavior and defines the type of the distribution: for $\xi > 0$ the GEV distribution is a Fréchet distribution with $\alpha = 1/\xi$, while for $\xi < 0$ it is a Weibull distribution with $\alpha = -1/\xi$. The case $\xi = 0$ is obtained in the limit $\xi \to 0$ and represents a Gumbel distribution. The probability density and the cumulative distribution function of the GEV distribution are shown in figure 6.1 for the three cases $\xi = 1/2$, $\xi = 0$ and $\xi = -1/2$, corresponding to the Fréchet, Gumbel and Weibull distribution types, respectively. The three distribution types have distinct forms. We observe that the Weibull distribution has a finite upper endpoint, while the Gumbel and the Fréchet distributions have infinite upper endpoints. However, the density decays exponentially for the Gumbel distribution and polynomially for the Fréchet distribution.

Since the extreme values of a random variable are associated with the tails of the distribution, the type of the limit distribution is determined by the tail behavior of the initial distribution F. Necessary and sufficient conditions for a distribution F to belong to the domain of attraction of one of the three limit distribution types can be found in references [157, 165, 166]. Here, we discuss some examples. Distributions whose tails decay exponentially give rise to a Gumbel distribution for the maximum. The Gumbel class contains a great variety of distributions ranging from light-tailed to moderate heavy-tailed distributions. It includes, for instance, the exponential, normal, log-normal, logistic, gamma distributions as well as the hyperbolic and the generalized hyperbolic distributions (with exception of the Student's t-distribution) which arise in models for financial returns [123]. Distributions whose tails decay like a power function lead to a Fréchet distribution for the maximum. These

distributions are particularly important for financial applications as they exhibit heavy tails. The Fréchet class includes, for instance, the Cauchy, Student's t, Pareto, inverse gamma and log-gamma distributions. Light-tailed distributions with finite upper endpoint lead to a Weibull distribution for the maximum. The Weibull class includes, for instance, the uniform and beta distributions.

6.2.2 Normalizing constants

The Fisher-Tippett-Gnedenko theorem requires the existence of normalizing constants a_N and b_N which depend on the sample size N. How can we determine these normalizing constants for an initial distribution F? In the following, we calculate some examples to demonstrate the construction of normalizing constants. Note that the choice of a_N and b_N is not unique. The interested reader is referred to references [157, 165, 166] for further details on the determination of the normalizing constants.

Example 1: Normal distribution. Consider iid random variables drawn from a normal distribution with the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) . \tag{6.12}$$

The corresponding distribution function is

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/\sigma} \mathrm{e}^{-y^2/2} dy = \Phi\left(\frac{x}{\sigma}\right) . \tag{6.13}$$

According to equation (6.4), the maximum of a sample of size N has the distribution

$$Q_N(x) = F^N(x) = \Phi^N\left(\frac{x}{\sigma}\right) . \tag{6.14}$$

To study the limit $N \to \infty$, we have to find sequences of constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that

$$Q_N(b_N + a_N x) = \Phi^N\left(\frac{b_N + a_N x}{\sigma}\right) = \left(1 - \left(1 - \Phi\left(\frac{b_N + a_N x}{\sigma}\right)\right)\right)^N \tag{6.15}$$

converges. The expression converges if

$$1 - \Phi\left(\frac{b_N + a_N x}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{\frac{(b_N + a_N x)}{\sigma}}^{\infty} \mathrm{e}^{-y^2/2} dy = \frac{g(x)}{N} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{6.16}$$

with some arbitrary function $g(x)$, which does not depend on N. The limit is then

$$\lim_{N\to\infty}\left(1-\frac{g(x)}{N}\right)^N = \exp\left(-g(x)\right) \ . \tag{6.17}$$

At this point, it is convenient to use an approximation of the integral in (6.16) for large x, namely

$$\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{(b_N+a_Nx)}{\sigma}}^{\infty} \mathrm{e}^{-y^2/2}dy \ \approx \ \frac{\sigma}{\sqrt{2\pi}(b_N+a_Nx)}\exp\left(-\frac{(b_N+a_Nx)^2}{2\,\sigma^2}\right) \ . \tag{6.18}$$

Using this approximation, our problem simplifies to solving

$$\frac{g(x)}{N} = \frac{\sigma}{\sqrt{2\pi}(b_N+a_Nx)}\exp\left(-\frac{(b_N+a_Nx)^2}{2\,\sigma^2}\right) \ . \tag{6.19}$$

We can find b_N by setting $x=0$. Thus, we obtain

$$\frac{\sigma}{\sqrt{2\pi}b_N}\exp\left(-\frac{b_N^2}{2\,\sigma^2}\right) = \frac{g(0)}{N} \ . \tag{6.20}$$

Making the ansatz $b_N/\sigma = \sqrt{2\log N}+c_N$ and choosing $g(0)=1$ leads to

$$c_N = -\frac{\log\,\log\,N+\log\,4\pi}{2\sqrt{2\log\,N}} \ . \tag{6.21}$$

With $a_N/\sigma = 1/\sqrt{2\log\,N}$, we find $g(x)=\mathrm{e}^{-x}$, so that the distribution of the normalized maximum converges towards a Gumbel distribution as N goes to infinity

$$Q_N(b_N+a_Nx) = \Phi^N\left(\frac{b_N+a_Nx}{\sigma}\right) = \left(1-\frac{\mathrm{e}^{-x}}{N}\right)^N \xrightarrow[N\to\infty]{} \exp\left(-\mathrm{e}^{-x}\right) \ . \tag{6.22}$$

Example 2: Exponential distribution. Consider iid random variables drawn from an exponential distribution with the probability density function

$$f(x) = \lambda\mathrm{e}^{-\lambda x} \ , \ x\geq 0 \ . \tag{6.23}$$

The corresponding distribution function is

$$F(x) = 1-\mathrm{e}^{-\lambda x} \ . \tag{6.24}$$

Again, our task is to find constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that

$$Q_N(b_N + a_N x) = F^N(b_N + a_N x) = \big(1 - \exp\left(-\lambda(b_N + a_N x)\right)\big)^N \tag{6.25}$$

converges for $N \to \infty$. Thus, we have to solve

$$\exp\left(-\lambda(b_N + a_N x)\right) = g(x)/N \ . \tag{6.26}$$

Setting $x = 0$, we find $b_N = \log N/\lambda$. With $a_N = 1/\lambda$, the distribution of the normalized maximum converges towards a Gumbel distribution as N goes to infinity

$$Q_N(b_N + a_N x) = \big(1 - \exp\left(-\lambda(b_N + a_N x)\right)\big)^N = \left(1 - \frac{\mathrm{e}^{-x}}{N}\right)^N \xrightarrow[N\to\infty]{} \exp\left(-\mathrm{e}^{-x}\right) \ . \tag{6.27}$$

Example 3: Cauchy distribution. Consider iid random variables drawn from a Cauchy distribution with the probability density function

$$f(x) = \frac{\kappa}{\pi}\frac{1}{\kappa^2 + x^2}, \ \kappa > 0, \ x \in \mathbb{R} \ . \tag{6.28}$$

The corresponding distribution function is

$$F(x) = \frac{1}{2} - \frac{1}{\pi}\arctan\frac{x}{\kappa} \ . \tag{6.29}$$

For large x it is convenient to use the series expansion

$$\arctan\, x = \frac{\pi}{2} - \frac{1}{x} + \mathcal{O}\left(\frac{1}{x^3}\right) \ . \tag{6.30}$$

Again, we want to find constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that

$$Q_N(b_N + a_N x) = F^N(b_N + a_N x) = \left(1 - \frac{\kappa}{\pi}\frac{1}{b_N + a_N x}\right)^N \tag{6.31}$$

converges for $N \to \infty$. Setting $x = 0$, we find $b_N = N\kappa/\pi$. With $a_N = N\kappa/\pi$, the distribution of the normalized maximum converges towards a Fréchet distribution as N goes to infinity

$$Q_N(b_N + a_N x) = \left(1 - \frac{(1+x)^{-1}}{N}\right)^N \xrightarrow[N\to\infty]{} \exp\left(-(1+x)^{-1}\right) \ . \tag{6.32}$$

Example 4: Pareto distribution. Consider iid random variables drawn from a Pareto distribution with the probability density function

$$f(x) = \frac{\alpha\ c^\alpha}{x^{1+\alpha}},\ x \geq c,\ c > 0,\ \alpha = 1/\xi\ . \tag{6.33}$$

The corresponding distribution function is

$$F(x) = 1 - \left(\frac{c}{x}\right)^{1/\xi}\ . \tag{6.34}$$

We have to find constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that

$$Q_N(b_N + a_N x) = F^N(b_N + a_N x) = \left(1 - \frac{c^{1/\xi}}{(b_N + a_N x)^{1/\xi}}\right)^N \tag{6.35}$$

converges for $N \to \infty$. Setting $x = 0$, we find $b_N = cN^\xi$. With $a_N = c\,\xi N^\xi$, the distribution of the normalized maximum converges towards a Fréchet distribution as N goes to infinity

$$Q_N(b_N + a_N x) = \left(1 - \frac{(1+\xi x)^{-1/\xi}}{N}\right)^N \xrightarrow[N\to\infty]{} \exp\left(-(1+\xi x)^{-1/\xi}\right)\ . \tag{6.36}$$

Example 5: Uniform distribution. Consider iid random variables drawn from a uniform distribution with the probability density function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}\ . \tag{6.37}$$

The corresponding distribution function is

$$F(x) = 1 - \left(1 - \frac{x-a}{b-a}\right)\ . \tag{6.38}$$

We have to find constants $a_N > 0$ and $b_N \in \mathbb{R}$ such that

$$Q_N(b_N + a_N x) = F^N(b_N + a_N x) = \left(1 - \left(1 - \frac{b_N + a_N x - a}{b-a}\right)\right)^N \tag{6.39}$$

converges for $N \to \infty$. Setting $x = 0$, we find $b_N = \left(b(N-1) + a\right)/N$. With $a_N = (b-a)/N$, the distribution of the normalized maximum converges towards a

Domain of attraction	Initial distribution	Parameter
Gumbel	Normal	$(\mu, \sigma) = (0, 1)$
	Log-normal	$(\mu, \sigma) = (0, 1)$
	Exponential	$\lambda = 2$
Fréchet	Cauchy	$\kappa = 2$
	Pareto	$c = 1, \alpha = 1, 2, 3$
Weibull	Uniform	$(a, b) = (0, 1)$

Table 6.1: Distributions used in the simulation study on the convergence rate of the normalized maximum.

Weibull distribution as N goes to infinity

$$Q_N(b_N + a_N x) = \left(1 - \frac{1-x}{N}\right)^N \xrightarrow[N\to\infty]{} \exp\left(-(1-x)\right) \ . \tag{6.40}$$

We note that it may not always be possible to find appropriate normalizing constants and thus a limiting distribution for the maximum. This is the case, for instance, for the Poisson and Bernoulli distributions.

6.2.3 Convergence rate of extremes: A simulation study

One interesting problem in extreme value theory concerns the convergence rate of the distribution of the normalized maximum m_N^* towards the appropriate limit distribution. Unfortunately, there exist no analogue to the Berry-Esseen theorem which gives a single rate of convergence under very general conditions in the central limit theorem. In extreme value theory, the rate of convergence depends strongly on the right tail of the initial distribution F and on the choice of the normalizing constants a_N and b_N. The topic was first raised by Fisher and Tippett (1928) who remarked the slow convergence rate of the normal extremes. Since then, the convergence rate of extremes has been addressed by several authors [169–174]. For a collection of results see references [165, 175].

Here, we perform a simulation study on the convergence rate of the normalized maximum m_N^* for a finite sample size N. We consider different initial distributions summarized in table 6.1. For each of these distributions, we generate samples of different size $N = 5, 10, 20, 50, 100, 1000, 10000$. For each sample size we proceed as follows: We generate $K = 5000$ replications, determine the maximum of each replication and normalize it with the appropriate normalizing constants. The

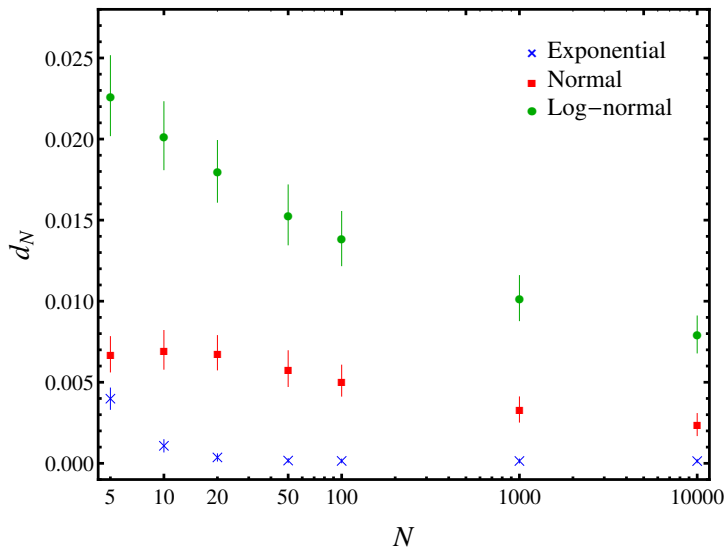


Figure 6.2: Distance d_N between the empirical distribution of the normalized maximum $\widehat{Q}_N$ and the Gumbel distribution G_ξ with $\xi = 0$ for exponential, normal and log-normal initial distributions versus the sample size N. Each point represents the distance d_N for a fixed N averaged over 500 simulation runs. The error bars represent the standard deviations.

normalizing constants for each distribution can be found in the examples discussed in the previous section. For the log-normal distribution we use the normalizing constants given in reference [176]. After obtaining the normalized maximum values, we estimate their empirical distribution function according to

$$\widehat{Q}_N(y) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{1}\{y_k \leq y\} \ , \tag{6.41}$$

where $y = m_N^*$ and $\mathbf{1}$ denotes the indicator function. We then compare the empirical distribution $\widehat{Q}_N(y)$ with the extreme value distribution G_ξ (6.11). To quantify the difference, we use the distance

$$d_N = \int_{-\infty}^{\infty} \left(\widehat{Q}_N(y) - G_\xi(y) \right)^2 \, \mathrm{d}y \ . \tag{6.42}$$

First, we focus on the Gumbel domain of attraction considering the exponential, normal and log-normal initial distributions. The corresponding distribution pa-

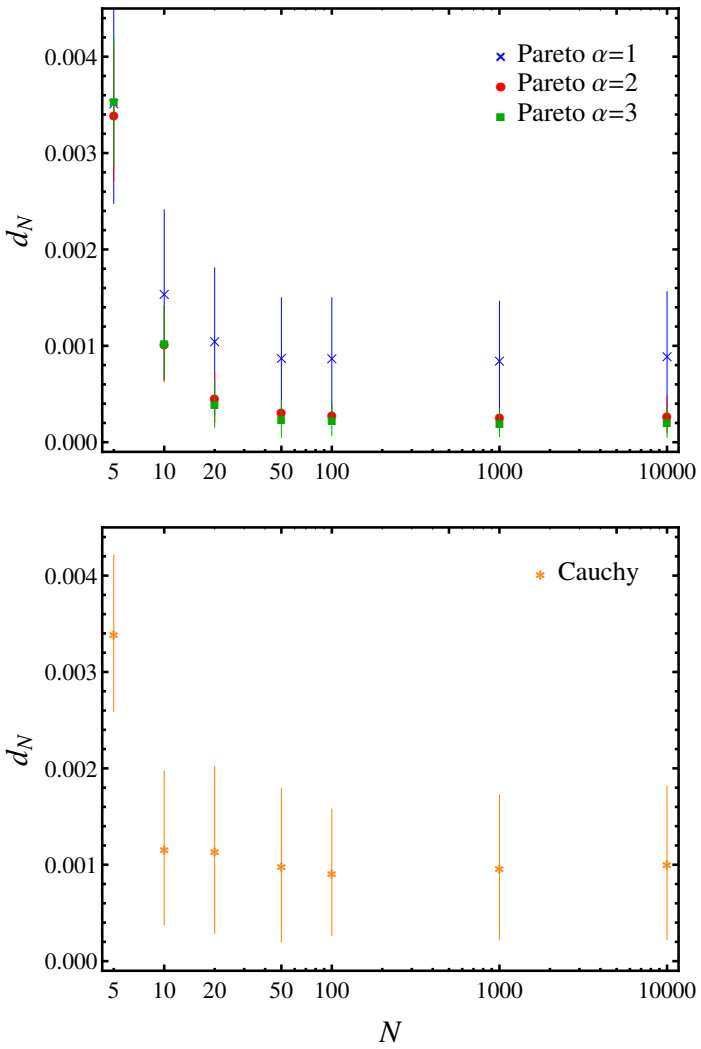


Figure 6.3: Distance d_N between the empirical distribution of the normalized maximum $\widehat{Q}_N$ and the Fréchet distribution (top) G_ξ with $\xi = 1/\alpha$ for Pareto initial distribution, and (bottom) G_ξ with $\xi = 1$ for Cauchy initial distribution versus the sample size N. Each point represents the distance d_N for a fixed N averaged over 500 simulation runs. The error bars represent the standard deviations.

