Technische Universität München
Zentrum Mathematik
HVB-Stiftungsinstitut für Finanzmathematik

Extendibility of Marshall-Olkin distributions via Lévy subordinators and an application to portfolio credit risk

Jan-Frederik Mai

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. Claudia Czado, Ph.D.
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Rudi Zagst
	2.	Univ.-Prof. Dr. Rüdiger Kiesel Universität Duisburg-Essen
	3.	Prof. Nick H. Bingham, Ph.D. Imperial College London / UK (schriftliche Beurteilung)

Die Dissertation wurde am 17.03.2010 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 22.06.2010 angenommen.

Für Otto

Zusammenfassung

Die Familie der austauschbaren Marshall-Olkin Verteilungen wird untersucht. Aus analytischer Sicht werden Querverbindungen mit vollständig monotonen Zahlenfolgen aufgezeigt. Aus wahrscheinlichkeitstheoretischer Sicht wird eine alternative Konstruktion erweiterbarer Marshall-Olkin Verteilungen mittels Lévy Subordinatoren hergeleitet. Dieses Resultat wird verwendet um effiziente Simulationsalgorithmen und ein Bewertungsmodell für Portfolio-Kreditderivate zu entwickeln.

Abstract

The family of exchangeable Marshall-Olkin distributions is investigated. From an analytical perspective, coherences with completely monotone sequences are revealed. From the viewpoint of probability theory, an alternative construction of extendible Marshall-Olkin distributions via Lévy subordinators is derived. This result is used to develop efficient simulation algorithms and a pricing model for portfolio credit derivatives.

Acknowledgements

Foremost, my thanks go to my advisor Rudi Zagst for giving me the opportunity to write this thesis. He provided me with excellent research conditions during the last years, including financial support by means of a teaching position, adequate allocation of resources to carry out my research, and the opportunity to participate in several scientific conferences. Moreover, I would like to thank Matthias Scherer for his support. He encouraged me to earn my Ph.D.-degree in Munich.

Furthermore, I am very grateful to Rüdiger Kiesel and to Nick Bingham for agreeing to serve as referees for this thesis.

I would like to thank Thorsten Schmidt, who drew my attention to copulas when he saw my presentation of the "Lévy-frailty default model" in a very early and half-baked version at the beginning of my Ph.D.-studies. I would also like to thank Fabrizio Durante, who helped me out with his copula references from time to time and encouraged me to solve an interesting problem which is now stated as one of the major findings of this thesis (Theorem 3.5.3). Furthermore, I would like to thank several people for reading previous versions of this manuscript or research papers, and giving me fruitful feedback. These are Nick Bingham, Fabrizio Durante, Christian Hering and Marius Hofert.

Finally, I thank my mother Brigitte and my sister Mirja for their ongoing faith in me, and especially my father Otto who always wanted the best for me. I know he would be as proud of me now as he always was. Last but not least, great thanks go to my wife Jasna. Taking into account the fact that I met her just before I started my studies of mathematics, I'm not discounting the fact that, by causing my emotional well-being, her love and support have the lion's share of my academic success.

Contents

1 Introduction **9**

2 Mathematical Background **13**

2.1 Notations 13

2.2 Copulas 14

2.2.1 A Bivariate Example 16

2.2.2 Sklar's Theorem and Survival Copulas 20

2.2.3 Dependence Measures 26

2.3 Important Copula Families 30

2.3.1 Extreme-Value Copulas 30

2.3.2 Marshall-Olkin Survival Copulas 33

2.3.3 Archimedean Copulas 37

2.4 Lévy Subordinators 42

2.4.1 Examples 47

2.5 Moment Problems and Completely Monotone Sequences 53

2.5.1 Hausdorff's Moment Problem 54

2.5.2 Moment Problem for Convex Distributions 57

3 An Analytical Study of the Marshall-Olkin Distribution **61**

3.1 Exchangeable Marshall-Olkin Survival Copulas 61

3.2 Motivation 66

3.3 d-Monotone Sequences 68

3.4 Reparameterization 74

3.5 A Characterization Theorem 76

3.6 Analogy with Archimedean Copulas 84

4 Construction of Infinite Marshall-Olkin Sequences **89**

4.1 Analytical Preliminaries 90

4.2 The Subclass of Lévy-Frailty Copulas 94

4.3 Concrete Examples 101
4.4 Distributional Properties of the Class eMO 107

5 Sampling Applications and Non-Exchangeable Structures 117
5.1 Sampling Exchangeable Cuadras-Augé Copulas 119
5.2 Hierarchical Lévy-Frailty Copulas 124
5.2.1 Construction and Sampling Algorithm 125
5.2.2 Hierarchical Marshall-Olkin Distribution 133

6 The Lévy-Frailty Default Model 135
6.1 Portfolio Credit Derivatives 135
6.2 Literature Overview: Default Models 142
6.3 Definition of the Model 149
6.4 The Portfolio Loss Distribution 156
6.5 Calibration to CDO Market Quotes 164
6.6 Comparison with Existing Models 169

7 Conclusion 177

A Appendix 179
A.1 Proof sketch for Theorem 2.3.2 179
A.2 Proof sketch for Theorem 2.4.3 181

Bibliography 189

1 Introduction

> *"In most sciences one generation tears down what another has built and what one has established another undoes. In mathematics alone each generation adds a new story to the old structure."*
>
> Hermann Hankel, German mathematician.

The initial motivation for this thesis stems from the field of *financial mathematics*, or, more precisely, from the subject of *portfolio credit risk modeling*. Considering a portfolio of d companies, their unknown future lifetimes are denoted by $\tau_1, \dots, \tau_d$. Since these numbers are random, their treatment requires the tool box of probability theory, and the core question is:

"what is an appropriate mathematical model for these lifetimes?"

As one can easily imagine, the word "appropriate" heavily depends on the specific application of the model. The application motivating this dissertation is the issue of pricing so-called *Collateralized Debt Obligations (CDOs)*. CDOs are financial contracts - typically traded between globally active financial institutions - offering insurance against company defaults to investors. The assesment of fair insurance premia for these contracts requires a viable mathematical model for the default times $\tau_1, \dots, \tau_d$. Let us suppose, for a moment, that financial institutions use inappropriate models which systematically underestimate the risks involved in CDOs. If new contracts are settled, the resulting insurance premia are systematically too small. The consequence might be an excessive growth of market volume of these contracts, spreading default risk all over the globe. Such a scenario is dangerous for the stability of the global economy. In fact, many experts in economic research argued that the mispricing of CDOs (and related contracts) played the role of a fire accelerant in the recent credit crisis.

To avoid such a scenario it is important to find an appropriate mathematical setup for $\tau_1, \dots, \tau_d$. Specifically when pricing CDOs, there are two fundamental difficulties the

model has to cope with: on the one hand, there is empirical evidence that company defaults are far from occuring independently of each other. This necessitates the use of a model which is based on a flexible multivariate distribution. On the other hand, the number of firms in consideration is very large - a typical convention is $d = 125$. This large dimensionality massively restricts the class of appropriate multivariate distributions, since a high level of mathematical viability must be guaranteed. The derivation of closed formulas or efficient approximations for insurance premia requires "simple" models, whereas an adequate treatment of the dependence structure between the default times requires "advanced" models. As often in applied mathematics, the appropriate trade-off between "simple" and "advanced" is a delicate question.

The present thesis proposes to apply a popular model from the field of reliability theory: the so-called *Marshall-Olkin distribution*, introduced in and named after the seminal reference [Marshall, Olkin (1967)]. The idea behind this multivariate distribution is an economy in which (joint) company defaults are triggered by exogenous shocks. These shocks might be interpreted as economy crises affecting one, two, three or more firms at a time. This interpretation renders the Marshall-Olkin distribution a reasonable choice for a CDO pricing model. For instance, [Andersen, Sidenius (2005), Burtschell et al. (2009)] propose the use of this approach in a simplistic special case, and [Embrechts et al. (2003), Giesecke (2003), Lindskog, McNeil (2003)] apply it in a more general form. However, the disadvantage of this choice is that the Marshall-Olkin distribution in its original form is very inconvenient to work with in large dimensions. In other words, non-trivial subclasses of this distribution are not simple enough to be of true practical value in dimension $d = 125$. When applying it to CDO pricing, it is difficult to efficiently compute the required insurance premia, and even a Monte Carlo simulation of the model is expensive - not to say impossible - on a standard computer. One reason for this drawback is that standard approximation techniques fail in the Marshall-Olkin model. In other popular state-of-the-art CDO pricing models used in practice, such as e.g. [Li (2000), Albrecher et al. (2007)], a useful method is to approximate the finite portfolio size of $d = 125$ by an infinite portfolio size, i.e. to consider the limiting behavior of the model as $d \to \infty$. However, these approximations are based on a latent factor representation which is not obvious in the Marshall-Olkin model, unless one considers a very simplistic special case as [Andersen, Sidenius (2005), Burtschell et al. (2009)].

The main contribution of the present dissertation is to overcome this drawback and to formalize the aforementioned limiting process $d \to \infty$ for exchangeable Marshall-Olkin

distributions. Put differently, we derive a latent factor representation for exchangeable and extendible Marshall-Olkin distributions. In a first step a purely analytical point of view is taken. If rearranged, a d-dimensional exchangeable Marshall-Olkin distribution can be identified with an associated sequence $(a_0, \ldots, a_{d-1})'$ of non-negative numbers. The latter satisfies a monotonicity property that is linked to the dimension d - termed *d-monotonicity*. Letting the dimension d tend to infinity then corresponds to the respective infinite sequence $\{a_k\}_{k\in\mathbb{N}_0}$ being *completely monotone* - an analytical notion well-studied in the academic literature for decades, see e.g. [Hausdorff (1921)]. If the associated sequence $(a_0, \ldots, a_{d-1})'$ is the initial part of an infinite completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$, it is formally possible to make sense of the limiting process $d \to \infty$. In this case, the Marshall-Olkin distribution is called *extendible*. As a byproduct, these analytical findings imply blatant similarities with well-known results in the related context of *Archimedean copulas*.

A further highlight of this thesis is the construction of a probability space on which one can explicitly work out the limiting process $d \to \infty$ for exchangeable Marshall-Olkin distributions. To this end, a result of [Gnedin, Pitman (2008)] is used to establish a one-to-one correspondence between completely monotone sequences and *Lévy subordinators*- i.e. non-decreasing stochastic processes with independent and stationary increments. By virtue of the aforementioned analytical results, a composition argument manifests a one-to-one correspondence between Lévy subordinators and extendible Marshall-Olkin distributions. Indeed, a transformation of the path of a Lévy subordinator to the unit interval $[0, 1]$ is interpreted as the random choice of a distribution function on $(0, \infty)$. Letting $\tau_1, \ldots, \tau_d$ be drawn independently and identically distributed from this predetermined distribution function, it is shown that $(\tau_1, \ldots, \tau_d)'$ has an exchangeable Marshall-Olkin distribution. Its associated sequence $(a_0, \ldots, a_{d-1})'$ is given in terms of the parameters of the underlying Lévy subordinator. The resulting link between Lévy subordinators, completely monotone sequences, and Marshall-Olkin distributions is explored thoroughly and several specific examples are provided to illustrate the findings.

The developed Bayesian two-step construction of Marshall-Olkin distributions via Lévy subordinators is not only new and interesting, but also more convenient than the original Marshall-Olkin model in large dimensions. Efficient Monte Carlo simulations are now available, and approximation techniques based on an infinite portfolio size can now be formalized. Consequently, the application of Marshall-Olkin distributions to the

pricing of CDOs is now possible. Moreover, when using this approach for real-world applications such as a calibration to CDO market quotes, it is simple enough to allow for an efficient implementation on a standard computer. In particular, it does not rely on time-consuming Monte-Carlo techniques when used for pricing.

Concerning the mathematical framework, the language of *copula theory* is used throughout this dissertation. In particular, the Marshall-Olkin distribution is treated by means of its so-called *survival copula*, which is more convenient to explore. The thesis is divided into seven chapters, the first of which is the present introduction. Chapter 2 recalls the required mathematical notions. Chapter 3 studies monotonicity properties of sequences and their relation to the class of exchangeable Marshall-Olkin survival copulas. The two main findings in this chapter are Theorems 3.4.1 (page 75) and 3.5.3 (page 78), which reveal how d-monotone sequences occur naturally in the present context. Chapter 4 defines so-called *Lévy-frailty copulas* and discusses their properties. The main result of this chapter is Theorem 4.2.2 (page 95), which provides a probabilistic construction. The theoretical value of Lévy-frailty copulas is demonstrated in Chapters 5 and 6, both of which may be viewed as applications of Theorem 4.2.2. Chapter 5 is concerned with random number generation and Chapter 6 deals with portfolio credit risk modeling. Finally, Chapter 7 overviews the findings.

The results in Chapter 2 are well-known, whereas the findings of Chapters 3 to 6 are the author's original work. Most of the results have also already been published in peer-refereed international journals, see [Mai, Scherer (2009a), Mai, Scherer (2009b), Mai, Scherer (2009c), Mai, Scherer (2009d), Mai, Scherer (2010)]. This thesis is self-contained in the sense that at least sketches of proofs are provided for those statements in Chapter 2 on which subsequent results rely.

2 Mathematical Background

"The dwarf sees farther than the giant, when he has the giant's shoulder to mount on."

Samuel Taylor Coleridge, English poet, in *The friend*, 1828.

This chapter introduces mathematical notions which are used in the present dissertation. Section 2.2 outlines the general idea of a copula. Section 2.3 continues by presenting specific examples of copulas, which are important to understand the results of this thesis. Section 2.4 introduces Lévy subordinators. These are stochastic processes that are applied later on to construct a class of multivariate distributions, which will be termed Lévy-frailty copulas. Finally, Section 2.5 is concerned with so-called moment problems for distributions on the unit interval. The corresponding results are useful to construct and sample several copulas.

2.1 Notations

Before we start, let us clarify some notations which are used throughout this dissertation.

- **Important sets:** $\mathbb{N}$ denotes the natural numbers $1, 2, \ldots$, and $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. $\mathbb{R}$ denotes the real numbers. Moreover, for $d \in \mathbb{N}$, $\mathbb{R}^d$ denotes the set of all d-dimensional column vectors with entries in $\mathbb{R}$. For a vector $\vec{v} \in \mathbb{R}^d$, we denote by $\vec{v}'$ its transpose. For some set A, we denote by $\mathcal{B}(A)$ the corresponding Borel σ-algebra, which is generated by all open subsets of A. The cardinality of a set A is denoted by $|A|$.

- **Important distributions:** Some frequently used probability distributions are sometimes abbreviated. These comprise the exponential distribution with mean $1/\lambda > 0$, which is denoted by $Exp(\lambda)$, and the uniform distribution on the unit

interval $[0,1]$, abbreviated by $U[0,1]$. Additionally, the Poisson distribution with mean $\beta > 0$ is denoted by $Poi(\beta)$. Abbreviations for other distributions are introduced when they first appear. The symbol $\sim$ means "distributed according to", e.g. $E \sim Exp(1)$ means that E is an exponential random variable with unit mean.

- **Functions:** Univariate as well as d-dimensional distribution functions are denoted by capital letters, mostly F, G. As an exception, a copula is denoted by the letter C; its arguments are denoted by $(u_1, \ldots, u_d)' \in [0,1]^d$. Moreover, the n-th derivative of a real-valued function f is abbreviated by $f^{(n)}$.

- **Probability spaces:** A probability space is denoted by $(\Omega, \mathcal{F}, \mathbb{P})$, with σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. The corresponding expectation operator is denoted by $\mathbb{E}$. Random variables (or vectors) are mostly denoted by the Greek letter τ (respectively $(\tau_1, \ldots, \tau_d)'$). As an exception, we write $(U_1, \ldots, U_d)'$ for a d-dimensional random vector with a copula as joint distribution function[1]. If two random variables τ_1, τ_2 are equal in distribution, we write $\tau_1 \stackrel{d}{=} \tau_2$. Similarly, $\stackrel{d}{\to}$ denotes convergence in distribution. Elements of the space Ω, usually denoted by ω, are almost always omitted as arguments of random variables, i.e. instead of writing $\tau(\omega)$, we simply write τ. Finally, the shortcut i.i.d. stands for "**i**ndependent and **i**dentically **d**istributed" and is used from time to time.

- **Stochastic processes:** A stochastic process $X : \Omega \times [0,\infty) \to \mathbb{R}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is denoted by X or by $\{X_t\}_{t \geq 0}$. I.e. we omit the argument $\omega \in \Omega$, and the "time argument" t is written as a subindex, i.e. X_t instead of $X(t)$. This is in order to avoid confusion with deterministic functions f, whose arguments are written in brackets, i.e. $f(x)$.

2.2 Copulas

The law of a d-dimensional random vector $(\tau_1, \ldots, \tau_d)'$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is usually studied from its *distribution function*

$$F(t_1, \ldots, t_d) := \mathbb{P}(\tau_1 \leq t_1, \ldots, \tau_d \leq t_d), \quad t_1, \ldots, t_d \in \mathbb{R}.$$

[1]The letter U indicates that $U_1, \ldots, U_d$ are uniformly distributed on the unit interval, see Definition 2.2.1 below.

For $i = 1, \ldots, d$ the univariate *marginal law* or *margin* F_i of the random variable τ_i can be retrieved from F via

$$F_i(t_i) := \mathbb{P}(\tau_i \leq t_i) = F(\infty, \ldots, \infty, t_i, \infty, \ldots, \infty), \quad t_i \in \mathbb{R}.$$

Using ∞ as an argument of F is a short-hand notation for taking the limits as the arguments $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_d$ tend to infinity. It is important to mention that it is not enough to know the margins $F_1, \ldots, F_d$ in order to determine F. Additionally it is required to know how the marginal laws are coupled. This is achieved by means of a *copula* of $(\tau_1, \ldots, \tau_d)'$. Generally speaking, knowing the margins and a copula is equivalent to knowing the distribution. It is now appropriate to give the definition of a copula.

Definition 2.2.1 (Copula)

(a) A function $C : [0,1]^d \to [0,1]$ is called a (d-dimensional) *copula*, if there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random vector $(U_1, \ldots, U_d)'$ such that $U_i \sim U[0,1]$ for all $i = 1, \ldots, d$ and

$$C(u_1, \ldots, u_d) = \mathbb{P}(U_1 \leq u_1, \ldots, U_d \leq u_d), \quad u_1, \ldots, u_d \in [0,1].$$

(b) On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $(U_1, \ldots, U_d)'$ be a random vector on $[0,1]^d$ whose joint distribution function is a copula $C : [0,1]^d \to [0,1]$. For $i = 2, \ldots, d$ and indices $1 \leq j_1 < \ldots < j_i \leq d$ the notation $C_{j_1,\ldots,j_i} : [0,1]^i \to [0,1]$ is introduced for the joint distribution function of the random vector $(U_{j_1}, \ldots, U_{j_i})'$. It is itself a copula and called an *i-margin* of C.

For a random vector $(U_1, \ldots, U_d)' \in [0,1]^d$ on the d-dimensional unit cube the values of its distribution function on $\mathbb{R}^d \setminus [0,1]^d$ are completely determined by its values on $[0,1]^d$. Thus, copulas are defined on $[0,1]^d$ only. The two simplest examples of copulas are the *independence copula* Π and the *copula of complete comonotonicity* (or *upper-Fréchet-Hoeffding bound*) M, defined by

$$\Pi(u_1, \ldots, u_d) := \prod_{i=1}^{d} u_i, \quad M(u_1, \ldots, u_d) := \min\{u_1, \ldots, u_d\}.$$

A random vector with joint distribution function Π is a vector of d independent random variables which are uniformly distributed on the unit interval. In contrast, a random

vector $(U_1, \ldots, U_d)'$ with joint distribution function M is equal in distribution to a random vector $(U, \ldots, U)'$ where $U \sim U[0,1]$, i.e. all components are identical almost surely.

Univariate distribution functions can be characterized by analytical properties. It is well-known that a function $F : \mathbb{R} \to [0,1]$ is the distribution function of a random variable τ if and only if it is right-continuous, non-decreasing, $\lim_{t\to-\infty} F(t) = 0$, and $\lim_{t\to\infty} F(t) = 1$, see e.g. [Billingsley (1995), Theorem 12.4, p. 176]. There is an analogous characterization of multivariate distribution functions via analytical properties. In particular, a (d-dimensional) copula is often defined as a d-increasing function subject to some boundary conditions, see e.g. [Nelsen (1999), page 40]. However, this analytical notion is not needed in the present thesis.

A copula is basically a multivariate distribution function with standardized margins. The choice of the $U[0,1]$-distribution is arbitrary to some degree. Sometimes it is more natural to work with another marginal law. For example a multivariate extreme-value distribution should have extreme-value margins as well, see e.g. the proof of Theorem 2.3.2 below. However, Sklar's Theorem, see Theorem 2.2.3 below, shows that there is a good reason for using the $U[0,1]$-law. It shows that an arbitary multivariate distribution can always be split into a copula and its margins. Sometimes such a splitting is quite useful. For instance, to comprehend the dependence structure behind a multivariate distribution it is often illuminating to get rid of the margins. This is exactly what a copula does. Furthermore, in the theory of dependence modeling there is some useful vocabulary which is based on the notion of a copula. Finally, in some applications a separation of dependence structure and marginal laws leads to a complexity reduction, since the one big problem of fitting a multivariate distribution to data can be split into two smaller problems: (1) fitting the margins and (2) fitting a copula. For these reasons, the language of copula theory is used throughout this thesis. Standard textbooks in this field are [Joe (1997), Nelsen (1999), McNeil et al. (2005)].

2.2.1 A Bivariate Example

To get a feeling for copulas, this short section introduces a specific bivariate copula which is of fundamental interest during this thesis. It was first introduced in [Cuadras, Augé (1981)].

Definition 2.2.2 (Bivariate Cuadras-Augé Copula)
For each $\alpha \in [0,1]$ the function $C_\alpha : [0,1]^2 \to [0,1]$, defined by setting $C_\alpha(u_1,u_2) = \min\{u_1,u_2\}\,\max\{u_1,u_2\}^{1-\alpha}$, is called *bivariate Cuadras-Augé copula*.

C_α is a geometric mean of the independence copula Π and the comonotonicity copula M, since $C_\alpha = \Pi^{1-\alpha}\,M^\alpha$. As this distribution occurs several times later on, it is discussed in some detail subsequently. First of all, to see that C_α actually defines a copula, a bivariate random vector with joint distribution function C_α is constructed, following an approach of [Marshall, Olkin (1967)]. If $\alpha \in \{0,1\}$, then $C_\alpha \in \{\Pi, M\}$, so it defines a copula. For $\alpha \in (0,1)$ let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space on which three independent exponential random variables $E_1, E_2, E_{1,2}$ are defined. We assume $E_1, E_2 \sim Exp\big((1-\alpha)/\alpha\big)$ and $E_{1,2} \sim Exp(1)$. Define the two random variables U_1 and U_2 by

$$U_1 := \exp\Big(-\frac{1}{\alpha}\,\min\big\{E_1,E_{1,2}\big\}\Big), \quad U_2 := \exp\Big(-\frac{1}{\alpha}\,\min\big\{E_2,E_{1,2}\big\}\Big).$$

Then, it follows for $0 < u_1, u_2 < 1$ that

$$\begin{aligned}
&\mathbb{P}(U_1 \le u_1, U_2 \le u_2)\\
&\quad = \mathbb{P}\Big(\min\big\{E_1,E_{1,2}\big\} \ge -\alpha\,\log(u_1), \min\big\{E_2,E_{1,2}\big\} \ge -\alpha\,\log(u_2)\Big)\\
&\quad = \mathbb{P}\Big(E_1 \ge -\alpha\,\log(u_1), E_2 \ge -\alpha\,\log(u_2), E_{1,2} \ge \max\big\{-\alpha\,\log(u_1), -\alpha\,\log(u_2)\big\}\Big)\\
&\quad = \mathbb{P}\big(E_1 \ge -\alpha\,\log(u_1)\big)\,\mathbb{P}\big(E_2 \ge -\alpha\,\log(u_2)\big)\,\mathbb{P}\big(E_{1,2} \ge -\alpha\,\log(\min\{u_1,u_2\})\big)\\
&\quad = u_1^{1-\alpha}\,u_2^{1-\alpha}\,\min\{u_1,u_2\}^\alpha = C_\alpha(u_1,u_2).
\end{aligned}$$

Furthermore for $u \in (0,1)$ and $i = 1,2$ it holds that

$$\begin{aligned}
\mathbb{P}(U_i \le u) &= \mathbb{P}\big(\min\{E_i,E_{1,2}\} \ge -\alpha\,\log(u)\big)\\
&= \mathbb{P}\big(E_i \ge -\alpha\,\log(u)\big)\,\mathbb{P}\big(E_{1,2} \ge -\alpha\,\log(u)\big) = u^{1-\alpha}\,u^\alpha = u.
\end{aligned}$$

Hence U_1, U_2 are both $U[0,1]$-distributed and C_α is a copula. Figure 2.1 illustrates the function C_α for different choices of α.

Unless $\alpha = 0$, the probability measure induced by C_α is not absolutely continuous with respect to the Lebesgue measure on $[0,1]^2$. In analytical terms this means that C_α is not differentiable on the diagonal of the unit square. It can be shown, see e.g. [Cuadras, Augé (1981)], that for $B \in \mathcal{B}\big([0,1]^2\big)$ the random vector $(U_1,U_2)'$ constructed

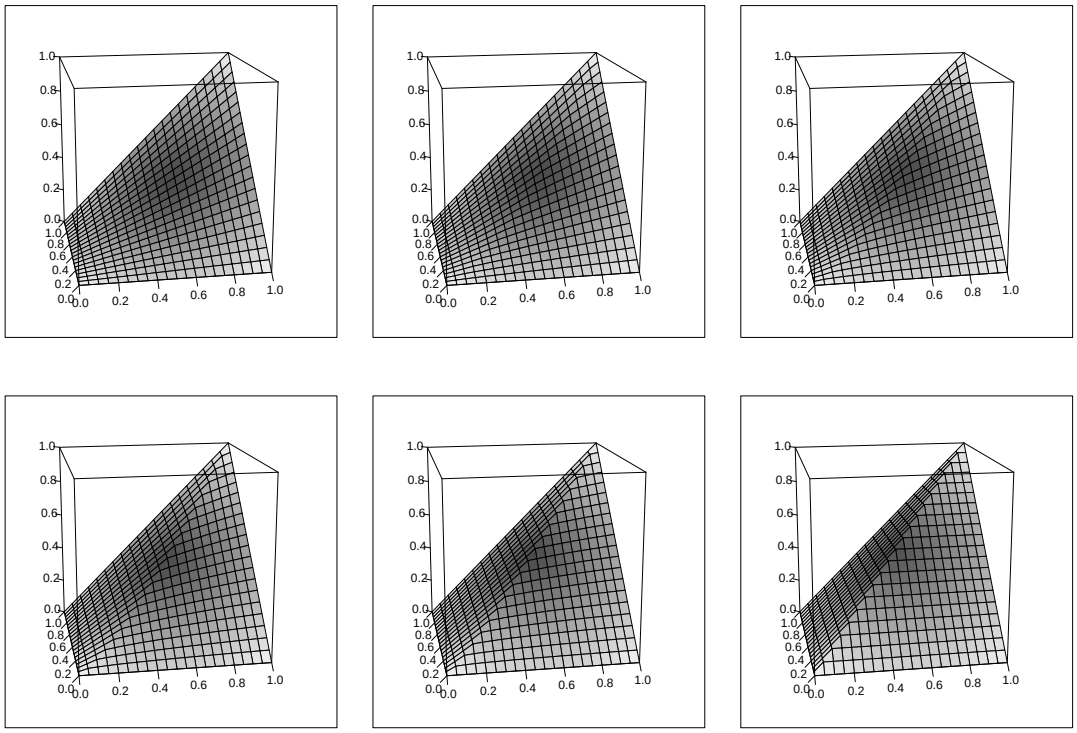


Figure 2.1 Function plots of $C_\alpha(u_1, u_2)$ for $\alpha = 0$ (upper left, independence case), $\alpha = 0.2$, $\alpha = 0.4$, $\alpha = 0.6$, $\alpha = 0.8$, and $\alpha = 1$ (lower right, case of complete comonotonicity). One observes that the bend on the diagonal emerges with increasing dependence parameter α.

above satisfies

$$\mathbb{P}\big((U_1, U_2)' \in B\big) = \iint_B (1-\alpha)\, \max\{u_1, u_2\}^{-\alpha}\, d(u_1, u_2) + \int_{\{u \in [0,1] \,|\, (u,u) \in B\}} \alpha\, u^{1-\alpha}\, du.$$

In particular it holds that

$$\mathbb{P}(U_1 = U_2) = \int_0^1 \alpha\, u^{1-\alpha}\, du = \frac{\alpha}{2-\alpha}. \tag{2.1}$$

This distributional property is visualized in Figure 2.2 by means of scatterplots, that is by generating $n \in \mathbb{N}$ i.i.d. samples of a random vector with joint distribution function C_α and plotting all these n points into one coordinate system.
Concerning the parameter estimation for C_α, [Ocana, Ruiz-Rivas (1990)] show that

moment-type estimators are outperformed by a maximum likelihood estimator for α, which is derived in [Ruiz-Rivas, Cuadras (1988)]. It is based on the notion of *curved exponential families* as introduced in [Efron (1975)]. The maximum likelihood estimator for α, based on n i.i.d. samples $(U_1^{(1)}, U_2^{(1)})', \ldots, (U_1^{(n)}, U_2^{(n)})'$ from C_α, is given by

$$\hat{\alpha} := \frac{T - n + \sqrt{(n-T)^2 + 4\,n_c\,T}}{2\,T}, \tag{2.2}$$

where

$$n_c := \left| \left\{ i \in \{1, \ldots, n\} \,\middle|\, U_1^{(i)} = U_2^{(i)} \right\} \right|, \quad T := -\sum_{i=1}^{n} \log\left(\max\left\{ U_1^{(i)}, U_2^{(i)} \right\} \right).$$

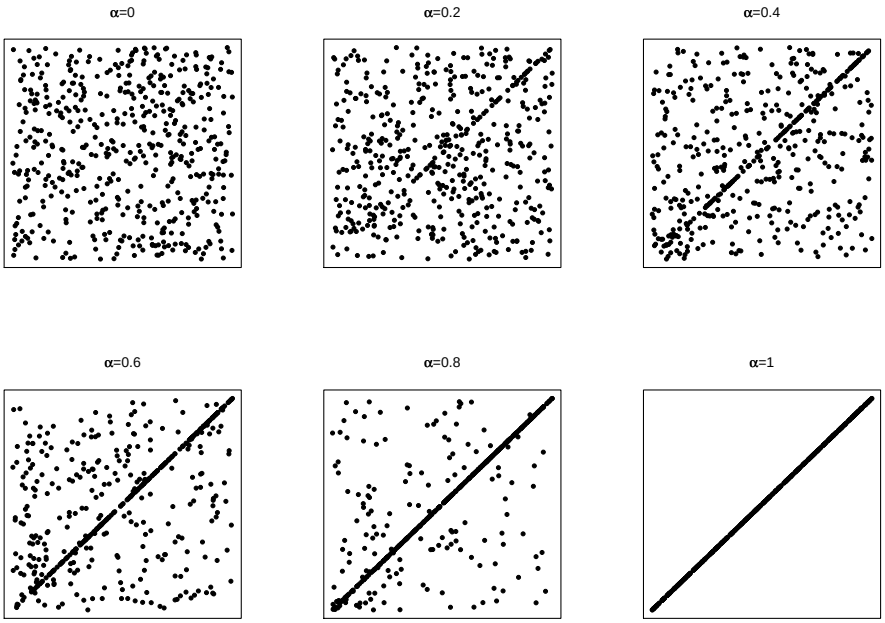


Figure 2.2 Scatterplots of $n = 500$ samples of $(U_1, U_2)' \sim C_\alpha$ with different choices of the parameter α. With increasing α, i.e. increasing dependence, more points are on the diagonal. It is also observed that $\alpha = 0$ implies independence, since no clusters are present, and $\alpha = 1$ implies comonotonicity, since all points are on the diagonal.

