Credit Portfolio Modelling with Elliptically Contoured Distributions
Approximation, Pricing, Dynamisation

Dissertation
zur Erlangung des Doktorgrades
Dr. rer. nat.
der Fakultät für Mathematik und Wirtschaftswissenschaften
der Universität Ulm

vorgelegt von
Dipl.-Math. oec. Clemens Prestele
aus Erlangen

Ulm, 2007
Amtierender Dekan:  Professor Dr. Frank Stehling

1. Gutachter:  Professor Dr. Rüdiger Kiesel, Universität Ulm
2. Gutachter:  Professor Dr. Ulrich Stadtmüller, Universität Ulm

My ventures are not in one bottom trusted,  
Nor to one place; nor is my whole estate  
Upon the fortune of this present year;  
Therefore my merchandise makes me not sad.

The Merchant of Venice, Act 1 Scene 1  
by William Shakespeare
Acknowledgements

First of all, I wish to express my sincere appreciation to Prof. Dr. Rüdiger Kiesel for offering me the opportunity to write this thesis under his supervision at the Institute of Mathematical Finance. I am very thankful for his great guidance, valuable suggestions and comprehensive support during the years of my doctoral research. I always enjoyed to contribute to the teaching duties and the “life” at the institute and highly appreciated the encouragement and financial assistance e.g. for taking part in seminars and conferences.

I extend my great gratitude to Prof. Dr. Ulrich Stadtmüller of the Institute of Number Theory and Probability Theory for his readiness to advise the thesis as second assessor. He has raised my interest in probability theory during his challenging and motivating lecture in stochastics, which I attended early in my studies, and I am thankful for the beneficial knowledge I gained from assisting him in a course in probability theory.

Grateful thoughts go to Prof. Dr. Nick H. Bingham for proofreading my English\footnote{Of course, all the remaining errors are entirely the author’s responsibility.} and for his valuable comments and suggestions. I want to express my warmest regards to Prof. Dr. Frank Stehling for his great cordiality and for numerous precious discussions, and to Prof. Dr. Werner Kratz for having been a great teacher and mentor to me from the early beginning of my studies.

I would like to thank my friends and colleagues Dr. Björn Böttcher, Gregor Mummenhoff, Dr. Hartmut Lanzinger, Dr. Martin Riesner, Matthias Lutz, Dr. Matthias Scherer, Monika Thalmaier, Dr. Peter N. Posch, Reik Börger, Sebastian Singer and Dr. Stefan Kassberger for great discussions, wonderful seminars and for their valuable assistance whenever I asked them for advice.

I want to express my special gratitude to my parents, my sister Sabine and my twin brother Benjamin for their loving empathy, continual support and perpetual encouragement throughout the entire time of my studies.

Finally, I am very grateful for the scholarship, which I received from the federal state of Baden-Württemberg (LGFG Baden-Württemberg) during the time of my doctoral studies.

Ulm, 17. September 2007

Clemens Prestele
# Contents

1 Introduction ............................................. 5
   1.1 Recent developments of the credit derivatives markets ................. 5
   1.2 Collateralized Debt Obligations .................................... 6
   1.3 Overview over the existing literature .................................. 7
   1.4 The contribution and the aim of this thesis .............................. 11
   1.5 Organization of this thesis ........................................... 17

2 Collateralized Debt Obligations and the pricing problem .............. 19
   2.1 The basic structure of a CDO ......................................... 19
   2.2 Challenges with modelling and pricing CDOs .......................... 22

3 Fundamental theoretical concepts ....................................... 25
   3.1 Elliptical distributions ............................................... 25
   3.2 Regular variation and Karamata’s Theorem ............................ 39
   3.3 Copulae ............................................................... 44
   3.4 A limit theorem for martingales ..................................... 48

4 Credit portfolio models ........................................... 51
   4.1 Basic setup of structural models .................................... 51
   4.2 The Merton model .................................................... 52
   4.3 A multivariate Merton model ....................................... 53
## Contents

4.4 One-period credit portfolio models ............................................ 54

4.5 One-period factor models ..................................................... 55
  4.5.1 Gaussian factor models: Vasicek, KMV and CreditMetrics ............ 55
  4.5.2 Gaussian one-factor model, implied and base correlations .......... 58
  4.5.3 Double $t$-distribution copula model .................................. 61
  4.5.4 Normal Inverse Gaussian factor model ................................ 61

5 An elliptical distributions credit portfolio model .......................... 63
  5.1 Introduction ................................................................. 63
  5.2 The setup - extension of the Gaussian case ............................... 66
  5.3 Introducing tail dependence in the elliptical distributions model .... 72
    5.3.1 Tail dependence in mixtures of Normal distributions .............. 73
    5.3.2 Examples for mixtures of Normal distributions .................... 81
    5.3.3 Effects of the distributions on the dependence structure ......... 88
    5.3.4 Tail dependence in linear factor models .......................... 90

6 Large portfolio approximation in the elliptical distributions model .... 93
  6.1 Credit ratings and credit losses ........................................ 94
    6.1.1 Credit ratings and rating thresholds ................................ 94
    6.1.2 Losses due to rating migrations and defaults .................... 96
  6.2 Large portfolio approximation ............................................ 97
  6.3 First example of the large portfolio approximation .................... 102

7 Application to the valuation of credit derivatives ....................... 107
  7.1 The modelling of the CDO structure .................................... 108
  7.2 The valuation of a CDO .................................................. 109
  7.3 Defaults and their probabilities ....................................... 112
7.4 Approximation of the portfolio losses ................................................. 113
7.5 Key quantities revisited within the models based on mixture distributions 117
  7.5.1 Default probabilities and default thresholds ............................... 117
  7.5.2 Expected tranche losses ......................................................... 120
  7.5.3 The number of defaults ......................................................... 123
7.6 Application of the models on CDO data ........................................... 124
  7.6.1 Sequence of computations for the valuation of CDOs .................... 124
  7.6.2 The data basis we use .......................................................... 124
  7.6.3 Calibrating the models ......................................................... 125
  7.6.4 Simulation study ................................................................. 126
  7.6.5 Performance of the models on the iTraxx data ............................ 129

8 Dynamic elliptical distributions model .............................................. 139
  8.1 Introduction .................................................................................. 139
  8.2 Dynamic setup of elliptical distributions factor model .................... 140
  8.3 Effects of the dynamical setup ..................................................... 141
  8.4 The discrete-time case .................................................................. 144
    8.4.1 Employing time-series models for the scaling process ............... 145
    8.4.2 Applying discrete-time volatility models .................................. 151
    8.4.3 Continuous-time results for the discrete-time case .................... 155
  8.5 The continuous-time case ............................................................. 158
    8.5.1 Continuous-time interest-rate models for the scaling process .... 158
    8.5.2 Dynamic setup via time-changed Brownian motions .................. 164
    8.5.3 Dynamic setup via subordinated Lévy processes ....................... 168
    8.5.4 Construction of a Lévy process with the Exp-Exp Law? ............ 172
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Supplementary lemma</td>
<td>179</td>
</tr>
<tr>
<td>A Supplementary data</td>
<td>181</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>183</td>
</tr>
<tr>
<td>Bibliography</td>
<td>185</td>
</tr>
<tr>
<td>List of Tables</td>
<td>193</td>
</tr>
<tr>
<td>List of Figures</td>
<td>195</td>
</tr>
<tr>
<td>Zusammenfassung</td>
<td>197</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Recent developments of the credit derivatives markets

In recent years, derivative securities have rapidly gained in importance among the participants of the various financial markets. This is especially true for the rather young market of credit derivatives, which has experienced an enormous market growth over the last years. In its Credit Derivatives Report 2006\(^1\), the British Bankers’ Association (BAA) revealed that while the global market of credit derivatives was at $180 billion\(^2\) in 1996, and first surpassed the $1 trillion threshold in the year 2001, it was estimated to surge to slightly over $20 trillion at the end of 2006 and is expected to reach $33 trillion by the year 2008. The BBA’s report showed that the dominance of the single-name Credit Default Swaps, in short CDSs, has decreased from 51% of the overall credit derivatives market to around 30% in 2006, with a tendency to drop slightly further until 2008, while the market of synthetic Collateralized Debt Obligations, in short CDOs, has remained constant at around 16% in the last years, and is also expected to stay at this level until 2008. However, as the overall market rapidly develops, the market for multi-name credit derivatives and especially the market for Collateralized Debt Obligations has also experienced an enormous growth.

After the first CDOs had been structured in the year 1987, CDOs started to gain popularity in the late 1990s before finally establishing themselves very well in the credit derivatives market in the early half of the first decade of this century. While in the year 2004 there had been global CDO sales reaching almost $231 billion worldwide, which includes both the issuance of cash CDOs (at $107 billion) and of synthetic CDOs (at $125 billion), these figures have risen to $511 billion for the year 2005 (cash CDOs at $191 billion and synthetic CDOs at $320 billion) and have reached the enormous amount of $994 billion in the year 2006 (cash CDOs at $470 billion and synthetic CDOs at $524

\(^1\) The Credit Derivatives Report 2006 was published in September 2006 by the British Bankers’ Association. A executive summary of this report is freely available at www.bba.org.uk.

\(^2\) One billion equals \(1 \times 10^9\), one trillion equals \(1 \times 10^{12}\).
1.2. Collateralized Debt Obligations

Collateralized Debt Obligations belong to the class of portfolio credit derivatives. As such they are derivative securities on a portfolio of credit-risky assets such as loans, bonds or on a portfolio of single-name credit derivatives. After having invested in credit-risky bonds or after having extended loans to some counterparties (e.g. a company such as Daimler Chrysler), a bank might decide to restructure the risk inherent in this portfolio of credit-risky assets with a CDO structure and to sell relevant portions of the risk to financial investors interested in engaging themselves with credit risk. These investors could be commercial banks, asset managers, hedge funds or insurance companies, which might seek for arbitrage opportunities, profitable leveraged returns, or for a new type of highly rated investment if e.g. law only permits one to engage oneself in investment-grade assets. On the other hand, the reasons for banks or other institutions to issue CDOs can also be manifold. They range from exploiting arbitrage opportunities, over the desire to shrink the balance sheets, to the aim of reducing the required regulatory or the economic capital. Therefore, there are also many different ways in which such a CDO restructuring can take place, depending on the aim of this restructuring or securitization process, and usually also on the risk appetite of the investors to whom the slices of the credit risk in the underlying portfolio are sold. There are also numerous taxation, accounting and legal issues involved in the choice of the setup.

In principle, there are two ways in which an investor that holds a portfolio of credit-risky assets can dispose himself of the credit risk inherent in the underlying assets. He can either sell the assets themselves to another investor or can use credit derivatives to transfer only the credit risk, but keep the bonds, loans or other credit-risky assets in his portfolio. The same holds with the structuring of CDOs. Often, the investor does not create the CDO tranches himself, but transfers the credit risk inherent in the assets on a single-name basis to another company, the so-called Special Purpose Vehicle (SPV), either by actually selling the assets to the SPV or by transferring the risk synthetically with the use of Credit Default Swaps. The construction of CDO tranches on the basis of these single-name CDS contracts is then called a synthetic CDO, which we will focus on in this thesis for simplicity.

---

3 These figures were taken from Chapter VI. in [Ban07] and the associated statistical data sets. They account for the CDO sales which have been settled in US dollars, as well as those agreed on in euros.
4 See for example [Luc01] and [FAA + 04] for a detailed discussion of CDO structures.
1.3. Overview over the existing literature

Historically, CDO sales have belonged to the over-the-counter market, as the large variety of possible CDO structures makes a standardisation of CDO contracts a difficult task. Such a standardisation is needed for the purpose of trading these products on stock exchanges, as there are many ways in which CDOs can be structured and as in principle one can choose any kind of underlying portfolio to structure a CDO upon. Only in 1996 the International Swaps and Derivatives Association (ISDA) published the first definition of a credit derivatives contract, which was altered and extended in 1999.\(^5\)

In 2004 the International Index Company (IIC) started with the so-called *iTraxx CDS indices*, which were introduced to set market standards and thus to increase the transparency and the efficiency of the credit derivatives market. They especially aim for European and Asian markets, and the goal is to attract further market participants as a result of improved market conditions. Having been founded in 2001 as a joint venture by several large investment banks, the IIC claims to be an “independent index supplier” for the fixed income, credit derivatives and FX market.\(^6\) More than 35 large banks\(^7\) have become licensed market makers for the *iTraxx Europe Indices*, and therefore contribute to the indices' popularity and liquidity. A CDS index is based on a portfolio of 125 very liquid CDS reference entities, and for each portfolio there is a CDS index with maturities of 3, 5, 7 and 10 years. The reference portfolio for the *iTraxx Europe CDS indices* for example consists of 125 European CDS entities, whose index membership is decided on the basis of liquidity rankings established by the market makers. Based on the reference CDS portfolio, the IIC has introduced and set standards for several derivative products such as options or futures on the *iTraxx* indices, as well as CDO tranches on the entire CDS portfolio. This so-called *tranced iTraxx* was the first standardised CDO to be introduced in the European markets, it consists of 5 tranches and is available with maturities of 5, 7 and 10 years.

1.3 Overview over the existing literature

The theory of credit risk models can essentially be divided into two large groups of models: the so-called *structural* or *firm-value models* and the so-called *intensity-based* or *reduced-form models*. While the former class of models has a direct economic interpretation as one analyses the value of a firm with respect to its liabilities, the formulation of the latter type of model is more abstract.

**Intensity-based models**

At the center of the intensity-based models lies the time when a default of a specific company takes place. This time of a default, in short default time, is modelled as the first jump of a counting process. Examples for such counting processes are the Poisson

\(^5\) See \url{www.isda.org}.

\(^6\) See the website \url{www.indexco.com} for a self-description of the International Index Company.

\(^7\) They include ABN AMRO, Bank of America, Barclays Capital, Deutsche Bank, HypoVereinsbank, JP Morgan, Landesbank Baden-Württemberg, Merrill Lynch, and several others.
process, where the intensity is a constant parameter, or the Cox process where the intensity is modelled as a stochastic function in time. This intensity function represents the probability for a default within the next very short time-interval. With this setup, the default of a company happens as a total surprise and is thus not predictable.\textsuperscript{8} The construction of these models is rather abstract, as one does not try to describe fundamental economic quantities such as firm values or liabilities, which might help one to understand the mechanism of a default of a company. Therefore, one loses the interpretation of why a default happens at a particular time. This modelling approach was used by Jarrow and Turnbull [JT95], Jarrow et al. [JLT97], Duffie and Singleton [DS99] and others in the context of single-name credit derivatives such as defaultable bonds, and by Duffee [Duf99], Davis and Lo [DL99], Jarrow and Yu [JY01] and others for the modelling of portfolios with credit-risky securities. Duffie [Duf98] was among the first to model and price portfolio credit derivatives such as first-to-default swaps. However, this is not so easily extended to other portfolio credit derivatives such as CDOs. Other ways of introducing dependence between the default times has been achieved by correlating the stochastic intensity functions as e.g. in Duffie and Gârleanu [DG01], or by modelling the dependence structure of the default times directly with the use of copula functions such as in Li [Li00]. Further analysis on the copula functions which are implied for example by a Cox process setup are analysed in Schönbucher and Schubert [SS01] or in Gregory and Laurent [GL05]. The valuation of CDOs within these setups usually relies on Monte Carlo simulations which, from a computational point of view, can become quite burdensome, especially if the portfolio size becomes large.

\textbf{Structural models}

The rationale behind the structural models is that the financial stability of a firm and its likelihood to default can be explained in terms of the value of the firm in comparison with its liabilities. If the liabilities exceed the assets of a firm at a certain time, then the firm is incapable of paying back its debt. The variations between the different structural models, that one finds in the literature, relate to the way the firm or asset values are modelled, whether the liabilities are modelled as a deterministic default barrier or as a stochastic process on its own, how the flow of information is modelled, and at which points in time a default can take place. For structural models that are used in a multidimensional setup, one also has to consider of how correlations can be introduced for the defaults.

The theory of structural models goes back to the seminal paper by Merton [Mer74], where he derives fair prices for defaultable bonds. Central to his setup is the assumption, that the value of a firm can be modelled via a geometric Brownian motion, and of how this firm-value process behaves in comparison with the known amount of debt that has to be paid back at the maturity of the bond. Merton’s model is considered a one-period model, as he assumes that a default of the firm can only happen at the maturity date. He then relates the payoff from holding the debt to the payoff of a European put option, so that he can apply the Black-Scholes theory [BS73], which was established for pricing European options.

\textsuperscript{8} See e.g. [Pro04], III.2, p. 103 for a definition of a predictable or a totally inaccessible stopping time
Since Merton [Mer73], there has been a lot of research in several different directions in order to try to overcome the limitations of this model, such as the problem of vanishing credit spreads for a very short time to maturity, predictability of the default times, or the assumption on the input parameters being constant over time. Extensions of Merton’s model to the multivariate case or to allow for a default time to happen continuously over time have also been pursued.

In 1976, Black and Cox [BC76] model the underlying firm values which trigger the defaults with geometric Brownian motions as before, but now instead of considering just one point in time where the default takes place, they allow that a firm can default at any time in the relevant time interval: when the firm value hits a default barrier for the first time, then the default occurs. This setup therefore relies on first hitting times, whose distribution functions are known when the process hitting the barrier is a Brownian motion. However, as soon as one models the firm values with a more complex process, one cannot expect to be able to obtain a closed-form expression for the distribution function of the first hitting time, which is needed for the valuation of credit-risky bonds or more complex derivatives. Within a multidimensional firm-value process setup, this issue becomes even more challenging. There have been several extensions of the Black and Cox model for the modelling of defaultable bonds, which try to incorporate more realistic aspects such as stochastic interest rates, bankruptcy costs, or stochastic default barriers. Bielecki and Rutkowski [BR02] present several structural first-passage models, such as those introduced by Leland [Lel94], Leland and Toft [LT96], Longstaff and Schwartz [LS95], Kim et al. [KRS93], Briys and de Varenne [BdV97], and depict their increasing analytical complexity.

In order to introduce multivariate structural models that entail correlated default times, one usually resorts to correlating the asset-value processes. Using correlated Brownian motions in a Black and Cox-type setup has been considered by Zhou [Zho01], where he studied the default correlations between two firms, and by Hull et al. [HPW06], where a factor structure with Brownian motions is studied for the pricing of synthetic CDOs.

However, multivariate versions of the Merton model have been considered much earlier, as one essentially needs to model the stochastic behaviour of the firm values at only one point in time, namely the maturity of the bonds. Indeed, Vasicek [Vas87] introduced a one-factor structure for the asset values of the firms which underlie the loans in a portfolio, where he assumes the factor and the idiosyncratic risk to be independent and normally distributed. From a computational point of view, the great advantage of a factor setup for large portfolios lies in the fact, that default events of the different firms become independent, if we condition on the factor. Under some homogeneity assumptions, Vasicek [Vas91] then derived his famous approximation result, where he considers the percentage gross loss distribution if the portfolio size in the loan portfolio goes to infinity. This type of approximation is called large homogeneous portfolio approximation. A Gaussian factor setup has also been used by several industry models such as the Global Correlation Model by Moody’s KMV (see e.g. Chapter 9 in [CGM01]), or the factor model by CreditMetrics (see [BFG97]). In 2004, McGinty et al. at J.P. Morgan (see [ABMW04]) introduced the
1.3. **Overview over the existing literature**

The notion of base correlations in the context of CDO pricing, which provides one with a simple tool to price bespoke CDOs, that is non-standard CDOs, and even more to analyse the ability of a model to produce high levels of joint default probabilities and a so-called correlation skew. This concept of base correlations relies on a very simple one-factor Gaussian model. A multi-factor Gaussian credit portfolio model was suggested by Klaassen et al. [KLSS01], that also allows to incorporate credit losses due to changes in the firms’ ratings. Klaassen et al. provide a large homogeneous portfolio approximating result within their setup and analyse the sensitivity of high credit-loss quantiles with respect to the $t$-distribution.

For the pricing of portfolio credit derivatives and in particular CDOs, the Gaussian one-factor model is widely used as a benchmark model, but it has great deficiencies with respect to reproducing the observed CDO market prices. This is usually attributed to the low probability, the multivariate Normal distribution assigns to extreme values jointly appearing in two or more components. In the past years, there have thus been made various suggestions for structural models by different authors, which mainly rely on the one-period assumption by Merton. Hull and White [HW04] were among the first to leave the Gaussian framework and to introduce other distributions for the factor and the idiosyncratic risk components within a one-factor structure. In order to produce higher levels of joint extremal events, they used univariate $t$-distributions within a one-factor framework and discussed an numerically efficient way of pricing credit derivatives such as CDOs. Instead of using $t$-distributions, Albrecher et al. [ALS06] propose using Lévy processes, such as Gamma, Inverse-Gaussian, Variance Gamma, Normal Inverse Gaussian or Meixner processes, so that the factor and the idiosyncratic risk both essentially follow the marginal distribution of such a process at time one. Essentially, these models then also become one-period models, even though the distributions stem from Lévy processes. The infinitely divisible Normal Inverse Gaussian (NIG) distribution has also been used by Kalemanova et al. [KSW07] in a one-factor setup. They choose their parameters in such a way that the distribution function of each latent variable, which by construction is the convolution of two NIG distributions, then also becomes NIG.

Another approach to including a stronger dependence among the latent variables was pursued by Andersen and Sidenius [AS04], as they introduced additional stochastic components into the Gaussian factor framework by explicitly modelling the recovery rate to depend on the factor and another independent idiosyncratic quantity, or by letting the factor loadings be stochastic and also a function of the factor.

Let us briefly mention that Frey et al. [FMN01] discussed models, where one directly models the latent variables vector with a Gaussian mixture structure without taking the detour via a factor structure. By that, they avoid the problem that the distribution of the linear combination of the factor and the idiosyncratic risk following arbitrary distributions might not be known.
1.4 The contribution and the aim of this thesis

The aim of this thesis is to introduce a new set of factor models for the pricing of portfolio credit derivatives, which match the observations on the derivatives markets better than the standard Gaussian model. These new models will be based on elliptical distributions, and will extend the Gaussian model in a consistent way. We will identify that mixtures of Normal distributions must be used within the class of elliptical distributions, and will classify which of these distributions can be used to produce more adequate probabilities for joint defaults in the underlying portfolio. On the basis of this elliptical factor setup, we will derive a very general large homogeneous portfolio approximation result, which helps us to deal with multi-name credit derivatives written on a large portfolio. We will provide a diligent study on the impact of our modelling framework to the pricing of Collateralized Debt Obligations and apply this to the valuation of iTraxx tranches. Finally, sundry possibilities to generate stochastic processes with Gaussian mixture marginals are analysed with regard to their applications within the previously discussed elliptical factor model.

An elliptical distributions credit portfolio model

The model by Klaassen et al. [KLSS01] has served us as a starting point for our analysis. They assume a multi-factor setup with centered Gaussian factor and individual components. This setup thus involves to specify the matrix of factor loadings and the variances of the independent individual risk variables. However, despite the additional amount of parameters that can be used for calibration, this model - just like the one-factor Gaussian setup - displays the fundamental deficiencies of the Gaussian distribution with respect to its weak dependence structure. The result of this is that the Gaussian models produce CDO tranche prices that do not correspond to the market values (see a discussion of the pricing problem with Gaussian distributions in Section 2.2 and especially Table 2.1). We therefore need to search for distributions that can produce a stronger dependence in the latent variables vector. However, as CDO reference portfolios can be very large, we still want to rely on a factor structure, as this reduces the dimension of the integrals, one has to compute e.g. for obtaining aggregate quantities such as the expected losses, from the size of the underlying portfolio to the size of the factor. In order not just to replace the Gaussian distribution with an arbitrary other specific distribution function in a rather ad-hoc fashion, we therefore modelled both the factor and the idiosyncratic risk vector with the use of the class of multivariate elliptical distributions (see Section 3.1 for a thorough introduction to elliptical distributions) without at first further specifying a representative from this class (see Section 5.2). We found out that employing elliptical distribution functions has various positive effects, which on the one hand help to keep the complexity of the model low, and on the other hand allow for a large degree of flexibility. One of the most important effects of using elliptical distributions, in particular for modelling the idiosyncratic risk vector, lies in the property that the dependence structure within the idiosyncratic risk vector, and thus within the latent variables vector, is remarkably strengthened. The idiosyncratic risk variables are in general no longer in-
1.4. The contribution and the aim of this thesis

dependent as in the Gaussian case, but only uncorrelated. This additionally entails that
the latent variables become uncorrelated, but in general not independent, if we condition
on a realization of the factor. One of the properties that we could make particular use
of in our setup was the reasonable number of parameters of the elliptical distributions,
as one can entirely determine an elliptical distribution by its mean vector, its (pseudo-)
covariance matrix and its scalar characteristic generator function. Furthermore, as the
multivariate Gaussian distribution is the canonical representative from this class, we were
able to construct our elliptical distributions model in such a way that it directly general-
ises the Gaussian models. The elliptical distributions are employed in such a way that
the first two parameters, i.e. the mean vector and the covariance matrix, coincide with
the parameters in the Gaussian model. However there is one additional “parameter”, the
generator function, with which one can modulate the degree of dependence. We could
identify the mild constraints on the generator functions, under which the covariance ma-
trices of the factor and the idiosyncratic vector, and thus also the variance/covariance
structure of the resulting latent variables remain unchanged. This eases the comparison
of the resulting models that we get on different choices of the generator function. How-
ever, despite the parsimony with respect to the number of parameters, we recognized the
class of elliptical distributions to be large enough to encompass many different known
distribution functions with diverse characteristics. We vigorously believe that the ellip-
tical distributions model therefore maintains a large flexibility for choosing appropriate
distributions, which helps modelling CDO portfolios in an entirely satisfying way. This
view is strongly supported in particular by our numerical studies in Chapter 7.

In Chapter 6 we derive an approximation result (cf. Theorem 6.2.3) as the portfolio size
tends to infinity, so that by consequence of Schönberg’s Theorem (cf. Theorem 3.1.13)
it is compulsory to work within the subclass of mixtures of Normal distributions. The
effects of relying on this subclass on the entire factor structure are the same as those
just outlined for the class of elliptical distributions, as this subclass is still very large and
comprises many of the important elliptical distribution functions.

After we have introduced this elliptical setup and discussed elementary properties, we
set about to analyse which characteristics the mixtures of Normal distributions need
to possess in order for the factor model to display a strong dependence structure, that
might help to overcome the mispricing problems of the Gaussian framework. One of the
concepts of measuring the dependence structure in a distressed situation is the property
of tail-dependence, which we comprehensively analyse within our setup. Every elliptically
distributed vector of dimension \( n \), say, can be decomposed into the independent product
of a positive scaling random variable and a vector which is uniformly distributed on the
\( n \)-th unit sphere. It is important to note that this scaling variable does not coincide
with the mixture variable that we have in the case of mixtures of Normal distributions.
Schmidt [Sch02], Hult & Lindskog [HL02] and Frahm et al. [FJS03] have analysed the
property of tail-dependence in an elliptically distributed vector in characterising the
scaling variable of the classical decomposition. However, as we are obliged to work within
a mixtures of Normal distributions setup to obtain the analytical approximation result
and as we want to directly model the distribution of the mixture variable without taking
The contribution and the aim of this thesis

The detour passed the classical decomposition, we have studied and proved the necessary and sufficient conditions on the mixing distributions, such that the resulting Gaussian mixture vectors display this property of tail-dependence (see Section 5.3.1). Based on this characterisation on the mixing distributions, in Section 5.3.2 we then construct various new mixtures of Normal distributions that possess this property of tail-dependence. These distribution functions will then later be studied in more detail in the context of the pricing of CDOs.

The approximation of portfolio losses within the elliptical framework

The higher the dimension of the reference credit portfolio and the more complex the assumed dependence structure between the individual exposures, the more problems occur in efficiently pricing the claims contingent on this large portfolio or in efficiently performing a sensitivity analysis of the resulting prices with respect to input quantities such as the individual default probabilities. In practice, there are often Monte-Carlo simulations used for the construction of the aggregate loss distribution. Yet, Monte Carlo simulations can be very time-consuming, especially with large portfolio sizes. Portfolio sizes of several hundreds of exposures are easily reached in various collateral structures, for example in the case where outstanding payments on credit cards are being securitised. Alternatives to Monte-Carlo simulations are analytical approximations such as the fundamental result by Vasicek [Vas91]. The efficiency and computational advantages of the analytical approximations over Monte-Carlo simulations are partially rooted in the use of factor structures, where the dimension of the factor should be a lot smaller than the dimension of the portfolio, in order to largely reduce the complexity of the model. These analytical approximations are also called large homogeneous portfolio approximations, and they describe how with growing portfolio size the individual risks are diversified away and thus only the systematic factor components remain.

At the center of Chapter 6 lies the large homogeneous portfolio approximation that we present in Theorem 6.2.3 and which we have been able to prove within the elliptical distributions framework. Indeed, we were able to derive this result in the more general setup, where only the distribution of the idiosyncratic risk vector belongs to the subclass of mixtures of Normal distributions, while the factor vector can follow any (not necessarily elliptical) distribution. We even allow credit losses not only to occur when a company defaults, but also when the likelihood to default changes, which can seen by a change of the credit rating (cf. Section 6.1). Our setup and Theorem 6.2.3, which was constructed hereupon, thus clearly generalises the CreditMetrics framework [BFG97] or the setup by Klaassen et al. [KLSS01]. Additionally, the losses may also vary according to an additional random variable that can be used to incorporate effects such as stochastic recovery rates. Section 6.2 will then deal with the assumptions that need to be satisfied for the large-portfolio approximation to hold and the approximation result itself. The assumptions on the homogeneity of the portfolio is very mild, as solely the second centered moments of the credit losses may not add up too rapidly with growing portfolio size. A remarkable property of this approximation result is that, while in factor structures with independent idiosyncratic risk variables, such as the Gaussian models, only the factor cannot be
1.4. The contribution and the aim of this thesis

diversified away, in the Gaussian mixture framework we additionally have the mixing variable that appears as a systematic component of the limit aggregate portfolio losses (cf. Corollary 6.2.4). This is clearly in line with intuition, as the mixing variable amplifies or diminishes every component of the idiosyncratic vector entirely in the same way.

The approximation result is then applied in Section 6.3, where we have made some homogeneity assumptions in order to simplify the analysis. We present a first numerical example, where we illustrate the limit portfolio loss distributions with the differing Gaussian mixture specifications of Section 5.3.2.

The results we have obtained in Chapter 6 were independent of the underlying probability measure, which could be a risk-neutral or pricing measure, as well as the historical measure. The approximation result can therefore be used for risk management purposes such as the minimum capital requirements in the context of Basel II (see [Bas06]), as well as for the pricing of multi-name credit derivatives.

Application to the valuation of credit derivatives

In a CDO pricing model, one usually assumes that each company can only be in two different states, either in the default or in the non-default state. The elliptical distributions framework, and in particular the mechanism of credit losses as it has been introduced in its full generality in Chapter 6, allows to incorporate many more aspects, however, one can strip down the model to the essential parts required for the valuation of CDOs. In Chapter 7, after we have introduced the basic notions and assumptions needed for the modelling of the CDOs, we step forward to investigate how the pricing of portfolio credit derivatives can be achieved within the elliptical distributions model. To this end, in Section 7.4 we first examine how the approximation result derived in the general elliptical framework in Theorem 6.2.3 can be applied on the CDO structure under additional homogeneity assumptions. We hereby study which terms are relevant within the elliptical context for evaluating CDOs, and derive tractable closed form expressions for each of these terms. In particular, the expected losses in a specific tranche, the default probabilities of the individual firms or the distributions of the limit overall losses are thoroughly investigated for the various Gaussian mixture specifications of Section 5.3.2. After all the analytical investigations, one needs to search for efficient ways of implementing the elliptical distributions model. In Section 7.5 we therefore assess how the key quantities such as the individual default probabilities or the expected losses on a tranche can be numerically exploited for the various Gaussian mixture specifications. In general one would expect to be obliged to use very different numerical approaches for each of them. However, regardless of which mixture specification of Section 5.3.2 we use, we show that all of the integrals appearing in the key expressions can be transformed into integrals, on which we can directly apply Gaussian quadrature formulas to compute their values. This thus opens the road for efficient numerical pricing procedures.

While the previous chapters and sections have demonstrated the multiple advantages of the elliptical distributions model from a theoretical point of view, the rest of Chapter 7 is dedicated to present the results we have gained from extensively testing the model
with its numerous variations on actual CDO data. When implementing a CDO pricing model, one has to be particularly careful as one needs to set up an environment that can handle the large amount of data involved for the pricing procedures, and that at the same time allows for computationally fast routines. We have therefore implemented the various multivariate Gaussian mixture models of Section 5.3.2, such as the models based on the multivariate $t$-distribution, the Power Law, the Power Log Law and the Exp-Exp Law, as well as the standard Gaussian model in a C++ environment. However, we compile the C++ code such that we obtain a dynamic link library\textsuperscript{9}, which can be imported as an “add-in” into Microsoft’s Excel, where the C++ pricing and calibration functions can then directly be used on the data given in a spreadsheet. With this setup, we therefore combine the efficient way in which one can handle large amount of data in Excel with the computationally fast programming language C++.

For each of the various candidates of the elliptical distributions model we thoroughly investigate the behaviour of the resulting CDO prices for a wide range of possible parameter settings. In particular, the Gaussian mixture specifications allow one to specify one additional parameter for each of the two components, that is, one for the factor and one for the idiosyncratic risk vector, which are related to the tail-dependence behaviour of the corresponding distributions.

In a first step we perform a simulation study where we analyse the different correlation levels one needs to suppose for the different model specifications, in order to reproduce various equity tranche prices, which were computed with the standard Gaussian model. In a second step, we closely study the models with respect to their ability to match CDO market data. As the tranched iTraxx represents the most liquid CDO in the European market, we apply our elliptical distributions setup with its various specifications onto iTraxx data. The prices we use stem from different dates, ranging from January 2006 to the end of May 2006 (see Tables A.1 and A.2). At first, in order to focus on a detailed investigation of the various models we limit ourselves to one date, the 6th of April 2006 (see Table 7.5), for which we closely examine the varying iTraxx tranche prices these models produce on different parameter settings. Each combination of the two parameters can lead to very diverse results, even within one mixture specification (see Tables 7.6 to 7.8).

This extensive analysis with varying parameters and mixture distributions shows that the elliptical distributions modelling framework is flexible enough to be able to match the market prices very well. Among the Gaussian mixture specifications that we studied in Section 7.6.5, in particular the elliptical distributions model based on the multivariate Exp-Exp Law seems to fit the market data the best and is able to produce a correlation skew similar to that induced by the market values (cf. Figure 7.2). Especially in comparison with the benchmark one-factor Gaussian model (see Section 4.5.2), the multivariate Exp-Exp Law model with the right choice of parameters always presents itself as superior (cf. Table 7.7).

\textsuperscript{9} See e.g. \url{http://support.microsoft.com/kb/87934/en-us/} for a short introduction to .DLL files and \url{http://support.microsoft.com/kb/178474/en-us} for how to build an add-in for Excel.
We close this chapter with an illustration of the changes over time of the correlation parameters within a calibrated model and of the resulting sums of absolute errors, that these calibrated models produce in comparison to the market data.

**Dynamic elliptical distributions model**

As for portfolio credit derivatives it is of crucial importance to carefully model the dependence structure within the portfolio, our focus lies on this issue in the first chapters of this thesis. The one-period models allow to precisely concentrate on this particular issue, while they at the same time implicitly ask for the asset-value process to be strictly stationary. In the application of our elliptical one-period model in Chapter 7, we see that this assumption leads to satisfactory results when we consider CDOs with just one maturity date, e.g. in 5 years time. One can argue that the distributional differences, that might exist during the lifetime of such a CDO, are leveled out under the one-period assumption.

However, if one wants to model CDOs which are written on the same reference portfolio, but with different maturities, it can make sense to allow the distributions of the asset values to actually vary over time. Usually, iTraxx tranches with a maturity of 5 years represent the most liquid of the traded iTraxx tranches, but there are also tranches with 7 and with 10 years of maturity available. Therefore, when one wants to model CDOs with different maturities consistently within one model at the same time, one might need to relax the condition of strict stationarity, which is studied in Chapter 8.

The aim of Chapter 8 is therefore to analyse how the static elliptical distributions framework of Chapter 5 can be extended, to allow for a dynamic setting and changing distributional characteristics over time. We present multiple possibilities that can be used to accomplish the transmission from a static to a dynamic setting, e.g. by using discrete-time time series models, continuous-time short-rate models from the interest rate theory, time-changed Brownian motions and, at last, subordinated Lévy processes. All the time, we show how the suggested processes can be adapted in such a way, that they perfectly fit into the static modelling framework. This allows us to directly apply the results obtained earlier as the large homogeneous portfolio approximation or the discussion about tail dependence.
1.5 Organization of this thesis

This thesis is organized as follows. After this first introductory chapter, in Chapter 2 we present a typical way of how CDOs can be structured and talk about the fundamental challenges when modelling and pricing CDOs. Chapter 3 provides us with the fundamental concepts that we need in the rest of the thesis, such as the notions of elliptical distributions, regular variation and Copulae, as well as some important theorems accompanying these concepts. A detailed overview of several of the prominent credit portfolio models is given in Chapter 4, some of which will serve as benchmarks later on. The elliptical factor model is presented in detail in Chapter 5, along with a thorough discussion of tail-dependence present in mixtures of Normal distributions. The subsequent Chapter 6 details how credit losses in a portfolio can originate and then centers around the large homogeneous portfolio approximation within the elliptical factor structure. The entire theory obtained so far is then applied in Chapter 7 on the pricing of CDOs, where also the results of a numerical study of the Gaussian mixture specifications applied on iTraxx data are given. Finally, the last Chapter 8 concentrates on the issue of embedding the elliptical factor model into a dynamic setup, by generating stochastic processes with Gaussian mixture marginals with the use of discrete-time, as well as continuous-time processes.
1.5. Organization of this thesis
Chapter 2

Collateralized Debt Obligations and the pricing problem

The aim of this chapter is to introduce the basic structure of a Collateralized Debt Obligation and to discuss the fundamental problems, which arise with the modelling and the pricing of these structures.

2.1 The basic structure of a CDO

We want to consider a CDO with a structure as displayed in Figure 2.1:

The risk seller, e.g. a bank, holds a portfolio of credit-risky assets whose credit risk shall
2.1. The basic structure of a CDO

be sold to different investors with the use of a so-called Special Purpose Vehicle (SPV). This special legal entity was set up by the risk seller for the sole purpose of restructuring the risk which is rooted in the single-name credit exposures and of placing this risk with investors. At first, the risk is transferred from the bank to the SPV on a single-name basis via Credit Default Swaps (CDS), where one CDS corresponds to one credit-risky asset in the risk seller’s portfolio. The SPV then pools the risk and sells certain portions of the risk to investors according to their risk appetite. These portions of risk are denoted in percentages of the possible portfolio losses (e.g. the sum of all notional amounts) and are called CDO tranches. The tranches usually take the names equity tranche, junior tranche, mezzanine tranche, senior tranche and super senior tranche. An investor in a specific tranche will then have to offset all the credit losses which lie within a specific predefined range of the possible portfolio losses that corresponds to his tranche. The investor in the so-called equity tranche will be vulnerable to the first so many percent losses in the credit-risky portfolio (e.g. 0% to 3% of possible portfolio losses), the investor in the junior tranche will then be obliged to cover the losses if the portfolio sustains more losses than the equity tranche investor has covered (e.g. 3% to 6% of possible portfolio losses). If there are even more losses in the underlying portfolio, then the investor in the mezzanine tranche needs to offset the losses for the risk seller, then comes the investor of the senior tranche, and at last the investor of the super senior tranche is asked to compensate for the losses of the risk seller. Therefore, the super senior tranche is only hit by the losses in the underlying portfolio if they reach such a high level that all the other, more junior CDO tranches have already been entirely absorbed. The senior tranche is therefore protected from below, namely only by the equity and the mezzanine tranche. Thus, it also needs quite some losses in the underlying portfolio for the senior tranche to be affected and for its investor to be obliged to compensate the risk seller for possible losses. The mezzanine tranche is less well protected from below, namely only by the equity and the junior tranches, and the junior tranche only has the equity tranche as a kind of cushion before it gets hit by the losses. The equity tranche is very likely to be affected by losses as for this only one underlying exposure is needed to cause credit losses.

On the other hand, the SPV will receive premia from the risk seller as part of the contractual commitments of the various CDSs, and will transfer this money to the investors, as they need to be compensated for taking on the risk. The height of the premia that the investors in the various tranches receive depends on the level of risk they have adopted and the bandwidth of the respective tranche. If the bandwidths are the same, then the equity tranche investor bears the largest risk and will thus also be attributed the largest premium, the junior tranche investor will receive the second largest premium and so on. The super senior tranche investor finally receives the lowest premium.

The offsetting default payments within a CDO structure are contingent payments as they are only exchanged once losses occur and only affect specific tranches and their investors at a particular credit event. The height of these payments are determined by the actual losses that occur, and the level of the aggregate losses in the portfolio determines who has to cover these losses. On the other hand, the premium payments are certain to be
2.1. The basic structure of a CDO

exchanged between the bank and the investors via the SPV. Yet, once a specific tranche has been hit by the portfolio losses, the risk premium on a tranche is reduced by the percentage of how much of the tranche has already been absorbed by the credit losses. For example, if a tranche covers the range between 3% and 6% of the portfolio losses and if the losses already add up to 4.5% of the portfolio’s notional amounts, then only half of the original premium is being paid. If the portfolio has sustained losses that make up 6% of the possible portfolio losses or more, then no premium is paid to the investor in this example.

Each CDO tranche corresponds to a specific risk level and is sometimes attributed a credit rating by a rating agency. This makes the investment easier to assess by potential investors, and allows investors (e.g. insurance companies) that might by law be forced to only invest in highly rated assets to participate in the credit-risk market. By first transferring the risk from the bank to the SPV and then letting the SPV sell it further on to the investors, the credit ratings of the CDO tranches become independent of the rating of the bank, which might be poor. The ratings then depend on the credit standing of the SPV and therefore on the underlying portfolio, as well as on the seniority of the respective tranche. Often, a CDO is structured in such a way that the needs and the risk appetite of one or more investors are met, and have therefore become very popular in the last years.

In a purely synthetic CDO such as the iTraxx tranches, there is no longer a bank that actually needs to hold this portfolio of credit-risky exposures and that wants to sell the risk. Instead, one focuses on a portfolio of CDS contracts on reference entities, but doesn’t even need to hold the underlying contracts in order to sell a CDO tranche to an interested investor. Yet, all the payments between the tranche seller and the tranche buyer have to be met as if the seller owned the underlying portfolio of CDS contracts. Therefore, especially with synthetic CDOs, there might be a noticeable counterparty risk\(^1\) present. However, we will neglect this risk for our further analysis.

When one talks about pricing a CDO, one actually means to determine the various premia that have to be paid to the investors in the different CDO tranches and which are considered to be fair. The height of the premia - the so-called credit spreads or just spreads - must be determined under the assumption of no arbitrage with the use of an adequate pricing model.

\(^1\) Counterparty risk describes the risk that one of the counterparties in the tranche deal, the tranche seller or the tranche buyer, will not be able to meet his contractual commitments.
2.2 Challenges with modelling and pricing CDOs

There are many strong challenges one faces when trying to accurately model CDOs. One of the prevailing challenges lies in the fact that the dimension of the portfolio underlying the CDO can be quite large. While the iTraxx reference portfolio consists of 125 reference entities, in principle CDOs can be structured on portfolios of an arbitrary dimension. But already the 125 iTraxx reference entities cause quite some complexity from a statistical point of view, especially if one tries to estimate their dependence structure. In a Gaussian model which we will review in Chapter 4, one relies entirely on multivariate Gaussian distributions which - in its most general form - already requires \( \frac{125 \cdot 126}{2} = 7875 \) correlation parameters to be estimated or to be found via calibration of the model to present market data.

In order to reduce the dimension, one usually relies on a one-factor setup, as a consequence of which the number of unknown correlation parameters reduces to 125 in the iTraxx case. But already this number of parameters can cause quite some difficulty due to either sparseness of data for applying robust statistical methods or, even more so, for the fact that the quantities of which one would need to estimate the correlation coefficients are often unobservable. On the other hand, calibrating a one-factor model with 125 different correlation parameters to the iTraxx tranches would mean that at each point in time one simultaneously calibrates 125 correlation parameters to only 5 data points as there are only 5 such tranches. Therefore, one often limits oneself to using only one single correlation parameter for the pricing of CDOs that simultaneously describes the dependence of each of the portfolio’s components with the single macro-factor. However, from the study of copula functions (see e.g. [Nel99] or the brief introduction in Section 3.3) we know that the concept of linear correlations does not entirely determine the dependence structure in a stochastic vector and that there are still infinitely many different multivariate distribution functions, that match the correlation structure one might have estimated.

While Brownian motions and thus also Normal distributions have extensively been used in various fields of financial modelling, empirical studies have shown that many financial quantities such as stock returns are usually not normally distributed. In these cases, one therefore tries to employ stochastic processes whose marginal distributions display characteristics such as heavy tails, skewness or tail-dependence. The standard and therefore benchmark pricing model for CDOs also uses multivariate Gaussian distributions, which has the advantage of being parsimonious in the number of appearing parameters. Additionally, key properties, such as that the sum of independent Gaussian random variables is still Gaussian, makes the derivation and then also computation of relevant quantities tractable, not only as there exist a number of numerical tools for the exploitation of Gaussian integrals. However, one of the important disadvantages of this Gaussian setup is once more rooted in the fact that one cannot model heavy tails or tail-dependence with the multivariate Gaussian distribution. In order to illustrate the effect of using such distribution functions for the pricing of CDOs, we have calibrated the standard Gaussian
model to the market values of the iTraxx tranches observed on the 6th of April 2006 in such a way that the equity tranche, that is the tranche covering the first 3% of the portfolio losses, is perfectly calibrated. This was attained for a correlation parameter of $\rho = 39.7\%$. As we can see in Table 2.1, at this level of correlation the 3% - 6% tranche is clearly overpriced while the most senior tranche, which is the tranche responsible for 12% to 22% of the portfolio losses, is underpriced.

<table>
<thead>
<tr>
<th>Tranche ranges</th>
<th>0% - 3%</th>
<th>3% - 6%</th>
<th>6% - 9%</th>
<th>9% - 12%</th>
<th>12% - 22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>iTraxx Market</td>
<td>20.04</td>
<td>56.04</td>
<td>16.65</td>
<td>9.17</td>
<td>2.57</td>
</tr>
<tr>
<td>Gaussian Model</td>
<td>20.04</td>
<td>120.13</td>
<td>21.28</td>
<td>5.51</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Table 2.1: The market values of the five iTraxx Europe Series 3 tranches on the 6th of April 2006 compared with the one-factor Gaussian model where the single correlation parameter $\rho$ was chosen such that the equity tranche is perfectly hit ($\rho = 39.7\%$).

The reasons for the overpricing of the more junior tranches and the underpricing of the senior tranches within the Gaussian model is usually attributed to the deficiency that the Normal distribution cannot produce as high a likelihood of common defaults as would be needed: the more senior tranches are only hit once there have been many defaults in the portfolio whose likelihood is positively related with the likelihood of joint defaults.
2.2. Challenges with modelling and pricing CDOs
Chapter 3

Fundamental theoretical concepts

3.1 Elliptical distributions

In practice, one of the most commonly used distribution functions for the modelling of key components in financial markets is still the Normal distribution. Most certainly, one of the main reasons for the ongoing popularity of this function lies in the fact that it is completely described by only two parameters in the univariate case and comparatively few parameters in the multidimensional case, which are all relatively easy to interpret. Yet, the Normal distribution cannot display important properties such as heavy tails or - in the multivariate case - tail-dependence which seem to be needed to appropriately reproduce specific aspects that are observed in stock markets and also in markets for credit derivatives (see also Section 2.2).

There are several families of distribution functions that try to overcome these deficiencies, that contain the Normal distribution as a kind of canonical representative, and that are used for modelling purposes in the wide realm of mathematical finance. Among these is the family of elliptically contoured distributions, in short elliptical distributions. It represents a family of multidimensional distribution functions which contains many other important distribution functions such as the $t$-distribution or the generalised inverse Gaussian distribution.

In this section we want to introduce this family as it will play a crucial role within our model for the valuation of CDOs, and we want to discuss some of its important properties which will be needed in the subsequent chapters. We will see that this family can be fully described by few parameters plus a special function, the so-called characteristic generator. Our introduction to elliptical distributions is based on the monograph by Fang, Kotz and Ng [FKN90].

We denote by $O(n)$ the group of orthogonal matrices in $\mathbb{R}^{n \times n}$, i.e.

$$O(n) = \{ \Gamma \in \mathbb{R}^{n \times n} | \Gamma^T \Gamma = I_n \}.$$
Definition 3.1.1 (Spherical distributions)
An $n$-dimensional random vector $X$ is said to follow a spherically symmetric distribution, in short a spherical distribution, if $\Gamma X \overset{d}{=} X$, for all $\Gamma \in \mathcal{O}(n)$.

Theorem 3.1.2 Let an $n$-dimensional random vector $X$ and its characteristic function $\psi_X$ be given. $X$ follows a spherical distribution iff there exists a scalar function $\phi : \mathbb{R}^+ \to \mathbb{R}$ such that $\psi_X(t) = \phi(tT t)$ for all $t \in \mathbb{R}^n$.

Proof: Let $X$ follow a spherical distribution. Then $\psi_X(\Gamma t) = \psi_{\Gamma^T X}(t) = \psi_X(t)$, for all $t \in \mathbb{R}^n$ and $\Gamma \in \mathcal{O}(n)$, as with $\Gamma \in \mathcal{O}(n)$ also $\Gamma^T \in \mathcal{O}(n)$. Let $s, t \in \mathbb{R}^n$ be arbitrary, but such that $s^T s = t^T t$. Then there exists a $\Gamma \in \mathcal{O}(n)$ such that $\Gamma s = t$, and so $\psi_X(t) = \psi_X(\Gamma s) = \psi_X(s)$. Hence, $\psi_X$ depends on an arbitrary variable $t \in \mathbb{R}^n$ only through the quantity $t^T t$. Thus, there exists a scalar function $\phi$ such that $\psi_X(t) = \phi(t^T t)$, for all $t \in \mathbb{R}^n$. Note that $-I_n \in \mathcal{O}(n)$, so that $\overline{\psi}_X(t) = \psi_X(t) = \psi_X(t) = \phi(t^T t)$, for all $t \in \mathbb{R}^n$, and the scalar function $\phi$ therefore maps from $\mathbb{R}^+$ to $\mathbb{R}$. Conversely, if there exists such a scalar function $\phi : \mathbb{R}^+ \to \mathbb{R}$ with $\psi_X(t) = \phi(t^T t)$ for $t \in \mathbb{R}^n$, then

$$\psi_{\Gamma X}(t) = \psi_X(\Gamma^T t) = \phi(t^T \Gamma^T \Gamma t) = \phi(t^T t) = \psi_X(t),$$

for all $t \in \mathbb{R}^n$ and $\Gamma \in \mathcal{O}(n)$. Therefore, $\Gamma X \overset{d}{=} X$, for all $\Gamma \in \mathcal{O}(n)$. \hfill \qed

A scalar function $\phi : \mathbb{R}^+ \to \mathbb{R}$ as in Theorem 3.1.2 is called the characteristic generator of the spherical distribution. In the following, we will write $X \sim S_n(\phi)$ to mean that the $n$-dimensional vector $X$ possesses a characteristic function of the form $t \mapsto \phi(t^T t)$, for $t \in \mathbb{R}^n$. For an $n$-dimensional random vector we want to denote the set of all possible characteristic generators by $\Phi_n$, that is

$$\Phi_n := \{ \phi : \mathbb{R}^+ \to \mathbb{R} \text{ such that } t \mapsto \phi(t^T t), t \in \mathbb{R}^n, \text{ is a characteristic function} \}.$$

Observe that these sets are decreasing in $n \in \mathbb{N}$, that is $\Phi_n \supset \Phi_{n+1}$ for $n \in \mathbb{N}$:

for an arbitrary $\phi$ in $\Phi_{n+1}$ there exists an $(n+1)$-dimensional random vector $X^{(n+1)}$ such that $\phi(t^{(n+1)} t^{(n+1)}) = \mathbb{E} \left( \exp \left( i t^{(n+1)} X^{(n+1)} \right) \right)$ for all $t^{(n+1)} \in \mathbb{R}^{n+1}$. For every $t^{(n)} \in \mathbb{R}^n$ we denote by $\tilde{t}^{(n+1)}$ the $(n+1)$-dimensional vector where the first $n$ components are filled with $t^{(n)}$ and the last component is equal to zero, i.e. $\tilde{t}^{(n+1)} := (t^{(n)}, 0)$. Let us denote by $X^{(n)}$ the $n$-dimensional random subvector of $X^{(n+1)}$ comprising the first $n$ components of $X^{(n+1)}$. Then

$$\phi \left( t^{(n)} t^{(n)} \right) = \phi \left( \tilde{t}^{(n+1)} \tilde{t}^{(n+1)} \right) = \mathbb{E} \left( \exp \left( i \tilde{t}^{(n+1)} X^{(n+1)} \right) \right) = \mathbb{E} \left( \exp \left( i t^{(n)} X^{(n)} \right) \right),$$

for every $t^{(n)} \in \mathbb{R}^n$, and therefore $t^{(n)} \mapsto \phi(t^{(n)} t^{(n)})$, with $t^{(n)} \in \mathbb{R}^n$, is the characteristic function of $X^{(n)}$. Hence, $\phi$ is also in $\Phi_n$.

Thus, the setting $\Phi_\infty := \cap_{n=1}^{\infty} \Phi_n$ becomes meaningful.

\footnote{We will use this acronym instead of the longer expression “if and only if".}
Let us analyse the shape of the set $\Phi_n$ of possible characteristic generators. To this end, we make use of Bochner’s Theorem.

**Theorem 3.1.3 (Bochner’s Theorem)**

A function $\psi : \mathbb{R}^n \to \mathbb{C}$ is a characteristic function iff the following properties hold:

- $\psi$ is continuous,
- $\psi(0) = 1$, and
- $\psi$ is positive semi-definite, i.e. for every $N \in \mathbb{N}$, every $x_1, \ldots, x_N \in \mathbb{R}^n$ and every $\lambda_1, \ldots, \lambda_N \in \mathbb{C}$ the sum
  \[
  \sum_{i,j=1}^N \psi(x_i - x_j)\lambda_i\overline{\lambda_j}
  \]
  is real and $\geq 0$ (where $\overline{\lambda_j}$ denotes the complex conjugate of $\lambda_j$).

**Proof:** See Theorem 22.3, p.189, and the subsequent remark in [Bau91], Theorem 3.2.3, p.58, in [Boc55] (for Euclidian spaces), or Bochner’s Theorem in Section 1.4.3, p.19, in [Rud62] (for locally compact abelian groups). As mentioned in [Rud62] (p.19), the result for $n = 1$ goes back to Bochner [Boc33] and the general result is due to Weil [Wei38]. □

**Corollary 3.1.4** A function $\phi : \mathbb{R}^+ \to \mathbb{R}$ is in $\Phi_n$ iff $\phi$ satisfies the following conditions:

- $\phi$ is continuous,
- $\phi(0) = 1$, and
- the mapping $t \in \mathbb{R}^n \to \phi(t't)$ is positive semi-definite, i.e. for every $N \in \mathbb{N}$, every $x_1, \ldots, x_N \in \mathbb{R}^n$ and every $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ the sum
  \[
  \sum_{i,j=1}^N \phi((x_i - x_j)'(x_i - x_j))\lambda_i\lambda_j
  \]
  is non-negative.

**Proof:** This corollary is a direct consequence of Bochner’s Theorem as $\phi$ maps into the real numbers: let us write $\lambda_j \in \mathbb{C}$ in condition three of Bochner’s Theorem as $\lambda_j = u_j + iv_j$ with $u_j, v_j \in \mathbb{R}$, for $j = 1, \ldots, N$. Then $\lambda_j\overline{\lambda_j} = (u_ju_j + v_jv_j) + i(v_ju_j - u_jv_j)$
3.1. Elliptical distributions

and

\[ \sum_{i,j=1}^{N} \phi((x_i - x_j)'(x_i - x_j)) \lambda_i \lambda_j = \sum_{i,j=1}^{N} \phi((x_i - x_j)'(x_i - x_j))(u_i u_j + v_i v_j) \]

\[ + i \sum_{i,j=1}^{N} \phi((x_i - x_j)'(x_i - x_j))(v_i u_j - u_i v_j) \]

\[ = \sum_{i,j=1}^{N} \phi((x_i - x_j)'(x_i - x_j))(u_i u_j + v_i v_j), \]

as \( \phi((x_i - x_j)'(x_i - x_j)) = \phi((x_j - x_i)'(x_j - x_i)) \) for all \( i, j \). This expression is always real and it suffices to check

\[ \sum_{i,j=1}^{N} \phi((x_i - x_j)'(x_i - x_j)) \lambda_i \lambda_j \]

with \( \lambda_i, \lambda_j \in \mathbb{R} \) for being non-negative.

\[ \Box \]

**Remark 3.1.5** The third condition on \( \phi \) in the preceding corollary cannot be made independent of the dimension \( n \) appearing in \( \Phi_n \), even though \( \phi \) is a scalar function from \( \mathbb{R}^+ \) into \( \mathbb{R} \), as \( \Phi_{n+1} \) is a proper subset of \( \Phi_n \), for \( n \geq 1 \) (see the comment right after Example 3.1.6), and thus one of the three conditions has to take the dimension into account. Therefore, the positive semi-definiteness of \( \phi \) is not equivalent to the last condition.

**Example 3.1.6** Let \( u^{(n)} \) be an \( n \)-dimensional random vector which is uniformly distributed on the unit sphere surface in \( \mathbb{R}^n \). Its characteristic function \( \psi_n \) then becomes

\[ \psi_n(t) = \int_{S \cap x^T x = 1} \exp(it^T x) dS/S_n, \]

for \( t \in \mathbb{R}^n \), with \( S_n := \frac{2\pi^{n/2}}{\Gamma(n/2)} \) being the area of the unit sphere surface in \( \mathbb{R}^n \). The vector \( u^{(n)} \) is spherically distributed as \( u^{(n)} = \Gamma u^{(n)} \), for all \( \Gamma \in \mathcal{O}(n) \), which is clear from the geometric properties of orthogonal matrices (as they preserve lengths and angles when the Euclidean vector space \( \mathbb{R}^n \) is furnished with the canonical scalar product \( \langle ., . \rangle \) with \( \langle x, y \rangle := x^T y, \ x, y \in \mathbb{R}^n \)), or which can be seen from the fact that the characteristic function \( \psi_n \) takes the form

\[ \psi_n(t) = \Omega_n(t) := \frac{1}{B((n-1)/2, 1/2)} \int_0^\pi \exp(i\sqrt{t} \cos(s)) \sin^{n-2}(s) ds, \quad \text{for } t \in \mathbb{R}^n, \]

according to Theorem 3.1, p.70, in [FKN90]. Here, the function \( B(\cdot, \cdot) \) is the beta function with \( B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta), \) for \( \alpha, \beta > 0 \).
The question that was raised in Remark 3.1.5 was whether $\Phi_{n+1}$ is a proper subset of $\Phi_n$. Theorem 3.2. in [FKN90] answers this question in a positive way as the function $t \mapsto \Omega_n(t^T t)$, for $t \in \mathbb{R}^{n+1}$, is not an $(n+1)$-dimensional characteristic function.

There is also another characterisation for $\Phi_n$ which goes back to Schönberg [Sch38] and which is a basis for a standard decomposition of a spherically distributed random vector into the product of a non-negative random variable $r$ and a random vector $u^{(n)}$ which is uniformly distributed on the unit sphere surface in $\mathbb{R}^n$. This decomposition will be discussed in the corollary of the following theorem. Yet, this decomposition will only play a minor role within our setup, as we will make use of another decomposition which assumes an additional structure.

**Theorem 3.1.7** A function $\phi$ lies in $\Phi_n$ iff there is a distribution function $F$ on $[0, \infty)$ such that

$$
\phi(x) = \int_0^\infty \Omega_n(xr^2)dF(r),
$$

where $\Omega_n(t^T t)$ is the characteristic function of a random vector $u^{(n)}$ that is uniformly distributed on the unit sphere surface in $\mathbb{R}^n$ as in Example 3.1.6.

**Proof:** As $\phi \in \Phi_n$ there must exist an $n$-dimensional random vector $Y$ with characteristic function $t \mapsto \phi(t^T t)$, for $t \in \mathbb{R}^n$, and distribution function $H$ on $\mathbb{R}^n$. Using some straightforward transformations one shows that $\phi$ must be of the above form (3.1) where $F(r) := \int_{y \in \mathbb{R}^n : \|y\| \leq r} dH(y)$, for $r \in \mathbb{R}^+$. For the converse, one starts with the representation (3.1) and assumes that the non-negative random variable $r$ follows the distribution function $F$. Additionally, $r$ shall be independent of an $n$-dimensional random vector $u^{(n)}$ which shall be uniformly distributed on the unit sphere surface in $\mathbb{R}^n$. Then the characteristic function of the product $ru^{(n)}$ evaluated at $t \in \mathbb{R}^n$ becomes $\phi(t^T t)$, which shows that $\phi \in \Phi_n$. For the more detailed proof we refer to Theorem 2.2 in [FKN90].

**Corollary 3.1.8** Let $X$ be an $n$-dimensional random vector with characteristic function $t \mapsto \phi(t^T t)$, $t \in \mathbb{R}^n$, where $\phi \in \Phi_n$. Then there exist a non-negative random variable $r \geq 0$ and an $n$-dimensional random vector $u^{(n)}$ which is independent of $r$ such that

$$
X \overset{d}{=} ru^{(n)},
$$

where $u^{(n)}$ is uniformly distributed on the unit sphere surface in $\mathbb{R}^n$ and where $r \sim F$ with $F$ being related to $\phi$ via (3.1).

**Proof:** The proof of this corollary goes exactly along the lines of the second part of the proof of Theorem 3.1.7
**3.1. Elliptical distributions**

**Theorem 3.1.9** If \( X = ru(n) \sim S_n(\phi) \) with \( \phi \in \Phi_n \), then
\[
\|X\| \overset{d}{=} r.
\]

If additionally \( P(X = 0) = 0 \), then
\[
\frac{X}{\|X\|} \overset{d}{=} u(n);
\]

moreover, \( \|X\| \) and \( \frac{X}{\|X\|} \) are independent.

**Proof:** The proof is adapted from Theorem 2.3, p. 30, in [FKN90]: since \( P(X = 0) = P(r = 0) = 0 \), we define the function \( f : \mathbb{R}^n \rightarrow \mathbb{R}^{n+1} \) via \( f_1(x) := (x^T x)^{1/2} \) and \( f_2, \ldots, n+1(x) := \frac{x}{\|x\|} \). Then
\[
\begin{pmatrix} \|X\| \\ \frac{X}{\|X\|} \end{pmatrix} \overset{d}{=} \begin{pmatrix} f_1(X) \\ f_2, \ldots, n+1(X) \end{pmatrix} = \begin{pmatrix} f_1(r u(n)) \\ f_2, \ldots, n+1(r u(n)) \end{pmatrix} = \begin{pmatrix} r \\ u(n) \end{pmatrix}.
\]
As \( r, u(n) \) are independent the claim follows. \( \square \)

**Definition 3.1.10 (Elliptical distributions)**

An \( n \)-dimensional random vector \( X \) is said to follow an elliptically contoured distribution, in short an elliptical distribution, with parameters \( \mu \in \mathbb{R}^n, \Sigma \in \mathbb{R}^{n \times n} \) and a scalar function \( \phi \in \Phi_k \), in notation \( X \sim EC_n(\mu, \Sigma, \phi) \), if \( X \) is of the form
\[
X = \mu + A^T Y,
\]
where \( k = \text{rank}(\Sigma) \leq n, \ A \in \mathbb{R}^{k \times n}, \ A^T A = \Sigma \) and \( Y \sim S_k(\phi) \).

Note that in the previous definition \( \phi \) must be in \( \Phi_k \), but not necessarily in \( \Phi_n \). Recall that when \( k < n \) we have \( \Phi_n \subset \Phi_k \), but \( \Phi_n \neq \Phi_k \) (again, see the comment right after Example 3.1.6).