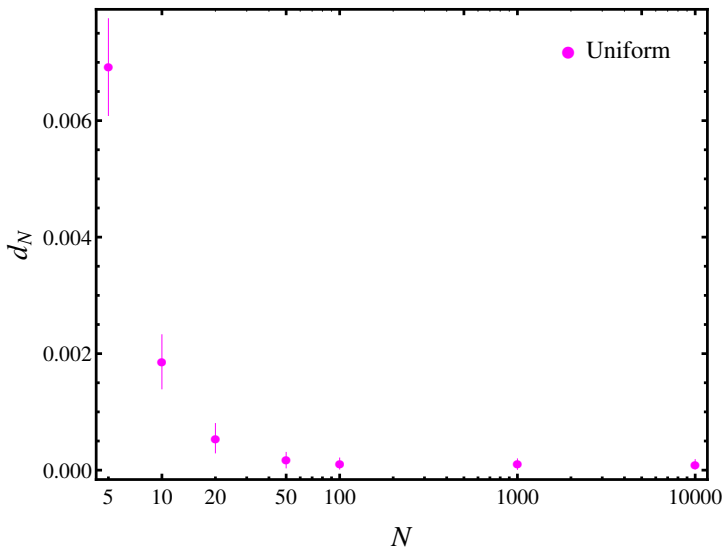


Figure 6.4: Distance d_N between the empirical distribution of the normalized maximum $\widehat{Q}_N$ and the Weibull distribution G_ξ with $\xi = -1$ for uniform initial distribution versus the sample size N. Each point represents the distance d_N for a fixed N averaged over 500 simulation runs. The error bars represent the standard deviations.

rameters can be found in table 6.1. Figure 6.2 shows the distance d_N between the empirical distribution of the normalized maximum $\widehat{Q}_N$ and the Gumbel distribution G_ξ with $\xi = 0$ for each sample size N. We observe a fast rate of convergence in the case of the exponential distribution: already for $N = 10$ the distribution of the normalized maximum is close to the Gumbel distribution, while for $N = 50$ they are almost indistinguishable. In contrast to this rapid rate of convergence, the normalized maximum of a sample of normal random variables converges extremely slowly to the limit distribution. Even worse is the rate of convergence of the normalized maximum of the log-normal random variables. The convergence rate does not depend on the parameters of the initial distribution, at least for the normalizing constants used in this study.

Next, we focus on the Fréchet domain of attraction. Here, we consider the Cauchy and Pareto initial distributions, see table 6.1. Figure 6.3 shows the distance d_N between the empirical distribution of the normalized maximum $\widehat{Q}_N$ and the Fréchet distribution G_ξ with $\xi > 0$. We observe a fast rate of convergence in both cases. For a sample drawn from a Pareto initial distribution we observe a faster rate of convergence, the higher the parameter α, i.e., the heavier the tails of the initial

distribution.

Finally, we look at the Weibull domain of attraction considering the uniform initial distribution, see table 6.1. Again, we observe a fast rate of convergence, see figure 6.4. Beyond $N = 50$ the empirical distribution $\widehat{Q}_N$ and the Weibull distribution G_ξ with $\xi = -1$ are almost indistinguishable.

6.3 Extreme value statistics of correlated random variables

So far, we have considered samples of independent random variables. In many applications, however, data are correlated. Here, we study the distribution of the maximum of equally correlated random variables. We begin with the normal case in section 6.3.1 and extend the results to the non-normal case in section 6.3.2. The results are verified in numerical simulations in section 6.3.3. Finally, we perform a comparison with financial data in section 6.3.4.

6.3.1 Maximum distribution of correlated normal random variables

Consider a sample of N correlated random variables $x = (x_1, \ldots, x_N)$ with the joint probability density function

$$p(x|\Sigma) = \frac{1}{\sqrt{\det 2\pi\Sigma}} \exp\left(-\frac{1}{2} x^\dagger \Sigma^{-1} x\right) \tag{6.43}$$

and covariance matrix Σ. We assume a covariance matrix of the form

$$\Sigma = (\sigma^2 - \rho^2)\mathbb{1}_N + \rho^2 g g^\dagger \ , \tag{6.44}$$

where $\mathbb{1}_N$ is the $N \times N$ unit matrix and $g = (1, \ldots, 1)$ a N component vector with unity in all entries. All diagonal elements are equal to σ^2, i.e., $\Sigma_{kk} = \sigma^2$, which represents the variance, and all off-diagonal elements are equal to ρ^2, i.e., $\Sigma_{kl} = \rho^2, k \neq l$, which represents the covariance. Furthermore, we require $\sigma^2 > \rho^2$.

In the following, we derive a distribution for the maximum of the correlated sample. The cumulative distribution function (cdf) of the maximum is defined as the probability to find the maximum in the interval between $-\infty$ and t, which is also the probability to find all random variables in the same interval, i.e.,

$$P_N^{\max}(t) = \Pr[x_1 \leq t, \ldots, x_N \leq t] = \int\limits_{x_n \leq t} \mathrm{d}[x]\ p(x|\Sigma) \ . \tag{6.45}$$

Thus, we have to compute the following integral

$$P_N^{\max}(t) = \frac{1}{\sqrt{\det 2\pi\Sigma}} \int\limits_{x_n \leq t} \mathrm{d}[x]\, \exp\left(-\frac{1}{2} x^\dagger \Sigma^{-1} x\right) \,. \tag{6.46}$$

Again, it is advantageous to use the Fourier transform of the multivariate normal distribution (6.43)

$$p(x|\Sigma) = \frac{1}{(2\pi)^N} \int \mathrm{d}[\omega]\, \mathrm{e}^{-\mathrm{i}\omega\cdot x} \exp\left(-\frac{1}{2}\omega^\dagger \Sigma\, \omega\right) \,, \tag{6.47}$$

with a N component real vector ω. With the covariance matrix (6.44), we can express the characteristic function in equation (6.47) in the following way

$$\begin{aligned}
\exp\left(-\frac{1}{2}\omega^\dagger \Sigma\, \omega\right) &= \exp\left(-\frac{1}{2}\omega^\dagger \left((\sigma^2-\rho^2)\mathbb{1}_N + \rho^2 g g^\dagger\right)\omega\right) \\
&= \exp\left(-\frac{\sigma^2-\rho^2}{2}\omega^\dagger\omega\right) \exp\left(-\frac{\rho^2}{2}(g\cdot\omega)^2\right) \\
&= \exp\left(-\frac{\sigma^2-\rho^2}{2}\omega^\dagger\omega\right) \frac{1}{\sqrt{2\pi\rho^2}} \int\limits_{-\infty}^{\infty} \mathrm{d}s\ \exp\left(-\frac{s^2}{2\rho^2} + \mathrm{i}sg\cdot\omega\right) \,,
\end{aligned} \tag{6.48}$$

where in the last step we express the second exponential function as a Gaussian integral. From now on, we require $\rho^2 > 0$ and $\sigma^2 > \rho^2 > 0$. Inserting the Fourier transform (6.47) with the characteristic function (6.48) into equation (6.46) and exchanging the order of integrations leads to

$$\begin{aligned}
P_N^{\max}(t) = \frac{(2\pi)^{-N}}{\sqrt{2\pi\rho^2}} &\int\limits_{-\infty}^{\infty} \mathrm{d}s \exp\left(-\frac{s^2}{2\rho^2}\right) \\
&\times \int\limits_{x_n\leq t} \mathrm{d}[x] \int \mathrm{d}[\omega] \exp\left(-\frac{\sigma^2-\rho^2}{2}\omega^\dagger\omega + \mathrm{i}(sg-x)\cdot\omega\right) \,.
\end{aligned} \tag{6.49}$$

The ω integral is Gaussian which gives

$$P_N^{\max}(t) = \frac{\left(2\pi(\sigma^2-\rho^2)\right)^{-N/2}}{\sqrt{2\pi\rho^2}} \int\limits_{-\infty}^{\infty} \mathrm{d}s \exp\left(-\frac{s^2}{2\rho^2}\right) \int\limits_{x_n\leq t} \mathrm{d}[x] \exp\left(-\frac{(sg-x)^2}{2(\sigma^2-\rho^2)}\right) \,. \tag{6.50}$$

We write the N dimensional integral in components

$$P_N^{\max}(t) = \frac{\left(2\pi(\sigma^2-\rho^2)\right)^{-N/2}}{\sqrt{2\pi\rho^2}} \int_{-\infty}^{\infty} \mathrm{d}s \exp\left(-\frac{s^2}{2\rho^2}\right) \prod_{n=1}^{N} \int_{-\infty}^{t} \mathrm{d}x_n \exp\left(-\frac{(s-x_n)^2}{2(\sigma^2-\rho^2)}\right) . \tag{6.51}$$

Performing the x_n integral leads to

$$P_N^{\max}(t) = \frac{1}{\sqrt{2\pi\rho^2}} \int_{-\infty}^{\infty} \mathrm{d}s \exp\left(-\frac{s^2}{2\rho^2}\right) \left(\frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{s-t}{\sqrt{2}\sqrt{\sigma^2-\rho^2}}\right)\right)^N . \tag{6.52}$$

Here, $\mathrm{erf}(\cdot)$ represents the error function defined as

$$\mathrm{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \mathrm{d}\tau \, \exp\left(-\tau^2\right) . \tag{6.53}$$

The error function is an odd function, i.e., $\mathrm{erf}(-y) = -\mathrm{erf}(y)$. Further, we recall the relation between the error function and the cumulative distribution Φ

$$\Phi(y) = \frac{1}{2} + \frac{1}{2} \, \mathrm{erf}\left(\frac{y}{\sqrt{2}}\right) . \tag{6.54}$$

Thus, we finally obtain

$$P_N^{\max}(t) = \frac{1}{\sqrt{2\pi\rho^2}} \int_{-\infty}^{\infty} \mathrm{d}s \exp\left(-\frac{s^2}{2\rho^2}\right) \Phi^N\left(\frac{t-s}{\sqrt{\sigma^2-\rho^2}}\right) \tag{6.55}$$

for the maximum distribution of correlated normally distributed random variables. We observe that the distribution of the maximum is a convolution of the distribution for the uncorrelated case $\Phi^N(t/\sigma)$ with a Gaussian function. The variance of the Gaussian function is the covariance ρ^2. In the limit $\rho \to 0$, we have

$$\lim_{\rho\to 0} P_N^{\max}(t) = \int_{-\infty}^{\infty} \mathrm{d}s \; \delta(s) \; \Phi^N\left(\frac{t-s}{\sigma}\right) = \Phi^N\left(\frac{t}{\sigma}\right) , \tag{6.56}$$

i.e., the maximum distribution for the uncorrelated case, see section 6.2.2.

We note that the distribution (6.55) depends on the variance σ^2 and the covariance ρ^2 of the random sample. We can alternatively express this result also in terms of the correlation coefficient $c = \rho^2/\sigma^2$. Substituting $\tilde{s} = s/\sigma$ into equation (6.55), we

obtain

$$P_N^{\max}(t) = \frac{1}{\sqrt{2\pi c}} \int_{-\infty}^{\infty} \mathrm{d}\tilde{s} \exp\left(-\frac{\tilde{s}^2}{2c}\right) \Phi^N\left(\frac{t/\sigma - \tilde{s}}{\sqrt{1-c}}\right) \ . \tag{6.57}$$

The corresponding probability density function (pdf) is obtained by taking the derivative

$$\begin{aligned} p_N^{\max}(t) &= \frac{\mathrm{d}P_N^{\max}(t)}{\mathrm{d}t} \\ &= \frac{1}{2\pi\sigma\sqrt{c(1-c)}} \int_{-\infty}^{\infty} \mathrm{d}\tilde{s} \exp\left(-\frac{\tilde{s}^2}{2c} - \frac{(t/\sigma - \tilde{s})^2}{2(1-c)}\right) N\ \Phi^{N-1}\left(\frac{t/\sigma - \tilde{s}}{\sqrt{1-c}}\right) \ . \end{aligned} \tag{6.58}$$

Figure 6.5 shows the pdf (6.58) and the cdf (6.57) for constant $\sigma = 0.1$ and $c = 0.3$ and different N values. As N increases the distribution of the maximum shifts to the right while the variance decreases. Figure 6.6 shows the pdf for a constant sample size $N = 300$ and different c and σ values.

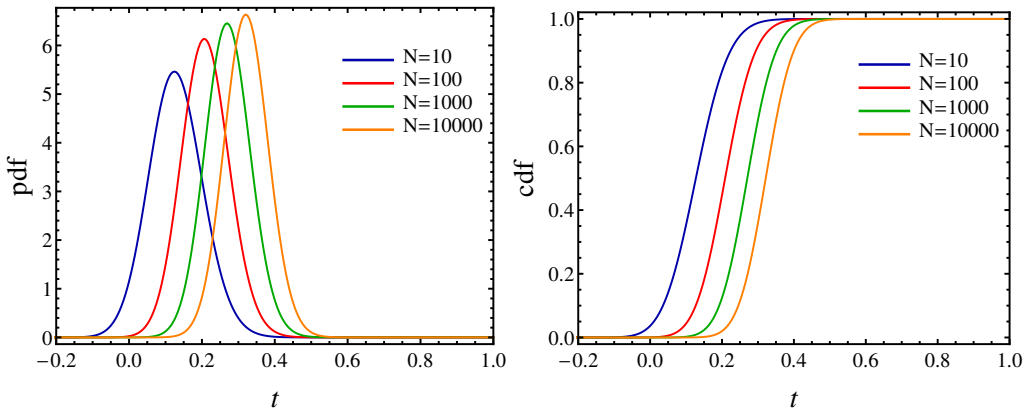


Figure 6.5: Probability density function (left) and cumulative distribution function (right) for fixed correlation coefficient $c = 0.3$, standard deviation $\sigma = 0.1$ and different sample size N.