2.2.2 Sklar's Theorem and Survival Copulas

At the heart of copula theory stands the seminal Theorem of Sklar.

Theorem 2.2.3 (Sklar (1959))
Let F be a d-dimensional distribution function with margins $F_1, \ldots, F_d$. Then there exists a d-dimensional copula C such that for all $(t_1, \ldots, t_d)' \in \mathbb{R}^d$ it holds that

$$F(t_1, \ldots, t_d) = C\big(F_1(t_1), \ldots, F_d(t_d)\big). \tag{2.3}$$

If $F_1, \ldots, F_d$ are continuous, then C is unique. Conversely, if C is a d-dimensional copula and $F_1, \ldots, F_d$ are univariate distribution functions, then the function F defined via (2.3) is a d-dimensional distribution function.

Proof
The original reference is [Sklar (1959)]. For a proof see also [Nelsen (1999), p. 18] or [Schweizer, Sklar (1983)]. In the case of continuous margins, we provide a proof of the survival analog of Sklar's Theorem, see Theorem 2.2.5 below, which relies on the same idea. □

Since we always deal with random vectors $(\tau_1, \ldots, \tau_d)'$ having continuous margins, the copula in Theorem 2.2.3 is unique, and we refer to it as **the** copula of $(\tau_1, \ldots, \tau_d)'$.

Sklar's Theorem allows to conveniently construct multivariate distribution functions in two steps. In a first step one may choose the univariate margins, and in a second step a copula. Thus, having a repertoire of parametric models for the margins and the copula, it is possible to fit a multivariate distribution to given data by first fitting the parameters of the margins and subsequently the copula parameters. This is the main reason for the popularity of copulas in statistical modeling. We exploit this idea in Chapter 6.

Example 2.2.4 (Gaussian Copula)
On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $(\tau_1, \ldots, \tau_d)'$ be a normally distributed random vector with joint distribution function

$$F(t_1, \ldots, t_d) := \iint_{\times_{i=1}^d (-\infty, t_i]} (2\,\pi)^{-\frac{d}{2}} \det\big(\Sigma\big)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} (\vec{s} - \vec{\mu})' \, \Sigma^{-1} \, (\vec{s} - \vec{\mu}) \Big) \, d\vec{s},$$

for a symmetric, positive-definite matrix Σ and a mean vector $\vec{\mu} = (\mu_1, \ldots, \mu_d)' \in \mathbb{R}^d$, where $\vec{s} := (s_1, \ldots, s_d)'$; and $\det(\Sigma)$ is the determinant of Σ. Denoting by $\sigma_1^2, \ldots, \sigma_d^2 > 0$

the diagonal entries of Σ, the marginal law F_i of τ_i is a normal distribution with mean μ_i and variance σ_i^2, $i = 1, \ldots, d$. The copula C of $(\tau_1, \ldots, \tau_d)'$ is called a *Gaussian copula* and is given by

$$C(u_1, \ldots, u_d) = F\big(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\big). \tag{2.4}$$

The copula of a multivariate distribution F with strictly increasing continuous margins $F_1, \ldots, F_d$ is always implicitly given by (2.4), but sometimes this expression can be computed explicitly. In the Gaussian case however, this is not possible due to the fact that no closed-form antiderivatives of normal densities are known.

A d-dimensional copula C induces a probability measure dC on the unit cube $[0,1]^d$. More clearly, if the random vector $(U_1, \ldots, U_d)'$ on $[0,1]^d$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function C, then

$$dC(B) := \mathbb{P}\big((U_1, \ldots, U_d)' \in B\big), \quad B \in \mathcal{B}\big([0,1]^d\big).$$

This probability measure, and also the copula C itself, is called *absolutely continuous* if there exists a density function $c : [0,1]^d \to [0, \infty)$ such that

$$C(u_1, \ldots, u_d) = \int_0^{u_1} \ldots \int_0^{u_d} c(s_1, \ldots, s_d)\, ds_d \ldots ds_1, \quad u_1, \ldots, u_d \in [0,1].$$

For example, the Gaussian copula from Example 2.2.4 above is absolutely continuous. If C is absolutely continuous, this means that $\mathbb{P}\big((U_1, \ldots, U_d)' \in B\big) = 0$ for every Borel set $B \in \mathcal{B}\big([0,1]^d\big)$ having zero (d-dimensional) Lebesgue measure. This dissertation deals mostly with copulas that are **not** absolutely continuous. We consider copulas C whose induced measure dC can be decomposed into $dC = dC^{abs} + dC^{sing}$, where dC^{abs} is absolutely continuous with respect to the Lebesgue measure, but dC^{sing} is not. In this case, C is said to have a *singular component*. One example of such a copula was already presented in Subsection 2.2.1 above: the bivariate Cuadras-Augé copula assigns positive mass to the diagonal of the unit square, a set with zero 2-dimensional Lebesgue measure. A singular component often causes analytical difficulties when working with the respective copula. Nevertheless, the results of the present dissertation show that such distributions still can have interesting and useful properties.

Sometimes it is more convenient to describe the distribution of $(\tau_1, \ldots, \tau_d)'$ by means of

its survival function instead of its distribution function. Especially when the components τ_i are interpreted as lifetimes, this description is more intuitive. Letting $(\tau_1, \ldots, \tau_d)'$ be defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, its survival function is defined as

$$\bar{F}(t_1, \ldots, t_d) := \mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d), \quad t_1, \ldots, t_d \in \mathbb{R}.$$

For $i = 1, \ldots, d$ the univariate marginal survival function $\bar{F}_i := 1 - F_i$ of τ_i can be retrieved from $\bar{F}$ via

$$\bar{F}_i(t_i) = \mathbb{P}(\tau_i > t_i) = \bar{F}(-\infty, \ldots, -\infty, t_i, -\infty, \ldots, -\infty), \quad t_i \in \mathbb{R}.$$

Using $-\infty$ as arguments of $\bar{F}$ is a short-hand notation for taking the limits as the arguments $t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_d$ tend to $-\infty$. Analogously to Sklar's Theorem 2.2.3 a d-dimensional survival function can be decomposed into a copula and its marginal survival functions. We state this result for continuous margins only, since we only need this case.

Theorem 2.2.5 (Survival Analog of Sklar's Theorem)
Let $\bar{F}$ be a d-dimensional survival function with continuous marginal survival functions $\bar{F}_1, \ldots, \bar{F}_d$. Then there exists a unique d-dimensional copula $\hat{C}$ such that for all $(t_1, \ldots, t_d)' \in \mathbb{R}^d$ it holds that

$$\bar{F}(t_1, \ldots, t_d) = \hat{C}\big(\bar{F}_1(t_1), \ldots, \bar{F}_d(t_d)\big). \tag{2.5}$$

Conversely, if $\hat{C}$ is a d-dimensional copula and $\bar{F}_1, \ldots, \bar{F}_d$ are univariate continuous survival functions, then the function $\bar{F}$ defined via (2.5) is a d-dimensional survival function.

Proof
This well-known result is noted e.g. in [McNeil et al. (2005), p. 195-196]. Since the statement is very important for the present thesis, we provide a detailed proof in the sequel. For the univariate distribution functions F_i their *generalized inverses* F_i^{-1} are defined by

$$F_i^{-1} : (0,1) \to \mathbb{R}, \quad F_i^{-1}(y) := \inf\{x \in \mathbb{R} : F_i(x) \geq y\}.$$

The following properties of the generalized inverse are required for the proof:

($*$) F_i continuous $\Rightarrow F_i \circ F_i^{-1}$ is the identity, see [McNeil et al. (2005), p.494, Propo-

sition A.3 (viii)]: we have to show that for all $y \in (0,1)$ it holds that

$$F_i\big(\inf\{x \in \mathbb{R} : F_i(x) \geq y\}\big) = y.$$

However, "$\geq$" follows trivially, and "$>$" would contradict the continuity of F_i.

$(**)$ F_i continuous $\Rightarrow F_i^{-1}$ strictly increasing, see [McNeil et al. (2005), p.494, Proposition A.3 (ii)]: by monotonicity of F_i, also F_i^{-1} is non-decreasing. Assume there exist $0 < y_1 < y_2 < 1$ such that $F_i^{-1}(y_2) = F_i^{-1}(y_1)$. Then by $(*)$ we have that

$$y_2 = F_i \circ F_i^{-1}(y_2) = F_i \circ F_i^{-1}(y_1) = y_1,$$

which is a contradiction.

$(***)$ If $\tau_i \sim F_i$, then $\tau_i \overset{d}{=} F_i^{-1} \circ F_i(\tau_i)$, see [McNeil et al. (2005), p.495, Proposition A.4]: for each $x \in \mathbb{R}$ it is trivial that

$$F_i^{-1} \circ F_i(x) = \inf\{z \in \mathbb{R} : F_i(z) \geq F_i(x)\} \leq x.$$

Furthermore, if τ_i is defined on $(\Omega, \mathcal{F}, \mathbb{P})$ it holds that

$$\begin{aligned}
\mathbb{P}\big(F_i^{-1} \circ F_i(\tau_i) < \tau_i\big) &= \mathbb{P}\big(\inf\{z \in \mathbb{R} : F_i(z) \geq F_i(\tau_i)\} < \tau_i\big) \\
&\leq \mathbb{P}\Big(\bigcup_{n\in\mathbb{N}} \big\{F_i(\tau_i - 1/n) = F_i(\tau_i)\big\}\Big) \\
&\leq \sum_{n\in\mathbb{N}} \mathbb{P}\big(F_i(\tau_i - 1/n) = F_i(\tau_i)\big) \\
&= \sum_{n\in\mathbb{N}} \mathbb{P}\Big(\big\{\omega \in \Omega \,\big|\, \mathbb{P}\big(\tau_i \in (\tau_i(\omega) - 1/n, \tau_i(\omega)]\big) = 0\big\}\Big) = 0.
\end{aligned}$$

This implies the claim.

Using $(*)-(***)$ above we can now prove the theorem. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $(\tau_1, \ldots, \tau_d)'$ have joint survival function $\bar{F}$ and continuous univariate survival functions $\bar{F}_1, \ldots, \bar{F}_d$. Then, it follows for each $i = 1, \ldots, d$ and $u \in (0,1)$ that

$$\begin{aligned}
\mathbb{P}\big(\bar{F}_i(\tau_i) \geq u\big) = \mathbb{P}\big(F_i(\tau_i) \leq 1-u\big) &\overset{(**)}{=} \mathbb{P}\big(F_i^{-1} \circ F_i(\tau_i) \leq F_i^{-1}(1-u)\big) \\
&\overset{(***)}{=} \mathbb{P}\big(\tau_i \leq F_i^{-1}(1-u)\big) = F_i \circ F_i^{-1}(1-u) \overset{(*)}{=} 1-u.
\end{aligned}$$

Hence, each $\bar{F}_i(\tau_i)$ is $U[0,1]$-distributed. Denote by $\hat{C}$ the joint distribution function of $\big(\bar{F}_1(\tau_1),\dots,\bar{F}_d(\tau_d)\big)'$, which is hence a copula by definition. Moreover, we compute for $t_1,\dots,t_d \in \mathbb{R}$ that

$$\begin{aligned}
\hat{C}\big(\bar{F}_1(t_1),\dots,\bar{F}_d(t_d)\big) &= \mathbb{P}\big(\bar{F}_i(\tau_i) \le \bar{F}_i(t_i)\ \forall\, i = 1,\dots,d\big) \\
&= \mathbb{P}\big(F_i(\tau_i) \ge F_i(t_i)\ \forall\, i = 1,\dots,d\big) \\
&= \mathbb{P}\big(F_i(\tau_i) \ge F_i(t_i)\ \forall\, i = 1,\dots,d,\ \tau_i \ge t_i\ \forall\, i = 1,\dots,d\big) \\
&\qquad + \mathbb{P}\big(F_i(\tau_i) \ge F_i(t_i)\ \forall\, i = 1,\dots,d,\ \exists\, j \in \{1,\dots,d\} : \tau_j < t_j\big) \\
&= \mathbb{P}\big(\tau_i \ge t_i\ \forall\, i = 1,\dots,d\big) \\
&\qquad + \mathbb{P}\big(\exists\, j \in \{1,\dots,d\} : F_i(\tau_i) \ge F_i(t_i)\ \forall\, i \ne j,\ \tau_j < t_j,\ F_j(\tau_j) = F_j(t_j)\big) \\
&= \mathbb{P}\big(\tau_i \ge t_i\ \forall\, i = 1,\dots,d\big) + 0 = \bar{F}(t_1,\dots,t_d).
\end{aligned}$$

The fourth equality in the above computation uses the fact that each F_i is non-decreasing, the fifth equality holds true, since $F_j(\tau_j)$ is $U[0,1]$-distributed and has thus no atoms, and the last equality holds, since all F_i are continuous and hence $\mathbb{P}(\tau_i = t_i) = 0$ for all i. Uniqueness of the copula $\hat{C}$ is shown as follows: assume there is another copula $\tilde{C}$ satisfying (2.5). Let $u_1,\dots,u_d \in [0,1]$, then by continuity of the margins there exist $t_1,\dots,t_d$ such that $\bar{F}_i(t_i) = u_i$ for all i. Accordingly, it follows that

$$\hat{C}(u_1,\dots,u_d) = \hat{C}\big(\bar{F}_1(t_1),\dots,\bar{F}_d(t_d)\big) = \tilde{C}\big(\bar{F}_1(t_1),\dots,\bar{F}_d(t_d)\big) = \tilde{C}(u_1,\dots,u_d).$$

Thus $\tilde{C} = \hat{C}$ and the first part of the theorem is established. Conversely, let $\hat{C}$ be a given copula and $\bar{F}_1,\dots,\bar{F}_d$ continuous univariate marginal survival functions. Consider a probability space $(\Omega,\mathcal{F},\mathbb{P})$, on which a random vector $(U_1,\dots,U_d)' \sim \hat{C}$ is defined. Then, we define the random vector

$$(\tau_1,\dots,\tau_d)' := \big(F_1^{-1}(1-U_1),\dots,F_d^{-1}(1-U_d)\big)'.$$

For each $i = 1,\dots,d$ and $t \in \mathbb{R}$ it holds almost surely that

$$\tau_i > t \;\Leftrightarrow\; \inf\{x \in \mathbb{R} : F_i(x) \ge 1 - U_i\} > t \;\Leftrightarrow\; F_i(t) < 1 - U_i.$$

Hence, it follows for $t_1,\dots,t_d \in \mathbb{R}$ that

$$\mathbb{P}\big(\tau_i > t_i\ \forall\, i\big) = \mathbb{P}\big(U_i < \bar{F}_i(t_i)\ \forall\, i\big) = \hat{C}\big(\bar{F}_1(t_1),\dots,\bar{F}_d(t_d)\big),$$

establishing the claim. □

Due to uniqueness, the copula $\hat{C}$ in Theorem 2.2.5 is called **the** *survival copula* of the random vector $(\tau_1, \ldots, \tau_d)'$ with survival function $\bar{F}$. It is important to stress that the survival copula $\hat{C}$ is a proper copula, i.e. a distribution function and not a survival function. Knowing the copula of a random vector allows to compute its survival copula, and vice versa. This computation is accomplished by the principle of inclusion and exclusion given below. It is a standard result from probability calculus and shows how to compute the probabilty of a union of events in terms of probabilities of intersections.

Lemma 2.2.6 (Principle of Inclusion and Exclusion)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. With $A_1, \ldots, A_n \in \mathcal{F}$ it holds that

$$\mathbb{P}\big(\cup_{i=1}^n A_i\big) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \le i_1 < \ldots < i_k \le n} \mathbb{P}\big(\cap_{j=1}^k A_{i_j}\big).$$

Proof
See e.g. [Billingsley (1995), p. 24]. □

Given the copula C of a random vector $(\tau_1, \ldots, \tau_d)'$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with continuous margins $F_1, \ldots, F_d$, its survival copula $\hat{C}$ is computed as follows: as in the proof of Theorem 2.2.5 we introduce the notation $F_i^{-1}(t) := \inf\{x \in \mathbb{R} \,|\, F_i(x) \ge t\}$, $t \in (0,1)$, for the *generalized inverse* of F_i, $i = 1, \ldots, d$. The continuity of the margins implies that $F_i \circ F_i^{-1}(t) = t$ for all $t \in (0,1)$ and $i = 1, \ldots, d$, see $(*)$ in the proof of Theorem 2.2.5. For $u_1, \ldots, u_d \in (0,1)$ it follows that

$$\begin{aligned}
\hat{C}(u_1, \ldots, u_d) &= \hat{C}\Big(\bar{F}_1\big(F_1^{-1}(1-u_1)\big), \ldots, \bar{F}_d\big(F_d^{-1}(1-u_d)\big)\Big) \\
&= \mathbb{P}\big(\tau_1 > F_1^{-1}(1-u_1), \ldots, \tau_d > F_d^{-1}(1-u_d)\big) \\
&= \mathbb{P}\Big(\cap_{k=1}^d \big\{\tau_k > F_k^{-1}(1-u_k)\big\}\Big) = 1 - \mathbb{P}\Big(\cup_{k=1}^d \big\{\tau_k \le F_k^{-1}(1-u_k)\big\}\Big) \\
&= 1 - \sum_{k=1}^d (-1)^{k+1} \sum_{1 \le j_1 < \ldots < j_k \le d} \mathbb{P}\big(\cap_{i=1}^k \big\{\tau_{j_i} \le F_{j_i}^{-1}(1-u_{j_i})\big\}\big) \\
&= 1 + \sum_{k=1}^d (-1)^k \sum_{1 \le j_1 < \ldots < j_k \le d} C_{j_1, \ldots, j_k}\Big(F_{j_1}\big(F_{j_1}^{-1}(1-u_{j_1})\big), \ldots, F_{j_k}\big(F_{j_k}^{-1}(1-u_{j_k})\big)\Big)
\end{aligned}$$

$$= 1 + \sum_{k=1}^{d} (-1)^k \sum_{1 \le j_1 < \ldots < j_k \le d} C_{j_1,\ldots,j_k}\big(1 - u_{j_1}, \ldots, 1 - u_{j_k}\big). \tag{2.6}$$

In the above computation, the second and the fifth equality follow from Theorems 2.2.5 and 2.2.3, respectively. Interchanging the roles of $\hat{C}$ and C yields by a similar computation that

$$C(u_1, \ldots, u_d) = 1 + \sum_{k=1}^{d} (-1)^k \sum_{1 \le j_1 < \ldots < j_k \le d} \hat{C}_{j_1,\ldots,j_k}\big(1 - u_{j_1}, \ldots, 1 - u_{j_k}\big).$$

An alternative view on the copula C and the survival copula $\hat{C}$ of a random vector $(\tau_1, \ldots, \tau_d)'$ with continuous margins $F_1, \ldots, F_d$ can be extracted from the proofs of Theorems 2.2.3 and 2.2.5: C is the distribution function of $\big(F_1(\tau_1), \ldots, F_d(\tau_d)\big)'$ and $\hat{C}$ is the distribution function of $\big(\bar{F}_1(\tau_1), \ldots, \bar{F}_d(\tau_d)\big)'$. Finally, the copula C (and hence also the survival copula $\hat{C}$ by the above computation) of a random vector $(\tau_1, \ldots, \tau_d)'$ with continuous margins is invariant under strictly increasing transformations. More precisely, for strictly increasing functions $g_1, \ldots, g_d : \mathbb{R} \to \mathbb{R}$, the copula of $\big(g_1(\tau_1), \ldots, g_d(\tau_d)\big)'$ is again C, see [Embrechts et al. (2003), Theorem 2.6]. This fact is often used to transform or standardize the marginal laws without changing the copula.

2.2.3 Dependence Measures

A bivariate copula can be visualized by a three-dimensional function plot. In two or three dimensions it is still possible to visualize properties of a copula by means of a two- or three-dimensional scatterplot. In larger dimensions, however, it is difficult to visualize and understand the dependence structure of a given copula. In this regard, it is convenient to know a probabilistic interpretation of the copula which accompanies the analytical form and explains how the dependence structure emerges naturally. In fact, we have not even defined a copula analytically in this thesis, but rather directly as a distribution function. Alternatively, properties of (high-dimensional) copulas can be quantified via certain dependence measures. Popular examples for multivariate dependence measures are d-dimensional extensions of the bivariate Spearman's Rho, Kendall's Tau, and Blomqvist's Beta, see e.g. the references [Wolff (1980), Schmid, Schmidt (2006), Schmid, Schmidt (2007a), Schmid, Schmidt (2007b)], or extremal dependence coefficients as considered e.g. in [Frahm (2006)]. We introduce some

of them in the sequel, since they are applied later on. The following notion is an intuitive measure of "extremal" dependence.

Definition 2.2.7 (Upper- and Lower-Tail Dependence)
For a bivariate copula C the coefficients of *upper- and lower-tail dependence* UTD_C and LTD_C are defined as

$$UTD_C := \lim_{u\uparrow 1} \frac{C(u,u) - 2\,u + 1}{1-u}, \quad LTD_C := \lim_{u\downarrow 0} \frac{C(u,u)}{u},$$

if the respective limit exists.

The intuition behind Definition 2.2.7 is that for a random vector $(U_1, U_2)'$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose joint distribution function is the copula C, it holds that

$$\begin{aligned} UTD_C &= \lim_{u\uparrow 1} \frac{C(u,u) - 2\,u + 1}{1-u} \\ &= \lim_{u\uparrow 1} \frac{\mathbb{P}(U_1 \le u, U_2 \le u) - \mathbb{P}(U_1 \le u) - \mathbb{P}(U_2 \le u) + 1}{\mathbb{P}(U_2 > u)} \\ &\overset{L.2.2.6}{=} \lim_{u\uparrow 1} \frac{\mathbb{P}(U_1 > u, U_2 > u)}{\mathbb{P}(U_2 > u)} = \lim_{u\uparrow 1} \mathbb{P}(U_1 > u \,|\, U_2 > u). \end{aligned}$$

Thus, the coefficient of upper-tail dependence equals the probability that U_1 is large given U_2 is large. Similarly, LTD_C equals the probability that U_1 is small given U_2 is small. Positive upper- or lower-tail dependence are desirable in stochastic models that support extreme scenarios. A bivariate Gaussian copula is a popular example for a distribution whose tail dependencies are both zero, see [McNeil et al. (2005), p. 211]. As a consequence, models based on normality assumptions are often critizised for their lack of extreme scenarios.

Remark 2.2.8 (Tail Dependence for Arbitrary Vectors)
The notions of upper- and lower-tail dependence are more generally defined for arbitrary bivariate random vectors $(\tau_1, \tau_2)'$ with continuous, but not necessarily uniform, marginal laws F_1, F_2. The definition in this more general case is the same as above, where the copula C is the copula of the corresponding random vector. More clearly, the lower-tail dependence coefficient of the bivariate random vector $(\tau_1, \tau_2)'$, defined on $(\Omega, \mathcal{F}, \mathbb{P})$, is given by

$$\lim_{t\downarrow 0} \mathbb{P}\big(\tau_1 \le F_1^{-1}(t) \,\big|\, \tau_2 \le F_2^{-1}(t)\big) = \lim_{u\downarrow 0} \frac{C(u,u)}{u} = LTD_C,$$

provided existence of the limit. For the upper-tail dependence coefficient the corresponding similar definition of UTD_C applies.

Remark 2.2.9 (Tail Dependence of Survival Copula)
Existence provided, the upper-tail dependence coefficient of the survival copula $\hat{C}$ of a random vector equals the lower-tail dependence coefficient of the respective copula, i.e. $UTD_{\hat{C}} = LTD_C$. This follows from the computation (2.6):

$$\begin{aligned} UTD_{\hat{C}} &= \lim_{u\uparrow 1} \frac{\hat{C}(u,u) - 2\,u + 1}{1-u} \overset{(2.6)}{=} \lim_{u\uparrow 1} \frac{1 + C(1-u,1-u) - 2\,(1-u) - 2\,u + 1}{1-u} \\ &= \lim_{u\downarrow 0} \frac{C(u,u)}{u} = LTD_C. \end{aligned}$$

Similarly, $LTD_{\hat{C}} = UTD_C$, existence provided.

Example 2.2.10 (Tail Dependence of Cuadras-Augé Copula)
In the case of a bivariate Cuadras-Augé copula C_α as introduced in Definition 2.2.2 one may check, using the rule of L'Hospital, that $UTD_{C_\alpha} = \alpha$, since

$$\begin{aligned} UTD_{C_\alpha} &= \lim_{u\uparrow 1} \frac{C_\alpha(u,u) - 2\,u + 1}{1-u} \\ &= \lim_{u\uparrow 1} \frac{u^{1+1-\alpha} - 2\,u + 1}{1-u} = \lim_{u\uparrow 1} \frac{(2-\alpha)\,u^{1-\alpha} - 2}{-1} = \alpha. \end{aligned}$$

In contrast, C_α has zero lower-tail dependence (unless $\alpha = 1$):

$$LTD_{C_\alpha} = \lim_{u\downarrow 0} \frac{u^{1+1-\alpha}}{u} = \lim_{u\downarrow 0} u^{1-\alpha} = \mathbf{1}_{\{\alpha=1\}}.$$

Asymmetric tail dependence parameters, i.e. $UTD_{C_\alpha} \neq LTD_{C_\alpha}$, are sometimes desirable. For example if C_α is the survival copula of two companies' bankruptcy times, then positive upper-tail dependence of C_α - correspondingly lower-tail dependence of the bankruptcy times by Remark 2.2.9 - has an intuitive interpretation. It implies that an early default of one firm is likely to coincide with an early default of the other firm as well. In contrast, zero lower-tail dependence of C_α means that an extraordinary long survival of one firm does not automatically induce a long life of the other.

One possible extension of the bivariate concept of tail dependence to dimensions $d > 2$ is introduced by [Frahm (2006)] and is presented below.

Definition 2.2.11 (Upper- and Lower-Extremal Dependence Coefficient)
For a d-dimensional copula $C : [0,1]^d \to [0,1]$ the *upper- and lower-extremal dependence coefficients* $UEDC_C$ and $LEDC_C$ are defined by

$$UEDC_C := \lim_{u \uparrow 1} \frac{1 + \sum_{i=1}^{d} (-1)^i \sum_{1 \le j_1 < \ldots < j_i \le d} C_{j_1,\ldots,j_i}(u,\ldots,u)}{1 - C(u,\ldots,u)},$$
$$LEDC_C := \lim_{u \downarrow 0} \frac{C(u_1,\ldots,u_d)}{\sum_{i=1}^{d} (-1)^{i+1} \sum_{1 \le j_1 < \ldots < j_i \le d} C_{j_1,\ldots,j_i}(u,\ldots,u)},$$

if the corresponding limit exists.