**Lemma 3.1.11** Let \( \mu \in \mathbb{R}^n, \ \Sigma \in \mathbb{R}^{n \times n} \) and \( \phi \in \Phi_k \), where we set \( k = \text{rank}(\Sigma) \). An \( n \)-dimensional random vector \( X \) follows an elliptical distribution \( X \sim EC_n(\mu, \Sigma, \phi) \) iff its characteristic function \( \psi_X \) takes the form
\[
\psi_X(t) = \exp(it^T \mu) \phi(t^T \Sigma t), \quad \text{for all} \ t \in \mathbb{R}^n.
\] (3.2)

**Proof:** If \( X \sim EC_n(\mu, \Sigma, \phi) \), then by definition there exists a \( k \)-dimensional vector \( Y \sim \Phi_k(\phi) \) and a matrix \( A \in \mathbb{R}^{k \times n} \), such that \( A^T A = \Sigma \) and \( X \overset{d}{=} \mu + A^T Y \). Therefore, with \( \psi_Y \) denoting the characteristic function of \( Y \), we obtain
\[
\psi_X(t) = \exp(it^T \mu) \psi_Y(t) = \exp(it^T \mu) \phi(t^T \Sigma t),
\]
for \( t \in \mathbb{R}^n \), due to Theorem 3.1.2 and the discussion directly thereafter. The converse follows from the fact that for any \( \Sigma \in \mathbb{R}^{n \times n} \) with \( k = \text{rank}(\Sigma) \) there exists a \( A \in \mathbb{R}^{k \times n} \) such that \( A^T A = \Sigma \) and there exists a \( k \)-dimensional random vector \( Y \) with characteristic function \( \mathbb{R}^k \ni t \mapsto \phi(t^T t) \) and as such \( Y \sim S_k(\phi) \). Then \( X \) equals \( \mu + A^T Y \) in distribution, which entails that \( X \sim EC_n(\mu, \Sigma, \phi) \).

Of course, the spherical distributions represent a subclass of the class of elliptical distributions, as an \( n \)-dimensional random vector \( X \) with \( X \sim S_n(\phi) \) also follows an \( EC_n(0, I_n, \phi) \) distribution.

**Example 3.1.12** If we have \( X \sim EC_n(\mu, \Sigma, \phi) \) where \( \phi(x) = \exp(-\frac{1}{2}x) \), for all \( x \in \mathbb{R} \), then we are in the special case that \( X \) follows a Normal distribution \( N_n(\mu, \Sigma) \) as the characteristic function of such a random vector is given via

\[
\psi(t) = \exp(it^T \mu) \exp(-\frac{1}{2}t^T \Sigma t),
\]

for \( t \in \mathbb{R}^n \).

Later on, we will need to look at random vectors \( X = X^{(n)} \) where the dimension \( n \) of the vector increases. Thus, if we want to retain the mapping \( t \mapsto \phi(t^T t) \) to be a characteristic function for every \( n \in \mathbb{N} \) we need to postulate that \( \phi \in \Phi_\infty \). But then \( \phi \) must be of a special form:

**Theorem 3.1.13** A function \( \phi : \mathbb{R}^+ \to \mathbb{R} \) lies in \( \Phi_\infty \) iff there exists a distribution function \( G_\infty \) with support \( [0, \infty) \) such that

\[
\phi(x) = \int_0^\infty \exp\left(-\frac{1}{2}x r^2\right) dG_\infty(r), \tag{3.3}
\]

for \( x \in \mathbb{R}^+ \).

**Proof:** As this result is central for the further development of our model, we would like to present its proof which is along the lines of the proof given in [FKN90], which itself was adapted from [Kin72].

Let \( \phi \in \Phi_\infty \). Then there exists a sequence \( (X_n)_{n \in \mathbb{N}} \) of random variables such that, for every \( n \in \mathbb{N} \), the vector \( X^{(n)} = (X_1, \ldots, X_n) \) possesses the characteristic function \( t \mapsto \phi(t^T t) \), for \( t \in \mathbb{R}^n \), and thus follows a spherical distribution. For an arbitrary \( n \in \mathbb{N} \), the permutation matrices, which are matrices \( H \in \mathbb{R}^{n \times n} \) with elements that are only 0 or 1 and with \( H^T H = I_n \), belong to the orthogonal group \( O(n) \), so that \( H X^{(n)} \overset{d}{=} X^{(n)} \) for every permutation matrix \( H \). Thus, the sequence \( (X_n)_{n \in \mathbb{N}} \) is exchangeable, which by definition means that for any \( n \in \mathbb{N} \) the joint distributions of all possible permutations

\(^2\) This is a direct consequence of the Cholesky decomposition, see [HJ90].
of \((X_1, \ldots, X_n)\) coincide. The infinite version of de Finetti’s Theorem (see e.g. Theorem 1.49 in Subsection 1.4.2.2 and also the quantities appearing in its proof in Subsection 1.5.4.2 in [Sch95]) guarantees that there exists a \(\sigma\)-field \(\mathcal{F}\), which can be chosen to be the tail \(\sigma\)-field generated by the sequence \((X_n)_{n \in \mathbb{N}}\), that is \(\mathcal{F} = \bigcap_{n \geq 1} \sigma(X_n, X_{n+1}, \ldots)\), such that conditional on \(\mathcal{F}\) the \(X_i, i \in \mathbb{N}\), become independent. Define

\[
\Phi(t) := \mathbb{E}(\exp(it X_1)|\mathcal{F}),
\]

for \(t \in \mathbb{R}\), so that \(\Phi\) becomes \(\mathcal{F}\)-measurable for every fixed \(t \in \mathbb{R}\), and continuous in \(t \in \mathbb{R}\), with probability one. Additionally, we have

\[
\Phi(-t) = \overline{\Phi(t)}, \quad |\Phi(t)| \leq 1, \quad \Phi(0) = 1, \quad (3.4)
\]

with probability one, for all \(t \in \mathbb{R}\). Let us fix an arbitrary \(n \in \mathbb{N}\) for the rest of the proof. The conditional independence of the \(X_i, i \in \mathbb{N}\), entails that

\[
\mathbb{E}\left(\exp\left(\sum_{i=1}^{n} t_i X_i\right)\big|\mathcal{F}\right) = \prod_{i=1}^{n} \Phi(t_i), \quad (3.5)
\]

with probability one, for all \((t_1, \ldots, t_n)^T \in \mathbb{R}^n\), as a consequence of which

\[
\phi(t^T t) = \mathbb{E}\left(\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right) = \mathbb{E}\left(\prod_{i=1}^{n} \Phi(t_i)\right), \quad (3.6)
\]

for all \(t = (t_1, \ldots, t_n)^T \in \mathbb{R}^n\). Let \(u, v\) be arbitrary real numbers and define \(s = (u^2 + v^2)^{1/2}\). Then

\[
\mathbb{E}(|\Phi(s) - \Phi(u)\Phi(v)|^2) = \mathbb{E}(|\Phi(s) - \Phi(u)\Phi(v)|^2|\Phi(s) - \Phi(u)\Phi(v)|) \tag{3.4}
\]

\[
= \mathbb{E}(|\Phi(s) - \Phi(u)\Phi(v)|^2|\Phi(-s) - \Phi(-u)\Phi(-v)|) 
\]

\[
= \mathbb{E}(\Phi(s)\Phi(-s)) - \mathbb{E}(\Phi(s)\Phi(-u)\Phi(-v)) 
- \mathbb{E}(\Phi(u)\Phi(v)\Phi(-s)) + \mathbb{E}(\Phi(u)\Phi(v)\Phi(-u)\Phi(-v)).
\]

Equation (3.6) allows us to compute the four terms on the lower right-hand side of the above expression:

\[
\mathbb{E}(\Phi(s)\Phi(-s)) = \phi((s, -s)^T (s, -s)) = \phi(2s^2), 
\]

\[
\mathbb{E}(\Phi(s)\Phi(-u)\Phi(-v)) = \phi((s, -u, -v)^T (s, -u, -v)) = \phi(s^2 + u^2 + v^2) = \phi(2s^2), 
\]

\[
\mathbb{E}(\Phi(u)\Phi(v)\Phi(-s)) = \phi((u, v, -s)^T (u, v, -s)) = \phi(u^2 + v^2 + s^2) = \phi(2s^2), 
\]

\[
\mathbb{E}(\Phi(u)\Phi(v)\Phi(-u)\Phi(-v)) = \phi((u, v, -u, -v)^T (u, v, -u, -v)) 
\]

\[
= \phi(u^2 + v^2 + u^2 + v^2) = \phi(2s^2). 
\]

As all these terms coincide we obtain

\[
\mathbb{E}(|\Phi(s) - \Phi(u)\Phi(v)|^2) = 0,
\]
3.1. Elliptical distributions

which entails that

\[ \Phi((u^2 + v^2)^{1/2}) = \Phi(s) = \Phi(u)\Phi(v), \]

with probability one. Therefore,

\[ \mathbb{P}\left( \Phi((u^2 + v^2)^{1/2}) = \Phi(u)\Phi(v) \right) = 1, \]

for any \( u, v \in \mathbb{R} \). As the rational numbers are countably many,

\[ \mathbb{P}\left( \Phi((u^2 + v^2)^{1/2}) = \Phi(u)\Phi(v), \text{ for any } u, v \in \mathbb{Q} \right) = 1, \]

and due to the continuity of \( \Phi \), it even holds that

\[ \mathbb{P}\left( \Phi((u^2 + v^2)^{1/2}) = \Phi(u)\Phi(v), \text{ for any } u, v \in \mathbb{R} \right) = 1. \]

With probability one, we then also have that

\[ \Phi(u) = \Phi(u)\Phi(0) = \Phi((u^2 + 0^2)^{1/2}) = \Phi(|u|), \]

and hence \( \Phi(u) = \Phi(-u) \), for all \( u \in \mathbb{R} \), as a consequence of which

\[ \Phi(u) = \Phi(-u) = \Phi(u), \]

for any \( u \in \mathbb{R} \), so that \( \Phi \) becomes a real-valued random function. Lemma A.1.1 then implies that there exists an \( a \in \mathbb{R}^+ \) such that

\[ \Phi(u) = \exp(-\frac{1}{2}au^2), \]

for all \( u \in \mathbb{R} \) with probability one. As \( a = -2 \log(\Phi(1)) \geq 0 \), \( a \) must be an non-negative \( \mathcal{F} \)-measurable random variable. Therefore,

\[ \mathbb{E}\left( \exp\left( i \sum_{i=1}^{n} t_i X_i \big| a \right) \bigg| a \right) = \mathbb{E}\left( \exp\left( i \sum_{i=1}^{n} t_i X_i \bigg| \mathcal{F} \right) \bigg| a \right) \]

\[ = \mathbb{E}\left( \prod_{i=1}^{n} \Phi(t_i) \bigg| a \right) \]

\[ = \mathbb{E}\left( \prod_{i=1}^{n} \exp\left( -\frac{1}{2}at_i^2 \bigg| a \right) \right) \]

\[ = \exp\left( -\frac{1}{2}a \sum_{i=1}^{n} t_i^2 \right), \]

for any \( t_1, \ldots, t_n \in \mathbb{R} \). Therefore, conditional on \( a \) the \( X_1, \ldots, X_n \) are independent and identically distributed as \( N(0,a) \). Let \( G \) be the distribution function of \( a \). Then

\[ \phi(t^T t) = \mathbb{E}\left( \exp\left( i \sum_{i=1}^{n} t_i X_i \right) \right) \]

\[ = \int_{0}^{\infty} \exp\left( -\frac{1}{2}r \sum_{i=1}^{n} t_i^2 \right) dG(r), \]

for all \( t = (t_1, \ldots, t_n)^T \in \mathbb{R}^n \), and therefore

\[ \phi(x) = \int_{0}^{\infty} \exp\left( -\frac{1}{2}rx \right) dG(r) = \int_{0}^{\infty} \exp\left( -\frac{1}{2}r^2x \right) dG_{\infty}(r), \]
for all \( x \in \mathbb{R} \), where the distribution function \( G_\infty \) is defined via \( G_\infty(r) := G(\sqrt{r}) \), for \( r \geq 0 \). For the converse of the proof let \( n \in \mathbb{N} \) be arbitrary, assume that there are \( n + 1 \) independent univariate random variables \( R, Y_1, \ldots, Y_n \) where \( Y_1, \ldots, Y_n \) are standard Gaussian and \( R \) is non-negative with distribution \( G_\infty \). Define \( X^{(n)} = (X_1, \ldots, X_n)^T \) via \( X_i := R Y_i \) for \( i = 1, \ldots, n \). Then \( X^{(n)} \) possesses \( t \mapsto \phi(t^T t) \) as characteristic function with \( \phi \) as in (3.3) and therefore \( \phi \in \Phi_n \). As \( n \) was arbitrarily chosen, we conclude that \( \phi \in \Phi_\infty \).

If we want to use elliptical distribution functions for the modelling of the relevant quantities in our credit setup, we need to make sure that the function \( \phi \) remains in the class of generators \( \Phi_n \), where \( n \in \mathbb{N} \) denotes the dimension of the underlying portfolio. After we have set up the general model in Section 5.2, in the sequel we will then be interested in deriving an approximation for the portfolio losses as the dimension \( n \) tends towards infinity. This approximation will help us to render the pricing of credit derivatives more tractable. Theorem 3.1.13 states that if we want to employ elliptical distributions within our model with an arbitrarily large dimension, it is necessary and sufficient to limit oneself to that subfamily of elliptical distributions which possesses a generator function \( \phi \) of the special form (3.3). Our focus will therefore lie on the so-called mixtures of Normal distributions, which is large enough to allow to model different characteristics such as heavy tails and tail dependence.

**Definition 3.1.14 (Mixtures of Normal distributions)**

An \( n \)-dimensional random vector \( X \) is said to follow a mixture of Normal distributions, if there exist a univariate, non-negative random variable \( R \geq 0 \), a covariance matrix \( \Sigma \in \mathbb{R}^{n \times n} \) and a Gaussian random vector \( Y \) with \( Y \sim N_n(0, \Sigma) \), such that \( R, Y \) are independent and \( X \overset{d}{=} RY \).

Indeed, the characteristic function \( \psi_X \) of an \( n \)-dimensional random vector \( X \) following a mixture of Normal distributions as in Definition 3.1.14 takes the form

\[
\psi_X(t) = \int_0^{\infty} \exp \left( -\frac{1}{2} t^T \Sigma t \right) dG(r), \quad t \in \mathbb{R}^n,
\]

for \( t \in \mathbb{R}^n \), where \( G \) denotes the distribution function of \( R \). Thus, we infer that \( X \) follows an elliptical distribution \( EC_n(0, \Sigma, \phi) \), where

\[
\phi(x) := \int_0^{\infty} \exp \left( -\frac{1}{2} r^2 x \right) dG(r), \quad x \in \mathbb{R},
\]

and therefore with \( \phi \in \Phi_\infty \) due to Theorem 3.1.13.

**Example 3.1.15** There are several well-known multivariate distribution functions that belong to the class of mixtures of Normal distributions: Let \( Y \sim N_n(0, \Sigma) \) for a covariance matrix \( \Sigma \in \mathbb{R}^{n \times n} \).
3.1. Elliptical distributions

1. Certainly, the centered multivariate Normal distribution functions belong to the mixtures of Normal distributions as one can simply set \( R := 1 \) in the previous definition and obtains \( X := RY = Y \sim N_n(0, \Sigma) \).

2. On setting \( R := \sqrt{m/S} \) where \( S \sim \chi_m \) is independent of \( Y \), for an arbitrary \( m \in \mathbb{N} \), we obtain the multivariate Student \( t \)-distribution with \( m \) degrees of freedom\(^3\) for \( X \), i.e.

\[
X := RY = \sqrt{m/Y} \sim Mt_n(m, 0, \Sigma).
\]

3. The centered generalised hyperbolic distributions also belong to the class of mixtures of Normal distributions. There, one assumes that \( R = \sqrt{U} \), where \( U \) is distributed according to a generalised inverse Gaussian distribution. More precisely, the density function \( f \) of \( R \) is defined as

\[
f(r) := \frac{(\gamma/\delta)^{\lambda}}{K_{\lambda}(\delta \gamma)} r^{2\lambda - 1} \exp \left( -\frac{1}{2} \left( \gamma^2 r^2 + \delta^2 \frac{1}{r^2} \right) \right), \text{ for } r > 0,
\]

where the Bessel function of the third kind (or MacDonald function) \( K_{\lambda} \) can be written as

\[
K_{\lambda}(x) = \frac{1}{2} \int_0^{\infty} u^{\lambda - 1} \exp \left( -\frac{1}{2} x \left( u + \frac{1}{u} \right) \right) du, \text{ for } x > 0,
\]

and where \( \lambda \in \mathbb{R}, \; \delta, \gamma > 0 \) (an in-depth discussion of Bessel functions can be found in [Wat66]). Then \( X := RY \) follows a centered generalised hyperbolic distribution, where the special case of \( \lambda = -\frac{1}{2} \) corresponds to the Normal Inverse Gaussian distributions. Note that the subclass of the Normal Inverse Gaussian distributions is stable under convolution if the parameters \( \beta \) and \( \gamma \) coincide for the distributions that are to be convoluted. In general this is not true for other subclasses of the class of generalised hyperbolic distributions.\(^4\)

In general, an elliptically distributed random vector \( X \sim EC_n(\mu, \Sigma, \phi) \) does not need to possess finite second moments. Indeed, the multivariate Student \( t \)-distribution with 2 df belongs to this class but does not possess finite second moments. Yet, existing second moments of the scaling variables guarantee the existence of the second moments of the elliptical vector.

**Theorem 3.1.16** Let \( X \) be spherically distributed with \( X = ru^{(n)} \sim S_n(\phi), \; r, u^{(n)} \) independent and \( \phi \in \Phi_n \). Then \( \text{Cov}(u^{(n)}) = \frac{1}{n} \) and \( \mathbb{E}(r^2) < \infty \) yields that \( X \) possesses finite first and second moments and

\[
\mathbb{E}(X) = 0 \quad \text{and} \quad \text{Cov}(X) = \frac{\mathbb{E}(r^2)}{n} I_n.
\]

---

\(^3\) In the following, the expression “degrees of freedom” will be shortened to “df”.

\(^4\) Generalized hyperbolic distributions were first introduced by Barndorff-Nielsen [Bar77]; see e.g. [BK02] for further details.
If $X$ follows a mixture of Normal distributions with $X = RY \sim S_n(\phi)$, $\phi \in \Phi_\infty$, $R, Y$ independent and $Y \sim N_n(0, I_n)$, then $\mathbb{E}(R^2) < \infty$ yields that $X$ possesses finite first and second moments and

$$\mathbb{E}(X) = 0 \quad \text{and} \quad \text{Cov}(X) = \mathbb{E}(R^2)I_n.$$ 

**Proof:** For the first part, let $Z \sim N_n(0, I_n)$ be a Gaussian vector. Then, as $\mathbb{P}(Z = 0) = 0$, Theorem 3.1.9 yields that $Z \overset{d}{=} ||Z||u^{(n)}$ with $||Z||$ being independent of $u^{(n)}$. We know that $||Z||^2 \sim \chi_n^2$ follows a Chi-Squared distribution with $n$ df. Thus, we have that $\mathbb{E}(||Z||^2) = n$. As $\mathbb{E}(||Z||) > 0$ and $\mathbb{E}(Z) = 0$, the independence of $||Z||$ and $u^{(n)}$ entails that

$$\mathbb{E}(X) = \mathbb{E}(r)\mathbb{E}(u^{(n)}) = \mathbb{E}(r)\frac{\mathbb{E}(Z)}{\mathbb{E}(||Z||)} = 0.$$ 

Additionally employing that $\mathbb{E}(u^{(n)}) = \mathbb{E}(Z) = 0$ and

$$I_n = \text{Cov}(Z) = \mathbb{E}(||Z||^2)\text{Cov}(u^{(n)}) = n\text{Cov}(u^{(n)})$$

yields

$$\text{Cov}(u^{(n)}) = \frac{1}{n},$$

and therefore

$$\text{Cov}(X) = \mathbb{E}(r^2)\text{Cov}(u^{(n)}) = \frac{\mathbb{E}(r^2)}{n}I_n.$$ 

For the second part, the independence of $R$ and $Y$ entails that

$$\mathbb{E}(X) = \mathbb{E}(R)\mathbb{E}(Y) = 0.$$ 

By the same independence, $\mathbb{E}(Y) = 0$ also yields

$$\text{Cov}(X) = \mathbb{E}(R^2)\text{Cov}(Y) = \mathbb{E}(R^2)I_n.$$

\[\square\]

**Theorem 3.1.17** Let the $n$-vector $X$ be elliptically distributed with $X = \mu + rA^Tu^{(k)} \sim EC_n(\mu, \Sigma, \phi)$, where $k = \text{rank}(\Sigma)$, $A \in \mathbb{R}^{k \times n}$, $A^T A = \Sigma$ and $\phi \in \Phi_n$. Then $\mathbb{E}(r^2) < \infty$ yields that $X$ possesses finite first and second moments and

$$\mathbb{E}(X) = \mu \quad \text{and} \quad \text{Cov}(X) = \frac{\mathbb{E}(r^2)}{\text{rank}(\Sigma)}\Sigma = -2\phi'(0)\Sigma.$$ 

If $X$ follows a mixture of Normal distributions with $X = RY \sim EC_n(0, \Sigma, \phi)$, $\phi \in \Phi_\infty$ and $Y \sim N_n(0, \Sigma)$, then $\mathbb{E}(R^2) < \infty$ yields that $X$ possesses finite first and second moments and

$$\mathbb{E}(X) = 0 \quad \text{and} \quad \text{Cov}(X) = \mathbb{E}(R^2)\Sigma = -2\phi'(0)\Sigma.$$
3.1. Elliptical distributions

**Proof:** The $k$-dimensional vector $Y$ defined by $Y := ru^{(k)} \sim \Phi_k(\phi)$ follows a spherical distribution and $X = \mu + A^T Y$. According to Theorem 3.1.16 it thus follows that

$$E(X) = \mu + A^T E(Y) = \mu$$

and

$$Cov(X) = A^T Cov(Y) A = \frac{E(r^2)}{n} \Sigma.$$

By $u_1$ we denote the first component of the $k$-vector $u^{(k)}$ and similarly by $Y_1$ the first component of $Y$. The characteristic function of $Y$ is then given via

$$\psi_Y(t) = \phi(t^T t) = \phi \left( \sum_{i=1}^{k} t_i^2 \right),$$

for $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$. As the second moments of $Y$ exist, we have

$$\frac{\partial^2}{\partial t_1^2} \psi_Y(t) = \frac{\partial^2}{\partial t_1^2} \phi \left( \sum_{i=1}^{k} t_i^2 \right) = 4t_1^2 \phi''(t^T t) + 2\phi'(t^T t),$$

for $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$. Therefore,

$$E(Y_1^2) = \frac{1}{t^2} \frac{\partial^2}{\partial t_1^2} \psi_Y(t) \bigg|_{t=0} = -2\phi'(0).$$

From Theorem 3.1.16 we know that $E(u_1) = 0$ and $E(u_1^2) = \text{Var}(u_1) = \frac{1}{k}$. Therefore,

$$E(r^2) \frac{1}{k} = E(r^2) E(u_1^2) = E(Y_1^2) = -2\phi'(0),$$

which concludes with

$$Cov(X) = -2\phi'(0) \Sigma.$$ 

In the special case of mixtures of Normal distributions we additionally obtain

$$Cov(X) = A^T Cov(Y) A = E(R^2) \Sigma.$$ 

\[ \square \]

In Section 5.3.1 we will derive a characterisation for mixtures of Normal distributions to display tail-dependence which will make use of results obtained before for the so-called density generators.

**Definition 3.1.18 (Density generators)**

Assume that the $n$-dimensional random vector $X \sim EC_n(0, \Sigma, \phi)$ with positive definite $\Sigma$ possesses a density function $h$ of the form $h(x) = |\Sigma|^{-\frac{1}{2}} \cdot \gamma(x^T \Sigma^{-1} x)$, for $x \in \mathbb{R}^n$, where $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ is a Borel measurable function. Then $\gamma$ is called the density generator of $X$. In this case, we also write $X \sim EC_n(0, \Sigma, \gamma)$. 


3.1. Elliptical distributions

Remark 3.1.19 If \(X \sim EC_n(0, \Sigma, \phi)\) possesses a density function it must necessarily be of the form in the above definition due to the structure imposed by elliptically contoured distributions: if \(X \sim EC_n(0, I_n, \phi)\) possesses a density \(h\) then the spherical property for the characteristic function \(\psi_X\), that is \(\psi_X(\Gamma t) = \psi_X(t) = \int_{\mathbb{R}^n} \exp(i\gamma^T x)h(x)dx\), for all \(\Gamma \in O(n)\) and all \(t \in \mathbb{R}^n\), implies that also \(h(\Gamma x) = h(x)\), for all \(\Gamma \in O(n)\) and all \(x \in \mathbb{R}^n\) (using the fact that \(|\det(\Gamma)| = 1\) after e.g. substituting \(y = \Gamma x\) in the above integral for \(\psi_X\) and by the uniqueness of the characteristic function). For any \(x, y \in \mathbb{R}^n\) with \(x^T x = y^T y\) there exists a \(\Gamma \in O(n)\) such that \(\Gamma x = y\), and for these values we obtain \(h(y) = h(\Gamma x) = h(x)\). Thus, \(x \mapsto h(x)\) is a function of \(x^T x\), denoted by \(\gamma\). The step from \(EC_n(0, I_n, \phi)\) to \(EC_n(0, \Sigma, \phi)\) is then simply taken via a density transformation.

Lemma 3.1.20 If \(X \overset{d}{=} RY \sim EC_n(0, \Sigma, \phi)\) with \(R \geq 0\), \(P(R = 0) = 0\), \(R \sim G\) independent of \(Y \sim N_n(0, \Sigma)\) and \(\Sigma\) positive definite, then the density generator of \(X\) becomes

\[
\gamma(t) = (2\pi)^{-n/2} \int_0^\infty \exp\left(-\frac{1}{2} \frac{t}{r^2}\right) dG(r), \text{ for } t \in \mathbb{R}^+,
\]

if \(E(R^n) < \infty\) holds true.

Proof: The independence of \(R, Y\) together with the Normal distribution of \(Y\) yields

\[
P(X \leq x) = P(RY \leq x) = E(P(RY \leq x | R))
\]

\[
= \int_0^\infty P(rY \leq x) dG(r)
\]

\[
= \int_0^\infty \int_{(-\infty, x]} n_n(y; 0, r^2\Sigma) dy dG(r)
\]

\[
= \int_{(-\infty, x]} ((2\pi)^n|\Sigma|)^{-\frac{1}{2}} \int_0^\infty \exp\left(-\frac{1}{2} \frac{y^T \Sigma^{-1} y}{r^2}\right) dG(r) dy.
\]

Therefore, the density function of \(X\) becomes

\[
h(x) := ((2\pi)^n|\Sigma|)^{-\frac{1}{2}} \int_0^\infty \exp\left(-\frac{1}{2} \frac{x^T \Sigma^{-1} x}{r^2}\right) dG(r)
\]

and thus the density generator arises as stated.

Additionally, \(h \in L_1\) if and only if \(\int_{\mathbb{R}^n} \gamma(x^T x)dx < \infty\). By virtue of Lemma 1.4 in
3.2. Regular variation and Karamata’s Theorem

[FKN90], p. 22, and substitution of \( y = 2ur^2 \) we obtain

\[
\int_{\mathbb{R}^n} \gamma(x^T x) dx = \frac{\pi^{n/2}}{\Gamma(n/2)} \int_0^\infty y^{n/2-1} \gamma(y) dy
\]

\[
= \frac{2^{-n/2}}{\Gamma(n/2)} \int_0^\infty \int_0^\infty y^{n/2-1} \exp(-\frac{1}{2} y/r^2) dy dG(r)
\]

\[
= \frac{1}{\Gamma(n/2)} \int_0^\infty r^n \int_0^\infty u^{n/2-1} \exp(-u) du dG(r)
\]

\[
= \int_0^\infty r^n dG(r)
\]

which is finite as demanded and thus \( h \) remains integrable.

3.2 Regular variation and Karamata’s Theorem

In this section we introduce the notion of regular variation. Its multivariate counterpart is the so-called tail-dependence which, as mentioned before, will be discussed in detail in the context of mixtures of Normal distributions in Section 5.3.1. The notion of regular variation goes back to Karamata [Kar30] (see [BGT89] for a monograph treatment of regular variation) and it represents one of the main concepts in the extreme value theory (see e.g. [Res87] or [MFE05]). After defining regularly varying functions we discuss several theorems that we will need in Section 5.3.1. Among these theorems is the well-known Karamata’s Theorem that helps to compare absolutely continuous distribution functions with their density functions in their tail behaviour at infinity.

**Definition 3.2.1 (Regular variation)**

Let \( f : \mathbb{R} \to \mathbb{R}^+ \) be a measurable function and \( X \) a random variable with distribution function \( F : \mathbb{R} \to [0,1] \). Then

1. \( f \) is called regularly varying at \( \infty \) with index \( \alpha \in \mathbb{R} \), in notation \( f \in RV_\infty^\alpha \), if for any \( t > 0 \)

\[
\lim_{x \to \infty} \frac{f(tx)}{f(x)} = t^\alpha;
\]

if \( \alpha = 0 \) then the function \( f \in RV_\infty^0 \) is called slowly varying at \( \infty \);

2. \( f \) is called regularly varying at \( 0 \) with index \( \alpha \in \mathbb{R} \), \( f \in RV_0^\alpha \), if for any \( t > 0 \)

\[
\lim_{x \to 0^+} \frac{f(tx)}{f(x)} = t^\alpha;
\]

if \( \alpha = 0 \) then the function \( f \in RV_0^0 \) is called slowly varying at \( 0 \);
3.2. Regular variation and Karamata’s Theorem

3. $X$ is called regularly varying at $\infty$ with index $\alpha \geq 0$, if $\bar{F} = 1 - F \in RV_{\alpha}^\infty$.

If $\alpha = 0$ then the random variable $X$ with $\bar{F} \in RV_0^\infty$ is called slowly varying at $\infty$.

4. $f$ is called O-regularly varying at $\infty$, in notation $f \in OR$, if for any $t \geq 1$

$$0 < \liminf_{x \to \infty} \frac{f(tx)}{f(x)} \leq \limsup_{x \to \infty} \frac{f(tx)}{f(x)} < \infty.$$ 

From the previous definition it directly follows that every regularly varying function is also O-regularly varying.

The following theorem is the famous Karamata’s Theorem, which was proved in 1930 by Karamata under assumptions of continuity (see [Kar30]) and by de Haan in 1970 for not necessarily continuous but at least measurable functions (see [dH70]). More precisely, the theorem stated here is only the first (easier) half of Karamata’s Theorem, as we won’t make use of the converse where one shows that only slowly varying functions behave in the way as stated below. Note, that we assume that all appearing functions are locally integrable, also on intervals including 0. More details on Karamata’s Theorem can for example be found in [BGT89] (Theorem 1.5.11, p. 28, and Theorem 1.6.1, p. 30) or in [Res87] (Karamata’s Theorem 0.6, p. 17).

**Theorem 3.2.2 (Karamata’s Theorem; direct half)**

Let a measurable function $l : \mathbb{R}^+ \to \mathbb{R}_0^+$ be slowly varying at infinity.

1. If $\alpha < -1$ then $\int_1^\infty t^{\alpha}l(t)dt$ is finite, for any $x \in \mathbb{R}^+$, $\int_x^\infty t^{\alpha}l(t)dt \in RV_{\alpha+1}^\infty$ and

$$\frac{x^{\alpha+1}l(x)}{\int_x^\infty t^{\alpha}l(t)dt} \to -\alpha - 1 \quad \text{if } x \to \infty.$$

2. If $\alpha \geq -1$ then $\int_0^x t^{\alpha}l(t)dt \in RV_{\alpha+1}^\infty$ and

$$\frac{x^{\alpha+1}l(x)}{\int_0^x t^{\alpha}l(t)dt} \to \alpha + 1 \quad \text{if } x \to \infty.$$

**Proof:** We want to give a proof for 1., which we have adapted from the proof of Proposition 1.5.10 in [BGT89], p. 27: set $\rho := \frac{1}{2}(\alpha + 1) < 0$ and $f(x) := x^{\frac{1}{2}(\alpha + 1)}l(x)$, for $x \in \mathbb{R}$. Then $f \in RV_{\rho}^\infty$ and

$$\frac{\int_x^\infty t^{\alpha}l(t)dt}{x^{\alpha+1}l(x)} + \frac{1}{\alpha + 1} = \int_1^\infty \left(\frac{f(ux)}{f(x)} - u^\rho\right) u^{\rho - 1}du. \quad (3.10)$$

One can show that if $f$ is a regularly varying function at $\infty$ with index $\rho < 0$, then $f(ux)/f(x)$ converges towards $u^\rho$ uniformly in $u$ on $[1, \infty)$ if $x \to \infty$ (see Theorem 1.5.2 in [BGT89], p. 22). As $\int_1^\infty u^{\rho - 1}du$ exists and represents an upper bound of the
3.2. Regular variation and Karamata’s Theorem

41

integral on the right-hand side of (3.10) if only \( x \) is large enough, the left-hand side of (3.10) tends to zero by dominated convergence. 

There exists a representation for every slowly varying function. This will be the issue of the next theorem, which will be used to compare the behaviour of slowly varying functions with power functions.

**Theorem 3.2.3 (Karamata’s Representation Theorem)**

A measurable function \( l : \mathbb{R}^+ \to \mathbb{R}_0^+ \) is slowly varying at infinity iff \( l \) can be represented as

\[
l(x) = c(x) \exp \left( \int_1^x s^{-1} \varepsilon(s) ds \right),
\]

for \( x > 0 \), where \( c : \mathbb{R}^+ \to \mathbb{R}_0^+ \), \( \varepsilon : \mathbb{R}^+ \to \mathbb{R}_0^+ \) are measurable functions and

\[
\lim_{x \to \infty} c(x) = c \in (0, \infty) \quad \lim_{x \to \infty} \varepsilon(x) = 0.
\]

**Proof:** If \( l \) satisfies the above representation, then for \( x > 0 \) and \( t > 0 \) we have

\[
\frac{l(tx)}{l(t)} = \frac{c(tx)}{c(t)} \exp \left( \int_t^{tx} s^{-1} \varepsilon(s) ds \right).
\]

From \( \lim_{x \to \infty} \varepsilon(x) = 0 \) we can deduce that for any \( \varepsilon > 0 \) there exists a \( t_0 > 0 \) such that

\[
\left| \int_t^{tx} s^{-1} \varepsilon(s) ds \right| \leq \varepsilon \left| \log(tx) - \log(t) \right| = \varepsilon \left| \log(x) \right|,
\]

for all \( t \geq t_0 \), so that the integral converges to 0 for \( t \to \infty \). Thus, \( l \) is slowly varying.

The converse makes use of part 2. of Karamata’s Theorem 3.2.2, as for \( l \) being slowly varying

\[
b(x) := \frac{xl(x)}{\int_0^x l(s)ds} \to 1 \quad \text{for} \quad x \to \infty.
\]

The function \( \varepsilon \) can then be defined via \( \varepsilon(x) := b(x) - 1 \), for \( x > 0 \), and thus tends to 1 for \( x \to \infty \). The function \( c \) is defined via \( c(x) := b(x) \int_0^1 l(s)ds \) so that \( c(x) \to c := \int_0^1 l(s)ds \) for \( x \to \infty \). With these settings one can deduce the desired representation. See the proof of the Karamata Representation in [Res87], pp. 17 - 18, for more details on the last step. \[\square\]

**Corollary 3.2.4** Let a measurable function \( l : \mathbb{R}^+ \to \mathbb{R}_0^+ \) be slowly varying at infinity and let \( \rho > 0 \) be arbitrary. Then \( \lim_{x \to \infty} x^\rho l(x) = \infty. \)

**Proof:** This is a direct consequence of Karamata’s Representation Theorem:

According to Theorem 3.2.3 there exist measurable functions \( c : \mathbb{R}^+ \to \mathbb{R}_0^+ \) and \( \varepsilon : \mathbb{R}^+ \to \mathbb{R}_0^+ \)
3.2. Regular variation and Karamata’s Theorem

\( \mathbb{R}^+ \to \mathbb{R}^+ \) with \( \lim_{x \to \infty} c(x) = c \in (0, \infty) \) and \( \lim_{x \to \infty} \varepsilon(x) = 0 \) such that

\[
l(x) = c(x) \exp \left( \int_1^x s^{-1} \varepsilon(s) ds \right),
\]

for \( x > 0 \). Then

\[
\frac{\log(x^\rho l(x))}{\log(x)} = \rho + \frac{\log(c(x))}{\log(x)} + \frac{\int_1^x s^{-1} \varepsilon(s) ds}{\log(x)},
\]

for \( x > 0 \). The second term on the right-hand side goes to 0 as \( \lim_{x \to \infty} c(x) = c \in (0, \infty) \).

For the third term, let \( \varepsilon > 0 \) be arbitrarily chosen. Then, as \( \lim_{x \to \infty} \varepsilon(x) = 0 \) there exists an \( x_0 > 1 \) such that \( 0 \leq \varepsilon(x) < \frac{\varepsilon}{2} \), for all \( x > x_0 \). Hence,

\[
0 \leq \int_{x_0}^x s^{-1} \varepsilon(s) ds < \frac{\varepsilon}{2} (\log(x) - \log(x_0)),
\]

for all \( x > x_0 \). Further, there exists an \( x_1 > 1 \) such that \( \int_1^{x_0} s^{-1} \varepsilon(s) ds / \log(x) < \frac{\varepsilon}{2} \), for all \( x > x_1 \). Set \( x_2 := \max\{x_0, x_1\} \). Therefore, for all \( x > x_2 \) we obtain

\[
0 \leq \frac{\int_1^x s^{-1} \varepsilon(s) ds}{\log(x)} \leq \frac{\int_1^{x_0} s^{-1} \varepsilon(s) ds}{\log(x)} + \frac{\varepsilon}{2} \left( 1 - \frac{\log(x_0)}{\log(x)} \right) \leq \varepsilon,
\]

as a consequence of which

\[
\frac{\log(x^\rho l(x))}{\log(x)} \to \rho \quad \text{or equivalently} \quad \frac{\log(x^\rho l(x))}{\log(x^\rho)} \to 1 \quad \text{if} \ x \to \infty.
\]

Then, for any \( 1 > \varepsilon > 0 \) there exists an \( x_3 > 1 \) such that \( 0 < (1 - \varepsilon) \leq \frac{\log(x^\rho l(x))}{\log(x^\rho)} \), which entails that \( (1 - \varepsilon) \log(x^\rho) \leq \log(x^\rho l(x)) \), both for all \( x > x_3 \). As \( \log(x^\rho) \to \infty \) if \( x \to \infty \), the claim follows.

We have named the following theorem a version of Karamata’s Tauberian Theorem as in the usual Karamata’s Tauberian Theorem the roles of the origin and infinity are interchanged. As it is pointed out in the remark after Theorem 2 in chapter XIII, §5, p.445, of [Fel71], the following theorem usually appears in two separate parts. While the implication from the Laplace-Stieltjes transform to the measure inducing function \( F \) is considered a Tauberian Theorem, the converse is usually related to as being an Abelian Theorem. In Section 5.3.1 we will need both parts for the discussion of tail-dependence.

**Theorem 3.2.5 (A version of Karamata’s Tauberian Theorem)**

Let \( F : \mathbb{R} \to \mathbb{R}^+ \) be a non-decreasing, right-continuous measurable function with \( F(x) = 0 \) for all \( x \leq 0 \). Then, for any \( \varepsilon > 0 \) there exists an \( x_1 > 1 \) such that \( 0 \leq \varepsilon(x) < \frac{\varepsilon}{2} \), for all \( x > x_1 \). Therefore, for all \( x > x_2 \) we obtain

\[
0 \leq \frac{\int_1^x s^{-1} \varepsilon(s) ds}{\log(x)} \leq \frac{\int_1^{x_0} s^{-1} \varepsilon(s) ds}{\log(x)} + \frac{\varepsilon}{2} \left( 1 - \frac{\log(x_0)}{\log(x)} \right) \leq \varepsilon,
\]

as a consequence of which

\[
\frac{\log(x^\rho l(x))}{\log(x)} \to \rho \quad \text{or equivalently} \quad \frac{\log(x^\rho l(x))}{\log(x^\rho)} \to 1 \quad \text{if} \ x \to \infty.
\]

Then, for any \( 1 > \varepsilon > 0 \) there exists an \( x_3 > 1 \) such that \( 0 < (1 - \varepsilon) \leq \frac{\log(x^\rho l(x))}{\log(x^\rho)} \), which entails that \( (1 - \varepsilon) \log(x^\rho) \leq \log(x^\rho l(x)) \), both for all \( x > x_3 \). As \( \log(x^\rho) \to \infty \) if \( x \to \infty \), the claim follows. \( \square \)
0 for all \( x < 0 \), and with Laplace-Stieltjes transform \( \pi_F \), i.e. \( \pi_F(s) = \int_0^{\infty} \exp(-sx)dF(x) \), which shall be finite for all large \( s \). Let a measurable function \( l : \mathbb{R} \rightarrow \mathbb{R}_0^+ \) be in \( RV_0^\infty \), and let \( c \geq 0, \rho \geq 0 \) be non-negative constants. Then the following statements are equivalent:

\[
F(x) \sim cx^\rho l(1/x)/\Gamma(1 + \rho) \quad \text{for } x \to 0^+,
\]
\[
\pi_F(s) \sim cs^{\rho - 1}l(s) \quad \text{for } s \to \infty.
\]

**Proof:** The theorem stems from [BGT89], p.38, Theorem 1.7.1’, which is essentially Theorem 3 in chapter XIII, §5, p.445, of [Fel71].

While the previous theorem allows one to compare the behaviour of a function with the behaviour of its Laplace-Stieltjes transform, the next theorem relates a density to the integral over this density.

**Theorem 3.2.6 (Monotone Density Theorem)**

Let \( U \) be given on \((0, \infty)\) by \( U(x) := \int_0^x u(y)dy \) for some \( u \in L^1(0, \infty) \). Let \( U(x) \sim cx^\rho l(x) \), for \( x \to 0^+ \), where \( c \in \mathbb{R}, \rho \geq 0, l \in RV_0^\rho \), and if \( u \) is non-decreasing on \((0, K)\) for some \( K > 0 \), then

\[
u(x) \sim cp^{\rho - 1}l(x) \quad \text{for } x \to 0^+.
\]

**Proof:** The proof is quite parallel to the proof of Theorem 1.7.2 in [BGT89], p. 39, which is the Monotone Density Theorem for \( x \to \infty \). Suppose that \( u \) is non-decreasing on \((0, K)\). If \( 0 < a < b < \infty \), then

\[
U(bx) - U(ax) = \int_{ax}^{bx} u(y)dy,
\]

so, for all \( x < K/\xi < K \),

\[
\frac{(b-a)u(ax)}{x^\rho l(x)} \leq U(bx) - U(ax) \leq \frac{(b-a)u(bx)}{x^\rho l(x)}.
\]

(3.11)

The middle term in (3.11) is equal to

\[
\frac{U(bx) - U(ax)}{x^\rho l(x)} = \frac{U(bx)}{(bx)^\rho l(x)} l(x) b^\rho - \frac{U(ax)}{(ax)^\rho l(x)} l(x) a^\rho \to c(b^\rho - a^\rho),
\]

as \( x \to 0^+ \), due to the assumptions on \( U \) and \( l \). Thus, the left inequality in (3.11) yields

\[
\limsup_{x \to 0^+} \frac{u(ax)}{x^{\rho - 1} l(x)} \leq \frac{c b^\rho - a^\rho}{b - a},
\]

(3.12)
and the right inequality in (3.11) yields
\[
\liminf_{x \to 0^+} \frac{u(bx)}{x^{\rho-1}I(x)} \geq \frac{b^\rho - a^\rho}{b - a}.
\] (3.13)
Setting \( a := 1 \) and letting \( b \to 1^+ \) in (3.12) yields
\[
\limsup_{x \to 0^+} \frac{u(x)}{x^{\rho-1}I(x)} \leq \epsilon \rho,
\]
whereas setting \( b := 1 \) and letting \( a \to 1^- \) in (3.13) gives
\[
\liminf_{x \to 0^+} \frac{u(x)}{x^{\rho-1}I(x)} \geq \epsilon \rho,
\]
and thus
\[
\lim_{x \to 0^+} \frac{u(x)}{x^{\rho-1}I(x)} = \epsilon \rho.
\]
Thus, the claim follows.

\[\square\]

3.3 Copulae

In this section we want to give a short introduction to copulae or copula functions, which will follow along the lines of Section 2.10 in [Nel99]. Copulae will not be used extensively in our setup, but will illuminate certain arguments brought forward in the following chapters.

Definition 3.3.1 An \( n \)-place real function \( H \) is a function whose domain, \( \text{Dom} H \), is a subset of \( \mathbb{R}^n \) and whose range, \( \text{Ran} H \), is a subset of \( \mathbb{R} \).

Definition 3.3.2 Let \( S_1, S_2, \ldots, S_n \) be nonempty subsets of \( \mathbb{R} \), and let \( H \) be an \( n \)-place real function such that \( \text{Dom} H = S_1 \times \cdots \times S_n \). Let \( B = [a, b] \) be an \( n \)-box all of whose vertices are in \( \text{Dom} H \). Then the \( H \)-volume of \( B \) is given by
\[
V_H(B) := \sum \text{sgn}(c)H(c),
\]
where the sum is taken over all vertices \( c \) of \( B \), and \( \text{sgn} \) is given by
\[
\text{sgn}(c) = \begin{cases} 
1, & \text{if } c_k = a_k \text{ for an even number of } k's \\
-1, & \text{if } c_k = a_k \text{ for an odd number of } k's.
\end{cases}
\]

Definition 3.3.3 An \( n \)-place real function \( H \) is \( n \)-increasing if \( V_H(B) \geq 0 \) for all \( n \)-boxes \( B \) whose vertices lie in \( \text{Dom} H \).
Definition 3.3.4 Let $S_1, S_2, \ldots, S_n$ be nonempty subsets of $\mathbb{R}$, and suppose that every $S_k$ has a least element $a_k$, $k = 1, \ldots, n$. We say that an $n$-place real function $H$ with domain $S_1 \times \cdots \times S_n$ is grounded if $H(t) = 0$ for all $t$ in $\text{Dom} H$ such that $t_k = a_k$ for at least one $k$.

Definition 3.3.5 Let $S_1, S_2, \ldots, S_n$ be nonempty subsets of $\mathbb{R}$, and suppose that every $S_k$ has a greatest element $b_k$, $k = 1, \ldots, n$. We then say that an $n$-place real function $H$ with $\text{Dom} H = S_1 \times \cdots \times S_n$ has margins, and the one-dimensional margins (or briefly margins) of $H$ are the functions $H_k$ given by:

$\text{Dom} H_k = S_k$, and $H_k(x) := H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n)$ for all $x$ in $S_k$.

Higher dimensional margins (briefly $k$-margins for $k \geq 2$) of $H$ are defined by fixing fewer places in $H$.

We can now define the class of copulae:

Definition 3.3.6 (Copulae) An $n$-dimensional copula, or briefly copula, is a function $C$ with the following properties:

1. $\text{Dom} C = I^n$

2. $C$ has (one-dimensional) margins $C_k$, $k = 1, \ldots, n$, which satisfy $C_k(u) = u$ for all $u$ in $S_k$, i.e. for every $u = (u_1, \ldots, u_n)$ in $I^n$, $C(u) = u_k$, if all coordinates of $u$ are 1 except $u_k$.

3. $C$ is grounded, i.e. for every $u = (u_1, \ldots, u_n)$ in $I^n$, $C(u) = 0$, if at least one coordinate of $u$ is 0.

4. $C$ is $n$-increasing, i.e. for every $a, b$ in $I^n$ such that $a \leq b$, $V_C([a,b]) \geq 0$.

Remark 3.3.7 With an appropriate extension of its domain to $\mathbb{R}^n$, every $n$-dimensional copula is an $n$-dimensional distribution function with margins which are uniform on $I$.

Definition 3.3.8 A general $n$-dimensional distribution function is a function $H$ with domain $\mathbb{R}^n$ such that

1. $H$ is $n$-increasing,

2. $H(t) = 0$ for all $t = (t_1, \ldots, t_n)$ in $\mathbb{R}^n$ such that $t_k = -\infty$ for at least one $k$, and $H(\pm \infty, \ldots, \pm \infty) = 1$. 
3.3. Copulae

**Theorem 3.3.9 (Sklar’s Theorem)**

Let $H$ be a general $n$-dimensional distribution function with margins $F_1, \ldots, F_n$. Then there exists an $n$-copula $C$ such that for all $x = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$,

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

(3.14)

If $F_1, \ldots, F_n$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $\text{Ran}F_1 \times \cdots \times \text{Ran}F_n$.

Conversely, if $C$ is a copula and $F_1, \ldots, F_n$ are distribution functions, then the function $H$ defined by (3.14) is a general $n$-dimensional distribution function with margins $F_1, \ldots, F_n$.

**Proof:** The proof uses two lemmata, where the first lemma establishes the statement for the so-called $n$-dimensional subcopulae, which are almost copulae but whose domains just need to be subsets of $I^n$. The second lemma shows that an $n$-dimensional subcopula can non-uniquely be extended to an $n$-dimensional copula. This extension is based on the so-called “multi-linear interpolation”. See for example [Nel99] or [Skl96] for more details.

One can “invert” Equation (3.14) and can thus express copulae in terms of an $n$-dimensional distribution function and the inverse functions of the margins.

**Definition 3.3.10** Let $F$ be a one-dimensional general distribution function. Then a quasi-inverse of $F$ is any function $F^{-}$ with domain $I$ with the following properties:

1. if $t$ is in $\text{Ran}F$, then $F^{-}(t)$ is any number $x$ in $\mathbb{R}$ such that $F(x) = t$, i.e., for all $t$ in $\text{Ran}F$, $F(F^{-}(t)) = t$;

2. if $t$ is not in $\text{Ran}F$, then $F^{-}(t) = \inf\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}$.

**Corollary 3.3.11** Let $H$ be a general $n$-dimensional distribution function with continuous margins $F_1, \ldots, F_n$, and choose the copula $C$ as in Equation (3.14) in Sklar’s Theorem. Let $F_1^{-}, \ldots, F_n^{-}$ be quasi-inverses of the functions $F_1, \ldots, F_n$, respectively. Then, for any $u = (u_1, \ldots, u_n)$ in $I^n$,

$$C(u) = H(F_1^{-}(u_1), \ldots, F_n^{-}(u_n)).$$

(3.15)

Let us consider $n$ random variables $X_1, \ldots, X_n$ on a common probability space $(\Omega, \Sigma, Q)$, with $n \geq 2$. We let $F_1, \ldots, F_n$ denote the distribution functions of the random variables $X_1, \ldots, X_n$, respectively. Sklar’s Theorem in terms of random variables and their distribution functions is a direct consequence of Sklar’s Theorem 3.3.9 and states the following:
3.3. Copulae

**Theorem 3.3.12** Let $X_1, \ldots, X_n$ be random variables with joint distribution function $H$ and distribution functions $F_1, \ldots, F_n$, respectively. Then there exists an $n$-dimensional copula $C$ such that

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \quad (3.16)$$

If $F_1, \ldots, F_n$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $\text{Ran} F_1 \times \cdots \times \text{Ran} F_n$.

**Definition 3.3.13** The copula $C$ in Theorem 3.3.12 will be called the copula of $X_1, \ldots, X_n$, and denoted $C_{X_1 \ldots X_n}$ when we explicitly want to identify the copula $C$ with the random variables $X_1, \ldots, X_n$.

**Theorem 3.3.14** Let $X_1, \ldots, X_n$ be random variables with continuous distribution functions $F_1, \ldots, F_n$, respectively, and with the copula $C_{X_1 \ldots X_n}$. If $\alpha_1, \ldots, \alpha_n$ are strictly increasing functions on $\text{Ran} X_1, \ldots, \text{Ran} X_n$, respectively, then $C_{\alpha_1(X_1) \ldots \alpha_n(X_n)} = C_{X_1 \ldots X_n}$. That is, $C_{X_1 \ldots X_n}$ is invariant under strictly increasing transformations of $X_1, \ldots, X_n$.

**Proof:** Let $G_1, \ldots, G_n$ denote the distribution functions of $\alpha_1(X_1), \ldots, \alpha_n(X_n)$, respectively. Since $\alpha_1, \ldots, \alpha_n$ are strictly increasing, $G_i(x_i) = Q(\alpha_i(X_i) \leq x_i) = Q(X_i \leq \alpha_i^{-1}(x_i)) = F_i(\alpha_i^{-1}(x_i))$, for every $x_i$ in $\mathbb{R}$ and every $i = 1, \ldots, n$. Hence, the distribution functions of $\alpha_1(X_1), \ldots, \alpha_n(X_n)$ are also continuous, and, by virtue of 3.3.12, the copula $C_{\alpha_1(X_1) \ldots \alpha_n(X_n)}$ uniquely exists. Thus, for any $x_1, \ldots, x_n$ in $\mathbb{R}$,

$$C_{\alpha_1(X_1) \ldots \alpha_n(X_n)}(G_1(x_1), \ldots, G_n(x_n)) = Q(\alpha_1(X_1) \leq x_1, \ldots, \alpha_n(X_n) \leq x_n)$$

$$= Q(X_1 \leq \alpha_1^{-1}(x_1), \ldots, X_n \leq \alpha_n^{-1}(x_n))$$

$$= C_{X_1 \ldots X_n}(F_1(\alpha_1^{-1}(x_1)), \ldots, F_n(\alpha_n^{-1}(x_n)))$$

$$= C_{X_1 \ldots X_n}(G_1(x_1), \ldots, G_n(x_n)).$$

As $G_1, \ldots, G_n$ are continuous, $\text{Ran} G_1 = \cdots = \text{Ran} G_n = I$, whence it follows that $C_{\alpha_1(X_1) \ldots \alpha_n(X_n)} = C_{X_1 \ldots X_n}$ on $I^n$. \hfill \Box
3.4 A limit theorem for martingales

Lemma 3.4.1 (Kronecker’s Lemma)
Let \((b_k)_{k \in \mathbb{N}}\) and \((x_k)_{k \in \mathbb{N}}\) be two real sequences, where \((b_k)_{k \in \mathbb{N}}\) is positive and tends monotonously towards infinity for \(k \to \infty\). Then, the convergence of
\[
\left( \sum_{k=1}^{n} b_k^{-1} x_k \right)_{n \in \mathbb{N}_0}
\]
for \(n \to \infty\) implies that \(\left( b_n^{-1} \sum_{k=1}^{n} x_k \right)_{n \in \mathbb{N}_0}\) converges towards zero for \(n \to \infty\).

Proof: The following proof is based on the proof of Lemma 1, Chapter VII.8, p. 239 in [Fel71]: Let us define the sequence \((\rho_n)_{n \in \mathbb{N}}\) by
\[
\rho_0 := 0 \quad \text{and} \quad \rho_n := \sum_{k=n}^{\infty} b_k^{-1} x_k,
\]
which is well defined due to the convergence of \(\left( \sum_{k=1}^{n} b_k^{-1} x_k \right)_{n \in \mathbb{N}_0}\). Then, for \(n \in \mathbb{N}\), we have
\[
b_n (\rho_{n-1} - \rho_n) = b_n \left( \sum_{k=n-1}^{\infty} b_k^{-1} x_k - \sum_{k=n}^{\infty} b_k^{-1} x_k \right) = x_n
\]
which entails that
\[
b_n^{-1} \sum_{k=1}^{n} x_k = b_n^{-1} \sum_{k=1}^{n} b_k (\rho_{k-1} - \rho_k) = b_n^{-1} \left( \sum_{k=0}^{n-1} b_{k+1} \rho_k - \sum_{k=1}^{n} b_k \rho_k \right)
\]
\[
= b_n^{-1} \left( \sum_{k=0}^{n-1} \rho_k (b_{k+1} - b_k) - b_n \rho_n + b_0 \rho_0 \right)
\]
\[
= -\rho_n + b_n^{-1} \sum_{k=0}^{n-1} \rho_k (b_{k+1} - b_k).
\]
Let \(\varepsilon > 0\) be arbitrarily chosen. As \(\rho_1 = \sum_{k=1}^{\infty} b_k^{-1} x_k\) converges according to the hypothesis, there must exist an \(N_\varepsilon \in \mathbb{N}\) such that \(|\rho_n| < \varepsilon/3\), for all \(n \geq N_\varepsilon\). Then, as \(b_n \to \infty\) there also exists an \(\tilde{N}_\varepsilon \in \mathbb{N}\) such that
\[
b_n^{-1} \sum_{k=0}^{N_\varepsilon-1} |\rho_k| (b_{k+1} - b_k) < \frac{\varepsilon}{3},
\]
for all \(n \geq \tilde{N}_\varepsilon\). Then the monotonicity of \((b_k)_{k \in \mathbb{N}}\) once more entails that, for all \(n > \tilde{N}_\varepsilon \lor N_\varepsilon\),
\[
\left| b_n^{-1} \sum_{k=1}^{n} x_k \right| \leq |\rho_n| + b_n^{-1} \sum_{k=0}^{N_\varepsilon-1} |\rho_k| (b_{k+1} - b_k) + b_n^{-1} \sum_{k=N_\varepsilon}^{n-1} |\rho_k| (b_{k+1} - b_k)
\]
\[
< 2 \frac{\varepsilon}{3} + b_n^{-1} (b_n - b_{N_\varepsilon}) \frac{\varepsilon}{3} = 2 \frac{\varepsilon}{3} + (1 - b_n^{-1} b_{N_\varepsilon}) \frac{\varepsilon}{3} < 3 \frac{\varepsilon}{3} = \varepsilon.
\]
Theorem 3.4.2 Let \((X_n)_{n \in \mathbb{N}_0}\) be a sequence of square-integrable random variables with 
\(E(X_{n+1}|X_n, \ldots, X_1) = 0\), for all \(n \in \mathbb{N}_0\), and define the martingale \((S_n)_{n \in \mathbb{N}_0}\) via \(S_n := \sum_{k=1}^{n} X_k\), for \(n \in \mathbb{N}_0\). If there exists a positive sequence \((b_k)_{k \in \mathbb{N}}\) which is monotonously increasing towards infinity, and if 
\[\sum_{k=1}^{\infty} b_k^{-2} E(X_k^2)\] 
converges then

1. the sequence 
\[\left(\sum_{k=1}^{n} b_k^{-1} X_k\right)_{n \in \mathbb{N}_0}\] 
converges almost surely for \(n \to \infty\), and

2. the sequence 
\[\left(b_n^{-1} \sum_{k=1}^{n} X_k\right)_{n \in \mathbb{N}_0}\] 
converges almost surely towards zero for \(n \to \infty\).

Proof: The proof can e.g. be found in Feller [Fel71], Theorem 3 of Chapter VII.9. The first part is a direct consequence of the martingale convergence theorem (Theorem 2 of Section VII.9 in [Fel71]) and the second part immediately follows from the first part via Kronecker’s Lemma 3.4.1. \(\square\)
3.4. A limit theorem for martingales
Chapter 4

Credit portfolio models

In the following we give an introduction to credit portfolio models which can also be used for the pricing of multi-name credit structures. Hereby, we will only focus on the firm-value or structural models, as the model we will present in the next chapters belongs to the class of multivariate structural models. We first introduce the general setup of this type of models, then review the fundamental Merton model [Mer74] and finally discuss some prominent CDO models that will serve us as benchmarks for our model in later chapters.

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space, where we consider \(\mathbb{P}\) to be a pricing measure. For the moment, let us consider a portfolio which contains \(n\) exposures, that is, \(n\) credit-risky securities which are assumed to stem from \(n\) different companies with \(n\) being quite large. These securities could be credit-risky bonds, loans or also single-name credit derivatives such as Credit Default Swaps. As we have detailed in Section 2.2, the central difficulty in modelling and pricing credit derivatives which are written on a portfolio with a large number of securities is rooted in the fact that one must model not only the behaviour of these many individual underlyings, but also how they interact with each other, that is, their dependence structure.

4.1 Basic setup of structural models

Common to all structural or firm-value models is the assumption that for each company \(j\) in question there exists a stochastic process \((A_{j,t})_{t\geq 0}\) that can be interpreted as the value of the company or of its assets over time. Generally speaking, this process triggers the company’s default if it falls too low, for example below a certain default threshold \((\bar{c}_j)_{t\geq 0}\), which can be a stochastic process itself. The point in time where the default takes place is usually called default time and is denoted by \(\tau_j\), for \(j = 1, \ldots, n\).

As we want to derive a model which enables us to price derivatives on a portfolio of credit-risky securities, we are interested in modelling the behaviour of the \(n\)-dimensional
4.2. The Merton model

The Merton model [Mer74] was the first model to use stochastic processes for the modelling of asset or firm values in order to present closed-form expressions for prices of credit-risky securities. Even though the Merton model was originally meant as a model for the pricing of defaultable zero coupon bonds, its principal idea, of modelling the firm value with a stochastic process that is triggering a default, remains the basis of all the structural models.

The main assumptions in Merton’s model were that the value process \((A_t)_{t \geq 0}\) of a specific company follows a geometric Brownian motion and that a company is only financed by one type of equity and one type of debt, namely the zero coupon bond which is to be priced. At the maturity \(T\) of the bond the company is obliged to pay back the notional amount \(K\) to the holder of the bond. It is natural to assume that this is only possible if the value of the firm at maturity is still at least as much as this amount \(K\). Therefore, the creditor receives either the notional amount \(K\) or the value of the firm \(A_T\) at time \(T\), whichever is less. From the point of view of the equity holder, he receives whatever is left after the debt has been repaid, that is the difference between the firm value \(A_T\) and the notional amount \(K\), if this value is positive, or nothing in the other case: \(\max\{A_T - K, 0\}\). The process \((A_t)_{t \geq 0}\) is modelled as a geometric Brownian motion, i.e. 

\[ A_t = \exp\left( r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t, \]

where \((W_t)_{t \geq 0}\) is a one-dimensional standard Brownian motion under the martingale measure \(\mathbb{P}\), \(r\) is the risk-free interest rate and \(\sigma\) denotes the volatility of \((A_t)_{t \geq 0}\). So the payment to the equity holder is exactly the payoff to a long position in a European call, one can employ the Black-Scholes options pricing theory. Herewith, one obtains a price for the value of the equity position and then also for the defaultable bond.
4.3 A multivariate Merton model

A crucial assumption underlying the Merton setup lies in the fact that a default of the company is only possible at the maturity date $T$. Even though the firm value was modelled over the whole period $[0, T]$ by a geometric Brownian motion which lives in the entire time interval $[0, T]$, the essential point in time is at the end $T$ of the period. Thus, the distributional behaviour of the process $(A_t)_{t \geq 0}$ is only relevant at time $T$. Such models are therefore often referred to as one-period models. The default time $\tau$ in this case then becomes

$$\tau := T \cdot \mathbb{1}_{\{A_T \leq K\}} + \infty \cdot \mathbb{1}_{\{A_T > K\}}.$$

4.3 A multivariate Merton model

If one wants to extend the Merton model to the multivariate case, one can assume that each of the firm-value processes $(A_{j,t})_{t \geq 0}$, $j = 1, \ldots, n$, follows a geometric Brownian motion, i.e.

$$A_{j,t} = \exp\left(\left(r - \frac{1}{2} \sigma_j^2\right) t + \sigma_j W_{j,t}\right),$$

where $(W^{(n)}_t = (W_{1,t}, \ldots, W_{n,t})^T)_{t \geq 0}$ is an $n$-dimensional Brownian motion under the martingale measure $\mathbb{P}$ with $\text{Cov}(W^{(n)}_t) = tC$, where $C$ is a correlation matrix, $r$ is again the risk-free interest rate and $\sigma_j$ denotes the volatility of $(A_{j,t})_{t \geq 0}$. Then the bond prices for the individual companies can be obtained in the very same way as in the univariate Merton case, regardless of the correlation structure underlying the multivariate Brownian motion. Likewise, only the marginal distributions of the processes $(A_{j,t})_{t \geq 0}$ at the possibly varying maturities $T_1, \ldots, T_n$ of the bonds are relevant for the valuation of these credit-risky assets.

Only if we want to model the portfolio underlying a credit derivative product such as a first-to-default swap or a CDO, the correlation between the components of $(W^{(n)}_t = (W_{1,t}, \ldots, W_{n,t})^T)_{t \geq 0}$ becomes crucial. Let $T$ be the maturity of such a portfolio credit derivative; then the default time of company $j$ is supposed to equal

$$\tau_j := T \cdot \mathbb{1}_{\{A_{j,T} \leq K_j\}} + \infty \cdot \mathbb{1}_{\{A_{j,T} > K_j\}},$$

where $K_j$ is the corresponding obligation that needs to be paid back at time $T$, for $j = 1, \ldots, n$. In this case, the modelling of the vector $A^{(n)}_T = (A_{1,T}, \ldots, A_{n,T})$ becomes relevant, whose components follow a Log-Normal distribution, thus

$$S_{j,T} = \log(A_{j,T}) \sim N\left(\left(r - \frac{1}{2} \sigma_j^2\right) T, \sigma_j^2 T\right)$$

for all $j = 1, \ldots, n$. Therefore,

$$S^{(n)}_T = (S_{1,T}, \ldots, S_{n,T}) \sim N_n(\mu_S T, \Sigma_S T),$$

where $\mu_S := (r - \frac{1}{2} \sigma_1^2, \ldots, r - \frac{1}{2} \sigma_n^2)^T$ and $\Sigma_S := (\sigma_1^2, \ldots, \sigma_n^2) C(\sigma_1^2, \ldots, \sigma_n^2)^T$. This distribution function can then be used to approach the pricing of various portfolio credit derivatives.
Finally, the correlations between the underlying Brownian motions \((W_{1,t})_{t \geq 0}, \ldots, (W_{n,t})_{t \geq 0}\) are often referred to as asset correlations, even though these are certainly not equal to the correlations between the asset-value processes \((A_{1,t})_{t \geq 0}, \ldots, (A_{n,t})_{t \geq 0}\).