We point out that the result derived here is consistent with earlier results of Gupta *et al.* [177, 178] for the case $\sigma = 1$. The distribution of the maximum of equally correlated normal variables has been extensively studied in the 1960s and 1970s, see e.g., references [178–180]. For a discussion on some special cases of unequally correlated random variables see e.g., references [181–183].

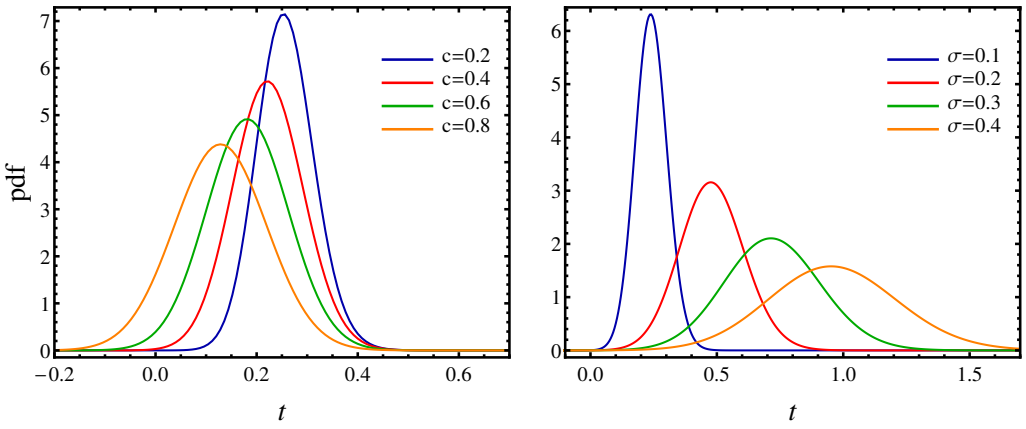


Figure 6.6: Left: Maximum pdf $p_N^{\max}(t)$ for fixed sample size $N = 300$, standard deviation $\sigma = 0.1$ and different c values. Right: Maximum pdf $p_N^{\max}(t)$ for fixed sample size $N = 300$, correlation coefficient $c = 0.3$ and different σ values.

6.3.2 Maximum distribution of correlated non-normal random variables

In the previous section, we derived a maximum distribution for a sample of equally correlated normal random variables. Here, we study non-normally distributed random variables.

Consider a sample of N correlated random variables $x = (x_1, \ldots, x_N)$ drawn from an arbitrary non-normal distribution $\widehat{p}(x|\Sigma)$. The distribution of the maximum can be obtained as in the normal case by solving the integral

$$\widehat{P}_N^{\max}(t) = \int\limits_{x_n \leq t} \mathrm{d}[x]\ \widehat{p}(x|\Sigma)\ . \tag{6.59}$$

Depending on the form of the distribution $\widehat{p}(x|\Sigma)$, the calculation of this integral could be extremely complicated. Here, we choose an alternative approach which allows us to use the results derived in the previous section. To this end, we introduce a deformation function which alters the multivariate normal distribution (6.43). Hence, the distribution $\widehat{p}(x|\Sigma)$ can be expressed as an average over the normal distribution

$$\widehat{p}(x|\Sigma) = \int\limits_0^\infty \mathrm{d}z\, h(z)\, p\left(x \middle| \frac{\Sigma}{z}\right) \tag{6.60}$$

with a deformation function $h(z)$. The deformation function fulfills the conditions

$$\int_{-\infty}^{\infty} h(z)\mathrm{d}z = 1 \quad \text{and} \quad h(z) \geq 0 \; . \tag{6.61}$$

For an arbitrary deformation function, the maximum distribution of a sample of correlated non-normal random variables can be written as

$$\widehat{P}_N^{\max}(t) = \int\limits_{x_n \leq t} \mathrm{d}[x]\, \widehat{p}(x|\Sigma) = \int_0^{\infty} \mathrm{d}z\, h(z) \int\limits_{x_n \leq t} \mathrm{d}[x]\, p\left(x \middle| \frac{\Sigma}{z}\right) \; . \tag{6.62}$$

The N dimensional integral represents the maximum distribution in the normal case where we replace σ^2 with σ^2/z and ρ^2 with ρ^2/z. Thus, we obtain

$$\widehat{P}_N^{\max}(t) = \int_0^{\infty} \mathrm{d}z\, \sqrt{z}\, h(z) \int_{-\infty}^{\infty} \mathrm{d}s \frac{\mathrm{e}^{-zs^2/2\rho^2}}{\sqrt{2\pi\rho^2}}\, \Phi^N\left(\frac{\sqrt{z}(t-s)}{\sqrt{\sigma^2-\rho^2}}\right) \; , \tag{6.63}$$

or in terms of the correlation coefficient $c = \rho^2/\sigma^2$

$$\widehat{P}_N^{\max}(t) = \int_0^{\infty} \mathrm{d}z\, \sqrt{z}\, h(z) \int_{-\infty}^{\infty} \mathrm{d}\tilde{s} \frac{\mathrm{e}^{-z\tilde{s}^2/2c}}{\sqrt{2\pi c}}\, \Phi^N\left(\frac{\sqrt{z}(t/\sigma-\tilde{s})}{\sqrt{1-c}}\right) \; . \tag{6.64}$$

To demonstrate our approach, we calculate an example. As a deformation function we choose a χ^2 distribution with L degrees of freedom given by

$$h(z) = \chi_L^2(z) = \frac{1}{2^{L/2}\Gamma(L/2)} z^{L/2-1} \exp\left(-\frac{z}{2}\right) \; . \tag{6.65}$$

Another example is discussed in appendix D. Inserting the χ^2 distribution (6.65) and the multivariate normal distribution (6.43) into equation (6.60), we obtain the deformed distribution

$$\begin{aligned} \widehat{p}(x|\Sigma) &= \frac{1}{2^{L/2}\Gamma(L/2)} \int_0^{\infty} \mathrm{d}z\; z^{L/2-1} \exp\left(-\frac{z}{2}\right) \frac{1}{\sqrt{\det 2\pi\Sigma/z}} \exp\left(-\frac{z}{2} x^{\dagger}\Sigma^{-1}x\right) \\ &= \frac{1}{2^{L/2}\Gamma(L/2)\sqrt{\det 2\pi\Sigma}} \int_0^{\infty} \mathrm{d}z\; z^{(L+N)/2-1} \exp\left(-\frac{z}{2}\left(1 + x^{\dagger}\Sigma^{-1}x\right)\right) \; . \end{aligned} \tag{6.66}$$

To simplify the z integral, we substitute

$$\rho = z\left(1 + x^\dagger \Sigma^{-1} x\right) \ . \tag{6.67}$$

This leads to

$$\begin{aligned}\widehat{p}(x|\Sigma) &= \frac{1}{2^{L/2}\Gamma(L/2)\sqrt{\det 2\pi\Sigma}} \int_0^\infty \mathrm{d}\rho \ \frac{\rho^{(L+N)/2-1}}{\left(1 + x^\dagger \Sigma^{-1} x\right)^{(L+N)/2}} \exp\left(-\frac{\rho}{2}\right) \\ &= \frac{1}{2^{L/2}\Gamma(L/2)\sqrt{\det 2\pi\Sigma}} \frac{1}{\left(1 + x^\dagger \Sigma^{-1} x\right)^{(L+N)/2}} \int_0^\infty \mathrm{d}\rho \ \rho^{(L+N)/2-1} \exp\left(-\frac{\rho}{2}\right) \ .\end{aligned} \tag{6.68}$$

We perform the ρ integral and find

$$\widehat{p}(x|\Sigma) = \frac{\Gamma((L+N)/2)}{\Gamma(L/2)\pi^{N/2}\sqrt{\det\Sigma}} \frac{1}{\left(1 + x^\dagger \Sigma^{-1} x\right)^{(L+N)/2}} \ , \tag{6.69}$$

which represents a multivariate Student's t-distribution with L degrees of freedom and covariance matrix Σ/L. With equation (6.64), we obtain

$$\widehat{P}_N^{\max}(t) = \frac{1}{2^{L/2}\Gamma(L/2)} \int_0^\infty \mathrm{d}z \ z^{(L-1)/2} \mathrm{e}^{-z/2} \int_{-\infty}^\infty \mathrm{d}\tilde{s} \ \frac{\mathrm{e}^{-z\tilde{s}^2/2c}}{\sqrt{2\pi c}} \ \Phi^N\left(\frac{\sqrt{z}(t/\sigma - \tilde{s})}{\sqrt{1-c}}\right) \tag{6.70}$$

for the maximum distribution of correlated t-distributed random variables. Figure 6.7 shows the maximum pdf

$$\widehat{p}_N^{\max}(t) = \frac{\mathrm{d}\widehat{P}_N^{\max}(t)}{\mathrm{d}t} \tag{6.71}$$

for a fixed sample size $N = 100$, standard deviation $\sigma = 0.1$, and different c and L values. Compared with the maximum distribution for the normal sample (6.58), the pdf is asymmetric with a longer right tail. It depends on an additional parameter, namely the degrees of freedom of the Student's t-distribution L. The smaller L, the heavier is the right tail of the maximum distribution.

We note that the result (6.64) is consistent with a similar result derived by Gupta *et al.* for the maximum distribution of correlated t-distributed random variables, see reference [184].

More recent work on the exact distribution of the maximum of correlated random variables can be found in reference [185].

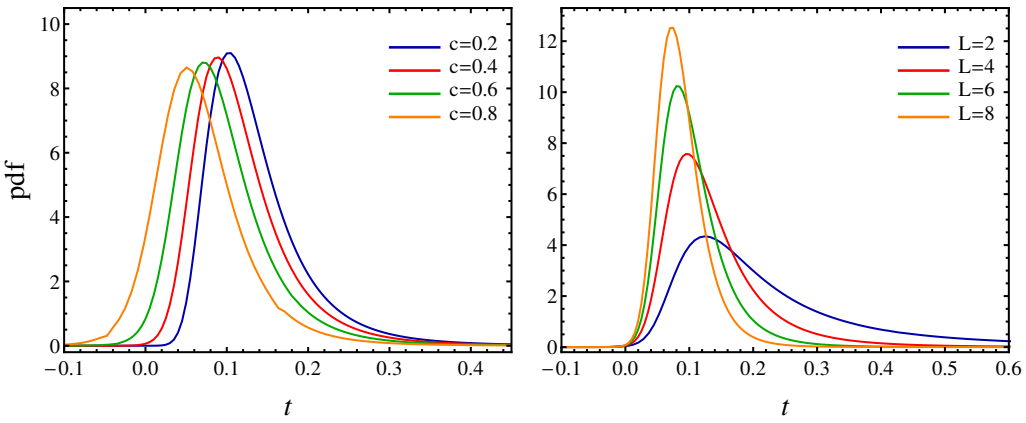


Figure 6.7: Left: Maximum pdf $\widehat{p}_N^{\max}(t)$ for fixed sample size $N = 300$, standard deviation $\sigma = 0.1$, $L = 5$ and different c values. Right: Maximum pdf $\widehat{p}_N^{\max}(t)$ for fixed sample size $N = 300$, standard deviation $\sigma = 0.1$, correlation coefficient $c = 0.4$ and different L values.

6.3.3 Comparison with numerical simulations

To confirm the validity of the results derived in the previous sections, we now compare them with numerical simulations. First, we consider the normal case. We generate $K = 5000$ random samples of size $N = 300$ from a multivariate normal distribution (6.43) with the covariance matrix

$$\Sigma = \sigma^2 \left((1-c)\mathbb{1}_N + cgg^\dagger \right) \ . \tag{6.72}$$

We carry out the simulations for three different correlation coefficients $c = 0.3, 0.5, 0.7$ and fixed standard deviation $\sigma = 0.1$. For each c value, we determine the maximum of each sample

$$m_{N,k} = \max(x_{1,k}, \ldots, x_{N,k}), \quad k = 1, \ldots, K \tag{6.73}$$

and compare the histogram of the maximum values with the maximum pdf $p_N^{\max}(t)$ (6.58). The results are depicted in figure 6.8. We find a good agreement between analytical and numerical results.

Second, we consider the non-normal case. We generate $K = 5000$ random samples of size $N = 300$ from a multivariate Student's t-distribution of the form (6.69) with the covariance matrix (6.72). We carry out the simulations for different correlation coefficients $c = 0.3, 0.5, 0.7$, fixed standard deviation $\sigma = 0.1$ and degrees of freedom $L = 5$. The histograms of the maximum values are compared with the maximum pdf

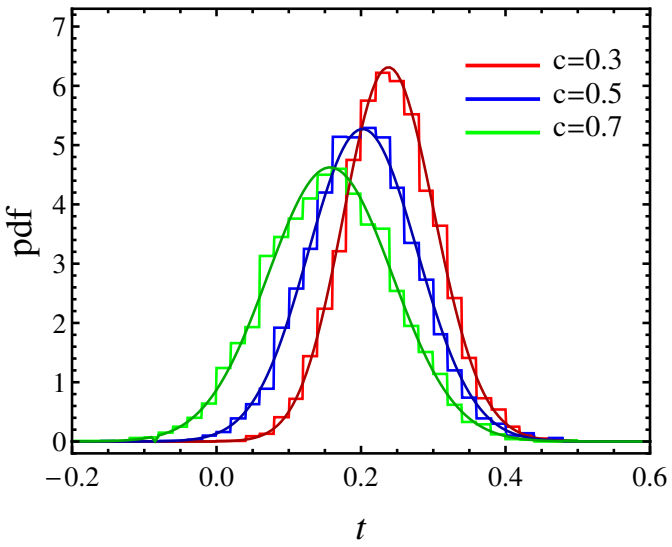


Figure 6.8: Comparison of the maximum pdf $p_N^{\max}(t)$ for correlated normal random variables (continuous lines) with numerical simulations for fixed sample size $N = 300$, $\sigma = 0.1$ and different correlation coefficients c.