Using the principle of inclusion and exclusion it is possible to understand the intuition behind this definition. Letting $(U_1,\ldots,U_d)'$ be defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and have as joint distribution function the copula C, it holds that

$$\begin{aligned}
UEDC_C &= \lim_{u \uparrow 1} \frac{1 + \sum_{i=1}^{d} (-1)^i \sum_{1 \le j_1 < \ldots < j_i \le d} C_{j_1,\ldots,j_i}(u,\ldots,u)}{1 - C(u,\ldots,u)} \\
&= \lim_{u \uparrow 1} \frac{1 + \sum_{i=1}^{d} (-1)^i \sum_{1 \le j_1 < \ldots < j_i \le d} \mathbb{P}(U_{j_1} \le u,\ldots,U_{j_i} \le u)}{1 - \mathbb{P}(U_1 \le u,\ldots,U_d \le u)} \\
&\overset{L.2.2.6}{=} \lim_{u \uparrow 1} \frac{\mathbb{P}(U_1 > u,\ldots,U_d > u)}{\mathbb{P}\big(\max\{U_1,\ldots,U_d\} > u\big)} \\
&= \lim_{u \uparrow 1} \mathbb{P}\big(\min\{U_1,\ldots,U_d\} > u \,\big|\, \max\{U_1,\ldots,U_d\} > u\big).
\end{aligned}$$

Thus, the $UEDC$ gives the probability that all components $U_1,\ldots,U_d$ are large given at least one of them is large. Similarly, the $LEDC$ gives the probability that all components are small given at least one is small. Let us point out that in the case $d = 2$ the $UEDC$ is **not** the same as the UTD, the difference being that one conditions on the maximum of U_1, U_2 to be greater than u, instead of just conditioning on U_2. There are other versions of upper-extremal dependence coefficients which really extend the bivariate UTD, compare e.g. [Li (2008)], however the $UEDC$ as defined above seems most adequate for our purpose. Concluding, the UTD as well as the $UEDC$ are limits of probabilities, hence they take values in $[0,1]$, and the closer they are to one, the more likely are extreme events. Similar interpretations hold for LTD and $LEDC$.

The following multivariate dependence measure, which is a multivariate rank correlation coefficient, is due to [Wolff (1980)].

Definition 2.2.12 (Multivariate Spearman's Rho)
For a given copula $C : [0,1]^d \to [0,1]$ *Spearman's multivariate Rho* is defined by

$$\rho_C^d := \frac{\int_{[0,1]^d} C(u_1,\ldots,u_d)\, d(u_1,\ldots,u_d) - \int_{[0,1]^d} \Pi(u_1,\ldots,u_d)\, d(u_1,\ldots,u_d)}{\int_{[0,1]^d} M(u_1,\ldots,u_d)\, d(u_1,\ldots,u_d) - \int_{[0,1]^d} \Pi(u_1,\ldots,u_d)\, d(u_1,\ldots,u_d)}.$$

Solving the integrals over Π and M in the definition of ρ_C^d, it follows that

$$\rho_C^d = \frac{d+1}{2^d-(d+1)}\Big(2^d \int_{[0,1]^d} C(u_1,\ldots,u_d)\, d(u_1,\ldots,u_d) - 1\Big),$$

see e.g. [Schmid, Schmidt (2006)]. The value ρ_C^d can be interpreted as the normalized average distance between the copula C and the independence copula Π. The normalization above makes sense, since it can be shown that any copula C satisfies $C \leq M$ pointwise, see e.g. [Embrechts et al. (2003), Theorem 2.3, p. 5]. The case $d = 2$ agrees with the commonly used Pearson's correlation coefficient, see [Nelsen (1999), p. 137]. Spearman's Rho ρ_C^d is a so-called *concordance measure*, see e.g. [Embrechts et al. (2003), p. 12] for further details.

2.3 Important Copula Families

Three types of copulas are introduced: extreme-value copulas, Marshall-Olkin survival copulas and Archimedean copulas. All of them are important for later results.

2.3.1 Extreme-Value Copulas

The notion of an extreme-value copula is interesting for the present thesis, since we study a particular class of copulas, which satisfies the extreme-value property.

Definition 2.3.1 (Extreme-Value Copula)
A copula $C : [0,1]^d \to [0,1]$ is called *extreme-value copula* if it satisfies the extreme-value property $C(u_1^t,\ldots,u_d^t) = C(u_1,\ldots,u_d)^t$ for all $t > 0$, $u_1,\ldots,u_d \in [0,1]$.

Specific examples of extreme-value copulas include the independence copula Π, the upper-Fréchet-Hoeffding bound M, as well as Marshall-Olkin survival copulas as discussed in Subsection 2.3.2 below. Such copulas occur in multivariate extreme-value the-

ory as possible limit copulas. More precisely, let F be a multivariate distribution function with continuous margins $F_1, \ldots, F_d$. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting i.i.d. random vectors $\{(\tau_1^{(n)}, \ldots, \tau_d^{(n)})'\}_{n\in\mathbb{N}}$ with distribution function F. For each $n \in \mathbb{N}$ one considers the vector of componentwise maxima $(M_{1:n}, \ldots, M_{d:n})'$, where $M_{k:n} := \max\{\tau_k^{(1)}, \ldots, \tau_k^{(n)}\}$, $k = 1, \ldots, d$. If there exist sequences $\{a_{1:n}\}_{n\in\mathbb{N}}, \ldots, \{a_{d:n}\}_{n\in\mathbb{N}}$, $\{b_{1:n}\}_{n\in\mathbb{N}}, \ldots, \{b_{d:n}\}_{n\in\mathbb{N}}$ and a random vector $(\tau_1, \ldots, \tau_d)'$ such that

$$\Big(\frac{M_{1:n} - a_{1:n}}{b_{1:n}}, \ldots, \frac{M_{d:n} - a_{d:n}}{b_{d:n}}\Big)' \xrightarrow{d} (\tau_1, \ldots, \tau_d)', \quad \text{as } n \to \infty, \tag{2.7}$$

then the copula of the limit random vector $(\tau_1, \ldots, \tau_d)'$ must be of extreme-value kind. A proof of this fact can be found e.g. in [Joe (1997), p. 172-174]. Conversely, given an extreme-value copula C, it may occur as the copula of a limit random vector as in (2.7): consider i.i.d. random vectors $\{(U_1^{(n)}, \ldots, U_d^{(n)})'\}_{n\in\mathbb{N}}$ with joint distribution function C on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $n \in \mathbb{N}$ denote by $M_{k:n} := \max\{U_k^{(1)}, \ldots, U_k^{(n)}\}$ the componentwise maxima, for $k = 1, \ldots.d$. Choosing $a_{k:n} \equiv 1$ and $b_{k:n} = 1/n$ for all $k \in 1, \ldots, d$ and $n \in \mathbb{N}$, it is observed for $t_1, \ldots, t_d \geq 0$ that

$$\begin{aligned}
&\mathbb{P}\Big(\frac{M_{1:n} - 1}{1/n} \leq -t_1, \ldots, \frac{M_{d:n} - 1}{1/n} \leq -t_d\Big) \\
&\quad = \mathbb{P}\Big(U_1^{(1)} \leq -\frac{t_1}{n} + 1, \ldots, U_d^{(1)} \leq -\frac{t_d}{n} + 1\Big)^n = C\Big(-\frac{t_1}{n} + 1, \ldots, -\frac{t_d}{n} + 1\Big)^n \\
&\quad = C\Big(\Big(-\frac{t_1}{n} + 1\Big)^n, \ldots, \Big(-\frac{t_d}{n} + 1\Big)^n\Big) \quad \overset{n\to\infty}{\longrightarrow} \quad C\big(e^{-t_1}, \ldots, e^{-t_d}\big).
\end{aligned}$$

The third equality in the computation above requires the extreme-value property of C. Notice that the entries $\exp(-t_k)$, $k = 1, \ldots, d$, of C in the limit constitute univariate extreme-value distribution functions of the Weibull kind as t_k ranges in $[0, \infty)$, see [Joe (1997), p. 170]. Hence, C occurs as a possible limit copula in (2.7).

From an analytical perspective, each extreme-value copula admits a so-called *Pickands representation*. This characterizes an extreme-value copula by a measure on the d-dimensional unit simplex subject to certain boundary conditions.

Theorem 2.3.2 (De Haan, Resnick (1977), Pickands (1981))
$C(u_1, \ldots, u_d)$ is an extreme-value copula if and only if there exists a (positive) finite measure δ on the d-dimensional unit simplex

$$S_d := \big\{(u_1, \ldots, u_d)' \in [0,1]^d \,\big|\, u_1 + \ldots + u_d = 1\big\},$$

subject to the conditions $\int_{S_d} u_j \, \delta(du_1, \ldots, du_d) = 1$, $j = 1, \ldots, d$, such that

$$C(u_1, \ldots, u_d) = \Big(\prod_{i=1}^{d} u_i \Big)^{P\left(\frac{\log u_1}{\sum_{k=1}^{d} \log u_k}, \ldots, \frac{\log u_d}{\sum_{k=1}^{d} \log u_k} \right)},$$

where P, called *Pickands dependence function*, is defined on S_d and given by

$$P(w_1, \ldots, w_d) = \int_{S_d} \max\{u_1 \, w_1, \ldots, u_d \, w_d\} \, \delta(du_1, \ldots, du_d).$$

Proof
This theorem is named after [Pickands (1981)], even though it is in turn based on [De Haan, Resnick (1977)]. Since the proof is important for our purpose but requires much effort, we do not skip it but only sketch the basic idea. This proof sketch can be found in the Appendix. □

Theorem 2.3.2 can be used to derive expressions for tail dependence parameters of general extreme-value copulas based on the measure δ, see [Li (2009)]. Moreover, Theorem 2.3.2 is useful to construct parametric families of extreme-value copulas. It is noted in [Joe (1997), p. 175] that for statistical applications "a goal is to find finite-dimensional parametric subfamiles that cover well the entire [infinite-dimensional] family [of extreme-value distributions]". The present dissertation deals with one specific parametric subfamily, called Marshall-Olkin survival copulas, see Subsection 2.3.2 below.

Example 2.3.3 (Pickands Dependence Function of Cuadras-Augé Copulas)
Consider the bivariate copula C_α from Definition 2.2.2. The extreme-value property is easily verified to hold for C_α, so in regard of Theorem 2.3.2 it is natural to ask what the measure δ and the corresponding dependence function P look like. It is derived in [Falk et al. (2004), Example 4.3.2, p. 124] that δ is a discrete measure, whose mass is concentrated on three points. More precisely, it is determined by

$$\delta\big(\{(1,0)'\}\big) = \delta\big(\{(0,1)'\}\big) = 1 - \alpha, \quad \delta\big(\{(1/2,1/2)'\}\big) = 2\,\alpha.$$

One verifies that the boundary conditions are valid, i.e.

$$\int_{S_2} u_1 \, \delta(du_1, du_2) = \int_{S_2} u_2 \, \delta(du_1, du_2) = (1-\alpha) \cdot 1 + 2\,\alpha \cdot \frac{1}{2} = 1.$$

Moreover, the Pickands dependence function P is computed to have the form

$$\begin{aligned} P(w_1, w_2) &= \int_{S_2} \max\{u_1\, w_1, u_2\, w_2\}\, \delta(du_1, du_2) \\ &= (1-\alpha)\, w_1 + (1-\alpha)\, w_2 + 2\,\alpha\, \frac{1}{2}\, \max\{w_1, w_2\} \\ &= \max\{w_1, w_2\} + (1-\alpha)\, \min\{w_1, w_2\}. \end{aligned}$$

Indeed, one may check that for $u_1, u_2 \in (0,1)$

$$P\Big(\frac{\log u_1}{\log(u_1\, u_2)}, \frac{\log u_2}{\log(u_1\, u_2)}\Big) = \frac{\log\big(\min\{u_1, u_2\}\big)}{\log(u_1\, u_2)} + (1-\alpha)\, \frac{\log\big(\max\{u_1, u_2\}\big)}{\log(u_1\, u_2)}.$$

Hence, it is verified that

$$(u_1\, u_2)^{P\left(\frac{\log u_1}{\log(u_1\, u_2)}, \frac{\log u_2}{\log(u_1\, u_2)}\right)} = e^{\log\left(\min\{u_1, u_2\}\right) + \log\left(\max\{u_1, u_2\}^{1-\alpha}\right)} = C_\alpha(u_1, u_2).$$

Since the Pickands dependence function $P(w_1, w_2)$ is defined only for nonnegative w_1, w_2 with $w_1 + w_2 = 1$, one may alternatively parameterize it by $w \in [0,1]$ setting $\tilde{P}(w) := P(w, 1-w)$. The function $\tilde{P}$ can then easily be visualized for different choices of $\alpha \in [0,1]$, see Figure 2.3. The fact that $\tilde{P}$ is not differentiable in $w = 1/2$ for $\alpha > 0$ indicates that C_α has a singular component in this case, see [Joe (1997), Theorem 6.5, p. 176].

2.3.2 Marshall-Olkin Survival Copulas

It is a well-known fact that a random variable τ with support $(0, \infty)$, defined on $(\Omega, \mathcal{F}, \mathbb{P})$, is exponentially distributed if and only if for all $s, t > 0$ it holds that

$$\mathbb{P}(\tau > s + t \,|\, \tau > t) = \mathbb{P}(\tau > s), \tag{2.8}$$

see e.g. [Billingsley (1995), p. 190]. In other words, the so-called *lack of memory property* (2.8) characterizes the exponential distribution. In fact, it is precisely this property which renders the exponential law one of the most popular and most tractable probability distributions. For instance it is noted in [David, Nagaraja (1970), p. 121] that "the exponential [distribution] occupies as commanding a position in life testing as does the normal [distribution] elsewhere in parametric theory." The article [Marshall, Olkin (1967)] is concerned with a multivariate analog of the lack of memory property. In general-

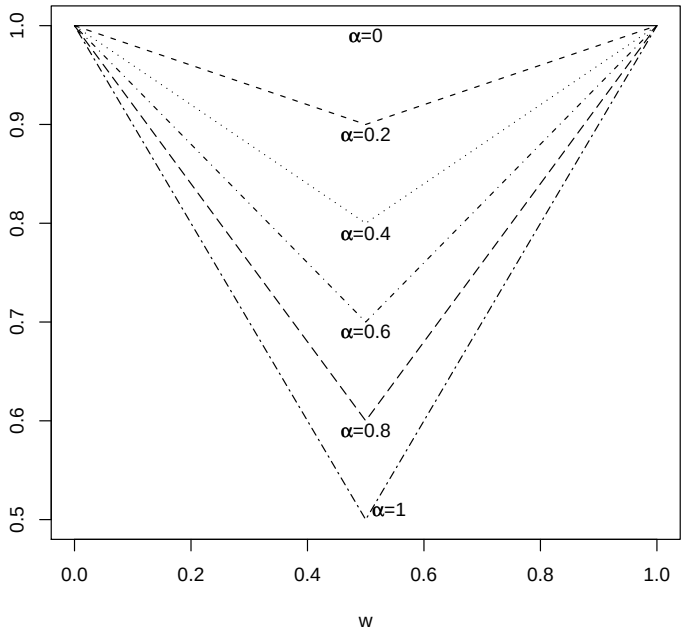


Figure 2.3 For $\alpha = 0, 0.2, 0.4, 0.6, 0.8, 1$, the Pickands dependence function $\tilde{P}(w) = \max\{w, 1-w\} + (1-\alpha)\min\{w, 1-w\}$ of a bivariate Cuadras-Augé copula is illustrated.

ization of (2.8), they consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a random vector $(\tau_1, \ldots, \tau_d)'$ with support $(0, \infty)^d$ is defined, which satisfies the following property: for all $s_1, \ldots, s_d$, $t \geq 0$ it holds that

$$\mathbb{P}\big(\tau_1 > s_1 + t, \ldots, \tau_d > s_d + t \,\big|\, \tau_1 > t, \ldots, \tau_d > t\big) = \mathbb{P}\big(\tau_1 > s_1, \ldots, \tau_d > s_d\big). \qquad (2.9)$$

Interpreting $\tau_1, \ldots, \tau_d$ as lifetimes, (2.9) intuitively means that the residual lifetimes are independent of age. Similar as in the univariate case, [Marshall, Olkin (1967)] show that there is precisely one parametric family of multivariate distributions satisfying the lack of memory property. More clearly, (2.9) has to be read iteratively as follows: in dimension $d = 1$ we already know that (2.9) implies that τ_1 is exponential. In dimension $d = 2$ one therefore postulates that τ_1, τ_2 both are exponential and (2.9) holds. It is shown in [Marshall, Olkin (1967), Lemma 2.2] that this implies the existence of parameters

$\lambda_{\{1\}}, \lambda_{\{2\}}, \lambda_{\{1,2\}} \geq 0$, with $\lambda_{\{k\}} + \lambda_{\{1,2\}} > 0$, $k = 1, 2$, such that for $t_1, t_2 \geq 0$ one has

$$\bar{F}(t_1, t_2) := \mathbb{P}(\tau_1 > t_1, \tau_2 > t_2) = \exp\big(- \lambda_{\{1\}} t_1 - \lambda_{\{2\}} t_2 - \lambda_{\{1,2\}} \max\{t_1, t_2\}\big).$$

This distribution is called the bivariate *Marshall-Olkin distribution*. Inductively, in dimension $d \geq 2$ one can show that if all $(d-1)$-dimensional subvectors of $(\tau_1, \ldots, \tau_d)'$ have a Marshall-Olkin distribution, and if (2.9) is satisfied, then it follows that there exist parameters $\lambda_I \geq 0$, $\emptyset \neq I \subset \{1, \ldots, d\}$, with $\sum_{I: k \in I} \lambda_I > 0$, $k = 1, \ldots, d$, such that for all $t_1, \ldots, t_d \geq 0$ one has

$$\bar{F}(t_1, \ldots, t_d) := \mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d) = \exp\Big(- \sum_{\emptyset \neq I \subset \{1, \ldots, d\}} \lambda_I \max_{i \in I}\{t_i\}\Big), \qquad (2.10)$$

see [Marshall, Olkin (1967), p. 39]. This distribution is called (d-dimensional) *Marshall-Olkin distribution*. Other multivariate exponential distributions are proposed in the literature, see e.g. [Gumbel (1960)], which do not share the lack of memory property. In fact, plugging exponential marginal laws into an arbitrary copula yields a multivariate distribution by virtue of Sklar's Theorem, which one might call "exponential". However, the previous motivation by means of the lack of memory property suggests that not every distribution with exponential margins deserves the name "multivariate exponential distribution". From this perspective, the Marshall-Olkin exponential distribution is the "right" multivariate exponential distribution. For a nice treatment of the characterization of the Marshall-Olkin distribution by the lack of memory property, the interested reader is referred to the exposé [Galambos, Kotz (1978), p. 103-132].

So far, the Marshall-Olkin distribution is only introduced analytically. It is also shown in [Marshall, Olkin (1967)] how to construct such a distribution. The intuition is a system of initially fully functional components which are affected by exogenous shocks destroying them. The random vector of extinction times of the components exhibits the Marshall-Olkin distribution. A shock can hit one or more components at the same time, rendering these extinction times dependent. In particular, when a shock hits e.g. five components at a time, then all these five extinction times have the same value, i.e. the distribution has a singular component. This property together with the intuitive interpretation makes this kind of distributions interesting in financial applications such as risk management and credit risk modeling, see for instance [Embrechts et al. (2003), Giesecke (2003), Lindskog, McNeil (2003)]- and Chapter 6 of the present thesis.

In order to outline the construction of [Marshall, Olkin (1967)] we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each non-empty subset $\emptyset \neq I \subset \{1, \dots, d\}$ let E_I be an exponential random variable with mean $1/\lambda_I > 0$, and assume that all these $2^d - 1$ random variables are independent. Some λ_I are allowed to be zero, in which case we mean that $E_I \equiv \infty$ with probability one. However, we must guarantee that $\sum_{I:k\in I} \lambda_I > 0$ for all $k = 1, \dots, d$. This means that for each $k = 1, \dots, d$ there is at least one subset $I \subset \{1, \dots, d\}$ containing k such that λ_I is strictly positive. In this case, the following random variables are well-defined in $(0, \infty)$:

$$\tau_k := \min\big\{E_I \,\big|\, I \subset \{1, \dots, d\},\, k \in I\big\}, \quad k = 1, \dots, d. \tag{2.11}$$

The random vector $\big(\tau_1, \dots, \tau_d\big)'$ has the Marshall-Olkin distribution given by (2.10), which can be seen from the following computation with $t_1, \dots, t_d \geq 0$:

$$\begin{aligned}\bar{F}(t_1, \dots, t_d) &:= \mathbb{P}(\tau_1 > t_1, \dots, \tau_d > t_d) = \mathbb{P}\big(E_I > \max_{i\in I}\{t_i\}\ \forall\, \emptyset \neq I \subset \{1, \dots, d\}\big) \\ &= \prod_{\emptyset\neq I\subset\{1,\dots,d\}} e^{-\lambda_I \max_{i\in I}\{t_i\}} = \exp\Big(-\sum_{\emptyset\neq I\subset\{1,\dots,d\}} \lambda_I \max_{i\in I}\{t_i\}\Big).\end{aligned}$$

For each $k = 1, \dots, d$ the distribution of τ_k is exponential with parameter $O_k := \sum_{I:k\in I} \lambda_I > 0$. This follows from the fact that the minimum of independent exponential random variables is again exponentially distributed, and the parameters are simply added up, see also the proof of Lemma 2.3.4 below. Intuitively, the random variable E_I is interpreted as the arrival time of an exogenous shock affecting those components of $(\tau_1, \dots, \tau_d)'$ which are indexed by a number in I. Accordingly, the k-th component is destroyed when hit by the first shock E_I with $k \in I$, motivating the definition (2.11). The survival copula $\hat{C}$ of this random vector is computed in [Li (2008), Proposition 1]. For our purpose however, a slightly different expression is more appropriate, which can be deduced directly from (2.10), see also [Joe (1997), p. 192].

Lemma 2.3.4 (Survival Copula of the Marshall-Olkin Distribution)
The survival copula $\hat{C}$ of the random vector $(\tau_1, \dots, \tau_d)'$ as defined in (2.11) is given by

$$\hat{C}(u_1, \dots, u_d) = \prod_{\emptyset\neq I\subset\{1,\dots,d\}} \min_{k\in I}\Big\{u_k^{\frac{\lambda_I}{O_k}}\Big\}, \quad u_1, \dots, u_d \in [0, 1].$$

Proof
In a first step one verifies that the marginal laws are exponential: for the k-th marginal survival function $\bar{F}_k$ of τ_k, it holds that

$$\bar{F}_k(t) = \mathbb{P}(\tau_k > t) = \mathbb{P}\big(E_I > t \,\forall\, I \subset \{1,\dots,d\} \,:\, k \in I\big) = \exp(-O_k\, t), \quad t \geq 0.$$

Hence, it follows from (2.10) that

$$\bar{F}(t_1,\dots,t_d) = \prod_{\emptyset \neq I \subset \{1,\dots,d\}} \min_{k \in I} \Big\{ \bar{F}_k(t_k)^{\frac{\lambda_I}{O_k}} \Big\}, \quad t_1,\dots,t_d \geq 0.$$

By an application of Theorem 2.2.5 the claim is thus established. □

Marshall-Olkin survival copulas are the main source of interest during this dissertation. Their analytical form is further studied in Chapter 3. The probabilistic intuition behind Marshall-Olkin distributions is revisited in Chapter 4 from an alternative perspective.

2.3.3 Archimedean Copulas

A function $\varphi : [0,\infty) \to [0,1]$ is called a *possible Archimedean generator* if it satisfies $\varphi(0) = 1$, $\lim_{x\to\infty} \varphi(x) = 0$, and if φ is strictly decreasing on $\big[0, \inf\{x \geq 0 \,:\, \varphi(x) = 0\}\big)$.[2] A copula $C : [0,1]^d \to [0,1]$ is called *Archimedean* if it admits the functional form

$$C(u_1,\dots,u_d) = \varphi\big(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_d)\big), \tag{2.12}$$

with a possible Archimedean generator φ. For φ and its inverse $\varphi^{-1} : (0,1] \to [0,\infty)$ the conventions $\varphi(\infty) := 0$ and $\varphi^{-1}(0) := \inf\{x \geq 0 \,:\, \varphi(x) = 0\}$ are applied. One immediately recognizes that an Archimedean copula is invariant under permutations of its arguments.

A function φ has to be a possible Archimedean generator in order for equation (2.12) to possibly define a copula. However, for a given possible Archimedean generator φ the function C from equation (2.12) is not always a proper copula. A necessary and

[2] [McNeil, Nešlehová (2009)] call such a function an *Archimedean generator*, but in regard of Theorems 2.3.8 and 2.3.9 below this terminology is unintuitive. Therefore, we add the word *possible* to indicate that not all these functions are related to an Archimedean copula.

sufficient condition on φ is derived in [McNeil, Nešlehová (2009)] and uses the following definition.

Definition 2.3.5 (d-Monotone Function)
For $d \geq 2$, a function $\varphi : [0, \infty) \to \mathbb{R}$ is called *d-monotone* if it is differentiable on $(0, \infty)$ up to the order $d-2$, if the derivatives satisfy

$$(-1)^k \, \varphi^{(k)}(x) \geq 0, \quad k = 0, 1, \ldots, d-2, \quad \forall x > 0,$$

if further $(-1)^{d-2} \, \varphi^{(d-2)}$ is non-increasing and convex on $(0, \infty)$, and if φ is continuous at zero[3].

A function $\varphi : [0, \infty) \to \mathbb{R}$ is 2-monotone if and only if it is non-negative, non-increasing, convex and continuous at zero. Moreover, for $d \geq 3$, d-monotonicity implies $(d-1)$-monotonicity. Of particular interest are functions that are d-monotone for all $d \geq 2$, see the following definition.

Definition 2.3.6 (Completely Monotone Function)
A function $\varphi : [0, \infty) \to \mathbb{R}$ is called *completely monotone*, if it is continuous at zero, has derivatives of all orders on $(0, \infty)$ and

$$(-1)^k \, \varphi^{(k)}(x) \geq 0, \quad k \in \mathbb{N}_0, \quad \forall x > 0,$$

i.e. if φ is d-monotone for all $d \geq 2$.

Completely monotone functions arise naturally in probability theory as Laplace transforms of non-negative random variables, see Theorem 2.3.7 below.

Theorem 2.3.7 (Bernstein (1929))
A function $\varphi : [0, \infty) \to \mathbb{R}$ is completely monotone with $\varphi(0) = 1$ if and only if there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a non-negative random variable $W \in [0, \infty)$ with $\varphi(x) = \mathbb{E}[\exp(-x\,W)]$, $x \geq 0$, i.e. φ is the Laplace transform of W.

Proof
Originally in [Bernstein (1929)]. See also [Feller (1966), Theorem 1, p. 439]. □

In the setup of Theorem 2.3.7 above, $\mathbb{P}(W = 0) > 0$ implies that $\varphi(x) \geq \mathbb{P}(W = 0) > 0$ for all $x \geq 0$. In order for φ to define a possible Archimedean generator one has to have

[3] Continuity on $(0, \infty)$ is implied by the other conditions anyway.

$\lim_{x\to\infty} \varphi(x) = 0$, i.e. $\mathbb{P}(W = 0) = 0$. Also let us stress that $W \in [0, \infty)$, so in particular $W = \infty$ is not allowed in Theorem 2.3.7. This assumption would introduce a point of discontinuity of φ at zero.

Using Definitions 2.3.5 and 2.3.6, the following two theorems are the cornerstone of the theory on Archimedean copulas.

Theorem 2.3.8 (Kimberling (1974))
Let φ be a possible Archimedean generator. Then equation (2.12) defines a copula for all $d \geq 2$ if and only if φ is completely monotone.

Proof
Originally in [Kimberling (1974)]. It also follows from the more general Theorem 2.3.9, a proof of which can be found in [McNeil, Nešlehová (2009), Theorem 2]. □

Whereas Theorem 2.3.8 considers all dimensions $d \geq 2$ simultaneously, the following result keeps $d \geq 2$ fixed, which leads to a more general statement.

Theorem 2.3.9 (McNeil, Nešlehová (2009))
Let φ be a possible Archimedean generator. Then equation (2.12) defines a copula if and only if φ is d-monotone.

Proof
See [McNeil, Nešlehová (2009), Theorem 2]. □

In our opinion, there are three major reasons why Archimedean copulas enjoy great popularity in many applications - in particular in large dimensions $d \gg 2$.

- The first reason is their great level of flexibility. The distributional properties of two Archimedean copulas with different generators can be quite unequal. For example, some bivariate Archimedean copulas exhibit upper-tail dependence but no lower-tail dependence, whereas other bivariate Archimedean copulas have both. Moreover, since they are parameterized by a function, there are infinitely many degrees of freedom when specifying an Archimedean copula - at least in theory.

- The second reason is their convenient parametric form (2.12), which is analytically tractable whenever its generator φ is nice. In particular, distributional properties of an Archimedean copula often correspond to analytical properties of its generator φ, which are sometimes easier to explore. This is a striking argument, since

for general high-dimensional copulas a nice analytical representation is not often available. For example, the Marshall-Olkin survival copula from Lemma 2.3.4 has a quite complicated form. The Gaussian copula from Example 2.2.4 does not even have a closed-form expression.

- The third reason is that many Archimedean copulas allow for efficient sampling routines in each dimension $d \geq 2$, as will be outlined briefly. For a given completely monotone possible Archimedean generator φ one may construct a random vector with copula given by (2.12) as follows.