### 4.4 One-period credit portfolio models

One of the main assumptions of the Merton model in the previous two sections was to assume that a default of company \(j\) can only take place at the maturity \(T\) of the credit derivative in question. In the following, we will refer to this assumption as the one-period assumption. We have seen that under this assumption the distribution of the vector of the log-returns \(S_T^{(n)}\) at maturity \(T\) is all that matters. However, having assumed that the underlying asset-value processes \((A_{j,t})_{t \geq 0}, j = 1, \ldots, n,\) all follow geometric Brownian motions, the resulting distribution of the vector \(S_T^{(n)}\) is already fixed and does not allow for further flexibility in specifying other distributional characteristics. Therefore, in the one-period models one usually does not specify a dynamic for the processes \((A_{1,t})_{t \geq 0}\) or \((S_{j,t})_{t \geq 0}, j = 1, \ldots, n,\) but rather directly assumes a specific structure and distribution solely for the log-return vector \(S_T^{(n)}\), which then also determines the distribution of the asset vector \(A_T^{(n)}\). The length of the period is often simplified to being one, that is \(T = 1,\) and often the time index \(T\) will be dropped in order to emphasize that one concentrates on the distribution of \(S_T^{(n)} = S^{(n)} = (S_1, \ldots, S_n)\) rather than on the dynamics of the entire process \((S_t^{(n)})_{t \geq 0}\). In order to ease the computations, one usually standardises the distribution functions of the log-returns \(S_j, j = 1, \ldots, n,\) to have zero means and unit variances. This standardisation does not have any impact on the mechanism that triggers the defaults, as the default thresholds can then likewise be shifted and scaled as the log-returns. Yet, the standardisation also simplifies the comparison of different specifications of the distributions of \(S^{(n)}\). The random variables \(S_1, \ldots, S_n\) are also referred to as latent variables that drive the defaults of the corresponding firms, as they usually cannot be observed directly or as their interpretation as log-returns should not be stressed too much.

**General multivariate Gaussian models**

The easiest generalisation of the Merton setup to a multivariate one-period model is to allow the latent variables vector \(S^{(n)}\) to follow an arbitrary multivariate Normal distribution, that is

\[
S^{(n)} \sim N_n(\mu, \Sigma)
\]

for appropriate values for the mean vector \(\mu \in \mathbb{R}^n,\) usually \(\mu = 0,\) and the covariance matrix \(\Sigma_n \in \mathbb{R}^{n \times n},\) which is usually equal to a correlation matrix with unit values on the main diagonal. Of course, this also encompasses the Merton setup of the previous sections. Yet, since the number of correlation parameters can be many to estimate or
to calibrate, one usually relies on factor models which reduce the complexity of the variance/covariance structure.

4.5 One-period factor models

In one-period structural models, if the dimension of the portfolio underlying a credit derivative becomes very large, one has to be especially careful with the modelling of the interaction of the many individual names in the portfolio with each other. Already the estimation of the asset correlations can require quite an effort with increasing portfolio size as the number of correlation parameters rises quadratically \((n^2 + n)/2\) in the number of exposures \((n)\). As we have mentioned before in Section 2.2, the reference portfolio of the liquid and widely used iTraxx tranches consist of 125 names, so that one already would have to specify 7875 correlation parameters. Additionally, it is reasonable to assume that many of the obligors are dependent on each other or are influenced by the behaviour of common macro variables. Therefore, one often applies factor models for the modelling of the latent variables, which serve as a means of reducing the complexity of the dependence structure by reducing the number of independent components that drive the behaviour of the latent variables.

Within such a factor model, one usually assumes the following linear regression structure for the latent variables:

\[
S_j = \beta_j^T M + \varepsilon_j, \tag{4.1}
\]

where \(M\) is an \(m\)-dimensional vector, which is called a factor and is interpreted to describe a systematic behaviour or macro variable and whose influence on each return value \(S_j\) is determined by the deterministic factor loading \(\beta_j \in \mathbb{R}^m\), and where \(\varepsilon_j\) represents the firm-specific or idiosyncratic stochastic risk component influencing each company \(j\), for \(j = 1, \ldots, n\). The factor \(M\) and the vector of the idiosyncratic risk variables \((\varepsilon_1, \ldots, \varepsilon_n)\) are considered to be independent, so that the variance of \(S_j\) simplifies to the sum of the variances of the two additive components \(\beta_j^T M\) and \(\varepsilon_j\). The number of the factors can vary widely and can depend greatly on how such factor models are applied. If one actually uses estimates of common underlying industry or country indices as e.g. in the KMV model, one allows for as many factors as are needed to explain the individual behaviour, while one needs to reduce the number of factors to a minimum when trying to calibrate the model to a very limited number of correlation products such as the five iTraxx tranches.

4.5.1 Gaussian factor models: Vasicek, KMV and CreditMetrics

The use of factor models with a Gaussian specification in the context of portfolio loss distributions became particularly popular due to the fundamental papers by Vasicek [Vas87] and [Vas91]. Vasicek’s model [Vas87] is based on the extension of the Merton model to the multivariate case as in Section 4.3, where he assumes that the \(n\) Brownian
motions \((W_{1,t})_{t\geq 0}, \ldots, (W_{n,t})_{t\geq 0}\) are correlated with the same correlation value \(\rho^2\). Due to this assumption, he can express each Brownian motion as the linear combination of two independent Brownian motions, that is
\[
W_{1,t} = \rho M_t + \sqrt{1 - \rho^2} \varepsilon_{1,t},
\]
where \((M_t)_{t\geq 0}, (\varepsilon_{1,t})_{t\geq 0}, \ldots, (\varepsilon_{n,t})_{t\geq 0}\) are independent Brownian motions. Given this structure and under homogeneity assumptions, he then derives the probability that the percentage gross loss on a portfolio at a specific time \(T\) equals a certain fraction, where he makes use of the independence and the identical distribution of the standard Gaussian random variables \(T^{-1/2}\varepsilon_{1,T}, \ldots, T^{-1/2}\varepsilon_{n,T}\). In his continuation in [Vas91], he then analysed the convergence of the percentage loss distribution when the portfolio size tends towards infinity. Again, this was performed in a one-factor Gaussian setup as just described.

The Gaussian factor model has also been used by several industry models such as the Global Correlation Model by Moody’s KMV or the factor model by CreditMetrics. The Global Correlation Model (see e.g. Chapter 9 in [CGM01] or Chapter 1 in [BOW03]) assumes that the log-return of company \(j\) can be decomposed as in Equation (4.1), but where the factor \(M\) can also differ from company to company, that is
\[
S_j = \beta_j M_j + \varepsilon_j,
\]
where \(M_j\) and \(\varepsilon_1, \ldots, \varepsilon_n\) are assumed to be independent centered Gaussian random variables with \(\text{Var}(M_j) = 1\) and \(\text{Var}(\varepsilon_j) = 1 - \beta_j^2\), such that the \(S_j\)’s are standard Gaussian random variables. The factor \(M_j\) is called the composite factor of firm \(j\), as one supposes that it is the weighted sum of indices that describe various industries or countries and which do not differ with the individual firms. These indices are then further decomposed (e.g. via a principal component analysis) into a representation where each of these industry or country indices can be expressed as the weighted sum of independent so-called global factors plus a residual component. Note that the industry and country indices do not need to be independent: we only have this property on the level of the global factors. Once more, if the asset values of a company \(j\) lie below a certain threshold, the company is considered to have defaulted. The default threshold in the KMV terminology is called the default point of a firm.

The model of CreditMetrics (see [BFG97]) works in a similar way as the KMV model, since it also assumes a Gaussian factor structure for the latent variables. Yet, the factors used within the CreditMetrics model are indices that incorporate both a specific industry component and a specific country component (see pp. 166 – 170 in [BFG97]). While the KMV model actually tries to calibrate their parameters such as the asset correlations to actual asset-value processes, CreditMetrics approximates these asset correlations with the use of equity correlations by calibrating the model to equity processes. These equity processes are more easily available even for smaller firms, while within the KMV setup one has to derive the unobservable asset-value processes from equity processes and other relevant market information. The documentation for the CreditMetrics model has been
4.5. One-period factor models

freely accessible, so that it has influenced the developments of several other portfolio credit-risk models.

Klaassen et al. [KLSS01] also placed themselves within a setup similar to the CreditMetrics model. They assumed that the latent variables have the factor representation of Equation (4.1) where the \( n \)-dimensional factor \( M \) follows a Normal distribution \( N_m(0, \Omega_M) \) with \( \Omega_M \) being an appropriate covariance matrix, and the idiosyncratic variables \( \varepsilon_1, \ldots, \varepsilon_n \) following univariate Normal distributions with \( \text{Var}(\varepsilon_j) = \omega_j > 0 \), for all \( j \). Klaassen et al. not only allow for the two states default or non-default, but also analyse credit losses due to rating migrations in a similar way as Bhatia et al. in their CreditMetrics model (see Section 8.4 in [BFG97]). They then derive a large homogeneous portfolio approximation result as an extension to the above discussed result by Vasicek [Vas91], but now within this multi-factor Gaussian structure. In the rest of this document, we will use some of the notation of this paper for our analysis, and will largely extend this model with the use of elliptical distributions.

Li [Li00] analysed the dependence structure of the default times as they follow from the CreditMetrics framework. If the default time of company \( j \) is given via

\[
\tau_j = T \mathbb{1}_{\{S_j \leq c^j\}} + \infty \mathbb{1}_{\{S_j > c^j\}},
\]

with appropriate default thresholds \( c^1, \ldots, c^n \), where the vector \( S^{(n)} = (S_1, \ldots, S_n) \) follows an \( n \)-dimensional Gaussian distribution \( N_n(0, \Sigma_n) \) with \( \Sigma_n \) being a correlation matrix, then the Gaussian dependence of the latent variables is partially preserved with the default times. More precisely, as \( \{\tau_j \leq T\} = \{S_j \leq c^j\} \), for all \( j \), it follows that the joint probability that all the default times lie before time \( T \) can be fully described by a Gaussian copula which is evaluated at the individual probabilities that each single variable defaults before time \( T \):

\[
P(\tau_1 \leq T, \ldots, \tau_n \leq T) = C_{\text{Gauss}, \Sigma_n}(P(\tau_1 \leq T), \ldots, P(\tau_n \leq T)),
\]

where \( C_{\text{Gauss}, \Sigma_n} \) is the Gaussian copula with the asset correlation matrix \( \Sigma_n \). Li therefore proposes generating default times not only for the time horizon \( T \), but also for any other time \( t \), by the use of Gaussian copulae whose correlation matrix stems from the asset correlations as above. By analysing the effects of the CreditMetrics framework on the default times, Li established a link between the structural Gaussian factor model and the intensity-based model where default times are linked by a Gaussian copula. However, his analysis can be extended to models that use copulae other than the Gaussian copula.

The models that we have discussed in this section, especially the models by Moody’s KMV and CreditMetrics, are models that are especially suitable for estimating the credit risk within a portfolio. They were not introduced for the main purpose of pricing portfolio credit derivatives such as CDOs. However, central components of these models were used for the development of pricing models which are more suited to describing the behaviour of such credit derivatives.
4.5. One-period factor models

4.5.2 Gaussian one-factor model, implied and base correlations

The multi-factor credit portfolio models of the previous section are widely used for risk-management purposes. However, the Gaussian one-factor model, where one additionally assumes that only one single correlation parameter describes the dependence structure sufficiently well, has become the standard model for the pricing of CDOs:

$$S_j = \rho M + \sqrt{1 - \rho^2} \varepsilon_j, \quad \text{for } j = 1, \ldots, n,$$

where $\rho = \text{Cov}(S_j, M) = \text{Cor}(S_j, M) \in [0, 1]$ is the constant correlation parameter between the latent variables $S_j$ and $M$, where $M, \varepsilon_1, \ldots, \varepsilon_n$ are independent standard Gaussian random variables, and thus also the univariate distributions of the latent variables are set with $S_j \sim N(0,1)$, for all $j$. Note that in this case $\text{Cov}(S_i, S_j) = \text{Cor}(S_i, S_j) = \rho^2$ holds, for all $i \neq j$. If one then wants to calibrate this model to given univariate default probabilities and correlation products such as the five iTraxx CDO tranches, one can make use of the fact that the univariate distribution function of the $S_j$’s are independent of the correlation parameter $\rho$:

$$pd_j = P(S_j \leq c_j) = \Phi(c_j),$$

and thus we obtain $c_j = \Phi^{-1}(pd_j)$, where $pd_j$ is the probability that firm $j$ will default sometime between now and the planning horizon. It then remains to choose the correlation parameter $\rho$, which is a characteristic of the portfolio underlying the CDO, in such a way that the one-factor Gaussian model correctly reproduces all the prices of the various tranches (five tranches in the case of the tranched iTraxx) that one observes at a specific point in time on the market. As one cannot expect to perfectly match all tranches at the same time with the Gaussian model and just one correlation parameter, one therefore must rely on particular distance measures in order to judge the fit of the model with a particular choice of the parameter $\rho$ to the market prices. This distance measure could for example be given by the sum of the squared (absolute, resp.) errors

$$\sum_{i=1}^5 \left( \zeta_{i, \text{market}} - \zeta_{i, \rho} \right)^2 \left( \sum_{i=1}^5 \left| \zeta_{i, \text{market}} - \zeta_{i, \rho} \right|, \text{resp.} \right),$$

where $\zeta_{i, \text{market}}$ denotes the market quote for the $i$-th tranche and $\zeta_{i, \rho}$ the model price for this tranche which depends on $\rho$, for $i = 1, \ldots, 5$, or could also be in relative terms such as $\sum_{i=1}^5 \left| \zeta_{i, \text{market}} - \zeta_{i, \rho} \right| / \zeta_{i, \text{market}}$. The choice of $\rho$ that minimizes one of these distances is considered the best fit with respect to this distance measure.

Similar to the concept of implied volatilities of options under the Black-Scholes model, one can introduce the concept of implied correlations under the standard Gaussian one-factor model that we have just described. Here, for every single tranche $i$ one separately chooses the optimal model parameter $\rho$ that replicates the market price for that particular tranche the best within the standard Gaussian model. The respective value $\rho^2$ is then called the implied correlation for that particular tranche. Even though the correlation parameter is a parameter that describes the underlying portfolio and should not differ from tranche to tranche, we usually obtain a non-constant correlation parameter for the
4.5. One-period factor models

various tranches, which is an indication that the Gaussian model is not the correct model for the pricing of CDOs.

<table>
<thead>
<tr>
<th>Tranche ranges</th>
<th>0% - 3%</th>
<th>3% - 6%</th>
<th>6% - 9%</th>
<th>9% - 12%</th>
<th>12% - 22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>iTraxx Market</td>
<td>20.04</td>
<td>56.04</td>
<td>16.65</td>
<td>9.17</td>
<td>2.57</td>
</tr>
<tr>
<td>Implied $\rho$</td>
<td>43.28%</td>
<td>28.98%</td>
<td>36.94%</td>
<td>43.68%</td>
<td>46.40%</td>
</tr>
<tr>
<td>Implied correlation</td>
<td>18.73%</td>
<td>8.40%</td>
<td>13.65%</td>
<td>19.08%</td>
<td>21.53%</td>
</tr>
</tbody>
</table>

Table 4.1: The implied $\rho$ and the implied correlation $\rho^2$ for each of the five iTraxx Europe Series 3 tranches on the 6th of April 2006.

Table 4.1 shows a typical example of how the implied correlations vary over the 5 tranches: while the equity tranche price implies a moderate value for $\rho^2$ of 18.73%, the junior tranche (3% - 6%) requires a relatively low $\rho^2$ of 8.4%. From here on, the implied correlation curve is increasing over 13.65% and 19.08%, to finally reach its highest point with the most senior tranche (12% - 22%) at 21.53%. This again reflects the deficiencies that we have mentioned in Section 2.2: when we use the Gaussian model with just one correlation parameter for all tranches which could be chosen such that one of the distance measures above is minimized or such that the equity tranche is perfectly fitted, there is a tendency to overestimate the more mezzanine tranches, i.e., the 3% - 6% and 6% - 9% tranches, and to underestimate the more senior tranches, i.e. the 9% - 12% and 12% - 22% tranches. The first tranche implies a correlation of 18.73% which is too high for the 3% - 6% and 6% - 9% tranches, and too low for the 9% - 12% and 12% - 22% tranches.\footnote{Note that for the concepts of implied correlations and base correlations one usually assumes that the recovery rate is constant at 40% for all obligors, which is consistent with the rules of the iTraxx contracts, and that the default probabilities do not differ from firm to firm, but can be deduced from the average market spreads of the single-name CDSs in the iTraxx reference portfolio. Additionally, one usually sets the risk-free interest rate used for discounting equal to 0% (see p. 24 in [ABMW04] for a table of the standard assumptions for the computation of base correlations).}

In 2004, McGinty et al. at J.P. Morgan (see [ABMW04]) introduced the concept of base correlations. They argued that the procedure of finding an implied correlation could lead to problems, as there might be two possible solutions for the implied correlation, especially for the mid-level tranches such as the mezzanine tranche. Additionally, the concept of implied correlations does not lead to a straightforward procedure of how to price bespoke CDOs on the same underlying portfolio, such as tranches covering e.g. 4% - 7% of the portfolio losses. Instead of computing the implied correlations from the prices of tranches of the form $\alpha_{i-1} - \alpha_i$, they suggested computing the implied correlations of the prices of “base” tranches of the form 0% - $\alpha_i$, which they then call base correlations (note that the $\alpha_i$’s are values in [0,1] and again represent the respective percentages of the portfolio losses). Even though, these base tranches are not traded and there are thus no prices quoted for them, one can bootstrap the prices of such base tranches 0% - $\alpha_i$ from the prices of the traded tranches. As we will see later on in Chapter 7, for a specific tranche that covers between $\alpha_{i-1}$ and $\alpha_i$ of the portfolio losses, the expected portfolio losses that fall into this range of $(\alpha_{i-1}, \alpha_i)$ are the main building blocks for
pricing this tranche. The bootstrapping method then makes use of the property that the expected losses in \((\alpha_{i-1}, \alpha_i]\) can be decomposed into the expected losses in the base tranche \((0, \alpha_i]\) minus those in the base tranche \((0, \alpha_{i-1}\), as

\[
(\min\{x, \alpha_i\} - \alpha_{i-1})^+ = (\min\{x, \alpha_i\} - 0)^+ - (\min\{x, \alpha_{i-1}\} - 0)^+
\]

\[
= \min\{x, \alpha_i\} - \min\{x, \alpha_{i-1}\},
\]

for all \(x \in [0, 1]\) (see Chapter 7 for more details on the notation). We can therefore start with the traded base tranche \((0, \alpha_1]\), then use this property for the traded tranche \((\alpha_1, \alpha_2] = (0, \alpha_2] \setminus (0, \alpha_1]\) to obtain the value of the non-traded tranche \((0, \alpha_2]\), then for \((\alpha_2, \alpha_3] = (0, \alpha_3] \setminus (0, \alpha_2]\) and so on. When we consider such a base tranche \((0, \alpha_i]\), the upper tranche limit \(\alpha_i\) is called a detachment point. Usually, the base correlation curve is a strictly increasing function in the detachment points, and thus avoids the problem of non-uniqueness. The concept of base correlations then allows one to give prices for bespoke tranches such as the 4\% - 7\% tranche, by taking the interpolated base correlation values for the 0\% - 4\% and the 0\% - 7\% tranches and then performing the bootstrapping method backwards.

Table 4.2 gives the values of the base correlations for the iTraxx tranche prices that we have already utilized before. We can clearly see, that the base correlations increase in the detachment points.

<table>
<thead>
<tr>
<th>Detachment Points</th>
<th>3%</th>
<th>6%</th>
<th>9%</th>
<th>12%</th>
<th>22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base (\rho)</td>
<td>43.28%</td>
<td>53.99%</td>
<td>61.28%</td>
<td>67.40%</td>
<td>78.53%</td>
</tr>
<tr>
<td>Base correlation</td>
<td>18.73%</td>
<td>29.15%</td>
<td>37.55%</td>
<td>45.42%</td>
<td>61.67%</td>
</tr>
</tbody>
</table>

Table 4.2: The base \(\rho\) and the base correlation \(\rho^2\) for each of the five iTraxx Europe Series 3 tranches on the 6th of April 2006.

Figure 4.1: The implied and the base correlations curves for the five iTraxx Europe Series 3 tranches on the 6th of April 2006.

If the Gaussian one-factor model with one single correlation parameter perfectly fitted the observed data and the interest rates vanished, the implied correlation curve as well as the base correlation curve would be flat for all tranches, resp. for all detachment points.
4.5. One-period factor models

Figure 4.1, which holds the two curves, the implied correlation curve on the left-hand side and the base correlation curve on the right-hand side, clearly indicates that this is not true. Due to the way these curves are formed, the shape of the implied correlation curve is referred to as (implied) correlation smile, and the shape of the base correlation curve as (base) correlation skew. Later on in Chapter 7, the ability of a model that one has calibrated to the iTraxx market data to produce such similar correlation smiles or skews will be analysed, as they indicate whether a model produces tranche prices which are similar to the market quotes.

4.5.3 Double $t$-distribution copula model

Hull and White [HW04] proposed a “double $t$-distribution copula” model which primarily uses univariate Student $t$-distributions. They assume that the latent variables are given via the following factor model:

$$S_j = \beta_j M + \varepsilon_j,$$

where $M$ and $\varepsilon_1, \ldots, \varepsilon_n$ are assumed to be independent centered random variables with $\text{Var}(M) = 1$ and $\text{Var}(\varepsilon_j) = 1 - \beta_j^2$. They then analysed the effects of using combinations of univariate Normal distributions and univariate Student $t$-distributions with different df for the factor $M$ and the idiosyncratic risks $\varepsilon_1, \ldots, \varepsilon_n$, when applying the model on the pricing of $n$-th-to-default swaps or CDOs. The disadvantage of using independent Student $t$-distributions for both components is that their linear combination will not fall into a known distribution class and will certainly not be $t$-distributed again, which entails some additional computational effort when applying this model to credit derivatives. However, the reason for their rather ad-hoc usage of Student $t$-distributed random variables, especially for the factor $M$, was that the univariate $t$-distribution displays heavy tails and that they observed that they can produce a “greater likelihood of a clustering of early defaults for several companies”.

4.5.4 Normal Inverse Gaussian factor model

In their Normal Inverse Gaussian (NIG) model, Kalemanova et al. [KSW07] proposed to use Normal Inverse Gaussian distributions for the factor and the idiosyncratic component, that is

$$S_j = \beta_j M + \varepsilon_j,$$

Note that we have changed the notation to make the setup comparable to our notation. The distributional properties of the appearing quantities remain the same.

See page 7 in [HW04].

A random variable $X$ follows an NIG distribution with real parameters $a, b, \mu, \delta$, i.e. $X \sim \text{NIG}(a, b, \mu, \delta)$, where $0 \leq |b| < a$, and $\delta > 0$, if its moment generating function $m : \mathbb{R} \to \mathbb{R}_+^*$ is given via $m(t) = \mathbb{E}(\exp(tX)) = \exp \left( \mu t + \delta (\gamma - \sqrt{\alpha^2 - (b + t)^2}) \right)$, where $\gamma := \sqrt{\alpha^2 - b^2}$. See also Example 3.1.15 for the multivariate NIG distribution.

Note that we have changed the notation to make the setup comparable to our notation. The distributional properties of the appearing quantities remain the same.
where $M$ and $\varepsilon_1, \ldots, \varepsilon_n$ are assumed to be independent centered Normal Inverse Gaussian random variables with $\text{Var}(M) = 1$ and $\text{Var}(\varepsilon_j) = 1 - \beta_j^2$. More precisely, they assumed that for real parameters $a, b$, with $0 \leq |b| < a$, we have $M \sim \text{NIG}(a, b, \mu, \delta)$ and $\varepsilon_j \sim \text{NIG}\left(\frac{a}{\beta_j}, \frac{b}{\beta_j}, \frac{1-\beta_j^2}{\beta_j} \mu, \frac{1-\beta_j^2}{\beta_j} \delta\right)$, where $\mu := -\frac{b\gamma^2}{a^2}, \delta := \frac{\gamma^3}{a^2}$ and $\gamma := \sqrt{a^2 - b^2}$.

The parameters $\mu, \delta$ and those for the distribution of $\varepsilon_j$ were chosen in such a way that $\mathbb{E}(M) = \mu + \delta b = -\frac{b\gamma^2}{a^2} + \frac{b\gamma^2}{a^2} = 0$, $\mathbb{E}(\varepsilon_j) = \frac{1-\beta_j^2}{\beta_j} \mu + \frac{1-\beta_j^2}{\beta_j} \delta \frac{b\beta_j}{\beta_j} = \frac{1-\beta_j^2}{\beta_j} \mathbb{E}(M) = 0$, $\text{Var}(M) = \delta \frac{\gamma^2}{\beta^2} = 1$, and $\text{Var}(\varepsilon_j) = \frac{1-\beta_j^2}{\beta_j^2} \delta \frac{\gamma^3}{\beta_j^2 \gamma^2} = (1 - \beta_j^2) \text{Var}(M) = 1 - \beta_j^2$, for all $j$. As, when the first two parameters coincide, the class of NIG distributions is stable under convolution, this setup yields $S_j \sim \text{NIG}\left(\frac{a}{\beta_j}, \frac{b}{\beta_j}, \frac{\mu}{\beta_j}, \frac{\delta}{\beta_j}\right)$, hence $\mathbb{E}(S_j) = 0$ and $\text{Var}(S_j) = 1$. The advantage of the NIG model lies in exactly this property that the class of NIG distributions is stable under convolution. However, the usage of NIG distributions also happens in a rather ad-hoc fashion and the way they are used is entirely rooted in the goal to achieve a setup, which is stable under convolution.
Chapter 5

An elliptical distributions credit portfolio model

5.1 Introduction

As the Merton model [Mer74] (see Section 4.2) made particular use of the results that have been obtained for the pricing of European options within the Black-Scholes framework [BS73], [Mer73], it has been central for this setup to model the dynamics of the asset-value process \( (A_t)_{t \geq 0} \) according to geometric Brownian motions. This thus entails that the marginal distributions of the log-process \( (S_t)_{t \geq 0} \) with \( S_t = \log(A_t) \) are all Gaussian. In fact, as Merton’s model is a one-period model, the distribution of the process \( (S_t)_{t \geq 0} \) at the maturity \( T \) of the credit derivative in question becomes the relevant modelling issue. Also Merton’s multivariate version (see Section 4.3) and its variations, such as those discussed in Section 4.5.1, yield multivariate Gaussian distributions for the log-process \( (S_t^{(n)})_{t \geq 0} \), where again \( S_t^{(n)} = (S_{1,t}, \ldots, S_{n,t})^T \) and \( S_{j,t} = \log(A_{j,t}) \), for \( j = 1, \ldots, n \) and \( t \geq 0 \) (see Section 4.1 for more details on the notation).

However, despite its advantages with respect to computability and parsimony in the number of parameters, the Gaussian distribution displays characteristics, such as thin tails, which are inconsistent with the observations one obtains from the financial markets. Within the Merton model this property of thin tails results in an underestimation of the default probability of a company, as the likelihood to cross the default threshold is considered too low when the distance to default is still large. The multivariate extensions of the Merton model, where one employs a multivariate Brownian motion and thus ends up with multivariate Gaussian distributions, not only display this deficiency for the marginals, but do also imply a dependence structure which proves to be too weak for the modelling of portfolios underlying multi-name credit derivatives. Our aim here is to introduce elliptical distributions as an extension of the multivariate Gaussian models in order to allow for a larger flexibility in the distributions and their specific properties, such as the univariate heavy-tails and the multivariate tail-dependence. As we have seen in Section 3.1, the family of elliptical distributions is large enough to embrace
many important distribution functions with quite different characteristics, such as the multivariate Normal distribution, the Student $t$-distribution or the class of centered generalised hyperbolic distributions.

As before in Chapter 4, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where we consider $\mathbb{P}$ to be a pricing measure. Again, $n$ denotes the number of credit-risky securities in the portfolio that underlies the credit derivative that we want to analyse. We assume these securities to stem from $n$ distinct companies.

Similar to the Merton model, we assume that there exists an $n$-dimensional firm value or asset-value process $(A_j^{(n)})_{t\geq 0}$, where the $j$-th component $(A_{j,t})_{t\geq 0}$ triggers a default of company $j$ if the process falls below a predefined threshold process $(c_j^{(n)})_{t\geq 0}$. As we have already outlined in the general discussion of structural models in Section 4.1, we will not model the process $(A_j^{(n)})_{t\geq 0}$ directly but rather the transformed $n$-dimensional process $S_j^{(n)} = (S_{1,t}, \ldots, S_{n,t})^T$ of latent variables $S_{j,t} = \log(A_{j,t})$, for $j = 1, \ldots, n$ and $t \geq 0$. The process $(S_{j,t})_{t\geq 0}$ will then of course not be compared with the process $(c_j^{(n)})_{t\geq 0}$ in order to judge whether a default has happened, but instead with the threshold process $(c_j^{(n)})_{t\geq 0}$ where $c_j^t := \log(c_j^{(n)})$, for $j = 1, \ldots, n$ and $t \geq 0$. Then the strict monotonicity of the logarithmic function entails that $A_{j,t} \leq \tilde{c}_j^t$ if and only if $S_{j,t} \leq c_j^t$, for any $j = 1, \ldots, n$ and $t \geq 0$.

Let us recall that the Merton model suggests interpreting the processes $(A_j^{(n)})_{t\geq 0}$, $j = 1, \ldots, n$, as the firm values of the respective companies. With this interpretation, it is readily understandable how the mechanism of default functions: if the value of the company is less than its liabilities, the company goes into default. However, we would like to see these processes $(A_j^{(n)})_{t\geq 0}$, $j = 1, \ldots, n$, in a more abstract way without necessarily giving them the interpretation of being the values of the respective firms. They shall solely be understood as specific, possibly unobservable processes which are responsible for triggering the defaults of the firms if they fall too low and which we would like to characterise with a stochastic model.

As we have outlined before in Section 4.2, the classical Merton model, that was first introduced for the pricing of bonds, is essentially a one-period model in that one only models the behaviour of the underlying asset process at the maturity date: a default of a company can only take place at maturity when the obligor has to pay back the principal. All the one-period models such as the multivariate extensions of the Merton model (see Section 4.3), the models by Klaassen et al. [KLSS01] (see Section 4.5.1), by Hull & White [HW04] (see Section 4.5.3) or by Kalemanova et al. [KSW07] (see Section 4.5.4) then focus on the modelling of the latent variables process $(S_t^{(n)})_{t\geq 0}$ at some fixed time $T$ in the future.

In the context of these one-period models, the default time of company $j$ is then given
5.1. Introduction

in the following way:

\[
\tau_j := T \mathbb{1}_{\{A_{j,T} < c\}} + \infty \mathbb{1}_{\{A_{j,T} \geq c\}},
\]

\[
= T \mathbb{1}_{\{S_{j,T} < \tilde{c}_T\}} + \infty \mathbb{1}_{\{S_{j,T} \geq \tilde{c}_T\}},
\]

for \( j = 1, \ldots, n \). In the following we also want to setup a one-period model and therefore need to further characterise the quantities \( A_{j,T}, S_{j,T}, \tilde{c}_T \) and \( c_T \).

However, it is important to note that it is not sufficient for the pricing of portfolio credit derivatives such as CDOs to know the behaviour of the process \( (S_t^{(n)})_{t \geq 0} \) only at the maturity \( T \). In particular, one additionally needs to know the characteristics at every point in time where there are premium payments to be exchanged. Yet, as with portfolio credit derivatives, the joint behaviour of the underlying credit-risky assets is the key modelling component, and as the premium payments are usually only exchanged every quarter of a year, the focus must rather lie on the dependence structure and not so much on the dynamics. Therefore, the usual assumption made is that the multi-period models are essentially combinations of one-period setups: at first one relies on modelling the distributional properties of the log-asset vector \( S_T^{(n)} \) at the time of maturity \( T \). To simplify notation, one usually drops the time parameter \( T \) and only talks about \( S^{(n)} \).

Then, as soon as one wants to price credit derivatives such as CDOs with payments at times other than maturity, one understands each payment date \( t \) as the end-date of a one-period model over the time interval \([0, t]\), where the distributional properties of \( S_t^{(n)} \) coincide with the distributional properties of \( S^{(n)} \). This means that the distance to the payment date is not directly incorporated into the distributional behaviour of the vector \( S_t^{(n)} \), and the process \( (S_t^{(n)})_{t \geq 0} \) is therefore modelled as being strictly stationary.

Still, the default and also the migration probabilities that differ over time have to be met: The further away the point in time with a credit-risky payment lies, the greater is the uncertainty about a company’s stability and thus the higher is the probability of a default. This will be accomplished via a time-dependent adjustment of the default or migration barriers. More details on this procedure will be given when we discuss the pricing of CDOs in greater depth in Chapter 7.

This chapter is organized as follows: first we will introduce a factor model for the modelling of the default mechanisms in a credit portfolio, where we extend the results of Klaassen et al. [KLSS01] (see Section 4.5.1) to the case of elliptical distributions. A direct consequence of this setup will be that we need to work within a particular subclass of the elliptical distributions, namely the mixtures of Normal distributions. Subsequently, we will discuss which candidates from the class of mixtures of Normal distributions display the property of tail-dependence and will introduce some new distribution functions based on this tail-dependence characterisation.

Imposing elliptical distributions and discussing their behaviour for our applications not only gives us a way of extending the Gaussian setup in a very consistent way without just choosing single distribution functions in a rather ad-hoc fashion as seems to have been done in several occasions in the literature, e.g. with the Hull & White model or the NIG model (see Sections 4.5.3 and 4.5.4). It also allows us to consider stochastic vectors \( S^{(n)} \)
5.2 The setup - extension of the Gaussian case

Let us turn to the modelling of the process \((S^{(n)}_t)_{t \geq 0}\) for which we want to assume a factor structure as \(n\) can be quite large. This factor structure allows us to reduce the dimensionality of the integrals that we have to compute for the valuation of portfolio credit derivatives later on. The factor structure entails that the latent variables processes become uncorrelated (but not independent) when we condition on the factor process. To this end, we suppose that there exists an \(m\)-dimensional risk-factor process \((M_t)_{t \geq 0}\) which represents a systematic or macro-economical component and can be interpreted to reflect the influence of the macro economy onto the default behaviour of the various companies. The dimension \(m\) of this factor is assumed to be much smaller than the dimension \(n\) of the portfolio. At every time \(t\) the \(S_{j,t}, j = 1, \ldots, n,\) will be influenced by this factor and by an individual risk component, that is, for every time \(t \geq 0,\) we assume the structure

\[
S^{(n)}_t := (S_{1,t}, \ldots, S_{n,t})^T = \beta^{(n)}_t M_t + \varepsilon^{(n)}_t, \tag{5.1}
\]

where \(\beta^{(n)}_t = (\beta_{1,t}, \ldots, \beta_{n,t})^T \in \mathbb{R}^{n \times m}\) is the matrix of factor-loadings \(\beta_{j,t} \in \mathbb{R}^m,\) which control the influence of the systematic factor \(M_t\) on each exposure \(j, j = 1, \ldots, n,\) and \(\varepsilon^{(n)}_t := (\varepsilon_{1,t}, \ldots, \varepsilon_{n,t})^T\) represents the vector of the individual behaviour or idiosyncratic risk \(\varepsilon_{j,t}\) of each exposure \(j, j = 1, \ldots, n.\) The vector \(\varepsilon^{(n)}_t\) is assumed to be independent of the factor \(M_t,\) for every \(t \geq 0.\) The distribution of the \(j\)-th component of \(S^{(n)}_t\) is denoted by \(F_{j,t},\) that is \(F_{j,t}(s) = \mathbb{P}(S_{j,t} \leq s),\) for all \(s \in \mathbb{R}\) and all \(1 \leq j \leq n.\)

Specifying the underlying stochastic structure

As mentioned in the introduction of this chapter, in the one-period models one of the key assumptions for the valuation of portfolio credit derivatives is that the process \((S^{(n)}_t)_{t \geq 0}\) is strictly stationary and thus that the distributions of the vectors \(S^{(n)}_t, t > 0,\) coincide for all relevant points in time \(t.\) Thus, it is sufficient to specify the distributional properties of this process at the maturity \(T,\) which is often assumed to be equal to \(T = 1.\) For this vector \(S^{(n)}_T\) we will drop the time parameter, so that the factor structure in Equation (5.1) then turns into

\[
S^{(n)} = (S_1, \ldots, S_n)^T = \beta^{(n)} M + \varepsilon^{(n)}, \tag{5.2}
\]

where the factor vector \(M\) and the idiosyncratic risk vector \(\varepsilon^{(n)} = (\varepsilon_1, \ldots, \varepsilon_n)^T\) are still independent. The distribution of the \(j\)-th component of \(S^{(n)}\) is then denoted by \(F_j,\) that
is, $F_j(s) = \mathbb{P}(S_j \leq s)$, for all $s \in \mathbb{R}$ and all $1 \leq j \leq n$. For now, we suppress the time parameter in the notation for the matrix of the factor loadings $\beta^{(n)} = (\beta_1, \ldots, \beta_n)^T \in \mathbb{R}^{n \times m}$. Both stochastic vectors will follow elliptical distributions:

$$\varepsilon^{(n)} \sim EC_n(0, \Sigma_n, \phi), \quad \text{and} \quad M \sim EC_m(0, \Omega_M, \phi_M). \quad (5.3)$$

where the covariance matrices of both vectors exist and are positive definite with $\text{Cov}(\varepsilon^{(n)}) = \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n}$, $\omega_i > 0$ and with $\text{Cov}(M) = \Omega_M$. In general, the parameter $\Sigma$ of an elliptically distributed vector $X \sim EC_n(0, \Sigma, \phi)$ is not necessarily also its covariance matrix. In fact, the covariance matrix need not even exist.

For example, the multivariate Student $t$-distribution with one degree of freedom belongs to the class of elliptical distributions (cf. Example 3.1.15), but does not possess second or higher moments. Yet, if the covariance matrix exists then, according to Theorem 3.1.17, we have $\text{Cov}(X) = -2\phi'(0)\Sigma$, so that $\text{Cov}(\varepsilon^{(n)}) = \Sigma_n$ and $\text{Cov}(M) = \Omega_M$ are essentially assumptions on the derivatives of the generator functions $\phi$ and $\phi_M$ at zero.

With this setup, we can easily incorporate the multivariate Gaussian model setup of e.g. Klaassen et al. [KLSS01] if we simply set $\phi_M(x) = \phi(x) = \exp(-\frac{1}{2}x)$, for all $x \in \mathbb{R}$. But even if we employ other characteristic generators, the correlation structure of $S^{(n)}$ is preserved from the Gaussian case:

$$\text{Cov}(S^{(n)}) = \beta^{(n)}\Omega_M\beta^{(n) T} + \Sigma_n. \quad (5.4)$$

In particular, the components of the vector $S^{(n)}$ remain uncorrelated if we condition on a realization of the factor $M$, as the covariance matrix $\Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n)$ is diagonal. However, the components of $S^{(n)}$ are in general far from being independent when conditioning on $M$: according to Definition 3.1.10 and Corollary 3.1.8, an elliptically distributed vector $\varepsilon^{(n)} \sim EC_n(0, \Sigma_n, \phi)$ can be decomposed into $\varepsilon^{(n)} \overset{d}{=} r\Gamma_n^{1/2}u^{(n)}$, where $\Gamma_n = \Sigma_n^{1/2} = \text{diag}(\omega_1^{1/2}, \ldots, \omega_n^{1/2})$, where $ru^{(n)} \sim S_n(\phi)$ and where the non-negative random variable $r \geq 0$ appears in every component of $\varepsilon^{(n)} \overset{d}{=} r\Gamma_n^{1/2}u^{(n)}$. This thus enables us to model the desired property of a much stronger conditional and unconditional dependence between the components of the underlying asset processes that can trigger defaults or rating migrations.

In the next chapter, we will derive an approximation result for the credit portfolio losses at any specific point in time. This will provide us with a tool which helps us to analyse credit derivatives more easily. Within this approximation, the dimension of the underlying vector $n$ will tend towards infinity. Therefore, if with “growing” $n$ we want to continue modelling the $n$-dimensional idiosyncratic vector $\varepsilon^{(n)}$ within the class of elliptical distributions and with the characteristic generator function $\phi$ which stays unchanged in $n$, then $\phi$ must lie in $\Phi_\infty$. This set $\Phi_\infty$ is the class of scalar functions that are characteristic generators of spherical random vectors of arbitrary dimensions (see the notions given after Theorem 3.1.2). In Theorem 3.1.13 we have established the important statement, that in the case where $\phi \in \Phi_\infty$, the distribution of $\varepsilon^{(n)}$ must belong to the class of mixtures of Normal distributions.
Thus, we need to suppose that there exist a non-negative random variable \( R \) and a Gaussian random vector \( Y^{(n)} = (Y_1, \ldots, Y_n)^T \sim N_n(0, \Sigma_n) \) such that
\[
\varepsilon^{(n)} \overset{d}{=} R \cdot Y^{(n)} \sim EC_n(0, \Sigma_n, \phi),
\]
with \( \phi(x) := \int_0^\infty \exp\left(-\frac{1}{2} r^2 x\right) dG_2(r) \), for \( x \in \mathbb{R} \), with \( G_2 \) a distribution function with support on the non-negative real line \([0, \infty)\), \( R \geq 0 \), \( R \sim G_2 \), and where \( M, R \) and \( Y^{(n)} \) are independent. In order to obtain \( \Sigma_n \) as the covariance matrix of \( \varepsilon^{(n)} \) as before, i.e. \( \text{Cov}(\varepsilon^{(n)}) = \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n} \) with \( \omega_i > 0 \), the conditions on the second moment of \( R \) and on the derivative of \( \phi \) turn into \( \mathbb{E}(R^2) = -2\phi'(0) = 1 \), as in the comment directly after Equation (5.3) and Theorem 3.1.17.

**Remark 5.2.1** When as here \( \varepsilon^{(n)} \) follows a mixture of Normal distributions one can easily understand that the components of the vector \( \varepsilon^{(n)} \) are uncorrelated but certainly not independent: due to Theorem 3.1.17 we surely still have that \( \text{Cov}(\varepsilon^{(n)}) = \mathbb{E}(R^2)\text{Cov}(Y^{(n)}) = \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \), and thus the components of \( \varepsilon^{(n)} \) are uncorrelated as desired. On the other hand, the scaling variable \( R \) appears in every component of \( \varepsilon^{(n)} = RY^{(n)} \), so that these components are not independent in general. Yet, if the random variable \( R \) is constant (e.g. \( R \equiv 1 \) ), then the vector \( \varepsilon^{(n)} \) follows a multivariate Gaussian distribution, and as the components of \( \varepsilon^{(n)} \) are uncorrelated they are also independent in this special case.

In order to obtain the approximation result for the portfolio losses, we can allow the factor \( M \) to follow any \( m \)-dimensional distribution. Yet, in some cases it can be helpful to restrict the distribution function of \( M \) to also stem from the class of mixtures of Normal distributions, for instance when we want to compute the distributions of the \( S_j \)’s and thereafter possibly the default thresholds that correspond to these distributions and given default probabilities.

**Lemma 5.2.2** For independent random variables \( R_1, R_2, Z_1 \) and \( Z_2 \) where \( R_i \geq 0 \), \( R_i \sim G_{R_i} \), \( \mathbb{P}(R_1 + R_2 = 0) = 0 \), and \( Z_i \sim N(0, \sigma_i^2) \), \( i = 1, 2 \), the distribution function \( H_{\sigma_1^2, \sigma_2^2} \) of the univariate random variable \( R_1Z_1 + R_2Z_2 \) is given by
\[
H_{\sigma_1^2, \sigma_2^2}(x) = \int_0^\infty \int_0^\infty \Phi\left(\frac{x}{\sqrt{r^2\sigma_1^2 + s^2\sigma_2^2}}\right) dG_{R_1}(r) dG_{R_2}(s),
\]
for \( x \in \mathbb{R} \), and its density function becomes
\[
h_{\sigma_1^2, \sigma_2^2}(x) = \int_0^\infty \int_0^\infty n(x; 0, r^2\sigma_1^2 + s^2\sigma_2^2) dG_{R_1}(r) dG_{R_2}(s),
\]
for \( x \in \mathbb{R} \), where \( n(\cdot; \mu, \sigma^2) \) denotes the one-dimensional Normal density function with parameters \( (\mu, \sigma^2) \).
**Proof:** For an arbitrary $x \in \mathbb{R}$, we have

\[
P(R_1Z_1 + R_2Z_2 \leq x) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{R_1Z_1 + R_2Z_2 \leq x\}} \mid R_1, R_2\right)\right)
\]

\[
= \int_0^\infty \int_0^\infty P(rZ_1 + sZ_2 \leq x) \, dG_{R_1}(r)dG_{R_2}(s)
\]

\[
= \int_0^\infty \int_0^\infty \Phi\left(\frac{x}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \, dG_{R_1}(r)dG_{R_2}(s),
\]

where the second equation is due to the independence of the $R_1, R_2, Z_1$ and $Z_2$. The third equation is a consequence of the hypothesis that the event of $\{R_1 = R_2 = 0\}$ is a null-set, which entails that the Gaussian variable $rZ_1 + sZ_2 \sim N(0, r^2\sigma_1^2 + s^2\sigma_2^2)$ is well defined, for any $r, s \geq 0$ with $r + s > 0$, as in this case $r^2\sigma_1^2 + s^2\sigma_2^2$ remains strictly positive.

The last expression in Equation (5.6) then equals

\[
\int_0^\infty \int_0^\infty \int_{-\infty}^{x} n(y; 0, r^2\sigma_1^2 + s^2\sigma_2^2) \, dy \, dG_{R_1}(r)dG_{R_2}(s),
\]

which concludes the proof by virtue of Tonelli’s Theorem for interchanging integrals with non-negative integrands.

**Remark 5.2.3** We note that even though the factor $M$ might follow a mixture of Normal distributions as in Lemma 5.2.4, the distribution of $S^{(n)}$ does not fall back into the class of elliptical distributions in general (except e.g. for $R = R_1 = 1$). However, if the factor $M \sim EC_m(0, \Omega_M, \phi_M)$ follows a mixture of Normal distributions we can assume $M$ to be one-dimensional without loss of generality: if $M$ is $m$-dimensional then there exist a random variable $R_1 \geq 0$ and an $m$-dimensional Gaussian vector $W \sim N_m(0, \Omega_M)$, independent of $R_1$, such that $M \overset{d}{=} R_1W$ and $\mathbb{E}(R_1^2) = 1$. Thus, there also exists a standard Gaussian variable $Z_1 \sim N(0, 1)$, independent of $R_1$ and of $\varepsilon^{(n)}$, such that for every factor loading $\beta_j \in \mathbb{R}^m$ we have

\[
\beta_j^T M \overset{d}{=} R_1\beta_j^T W \overset{d}{=} \rho_j R_1 Z_1 = \rho_j \tilde{M},
\]

where $\rho_j := \sqrt{\beta_j^T \Omega_M \beta_j}$ and $\tilde{M} := R_1Z_1 \sim EC_1(0, \phi_M) = S_1(\phi_M)$. Recall that $S_1(\phi_M)$ denotes the spherical distribution functions with characteristic generator $\phi_M$.

**Lemma 5.2.4** If the $S_j$‘s have the above factor representation (5.2) - (5.5), and if $M \overset{d}{=} R_1W$ also follows a mixture of Normal distributions $EC_m(0, \Omega_M, \phi_M)$ with

- a random variable $R_1 \geq 0$, and an $m$-dimensional Gaussian vector $W \sim N_m(0, \Omega_M)$, which is independent of $R_1$,
5.2. The setup - extension of the Gaussian case

- $\phi_M(x) := \int_0^\infty \exp\left(-\frac{1}{2} r^2 x\right) dG_1(r)$, for $x \in \mathbb{R}$, where $G_1$ is the distribution function of $R_1$ with $\int_0^\infty r^2 dG_1(r) = 1$.
- $\text{Cov}(M) = \text{Cov}(W) = -2\phi'_M(0)\Omega_M = \Omega_M$ being positive definite,
- and with $\mathbb{P}(R_1 + R = 0) = 0$,

then

$$S_j \sim F_j := H_{\beta_j^T \omega_j}, \quad j = 1, \ldots, n,$$

where the distribution functions $G_{R_1}$ and $G_{R_2}$ in Lemma 5.2.2 are substituted by $G_1$ and $G_2$, respectively, and $\rho_j := \sqrt{\beta_j^T \Omega_M \beta_j}$.

Proof: This lemma is a direct consequence of Lemma 5.2.2 and Remark 5.2.3.

From now on we will assume that the dimension of the factor $M$ and thus of the factor process $(M_t)_{t \geq 0}$ equals $m = 1$. As we have outlined in Example 3.1.15, the special case where $M$ and $\varepsilon^{(n)}$ follow multivariate $t$-distributions represents an example which fits into our elliptical distributions model. On the other hand, the “double $t$-distribution copula” model by Hull and White [HW04] which we have reviewed in Subsection 4.5.3 also gives a factor model with $t$-distributed quantities. However, even though our model seems very close to Hull & White’s double $t$-model, there is an important difference which is rooted in the different assumptions on the distribution of the components of $\varepsilon^{(n)} = (\varepsilon_1, \ldots, \varepsilon_n)$. While for Hull and White the $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d random variables, each of them following a univariate $t$-distribution, in our model we assume that the whole vector $\varepsilon^{(n)}$ follows a multivariate $t$-distribution. While this also entails that each component of the vector $\varepsilon^{(n)}$ follows a univariate $t$-distribution, they are however not independent, as we have stressed in Remark 5.2.1. Indeed, the components display a very special dependence structure which can be captured with the $t$-copula, while the dependence structure in Hull & White’s model is actually captured by the product copula, which represents the independence case.

Example 5.2.5 We want to highlight the effects of allowing different mixtures of Normal distributions for the stochastic behaviour of the factor and of the idiosyncratic risk vector. To this end, we compare 5000 realizations of the two-dimensional vector of $S^{(2)} = (S_1, S_2) = (\beta_1, \beta_2)^T M + \varepsilon^{(2)}$ in the case, where the factor $M$ and the idiosyncratic risk vector $\varepsilon^{(2)}$ follow multivariate Normal distributions, with 5000 realizations of $S^{(2)}$, when taking multivariate Student $t$-distributions. In both cases, the first two moments shall coincide and be given by

$$\mathbb{E}(S^{(2)}) = 0 \text{ and } \text{Cov}(S^{(2)}) = (\beta_1, \beta_2)^T \text{Var}(M)(\beta_1, \beta_2) + \text{Cov}(\varepsilon^{(2)}),$$
where \( \text{Cov}(\varepsilon^{(2)}) = \Sigma := \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \) and with factor loadings \( \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} := \begin{pmatrix} 1 \\ 2 \end{pmatrix} \).

For the first case we consider

\[ M \sim N(0, 1) \text{ and } \varepsilon^{(2)} \sim N_2(0, \Sigma), \]

whereas in the second case we assume

\[ M \sim Mt_1 \left(4, 0, \frac{1}{2}\right) \text{ and } \varepsilon^{(2)} \sim Mt_2 \left(6, 0, \frac{2}{3} \Sigma_2\right). \]

Therefore,

\[ \text{Cov}(S^{(2)}) = \begin{pmatrix} 1.5 & 2 \\ 2 & 4.5 \end{pmatrix}. \]

Figure 5.1: Scatter plot with 5000 realizations of the two-dimensional latent variables vector \( S^{(2)} \) in the case of the standard Gaussian factor model (left) and the factor model based on the multivariate t-distribution (right).

In Figure 5.1 we can see that even though means, variances and correlations coincide in both cases, we have a lot more points in the far upper right and in the far lower left corners in the model with the Student t-distributions than in the Gaussian model.

If we use these two very different multivariate distribution specifications for the process \( (S_t^{(2)})_{t\geq0} \) which triggers the possible defaults of the two companies, then we anticipate a lot more joint extremal events such as joint defaults of both obligors to take place in the multivariate double t-distribution copula model as opposed to the pure Gaussian case.

When using such a multivariate double t-distribution copula setup for the modelling of CDOs, we expect the spreads being paid for junior tranches to be lower, and those for senior tranches to be higher than for those within the Gaussian model. This will turn out to be true as we will discuss in Chapter 7.
5.3 Introducing tail dependence in the elliptical distributions model

The Gaussian one-period models such as the multivariate extension of the Merton model or the Gaussian copula model are the standard models which are widely used in the bank industry for the pricing of multi-name credit derivatives such as CDOs or at least as benchmarks for other models. Yet, using a Gaussian model leads to the problem that it has the tendency to underprice the upper or more senior tranches in comparison to the lower or more junior tranches in a CDO as seen in Section 2.2. The junior tranches are hit by the first losses that are inflicted upon the underlying portfolio and therefore serve as a kind of cushion or protection for the more senior tranches. These are only hit if many defaults occur, which then entails that the portfolio experiences large overall losses. The occurrence of many defaults in this structural one-period model would necessitate that the likelihood of joint large downward movements of the asset values \( A_1, \ldots, A_n \), and thus the log-asset values \( S_1, \ldots, S_n \) that trigger the defaults of the various companies, is sufficiently large. However, when modelling the log-asset value vector \( S^{(n)} \) by multivariate Gaussian distribution functions, this likelihood of large joint downward movements is very low, which is understood to be one of the major sources of the underpricing of the more mezzanine or senior tranches. A parameter that measures this likelihood of large joint values of the components in a multivariate vector \( X \) is the so-called tail-dependence coefficient, which we will introduce shortly. Multivariate Gaussian distributions display a tail-dependence coefficient of zero, which roughly means that the occurrence of a large value in one component of a Gaussian vector \( X \) gives us a probability of zero for another component of the same vector \( X \) to be also large. The dependence structure between the components of this vector \( X \) induced by a Gaussian model is thus too weak to be suitable for pricing the more senior tranches.

We continue to assume that the log-asset values \( S_j \)'s follow the above factor representation (5.2) - (5.5), where the idiosyncratic risk vector \( \varepsilon^{(n)} \) follows a mixture of Normal distributions and where the factor \( M \) follows an arbitrary elliptical distribution, possibly also a mixture of Normal distributions. Stepping from a purely Gaussian setup to a setup where we employ elliptical or mixtures of Normal distributions already enforces the dependence structure between the components of the log-asset value vector \( S^{(n)} \). In the pure Gaussian setup the property that the components of \( \varepsilon^{(n)} \) are uncorrelated also induces that they are independent, and thus entails that conditional on a realization of the factor \( M \) we are left with independent components \( S_1, \ldots, S_n \). The same conditional independence property is valid in the models by Hull & White (see Section 4.5.3) and by Kalemanova et al. [KSW07]. However, this independence conditional on the factor \( M \) no longer holds in our model as the components of an elliptically distributed vector are in general still dependent even if uncorrelated (see the explanations directly after Equation (5.4) or Remark 5.2.1). Employing elliptical and, more specifically, mixtures of Normal distributions for the idiosyncratic vector \( \varepsilon^{(n)} \) thus opens up the road for better matching the senior tranches in a CDO. Yet, as the various representatives from the class of multivariate mixtures of Normal distributions can themselves display very different
dependence structures, we need to explore this class of distribution functions further with respect to their dependence structure.

The aim of this section is thus to analyse which representatives from the mixtures of Normal distributions display the property of tail-dependence and can thus be used in our elliptical distributions model to finally obtain a stronger dependence between the components of the log-asset value vector \( S^{(n)} \). Once we choose \( M \) to also follow a mixture of Normal distributions, Remark 5.2.3 clarifies that it is sufficient for the factor \( M \) to be one-dimensional. Therefore, the analysis of the tail-dependence property really aims for the vector \( \varepsilon^{(n)} \). While the study performed in this section helps to introduce a stronger dependence structure in the log-asset value vector \( S^{(n)} \), it also helps us to compose new distribution functions from the class of mixtures of Normal distributions that exactly fulfill the characteristics needed for the distribution function to display tail-dependence. We close this section by using the here obtained characterisations to construct three new mixtures of Normal distributions that will also be employed in our numerical study in Chapter 7, where we compare various specifications of our elliptical distributions model with benchmark models such as the Gaussian model and the model by Hull & White.

### 5.3.1 Tail dependence in mixtures of Normal distributions

Let an \( n \)-dimensional vector \( X \) follow a mixture of Normal distributions

\[
X \overset{d}{=} R \cdot W \sim EC_n(0, \Sigma, \phi),
\]

where \( R \sim G \) is a non-negative random variable with distribution \( G \), \( W \sim N_n(0, \Sigma) \) is a Gaussian random vector,

\[
\phi(x) := \int_{0}^{\infty} \exp \left( -\frac{1}{2} r^2 x \right) dG(r), \text{ for } x \in \mathbb{R},
\]

with \( \mathbb{E}(R^2) = -2\phi'(0) = 1 \) and independent \( R, W \). The notion of tail-dependence refers to the dependence structure between two components of an arbitrary vector. Therefore, in the sequel we will focus on the attributes we have to impose on the distribution \( G \) of \( R \) in order for \( X \) to possess bivariate margins with tail dependence.

**Definition 5.3.1 (Tail dependence)**

A two-dimensional random vector \( X = (X_1, X_2)^T \) is said to be:

1. upper tail-dependent if

\[
\lambda_U := \lim_{v \to 1^-} \mathbb{P} \left( X_1 > F_1^{-1}(v) \mid X_2 > F_2^{-1}(v) \right) > 0,
\]

where \( F_1^{-1}, F_2^{-1} \) denote the generalised inverse distribution functions of \( X_1, X_2 \). Accordingly, \( X = (X_1, X_2)^T \) is said to be upper tail-independent if \( \lambda_U \) is equal to 0. The parameter \( \lambda_U \) is called the upper tail-dependence coefficient.
5.3. Introducing tail dependence in the elliptical distributions model

2. lower tail-dependent if

\[ \lambda_L := \lim_{v \to 0^+} \mathbb{P} \left( X_1 \leq F_1^{-1}(v) \mid X_2 \leq F_2^{-1}(v) \right) > 0, \] (5.9)

where \( F_1^{-1}, F_2^{-1} \) denote the generalised inverse distribution functions of \( X_1, X_2 \).

Accordingly, \( X = (X_1, X_2)^T \) is said to be lower tail-independent if \( \lambda_L \) is equal to 0. The parameter \( \lambda_L \) is called the lower tail-dependence coefficient.

Note that slowly and regularly varying functions (see Definition 3.2.1) will take a central role in the discussion about the tail-dependence property of a random vector.

Define the function \( \gamma : \mathbb{R}^+ \to \mathbb{R}^+ \) by

\[ \gamma(t) := (2\pi)^{-n/2} \int_0^\infty \exp \left( -\frac{1}{2} \frac{t}{r^2} \right) dG(r), \text{ for } t \in \mathbb{R}^+. \]

For our setup, we only need to require \( R \) to be square-integrable. However, if even \( \mathbb{E}(R^n) < \infty \) is satisfied and if the covariance matrix \( \Sigma \) is positive definite then \( \gamma \) represents the density generator of \( X \), as we have seen in Lemma 3.1.20.

Let us assume that \( R \) admits a density function, i.e. there exists an integrable function \( g : \mathbb{R}^+ \to \mathbb{R}^+ \) such that

\[ G(r) = \int_0^r g(s) ds, \text{ for } r \geq 0, \text{ and } G(r) = 0, \text{ for } r < 0. \] (5.10)

Define

\[ \tilde{\gamma}(t) := \int_0^\infty \exp \left( -\frac{1}{2} \frac{t}{r^2} \right) dG(r); \]

then

\[ \tilde{\gamma}(t) = \int_0^\infty \exp \left( -\frac{1}{2} \frac{t}{r^2} \right) g(r) \, dr = \int_0^\infty \exp (-tu) g \left( \frac{1}{\sqrt{2u}} \right) \frac{1}{(2u)^{3/2}} \, du = \int_0^\infty \exp (-tu) \tilde{g}(u) \, du = \int_0^\infty \exp (-tu) \, d\tilde{G}(u), \]

where for the second equality we have substituted \( r = \frac{1}{\sqrt{2u}} \) and where we have defined the density function \( \tilde{g} \) and the distribution function \( \tilde{G} \) via \( \tilde{g}(u) := \frac{g(\sqrt{2u})}{(2u)^{3/2}}, \tilde{G}(u) := \int_0^u \tilde{g}(s) ds, \) for \( u > 0, \) and \( \tilde{g}(u) := 0, \tilde{G}(u) := 0, \) for \( u \leq 0. \) The function \( \tilde{\gamma} \) therefore becomes the Laplace-Stieltjes transform of \( \tilde{G} \), and we obtain the following result.
Lemma 5.3.2 For \( l \in RV_0^{\infty} \) slowly varying and \( \alpha > 0 \) and \( c \geq 0 \) constants, the following statements are equivalent:

1. \( \tilde{G}(x) \sim cx^\alpha l(1/x)/\Gamma(1+\alpha) \) for \( x \to 0^+ \)
2. \( \gamma(t) \sim c(2\pi)^{-n/2}t^{-\alpha}l(t) \) for \( t \to \infty \).

Proof: The equivalence follows directly from Theorem 3.2.5 (a version of Karamata’s Tauberian theorem) as \( \tilde{G} \) is non-decreasing and continuous on \( \mathbb{R} \), \( \tilde{G}(u) = 0 \), for \( u < 0 \), and \( \gamma(t) = (2\pi)^{-n/2}\tilde{\gamma}(t) < \infty \), for \( t \in \mathbb{R}^+ \).

Theorem 5.3.3 If there exist a slowly varying function \( l \in RV_0^{\infty} \), and constants \( \alpha > 0 \) and \( c \geq 0 \), such that

\[
g(y) \sim c2^{1-\alpha}y^{-2\alpha-1}\tilde{l}(y)/\Gamma(\alpha) \quad \text{for } y \to \infty
\]

where \( \tilde{l}(y):=l(2y^2) \in RV_0^{\infty} \) for \( y \in \mathbb{R}^+ \), then

\[
\gamma(t) \sim c(2\pi)^{-n/2}t^{-\alpha}l(t) \quad \text{for } t \to \infty.
\]

Proof: We define \( \tilde{g}(y) := \frac{\tilde{g}(1/y)}{y^2} \) for \( y \in \mathbb{R}^+ \). Thus, \( \tilde{g}(y) = \tilde{g}(1/y)y^{-2} \) and as \( g(y) = \tilde{g}(1/y)y^{-2} = 4\tilde{g}(2y^2)y \), we have the following equivalences:

\[
\begin{align*}
g(y) & \sim c2^{1-\alpha}y^{-2\alpha-1}\tilde{l}(y)/\Gamma(\alpha) \quad \text{for } y \to \infty \\
\iff \tilde{g}(y) & \sim c2^{1-\alpha}y^{-2\alpha+2}\tilde{l}(y)/\Gamma(\alpha) \quad \text{for } y \to \infty \\
\iff \tilde{g}(y) & \sim cy^{\alpha-1}\tilde{l}(1/y)/\Gamma(\alpha) \quad \text{for } y \to 0^+ \\
\iff \tilde{\tilde{g}}(1/y) & \sim cy^{\alpha+1}\tilde{l}(1/y)/\Gamma(\alpha) \quad \text{for } y \to 0^+ \\
\iff \tilde{g}(y) & \sim cy^{-\alpha-1}\tilde{l}(1/y)/\Gamma(\alpha) \quad \text{for } y \to \infty.
\end{align*}
\]

By virtue of Theorem 3.2.2 the above implies that

\[
\int_0^{1/x} \tilde{g}(y)dy = \int_x^{\infty} \frac{\tilde{g}(1/y)}{y^2}dy = \int_x^{\infty} \tilde{g}(y)dy \sim cx^{-\alpha}l(x)/\Gamma(\alpha + 1) \quad \text{for } x \to \infty,
\]

which is equivalent to

\[
\tilde{G}(x) = \int_0^{x} \tilde{g}(y)dy \sim cx^\alpha l(1/x)/\Gamma(1+\alpha) \quad \text{for } x \to 0^+.
\]

Making use of the Abelian part of Lemma 5.3.2, i.e. the inference from the distribution function to the Laplace-Stieltjes transform, yields the claim.