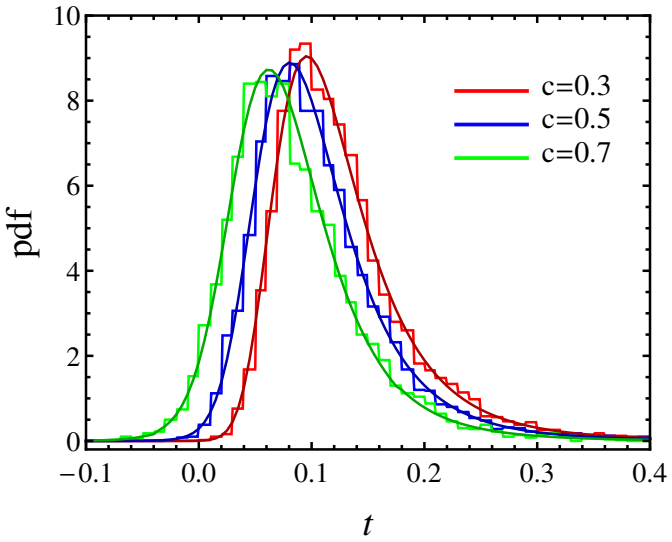


Figure 6.9: Comparison of the maximum pdf $\widehat{p}_N^{\max}(t)$ for correlated t-distributed random variables (continuous lines) with numerical simulations for fixed sample size $N = 300$, $\sigma = 0.1$, $L = 5$ and different correlation coefficients c.

$\widehat{p}_N^{\max}(t)$ (6.71), see figure 6.9. Again, we find a good agreement between analytical results and simulations.

6.3.4 Comparison with financial data

Finally, we compare our results with empirical stock returns. We consider a dataset consisting of $N = 301$ stocks from the S&P 500 index traded in the time period January 1992 to December 2014. From the price time series we calculate the returns

$$r_n(t) = \frac{S_n(t + \Delta t) - S_n(t)}{S_n(t)}, \quad n = 1, \ldots, N \ , \tag{6.74}$$

where $S_n(t)$ is the price of the stock n at time t and $\Delta t = 1$ trading day. Thus, we obtain $T = 5793$ return vectors $r(t)$. For each return vector we determine the maximum

$$m_N(t) = \max(r_1(t), \ldots, r_N(t)), \quad t = 1, \ldots, T \ . \tag{6.75}$$

In section 2.3, we assumed normally distributed return vectors at each time t and verified this assumption in an empirical study in section 2.4. Therefore, we compare the histogram of the empirical maximum values $m_N(t)$ with the maximum pdf $p_N^{\max}(t)$ (6.58) for the normal case. The maximum pdf depends on two parameters, namely the standard deviation σ and the correlation coefficient c. In the numerical simulations, we had all information about the parameters. Here, we have to extract them from the empirical data. We estimate the correlation coefficient c by averaging over the off-diagonal elements of the correlation matrix evaluated over the whole observation period, and the standard deviation σ by averaging over the standard deviations of all stocks estimated over the whole observation period. The comparison reveals no agreement between theory and data, see figure 6.10. The histogram is asymmetric with a longer right tail and cannot be described by the maximum pdf even by fitting the parameters σ and c.

We point out that we derived the maximum distribution $p_N^{\max}(t)$ for a fixed covariance matrix Σ. However, as we discussed in the previous chapters, return time series are non-stationary. Both standard deviation and correlation change significantly over time. Thus, it would be interesting to study the effect of the non-stationarity on the distribution of the maximum. On the other hand, we assumed equally correlated random variables. This is also hardly the case for stock returns. Extending the result to the unequally correlated case could lead to a more realistic description of financial data. These are subjects for future research.

In addition, we compare the histogram of the empirical maximum values $m_N(t)$ with the maximum pdf $\widehat{p}_N^{\max}(t)$ (6.71) for the non-normal case which also exhibits a longer right tail. We note that the maximum distribution $\widehat{p}_N^{\max}(t)$ depends on the parameters σ and c of the normal distribution $p(x|\Sigma)$. From the data we can

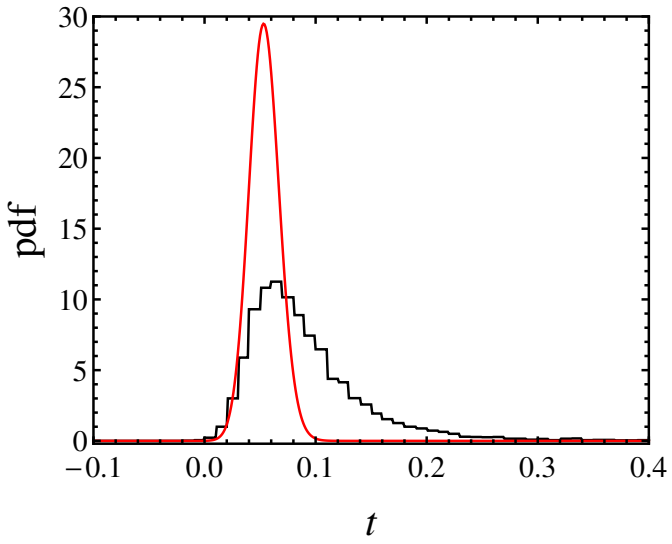


Figure 6.10: Histogram of the empirical maximum values (black) compared with the maximum pdf $p_N^{\max}(t)$ (red) for $N = 301$, where $\sigma = 0.022$ and $c = 0.28$ are estimated from the data.

only obtain the parameters of the deformed distribution $\widehat{p}(x|\Sigma)$. The corresponding correlation matrix $\Sigma^{(d)}$ slightly differs from Σ. We calculate $\Sigma^{(d)}$ as follows

$$\begin{aligned}\Sigma^{(d)} = \langle xx^\dagger \rangle &= \int \mathrm{d}[x]\, xx^\dagger\, \widehat{p}(x|\Sigma) \\ &= \int \mathrm{d}[x]\, xx^\dagger\, p(x|\Sigma) \int_0^\infty \mathrm{d}z\, \frac{h(z)}{z} \\ &= \Sigma\, \overline{z^{-1}}\ . \end{aligned} \tag{6.76}$$

The covariance matrices differ by the average of $1/z$. For the case $h(z) = \chi^2_L(z)$ we obtain

$$\Sigma^{(d)} = \frac{\Sigma}{L-2}\ . \tag{6.77}$$

Thus, we have $c = c^{(d)}$ and $\sigma = \sigma^{(d)}\sqrt{L-2}$, where $c^{(d)}$ and $\sigma^{(d)}$ are the average parameters obtained from the data. Here, we have $c^{(d)} = 0.28$ and $\sigma^{(d)} = 0.022$. The remaining parameter L is fitted to the empirical distribution. Again, the comparison shows a poor agreement when determining the parameters σ and c from the data, see figure 6.11. Fitting all parameters improves the agreement considerably.

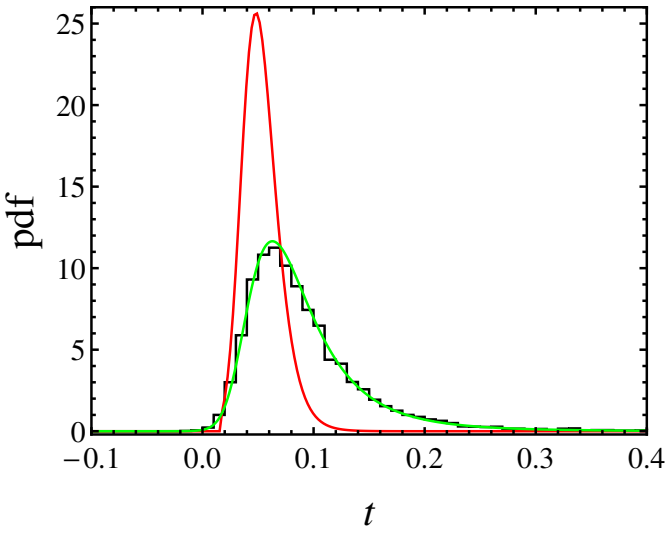


Figure 6.11: Histogram of the empirical maximum values (black) compared with the maximum pdf $\widehat{p}_N^{\max}(t)$ for $N = 301$, (red) $\sigma = 0.082$ and $c = 0.28$ estimated from the data and $L = 16$ fitted to the data, and (green) $\sigma = 0.046$, $c = 0.4$ and $L = 4$ fitted to the data.

6.4 Summary

In this chapter, we presented some preliminary work and first results on the extreme value statistics of correlated random variables. First, we recapitulated the theory for iid random variables and studied the convergence rate of the properly normalized maximum towards the limit distributions for a finite sample size. We found that the convergence rate depends strongly on the form of the initial distribution. In particular, the normalized maximum of a sample of normal and log-normal random variables converges extremely slowly towards the Gumbel distribution. In contrast to this slow rate of convergence, we observe a fast convergence rate in case of the exponential distribution. We found a fast rate of convergence also for the Pareto, Cauchy and uniform initial distributions.

Next, we addressed the correlated case. We derived a maximum distribution for a sample of equally correlated normal random variables, which is a convolution of the distribution of the maximum for the uncorrelated case and a Gaussian function. Introducing a deformation function, we were able to generalize this result to the non-normal case. In particular, we derived a maximum distribution for a sample of equally correlated t-distributed random variables. We verified our results in numerical simulations.

For future research, there are several directions that one might pursue. So far,

we have considered only equally correlated random variables. Therefore, a next step would be to extend our approach to the unequally correlated case. On the other hand, the maximum distributions were derived for a fixed covariance matrix, i.e., they hold only in a stationary setting. Taking the non-stationarity of the covariances into account could be beneficial for financial applications.

Furthermore, it would be interesting to study if our results lead to an universal distribution as in the uncorrelated case. To this end, one has to find normalizing constants which stabilize the maximum, if they do exist, and study the large sample limit. This will be an important contribution to the extreme value theory of correlated random variables.

Chapter 7

Conclusion and outlook

Non-stationarity is a central aspect of financial markets. It manifests itself, on the one hand, in the fluctuating volatilities of single stocks, and on the other hand, in the time-varying correlations between different stocks. Here, we focused on the non-stationarity of correlations.

In order to model this non-stationarity, we proposed an approach based on random matrices. Applications of random matrix theory in finance are not new. Typically, they address the statistical properties of empirical correlation matrices and model the estimation errors due to the finiteness of the financial time series. Here, we put forward a new application of random matrix theory. In contrast to other applications, which are interested in the statistics of an individual large empirical correlation matrix, our approach considers an ensemble of empirical correlation matrices created by the fluctuating correlations. This ensemble is modeled by a Wishart random matrix ensemble constructed such that it has on average the same correlation structure as the ensemble of empirical correlation matrices. Further, we assumed return vectors following a multivariate normal distribution with different correlation or covariance matrices. Averaging the multivariate normal distribution over the random matrix ensemble led to a correlation averaged distribution, which yields a realistic quantitative description of heavy-tailed multivariate return distributions. Our model demonstrated that the non-stationarity of the correlations between stocks leads to heavy tails in the multivariate return distribution of a stock market.

After establishing the random matrix approach, we studied some applications, focusing in particular on stock markets. Recently, the random matrix approach was also applied in the context of credit risk which led to the computation of an average loss distribution for credit portfolios [5]. We note that although derived having financial applications in mind, our random matrix approach can also be applied to other complex systems exhibiting non-stationary behavior.

We applied the random matrix approach to investigate the effect of non-stationarity on the distribution of portfolio returns. We derived a heavy-tailed distribution for the return of a portfolio which we compared with the portfolio returns of randomly selected portfolios. We found that the average portfolio return distribution yields a good description of empirical portfolio returns, in particular in the central part of the distribution. Still, we found some deviations in the tails. The choice of the Wishart distribution for the random correlation matrix ensemble is a reasonable assumption, but it obviously cannot capture all empirical details. Recently, Meudt *et al.* [186] extended our random matrix approach by introducing a deformation of the Wishart ensemble. They found an ensemble characterized by an algebraic distribution which improved the overall agreement with financial data.

Further, we used the random matrix approach to study the stability of the correlation structure of market states. We identified market states as clusters of correlation matrices with similar correlation structures. As fluctuations due to measurement noise compete with actual fluctuations due to non-stationarity, we cannot study the correlation structure stability directly from the empirical correlation matrices. The random matrix approach provides an alternative. It reduces the complexity of the financial market to a single parameter which measures the fluctuations of true correlations due to non-stationarity. This parameter characterizes the stability of the correlation structure and most importantly, it can be determined directly from the empirical return distributions. We found clear indications for correlation fluctuations within market states. Furthermore, our analysis revealed an intriguing relationship between average correlation and correlation fluctuations. The strongest fluctuations occur during periods with high average correlation, which is in particular the case in times of crisis.

As correlations measure only the linear dependence between time series, we next used a copula approach to study the dependence structure of market states. To this end, we estimated empirical pairwise copulas for each market state. We considered both original returns, which exhibit time-varying trends and volatilities, as well as locally normalized ones, where the non-stationarity had been removed. This allowed us to study the dependence structure on two different scales: a global and a local scale. We found evidence for asymmetric dependencies between financial returns. Empirical copulas exhibit an asymmetry in the positive tail dependence. We observed a stronger lower tail dependence, indicating a larger probability for simultaneous extreme negative returns, in particular in times of crisis. Furthermore, we compared the empirical pairwise copulas for each state with a bivariate K-copula calculated from the correlation averaged multivariate return distribution. We found that the K-copula captures the overall empirical dependence structure of market states rather well, in particular on a local scale. Thus, we arrived at a consistent picture within our random matrix model: The correlation averaged multivariate normal distribution describes the heavy-tailed empirical return distributions and

the K-copula captures the empirical dependence structure.

A further focus of this thesis was the extreme value statistics of correlated random variables. In particular, we studied the distribution of the maximum of correlated samples. Assuming equally correlated random variables, we derived a maximum distribution for a sample drawn from a multivariate normal distribution. Introducing a deformation function, we further extended this result to the non-normal case and derived a maximum distribution for a sample of equally correlated t-distributed random variables. Numerical simulations verified our results. For future research, there are several directions that one might pursue. On the one hand, a natural next step would be to study the maximum distribution of unequally correlated random variables. On the other hand, it would be interesting to investigate the effect of non-stationarity on the distribution of the maximum. This could be beneficial for financial applications. In addition, it would be interesting to study if our results lead to an universal distribution as in the uncorrelated case. This would be an important contribution to the extreme value theory of correlated random variables and could shed some light on the issue of the universality classes in the correlated case.

Appendix A

Supplementary material to chapter 2

Here, we present some details of the random matrix approach introduced in chapter 2.

A.1 The case $N < K$

We show that the pdf (2.1) is well defined for $N < K$ in terms of proper delta functions. To this end, we diagonalize the covariance matrix in the following way

$$\Sigma = U^\dagger \Lambda \, U \tag{A.1}$$

with an orthogonal matrix $U = U^\dagger$ and a diagonal matrix of the eigenvalues Λ. In the case $N < K$, the covariance matrix has N positive non-zero and $K - N$ zero eigenvalues

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N, 0, \ldots, 0) \; . \tag{A.2}$$

We rotate the vector x into the eigenbasis of the covariance matrix, write $v = Ux$, and obtain the pdf

$$\begin{aligned} g(v|\Lambda) &= \frac{1}{\sqrt{2\pi}^K} \frac{1}{\sqrt{\det \Lambda}} \exp\left(-\frac{1}{2} v^\dagger \Lambda^{-1} v\right) \\ &= \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\lambda_k}} \exp\left(-\frac{v_k^2}{2\lambda_k}\right) \prod_{k=N+1}^{K} \delta(v_k) \; , \end{aligned} \tag{A.3}$$

which clearly is well defined. A corresponding line of reasoning can be used for the Fourier transform (2.9).