Theorem 2.3.10 (Marshall, Olkin (1988))
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a positive random variable $W > 0$ with Laplace transform φ and, independently, i.i.d. random variables $E_1, \ldots, E_d$ with $E_1 \sim Exp(1)$ are defined. Then, the survival copula of the random vector $(\tau_1, \ldots, \tau_d)'$, defined by $\tau_k := E_k/W$ for each $k = 1, \ldots, d$, is given by equation (2.12). Moreover, the survival function $\bar{F}_k$ of τ_k equals φ, for $k = 1, \ldots, d$. In mathematical terms, with C given by (2.12), it holds that

$$\mathbb{P}\Big(\frac{E_1}{W} > t_1, \ldots, \frac{E_d}{W} > t_d\Big) = C\big(\varphi(t_1), \ldots, \varphi(t_d)\big), \quad t_1, \ldots, t_d \geq 0.$$

Proof
This is due to [Marshall, Olkin (1988)] and the proof works as follows. For positive $t_1, \ldots, t_d > 0$ it holds that

$$\begin{aligned}\mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d) &= \mathbb{P}(E_1 > t_1 W, \ldots, E_d > t_d W) \\ &= \mathbb{E}\Big[\prod_{i=1}^{d} \mathbb{P}(E_i > t_i W \,|\, W)\Big] = \mathbb{E}\Big[\prod_{i=1}^{d} e^{-t_i W}\Big] = \mathbb{E}\Big[e^{-W \sum_{i=1}^{d} t_i}\Big] = \varphi\Big(\sum_{i=1}^{d} t_i\Big).\end{aligned}$$

For each $k = 1, \ldots, d$ the random variable τ_k has survival function φ since

$$\mathbb{P}(\tau_k > t) = \mathbb{P}(E_k > t W) = \mathbb{E}\big[\mathbb{P}(E_k > t W \,|\, W)\big] = \mathbb{E}[e^{-t W}] = \varphi(t), \quad t > 0.$$

Thus, Theorem 2.2.5 implies that the survival copula $\hat{C}$ of $(\tau_1, \ldots, \tau_d)'$ is given by

$$\hat{C}(u_1, \ldots, u_d) = \varphi\Big(\sum_{i=1}^{d} \varphi^{-1}(u_i)\Big).$$

This establishes the claim. □

The major application of this probabilistic construction is that it shows how to simulate random vectors which have an Archimedean copula with generator φ as joint distribution function. Using Theorem 2.3.10, this is possible whenever the random variable W with Laplace transform φ can be simulated. To generate a random vector $(U_1, \ldots, U_d)'$ with joint distribution function C from equation (2.12) one needs to simulate the random vector $(\tau_1, \ldots, \tau_d)'$ from Theorem 2.3.10 and set $U_i := \varphi(\tau_i)$, $i = 1, \ldots, d$. Note that one has to be able to evaluate the Laplace transform φ. The computational efficiency of this sampling routine grows linearly in the dimension d, since it requires one to only simulate d exponential random variables and once the random variable W. For a list of specific examples of sampling algorithms we refer to [Hofert (2007)].

A list of d-monotone Archimedean generators can be found in [Nelsen (1999)], or in [Charpentier, Segers (2009)]. We present two examples in the sequel to illustrate the previous statements.

Example 2.3.11 (Clayton Copula)
For $\theta > 0$ the function $\varphi(x) = (1 + x)^{-1/\theta}$, $x \geq 0$, is completely monotone. To see this using Theorem 2.3.7, consider a random variable W which is $\Gamma(1, 1/\theta)$-distributed, i.e. its density is given by

$$f_W(x) = \mathbf{1}_{\{x>0\}} \frac{1}{\Gamma(1/\theta)} x^{\frac{1}{\theta}-1} e^{-x}, \quad \Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \, dt.$$

Then the Laplace transform of W is given by $\mathbb{E}[\exp(-x\,W)] = \varphi(x)$, $x \geq 0$, see [Feller (1966), p. 503]. One easily observes that $\varphi^{-1}(y) = (1/y)^\theta - 1$, $y \in [0, 1]$. The resulting one-parametric family of Archimedean copulas is called the *Clayton family*, named after [Clayton (1978)], and has the form

$$C(u_1, \ldots, u_d) = \left(1 - d + \sum_{i=1}^d u_i^{-\theta}\right)^{-\frac{1}{\theta}}.$$

Using the sampling strategy from Theorem 2.3.10, scatterplots from a two- and three-dimensional Clayton copula with parameter $\theta = 2$ are presented in Figure 2.4.

Example 2.3.12 (Proper d-Monotone Generator)
In [McNeil, Nešlehová (2009)] one can find examples of d-monotone functions which are not completely monotone, i.e. which are not Laplace transforms of a positive random variable. It is carried out that an Archimedean generator φ is d-monotone if and only if

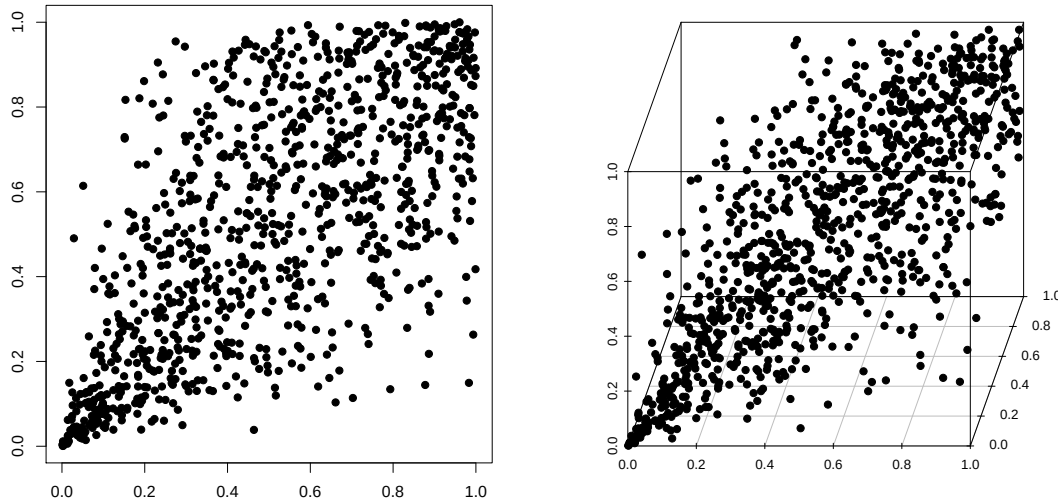


Figure 2.4 Scatterplots of 1000 samples from a Clayton copula in two and three dimensions, respectively. The parameter used is $\theta = 2$.

there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a positive random variable $W > 0$ such that $\varphi(x) = \mathbb{E}\big[\max\{1 - x/W, 0\}^{d-1}\big]$, $x \geq 0$. This probabilistic interpretation of d-monotone functions is due to [Williamson (1956)]. For example, $\varphi(x) = \max\{1 - x, 0\}$ is 2-monotone, but not d-monotone for $d \geq 3$. To generate random vectors from the corresponding two-dimensional Archimedean copula requires one to use a different probabilistic construction than that of Theorem 2.3.10. Since this approach is not applied in this thesis, the interested reader is referred to [McNeil, Nešlehová (2009)].

2.4 Lévy Subordinators

Lévy processes are continuous-time analoges of discrete-time random walks. The increments of a Lévy process are stationary and do not depend on the past, not even on the current value of the process. Stochastic processes of this type are used as building blocks for probabilistic models in many applications, e.g. in financial engineering. Prominent examples of Lévy processes are the Brownian motion and the Poisson process. If a Lévy process has almost surely non-decreasing paths, it is called a

Lévy subordinator. For further background on Lévy processes we refer the reader to the standard textbooks [Bertoin (1996), Bertoin (1999), Sato (1999), Schoutens (2003), Applebaum (2004), Cont, Tankov (2004)].

Definition 2.4.1 (Classical Lévy Subordinator)
A $[0,\infty)$-valued stochastic process $\Lambda = \{\Lambda_t\}_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a *classical Lévy subordinator* if it is a non-decreasing Lévy process, i.e. $\Lambda_0 = 0$ holds $\mathbb{P}$-almost surely, Λ has càdlàg paths[4], and the following conditions are satisfied:

(1) Λ is stochastically continuous, i.e.

$$\forall t \geq 0,\ \forall \epsilon > 0 \text{ it holds that } \lim_{h \downarrow 0} \mathbb{P}\big(|\Lambda_{t+h} - \Lambda_t| \geq \epsilon\big) = 0.$$

(2) Λ has independent increments, i.e. for all $t_0 \leq t_1 \leq \ldots \leq t_n$ the random variables $\Lambda_{t_0}, \Lambda_{t_1} - \Lambda_{t_0}, \ldots, \Lambda_{t_n} - \Lambda_{t_{n-1}}$ are stochastically independent.

(3) Λ has stationary increments, i.e. the law of $\Lambda_{t+h} - \Lambda_t$ is independent of $t \geq 0$ for each $h \geq 0$, i.e. $\Lambda_{t+h} - \Lambda_t \stackrel{d}{=} \Lambda_h$.

(4) $t \mapsto \Lambda_t$ is almost surely non-decreasing.

It is important to note that condition (1) in the definition above does not imply that Λ has continuous paths. It basically means that the jump times of paths of Λ are not allowed to be deterministic. The simplest example of a classical Lévy subordinator is a *(homogeneous) Poisson process* $N = \{N_t\}_{t\geq 0}$. It can be constructed as follows: on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $\{E_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables with $E_1 \sim Exp(\beta)$ for a parameter $\beta > 0$. Then, N is defined via

$$N_t := \sum_{n=1}^{\infty} \mathbf{1}_{\{E_1 + \ldots + E_n \leq t\}}, \quad t \geq 0. \tag{2.13}$$

Thus, a path of N starts at $N_0 = 0$, remains there until time E_1, and then jumps to one. It remains in state one until time $E_1 + E_2$ and then jumps to two, and so on. Thus, the state space of N is $\mathbb{N}_0$. The fact that N has i.i.d. increments is heavily related to the (lack of memory property of the) exponential distribution of the random variables $\{E_i\}_{i\in\mathbb{N}}$. Their parameter β is called the *intensity* of the Poisson process N. The nomenclature is justified by the fact that N_t is $Poi(\beta\, t)$-distributed for all $t > 0$.

[4]This means that $\mathbb{P}$-almost surely $\lim_{u\uparrow t} \Lambda_u$ exists $\forall\, t > 0$, and $\lim_{u\downarrow t} \Lambda_u = \Lambda_t\ \forall\, t \geq 0$.

For the purpose of this work it is convenient to extend Definition 2.4.1 to include the (absorbing) state infinity as possible value for Λ_t, $t > 0$.

Definition 2.4.2 (Lévy Subordinator)
A $[0, \infty) \cup \{\infty\}$-valued stochastic process $\Lambda = \{\Lambda_t\}_{t \geq 0}$ is called *Lévy subordinator* if it is defined for $t \geq 0$ by $\Lambda_t := \tilde{\Lambda}_t + \infty \cdot \mathbf{1}_{\{N_t \geq 1\}}$, where $\tilde{\Lambda} = \{\tilde{\Lambda}_t\}_{t \geq 0}$ is a classical ($[0, \infty)$-valued) Lévy subordinator and $N = \{N_t\}_{t \geq 0}$ is an independent Poisson process. The intensity of the Poisson process N is called the *killing rate* of Λ. As a convention, it is allowed to be zero in which case we mean that $\Lambda = \tilde{\Lambda}$.

In the literature, e.g. in [Applebaum (2004), Bertoin (1996)], a process Λ according to Definition 2.4.2 is sometimes called *killed subordinator*[5]. This emphasizes the intuitive interpretation that the process is "killed" when it jumps to infinity, since it remains in this state due to the non-decreasing paths. Adding the state ∞, sometimes called *cemetery state*, leads to a *compactification* of the state space $[0, \infty)$. Such an analytical technique is sometimes useful to include "marginal cases" in derivations; in fact, this will later be the case, compare e.g. Theorem 2.5.8. Due to such technical reasons, for the rest of this dissertation by the term "Lévy subordinator" we always refer to the extended Definition 2.4.2, i.e. we actually mean a killed subordinator.

Lévy processes are most easily treated by means of their characteristic function. Due to positiveness in case of a Lévy subordinator, i.e. a non-decreasing Lévy process, it is even more convenient to consider the (existing) Laplace transform. The analytical form of it is nowadays known as the *Lévy-Khinchin representation*, see Theorem 2.4.3 below. It relies on the fact that for each $t > 0$ the distribution of Λ_t, Λ being a classical Lévy subordinator, is *infinitely divisible*. More precisely, Definition 2.4.1 implies for all $n \in \mathbb{N}$ and $t > 0$ that

$$\Lambda_t = \Lambda_{\frac{t}{n}} + \left(\Lambda_{\frac{2\,t}{n}} - \Lambda_{\frac{t}{n}}\right) + \ldots + \left(\Lambda_{\frac{n\,t}{n}} - \Lambda_{\frac{(n-1)\,t}{n}}\right) \stackrel{d}{=} \Lambda^{(1)}_{\frac{t}{n}} + \Lambda^{(2)}_{\frac{t}{n}} + \ldots + \Lambda^{(n)}_{\frac{t}{n}},$$

where $\Lambda^{(i)}$, for $i \in \mathbb{N}$, are independent copies of Λ. I.e. the random variable Λ_t can be represented in distribution as the sum of n i.i.d. random variables for each n, a distributional property called "infinite divisibility", see [Sato (1999)] for further details. Conversely, given an arbitrary infinitely divisible distribution on $[0, \infty)$, there exists a

[5] But not always. We use the terminology of [Bertoin (1999), Gnedin, Pitman (2005)] who omit the term "killed".

classical Lévy subordinator Λ such that Λ_1 is distributed according to it. This correspondence allows to transfer results about infinitely divisible distributions to Lévy processes. For instance, Theorem 2.4.3 below characterizes a Lévy subordinator by means of a constant $\mu \geq 0$ and a measure ν on $(0, \infty]$.

Theorem 2.4.3 (Lévy (1934), Khinchin (1937))
The Laplace transforms of a Lévy subordinator Λ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ admit the functional form

$$\mathbb{E}[e^{-x\,\Lambda_t}] = e^{-t\,\Psi(x)}, \quad x \geq 0, \quad t \geq 0, \tag{2.14}$$

where the function $\Psi : [0, \infty) \to [0, \infty)$ is called the *Laplace exponent* of Λ. Moreover, there is a unique non-negative drift $\mu \geq 0$ and a unique positive measure ν on $(0, \infty]$, called the *Lévy measure* of Λ, such that

$$\Psi(x) = \mu\, x + \int_{(0,\infty]} (1 - e^{-t\,x})\, \nu(dt), \quad x \geq 0. \tag{2.15}$$

The Lévy measure ν satisfies the conditions

$$\int_{(0,1]} t\, \nu(dt) < \infty, \quad \nu\big((\epsilon, \infty]\big) < \infty, \quad \text{for all } \epsilon > 0. \tag{2.16}$$

Conversely, given a drift $\mu \geq 0$ and a measure ν on $(0, \infty]$ satisfying (2.16), there exists a Lévy subordinator with drift μ and Lévy measure ν. Thus, the distributional properties of a Lévy subordinator are completely characterized by its so-called *characteristics* (μ, ν).

Remark 2.4.4 (Mass at Infinity)
The right hand side of equation (2.15) is a short hand notation for

$$\Psi(x) = \mu\, x + \int_{(0,\infty)} (1 - e^{-x\,t})\, \nu(dt) + \mathbf{1}_{\{x>0\}}\, \nu(\{\infty\}), \quad x \geq 0. \tag{2.17}$$

It is justified by using the conventions $0 \cdot \infty = 0$ and $\exp(-\infty) = 0$, which imply that $\Psi(0) = 0$ even though one might have $\nu(\{\infty\}) > 0$. Positive mass of ν at infinity introduces a discontinuity of Ψ at zero. However, Ψ is smooth on $(0, \infty)$. It is a so-called *Bernstein function*, i.e. Ψ is infinitely often differentiable on $(0, \infty)$ and the first derivative $\Psi^{(1)}$ is completely monotone, see e.g. [Applebaum (2004), Theorem 1.3.23(2), p.

52]. The number $\nu(\{\infty\}) \in [0, \infty)$ is precisely the killing rate. Hence, in case of a classical Lévy subordinator it holds that $\nu(\{\infty\}) = 0$ and thus the last term $\mathbf{1}_{\{x>0\}}\,\nu(\{\infty\})$ in (2.17) vanishes.

Proof (of Theorem 2.4.3)
Originally due to [Lévy (1934), Khinchin (1937), Khinchin (1938)]. Since this representation is essential for the present dissertation, at least a sketch of the proof is provided in the Appendix. □

Example 2.4.5 (Characteristics of a Poisson Process)
On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, consider a homogeneous Poisson process $N = \{N_t\}_{t\geq 0}$ with intensity $\beta > 0$, as defined in (2.13). Then for each $t > 0$ the random variable N_t is $Poi(\beta\, t)$-distributed, i.e. the Laplace transform of N_t is computed as

$$\mathbb{E}\big[e^{-x\,N_t}\big] = \sum_{k=0}^{\infty} e^{-x\,k}\,\frac{(\beta\, t)^k}{k!}\,e^{-\beta\, t} = e^{-t\,\beta\,(1-e^{-x})}, \quad x \geq 0.$$

Thus, the Laplace exponent of N is $\Psi(x) = \beta\,\big(1 - \exp(-x)\big)$, $x \geq 0$. Obviously, N has zero drift $\mu = 0$, and the Lévy measure ν is a one-point mass. More precisely, $\nu(B) = \beta\,\mathbf{1}_{\{1\in B\}}$, for $B \in \mathcal{B}\big((0,\infty]\big)$.

Intuitively, for any Borel set $B \in \mathcal{B}\big((0,\infty]\big)$ the value $\nu(B)$ is equal to the expected number of jumps of Λ within one unit of time whose size is in B, i.e.

$$\nu(B) = \mathbb{E}\Big[\Big|\big\{t \in (0,1] \,:\, \Lambda_t - \lim_{u\uparrow t}\Lambda_u \in B\big\}\Big|\Big], \tag{2.18}$$

see e.g. [Cont, Tankov (2004), Definition 3.4, p. 76]. Hence, ν bears the information about size and frequency of the jumps of Λ. Finally, new Lévy subordinators can be constructed from known ones by the idea of subordination: given two independent Lévy subordinators $\Lambda^{(1)}$ and $\Lambda^{(2)}$, it holds that

$$\Lambda = \{\Lambda_t\}_{t\geq 0}, \quad \Lambda_t := \Lambda^{(2)}_{\Lambda^{(1)}_t}, \quad t \geq 0,$$

is again a Lévy subordinator, see e.g. [Bertoin (1999), Proposition 8.6]. One can even compute its characteristics, a result which is due to [Huff (1969)]. The original idea of subordination of a process is due to [Bochner (1955)]. Moreover, it is an easy exercise to check that the sum of two independent Lévy subordinators is again a Lévy subordinator.

Thus, it is theoretically possible to construct a huge repertoire of parametric families of Lévy subordinators from known ones.

2.4.1 Examples

Four of the most popular Lévy subordinators are introduced in the sequel. For a list of more infinitely divisible distributions on the positive half-axis see e.g. [Sato (1999), Remark 8.12, p. 46] and the references therein.

2.4.1.1 Compound Poisson Subordinators With Drift

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which $\{J_i\}_{i\in\mathbb{N}}$ are i.i.d. positive random variables and $N = \{N_t\}_{t\geq 0}$ is an independent Poisson process with intensity $\beta > 0$. With a non-negative drift $\mu \geq 0$ defining

$$\Lambda_t := \mu\, t + \sum_{i=1}^{N_t} J_i, \quad t \geq 0,$$

it follows from [Cont, Tankov (2004), Proposition 3.3, p. 71] that $\Lambda = \{\Lambda_t\}_{t\geq 0}$ is a Lévy subordinator, called a *compound Poisson subordinator (with drift)*. Furthermore it follows from [Cont, Tankov (2004), Proposition 3.5, p. 75] that the Lévy measure ν of Λ has the special form

$$\nu(B) = \beta\, \mathbb{P}(J_1 \in B), \quad B \in \mathcal{B}\big((0,\infty)\big).$$

This implies that the Laplace exponent Ψ of Λ is given by

$$\Psi(x) = \mu\, x + \beta\, \mathbb{E}\Big[\big(1 - e^{-x\, J_1}\big)\Big], \quad x \geq 0. \tag{2.19}$$

Intuitively, Λ grows linearly with constant drift μ, it jumps whenever the Poisson process N jumps, and the i-th jump has random jump size J_i. Thus, in a bounded time interval $[s,t]$, for $0 \leq s < t$, compound Poisson subordinators (with drift) almost surely exhibit only finitely many jumps. More precisely, the number of jumps of Λ in $[s,t]$ is $Poi\big(\beta\,(t-s)\big)$-distributed. A typical path of such a stochastic process is illustrated in Figure 2.5. It is worth mentioning that a Lévy subordinator is of compound Poisson type (possibly with drift) if and only if it has almost surely finitely many jumps on any bounded time

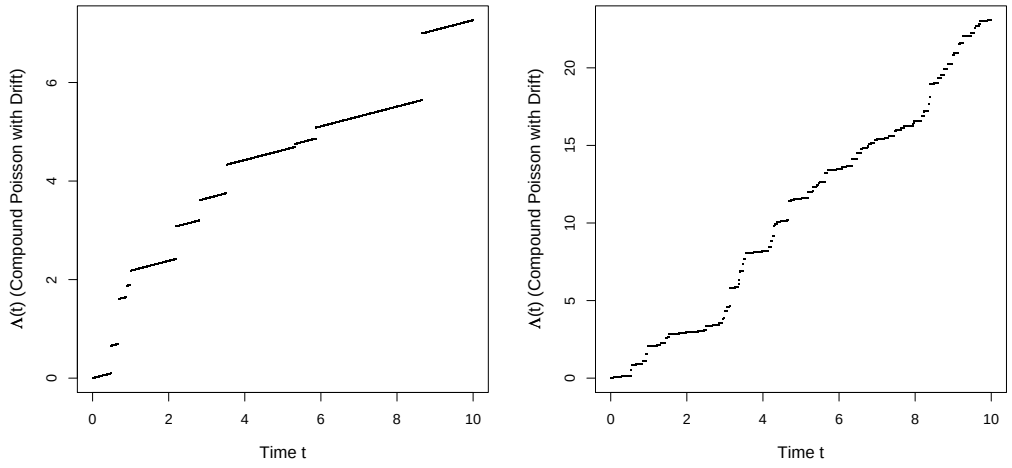


Figure 2.5 Simulated paths of a compound Poisson process with drift and $Exp(\eta)$-distributed jump sizes. In the left graph the jump intensity is $\beta = 1$, the drift $\mu = 0.2$, and $\eta = 2$. Parameters in the right graph are $(\mu, \eta, \beta) = (0.2, 4, 8)$. Both paths are simulated up to time $T = 10$.

interval. Compound Poisson subordinators are thus said to be *finite activity processes*. Analytically, a Lévy subordinator with Lévy measure ν is of compound Poisson type if and only if there is an $\epsilon > 0$ such that $\nu\big((0,\epsilon)\big) < \infty$. This follows more or less from (2.16) and (2.18).

2.4.1.2 Gamma Subordinator

Following [Schoutens (2003), p. 52], a Lévy subordinator Λ is called *Gamma subordinator*, if it has zero drift and its Lévy measure ν, parameterized by $(\beta, \eta) \in (0,\infty)^2$, is absolutely continuous with respect to the Lebesgue measure on $(0,\infty)$ and has the special form

$$\nu(dt) = \beta\, e^{-\eta\, t}\, \frac{1}{t}\, \mathbf{1}_{\{t>0\}}\, dt.$$

It is easy to check that the measure ν defined in this way satisfies (2.16) and hence defines a Lévy subordinator. It is possible to compute the Laplace exponent Ψ in closed form. Following [Tricomi (1951)], for continuous functions $f : (0,\infty) \to \mathbb{R}$ with existing

limits $\lim_{t\downarrow 0} f(t) \in \mathbb{R}$ and $\lim_{t\to\infty} f(t) \in \mathbb{R}$ the so-called *Frullani Theorem* states that

$$\int_{(0,\infty)} \big(f(a\,t) - f(b\,t)\big)\,\frac{1}{t}\,dt = \big(\lim_{t\downarrow 0} f(t) - \lim_{t\to\infty} f(t)\big)\,\log\Big(\frac{b}{a}\Big).$$

In particular, the function $f(t) := \exp(-t)$ is admissible in this formula and with $a := \eta$ and $b := \eta + x$ one may deduce

$$\Psi(x) = \int_{(0,\infty)} \big(1 - e^{-x\,t}\big)\,\beta\,e^{-\eta\,t}\,\frac{1}{t}\,dt = \beta\,\log\Big(1 + \frac{x}{\eta}\Big).$$

Therefore, the Lévy-Khinchin representation implies that the random variable Λ_t has Laplace transform

$$\mathbb{E}[e^{-x\,\Lambda_t}] = e^{-t\,\beta\,\log\left(1+\frac{x}{\eta}\right)} = \Big(1 + \frac{x}{\eta}\Big)^{-\beta\,t}, \quad x > 0,\, t > 0.$$

This is known to be the Laplace transform of a Gamma distribution, which explains the nomenclature of the process Λ. More precisely, for $t > 0$ the density f_{Λ_t} of Λ_t is given by

$$f_{\Lambda_t}(x) = \frac{\eta^{\beta\,t}}{\Gamma(\beta\,t)}\,x^{\beta\,t-1}\,e^{-\eta\,x}\,\mathbf{1}_{\{x>0\}}, \quad \Gamma(x) = \int_0^\infty e^{-s}\,s^{x-1}\,ds,$$

and we denote this Gamma distribution by $\Gamma(\beta\,t, \eta)$ in the sequel. Note that for each $\epsilon > 0$ it holds that $\nu\big((0,\epsilon)\big) = \infty$, meaning that a Gamma subordinator jumps almost surely infinitely often in a finite time interval. Lévy subordinators with this property are thus said to exhibit *infinite activity*. A typical path of a Gamma subordinator is illustrated in Figure 2.6. Due to the infinite activity, simulating a path of $\{\Lambda_t\}_{t\in[0,T]}$ is impossible without discretization bias. The simulation is accomplished by defining a grid $0 < T/n < 2\,T/n < \ldots < (n-1)\,T/n < T$ and accumulating i.i.d. random variables which are $\Gamma(\beta\,T/n, \eta)$-distributed.

2.4.1.3 Inverse Gaussian Subordinator

Following [Schoutens (2003), p. 53], a Lévy subordinator Λ is called *Inverse Gaussian subordinator*, if it has zero drift and its Lévy measure ν, parameterized by $(\beta, \eta) \in (0,\infty)^2$, is absolutely continuous with respect to the Lebesgue measure on $(0,\infty)$ and

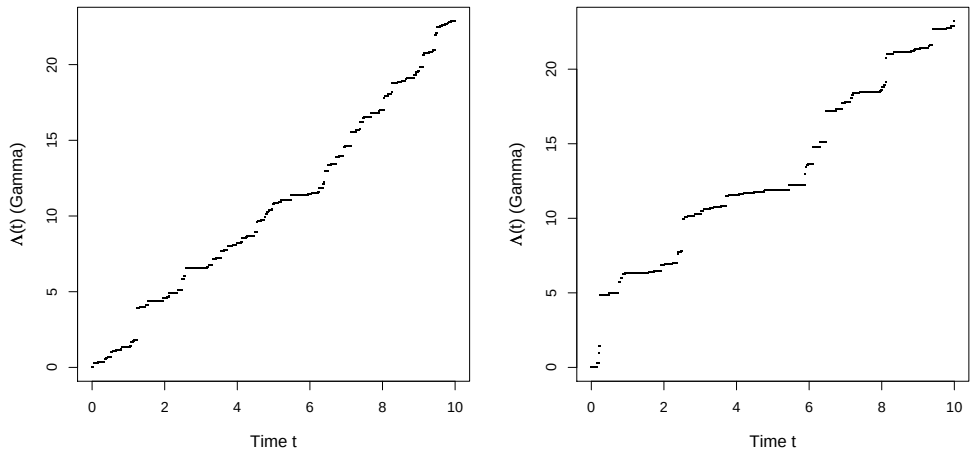


Figure 2.6 Simulated paths of a Gamma subordinator. Parameters used are $(\eta, \beta) = (2, 4.725)$ (left) and $(\eta, \beta) = (1, 2)$ (right). Both paths are simulated up to time $T = 10$ using $n = 1000$ grid points.

has the form

$$\nu(dt) = \frac{1}{\sqrt{2\,\pi}}\,\frac{\beta}{t^{\frac{3}{2}}}\,e^{-\frac{1}{2}\,\eta^2\,t}\,\mathbf{1}_{\{t>0\}}\,dt.$$

For each $t > 0$ the density f_{Λ_t} of the random variable Λ_t is known to be

$$f_{\Lambda_t}(x) = \frac{\beta\,t}{\sqrt{2\,\pi}}\,e^{\eta\,\beta\,t}\,\frac{1}{x^{\frac{3}{2}}}\,e^{-\frac{1}{2}\left(\beta^2\,t^2\,\frac{1}{x}+\eta^2\,x\right)}\,\mathbf{1}_{\{x>0\}},$$

see [Applebaum (2004), p. 51]. The distribution given by this density is called Inverse Gaussian distribution with parameters $(\beta\,t, \eta)$, denoted by $IG(\beta\,t, \eta)$. The Laplace exponent of Λ is given by[6]

$$\Psi(x) = \beta\left(\sqrt{2\,x + \eta^2} - \eta\right), \quad x \geq 0. \tag{2.20}$$

The name "Inverse Gaussian" stems from the fact that Λ may be constructed as

$$\Lambda_t := \inf\left\{s > 0 \,:\, \eta\,s + X_s = \beta\,t\right\},$$

[6]See [Applebaum (2004), p. 51].

where $X = \{X_t\}_{t\geq 0}$ is a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, see [Applebaum (2004), p. 51]. Thus, Λ_t can be interpreted as the first hitting-time of the level $\beta\, t$ of a Brownian motion with drift. The resulting relation to the normal distribution justifies the naming. Moreover, [Shuster (1968)] shows how to express the distribution function of Λ_t in terms of the standard normal distribution function Φ: for all $t > 0$, $x \geq 0$ it holds that

$$\mathbb{P}(\Lambda_t \leq x) = \Phi\Big(\eta\,\sqrt{x} - \frac{\beta\, t}{\sqrt{x}}\Big) + e^{2\,\beta\, t\,\eta}\,\Phi\Big(-\eta\,\sqrt{x} - \frac{\beta\, t}{\sqrt{x}}\Big), \quad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{s^2}{2}}\, ds.$$

Like the Gamma subordinator, an Inverse Gaussian subordinator exhibits infinite activity, since $\nu\big((0,\epsilon)\big) = \infty$ for each $\epsilon > 0$. Figure 2.7 illustrates typical paths of such a process. The sampling is done similar as in the case of a Gamma subordinator.