Theorem 5.3.4 Let \( X \overset{\text{d}}{=} R \cdot W \sim EC_n(0, \Sigma, \gamma), n \geq 2, \) with a positive definite matrix \( \Sigma, \) follow a mixture of Normal distributions, and where the random variable \( R \) has a
density $g$ as in Equation (5.10). If there exist a slowly varying function $l \in RV_0^\infty$, constants $\alpha > n/2$ and $c \geq 0$, such that

$$g(y) \sim c 2^{1-\alpha} y^{-2\alpha-1} \tilde{l}(y)/\Gamma(\alpha) \quad \text{for } y \to \infty$$

where $\tilde{l}(y) := l(2y^2) \in RV_0^\infty$ for $y \in \mathbb{R}^+$, then

1. all bivariate margins of $X$ are tail-dependent;
2. the upper and the lower tail-dependence coefficients coincide; for $n = 2$, i.e. for $X = (X_1, X_2)^T \sim EC_n\left(0, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \gamma \right)$, they are given via

$$\lambda = \frac{\int_0^{h(\rho)} \frac{\alpha^n \cdot \sqrt{1-u^2}}{u} du}{\int_0^1 \frac{\alpha^n \cdot \sqrt{1-u^2}}{u} du}$$

where $\rho := \frac{\sigma_{12}}{\sqrt{\sigma_{11} \sigma_{22}}}$, $h(\rho) := \left(1 + \frac{(1-\rho)^2}{1-\rho^2}\right)^{-\frac{1}{2}}$.

![Figure 5.2: Tail-dependence coefficient $\lambda$ as a function of the tail index $\alpha$, $\alpha$ going from 1 to 30, where we set $\rho = 0$.](image)

**Proof:** Under the hypothesis that

$$g(y) \sim c 2^{1-\alpha} y^{-2\alpha-1} \tilde{l}(y)/\Gamma(\alpha) \quad \text{for } y \to \infty,$$

with $\tilde{l}(y) = l(2y^2) \in RV_0^\infty$ for $y \in \mathbb{R}^+$, from Theorem 5.3.3 we deduce that

$$\gamma(t) \sim c(2\pi)^{-n/2} t^{-\alpha} l(t) \quad \text{for } t \to \infty.$$  

Thus, $\gamma$ is in $RV_\infty^\infty$. A slight modification of Theorem 5.5 in [Sch02], p. 322, then yields the stated results.  

\[ \square \]
Another approach for analysing the tail dependence of mixtures of Normal distributions can be based on results by Hult and Lindskog (see e.g. [HL02]). To this end let us again consider an \( n \)-dimensional vector \( X \) which follows a mixture of Normal distributions, i.e.
\[
X \overset{d}{=} R \cdot W \sim EC_n(0, \Sigma, \phi),
\]
(5.12)
where \( R \sim G \) is a non-negative random variable with distribution \( G \), \( W \sim N_n(0, \Sigma) \) is a Gaussian random vector,
\[
\phi(x) := \int_0^\infty \exp \left( -\frac{1}{2} r^2 x \right) dG(r), \text{ for } x \in \mathbb{R},
\]
with \( \mathbb{E}(R^2) = -2\phi'(0) = 1 \) and independent \( R, W \). Note that for this approach we do not need to assume the scaling variable \( R \) to possess a density as in the discussion before.

Clearly, due to the Cholesky decomposition there exist a matrix \( A \in \mathbb{R}^{k \times n} \) with \( A^T A = \Sigma \) and \( k = \text{rank}(\Sigma) \), and an \( k \)-dimensional Gaussian random vector \( N^{(k)} \sim N_k(0, I_k) \) independent of \( R \) such that \( W = A^T N^{(k)} \). In the following, we will again assume that \( n = k \), i.e. that \( \Sigma > 0 \), whence \( A = \Sigma^{1/2} \) follows.

Besides the Gaussian mixture decomposition given in Equation (5.12), \( X \) being elliptically distributed also has the standard representation as outlined in Corollary 3.1.8, so that \( X \sim EC_n(0, \Sigma, \phi) \) can be decomposed into the product
\[
X \overset{d}{=} r A^T u^{(n)},
\]
where \( r \geq 0 \) is a univariate and non-negative random variable, \( u^{(n)} \) is a random vector which is uniformly distributed on the unit sphere surface in \( \mathbb{R}^n \), \( r \) and \( u^{(n)} \) being independent.

As the second moments of the components of \( X \) exist, Theorem 3.1.16 yields that \( \mathbb{E}(u^{(n)}) = \frac{1}{n} I_n \), and as \( \mathbb{E}(r^2) \mathbb{C}ov(u^{(n)}) = \mathbb{C}ov(ru^{(n)}) = I_n \), we obtain \( \mathbb{E}(r^2) = n \). Let us define \( Y := ru^{(n)} \sim S_n(\phi) \). By uniqueness of the two representations, \( Y \overset{d}{=} RN^{(n)} \), and by virtue of Theorem 3.1.9 we thus have
\[
r \overset{d}{=} \| Y \| \overset{d}{=} R \| N^{(n)} \| \quad (5.13)
\]
and
\[
u^{(n)} \overset{d}{=} \frac{Y}{\| Y \|} \overset{d}{=} \frac{RN^{(n)}}{\| N^{(n)} \|} = \frac{N^{(n)}}{\| N^{(n)} \|},
\]
if \( \mathbb{P}(R = 0) = \mathbb{P}(r = 0) = 0 \). Returning to the tail-dependence problem, we state a known result that deals with a characterisation on the scaling variate \( r \) in the standard representation of an elliptically distributed random vector \( X \). This theorem is adapted from Theorem 4.3 in [HL02] and the discussions after Remark 3.6 in [FJS03].
Theorem 5.3.5 Let \( X \stackrel{d}{=} rA^T u^{(k)} \sim EC_n(0, \Sigma, \phi) \) with \( \Sigma_{ii} > 0 \) for \( i = 1, \ldots, n \), \( |\rho_{ij}| < 1 \) for all \( i \neq j \), with \( u^{(k)} \) a \( k \)-vector which is uniformly distributed on the unit sphere surface in \( \mathbb{R}^k \), \( r \) a non-negative random variable independent of \( u^{(k)} \), and \( A \in \mathbb{R}^{k \times n} \) with \( A^T A = \Sigma \) and \( k = \text{rank}(\Sigma) \). Then the following statements are equivalent:

1. \( r \) is regularly varying at \( \infty \) with index \( \alpha > 0 \).
2. For all \( i \neq j \), \( (X_i, X_j)^T \) has tail dependence.

Moreover, if \( r \) is regularly varying at \( \infty \) with index \( \alpha > 0 \), then the upper and the lower tail-dependence coefficients coincide. Then, for any \( i \neq j \), the tail-dependence coefficients of \( (X_i, X_j)^T \) with \( \text{Cov}((X_i, X_j)^T) = \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} \) are given by

\[
\lambda = 2\bar{t}_{\alpha+1} \left( \sqrt{\alpha + 1} \frac{1 - \rho_{ij}}{1 + \rho_{ij}} \right),
\]

where \( \rho_{ij} := \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \), and \( \bar{t}_{\alpha+1} \) the survival function of the univariate Student t-distribution with \( \alpha + 1 \) df.

We now want to use the distributional relationship (5.13) to derive a characterisation for the scaling variate \( R \) in the mixture of Normal distributions representation of \( X \). To this end we must analyse the relationship between the tail behaviour of \( R \) and \( r \).

Lemma 5.3.6 Let \( R \) and \( r \) be two non-negative random variables and \( N^{(n)} \sim N_n(0, I_n) \) an \( n \)-dimensional Gaussian vector which is independent of \( R \), such that \( r = R \parallel N^{(n)} \parallel \). Then \( r \) is regularly varying at \( \infty \) with index \( \alpha > 0 \) if and only if \( R \) is regularly varying at \( \infty \) with index \( \alpha > 0 \). In either case, we have the following relationship

\[
\mathbb{P}(r > x) \sim \frac{\frac{\alpha}{2} \Gamma\left(\frac{\alpha}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \mathbb{P}(R > x) \quad \text{for } x \to \infty.
\]  

(5.14)

Proof: If \( R \) is regularly varying at \( \infty \) with index \( \alpha > 0 \), then there exists a slowly varying function \( L \) such that \( \mathbb{P}(R > x) = x^{-\alpha} L(x) \). As \( xL(x) \in RV^\infty_1 \) we deduce by virtue of Corollary 3.2.4 that \( xL(x) \to \infty \) as \( x \to \infty \). Therefore, there exists an \( \tilde{x} > 0 \) such that \( xL(x) > c_n \) for all \( x \geq \tilde{x} \).

Now, let \( \varepsilon > 0 \) be arbitrarily chosen. Then, for all \( \beta > 0 \) there exists an \( x_\beta > 0 \) such that \( 0 < x^\beta \exp\left(-x^2/4\right) < \frac{\varepsilon}{\sqrt{\pi}} \) for all \( x \geq x_\beta \). For \( \beta := \alpha + n \) and every \( x \geq x_{\alpha+n} \) we therefore obtain

\[
x^{\alpha+1} y^{n-1} \exp\left(-y^2/4\right) \leq y^{\alpha+n} \exp\left(-y^2/4\right) < \frac{\varepsilon}{\sqrt{\pi}}
\]
5.3. Introducing tail dependence in the elliptical distributions model

for all \( y \in [x, \infty) \), and as a consequence of which

\[
x^{\alpha+1} \int_x^\infty y^{n-1} \exp\left(-y^2/2\right) \, dy < \frac{\varepsilon}{\sqrt{n}} \int_x^\infty \exp\left(-y^2/4\right) \, dy < \varepsilon
\]

for all \( x \geq x_{\alpha+n} \).

Therefore,

\[
\frac{\mathbb{P}(\| N^{(n)} \| > x)}{\mathbb{P}(R > x)} = \frac{c_n \int_x^\infty y^{n-1} \exp\left(-y^2/2\right) \, dy}{x^{-\alpha}L(x)} = \frac{x^{\alpha+1} \int_x^\infty y^{n-1} \exp\left(-y^2/2\right) \, dy}{xL(x)} < c_n \frac{\varepsilon}{c_n} = \varepsilon
\]

for all \( x \geq \max\{\tilde{x}, x_{\alpha+n}\} \).

Thus, \( \mathbb{P}(\| N^{(n)} \| > x) = o(\mathbb{P}(R > x)) \) for \( x \to \infty \). The corollary after Theorem 3 in [EG80] then establishes the regular variation of \( r \) with index \( \alpha > 0 \).

The proof for the reverse is a generalisation of the proof for a similar statement given in [BDM02a]. In the statement in [BDM02a] the dimension of the Normal distribution equals \( n = 1 \).

Let us assume that \( r \) is regularly varying at \( \infty \) with index \( \alpha > 0 \). Then there exists a slowly varying function \( L \in RV^\infty \) such that

\[
L(x)x^{-\alpha} = \mathbb{P}(r > x) = \mathbb{P}(R \| N^{(n)} \| > x) = \int_0^\infty \mathbb{P}\left(R > \frac{x}{s}\right) g_n(z) \, dz
\]

\[
= c_n x^n \int_0^\infty \frac{\mathbb{P}\left(R > \frac{x}{\sqrt{2s}}\right)}{(2s)^{1+\frac{\alpha}{2}}} \exp\left(-x^2 s\right) \, ds \quad \text{for } x > 0,
\]

where \( g_n \) denotes the density function of \( \| N^{(n)} \| = \sqrt{\sum_{i=1}^n N_i^2} \) with \( N_i \sim N(0, 1) \) iid, i.e.

\[
g_n(x) = c_n x^{n-1} \exp\left(-x^2/2\right) \mathbb{1}_{\{x \geq 0\}},
\]

and where we define the constant \( c_n := \frac{1}{\Gamma(\frac{\alpha}{2})^{2\frac{n}{2} - 1}} > 0 \).

This entails that

\[
\hat{U}(x) = \int_0^\infty \exp(-xs) \, dU(s) \sim c_n^{-1} x^{-\frac{\alpha+n}{2}} L(\sqrt{x}) \quad \text{for } x \to \infty,
\]

where

\[
U(z) = \int_0^z \frac{\mathbb{P}\left(R > \frac{x}{\sqrt{2s}}\right)}{(2s)^{1+\frac{\alpha}{2}}} \, ds, \quad \text{for } z \geq 0.
\]

Making use of a Tauberian theorem, namely the part from the Laplace-Stieltjes transform to the distribution function in Theorem 3.2.5, we deduce that

\[
U(s) \sim \frac{c_n^{-1} s^{\frac{\alpha+n}{2}} L(\frac{1}{\sqrt{s}})}{\Gamma\left(1 + \frac{\alpha+n}{2}\right)} \quad \text{for } s \to 0^+,
\]

(5.15)
as we can set $U(z) := 0$ for all $z < 0$ and then obtain a function $U$ which is non-decreasing and continuous on the real line $\mathbb{R}$. Let us rewrite $U$ as

$$U(z) = \int_0^z P\left(R > \frac{1}{\sqrt{2s}}\right) ds = \frac{1}{n} \int_0^{(2z)^{\frac{n}{2}}} P\left(R > y^{-\frac{1}{n}}\right) dy$$

and introduce the closely related function $\tilde{U}$ by

$$\tilde{U}(s) := n U\left(\frac{s^{\frac{n}{2}}}{2}\right) = \int_0^s P\left(R > y^{-\frac{1}{n}}\right) dy. \quad (5.16)$$

Then (5.15) implies that

$$\tilde{U}(s) \sim \frac{n c_n^{-1} 2^{-\frac{(n+\frac{\alpha}{2})}{2}} s^{1+\frac{n}{2}} L(\sqrt{2} s^{-\frac{1}{n}})}{\Gamma(1+\frac{\alpha}{2}+\frac{n}{2})} \quad \text{for } s \to 0^+.$$ 

The integrand on the right-hand side of (5.16) is non-decreasing in $y$, so that Theorem 3.2.6 gives

$$P\left(R > y^{-\frac{1}{n}}\right) \sim \frac{n 2^{-\frac{\alpha+n}{2}} \alpha + n}{c_n \Gamma(1+\frac{\alpha}{2}+\frac{n}{2})} \frac{\alpha + n}{n} y^{\frac{\alpha}{n}} L(\sqrt{2} y^{-\frac{1}{n}})$$

$$= 2^{-\frac{\alpha}{2}} \frac{\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{\alpha}{2}+\frac{n}{2})} y^{\frac{\alpha}{n}} L(\sqrt{2} y^{-\frac{1}{n}}) \quad \text{for } y \to 0^+,$$

or

$$P\left(R > x\right) \sim 2^{-\frac{\alpha}{2}} \frac{\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{\alpha}{2}+\frac{n}{2})} x^{-\alpha} L(\sqrt{2} x) \quad \text{for } x \to \infty. \quad (5.17)$$

Thus, we obtain the regular variation of $R$ with index $\alpha$. Additionally, Equation (5.17) also yields the relationship stated in Equation (5.14) as $L \in RV_0^\infty$ and therefore

$$x^{-\alpha} L(\sqrt{2} x) = P\left(r > x\right) \frac{L(\sqrt{2} x)}{L(x)} \sim P\left(r > x\right) \quad \text{for } x \to \infty.$$

\[\square\]

**Remark 5.3.7** If we define the density function $g_n$ and the positive constant $c_n$ as in the previous proof, for any $\alpha > 0$ we then obtain

$$E\left(\|N^{(n)}\|^{\alpha}\right) = \int_0^\infty x^\alpha g_n(x) dx = c_n \int_0^\infty x^{\alpha+n-1} \exp(-x^2/2) dx$$

$$= \frac{2^{\frac{\alpha}{2}+\frac{n}{2}-1}}{\Gamma(\frac{\alpha}{2}) 2^{\frac{n}{2}-\frac{1}{2}}} \int_0^\infty x^{\alpha+n-1} \exp(-z) dz = \frac{2^{\frac{\alpha}{2}} \Gamma(\frac{\alpha}{2}+\frac{n}{2})}{\Gamma(\frac{\alpha}{2})} < \infty,$$

which is exactly the term appearing in the statement of the previous lemma (cf. also Proposition 3 in [Bre65] and Proposition A.1 in [BDM02b]).
We can therefore conclude with the following theorem, which is a direct consequence of Theorem 5.3.5 and Lemma 5.3.6.

**Theorem 5.3.8** Let \( X \overset{d}{=} R \cdot W \sim EC_n(0, \Sigma, \phi) \) with a positive definite matrix \( \Sigma \) follow a mixture of Normal distributions, i.e. \( W \sim N_n(0, \Sigma) \) and \( R \geq 0 \) independent with \( \mathbb{E}(R^2) = -2\phi'(0) = 1 \). Then the following statements are equivalent:

1. \( R \) is regularly varying at \( \infty \) with index \( \alpha > 0 \).
2. For all \( i \neq j \), \( (X_i, X_j)^T \) has tail dependence.

Moreover, if \( R \) is regularly varying at \( \infty \) with index \( \alpha > 0 \), then the upper and the lower tail-dependence coefficients coincide. Then, for any \( i \neq j \), the tail-dependence coefficients of \( (X_i, X_j)^T \) with \( \text{Cov}((X_i, X_j)^T) = \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} \) are given by

\[
\lambda = 2\hat{t}_{\alpha+1} \left( \sqrt{\alpha + 1} \sqrt{\frac{1 - \rho_{ij}}{1 + \rho_{ij}}} \right),
\]

where \( \rho_{ij} := \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \), and \( \hat{t}_{\alpha+1} \) the survival function of the univariate Student \( t \)-distribution with \( \alpha + 1 \) df.

### 5.3.2 Examples for mixtures of Normal distributions

Now that we have the desired tail-dependence characterisation for the scaling variable \( R \) of a Gaussian mixture vector \( X \overset{d}{=} RW \) to hand, we can use this to revisit common multivariate distribution functions, where it is known that they display the property of tail-dependence (as with the multivariate \( t \)-distribution), and to construct new multivariate distribution functions, that also have this tail-dependence property. At the same time, we want to assure that also the other requirements of our setup on the distribution of \( R \) such as \( \mathbb{E}(R^2) = 1 \) are satisfied.

For the construction of new tail-dependent mixtures of Normal distributions, we need to specify an appropriate distribution function for \( R \). This can either be accomplished by analysing known univariate distribution functions, whose tails are regularly varying, and subsequently transforming them to fit into our setup, or by employing regularly varying functions for the construction of new univariate distribution functions. The Pareto distribution\(^1\), for example, is one of the canonical candidates for distributions whose tails are

\[^1\] The distribution function of the Pareto distribution \( \text{Pareto}(\beta) \) with parameter \( \beta > 0 \) is given by

\[
F(x) = \begin{cases} 
0, & \text{if } x < 1; \\
1 - x^{-\beta}, & \text{if } x \geq 1.
\end{cases}
\]

Second moments exist for \( \beta > 2 \), and in this case its raw second moment becomes \( \int_0^\infty x^2 dF(x) = \frac{\beta}{\beta-2} \).
regularly varying. However, it is only supported on the interval $[1, \infty)$ and does not possess unit second moments. Both of these unwanted properties are taken care of in what we called the ‘Power Law’ in 2. (see below). Let us consider the following exemplary specifications for the mixture distribution $G$ with $R \sim G$ and density $g$:

1. (The multivariate $t$-distribution (MV-T))

If $X \stackrel{d}{=} R \cdot W \sim EC_n(0, \Sigma, \phi)$ follows a multivariate distribution, i.e. $X \sim Mt_n(\nu, 0, \frac{\nu - 2}{\nu} \cdot \Sigma)$ with $\Sigma$ positive definite and $\nu > 2$, the density function $g$ of $R$ is given by

$$g(r) := \frac{(\nu - 2)^{\nu/2}}{2^{\nu/2-1} \Gamma(\nu/2)} r^{-\nu-1} \exp\left(-\frac{\nu - 2}{2r^2}\right), \text{ for } r > 0.$$ 

Identify $\alpha := \frac{\nu}{2}$, set $c_\nu := \frac{(\nu - 2)^{\nu/2}}{\Gamma(\nu/2)}$ and $l(y) := \exp\left(-\frac{\nu - 2}{y}\right)$, for $y > 0$, then $l \in RV_{\nu}^\infty$, as for any $t > 0$ we have

$$\lim_{y \to \infty} \frac{l(ty)}{l(y)} = \lim_{y \to \infty} \frac{\exp\left(-\frac{\nu - 2}{ty}\right)}{\exp\left(-\frac{\nu - 2}{y}\right)} = \lim_{y \to \infty} \frac{\exp\left(-\frac{(\nu - 2)(t - 1)}{ty}\right)}{1} = 1,$$

and therefore

$$g(r) = c_\nu r^{1-\frac{\nu}{2} - \frac{\nu}{2} - 1} \frac{l(r)}{\Gamma(\nu/2)} \in RV_{\nu - 1}^\infty,$$

as in Theorem 5.3.4, where again $\tilde{l}(r) := l(2r^2)$, for $r > 0$, and $\tilde{l} \in RV_{\nu}^\infty$. This density function $g$ is independent of the dimension $n$ of $X$. And since for the tail-dependence property we are only interested in the bivariate margins of the $n$-dimensional vector $X$, it suffices to consider the density generator for every two-dimensional sub-vector $(X_i, X_j)^T$ of $X = (X_1, \ldots, X_n)^T$, which is independent of the choice of $i, j$. Therefore, it is sufficient to choose $n = 2$ in Theorem 5.3.4. The assumption $\nu > 2$ for the existence of the second moment of $R$ implies that $\alpha = \frac{\nu}{2} > \frac{n}{2} = 1$ and by Theorem 5.3.4 all bivariate margins of $X$ are tail-dependent. Recall that the condition $\alpha = \frac{\nu}{2} > \frac{n}{2} = 1$ is required for the existence of the probability generator function $\gamma$, as it assures the common density of $(X_i, X_j)^T$, i.e. $h_{ij}(x) = |\Sigma_{ij}|^{-\frac{1}{2}} \cdot \gamma\left(x^T \Sigma_{ij}^{-1} x\right)$ with $\Sigma_{ij} := \begin{pmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{pmatrix}$, for $x \in \mathbb{R}^2$, to be integrable (see Lemma 3.1.20 and its proof for more details).

In the limit, we have the following well known behaviour: Due to Stirling’s formula, $\Gamma(\frac{\nu}{2}) \sim \frac{(\frac{\nu}{2} - 1)^{\frac{\nu}{2} - 1}}{\sqrt{2\pi(\frac{\nu}{2} - 1)}}$ for $\nu \to \infty$, and thus, for any $r > 0$,

$$g(r) \stackrel{\nu \to \infty}{\sim} \frac{(\nu - 2)^{\nu/2}}{2^{\nu/2-1} \Gamma(\nu/2)} r^{-\nu-1} \exp\left(-\frac{\nu - 2}{2r^2}\right)$$

$$\sim \frac{2^{\nu/2}(\frac{\nu}{2} - 1)^{\nu/2} \exp(\frac{\nu}{2} - 1)}{2^{\nu/2-1}(\frac{\nu}{2} - 1) \frac{\nu}{2} - 1} \frac{\sqrt{2\pi(\frac{\nu}{2} - 1)}}{r^{-\nu-1} \exp\left(-\frac{\nu - 2}{2r^2}\right)}$$
5.3. Introducing tail dependence in the elliptical distributions model

\[
= \frac{\sqrt{2(\nu^2 - 1)}}{\sqrt{\pi}} r^{-\nu - 1} \exp \left( \left( \frac{\nu}{2} - 1 \right) \left( 1 - \frac{1}{r^2} \right) \right)
\]

\[
= \frac{\sqrt{2(\nu^2 - 1)}}{\sqrt{\pi}} \exp \left( \nu \left( \frac{1}{2} - \log (r) - \frac{1}{2r^2} \right) - \log (r) - 1 + \frac{1}{r^2} \right)
\]

\[
\nu \to \infty \rightarrow \begin{cases} 
0, & \text{if } r \neq 1; \\
\infty, & \text{if } r = 1;
\end{cases}
\]

as \( \frac{1}{2} - \log (r) - \frac{1}{2r^2} \leq 0 \), for all \( r > 0 \).

Therefore, we have recovered the well known property that \( X \xrightarrow{d} 1 \cdot W = W \sim N_n(0, \Sigma) \), as \( \nu \to \infty \): Student \( t \)-distributions tend to normality as the degrees of freedom increase.

2. (A multivariate ‘Power Law’ (MV-Pow))

Assume \( X^{d} = R \cdot W \sim EC_n(0, \Sigma, \phi), n \geq 2 \), with \( \Sigma \) positive definite, where we define the density function \( g \) of \( R \) by\(^2\)

\[
g(r) := \frac{2\alpha}{c_\alpha} \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha - 1}, \text{ for } r > 0,
\]

with \( \alpha > 1 \) and \( c_\alpha := \sqrt{(\alpha - 1)(2\alpha - 1)} \). Then, for \( x \geq 0 \),

\[
G(x) := \int_0^x g(r) \, dr = - \left[ \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha} \right]_0^x = 1 - \left( 1 + \frac{x}{c_\alpha} \right)^{-2\alpha} \xrightarrow{r \to \infty} 1
\]

and integrating by parts twice yields

\[
\int_0^\infty r^2 g(r) \, dr = - \left[ \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha} \right]_0^\infty + 2 \int_0^\infty \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha} \, r \, dr
\]

\[
= \frac{2c_\alpha}{-2\alpha + 1} \left[ \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha + 1} \right]_0^\infty - \frac{2c_\alpha}{-2\alpha + 1} \int_0^\infty \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha + 1} \, dr
\]

\[
= - \frac{2c_\alpha}{(-2\alpha + 1)(-2\alpha + 2)} \left[ \left( 1 + \frac{r}{c_\alpha} \right)^{-2\alpha + 2} \right]_0^\infty = 1.
\]

Hence, \( g \) is truly a density function on \((0, \infty)\) and \( \mathbb{E}(R^2) = 1 \).

Comparing \( g \) with the condition in (5.11) we obtain

\[
g(r) \sim c^{2(1-\alpha)} r^{-2\alpha - 1} l(r)/\Gamma(\alpha) \text{ for } r \to \infty,
\]

where \( c := \alpha c_\alpha^{2\alpha} \Gamma(\alpha) 2^{\alpha} \) and \( l(r) := 1 \in \mathbb{R}_0^\infty \).

When again considering an arbitrary two-dimensional sub-vector \((X_i, X_j)^T\) of \( X = (X_1, \ldots, X_n)^T \), by virtue of Theorem 5.3.4 we again conclude that with \( \alpha > 1 \) all bivariate margins of \( X \) are tail-dependent.

\(^2\) With this setting, we have \( R \equiv c_\alpha (P - 1) \) where \( P \sim \text{Pareto}(2\alpha) \).
For growing $\alpha$ we obtain the following behaviour: $c_\alpha \sim \sqrt{2\alpha}$ and thus

$$g(r) = \frac{2\alpha}{c_\alpha} \left(1 + \frac{r}{c_\alpha}\right)^{-2\alpha - 1}$$

$$\sim \frac{\sqrt{2}}{\alpha} \left(1 + \frac{r/\sqrt{2}}{\alpha}\right)^{-2\alpha - 1}$$

$$\sim \sqrt{2} \cdot \exp(-\sqrt{2}r), \text{ for any } r > 0.$$ 

Therefore, we unfortunately do not have the property that $X \overset{d}{=} 1 \cdot W = W \sim N_n(0, \Sigma)$, as $\alpha \to \infty$.

3. (A multivariate ‘Power Log Law’ (MV-Pow-Log))

Define a density $g$ via

$$g(r) := \frac{4\alpha^2}{c_\alpha} \left(1 + \frac{r}{c_\alpha}\right)^{-2\alpha - 1} \cdot \log \left(1 + \frac{r}{c_\alpha}\right), \text{ for } r > 0,$$

with $\alpha > 1$ and $c_\alpha = \sqrt{(1-\alpha)^2(1-2\alpha)^2/6\alpha^2 - 6\alpha + 1}$. With this setting, via a substitution of $y = \log \left(1 + \frac{r}{c_\alpha}\right)$ and subsequent integration by parts, for $x \geq 0$, we find

$$G(x) := \int_0^x g(r) \, dr = 4\alpha^2 \int_0^x \log\left(1 + \frac{r}{c_\alpha}\right) \exp(-2\alpha y) y \, dy$$

$$= -2\alpha \left[ \exp(-2\alpha y) y \log\left(1 + \frac{r}{c_\alpha}\right) + 2\alpha \int_0^y \log\left(1 + \frac{r}{c_\alpha}\right) \exp(-2\alpha y) \, dy \right]$$

$$= 1 - \left(1 + 2\alpha \log \left(1 + \frac{x}{c_\alpha}\right) \right) \left(1 + \frac{x}{c_\alpha}\right)^{-2\alpha} x \to \infty 1$$

and

$$\int_0^\infty r^2 g(r) \, dr = 4\alpha^2 c_\alpha^2 \int_0^\infty \left(\exp(2y) - 2 \exp(y) + 1\right) \exp(-2\alpha y) y \, dy$$

$$= 4\alpha^2 c_\alpha^2 \int_0^\infty \left(\exp(2(1 - \alpha)y) - 2 \exp((1 - 2\alpha)y) + \exp(-2\alpha y)\right) y \, dy$$

$$= 4\alpha^2 c_\alpha^2 \left(\frac{1}{4(1-\alpha)^2} - \frac{2}{(1-2\alpha)^2} + \frac{1}{4\alpha^2}\right) = c_\alpha^2 - \frac{6\alpha^2 - 6\alpha + 1}{(1-\alpha)^2(1-2\alpha)^2} = 1.$$ 

Hence, $g$ is truly a density function on $(0, \infty)$ and $E(R^2) = 1$.

Again, comparing $g$ with the condition in (5.11), we obtain

$$g(r) \sim c 2^{1-\alpha} r^{-2\alpha - 1} \tilde{l}(r)/\Gamma(\alpha) \text{ for } r \to \infty,$$

where $c := \alpha^2 c_\alpha^2 \Gamma(\alpha) 2^{\alpha - 1}$ and $\tilde{l}(r) := \log \left(1 + \frac{r}{c_\alpha}\right)$, for $r > 0$, thus $\tilde{l} \in RV_0^\infty$, as for any $t > 0$ we have

$$\lim_{y \to \infty} \frac{\tilde{l}(ty)}{\tilde{l}(y)} = \lim_{y \to \infty} \frac{\log \left(1 + \frac{ty}{c_\alpha}\right)}{\log \left(1 + \frac{y}{c_\alpha}\right)} = t \cdot \lim_{y \to \infty} \frac{\left(1 + \frac{y}{c_\alpha}\right)}{\left(1 + \frac{ty}{c_\alpha}\right)} = t \cdot \frac{1}{t/c_\alpha} = 1.$$
As with the Power Law in 2., all bivariate margins of $X^d = R \cdot W \sim EC_n(0, \Sigma, \phi)$, $n \geq 2$, with $\Sigma$ positive definite and with density function $g$ for $R$ display tail dependence.

For growing $\alpha$ we obtain the following behaviour: $c_\alpha \sim \sqrt{2/3\alpha}$ and thus

$$g(r) = \frac{4\alpha^2}{c_\alpha} \cdot \left(1 + \frac{r}{c_\alpha}\right)^{-2\alpha-1} \cdot \log \left(1 + \frac{r}{c_\alpha}\right)^{-1},$$

$$\alpha \sim \sqrt{6\alpha} \cdot \left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{-2\alpha-1} \cdot \log \left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{-1} \sim \exp(-\sqrt{6r}), \text{ for any } r > 0,$$

as

$$\left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{-2\alpha-1} = \frac{\left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{-1}}{\left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{\alpha}} \sim \exp(-\sqrt{6r})$$

and

$$\alpha \cdot \log \left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right) = \log \left(1 + \frac{\sqrt{2\alpha}}{\alpha}\right)^{\alpha} \sim \sqrt{3/2r}.$$
and

\[
\int_0^\infty r^2 g(r)dr = c_\alpha^\alpha \frac{c_\alpha}{\Gamma} \int_0^\infty \left( \frac{r}{c_\alpha} \right)^{-2\alpha+1} l_\alpha \left( \frac{r}{c_\alpha} \right) dr = \frac{\alpha}{\tilde{c}_\alpha} \int_0^\infty y^{-2\alpha+1} l_\alpha (y) dy = 1.
\]

Hence, \( g \) is truly a density function on \((0, \infty)\) and \( \mathbb{E}(R^2) = 1 \).

Once more, comparing \( g \) with the condition in (5.11) we obtain

\[
g(r) \sim c_\alpha^2 r^{-2\alpha} \tilde{l}(r)/\Gamma(\alpha) \text{ for } r \to \infty,
\]

where \( c := \frac{\alpha c_\alpha^\alpha \Gamma(\alpha) 2^{\alpha-1}}{\tilde{c}_\alpha} \) and \( \tilde{l}(r) := l_\alpha (r) \in RV_0^\infty \), as for any \( t > 0 \) we have

\[
\lim_{y \to \infty} \frac{\tilde{l}(ty)}{\tilde{l}(y)} = \lim_{y \to \infty} \frac{\exp(-(ty)^{-\alpha}) \cdot \exp(-\exp(-(ty)^{-\alpha}))}{\exp(-y^{-\alpha}) \cdot \exp(-\exp(-y^{-\alpha}))}
= \lim_{y \to \infty} \exp(-(t^{-\alpha} - 1) y^{-\alpha}) \cdot \exp(-\exp(-(ty)^{-\alpha}) + \exp(-y^{-\alpha}))
= 1 \cdot \exp(-1 + 1) = 1.
\]

As with the previous laws, all bivariate margins of \( X \overset{d}{=} R \cdot W \sim EC_n(0, \Sigma, \phi), \ n \geq 2 \), with \( \Sigma \) being positive definite and with density function \( g \) for \( R \) are tail-dependent.

Letting \( \alpha \) approach infinity, we obtain

\[
\tilde{c}_\alpha = \frac{1}{\Gamma} \int_0^\infty z^{1-\frac{2}{\alpha}} l_{-1} (z) dz \overset{\alpha \to \infty}{\to} \frac{1}{\Gamma} \int_0^\infty z l_{-1} (z) dz = 1,
\]

by monotone convergence, hence \( c_\alpha = 1/\sqrt{\tilde{c}_\alpha} \overset{\alpha \to \infty}{\to} 1 \), and thus

\[
c_\alpha g(c_\alpha r) = \frac{\alpha}{\Gamma} r^{-2\alpha-1} l_\alpha (r)
= \frac{\alpha}{\Gamma} r^{-2\alpha-1} \exp(-r^{-\alpha}) \cdot \exp(-\exp(-r^{-\alpha}))
\overset{\alpha \to \infty}{\to}
\begin{cases} 
0, & \text{if } r \neq 1; \\
\infty, & \text{if } r = 1.
\end{cases}
\]

This yields

\[
g(r) \overset{\alpha \to \infty}{\to}
\begin{cases} 
0, & \text{if } r \neq 1; \\
\infty, & \text{if } r = 1,
\end{cases}
\]

and therefore, \( X \overset{d}{=} 1 \cdot W = W \sim N_n(0, \Sigma) \), as \( \alpha \to \infty \).

Figure 5.3 illustrates the different scaling densities that we have just discussed. One can also observe in the graphs that for growing \( \alpha \) the multivariate \( t \)-distribution and the multivariate Exp-Exp Law concentrate their probability mass around the value one, and that both the Power Law and the Power Log Law are far away from displaying this convergence property.
5.3. Introducing tail dependence in the elliptical distributions model

Remark 5.3.9 In Example 3.1.15 we have explained that also the class of centered generalised hyperbolic distributions belongs to the mixtures of Normal distributions: Consider a non-negative random variable \( R := \frac{\sqrt{U}}{\sqrt{E(U)}} \) with \( U \) following a generalised inverse Gaussian distribution \( GIG(\lambda, \delta, \gamma) \), \( \lambda \in \mathbb{R} \), \( \delta, \gamma > 0 \), then the density \( g \) of \( R \) is given as

\[
g(r) := \frac{K_\lambda^{\lambda+1}(\delta\gamma)}{K_\lambda^{\lambda+1}(\delta\gamma)} r^{2\lambda-1} \exp \left( -\frac{1}{2} \left( \gamma^2 c_{\lambda, \delta, \gamma} r^2 + \frac{\delta^2}{c_{\lambda, \delta, \gamma} r^2} \right) \right),
\]

for \( r > 0 \), where \( c_{\lambda, \delta, \gamma} := \frac{\delta}{\gamma} K_\lambda^{\lambda+1}(\delta\gamma) > 0 \) and \( K_\lambda \) is the Bessel function of the third kind as in Equation (3.8). Substituting \( y = \frac{\gamma}{3} c_{\lambda, \delta, \gamma} r^2 \), this setup yields

\[
\int_0^\infty r^2 g(r) dr = \left( \frac{\delta}{\gamma c_{\lambda, \delta, \gamma}} \right)^{3/2} \frac{1}{2} \int_0^\infty y^{1/2} g \left( \frac{\delta y}{\gamma c_{\lambda, \delta, \gamma}} \right) ^{1/2} dy
\]

\[
= \left( \frac{\delta}{\gamma c_{\lambda, \delta, \gamma}} \right)^{\lambda+1} \frac{K_\lambda^{\lambda+1}(\delta\gamma)}{K_\lambda^{\lambda+1}(\delta\gamma)} \frac{1}{2} \int_0^\infty y^\lambda \exp \left( -\frac{1}{2} \delta \gamma \left( y + \frac{1}{y} \right) \right) dy = 1
\]

as desired. However, the density function \( g \) does not fulfill condition (5.11) in Theorem 5.3.4 as for \( l(r) := \exp \left( -\frac{1}{2} \left( \gamma^2 c_{\lambda, \delta, \gamma} r^2 + \frac{2\delta^2}{c_{\lambda, \delta, \gamma} r^2} \right) \right) \) and \( \bar{l}(r) := l(2r^2) \), \( r > 0 \), we have

\[
g(r) = \frac{K_\lambda^{\lambda+1}(\delta\gamma)}{K_\lambda^{\lambda+1}(\delta\gamma)} r^{2\lambda-1} \bar{l}(r) \text{ but } l \not\in RV^\infty_0. \text{ Indeed, for } t > 0 \text{ we have}
\]

\[
\frac{l(tr)}{l(r)} = \exp \left( -\frac{1}{2} \left( \gamma^2 c_{\lambda, \delta, \gamma} (t^2-1) r + \frac{2\delta^2}{c_{\lambda, \delta, \gamma} r^2} \left( \frac{1}{t} - \frac{1}{r} - 1 \right) \right) \right) \rightarrow \left\{ \begin{array}{ll}
\infty, & \text{if } t < 1; \\
1, & \text{if } t = 1; \\
0, & \text{if } t > 1.
\end{array} \right.
\]

In fact, the density \( g \) and therefore also the random variable \( R \) are not regularly varying. Theorem 5.3.8 then states that the bivariate margins of the random vector \( X \) with \( X \sim R \cdot W \sim EC_n(0, \Sigma, \phi) \) must be tail-independent. These observations are in line with the observations by Schmidt [Sch92], which are obtained with the use of characterisations on the density generator functions of elliptical distributions.
5.3.3 Effects of the distributional specifications on the dependence structure

We want to further illustrate the effects on the dependence structure of the log-asset vector $S^{(n)}$ when allowing mixtures of Normal distributions for the stochastic behaviour of the factor and of the idiosyncratic risk components. To this end, we compare 5000 realizations of the two-dimensional vector of latent variables $S^{(2)} = (S_1, S_2)$ assuming that this vector is modelled via a factor setup as in (5.2) - (5.5). We alter the distribution functions of the mixing variables $R$ and $R_1$ according to the distribution functions introduced in Section 5.3.2. As benchmarks we use a Gaussian factor setup, where the factor and the idiosyncratic vectors follow Gaussian distributions, and the factor setup underlying the Hull & White model that we have reviewed in Section 4.5.3. Note that the Hull & White model does not belong to our elliptical distributions framework.

As outlined in Remark 5.2.3, without loss of generality we can assume that the dimension of the factor $M$ is equal to 1. In order to simplify the comparison of the models, we scale the distributions of the relevant random variables so that $\text{Var}(M) = \Omega_M = 1$ and $\text{Var}(\varepsilon_1) = \omega_1 = 1 - \beta_1^2$, $\text{Var}(\varepsilon_2) = \omega_2 = 1 - \beta_2^2$, which entail that $\text{Var}(S_1) = \text{Var}(S_2) = 1$. With this setup we therefore obtain

$$\mathbb{E}(S^{(2)}) = 0, \quad \text{Cov}(\varepsilon^{(2)}) = \Sigma_2 = \begin{pmatrix} 1 - \beta_1^2 & 0 \\ 0 & 1 - \beta_2^2 \end{pmatrix}$$

and

$$\text{Cov}(S^{(2)}) = (\beta_1, \beta_2)^T \text{Var}(M)(\beta_1, \beta_2) + \text{Cov}(\varepsilon^{(2)}) = \begin{pmatrix} 1 & \beta_1 \beta_2 \\ \beta_1 \beta_2 & 1 \end{pmatrix}.$$ 

For the following Figure 5.4 we have chosen the factor loadings $\beta_1$ and $\beta_2$ to both equal $\sqrt{0.6}$. The remaining distribution specific parameters where chosen as follows:

1. Hull & White model with df $\nu = \kappa = 5$ for $M$ & $\varepsilon^{(n)}$;
2. MV-T model with df $\nu = 5$ for $M$ & $\kappa = 30$ for $\varepsilon^{(n)}$;
3. MV-Exp-Exp model with $\alpha_1 = 2.5$ for $M$ & $\alpha_2 = 2.5$ for $\varepsilon^{(n)}$;
4. MV-Pow and MV-Pow-Log model with $\alpha_1 = 5$ for $M$ & $\alpha_2 = 5$ for $\varepsilon^{(n)}$.

We can see that even though means, variances and correlations coincide in all cases, we have a lot more points in the far upper right and in the far lower left corners in the different “non-trivial” mixture models with stochastic scaling variables $R$ and $R_1$ than in the “trivial” Gaussian model where $R = R_1 = 1$. Therefore, when we use these distributions for the modelling of the latent variables driving the defaults, we can thus expect the likelihood for joint extremal events such as joint defaults of several obligors to be significantly higher than in a Gaussian model.
5.3. Introducing tail dependence in the elliptical distributions model

Figure 5.4: Scatter plots for the two-dimensional vector $(S_1, S_2)$ following different one-factor models with factor loadings $\beta_1 = \beta_2 = \sqrt{0.6}$. 
5.3.4 Tail dependence in linear factor models

As we have noted before in Section 5.2, the various multivariate Gaussian models such as the one by Klaassen et al. [KLSS01] (see Section 4.5.1) can easily be incorporated into our elliptical distributions model. As the Normal distribution is closed under convolution, with a normally distributed factor $M$ and a normally distributed idiosyncratic vector $\varepsilon^{(n)}$ then also the latent variables vector $S^{(n)}$ becomes normally distributed. This entails, that there is not only a weak dependence structure within the Gaussian vector $\varepsilon^{(n)}$, but also the same weak dependence within the resulting Gaussian vector $S^{(n)}$. Yet, as the elliptical distributions are in general not closed under convolution (see Remark 5.2.3), studying the tail-dependence property between the components of the vector $S^{(n)}$ which is characterised by the factor structure as in Equation (5.2) is of course quite a different matter from studying the tail dependence between the components of the vector $\varepsilon^{(n)}$.

The last subsection has helped us to identify which of the mixtures of Normal distributions display the property of tail-dependence in order to be able to incorporate a strong dependence in the vector $S^{(n)}$. Especially, if we condition on a realization of the common factor $M$, we do not end up with independent components of $S^{(n)}$ as in the usual factor model setup as reviewed in Section 4.5, but instead retain a strong dependence that will help us to tackle the problem of underpriced senior tranches in a CDO. Additionally, the tail-dependence characterisation has helped us to identify promising representatives from the large class of mixtures of Normal distributions and to construct new distribution functions from this class that all display this tail-dependence property.

Yet, we briefly want to review what governs the tail-dependence property of the components of $S^{(n)}$ when we use a linear factor model setup. These results are adapted from Malevergne and Sornette [MS04].

We assume that the two-dimensional linear factor model takes the following form:

$$
S_1 = \beta_1 M + \varepsilon_1, \\
S_2 = \beta_2 M + \varepsilon_2,
$$

where $\beta_1, \beta_2$ are positive non-stochastic constants, and where the factor $M$ is independent of $\varepsilon_1$ and $\varepsilon_2$. Let $F_1, F_2, F_M, F_{\varepsilon_1}$ and $F_{\varepsilon_2}$ denote the distribution functions of $S_1$, $S_2$, $M$, $\varepsilon_1$, $\varepsilon_2$ and $Y$, respectively. The factor $M$ shall also posses a density $f_M$.

**Theorem 5.3.10** Under the assumptions that

1. the distribution functions of $M, \varepsilon_1$ and $\varepsilon_2$ have infinite support;
2. for all $x \geq 1$, we have

$$
\lim_{t \to \infty} \frac{tf_M(tx)}{F_M(t)} = f(x),
$$

where the limit function $f$ is defined on $[1, \infty)$ via the above relationship;
3. There exist real numbers $t_0 > 0$, $\delta > 0$ and $A > 0$ such that for all $t \geq t_0$ and all $x \geq 1$ we have
\[
\frac{F_M(tx)}{F_M(t)} \leq A x^\delta;
\]

4. And there are two constants $l_1, l_2 \in \mathbb{R}^+$ such that
\[
\lim_{u \to 1^-} \frac{F_M^{-1}(u)}{F_M^{-1}(t)} = l_1 \quad \text{and} \quad \lim_{u \to 1^-} \frac{F_M^{-1}(u)}{F_M^{-1}(t)} = l_2;
\]
then the coefficient of upper tail dependence of $(S_1, S_2)$ is given by
\[
\lambda_U = \int_{\max\left\{ \frac{l_1}{M}, \frac{l_2}{M} \right\}}^\infty f(x)dx.
\]

**Proof:** The theorem and its proof are given in Appendix A.2 in [MS04].

In their paper, Malevergne and Sornette first establish the tail dependence between one component $S_i$ and the factor $M$. Having Theorem 5.3.10 then also to hand, they conclude that the tail dependence of $(S_1, S_2)$ is the minimum of the tail dependence of $(S_1, M)$ and that of $(S_2, M)$. Furthermore, Malevergne and Sornette show that if the distribution function of the factor $M$ is rapidly varying, then the upper tail-dependence coefficient equals zero. The prime examples of rapidly varying distribution functions are the Normal and the Exponential distribution function. On the other hand, if both the coefficient equals zero. The prime examples of rapidly varying distribution functions are mixed mixture of Normal distributions, the upper and the lower tail-dependence coefficients coincide and the tail-dependence coefficient $\lambda$ is given by the just mentioned formulas for the upper tail-dependence coefficient, i.e. $\lambda = \lambda_U$ and $\lambda_U$ as above.
5.3. Introducing tail dependence in the elliptical distributions model
Chapter 6

Large portfolio approximation in the elliptical distributions model

Large homogeneous portfolio approximations are widely used for the pricing of complex credit structures such as CDOs, as they allow one to compute portfolio loss distribution functions and quantities derived therefrom such as expected losses on a tranche, and therefore also CDO prices, in a fast and reliable way. Additionally, the analytical expressions for the prices of the credit derivatives so obtained give way to efficient sensitivity analyses, that is, to investigations of how changes of the input variables influence the resulting prices of the credit derivatives. The results of these investigations can then be used for hedging purposes.

The idea of such approximations go back to the seminal paper by Vasicek [Vas91], where he discussed the limiting loan loss distribution when the size of a bond portfolio tends towards infinity. Vasicek presented his limit result for a Gaussian one-factor model with one correlation parameter (see also Section 4.5.1). In the following we will step forward to deriving a much more general limit result within the elliptical distributions framework that we have laid out in the previous chapters. For this, we first discuss the use of credit ratings and how credit-rating migrations will generate credit losses. In the center of this chapter then lies the large homogeneous portfolio approximation result and the assumptions leading there. We finally complete the chapter with a fundamental example where we combine the approximation result with the miscellaneous mixture specifications of Section 5.3.2 and discuss how the different choices lead to sundry limiting loss distributions.
6.1 Credit ratings and credit losses

6.1.1 Credit ratings and rating thresholds

In the Merton model one is solely interested in whether a company defaults or not. Therefore, the default probability of a firm is one of the key parameters in this model. This default probability is naturally very closely linked to the present rating, which might be assigned to a company by rating agencies such as Standard and Poor’s, Moody’s, or Fitch. In fact, one might consider calibrating a credit model with its default probabilities to data published by these rating agencies concerning the default frequencies that have been observed in the past for each of the rating categories (see e.g. [KHB00] and Section 1.1.1.2 in [BOW03]). Additional to losses that the investor in a credit derivative might experience on a default of a company, the investor can also be vulnerable to changes of the credit ratings of the companies which are in the portfolio underlying the credit derivative. Indeed, as the credit rating of one of the underlying obligors changes, the likelihood that default payments have to be exchanged changes and thus the market value of the credit derivative itself. In the following we will therefore also include the present credit rating of each company in the model. We will assume that two companies which have the same present credit rating have the same probability of default, as well as the same probability of a change in their rating.

We want to assume the existence of \( r \) rating categories, which we denote by \( \{1, \ldots, r\} \). These rating categories might for example represent the ratings AAA, AA, A, \ldots, D in the notation of Standard and Poor’s ratings or Aaa, Aa1, Aa2, \ldots, C for Moody’s classification. The present credit rating of company \( j \) is denoted by \( u_j \), which is a known, non-random value in the set \( \{1, \ldots, r\} \), where \( j = 1, \ldots, n \). The probability of a company with a credit rating of \( u \) at time 0 having a rating \( v \) at a future time \( t \) is denoted by \( p_{uv,t} \). We assume that these probabilities are given in a stochastic matrix \( P_t := (p_{uv,t})_{1 \leq u,v \leq r} \), for \( t > 0 \), such that \( \sum_{v=1}^{r} p_{uv,t} = 1 \) for all \( u \) and all \( t \). Rating category 1 will be interpreted as the state of default and \( r \) as the best rating category.

Remark 6.1.1 In the previous Chapter 5, we have introduced elliptical distributions within a one-period framework. However, the results of the present chapter hold in more general terms, especially when we want to make the distribution function of the vector of log-returns \( S(t) \) time-dependent as in Chapter 8. In the following, we will thus index the appearing processes and distribution functions with the relevant time indices.

Definition 6.1.2 For any time \( t > 0 \) and any company \( j \) with rating \( u_j \) at time 0, \( 1 \leq j \leq n \), we define rating thresholds \( c_{uv,t}^j \), \( 0 \leq v \leq r \), as follows:

\[
\begin{align*}
  c_{u_j,0,t}^j := -\infty, & \quad c_{u_j,r,t}^j := +\infty,
\end{align*}
\]
6.1. Credit ratings and credit losses

for the left and the right end points, and

\[ c_{u_j,v,t}^j := F_{j,t}^{-1} \left( \sum_{i=1}^{v} p_{u_j,i,t} \right), \quad \text{for all } 1 \leq v \leq r - 1. \quad (6.1) \]

Here, \( F_{j,t} \) denotes the distribution function of the \( j \)-th component of \( S^{(n)}_t \) as in Section 5.2, and \( F_{j,t}^{-1} \) denotes the inverse function of \( F_{j,t} \) if \( F_{j,t} \) is invertible, and the generalised inverse function of \( F_{j,t} \) if \( F_{j,t} \) is not invertible but continuous.

**Remark 6.1.3** In the one-period framework, the thresholds \( c_{u_j,v,t}^j \) can be computed in a slightly easier way, as the distribution functions \( F_{j,t} \) become time-independent. Equation (6.1) then simplifies to

\[ c_{u_j,v,t}^j = F_j^{-1} \left( \sum_{i=1}^{v} p_{u_j,i,t} \right), \quad \text{for all } 1 \leq v \leq r - 1. \]

Even though the one-period setup entails that the distribution functions of the \( S_t^{(n)} \), \( t > 0 \), are the same for every relevant point in time \( t \in (0,T] \), the default and migration probabilities for the different time horizons have to be met within the model. Therefore, by introducing the time-dependent default and rating thresholds as above, one incorporates the time component while still retaining the concept of using successive one-period models with the same distributional properties.

**Lemma 6.1.4** The thresholds \( c_{u_j,v,t}^j \in \mathbb{R} \) defined in Definition 6.1.2 correspond to the probabilities of company \( j \) migrating from credit rating \( u_j \) at time 0 to an arbitrary rating \( v \) at time \( t > 0 \), as they satisfy

\[ p_{u_j,v,t} = \mathbb{P} \left( c_{u_j,0,t}^j < S_{j,t} \leq c_{u_j,v,t}^j \right) \]

for all \( 1 \leq v \leq r \), \( 1 \leq j \leq n \) and all \( t > 0 \).

**Proof:** By definition, for \( 2 \leq v \leq r - 1 \), \( 1 \leq j \leq n \) and \( t > 0 \), we have

\[ p_{u_j,v,t} = \sum_{i=1}^{v} p_{u_j,i,t} - \sum_{i=1}^{v-1} p_{u_j,i,t} = F_{j,t} \left( c_{u_j,v,t}^j \right) - F_{j,t} \left( c_{u_j,v-1,t}^j \right) = \mathbb{P} \left( c_{u_j,v-1,t}^j < S_{j,t} \leq c_{u_j,v,t}^j \right). \]

Additionally, \( p_{u_j,1,t} = \sum_{i=1}^{1} p_{u_j,i,t} - 0 = F_{j,t} \left( c_{u_j,1,t}^j \right) - F_{j,t} \left( -\infty \right) = \mathbb{P} \left( c_{u_j,0,t}^j < S_{j,t} \leq c_{u_j,1,t}^j \right) \), and for \( v = r \) we obtain

\[ p_{u_j,r,t} = 1 - \sum_{i=1}^{r-1} p_{u_j,i,t} = F_{j,t} \left( \infty \right) - F_{j,t} \left( c_{u_j,r-1,t}^j \right) = \mathbb{P} \left( c_{u_j,r-1,t}^j < S_{j,t} \leq c_{u_j,r,t}^j \right). \]

\( \square \)
With the use of the rating thresholds we can now define the unknown, that is, stochastic credit rating of company \( j \) at time \( t > 0 \) which corresponds to the view taken on the company’s future log-asset value \( S_{j,t} \) and the credit migration probabilities:

**Definition 6.1.5** The credit rating \( V_{j,t} \) of company \( j \) at time \( t > 0 \) is defined via

\[
V_{j,t} := \sum_{v=1}^{r} 1_{\{c^{j}_{u_{j},v-1,t} < S_{j,t} \leq c^{j}_{u_{j},v,t}\}} \cdot v,
\]

for \( 1 \leq j \leq n \).

**Remark 6.1.6** Under the one-period assumption, the credit rating \( V_{j,t} \) of company \( j \) at time \( t > 0 \) is in distribution equal to

\[
V_{j,t} \overset{d}{=} \sum_{v=1}^{d} 1_{\{c^{j}_{u_{j},v-1,t} < S_{j,t} \leq c^{j}_{u_{j},v,t}\}} \cdot v,
\]

for \( 1 \leq j \leq n \).

### 6.1.2 Losses due to rating migrations and defaults

We can now turn to the central modelling components which are needed for the pricing of credit derivatives: the credit losses. In the following we will define these stochastic credit losses that occur if a company \( j \) defaults or undergoes a credit migration at a certain time \( t \). The magnitude of the credit losses differs from firm to firm and depends on the initial rating \( u_{j} \), as well as the rating \( V_{j,t} = v_{j,t} \) at time \( t \). Additionally we allow for another \( d \)-dimensional stochastic component \( (\Psi_{t})_{t \geq 0} \) in the losses that can be seen as some important economic variable process influencing the height of the default or migration losses at any time \( t \).

**Definition 6.1.7** For a company \( j \) with credit rating \( u_{j} \) at time 0, we denote the credit loss at time \( t \) by \( \pi_{t}(j, u_{j}, v_{j,t}, \psi_{t}) \) that occurs if the rating of the company turns out to be \( V_{j,t} = v_{j,t} \) at time \( t \) and if the economic variable at time \( t \) becomes \( \Psi_{t} = \psi_{t} \).

Summarizing the above, we want to characterise each exposure \( j \) at time \( t > 0 \) by the vector

\[(S_{j,t}, u_{j}, V_{j,t}, \pi_{t}(j, u_{j}, V_{j,t}, \Psi_{t})).\]

The main building block that we need for valuing portfolio credit derivatives is the distribution of the overall losses in a portfolio at a specific time. As we later on want to derive an approximation of these overall losses with increasing portfolio size \( n \), we need to index the overall portfolio loss variable with this parameter \( n \):
6.2. Large portfolio approximation

Definition 6.1.8 We define the overall credit losses $C_{n,t}$ of a credit portfolio consisting of $n$ exposures at time $t > 0$ by

$$C_{n,t} := \sum_{j=1}^{n} \pi_t(j, u_j, V_{j,t}, \Psi_t) = \sum_{j=1}^{n} \sum_{v_j=1}^{r} \pi_{j,v,t} \cdot Z_{j,v,t},$$

where $Z_{j,v,t} := \mathbb{1}_{\{ c_{u_j,v-1,t} < S_{j,t} \leq c_{u_j,v,t} \}}$ and $\pi_{j,v,t} := \pi_t(j, u_j, v, \Psi_t)$, for $1 \leq j \leq n$ and $1 \leq v \leq r$.

6.2 Large portfolio approximation

The aim of the following is to approximate the overall credit losses $C_{n,t}$ at any fixed time $t$ by a simpler expression, as the number $n$ of the exposures in the portfolio tends towards infinity. This simpler expression can then be used to e.g. give approximate prices for complex derivatives written on a large portfolio with a complex credit-loss structure.

Definition 6.2.1 For any time $t$, and given the overall credit losses $(C_{n,t})_{n \geq 0}$ as in Definition 6.1.8, which we assume to be integrable, we define the sequence of conditional credit losses $(B_{n,t})_{n \geq 0}$ by

$$B_{n,t} = \mathbb{E}(C_{n,t} | R_t, M_t, \Psi_t) = \sum_{j=1}^{n} \mathbb{E}(\pi_t(j, u_j, V_{j,t}, \Psi_t) | R_t, M_t, \Psi_t)$$

$$= \sum_{j=1}^{n} \sum_{v_j=1}^{r} \mathbb{E}(\pi_{j,v,t} \cdot Z_{j,v,t} | R_t, M_t, \Psi_t),$$

where again $Z_{j,v,t} := \mathbb{1}_{\{ c_{u_j,v-1,t} < S_{j,t} \leq c_{u_j,v,t} \}}$ and $\pi_{j,v,t} = \pi_t(j, u_j, v, \Psi_t)$.

Before we go on, let us introduce the notation of the above factor representation (5.1) - (5.5), where we now let the log-asset values $S_{j,t}$’s display the time parameter $t$ for all the components. That is, for an arbitrary $n \in \mathbb{N}$, we assume that

$$S_{(n)}(t) = (S_{1,t}, \ldots, S_{n,t})^T = \beta_t M_t + \varepsilon_t,$$

where every $\varepsilon_t^{(n)}$, $t > 0$, follows a mixture of Normal distributions:

$$\varepsilon_t^{(n)} \overset{d}{=} R_t \cdot Y_t^{(n)} \sim EC_{n}(0, \Sigma_n, \phi_t),$$

where $R_t \geq 0$ is a non-negative random variable, $Y_t^{(n)} = (Y_{1,t}, \ldots, Y_{n,t})^T \sim N_n(0, \Sigma_n)$ is a Gaussian random vector, $\phi_t(x) := \int_{0}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dG_{2,t}(r)$, for $x \in \mathbb{R}$, with $G_{2,t}$ being a distribution function on the non-negative real line $[0, \infty)$, $R_t \geq 0$, $R_t \sim G_{2,t}$, $\mathbb{E}(R_t^2) = -2\phi_t'(0) = 1$, and where $M_t$, $R_t$ and $Y_t^{(n)}$ are independent. Again, the
covariance matrix $\Sigma_n$ is of the diagonal form with $\text{Cov}(\varepsilon^{(n)}) = \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n}$ and $\omega_i > 0$. For the process $(\Psi_t)_{t>0}$ we will additionally assume that for every $t > 0$ and any $n \in \mathbb{N}$, $(\Psi_t, M_t)$ and $\varepsilon_t^{(n)}$ are independent, while $\Psi_t$ and $M_t$ can be dependent.

In order to be able to derive the approximating result for the portfolio losses, we will need to make use of some assumptions which are given in the following.

**Assumption 1:**

\begin{align*}
\text{(6.4)}
\end{align*}

a.) For every time $t > 0$, there exist (non-random) Borel functions $f_{j,t}$ with $f_{j,t} : \mathbb{R}^{2+m+d} \to \mathbb{R}$ such that

$$
\pi_t(j, u_j, V_{j,t}, \Psi_t) = f_{j,t}(Y_{j,t}, R_t, M_t, \Psi_t), \quad \text{for all } j \in \mathbb{N}.
$$

b.) At any time $t > 0$, there exist strictly increasing sequences of real constants $(b_{n,t})_{n \geq 0}$ with $b_{n,t} \to \infty$, for $n \to \infty$, such that

$$
\sum_{n=1}^{\infty} \frac{1}{b_{n,t}^2} \mathbb{E}
\left[
\left(
\pi_{n,V_{n,t,t}} - \mathbb{E}(\pi_{n,V_{n,t,t}} | R_t, M_t, \Psi_t)
\right)^2
\right] < \infty.
$$

\[ \square \]

**Remark 6.2.2** We give some clarifying remarks about assumption 1:

1. Part a.) of assumption 1 means, that the height of loss $\pi_t(j, u_j, V_{j,t}, \Psi_t)$ for exposure $j$ and for given values of $j, u_j, V_{j,t}$ and $\Psi_t$ can still be stochastic, but the randomness can only stem from the random variables that are already involved, namely from $Y_{j,t}, R_t, M_t$ and $\Psi_t$ for exposure $j$. Recall that at any time $t > 0$ the $(M_t, \Psi_t), R_t, Y_{1,t}, Y_{2,t}, \ldots$ are independent in our factor setup.

2. In our examples later on (e.g. in Section 6.3), the elements of the sequence $(b_{n,t})_{n \geq 0}$ in part b.) of assumption 1 can be chosen as the sum of all notionals up to exposure $n$.

3. If, for an arbitrary $t$, we find a strictly increasing sequence of real constants $(b_{n,t})_{n \geq 0}$ with $b_{n,t} \to \infty$, for $n \to \infty$, where

$$
\sum_{n=1}^{\infty} \frac{1}{b_{n,t}^2} \mathbb{E}(\pi_{n,V_{n,t,t}}^2) < \infty,
$$
then this sequence also fulfills the requirements in part b.) of assumption 1. This follows from the fact that

\[
\begin{align*}
  &\mathbb{E}\left(\left(\pi_{n,V_{n,t},t} - \mathbb{E}(\pi_{n,V_{n,t},t} | R_t, M_t, \Psi_t)\right)^2\right) \\
  &\leq 2\mathbb{E}\left(\left(\pi_{n,V_{n,t},t}\right)^2 + \left(\mathbb{E}(\pi_{n,V_{n,t},t} | R_t, M_t, \Psi_t)\right)^2\right) \\
  &\leq 4\mathbb{E}\left(\left(\pi_{n,V_{n,t},t}\right)^2\right),
\end{align*}
\]

where the last inequality follows from the conditional Jensen’s inequality.

**Assumption 2:**

\((6.5)\)

a.) The credit losses \(\pi_t(j,u_j,v_{j,t},\Psi_t)\) are measurable with respect to \(\sigma(R_t, M_t, \Psi_t)\), for all \(1 \leq v_{j,t} \leq r\), all \(j \in \mathbb{N}\) and all \(t > 0\).

b.) For every time \(t > 0\), one of the following two cases is fulfilled:

- **Case 1:** \(\Psi_t\) is independent of \(\sigma(R_t, M_t) \vee \sigma(S_{j,t}) \subseteq \sigma(R_t, M_t, Y_{j,t})\), for all \(j \in \mathbb{N}\),
- **or** case 2: \(\sigma(R_t, M_t, \Psi_t) = \sigma(R_t, M_t)\).

We now turn to our central approximation result. This result generalises the approximation result that was obtained so far for the Normal distribution function as in [KLSS01] to the general class of elliptical distributions.

**Theorem 6.2.3** Under Assumption 1.) the following approximation holds for the overall credit losses in our elliptical distributions factor setup:

\[
\frac{C_{n,t} - B_{n,t}}{b_{n,t}} \rightarrow 0, \quad \text{as } n \to \infty, \quad \text{almost surely,}
\]

for any time \(t > 0\).