A.2 χ^2 representation of equation (2.18)

We show the appearance of the χ^2 distribution in equation (2.18), see [1]. Using the formula

$$\frac{1}{a^\eta} = \frac{1}{2^\eta \Gamma(\eta)} \int_0^\infty z^{\eta-1} \exp\left(-\frac{a}{2}z\right)\, dz\ , \tag{A.4}$$

we can cast equation (2.16) into the form

$$\begin{aligned}\langle g\rangle(x|\Sigma,N) = \frac{1}{2^{N/2}\ \Gamma(N/2)} \int_0^\infty \mathrm{d}z\ z^{N/2-1} \exp\left(-\frac{z}{2}\right) \\ \times \int \frac{\mathrm{d}[\omega]}{(2\pi)^K}\ \exp\left(-\mathrm{i}\omega^\dagger x - \frac{z}{2N}\omega^\dagger \Sigma \omega\right)\end{aligned} \tag{A.5}$$

with $\Sigma = \sigma C \sigma$. The ω integral yields a multivariate Gaussian with the covariance matrix $z\Sigma/N$. Hence, we arrive at

$$\langle g\rangle(x|\Sigma,N) = \int_0^\infty \chi^2_N(z)\ g\left(x\middle|\frac{z}{N}\Sigma\right) dz \tag{A.6}$$

with the χ^2 distribution of N degrees of freedom

$$\chi^2_N(z) = \frac{1}{2^{N/2}\Gamma(N/2)} z^{N/2-1} \exp\left(-\frac{z}{2}\right) \tag{A.7}$$

for $z \geq 0$ and zero otherwise.

A.3 Equivalence between $w(W|C,N)$ and $\widehat{w}(A|\Sigma,N)$

We argue that

$$\widehat{w}(A|\Sigma,N) = \sqrt{\frac{N}{2\pi}}^{KN} \frac{1}{\sqrt{\det\Sigma}^N}\ \exp\left(-\frac{N}{2}\operatorname{tr} A^\dagger \Sigma^{-1} A\right) \tag{A.8}$$

is equivalent to $w(W|C,N)$, where $\Sigma = \sigma C \sigma$ and $A = \sigma W$. To this end, we have to show that

$$\int d[A]\ \widehat{w}(A|\Sigma,N)\ f(A) = \int d[W]\ w(W|C,N)\ f(\sigma W) \tag{A.9}$$

with an arbitrary test function $f(A)$. Changing the variables according to $A = \sigma W$, we find

$$\int d[A]\ \widehat{w}(A|\Sigma,N)\ f(A) = (\det\sigma)^N \int d[W]\ \widehat{w}(\sigma W|\Sigma,N)\ f(\sigma W) \tag{A.10}$$

$$= \int d[W]\ w(W|C,N)\ f(\sigma W)\ , \tag{A.11}$$

which proves the assertion.

Appendix B

Stock data

B.1 Database

Yahoo! Finance [106] provides free historical (daily) data for stocks and indices. To access the quotes, we enter the stock symbol, also called ticker symbol, and choose Historical Prices in the Quotes menu, see figure B.1.

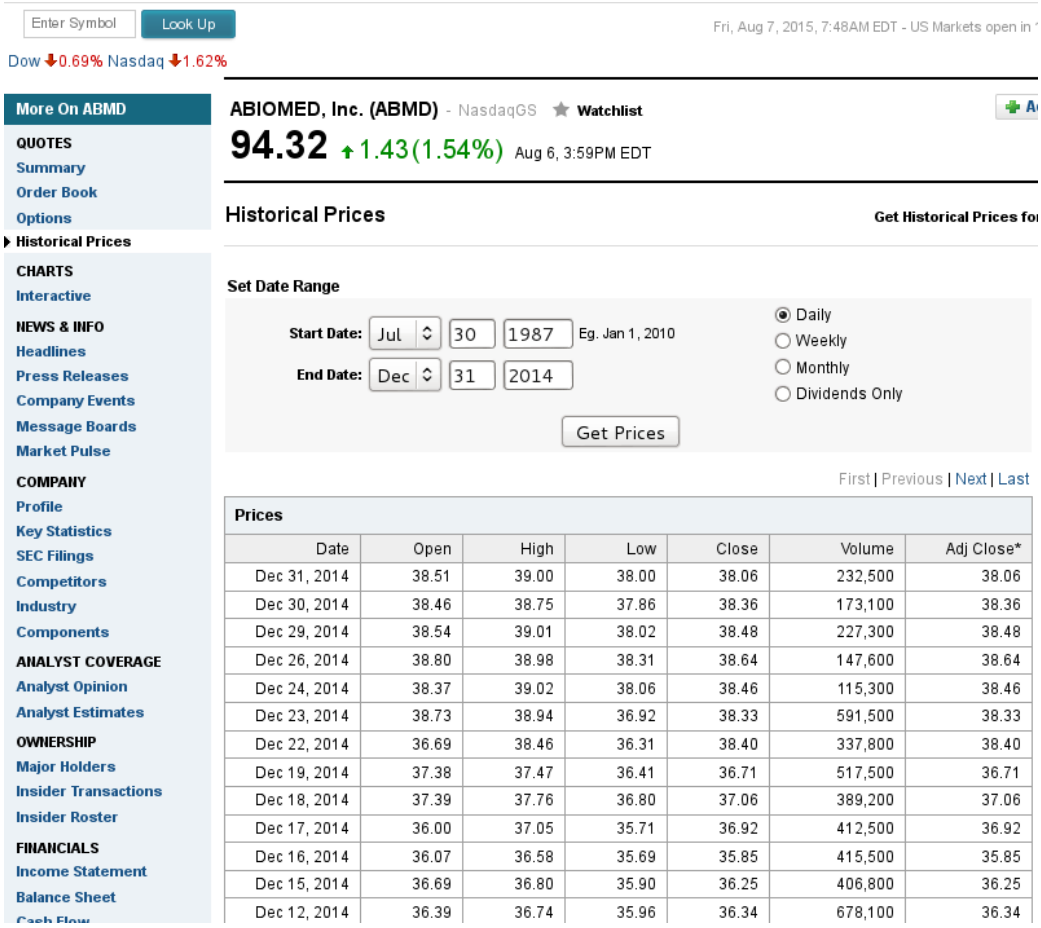
On the Historical Prices page of a stock, we can filter historical prices by a custom date range, or with the available preset filters. For daily prices the open, high, low, close, and volume for each trading day are shown. For the data analysis we are interested in the last column containing the adjusted close, which provides the closing price for the requested day adjusted for all applicable splits and dividend distributions.

B.2 Download

To download the historical prices for a set of stocks, we need a list of their ticker symbols. The list serves as an input for the download script

```
while read ticker
do
echo $ticker
curl "http://ichart.finance.yahoo.com/table.csv?s="$ticker"&a=00&b=3
&c=2000&d=11&e=30&f=2012&g=d&ignore=.csv" | awk -F',' '{ print $1" "$7 }'
| grep -v Date > $ticker.txt
done
```

For each stock we obtain a file named `ticker.txt` which contains a list of dates and daily adjusted closing prices in the specified time period, here January 3, 2000



Date	Open	High	Low	Close	Volume	Adj Close*
Dec 31, 2014	38.51	39.00	38.00	38.06	232,500	38.06
Dec 30, 2014	38.46	38.75	37.86	38.36	173,100	38.36
Dec 29, 2014	38.54	39.01	38.02	38.48	227,300	38.48
Dec 26, 2014	38.80	38.98	38.31	38.64	147,600	38.64
Dec 24, 2014	38.37	39.02	38.06	38.46	115,300	38.46
Dec 23, 2014	38.73	38.94	36.92	38.33	591,500	38.33
Dec 22, 2014	36.69	38.46	36.31	38.40	337,800	38.40
Dec 19, 2014	37.38	37.47	36.41	36.71	517,500	36.71
Dec 18, 2014	37.39	37.76	36.80	37.06	389,200	37.06
Dec 17, 2014	36.00	37.05	35.71	36.92	412,500	36.92
Dec 16, 2014	36.07	36.58	35.69	35.85	415,500	35.85
Dec 15, 2014	36.69	36.80	35.90	36.25	406,800	36.25
Dec 12, 2014	36.39	36.74	35.96	36.34	678,100	36.34

Figure B.1: Screenshot Yahoo! Finance historical prices for the ABIOMED stock.

to December 30, 2012.

Alternatively, one can use the `FinancialData` function of Mathematica

```
Table[FinancialData[TickerList[[k]], {{2000, 1, 3}, {2012, 12, 30}}],
      {k,1,K}]
```

where `TickerList` contains the ticker symbols of the K stocks.

B.3 Datasets

In addition, we present a full list of the stocks used in the empirical studies in this thesis. The stocks are listed by their ticker symbols ordered in alphabetical order.

The following stocks were used in the empirical study in section 2.4.1:

S&P 500 1992-2012: AA, AAPL, ABT, ADBE, ADI, ADM, ADP, ADSK, AEP, AES, AET, AFL, AGN, AIG, ALTR, AMAT, AMD, AMGN, AON, APA, APC, APD, APH, ARG, AVP, AVY, AXP, AZO, BA, BAC, BAX, BBT, BBY, BCR, BDX, BEN, BF.B, BHI, BIG, BIIB, BK, BLL, BMC, BMS, BMY, C, CA, CAG, CAH, CAT, CB, CCE, CCL, CELG, CERN, CI, CINF, CL, CLF, CLX, CMA, CMCSA, CMI, CMS, CNP, COG, COP, COST, CPB, CSC, CSCO, CSX, CTAS, CTL, CVH, CVS, CVX, D, DD, DE, DELL, DHR, DIS, DNB, DOV, DOW, DTE, DUK, ECL, ED, EFX, EIX, EMC, EMR, EOG, EQT, ETN, ETR, EXC, F, FAST, FDO, FDX, FHN, FISV, FITB, FLS, FMC, FRX, GAS, GCI, GD, GE, GIS, GLW, GPC, GPS, GT, GWW, HAL, HAS, HBAN, HCP, HD, HES, HNZ, HOG, HON, HOT, HP, HPQ, HRB, HRL, HRS, HST, HSY, HUM, IBM, IFF, IGT, INTC, IP, IPG, IR, ITW, JCI, JCP, JEC, JNJ, JPM, JWN, K, KEY, KIM, KLAC, KMB, KO, KR, L, LEG, LEN, LH, LLTC, LLY, LM, LMT, LNC, LOW, LSI, LTD, LUK, LUV, MAS, MAT, MCD, MDT, MHP, MKC, MMC, MMM, MO, MOLX, MRK, MRO, MSFT, MSI, MTB, MU, MUR, MWV, MYL, NBL, NBR, NE, NEE, NEM, NI, NKE, NOC, NSC, NTRS, NU, NUE, NWL, OI, OKE, OMC, ORCL, OXY, PAYX, PBCT, PBI, PCAR, PCG, PCL, PCP, PEP, PFE, PG, PGR, PH, PHM, PLL, PNC, PNW, POM, PPG, PPL, PSA, QCOM, R, RDC, RF, ROK, ROST, RRD, RSH, RTN, S, SCG, SCHW, SEE, SHW, SIAL, SLB, SLM, SNA, SO, SPLS, STI, STJ, STT, SVU, SWK, SWN, SWY, SYK, SYMC, SYY, T, TAP, TE, TEG, TER, TGT, THC, TIF, TJX, TLAB, TMK, TMO, TROW, TRV, TSN, TSO, TXN, TXT, TYC, UNH, UNP, USB, UTX, VAR, VFC, VLO, VMC, VNO, VZ, WAG, WDC, WEC, WFC, WFM, WHR, WM, WMB, WMT, WPO, WY, X, XEL, XL, XLNX, XOM, XRAY, XRX, ZION

The following stocks were used in the empirical studies in sections 2.4.2 and 2.4.3, and chapters 4 and 5:

NASDAQ Composite 1992-2013: AAPL, ABMD, ACAT, ACET, ACXM, ADBE, ADI, ADP, ADSK, AEGN, AGII, AGYS, AIRM, ALCO, ALKS, ALOG, ALOT, ALTR, AMAG, AMAT, AMGN, AMSC, AMSWA, AMWD, ANAT, APOG, ARKR, AROW, ASBC, ASEI, ASMI, ASNA, ASRV, ASTE, ATAX, ATML, ATNI, ATRI, AVNW, BCPC, BEAV, BELFA, BIIB, BOBE, BONT, BOOM, BOTA, BPOP, BSET, BTUI, CA, CAR, CASY, CATY, CBRL, CBRX, CBSH, CDNS, CELG, CERN, CGNX, CINF, CLDX, CMCSA, CMCSK, CNMD, COBR, COHR, COHU, COKE, COST, CRUS, CRVL, CSCO, CSWC, CTAS, CTG, CY, CYBE, CYTR, DAIO, DGAS, DGII, DORM, DXYN, EA, ECOL, EMCI, ENZN, ERIC, ESIO, EXPD, EXPO, EZPW, FAST, FFBC, FISV, FITB, FIZZ, FMBI, FMER, FOX, FSTR, FTR, FULT, FWLT, GIII, GK, GLDC, GNTX, GT, HAS, HBAN, HBHC, HCSG, HELE, HOLX, HRTX, HTCH, HTLD, HURC, HWAY, IDCC, IDTI, IDXX, IIN, IMGN, IMKTA, IMMU, INDB, INPH, INTC, IPAR, ISIS, JBHT, JBSS, JCS, JJSF, JKHY, JOUT, KBALB, KELYA, KLAC, KLIC, KTCC, LANC, LAWS, LCUT, LLTC, LNCE, LRCX, LSCC, LYTS, MAT, MCRS, MDCI, MENT, MERC, MGEE, MGIC, MGPI, MGRC, MLHR, MOCO, MSCC, MSEX, MSFT, MTSC, MU, MXIM, MXWL, MYL, NDSN,

NEWP, NPBC, NTRS, NWLI, ODFL, ONTY, ORBK, ORBT, OTTR, PATK, PAYX, PBCT, PCAR, PCH, PENX, PERF, PICO, PKOH, PLAB, PLXS, PMCS, PNRA, POWL, PPC, PRGS, PTC, QCOM, QLTI, RAVN, REGN, RGEN, ROST, RRD, SAFM, SBCF, SEIC, SHLM, SIAL, SIGI, SIGM, SIVB, SKYW, SLM, SONC, SPAN, SPAR, SPEX, SPLS, STFC, SUSQ, SWKS, SYMC, SYNL, TECD, TECU, TILE, TROW, TRST, TTEK, TWIN, TXN, UBSI, UMBF, USEG, USTR, VICR, VOD, VRTX, VTSS, WAFD, WDC, WDFC, WERN, WPPGY, WRLD, WSFS, XCRA, XLNX, XOMA, XRAY, YRCW, ZBRA, ZION, ZIXI, ZOOM

The following stocks were used in the empircal study in chapter 3:

NASDAQ Composite 1992-2012: AAPL, ABMD, ACAT, ACET, ADBE, ADI, ADP, ADSK, AEGN, AGYS, AIRT, ALCO, ALKS, ALOT, ALRN, ALTR, AMAG, AMAT, AMGN, AMSC, AMSWA, AMWD, APAGF, ARCI, ARTW, ASBC, ASBI, ASNA, ASRV, ASYS, ATAX, ATML, ATRO, ATX, BANF, BCPC, BEAV, BIIB, BKSC, BMTC, BOBE, BPOP, BRID, BSET, BWINA, BWINB, CASY, CA, CATY, CBRL, CBRX, CBSH, CDNS, CELG, CERN, CFNB, CGNX, CINF, CLDX, CMCSA, CMCSK, CNBKA, CNMD, COHR, COHU, COST, CRBC, CRRC, CRUS, CSCO, CSWC, CTAS, CYTR, CY, DELL, DGAS, DGICB, DGII, DIAL, EA, ECOL, ELSE, EMCI, ENZN, EPHC, ERIC, EXAR, FARM, FAST, FELE, FFBC, FISV, FITB, FLOW, FMER, FSBI, FWLT, GLCH, GLDC, GNTX, GSBC, HAS, HBAN, HBHC, HOLX, HTCH, HWAY, IDCC, IDTI, IDXX, IFSIA, IMGN, IMKTA, IMMU, INCB, INTC, IPAR, ISIS, ITIC, JBHT, JCS, JKHY, KELYA, KLAC, KLIC, KSWS, KTCC, LANC, LAWS, LCUT, LLTC, LNCE, LRCX, LSCC, LTXC, LYTS, MAG, MAT, MCRS, MENT, MERC, MGEE, MGRC, MKTAY, MLHR, MOLX, MSEX, MSFG, MSFT, MU, MXIM, MXWL, MYL, NABI, NAFC, NAVG, NBBC, NDSN, NHTB, NPBC, NSEC, NTRS, NTSC, NWK, NWS, ORCL, OTTR, PAYX, PBCT, PCAR, PCH, PENX, PLAB, PLFE, PLXS, PMCS, PMFG, PMTC, PNRA, POPE, PRGO, PRGS, PRST, PTSI, QCOM, REGN, RELL, REXI, RGEN, ROST, RRD, SAFM, SBCF, SHLM, SIAL, SIGM, SKYW, SLM, SMSC, SPAN, SPAR, SPLS, STEI, STFC, SUSQ, SVNT, SWKS, SYMC, SYMM, TECD, TECUB, TLAB, TROW, TRST, TTEK, TWIN, TXN, UBSI, UFCS, UMBF, UNAM, USLM, VICR, VLGEA, VOD, VRTX, VTSS, WAFD, WDC, WDFC, WERN, WEYS, WPPGY, WSB, WSCI, XLNX, XOMA, XRAY, ZBRA, ZIXI, ZOOM

The following stocks were used in the empircal study in chapter 6:

S&P 500 1992-2014: AA, AAPL, ABT, ADBE, ADI, ADM, ADP, ADSK, AEP, AES, AET, AFL, AIG, ALTR, AMAT, AME, AMGN, AON, APA, APC, APD, APH, ARG, AVP, AVY, AXP, AZO, BA, BAC, BAX, BBT, BBY, BCR, BDX, BEN, BHI, BIIB, BK, BLL, BMS, BMY, C, CA, CAG, CAH, CAT, CB, CCE, CCL, CELG, CERN, CI, CINF, CL, CLF, CLX, CMA, CMCSA, CMI, CMS, CNP, COG, COP, COST, CPB, CSC, CSCO, CSX, CTAS, CTL, CVS, CVX, D, DD, DE, DHR, DIS, DNB, DOV, DOW, DTE, DUK, EA, ECL, ED, EFX, EIX, EMC, EMR, EOG, EQT, ESV, ETN, ETR, EXC, EXPD, F, FAST, FDO, FDX, FISV, FITB, FLS, FMC, FTR, GAS, GD, GE, GHC, GIS, GLW, GPC, GPS, GT, GWW, HAL, HAR, HAS,

HBAN, HCP, HD, HES, HOG, HON, HOT, HP, HPQ, HRB, HRL, HRS, HST, HSY, HUM, IBM, IFF, IGT, INTC, IP, IPG, IR, ITW, JCI, JEC, JNJ, JPM, JWN, K, KEY, KIM, KLAC, KMB, KO, KR, KSU, L, LB, LEG, LEN, LH, LLTC, LLY, LM, LMT, LNC, LOW, LRCX, LUK, LUV, MAS, MAT, MCD, MDT, MHFI, MKC, MMC, MMM, MO, MOS, MRK, MRO, MSFT, MSI, MTB, MU, MUR, MWV, MYL, NBL, NBR, NE, NEE, NEM, NI, NKE, NOC, NSC, NTRS, NUE, NWL, OI, OKE, OMC, ORCL, OXY, PAYX, PBCT, PBI, PCAR, PCG, PCL, PCP, PEG, PEP, PFE, PG, PGR, PH, PHM, PKI, PLL, PNC, PNR, PNW, POM, PPG, PPL, PRGO, PSA, PVH, QCOM, R, RDC, REGN, RF, ROK, ROST, RTN, SCG, SCHW, SEE, SHW, SIAL, SLB, SLM, SNA, SO, SPLS, STI, STJ, STT, SWK, SWN, SYK, SYMC, SYY, T, TAP, TE, TEG, TGT, THC, TIF, TJX, TMK, TMO, TROW, TRV, TSN, TSO, TSS, TXN, TXT, TYC, UNH, UNM, UNP, USB, UTX, VAR, VFC, VLO, VMC, VNO, VRTX, VZ, WDC, WEC, WFC, WHR, WM, WMB, WMT, WY, X, XEL, XL, XLNX, XOM, XRAY, XRX, ZION

Appendix C

Comparison of empirical copulas with a Student's t-copula

Here, we compare the empirical pairwise copula densities for each market state with the Student's t-copula density. We estimate the Student's t-copula density $\mathrm{cop}_{\bar{c},\nu}(u,v)$ in the same way as the K-copula density in section 5.4.2

$$\begin{aligned}\mathrm{cop}_{\bar{c},\nu}(u,v) = \mathrm{Cop}_{\bar{c},\nu}(u,v) - \mathrm{Cop}_{\bar{c},\nu}(u,v-\Delta v) - \mathrm{Cop}_{\bar{c},\nu}(u-\Delta u,v) \\ + \mathrm{Cop}_{\bar{c},\nu}(u-\Delta u,v-\Delta v)\ ,\end{aligned} \tag{C.1}$$

where $\mathrm{Cop}_{\bar{c},\nu}(u,v)$ is the Student's t-copula (5.12). We then compute the difference between empirical and analytical copula density for each state

$$\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\bar{c},\nu}(u,v)\ , \qquad i = 1,\ldots,6\ , \tag{C.2}$$

where $\bar{c}$ is the average correlation coefficient of all $K(K-1)/2$ stock pairs for the considered state. The free parameter ν is determined by a fit which minimizes the sum of squared differences between empirical and analytical copula density. The differences between the empirical copula density and the Student's t-copula density for each state are presented in figure C.1 for the original and in figure C.2 for the locally normalized returns. The parameter values are summarized in table C.1. We find an overall good agreement with deviations mainly in the corners of the copula densities. However, the K-copula provides a slightly better fit than the Student's t-copula reflected in the smaller least squares, see tables 5.2 and C.2.

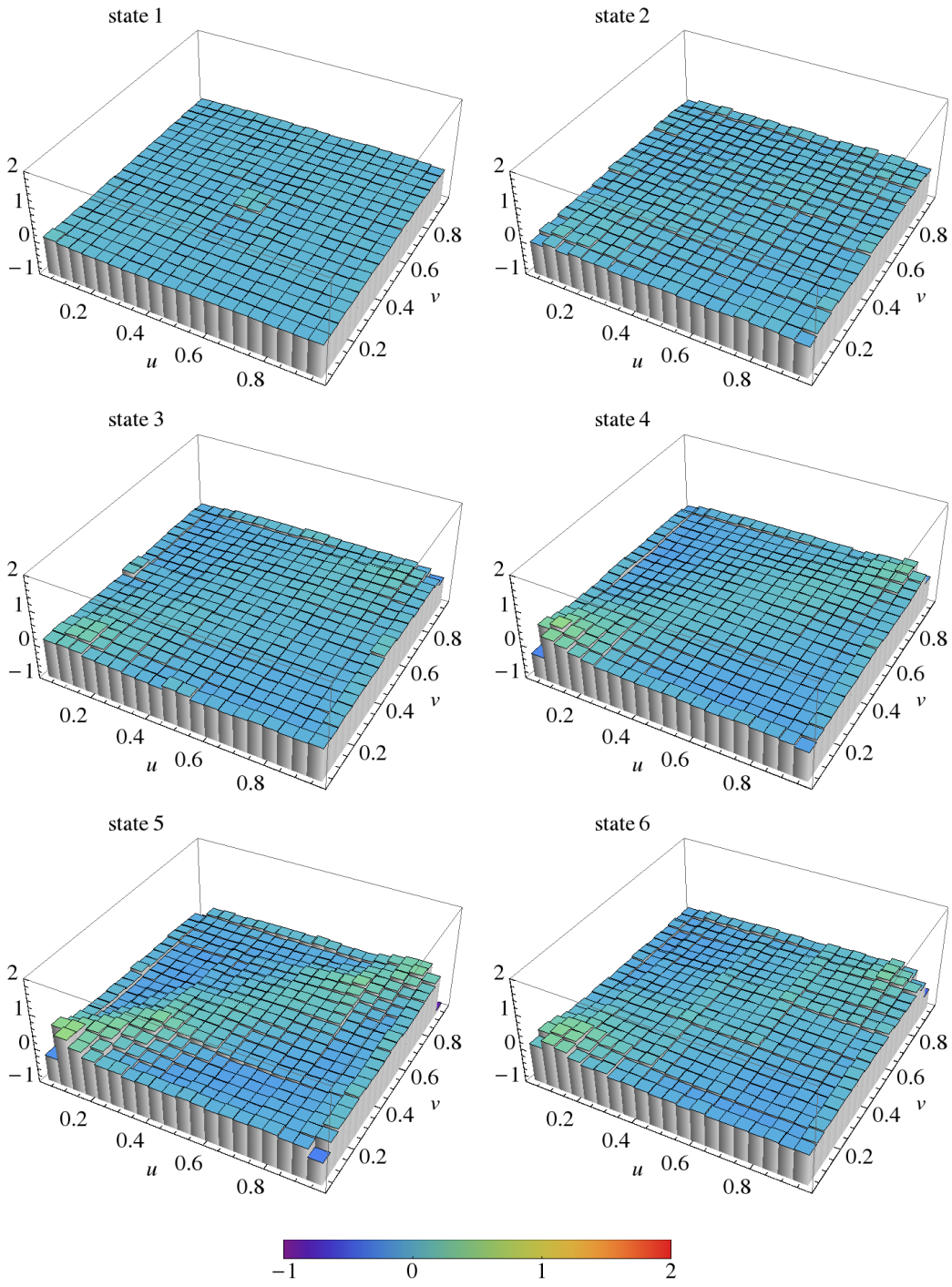


Figure C.1: Difference between the empirical copula density and the Student's t-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\tilde{c},\nu}(u,v)$ for the original returns.

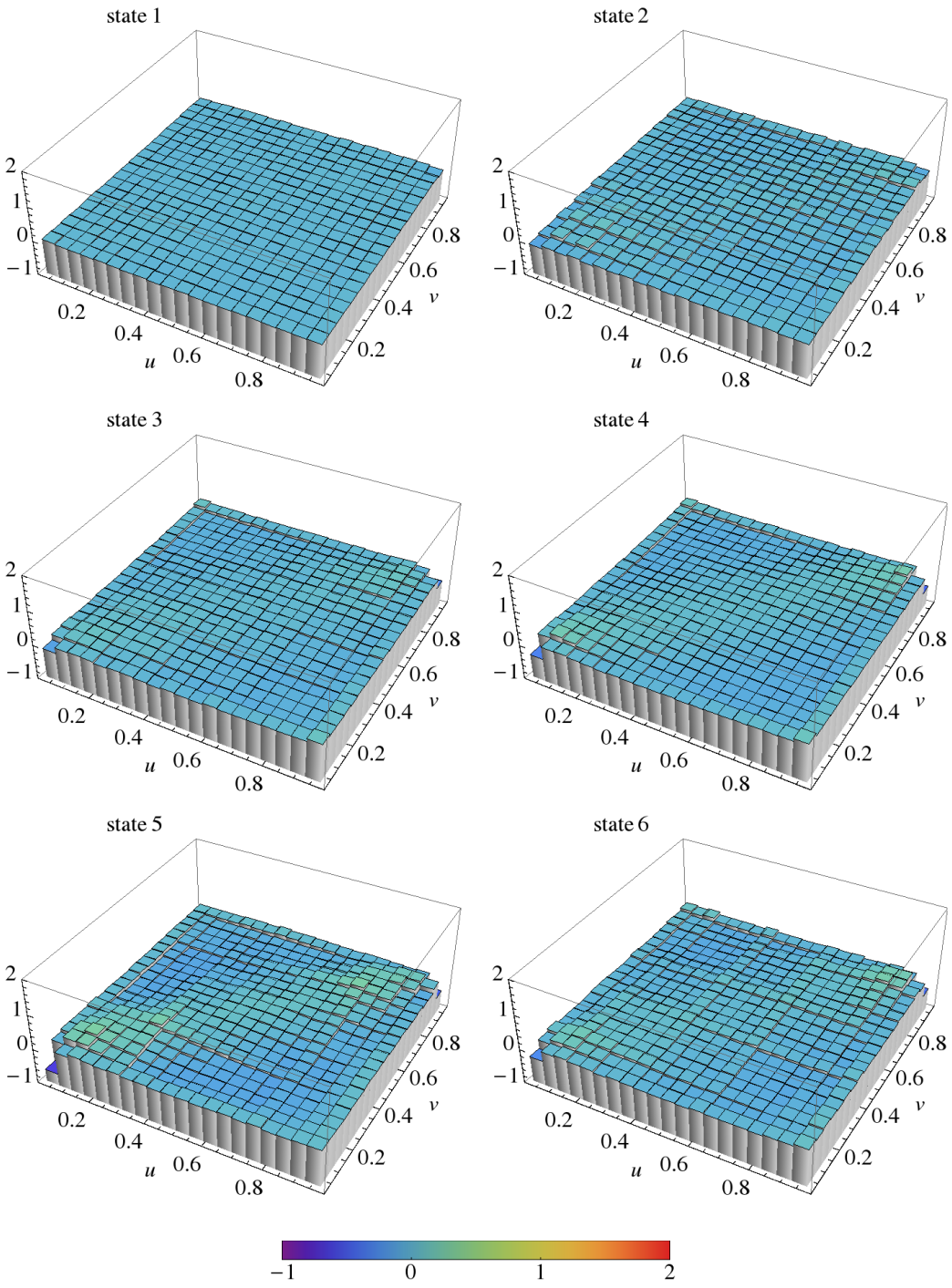


Figure C.2: Difference between the empirical copula density and the Student's t-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\tilde{c},\nu}(u,v)$ for the locally normalized returns.

returns		state 1	state 2	state 3	state 4	state 5	state 6
original	$\bar{c}$	0.046	0.13	0.17	0.25	0.42	0.22
	ν	42.7	13.3	9.8	7.2	3.4	13.2
loc. normalized	$\bar{c}$	0.048	0.13	0.19	0.28	0.43	0.25
	ν	98.0	29.3	33.2	18.7	9.5	24.4

Table C.1: Parameters of the Student's t-copula density for the original and the locally normalized returns.

returns	state 1	state 2	state 3	state 4	state 5	state 6
original	0.11	0.46	1.25	2.99	6.30	3.04
loc. normalized	0.043	0.39	0.72	1.52	3.43	1.54

Table C.2: Least sum of squared differences between empirical and Student's t-copula densities.