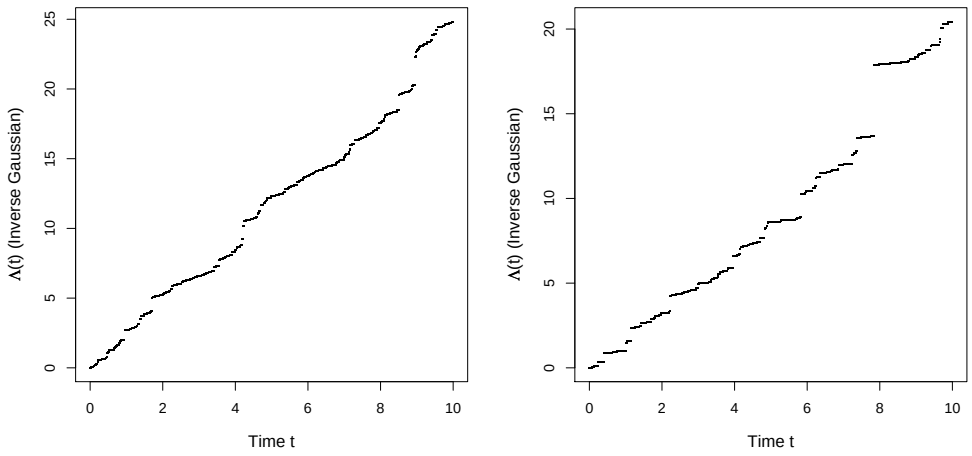


Figure 2.7 Simulated paths of an Inverse Gaussian subordinator. The parameters are $(\eta, \beta) = (2, 4.725)$ (left) and $(\eta, \beta) = (1, 2)$ (right). Both paths are simulated up to time $T = 10$ using $n = 1000$ grid points.

2.4.1.4 Stable Subordinator

A Lévy subordinator Λ is called an *α-stable subordinator* with parameter $\alpha \in (0, 1)$, if it has zero drift $\mu = 0$ and its Lévy measure ν is absolutely continuous with respect to

the Lebesgue measure and is defined via

$$\nu(dt) = \frac{\alpha}{\Gamma(1-\alpha)} \frac{1}{t^{1+\alpha}} \mathbf{1}_{\{t>0\}}\, dt.$$

One immediately checks that $\nu\big((0,\epsilon)\big) = \infty$ for $\epsilon > 0$. Hence, Λ is another infinite activity process. It can be verified by an application of Fubini's theorem, see e.g. [Applebaum (2004), p. 69], that the Laplace exponent of Λ is given by $\Psi(x) = x^\alpha$, $x \geq 0$. The distribution of Λ_t, $t > 0$, is called an α-stable distribution. We denote $\Lambda_t \sim S\big(\alpha, 1, \big(\cos(\pi\,\alpha/2)\,t\big)^{\frac{1}{\alpha}}, 0; 1\big)$, using the notation[7] of [Hofert (2007)]. The density f_{Λ_t} of Λ_t is not known in closed form, but [Nolan (1997)] uses Fourier inversion techniques to compute a numerically convenient form, which is given by

$$f_{\Lambda_t}(x) = t^{-\frac{1}{\alpha}}\, f_{\Lambda_1}\big(t^{-\frac{1}{\alpha}}\, x\big),$$

where

$$\begin{aligned} f_{\Lambda_1}(x) &= \mathbf{1}_{\{x>0\}} \frac{\alpha\,\big(\frac{x}{\gamma}\big)^{\frac{1}{\alpha-1}}}{\gamma\,\pi\,(1-\alpha)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} g_\alpha(u)\, e^{-\big(\frac{x}{\gamma}\big)^{\frac{\alpha}{\alpha-1}}\, g_\alpha(u)}\, du, \\ g_\alpha(u) &= \Big(\cos\Big(\frac{\pi\,\alpha}{2}\Big)\Big)^{\frac{1}{\alpha-1}} \Big(\frac{\cos u}{\sin\big(\alpha\big(\frac{\pi}{2}+u\big)\big)}\Big)^{\frac{\alpha}{\alpha-1}} \frac{\cos\big(\frac{\pi}{2}\,\alpha + (\alpha-1)\,u\big)}{\cos u}, \\ \gamma &= \Big(\cos\Big(\frac{\pi\,\alpha}{2}\Big)\Big)^{\frac{1}{\alpha}}. \end{aligned} \tag{2.21}$$

An α-stable subordinator has a heavy-tailed distribution. For instance, it is shown in [Wolfe (1975)] that

$$\mathbb{E}[\Lambda_t^\beta] = \begin{cases} \frac{t^{\frac{\beta}{\alpha}}\,\Gamma\left(1-\frac{\beta}{\alpha}\right)}{\Gamma(1-\beta)} & , \beta \in (0,\alpha) \\ \infty & , \beta \geq \alpha \end{cases}, \quad t > 0.$$

In particular, such distributions are standard examples for random variables without first moment. This property together with the convenient functional form of the Laplace exponent make this process interesting in many applications. Notice in particular that $\Psi(1) = 1$, independently of α, an important property for later results. Typical paths of an α-stable subordinator are illustrated in Figure 2.8. The sampling is again accomplished by accumulating n i.i.d. $S\big(\alpha, 1, \big(\cos(\pi\,\alpha/2)\,T/n\big)^{\frac{1}{\alpha}}, 0; 1\big)$-distributed random variables.

[7] Although appearing a little awkward, this notation is standard in the literature on stable distributions.

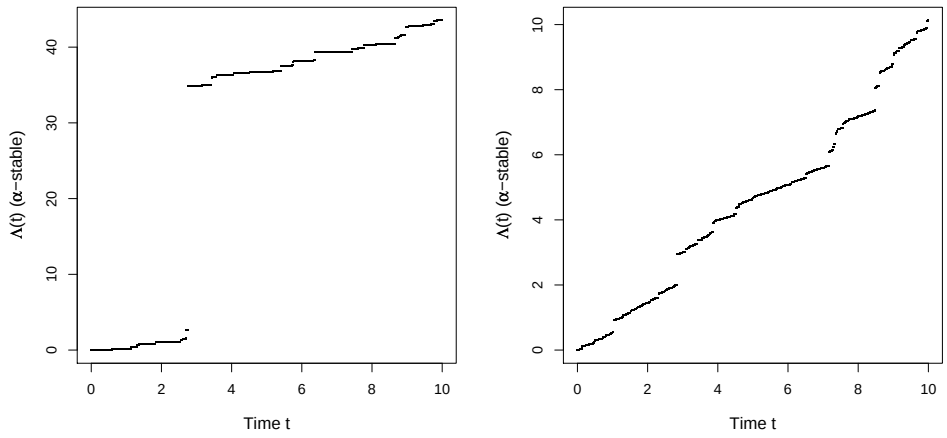


Figure 2.8 Simulated paths of an α-stable subordinator with $\alpha = 0.5$ (left) and $\alpha = 0.8$ (right). Both paths are simulated up to time $T = 10$ using $n = 1000$ grid points.

2.5 Moment Problems and Completely Monotone Sequences

Given a sequence of real numbers $\{a_k\}_{k\in\mathbb{N}_0}$, moment problems deal with the question whether there exists a random variable τ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $a_k = \mathbb{E}[\tau^k]$ for all $k \geq 0$. If the search is restricted to random variables τ on the positive half-axis $[0, \infty)$, then one speaks of the *Stieltjes moment problem*, whereas the *Hamburger moment problem* considers arbitrary $\tau \in \mathbb{R}$. In both cases it is possible that two random variables τ_1 and τ_2, which are not equal in distribution, have the same sequence of moments $a_k = \mathbb{E}[\tau_1^k] = \mathbb{E}[\tau_2^k]$, $k \geq 0$; for an example see [Feller (1966), p. 227]. Thus, the moment problem may have more than one solution and conditions on the sequence $\{a_k\}_{k\in\mathbb{N}_0}$ need to be derived that guarantee uniqueness. In contrast, if one restricts one's search to random variables τ on the compact set $[0, 1]$, then one can show that there exists at most one solution. This special case is called the *Hausdorff moment problem*; or sometimes the *little moment problem*. Sequences of moments of random variables on $[0, 1]$ can be characterized by a specific property, called *complete monotonicity*. These sequences are of fundamental interest during this dissertation. Moreover, the similar concept of *completely alternating sequences* is related to Lévy

subordinators and also to moments of convex distribution functions. These notions and coherences are carried out in this section.

2.5.1 Hausdorff's Moment Problem

The Hausdorff moment problem characterizes distributions on the unit interval $[0,1]$ by their moments. To state this theorem the following definition is needed. Standard references on completely monotone sequences are e.g. [Widder (1946), Feller (1966), Karlin (1968), Lorch, Newman (1983)]. The notation of [Gnedin, Pitman (2008)] is adopted in the following.

Definition 2.5.1 (Difference Operator)
For a sequence $\{a_k\}_{k\in\mathbb{N}_0}$ of real numbers the *difference operator* ∇ is defined by $\nabla a_k := a_k - a_{k+1}$ for all $k \in \mathbb{N}_0$.

The difference operator ∇ is a linear operator in the sense that for sequences $\{a_k\}_{k\in\mathbb{N}_0}$, $\{b_k\}_{k\in\mathbb{N}_0}$ and $\alpha \in \mathbb{R}$ one has that

$$\begin{aligned}\nabla(\alpha\, a_k + b_k) &= (\alpha\, a_k + b_k) - (\alpha\, a_{k+1} + b_{k+1}) \\ &= \alpha\,(a_k - a_{k+1}) + (b_k - b_{k+1}) = \alpha\,\nabla a_k + \nabla b_k.\end{aligned}$$

In the sequel, ∇ is often applied iteratively. For example we write $\nabla^2 a_k$ for the expression $\nabla(\nabla a_k) = \nabla(a_k - a_{k+1}) = \nabla a_k - \nabla a_{k+1} = a_k - 2\,a_{k+1} + a_{k+2}$. More generally, we write $\nabla^j a_k$ when ∇ is applied j times to a_k. Notice in particular that the expression $\nabla^j a_k$ involves the $j+1$ numbers $a_k, \ldots, a_{k+j}$. Moreover, it is convenient to introduce the notation $\nabla^0 a_k := a_k$.

The following lemma is included for the sake of completeness; it is easily seen to be true by induction. Since it is extensively used later on without further reference, it is recommended that the reader remembers the result of Lemma 2.5.2 like a definition of the expression $\nabla^j a_k$.

Lemma 2.5.2 (Iterating the Difference Operator)
By iterating the difference operator it follows that

$$\nabla^j a_k = \sum_{i=0}^{j} (-1)^i \binom{j}{i} a_{k+i}, \quad k \in \mathbb{N}_0,\, j \in \mathbb{N}_0. \tag{2.22}$$

Proof

Statement (2.22) is proved by induction over j. For $j = 0$ the statement is clear. Now the induction hypothesis (IH) is that (2.22) is true for $j \in \mathbb{N}_0$. Then,

$$\begin{aligned}
&\sum_{i=0}^{j+1}(-1)^i\binom{j+1}{i}a_{k+i}\\
&\quad= a_k + \sum_{i=1}^{j}(-1)^i\binom{j+1}{i}a_{k+i} + (-1)^{j+1}a_{k+j+1}\\
&\quad\overset{(*)}{=} a_k + \sum_{i=1}^{j}(-1)^i\binom{j}{i}a_{k+i} + \sum_{i=1}^{j}(-1)^i\binom{j}{i-1}a_{k+i} + (-1)^{j+1}a_{k+j+1}\\
&\quad= \sum_{i=0}^{j}(-1)^i\binom{j}{i}a_{k+i} + \sum_{i=0}^{j-1}(-1)^{i+1}\binom{j}{i}a_{k+i+1} + (-1)^{j+1}\binom{j}{j}a_{k+j+1}\\
&\quad= \sum_{i=0}^{j}(-1)^i\binom{j}{i}a_{k+i} - \sum_{i=0}^{j}(-1)^i\binom{j}{i}a_{k+i+1}\\
&\quad= \sum_{i=0}^{j}(-1)^i\binom{j}{i}\nabla a_{k+i} = \nabla\sum_{i=0}^{j}(-1)^i\binom{j}{i}a_{k+i} \overset{(IH)}{=} \nabla\nabla^j a_k = \nabla^{j+1}a_k,
\end{aligned}$$

where we have used in $(*)$ that

$$\binom{j+1}{i} = \binom{j}{i} + \binom{j}{i-1}, \quad 1 \le i \le j. \qquad \square$$

Analogously to the concept of a completely monotone function, compare Definition 2.3.6, the notion of a completely monotone sequence is introduced below.

Definition 2.5.3 (Completely Monotone Sequence)

The sequence $\{a_k\}_{k\in\mathbb{N}_0}$ is called *completely monotone* if $\nabla^j a_k \ge 0$ for all $k \in \mathbb{N}_0$, and for all $j \in \mathbb{N}_0$.

Now suppose $\tau : \Omega \to [0,1]$ is a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denoting its k-th moment by $a_k := \mathbb{E}[\tau^k]$ for all $k \in \mathbb{N}_0$, one easily verifies that

$$\nabla^j a_k = \sum_{i=0}^{j}\binom{j}{i}(-1)^i\,\mathbb{E}[\tau^{k+i}] = \mathbb{E}\Big[\tau^k\sum_{i=0}^{j}\binom{j}{i}(-1)^i\,\tau^i\Big] = \mathbb{E}\Big[\tau^k\,(1-\tau)^j\Big] \ge 0.$$

Thus, the sequence of moments of τ is completely monotone. More challenging to derive is the fact that the converse implication is also true. This is the content of the following

theorem.

Theorem 2.5.4 (Hausdorff (1921))
The sequence $\{a_k\}_{k\in\mathbb{N}_0}$ is completely monotone and $a_0 = 1$ if and only if there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random variable $\tau : \Omega \to [0,1]$ such that $a_k = \mathbb{E}[\tau^k]$ for all $k \in \mathbb{N}_0$. Moreover, the random variable τ is uniquely determined in distribution by its moments.

Proof
Originally in [Hausdorff (1921), Hausdorff (1923)]. See also [Feller (1966), p. 225] for a proof. Note that the uniqueness of τ heavily relies on the boundedness of the interval $[0,1]$ and follows from a classical result of Müntz and Szász, see [Müntz (1914), Szász (1916)]. □

The idea of the proof of Theorem 2.5.4 is to define a discrete distribution on the grid $\{i/d \,|\, i = 0, \ldots, d\}$ for each d from the given completely monotone sequence, and then let d tend to infinity. One can prove that the limit distribution on $[0,1]$ has the given sequence as moments. The construction of the discrete distributions uses a summation identity which is required several times later on. We extract it from the proof as a lemma for later reference.

Lemma 2.5.5 (Summation Identity)
Let $(a_0, \ldots, a_{d-1})'$ be an arbitrary finite sequence. Then it holds that

$$a_0 = \sum_{i=0}^{d-1} \binom{d-1}{i} \nabla^{d-1-i} a_i.$$

Proof
Replacing $k = i + j$ in the second equality and interchanging sums in the third one computes

$$\begin{aligned}
\sum_{i=0}^{d-1} \binom{d-1}{i} \nabla^{d-1-i} a_i &= \sum_{i=0}^{d-1} \binom{d-1}{i} \sum_{j=0}^{d-1-i} \binom{d-1-i}{j} (-1)^j \, a_{i+j} \\
&\overset{(k=i+j)}{=} \sum_{i=0}^{d-1} \binom{d-1}{i} \sum_{k=i}^{d-1} \binom{d-1-i}{k-i} (-1)^{k-i} \, a_k \\
&= \sum_{k=0}^{d-1} (-1)^k \, a_k \sum_{i=0}^{k} (-1)^i \binom{d-1}{i} \binom{d-1-i}{k-i}
\end{aligned}$$

$$\begin{aligned}&\overset{(*)}{=} \sum_{k=0}^{d-1}(-1)^k \binom{d-1}{k} a_k \sum_{i=0}^{k}(-1)^i \binom{k}{i} \\ &= \sum_{k=0}^{d-1}(-1)^k \binom{d-1}{k} a_k \,(1-1)^k = a_0.\end{aligned}$$

For equality $(*)$ one uses that

$$\binom{d-1}{i}\binom{d-1-i}{k-i} = \binom{d-1}{k}\binom{k}{i}. \qquad \square$$

Comparing Theorems 2.5.4 and 2.3.7, one observes a connection between completely monotone functions and completely monotone sequences: for a given completely monotone function φ with $\varphi(0) = 1$, the sequence $\{\varphi(k)\}_{k\in\mathbb{N}_0}$ is completely monotone as well: as φ is completely monotone, $\varphi(x) = \mathbb{E}[\exp(-x\,W)]$ for some random variable $W \geq 0$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by Theorem 2.3.7. Hence $\varphi(k)$ is the k-th moment of the random variable $\exp(-W) \in [0,1]$. Further similarities between both concepts are studied in [Lorch, Newman (1983)] and in Section 3.6.

2.5.2 Moment Problem for Convex Distributions

A key reference for the results in this dissertation is the paper [Gnedin, Pitman (2008)]. In this reference a characterization of convex distribution functions on $[0,1]$ is derived and a connection to the Lévy-Khinchin representation, compare Theorem 2.4.3, is observed. This connection turns out to be useful to derive an alternative probabilistic model for a subfamily of Marshall-Olkin distributions. The following definition is required.

Definition 2.5.6 (Completely Alternating Sequence)
A sequence $\{c_k\}_{k\in\mathbb{N}_0}$ is called *completely alternating* if $\nabla^j c_k \leq 0$ for all $k \in \mathbb{N}_0$, and for all $j \in \mathbb{N}$.

Note that $\nabla^0 c_k = c_k$ need not be ≤ 0. [Gnedin, Pitman (2008)] study the sequence of moments of random variables on $[0,1]$ whose distribution functions are convex[8]. More

[8] We mean convex on $[0,1]$. Of course, a distribution function cannot be convex on $\mathbb{R}$.

clearly, the authors consider distribution functions F which satisfy $F(0) = 0$, $F(1) = 1$, and are convex on $[0, 1]$. These can be written as

$$F(x) = \int_{(0,x)} f(t)\, dt + \mathbf{1}_{\{x=1\}}, \quad x \in [0, 1], \tag{2.23}$$

for some non-negative, non-decreasing function f. In particular, these distribution functions are continuous on $[0, 1)$ and have a possible jump at one. The following lemma clarifies the relation between moments of convex distribution functions on $[0, 1]$ and completely alternating sequences.

Lemma 2.5.7 (Gnedin, Pitman (2008))
A sequence $\{a_k\}_{k\in\mathbb{N}_0}$ is the sequence of moments of a convex distribution function on $[0, 1]$ if and only if the sequence $\{c_k\}_{k\in\mathbb{N}_0}$ defined by $c_0 = 0$, $c_k = k\, a_{k-1}$, $k \geq 1$, is completely alternating with $c_1 = 1$.

Proof
See [Gnedin, Pitman (2008), Theorem 1.7]. The proof uses Theorem 2.5.4 and some technical computations. □

The important result for our purpose is the following link between Lévy subordinators and completely alternating sequences, which can be derived from Lemma 2.5.7.

Theorem 2.5.8 (Gnedin, Pitman (2008))
The sequence $\{c_k\}_{k\in\mathbb{N}_0}$ with $c_0 = 0$ is completely alternating if and only if there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a Lévy subordinator Λ with Laplace exponent Ψ such that $c_k = \Psi(k)$ for all $k \in \mathbb{N}_0$. Moreover, Λ is uniquely determined in distribution by this sequence.

Proof
See [Gnedin, Pitman (2008), Corollary 4.2]. Uniqueness follows from the classical result of Müntz and Szász, see [Müntz (1914), Szász (1916)]. Since this result is very important for the remaining work and it is interesting to recognize the connection between Lévy subordinators and convex distributions on $[0, 1]$, the proof is given in the sequel. We start by proving sufficiency. If $\Lambda_t \equiv 0$, the statement is clear, hence we may assume that Λ is non-trivial. Theorem 2.4.3 shows that there is a drift $\mu \geq 0$ and a measure ν

on $(0,\infty]$ such that the Laplace exponent of Λ has the form

$$\begin{aligned}\Psi(x) &= \mu\, x + \int_{(0,\infty]} (1 - e^{-t\,x})\, \nu(dt) = \mu\, x + x \int_{(0,\infty]} \int_{(0,t)} e^{-u\,x}\, du\, \nu(dt) \\ &= \mu\, x + x \int_{(0,\infty]} \int_{(0,\infty)} \mathbf{1}_{\{u<t\}}\, e^{-u\,x}\, du\, \nu(dt) = \mu\, x + x \int_{(0,\infty)} \int_{(0,\infty]} \mathbf{1}_{\{u<t\}}\, \nu(dt)\, e^{-u\,x}\, du \\ &= \mu\, x + x \int_{(0,\infty)} \nu\big((u,\infty]\big)\, e^{-u\,x}\, du = \mu\, x + x \int_{(0,1)} \nu\big((-\log y,\infty]\big)\, y^{x-1}\, dy. \qquad (2.24)\end{aligned}$$

The fourth equality above relies on Tonelli's theorem and the last equality uses the substitution $y = \exp(-u)$. For $x = 1$ this implies that

$$\int_{(0,1)} \nu\big((-\log y,\infty]\big)/\Psi(1)\, dy = 1 - \mu/\Psi(1) \leq 1. \qquad (2.25)$$

Division by $\Psi(1)$ is possible since Λ is non-trivial, implying $\Psi(1) > 0$. Defining the function $F : [0,1] \to [0,1]$ by $F(x) := \int_0^x \nu\big((-\log y,\infty]\big)/\Psi(1)\, dy + \mathbf{1}_{\{x=1\}}$, one observes that F is a convex distribution function on $[0,1]$. This can be seen from equation (2.25) and the fact that $y \mapsto \nu\big((-\log y,\infty]\big)/\Psi(1)$ is non-decreasing (hence F is convex). Moreover, equation (2.25) implies that the probability measure dF induced by F assigns the mass $\mu/\Psi(1)$ to the right endpoint of the unit interval. Thus, the above computation shows that for $k \in \mathbb{N}$

$$\Psi(k) = \Psi(1) \left(k\, \frac{\mu}{\Psi(1)} + k \int_{(0,1)} y^{k-1}\, dF(y) \right) = \Psi(1)\, k \int_{(0,1]} y^{k-1}\, dF(y).$$

Using Lemma 2.5.7, it is shown that the sequence $\{\Psi(k)/\Psi(1)\}_{k\in\mathbb{N}_0}$ is completely alternating. However, this implies that the sequence $\{\Psi(k)\}_{k\in\mathbb{N}_0}$ is also completely alternating. Necessity is proved similarly: assume we are given a completely alternating sequence $\{c_k\}_{k\in\mathbb{N}_0}$ which is not identically zero (in this case the statement is true setting $\Lambda \equiv 0$). We first assume that $c_1 = 1$. Lemma 2.5.7 implies that there is a convex distribution function F whose moments satisfy

$$c_k = k \int_{(0,1]} y^{k-1}\, dF(y), \quad k \in \mathbb{N}.$$

Recall that F is given as in equation (2.23), so there is a non-decreasing density f on $(0,1)$ and the probability mass at 1 given by $\mu := 1 - \lim_{x\uparrow 1} F(x)$. Define a measure ν on $(0,\infty]$ by $\nu\big(\{\infty\}\big) := \lim_{y\downarrow 0} f(y)$, $\nu\big((y,\infty]\big) := f\big(\exp(-y)\big)$. First observe that ν defined in this way actually gives a Lévy measure, i.e. it satisfies (2.16). To see this, an

application of Fubini's Theorem shows

$$\begin{aligned}\int_{(0,1)} t\,\nu(dt) &= \int_{(0,1)}\int_{(0,1)} \mathbf{1}_{\{u<t\}}\,du\,\nu(dt) = \int_{(0,1)}\int_{(0,1)} \mathbf{1}_{\{u<t\}}\,\nu(dt)\,du \\ &= \int_{(0,1)} \nu\big((u,1)\big)\,du \le \int_{(0,1)} \nu\big((u,\infty]\big)\,du = \int_{(0,1)} f\big(\exp(-u)\big)\,du \\ &= \int_{(e^{-1},1)} f(y)\,\frac{1}{y}\,dy \le \int_{(e^{-1},1)} f(y)\,e\,dy \le e < \infty. \\ \nu\big((\epsilon,\infty]\big) &= f\big(\exp(-\epsilon)\big) < \infty, \quad \epsilon > 0.\end{aligned}$$

Hence, there exists a Lévy subordinator Λ with Lévy measure ν and drift $\mu \ge 0$. The computation (2.24) shows that the Laplace exponent Ψ of Λ is given by

$$\begin{aligned}\Psi(x) &= \mu\,x + x\int_{(0,1)} \nu\big((-\log y,\infty]\big)\,y^{x-1}\,dy \\ &= \mu\,1^{x-1}\,x + x\int_{(0,1)} f(y)\,y^{x-1}\,dy = x\int_{(0,1]} y^{x-1}\,dF(y).\end{aligned}$$

Hence, for $k \in \mathbb{N}$ it holds that $\Psi(k) = c_k$. Finally, if $c_1 \neq 1$, then $c_1 \neq 0$ because otherwise $\{c_k\}_{k\in\mathbb{N}_0}$ was identically zero[9]. Hence, the sequence $\{\tilde{c}_k\}_{k\in\mathbb{N}_0} := \{c_k/c_1\}_{k\in\mathbb{N}_0}$ is well-defined and completely alternating with $\tilde{c}_0 = 0$ and $\tilde{c}_1 = 1$. The above proof implies the existence of a Lévy subordinator $\tilde{\Lambda}$ with Laplace exponent $\tilde{\Psi}$ satisfying $\tilde{c}_k = \tilde{\Psi}(k)$, $k \in \mathbb{N}_0$. Denote the characteristics of $\tilde{\Lambda}$ by $(\tilde{\mu},\tilde{\nu})$. But then the function $\Psi := c_1\,\tilde{\Psi}$ is also the Laplace exponent of a Lévy subordinator (with drift $c_1\,\tilde{\mu}$ and Lévy measure $c_1\,\tilde{\nu}$), hence the claim is established. □

Example 2.5.9 (Single-Shock Subordinator)
If $c_0 = 0$ and $c_k = \alpha > 0$ for all $k \ge 1$, then $\{c_k\}_{k\in\mathbb{N}_0}$ is completely alternating and the associated Lévy subordinator is determined by zero drift and a Lévy measure which concentrates all mass at infinity. From a probabilistic point of view this means that the Lévy subordinator is identically zero until it jumps to the absorbing state infinity. The expected number of such jumps per unit of time is α. From an analytical point of view, the Laplace exponent in this case is given by $\Psi(x) = \alpha\,\mathbf{1}_{\{x>0\}}$, $x \ge 0$.

[9] Since $\nabla c_1 \le 0$ and $c_1 = 0$ imply $c_2 \ge 0$, and $\nabla^2 c_0 \le 0$ and $c_0 = c_1 = 0$ imply $c_2 \le 0$, it follows that $c_2 = 0$. Then proceed iteratively.

3 An Analytical Study of the Marshall-Olkin Distribution

"Die Mathematik sieht vollständig ab von der actualen Bedeutung, die man ihren Begriffen geben, von der actualen Gültigkeit, die man ihren Sätzen zusprechen kann. Ihre indefinablen Begriffe sind willkürlich gewählte Denkobjecte, ihre Axiome willkürlich, jedoch widerspruchsfrei gewählte Beziehungen zwischen diesen Objecten. Die Mathematik ist Wissenschaft des reinen Denkens, gleich der formalen Logik."

F. Hausdorff, German mathematician.

The goal of this chapter is to understand the functional form of the Marshall-Olkin distribution in the exchangeable special case. To this end, the exchangeable subclass of the Marshall-Olkin distribution is introduced in Section 3.1. After a motivation in Section 3.2, d-monotone sequences are introduced in Section 3.3. Sections 3.4 and 3.5 propose to use these sequences to parameterize the exchangeable Marshall-Olkin distribution. Finally, Section 3.6 reveals interesting analytical similarities between exchangeable Marshall-Olkin survival copulas and Archimedean copulas.

3.1 Exchangeable Marshall-Olkin Survival Copulas

A random vector $(\tau_1, \ldots, \tau_d)'$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *exchangeable* if

$$(\tau_1, \ldots, \tau_d)' \stackrel{d}{=} (\tau_{\pi(1)}, \ldots, \tau_{\pi(d)})'$$

holds for all permutations π on $\{1, \ldots, d\}$. Alternatively, it is immediately clear that $(\tau_1, \ldots, \tau_d)'$ is exchangeable if and only if its survival function is invariant with respect

to permutations of its arguments, i.e. for each permutation π on $\{1,\ldots,d\}$ one has

$$\mathbb{P}(\tau_1 > t_1,\ldots,\tau_d > t_d) = \mathbb{P}(\tau_1 > t_{\pi(1)},\ldots,\tau_d > t_{\pi(d)}), \quad t_1,\ldots,t_d \in \mathbb{R}.$$

The following lemma clarifies which Marshall-Olkin distributions are exchangeable.