**Proof:** Fix \(t > 0\). For notational reasons, we set

\[
X_{j,t} := \pi_t(j,u_j,v_{j,t},\Psi_t) - \mathbb{E}(\pi_t(j,u_j,V_{j,t},\Psi_t) | R_t, M_t, \Psi_t)
\]

and define the \(\sigma\)-algebras \(\mathcal{A}_t, \mathcal{A}_{n,t}\) and \(\tilde{\mathcal{A}}_{n,t}\) by

\[
\mathcal{A}_t := \sigma(R_t, M_t, \Psi_t), \quad \mathcal{A}_{n,t} := \sigma(Y_{1,t}, \ldots, Y_{n,t}, R_t, M_t, \Psi_t)
\]

and

\[
\tilde{\mathcal{A}}_{n,t} := \sigma(X_{1,t}, \ldots, X_{n,t}),
\]
for \( n \in \mathbb{N}_0 \). Then, under Assumption 1, we obtain
\[
X_{n,t} = f_{n,t}(Y_{n,t}, R_t, M_t, \Psi_t) - \mathbb{E}(f_{n,t}(Y_{n,t}, R_t, M_t, \Psi_t) | A_t),
\]
\( A_t \subset A_{n,t} \) and \( \tilde{A}_{n,t} \subset A_{n,t} \), for \( n \in \mathbb{N}_0 \). As \( Y_{n+1,t}, Y_{n,t}, \ldots, Y_{1,t}, (R_t, M_t, \Psi_t) \) are independent, we obtain that
\[
\mathbb{E}(f_{n+1,t}(Y_{n+1,t}, R_t, M_t, \Psi_t) | A_{n,t}) = \mathbb{E}(f_{n+1,t}(Y_{n+1,t}, R_t, M_t, \Psi_t) | A_t) \quad (6.6)
\]
almost surely, and therefore
\[
\mathbb{E}(X_{n+1,t} | A_{n,t}) = \mathbb{E}(\mathbb{E}(X_{n+1,t} | A_{n,t}) | \tilde{A}_{n,t}) = 0.
\]

Then, as
\[
C_{n,t} - B_{n,t} = \sum_{j=1}^{n} X_{j,t}
\]
and as by assumption 1 b.) there exists a strictly increasing sequence \((b_{n,t})_{n \geq 0}\) with \( b_{n,t} \to \infty \), for \( n \to \infty \), such that
\[
\sum_{j=1}^{\infty} \frac{1}{b_{j,t}} \mathbb{E} \left( \left[ \pi_{j,v_{j,t},t} - \mathbb{E}(\pi_{j,v_{j,t},t} | R_t, M_t, \Psi_t) \right]^2 \right) \leq \sum_{j=1}^{\infty} \frac{1}{b_{j,t}} \mathbb{E}(X_{j,t}^2) < \infty,
\]
we can make use of an almost-sure convergence theorem for martingales (see Theorem 3.4.2), and thus deduce that \( \frac{C_{n,t} - B_{n,t}}{b_{n,t}} = \frac{1}{b_{n,t}} \sum_{j=1}^{n} X_{j,t} \to 0 \) for \( n \to \infty \), almost surely.

\[\square\]

**Corollary 6.2.4** Under the Assumptions 1 and 2, and within the elliptical distributions factor framework, we have
\[
\frac{C_{n,t} - B_{n,t}}{b_{n,t}} \to 0, \quad \text{as } n \to \infty, \text{ almost surely},
\]
for any time \( t \geq 0 \), and the elements of the approximating sequence \( (B_{n,t})_{n \in \mathbb{N}} \) simplify to
\[
B_{n,t} = \sum_{j=1}^{n} \sum_{v_{j,t}=1}^{r} \pi_t(j, u_j, v_{j,t}, \Psi_t) \cdot \mathbb{E} \left( Z_{j,v_{j,t},t} | R_t, M_t, \Psi_t \right)
= \sum_{j=1}^{n} \sum_{v_{j,t}=1}^{r} \pi_t(j, u_j, v_{j,t}, \Psi_t) \cdot \Phi_{j,v_{j,t},t} (M_t, R_t)
\]
with
\[ \Phi_{j,v,t}(m,s) := \Phi \left( \frac{c_{u_j,v,j,t}^j - \beta_j^T m}{\sqrt{\omega_j}} \right) - \Phi \left( \frac{c_{u_j,v-1,j,t}^j - \beta_j^T m}{\sqrt{\omega_j}} \right), \]
for \( m \in \mathbb{R}^m, s \in \mathbb{R}^+ \), \( j \in \mathbb{N} \) and \( 1 \leq v \leq r \)

**Proof:** Let us again fix \( t > 0 \). By Definition 6.2.1 and Assumption 2 a.), we directly deduce that
\[ B_{n,t} = \sum_{j=1}^n \sum_{v_{j,t}=1}^r \mathbb{E} \left( \pi_t(j,u_j,v_{j,t},\Psi_t) \cdot Z_{j,v_{j,t},t} \mid R_t, M_t, \Psi_t \right) \]
\[ = \sum_{j=1}^n \sum_{v_{j,t}=1}^r \pi_t(j,u_j,v_{j,t},\Psi_t) \cdot \mathbb{E} \left( Z_{j,v_{j,t},t} \mid R_t, M_t, \Psi_t \right). \]
In either case of Assumption 2 b.) it follows that
\[ \mathbb{E} \left( Z_{j,v_{j,t},t} \mid R_t, M_t, \Psi_t \right) = \mathbb{E} \left( Z_{j,v_{j,t},t} \mid R_t, M_t \right) \]
almost surely, which equals
\[ \mathbb{P} \left( c_{u_j,v_{j,t},t}^j < S_{j,t} \leq c_{u_j,v_{j,t},t}^j \mid R_t, M_t \right) \]
\[ = \mathbb{P} \left( c_{u_j,v_{j,t},t-1}^j < \beta_j^T M_t + R_t Y_{j,t} \leq c_{u_j,v_{j,t},t}^j \mid R_t, M_t \right) \]
\[ = \mathbb{P} \left( \frac{c_{u_j,v_{j,t},t-1}^j - \beta_j^T M_t}{\sqrt{\omega_j R_t}} < \frac{Y_{j,t}}{\sqrt{\omega_j}} \leq \frac{c_{u_j,v_{j,t},t}^j - \beta_j^T M_t}{\sqrt{\omega_j R_t}} \mid R_t, M_t \right) \]
\[ = \hat{\Phi}_{j,v,t}(M_t, R_t), \]
almost surely, as \( Y_{j,t} \sim N(0,\omega_j) \) which is independent of \( (R_t, M_t) \). \( \square \)

**Remark 6.2.5** As we have outlined in the introduction of this chapter, large homogeneous portfolio approximations, in particular in the context of Gaussian factor models, have been discussed by several authors. One has to be particularly careful however with the way one approaches the proof of such an approximation. Some authors suggest to use the strong law of large numbers for the sums, which appear in their setups. However, the quantities which are added up are usually only independent, if we condition on a particular sigma algebra, which differs for the various approximation results. In order to be able to use the strong law of large numbers in these cases, one would however need to assure the existence of particular regular conditional probability measures under which the appearing quantities are independent. A priori, the existence of such regular conditional probability measures is not evident and some authors lack to mention the difficulty which lies in this argument. In our proof we do not have to guarantee the existence of such probability measures and thus avoid these technical problems, even though our setup is very general.
6.3 First example of the large portfolio approximation

In the present section we discuss the above methodology on a stylised loan portfolio. Once more, let us fix an arbitrary point in time \( t > 0 \) and analyse the losses in the one period \((0,t]\). The large homogeneous portfolio approximation result in the previous section was derived with the factor \( M_t \) following an arbitrary distribution, while only the idiosyncratic vector \( \varepsilon^{(n)} \) had to follow a mixture of Normal distributions. Yet, in order to simplify the analysis, we will also assume that the factor \( M_t \) follows such a mixture of Normal distributions and will thus place ourselves within the setup of Lemma 5.2.2. We want to assume that the factor loadings \( \beta_{j,t} = \beta_j \) are constant in time, as well as homogeneous across the portfolio with \( \rho \equiv \rho_j = \sqrt{\beta_j^T \Omega_M \beta_j} \) and that \( \omega_j \equiv 1 - \rho^2 \), for all \( j = 1, \ldots, n \).

In the one-period setup and with Remark 5.2.3 we can further suppose that
\[
S_{j,t} = \beta_j^T M_t + \varepsilon_{j,t} \overset{d}{=} \rho_j \tilde{M} + \varepsilon_j = \rho \tilde{M} + \varepsilon_j, \quad \text{for all } j,
\]
with \( \tilde{M} = R_1 Z_1 \sim EC_1(0, 1, \phi_M) = S_1(\phi_M), \) \( \text{Var} (\tilde{M}) = 1 \). Furthermore, we focus on only two rating classes, namely \( v = 1 \) for the default case and \( v = 2 \) for non-default. At the beginning, that is at time \( t = 0 \), there are only non-defaulted loans in the portfolio and the losses upon rating migrations are assumed to be given by
\[
\pi_t(j, 2, v_{j,t}, \Psi_t) = \begin{cases} 
N_j, & \text{for } v_{j,t} = 1; \\
0, & \text{for } v_{j,t} = 2.
\end{cases}
\]
Here, \( N_j \) is supposed to be the deterministic size of the \( j \)-th loan, such that there is no recovery upon default and hence a loss given default (LGD) of a 100%. Under Assumptions 1 and 2, the conditional overall losses \( B_{n,t} \) then become
\[
B_{n,t} = \sum_{j=1}^{n} \sum_{v_{j}=1}^{r} \pi_t(j, u_{j}, v_{j}, \Psi_t) \cdot \hat{\Phi}_{j,v_{j},t}(M_t, R_t) = \sum_{j=1}^{n} N_j \cdot \hat{\Phi}_{j,1,t}(M_t, R_t),
\]
where the one-period model assumptions once more comes into play in the expression
\[
\hat{\Phi}_{j,1,t}(M_t, R_t) = \Phi \left( \frac{c_{2,1,t} - \beta_j^T M_t}{\sqrt{\omega_j R_t}} \right) - 0 \overset{d}{=} \Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right),
\]
with, due to Definition 6.1.2 and Lemma 5.2.2,
\[
c_{2,1,t} = H^{-1}_{\beta_j^T \Omega_M \beta_j \omega_j} \left( \sum_{i=1}^{1} p_{2,i,t} \right) = H^{-1}_{\rho^2,1-\rho^2} (p_{2,1,t}) =: c_{2,1,t},
\]
for all \( j \). Note, that \( p_{2,1,t} \) is the probability of an arbitrary firm in the portfolio to have defaulted up to time \( t \). Consequently,
\[
B_{n,t} \overset{d}{=} \Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right) \cdot \sum_{j=1}^{n} N_j, \quad \text{for all } n \in \mathbb{N}.
\]
Therefore, in this basic example we can approximate the overall losses $C_{n,t}$ in the portfolio by an expression which is essentially a fraction of the weighted average loan size $\bar{\pi}_{n,t} := \frac{1}{b_{n,t}} \sum_{j=1}^{n} N_j$, where the positive sequence $(b_{n,t})_{n \geq 0}$ stems from Assumption 1.

More precisely, under Assumption 1 we obtain

$$C_{n,t} b_{n,t} - \bar{\pi}_{n,t} \Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty, \text{ almost surely.}$$

From this result we can directly deduce a natural candidate for the sequence $(b_{n,t})_{n \geq 0}$, namely $b_{n,t} := \sum_{j=1}^{n} N_j$, the sum of all notionals of the $n$ loans in the portfolio, which is independent of $t$. In this case, $\bar{\pi}_{n,t} \equiv 1$, and

$$C_{n,t} b_{n,t} - \Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty, \text{ almost surely.}$$

This last expression gives rise to an evaluation of the asymptotic default losses with the various mixtures of Normal distributions of Section 5.3.2. We assume that the distribution function $G_1$ of the scaling variable $R_1$ has a density function $g_1$ and so does the distribution function $G_2$ of $R$ with density function $g_2$. Then the density function of the limit losses

$$\Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right)$$

is given by

$$h(x) := \int_0^{\infty} \int_0^{\infty} \frac{\varphi \left( \left( \sqrt{1 - \rho^2 \Phi^{-1}(x)} s - c_{2,1,t} \right) / (\rho r) \right)}{\varphi \left( \Phi^{-1}(x) \right)} \frac{1 - \rho^2 s}{\rho r} g_1(r) g_2(s) dr ds,$$

for $x \in [0,1]$, as by independence of $Z_1, R_1$ and $R$ we have

$$P \left( \Phi \left( \frac{c_{2,1,t} - \rho \tilde{M}}{\sqrt{1 - \rho^2 R}} \right) \leq x \right)$$

$$= \int_0^{\infty} \int_0^{\infty} \Phi \left( \frac{\sqrt{1 - \rho^2 s \Phi^{-1}(x)} - c_{2,1,t}}{\rho r} \right) g_1(r) g_2(s) dr ds$$

$$= \int_0^{\infty} \int_0^{\infty} \Phi \left( \sqrt{1 - \rho^2 s \Phi^{-1}(x)} - c_{2,1,t} \right) g_1(r) g_2(s) dr ds$$

$$= \int_0^{x} h(z) dz,$$

for $x \in [0,1]$, where the last equality follows by substitution of $z = \Phi \left( \frac{\rho r y + c_{2,1,t}}{\sqrt{1 - \rho^2 s}} \right)$.

For our analysis of the loss distribution, we use the different tail-dependent mixtures of Normal distributions that we have introduced in Section 5.3.2 for the factor, as well
as for the idiosyncratic risk vector. For all these distributions, we assume the scaling variable $R_1$ in the factor to display a tail-index coefficient of $\alpha_1 = 1.5$ and the scaling variable $R$ in the idiosyncratic component to have a tail-index coefficient of $\alpha_2 = 20$. In the multivariate $t$-distribution case this translates into $\nu = 3$ df for $R_1$, resp. $\kappa = 40$ df for $R$. The correlation coefficient $\rho$ is supposed to be equal to 0.40 and the default probability, that is the probability of migrating from credit rating 2 to 1, is equal to $p_{2,1,t} = 5\%$. With these settings, the default thresholds $c_{2,1,t} = H^{-1}_{\rho^2,1-p^2} (p_{2,1,t})$ take the values as displayed in the following table:\textsuperscript{1}

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>-1.6449</td>
<td>-1.6449</td>
</tr>
<tr>
<td>MV-T</td>
<td>-1.6371</td>
<td>-1.6190</td>
</tr>
<tr>
<td>MV-Exp-Exp</td>
<td>-1.6415</td>
<td>-1.6230</td>
</tr>
<tr>
<td>MV-Pow</td>
<td>-1.4340</td>
<td>-1.4333</td>
</tr>
<tr>
<td>MV-Pow-Log</td>
<td>-1.5367</td>
<td>-1.5152</td>
</tr>
</tbody>
</table>

The density and distribution functions of the relative limit portfolio losses $\Phi\left(\frac{c_{2,1,t}-\rho M}{\sqrt{1-\rho^2} R}\right)$, which result from the different distributional specifications are then given in Figure 6.1.

We can see that the Power Law and the Power Log Law specifications display a behaviour that strongly differs from the Gaussian, the $t$-distribution and the Exp-Exp Law cases. Even though the individual default probabilities are at 5\% for all cases, the Power Law and the Power Log Law produce a remarkably smaller probability for large portfolio losses than the other distributions. The Gaussian, the $t$-distribution and the Exp-Exp Law specifications seem to behave similarly, in particular the graphs in the last row of Figure 6.1 however depict the differences between these distributional assumptions. These variations also lead to substantially different CDO tranche prices for the Gaussian, the $t$-distribution and the Exp-Exp Law specifications as we will see in Chapter 7.

\textsuperscript{1} It is important to recall that the names of these models only indicate which distributions are used for the factor and the idiosyncratic risk vector. However, they are in general no indication on the distributions of the resulting log-asset value vector (cf. Lemma 5.2.2).
6.3. First example of the large portfolio approximation

Figure 6.1: The densities and distribution functions of the relative limit portfolio loss distributions for the various Gaussian mixture models with $\rho = 0.2$ (left column) and $\rho = 0.4$ (right column). The other parameters were chosen as $\alpha_1 = 1.5$, $\alpha_2 = 20$, and $p_{2,1,t} = 5\%$. 
6.3. First example of the large portfolio approximation
Chapter 7

Application to the valuation of credit derivatives

In this chapter we will finally approach the pricing of Collateralized Debt Obligations (CDOs). As we have seen in Section 2.2, the standard Gaussian model displays unwanted results for the prices of such CDO tranches. We believe that one of the main defects in the Gaussian model lies in the fact that the multivariate Gaussian distribution function does not possess the property of tail-dependence and can therefore not reproduce the high levels of likelihood for joint defaults within the portfolio underlying the CDOs.

In the following, we therefore want to analyse how the elliptical distributions model performs in the modelling and pricing of CDOs in comparison to the standard Gaussian model and the model by Hull & White (see Section 4.5.3). As we have seen, the elliptical distributions model generalises the Gaussian model in a consistent and natural way, as the Gaussian distribution is the simplest representative from the class of mixtures of Normal distributions. Additionally, the elliptical distributions model displays a very flexible setup, as one can choose the scaling variables $R$ and $R_1$ and their distribution functions according to need, especially with respect to the desired level of tail-dependence.

After we have presented the notions and definitions that we use for the CDO structure in Section 7.1, we will apply the theoretical results we have obtained in all the previous chapters. Hereby, the result concerning the approximation of portfolio losses from Chapter 6 will play the central role for our analysis, as it provides us with tractable analytical expressions as approximations for the distribution of the portfolio losses and enables us to compute the relevant quantities, such as the expected losses in a specific tranche. After first discussing these quantities for the general setup with arbitrary mixtures of Normal distributions, we then focus on the possible specifications of the elliptical distributions model as presented in Section 5.3.2, which we want to use for the pricing procedures later on in this chapter. We further detail how standard numerical procedures can be employed in order to compute such components as the default probabilities and the expected tranche losses, which are important for deriving the fair spreads of the various tranches. Finally we will discuss in detail which tranche prices these distribution
functions produce. This will show us how the elliptical distributions model helps us to overcome the mispricing of CDOs which we have observed with the Gaussian model in Section 2.2.

Even though we focus on the modelling and the valuation of CDOs, let us note however that the entire model and many of the quantities we derive within this model in this chapter can be applied directly for the valuation of other portfolio credit derivatives, such as \( n \)-th-to-default swaps.

### 7.1 The modelling of the CDO structure

Consider a synthetic CDO as described in Section 2.1 with the following characteristics and notation:

1. The CDO runs between time 0 and time \( T \), the maturity of the CDO, where time is denoted in years;
2. The underlying portfolio consists of \( n \) contracts (bonds, loans, etc.), \( n \in \mathbb{N} \), where the \( j \)-th contract has a notional amount of \( N_j \), \( j = 1, \ldots, n \);
3. The sum of all notional amounts in the portfolio is denoted by \( \bar{N}_n \), i.e. \( \bar{N}_n := \sum_{j=1}^{n} N_j \);
4. The overall losses in the underlying portfolio consisting of \( n \) contracts at time \( t \in [0, T] \) is denoted by \( C_{n,t} \) with \( 0 \leq C_{n,t} \leq \bar{N}_n \) (the specification of the portfolio losses \( C_{n,t} \) will be made more precise in Section 7.4);
5. The relative overall losses in the underlying portfolio with \( n \) contracts at time \( t \in [0, T] \) is given by \( L_n(t) := \frac{C_{n,t}}{\bar{N}_n} \), which represents a value in \([0, 1]\);
6. There are \( q \) CDO tranches, whose boundaries are understood as a fraction of the sum of the overall notional amounts \( \bar{N}_n \). The tranches can thus be written as a partition of \([0, 1]\) : there are detachment points \( \alpha_0, \alpha_1, \ldots, \alpha_q \) such that \( 0 = \alpha_0 < \alpha_1 < \ldots < \alpha_q = 1 \) and where the interval \( I_i := [\alpha_{i-1}, \alpha_i) \) corresponds to the \( i \)-th tranche, \( i = 1, \ldots, q \);
7. The relative losses for tranche \( i \) at time \( t \) are denoted by \( L_{\alpha_{i-1}}(t) \), that is, if the relative portfolio losses equal \( L_n(t) \), then the relative losses for tranche \( i \) are given via \( L_{\alpha_i} \) \( (t) := \Theta_i(L_n(t)) \) \([0, 1]\), where

\[
\Theta_i(x) := \frac{1}{\alpha_i - \alpha_{i-1}} \cdot \min\{(x - \alpha_{i-1})^+, \alpha_i - \alpha_{i-1}\}
\]

\[
= \frac{1}{\alpha_i - \alpha_{i-1}} \cdot (\min\{x, \alpha_i\} - \alpha_{i-1})^+,
\]
7.2. The valuation of a CDO

The payments that are related to a specific CDO tranche are usually divided into the premium leg and the default leg. The premium leg consists of all the premium payments from the bank (the risk seller) via the special purpose vehicle to the investor in this tranche (the risk buyer) that needs to be compensated for accepting the risk. On the other hand, the default leg consists of all the default payments that are transferred from the investor (the protection seller) via the special purpose vehicle to the bank (the risk buyer): whenever there has been a default in the underlying portfolio that affects this specific CDO tranche, then the investor has to offset the losses of the bank. In the following we will further discuss these two legs separately and use the notions and
7.2. The valuation of a CDO

definitions given for the CDO structure in the previous section. Under the assumption that there is no arbitrage possible, the present value of the premium leg has to equal the present value of the default leg.

The premium leg

At every of the \( M \) spread payment days \( t_1, \ldots, t_M \in (0, T] \), there is a spread payment for each tranche \( i \) of
\[
\varsigma_i (1 - L_{\alpha_i - 1}^\varsigma_i (t_m)) \cdot \Delta t_m.
\]
Additionally, there can also be an upfront premium payment \( \varsigma^U_i \) being paid to the investor at time \( t_0 = 0 \).

Thus, the present value of the premium leg of tranche \( i \), depending on the constant spread \( \varsigma_i \) and the upfront payment \( \varsigma^U_i \), becomes
\[
PVP_{PL_i}(\varsigma^U_i, \varsigma_i) := \varsigma^U_i + \mathbb{E} \left[ \sum_{m=1}^{M} \varsigma_i \cdot \left( 1 - L_{\alpha_i - 1}^\varsigma_i (t_m) \right) \cdot \Delta t_m \cdot B(0, t_m) \right]
\]
\[
= \varsigma^U_i + \varsigma_i \cdot \sum_{m=1}^{M} \Delta t_m \cdot \left( 1 - \mathbb{E} \left( L_{\alpha_i - 1}^\varsigma_i (t_m) \right) \right) \cdot B(0, t_m)
\]
\[
= \varsigma^U_i + \varsigma_i \cdot PVP_{PL_i}(0, 1),
\]
where, as stated before, we suppose that the risk-free interest rate and the quantities underlying the portfolio losses \( L_n(t) \) are independent, whence we can simply use the time- \( t \)-value \( B(0, t) \) of the non-stochastic term-structure in order to discount payments at a future point in time \( t \) to today.

The default leg

The losses affecting a specific tranche, \( i \) say, which occur in the time interval \([0, T]\) are assumed to be immediately offset by the investor holding this specific tranche. Thus, the default payments for tranche \( i \), seen from today \( t_0 = 0 \), equal
\[
\int_0^T B(0, s) dL_{\alpha_i - 1}^\varsigma_i(s).
\]
As this integral can be interpreted as a Riemann-Stieltjes integral and as \( B(0, \cdot) \) is deterministic, the present value of the default leg of tranche \( i \) thus becomes
\[
PVD_{DL_i} := \mathbb{E} \left( \int_0^T B(0, s) dL_{\alpha_i - 1}^\varsigma_i(s) \right)
\]
\[
= \int_0^T B(0, s) d \mathbb{E} \left( L_{\alpha_i - 1}^\varsigma_i(s) \right)
\]
\[
= \lim_{k \to \infty} \sum_{l=1}^{k} B(0, s_l^k) \cdot \left[ \mathbb{E} \left( L_{\alpha_i - 1}^\varsigma_i(s_{l+1}^k) \right) - \mathbb{E} \left( L_{\alpha_i - 1}^\varsigma_i(s_{l}^k) \right) \right],
\]
for arbitrary partitions $P_k = \{ s^k_1, \ldots, s^k_k : 0 \leq s^k_1 < \ldots < s^k_k \leq T \}$ of $[0, T]$ with $|P_k| \to 0$ as $k \to \infty$.

**Determining the spread of a CDO tranche**

Under the assumption of no arbitrage, the present values of the premium leg and of the default leg of any tranche $i$ as given above have to be equal. The price of a specific CDO tranche $i$ is exactly the spread $\varsigma_i$ that makes these two present values equal to each other. This spread $\varsigma_i$ is then called the *fair spread* for tranche $i$. Setting the above expressions equal to each other and solving for $\varsigma_i$ yields

\[
PVPL_i(\varsigma^U_i, \varsigma_i) = PVDL_i
\iff \varsigma^U_i + \varsigma_i \cdot PVPL_i(0,1) = PVDL_i
\iff \varsigma_i = \frac{PVDL_i - \varsigma^U_i}{PVPL_i(0,1)}.
\]

Once the height of the upfront payment $\varsigma^U_i$ is agreed on, the two expressions $PVDL_i$ and $PVPL_i(0,1)$ entirely determine the fair spread $\varsigma_i$.

For the equity tranche within the standardised tranched iTraxx the spread $\varsigma_1$ is fixed at 500 bps, so that the upfront payment $\varsigma^U_1$, which is usually quoted in percent, needs to be determined:

\[
PVPL_i(\varsigma^U_1, \varsigma_1) = PVDL_i
\iff \varsigma^U_1 = PVDL_i - \varsigma_1 \cdot PVPL_i(0,1).
\]

For the other iTraxx tranches there are only regular payments being exchanged but no upfront payments. Once more, the key values are the two expressions $PVDL_i$ and $PVPL_i(0,1)$.

For these expressions however we mainly need to compute the expected losses $E(L_{n-1}^\alpha(t)) = E(\Theta_i(L_n(t)))$ at any relevant point in time $t > 0$. Therefore, for the pricing of a CDO it is of crucial importance to accurately model the losses in the underlying portfolio $L_n(t)$, while the model must still permit us to compute the expected value of the losses for each individual tranche. In the following we will see how we can determine the expected losses of a specific tranche and related quantities within the elliptical distributions model and its specifications that we have introduced in Section 5.3.2.
7.3 Defaults and their probabilities

We assume that the \( n \) exposures in the underlying portfolio all stem from different companies (they have for example taken out a loan or issued a bond), that is, there are \( n \) distinct obligors that can be responsible for triggering a credit event which entails credit losses for the holder of the portfolio. These losses then have to be offset by the respective CDO investor if this part of the portfolio’s credit risk has been sold to a protection seller. A common assumption being made for the pricing of CDOs is that there are only two credit states in focus. The obligors behind the \( n \) exposures in the underlying portfolio are assumed to be either in a default or in a non-default state. Again, at the inception of the CDO none of the obligors shall have previously defaulted.

Here, compared to the setup in the previous chapter, we have a reduction of rating classes to just two classes, default and non-default. Therefore, we want to introduce additional flexibility for the modelling of the individual likelihood to default by letting \( p_j(t) \) denote the risk-neutral probability for obligor \( j \) to default up to and including time \( t \in [0,T] \), \( j = 1, \ldots, n \). If the default probabilities \( p_j(t) \) coincide for all obligors \( j = 1, \ldots, n \), we are back in the notation of the example in Section 6.3 and therefore \( p_j(t) = p_{2,1,t} \) if we again let \( v = 2 \) denote the non-default state and \( v = 1 \) the default state.

As in Chapter 5, obligor \( j \) has defaulted at an arbitrary time \( t > 0 \) if the log-asset value \( S_{j,t} \) of obligor \( j \) at time \( t \) lies below or at the default threshold \( c_t \). Again, for every time \( t > 0 \) the asset values

\[
S_t^{(n)} = (S_{1,t}, \ldots, S_{n,t})^T = \beta^{(n)}_t M_t + \epsilon_t^{(n)}
\]

are modelled under the one-period assumption (see Section 5.2) with

\[
\epsilon_t^{(n)} \overset{d}{=} R_t \cdot Y_t^{(n)} \sim EC_n(0, \Sigma_n, \phi),
\]

with \( \phi(x) := \int_0^\infty \exp \left( -\frac{1}{2} x^2 \right) dG_2(r), \) for \( x \in \mathbb{R} \), where \( G_2 \) is a distribution function with support on the non-negative real line \([0, \infty)\), \( R_t \geq 0 \) being independent of \( Y_t^{(n)} \sim N_n(0, \Sigma_n) \), \( R_t \sim G_2 \), \( \mathbb{E}(R_t^2) = -2\phi'(0) = 1 \) and \( \text{Cov}(\epsilon_t^{(n)}) = \text{Cov}(Y_t^{(n)}) = \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n} \) with \( \omega_i > 0 \). Additionally, the factor loadings \( \beta_t^{(n)} = \beta^{(n)} = (\beta_1, \ldots, \beta_n)^T \) shall be constant over time and for every \( t > 0 \) the factor shall also follow a mixture of Normal distributions

\[
M_t \overset{d}{=} R_{1,t} W_t \sim EC_m(0, \Omega_M, \phi_M),
\]

with \( \phi_M(x) := \int_0^\infty \exp \left( -\frac{1}{2} x^2 \right) dG_1(r), \) for \( x \in \mathbb{R} \), where \( G_1 \) is a distribution function with support on the non-negative real line \([0, \infty)\), \( R_{1,t} \geq 0 \) being independent of \( W_t \sim N_m(0, \Omega_M) \), \( R_{1,t} \sim G_2 \), \( \mathbb{E}(R_{1,t}^2) = -2\phi'_M(0) = 1 \) and \( \text{Cov}(M_t) = \text{Cov}(W_t) = \Omega_M \) with \( \Omega_M \) being positive definite.

We also want \( \mathbb{P}(R_{1,t} + R_t = 0) = 0 \) to be satisfied (see Lemma 5.2.4). As in Remark 5.2.3, for every \( t > 0 \) there exists a standard Normal random variable \( Z_{1,t} \sim \mathbb{N}_(0,1) \).
7.4 Approximation of the portfolio losses

\( N(0,1) \) such that \( R_{1,t}, Z_{1,t}, R_t, Y_{1,t}, \ldots, Y_{n,t} \) are independent and such that \( \beta_j^T M_t \overset{d}{=} R_{1,t} \beta_j^T W_t = \rho_j R_{1,t} Z_{1,t} \) with \( \rho_j := \sqrt{\beta_j^T \Omega_M \beta_j} \), for \( j = 1, \ldots, n \).

We define the Bernoulli random variables \( D_{j,t} \) by \( D_{j,t} := \mathbb{1}_{\{S_{j,t} \leq c^j_t\}} \), for \( j = 1, \ldots, n \) and \( t > 0 \). Again (cf. Definition 6.1.2 and Lemma 6.1.4), each threshold \( c^j_t \) corresponds to the \( j \)-th risk-neutral default probability \( p_j(t) \) at time \( t \):

\[
\{D_{j,t} = 1\} = \{S_{j,t} \leq c^j_t\} \quad \text{and} \quad p_j(t) = \mathbb{P}\left(S_{j,t} \leq c^j_t\right) = F_j \left(c^j_t\right),
\]
due to the one-period assumption, where \( F_j \) denotes the distribution function of \( S_{j,t} \). Therefore, \( c^j_t = F_j^{-1}(p_j(t)) \), where \( F_j^{-1} \) is the inverse function of \( F_j \) if \( F_j \) is invertible, and the generalised inverse function of \( F_j \) if \( F_j \) is not invertible but continuous (cf. Lemma 5.2.2 for a closer analysis of \( F_j \)).

7.4 Approximation of the portfolio losses

We assume that on default of obligor \( j \), the defaulted company will still be able to pay back some fraction \( \delta_j \) of its obligations \( N_j \). The so-called recovery rate \( \delta_j \in [0,1] \) is assumed to be non-stochastic, and on default there will thus be a recovery payment of \( \delta_j \cdot N_j \) to the owner of the portfolio with the credit-risky assets underlying the CDO. Thus, the loss on default is assumed to equal \( \pi(j) := (1 - \delta_j)N_j \). In the above language of Definition 6.1.7 and the example in Section 6.3 we thus obtain

\[
\pi_t(j,2,\omega_j,\Psi_t) = \begin{cases} 
\pi(j), & \text{for } \omega_j = 1; \\
0, & \text{for } \omega_j = 2.
\end{cases}
\]

The overall credit losses \( C_{n,t} \) at time \( t \in [0,T] \) of this \( n \)-dimensional credit portfolio as introduced in Definition 6.1.8 then becomes

\[
C_{n,t} = \sum_{j=1}^{n} \pi(j) \cdot D_{j,t},
\]

with \( D_{j,t} = \mathbb{1}_{\{S_{j,t} \leq c^j_t\}} \) and \( c^j_t = F_j^{-1}(p_j(t)) \) as above. Its conditional counterpart \( B_{n,t} \) of Definition 6.2.1 turns into

\[
B_{n,t} = \mathbb{E}\left(C_{n,t} | M_t, R_t\right) = \sum_{j=1}^{n} \pi(j) \cdot \mathbb{E}\left(D_{j,t} | M_t, R_t\right)
= \sum_{j=1}^{n} \pi(j) \cdot \mathbb{P}\left(S_{j,t} \leq c^j_t | M_t, R_t\right) = \sum_{j=1}^{n} \pi(j) \cdot \Phi_{j,1,t} (M_t, R_t),
\]
as Assumption 2 in Section 6.2 is satisfied and where the expression \( \Phi_{j,1,t} (M_t, R_t) \) of Corollary 6.2.4 simplifies to

\[
\Phi_{j,1,t} (M_t, R_t) = \Phi\left(\frac{c^j_t - \beta_j^T M_t}{\sqrt{\omega_j} R_t}\right) \overset{d}{=} \Phi\left(\frac{c^j_t - \rho_j R_{1,t} Z_{1,t}}{\sqrt{\omega_j} R_t}\right) = \hat{\Phi}_{j,t} (R_{1,t}, Z_{1,t}, R_t).
\]
Similarly, the conditional relative overall losses \( (K_n(t))_{n \geq 0} \) at time \( t \in [0, T] \) are defined via

\[
K_n(t) := \mathbb{E} (L_n(t) | M_t, R_t) = \frac{B_{n,t}}{N_n},
\]

where \( L_n(t) \) are the unconditional relative overall losses at time \( t \in [0, T] \) in the \( n \)-dimensional portfolio as defined in the CDO structure in Section 7.1.

Under weak constraints on the notionals and the recovery rates (for example \( N_j \in [N_{\min}, N_{\max}] \) with \( N_{\min} > 0, N_{\max} < \infty \) and \( \delta_j \geq \delta_{\min} > 0 \) for all \( j \)), the condition

\[
\sum_{n=1}^{\infty} \frac{1}{N_n^2} (1 - \delta_n)^2 N_n^2 < \infty
\]

is satisfied and thus also Assumption 1 in Section 6.2. Then Theorem 6.2.3 and its Corollary 6.2.4 state that for every \( t \in (0, T] \) we have

\[
L_n(t) - K_n(t) = C_{n,t} \cdot \frac{B_{n,t}}{N_n} \to 0 \quad \text{as} \quad n \to \infty, \quad \text{almost surely.}
\]

Even though this large-portfolio approximation result holds in very general terms, one often assumes that the following homogeneity assumptions are satisfied in the portfolio:

**Homogeneous Portfolio Assumptions:**

1. \( \delta_j = \delta, \beta_j = \beta \) and \( \omega_j = \omega, \) for all \( j = 1, \ldots, n, \)
2. \( p_j(t) = p(t), \) for all \( j = 1, \ldots, n. \)

These then entail that also \( \rho_j = \rho \) and \( \pi(j) = (1-\delta_j)N_j = (1-\delta)N_j, \) for all \( j = 1, \ldots, n. \)

Therefore, the conditional relative overall losses \( K_n(t) \) simplify to

\[
K_n(t) = \frac{1}{N_n} \sum_{j=1}^{n} \pi(j) \cdot \Phi \left( \frac{c^j - \beta^T M_t}{\sqrt{\omega^2 R_t}} \right) = \Phi \left( \frac{c^1 - \beta^T M_t}{\sqrt{\omega^2 R_t}} \right) \cdot \frac{1}{N_n} \sum_{j=1}^{n} (1 - \delta)N_j = (1 - \delta) \cdot \Phi \left( \frac{c^1 - \rho R_{1,t}Z_{1,t}}{\sqrt{\omega^2 R_t}} \right),
\]

for all \( n \in \mathbb{N}. \)

In this case, the distribution function \( F_{K_n(t)} \) of \( K_n(t) \) no longer depends on \( n \) and
becomes
\[
F_{K_n(t)}(y) = \mathbb{P}\left( (1 - \delta) \cdot \Phi \left( \frac{c_l - \rho R_{1,t} Z_{1,t}}{\sqrt{\omega R_t}} \right) \leq y \right) \\
= \mathbb{P}\left( \frac{c_l - \rho R_{1,t} Z_{1,t}}{\sqrt{\omega R_t}} \leq \Phi^{-1}\left( \frac{y}{1 - \delta} \right) \right) \\
= \mathbb{P}\left( \frac{c_l - \Phi^{-1}\left( \frac{y}{1 - \delta} \right) \sqrt{\omega R_t}}{\rho R_{1,t}} \leq Z_{1,t} \right) \\
= \int_0^\infty \int_0^\infty \Phi\left( \frac{\Phi^{-1}\left( \frac{y}{1 - \delta} \right) \sqrt{\omega R_t} - c_l}{\rho r_1} \right) dG_1(r_1) dG_2(r) \\
= \int_0^\infty \int_{-\infty}^\infty \frac{\phi^{-1}\left( \frac{x-r}{\rho r_1} \right) \sqrt{\omega R_t} - c_l}{\rho r_1} \varphi(x) dx dG_1(r_1) dG_2(r) \\
= \int_0^\infty \int_{-\infty}^\infty \frac{\Phi^{-1}\left( \frac{y}{1 - \delta} \right) \sqrt{\omega R_t} - c_l}{\rho r_1} \cdot \frac{1}{1 - \delta} \frac{\sqrt{\omega R_t}}{\rho r_1} dz dG_1(r_1) dG_2(r),
\]
where the change of variables for the last equality was performed via the setting \( z := (1 - \delta) \Phi\left( \frac{x-r_{1,t}^+ + c_l^+}{\sqrt{\omega R_t}} \right) \). Hence, the density function \( f_{K_n(t)} \) of \( K_n(t) \) becomes
\[
f_{K_n(t)}(y) := \int_0^\infty \int_{-\infty}^\infty \frac{\Phi^{-1}\left( \frac{y}{1 - \delta} \right) \sqrt{\omega R_t} - c_l}{\rho r_1} \cdot \frac{1}{1 - \delta} \frac{\sqrt{\omega R_t}}{\rho r_1} dz dG_1(r_1) dG_2(r),
\]
for \( y \in \mathbb{R} \). The functions \( \Theta_i : [0, 1] \rightarrow [0, 1] \) with \( \Theta_i(x) = \frac{1}{\alpha_i - \alpha_{i-1}} \cdot \min\{0, x - \alpha_i \} \) as in Section 7.1 are increasing, piecewise linear and Lipschitz-continuous functions with Lipschitz constant \( c_{\Theta_i} := \frac{1}{\alpha_i - \alpha_{i-1}} \), as for any \( x, y \in [0, 1] \) with \( x \leq y \) we have
\[
|\Theta_i(y) - \Theta_i(x)| \leq c_{\Theta_i} \left\{ \begin{array}{ll}
0, & \text{if } x \geq \alpha_i \text{ or } y \leq \alpha_{i-1}; \\
y - x, & \text{if } x \in [\alpha_{i-1}, \alpha_i] \text{ and } y \in [\alpha_i, \alpha_{i-1}]; \\
y - \alpha_{i-1}, & \text{if } x \leq \alpha_{i-1} \text{ and } y \in [\alpha_{i-1}, \alpha_i]; \\
\alpha_i - x, & \text{if } x \in [\alpha_i, \alpha_{i-1}] \text{ and } y \leq \alpha_i; \\
\alpha_i - \alpha_{i-1}, & \text{if } x \leq \alpha_{i-1} \text{ and } y \geq \alpha_i;
\end{array} \right.
\]
\[
\leq c_{\Theta_i}(y - x).
\]
As a consequence of the Lipschitz continuity, if
\[
L_n(t) - K_n(t) \rightarrow 0
\]
holds almost surely for \( n \rightarrow \infty \), then also
\[
\mathbb{E}(\Theta_i(L_n(t))) - \mathbb{E}(\Theta_i(K_n(t))) \rightarrow 0
\]
7.4. Approximation of the portfolio losses

holds for $n \to \infty$. Indeed, we have

$$|E[\Theta_i(L_n(t)) - \Theta_i(K_n(t))]| \leq E[|\Theta_i(L_n(t)) - \Theta_i(K_n(t))|] \leq c_{\Theta_i} \cdot E(|L_n(t) - K_n(t)|),$$

which tends towards zero for $n \to \infty$ by dominated convergence, as $|L_n(t) - K_n(t)| \leq 2$ almost surely for all $t > 0$ and $n \in \mathbb{N}_0$.

Therefore, the expected relative losses on tranche $i$, i.e. $E(\Theta_i(L_n(t)))$, can for large $n$ be approximated by the simpler expression

$$E[\Theta_i(K_n(t))] = \int_0^1 \Theta_i(y) f_{K_1(t)}(y) \, dy = \int_{\alpha_{i-1}}^1 \Theta_i(y) f_{K_1(t)}(y) \, dy$$

(7.5)

which no longer depends on $n \in \mathbb{N}_0$. The last equality is obtained directly by virtue of Equation (7.4). In Section 7.5.2, we will further analyse the expected relative tranche losses for the various specifications of the mixtures of Normal distributions.

Under the Homogeneous Portfolio Assumptions (7.3), we can also easily approximate the probability of tranche $i$ being hit at a specific point in time $t > 0$ via

$$q_i := \mathbb{P}(L_n(t) > \alpha_{i-1}) \approx \mathbb{P}(K_n(t) > \alpha_{i-1}) = \int_{\alpha_{i-1}}^1 f_{K_1(t)}(y) \, dy$$

$$= \int_0^\infty \int_0^\infty \Phi \left( \frac{c_1^1 - \Phi^{-1} \left( \frac{\alpha_{i-1}}{1-\delta} \right)}{\sqrt{\omega}} \right) \varphi(x) \, dx \, dG_1(r_1) \, dG_2(r),$$

$$=: \tilde{q}_i.$$

For the loss given default on tranche $i$ at time $t > 0$, we then obtain

$$LGD_i := E[\Theta_i(L_n(t))|L_n(t) > \alpha_{i-1}] = \frac{1}{\mathbb{P}(L_n(t) > \alpha_{i-1})} \cdot E \left[ 1_{\{L_n(t) > \alpha_{i-1}\}} \Theta_i(L_n(t)) \right]$$

$$= \frac{E[\Theta_i(L_n(t))]}{q_i} \approx \frac{E[\Theta_i(K_n(t))]}{\tilde{q}_i} = \frac{\int_{\alpha_{i-1}}^1 \Theta_i(y) f_{K_1(t)}(y) \, dy}{\int_{\alpha_{i-1}}^1 f_{K_1(t)}(y) \, dy}.$$
7.5 Key quantities revisited within the models based on mixtures of Normal distributions

With the above CDO structure, the factor setup based on the mixtures of Normal distributions and the approximation to hand, we can now step towards analysing the expressions which are needed to price the CDO tranches, to calibrate the model to market data or to benchmark the model with other models. We will do this within the various Gaussian mixture specifications introduced in Section 5.3.2.

7.5.1 Default probabilities and default thresholds

Knowing the default probability $p_j(t)$ of a specific obligor $j$, we then need to compute the default threshold $c_j^t$ which corresponds to this default level under the risk-neutral probability measure $\mathbb{P}$, that is, such that $p_j(t) = \mathbb{P}(S_j,t \leq c_j^t)$. As, in general, one cannot give an explicit expression for the inverse function of $F_{j,t}$, we need to find this value $c_j^t$ via a numerical procedure. To this end we will have to evaluate the distribution function $F_{j,t}$ at possible threshold values $c_j^t$ in order to test whether these values correspond to the given default probability:

$$p_j(t) = F_{j,t}(c_j^t) = \mathbb{P}(S_j,t \leq c_j^t) = \mathbb{P}\left(\beta_j^T R_{1,t}W_t + R_{t}Y_{j,t} \leq c_j^t\right)$$

$$= \int_0^\infty \int_0^\infty \Phi\left(\frac{c_j^t}{\sqrt{\sigma_1^2 r_1^2 + \omega_j r^2}}\right) dG_1(r_1) dG_2(r),$$

where we have used Lemma 5.2.2 with $\sigma_1^2 = \rho_j^2$ and $\sigma_2^2 = \omega_j$.

In the following we want to apply the various mixtures of Normal distributions of Section 5.3.2 on the pricing of CDOs. To this end, we will have default probabilities given and then need to find the appropriate default thresholds which correspond to these default probabilities. Therefore, we will need to compute the double integrals in Equation (7.6) numerically with the different specifications:

In the multivariate t-distribution case:

Let $R_{1,t}$ possess a density $g_1$ with

$$g_1(r_1) := c_1 r_1^{-\nu-1} \exp\left(-\frac{\nu - 2}{2r_1^2}\right),$$
for $r_1 > 0$, with $\nu > 2$ and $c_1 := \frac{(\nu-2)^{\nu/2}}{2^\nu\Gamma(\nu/2)^2}$, and $R_t$ the density $g_2$ with

$$g_2(r) := c_2 r^{-\kappa-1} \exp \left( -\frac{\kappa-2}{2r^2} \right),$$

for $r > 0$, with $\kappa > 2$ and $c_2 := \frac{(\kappa-2)^{\kappa/2}}{2^{\kappa/2} \Gamma(\kappa/2)^2}$, as discussed in Section 5.3.2. Then integration by substitution with $p = \frac{\nu-2}{2r^2}$ and $q = \frac{\kappa-2}{2r^2}$, thus $r_1^2 = \frac{\nu-2}{2p}$, $r^2 = \frac{\kappa-2}{2q}$,

$$dr_1 = -\frac{1}{2} p^{-3/2} \left( \frac{\nu-2}{2} \right)^{1/2} dp \quad \text{and} \quad dr = -\frac{1}{2} q^{-3/2} \left( \frac{\kappa-2}{2} \right)^{1/2} dq$$

yields

$$\int_0^\infty \int_0^\infty \Phi \left( \frac{c_1^2}{\sqrt{\rho_j r_1^2 + \omega_j r^2}} \right) g_1(r_1)g_2(r) dr_1 dr = \int_0^\infty \int_0^\infty f_j(p, q)p^{\nu-1}e^{-p}q^{\kappa-1}e^{-q} dp dq,$$

with

$$f_j(p, q) := \frac{1}{\Gamma(\nu/2)\Gamma(\kappa/2)} \Phi \left( \frac{c_1^2}{\sqrt{\rho_j \frac{\nu-2}{2p} + \omega_j \frac{\kappa-2}{2q}}} \right)$$

as

$$c_1 c_2 \frac{1}{4} \left( \frac{\nu-2}{2} \right)^{-\frac{\nu}{2}} \left( \frac{\kappa-2}{2} \right)^{-\frac{\kappa}{2}} = \frac{1}{\Gamma(\nu/2)\Gamma(\kappa/2)}.$$

### In the multivariate Power Law case:

Let $R_{1,t}$ possess a density $g_1$ with

$$g_1(r_1) := \frac{2\alpha_1}{\alpha_1 c_1} \left( 1 + \frac{r_1}{c_1} \right)^{-2\alpha_1-1},$$

for $r_1 > 0$, with $\alpha_1 > 1$ and $c_1 := \sqrt{(\alpha_1 - 1)(2\alpha_1 - 1)}$, and $R_t$ the density $g_2$ with

$$g_2(r) := \frac{2\alpha_2}{\alpha_2 c_2} \left( 1 + \frac{r}{c_2} \right)^{-2\alpha_2-1},$$

for $r > 0$, with $\alpha_2 > 1$ and $c_2 := \sqrt{(\alpha_2 - 1)(2\alpha_2 - 1)}$, as discussed in Section 5.3.2. Then integration by substitution with again $p = 2\alpha_1 \log \left( 1 + \frac{r_1}{c_1} \right)$ and $q = 2\alpha_2 \log \left( 1 + \frac{r}{c_2} \right)$, yields

$$\int_0^\infty \int_0^\infty \Phi \left( \frac{c_1^2}{\sqrt{\rho_j r_1^2 + \omega_j r^2}} \right) g_1(r_1)g_2(r) dr_1 dr = \int_0^\infty \int_0^\infty f_j(p, q)e^{-p}e^{-q} dp dq,$$
Define as before in Section 5.3.2 the constant $I$ and the function $l_{\alpha} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$ via

\[ I := \int_0^\infty \exp(-\exp(-y)) \cdot y \exp(-y) dy \quad \text{and} \quad l_{\alpha}(y) := \exp(-y^{-\alpha}) \cdot \exp(-\exp(-y^{-\alpha})) \],

for any $y \in \mathbb{R}_0^+$ and $\alpha \in \mathbb{R}$.

In the multivariate Power Log Law case:

Let $R_{1,t}$ possess a density $g_1$ with

\[ g_1(r_1) := \frac{4\alpha_1^2}{c_{\alpha_1}} \left(1 + \frac{r_1}{c_{\alpha_1}}\right)^{-2\alpha_1 - 1} \cdot \log \left(1 + \frac{r_1}{c_{\alpha_1}}\right), \]

for $r_1 > 0$, with $\alpha_1 > 1$ and $c_{\alpha_1} := \sqrt{\frac{(1-\alpha_1)^2(1-2\alpha_1)^2}{6\alpha_1^2 - 6\alpha_1 + 1}}$, and $R_t$ the density $g_2$ with

\[ g_2(r) := \frac{4\alpha_2^2}{c_{\alpha_2}} \left(1 + \frac{r}{c_{\alpha_2}}\right)^{-2\alpha_2 - 1} \cdot \log \left(1 + \frac{r}{c_{\alpha_2}}\right), \]

for $r > 0$, with $\alpha_2 > 1$ and $c_{\alpha_2} := \sqrt{\frac{(1-\alpha_2)^2(1-2\alpha_2)^2}{6\alpha_2^2 - 6\alpha_2 + 1}}$, as discussed in Section 5.3.2. Then integration by substitution with $p = 2\alpha_1 \log \left(1 + \frac{r_1}{c_{\alpha_1}}\right)$ and $q = 2\alpha_2 \log \left(1 + \frac{r}{c_{\alpha_2}}\right)$, thus $r_1^2 = c_{\alpha_1}^2 \left(\exp\left(\frac{p}{2\alpha_1}\right) - 1\right)^2$, $r^2 = c_{\alpha_2}^2 \left(\exp\left(\frac{q}{2\alpha_2}\right) - 1\right)^2$, $dr_1 = \frac{c_{\alpha_1}}{2\alpha_1} \exp\left(\frac{p}{2\alpha_1}\right) dp$ and $dr = \frac{c_{\alpha_2}}{2\alpha_2} \exp\left(\frac{q}{2\alpha_2}\right) dq$ yields

\[
\int_0^\infty \int_0^\infty \Phi \left( \frac{c^j}{\sqrt{\rho_j^2 + \omega_j r^2}} \right) g_1(r_1)g_2(r) dr_1 dr = \int_0^\infty \int_0^\infty f_j(p,q)pe^{-p}qe^{-q}dpdq,
\]

with

\[ f_j(p,q) := \Phi \left( \frac{c^j}{\sqrt{\rho_j^2 + \omega_j}} \right). \]

In the multivariate Exp-Exp Law case:
Let then \( R_{1,t} \) possess a density \( g_1 \) with
\[
g_1(r_1) := \frac{\alpha_1}{c_{\alpha_1} l_1} \left( \frac{r_1}{c_{\alpha_1}} \right)^{-2\alpha_1 - 1} l_1 \left( \frac{r_1}{c_{\alpha_1}} \right),
\]
for \( r_1 > 0 \), with \( \alpha_1 > 1 \) and \( c_{\alpha_1} := \left( \frac{1}{7} \int_0^\infty \exp(-\exp(-y)) \cdot y^{1-\frac{2}{\alpha_1}} \exp(-y) dy \right)^{-\frac{1}{2}} \), and \( R_t \) the density \( g_2 \) with
\[
g_2(r) := \frac{\alpha_2}{c_{\alpha_2} l_2} \left( \frac{r}{c_{\alpha_2}} \right)^{-2\alpha_2 - 1} l_2 \left( \frac{r}{c_{\alpha_2}} \right),
\]
for \( r > 0 \), with \( \alpha_2 > 1 \) and \( c_{\alpha_2} := \left( \frac{1}{7} \int_0^\infty \exp(-\exp(-y)) \cdot y^{1-\frac{2}{\alpha_2}} \exp(-y) dy \right)^{-\frac{1}{2}} \). Then
integration by substitution with \( p = \left( \frac{c_{\alpha_1}}{r_1} \right)^{\alpha_1} \) and \( q = \left( \frac{c_{\alpha_2}}{r} \right)^{\alpha_2} \), thus \( r_1^2 = c_{\alpha_1}^2 p^{\frac{2}{\alpha_1}} \), \( r^2 = c_{\alpha_2}^2 q^{\frac{2}{\alpha_2}} \), \( dr_1 = -\frac{c_{\alpha_1} p^{-\frac{1}{\alpha_1} - 1}}{\alpha_1} dp \) and \( dr = -\frac{c_{\alpha_2} q^{-\frac{1}{\alpha_2} - 1}}{\alpha_2} dq \) yields
\[
\int_0^\infty \int_0^\infty \Phi \left( \frac{c_1}{\sqrt{p_j^2 r_1^2 + \omega_j r^2}} \right) g_1(r_1) g_2(r) dr_1 dr = \int_0^\infty \int_0^\infty f_j(p, q) p e^{-p} q e^{-q} dp dq,
\]
with
\[
f_j(p, q) := \frac{1}{T^2} \Phi \left( \frac{c_1}{\sqrt{p_j^2 c_{\alpha_1} p^{-\frac{2}{\alpha_1}} + \omega_j c_{\alpha_2} q^{-\frac{2}{\alpha_2}}}} \right) \exp(-\exp(-p) - \exp(-q)).
\]

The integrals in all the above cases are of the form \( \int_0^\infty f(x) x^\beta e^{-x} dx \), where \( \beta \geq 0 \) and with \( f \) being a continuous function. This gives rise to a numerical integration procedure via the Gauss-Laguerre quadrature, where one approximates the integrals of the above form via the finite sum \( \sum_{j=1}^N w_j f(x_j) \) for a sufficiently large \( N \in \mathbb{N}_0 \), appropriate weights \( w_j \) and appropriate abscissas \( x_j \), \( j = 1, \ldots, N \) (see e.g. Section 2.3 in [SS66] for more details on the Gauss-Laguerre quadrature). The standard Normal distribution function \( \Phi(\cdot) \) can be evaluated either via the Gauss-Hermite quadrature, where one approximates integrals of the form \( \int_0^\infty f(x) e^{-x^2} dx \) by a finite sum as in the previous Gauss-Laguerre case (see also e.g. Section 2.3 in [SS66] for more details on the Gauss-Hermite quadrature), or by using numerical procedures that were developed for evaluating the error function \( \text{erf}(y) := \frac{2}{\sqrt{\pi}} \int_0^y e^{-x^2} dx = 2\Phi(\sqrt{2}y) - 1 \), for \( y \geq 0 \) (see e.g. [Cod69]).

### 7.5.2 Expected tranche losses

As we have discussed in Section 7.2, one of the key quantities needed for the valuation of a CDO is given by the expected losses \( E(L_{a_{i-1}}^\alpha(t)) = E(\Theta_i(L_n(t))) \) on tranche \( i \).
at a specific time \( t \). By the approximation result given in Section 6.2 and adapted to our homogeneous portfolio specifications in Section 7.4, we can approximate this quantity by 
\[
E[\Theta_i(K_n(t))] = \int_{\alpha_i-1}^{1} \Theta_i(y)f_{K_1(t)}(y) \, dy
\]
(7.11)

and Equation (7.4) and the preceding Homogeneous Portfolio Assumptions (7.3) yield
\[
E[\Theta_i(K_n(t))] = \int_{0}^{\infty} \int_{0}^{\infty} \int_{-\infty}^{\infty} \Theta_i \left( 1 - \delta \right) \cdot \Phi \left( \frac{c_1^2 - \rho r_{1x}}{\sqrt{\omega r}} \right) \varphi(x) \, dx \, dG_1(r_1) \, dG_2(r).
\]

In order to apply the elliptical distributions model for the valuation of CDOs, we want to compute the last expression in Equation (7.11) numerically for the various Gaussian mixture specifications of Section 5.3.2. Once more, we aim at transforming these integrals in such a way that we can exploit them with the use of standard integration techniques.

### In the multivariate t-distribution case:

Let \( R_{1,t} \) and \( R_t \) possess densities \( g_1 \) and \( g_2 \) as in the multivariate t-distribution case in Section 7.5.1. Integrating the right-hand side of Equation (7.11) by substitution of \( p = \frac{\nu - 2}{2r_1}, \quad q = \frac{\kappa - 2}{2r_2} \) and, additionally to the procedure before, with \( y = \frac{x}{\sqrt{2}} \) yields
\[
E[\Theta_i(K_n(t))] = \int_{0}^{\infty} \int_{0}^{\infty} \int_{-\infty}^{\infty} f_j(y,p,q) e^{-y^2} p^{\frac{\nu}{2}-1} e^{-p} q^{\frac{\kappa}{2}-1} e^{-q} \, dy \, dp \, dq,
\]
(7.12)
where
\[
f_j(y,p,q) := \frac{1}{\Gamma(\nu/2)\Gamma(\kappa/2)\sqrt{\pi}} \Theta_i \left( 1 - \delta \right) \cdot \Phi \left( \frac{c_1^2 - \rho (\nu - 2)^{1/2} p^{-1/2} y}{\sqrt{\omega} (\frac{\kappa}{2})^{1/2} q^{-1/2}} \right),
\]
as
\[
c_1c_2^2 \frac{1}{4} \left( \frac{\nu - 2}{2} \right)^{-\nu} \left( \frac{\kappa - 2}{2} \right)^{-\kappa} \sqrt{2} \Gamma(\nu/2)\Gamma(\kappa/2)\sqrt{\pi} = 1
\]

### In the multivariate Power Law case:

Let \( R_{1,t} \) and \( R_t \) possess densities \( g_1 \) and \( g_2 \) as in the Power Law case in Section 7.5.1. Integrating the right-hand side of Equation (7.11) by substitution of \( p = 2\alpha_1 \log \left( 1 + \frac{x}{c_1} \right) \) and \( q = 2\alpha_2 \log \left( 1 + \frac{x}{c_2} \right) \) and, additionally to the procedure before, with \( y = \frac{x}{\sqrt{2}} \) yields
\[
E[\Theta_i(K_n(t))] = \int_{0}^{\infty} \int_{0}^{\infty} \int_{-\infty}^{\infty} f_j(y,p,q) e^{-y^2} e^{-p} e^{-q} \, dy \, dp \, dq,
\]
(7.13)
7.5. Key quantities revisited within the models based on mixture distributions

where

\[ f_j(y, p, q) := \frac{1}{\sqrt{\pi}} \Theta_i \left( (1 - \delta) \cdot \Phi \left( \frac{c_1^t - \sqrt{2} \rho c_{\alpha_1} \left( \exp \left( \frac{p}{2\alpha_1} \right) - 1 \right)}{\sqrt{\omega c_{\alpha_2}} \left( \exp \left( \frac{q}{2\alpha_2} \right) - 1 \right)} \right) \right) \, . \]

In the multivariate Power Log Law case:

Let \( R_{1,t} \) and \( R_t \) possess densities \( g_1 \) and \( g_2 \) as in the Power Log Law case in Section 7.5.1. Integrating the right-hand side of Equation (7.11) by substitution of \( p = 2\alpha_1 \log \left( 1 + \frac{r_1}{c_{\alpha_1}} \right) \) and \( q = 2\alpha_2 \log \left( 1 + \frac{r}{c_{\alpha_2}} \right) \) and, additionally to the procedure before, with \( y = \frac{x}{\sqrt{2}} \) yields

\[ \mathbb{E}[\Theta_i(K_n(t))] = \int_0^\infty \int_0^\infty \int_{-\infty}^\infty f_j(y, p, q) e^{-y^2} p e^{-p} q e^{-q} dy dp dq, \quad (7.14) \]

where

\[ f_j(y, p, q) := \frac{1}{\sqrt{\pi}} \Theta_i \left( (1 - \delta) \cdot \Phi \left( \frac{c_1^t - \sqrt{2} \rho c_{\alpha_1} \left( \exp \left( \frac{p}{2\alpha_1} \right) - 1 \right)}{\sqrt{\omega c_{\alpha_2}} \left( \exp \left( \frac{q}{2\alpha_2} \right) - 1 \right)} \right) \right) \, . \]

In the multivariate Exp-Exp Law case:

Let \( R_{1,t} \) and \( R_t \) possess densities \( g_1 \) and \( g_2 \) as in the multivariate Exp-Exp Law case in Section 7.5.1. Integrating the right-hand side of Equation (7.11) by substitution of \( p = \left( \frac{\alpha_1}{c_{\alpha_1}} \right)^{\alpha_1} \) and \( q = \left( \frac{\alpha_2}{c_{\alpha_2}} \right)^{\alpha_2} \) and, additionally to the procedure before, with \( y = \frac{x}{\sqrt{2}} \) yields

\[ \mathbb{E}[\Theta_i(K_n(t))] = \int_0^\infty \int_0^\infty \int_{-\infty}^\infty f_j(y, p, q) e^{-y^2} p e^{-p} q e^{-q} dy dp dq, \quad (7.15) \]

where

\[ f_j(y, p, q) := \frac{1}{T^2 \sqrt{\pi}} \Theta_i \left( (1 - \delta) \cdot \Phi \left( \frac{c_1^t - \sqrt{2} \rho c_{\alpha_1} \left( \exp \left( \frac{p}{2\alpha_1} \right) - 1 \right)}{\sqrt{\omega c_{\alpha_2}} q^{-\frac{1}{\alpha_2}} y} \right) \exp(-e^{-p} - e^{-q}). \]

The three-dimensional integrals in all the above expressions consist of nested one-dimensional integrals that are either of the form \( \int_0^\infty f(x)x^\beta e^{-x^2} dx \), where \( \beta \geq 0 \), or of the form \( \int_{-\infty}^\infty f(x)e^{-x^2} dx \), with \( f \) being a continuous function. In order to compute the former type of integral numerically, we can once more use the Gauss-Laguerre quadrature, and for the latter one the Gauss-Hermite quadrature (see e.g. Section 2.3 in [SS66] for more details on both types of quadrature procedures).
7.5. Key quantities revisited within the models based on mixture distributions

7.5.3 The number of defaults

The Bernoulli random variables \( D_{j,t} = \mathbb{1}_{\{S_{j,t} \leq c_j\}} \), \( j = 1, \ldots, n \), are only independent if we condition on a realization of the vector \((M_t, R_t)\). Under the Homogeneous Portfolio Assumptions (7.3), the number of defaults in the portfolio being \( k \) has the following probability

\[
P(D_1 + \ldots + D_n = k) = \mathbb{E}(P(D_1 + \ldots + D_n = k | R_{1,t}, Z_{1,t}, R_t))
\]

\[
= \binom{n}{k} \int_0^\infty \int_0^\infty \int_{-\infty}^{\infty} g(r_1, x, r)^k (1 - g(r_1, x, r))^{n-k} \varphi(x) \, dx \, dG_1(r_1) \, dG_2(r),
\]

where

\[
g(r_1, x, r) := P(D_1 = 1 | R_{1,t} = r_1, Z_{1,t} = x, R_t = r)
\]

\[
= P \left( \frac{Y_1}{\sqrt{\omega}} \leq \frac{F_1^{-1}(p) - \rho r_1 x}{\sqrt{\omega} r} \right)
\]

\[
= \Phi \left( \frac{F_1^{-1}(p) - \rho r_1 x}{\sqrt{\omega} r} \right).
\]

When specifying the distribution functions \( G_1 \) and \( G_2 \) to be of the previously discussed mixtures of Normal distributions, the preceding probability can likewise be transformed via subsequent substitutions into nested integrals of the forms suitable for Gaussian quadrature formulas, as for the default probabilities in Section 7.5.1 and the expected tranche losses in Section 7.5.2.
7.6 Application of the models on CDO data

In this section we will investigate the previously discussed models with respect to their behaviour when pricing Collateralized Debt Obligations. In a first step we will point out important issues, such as which data we use and how the calibration takes place. We will then provide a simulation study to analyse the different price levels the models produce on different choices for the correlation and the distributional parameters, and will finally apply the models on actual market data of credit derivatives, namely on the prices of the standard iTraxx indices.