Appendix D

Deformation with a beta prime distribution

Here, we consider another example for the deformation function, namely a beta prime distribution

$$h(z) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{z^{a-1}}{(1+z)^{a+b}} \tag{D.1}$$

with the parameters $a, b > 0$. Recently, beta prime distribution was also used by Meudt *et al.* [186] as a return distribution deformation function.

The deformed distribution is calculated in the following way

$$\begin{aligned}
\widehat{q}(x|\Sigma) &= \int_0^\infty \mathrm{d}z\, h(z)\; p\left(x \middle| \frac{\Sigma}{z}\right) \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^\infty \mathrm{d}z\, \frac{z^{a-1}}{(1+z)^{a+b}} \frac{1}{\sqrt{\det 2\pi\Sigma/z}}\; \exp\left(-\frac{z}{2} x^\dagger \Sigma^{-1} x\right) \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\sqrt{\det 2\pi\Sigma}} \int_0^\infty \mathrm{d}z\, \frac{z^{a-1} z^{N/2}}{(1+z)^{a+b}}\; \exp\left(-\frac{z}{2} x^\dagger \Sigma^{-1} x\right)\ .
\end{aligned} \tag{D.2}$$

We substitute

$$\rho = z\, x^\dagger \Sigma^{-1} x \tag{D.3}$$

and obtain

$$\begin{aligned}\widehat{q}(x|\Sigma) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\sqrt{\det 2\pi\Sigma}}\,\frac{1}{(x^\dagger\Sigma^{-1}x)^{a+N/2}}\int\limits_0^\infty \mathrm{d}\rho\,\frac{\rho^{a-1+N/2}\;\mathrm{e}^{-\rho/2}}{\left(1+\frac{\rho}{x^\dagger\Sigma^{-1}x}\right)^{a+b}} \\ &= \frac{2^{N/2-b}\,\Gamma(a+b)\Gamma(a+N/2)}{\Gamma(a)\Gamma(b)\sqrt{\det 2\pi\Sigma}\,(x^\dagger\Sigma^{-1}x)^{N/2-b}}\;\mathcal{U}\left(a+b,1+b-\frac{N}{2},\frac{x^\dagger\Sigma^{-1}x}{2}\right)\;.\end{aligned} \tag{D.4}$$

Here, $\mathcal{U}(x,y,z)$ is the confluent hypergeometric function [187] with the integral representation

$$\mathcal{U}(x,y,z) = \frac{1}{\Gamma(x)}\int\limits_0^\infty \mathrm{d}t\; t^{x-1}(1+t)^{y-x-1}\exp\left(-zt\right) \tag{D.5}$$

for positive real parts of x and z. According to equation (6.64), the maximum distribution is given by

$$\widehat{Q}_N^{\max}(t) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int\limits_0^\infty \mathrm{d}z\;\frac{z^{a-1/2}}{(1+z)^{a+b}}\int\limits_{-\infty}^\infty \mathrm{d}\tilde{s}\,\frac{\mathrm{e}^{-z\tilde{s}^2/2c}}{\sqrt{2\pi c}}\;\Phi^N\left(\frac{\sqrt{z}(t/\sigma-\tilde{s})}{\sqrt{1-c}}\right)\;. \tag{D.6}$$

In addition, we compare the histogram of the empirical maximum values $m_N(t)$ with the maximum pdf

$$\widehat{q}_N^{\max}(t) = \frac{\mathrm{d}\widehat{Q}_N^{\max}(t)}{\mathrm{d}t}\;. \tag{D.7}$$

From equation (6.76), we obtain

$$\Sigma^{(d)} = \frac{b}{a-1}\Sigma\;. \tag{D.8}$$

Thus, we determine the parameters c and σ from $c = c^{(d)}$ and $\sigma = \sigma^{(d)}\sqrt{(a-1)/b}$, where $c^{(d)}$ and $\sigma^{(d)}$ are the average parameters estimated from the data. Here, we have $c^{(d)} = 0.28$ and $\sigma^{(d)} = 0.022$. The remaining parameters a and b are fitted to the empirical distribution. The results are shown in figure D.1. Again, we find a poor agreement between theory and data. Fitting all parameters improves the agreement.

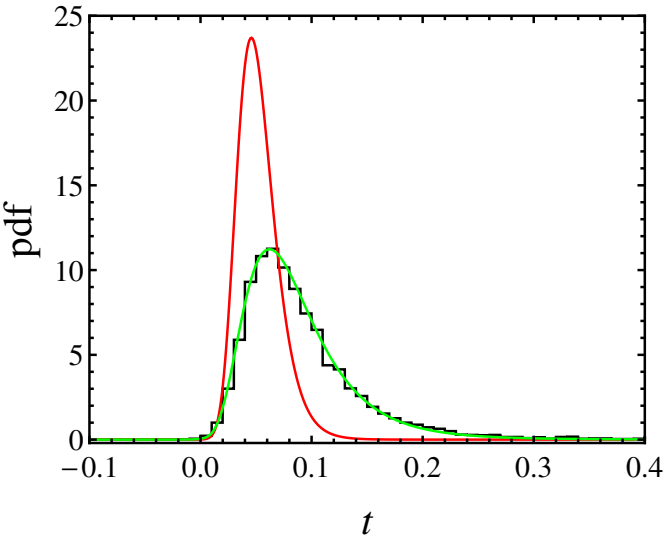


Figure D.1: Histogram of the empirical maximum values (black) compared with the maximum pdf $\widehat{q}_N^{\max}(t)$ for $N = 301$, (red) $\sigma = 0.02$ and $c = 0.28$ estimated from the data, $a = 9$ and $b = 10$ fitted to the data, and (green) $\sigma = 0.042$, $c = 0.37$, $a = 3$ and $b = 4$ fitted to the data.

List of Figures

List of Tables

Bibliography

[1] T. A. Schmitt, D. Chetalova, R. Schäfer, and T. Guhr. Non-stationarity in financial time series: Generic features and tail behavior. *Europhysics Letters*, 103:58003, 2013.

[2] D. Chetalova, T. A. Schmitt, R. Schäfer, and T. Guhr. Portfolio return distributions: Sample statistics with stochastic correlations. *International Journal of Theoretical and Applied Finance*, 18:1550012, 2015.

[3] D. Chetalova, R. Schäfer, and T. Guhr. Zooming into market states. *Journal of Statistical Mechanics*, P01029, 2015.

[4] D. Chetalova, M. Wollschläger, and R. Schäfer. Dependence structure of market states. *Journal of Statistical Mechanics*, P08012, 2015.

[5] T. A. Schmitt, D. Chetalova, R. Schäfer, and T. Guhr. Credit risk and the instability of the financial system: An ensemble approach. *Europhysics Letters*, 105:38004, 2014.

[6] A. Bunde, J. Kropp, and H. J. Schellnhuber. *The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes*. Springer, Berlin, 1999.

[7] Z. Burda, J. Jurkiewicz, and M. A. Nowak. Is econophysics a solid science? *Acta Physica Polonica*, B 34:87–131, 2003.

[8] C. Schinckus. Is econophysics a new discipline? A neopositivist argument. *Physica A*, 389:3814–3821, 2010.

[9] J. D. Farmer, M. Shubik, and E. Smith. Is economics the next physical science? *Physics Today*, 58:37–42, 2005.

[10] L. Bachelier. Theoriè de la spèculation. *Annales Scientifiques de l'Ecole Normale Superieure*, III(17):21–86, 1900.

[11] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 17:549–560, 1905.

[12] M. F. M. Osborne. Brownian motion in the stock market. *Operations Research*, 7:145–173, 1959.

[13] P. A. Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6:13–29, 1965.

[14] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economics*, 81:637–659, 1973.

[15] R. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[16] B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36:394–419, 1963.

[17] R. N. Mantegna and H. E. Stanley. Turbulence and financial markets. *Nature*, 383:587–588, 1996.

[18] R. N. Mantegna and H. E. Stanley. Stock market dynamics and turbulence: parallel analysis of fluctuation phenomena. *Physica A*, 239:255–266, 1997.

[19] D. Chowdhury and D. Stauffer. A generalized spin model of financial markets. *The European Physical Journal B*, 8:477–482, 1999.

[20] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley. Universal and non-universal properties of cross correlations in financial time series. *Physical Review Letters*, 83:1471–1474, 1999.

[21] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397:498–500, 1999.

[22] R. N. Mantegna and H. E. Stanley. Scaling behavior in the dynamics of an economic index. *Nature*, 376:46–49, 1995.

[23] A. Pagan. The econometrics of financial markets. *Journal of Empirical Finance*, 3:15–102, 1996.

[24] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[25] Y. Liu, P. Cizeau, V. Plerou, M. Meyer, C. K. Peng, and H. E. Stanley. Correlations in economic time series. *Physica A*, 245:437–440, 1997.

[26] L. Laloux, P. Cizeau, J. P. Bouchaud, and M. Potters. Noise dressing of financial correlation matrices. *Physical Review Letters*, 83:1467–1470, 1999.

[27] J. D. Noh. Model for correlations in stock markets. *Physical Review E*, 61: 5981, 2000.

[28] P. Cizeau, M. Potters, and J. P. Bouchaud. Correlation structure of extreme stock returns. *Quantitative Finance*, 1:217–222, 2001.

[29] B. LeBaron. Agent based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, 24:679–702, 2000.

[30] M. Levy, H. Levy, and S. Solomon. *Microscopic Simulation of Financial Markets*. Academic Press, New York, 2000.

[31] B. LeBaron. A builder's guide to agent based financial markets. *Quantitative Finance*, 1:254–261, 2001.

[32] E. Aurell, R. Baviera, O. Hammarlid, M. Serva, and A. Vulpiani. A general methodology to price and hedge derivatives in incomplete markets. *International Journal of Theoretical and Applied Finance*, 3:1–24, 2000.

[33] J. P. Bouchaud and M. Potters. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. Cambridge University Press, Cambridge, 1999.

[34] G. Montagna, O. Nicrosini, and A. Moreni. A path integral way of option pricing. *Physica A*, 310:450–466, 2002.

[35] M. Marsili, S. Maslov, and Y.-C. Zhang. Dynamical optimization theory of a diversified portfolio. *Physica A*, 253:403–418, 1998.

[36] D. Sornette. Large deviations and portfolio optimization. *Physica A*, 261: 251–283, 1998.

[37] S. Sharifi, M. Crane, A. Shamaie, and H. Ruskin. Random matrix theory for portfolio optimization: a stability approach. *Physica A*, 335:629–643, 2004.

[38] J. J. Ramsden and G. Kiss-Haypal. Company size distribution in different countries. *Physica A*, 277:220–227, 2000.

[39] R. L. Axtell. Zipf distribution of U.S. firm sizes. *Science*, 293:1818–1820, 2001.

[40] A. Dragulescu and V. M. Yakovenko. Statistical mechanics of money. *The European Physical Journal B*, 17:723–729, 2000.

[41] A. Dragulescu and V. M. Yakovenko. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. *Physica A*, 299:213–221, 2001.

[42] V. M. Yakovenko and J. B. Rosser. Colloquium: Statistical mechanics of money, wealth, and income. *Reviews of Modern Physics*, 81:1703, 2009.

[43] A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel. Econophysics review: I. Empirical facts. *Quantitative Finance*, 11:991–1012, 2011.

[44] A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel. Econophysics review: II. Agent-based models. *Quantitative Finance*, 11:1013–1041, 2011.

[45] R. N. Mantegna and H. E. Stanley. *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge, 1999.

[46] W. Paul and J. Baschnagel. *Stochastic Processes From Physics to Finance*. Springer, Berlin, 1999.

[47] J. Voit. *The Statistical Mechanics of Financial Markets*. Springer, Berlin, 2005.

[48] N. Wiener. Differential space. *Journal of Mathematics and Physics*, 2:131–174, 1923.

[49] W. C. Mitchell. The making and using of index numbers. *Bulletin of the US Bureau of Labor Statistics*, 173, 1915.

[50] M. Oliver. *Les nombres indices de la variation des prix*. M. Giard, Paris, 1927.

[51] F. C. Mills. The behavior of prices. *National Bureau of Economic Research*, 1927.

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

[53] E. Fama. The behavior of stock market prices. *The Journal of Business*, 38: 34–105, 1965.

[54] R. C. Blattberg and N. Gonedes. A comparison of the stable and student distributions as models for stock prices. *The Journal of Business*, 47:244–280, 1974.

[55] A. Peiro. The distribution of stock returns: international evidence. *Applied Financial Economics*, 4:431–439, 1994.

[56] S. Kon. Models of stock returns – a comparison. *The Journal of Finance*, 39: 147–165, 1984.

[57] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1: 289–299, 1995.

[58] O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and modeling of stock returns. *Scandinavian Journal of Statistics*, 24:1–13, 1997.

[59] F. Black. Studies of stock price volatility changes. In *1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, pages 177–181, Washington DC, 1976. American Statistical Association.

[60] G. W. Schwert. Why does stock market volatility change over time? *The Journal of Finance*, 44:1115–1153, 1989.

[61] R. Cont. Volatility clustering in financial markets : Empirical facts and agent-based models. In *Long Memory in Economics*, pages 289–309, London, 2007. Springer.

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

[63] S. J. Taylor. Financial returns modelled by the product of two stochastic processes – a study of daily sugar prices 1961-79. In *Time Series Analysis: Theory and Practice 1*, pages 203–226, Amsterdam: North-Holland, 1982. Anderson, O. D.

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

[65] L. Bauwens, C. M. Hafner, and S. Laurent. *Handbook of Volatility Models and Their Applications.* Wiley, Chichester, 2012.

[66] G. Bekaert and C. R. Harvey. Time-varying world market integration. *The Journal of Finance*, L:403–44, 1995.

[67] F. M. Longin and B. Solnik. Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance*, 14:3–26, 1995.

[68] D. J. Fenn, M. A. Porter, S. Williams, M. McDonald, N. F. Johnson, and N. S. Jones. Temporal evolution of financial-market correlations. *Physical Review E*, 84:026109, 2011.

[69] M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, and H. E. Stanley. Identifying states of a financial market. *Scientific Reports*, 2 644, 2012.

[70] L. Laloux, P. Cizeau, and M. Potters. Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3: 391–397, 2000.

[71] T. Guhr and B. Kälber. A new method to estimate the noise in financial correlation matrices. *Journal of Physics A*, 36:3009–3032, 2003.