Lemma 3.1.1 (Exchangeable Marshall-Olkin distribution)
On a probability space $(\Omega,\mathcal{F},\mathbb{P})$ let $(\tau_1,\ldots,\tau_d)'$ be a random vector with Marshall-Olkin distribution, i.e. with survival function (2.10) for parameters $\lambda_I \geq 0$, $\emptyset \neq I \subset \{1,\ldots,d\}$, such that $\sum_{I:k\in I}\lambda_I > 0$, $k = 1,\ldots,d$. Then $(\tau_1,\ldots,\tau_d)'$ is exchangeable if and only if the parameters satisfy the following condition:

$$|I| = |\tilde{I}| \;\Rightarrow\; \lambda_I = \lambda_{\tilde{I}}. \tag{3.1}$$

Proof
First suppose that (3.1) is valid. Without loss of generality we may assume that $(\Omega,\mathcal{F},\mathbb{P})$ is the probability space from the original construction of [Marshall, Olkin (1967)], on which $(\tau_1,\ldots,\tau_d)'$ is constructed by (2.11). Rewriting this definition we observe that

$$\tau_k := \min_{i=1,\ldots,d}\Big\{\min\big\{E_I \,\big|\, I\subset\{1,\ldots,d\},\, k\in I,\, |I| = i\big\}\Big\}, \quad k = 1,\ldots,d.$$

For all $i,k = 1,\ldots,d$ there are precisely $d-1$ choose $i-1$ subsets I of $\{1,\ldots,d\}$ with i elements containing k. By assumption their associated parameters λ_I are identical, and in particular independent of k. It follows for $\{i,k\}\subset\{1,\ldots,d\}$ that the distribution of

$$\min\big\{E_I \,\big|\, I\subset\{1,\ldots,d\},\, k\in I,\, |I| = i\big\},$$

and therefore the distribution of τ_k, is independent of k. This implies that $(\tau_1,\ldots,\tau_d)'$ is exchangeable.

Conversely, assume that $(\tau_1,\ldots,\tau_d)'$ is exchangeable. This means that the survival function (2.10) is invariant with respect to its arguments. In order to simplify notations, we write $\bar{F}(\vec{t})$ instead of $\bar{F}(t_1,\ldots,t_d)$, where $\vec{t} := (t_1,\ldots,t_d)'$. Moreover, the i-th unit vector in $\mathbb{R}^d$ is denoted by $\vec{e}_i$. We prove (3.1) by induction over the cardinality of subsets of $\{1,\ldots,d\}$. To begin with, we verify $\lambda_{\{1\}} = \lambda_{\{2\}} = \ldots = \lambda_{\{d\}}$: for each $k = 2,\ldots,d$

exchangeability implies that

$$\sum_{\substack{\emptyset \neq I \subset \{1,\ldots,d\} \\ I \neq \{1\}}} \lambda_I = -\log \bar{F}\Big(\sum_{i=2}^{d} \vec{e}_i\Big) = -\log \bar{F}\Big(\sum_{\substack{i=1 \\ i \neq k}}^{d} \vec{e}_i\Big) = \sum_{\substack{\emptyset \neq I \subset \{1,\ldots,d\} \\ I \neq \{k\}}} \lambda_I.$$

Subtracting the sum of all parameters on both sides, this in turn verifies $\lambda_{\{1\}} = \lambda_{\{2\}} = \ldots = \lambda_{\{d\}}$. Now by induction hypothesis we assume that all parameters λ_I corresponding to subsets $I \subset \{1, \ldots, d\}$ of cardinality $|I| \leq k$ are identical. We prove then that all parameters λ_I corresponding to subsets $I \subset \{1, \ldots, d\}$ of cardinality $|I| = k + 1$ are identical. To this end, let I_0 be an arbitrary subset of $\{1, \ldots, d\}$ of cardinality $|I_0| = k+1$. Then

$$\sum_{\substack{\emptyset \neq I \subset \{1,\ldots,d\} \\ I \not\subseteq I_0}} \lambda_I = -\log \bar{F}\Big(\sum_{\substack{i=1 \\ i \notin I_0}}^{d} \vec{e}_i\Big) = -\log \bar{F}\Big(\sum_{i=k+2}^{d} \vec{e}_i\Big) = \sum_{\substack{\emptyset \neq I \subset \{1,\ldots,d\} \\ I \not\subseteq \{1,\ldots,k+1\}}} \lambda_I.$$

Subtracting the sum of all parameters on both sides, this implies

$$\lambda_{I_0} + \sum_{\substack{\emptyset \neq I \subset I_0 \\ |I| \leq k}} \lambda_I = \lambda_{\{1,\ldots,k+1\}} + \sum_{\substack{\emptyset \neq I \subset \{1,\ldots,k+1\} \\ |I| \leq k}} \lambda_I.$$

Using the induction hypothesis, this verifies that $\lambda_{\{1,\ldots,k+1\}} = \lambda_{I_0}$. Since I_0 was arbitrary with cardinality $k + 1$, we may conjecture that all parameters λ_I with $|I| = k + 1$ are identical. The claim is established. □

For the exchangeable subfamily of Marshall-Olkin distributions, Lemma 3.1.1 shows that the sets $\{\lambda_I \,|\, |I| = k\}$ are singletons for $k = 1, \ldots, d$. This means that, instead of $2^d - 1$ parameters λ_I, $\emptyset \neq I \subset \{1, \ldots, d\}$, an exchangeable Marshall-Olkin distribution is parameterized by only d parameters $\lambda_1, \ldots, \lambda_d \geq 0$, where $\lambda_k := \lambda_{\{1,\ldots,k\}}$, $k = 1, \ldots, d$. It is important to stress that the case $\lambda_1 = \ldots = \lambda_d = 0$ is excluded by the earlier assumptions on the λ_I's in order for construction (2.11) to make sense. Applying Lemma 2.3.4, the survival copula $\hat{C}$ of the Marshall-Olkin distribution can be massively simplified in the exchangeable case. This is the content of the following lemma.

Lemma 3.1.2 (Exchangeable Marshall-Olkin Survival Copula)
The survival copula of the random vector $(\tau_1, \ldots, \tau_d)'$ in the exchangeable case, i.e. with parameters $(\lambda_1, \ldots, \lambda_d)' \in [0, \infty)^d \setminus \{(0, \ldots, 0)'\}$, where $\{\lambda_I \,|\, |I| = i\} = \{\lambda_i\}$ for all

$i = 1, \ldots, d$, is given by

$$\hat{C}(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{\frac{\sum_{i=0}^{d-k} \binom{d-k}{i} \lambda_{i+1}}{\sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1}}}, \tag{3.2}$$

where $u_{(1)} \leq \ldots \leq u_{(d)}$ denotes the ordered list of $u_1, \ldots, u_d \in [0,1]$.

Proof

The statement is obtained by applying Lemma 2.3.4 in the exchangeable special case: it is observed in this case that $O_k := \sum_{I: k \in I} \lambda_I$ is independent of k, since

$$\sum_{I: k \in I} \lambda_I = \sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1} =: a_0.$$

This is due to the fact that for each index k there are precisely $d-1$ choose i subsets I of $\{1, \ldots, d\}$ with $i+1$ elements containing k, $i = 0, \ldots, d-1$. Hence, an application of Lemma 2.3.4 implies

$$\begin{aligned}\hat{C}(u_1, \ldots, u_d) &= \prod_{k=1}^{d} \prod_{1 \leq i_1 < \ldots < i_k \leq d} \Big(\min_{l=1,\ldots,k} \{u_{i_l}\} \Big)^{\frac{\lambda_k}{a_0}} \\ &= u_{(d)}^{\frac{\lambda_1}{a_0}} \, u_{(d-1)}^{\frac{\lambda_1}{a_0} + \frac{\lambda_2}{a_0}} \, u_{(d-2)}^{\frac{\lambda_1}{a_0} + 2\frac{\lambda_2}{a_0} \frac{\lambda_3}{a_0}} \cdots u_{(1)}^{\sum_{i=0}^{d-1} \binom{d-1}{i} \frac{\lambda_{i+1}}{a_0}} = \prod_{k=1}^{d} u_{(k)}^{\frac{1}{a_0} \sum_{i=0}^{d-k} \binom{d-k}{i} \lambda_{i+1}}.\end{aligned}$$

The second equation illustrates the required combinatorial observation: the k-th largest element $u_{(k)}$ of $u_1, \ldots, u_d$ is once the minimum of a set with one element (namely of $\{u_{(k)}\}$), $d-k$ times the minimum of a set with two elements (namely of $\{u_{(i)}, u_{(k)}\}$ for $i > k$), $d-k$ choose 2 times the minimum of a set with three elements, and so on. $\square$

The parametric family of copulas of the form (3.2) is denoted by eMO (standing for *exchangeable Marshall-Olkin*) in the sequel. This class is the main source of interest during this dissertation. For the sake of clarity, we explicitly define the class of exchangeable Marshall-Olkin survival copulas by

$$eMO := \left\{ \prod_{k=1}^{d} u_{(k)}^{\frac{\sum_{i=0}^{d-k} \binom{d-k}{i} \lambda_{i+1}}{\sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1}}} \;\middle|\; (0, \ldots, 0)' \neq (\lambda_1, \ldots, \lambda_d)' \in [0, \infty)^d \right\}.$$

Any $C \in eMO$ is invariant under permutations of its arguments. This implies that for

$2 \leq i \leq d$, all i-margins of C are of the same structural kind, which corresponds to the exchangeability of the distribution. In particular, 2-margins are bivariate Cuadras-Augé copulas with parameter $\lambda_2/(\lambda_1 + \lambda_2)$, compare Definition 2.2.2. Hence, the class eMO is a multivariate extension of bivariate Cuadras-Augé copulas. Figure 3.1 illustrates scatterplots from trivariate eMO-copulas. It is observed in the left plot that the measure dC induced by C assigns positive mass to the diagonal of the unit cube. More challenging to recognize in the plots is that dC also assigns positive mass to the planes $\{u_1 = u_2\}$, $\{u_1 = u_3\}$, and $\{u_2 = u_3\}$.

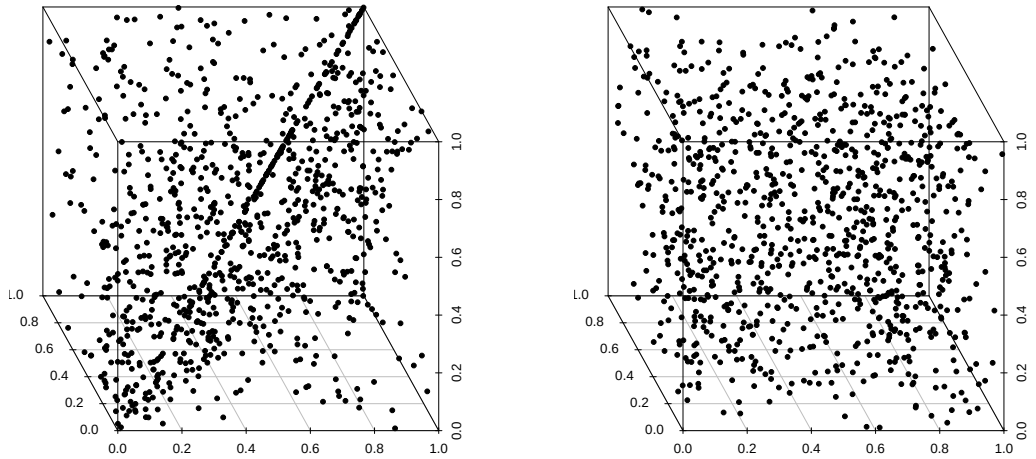


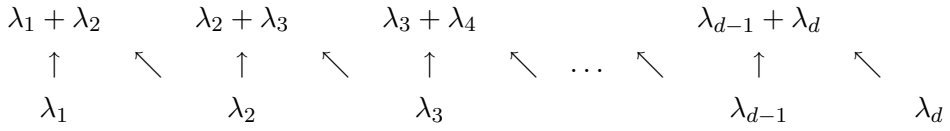
Figure 3.1 Scatterplots of 1000 samples from the copula $C(u_1, u_2, u_3) \in eMO$, corresponding to the parameters $(\lambda_1, \lambda_2, \lambda_3)' = (1.2, 1.5, 1)'$ (left) and $(\lambda_1, \lambda_2, \lambda_3)' = (5, 0.1, 0.1)'$ (right). The parameters of the right plot are chosen such that $\lambda_1 \gg \max\{\lambda_2, \lambda_3\}$ which intuitively means that $C \approx \Pi$.

3.2 Motivation

Some of the computations in this chapter are quite technical. Therefore, this short paragraph outlines where we are going and motivates the upcoming results. To this end, consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random vector $(\tau_1, \ldots, \tau_d)'$, which follows an exchangeable Marshall-Olkin distribution with parameters $(\lambda_1, \ldots, \lambda_d)' \in [0, \infty)^d \setminus \{(0, \ldots, 0)'\}$. Without loss of generality we let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space from the original construction (2.11) and recall that λ_k equals the intensity of exogenous shocks affecting k-dimensional subvectors of $(\tau_1, \ldots, \tau_d)'$. Now consider a $(d-1)$-dimensional subvector, without loss of generality $(\tau_1, \ldots, \tau_{d-1})'$. Rewriting the definition (2.11) of the Marshall-Olkin distribution implies for $k = 1, \ldots, d-1$ that

$$\begin{aligned}\tau_k &= \min\left\{E_I \,\middle|\, k \in I\right\} = \min\left\{E_I \,\middle|\, k \in I,\, d \notin I\right\} \cup \left\{E_I \,\middle|\, \{k, d\} \subset I\right\} \\ &= \min\left\{ \min\left\{E_I \,\middle|\, k \in I,\, d \notin I\right\}, \min\left\{E_I \,\middle|\, \{k, d\} \subset I\right\}\right\}.\end{aligned}$$

The first minimum corresponds to the definition of a $(d-1)$-dimensional exchangeable Marshall-Olkin distribution with parameters $(\lambda_1, \ldots, \lambda_{d-1})'$. However, the appearance of the second minimum suggests that exogenous shocks affecting the last component - which we now eliminated - have to be taken into account as well. More precisely, the components indexed by $I \subset \{1, \ldots, d-1\}$ are affected by E_I and $E_{I \cup \{d\}}$. Since $\min\{E_I, E_{I \cup \{d\}}\} \sim Exp(\lambda_{|I|} + \lambda_{|I|+1})$, this observation implies that all $(d-1)$-dimensional subvectors of $(\tau_1, \ldots, \tau_d)'$ follow a $(d-1)$-dimensional exchangeable Marshall-Olkin distribution corresponding to the parameters $(\lambda_1 + \lambda_2, \lambda_2 + \lambda_3, \ldots, \lambda_{d-1} + \lambda_d)' \in [0, \infty)^{d-1} \setminus \{(0, \ldots, 0)'\}$. Consequently, one can compute the parameters of the $(d-1)$-dimensional subvectors from $\lambda_1, \ldots, \lambda_d$ via the following geometric scheme:



One can now proceed iteratively and compute the following triangular scheme, the k-th row of which corresponds to the parameters of all k-dimensional subvectors of $(\tau_1, \ldots, \tau_d)'$:

$$
\begin{array}{lllllll}
\sum_{i=1}^{d}\binom{d-1}{i-1}\lambda_i & & & & & & \\
\sum_{i=1}^{d-1}\binom{d-2}{i-1}\lambda_i & \sum_{i=2}^{d}\binom{d-2}{i-2}\lambda_i & & & & & \\
\vdots & & \ddots & & & & \\
\lambda_1{+}2\,\lambda_2{+}\lambda_3 & \lambda_2{+}2\,\lambda_3{+}\lambda_4 & \lambda_3{+}2\,\lambda_4{+}\lambda_5 & \dots & \lambda_{d-2}{+}2\,\lambda_{d-1}{+}\lambda_d & & \\
\lambda_1+\lambda_2 & \lambda_2+\lambda_3 & \lambda_3+\lambda_4 & \dots & \lambda_{d-2}+\lambda_{d-1} & \lambda_{d-1}+\lambda_d & \\
\lambda_1 & \lambda_2 & \lambda_3 & \dots & \lambda_{d-2} & \lambda_{d-1} & \lambda_d
\end{array}
$$

In particular, the value at the tip of the triangle[1] corresponds to the exponential rate of the random variables $\tau_1, \dots, \tau_d$. This triangular scheme of non-negative numbers determines the distribution of $(\tau_1, \dots, \tau_d)'$, since already the bottom row does that. However, instead of computing the triangle from the bottom row we can alternatively start from the left column and compute the whole triangular scheme from it. Hence, the distribution of $(\tau_1, \dots, \tau_d)'$ is alternatively determined by the first column. More clearly, denoting the elements of the left column by

$$a_0 := \sum_{i=1}^{d}\binom{d-1}{i-1}\lambda_i,\ a_1 := \sum_{i=1}^{d-1}\binom{d-2}{i-1}\lambda_i, \dots,\ a_{d-2} := \lambda_1+\lambda_2,\ a_{d-1} := \lambda_1,$$

one can compute the same triangular scheme proceeding from the left as follows:

$$
\begin{array}{lllllll}
a_0 & & & & & & \\
 & \searrow & & & & & \\
a_1 & \to & a_0 - a_1 = \nabla a_0 & & & & \\
 & \searrow & & \searrow & & & \\
a_2 & \to & a_1 - a_2 = \nabla a_1 & \to & \nabla a_0 - \nabla a_1 = \nabla^2 a_0 & & \\
 & \searrow & & \searrow & & \searrow & \\
\vdots & & \vdots & & \vdots & & \ddots
\end{array}
$$

This chapter formalizes the idea outlined above and proposes to treat exchangeable Marshall-Olkin distributions via the parameters $(a_0, \dots, a_{d-1})'$ instead of the original

[1] This value was denoted by a_0 in the proof of Lemma 3.1.2 above.

parameters $(\lambda_1, \ldots, \lambda_d)'$. The benefit of this reparameterization is twofold: on the one hand, the new parameters $(a_0, \ldots, a_{d-1})'$ satisfy an interesting analytical property called d-monotonicity. This allows to derive results on eMO-copulas that resemble Theorems 2.3.8 and 2.3.9 on Archimedean copulas. On the other hand, it helps to study the extendibility of Marshall-Olkin distributions: we are particularly interested in such sequences $(a_0, \ldots, a_{d-1})'$ that are not only d-monotone, but which actually stem from a completely monotone sequence. The notion of complete monotonicity gives rise to a very useful link to Lévy subordinators. It is used to derive an alternative probabilistic model for extendible Marshall-Olkin distributions, which is quite convenient for applications, see the following Chapters 4, 5 and 6.

To avoid confusion, let us notice that the above motivation was given in terms of the exchangeable Marshall-Olkin distribution - without using copulas. For the remainder of this dissertation we prefer considering eMO-copulas as defined in Section 3.1 above. However, this reformulation is no loss of generality but rather a matter of convenience. The use of copulas has the effect that all numbers in the above triangular schemes are divided by $a_0 = \sum_{i=1}^{d} \binom{d-1}{i-1} \lambda_i$, which corresponds to the normalization to uniform marginals.

3.3 d-Monotone Sequences

For this dissertation it is convenient to consider a generalization of completely monotone sequences: the notion of d-monotonicity for finite sequences. Finite sequences of length d will be identified with vectors in $\mathbb{R}^d$. Intuitively, $(a_0, \ldots, a_{d-1})' \in \mathbb{R}^d$ is defined to be d-monotone if $\nabla^{j-1} a_k \geq 0$, whenever this expression is well-defined. Recall that the difference operator ∇ was introduced in Definition 2.5.1 and the expression $\nabla^{j-1} a_k = \sum_{i=0}^{j-1} \binom{j-1}{i} (-1)^i a_{k+i}$ involves the numbers $a_k, \ldots, a_{k+j-1}$, which are not defined when $j + k > d$.

Definition 3.3.1 (d-Monotone Sequence)
A finite sequence $(a_0, \ldots, a_{d-1})' \in \mathbb{R}^d$ is called *d-monotone* if it satisfies

$$\nabla^{j-1} a_k \geq 0, \quad k = 0, 1, \ldots, d-1,\ j = 1, 2, \ldots, d-k. \tag{3.3}$$

For later use the following notation is introduced:

$$\mathcal{M}_d := \Big\{ (a_0, \ldots, a_{d-1})' \in \mathbb{R}^d \,\Big|\, a_0 = 1,\ (a_0, \ldots, a_{d-1})' \text{ is } d\text{-monotone} \Big\}.$$

The sequence of the first d members of a completely monotone sequence is in particular d-monotone. In this regard, the concept of d-monotonicity is more general than that of complete monotonicity. Given $(a_0, \ldots, a_{d-1})' \in \mathbb{R}^d$, the following triangular scheme shows the indices j and k for which the expression $\nabla^{j-1} a_k$ is well-defined:

$\nabla^{j-1} a_k$	$j=1$	$j=2$	$\ldots$	$j=d-1$	$j=d$
$k=d-1$	$\nabla^0 a_{d-1}$	n.a.	$\ldots$	n.a.	n.a.
$k=d-2$	$\nabla^0 a_{d-2}$	$\nabla^1 a_{d-2}$	$\ldots$	n.a.	n.a.
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$k=1$	$\nabla^0 a_1$	$\nabla^1 a_1$	$\ldots$	$\nabla^{d-2} a_1$	n.a.
$k=0$	$\nabla^0 a_0$	$\nabla^1 a_0$	$\ldots$	$\nabla^{d-2} a_0$	$\nabla^{d-1} a_0$

Lemma 3.3.2 below states that condition (3.3) is partially redundant, i.e. d-monotonicity can equivalently be defined using fewer conditions. More precisely, it claims that if the terms on the diagonal of the above triangular scheme are all non-negative, then all terms below the diagonal are non-negative as well, i.e. the sequence $(a_0, \ldots, a_{d-1})'$ is d-monotone. For example, in the case $d = 2$ this statement is translated into

$$a_1 \geq 0 \text{ and } a_0 - a_1 \geq 0 \;\Rightarrow\; a_0 \geq 0,$$

which is obvious. A recursive argument is used to establish the general statement.

Lemma 3.3.2 (Alternative Characterization of *d*-Monotonicity)
The sequence $(a_0, \ldots, a_{d-1})' \in \mathbb{R}^d$ is d-monotone if and only if

$$\nabla^{d-1-k} a_k \geq 0, \quad k = 0, 1, \ldots, d-1.$$

Proof

Necessity is clear. For sufficiency, assume that all elements on the diagonal of the triangular scheme above are non-negative, i.e. $\nabla^{d-1-k} a_k \geq 0$, $k = 0, 1, \ldots, d-1$. We first show that this implies that all elements on the subdiagonal are non-negative as well, i.e. we show $\nabla^{d-2-k} a_k \geq 0$, $k = 0, 1, \ldots, d-2$. To this end, let $k \in \{0, \ldots, d-2\}$

and observe that

$$\nabla^{d-2-k} a_k = \nabla^{d-1-(k+1)} a_{k+1} + \nabla^{d-1-k} a_k.$$

Since both terms on the right hand side are non-negative by assumption, the elements on the subdiagonal are non-negative as well. We can now proceed iteratively by considering the "subsubdiagonal", the "subsubsubdiagonal" and so on. This establishes the claim.□

To prove that d-monotonicity is more general than complete monotonicity, a simple example for a 3-monotone sequence which is not a subsequence of a completely monotone sequence is provided.

Example 3.3.3 (Proper d-Monotone Sequence)
Consider the sequence $(a_0, a_1, a_2)' := (1, 1/2, \epsilon)'$ with $0 \leq \epsilon < 1/4$. Then by Lemma 3.3.2 it holds that $(1, 1/2, \epsilon)' \in \mathcal{M}_3$, since

$$\nabla^2 a_0 = 1 - 2\,\frac{1}{2} + \epsilon \geq 0, \quad \nabla^1 a_1 = \frac{1}{2} - \epsilon \geq 0, \quad \nabla^0 a_2 = \epsilon \geq 0.$$

However, there exists no (infinite) completely monotone sequence $\{b_k\}_{k\in\mathbb{N}_0}$ such that $(b_0, b_1, b_2)' = (1, 1/2, \epsilon)'$. If there was such a sequence, then by Theorem 2.5.4 there was a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random variable τ with values in $[0, 1]$ such that $b_k = \mathbb{E}[\tau^k]$. But Jensen's inequality would then imply that

$$\frac{1}{4} > \epsilon = b_2 = \mathbb{E}[\tau^2] \geq \mathbb{E}[\tau]^2 = b_1^2 = \frac{1}{4},$$

which is a contradiction. Hence, the finite sequence $(1, 1/2, \epsilon)'$ is a proper 3-monotone sequence in the sense that it cannot be extended to a completely monotone sequence.

It is obvious that every element $(1, a_1)'$ of $\mathcal{M}_2$ can be extended to a completely monotone sequence, e.g. by the sequence $\{a_1^k\}_{k\in\mathbb{N}_0}$, which equals the sequence of moments of a constant random variable. However, Example 3.3.3 above shows that for $d \geq 3$ this is no longer true in general. It follows from[2] [Dette, Studden (1997), Theorem 1.4.3, p. 20] that a sequence $(a_0, a_1, \ldots, a_{d-1})'$ can be extended to a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ if and only if the so-called *Hankel determinants* $\hat{H}_1, \check{H}_1, \hat{H}_2, \check{H}_2, \ldots, \hat{H}_{d-1}, \check{H}_{d-1}$

[2]This result is originally derived in the monograph [Karlin, Shapley (1953)].

are all non-negative, where for all $l \in \mathbb{N}$ with $2\,l \leq d-1$ and for all $k \in \mathbb{N}_0$ with $2\,k+1 \leq d-1$ one has

$$\hat{H}_{2\,l} := \det \begin{pmatrix} a_0 & \dots & a_l \\ \vdots & & \vdots \\ a_l & \dots & a_{2\,l} \end{pmatrix}, \quad \check{H}_{2\,l} := \det \begin{pmatrix} \nabla a_1 & \dots & \nabla a_l \\ \vdots & & \vdots \\ \nabla a_l & \dots & \nabla a_{2\,l-1} \end{pmatrix},$$

$$\hat{H}_{2\,k+1} := \det \begin{pmatrix} a_1 & \dots & a_{k+1} \\ \vdots & & \vdots \\ a_{k+1} & \dots & a_{2\,k+1} \end{pmatrix}, \quad \check{H}_{2\,k+1} := \det \begin{pmatrix} \nabla a_0 & \dots & \nabla a_k \\ \vdots & & \vdots \\ \nabla a_k & \dots & \nabla a_{2\,k} \end{pmatrix}.$$

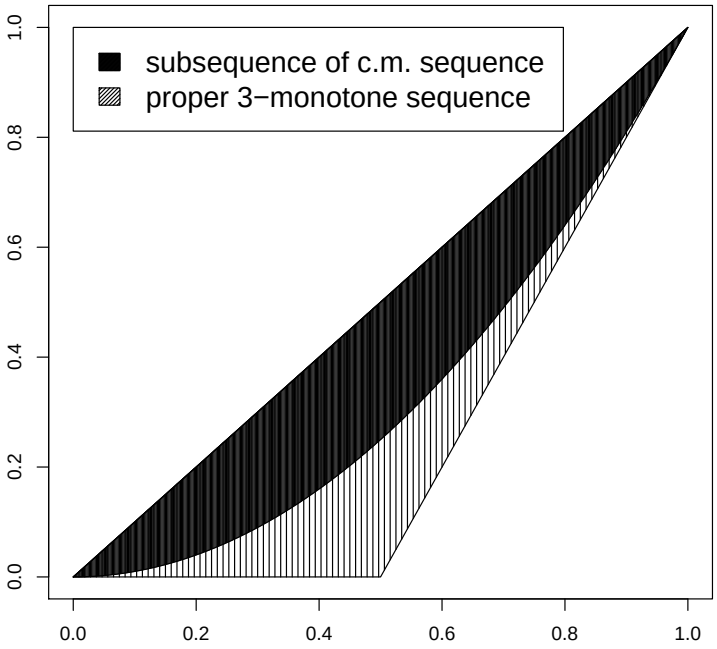


Figure 3.2 Illustration of all $(1, a_1, a_2)' \in \mathcal{M}_3$ (a_1 on the x-axis and a_2 on the y-axis). The set is subdivided into sequences that can be obtained as moments of a random variable and proper 3-monotone sequences. Computing the areas in the plot one may conclude that the share of sequences in $\mathcal{M}_3$, which can be extended to a completely monotone sequence, in all of $\mathcal{M}_3$ is 2/3.

For example, a sequence $(1, a_1, a_2)'$ is extendible to a completely monotone sequence if and only if $1 \geq a_1 \geq a_2 \geq a_1^2$. Figure 3.2 illustrates the set $\mathcal{M}_3$ and the subset of $\mathcal{M}_3$ which consists of all proper 3-monotone sequences.

In the sequel, it is shown how d-monotone sequences can be constructed. This is useful to derive a convenient parametric form of exchangeable Marshall-Olkin survival copulas. It was already mentioned that d-monotone sequences arise naturally as the sequences of moments of a random variable on $[0, 1]$. However, Example 3.3.3 shows that not every d-monotone sequence can be obtained by this approach. Therefore, the following lemma establishes a construction method which generates all members of $\mathcal{M}_d$ from d non-negative numbers $\lambda_1, \ldots, \lambda_d \geq 0$ with at least one being positive.

Lemma 3.3.4 (Construction From Non-Negative Numbers)
Define the mapping $\hat{\varphi}_d : [0, \infty)^d \setminus \{(0, \ldots, 0)'\} \to [0, \infty)^d$ by

$$\hat{\varphi}_d(\lambda_1, \ldots, \lambda_d)_k := \frac{\sum_{i=0}^{d-k-1} \binom{d-k-1}{i} \lambda_{i+1}}{\sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1}}, \quad k = 0, \ldots, d-1.$$

Then it holds that $\hat{\varphi}_d\big([0, \infty)^d \setminus \{(0, \ldots, 0)'\}\big) = \mathcal{M}_d$.