7.6.1 The sequence of computations needed for the valuation of CDOs

In order to price CDO tranches within our elliptical framework, we need to specify a correlation parameter $\rho$, we set $\omega = 1 - \rho^2$ as before, choose scaling distribution functions $G_1$ and $G_2$ such as those introduced in Section 5.3.2, and require risk-neutral default probability curves $(p_j(t))_{t \geq 0}$ (we will assume that these curves coincide for all $j = 1, \ldots, 125$ as in Assumption (7.3)). Then the following sequence of computations needs to be performed for the pricing of the CDO tranches:

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $\rho$, $G_1, G_2$, $(p_j(t))_{t \geq 0}$</td>
<td>$(c_j(t))_{t \geq 0}$</td>
<td>the default thresholds</td>
</tr>
<tr>
<td>2. $\rho$, $G_1, G_2$, $(c_j(t))_{t \geq 0}$</td>
<td>$E(\Theta_i(K_n(\cdot)))$</td>
<td>the expected tranche losses</td>
</tr>
<tr>
<td>3. $E(\Theta_i(K_n(\cdot)))$</td>
<td>$PVPL_i(0, 1)$</td>
<td>the premium legs</td>
</tr>
<tr>
<td>$E(\Theta_i(K_n(\cdot)))$</td>
<td>$PVDL_i$</td>
<td>the default legs</td>
</tr>
<tr>
<td>4. $PVDL_i$, $PVPL_i(0, 1)$</td>
<td>$\zeta_i$</td>
<td>the upfront payments &amp; spreads</td>
</tr>
</tbody>
</table>

For step 1. one has to invert the computations from the default thresholds to the default probabilities, which we have detailed in Section 7.5.1. Step 2. requires the computation of the three-dimensional integrals discussed in Section 7.5.2. Consult Section 7.2 for steps 3. and 4.

7.6.2 The data basis we use

Already in Section 1.2 and in Section 2.2 we have introduced the iTraxx Europe Indices which are based on a portfolio of 125 CDS reference entities. There are not only the 125 single-name CDS quotes available, but in particular also quotes for the five standardised iTraxx tranches, which range from 0% to 3%, 3% to 6%, 6% to 9%, 9% to 12%, and 12% to 22% of the portfolio losses. We obtained the CDS prices as well as the tranched
iTraxx prices for the iTraxx Europe 3 series from the Bloomberg system.\footnote{In 2005 the Landesbank Baden-Württemberg equipped the University of Ulm for educational purposes with a number of computer terminals, which are connected to the Bloomberg database. There is a vast amount of real-time and historical data available covering most of the financial markets and it provides one with numerous analytical tools.} Concerning the tranched iTraxx values, we focused on the prices of the five standard tranches observed in a time range spanning from January 2006 to May 2006 for the 5 year CDO, which reaches its maturity on the 20th of September 2010. In particular, we used the prices obtained for the 6th of April 2006 to perform an extensive analysis of the various models and then utilized fortnightly prices for investigations of the evolution of the models over time (see Table A.1). In order to build the individual default probability curves, we made use of the 3 year and the 5 year single-name CDS quotes of the 125 reference entities, which we likewise downloaded for these days during the time span from January 2006 to May 2006. Additionally, a Euro Benchmark Yield curve served us for building a discount curve, which we also obtained from the Bloomberg system for these days.

### 7.6.3 Calibrating the models

#### The individual default probabilities

Under the Homogeneous Portfolio Assumptions (7.3) one supposes that the variations between the individual default probabilities of the firms represented in the reference portfolio are negligible. At every particular day and for every maturity we therefore used the average of the 125 different CDS quotes as the hypothetical single CDS quote representing the entire portfolio (see Table A.2). On the 6th of April 2006, for example, we obtained an annual spread of 19.36 bps for a 3 year CDS and a spread of 34.47 bps for a CDS with a time to maturity of 5 years. In order to obtain a default probability curve from this set of CDS quotes, we performed a bootstrap algorithm and fitted an intensity model with a piecewise constant intensity function successively to the quotes (see e.g. Section 8.2.4 in [BR02] for this standard approach and the underlying assumptions). This then results in default probabilities at time \( t \in (0, 5] \) of \( pd(t) = 1 - \exp(-\lambda_1 \min(t, 3) - \lambda_2 (t - 3)^+) \), where \( \lambda_1 = 32 \) bps and \( \lambda_2 = 100 \) bps.

#### Fitting the correlation parameters

Once one has chosen a particular specification, such as the Exp-Exp Law, there are the tail-index parameters \( \alpha_1, \alpha_2 \) and the correlation parameter that one can use for fitting the model to given quotes of the CDO tranches, such as the previously discussed tranched iTraxx values. In order to reduce the dimension of this multidimensional optimisation problem we considered several combinations of possible tail index parameters, and then aimed at finding the correlation parameter \( \rho \) that minimizes the distance between the given market prices and the theoretical values obtained from the model with this choice.
of $\alpha_1, \alpha_2$ and $\rho$

$$\text{dist}(\zeta_{\text{market}}^1, \ldots, \zeta_{\text{market}}^q, \zeta_{\alpha_1, \alpha_2, \rho}^1, \ldots, \zeta_{\alpha_1, \alpha_2, \rho}^q)^T \xrightarrow{\rho} \min,$$

where $\zeta_{\text{market}}^i$ is the market price of the $i$-th tranche and $\zeta_{\alpha_1, \alpha_2, \rho}^i$ the theoretical model price for a particular parameter setting of $\alpha_1, \alpha_2$ and $\rho$. These values $\zeta_{\text{market}}^i, \zeta_{\alpha_1, \alpha_2, \rho}^i$ can either refer to the upfront payments as in the iTraxx equity tranche case or to the fair spreads being quarterly exchanged (see Section 7.2).

In principle there are several ways in which one can measure this distance as we have already outlined in Section 4.5.2. For values $x := (x_1, \ldots, x_q)^T$ and $y := (y_1, \ldots, y_q)^T$, one can for example use the sum of the squared errors $\text{dist}(x, y) := \sum_{i=1}^q (x_i - y_i)^2$, the sum of the absolute errors $\text{dist}(x, y) := \sum_{i=1}^q |x_i - y_i|$, or the sum of the relative errors $\text{dist}(x, y) := \sum_{i=1}^q \frac{|x_i - y_i|}{x_i}$ as distance measures. The former two distance measures, using absolute values, place a focus on the equity and more junior tranches in the optimisation procedure, as their prices are typically higher than those for the more senior tranches.

When there is an upfront payment for a tranche, such as for the iTraxx equity tranches, the difficulty lies in particular in the question of how to treat the upfront payment in comparison with the periodical spread payments of the other tranches. One possibility is to treat the upfront payment exactly as the other spread payments, but where all figures are dealt with e.g. in basis points (the iTraxx equity tranche is usually quoted in percent). In this case, the above distance measures could also be used. However, treating upfront payments the same as the fair spreads being paid on a regular basis can be viewed with a critical eye.

For the iTraxx case one therefore usually calibrates the model to perfectly fit only the equity tranche with an appropriate $\rho$. In order to analyse the quality of this now calibrated model to replicate the entire set of market quotes, one can then apply a distance measure on the periodically exchanged spreads of the remaining tranches. We will pursue this procedure also for the simulation study, that will follow in the next section.

### 7.6.4 Simulation study

In the following we present a simulation study, where we compare the different elliptical distributions models, which we have introduced in the previous sections and chapters, with the standard Gaussian model and the Hull & White model (cf. Section 4.5.3)\(^2\). We will analyse the correlation parameters which are needed within the various models to replicate specific equity tranche prices and will also investigate the resulting prices these models produce at different correlation levels for exemplary mezzanine and senior CDO tranches.

\(^2\) Recall that the standard Gaussian model falls into our framework based on elliptical distributions, but the Hull & White model does not belong to this newly introduced class of models.
7.6. Application of the models on CDO data

We assume to have a homogeneous portfolio of 100 names where each name is represented in the portfolio with the same notional. The recovery rates are all assumed to be $\delta = 40\%$, and we suppose that there are Credit Default Swaps with a time to maturity of 5 years for these names, which are all quoted at 40 bps per annum. The CDO constructed on this portfolio shall also have a time to maturity of 5 years and shall be split into tranches from 0% to 3%, from 3% to 10%, and from 10% to 100% of the portfolio losses. The discount factors were constructed from the Euro Benchmark Yield curve on the 6th of April 2006.

As the Gaussian model is the standard model for pricing portfolio credit derivatives, in a first step we computed the prices of the three tranches the Gaussian model produces for different correlation levels $\rho$. As for the mezzanine (3% - 10% tranche) and the senior tranches (10% - 100% tranche), also for the equity tranche (0% - 3% tranche) we computed the fair spread without allowing for an upfront payment. The resulting values can be found in Table 7.1. One can see that, while the price of the equity tranche declines and the value of the senior tranche increases in $\rho$, the price of the mezzanine tranche first increases in $\rho$, but finally decreases again for high values of $\rho$.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\rho^2$</th>
<th>0% - 3%</th>
<th>3% - 10%</th>
<th>10% - 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0%</td>
<td>0.0%</td>
<td>1973.12</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>30.0%</td>
<td>9.0%</td>
<td>1682.49</td>
<td>67.78</td>
<td>0.03</td>
</tr>
<tr>
<td>40.0%</td>
<td>16.0%</td>
<td>1456.15</td>
<td>115.40</td>
<td>0.43</td>
</tr>
<tr>
<td>50.0%</td>
<td>25.0%</td>
<td>1217.51</td>
<td>158.09</td>
<td>1.67</td>
</tr>
<tr>
<td>60.0%</td>
<td>36.0%</td>
<td>980.11</td>
<td>182.50</td>
<td>4.63</td>
</tr>
<tr>
<td>80.0%</td>
<td>64.0%</td>
<td>558.16</td>
<td>189.43</td>
<td>13.61</td>
</tr>
<tr>
<td>95.0%</td>
<td>90.3%</td>
<td>218.77</td>
<td>140.42</td>
<td>26.45</td>
</tr>
</tbody>
</table>

Table 7.1: CDO tranche spreads in bps within the Gaussian model.

In the sequel we calibrated various models in such a way that they perfectly matched the equity tranche prices produced by the Gaussian model. Table 7.2 shows which correlation level $\rho$ in the Gaussian model corresponds to which correlation in the other models. The table includes the values for the Hull & White model (HW) and the elliptical distributions models constructed with the use of the multivariate t-distribution (MVT) and the multivariate Exp-Exp Law (ExpExp) as introduced in Section 5.3.2. For all these models one needs to specify two parameters, one for the distribution of the factor and one for the idiosyncratic risk vector. These parameters are given in the table directly after the name of the model, where the first parameter corresponds to the factor distribution and the second to the idiosyncratic risk vector. Note that these parameters represent the tail index parameters in the ExpExp model and the degrees of freedom in the HW and the MVT models.$^3$ The choice of the parameters was based on the ability of the models with this set of parameters to reproduce the iTraxx tranche prices in the following section.$^4$

---

$^3$ The tail index parameter in the MVT model is just half the degrees of freedom.

$^4$ For the Hull & White model we used the parameters proposed by Hull and White [HW04] or discussed in the literature (see e.g. [BGL05]).
The MVT 3-20, the MVT 3-40 and the ExpExp 1.5-20 models were not able to match the equity price generated by the Gaussian model at zero correlation. However, the iTraxx prices we studied usually implied a correlation $\rho$ under the Gaussian model in the range of 30% to 60%. The elliptical distributions models based on the Power Law and on the Power Log Law showed an entirely different behaviour, could not reproduce the equity tranche prices for any choice of parameters and were thus not included in this study.

<table>
<thead>
<tr>
<th>Model</th>
<th>0.0%</th>
<th>30.0%</th>
<th>40.0%</th>
<th>50.0%</th>
<th>60.0%</th>
<th>80.0%</th>
<th>95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HW 4-4</td>
<td>6.4%</td>
<td>35.3%</td>
<td>45.9%</td>
<td>55.9%</td>
<td>65.3%</td>
<td>81.9%</td>
<td>95.1%</td>
</tr>
<tr>
<td>HW 5-5</td>
<td>4.8%</td>
<td>35.1%</td>
<td>45.7%</td>
<td>55.5%</td>
<td>64.8%</td>
<td>81.4%</td>
<td>95.0%</td>
</tr>
<tr>
<td>MVT 3-20</td>
<td>–</td>
<td>11.9%</td>
<td>31.6%</td>
<td>44.8%</td>
<td>56.3%</td>
<td>76.1%</td>
<td>92.7%</td>
</tr>
<tr>
<td>MVT 3-40</td>
<td>–</td>
<td>24.0%</td>
<td>36.2%</td>
<td>47.1%</td>
<td>57.4%</td>
<td>76.1%</td>
<td>92.6%</td>
</tr>
<tr>
<td>ExpExp 1.5-10</td>
<td>–</td>
<td>26.2%</td>
<td>37.1%</td>
<td>47.4%</td>
<td>57.5%</td>
<td>75.8%</td>
<td>92.4%</td>
</tr>
<tr>
<td>ExpExp 1.5-20</td>
<td>0.2%</td>
<td>27.9%</td>
<td>38.0%</td>
<td>48.0%</td>
<td>57.7%</td>
<td>75.7%</td>
<td>92.4%</td>
</tr>
<tr>
<td>ExpExp 1.5-50</td>
<td>0.1%</td>
<td>28.2%</td>
<td>38.2%</td>
<td>48.1%</td>
<td>57.8%</td>
<td>75.7%</td>
<td>92.4%</td>
</tr>
</tbody>
</table>

Table 7.2: The parameter $\rho$ needed in the various models to replicate the equity tranche prices produced by the Gaussian model.

One can observe that the proposed mixture models such as the MVT and the ExpExp model in general need lower correlations than the Gaussian model to reproduce the same value for the equity tranche. This holds for almost all correlation levels we used for the Gaussian model (except for the zero correlation case). On the other hand, the given specifications of the HW model require a higher correlation than the Gaussian model.

Using these so calibrated models, we can then price the other tranches of this exemplary CDO. The following two tables (Table 7.3 and Table 7.4) show the prices of the mezzanine and the senior tranche, that result from the previously calibrated models.

<table>
<thead>
<tr>
<th>Model</th>
<th>0.0%</th>
<th>30.0%</th>
<th>40.0%</th>
<th>50.0%</th>
<th>60.0%</th>
<th>80.0%</th>
<th>95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>0.00</td>
<td>67.78</td>
<td>115.40</td>
<td>158.09</td>
<td>182.50</td>
<td>189.43</td>
<td>140.42</td>
</tr>
<tr>
<td>HW 4-4</td>
<td>0.23</td>
<td>47.42</td>
<td>72.93</td>
<td>92.88</td>
<td>104.62</td>
<td>115.62</td>
<td>99.59</td>
</tr>
<tr>
<td>HW 5-5</td>
<td>0.04</td>
<td>53.26</td>
<td>83.00</td>
<td>105.08</td>
<td>120.44</td>
<td>126.72</td>
<td>104.19</td>
</tr>
<tr>
<td>MVT 3-20</td>
<td>–</td>
<td>66.13</td>
<td>98.75</td>
<td>120.93</td>
<td>136.03</td>
<td>141.84</td>
<td>110.48</td>
</tr>
<tr>
<td>MVT 3-40</td>
<td>–</td>
<td>58.58</td>
<td>87.36</td>
<td>111.52</td>
<td>129.18</td>
<td>139.30</td>
<td>109.95</td>
</tr>
<tr>
<td>ExpExp 1.5-10</td>
<td>–</td>
<td>51.49</td>
<td>81.05</td>
<td>106.57</td>
<td>128.03</td>
<td>140.68</td>
<td>115.03</td>
</tr>
<tr>
<td>ExpExp 1.5-20</td>
<td>0.51</td>
<td>48.22</td>
<td>77.95</td>
<td>104.05</td>
<td>126.93</td>
<td>140.24</td>
<td>115.22</td>
</tr>
<tr>
<td>ExpExp 1.5-50</td>
<td>0.00</td>
<td>47.54</td>
<td>77.23</td>
<td>103.35</td>
<td>126.69</td>
<td>140.23</td>
<td>115.29</td>
</tr>
</tbody>
</table>

Table 7.3: Resulting prices in bps for 3% - 10% tranche.
7.6. Application of the models on CDO data

<table>
<thead>
<tr>
<th>Model</th>
<th>0.0%</th>
<th>30.0%</th>
<th>40.0%</th>
<th>50.0%</th>
<th>60.0%</th>
<th>80.0%</th>
<th>95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>0.00</td>
<td>0.03</td>
<td>0.43</td>
<td>1.67</td>
<td>4.63</td>
<td>13.61</td>
<td>26.45</td>
</tr>
<tr>
<td>HW 4-4</td>
<td>0.00</td>
<td>1.36</td>
<td>3.32</td>
<td>6.17</td>
<td>10.00</td>
<td>18.94</td>
<td>29.69</td>
</tr>
<tr>
<td>HW 5-5</td>
<td>0.00</td>
<td>0.96</td>
<td>2.63</td>
<td>5.33</td>
<td>8.88</td>
<td>18.07</td>
<td>28.92</td>
</tr>
<tr>
<td>MVT 3-20</td>
<td>–</td>
<td>0.02</td>
<td>1.38</td>
<td>3.94</td>
<td>7.45</td>
<td>16.73</td>
<td>28.95</td>
</tr>
<tr>
<td>MVT 3-40</td>
<td>–</td>
<td>0.50</td>
<td>2.08</td>
<td>4.52</td>
<td>7.88</td>
<td>16.90</td>
<td>28.97</td>
</tr>
<tr>
<td>ExpExp 1.5-10</td>
<td>–</td>
<td>0.89</td>
<td>2.43</td>
<td>4.81</td>
<td>7.92</td>
<td>16.77</td>
<td>28.10</td>
</tr>
<tr>
<td>ExpExp 1.5-20</td>
<td>0.00</td>
<td>1.06</td>
<td>2.61</td>
<td>4.96</td>
<td>7.97</td>
<td>16.80</td>
<td>28.08</td>
</tr>
<tr>
<td>ExpExp 1.5-50</td>
<td>0.00</td>
<td>1.09</td>
<td>2.65</td>
<td>5.00</td>
<td>7.98</td>
<td>16.80</td>
<td>28.08</td>
</tr>
</tbody>
</table>

Table 7.4: Resulting prices in bps for 10% - 100% tranche.

At the extreme end for $\rho = 100\%$ the prices of the tranches across the various models must coincide, as they are all calibrated to the same default probability curve. Indeed, the expected losses take the following values regardless of the model being used:

\[
\mathbb{E}(\Theta_i(K_n(t))) = \mathbb{E}\left(\Theta_i\left(1 - \delta \mathbb{1}_{M_t \leq c_1}\right)\right) = \Theta_i (1 - \delta) \cdot \mathbb{E}\left(\mathbb{1}_{M_t \leq c_1}\right)
\]

\[
= \Theta_i (1 - \delta) \cdot \mathbb{P}\left(S_{t,1} \leq c_1\right) = \Theta_i (1 - \delta) \cdot p_1(t),
\]

for any $i = 1, \ldots, q$ and any $t > 0$.

However, especially in the important region for the correlation parameter $\rho$ between 30% and 50% the prices largely differ between the various models. While the Gaussian model assigns very high prices to the mezzanine 3% - 10% tranche in comparison with the models based on elliptical distributions, the picture is exactly the opposite for the senior 10% - 100% tranche, where the Gaussian model prices are by far less than those induced by the other models. This reflects the fact that the Gaussian model puts too little probability to joint extremal behaviour of the asset values triggering the defaults, in particular when compared to the tail-dependent elliptical distributions. As one uses the same level of individual default probabilities, in a Gaussian setup the single defaults in the portfolio must then be assigned a comparably high probability and therefore lead to relatively high prices for the tranches covering the lower part of the portfolio loss distribution, such as the 3% - 10% tranche.

7.6.5 Performance of the models on the iTraxx data

In this section we want to analyse how the elliptical distributions models perform with respect to the pricing of standardised and traded CDO tranches. As an exemplary date we chose the 6th of April 2006, when we observed the following market values for the five tranches of the iTraxx Europe 3 series with a maturity on the 20th of September 2010:
Table 7.5: The market values of the iTraxx Europe 3 series tranches on the 6th of April 2006.

Note that the value given for the equity tranche is denoted in percent and represents the upfront payment, which is exchanged between the counterparties besides the periodical spread payments of 500 bps. As mentioned before, the average CDS spreads in the underlying portfolio were at an annual spread of 19.36 bps for a 3 year CDS and at a spread of 34.47 bps for a 5 year CDS. The recovery rate is equal to $\delta = 40\%$ for all entities according to the iTraxx rules.

**Model based on the multivariate t-distribution**

The following table shows the prices for the iTraxx tranches, which were generated by the elliptical distributions model, where the factor and the idiosyncratic risk vector both follow multivariate $t$-distributions with various parameter combinations. For every model we specified two parameters (e.g. 3 - 40), of which the first value (e.g. 3) denotes the degrees of freedom (df) used for the factor distribution and the second (e.g., 40) those for the distribution of the idiosyncratic risk vector. The correlation parameter $\rho$ was chosen such that the absolute error in the equity tranche was minimal.

In half of the cases, which are displayed below, the first tranche can be perfectly matched, however within the MVT 5-5, the MVT 10-10, the MVT 10-5 and the MVT 50-5 models the market price of the equity tranche was not attained. In the table one can see that the models, where the df for the factor are lower than those for the idiosyncratic risk vector, yield better results in fitting the market data and therefore lower values for the absolute errors than the other models, where the situation is reversed.

By construction of the model in Section 5.3.2, the df have to be larger than two. When we fixed the df for the factor at the lowest possible integer value, i.e. at 3, which corresponds to a tail-index of 1.5, the price of the 3% - 6% was decreasing towards its market value and the prices for the other tranches were increasing with growing df for the idiosyncratic risk vector. As the value of the 3% - 6% is relatively large compared to the more senior tranches, the sum of the absolute errors therefore also decreased with growing df for the idiosyncratic risk vector. Additionally, one can observe that with growing df also the level of correlation increases, which is needed to match the equity tranche.
7.6. Application of the models on CDO data

<table>
<thead>
<tr>
<th></th>
<th>iTraxx</th>
<th>Gaussian</th>
<th>3 - 20</th>
<th>3 - 40</th>
<th>3 - 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td></td>
<td>39.67%</td>
<td>28.41%</td>
<td>32.88%</td>
<td>33.46%</td>
</tr>
<tr>
<td>$\rho^2$</td>
<td></td>
<td>0.00%</td>
<td>15.74%</td>
<td>8.07%</td>
<td>10.81%</td>
</tr>
<tr>
<td>0% - 3%</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
</tr>
<tr>
<td>3% - 6%</td>
<td>56.04</td>
<td>120.13</td>
<td>101.49</td>
<td>81.62</td>
<td>78.40</td>
</tr>
<tr>
<td>6% - 9%</td>
<td>16.65</td>
<td>21.28</td>
<td>20.85</td>
<td>22.79</td>
<td>23.01</td>
</tr>
<tr>
<td>9% - 12%</td>
<td>9.17</td>
<td>5.51</td>
<td>8.27</td>
<td>12.16</td>
<td>12.92</td>
</tr>
<tr>
<td>12% - 22%</td>
<td>2.57</td>
<td>0.55</td>
<td>3.28</td>
<td>5.40</td>
<td>5.71</td>
</tr>
<tr>
<td>abs. error</td>
<td>74.40</td>
<td>51.26</td>
<td>37.54</td>
<td>35.61</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>5 - 20</th>
<th>5 - 5</th>
<th>10 - 10</th>
<th>10 - 5</th>
<th>50 - 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>28.86%</td>
<td>0.00%</td>
<td>1.10%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>$\rho^2$</td>
<td>8.33%</td>
<td>0.00%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>0% - 3%</td>
<td>20.04</td>
<td>12.10</td>
<td>19.22</td>
<td>12.10</td>
<td>12.10</td>
</tr>
<tr>
<td>3% - 6%</td>
<td>112.59</td>
<td>209.41</td>
<td>134.11</td>
<td>209.41</td>
<td>209.41</td>
</tr>
<tr>
<td>6% - 9%</td>
<td>20.84</td>
<td>66.31</td>
<td>26.73</td>
<td>66.31</td>
<td>66.31</td>
</tr>
<tr>
<td>9% - 12%</td>
<td>6.29</td>
<td>40.28</td>
<td>4.51</td>
<td>40.28</td>
<td>40.28</td>
</tr>
<tr>
<td>12% - 22%</td>
<td>1.71</td>
<td>4.11</td>
<td>0.31</td>
<td>4.11</td>
<td>4.11</td>
</tr>
<tr>
<td>abs. error</td>
<td>64.48</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 7.6: Various specifications of the model based on the multivariate $t$-distribution. The iTraxx Europe Series 3 tranches are from the 6th of April 2006. All prices are denoted in basis points except for the equity tranche values, which are given in percent.

Model based on the Exp-Exp Law

The table below shows the prices for the iTraxx tranches, which were generated by the elliptical distributions model, where the factor and the idiosyncratic risk vector both follow a multivariate Exp-Exp Law. Here, we had to specify two tail-index parameters $\alpha_1, \alpha_2$ for every model (e.g. 1.5 - 20), of which again the first value (i.e. $\alpha_1 = 1.5$) refers to the factor distribution and the second (i.e. $\alpha_2 = 20$) to the distribution of the idiosyncratic risk vector. Once more, the correlation parameter $\rho$ was chosen such that the absolute error in the equity tranche is minimal.

In most of the cases, which we present below, the first tranche can be perfectly matched. Only with the ExpExp 3-3 we had difficulty in attaining the market price of the equity tranche. Similar to the setup with the multivariate $t$-distribution, those models yield the best results in fitting the market data, where the tail-index $\alpha_1$ of the factor is lower than the parameter $\alpha_2$ of the idiosyncratic risk vector. In these cases, the absolute errors are also significantly lower than with the other models, where the relation between $\alpha_1$ and $\alpha_2$ is reversed.

By construction of the model in Section 5.3.2, the tail-index parameters need to be larger
than one. When we fix the parameter $\alpha_1$ for the factor at 1.5, the price of the 3% - 6% is again decreasing towards its market value of 56.04 bps and the prices for the other tranches are increasing with growing parameter $\alpha_2$. Again, as the value of the 3% - 6% is relatively large compared to the more senior tranches, the sum of the absolute errors therefore also decreased with growing $\alpha_2$. Additionally, one can observe that with growing $\alpha_2$ also the level of correlation, which is needed to match the equity tranche, slightly increases.

<table>
<thead>
<tr>
<th></th>
<th>iTraxx</th>
<th>Gaussian</th>
<th>1.5 - 10</th>
<th>1.5 - 20</th>
<th>1.5 - 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>39.67%</td>
<td>33.28%</td>
<td>34.06%</td>
<td>34.23%</td>
<td></td>
</tr>
<tr>
<td>$\rho^2$</td>
<td>15.74%</td>
<td>11.08%</td>
<td>11.60%</td>
<td>11.72%</td>
<td></td>
</tr>
<tr>
<td>0% - 3%</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
<td></td>
</tr>
<tr>
<td>3% - 6%</td>
<td>56.04</td>
<td>120.13</td>
<td>71.45</td>
<td>65.38</td>
<td></td>
</tr>
<tr>
<td>6% - 9%</td>
<td>16.65</td>
<td>21.28</td>
<td>23.02</td>
<td>24.00</td>
<td></td>
</tr>
<tr>
<td>9% - 12%</td>
<td>9.17</td>
<td>5.51</td>
<td>11.94</td>
<td>12.44</td>
<td></td>
</tr>
<tr>
<td>12% - 22%</td>
<td>2.57</td>
<td>0.55</td>
<td>5.72</td>
<td>6.16</td>
<td></td>
</tr>
<tr>
<td>abs. error</td>
<td>74.40</td>
<td>27.69</td>
<td>24.71</td>
<td>24.03</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>5 - 20</th>
<th>3 - 3</th>
<th>5 - 5</th>
<th>5 - 50</th>
<th>40 - 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>38.59%</td>
<td>2.08%</td>
<td>31.40%</td>
<td>38.81%</td>
<td>32.05%</td>
</tr>
<tr>
<td>$\rho^2$</td>
<td>14.89%</td>
<td>0.04%</td>
<td>9.86%</td>
<td>15.06%</td>
<td>10.27%</td>
</tr>
<tr>
<td>0% - 3%</td>
<td>20.04</td>
<td>17.58</td>
<td>20.04</td>
<td>20.04</td>
<td>20.04</td>
</tr>
<tr>
<td>3% - 6%</td>
<td>113.26</td>
<td>144.82</td>
<td>121.77</td>
<td>112.71</td>
<td>123.92</td>
</tr>
<tr>
<td>6% - 9%</td>
<td>23.88</td>
<td>24.19</td>
<td>23.99</td>
<td>24.49</td>
<td>20.62</td>
</tr>
<tr>
<td>9% - 12%</td>
<td>6.97</td>
<td>8.86</td>
<td>4.04</td>
<td>7.21</td>
<td>3.40</td>
</tr>
<tr>
<td>12% - 22%</td>
<td>1.17</td>
<td>1.53</td>
<td>0.32</td>
<td>1.21</td>
<td>0.14</td>
</tr>
<tr>
<td>abs. error</td>
<td>68.05</td>
<td>–</td>
<td>77.83</td>
<td>67.33</td>
<td>80.05</td>
</tr>
</tbody>
</table>

Table 7.7: Various specifications of the model based on the multivariate Exp-Exp Law. The iTraxx Europe Series 3 tranches are from the 6th of April 2006. All prices are denoted in basis points except for the equity tranche values, which are given in percent.

In the above tables we could clearly observe that the models, which are based on elliptical distributions, display a much better fit to the analysed iTraxx data than the Gaussian standard model, as they show much smaller error values. Comparing the newly introduced models with each other, the model based on the Exp-Exp Law even outperforms the one based on the multivariate $t$-distribution. This is again transparent from the resulting error values.

Another way of comparing models with respect to their ability to match the market prices lies in the concepts of the implied and of the base correlations, which we introduced in Section 4.5.2. Once the various models have been calibrated to the equity tranche and once we have calculated the model prices for the other tranches, individually for each tranche we implied the correlation parameters, which are required within the Gaussian
7.6. Application of the models on CDO data

standard model to match these model prices. The results are illustrated in Figure 7.2, where we can see that the newly introduced models explain the market data much better than the Gaussian model. While the latter model produces an implied correlations curve, which is almost flat, the models based on the elliptical distributions are able to reproduce the characteristic correlation smile and the likewise typical base correlation skew, which are induced by the market values.

Figure 7.2: The implied and the base correlations of the various Gaussian mixture specifications within the elliptical distributions framework. The models were calibrated to the equity tranche of the iTraxx Europe series 3 data of the 6th of April 2006.

The previous figures not only illustrate the superiority of the elliptical distributions models based on the $t$-distribution and on the Exp-Exp Law over the Gaussian model, but also give another indication of which of these two mixtures of Normal distributions is better suited for our pricing purposes. We can see for both, for the implied correlations as well as for the base correlations, that while the graphs for the models based on the Exp-Exp Law show a very similar behaviour for the chosen sets of parameters, all of them are closer to the graphs for the market values than those for the multivariate $t$-distribution.
7.6. Application of the models on CDO data

Models based on the Power Law and the Power Log Law

Unlike the other elliptical distributions models, which we have just studied with respect to its applicability to iTraxx market data, the Power Law and the Power Log Law display a behaviour, which seems to be incompatible with the observed data. Below we give the prices of the standardised iTraxx tranches these two laws induce.

Compared to the market values of the iTraxx tranches, we can see that both laws assign too little value to the 0% - 3% tranche, but value the other tranches by far too high. This seems counterintuitive to the observations of Section 6.3. The prices of the equity tranche are decreasing in $\rho$ and the model prices of the equity tranche are still far too low compared to the market quotes, even for $\rho = 0$. Again, in the model specifications (e.g., Pow 1.5-50) the first value denotes the tail index $\alpha_1$ for the factor distribution and the second the tail index $\alpha_2$ for the idiosyncratic risk distribution.

Note that a negative upfront payment as e.g. in the Pow 1.5 - 5 for $\rho = 40\%$ is needed if the periodic spread payments of 500 bps for the equity tranche, which are part of the iTraxx standardisation rules, are too high.

<table>
<thead>
<tr>
<th></th>
<th>$\rho$</th>
<th>0% - 3%</th>
<th>3% - 6%</th>
<th>6% - 9%</th>
<th>9% - 12%</th>
<th>12% - 22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>iTraxx</td>
<td>20.04</td>
<td>56.04</td>
<td>16.65</td>
<td>9.17</td>
<td>2.57</td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>39.67%</td>
<td>20.04</td>
<td>120.13</td>
<td>21.28</td>
<td>5.51</td>
<td>0.55</td>
</tr>
<tr>
<td>Pow 1.5-5</td>
<td>0.00%</td>
<td>0.02</td>
<td>276.51</td>
<td>165.63</td>
<td>85.94</td>
<td>25.12</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>0.00</td>
<td>274.64</td>
<td>165.35</td>
<td>86.54</td>
<td>25.22</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>-0.20</td>
<td>268.40</td>
<td>165.11</td>
<td>86.55</td>
<td>25.80</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>-0.62</td>
<td>258.03</td>
<td>161.49</td>
<td>87.59</td>
<td>26.83</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>-1.26</td>
<td>243.10</td>
<td>154.31</td>
<td>88.91</td>
<td>28.47</td>
</tr>
<tr>
<td>Pow 1.5-50</td>
<td>0.00%</td>
<td>1.48</td>
<td>283.65</td>
<td>164.59</td>
<td>75.60</td>
<td>17.01</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>1.45</td>
<td>281.89</td>
<td>162.05</td>
<td>78.38</td>
<td>17.21</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>1.22</td>
<td>276.33</td>
<td>159.57</td>
<td>80.71</td>
<td>18.01</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>0.75</td>
<td>266.69</td>
<td>155.25</td>
<td>82.43</td>
<td>19.53</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>0.04</td>
<td>253.07</td>
<td>148.67</td>
<td>81.74</td>
<td>22.20</td>
</tr>
<tr>
<td>Pow 5-5</td>
<td>0.00%</td>
<td>0.02</td>
<td>276.51</td>
<td>165.63</td>
<td>85.94</td>
<td>25.12</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>0.05</td>
<td>274.97</td>
<td>164.76</td>
<td>87.47</td>
<td>25.09</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>-0.04</td>
<td>269.91</td>
<td>164.90</td>
<td>88.86</td>
<td>25.71</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>-0.45</td>
<td>259.80</td>
<td>162.05</td>
<td>90.44</td>
<td>27.38</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>-1.23</td>
<td>243.47</td>
<td>153.70</td>
<td>91.94</td>
<td>29.86</td>
</tr>
</tbody>
</table>
Table 7.8: Various specifications of the models based on the Power Law (Pow) and on the Power Log Law (PowLog) with different levels for the correlation parameter $\rho$. The iTraxx Europe Series 3 tranches are from the 6th of April 2006. All prices are denoted in basis points except for the equity tranche values, which are given in percent.

<table>
<thead>
<tr>
<th></th>
<th>$\rho$</th>
<th>0% - 3%</th>
<th>3% - 6%</th>
<th>6% - 9%</th>
<th>9% - 12%</th>
<th>12% - 22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PowLog 1.5-5</td>
<td>0.00%</td>
<td>4.27</td>
<td>256.82</td>
<td>133.01</td>
<td>72.44</td>
<td>15.07</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>4.00</td>
<td>260.62</td>
<td>133.84</td>
<td>72.26</td>
<td>15.28</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>3.52</td>
<td>261.15</td>
<td>134.18</td>
<td>71.80</td>
<td>16.26</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>2.78</td>
<td>256.73</td>
<td>134.40</td>
<td>71.13</td>
<td>18.18</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>1.73</td>
<td>247.43</td>
<td>133.72</td>
<td>70.93</td>
<td>20.89</td>
</tr>
<tr>
<td>PowLog 1.5-50</td>
<td>0.00%</td>
<td>6.46</td>
<td>271.96</td>
<td>122.27</td>
<td>50.62</td>
<td>6.10</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>6.34</td>
<td>269.74</td>
<td>123.14</td>
<td>51.38</td>
<td>6.63</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>5.96</td>
<td>264.56</td>
<td>123.40</td>
<td>52.65</td>
<td>8.13</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>5.22</td>
<td>256.87</td>
<td>122.44</td>
<td>54.35</td>
<td>10.77</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>4.08</td>
<td>246.78</td>
<td>119.90</td>
<td>56.19</td>
<td>14.36</td>
</tr>
<tr>
<td>PowLog 5-5</td>
<td>0.00%</td>
<td>4.27</td>
<td>256.82</td>
<td>133.01</td>
<td>72.44</td>
<td>15.07</td>
</tr>
<tr>
<td></td>
<td>10.00%</td>
<td>3.96</td>
<td>262.27</td>
<td>135.49</td>
<td>72.36</td>
<td>15.06</td>
</tr>
<tr>
<td></td>
<td>20.00%</td>
<td>3.59</td>
<td>264.08</td>
<td>138.25</td>
<td>72.92</td>
<td>15.74</td>
</tr>
<tr>
<td></td>
<td>30.00%</td>
<td>2.95</td>
<td>260.86</td>
<td>139.64</td>
<td>74.05</td>
<td>17.94</td>
</tr>
<tr>
<td></td>
<td>40.00%</td>
<td>1.84</td>
<td>251.21</td>
<td>137.73</td>
<td>75.90</td>
<td>21.68</td>
</tr>
</tbody>
</table>

The correlation levels and the sums of absolute errors over time

Up to now in this section we have focused on one particular date where the elliptical distributions model based on the Exp-Exp Law reproduced the market prices reasonably well. We want to close this chapter with an illustration of how the various models behave over time and in particular how the calibrated correlation parameters change. To this end we used the tranched iTraxx data, as well as the CDS prices that we observed on every second Wednesday between January 2006 and May 2006. The observed values can be found in Table A.1 and in Table A.2.

We then calibrated the Gaussian model, the model based on the multivariate $t$-distribution and the one based on the multivariate Exp-Exp Law to the equity tranche prices, which we observed on these Wednesdays. The so obtained correlation parameters $\rho$ are given in Figure 7.3. Note however, that the correlation value for the MVT 3-20 model on the 3rd of May 2006 is set to zero, even though the model was not able to reproduce the equity tranche price on this day, which was quite low compared to the values observed on other days. One can see that for all the models the correlation parameters change almost in a parallel way over time, and that the required correlation level in the Gaussian model is always above the correlation level of the other models. The ExpExp 1.5-10, the ExpExp 1.5-20, the ExpExp 1.5-50 and the MVT 3-40 model however induce very similar correlation parameters, which are greater than those required for the MVT
3-20 model at any point in time.

Figure 7.3: The correlation parameters $\rho$ over time, which are needed for the various models to perfectly match the equity tranche prices of the iTraxx Europe series 3 data.

Figure 7.4: The sums of absolute errors (in bps) resulting from the previously calibrated models.

Figure 7.4 shows how the calibrated models fit the iTraxx tranche prices over time. After perfectly fitting the equity tranche, we computed the sum of the absolute errors between the model prices and the market prices for each model specification and again for each second Wednesday between January and May 2006. As mentioned for the previous graph, all models could reproduce the equity tranche prices except for the MVT 3-20
model on the 3rd of May 2006 (note that we set the sum of absolute errors to zero in this case). While for this day the error sums for the remaining models are very close to each other, on the other days the Gaussian model displays much larger error values than the elliptical distributions models. Among the models based on the mixtures of Normal distributions of Section 5.3.2, the ExpExp 1.5-20 model and the ExpExp 1.5-50 model present themselves as the models fitting the observed market data the best, and this uniformly over all observation dates.
7.6. Application of the models on CDO data
Chapter 8

Dynamic elliptical distributions model

8.1 Introduction

As we have seen in the previous chapters, for portfolio credit derivatives it is of crucial importance to model the dependence structure within the portfolio correctly. Therefore, most of our focus so far has been laid on exactly this issue. Even though we have introduced the asset vector or the log-return vector as a stochastic process in Section 5.2, we have assumed that the log-return vector process \((S_t^{(n)})_{t \in \mathbb{R}^+}\) is strictly stationary. In order to concentrate on the dependence structure we have subsequently dropped the time index and modelled the asset values at only one unspecified point in time. We have seen in Section 7.6 that this one-period assumption leads to satisfactory results when we consider CDOs with just one maturity date, e.g. in 5 years time. In some sense, this setup levels out the differences that exist during the lifetime of such a CDO.

At the latest since the introduction of the iTraxx tranches, there are CDOs on the same underlying portfolio available where the only difference lies in the varying maturity dates. Even though the iTraxx tranches with a maturity of 5 years usually represent the most liquid of the traded iTraxx tranches, there are also iTraxx tranches with maturities of 7 and of 10 years available on the market. Therefore, when one wants to model CDOs with different maturities consistently within one model at the same time, one might need to relax the assumption of this strict stationarity and thus of constant distributions of the log-returns.

The aim of this chapter is therefore to analyse how the static elliptical distributions framework, that we have presented in the previous chapters, can be embedded into a dynamic setting which allows for changing distributional characteristics. We want to model an explicit temporal dependence in the asset-value process via the factor process \((M_t)_{t \geq 0}\) and the process of idiosyncratic risks \((\varepsilon_t^{(n)})_{t \geq 0}\), whose marginal distributions should remain in the class of elliptical distributions at any point in time. To this end we will
thoroughly discuss multiple possibilities that the various theories of stochastic processes provide us with. In particular, we will make use of discrete-time time series models, of continuous-time short-rate models from the interest rate theory, of time-changed Brownian motions and of subordinated Lévy processes. All the time, the discussed processes will be adapted in such a way that they perfectly fit into the static model of the previous chapters and so that we can use the results obtained therein as the large homogeneous portfolio approximation or the discussion about tail dependence.

8.2 Dynamic setup of elliptical distributions factor model

We work on a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\), where we interpret \(\mathbb{P}\) as a given risk-neutral probability or pricing measure. The filtration \(\mathbb{F} = (\mathcal{F}_t)_{t \in I}\) shall be indexed by a set \(I\) that can either be a subset of \(\mathbb{Z}\) or of \(\mathbb{R}\) (possibly only up to the maturity \(T\) with \(T \in \mathbb{N}\) or \(T \in \mathbb{R}\)). In the continuous-time case with \(I \subseteq \mathbb{R}\) we assume that \(\mathbb{F}\) satisfies the usual conditions of right-continuity and completeness, while in both cases \(\mathcal{F}_0\) shall be the trivial \(\sigma\)-field.

We will require the \(m\)-dimensional risk factor process \((M_t)_{t \in I}\), which represents the systematic component of the log-return vector, as well as the \(n\)-dimensional process \((\varepsilon_t^{(n)})_{t \in I}\), which represents the vector of the idiosyncratic risk components \(\varepsilon_{1,t}, \ldots, \varepsilon_{n,t}\), to be \(\mathbb{F}\)-adapted. Again, the vector \(\varepsilon_t^{(n)}\) is assumed to be independent of \((M_t, \Psi_t)\), for every \(t \in I\). Together, the factor and the idiosyncratic risk shall again constitute the process of log-returns \((S_t^{(n)})_{t \in I}\) via the factor structure as in Equation (6.2).

As we have motivated in the static setup in the previous chapters, it is necessary to assume that the idiosyncratic risk vector \(\varepsilon_t^{(n)}\) follows a mixture of Normal distributions at any point in time \(t \in I\), as soon as we let the portfolio size \(n\) to be arbitrarily large (cf. Section 5.2). We will therefore restrict our search for possible processes, for modelling the idiosyncratic risk vector process \((\varepsilon_t^{(n)})_{t \in I}\) to be of the structure as in Equation (6.3), where at any time \(t \in I\), the vector \(\varepsilon_t^{(n)}\) shall follow a mixture of Normal distributions, that is, there exist a process of non-negative random variables \((R_t)_{t \in I}\) and a process of Gaussian random vectors \((Y_t^{(n)})_{t \in I} = ((Y_{1,t}, \ldots, Y_{n,t})^T)_{t \in I}\) such that for every \(t \in I\)

\[
\varepsilon_t^{(n)} \overset{d}{=} R_t \cdot Y_t^{(n)} \sim \text{EC}_n(0, \Sigma_n, \phi_t), \quad (8.1)
\]

where

\[
\phi_t(x) := \int_0^\infty \exp \left( -\frac{1}{2} r^2 x \right) dG_{2,t}(r), \text{ for } x \in \mathbb{R},
\]

with a distribution function \(G_{2,t}\) on \((0, \infty)\), \(R_t \geq 0\), \(R_t \sim G_{2,t}\), \(E(R_t^2) = -2\phi'_t(0) = 1\), \(Y_t^{(n)} \sim \text{N}_n(0, \Sigma_n)\), \(\Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n}\), and \(\omega_i > 0\).

Even though our static setup and especially the approximation result of Chapter 6 did not oblige the distribution of the factor \(M_t\) to also belong to the class of mixtures of Normal distributions, we want to assume this property to hold in the following. In this
8.3. Effects of the dynamical setup

The main difference between the dynamic setup and the static setup lies in the different assumptions on the marginal distributions of the processes \((S_{j,t})_{t \in I}\), with \(j = 1, \ldots, n\). While in the static or one-period case, the distribution function \(F_{j,t}\) of the log-return \(S_{j,t}\) was assumed to be independent of the time, that is \(S_{j,t} \sim F_{j,t} = F_{j}\), for every time \(t \in I\), the dynamic case allows for variations of the distribution functions \(F_{j,t}\) in \(t \in I\). This has direct consequences for the computation of the rating thresholds \(c_{u_{j,v,t}}\) in Definition 6.1.2, as they are specific quantiles of the distribution functions \(F_{j,t}, j = 1, \ldots, n\) (see Remark 6.1.3 for the one-period case). Therefore, the quantities that are derived from the thresholds will also change, such as the credit ratings \(V_{j,t}\), or the credit loss \(\pi_t(j, u_j, v_j, \psi_t)\). However, the definitions of these quantities remain the same as in Chapter 6, since we have presented them there in a general form which applies for both cases, for distribution functions \(F_{j,t}\) which are constant or for those which vary over time. Likewise, the large homogeneous portfolio approximation of Section 6.2 remains untouched, so that also in the dynamic version of our model we can approximate the overall credit losses \(C_{n,t}\) by the more tractable conditional credit losses \(B_{n,t}\) if the portfolio size is large enough and under some further assumptions on the second moments of the credit losses.

When we want to apply the dynamic setup and consequently also the large homogeneous portfolio approximation for the pricing of CDOs with the structure presented in Section 2.1, we can pursue as in Section 7.4. In this section, we restricted ourselves to the two
rating classes non-default and default, and under the Homogeneous Portfolio Assumptions (7.3) we can approximate the expected value of the relative portfolio losses in a specific tranche \( i \) by the expected value of the conditional counterpart of the relative portfolio losses, which then represents the key quantity for the pricing of CDOs, as we have seen in Section 7.2.

As the large homogeneous portfolio approximation holds for the static case as for the dynamic case, we have chosen to present it in the more general notation in Section 6.2, where we have indexed the appearing quantities with the time parameter \( t \).

Nonetheless, in the static case the distribution functions \( F_{j,t}, G_{1,t}, G_{2,t} \) were actually independent of the time parameter \( t > 0 \), and only the default threshold \( c^1 = c^1_{2,1,t} = F^{-1}_1(p_{2,1,t}) \) introduced some time dependence into the model.

This is different in the dynamic setup. Indeed, the large homogeneous portfolio approximation result within the dynamic model yields that

\[
E(L_{\alpha_i}^\alpha(t)) = E(\Theta_i(L_n(t))) \approx E[\Theta_i(K_n(t))],
\]

but where the last expression now turns into

\[
= \int_0^\infty \int_0^\infty \int_0^\infty \Theta_i(1 - \delta) \cdot \Phi \left( \frac{c^1 - \rho r_1 x}{\sqrt{\omega r}} \right) \varphi(x) \, dx \, dG_{1,t}(r_1) \, dG_{2,t}(r),
\]

with \( c^1 := c^1_{2,1,t} = F^{-1}_1(p_{2,1,t}) \) being the time \( t \) default threshold corresponding to the default probability \( p_{2,1,t} \), that is to the probability of any company migrating from the non-default state 2 to the default state 1 until time \( t \). Note that this default threshold now not only depends on the time via \( p_{2,1,t} \), but also via the distribution function \( F_{1,t} \), and that the scaling distribution functions \( G_{1,t}, G_{2,t} \) in the expected losses in Equation (8.7) can now vary with time (compare with Equation (7.11)).

In fact, the distribution functions \( G_{1,t}, G_{2,t} \) now become the main building blocks of the dynamic model, and will be the center of our attention within our study of introducing stochastic processes with temporal dependence. Specifying the stochastic processes \( (R_{1,t})_{t \in I} \) and \( (R_{t})_{t \in I} \), and therefore also their marginal distribution functions \( G_{1,t} \) and \( G_{2,t} \), for \( t \in I \), fully determines the marginal distribution functions \( F_{j,t}, t \in I, \) of the log-return process \( (S_{j,t})_{t \in I} \), as

\[
F_{j,t}(x) = \int_0^\infty \int_0^\infty \Phi \left( \frac{x}{\sqrt{\rho^2_{j,t} r^2 + s^2 \omega_j}} \right) \, dG_{1,t}(r) \, dG_{2,t}(s),
\]

for \( x \in \mathbb{R} \) and \( t \in I \) (see also Lemma 5.2.2 and Equation (7.6)).

Nonetheless, we have to clarify that one of the main assumptions underlying the one-period models such as the Merton model remains also within this dynamic setup. To understand this, let us denote the deterministic default threshold curve corresponding to company \( j \) by \( (c^1_j)_{s \in I} \), which has to be calibrated to single-name CDS spreads as
8.3. Effects of the dynamical setup

we have outlined in the application of our model on the pricing of CDOs in Chapter 7. While the Merton model [Mer74] assumed that the default of a firm can only happen at one specific point in time, namely the maturity date $T$ of the credit-risky bond, Black & Cox [BC76] relaxed this assumption to allow the default to happen during the entire period $[0, T]$. They used the concept of first hitting times, where a default is triggered as soon as the asset-value process, which is modelled by a geometric Brownian motion, hits the default barrier the first time. When we transfer this concept to the log-return process $(S_{j,s})_{s \in I}$ and the default threshold curve $(c_j^s)_{s \in I}$, the time of default is then given via

$$\tau_j := \inf \{ s \geq 0 : S_{j,s} \leq c_j^s \}.$$

However, empirical studies haven shown that even for single-name credit derivatives the geometric Brownian motion framework of Merton [Mer74], as well as of Black & Cox [BC76], produces unrealistic prices. Unfortunately, as soon as one leaves the world of geometric Brownian motions in order to produce better fits for single-name credit derivatives or if one wants to incorporate a more realistic dependence structure between the underlying asset processes for the modelling of credit portfolios, it is mathematically very difficult to give analytical expressions for the first-hitting-time distributions. Especially for CDOs, where the accurate modelling of the dependence structure is seen to have a higher priority than the dynamics, one therefore usually resorts to the one-period framework, where one does not need the first-hitting-time distributions. This road was followed by the prominent CDO pricing models which we have reviewed in Section 4.4 and also by our one-period setup of Chapter 5. Even though our dynamic approach relaxes the condition that the log-return distributions remain unchanged over time, one still needs to assume that for a sufficiently smooth Borel-measurable function $f : \mathbb{R} \mapsto \mathbb{R}$ the following set of equations is exact:

$$\mathbb{E} \left( f(1_{\{\tau_j \leq t\}}) \right) = \mathbb{E} \left( f(1 - 1_{\{\tau_j > t\}}) \right) = \mathbb{E} \left( f \left( 1 - 1_{\{S_{j,s} > c_j^s, \forall s \in [0,t]\}} \right) \right) = \mathbb{E} \left( f \left( 1 - 1_{\{S_{j,t} > c_j^t\}} \right) \right) = \mathbb{E} \left( f \left( 1_{\{S_{j,t} \leq c_j^t\}} \right) \right).$$

In general, one would expect that the step from the first to the second line is only an approximation. The assumption of the validity of this set of equations is for example needed for the specific choice of $f(x) := x$, for $x \in \mathbb{R}$, where we obtain

$$\mathbb{P}(\tau_j \leq t) = 1 - \mathbb{P}(\tau_j > t) = 1 - \mathbb{P} \left( S_{j,t} > c_j^t \right) = F_{j,t}(c_j^t),$$

for $t \in I$. The importance of the validity of this very assumption also applies to the expected losses on a tranche, as only in this case the expected losses in Equation (8.6) are truly equal to the expression in Equation (8.7).

In the following we want to analyse two main directions for the specification of the dynamics, one where $I \subset \mathbb{Z}$ and the other where $I \subset \mathbb{R}$. The main requirement for the specification is the ability to evaluate the distribution functions $G_{1,t}$ and $G_{2,t}$ of the factor and the idiosyncratic component at any time $t \in I$, as these are the crucial building blocks for the computation of the expected losses in each tranche.
Most of the models that we will discuss in the next sections can be used for both the one-dimensional factor process and the \( n \)-dimensional idiosyncratic risk vector process. However, we will also make use of volatility models where only a one-dimensional specification fits into the desired framework, so that this structure can only be used on the factor process. In general, we will consider processes \((X_t)_{t \in I}\) with marginal mixtures of Normal distributions of the form
\[
X_t \overset{d}{=} R_t N_t, \quad \text{for every } t \in I,
\]
with \(R_t \geq 0\), \(R_t\) and \(N_t\) independent, where either \(N_t \sim \mathcal{N}(0, 1)\) or \(N_t \sim \mathcal{N}(0, \Sigma_n)\), \(\Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n)\).

The distribution function of \(R_t\) will be denoted by \(G_t\) for \(t \in I\).

In most of the setups we will not need to specify the temporal dependence of the process \((N_t)_{t \in I}\), as it usually does not have any impact on the quantities such as the expected losses.

### 8.4 The discrete-time case

In this section we want to understand the index set \(I\) as being the natural numbers including zero, that is \(I = \mathbb{N} = \{0, 1, 2, \ldots\}\), or the whole numbers, that is \(I = \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}\).

If we want to concentrate on the modelling of the dynamics of a one-dimensional \((X_t)_{t \in \mathbb{N}}\) with \(X_t \overset{d}{=} R_t N_t \sim \mathcal{E}C_1(0, 1, \phi_t) = S_1(\phi_t)\), for every \(t \in \mathbb{N}\), which can be used for the factor process \((M_t)_{t \in I}\), there are different routes we can pursue. We could either model the process \((R_t)_{t \in \mathbb{N}}\) via a time series specification that yields a univariate non-negative process which does not take any values of \((N_t)_{t \in \mathbb{N}}\) into account, or we could directly apply the theory of volatility models such as the ARCH and the GARCH models. In the second case, the process \((N_t)_{t \in \mathbb{N}}\) must be a one-dimensional white noise process, as a consequence of which this framework is restricted to the modelling of the factor process. However, in the first case the modelling of the process \((R_t)_{t \in \mathbb{N}}\) is entirely independent of the modelling of the process \((N_t)_{t \in \mathbb{N}}\) so that this setup can be used for the factor process, as well as for the idiosyncratic risk vector process. We will begin with this separate modelling of \((R_t)_{t \in \mathbb{N}}\).

But before we start, we first need some standard notions from the theory of time-series models. A more detailed discussion of the concepts used here can be found in the monographs by Brockwell & Davis [BD91], Hamilton [Ham94] or Kreiss & Neuhaus [KN06].

**Definition 8.4.1 (The autocovariance function)**

If \((Y_t)_{t \in I}\) is a process such that \(\text{Var}(Y_t) < \infty\) for all \(t\) in the index set \(I\), then the autocovariance function \(\gamma_Y(\cdot, \cdot)\) of \((Y_t)_{t \in I}\) is defined by
\[
\gamma_Y(r, s) = \text{Cov}(Y_r, Y_s), \quad \text{for } r, s \in I.
\]

**Definition 8.4.2 ((Strictly) stationary processes)**

The time series \((Y_t)_{t \in \mathbb{Z}}\) is said to be stationary if
8.4. The discrete-time case

1. $\mathbb{E}(Y_t^2) < \infty$ for all $t \in \mathbb{Z}$,
2. $\mathbb{E}(Y_t) = m$ for all $t \in \mathbb{Z}$, and
3. $\gamma_Y(r, s) = \gamma_Y(r + t, s + t)$ for all $r, s, t \in \mathbb{Z}$,

and $(Y_t)_{t \in \mathbb{Z}}$ is said to be strictly stationary if the joint distributions of $(Y_{t_1}, \ldots, Y_{t_k})$ and of $(Y_{t_1+h}, \ldots, Y_{t_k+h})$ coincide for all positive integers $k \in \mathbb{N}_0$ and for all $t_1, \ldots, t_k, h \in \mathbb{Z}$.

8.4.1 Employing time-series models for the scaling process

In the following, we want to discuss different possibilities of modelling the scaling process $R = (R_t)_{t \in \mathbb{N}}$ with discrete-time processes. The range of the models we will discuss in the following will start with the very simple white noise processes, will make use of moving average processes and will lead to autoregressive models.

White noise processes for the scaling process

Definition 8.4.3 (White noise processes)
A process $(Y_t)_{t \in J_q}$, with $q \in \mathbb{N} \cup \{\infty\}$, $J_q = \{j \in \mathbb{Z} : j \geq -q\}$ being a subset of the whole numbers and with $J_\infty = \mathbb{Z}$ being the entire whole numbers, is called white noise process if its elements have mean zero, constant variance $\sigma^2$ and are uncorrelated across time, that is $\mathbb{E}(Y_t Y_s) = 0$ for $t \neq s$. The white noise process is called independent if its elements are even independent across time.

Moving average processes for the scaling process

Definition 8.4.4 (MA processes)
A process $(Y_t)_{t \in \mathbb{N}}$ is called $q$-th order moving average process, denoted by $MA(q)$, with $q \in \mathbb{N}$, if there are a real number $\mu \in \mathbb{R}$, a white noise process $(Z_t)_{t \in J_q}$ and real (deterministic) numbers $\alpha_0, \ldots, \alpha_q$ such that

$$Y_t = \mu + \sum_{i=0}^{q} \alpha_i Z_{t-i},$$

for $t \in \mathbb{N}$.

We now want to analyse under which additional assumptions moving average processes can be applied in our setup. Either we could use such processes essentially directly for the scaling process and choose the appearing constants and the white noise process...
appropriately, or we can make use of a transformation of a moving average process for the scaling process.

The first option can be implemented as follows. Assume a scaled non-negative MA($q$) process for $(R_t)_{t \in \mathbb{N}}$, that is, let

$$R_t = \gamma \left( \mu + \sum_{i=0}^{q} \alpha_i Z_{t-i} \right), \quad \text{for } t \in \mathbb{N},$$

where $(Z_t)_{t \geq -q}$ is a white noise process with $Z_t \geq -K$, for all $t$, and with $K \in \mathbb{R}^+$ an arbitrary non-negative constant that allows for a $\mu \geq 0$ (e.g. $\mu = \sum_{i=0}^{q} \alpha_i K$) such that $R_t \geq 0$, for all $t$, where $\alpha_0, \ldots, \alpha_q \geq 0$ are non-negative real coefficients and $\gamma := [\mu^2 + \text{Var}(Z_1) \sum_{i=0}^{q} \alpha_i^2]^{-1/2}$, assumed to be finite. Then

$$\mathbb{E}(R_t) = \gamma \mu, \quad \mathbb{E}(R_t^2) = 1 \quad \text{and} \quad \text{Var}(R_t) = \gamma^2 \text{Var}(Z_1) \sum_{i=0}^{q} \alpha_i^2,$$

for every $t \in \mathbb{N}$. In general, the process $(R_t)_{t \in \mathbb{N}}$ is only stationary, but becomes strictly stationary, if we further assume that the $(Z_t)_{t \geq -q}$ are independent and identically distributed. The independence of the $(Z_t)_{t \geq -q}$ guarantees that we obtain the distribution function $G_t$ of $R_t$ via the computation of a $(q + 1)$-dimensional convolution, and the iid-property yields that the distribution $G_t$ does not depend on $t$.

**Remark 8.4.5** Note that in general the process $(R_t)_{t \in \mathbb{N}}$ is not necessarily strictly stationary if $(Z_t)_{t \geq -q}$ is only a white noise process with uncorrelated components, even if we assume that they are identically distributed:

*let two independent processes $(B_t)_{t \in \mathbb{Z}}$ and $(N_t)_{t \in \mathbb{Z}}$ be given, where $N_t \sim N(0,1)$ and $\mathbb{P}(B_t = 0) = \mathbb{P}(B_t = 1) = 0.5$ for all $t$, and assume that all $B_t$ and $N_t$, $t \in \mathbb{Z}$, are independent. Define a process $(Z_t)_{t \in \mathbb{Z}}$ by

$$Z_t := B_t \cdot N_t, \quad \text{for } t \in \mathbb{Z} \setminus \{1\}, \quad \text{and} \quad Z_1 := B_0 \cdot N_1.$$

Then the elements of $(Z_t)_{t \in \mathbb{Z} \setminus \{1\}}$ are independent and identically distributed, and so are the elements of $(Z_t)_{t \in \mathbb{Z} \setminus \{0\}}$, but $Z_0 = B_0 N_0$ and $Z_1 = B_0 N_1$ are only uncorrelated:

$$\text{Cov}(Z_0, Z_1) = \mathbb{E}(Z_0 Z_1) = \mathbb{E}(B_0^2 N_0 N_1) = 0,$$

yet

$$\mathbb{P}(B_0 N_0 \geq 0, B_0 N_1 \geq 0) = \mathbb{P}(N_0 \geq 0, N_1 \geq 0, B_0 = 1) + \mathbb{P}(B_0 = 0) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8},$$

while

$$\mathbb{P}(B_0 N_0 \geq 0) = \mathbb{P}(B_0 N_1 \geq 0) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

and $\frac{5}{8} \neq \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$.

However, $(Z_t)_{t \in \mathbb{Z}}$ is a white noise process which we want to use to define an MA process $(R_t)_{t \in \mathbb{N}}$ as above. Suppose that $q = 1$, $\alpha_0 = \alpha_1 = 1$, then

$$R_0 = \gamma (Z_0 + Z_{-1}) \quad \text{and} \quad R_1 = \gamma (Z_1 + Z_0),$$
for which we obtain that

\[
\mathbb{P}(R_0 \geq 0) = \mathbb{P}(B_0N_0 + B_{-1}N_{-1} \geq 0) \\
= \mathbb{P}(N_0 + N_{-1} \geq 0) \mathbb{P}(B_0 = B_{-1} = 1) + \mathbb{P}(B_0 = B_{-1} = 0) \\
+ \mathbb{P}(N_{-1} \geq 0) \mathbb{P}(B_0 = 0, B_{-1} = 1) + \mathbb{P}(N_0 \geq 0) \mathbb{P}(B_0 = 1, B_{-1} = 0) \\
= \frac{1}{8} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = \frac{5}{8}
\]

which is not equal to

\[
\mathbb{P}(R_1 \geq 0) = \mathbb{P}(B_0N_1 + B_{-1}N_{0} \geq 0) = \mathbb{P}(N_1 + N_{0} \geq 0) \mathbb{P}(B_0 = 1) + \mathbb{P}(B_0 = 0) \\
= \frac{1}{4} + \frac{1}{2} = \frac{3}{4}.
\]

Therefore, \(R_0\) and \(R_1\) do not have the same distribution, and \((R_t)_{t \in \mathbb{N}}\) is only a stationary, but not a strictly stationary process.

**Example 8.4.6** The simplest case for an \(MA(q)\) process is obtained, when \(\alpha_0 = \ldots = \alpha_q\). Then, if the \((Z_t)_{t \geq -q}\) are e.g. independent \(\Gamma(\delta, \lambda)\) random variables, we have that

\[
\frac{R_t}{\gamma_0} = \sum_{i=0}^{q} Z_{t-i} \sim \Gamma((q+1)\delta, \lambda), \quad t \in \mathbb{N},
\]

from which the distribution function \(G_t\) of \(R_t\) can be directly computed.