[72] R. Schäfer and T. Guhr. Local normalization: Uncovering correlations in non-stationary financial time series. *Physica A*, 389:3856–3865, 2010.

[73] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7:77–91, 1952.

[74] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20 A:32–52, 1928.

[75] E. Wigner. Characteristic vectors of bordered matrices with infinite dimension. *Annals of Mathematics*, 62:548–564, 1955.

[76] F. J. Dyson. Statistical theory of the energy levels of complex systems I. *Journal of Mathematical Physics*, 3:140, 1962.

[77] F. J. Dyson. Statistical theory of the energy levels of complex systems II. *Journal of Mathematical Physics*, 3:157, 1962.

[78] F. J. Dyson. Statistical theory of the energy levels of complex systems III. *Journal of Mathematical Physics*, 3:166, 1962.

[79] F. J. Dyson. The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. *Journal of Mathematical Physics*, 3:1200, 1962.

[80] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley. Random matrix approach to cross correlations in financial data. *Physical Review E*, 65:066126, 2002.

[81] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller. Random-matrix theories in quantum physics: common concepts. *Physics Reports*, 299:189–425, 1998.

[82] F. Mezzadri and N. C. Snaith. *Recent Perspectives in Random Matrix Theory and Number Theory*. Cambridge University Press, Cambridge, 2005.

[83] R. Couillet and M. Debbah. *Random Matrix Methods for Wireless Communications*. Cambridge University Press, Cambridge, 2011.

[84] K. Rajan and L. F. Abbott. Eigenvalue spectra of random matrices for neural networks. *Physical Review Letters*, 97:188104, 2006.

[85] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley. A random matrix theory approach to financial cross correlations. *Physica A*, 287:374–382, 2000.

[86] M. Potters, J. P. Bouchaud, and L. Laloux. Financial applications of random matrix theory: Old laces and new pieces. *Acta Physica Polonica B*, 36: 2767–2784, 2005.

[87] S. Drozdz, J. Kwapien, and P. Oswiecimka. Empirics versus RMT in financial cross-correlations. *Acta Physica Polonica B*, 58:4027–4039, 2007.

[88] A. M. Sengupta and P. P. Mitra. Distributions of singular values for some random matrices. *Physical Review E*, 60:3389, 1999.

[89] S. Pafka and I. Kondor. Estimated correlation matrices and portfolio optimization. *Physica A*, 343:623–634, 2004.

[90] S. Pafka, M. Potters, and I. Kondor. Exponential weighting and random-matrix-theory-based filtering of financial covariance matrices for portfolio optimization, 2004. arXiv: cond-mat/0402573.

[91] J. Daly, M. Crane, and H. J. Ruskin. Random matrix theory filters and currency portfolio optimization. *Journal of Physics: Conference Series*, 221: 012003, 2010.

[92] M. C. Münnix, R. Schäfer, and T. Guhr. A random matrix approach to credit risk. *PLoS ONE*, 9:e98030, 2014.

[93] R. F. Engle. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics*, 20:339–350, 2002.

[94] Y. K. Tse and A. K. C. Tsui. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business and Economic Statistics*, 20:351–362, 2002.

[95] C. van Emmerich. Modeling correlation as a stochastic process, 2006. Technical report, Bergische Universität Wuppertal.

[96] J. Ma. Pricing foreign equity options with stochastic correlation and volatility. *Annals of Economics and Finance*, 10:303–327, 2009.

[97] V. Golosnoy, B. Gribisch, and R. Liesenfeld. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167:211–223, 2002.

[98] A. G. Akritas, E. K. Akritas, and G. I. Malaschonok. Various proofs of Sylvester's (determinant) identity. *Mathematics and Computers in Simulation*, 42:585–593, 1996.

[99] S. D. Dubey. Compound gamma, beta and F distributions. *Metrika*, 16: 27–31, 1970.

[100] O. Barndorff-Nielsen, J. Kent, and M. Sørensen. Normal variance-mean mixtures and z distributions. *International Statistical Review*, 50:145–159, 1982.

[101] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 2007.

[102] S. Cambanis, S. Huang, and G. Simons. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11:368–385, 1981.

[103] K. T. Fang, S. Kotz, and K. W. Ng. *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London, 1990.

[104] W. Breymann, A. Dias, and P. Embrechts. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, 3:1–14, 2003.

[105] R. Mashal, M. Naldi, and A. Zeevi. Comparing the dependence structure of equity and asset returns. *Risk*, 16:82–87, 2003.

[106] Yahoo! Finance. URL: http://finance.yahoo.com.

[107] W. C. Parr. Minimum distance estimation: A bibliography. *Communications in Statistics-Theory and Methods*, A10:1205–1224, 1981.

[108] Ö. Öztürk and T. P. Hettmansperger. Generalized weighted Cramer-von Mises distance estimators. *Biometrica*, 84:283–294, 1997.

[109] T. W. Anderson and D. A. Darling. A test of goodness-of-fit. *Journal of the American Statistical Association*, 49:765–769, 1954.

[110] D. D. Boos. Minimum Anderson-Darling estimation. *Communications in Statistics-Theory and Methods*, 11:2747–2774, 1982.

[111] H. Schaller and S. Van Norden. Regime switching in stock market returns. *Applied Financial Economics*, 7:177–191, 1997.

[112] M. Marsili. Dissecting financial markets: Sectors and states. *Quantitative Finance*, 2:297–302, 2002.

[113] D.-M. Song, M. Tumminello, W.-X. Zhou, and R. N. Mantegna. Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Physical Review E*, 84:026108, 2011.

[114] G. Buccheri, S. Marmi, and R. N. Mantegna. Evolution of correlation structure of industrial indices of U.S. equity markets. *Physical Review E*, 88:012806, 2013.

[115] L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, New York, 1990.

[116] R. Tibshirani, G. Walther, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society B*, 63:411–423, 2001.

[117] E. P. Wigner. Random matrices in physics. *SIAM Review*, 9:1–23, 1967.

[118] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong. Random-matrix physics: Spectrum and strength fluctuations. *Reviews of Modern Physics*, 53:385–480, 1981.

[119] A. Sklar. Fonctions de rèpartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Universitè de Paris*, 8:229–231, 1959.

[120] A. Sklar. Random variables, joint distribution functions, and copulas. *Kybernetika*, 9:449–460, 1973.

[121] P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependency in risk management: Properties and pitfalls. In *Value at Risk and Beyond*, New York, 2002. Cambridge University Press.

[122] P. Embrechts, F. Lindskog, and A. McNeil. Modelling dependence with copulas and application to risk management. In *Handbook of heavy tailed distributions in finance*, Amsterdam, 2003. Elsevier.

[123] A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ, 2005.

[124] J. Rosenberg and T. Schuermann. A general approach to integrated risk management with skewed fat-tailed risks. *Journal of Financial Economics*, 79:569–614, 2006.

[125] E. Kole, K. Koedijk, and M. Verbeek. Selecting copulas for risk management. *Journal of Finance and Banking*, 31:2405–2423, 2007.

[126] D. Brigo, A. Pallavicini, and R. Torresetti. *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models*. Wiley Finance, Chichester, 2010.

[127] A. Meucci. A new breed of copulas for risk and portfolio management. *Risk*, 24:122–126, 2011.

[128] J. V. Rosenberg. Nonparametric pricing of multivariate contingent claims. *The Journal of Derivatives*, 10:9–26, 2003.

[129] M. N. Bennett and J. E. Kennedy. Quanto pricing with copulas. *The Journal of Derivatives*, 12:26–45, 2004.

[130] U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance*. Wiley, Chichester, 2004.

[131] R. W. J. van den Goorbergh, C. Genest, and B. J. M. Werker. Multivariate option pricing using using dynamic copula models. *Insurance: Mathematics and Economics*, 37:101–114, 2005.

[132] J. Hull and A. White. Valuing credit derivatives using an implied copula approach. *The Journal of Derivatives*, 14:8–28, 2006.

[133] M. Hofert and M. Scherer. CDO pricing with nested Archimedean copulas. *Quantitative Finance*, 11:775–787, 2011.

[134] D. A. Hennessy and H. E. Lapan. The use of Archimedean copulas to model portfolio allocations. *Mathematical Finance*, 12:143–154, 2002.

[135] A. J. Patton. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Economics*, 2:130–168, 2004.

[136] A. Di Clemente and C. Romano. Measuring and optimizing portfolio credit risk: a copula-based approach. *Economic Notes*, 33:325–367, 2004.

[137] H. Boubaker and N. Sghaier. Portfolio optimization in the presence of dependent financial returns with long memory: A copula based approach. *Journal of Banking and Finance*, 37:361–377, 2013.

[138] C. Genest, M. Gendron, and M. Bourdeau-Brien. The advent of copulas in finance. *The European Journal of Finance*, 15:609–618, 2009.

[139] A. J. Patton. Handbook of Economic Forecasting. In *Copula-Based Models for Financial Time Series*, pages 767–785, Berlin, 2012. Springer.

[140] T. Mikosch. Copulas: tales and facts. *Extremes*, 9:3–20, 2006.

[141] C. Genest and B. Rémillard. Discussion of "Copulas: tales and facts". *Extremes*, 9:27–36, 2006.

[142] H. Joe. Discussion of "Copulas: tales and facts". *Extremes*, 9:37–41, 2006.

[143] F. M. Longin and B. Solnik. Extreme correlation of international equity markets. *The Journal of Finance*, 56:649–676, 2001.

[144] A. Ang and J. Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63:443–494, 2002.

[145] Y. Hong, J. Tu, and G. Zhou. Asymmetries in stock returns: statistical tests and economic evaluation. *The Review of Financial Studies*, 20:1547–1581, 2007.

[146] H. Joe. *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall, London, 1997.

[147] R. B. Nelsen. *An Introduction to Copulas*. Springer, New York, 2006.

[148] Y. Fan and A. J. Patton. Copulas in econometrics. *Annual Review of Economics*, 6:179–200, 2014.

[149] M. Wollschläger and R. Schäfer. Impact of non-stationarity on estimating and modeling empirical copulas of daily stock returns, 2015. arXiv:1506.08054.

[150] S. Demarta and A. McNeil. The t copula and related copulas. *International Statistical Review*, 73:111–129, 2005.

[151] W. Sun, S. V. Stoyanov, S. Rachev, and F. J. Fabozzi. Multivariate skewed students's t copula in the analysis of nonlinear and asymmetric dependence in the German equity market. *Studies in Nonlinear Dynamics and Econometrics*, 12, 2008.

[152] M. Ammann and S. Süss. Asymmetric dependence patterns in financial time series. *The European Journal of Finance*, 15:703–719, 2009.

[153] M. R. Fréchet. Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*, 6:93, 1927.

[154] R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180–190, 1928.

[155] B. V. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, 44:423–453, 1943.

[156] E. J. Gumbel. *Statistics of Extremes*. Columbia University Press, New York, 1958.

[157] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.

[158] S. N. Majumdar and A. Comtet. Airy distribution function: From the area under a Brownian excursion to the maximal height of fluctuating interfaces. *Journal of Statistical Physics*, 119:777–826, 2005.

[159] D. S. Dean and S. N. Majumdar. Extreme-value statistics of hierarchically correlated variables deviation from Gumbel statistics and anomalous persistence. *Physical Review E*, 64:046121, 2001.

[160] S. N. Majumdar and A. Comtet. Exact maximal height distribution of fluctuating interfaces. *Physical Review Letters*, 92:225501, 2004.

[161] S. N. Majumdar and A. Pal. Extreme value statistics of correlated random variables, 2015. arXiv:1406.6768.

[162] T. Guhr. Extreme value statistics of correlated random variables, 2015. Unpublished notes.

[163] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer, London, 2001.

[164] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York, 1978.

[165] M. R. Leadbetter, G. Lindgren, and H. Rootzen. *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York, 1982.

[166] L. De Haan and A. Ferreira. *Extreme Value Theory*. Springer, New York, 2006.

[167] R. von Mises. La distribution de la plus grande de n valeurs. *American Mathematical Society*, 2:271–294, 1954.

[168] A. F. Jenkinson. The frequency distribution of the annual maximum (or minimum) values of meteorological events. *Quarterly Journal of the Royal Meteorological Society*, 81:158–172, 1955.

[169] P. Hall. On the rate of convergence of normal extremes. *Journal of Applied Probability*, 16:443–439, 1979.

[170] W. J. Hall and J. A. Wellner. The rate of convergence in law of the maximum of an exponential sample. *Statistica Neerlandica*, 33:151–154, 1979.

[171] R. L. Smith. Uniform rates of convergence in extreme value theory. *Advances in Applied Probability*, 14:600–662, 1982.

[172] R. Davis. The rate of convergence in distribution of the maxima. *Statistica Neerlandica*, 36:31–25, 1982.

[173] E. Omey and S. T. Rachev. On the rate of convergence in extreme value theory. *Theory of Probability and Its Applications*, 33:601–607, 1988.

[174] A. A. Balkema and L. De Haan. A convergence rate in extreme value theory. *Journal of Applied Probability*, 27:577–585, 1990.

[175] S. I. Resnik. *Extreme Values, Regular Variation, and Point Processes*. Springer, New York, 2007.

[176] R. Takahashi. Normalizing constants of a distribution which belongs to the domain of attraction of the Gumbel distribution. *Statistics & Probability Letters*, 5:197–200, 1987.

[177] S. S. Gupta. Probability integrals of multivariate normal and multivariate t. *Annals of Mathematical Statistics*, 34:792–828, 1963.

[178] S. S. Gupta, K. Nagel, and S. Panchapakesan. On the order statistics of equally correlated normal random variables. *Biometrica*, 60:403–413, 1973.

[179] M. Greig. Extremes in a random assembly. *Biometrica*, 54:273–282, 1967.

[180] T. R. Hoffmann and J. G. Saw. Distribution of the largest of a set of equicorrelated normal variables. *Communications in Statistics*, 4:49–53, 1975.

[181] J. O. Rawlings. Order statistics for a special class of unequally correlated multinormal variates. *Biometrics*, 32:875–887, 1976.

[182] W. G. Hill. Order statistics of correlated variables and implications in genetic selection programms. *Biometrics*, 32:889–902, 1976.

[183] K. K. Saha and B. C. Sutradhar. On the distribution of the extremes of unequally correlated normal variables with applications to antedependent cluster data. *Annals of the Institute of Statistical Mathematics*, 51:301–322, 1999.

[184] S. S. Gupta, S. Panchapakesan, and J. K. Sohn. On the distribution of the studentized maximum of equally correlated normal random variables. *Communications in Statistics-Simulation and Computation*, 14:103–135, 1985.

[185] R. B. Arellano-Valle and M. G. Genton. On the exact distribution of the maximum of absolutely continuous dependent random variables. *Statistics & Probability Letters*, 78:27–35, 2008.

[186] F. Meudt, M. Theissen, R. Schäfer, and T. Guhr. Constructing analytically tractable ensembles of non-stationary covariances with an application to financial data, 2015. arXiv:1503.01584.

[187] M. Abramowitz and I. A. Stegun. Confluent Hypergeometric Functions. In *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, pages 503–515, New York, 1972. Dover.