Proof
We first prove that $\hat{\varphi}_d\big([0, \infty)^d \setminus \{(0, \ldots, 0)'\}\big) \subset \mathcal{M}_d$. For given $(\lambda_1, \ldots, \lambda_d)' \in [0, \infty)^d \setminus \{(0, \ldots, 0)'\}$ define $a_k := \hat{\varphi}_d(\lambda_1, \ldots, \lambda_d)_k$ for $k = 0, \ldots, d-1$. By Lemma 3.3.2 it suffices to prove that $\nabla^{d-1-k} a_k \geq 0$ for $k = 0, \ldots, d-1$. Denote by $c := \sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1} > 0$ the positive denominator of the a_k, then

$$\begin{aligned}
\nabla^{d-1-k} a_k &= \sum_{i=0}^{d-1-k} \binom{d-1-k}{i} (-1)^i \, a_{k+i} \\
&= \frac{1}{c} \sum_{i=0}^{d-1-k} \binom{d-1-k}{i} (-1)^i \sum_{j=0}^{d-(k+i)-1} \binom{d-(k+i)-1}{j} \lambda_{j+1} \\
&= \frac{1}{c} \sum_{j=0}^{d-1-k} \lambda_{j+1} \sum_{i=0}^{d-k-j-1} (-1)^i \underbrace{\binom{d-1-k}{i} \binom{d-(k+i)-1}{j}}_{= \frac{(d-1-k)!}{(d-k-j-i-1)!\, j!\, i!}} \\
&= \frac{1}{c} \sum_{j=0}^{d-1-k} \lambda_{j+1} \sum_{i=0}^{d-k-j-1} (-1)^i \frac{(d-1-k)!\,(d-k-j-1)!}{(d-k-j-i-1)!\, j!\, i!\, (d-k-j-1)!}
\end{aligned}$$

$$= \frac{1}{c} \sum_{j=0}^{d-1-k} \lambda_{j+1} \binom{d-1-k}{j} \underbrace{\sum_{i=0}^{d-k-j-1} (-1)^i \binom{d-k-j-1}{i}}_{=(1-1)^{d-k-j-1}=\mathbf{1}_{\{j=d-k-1\}}}$$

$$= \frac{1}{c} \lambda_{d-k} \geq 0.$$

Thus, $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$.

Now we prove the reverse inclusion $\hat{\varphi}_d\big([0,\infty)^d \setminus \{(0,\ldots,0)'\}\big) \supset \mathcal{M}_d$. Consider a sequence $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$. Define the numbers $\lambda_k := \nabla^{k-1} a_{d-k}$ for $k = 1, \ldots, d$. By d-monotonicity it follows that $(\lambda_1, \ldots, \lambda_d)' \in [0,\infty)^d$. The condition $a_0 = 1$ implies that $(\lambda_1, \ldots, \lambda_d)' \neq (0, \ldots, 0)'$, since this would imply that $(a_0, \ldots, a_{d-1})' = (0, \ldots, 0)' \notin \mathcal{M}_d$. This choice of λ_k is such that the denominator in the definition of $\hat{\varphi}_d(\lambda_1, \ldots, \lambda_d)$ equals one, which is shown in the sequel. Using in the following order: the definition of λ_{i+1}, a change of the order of summation, the symmetry of the binomial coefficient, and the summation identity from Lemma 2.5.5, it holds that

$$\begin{aligned} c :&= \sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1} = \sum_{i=0}^{d-1} \binom{d-1}{i} \nabla^i a_{d-1-i} \\ &= \sum_{i=0}^{d-1} \binom{d-1}{d-1-i} \nabla^{d-1-i} a_i \\ &= \sum_{i=0}^{d-1} \binom{d-1}{i} \nabla^{d-1-i} a_i = a_0 = 1. \end{aligned}$$

Hence, the denominator c equals one and left to show is that the numerator satisfies

$$\sum_{i=0}^{d-k-1} \binom{d-k-1}{i} \underbrace{\nabla^i a_{d-i-1}}_{=\lambda_{i+1}} = a_k, \quad k = 0, \ldots, d-1.$$

To this end, define d mappings

$$f_k(a_0, \ldots, a_{d-1}) := a_k - \sum_{i=0}^{d-k-1} \binom{d-k-1}{i} \nabla^i a_{d-i-1}, \quad k = 0, \ldots, d-1.$$

Then each f_k is linear and it suffices to show that $f_k \equiv 0$ for all $k = 0, \ldots, d-1$. Considering the standard basis in $\mathbb{R}^d$, it is enough to verify that $f_k(e_0^{(l)}, \ldots, e_{d-1}^{(l)}) = 0$

for all $l = 1, \ldots, d$ where $e_j^{(l)} = \mathbf{1}_{\{l-1=j\}}$. First, we observe that

$$\begin{aligned}\nabla^i e_{d-i-1}^{(l)} &= \sum_{j=0}^{i} \binom{i}{j} (-1)^j \, e_{d-i-1+j}^{(l)} = \sum_{j=0}^{i} \binom{i}{j} (-1)^j \, \mathbf{1}_{\{j=l+i-d\}} \\ &= \binom{i}{l+i-d} (-1)^{l+i-d} \, \mathbf{1}_{\{i+l\geq d\}}.\end{aligned}$$

It follows that

$$\begin{aligned}f_k(e_0^{(l)}, \ldots, e_{d-1}^{(l)}) &= e_k^{(l)} - \sum_{i=0}^{d-k-1} \binom{d-k-1}{i} \binom{i}{l+i-d} (-1)^{l+i-d} \, \mathbf{1}_{\{i+l\geq d\}} \\ &= \mathbf{1}_{\{l-1=k\}} - \sum_{i=d-l}^{d-k-1} (-1)^{l+i-d} \binom{d-k-1}{i} \binom{i}{l+i-d} \\ &\overset{(\nu=d-i)}{=} \mathbf{1}_{\{l=k+1\}} - \sum_{\nu=k+1}^{l} (-1)^{l-\nu} \binom{d-k-1}{d-\nu} \binom{d-\nu}{l-\nu} \\ &\overset{(*)}{=} \mathbf{1}_{\{l=k+1\}} - (-1)^{-l+k+1} \binom{d-(k+1)}{d-l} \sum_{\nu=0}^{l-(k+1)} (-1)^{\nu} \binom{l-(k+1)}{\nu} \\ &= \mathbf{1}_{\{l=k+1\}} - (-1)^{l+k+1} \binom{d-(k+1)}{d-l} (1-1)^{l-(k+1)} = 0.\end{aligned}$$

Equality $(*)$ uses the fact that

$$\binom{d-k-1}{d-\nu} \binom{d-\nu}{l-\nu} = \binom{d-k-1}{d-l} \binom{l-k-1}{\nu-k-1}.$$

The claim is established. □

Remark 3.3.5 (Surjectivity of $\hat{\varphi}_d$)
It is important that $\hat{\varphi}_d$ is surjective. More precisely, the proof of Lemma 3.3.4 shows that each element $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$ can be represented as the image under $\hat{\varphi}_d$ of the vector $(\nabla^0 a_{d-1}, \nabla^1 a_{d-2}, \ldots, \nabla^{d-1} a_0)' \in [0,\infty)^d \setminus \{(0, \ldots, 0)'\}$. This surjectivity property is used in the proof of Theorem 3.4.1 below.

3.4 Reparameterization

The preceding findings on d-monotone sequences are used to derive a more convenient parameterization of the class eMO. This is the first major theorem of the present thesis.

Recall from Lemma 3.1.2 that each copula $C \in eMO$ is parameterized by d non-negative numbers $(\lambda_1, \ldots, \lambda_d)' \in [0,\infty)^d \setminus \{(0,\ldots,0)'\}$ and explicitly given by

$$C(u_1,\ldots,u_d) = \prod_{k=0}^{d-1} u_{(k+1)}^{\frac{\sum_{i=0}^{d-k-1}\binom{d-k-1}{i}\lambda_{i+1}}{\sum_{i=0}^{d-1}\binom{d-1}{i}\lambda_{i+1}}}, \tag{3.4}$$

where $u_{(1)} \leq u_{(2)} \leq \ldots \leq u_{(d)}$ denotes the ordered list of $u_1,\ldots,u_d \in [0,1]$.

Theorem 3.4.1 (Alternative Representation of the Class *eMO*)
The family eMO can alternatively be written as

$$eMO = \Big\{ \prod_{k=1}^{d} u_{(k)}^{a_{k-1}} \,\Big|\, (a_0,\ldots,a_{d-1})' \in \mathcal{M}_d \Big\}.$$

Moreover, the original parameters $(\lambda_1,\ldots,\lambda_d)' \in [0,\infty)^d \setminus \{(0,\ldots,0)'\}$ and the new parameters $(a_0,\ldots,a_{d-1})' \in \mathcal{M}_d$ are related via

$$\begin{aligned} a_k &= \hat{\varphi}_d(\lambda_1,\ldots,\lambda_d)_k = \frac{\sum_{i=0}^{d-k-1}\binom{d-k-1}{i}\lambda_{i+1}}{\sum_{i=0}^{d-1}\binom{d-1}{i}\lambda_{i+1}}, \quad k = 0,\ldots,d-1, \\ \lambda_k &= c\,\nabla^{k-1} a_{d-k}, \quad k = 1,\ldots,d, \end{aligned}$$

where $c > 0$ is an arbitrary positive constant.

Proof
At first, the inclusion "$\subset$" is investigated: Let $C(u_1,\ldots,u_d)$ denote the exchangeable Marshall-Olkin survival copula from equation (3.4), corresponding to the parameters $(\lambda_1,\ldots,\lambda_d)' \in [0,\infty)^d \setminus \{(0,\ldots,0)'\}$. It follows immediately from Lemma 3.3.4 that the exponents in (3.4) form a d-monotone sequence starting from one.
To prove the converse inclusion "$\supset$", let $(a_0,\ldots,a_{d-1})' \in \mathcal{M}_d$ and consider the function $C(u_1,\ldots,u_d) = \prod_{i=1}^{d} u_{(i)}^{a_{i-1}}$. Then, define $\lambda_k := \nabla^{k-1} a_{d-k}$ for $k \in \{1,\ldots,d\}$. By d-monotonicity it follows that $\lambda_1,\ldots,\lambda_d$ are non-negative. Since $a_0 = 1$, it is also easily seen that at least one $\lambda_k > 0$ is strictly positive. Moreover, the proof of Lemma 3.3.4 implies that

$$a_k = \hat{\varphi}_d(\lambda_1,\ldots,\lambda_d)_k = \frac{\sum_{i=0}^{d-k-1}\binom{d-k-1}{i}\lambda_{i+1}}{\sum_{i=0}^{d-1}\binom{d-1}{i}\lambda_{i+1}}, \quad k = 0,\ldots,d-1.$$

Hence, by definition the Marshall-Olkin distribution corresponding to the parameters

$\lambda_1, \ldots, \lambda_d$ has survival copula $C(u_1, \ldots, u_d)$, which establishes the claim. Obviously, the introduction of an arbitrary constant $c > 0$ does not affect the above computation. Finally, the claimed relations between the original and the new parameters are immediate from the proof of Lemma 3.3.4. □

Remark 3.4.2 (The Constant c in Theorem 3.4.1)
We know from Lemma 3.3.4 that $\hat{\varphi}_d$ is surjective. It is not injective however, since the parameters $c\,(\lambda_1, \ldots, \lambda_d)^{'} \in [0, \infty)^d \setminus \{(0, \ldots, 0)^{'}\}$ are mapped onto the same element of $\mathcal{M}_d$, independently of the choice of $c > 0$. Recall the definition (2.11) of the random vector $(\tau_1, \ldots, \tau_d)^{'}$ in the original Marshall-Olkin model associated with these parameters. Then the random variables $\tau_1, \ldots, \tau_d$ are exponentially distributed with parameter $c \sum_{i=0}^{d-1} \binom{d-1}{i} \lambda_{i+1}$. In contrast, c has no effect on the dependence structure, i.e. on the survival copula. Conversely, for a given element $(a_0, \ldots, a_{d-1})^{'} \in \mathcal{M}_d$, the choice $\lambda_k := c\,\nabla^{k-1} a_{d-k}$, $k = 1, \ldots, d$, implies that $\tau_1, \ldots, \tau_d$ are $Exp(c)$-distributed. A canonical choice is $c = 1$, as in the proof of Lemma 3.3.4.

Up to this point we have reparameterized a copula from the class eMO by a sequence $(a_0, \ldots, a_{d-1})^{'} \in \mathcal{M}_d$. This is useful, since the functional form $\prod_{k=1}^{d} u_{(k)}^{a_{k-1}}$ is convenient to derive further properties of the copula. Moreover, we will show in the next section that one can derive analytical results that are quite similar to Theorems 2.3.8 and 2.3.9 on Archimedean copulas.

3.5 A Characterization Theorem

We have seen in the previous section that for each finite sequence $(a_0, \ldots, a_{d-1})^{'} \in \mathcal{M}_d$ the function $C(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}}$ is in eMO. In particular the property of d-monotonicity of the sequence $(a_0, \ldots, a_{d-1})^{'}$ is sufficient to define a proper copula. In the present section we show that this property is also necessary, i.e. copulas in the class eMO are characterized by sequences in $\mathcal{M}_d$. As a corollary, we are then able to state an alternative version of Theorem 2.5.4. An alternative proof of sufficiency is also given, which allows as a byproduct to determine the Pickands representation for copulas of the family eMO, extending Example 2.3.3 to the multivariate case. Theorem 3.5.3 is the second major result of the present dissertation. Two technical lemmata are required for its proof.

Lemma 3.5.1 (Technical Lemma 1)
For $j \in \mathbb{N}$ and $i \in \{0, 1, \ldots, j-1\}$ the following equality holds:

$$j \sum_{l=i+1}^{j} (-1)^l \binom{j}{l} = (i+1)\,(-1)^{i+1} \binom{j}{i+1}.$$

Proof
Fix $j \in \mathbb{N}$. The claim is easily verified for $i = j-1$. Now we assume by induction hypothesis (IH) that the statement holds for $i \in \{1, \ldots, j-1\}$. It is then shown that this implies that the equality also holds for $i-1$.

$$\begin{aligned} j \sum_{l=i}^{j} (-1)^l \binom{j}{l} &= j \sum_{l=i+1}^{j} (-1)^l \binom{j}{l} + j\,(-1)^i \binom{j}{i} \\ &\overset{(IH)}{=} (i+1)\,(-1)^{i+1} \binom{j}{i+1} + j\,(-1)^i \binom{j}{i} \\ &= (-1)^i \left(\binom{j}{i} j - (i+1) \binom{j}{i+1} \right) \\ &\overset{(*)}{=} (-1)^i\, j \binom{j-1}{i-1} = (-1)^i\, i \binom{j}{i}. \end{aligned}$$

For $(*)$ it is used that

$$\binom{j}{i} = \binom{j-1}{i-1} + \binom{j-1}{i}, \quad (i+1)\binom{j}{i+1} = j\binom{j-1}{i}.$$

Thus, the claim is verified by a backward induction over i. □

Lemma 3.5.2 (Technical Lemma 2)
For an arbitrary sequence $(a_0, \ldots, a_{d-1})'$ of non-negative real numbers (not necessarily d-monotone), $w_1, \ldots, w_d \in \mathbb{R}$, and $d \in \mathbb{N}$ it holds that[3]

$$\sum_{i=1}^{d} a_{d-i}\, w_{(i)} = \sum_{j=0}^{d-1} \Big(\sum_{1 \le i_0 < \ldots < i_j \le d} \max\{w_{i_0}, \ldots, w_{i_j}\} \Big)\, \nabla^j a_{d-1-j}.$$

Proof
For $d = 1$ the claim is easily verified. Now assume the claim holds for $d \ge 1$. It is shown that the claim then also holds for $d+1$. So we consider sequences $(a_0, \ldots, a_d)' \in \mathbb{R}^{d+1}$, and $(w_1, \ldots, w_{d+1})' \in \mathbb{R}^{d+1}$. Firstly, without loss of generality assume $w_{d+1} = w_{(d+1)}$

[3] As before, $w_{(1)} \le \ldots \le w_{(d)}$ denotes the ordered list of the numbers $w_1, \ldots, w_d \in \mathbb{R}$.

(obviously, the claimed statement does not depend on the order of the w_i, so re-indexing is possible). Using the induction hypothesis (IH) on the sequences $(a_1, \ldots, a_d)'$ and $(w_1, \ldots, w_d)'$, one obtains

$$\begin{aligned}\sum_{i=1}^{d+1} a_{d+1-i}\, w_{(i)} &= \sum_{i=1}^{d} a_{d-i+1}\, w_{(i)} + a_0\, w_{d+1} \\ &\overset{(IH)}{=} \sum_{j=0}^{d-1} \Big(\sum_{1 \le i_0 < \ldots < i_j \le d} \max\{w_{i_0}, \ldots, w_{i_j}\} \Big) \nabla^j a_{d-j} + a_0\, w_{d+1}.\end{aligned}$$

Note that the first equality in the computation above uses the assumption $w_{d+1} = w_{(d+1)}$, which implies that the ordered list of the d numbers $w_1, \ldots, w_d$ equals the d smallest numbers in the ordered list of the $d+1$ numbers $w_1, \ldots, w_{d+1}$. Next it is observed that

$$\begin{aligned}\sum_{1 \le i_0 < \ldots < i_j \le d} \max\{w_{i_0}, \ldots, w_{i_j}\} &= \sum_{1 \le i_0 < \ldots < i_j \le d+1} \max\{w_{i_0}, \ldots, w_{i_j}\} \\ &\quad - \sum_{\substack{1 \le i_0 < \ldots < i_j \le d+1 \\ \text{at least one } i_l = d+1}} \max\{w_{i_0}, \ldots, w_{i_j}\}.\end{aligned}$$

Thus, establishing the claim is equivalent to showing

$$\begin{aligned}(\nabla^d a_0)\, w_{d+1} &\overset{!}{=} a_0\, w_{d+1} - \sum_{j=0}^{d-1} \Big(\sum_{\substack{1 \le i_0 < \ldots < i_j \le d+1 \\ \text{at least one } i_l = d+1}} \max\{w_{i_0}, \ldots, w_{i_j}\} \Big) \nabla^j a_{d-j} \\ &= a_0\, w_{d+1} - w_{d+1} \sum_{j=0}^{d-1} \binom{d}{j} \nabla^j a_{d-j}. \qquad (3.5)\end{aligned}$$

However, using the summation identity from Lemma 2.5.5, changing the order of summation, and using symmetry of the binomial coefficient, it follows that

$$a_0 = \sum_{j=0}^{d} \binom{d}{j} \nabla^{d-j} a_j = \sum_{j=0}^{d} \binom{d}{d-j} \nabla^j a_{d-j} = \sum_{j=0}^{d} \binom{d}{j} \nabla^j a_{d-j}.$$

The validity of equation (3.5) and thus the claim is now immediate. □

Theorem 3.5.3 (Characterization of the Class eMO)
Let $d \ge 2$ be fixed. The function $C(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}}$ is a copula if and only if $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$.

Proof

We prove Theorem 3.5.3 using Lemmata 3.5.1 and 3.5.2 above. We start by proving necessity. By the uniform marginals property of a copula one has $C(u,1,\ldots,1)=u$ for all $u\in[0,1]$. This implies $u^{a_0}=u$ for all $u\in[0,1]$. Hence, $a_0=1$. Now let $(U_1,\ldots,U_d)'$ be a random vector with distribution function $C(u_1,\ldots,u_d)$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. We define the sequence $c_0:=0$, $c_k:=\sum_{i=0}^{k-1}a_i$, $k=1,\ldots,d$. Equivalently to condition (3.3), it will be shown that $c_j>0$ and $(1/c_j)\,\nabla^{j-1}a_k\geq 0$ for all $k=0,1,\ldots,d-1$, $j=1,\ldots,d-k$. To this end, the first step is to show that $c_j>0$ for all $j=1,\ldots,d-1$ in order to guarantee that $1/c_j$ is well-defined: for $k=1,\ldots,d-1$ L'Hospital's rule implies that

$$\begin{aligned}
a_k &= \lim_{u\uparrow 1}\frac{1-\left(u^{a_0+\ldots+a_k}+(1-u^{a_0+\ldots+a_{k-1}})\right)}{1-u}\\
&= \lim_{u\uparrow 1}\frac{1-\left(\mathbb{P}(U_1\leq u,\ldots,U_{k+1}\leq u)+\mathbb{P}(\max\{U_1,\ldots,U_k\}>u)\right)}{1-u}\\
&= \lim_{u\uparrow 1}\frac{\mathbb{P}(U_1\leq u,\ldots,U_k\leq u,U_{k+1}>u)}{\mathbb{P}(U_{k+1}>u)}\geq 0.
\end{aligned}$$

This proves non-negativity of the a_k and thus shows that $c_j\geq c_1=a_0=1>0$ for all $j=1,\ldots,d$. Therefore, $1/c_j$ is well-defined and we can proceed to prove that $(1/c_j)\,\nabla^{j-1}a_k\geq 0$ for all $k=0,1,\ldots,d-1$, $j=1,\ldots,d-k$:

$$\begin{aligned}
\frac{1}{c_j}\nabla^{j-1}a_k &\overset{L.2.5.2}{=} -\frac{1}{c_j}\sum_{i=0}^{j-1}a_{i+k}\,(-1)^{i+1}\binom{j-1}{i}=-\frac{1}{j\,c_j}\sum_{i=0}^{j-1}a_{i+k}\,(-1)^{i+1}\binom{j}{i+1}(i+1)\\
&\overset{L.3.5.1}{=} -\frac{1}{c_j}\sum_{i=0}^{j-1}a_{i+k}\sum_{l=i+1}^{j}(-1)^l\binom{j}{l}=-\frac{1}{c_j}\sum_{l=1}^{j}(-1)^l\binom{j}{l}\sum_{i=0}^{l-1}a_{i+k}\\
&= -\frac{1}{c_j}\sum_{l=1}^{j}(-1)^l\binom{j}{l}(c_{l+k}-c_k)\\
&= -\frac{1}{c_j}\sum_{l=1}^{j}(-1)^l\binom{j}{l}c_{l+k}+\frac{1}{c_j}\left(\sum_{l=1}^{j}(-1)^l\binom{j}{l}\right)c_k\\
&= \lim_{u\uparrow 1}\frac{\sum_{l=1}^{j}(-1)^l\binom{j}{l}c_{l+k}\,u^{c_{l+k}-1}}{-c_j\,u^{c_j-1}}-\lim_{u\uparrow 1}\frac{-c_k\,u^{c_k-1}}{-c_j\,u^{c_j-1}}\\
&\overset{(*)}{=} \lim_{u\uparrow 1}\frac{1+\sum_{l=1}^{j}(-1)^l\binom{j}{l}u^{c_{l+k}}}{1-u^{c_j}}-\lim_{u\uparrow 1}\frac{1-u^{c_k}}{1-u^{c_j}}\\
&= \lim_{u\uparrow 1}\frac{1+\sum_{l=1}^{j}(-1)^l\sum_{1\leq m_1<\ldots<m_l\leq j}u^{c_{l+k}}}{1-u^{c_j}}-\lim_{u\uparrow 1}\frac{1-u^{c_k}}{1-u^{c_j}}.
\end{aligned}$$

In $(*)$ above the rule of L'Hospital is used. Furthermore, for $i \geq 2$ it holds that

$$u^{c_i} = u^{\sum_{s=0}^{i-1} a_s} = \prod_{s=1}^{i} u^{a_{s-1}} = C(\underbrace{u, \ldots, u}_{i}, \underbrace{1, \ldots, 1}_{d-i}) =: C^{[i]}(u, \ldots, u),$$

and by the principle of inclusion and exclusion, compare Lemma 2.2.6, with the sets

$$A_l := \{U_1 \leq u, \ldots, U_k \leq u\} \cap \{U_{k+l} \leq u\}, \quad l = 1, \ldots, j,$$

one obtains

$$\begin{aligned}
&\mathbb{P}\big(\{\max\{U_1, \ldots, U_k\} > u\} \cup \{U_{k+1} > u, \ldots, U_{k+j} > u\}\big) \\
&\quad = 1 - \mathbb{P}\Big(\cup_{l=1}^{j} A_l\Big) = 1 + \sum_{l=1}^{j} (-1)^l \sum_{1 \leq m_1 < \ldots < m_l \leq j} \mathbb{P}\Big(\cap_{z=1}^{l} A_{m_z}\Big) \\
&\quad = 1 + \sum_{l=1}^{j} (-1)^l \sum_{1 \leq m_1 < \ldots < m_l \leq j} C^{[l+k]}(u, \ldots, u) = 1 + \sum_{l=1}^{j} (-1)^l \sum_{1 \leq m_1 < \ldots < m_l \leq j} u^{c_{l+k}}.
\end{aligned}$$

Thus, it is concluded that

$$\begin{aligned}
\frac{1}{c_j} \nabla^{j-1} a_k &= \lim_{u \uparrow 1} \frac{1 + \sum_{l=1}^{j} (-1)^l \sum_{1 \leq m_1 < \cdots < m_l \leq j} u^{c_{l+k}}}{1 - u^{c_j}} - \lim_{u \uparrow 1} \frac{1 - u^{c_k}}{1 - u^{c_j}} \\
&= \lim_{u \uparrow 1} \left(\frac{\mathbb{P}\big(\{\max\{U_1, \ldots, U_k\} > u\} \cup \{U_{k+1} > u, \ldots, U_{k+j} > u\}\big)}{\mathbb{P}\big(\max\{U_1, \ldots, U_j\} > u\big)} \right. \\
&\qquad \left. - \frac{\mathbb{P}\big(\max\{U_1, \ldots, U_k\} > u\big)}{\mathbb{P}\big(\max\{U_1, \ldots, U_j\} > u\big)} \right) \\
&= \lim_{u \uparrow 1} \left(\frac{\mathbb{P}\big(\{U_1 \leq u, \ldots, U_k \leq u\} \cap \{U_{k+1} > u, \ldots, U_{k+j} > u\}\big)}{\mathbb{P}\big(\max\{U_1, \ldots, U_j\} > u\big)} \right). \qquad (3.6)
\end{aligned}$$

The computation proves non-negativity of $(1/c_j)\, \nabla^{j-1} a_k$, since probabilities are non-negative, which carries over to the limit. Concluding, it is proved that it is necessary for $(a_0, \ldots, a_{d-1})'$ to satisfy condition (3.3). Hence, it is d-monotone.

Now we turn to a new proof of sufficiency. For this, first observe that $a_0 = 1$ and d-monotonicity imply that

$$a_0 = 1, \quad \nabla^{d-1-j} a_j \geq 0, \quad j = 0, \ldots, d-1. \qquad (3.7)$$

For $j = 0, \ldots, d-1$ define $p_j^{(d-1)} := \binom{d-1}{j} \nabla^{d-1-j} a_j$. By (3.7) it follows that $p_j^{(d-1)} \geq 0$ for all $j = 0, \ldots, d-1$. Furthermore, using the summation identity of Lemma 2.5.5, it follows that the $p_j^{(d-1)}$ sum up to $a_0 = 1$. Therefore, changing the order of summation and using the symmetry of the binomial coefficient, it is verified that

$$1 = \sum_{j=0}^{d-1} p_{d-1-j}^{(d-1)} = \sum_{j=0}^{d-1} \binom{d-1}{d-1-j} \nabla^j a_{d-1-j} = \sum_{j=0}^{d-1} \binom{d-1}{j} \nabla^j a_{d-1-j}$$

is a convex combination of one (note in particular that all summands are non-negative). For $j = 0, \ldots, d-1$, on the d-dimensional unit simplex S_d we define the function

$$P_j(w_1, \ldots, w_d) := \sum_{1 \leq i_0 < \ldots < i_j \leq d} \binom{d-1}{j}^{-1} \max\{w_{i_0}, \ldots, w_{i_j}\}.$$

It is observed that P_j is a Pickands dependence function corresponding to the measure $\delta^{(j)}$, which is defined by

$$\delta^{(j)}\left(\left\{\frac{1}{j+1}(\vec{e}_{i_0} + \ldots + \vec{e}_{i_j})\right\}\right) := \frac{j+1}{\binom{d-1}{j}}, \quad 1 \leq i_0 < \ldots < i_j \leq d,$$

$\delta^{(j)}$ zero else, where $\vec{e}_l = (0, \ldots, 0, 1, 0 \ldots, 0)'$ denotes the l-th unit vector in $\mathbb{R}^d$. This can be seen from

$$\begin{aligned} &\int_{S_d} \max\{w_1 u_1, \ldots, w_d u_d\} \, \delta^{(j)}(du_1, \ldots, du_d) \\ &\quad = \sum_{1 \leq i_0 < \ldots < i_j \leq d} \frac{j+1}{\binom{d-1}{j}} \max\left\{\frac{w_{i_0}}{j+1}, \ldots, \frac{w_{i_j}}{j+1}\right\} = P_j(w_1, \ldots, w_d). \end{aligned}$$

Note in particular for all $k = 1, \ldots, d$ that

$$\int_{S_d} u_k \, \delta^{(j)}(du_1, \ldots, du_d) = \sum_{\substack{1 \leq i_0 < \ldots < i_j \leq d \\ \text{at least one } i_l = k}} \frac{j+1}{\binom{d-1}{j}} \frac{1}{j+1} = 1.$$

By [Falk et al. (2004), p. 123], Pickands dependence functions are stable under convex combinations. Hence, the function

$$P(w_1, \ldots, w_d) := \sum_{j=0}^{d-1} \binom{d-1}{j} \nabla^j a_{d-1-j} \, P_j(w_1, \ldots, w_d)$$

defines again a Pickands dependence function. Finally, it is shown in Lemma 3.5.2 that $P(w_1, \ldots, w_d) = \sum_{i=1}^{d} a_{d-i}\, w_{(i)}$. Therefore the Pickands Representation Theorem 2.3.2 shows that the function

$$\begin{aligned}\Big(\prod_{i=1}^{d} u_i\Big)^{P\left(\frac{\log u_1}{\sum_{l=1}^{d}\log u_l},\ldots,\frac{\log u_d}{\sum_{l=1}^{d}\log u_l}\right)} &= \exp\left(P\Big(\frac{\log u_1}{\sum_{l=1}^{d}\log u_l},\ldots,\frac{\log u_d}{\sum_{l=1}^{d}\log u_l}\Big)\sum_{l=1}^{d}\log u_l\right)\\ &= \exp\Big(\sum_{i=1}^{d} a_{d-i}\,\frac{\log u_{(d-i+1)}}{\sum_{l=1}^{d}\log u_l}\sum_{l=1}^{d}\log u_l\Big)\\ &= \exp\Big(\sum_{i=1}^{d} a_{i-1}\,\log u_{(i)}\Big) = \prod_{i=1}^{d} u_{(i)}^{a_{i-1}}\end{aligned}$$

is a copula. Thus, the claim is established. □

The following example illustrates Theorem 3.5.3 by writing out the condition of d-monotonicity on the sequence of parameters for the low-dimensional cases $d \in \{2, 3, 4\}$.