Instead of modelling the process \((R_t)_{t \in \mathbb{N}}\) directly, we could also opt to model a transform of the process \((R_t)_{t \in \mathbb{N}}\) as a not necessarily non-negative \(MA(q)\) process such that we obtain positive values for \((R_t)_{t \in \mathbb{N}}\) when taking the inverse of this transform. A natural choice is to consider \((\log(R_t))_{t \in \mathbb{N}}\). Let

\[
R_t = \gamma_t \left( \mu + \exp \left( \sum_{i=0}^{q} \alpha_i Z_{t-i} \right) \right), \quad \text{for } t \in \mathbb{N},
\]

with \(\mu \geq 0\), with coefficients \(\alpha_0, \ldots, \alpha_q\) in \(\mathbb{R}\), with an independent white noise process \((Z_t)_{t \geq -q}\) where the moment generating functions \(M_t(\cdot)\) of its elements are assumed to exist in all points \(\alpha_i, \ 2\alpha_i\), for \(i = 0, \ldots, q\), and with a deterministic process \(\gamma_t := [\mu^2 + 2\mu \prod_{i=0}^{q} M_{t-i}(\alpha_i) + \prod_{i=0}^{q} M_{t-i}(2\alpha_i)]^{-1/2}\). We obtain

\[
\mathbb{E}(R_t) = \gamma_t \left( \mu + \prod_{i=0}^{q} M_{t-i}(\alpha_i) \right), \quad \mathbb{E}(R_t^2) = 1
\]

and

\[
\mathbb{V}ar(R_t) = \gamma_t^2 \prod_{i=0}^{q} (M_{t-i}(2\alpha_i) - M_{t-i}(\alpha_i)),
\]

for every \(t \in \mathbb{N}\). In general, \((R_t)_{t \in \mathbb{N}}\) is not even stationary, but we can again obtain the distribution function \(G_t\) of every \(R_t\), \(t \in \mathbb{N}\), via the computation of a \((q+1)\)-dimensional convolution followed by a transformation of the convolution due to the performed mapping \(x \mapsto \gamma_t(\mu + \exp(x))\). If we additionally assume that the terms in
8.4. The discrete-time case

\[(Z_t)_{t \geq -q}\) are also identically distributed, then \((R_t)_{t \in \mathbb{N}}\) becomes even strictly stationary and \(\gamma_t = \gamma\) no longer depends on \(t\).

Briefly summarizing the above, moving average processes in the above two setups involve \(q + 1\) coefficients and additional parameters for the distribution functions of the \(Z_t\)'s. Under the assumption that we are using an independent white noise process, we can directly compute the distribution functions \(G_t\) of \(R_t\) at any time \(t\) via a \((q + 1)\)-dimensional convolution procedure. Yet, the way one employs an MA process for the non-negative scaling process has a crucial effect on whether or not one obtains a stationary process \((R_t)_{t \in \mathbb{N}}\). In any case, if we assume that the applied independent white noise process has identically distributed components, then even the strict stationarity of \((R_t)_{t \in \mathbb{N}}\) is satisfied and one falls back to the static one-period case again.

**Autoregressive moving average processes for the scaling process**

**Definition 8.4.7 (ARMA processes)**

Let a white noise process \((Z_t)_{t \in \mathbb{Z}}\) and coefficients \(\alpha_1, \ldots, \alpha_q \in \mathbb{R}\) and \(\delta_1, \ldots, \delta_p \in \mathbb{R}\) be given where \(p, q \in \mathbb{N}\) are fixed integers and \(\alpha_q \delta_p \neq 0\).

1. A process \((Y_t)_{t \in \mathbb{Z}}\) is called an autoregressive moving average process of order \((p, q)\), denoted by \(\text{ARMA}(p, q)\), if it represents a solution to the difference equation

   \[Y_t = c + \sum_{i=1}^{p} \delta_i Y_{t-i} + Z_t + \sum_{j=1}^{q} \alpha_j Z_{t-j}, \text{ for } t \in \mathbb{Z},\]  

   with a \(c \in \mathbb{R}\).

2. With the above given coefficients \(\alpha_1, \ldots, \alpha_q\) and \(\delta_1, \ldots, \delta_p\) we define the following two polynomials, the so-called corresponding \(z\)-transformations:

   \[A(z) := \sum_{i=0}^{p} d_i z^i \text{ and } B(z) := \sum_{j=0}^{q} \alpha_j z^j, \text{ for } z \in \mathbb{C},\]

   where

   \[d_i := \begin{cases} 1, & \text{for } i = 0; \\ -\delta_i, & \text{for } i = 1, \ldots, p, \end{cases}\]

   and \(\alpha_0 := 1\).

If the polynomials \(A\) and \(B\) have common zeros, then there exist polynomials \(A_0\), \(B_0\) and \(H\) where \(A_0\) and \(B_0\) are without common zeros, where \(A_0(0) = B_0(0) = 1\) and where \(A = A_0 \cdot H\) and \(B = B_0 \cdot H\). If \(A\) and \(B\) have no common zeros, we set \(H \equiv 1\).

With the previous definition and the introduction of the backward shift operator \(L\) via \(L^i X_t = X_{t-i}\) for any process \((X_t)\), any \(t \in \mathbb{Z}\) and an arbitrary \(i \in \mathbb{Z}\) (in particular,
Theorem 8.4.8 Let coefficients $\alpha_0, \ldots, \alpha_q \in \mathbb{R}$ and $\delta_1, \ldots, \delta_p \in \mathbb{R}$ be given where $p, q \in \mathbb{N}$ are fixed integers and $\alpha_q \delta_p \neq 0$.

1. For every white noise process $(Z_t)_{t \in \mathbb{Z}}$ there exists at least one stationary ARMA($p, q$) process $(Y_t)_{t \in \mathbb{Z}}$ that satisfies the Equation (8.8) if the reduced $z$-transformation $A_0$ has no zeros $z \in \mathbb{C}$ with $|z| = 1$. In this case, if we additionally have that $H = 1$, that is that the $z$-transformations $A$ and $B$ have no common zeros, then the stationary ARMA($p, q$) process $(Y_t)_{t \in \mathbb{Z}}$ is even unique.

2. Let a white noise process $(Z_t)_{t \in \mathbb{Z}}$ and a corresponding stationary ARMA($p, q$) process $(Y_t)_{t \in \mathbb{Z}}$ satisfying the Equation (8.8) be given and assume that $H = 1$. Then there exist unique real constants $(\psi_j)_{j \in \mathbb{Z}}$ and a $\mu \in \mathbb{R}$ such that $(Y_t)_{t \in \mathbb{Z}}$ can be expressed as

$$Y_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \text{ for } t \in \mathbb{Z}, \quad \text{and } \sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$ 

We have $\mu = \frac{c}{A(z)}$ and the coefficients $(\psi_j)_{j \in \mathbb{Z}}$ can be obtained from the Laurent expansion of $\frac{B(z)}{A(z)}$ around the center $z_0 = 0$ and on the annulus $\{z \in \mathbb{C} : \frac{1}{p} \leq |z| \leq \rho\}$ for an appropriately chosen $\rho > 0$.

3. Let a white noise process $(Z_t)_{t \in \mathbb{Z}}$ and a corresponding ARMA($p, q$) process $(Y_t)_{t \in \mathbb{Z}}$ satisfying the Equation (8.8) be given and assume that $H = 1$. There exist unique real constants $(\psi_j)_{j \in \mathbb{Z}}$ and a $\mu \in \mathbb{R}$ such that $(Y_t)_{t \in \mathbb{Z}}$ can be expressed as

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \text{ for } t \in \mathbb{Z}, \quad \text{and } \sum_{j=0}^{\infty} |\psi_j| < \infty,$$

iff $A(z) \neq 0$ in $\{z \in \mathbb{C} : |z| \leq 1\}$. The ARMA($p, q$) process $(Y_t)_{t \in \mathbb{Z}}$ is then called causal and it is also stationary, the coefficients $(\psi_j)_{j \in \mathbb{N}}$ are the coefficients of the Taylor expansion of $\frac{B(z)}{A(z)}$ around the center $z_0 = 0$, and again we have $\mu = \frac{c}{A(z)}$.

Proof: This theorem is essentially based on Theorem 7.4, Theorem 7.7 and Theorem 7.10 in [KN06].

Remark 8.4.9 Due to the previous theorem a causal, stationary autoregressive moving average process can thus be seen as a special case of a linear process:
8.4. The discrete-time case

A linear process is a process \( (Y_t)_{t \in \mathbb{Z}} \) with \( Y_t = \mu + \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \) for \( t \in \mathbb{Z}, \) where \( (Z_t)_{t \in \mathbb{Z}} \) is a stochastic process with \( \sup_{t \in \mathbb{Z}} \mathbb{E}(|Z_t|) < \infty, \) where \( \mu \in \mathbb{R} \) and where \( (\psi_j)_{j \in \mathbb{N}} \) are real coefficients with \( \sum_{j=0}^{\infty} |\psi_j| < \infty. \) Under these conditions, the series \( \sum_{j=0}^{\infty} \psi_j Z_{t-j} \) is almost surely absolutely convergent, as well as convergent in \( L_1. \)

In order to guarantee that the scaling process \( (R_t)_{t \in \mathbb{N}} \) stays non-negative, we could use such an autoregressive moving average process in combination with the exponential function: Let \( R_t := \gamma (\mu + \exp(Y_t)), \) for \( t \in \mathbb{N}, \) with \( \mu \in \mathbb{R}^+ \) and

\[
Y_t = \sum_{i=1}^{p} \delta_i Y_{t-i} + Z_t + \sum_{j=1}^{q} \alpha_j Z_{t-j}, \text{ for } t \in \mathbb{Z},
\]

where \( (Z_t)_{t \in \mathbb{Z}} \) is assumed to be an independent white noise process that additionally has identically distributed components, where \( \alpha_1, \ldots, \alpha_q \) and \( \delta_1, \ldots, \delta_p \) are some real coefficients and where \( \gamma \in \mathbb{R} \) is a not yet specified constant. If the conditions in the third part of the previous theorem are satisfied, then \( R_t \) can be represented in the form

\[
R_t = \gamma \left( \mu + \exp \left( \sum_{j=0}^{\infty} \psi_j Z_{t-j} \right) \right), \text{ for } t \in \mathbb{N},
\]

with coefficients \( (\psi_j)_{j \in \mathbb{N}} \) as above. Then \( (Y_t)_{t \in \mathbb{Z}} \) and thus also \( (R_t)_{t \in \mathbb{N}} \) are strictly stationary. If we assume that the Laplace transform of \( Z_1 \) is such that \( \mathbb{E}(R_t^2/\gamma^2) \) exists, then we can choose \( \gamma \) such that \( \mathbb{E}(R_t^2) = 1 \) for all \( t \in \mathbb{N}. \) In this case, the process \( (R_t)_{t \in \mathbb{N}} \) fits into our scaling process framework.

As in the moving average case, the autoregressive moving average process setup combined with an \( \exp \)-transform entails the strict stationarity of \( (R_t)_{t \in \mathbb{N}} \) once the white noise process is chosen appropriately. However, a simple white noise process \( (Z_t)_{t \in \mathbb{Z}} \) without any further restrictions yields a non-stationary scaling process in general. When working with an independent white noise process, Theorem 8.4.8 entails that we can compute the distribution \( G_t \) of \( R_t \) via an infinite dimensional convolution for any interesting point in time \( t. \) The workload is slightly reduced, once \( G_t \) does not depend on \( t \) anymore in the case of strict stationarity of \( (R_t)_{t \in \mathbb{N}}, \) that is when the \( Z_t, t \in \mathbb{Z} \) are iid random variables. Yet, this certainly still is a tedious task to do, as one cannot expect to obtain a closed-form analytical expression for this infinite convolution.

Instead of the transformation \( t \mapsto \gamma (\mu + \exp(z)) \) one could employ \( t \mapsto \gamma (\mu + z^2) \) (where the Cauchy product formula for infinite series yields a method of how to set the \( \gamma \) in order to obtain \( \mathbb{E}(R_t^2) = 1 \)) or \( z \mapsto \gamma (\mu + |z|) \) as transformations on the possibly negative ARMA process \( (Z_t)_{t \in \mathbb{Z}}. \) These transformations would then assure that the processes \( (R_t)_{t \in \mathbb{N}} \) so obtained again become non-negative processes. Another approach could involve only considering a non-negative white noise process, and only \( z \)-transforms \( A \) and \( B \) such that the coefficients \( (\psi_j)_{j \in \mathbb{N}} \) also remain non-negative. Even though these approaches then yield simpler conditions on the white noise process and the coefficients \( \alpha_1, \ldots, \alpha_q \) and \( \delta_1, \ldots, \delta_p \) for assuring \( R_t \in L_2, \) we still have to deal with the
infinite series in order to compute the distribution function \( G_t \) of \( R_t \), for \( t \in \mathbb{N} \). All of the standard time series models just discussed yield a strictly stationary scaling process \((R_t)_{t \in \mathbb{N}}\) if the underlying white noise process \((Z_t)_{t \in I}\) has iid-components. Computational tractability could make it necessary to assume that the components are independent, but not necessarily identically distributed. Yet, in the case of non-identical distributions, a clear structure of how the distribution functions of the \( Z_t \) change has to be put in place in order to be able to compute the relevant distribution functions \( G_t \) of \( R_t \) for any given time \( t \). This would involve arbitrary choices, as there are essentially only 5 data points at three different points in time (the 5 quotes for the different tranches of the 3-, 5- and the 10-year tranched iTraxx) onto which to base our temporal behaviour.

However, the assumption of iid-components and thus of a strict stationarity would yield that the distribution functions of the process \((X_t)_{t \in \mathbb{N}}\) at any time \( t \) no longer depend on time, even though we have imposed a temporal dependence structure. This would mean that for the computation of the expression for the expected losses in Equation (8.7) we would not really have left the static case in Chapter 7 (cf. Equation (7.11)). In fact, we would even have reduced the number of choices of distribution functions \( G_t \) as \( G \) now can only be the distribution function of an \( R_t \) produced via a time series model.

**Remark 8.4.10** Employing a white noise process \((Z_t)_{t \in I}\) with iid-components can still yield a non-stationary process, if we for example consider ARIMA processes: If \( d \in \mathbb{N} \) is a non-negative integer, then \((Y_t)_{t \in I}\) is called an autoregressive-integrated moving average process of order \((p,d,q)\), denoted by ARIMA\((p,d,q)\), if \((X_t)_{t \in I}\) given by \( X_t := (1-L)^d Y_t \), for \( t \in I \), is a stationary, causal ARMA\((p,q)\) process. This means that \((Y_t)_{t \in I}\) must satisfy \( A(L)(1-L)^d Y_t = B(L)Z_t \), where \((Z_t)_{t \in I}\) is a white noise process and \( A, B \) are polynomials of degree \( p \) and \( q \) respectively, and where \( A(z) \neq 0 \) for all \( z \) with \( |z| \leq 1 \). The ARIMA process \((Y_t)_{t \in I}\) is then stationary iff \( d = 0 \). Yet, if \( d \geq 1 \), the first and second moments \( \mathbb{E}(Y_t) \) and \( \mathbb{E}(Y_t^2) \) are not determined by the above difference equation \( A(L)(1-L)^d Y_t = B(L)Z_t \). This entails that we would need further assumptions on the structure of the process to continue with our analysis (which we do not pursue further here).

### 8.4.2 Applying discrete-time volatility models

In the following we want to discuss different possibilities of how to bring discrete-time volatility models such as ARCH and GARCH processes into play in our elliptical distributions framework. The volatility models will be used for the product process \((X_t)_{t \in \mathbb{N}} = (R_t N_t)_{t \in \mathbb{N}}\) and possess the property that at any time \( t \) the past of the process \((N_t)_{t \in \mathbb{N}}\), i.e. some of the values \( N_0, N_1, \ldots, N_{t-2}, N_{t-1} \), has an effect on the present value of \((R_t)_{t \in \mathbb{N}}\), that is on \( R_t \). Note that the process \((N_t)_{t \in \mathbb{N}}\) needs to be one-dimensional, so that this setup can only be used for the factor process. Additionally, in order to fit into our setup, in all the models we need to assume that \((N_t)_{t \in \mathbb{N}}\) is an iid sequence of standard Gaussian random variables, \( N_t \sim N(0,1) \), for all \( t \in \mathbb{N} \), and at any fixed time \( t \), the two variables \( R_t \) and \( N_t \) remain independent.
8.4. The discrete-time case

Autoregressive conditional heteroscedasticity processes

We will start with the simplest volatility model, the so-called ARCH model, which was introduced by Engle [Eng82] and which takes the following form:

**Definition 8.4.11 (ARCH processes)**

Let a white noise process \((Z_t)_{t \in \mathbb{Z}}\), a vector of parameters \(\alpha \in \mathbb{R}^k\) and a suitable Borel function \(h : \mathbb{R}^{p+k} \rightarrow \mathbb{R}_0^+\) be given where \(p, k \in \mathbb{N}\) are given integers. A process \((Y_t)_{t \in \mathbb{Z}}\) is called an autoregressive conditional heteroscedasticity process of order \(p\), denoted by \(ARCH(p)\), if it satisfies

\[
Y_t = h_t Z_t, \quad \text{where} \quad h^2_t = h(Y_{t-1}, \ldots, Y_{t-p}, \alpha)
\]

for \(t \in \mathbb{Z}\).

In the theory of ARCH processes, one often assumes that \(h\) is of the form

\[
h(y_1, \ldots, y_p, \alpha) = \alpha_0 + \sum_{i=1}^{p} \alpha_i y^2_i
\]

where \(\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_p) \in \mathbb{R}_0^+ \times (\mathbb{R}^+)^p\) is a \((p+1)\)-dimensional vector with non-negative coefficients \(\alpha_i \geq 0, \ i = 1, \ldots, p, \ \text{and} \ \alpha_0 > 0\).

**Theorem 8.4.12** Let an independent white noise process \((Z_t)_{t \in \mathbb{Z}}\) and a vector of parameters \(\alpha \in \mathbb{R}_0^+ \times (\mathbb{R}^+)^p\) be given where \(p \in \mathbb{N}\). The corresponding ARCH\((p)\) process \((Y_t)_{t \in \mathbb{Z}}\) with

\[
Y_t = h_t Z_t \quad \text{and} \quad h^2_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y^2_{t-i},
\]

for \(t \in \mathbb{Z}\), is stationary iff \(A(z) \neq 0\) for all \(z\) with \(|z| \leq 1\) where \(A(z) := 1 - \sum_{i=1}^{p} \alpha_i z^i\) for \(z \in \mathbb{C}\). The variance is then given by \(\text{Var}(Y_t) = \mathbb{E}(Y^2_t) = \mathbb{E}(h^2_t) = \alpha_0/(1 - \sum_{i=1}^{p} \alpha_i)\).

**Proof:** For a proof see e.g. [Eng82].

The process \((X_t)_{t \in \mathbb{N}}\) that satisfies

\[
X_t = R_t N_t \quad \text{and} \quad R^2_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i X^2_{t-i},
\]

where \(N_t \sim N(0, 1)\) are iid Normal for all \(t\), thus follows an ARCH\((p)\) model with an independent Gaussian white noise process \((N_t)_{t \in \mathbb{Z}}\). The condition \(\alpha_0 + \sum_{i=1}^{p} \alpha_i = 1\) is necessary for the desired property \(\mathbb{E}(R^2_t) = 1\), and it is also sufficient for the condition \(A(z) \neq 0\) for all \(z\) with \(|z| \leq 1\), and thus by virtue of Theorem 8.4.12 for the stationarity of \((X_t)_{t \in \mathbb{N}}\).
With the above settings, the ARCH process \((X_t)_{t \in \mathbb{N}}\) becomes a special white noise process, as \(\mathbb{E}(X_t) = 0\), \(\mathbb{E}(X_t^2) = 1\) and as, due to the independence of \(N_t\) of its past \(N_{t-1}, N_{t-2}, \ldots\), all the autocovariances are zero:

\[
\mathbb{E}(X_tX_{t-h}) = \mathbb{E}(N_t)\mathbb{E}(R_tN_{t-h}R_{t-h}) = 0,
\]

for \(h \in \mathbb{N}_0\) and \(t \in \mathbb{N}\). However, the process \((R_t^2)_{t \in \mathbb{N}}\) is certainly not a special MA\((p)\) process as described in Definition 8.4.4, as the white noise process \((X_t)_{t \in \mathbb{N}}\) does not enter linearly into the specification for \((R_t^2)_{t \in \mathbb{N}}\).

At any fixed time \(t\), \(R_t\) does not depend on \(X_t = R_tN_t\) and specifically not on \(N_t\). In fact, \(R_t\) and \(N_t\) need to be independent for our setup, at any time \(t\). This entails that in the first place we cannot use a general white noise process \((N_t)_{t \in \mathbb{Z}}\), but indeed have to assume the process \((N_t)_{t \in \mathbb{Z}}\) to be an independent white noise process. Yet, this does not imply that also \((X_t)_{t \in \mathbb{N}}\) becomes an independent white noise process in general. Thus, it becomes difficult to compute the distribution function \(G_t\) of \(R_t\) despite the possibly small number of parameters \(p + 1\) (resp. the number \(p\) of free parameters), due to the fact that the convolution procedure as discussed e.g. for the MA processes above cannot be applied in this framework.

Even though the dependence of \(R_t\) on the past of \((N_t)\) within an ARCH model displays a useful characteristic from a statistical point of view, it limits the choice of distribution functions \(G_t\) that we obtain for \(R_t\) and makes the computation of these \(G_t\) a laborious task. Clearly, the source of risk for \(R_t\) here can only come from standard Gaussian random variables \((N_t)\).

Extending the ARCH model as when going from an MA process setup to an ARMA process setup, we arrive at the GARCH\((p,q)\) models, that entail the following:

**Definition 8.4.13 (GARCH processes)**

Let a white noise process \((Z_t)_{t \in \mathbb{Z}}\), coefficients \(\alpha_1, \ldots, \alpha_p \in \mathbb{R}^+\) and \(\delta_1, \ldots, \delta_q \in \mathbb{R}^+\) and an \(\alpha_0 \in \mathbb{R}_0^+\) be given where \(p \in \mathbb{N}_0\) and \(q \in \mathbb{N}\) are fixed integers and \(\alpha_0\delta_q \neq 0\). A process \((Y_t)_{t \in \mathbb{Z}}\) is called a generalised autoregressive conditional heteroscedasticity process of order \((p,q)\), denoted by GARCH\((p,q)\), if it satisfies

\[
Y_t = R_tZ_t,
\]

where

\[
R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \delta_j Y_{t-j}^2,
\]

for \(t \in \mathbb{Z}\).

**Theorem 8.4.14** Let an independent Gaussian white noise process \((Z_t)_{t \in \mathbb{Z}}\), coefficients \(\alpha_1, \ldots, \alpha_p \in \mathbb{R}^+\) and \(\delta_1, \ldots, \delta_q \in \mathbb{R}^+\) and an \(\alpha_0 \in \mathbb{R}_0^+\) be given where \(p \in \mathbb{N}_0\) and \(q \in \mathbb{N}\) are fixed integers and \(\alpha_0\delta_q \neq 0\). The corresponding GARCH\((p,q)\) process \((Y_t)_{t \in \mathbb{Z}}\) with

\[
Y_t = R_tZ_t \quad \text{and} \quad R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \delta_j Y_{t-j}^2,
\]

for \(t \in \mathbb{Z}\),
8.4. The discrete-time case

is stationary with $E(Y_t) = 0$, $\text{Var}(Y_t) = E(R_t^2) = \alpha_0/(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \delta_j)$ and $\text{Cov}(Y_t, Y_{t-h}) = 0$ for $t \in \mathbb{Z}$, $h \in \mathbb{N}$ iff $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \delta_j < 1$.

**Proof:** For a proof see e.g. [Bol86].

If we use an independent Gaussian white noise process $(N_t)_{t \in \mathbb{Z}}$ for $(Z_t)_{t \in \mathbb{Z}}$ as in the previous Theorem 8.4.14, the process $(X_t)_{t \in \mathbb{N}}$ given by

$$X_t = R_t N_t$$

and

$$R_t^2 = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \delta_j} + \sum_{i=1}^p \alpha_i R_{t-i}^2 + \sum_{j=1}^q \delta_j X_{t-j}^2,$$

thus represents a $\text{GARCH}(p, q)$ process, where $R_t$ and $N_t$ are independent, for all $t$. In order to obtain unit second moment for $R_t$, that is $E(R_t^2) = 1$, according to Theorem 8.4.14 it is necessary and sufficient to demand that $\alpha_0 + \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \delta_j = 1$, as $\alpha_0 > 0$.

As in the $\text{ARCH}$ case and according to Theorem 8.4.14, $(X_t)_{t \in \mathbb{N}}$ becomes a particular white noise process. Yet, the process $(R_t^2)_{t \in \mathbb{N}}$ does not follow a special $\text{ARMA}$ process as specified in Definition 8.4.7, due to the fact that $(X_t)_{t \in \mathbb{N}}$ again does not enter linearly into the specification for $(R_t^2)_{t \in \mathbb{N}}$ and as $(X^2_t)_{t \in \mathbb{N}}$ does not become a white noise process in general.

Indeed, the big downside of this setup for our modelling purposes lies in the great difficulty of presenting a closed-form expression for the distribution function $G_t$ for any non-trivial choice of parameters, despite the possibly small number of $p + q + 1$ parameters (resp. $p + q$ free parameters). Even though we are able to incorporate non-Gaussian aspects such as heavy-tails in the characteristics of the resulting distribution function $G_t$, the flexibility for the resulting $G_t$ is still limited, as with this $\text{GARCH}$ specification the only source of risk stems from Gaussian random variables.

**Further autoregressive conditional heteroscedasticity processes**

Engle et al. [BEV05] propose a $\text{TARCH}(1, 1)$ model for the market returns that trigger the defaults. Roughly speaking, a $\text{TARCH}$ model is a $\text{GARCH}$ model with an additional asymmetric component in the specification of the scaling quantity $R_t^2$:

**Definition 8.4.15 (TARCH processes)**

Let a white noise process $(Z_t)_{t \in \mathbb{Z}}$, coefficients $\alpha_1, \ldots, \alpha_p, \delta_1, \ldots, \delta_q, \delta_1^*, \ldots, \delta_q^* \in \mathbb{R}^+$ and an $\alpha_0 \in \mathbb{R}_0^+$ be given, where $p \in \mathbb{N}_0$ and $q \in \mathbb{N}$ are fixed integers and $\alpha_0 \delta_q \neq 0$. A process $(Y_t)_{t \in \mathbb{Z}}$ is called a threshold autoregressive conditional heteroscedasticity process.
of order \((p, q)\), denoted by \(TARCH(p, q)\), if it satisfies
\[
Y_t = R_t Z_t,
\]
where
\[
R_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{j=1}^{q} \delta_j Y_{t-j}^2 + \sum_{j=1}^{q} \delta^*_j Y^2_{t-j} \mathbb{1}_{\{Y_{t-j} \leq 0\}}, \quad \text{for } t \in \mathbb{Z}.
\]

Specializing the white noise process \((Z_t)_{t \in \mathbb{Z}}\) once more to be an independent Gaussian white noise process \((N_t)_{t \in \mathbb{Z}}\), i.e. \(N_t \sim N(0, 1)\) are iid, then \((X_t)_{t \in \mathbb{Z}}\) with
\[
X_t = R_t N_t \quad \text{and} \quad R_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j R_{t-j}^2 + \sum_{i=1}^{q} \delta_i X_{t-i}^2 + \sum_{i=1}^{q} \delta^*_i X_{t-i}^2 \mathbb{1}_{\{X_{t-i} \leq 0\}}
\]
follows a \(TARCH(p, q)\) process and \(R_t\) and \(N_t\) become independent as desired. For the coefficients \(\alpha_1, \ldots, \alpha_p, \delta_1, \ldots, \delta_q, \delta^*_1, \ldots, \delta^*_q \geq 0\) and \(\alpha_0\) we additionally assume that
\[
\alpha_0 + \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \delta_j + \frac{1}{2} \sum_{j=1}^{q} \delta^*_j = 1, \quad \text{by consequence of which } \mathbb{E}(R_t^2) = 1 \quad \text{(the reasoning here is parallel to the reasoning in the } \text{GARCH} \text{ case). The factor } \frac{1}{2} \text{ in front of the sum with the } \delta^*_j \text{'s stems from the fact that } \mathbb{E}(N_t^2 \mathbb{1}_{\{N_t \leq 0\}}) = \frac{1}{2} \text{).}
\]

For the definition of the process \((R_t)_{t \in \mathbb{N}}\) in this \(TARCH(p, q)\) setup \(2p + q\) parameters are involved. Even more than in the \(GARCH\) case, the big downside for our modelling purposes lies again in the great difficulty of presenting a closed-form expression for the distribution function \(G_t\) for any non-trivial choice of parameters. Once again, the flexibility for the resulting \(G_t\) is still limited, as with this \(TARCH\) specification the only source of risk stems from Gaussian random variables.

We could certainly continue to introduce more complex properties within our time-series framework, for example adequate specifications of \(IGARCH\), \(EGARCH\) or other volatility models could be applied for the modelling of the process \((X_t)_{t \in \mathbb{N}}\). Yet, the more involved the models are from a statistical perspective, the higher is the tendency that the difficulty in giving closed-form expressions for the distribution function \(G_t\) will become disproportionate.

### 8.4.3 Step in between: using continuous-time results for the discrete-time case

On the one hand, it is easier to derive a likelihood function within the time series models than within systems of stochastic differential equations and thus to do forecasting and parameter estimation therein. This is especially the case if these differential equations are nonlinear. On the other hand, if we are interested in distributional properties of the processes at any time \(t\), in many cases the stochastic differential equations are easier to handle than the time series models. However, if we assume that the length of the discrete-time intervals is very small, the continuous-time diffusions can still provide us with some insight for the discrete-time time series models, as we want to briefly outline in the following. The results are discussed in further depth in [SV79], [Nel90] or in [Gou97].
The idea is to start off with a sequence of $d$-dimensional processes $((Y_t^{(h)})_{t \in \mathbb{R}^+})_h$ in discrete-time which are indexed by the time unit $h \in \mathbb{R}^+$ between two successive observations. We want to extend them to the continuous time setup by assuming them to be constant between two such successive observation times. That is, for the $h$-th process we assume the form

$$Y_t^{(h)} = Y_{kh}^{(h)}, \quad \text{for } kh \leq t \leq (k+1)h, k \in \mathbb{N},$$

almost surely.

For any $h$, let the process $((Y_t^{(h)})_{t \in \mathbb{N}})$ be Markov of order one, that is, assume that the value $Y_{kh}^{(h)}$ does only depend on the most recent value $Y_{(k-1)h}^{(h)}$. Denote by $\nu_h$ the initial distribution of $Y_0^{(h)}$ and let

$$\Pi_h(y, A) := \mathbb{P}\left(Y_{kh}^{(h)} \in A \mid Y_{(k-1)h}^{(h)} = y\right)$$

denote the conditional distribution of the $h$-th process given the last observed value $y \in \mathbb{R}^d$ where $A \in \mathcal{B}(\mathbb{R}^d)$. We set

$$b_h(y) = \frac{1}{h} \mathbb{E}\left(Y_{kh}^{(h)} - Y_{(k-1)h}^{(h)} \mid Y_{(k-1)h}^{(h)} = y\right)$$

and

$$a_h(y) = \frac{1}{h} \mathbb{E}\left(\left[Y_{kh}^{(h)} - Y_{(k-1)h}^{(h)} \mid Y_{(k-1)h}^{(h)} = y\right] \left[Y_{(k-1)h}^{(h)} - Y_{(k-1)h}^{(h)}\right]^T \mid Y_{(k-1)h}^{(h)} = y\right),$$

for $y \in \mathbb{R}^d$ and $h \in \mathbb{R}^+$. Under further technical assumptions the following theorem holds when we let the time unit $h$ tend towards zero:

**Theorem 8.4.16** If, for $h$ tending towards 0, the sequence of initial distributions $(\nu_h)_h$ tends towards a limit distribution $\nu_0$, and if $(b_h)_h$ and $(a_h)_h$ converge uniformly on compact sets towards well behaved functions $a$ and $b$, where $a(y)$ is a positive definite matrix for every $y \in \mathbb{R}^d$, then the sequence of processes $((Y_t^{(h)})_{t \in \mathbb{R}^+})_h$ converges in distribution to a process $Y = (Y_t)_{t \in \mathbb{R}^+}$ which is a solution of the stochastic differential equation

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t,$$

where, for every $y \in \mathbb{R}^d$, $\sigma(y)$ is a positive definite matrix such that $a(y) = \sigma(y)\sigma(y)^T$, where $W$ is a $d$-dimensional standard Brownian motion independent of $Y_0$, and where $Y_0$ has distribution $\nu_0$.

The convergence in probability of the sequence $((Y_t^{(h)})_{t \in \mathbb{R}^+})_h$ in the previous theorem does not mean that $(Y_t^{(h)})_h$ converges solely for a fixed $t \in \mathbb{R}^+$, but rather for any $0 \leq t \leq T$, where $T \in [0, \infty)$ is fixed. Let $T \in [0, \infty)$ be arbitrarily chosen, then the sequence of probability laws generating the entire sample paths of $((Y_t^{(h)})_{0 \leq t \leq T})_h$ tends towards the probability law generating the sample path of $(Y_t)_{0 \leq t \leq T}$. More on this type of convergence and on the technical details can be found in [Nel90].
Let us consider a slightly simplified \( \text{GARCH} - M(1,1) \) model in discrete-time for a moment, that is, assume that \((Y_t)_{t \in \mathbb{Z}}\) satisfies the following system of equations:

\[
Y_t = Y_{t-1} + R_t N_t, \\
R_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \delta_1 R_{t-1}^2 N_{t-1}^2,
\]

where \((N_t)_{t \in \mathbb{Z}}\) is an independent Gaussian white noise process and \(\alpha_0, \alpha_1, \delta_1 > 0\) are given parameters such that \(\alpha_0 + \alpha_1 + \delta_1 = 1\). The aim is to find a continuous approximation of such a discrete-time model. To this end, let us assume the following sequence of models, which again shall be indexed by the time unit \(h\):

\[
Y_{kh} = Y_{(k-1)h} + \sqrt{h} R_{(k-1)h}^{(h)} N_k^{(h)},
\]

\[
R_{kh}^{(h)^2} = \alpha_{0,h} + \alpha_{1,h} R_{(k-1)h}^{(h)^2} + h \delta_{1,h} R_{(k-1)h}^{(h)^2} N_{k-1}^{(h)^2},
\]

where, for every \(h \in \mathbb{R}^+\), \((N_t^{(h)})_{t \in \mathbb{Z}}\) is an independent Gaussian white noise process with \(N_t^{(h)} \sim N(0,1)\) and where \(\alpha_{0,h}, \alpha_{1,h}, \delta_{1,h} > 0\) are positive constants such that \(\alpha_{0,h} + \alpha_{1,h} + h \delta_{1,h} = 1\).

According to the previous Theorem 8.4.16 and under additional assumptions on the sequence of the distributions of the \(R_0^{(h)}\), \(h \in \mathbb{R}^+\), and on the sequences of the parameters \((\alpha_{0,h})_h\), \((\alpha_{1,h})_h\) and \((\delta_{1,h})_h\) such as

\[
\lim_{h \downarrow 0} \frac{\alpha_{0,h}}{h} = \lim_{h \downarrow 0} \frac{1 - \alpha_{1,h} - h \delta_{1,h}}{h} = \kappa \geq 0 \quad \text{and} \quad \lim_{h \downarrow 0} 2h \delta_{1,h}^2 = \delta^2
\]

the sequence of processes \(\left((Y_t^{(h)}, R_t^{(h)^2})_{t \in \mathbb{R}^+}\right)_h\) defined by the extension of the above sequence of \(\text{GARCH}(1,1) - M\) models to continuous time converges towards the processes \((Y_t, r_t^2)_{t \in \mathbb{R}^+}\) whose components satisfy the following system of stochastic differential equations:

\[
dY_t = R_t dW_{1,t},
\]

\[
dR_t^2 = \kappa(1 - R_t^2) dt + \delta R_t^2 dW_{2,t},
\]

where \((W_{1,t})_{t \in \mathbb{R}^+}\) and \((W_{2,t})_{t \in \mathbb{R}^+}\) are independent standard Brownian motions and where the process \((R_t^2)_{t \in \mathbb{R}^+}\) becomes strictly stationary with \(R_t^2\) following an inverse Gamma distribution\(^1\) \(\text{Inv}\Gamma(1 + 2\kappa/\delta^2, 2\kappa/\delta^2)\), for all \(t \in \mathbb{R}^+\). In particular, \(R_t\) also satisfies the desired properties of \(R_t \geq 0\) and \(E(R_t^2) = 1\). Even though the \(\text{GARCH} - M(1,1)\) model does not fit into our framework, as the marginal distributions of the process \((Y_t)_{t \in \mathbb{Z}}\) do not belong to the mixtures of Normal distributions, the approximation and derivation of the scaling process \((R_t)_{t \in \mathbb{R}^+}\) within this model can be used directly for the approximation of the distribution of the scaling process \((R_t)_{t \in \mathbb{R}^+}\) within a \(\text{GARCH}(1,1)\) model. Indeed, the difference between the \(\text{GARCH}(1,1)\) and the \(\text{GARCH} - M(1,1)\) model only lies in the impact of the process \((R_t)_{t \in \mathbb{R}^+}\) on \((Y_t)_{t \in \mathbb{R}^+}\),

\(^{1}\) A random variable \(X\) is said to follow an inverse Gamma distribution with parameters \(a, b > 0\), denoted by \(\text{Inv}\Gamma(a, b)\), if its density function is given via \(f(x) = \frac{b^a}{\Gamma(a)} x^{-a-1} \exp(-bx)\). This distribution function has mean \(\frac{b}{a-1}\) for \(a > 1\).
either via the setting \( Y_t = R_t N_t \) or via \( Y_t = Y_{t-1} + R_t N_t \). The specification of the process \((R_t)_{t \in \mathbb{R}^+}\) stays unchanged and thus the results from the above discussion can be directly transferred to the \(GARCH(1,1)\) case.

### 8.5 The continuous-time case

The following sections will now deal with the continuous-time case, where the index set \( I \) is a subset of the non-negative real numbers \( \mathbb{R}^+ \). Even though the step from a static model to a dynamic model seems to be more easily accomplished via a discrete-time evolution of the risk components, the obvious downside there lies in the fact that we often have great difficulty in giving closed-form expressions for the distribution function \( G_t \) of the scaling random variable \( R_t \) at any time \( t \in I \). A way of tackling this problem was to look at approximations of the discrete-time models such as the \(GARCH\) models by continuous-time stochastic processes where the key components can often be controlled more easily, as we have just seen in the previous Section 8.4. However, instead of starting with a discrete-time model and afterwards using results from the continuous-time world, in this section we directly want to analyse continuous-time models that fit into our elliptical distributions framework and provide us with the relevant building blocks needed for the computations of the expected losses in a specific tranche of a CDO.

As in the discrete-time case, we can have two possibilities. First, we can model only the scaling process \((R_t)_{t \in I}\) separately from the Gaussian process \((N_t)_{t \in I}\), which will be done in the next section and which can then be used for both components, the factor and the idiosyncratic component. The other approach is to model \((R_t)_{t \in I}\) and \((N_t)_{t \in I}\) together, which will be done in the subsequent sections. There, we have to be a bit more careful with applying this setup to the idiosyncratic risk process. In both cases, we have to make sure that \( R_t \geq 0 \) and that \( R_t, N_t \) are independent, for every \( t \in I \). In order to be able to apply the static setup directly, we are also interested to have that \( \mathbb{E}(R_t^2) = 1 \).

As outlined above, in this continuous-time case with \( I \subseteq \mathbb{R}^+ \) we assume that \( \mathcal{F} \) satisfies the usual conditions of right-continuity and completeness, while \( \mathcal{F}_0 \) is again the trivial \( \sigma \)-field. Again, all dynamics are observed under a risk-neutral pricing measure \( \mathbb{P} \).

#### 8.5.1 Employing continuous-time processes from the interest-rate theory for the scaling process

Parallel to the discrete-time setup, the most obvious procedure is to modify standard continuous-time processes for the scaling process \((R_t)_{t \in \mathbb{R}^+}\) so that they fit into our setup. Yet, the different specifications that we will discuss here for \((R_t)_{t \in \mathbb{R}^+}\) do not restrict the specification of the process \((N_t)_{t \in \mathbb{R}^+}\). Indeed, \((N_t)_{t \in \mathbb{R}^+}\) can be either one-dimensional or \( n \)-dimensional and can represent e.g. a continuous-time Gaussian white noise process or stem from a Brownian motion displaying the desired variance/covariance structure.
The interest-rate theory provides us with many well-discussed models that yield non-negative processes, such as the model by Cox, Ingersoll and Ross [CIR85], the Dothan model [Dot78] or the Black-Karasinski model [BK91]. In the following we will discuss how to apply these models within our elliptical distributions setup.

Using the Cox-Ingersoll-Ross model

Assume that the process \((r_t)_{t \in \mathbb{R}^+}\) evolves according to the following stochastic differential equation:
\[
dr_t = \kappa(\theta - r_t)dt + \sigma \sqrt{r_t}dW_t,
\]
where \((W_t)_{t \in \mathbb{R}^+}\) is a one-dimensional standard Brownian motion, which we assume to be independent of the process \((N_t)_{t \in \mathbb{R}^+}\), and \(r_0, \kappa, \theta, \sigma > 0\) are positive constants with 
\[2\kappa \theta > \sigma^2.\]
The last condition on the parameters assures that the process \((r_t)_{t \in \mathbb{R}^+}\) remains positive at any time \(t\). This type of process goes back to Cox, Ingersoll and Ross [CIR85], and is often used for the modelling of short rates. The distributional behaviour of \(r_t\), for an arbitrary \(t \in \mathbb{R}^+\), is well established (see e.g. [BM06]): \(r_t\) essentially follows a non-central Chi-Square distribution and possesses the following density function:
\[
p_{r_t}(x) = c_t p_{\chi^2}(\nu, \lambda_t)(c_t x),
\]
where
\[
c_t = \frac{4\kappa}{\sigma^2(1 - \exp(-\kappa t))},
\]
\[
\nu = \frac{4\kappa \theta}{\sigma^2},
\]
\[
\lambda_t = c_t r_0 \exp(-\kappa t),
\]
and where \(p_{\chi^2}(\nu, \lambda)\) denotes the density function of the non-central Chi-Square distribution\(^2\) with \(\nu\) df and non-centrality parameter \(\lambda\). From this we can deduce that
\[
E(r_t^2) = \theta^2 + \theta \frac{\sigma^2}{2\kappa} + a \exp(-\kappa t) + b \exp(-2\kappa t) > 0,
\]
for \(t \in \mathbb{R}^+\), where
\[
a := r_0 \frac{\sigma^2}{\kappa} + 2r_0 \theta - 2\theta^2 - \theta \frac{\sigma^2}{\kappa},
\]
\[
b := -r_0 \frac{\sigma^2}{\kappa} - 2r_0 \theta + r_0^2 + \theta^2 + \theta \frac{\sigma^2}{2\kappa} = (r_0 - \theta)^2 + \left(\theta - 2r_0\right) \frac{\sigma^2}{2\kappa}.
\]
\(^2\)The density function of the non-central Chi-Square distribution with \(\nu\) df and non-centrality parameter \(\lambda > 0\) is given via \(p_{\chi^2(\nu, \lambda)}(x) = \sum_{i=0}^{\infty} \frac{(-\lambda/2)^i}{i!} p_{\chi^2(\nu+2i)}(x)\), for \(x > 0\), where \(p_{\chi^2(\nu)}\) is the density of a central Chi-Square distribution with \(\nu\) df given via \(p_{\chi^2(\nu)}(x) := \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2 - 1} e^{-x/2}\), for \(x > 0\).
Then, as $\kappa > 0$, $\mathbb{E}(r_t^2)$ is constant in $t \in \mathbb{R}^+$ iff $a = b = 0$, which is equivalent to demanding

$$\frac{\gamma_0^2 - \theta^2}{\theta} = \frac{\sigma^2}{2\kappa} \quad \text{and} \quad \frac{(r_0 - \theta)^2(r_0 + 2\theta)}{\theta} = 0.$$ 

Yet, this would entail that $r_0 = \theta$, and would thus imply that $\sigma = 0$, which is a contradiction to the restriction on the parameters.

Therefore, we have to scale the process $(r_t)_{t \in \mathbb{R}^+}$ appropriately in order to obtain a process $(R_t)_{t \in \mathbb{R}^+}$ that matches our framework. Let $\gamma_t := 1/\sqrt{\mathbb{E}(r_t^2)}$ and $R_t := \gamma_t r_t$ for $t \in \mathbb{R}^+$. Then $R_t \geq 0$ and $\mathbb{E}(R_t^2) = 1$ for all $t \in \mathbb{R}^+$. Additionally, we have

$$dR_t = \gamma_t dr_t + r_t \gamma_t' dt = \left(\kappa \theta \gamma_t + R_t \left(\frac{\gamma_t'}{\gamma_t} - \kappa\right)\right) dt + \sqrt{\gamma_t R_t \sigma^2} dW_t,$$

and $R_t$ possesses the following density function:

$$p_{R_t}(x) = \frac{c_t}{\gamma_t} p_{\chi^2(\nu, \lambda_t)} \left(\frac{c_t x}{\gamma_t}\right),$$

for $x \in \mathbb{R}^+$, where $c_t, \lambda_t$ and $\nu$ are as above and $t \in \mathbb{R}^+$.

Using the Dothan model

Another popular model within the theory of short-rate models is where one essentially assumes a geometric Brownian motion for the evolution of the short-rate process. This short-rate model is also known as the Dothan model and was discussed in [Dot78]. Assume that the process $(r_t)_{t \in \mathbb{R}^+}$ evolves according to the following stochastic differential equation:

$$dr_t = \sigma \sqrt{r_t} dt + \sigma r_t dW_t,$$

where $(W_t)_{t \in \mathbb{R}^+}$ is a one-dimensional standard Brownian motion, which we assume to be independent of the process $(N_t)_{t \in \mathbb{R}^+}, \ r_0, \sigma > 0$ are positive constants and $a \in \mathbb{R}$.

One can solve this differential equation for $r_t$ and obtains

$$r_t = r_0 \exp \left(\frac{1}{2} \sigma^2 t + \sigma W_t\right).$$

Thus, the process $(r_t)_{t \in \mathbb{R}^+}$ stays positive, $r_t/r_0$ follows a Log-Normal distribution\(^3\) $r_t/r_0 \sim LN((a - \frac{1}{2} \sigma^2) t, \sigma^2 t)$ and $\mathbb{E}(r_t^2) = r_0^2 \exp((2a + \sigma^2) t)$, for $t \in \mathbb{R}^+$. Again, we set $\gamma_t := 1/\sqrt{\mathbb{E}(r_t^2)} = r_0^{-1} \exp(-a + \sigma^2/2) t$ and

$$R_t := \gamma_t r_t = \exp(-\sigma^2 t + \sigma W_t)$$

---

\(^3\) A random variable $X$ is said to follow a Log-Normal distribution with parameters $(\mu, \sigma^2)$, denoted by $LN(\mu, \sigma^2)$, if $\log(X) \sim N(\mu, \sigma^2)$. Its density function is then given via $p_{LN(\mu, \sigma^2)}(x) := \frac{1}{\sigma \sqrt{2\pi}} e^{-(\log x - \mu)^2/(2\sigma^2)}$, for $x \in \mathbb{R}^+$. 
for $t \in \mathbb{R}^+$. Then $R_0 = 1$, $R_t \geq 0$ and $\mathbb{E}(R_t^2) = 1$ for all $t \in \mathbb{R}^+$. Therefore, for this model it suffices to specify the volatility parameter $\sigma$. Additionally, we have

$$dR_t = \gamma_t dr_t + r_t \gamma_t' dt = -\frac{1}{2} \sigma^2 r_t dt + \sigma r_t dW_t,$$

and $R_t \sim LN(-\sigma^2 t, \sigma^2 t)$ for $t \in \mathbb{R}^+$.

**Using the Black-Karasinski model**

Another readily available type of model is the Black-Karasinski model [BK91] where we assume that the logarithm of the process $(r_t)_{t \in \mathbb{R}^+}$, that is the process $(y_t = \log(r_t))_{t \in \mathbb{R}^+}$, follows an Ornstein-Uhlenbeck process. More precisely, we assume that $(y_t)_{t \in \mathbb{R}^+}$ satisfies

$$dy_t = (\theta_t - ay_t) dt + \sigma dW_t,$$

where $(W_t)_{t \in \mathbb{R}^+}$ is a one-dimensional standard Brownian motion, which we assume to be independent of the process $(N_t)_{t \in \mathbb{R}^+}$, $a, \sigma > 0$ are positive constants and $(\theta_t)_{t \in \mathbb{R}^+}$ is a deterministic process with values in $\mathbb{R}^+$. The dynamics of $(r_t)_{t \in \mathbb{R}^+}$ are then specified via the stochastic differential equation

$$dr_t = \left( \theta_t + \frac{\sigma^2}{2} - a \log(r_t) \right) r_t dt + \sigma r_t dW_t,$$

for $t \in \mathbb{R}^+$, where $r_0 := \exp(y_0) > 0$. If we set $\theta_t \equiv \theta > 0$ we obtain the so-called Exponential Vasicek model. Once more, one can solve the above differential equation for $r_t$ and obtains

$$r_t = \exp \left( \log(r_0) e^{-at} + \int_0^t e^{-a(t-u)} \theta_u du + \sigma \int_0^t e^{-a(t-u)} dW(u) \right).$$

Thus, the process $(r_t)_{t \in \mathbb{R}^+}$ stays positive, $r_t$ follows a Log-Normal distribution $r_t \sim LN(\log(r_0) e^{-at} + \int_0^t e^{-a(t-u)} \theta_u du, \sigma^2 \int_0^t e^{-2a(t-u)} du)$ and

$$\mathbb{E}(r_t^2) = \exp \left( 2 \log(r_0) e^{-at} + 2 \int_0^t e^{-a(t-u)} \theta_u du + \frac{\sigma^2}{a} (1 - e^{-2at}) \right),$$

for $t \in \mathbb{R}^+$. Again, we set $\gamma_t := 1/\sqrt{\mathbb{E}(r_t^2)}$ and $R_t := \gamma_t r_t$, for $t \in \mathbb{R}^+$. Then $R_0 = 1$, $R_t \geq 0$ and $\mathbb{E}(R_t^2) = 1$, for all $t \in \mathbb{R}^+$. Additionally, we have

$$dR_t = \gamma_t dr_t + r_t \gamma_t' dt$$

$$= \left( \theta_t + \frac{\sigma^2}{2} + \gamma_t' + a \log(\gamma_t) - a \log(R_t) \right) R_t dt + \sigma R_t dW_t$$

and

$$R_t \sim LN \left( \log(r_0) e^{-at} + \log(\gamma_t) + \int_0^t e^{-a(t-u)} \theta_u du, \sigma^2 \int_0^t e^{-2a(t-u)} du \right),$$

for $t \in \mathbb{R}^+$. 

8.5. The continuous-time case
Using the Mercurio-Moraleda model

Mercurio and Moraleda [MM00] proposed a model where the logarithm of the process \((r_t)_{t \in \mathbb{R}^+}\), that is the process \((y_t = \log(r_t))_{t \in \mathbb{R}^+}\), also follows an Ornstein-Uhlenbeck process. Let us assume that \((y_t)_{t \in \mathbb{R}^+}\) satisfies
\[
dy_t = \left( \theta_t - \left( \lambda - \frac{\delta}{1 + \delta t} \right) y_t \right) dt + \sigma dW_t,
\]
where \((W_t)_{t \in \mathbb{R}^+}\) is a one-dimensional standard Brownian motion, which we assume to be independent of the process \((N_t)_{t \in \mathbb{R}^+}\), \(a, \sigma > 0\) are positive constants, \(\delta \geq 0\) and \((\theta_t)_{t \in \mathbb{R}^+}\) is a deterministic process with values in \(\mathbb{R}^+_0\). The dynamics of \((r_t)_{t \in \mathbb{R}^+}\) are then specified via the stochastic differential equation
\[
dr_t = \left( \theta_t + \frac{\sigma^2}{2} - \left( \lambda - \frac{\delta}{1 + \delta t} \right) \log(r_t) \right) r_t dt + \sigma r_t dW_t,
\]
for \(t \in \mathbb{R}^+\), where \(r_0 := \exp(y_0) > 0\). For \(\delta = 0\) we fall back into the Black-Karasinski model of the type we have just discussed in the last paragraph. Once more, one can solve the above differential equation for \(r_t\) and obtains
\[
r_t = \exp \left( \log(r_0) B(0, t) + \int_0^t B(u, t) \theta_u du + \sigma \int_0^t B(u, t) dW(u) \right),
\]
where \(B(u, t) := \frac{1 + \delta u}{1 + \delta t} e^{-(t-u)}\), for \(0 \leq u \leq t\). Thus, the process \((r_t)_{t \in \mathbb{R}^+}\) stays positive, \(r_t\) follows a Log-Normal distribution
\[
r_t \sim LN(\log(r_0) B(0, t) + \log(\gamma_t) + \int_0^t B(u, t) \theta_u du, \sigma^2 \int_0^t B^2(u, t) du),
\]
and
\[
E(r_t^2) = \exp \left( 2 \log(r_0) B(0, t) + 2 \int_0^t B(u, t) \theta_u du + 2\sigma^2 \int_0^t B^2(u, t) du \right),
\]
for \(t \in \mathbb{R}^+\). Again, we set \(\gamma_t := 1/\sqrt{E(r_t^2)}\) and \(R_t := \gamma_t r_t\), for \(t \in \mathbb{R}^+\). Then \(R_0 = 1\), \(R_t \geq 0\) and \(E(R_t^2) = 1\), for all \(t \in \mathbb{R}^+\). Additionally, we have
\[
dR_t = \gamma_t dr_t + r_t \gamma_t' dt
= \left( \theta_t + \frac{\sigma^2}{2} + \frac{\gamma_t'}{\gamma_t} + \left( \lambda - \frac{\delta}{1 + \delta t} \right) (\log(\gamma_t) - \log(R_t)) \right) R_t dt + \sigma R_t dW_t
\]
and
\[
R_t \sim LN \left( \log(r_0) B(0, t) + \log(\gamma_t) + \int_0^t B(u, t) \theta_u du, \sigma^2 \int_0^t B^2(u, t) du \right),
\]
for \(t \in \mathbb{R}^+\).
Discussion of the above models from the interest-rate theory

All of the above models are well understood in the context of short-rate models and many relevant properties of the resulting processes are readily available. While for our setup we do not need to make use of pricing formulas for standard contracts such as bonds (as we do not apply these models for such short-rate processes), the results about the distributional behaviour and some limit arguments can be employed directly within our context, as we have just seen. While the Cox-Ingersoll-Ross model essentially resulted in Chi-Square distributions for the scaling variables $R_t$, $t \in \mathbb{R}^+$, the other models, that is the Dothan, the Black-Karasinski and the Mercurio-Moraleda model, essentially yielded Log-Normal distributions for the $R_t$, $t \in \mathbb{R}^+$. As a result of this, all of these models present closed-form expressions for the distributions $G_t$, $t \in \mathbb{R}^+$, which are needed as the main building blocks for the valuation of the CDO tranches. Additionally, the four different models allow considerable flexibility as the number of parameters can be chosen according to the amount of data available. While in our specification the Dothan model made only use of one parameter, the Cox-Ingersoll-Ross model gets by with 3 parameters, and the two remaining models, i.e. the Black-Karasinski and the Mercurio-Moraleda model, provide us with the flexibility of choosing a deterministic function that suits our data best.

In the interest-rate theory, the problem arises that a log-normally distributed short-rate produces a so-called “explosion of the bank account”. In that, the expected value of the investment of a fixed amount into a bank account goes to infinity if we let the time interval, during which we remain invested in this position, tend to zero (see Section 3.2.2. in [BM06] for a discussion of this feature). However, this does not play any role in our context, as all we need to postulate is the existence of the second moment of the scaling variable at any time $t$, which is indeed satisfied for all these models.

We may wonder whether these short-rate models entail tail-dependence for the resulting Normal mixture vector $R_tN_t$, where $R_t$ stems from one of the above models and $N_t \sim N_n(0, \Sigma_n)$ with an $n$-dimensional (diagonal) covariance matrix $\Sigma_n$. As we haven seen in Section 5.3.1, all we need for this is to demand that the tail of the distribution function of $R_t$ is regularly varying at infinity. Yet, neither the Log-Normal distribution nor the Chi-Square distribution belong to this class:

In the case where $R_t$ follows a Log-Normal distribution $LN(\mu, \sigma)$, its tail can be represented via

$$1 - F_{LN(\mu, \sigma)}(y) = \Phi \left( \frac{\mu - \log(y)}{\sigma} \right) \quad \text{for } y \in \mathbb{R}_0^+.$$
Thus, by virtue of L’Hospital’s rule we have

\[
\lim_{x \to \infty} \frac{1 - F_{LN(\mu, \sigma)}(tx)}{1 - F_{LN(\mu, \sigma)}(x)} = \lim_{x \to \infty} \frac{\Phi\left(\frac{\mu - \log(tx)}{\sigma}\right)}{\Phi\left(\frac{\mu - \log(x)}{\sigma}\right)} = \lim_{x \to \infty} \frac{\varphi\left(\frac{\mu - \log(x) - \log(t)}{\sigma}\right)}{\varphi\left(\frac{\mu - \log(x)}{\sigma}\right)}
\]

\[
= \lim_{x \to \infty} \exp\left(-\frac{1}{2\sigma^2} \log(x) \log(t) + \frac{1}{\sigma^2} \log(t)(\log(t) - \mu)\right)
\]

\[
= \begin{cases} 
\infty, & \text{for } t < 1; \\
1, & \text{for } t = 1; \\
0, & \text{for } t > 1;
\end{cases}
\]

for \( t > 0 \). Therefore, the tail of the Log-Normal distribution cannot be regularly varying.

The same holds for the Chi-Square distribution, where already its central form fails this property: For \( t > 0 \), again L’Hospital’s rule entails that

\[
\lim_{x \to \infty} \frac{1 - F_{\chi^2(\nu)}(tx)}{1 - F_{\chi^2(\nu)}(x)} = \lim_{x \to \infty} \frac{\int_{xt}^{\infty} s^{\nu/2-1} \exp(-s/2)ds}{\int_{x}^{\infty} s^{\nu/2-1} \exp(-s/2)ds} = \lim_{x \to \infty} \frac{\nu/2 \exp(-xt/2)}{\exp(-x/2)} \]

\[
= \frac{\nu/2}{\exp(-1/2(t-1)x)} \begin{cases} 
\infty, & \text{for } t < 1; \\
1, & \text{for } t = 1; \\
0, & \text{for } t > 1.
\end{cases}
\]

Yet, the ease in computing the relevant objects such as the distribution functions \( G_t \) for any relevant time \( t \), and at the same time introducing several flexible roads on how to make our elliptical distributions model dynamic, are the strong advantages of using such continuous-time short-rate models. Additionally, they provide us with an intuition of how the dependence between the components in the vector \( R_t N_t \sim EC_n(0, \Sigma_n, \phi_t) \) changes in \( t \), for example if we choose for the process \((R_t)_{t \in \mathbb{R}^+}\) one of the above models which stems from a process \((r_t)_{t \in \mathbb{R}^+}\) that displays such a property as e.g. mean-reversion (for example the CIR or the Exponential Vasicek model display this property, but not the Dothan model).

### 8.5.2 Dynamic Gaussian mixture setup via time-changed Brownian motions

In the following, we want to discuss a way of constructing a process \((X_t)_{t \in I}\) via Brownian motions, such that its marginal distributions all follow mixtures of Normal distributions. Hereby, the components of \( X_t \), that are the scaling variable \( R_t \) and the Gaussian component \( N_t \), will be modelled together, for \( t \in I \). We will first start with one-dimensional processes, which then aims for modelling the factor process, and will finally end this section with \( n \)-dimensional processes, which can be considered for the modelling of the idiosyncratic vector process.
We assume the filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\), with \(\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}\), to be rich enough to allow for the existence of a standard \(\mathbb{F}\)-Brownian motion \(W\), and a family 
\[ C := \{C_s\}_{s \geq 0} \]
of stopping times such that the mapping
\[ s \mapsto C_s \]
is right-continuous and increasing.

A family \(C\) with these properties is called a **time change**.

**Proposition 8.5.1** If the mapping \(s \mapsto C_s\) is even continuous and \(\mathbb{E}(C_t) < \infty\) for all \(t \geq 0\), then the time-changed Brownian motion \(\hat{W}\) with \(\hat{W}_t := W_{C_t}\), for \(t \geq 0\), is a continuous martingale with respect to the filtration \(\hat{\mathbb{F}} := (\hat{\mathcal{F}}_t)_{t \geq 0}\), where \(\hat{\mathcal{F}}_t := \mathcal{F}_{C_t}\) for \(t \geq 0\), with quadratic variation \(<\hat{W}, \hat{W}>_t = C_t\) and \(\mathbb{E}(\hat{W}^2_t) = \mathbb{E}(C_t)\), for all \(t \geq 0\).

**Proof:** The Brownian motion is a continuous martingale and is thus also a continuous local martingale. As the mapping \(s \mapsto C_s\) is continuous, \(W\) naturally fulfills the weaker assumption of \(C\)-continuity, where one requests \(W\) to be constant on each interval \([C_t, C_t]\), \(t \geq 0\). The family \(C\) is almost surely finite, as \(\mathbb{E}(C_t) < \infty\) and \(C_t \in \mathbb{R}^+\) imply \(C_t < \infty\) almost surely, for every \(t \geq 0\). By virtue of Proposition 1.5 in Chapter V. of [RY94], the time-changed Brownian motion \(\hat{W}\) is a local martingale with respect to \(\hat{\mathbb{F}}\) and \(<\hat{W}, \hat{W}>_t = <W, W>_t = C_t\), for all \(t \geq 0\).

Taking into consideration that
\[ \mathbb{E}(<\hat{W}, \hat{W}>_t) = \mathbb{E}(C_t) < \infty, \quad t \geq 0,\]
and making use of Corollary 3 after Theorem 27 in Chapter II. of [Pro04], we conclude that \(\hat{W}\) is a martingale with \(\mathbb{E}(\hat{W}^2_t) = \mathbb{E}(<\hat{W}, \hat{W}>_t) = \mathbb{E}(C_t)\), for all \(t \geq 0\). 

Define the non-negative process \((\hat{R}_t)_{t \in I}\) and the process \((N_t)_{t \in I}\) as follows:
\[ \hat{R}_t := \sqrt{C_t} \quad \text{and} \quad N_t := \frac{W_{C_t}}{\sqrt{C_t}}, \quad \text{for} \; t \in I.\]

Then \(\hat{W}_t = W_{C_t} = \hat{R}_tN_t\), for any \(t \in I\), additionally we have \(N_t|_{C_t=c} = \frac{W_{C_t}}{\sqrt{C_t}}|_{C_t=c} \sim N(0, 1)\) for \(c \in \mathbb{R}^+\) due to the independence of \(W\) and \(C\), and we obtain for any \(r \geq 0\) and any \(x \in \mathbb{R}\) the following set of equations:
\[ \mathbb{P}(\hat{R}_t \leq r, N_t \leq x) = \mathbb{P}\left(\sqrt{C_t} \leq r, N_t \leq x\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{\sqrt{C_t} \leq r\}} \mathbb{1}_{\{N_t \leq x\}} | C_t\right)\right)\]
\[ = \mathbb{E}\left(\mathbb{1}_{\{\sqrt{C_t} \leq r\}} \mathbb{E}\left(\mathbb{1}_{\{N_t \leq x\}} | C_t\right)\right) = \int_0^r \mathbb{P}\left(\frac{W_c}{\sqrt{c}} \leq x\right) dF_{C_t}(c)\]
\[ = \int_0^\infty \mathbb{1}_{\{\sqrt{c} \leq r\}} \Phi(x) dF_{C_t}(c) = \mathbb{E}\left(\mathbb{1}_{\{\sqrt{C_t} \leq r\}} \Phi(x)\right)\]
\[ = \mathbb{P}(\hat{R}_t \leq r) \Phi(x),\]
where the fourth equality holds as $W, C$ are independent. If we let $r$ go to infinity, we see that $P(N_t \leq x) = \Phi(x)$ for $x \in \mathbb{R}$, and therefore $\tilde{R}_t$ and $N_t$ are independent with $N_t \sim N(0,1)$. Thus, $\tilde{X}_t := \tilde{R}_t N_t \sim EC_1(0,c_t,\tilde{\phi}_t)$, where $c_t := \text{Var}(\tilde{X}_t) = E(C_t)$ and

$$\tilde{\phi}_t(x) := \int_0^\infty \exp \left( -\frac{1}{2} r^2 x \right) d\tilde{G}_t(r), \quad x \in \mathbb{R},$$

with $\tilde{G}_t$ being the distribution function of $\tilde{R}_t = \sqrt{C_t}$.

As we want to fall back into our framework of the scaling variable $R_t$ having unit-variance, we need to set $R_t := \frac{1}{\sqrt{c_t}} \tilde{R}_t$ and thereafter $X_t := R_t N_t$. Then $X_t = \frac{1}{\sqrt{c_t}} \tilde{X}_t = \frac{1}{\sqrt{\pi}} W C_t$. Yet, we lose the martingale property of $X$, if $c_t$ remains to depend on $t$. The mapping $t \mapsto E(C_t) = c_t$ is increasing, as $C$ is an increasing family of stopping times. In general, $X$ therefore only remains an $\mathbb{F}$-supermartingale due to the decreasing monotonicity of $t \mapsto \frac{1}{\sqrt{c_t}}$, but which is closed as

$$\sup_{t \in \mathbb{R}} \mathbb{E}(|X_t|) = \sup_{t \in \mathbb{R}} \frac{1}{\sqrt{c_t}} \mathbb{E}(|W C_t|) \overset{\text{W. ind.}}{=} \sqrt{\frac{2}{\pi}} \sup_{t \in \mathbb{R}} \frac{1}{\sqrt{c_t}} \mathbb{E}(\sqrt{C_t}) \leq \sqrt{\frac{2}{\pi}} \sup_{t \in \mathbb{R}} \sqrt{c_t} = \sqrt{\frac{2}{\pi}} < \infty,$$

by Jensen’s inequality. If we apply this to the factor $M$, that is on setting $M_t := X_t = \frac{1}{\sqrt{c_t}} \tilde{X}_t = \frac{1}{\sqrt{\pi}} W C_t$, then $M$ fulfills the properties in (8.4) and represents a uniformly integrable $\mathbb{F}$-supermartingale.

**Remark 8.5.2** If instead we use the above construction directly and set $M := \tilde{X}$, that is $M_t := \tilde{X}_t = \tilde{R}_t N_t = W C_t$, for every $t \in I$, then $M$ becomes an $\mathbb{F} = (\mathcal{F}_t)_{t \in I} = (\mathcal{F}_C)_{t \in I}$-martingale. Yet, the conditions on the variance of $M_t$ in (8.4) are no longer satisfied, as $\text{Var}(M_t) = E(C_t) = c_t$, which must not equal one, for all $t \geq 0$ in general.

Extending the above one-dimensional case to $n$ dimensions, we can argue as follows: on the probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, let an $n$-dimensional $\mathbb{F}$-Brownian motion $B^{(n)}(t) = (B^{(n)}(t))_{t \in \mathbb{R}^+}$ with zero-drift and covariance matrix $\Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n}$, $\omega_i > 0$, be given where $B^{(n)}(t) = (B_{1,t}, \ldots, B_{n,t})$, $t \in \mathbb{R}^+$. Assume that there also exists a non-negative one-dimensional process $C = (C_t)_{t \in \mathbb{R}^+}$, independent of $B^{(n)}$ and with $0 < E(C_t) < \infty$, for all $t \in \mathbb{R}^+$. Define the non-negative process $(\tilde{R}_t)_{t \in I}$ and the processes $(Y_{t}^{(n)})_{t \in I}$, $(X_{t}^{(n)})_{t \in I}$ as follows:

$$\tilde{R}_t := \sqrt{C_t}, \quad Y_{t}^{(n)} := (Y_{1,t}, \ldots, Y_{n,t})^T := \frac{1}{\sqrt{C_t}} B_{C_t}^{(n)} = \left( \frac{B_{1,t}}{\sqrt{C_t}}, \ldots, \frac{B_{n,t}}{\sqrt{C_t}} \right)^T$$

and

$$X_{t}^{(n)} := \tilde{R}_t Y_{t}^{(n)} = B_{C_t}^{(n)}, \quad \text{for } t \in I.$$

Let us fix $t \in I$ for a moment. Then, as in the one-dimensional case we can condition on a realization of $C_t$, i.e. $C_t = c$ for $c \in \mathbb{R}^+$, and we obtain $Y_{t}^{(n)}|_{C_t=c} = \frac{1}{\sqrt{c_t}} B_{C_t}^{(n)}|_{C_t=c} \sim \mathcal{N}(0, C_t)$.
\[ N_n(0, \Sigma_n) \text{ with } \Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n), \]\n
where the components of \( Y^{(n)}_t \) are thus conditionally independent. The independence of \( B^{(n)} \) and \( C \) then yields for any \( r \geq 0 \) and any \( x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n \) the following set of equations:

\[
P\left( \sqrt{C_t} \leq r, Y^{(n)}_t \leq x \right) = E\left( E\left( \mathbb{1}_{\{\sqrt{C_t} \leq r\}} \mathbb{1}_{\{Y^{(n)}_t \leq x\}} \mid C_t \right) \right) = E\left( \mathbb{1}_{\{\sqrt{C_t} \leq r\}} E\left( \mathbb{1}_{\{Y^{(n)}_t \leq x_1, \ldots, Y^{(n)}_t \leq x_n\}} \mid C_t \right) \right)
\]

\[
B^{(n)}, C \overset{\text{ind.}}{=} \int_0^r 2 \mathbb{P} \left( B^{(n), c}_{1/c} \leq x_1, \ldots, B^{(n), c}_{n/c} \leq x_n \right) dF_{C_t}(c)
\]

\[
= \int_0^\infty \mathbb{1}_{\{\sqrt{C_t} \leq r\}} \prod_{i=1}^n \Phi \left( \frac{x_i}{\sqrt{\omega_i}} \right) dF_{C_t}(c)
\]

\[
= \mathbb{P}(\sqrt{C_t} \leq r) \prod_{i=1}^n \Phi \left( \frac{x_i}{\sqrt{\omega_i}} \right)
\]

If we again let \( r \) go to infinity, we see that

\[
P\left( Y^{(n)}_t \leq x \right) = \mathbb{P}(Y_{1,t} \leq x_1, \ldots, Y_{n,t} \leq x_n) = \prod_{i=1}^n \Phi \left( \frac{x_i}{\sqrt{\omega_i}} \right)
\]

for \( x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n \). Therefore, \( \tilde{R}_t = \sqrt{C_t} \) and \( Y^{(n)}_t \) are independent with \( Y^{(n)}_t \sim N_n(0, \Sigma_n) \). Then, with the settings \( c_t := E(C_t) \) and \( R_t := \frac{1}{\sqrt{c_t}} \tilde{R}_t \), for \( t \in \mathbb{R}^+ \), the process \( (X^{(n)}_t)_{t \in I} \) defined by

\[
X^{(n)}_t := \frac{1}{\sqrt{C_t}} \tilde{X}^{(n)}_t = R_t Y^{(n)}_t = \frac{1}{\sqrt{C_t}} \tilde{R}_t Y^{(n)}_t = \frac{1}{\sqrt{C_t}} B^{(n)}_t
\]

satisfies the desired conditions of \( \text{Cov}(X^{(n)}_t) = \Sigma_n \) and \( X^{(n)}_t \sim EC_n(0, \Sigma_n, \phi_t) \) where

\[
\phi_t(x) = \int_0^\infty \exp \left( -\frac{1}{2} r^2 x \right) dG_t(r), \text{ for } x \in \mathbb{R},
\]

with \( G_t \) being the distribution function of \( R_t \). Of course, one can specify the process \( C \) to be a continuous time change as discussed above. Then there is a similar argument as in Proposition 8.5.1 that renders \( (\tilde{X}^{(n)}_t)_{t \in I} \) an \( \tilde{F} := (\tilde{F}_t)_{t \geq 0} \) martingale, where \( \tilde{F}_t := FC_t \) for \( t \geq 0 \).

**Remark 8.5.3** Caution must be applied if we want to use time-changed Brownian motions for both the factor process and the process of the idiosyncratic risk vector. As the factor and the idiosyncratic vector need to be independent at any time \( t \), we cannot use
the same time change for both processes. Indeed, the two time changes have to be independent, while they are allowed to agree in distribution. Equality in distribution would then also imply that \( G_{1,t} = G_{2,t} \) for all \( t \in I \), so that the scaling weights for the factor and for the idiosyncratic components agree. However, each of the time-changed Brownian motion processes \((\tilde{X}_t)_{t \in I}\) is a martingale only with respect to the filtration \( \hat{F} = (\hat{F}_t)_{t \geq 0} = (F_{C_t})_{t \geq 0} \), where its own time change \( C \) has been used. Therefore, it does not make any sense to stick to the postulation that the processes \((\tilde{X}_t)_{t \in I}\) remain martingales, but indeed it becomes reasonable to neglect this property and to assure that the processes fit into our setup with \( \mathbb{E}(R_t^2) = 1 \).

### 8.5.3 Dynamic setup via subordinated Lévy processes

In the previous section we have discussed an extension of our model where the static Gaussian part within the mixtures of Normal distributions was made dynamic with the use of a Brownian motion, which is a process with stationary and independent increments. Processes with stationary and independent increments can also be introduced in a natural way for the entire mixture of Normal distributions, which leads to the discussion of infinitely divisible distributions and Lévy processes.

**Definition 8.5.4** A random vector \( X^{(n)} \) is said to be infinitely divisible if for every \( k \in \mathbb{N} \) there are \( k \) independent random vectors \( X_{k,1}^{(n)}, \ldots, X_{k,k}^{(n)} \) such that \( X_{k,j}^{(n)} \overset{d}{=} X_{k,1}^{(n)} \), for all \( j \), and

\[
X^{(n)} \overset{d}{=} X_{k,1}^{(n)} + \ldots + X_{k,k}^{(n)}.
\]

In this case, the distribution function, the density function (in the case of absolute continuity), and the transforms such as the characteristic function and the Laplace-Stieltjes transform of \( X^{(n)} \) are called infinitely divisible.

**Definition 8.5.5** A process \((X_t^{(n)})_{t \in \mathbb{R}^+}\) on \((\Omega, \mathcal{F}, \mathbb{P}, F)\) is called a Lévy process if \((X_t^{(n)})_{t \in \mathbb{R}^+}\) is a càdlàg\(^4\) \( \mathbb{F} \)-adapted process such that \( X_0^{(n)} = 0 \), where

1. the increments are independent, i.e. for all \( 0 \leq s \leq t \) the variable \( X_t^{(n)} - X_s^{(n)} \) is independent of \( \mathcal{F}_s \),

2. the increments are stationary, i.e. for all \( 0 \leq s \leq t \) we have \( X_t^{(n)} - X_s^{(n)} \overset{d}{=} X_{t-s}^{(n)} \), and where

3. the process is continuous in probability, i.e.

\[
\forall \varepsilon > 0, \quad \lim_{t \to s} \mathbb{P} \left( \| X_t^{(n)} - X_s^{(n)} \| \geq \varepsilon \right) = 0.
\]

\(^4\) This is an acronym and stands for the French expression continu à droite, limite à gauche, which means that the process is continuous from the right and possesses limits from the left at any point \( t \).
Proposition 8.5.6 Let \((X^{(n)}_t)_{t \in \mathbb{R}^+}\) be a Lévy process on \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\). Then, for every \(t \in \mathbb{R}^+\), \(X^{(n)}_t\) is infinitely divisible. Conversely, if there is a random vector \(Y^{(n)}\) which is infinitely divisible then there exists a Lévy process \((X^{(n)}_t)_{t \in \mathbb{R}^+}\) such that \(X^{(n)}_t \overset{d}{=} Y^{(n)}\).

Proof: This proposition is essentially Proposition 3.1 in [CT04] or Proposition 3.1 and 3.2 in [SvH04].

In the theory of Lévy processes and infinitely divisible distributions, Laplace-Stieltjes transforms and especially characteristic functions play a central role mainly due to the Lévy-Khintchine representation that we will state in a moment.

Definition 8.5.7 For a random vector \(X^{(n)}\) with distribution function \(F\)

1. the characteristic function \(\psi_F : \mathbb{R}^n \to \mathbb{C}\) is given via
   \[
   \psi_F(u) := \mathbb{E}(\exp(iu^T X^{(n)})) = \int_{\mathbb{R}^n} \exp(iu^T x) F(dx),
   \]
   for \(u \in \mathbb{R}^n\), and

2. the Laplace-Stieltjes transform \(\pi_F : \mathbb{R}^+ \to (0, \infty)\) is given via
   \[
   \pi_F(u) := \mathbb{E}(\exp(-u^T X^{(n)})) = \int_{\mathbb{R}^n} \exp(-u^T x) F(dx),
   \]
   for \(u \in \mathbb{R}^+\).

The following is essentially a combination of Lemma 7.5 and Lemma 7.6 in [Sat99].

Proposition 8.5.8 If \(X^{(n)}\) is an infinitely divisible random vector with distribution function \(F\) and characteristic function \(\psi_F\), then there is a unique continuous function \(\Psi_F : \mathbb{R}^n \to \mathbb{C}\) such that \(\Psi_F(0) = 0\) and \(\psi_F(z) = \exp(\Psi_F(z))\) for \(z \in \mathbb{R}^n\). Furthermore, for any \(k \in \mathbb{N}\) there is a unique continuous function \(\psi_{F,k} : \mathbb{R}^n \to \mathbb{C}\) such that \(\psi_{F,k}(0) = 1\) and \(\psi_{F,k}(z)^k = \psi_F(z)\) for all \(z \in \mathbb{R}^n\). Additionally, we have the relationship of \(\psi_{F,k}(z) = \exp(\Psi_F(z)/k)\) for \(z \in \mathbb{R}^n\) and \(k \in \mathbb{N}\).

If \(X^{(n)}\) is an infinitely divisible random vector with distribution function \(F\) and \(\Psi_F\) is the unique function out of the previous proposition, then for \(t \in \mathbb{R}^+\) we define the function \(\psi^t_F : \mathbb{R}^n \to \mathbb{C}\) via the setting \(\psi^t_F(z) := \exp(t \Psi_F(z))\), for \(z \in \mathbb{R}^n\). According to Lemma 7.9 in [Sat99], these functions \(\psi^t_F\) are again characteristic functions of some distribution functions, which we shall denote by \(F^{\ast t}\) (then, \(\psi^t_F = \psi_{F^{\ast t}}\), \(t \in \mathbb{R}^+\), and they are again infinitely divisible. The corresponding Laplace-Stieltjes transform also satisfies \(\pi_{F^{\ast t}} = \pi^t_F\), for \(t \in \mathbb{R}^+\), where the latter expression \(\pi^t_F\) is given in the usual way via \(\pi^t_F(u) = \exp(t \log(\pi_F(u)))\), \(t \in \mathbb{R}^+\) and \(u \in \mathbb{R}^{+n}\), as \(\pi : \mathbb{R}^{+n} \to \mathbb{R}_0^+\).
Having this to hand, there is an important way of characterising Lévy processes where the characterisation is based on the following theorem, which we have adapted from Theorem 3.2 in [CT04]:

**Theorem 8.5.9 (The Lévy-Khintchine representation)**

Let \((X_t^{(n)})_{t \in \mathbb{R}^+}\) be a Lévy process on \((\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F})\). Then, for any \(t \in \mathbb{R}^+\), the characteristic function of \(X_t^{(n)}\) can be represented as

\[
\mathbb{E}(\exp(iu^T X^{(n)})) = \exp(t \Psi(u)),
\]

with \(\Psi(u) = -\frac{1}{2} u^T A u - i\gamma^T u + \int_{\mathbb{R}^n} \left( \exp(iu^T x) - 1 - iu^T x 1_{\{\|x\|\leq 1\}} \right) \nu(dx),\)

for \(u \in \mathbb{R}^n\), where \(A \in \mathbb{R}^{n \times n}\) is a symmetric positive semi-definite matrix, \(\gamma \in \mathbb{R}^n\) and \(\nu\) is a positive Radon measure on \(\mathbb{R}_0^n\) satisfying

\[
\int_{\|x\|\leq 1} \|x\|^2 \nu(dx) < \infty \quad \text{and} \quad \int_{\|x\|\geq 1} \nu(dx) < \infty.
\]

The measure \(\nu\) is called the Lévy measure of \((X_t^{(n)})_{t \in \mathbb{R}^+}\), the matrix \(A\) the Gaussian covariance matrix, and \((A, \nu, \gamma)\) is named the characteristic triplet of the process \((X_t^{(n)})_{t \in \mathbb{R}^+}\). The components in the triplet uniquely determine the distribution of the process. Note, that the so-called characteristic exponent \(\Psi\) in the Lévy-Khintchine representation is independent of \(t\). As \(\Psi\) represents the cumulant generating function of \(X_1^{(n)}\), it is completely determined by the infinitely divisible distribution of \(X_1^{(n)}\). In the case of our mixtures of Normal distributions where \(X^{(n)} = R Y^{(n)} \sim F\), such that \(R, Y^{(n)}\) are independent, such that the univariate scaling variable is non-negative \(R \geq 0\) with \(R \sim G\) and such that \(Y^{(n)} \sim N_n(0, I_n)\), the characteristic function attains the following form:

\[
\psi_F(u) = \mathbb{E}(\exp(iu^T R Y^{(n)})) = \int_0^\infty \mathbb{E}(\exp(iru^T Y^{(n)})) dG(r) = \int_0^\infty \exp \left(- \frac{r^2 u^T u}{2} \right) dG(r) = \int_0^\infty \exp(-ru^T u) d\hat{G}(r) = \pi_G(u^T u),
\]

for \(u \in \mathbb{R}^n\), where the distribution function \(\hat{G}\) is defined via the setting \(\hat{G}(r) := G(\sqrt{2r})\) for \(r \in \mathbb{R}^+\).

**Proposition 8.5.10** If \(F\) is a mixture of Normal distributions with scaling distribution \(G\) and distribution function \(\hat{G}\) as just defined in the context of the previous set of equations, then if \(\hat{G}\) is infinitely divisible, so is \(F\).

\(^5\) For the Lévy-Khintchine theory and the Lévy-Itô decomposition see also [Ber96].
The continuous-time case

**Proof:** Under the assumption that $\hat{G}$ is infinitely divisible, Theorem 2.8 in Chapter III of [SvH04] guarantees that the Laplace-Stieltjes transform $\pi_{\hat{G}}$ does not have any zeroes, so that $\psi_F(u) = \pi_{\hat{G}}(u^T u) \neq 0$, for all $u \in \mathbb{R}^n$. According to Proposition 8.5.8 and the remarks thereafter we obtain

$$\psi_F^t(u) = \pi_{\hat{G}}^t(u^T u) = \int_0^\infty \exp(-ru^T u) d\hat{G}^s(r),$$

for all $u \in \mathbb{R}^n$ and $t \in \mathbb{R}^+$, which entails that $\psi_F^t$ is a characteristic function of a random vector $R^s Y^{(n)}$ following a mixture of Normal distributions with the univariate scaling variable $R^s$ satisfying $R^s \sim G^s$, where $G^s(r) := \hat{G}^s(r^2/2)$, for $r, t \in \mathbb{R}^+$. Hence, it follows that $F$ is infinitely divisible as, with that, $\psi_F^{1/k}$ is a characteristic function of a random vector $X^{(n)}_{k,1}$ and as $\psi_F(u) = (\psi_F^{1/k}(u))^k$, for all $u \in \mathbb{R}^n$ and $k \in \mathbb{N}$.

Thus, if we want to construct a multivariate Lévy process with the use of mixtures of Normal distributions and by dint of Proposition 8.5.6, it is sufficient to require that $\hat{G}$ is infinitely divisible.

The previous proposition serves us as a motivation for using appropriate Lévy processes at relevant places in the previous section. For instance, we could specify the general process $C$ in the subordinated Brownian motion setting (8.11) to be a subordinator, that is a one-dimensional increasing Lévy process, which is thus also non-negative at any point in time almost surely. Let us assume that such a subordinator $(Z_t)_{t \in \mathbb{R}^+}$ is given with characteristic triplet $(0, \rho, b)$ and thus with $l(u) = bu + \int_0^\infty (\exp(u \pi) - 1) \rho(dx)$, for $u \leq 0$, where the function $l : (-\infty, 0] \to \mathbb{R}$ defined via $l(u) := \log(\mathbb{E}(\exp(u Z_1)))$ is called the Laplace exponent of $(Z_t)_{t \in \mathbb{R}^+}$. Again, by $(B_t^{(n)})$ we denote an $n$-dimensional $\mathbb{F}$-Brownian motion with zero-drift and covariance matrix $\Sigma_n = \text{diag}(\omega_1, \ldots, \omega_n) \in \mathbb{R}^{n \times n}$, $\omega_i > 0$. By virtue of Theorem 4.2 in [CT04], the subordinated process $(\tilde{X}_t^{(n)})_{t \in \mathbb{R}^+}$ obtained via the setting

$$\tilde{X}_t^{(n)} := B_{Z_t}^{(n)}, \quad t \in \mathbb{R}^+,$$

also becomes a Lévy process, where the characteristic function of $\tilde{X}_t^{(n)}$ becomes

$$\mathbb{E} \left( \exp(\imath u^T \tilde{X}_t^{(n)}) \right) = \exp(tl(-\frac{1}{2} u^T \Sigma_n u)), $$

for all $t \geq 0$ and all $u \in \mathbb{R}^n$. Once more, we set $\tilde{z}_t := \mathbb{E}(Z_t) > 0$,

$$\tilde{R}_t := \sqrt{Z_t}, \quad R_t := \frac{1}{\sqrt{\tilde{z}_t}} \tilde{R}_t \quad \text{and} \quad Y_t^{(n)} := \left( B_{1,z_t}, \ldots, B_{n,z_t} \right)^T,$$

for $t \in \mathbb{R}^+$. Then the process $(X_t^{(n)})_{t \in \mathbb{R}^+}$ defined by

$$X_t^{(n)} := R_t Y_t^{(n)} = \frac{1}{\sqrt{\tilde{z}_t}} \tilde{X}_t^{(n)}$$
is no longer a Lévy process but, as we have seen in the previous section, satisfies the desired conditions of

\[ \text{Cov}(X_t^{(n)}) = \Sigma_n \quad \text{and} \quad X_t^{(n)} \sim \text{EC}_n(0, \Sigma_n, \phi_t), \]

where

\[ \phi_t(x) = \int_0^\infty \exp\left(-\frac{1}{2}r^2x\right) dG_{2,t}(r), \quad \text{for} \ x \in \mathbb{R}, \]

\( G_{2,t} \) being the distribution function of \( R_t \). Even though we lose the Lévy process property when going from \( (\tilde{X}_t^{(n)})_{t \in I} \) to \( (X_t^{(n)})_{t \in I} \) and also the stationarity of the distributions of the increments, all we need for the computation of the expected losses is the distribution function of \( R_t \). Therefore, we only have to make sure that we select a Lévy process \( (Z_t)_{t \in I} \) where the distribution functions of the increments are readily available.

### 8.5.4 Is the construction of a Lévy process with the Exp-Exp Law from the static model possible?

In the static model that we have discussed in the previous chapters, we developed a new distribution function, which we have called the multivariate Exp-Exp Law (see Section 5.3.2), that proved to be particularly useful for the pricing of CDOs.

By construction, the multivariate Exp-Exp Law belongs to the subclass of mixtures of Normal distributions which display the property of tail-dependence. Recall that for a random vector \( X^{(n)} \) of the form \( X = RY^{(n)} \) where \( Y^{(n)} \sim N_n(0, \Sigma) \) which follows this Exp-Exp Law, the distribution function \( G \) of the scaling variable \( R \) possesses the density function \( g \) with

\[ g(r) = \frac{\alpha}{c_\alpha I} \left( \frac{r}{c_\alpha} \right)^{-2\alpha-1} l_\alpha \left( \frac{r}{c_\alpha} \right), \quad \text{for} \ r > 0 \quad \text{and} \quad \alpha > 1, \]

with \( l_\alpha(y) = \exp(-y^{-\alpha}) \cdot \exp(-\exp(-y^{-\alpha})) \), for any \( \alpha \in \mathbb{R} \), and appropriate constants \( I \) and \( c_\alpha \) (cf. Section 5.3.2).

According to Proposition 8.5.10, we can check whether \( \tilde{G} \) defined via \( \tilde{G}(r) := G(\sqrt{2r}) \), for \( r \in \mathbb{R}^+ \), is infinitely divisible in order to deduce that the mixture of Normal distributions of \( X^{(n)} \) is infinitely divisible. Since if a random variable \( X \) is infinitely divisible, then so too is \( aX \), for \( a \in \mathbb{R} \), it suffices to analyse the distribution function \( \tilde{G} \) with \( \tilde{G}(r) := G(c_\alpha \sqrt{r}) \) with respect to infinite divisibility, where the density function \( \tilde{g} \) becomes

\[ \tilde{g}(r) := \frac{c_\alpha}{2\sqrt{r}} g(c_\alpha \sqrt{r}) \]

\[ = \frac{\alpha/2}{I} r^{-\alpha-1} l_{\alpha/2}(r) \]

\[ = \frac{\beta}{I} r^{-2\beta-1} l_{\beta}(r) \]

\[ = \frac{\beta}{I} r^{-2\beta-1} \exp(-r^{-\beta}) \exp(-\exp(-r^{-\beta})), \]
The continuous-time case

for \( r > 0 \) and \( \beta := \alpha/2 > 1/2 \).

The characteristic function of \( \hat{G} \) then becomes

\[
\psi_{\hat{G}}(z) = \int_{0}^{\infty} \exp(izs)\hat{g}(s) \, ds
\]

\[
= \frac{\beta}{T} \int_{0}^{\infty} \exp(izs)\exp(-s^{-\beta})\exp(-\exp(-s^{-\beta})) \cdot s^{-2\beta-1} \, ds
\]

\[
= \frac{1}{T} \int_{0}^{\infty} \exp(izs^{-1/\beta})\exp(-s)\exp(-\exp(-s)) \cdot s \, ds,
\]

for \( z \in \mathbb{R} \). Likewise, the Laplace-Stieltjes transform becomes

\[
\pi_{\hat{G}}(z) = \int_{0}^{\infty} \exp(-zs)\hat{g}(s) \, ds
\]

\[
= \frac{\beta}{T} \int_{0}^{\infty} \exp(-zs)\exp(-s^{-\beta})\exp(-\exp(-s^{-\beta})) \cdot s^{-2\beta-1} \, ds
\]

\[
= \frac{1}{T} \int_{0}^{\infty} \exp\left(-zs^{-1/\beta}\right)\exp(-s)\exp(-\exp(-s)) \cdot s \, ds,
\]

for \( z \in \mathbb{R}^+ \). Unfortunately, we were not able to give a closed-form expression for either \( \psi_{\hat{G}} \) or for \( \pi_{\hat{G}} \) where we could directly see whether \( \hat{G} \) is infinitely divisible or not.

There are several reasonings that can principally be used to check whether or not a given distribution function is infinitely divisible. In the following we want to state some of them and give some indications on how these reasonings behavied for our distribution function.

The most direct path to follow is to analyse whether one can derive a \( k \)-th convolution root for \( \hat{G} \) that would again be of the same form as \( \hat{G} \) but possibly with different parameters. This procedure can be well applied for instance for Gaussian distributions (see e.g. Examples 7.2 in [Sat99]). Neither our attempts via the Laplace-Stieltjes transform nor via finding a closed form for the convolution gave us any insight on how to advance further in this direction.

From the theory of infinitely divisible distributions we know the following about the zeroes of the characteristic function:

**Theorem 8.5.11** If \( F \) is an infinitely divisible distribution function on \( \mathbb{R}^d \), then its characteristic function \( \psi_F \) has no zero, that is \( \psi_F(z) \neq 0 \) for any \( z \in \mathbb{R}^d \).

**Proof:** See Lemma 7.5 in [Sat99]. \( \square \)

Unfortunately, this property of having no zeroes is not sufficient to prove the infinite divisibility of a distribution function. The binomial distribution \( B(n,p) \) provides us with a counterexample, as its characteristic function does not have a zero if \( p \neq 1/2 \), but it is not infinitely divisible (see the remark after Lemma 7.5 and Examples 7.2 in
8.5. The continuous-time case

[Sat99]). If the theorem provided us with an equivalence we would have proved the infinite divisibility of mixtures of Normal distributions directly, as in the set of Equations (8.12) we see that the characteristic function is strictly positive on $\mathbb{R}^n$. Yet, we were not able to derive a contradiction for the characteristic function of $\hat{G}$ on $\mathbb{R}^+$ to the necessary property for infinite divisibility of the previous theorem.

**Lemma 8.5.12** If $(M_k)_{k\in\mathbb{N}}$ is a sequence of infinitely divisible distribution functions that converges towards a distribution function $F$, then $F$ is also infinitely divisible.

**Proof:** See Lemma 7.8 in [Sat99].

Again, we were not able to make use of this lemma as we could not approximate the distribution function $\hat{G}$ by an appropriate sequence where the infinite divisibility of its terms was assured.

As $G$ is a distribution function on $(0, \infty)$ we did continue with the following two theorems which essentially stem from [Sat99]:

**Theorem 8.5.13** Let $G$ be an absolutely continuous distribution function on $(0, \infty)$ with density function $g$. If $g$ is positive on $(0, \infty)$ and log-convex in the sense that $\log(g(\cdot))$ is convex on $(0, \infty)$, then $G$ is infinitely divisible.

**Proof:** See Theorem 51.4 in [Sat99].

Unfortunately, $\hat{g}$ is not log-convex on $(0, \infty)$:

$$\frac{\partial^2}{\partial r^2} \log(\hat{g}(r)) = -r^{-\beta - 2} \left( \beta^2 + \beta - r^\beta (1 + 2\beta) \right)$$

$$+ \exp(-r^\beta)(\beta r^{-\beta - 2} - \beta^2 r^{-2\beta - 2} + \beta^2 r^{-\beta - 2})$$

$$\longrightarrow -\infty \quad \text{as } r \to 0^+,$$

and as a consequence of which $\frac{\partial^2}{\partial r^2} \log(\hat{g}(r))$ cannot be non-negative for all $r > 0$ which would be equivalent to $\hat{g}$ being log-convex on $(0, \infty)$. Another theorem going into the same direction as Theorem 8.5.13 is the following:

**Theorem 8.5.14** Let $G$ be an absolutely continuous distribution function on $(0, \infty)$ with density function $g$. If $g$ is positive and completely monotone on $(0, \infty)$, that is $g \in C^\infty(0, \infty)$ and $(-1)^k \frac{d^k}{dr^k} g(\cdot) \geq 0$ on $(0, \infty)$, for all $k \in \mathbb{N}$, then $G$ is infinitely divisible.

**Proof:** See Theorem 51.6 in [Sat99].
Unfortunately, $\hat{g}$ is not completely monotone:

$$\frac{\partial}{\partial r} \hat{g}(r) = -\hat{g}(r) r^{-1} \left(2\beta + 1 + \beta r^{-\beta} (\exp(-r^{-\beta}) - 1)\right)$$

$$\begin{cases} > 0, & \text{for } r > 0 \text{ sufficiently small;} \\ < 0, & \text{for } r > 0 \text{ sufficiently large,} \end{cases}$$

as $-\hat{g}(r) r^{-1} < 0$ for all $r \in (0, \infty)$ and as $r^{-\beta} \to \begin{cases} \infty, & \text{for } r \to 0^+; \\ 0, & \text{for } r \to \infty, \end{cases}$ yields

$$2\beta + 1 + \beta r^{-\beta} (\exp(-r^{-\beta}) - 1) \to \begin{cases} -\infty, & \text{for } r \to 0^+; \\ 2\beta + 1 \geq 2 > 0, & \text{for } r \to \infty. \end{cases}$$

Therefore, $(-1)^k \frac{d^k}{dr^k} \hat{g}(r) \geq 0$ does not hold for $k = 1$ and a sufficiently small $r \in (0, \infty)$.

**Lemma 8.5.15** Let $G$ be a distribution function on $(0, \infty)$ with Laplace-Stieltjes transform $\pi_G$. Then the following statements are equivalent:

1. $\pi_G$ is infinitely divisible;
2. $\pi_t^G$ is completely monotone for all $t > 0$;
3. the function $l_G : (0, \infty) \to \mathbb{R}$ with $l_G(r) := -\log(\pi_G(r))$, for $r > 0$, is a Bernstein function, i.e. $l_G \in C^\infty(0, \infty)$, $l(\cdot) \geq 0$ and $(-1)^k \frac{d^k}{dr^k} l_G(\cdot) \leq 0$ on $(0, \infty)$, for all $k \in \mathbb{N}_0$.

**Proof:** For the equivalence of the first and the second statement we refer to Proposition 2.3 in [SvH04]. According to Proposition 9.2 in [BF75], $l_G$ is a Bernstein function iff $l_G(\cdot) \geq 0$ and if the function $\exp(-tl_G(\cdot))$ is completely monotone, for all $t > 0$. Observe that $l_G(r) = -\log(\pi_G(r)) \geq 0$ is equivalent to $\pi_G(r) \leq 1$ which is always satisfied as $\pi_G(r) = \int_0^\infty \exp(-rt)dG(t) \leq \int_0^\infty 1 dG(t) = 1$ with $r > 0$. Thus, we only have to prove the equivalence of the complete monotonicity of $\exp(-tl_G(\cdot))$, for all $t > 0$, and of the second statement. Yet, it holds that $\exp(-tl_G(r)) = \pi_t^G(r)$, for all $t, r > 0$, by consequence of which the proof is accomplished.

The second property that $\pi_t^G$ is completely monotone, for all $t > 0$, is very hard to prove analytically in our context. Similarly difficult is proving that $l_G$ is a Bernstein function, as the relevant property demands that one looks at all derivatives of $l_G$. In the hope to obtain a contradiction already for a derivative of low order or in order to at least gain some insight we analysed this property numerically up to the fifth derivative. The graphical output of the numerical computations can be seen in Figures 8.1 - 8.3.
Figure 8.1: Graphs of the function $l^*_G$ and its first derivative for $\beta = 0.6, 1.0, 2.0, 3.0$ and 4.0.

Figure 8.2: Graphs of the second and the third derivative of the function $l^*_G$ for $\beta = 0.6, 1.0, 2.0, 3.0$ and 4.0.
While the function $l_\hat{G}$ itself, as well as its first, its third and its fifth derivatives stay non-negative for the chosen 5 different values of the parameter $\beta$, the second and the fourth derivatives assume only non-positive values for the given set of parameters $\beta$. This property is entirely in line with the definition of a Bernstein function, and strongly suggests that $\hat{G}$ is infinitely divisible. However, a mathematical proof of this conjecture remains still to be found.
8.5. *The continuous-time case*
Appendix A

Supplementary lemma

The following classical result is used in Chapter 3:

Lemma A.1.1 If a continuous function $f : \mathbb{R} \to \mathbb{R}$ satisfies

$$f((u^2 + v^2)^{1/2}) = f(u)f(v), \quad (A.1)$$

for all $u, v \in \mathbb{R}$, as well as $|f(u)| \leq 1$, for all $u \in \mathbb{R}$ and $f(0) = 1$, then there exists an $a \in \mathbb{R}^+$ such that

$$f(u) = \exp \left( -\frac{1}{2} au^2 \right),$$

for all $u \in \mathbb{R}$ and $a = -2 \log(f(1)) \geq 0$.

Proof: Bringing $f(0) = 1$ and the defining relation (A.1) together, we obtain

$$f(u) = f(u)f(0) = f((u^2 + 0^2)^{1/2}) = f(|u|),$$

so that $f(u) = f(-u)$, for all $u \in \mathbb{R}$. Via induction one can deduce from (A.1) that for any $n \in \mathbb{N}$ and any real numbers $u_1, \ldots, u_n$ we have

$$f \left( \left( \sum_{i=1}^{n} u_i^2 \right)^{1/2} \right) = \prod_{i=1}^{n} f(u_i),$$

so that also

$$f(\sqrt{n}u) = f(\sqrt{n}|u|) = f^n(u), \quad (A.2)$$

for all $u \in \mathbb{R}$ and any $n \in \mathbb{N}$. As $f(0) = 1$ and as $f$ is continuous there exists a $\delta > 0$ such that $f(u) > 0$ for all $t \in U_\delta(0)$. But as there exists an $n \in \mathbb{N}$ such that $\frac{1}{\sqrt{n}} \in U_\delta(0)$ we conclude that

$$f(1) = f \left( \sqrt{n} \frac{1}{\sqrt{n}} \right) = f^n \left( \frac{1}{\sqrt{n}} \right) > 0.$$
Thus, \( a := -2 \log(f(1)) \) is well defined and non-negative, as \( f(1) = |f(1)| \leq 1 \) according to the assumptions. Equation (A.2) further implies that

\[
f(u) = f^{1/n}(\sqrt{n}u) = f^{1/n} \left( \sqrt{k} \left( \frac{\sqrt{n}}{\sqrt{k}} u \right) \right) = f^{k/u} \left( \sqrt{\frac{n}{k}} u \right),
\]

therefore

\[
f \left( \sqrt{\frac{n}{k}} u \right) = f^{n/k}(u)
\]

for all \( k, n \in \mathbb{N} \) and \( u \in \mathbb{R} \), and henceforth also

\[
f (\sqrt{q}u) = f^q(u),
\]

for all \( q \in \mathbb{Q}_0^+ \) and \( u \in \mathbb{R} \). Due to the continuity of \( f \) we even have that

\[
f (tu) = f^{t^2}(u),
\]

for all \( t \in \mathbb{R}_0^+ \) and \( u \in \mathbb{R} \). From this follows that

\[
f(t) = f(t \cdot 1) = f^{t^2}(1) = \exp \left( -\frac{1}{2} at^2 \right),
\]

for all \( t \in \mathbb{R}_0^+ \). For \( t < 0 \) we have

\[
f(t) = f(-t \cdot (-1)) = f^{(-t)^2}(-1) = f^{t^2}(1) = \exp \left( -\frac{1}{2} at^2 \right)
\]

and as \( f(0) = 1 = \exp(0) \), the proof is established. \( \square \)
Supplementary data

The next two tables give the prices for various indices on the iTraxx Europe Series 3 reference portfolio on every second Wednesday from January to May 2006. The first table holds the 5 year tranched iTraxx values which mature on the 20th of September 2010, and the second table shows the average of all spreads of the CDS entities in the reference portfolio with the respective time to maturity (3, 5, 7, 10 years).

Table A.1: The prices of the iTraxx Europe Series 3 tranches on the different dates. All prices are denoted in basis points per annum except for the equity tranche values, which are given in percent.

<table>
<thead>
<tr>
<th>Date</th>
<th>0% - 3%</th>
<th>3% - 6%</th>
<th>6% - 9%</th>
<th>9% - 12%</th>
<th>12% - 22%</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.01.2006</td>
<td>22.96</td>
<td>68.50</td>
<td>21.95</td>
<td>11.16</td>
<td>5.87</td>
</tr>
<tr>
<td>25.01.2006</td>
<td>23.45</td>
<td>60.25</td>
<td>18.92</td>
<td>8.62</td>
<td>5.50</td>
</tr>
<tr>
<td>08.02.2006</td>
<td>23.04</td>
<td>65.69</td>
<td>20.43</td>
<td>10.09</td>
<td>4.66</td>
</tr>
<tr>
<td>22.02.2006</td>
<td>22.48</td>
<td>65.67</td>
<td>19.80</td>
<td>10.09</td>
<td>4.59</td>
</tr>
<tr>
<td>08.03.2006</td>
<td>22.60</td>
<td>61.25</td>
<td>18.58</td>
<td>10.68</td>
<td>2.94</td>
</tr>
<tr>
<td>22.03.2006</td>
<td>22.03</td>
<td>60.32</td>
<td>17.75</td>
<td>10.43</td>
<td>2.95</td>
</tr>
<tr>
<td>05.04.2006</td>
<td>20.05</td>
<td>55.56</td>
<td>16.66</td>
<td>9.18</td>
<td>2.55</td>
</tr>
<tr>
<td>19.04.2006</td>
<td>19.65</td>
<td>54.30</td>
<td>16.48</td>
<td>9.01</td>
<td>2.42</td>
</tr>
<tr>
<td>03.05.2006</td>
<td>18.03</td>
<td>36.99</td>
<td>12.06</td>
<td>5.62</td>
<td>2.41</td>
</tr>
<tr>
<td>17.05.2006</td>
<td>12.54</td>
<td>18.30</td>
<td>5.04</td>
<td>2.36</td>
<td>2.09</td>
</tr>
<tr>
<td>31.05.2006</td>
<td>15.30</td>
<td>28.79</td>
<td>7.91</td>
<td>3.61</td>
<td>2.59</td>
</tr>
</tbody>
</table>

Table A.2: The averages of all annual CDS spreads for the different maturities and on the different dates. All prices are denoted in basis points.

<table>
<thead>
<tr>
<th>Date</th>
<th>3Y-CDS</th>
<th>5Y-CDS</th>
<th>7Y-CDS</th>
<th>10Y-CDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.01.2006</td>
<td>25.43</td>
<td>41.12</td>
<td>55.35</td>
<td>63.40</td>
</tr>
<tr>
<td>25.01.2006</td>
<td>24.80</td>
<td>41.12</td>
<td>56.58</td>
<td>64.04</td>
</tr>
<tr>
<td>08.02.2006</td>
<td>25.82</td>
<td>41.85</td>
<td>56.93</td>
<td>65.05</td>
</tr>
<tr>
<td>22.02.2006</td>
<td>25.02</td>
<td>41.33</td>
<td>56.66</td>
<td>65.08</td>
</tr>
<tr>
<td>08.03.2006</td>
<td>23.17</td>
<td>39.52</td>
<td>54.56</td>
<td>63.36</td>
</tr>
<tr>
<td>22.03.2006</td>
<td>22.70</td>
<td>38.78</td>
<td>54.71</td>
<td>64.45</td>
</tr>
<tr>
<td>05.04.2006</td>
<td>19.31</td>
<td>34.49</td>
<td>47.69</td>
<td>57.28</td>
</tr>
<tr>
<td>19.04.2006</td>
<td>19.60</td>
<td>34.65</td>
<td>49.75</td>
<td>58.28</td>
</tr>
<tr>
<td>03.05.2006</td>
<td>15.76</td>
<td>29.55</td>
<td>43.31</td>
<td>50.98</td>
</tr>
<tr>
<td>17.05.2006</td>
<td>18.70</td>
<td>32.69</td>
<td>46.86</td>
<td>54.55</td>
</tr>
<tr>
<td>31.05.2006</td>
<td>19.60</td>
<td>34.02</td>
<td>48.15</td>
<td>55.80</td>
</tr>
</tbody>
</table>
**List of Abbreviations**

\( \mathbb{R} \) denotes the ordinary real line \((-\infty, +\infty)\).

\( \mathbb{R}^+ \) denotes the extended real line \([-\infty, +\infty]\).

\( \mathbb{R}^0_+ \) denotes the positive real line \((0, +\infty)\).

\( \mathbb{Q} \) denotes the rational numbers.

\( \mathbb{Q}^+ \) denotes the non-negative rational numbers \( \mathbb{Q} \cap [0, +\infty) \).

\( \mathbb{Q}^+_0 \) denotes the positive rational numbers \( \mathbb{Q} \cap (0, +\infty) \).

\( \mathbb{N} \) denotes the non-negative integers \( \{0, 1, 2, \ldots\} \).

\( \mathbb{N}_0 \) denotes the positive integers \( \{1, 2, 3, \ldots\} \).

\( \mathbb{R}^n \) denotes the extended real \( n \)-space \( \mathbb{R} \times \cdots \times \mathbb{R} \), where \( n \in \mathbb{N}_0 \).

\( a \leq b \) (\( a < b \), respectively), where \( a = (a_1, \ldots, a_n) \), \( b = (b_1, \ldots, b_n) \) in \( \mathbb{R}^n \), means \( a_k \leq b_k \) (\( a_k < b_k \), respectively) for all \( k \).

\([a, b] \), where \( a = (a_1, \ldots, a_n) \), \( b = (b_1, \ldots, b_n) \) in \( \mathbb{R}^n \) with \( a \leq b \), denotes the \( n \)-box \( B = [a_1, b_1] \times \cdots \times [a_n, b_n] \), i.e. the Cartesian product of \( n \) closed intervals.

The vertices of an \( n \)-box are the points \( c = (c_1, \ldots, c_n) \), where each \( c_k \) equals either \( a_k \) or \( b_k \).

The unit \( n \)-cube \( I^n \) is the product \( I \times \cdots \times I \) where \( I = [0, 1] \).

Let \( A \in \mathbb{R}^{m \times n} \), then \( A^T \in \mathbb{R}^{n \times m} \) denotes the transposed matrix of \( A \).

\( I_n \) denotes the \( n \)-dimensional unity matrix, for \( n \in \mathbb{N}_0 \).

\( (x)^+ = \max\{x, 0\} \).

\(([0, 1], \mathcal{B}([0, 1]), \lambda_{[0,1]}) \) denotes the Borel measure space, with \( \mathcal{B}([0, 1]) \) being the \( \sigma \)-algebra of Borel sets in \([0, 1]\), and \( \lambda_{[0,1]} \) the Lebesgue measure restricted on \([0, 1]\).

\( \Phi \) denotes the standard Normal distribution function and \( \varphi \) its density function.

\( d \) stands for equality in distribution.

\( f(x) \sim g(x) \) if \( x \to a \), where \( f, g \) are appropriate functions:

this means that \( \lim_{x \to a} \frac{f(x)}{g(x)} = 1 \).

\( X \sim F \), where \( X \) is a random variable and \( F \) a distribution function:

this means that \( X \) has distribution \( F \).

"iff" is the short form of "if and only if", "df" is the one of "degrees of freedom".
Bibliography


List of Tables

2.1 Comparison of tranched iTraxx Market values and one-factor Gaussian model prices for the 6th of April 2006. .................................................. 23

4.1 The implied correlations for the iTraxx Europe Series 3 tranches on the 6th of April 2006. ................................................................. 59

4.2 The base correlations for the five iTraxx Europe Series 3 tranches on the 6th of April 2006. ................................................................. 60

7.1 Simulation study: CDO tranche spreads within the Gaussian model. .... 127

7.2 Simulation study: required correlation parameters for various models to match equity tranche prices. .................................................. 128

7.3 Simulation study: resulting prices for mezzanine tranche. ................. 128

7.4 Simulation study: resulting prices for senior tranche. ...................... 129

7.5 The market values of the iTraxx Europe 3 series tranches on the 6th of April 2006. ................................................................. 130

7.6 Tranche prices within various specifications of the model based on the multivariate t-distribution. .................................................. 131

7.7 Tranche prices within various specifications of the model based on the multivariate Exp-Exp Law. .................................................. 132

7.8 Tranche prices within various specifications of the models based on the Power Law and on the Power Log Law. .................................. 135

A.1 The prices of the iTraxx Europe Series 3 tranches on the different dates. 181

A.2 The averages of the annual CDS spreads for different maturities and different dates. .................................................. 181
List of Tables
List of Figures

2.1 Typical structure of a CDO. .................................................. 19

4.1 The implied and the base correlations curves for the five iTraxx Europe
Series 3 tranches on the 6th of April 2006. ............................... 60

5.1 Scatter plot of latent variables vector for Gaussian model and for model
based on the multivariate t-distribution. ..................................... 71

5.2 Tail-dependence coefficient $\lambda$ as a function of the tail index $\alpha$. ....... 76

5.3 Densities of the different mixing distribution functions. ...................... 87

5.4 Scatter plots for the two-dimensional vector $(S_1, S_2)$ following different
one-factor models with factor loadings $\beta_1 = \beta_2 = \sqrt{0.6}$. .......... 89

6.1 The densities and distribution functions of the relative limit portfolio loss
distributions for the various Gaussian mixture models. .................. 105

7.1 Schematic view of $\Theta_i$. ....................................................... 109

7.2 The implied and the base correlations of the various elliptical distributions
models, which were calibrated to iTraxx Europe series 3 data. ............ 133

7.3 The correlation parameters $\rho$ within the various models for a perfect fit
to the iTraxx equity tranche prices over time. ............................. 136

7.4 The sums of absolute errors of the calibrated models over time. .......... 136

8.1 Graphs of the function $l_{\hat{G}}$ and its first derivative for various values of $\beta$. 176

8.2 Graphs of the second and third derivative of $l_{\hat{G}}$ for various values of $\beta$. 176

8.3 Graphs of the fourth and fifth derivative of $l_{\hat{G}}$ for various values of $\beta$. 177
Zusammenfassung

Das Ziel dieser Arbeit ist es, eine neue Gruppe von Faktormodellen einzuführen, die für die Bewertung von auf Portfolios basierenden Kreditderivaten verwendet werden und die die Beobachtungen auf den Finanzmärkten für diese Derivate besser abbilden sollen als die bisherigen Standardmodelle. Diese neuen Faktormodelle basieren auf elliptischen Verteilungsfunktionen und werden das Gauß’sche Standardmodell konsistent erweitern. Wir werden zeigen, dass die Teilklasse der Gauß’schen Mixturverteilungen (mixtures of Normal distributions) für unsere Modellierung mit elliptischen Verteilungsfunktionen eine zentrale Rolle spielen muss. Im Anschluss daran werden wir einordnen, welche dieser Mixturverteilungen dazu verwendet werden können, um angemessene Wahrscheinlichkeiten für gemeinsame Ausfälle im zugrunde liegenden Portfolio wiederzugeben. Auf Grundlage dieses elliptischen Faktormodells werden wir ein allgemeines so genanntes Large Homogeneous Portfolio Approximationsresultat herleiten, das uns hilft, mit Kreditderivaten umzugehen, die sich auf hochdimensionale Kreditportfolios beziehen. Sorgfältig werden wir die Auswirkungen unseres Modellansatzes auf die Bewertung von Collateralized Debt Obligations (CDOs) analysieren und werden die hier gewonnenen Erkenntnisse für die Bewertung von ITraxx Tranchen einsetzen. Schließlich werden diverse Möglichkeiten diskutiert, wie stochastische Prozesse erzeugt werden können, die Gauß’sche Mixturverteilungen als Randverteilungen besitzen, wobei ein besonderes Augenmerk auf ihre Verwendung im vorher eingehend besprochenen elliptischen Faktormodell gelegt wird.

Einleitung

In den letzten Jahren wuchs der Markt für Kreditderivate und insbesondere der für CDOs rasant an, was wir zu Beginn der Einleitung in Kapitel 1 in einem Überblick über die Entwicklungen der letzten Jahre verdeutlichen. Aber nicht nur das enorme Wachstum und die daraus resultierende ökonomische Relevanz, sondern gerade auch die Komplexität von strukturierten, kreditsensitiven Produkten wie CDOs machen es notwendig, intensiv an einer akkuraten Modellierung dieser Produkte zu arbeiten. Nach einer allgemeinen Einführung besprechen wir, wie CDOs konstruiert und weshalb solche Verbriefungen strukturiert und gehandelt werden. Wir führen aus, dass CDOs zwar bisher aufgrund der Vielfalt der möglichen Strukturen nur außerbörslich gehandelt wurden, jedoch durch die Einführung der ITraxx CDS Indices im Jahre 2004 eine Standardisierung des CDO Marktes angestoßen wurde.
Zusammenfassung


Der Rest des ersten Kapitels diskutiert das Ziel dieser Arbeit, umreißt die einzelnen Resultate, die wir im Rahmen dieser Arbeit erzielen konnten, und bespricht die Gliederung des vorliegenden Textes.

**Verschiedene Kreditportfoliomodelle und das Bewertungsproblem**

Kapitel zwei führt genauer in die Thematik der CDOs und der damit verbundenen Modellierungsprobleme ein. Anfänglich diskutieren wir eingehend eine typische Struktur eines CDOs und erläutern, wie die einzelnen Tranchen nacheinander von Verlusten in dem Portfolio betroffen werden, auf das sich der CDO bezieht. Zweiter Bestandteil dieses Kapitels ist eine Diskussion über die Herausforderungen bei der Modellierung von CDOs, speziell unter Verwendung des Gauß’schen Standardmodells. Wir legen anhand eines ersten Zahlenbeispiels dar, dass das Gauß’sche Modell zu einer Überbewertung der Mezzanine Tranchen und zu einer Unterbewertung der Senior Tranchen führt.

**Grundlegende Konzepte**

Das dritte Kapitel stellt die für unsere Arbeit grundlegenden Konzepte und Definitionen vor und diskutiert wichtige Eigenschaften der eingeführten Begriffe. Zuerst stehen die Klassen der sphärischen und die der elliptischen Verteilungen im Vordergrund, da diese die Basis unseres Modells bilden. Neben der Definition dieser Verteilungen geben wir äquivalente Charakterisierungen an, gehen genauer auf die charakteristische Erzeugendenfunktion ein und diskutieren wichtige Teilklassen, wie z.B. die Klasse der Gauß’schen Mixturverteilungen.

Ein zweiter Schwerpunkt dieses Kapitels liegt auf dem Begriff der regulären Variation (regular variation). In diesem Zusammenhang präsentieren wir eine Version des für uns später relevanten Theorems von Karamata.
Zusammenfassung


Einführung in strukturelle Kreditportfoliomodelle

Bevor wir unser auf elliptischen Verteilungen basierendes Modell einführen, besprechen wir in Kapitel 4 verschiedene Vertreter aus der Klasse der strukturellen Kreditportfoliomodelle genauer. Wir gehen zuerst auf die von uns später verwendete Notation ein, bevor dann das Merton Modell und seine mehrdimensionalen Erweiterungen Gegenstand der Diskussion sind. Im Anschluss daran durchleuchten wir verstärkt verschiedene Ein-Perioden-Modelle, die sich aus dem Merton Modell entwickelt haben. Insbesondere kommen hierbei die Ein-Faktor-Modelle zum Tragen, wobei hier sowohl die Gauß’schen Ein-Faktor-Modelle wie die von Vasicek [Vas87] oder CreditMetrics [BFG97], als auch die Ein-Faktor-Modelle mit anderen Verteilungsannahmen Betrachtung finden. Hierzu gehört unter anderem das Modell von Hull & White [HW04], das auf eindimensionalen t-Verteilungen beruht. Des Weiteren erläutern wir das Konzept der implied correlation und das der base correlation, die einen Hinweis für die Eignung eines Modells für die Bewertung von CDOs geben können.

Ein Kreditportfoliomodell mit elliptischen Verteilungen

In Kapitel 5 führen wir unser auf elliptischen Verteilungen beruhendes Portfoliomodell ein. Als Ausgangspunkt diente uns das Modell von Klaassen et al. [KLSS01], bei dem die Autoren eine Mehr-Faktor-Struktur annehmen, wobei sowohl der Faktor, als auch die idiosynkratischen Komponenten zentriert und normalverteilt sind. Für diesen Aufbau müssen wir daher nur die Matrix von Faktorladungen, die Kovarianzstruktur des mehrdimensionalen Faktors und die Varianzen der unabhängigen, individuellen Risikovariablen bestimmen. Im Vergleich zum Gauß’schen Ein-Faktor-Modell hat man in diesem Mehr-Faktor-Aufbau einige zusätzliche Freiheiten, jedoch können auch die zusätzlichen Parameter, die für ein Anpassen des Modells an Marktdaten verwendet werden können, nicht die grundlegenden Mängel der Normalverteilung beheben. Hierzu gehört die von der Normalverteilung induzierte schwache Abhängigkeitsstruktur, insbesondere in den Rändern der Verteilung. Dieses hat zur Folge, dass die Gauß’schen Modelle Preise für CDO Tranchen generieren, die nicht mit den Marktwerten übereinstimmen (siehe auch Abschnitt 2.2). Deshalb suchen wir nach Verteilungen, deren Anwendung auf die latenten Variablen eine stärkere Abhängigkeitsstruktur und somit höhere Wahrscheinlichkeiten für gemeinsame Ausfälle zur Folge hat. Da Portfolios, auf die CDOs konstruiert werden, sehr groß sein können, wollen wir weiterhin in der Klasse der Faktormodelle bleiben. Durch die Verwendung eines Faktoransatzes erreichen wir, dass die Dimension der für aggregierte
Größen, wie die erwarteten Verluste, zu berechnenden Integrale von der Dimension des Portfolios auf die Dimension des Faktors reduziert werden kann.

Wir modellieren sowohl den Faktor als auch den Vektor der idiosynkratischen Risiken mit Hilfe der Klasse der mehrdimensionalen elliptischen Verteilungsfunktionen (siehe Abschnitt 3.1). Dadurch stellen wir einerseits eine feste Struktur, mit der wir in der Folge arbeiten können, für mögliche Verteilungsfunktionen zur Verfügung, andererseits erhält uns diese Annahme die Möglichkeit, die Gauß'sche Wahrscheinlichkeitsverteilung aus dem Standardmodell mit Bedacht durch andere Verteilungen zu ersetzen, die uns nach theoretischen Überlegungen als geeignet erscheinen. Somit verzichten wir zuerst bewusst auf eine Festlegung auf einen speziellen Repräsentanten dieser Klasse (siehe Abschnitt 5.2).

wird insbesondere auch klar von unseren numerischen Beispielen aus Kapitel 7 gestützt.

In Kapitel 6 leiten wir ein Approximationsresultat her, bei dem die Portfoliogröße gegen Unendlich strebt. Eine direkte Konsequenz aus Schönbergs Theorem (vgl. Theorem 3.1.13) ist jedoch, dass wir für diesen Zweck fordern müssen, dass der Vektor der idiosynkratischen Risiken sogar einer Gauß'schen Mixturverteilung folgt. Die Auswirkungen, wie wir sie im letzten Paragraphen für den allgemeinen elliptischen Fall verdeutlicht haben, bleiben auch bei einer Beschränkung auf die Teilklassen der Gauß'schen Mixturverteilungen grundsätzlich die gleichen, da diese Teilklassen noch sehr groß ist und ebenso viele wichtige Verteilungsfunktionen wie die mehrdimensionale \( t \)-Verteilung beinhalten.

Nachdem wir dieses elliptische Modell eingeführt haben und hierfür grundlegende Eigenschaften diskutiert haben, gehen wir in Abschnitt 5.3.1 die Frage an, welche Merkmale die Gauß'schen Mixturverteilungen besitzen müssen, damit das daraus resultierende Faktormodell eine starke Abhängigkeitsstruktur aufweist. Diese soll dabei helfen, die Probleme der fehlerhaften Bewertung von CDOs unter Verwendung des Gauß'schen Modells zu überwinden. Ein Maß für den Grad der Abhängigkeit in einer angespannten Marktsituation ist die so genannte \textit{tail dependence}, die wir umfassend in unserer Anordnung analysieren. Jeder elliptisch verteilte Zufallsvektor, z.B. der Dimension \( n \), kann in das unabhängige Produkt einer nicht-negativen Skalierungsvariable und eines auf der \( n \)-dimensionalen Einheitskugel gleich verteilten Vektors zerlegt werden. Es ist wichtig anzumerken, dass diese Skalierungsvariable nicht mit der Mischvariable zusammenfällt, die wir im Falle der Gauß'schen Mixturverteilung haben. Schmidt [Sch02], Hult & Lindskog [HL02] und Frahm et al. [FJS03] haben die \textit{tail dependence} innerhalb eines elliptisch verteilten Vektors analysiert, wobei sie eine Charakterisierung der Skalierungsvariablen der klassischen Zerlegung untersucht haben. Da wir jedoch innerhalb der Gauß'schen Mixturverteilungen arbeiten müssen, um das analytische Approximationsresultat zu erhalten, und da wir direkt die Verteilung der Mixturvariablen modellieren wollen, ohne den Umweg über die klassische Zerlegung gehen zu müssen, haben wir in Theorem 5.3.4 die Bedingungen für eine Gauß’sche Mixturverteilung studiert, die dazu führen, dass ein Vektor mit dieser Mixturverteilung die Eigenschaft der \textit{tail dependence} aufweist. Basierend auf dieser Charakterisierung der Mixturverteilungen, konstruieren wir im Abschnitt 5.3.2 verschiedene neue Gauß’sche Mixturverteilungen mit genau dieser \textit{tail dependence}. Diese Verteilungsfunktionen werden dann später im Zusammenhang mit der Bewertung von CDOs genauer untersucht.

\textbf{Die Approximation von Verlusten im Portfolio innerhalb des elliptischen Modells}

Im Zentrum des 6. Kapitels steht ein Approximationsresultat, das wir in Theorem 6.2.3 angeben und das wir im Rahmen des elliptischen Modells zeigen konnten. Dieses Resultat lässt sich unter die von Vasicke [Vas91] begründete Klasse von \textit{Large Homogeneous Portfolio} Approximationen einordnen, bei denen man versucht, die durch Kreditereignisse hervorgerufenen Gesamtverluste in einem Referenzportfolio durch die Verluste in
Zusammenfassung


Im Abschnitt 6.3 wenden wir das Approximationsresultat dann auf ein Beispielportfolio an, wobei wir zur Vereinfachung einige zusätzliche Homogenitätsannahmen treffen. An einem ersten numerischen Beispiel illustrieren wir die Auswirkungen der verschiedenen Gauß’schen Mixturverteilungen aus Abschnitt 5.3.2 anhand der Verlustverteilung des Grenzportfolio..

Unsere Ergebnisse in diesem Kapitel sind unabhängig von dem zugrunde liegenden Wahrscheinlichkeitsmaß, das sowohl das risikoneutrale oder Bewertungsmaß, als auch das historische Maß sein kann. Das Approximationsresultat kann daher für Zwecke des Risikomanagements, wie z.B. im Rahmen der Anforderungen für Risikokapital gemäß Basel II (siehe [Bas06]), verwendet werden, als auch für die Modellierung und Bewertung von Kreditderivaten, die sich auf mehrere Schuldner beziehen.
Die Anwendung des elliptischen Modells auf Kreditderivate

In Kapitel 7 beschäftigen wir uns mit der Abbildung von Kreditderivaten mit Hilfe des elliptischen Modells und im Speziellen mit der Bewertung von CDOs. In einem Bewertungsmodell für CDOs nimmt man üblicherweise an, dass sich jede Firma nur in einem von zwei Zuständen befinden kann, entweder ist sie bereits ausgefallen oder sie ist nicht ausgefallen. Das elliptische Modell und im Speziellen der Mechanismus, wie es zu Kreditverlusten im Portfolio kommt und den wir allgemein in Kapitel 6 eingeführt haben, erlaubt neben der Verwendung zusätzlicher Zustände auch einige weitere allgemeine Modellierungsaspekte. Diese können jedoch auf die für die Bewertung von CDOs notwendigen Bestandteile reduziert werden.

Wir führen zuerst einige grundlegende Begriffe ein, treffen Annahmen für die Modellierung von CDOs und schließen danach eine Untersuchung darüber an, wie die Bewertung von Kreditderivaten auf Portfolios, wie z.B. CDOs, im Rahmen unseres elliptischen Modells durchgeführt werden kann. Zu diesem Zwecke prüfen wir in Abschnitt 7.4, wie das Approximationsresultat, das wir für das allgemeine elliptische Modell in Theorem 6.2.3 hergeleitet haben, unter zusätzlichen Homogenitätsannahmen auf CDO Strukturen angewandt werden kann. Hierbei bestimmen wir, welche Ausdrücke im elliptischen Verteilungskontext benötigt werden für die Bewertung von CDOs, und leiten für jeden dieser Ausdrücke geschlossene Formeln her. Hierzu gehören die Grenzverteilung der Portfolioverluste, die Ausfallwahrscheinlichkeiten der jeweiligen Firmen und die erwarteten Verluste in einer spezifischen Tranche.

Neben den analytischen Betrachtungen stellt sich auch die Frage einer effizienten Implementierung des elliptischen Modells. Im Abschnitt 7.5 untersuchen wir deshalb eingehend, wie die Schlüsselgrößen, wie die erwarteten Verluste in einer Tranche oder die Ausfallwahrscheinlichkeit einer einzelnen Firma, im Falle der verschiedenen Gauß'schen Mischverteilungen numerisch berechnet werden können. Wir zeigen, dass alle Integrale, die in den Schlüsselgrößen vorkommen, auf Integrale zurückgeführt werden können, zu deren Berechnung wir direkten Gauß'schen Quadraturverfahren der gleichen Art verwenden können, unabhängig von der Wahl der Mischspecifikation aus Abschnitt 5.3.2. Dieses eröffnet die Möglichkeit für effiziente numerische Bewertungsmethoden.

wird der C++ Programmcode in eine so genannte dynamic link library\(^1\) übersetzt, die als „Add-In“ in Microsofts Excel importiert werden kann. Dort können dann die C++ Bewertungs- und Kalibrierungsfunktionen direkt auf die in Excel abgelegten Daten verwendet werden. Mit diesem Aufbau verbinden wir den effizienten Umgang mit großen Daten unter Excel einerseits und die schnelle Programmiersprache C++ andererseits.


Mit der eingehenden Untersuchung in diesem Kapitel konnten wir zeigen, dass unser elliptisches Modell eine flexible Erweiterung der Gauß’schen Standardmodelle liefert, die gleichzeitig die Marktpreise gut widerspiegelt, und konnten somit einen wertvollen Beitrag zur besseren Modellierung von auf Portfolios basierenden Kreditderivaten wie CDOs liefern, insbesondere bezüglich der Frage der Abhängigkeitsstruktur zwischen den Papieren im zugrunde liegenden Kreditportfolio.

Dynamisches elliptisches Modell


In Kapitel 8 ergründen wir nun, wie man die Annahme der strikten Stationarität abschwächt kann. Eine solche Abschwächung ist insbesondere dann sinnvoll, wenn man verschiedene CDOs konsistent innerhalb eines Modells abbilden will, die zwar auf dem gleichen Referenzportfolio basieren, aber verschiedene Laufzeiten haben. Das Ziel ist somit die Beantwortung der Frage, wie das statische elliptische Modell aus Kapitel 5 dahingehend dynamisiert und erweitert werden kann, dass sich die Verteilungen der Firmenwerte über die Zeit hinweg verändern können.


Abschließend gehen wir der Frage nach, ob man mit der in Abschnitt 5.3.2 durch das Exp-Exp-Law Modell eingeführten Verteilung einen mehrdimensionalen Lévy-Prozess generieren kann. Viele der von uns untersuchten Eigenschaften der Verteilung deuten stark daraufhin, jedoch wird der rigoros mathematische Nachweis dieser Aussage nur sehr schwer zu führen sein.