Example 3.5.4 (Low-Dimensional Cases)

(a) In the bivariate case, Theorem 3.5.3 corresponds to

$$\min\{u_1, u_2\}\,\max\{u_1, u_2\}^{a_1} \text{ is a copula} \;\Leftrightarrow\; 0 \le a_1 \le 1.$$

This family coincides with the well-known bivariate Cuadras-Augé family, see [Cuadras, Augé (1981)] and Definition 2.2.2.

(b) In the trivariate case, using Lemma 3.3.2 and Theorem 3.5.3 implies

$$u_{(1)}\, u_{(2)}^{a_1}\, u_{(3)}^{a_2} \text{ is a copula} \;\Leftrightarrow\; 0 \le a_2 \le a_1,\; 2\,a_1 \le 1 + a_2.$$

(c) In the case $n = 4$, using Lemma 3.3.2 and Theorem 3.5.3 implies

$$\begin{aligned}u_{(1)}\, u_{(2)}^{a_1}\, u_{(3)}^{a_2}\, u_{(4)}^{a_3} \text{ is a copula} \;\Leftrightarrow\quad & 0 \le a_3 \le a_2,\; 2\,a_2 \le a_1 + a_3,\\ & 1 - 3\,a_1 + 3\,a_2 - a_3 \ge 0.\end{aligned}$$

Remark 3.5.5 (Alternative Version of Theorem 3.5.3)

Theorem 3.5.3 can be reformulated in terms of the survival function of the exchangeable Marshall-Olkin distribution. Denoting by $t_{(1)} \le \ldots \le t_{(d)}$ the ordered list of the numbers

$t_1, \ldots, t_d \geq 0$, the function

$$(t_1, \ldots, t_d) \mapsto \bar{F}(t_1, \ldots, t_d) := \exp\Big(- \sum_{k=1}^{d} t_{(d+1-k)}\, a_{k-1} \Big) \tag{3.8}$$

is a (d-dimensional) survival function on $[0, \infty)^d$ if and only if $(a_0, \ldots, a_{d-1})'$ is d-monotone and $a_0 > 0$. In this case, (3.8) defines an exchangeable Marshall-Olkin distribution. To prove this, one rewrites

$$\bar{F}(t_1, \ldots, t_d) = C\big(e^{-a_0\, t_1}, \ldots, e^{-a_0\, t_d}\big), \quad \text{with } C(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}/a_0},$$

and applies Theorem 3.5.3 together with the survival analog of Sklar's Theorem 2.2.5.

Finally, Theorem 3.5.3 allows to formulate an alternative version of Theorem 2.5.4 by means of copulas, which deserves to be stated as a corollary.

Corollary 3.5.6 (Hausdorff's Theorem for Copulas)
The following statements are equivalent.

(a) The sequence $\{a_k\}_{k \in \mathbb{N}_0}$ is completely monotone with $a_0 = 1$.

(b) There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random variable $\tau : \Omega \to [0, 1]$ such that $a_k = \mathbb{E}[\tau^k]$ for all $k \in \mathbb{N}_0$.

(c) The function $C(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}}$ is a copula for all $d \geq 2$.

Proof
The first equivalence is the classical result of [Hausdorff (1921), Hausdorff (1923)], which was already stated in Theorem 2.5.4. For the second equivalence it is enough to observe that a sequence $\{a_k\}_{k \in \mathbb{N}_0}$ is completely monotone if and only if the finite sequence $(a_0, \ldots, a_{d-1})'$ is d-monotone for every $d \geq 2$. Thus, Theorem 3.5.3 establishes the claim. □

Remark 3.5.7 (Sufficiency of Complete Monotonicity)
In total, three different proofs of the sufficiency statement "(a) ⇒ (c)" in Corollary 3.5.6 are provided in this dissertation. The first proof uses the reparameterization of exchangeable Marshall-Olkin survival copulas from Theorem 3.4.1. This proof is the most direct one given, in the sense that it relies on basic algebra. A second proof is carried out in Theorem 3.5.3 applying the Pickands representation, see Theorem 2.3.2,

which itself is a "deep result". Finally, Chapter 4 below highlights that a completely monotone sequence implicitly defines a Lévy subordinator, which can be used to construct a random vector having the copula in concern as joint distribution function, see Theorem 4.2.2. This third proof of sufficiency relies on Theorem 2.5.8, which itself uses the original result of Hausdorff and the Lévy-Khinchin formula. Hence, it also relies on "deep results".

3.6 Analogy with Archimedean Copulas

In this section we outline some interesting analogies between Archimedean copulas, compare Subsection 2.3.3, and copulas from the class eMO. Recall that an Archimedean copula is parameterized by a function $\varphi : [0, \infty) \to [0, 1]$ with $\varphi(0) = 1$, $\lim_{x\to\infty} \varphi(x) = 0$, and has the analytical form $\varphi\big(\varphi^{-1}(u_1) + \ldots + \varphi^{-1}(u_d)\big)$. On the contrary, a copula from eMO is parameterized by a finite sequence $(a_0, \ldots, a_{d-1})'$ with $a_0 = 1$ and has the functional form $u_{(1)}^{a_0} \cdot u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$. Even though both copula families appear to be very different regarding their distributional properties, they share many analytical similarities.

The derivative of a function φ (in case of existence) is defined to be the function $\varphi^{(1)}(x) := \lim_{h\downarrow 0} \frac{1}{h}\big(\varphi(x + h) - \varphi(x)\big)$. An analogous concept for sequences is given by the difference operator ∇. The "derivative" of a sequence $\{a_k\}_{k\in\mathbb{N}_0}$ might be defined as[4] $\{-\nabla a_k\}_{k\in\mathbb{N}_0} = \{a_{k+1} - a_k\}_{k\in\mathbb{N}_0}$. In both concepts, the notion of complete monotonicity means that a function, respectively a sequence, has derivatives of all orders, which alternate in sign, i.e.

$$\begin{aligned} \varphi \text{ completely monotone } &\Leftrightarrow (-1)^n \varphi^{(n)}(x) \geq 0, \, \forall n \in \mathbb{N}_0, \, x > 0, \\ \{a_k\}_{k\in\mathbb{N}_0} \text{ completely monotone } &\Leftrightarrow \nabla^n a_k \geq 0, \, \forall n \in \mathbb{N}_0, \, k \in \mathbb{N}_0. \end{aligned}$$

Interestingly, the analytical notion of complete monotonicity has a probabilistic interpretation in both cases. [Hausdorff (1921), Hausdorff (1923)] first discovered that a sequence $\{a_k\}_{k\in\mathbb{N}_0}$ with $a_0 = 1$ is completely monotone if and only if there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random variable τ on the unit interval such that $a_k = \mathbb{E}[\tau^k]$ for all $k \in \mathbb{N}_0$, compare Theorem 2.5.4. On the other hand, [Bernstein (1929)]

[4] In the literature one often considers the difference operator $\Delta := -\nabla$, however we adopt the notation of [Gnedin, Pitman (2008)] in order to avoid writing too many minus signs.

first observed that a function φ with $\varphi(0) = 1$ is completely monotone if and only if it is the Laplace transform of a non-negative random variable, i.e. $\varphi(x) = \mathbb{E}[\exp(-x\,W)]$ for $W \geq 0$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, compare Theorem 2.3.7. Regarding copulas, [Kimberling (1974)] shows that completely monotone functions (which start at one and tend to zero as their argument tends to infinity) can alternatively be characterized by Archimedean copulas, compare Theorem 2.3.8. In a similar spirit, the results in this dissertation show that there is an analogous result which characterizes completely monotone sequences by means of exchangeable Marshall-Olkin survival copulas, see Corollary 3.5.6. Both statements give sufficient conditions for a function, respectively a sequence, to define a proper copula in every dimension. However, for fixed dimension $d \geq 2$ the notion of complete monotonicity is not necessary to define a proper copula. [McNeil, Nešlehová (2009)] characterize Archimedean copulas in a fixed dimension $d \geq 2$, compare Theorem 2.3.9. For this purpose, the notion of a completely monotone function is relaxed to the condition of d-monotonicity, which is a little more technical, compare Definition 2.3.5. The analogous characterization for copulas from the class eMO is established in Theorem 3.5.3 and uses the condition of d-monotonicity of a (finite) sequence, which may also be viewed as a relaxation of complete monotonicity, compare Definition 3.3.1. Notice that in both cases the somewhat technical definitions of d-monotonicity may be justified by these "copula characterization theorems". Table 3.1 summarizes these analytical similarities between both copula families.

Finally, considering the subclasses of Archimedean copulas and of class eMO which are given by completely monotone functions, respectively completely monotone sequences, both families admit a similar probabilistic model. Having determined the positive random variable $W > 0$ corresponding to a Laplace transform φ, it is demonstrated in Theorem 2.3.10 that the random vector $(\tau_1, \ldots, \tau_d)' := (E_1/W, \ldots, E_d/W)'$ has the Archimedean survival copula $\varphi\big(\varphi^{-1}(u_1) + \ldots + \varphi^{-1}(u_d)\big)$, where $E_1, \ldots, E_d$ are independent of W and i.i.d. $Exp(1)$-distributed. Similarly, in the next chapter we will show that a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ implicitly defines a Lévy subordinator Λ with Laplace exponent Ψ. It will be shown that the random vector of first passage times of Λ across d i.i.d. $Exp(1)$-distributed random variables has survival copula $u_{(1)}^{a_0} \cdot u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$. Thus, the common factor inducing dependence among the initially independent exponential trigger variables is a non-negative random variable in the Archimedean case, and a Lévy subordinator in the eMO case. The fact that Λ is a stochastic process instead of a random variable induces the ordering which appears in the analytical form of the copula.

Both probabilistic models provide efficient sampling routines in each dimension. This is particularly useful in large dimensions, where sampling routines are usually hard to find. On a theoretical level, completely monotone parameters correspond to *extendibility* of the distribution, i.e. it is possible to view the respective d-dimensional random vector as the d initial members of an exchangeable sequence of infinitely many random variables, see also Remark 4.2.3 below, and [Aldous (1985)] for a detailed survey on this topic. In particular, a seminal theorem of De Finetti implies the existence of a σ-field, conditioned on which this sequence of random variables is i.i.d.. Indeed, in the Archimedean setup of Theorem 2.3.10 the random variables $\tau_1, \ldots, \tau_d$ are i.i.d. conditioned on the value of W. From a Bayesian perspective, $\tau_1, \ldots, \tau_d$ are i.i.d. exponentially distributed with randomly drawn parameter $W > 0$. The analogous result in the eMO-case is the content of the next chapter. For both families, a "conditionally i.i.d.-interpretation" is impossible in the case of proper d-monotone parameters.

Table 3.1 Comparison Archimedean copulas vs. copulas from eMO.

	Archimedean copula	Copula from eMO
analytical form	$\varphi\big(\varphi^{-1}(u_1) + \ldots + \varphi^{-1}(u_d)\big)$	$u_{(1)}^{a_0}\, u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$
parameter	function φ with $\varphi(0) = 1$, and $\lim_{x\to\infty} \varphi(x) = 0$	sequence $(a_0, \ldots, a_{d-1})'$ with $a_0 = 1$
derivative operator	$-\varphi^{(1)}(x) = \lim\limits_{h\downarrow 0} \frac{1}{h}\big(\varphi(x) - \varphi(x+h)\big)$	$\nabla a_k = a_k - a_{k+1}$
completely monotone	$(-1)^n\, \varphi^{(n)}(x) \geq 0$, $\forall n \in \mathbb{N}_0$, $x > 0$	$\nabla^n a_k \geq 0$, $\forall n, k \in \mathbb{N}_0$
$\Leftrightarrow$	φ Laplace transform of $W \geq 0$	$\{a_k\}$ sequence of moments of $\tau \in [0,1]$
d-monotone	$(-1)^n\, \varphi^{(n)}(x) \geq 0$, for $n = 1, 2, \ldots, d-2$, $(-1)^{d-2}\, \varphi^{(d-2)}(x) \searrow$ and convex	$\nabla^n a_k \geq 0$, $\forall n \leq d-1$, $k \leq d-1-n$
copula $\Leftrightarrow$	φ d-monotone	$(a_0, \ldots, a_{d-1})'$ d-monotone
copula for all $d \geq 2$ $\Leftrightarrow$	φ completely monotone	$\{a_k\}_{k\in\mathbb{N}_0}$ completely monotone

4 Construction of Infinite Marshall-Olkin Sequences

"Probability does not exist."

B. De Finetti, Italian mathematician.

Given a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$, Corollary 3.5.6 in the previous chapter provides us with an infinite family of exchangeable Marshall-Olkin distributions - one for each dimension $d \geq 2$. This result indicates the existence of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which an infinite sequence of random variables is defined, such that every finite subvector of length d has the exchangeable Marshall-Olkin distribution associated with $(a_0, \ldots, a_{d-1})'$. However, the original construction of the Marshall-Olkin distribution (2.11) does not provide such an infinite sequence. It is absolutely not clear how such a probability space - respectively such an infinite Marshall-Olkin sequence - should be constructed.

This chapter shows that a given completely monotone sequence can be associated uniquely with a Lévy subordinator. The infinite sequence of first passage times of this Lévy subordinator across an infinite sequence of i.i.d. exponential threshold levels is shown to define an infinite Marshall-Olkin sequence, as desired. This alternative model is of theoretical interest and advantageous in several ways, as shown later in this thesis.

The chapter is organized as follows. Section 4.1 uses Theorem 2.5.8 to constitute a useful connection between Lévy subordinators and random variables on $[0, 1]$. Section 4.2 defines the class of *Lévy-frailty copulas* as a subclass of eMO and provides an alternative probabilistic construction. Section 4.3 presents concrete examples of Lévy-frailty copulas, and in Section 4.4 several distributional properties of the class eMO are derived, partially making use of the new construction method.

4.1 Analytical Preliminaries

The goal of this chapter is to establish a connection between Lévy subordinators and the family eMO. As a first step, this section associates certain Lévy subordinators with random variables on $[0,1]$. To begin with, the following technical lemma constitutes an intimate connection between completely alternating and completely monotone sequences.

Lemma 4.1.1 (Completely Monotone vs. Alternating Sequences)

(a) Let $\{a_k\}_{k\in\mathbb{N}_0}$ be completely monotone. Then the sequence $\{c_k\}_{k\in\mathbb{N}_0}$, defined by $c_0 := 0$, $c_k := \sum_{i=0}^{k-1} a_i$, $k \geq 1$, is completely alternating.

(b) Conversely, let $\{c_k\}_{k\in\mathbb{N}_0}$ be completely alternating with $c_0 = 0$. Then the sequence $\{a_k\}_{k\in\mathbb{N}_0}$, defined by $a_k := -\nabla c_k$, $k \geq 0$, is completely monotone.

Combining (a) and (b) yields a bijection

$$\Big\{\{a_k\}_{k\in\mathbb{N}_0} \text{ completely monotone}\Big\} \overset{1-1}{\leftrightarrow} \Big\{\{c_k\}_{k\in\mathbb{N}_0} \text{ completely alternating} \,\Big|\, c_0 = 0\Big\}.$$

Proof

Denote by Φ_1, Φ_2 the mappings from the described construction, i.e.

$$\Phi_1\big(\{a_k\}_{k\in\mathbb{N}_0}\big) = \{c_k\}_{k\in\mathbb{N}_0}, \text{ where } c_0 = 0,\ c_k = \sum_{i=0}^{k-1} a_i, \quad k \geq 1,$$
$$\Phi_2\big(\{c_k\}_{k\in\mathbb{N}_0}\big) = \{a_k\}_{k\in\mathbb{N}_0}, \text{ where } a_k = -\nabla c_k, \quad k \geq 0.$$

First notice that both mappings are well-defined: for a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ the image sequence $\{c_k\}_{k\in\mathbb{N}_0} := \Phi_1\big(\{a_k\}_{k\in\mathbb{N}_0}\big)$ is completely alternating with $c_0 = 0$, since $\nabla^j c_k = \nabla^{j-1}(c_k - c_{k+1}) = -\nabla^{j-1} a_k \leq 0$ for all $j \in \mathbb{N}$. Conversely, for a completely alternating sequence $\{c_k\}_{k\in\mathbb{N}_0}$ with $c_0 = 0$ the image sequence $\{a_k\}_{k\in\mathbb{N}_0} := \Phi_2\big(\{c_k\}_{k\in\mathbb{N}_0}\big)$ is completely monotone, since $\nabla^j a_k = -\nabla^{j+1} c_k \geq 0$ for all $j \in \mathbb{N}_0$. It is left to show that $\Phi_1 \circ \Phi_2\big(\{c_k\}_{k\in\mathbb{N}_0}\big) = \{c_k\}_{k\in\mathbb{N}_0}$ and $\Phi_2 \circ \Phi_1\big(\{a_k\}_{k\in\mathbb{N}_0}\big) = \{a_k\}_{k\in\mathbb{N}_0}$. The reader can easily convince herself that this is true. □

Heuristically speaking and using the notations from previous chapters, there is a bijection

$$\begin{aligned}
&\Big\{\Lambda \text{ Lévy subordinator} \,\Big|\, \Psi(1)=1\Big\}\\
&\quad \overset{1-1}{\leftrightarrow} \Big\{\{c_k\}_{k\in\mathbb{N}_0} \text{ completely alternating} \,\Big|\, c_0=0,\, c_1=1\Big\} \qquad \text{(by Theorem 2.5.8)}\\
&\quad \overset{1-1}{\leftrightarrow} \Big\{\{a_k\}_{k\in\mathbb{N}_0} \text{ completely monotone} \,\Big|\, a_0=1\Big\} \qquad \text{(by Lemma 4.1.1)}\\
&\quad \overset{1-1}{\leftrightarrow} \Big\{\tau \text{ random variable on } [0,1]\Big\} \qquad \text{(by Theorem 2.5.4).}
\end{aligned}$$

More correctly, this mapping - denoted by $\mathcal{H}$ in the sequel - defines a bijection between Lévy characteristics and distribution functions, i.e.

$$\begin{aligned}
\mathcal{H}: &\Big\{(\mu,\nu) \,\Big|\, \mu\in[0,1],\, \nu \text{ measure on } (0,\infty] \text{ satisfying (2.16)},\, \mu+\int_{(0,\infty]}\big(1-e^{-t}\big)\,\nu(dt)=1\Big\}\\
&\overset{1-1}{\leftrightarrow} \Big\{F:[0,1]\to[0,1] \,\Big|\, F \text{ non-decreasing and right-continuous},\, F(1)=1\Big\}.
\end{aligned}$$

In order to simplify notations, we identify a distribution function F on $[0,1]$ with a random variable τ, and the Lévy characteristics (μ,ν) with a Lévy subordinator Λ - both being defined on a generic probability space $(\Omega,\mathcal{F},\mathbb{P})$. The term "generic" means that we do not really work on the probability space $(\Omega,\mathcal{F},\mathbb{P})$. Rather we are concerned with statements in distribution, i.e. the only point in considering $(\Omega,\mathcal{F},\mathbb{P})$ is that we want to write $\mathbb{E}[\tau]$ or $\mathbb{P}(\tau\in dt)$ rather than $\int_{[0,1]} t\,dF(t)$ or $dF(t)$. Moreover, we write $\mathcal{H}(\Lambda)=\tau$, meaning actually $\mathcal{H}\big((\mu,\nu)\big)=F$, where $\tau\sim F$ and (μ,ν) are the characteristics of Λ. Keeping this terminology in mind, $\mathcal{H}$ can now be defined as follows.

Definition 4.1.2 (The Bijection $\mathcal{H}$)
For a Lévy subordinator Λ with Laplace exponent Ψ satisfying $\Psi(1)=1$, $\mathcal{H}(\Lambda)$ is the unique random variable τ on $[0,1]$ satisfying $\mathbb{E}[\tau^k]=\Psi(k+1)-\Psi(k)$, $k\in\mathbb{N}_0$. Conversely, for a random variable τ on the unit interval, $\mathcal{H}^{-1}(\tau)$ is the unique Lévy subordinator whose Laplace exponent Ψ satisfies $\Psi(k+1)-\Psi(k)=\mathbb{E}[\tau^k]$, $k\in\mathbb{N}_0$.

The explicit form of the bijection $\mathcal{H}$ is determined in the following lemma.

Lemma 4.1.3 (Explicit Form of $\mathcal{H}$)
In the statements below we apply the conventions $-\log 0:=\infty$ and $\exp(-\infty):=0$.

(a) Let Λ be a Lévy subordinator whose Laplace exponent Ψ satisfies $\Psi(1)=1$, and

denote its drift by μ and its Lévy measure by ν. Then

$$\mathbb{P}\big(\mathcal{H}(\Lambda) \in B\big) = \mu\, \mathbf{1}_{\{1\in B\}} + \int_{\{-\log b \,|\, b\in B\setminus\{1\}\}} \big(1 - e^{-t}\big)\, \nu(dt), \quad B \in \mathcal{B}\big([0,1]\big). \tag{4.1}$$

(b) Conversely, let τ be a random variable on the unit interval $[0,1]$, and denote the drift of $\mathcal{H}^{-1}(\tau)$ by μ and its Lévy measure by ν. Then

$$\mu = \mathbb{P}(\tau = 1), \quad \nu(dt) = \big(1 - e^{-t}\big)^{-1}\, \mathbb{P}(-\log \tau \in dt) \text{ on } (0,\infty]. \tag{4.2}$$

In particular, the associated Laplace exponent Ψ satisfies $\Psi(1) = 1$.

Proof

(a) Denote by κ the measure defined by the right-hand side of (4.1) and verify that

$$\kappa\big([0,1]\big) = \mu + \int_{(0,\infty]} \big(1 - e^{-t}\big)\, \nu(dt) = \Psi(1) = 1,$$

hence κ actually defines a probability measure on $[0,1]$. Furthermore, define a measure $\tilde{\nu}$ on $[0,1)$ by setting $\tilde{\nu}(B) := \nu\big(\{-\log b \,|\, b \in B\}\big)$ for $B \in \mathcal{B}\big([0,1)\big)$. It follows from the change-of-variable formula, see e.g. [Billingsley (1995), Theorem 16.13, p. 216], that

$$\int_{(0,\infty]} f\big(e^{-t}\big)\, \nu(dt) = \int_{[0,1)} f(x)\, \tilde{\nu}(dx), \quad \forall\, f : [0,1) \to [0,1]. \tag{4.3}$$

In particular, this implies for $y \in [0,1)$ that

$$\begin{aligned}\kappa\big([0,y]\big) &= \int_{(0,\infty]} \big(1 - e^{-t}\big)\, \mathbf{1}_{\{t\in[-\log y,\infty]\}}\, \nu(dt) \\ &\overset{(4.3)}{=} \int_{[0,1)} (1-x)\, \mathbf{1}_{\{-\log x\in[-\log y,\infty]\}}\, \tilde{\nu}(dx) = \int_{[0,y]} (1-x)\, \tilde{\nu}(dx),\end{aligned}$$

and hence $\kappa(dy) = (1-y)\, \tilde{\nu}(dy)$ on $[0,1)$. Consequently, this verifies

$$\begin{aligned}\int_{[0,1]} y^k\, \kappa(dy) &= \mu + \int_{[0,1)} y^k\, (1-y)\, \tilde{\nu}(dy) \\ &\overset{(4.3)}{=} \mu + \int_{(0,\infty]} e^{-k\,t}\, \big(1 - e^{-t}\big)\, \nu(dt) = \Psi(k+1) - \Psi(k).\end{aligned}$$

The claim is thus established by the definition of $\mathcal{H}(\Lambda)$.

(b) The first step is to check that the measure ν defined in (4.2) actually is a Lévy measure, i.e. satisfies (2.16). With $\epsilon > 0$ this can be seen as follows:

$$\begin{aligned}\nu\big((\epsilon,\infty]\big) &= \mathbb{E}\Big[\frac{1}{1-\tau}\,\mathbf{1}_{\{-\log\tau\in(\epsilon,\infty]\}}\Big] = \mathbb{E}\Big[\frac{1}{1-\tau}\,\mathbf{1}_{\{\tau\in[0,e^{-\epsilon})\}}\Big] \le \frac{1}{1-e^{-\epsilon}} < \infty,\\ \int_{(0,1]} t\,\nu(dt) &= \int_{(0,1]} t\,\big(1-e^{-t}\big)^{-1}\,\mathbb{P}(-\log\tau\in dt)\\ &= \mathbb{E}\Big[\frac{-\log\tau}{1-\tau}\,\mathbf{1}_{\{\tau\in[e^{-1},1)\}}\Big] \overset{(*)}{\le} e < \infty.\end{aligned}$$

Equation $(*)$ holds, since using the series expansion $\log(1+y) = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\,y^k$ implies for $y\in[\exp(-1),1)$ that

$$\begin{aligned}\Big|\frac{\log y}{1-y}\Big| &= \Big|\frac{\log\big(1+(-(1-y))\big)}{1-y}\Big| = \Big|\sum_{k=1}^{\infty}\frac{-1}{k}\,(1-y)^{k-1}\Big|\\ &\le \Big(\sum_{k=0}^{\infty}(1-y)^k\Big) = \frac{1}{y} \le e.\end{aligned}$$

Denoting by Ψ the Laplace exponent corresponding to the characteristics (μ,ν) defined in (4.2), it is finally verified for $k\in\mathbb{N}_0$ that

$$\begin{aligned}\Psi(k+1)-\Psi(k) &= \mu + \int_{(0,\infty]} e^{-k\,t}\,\big(1-e^{-t}\big)\,\nu(dt)\\ &= \mathbb{P}(\tau=1) + \int_{(0,\infty]} e^{-k\,t}\,\big(1-e^{-t}\big)\,\big(1-e^{-t}\big)^{-1}\,\mathbb{P}(-\log\tau\in dt)\\ &= \mathbb{E}\big[\tau^k\,\mathbf{1}_{\{\tau=1\}}\big] + \mathbb{E}\big[\tau^k\,\mathbf{1}_{\{\tau<1\}}\big] = \mathbb{E}[\tau^k],\end{aligned}$$

establishing the claim (in particular $\Psi(1)=1$ follows). □

The following examples are special cases of Lemma 4.1.3, which the reader can immediately verify on her own.

Example 4.1.4 (Distributions With Atoms)
If Λ is a Lévy subordinator with drift μ and Lévy measure ν such that its Laplace exponent Ψ satisfies $\Psi(1)=1$, it holds that $\mathbb{P}\big(\mathcal{H}(\Lambda)=1\big)=\mu$ and $\mathbb{P}\big(\mathcal{H}(\Lambda)=0\big)=\nu(\{\infty\})$. Moreover, for $b\in(0,1)$ it holds that

$$\mathbb{P}\big(\mathcal{H}(\Lambda)=b\big) = (1-b)\,\nu\big(\{-\log b\}\big).$$

In particular, discrete random variables correspond to discrete Lévy measures.

Example 4.1.5 (Absolutely Continuous Random Variables)
Let τ be an absolutely continuous random variable on $[0,1]$ with density f_τ defined on $(0,1)$. Then, $\mathcal{H}^{-1}(\tau)$ has zero drift $\mu = 0$ and Lévy measure given by

$$\nu(dt) = \mathbf{1}_{\{t>0\}} \frac{e^{-t}}{1-e^{-t}} f_\tau\big(e^{-t}\big)\, dt.$$

Conversely, let Λ be a Lévy subordinator with $\Psi(1) = 1$, zero drift $\mu = 0$ and a Lévy measure $\nu(dt) = f_\nu(t)\, dt$, which is absolutely continuous with respect to the Lebesgue measure on $(0,\infty)$. Then, $\tau := \mathcal{H}(\Lambda)$ is absolutely continuous with density given by

$$f_\tau(x) = \frac{1-x}{x} f_\nu\big(-\log x\big), \quad x \in (0,1).$$

4.2 The Subclass of Lévy-Frailty Copulas

Next, the subclass of *Lévy-frailty copulas* is defined. On a high level, Lévy-frailty copulas determine the subclass of eMO which allow for a probabilistic model in which the components of the respective random vector are conditionally i.i.d., see Theorem 4.2.2 and Remark 4.2.3 below.

Definition 4.2.1 (Lévy-Frailty Copula)
For a given random variable τ on the unit interval denote the Laplace exponent of $\mathcal{H}^{-1}(\tau)$ by Ψ. For $d \geq 2$, we define the copula

$$C_\Psi(u_1,\ldots,u_d) := \prod_{i=1}^{d} u_{(i)}^{\mathbb{E}[\tau^{i-1}]} = \prod_{i=1}^{d} u_{(i)}^{\Psi(i)-\Psi(i-1)}, \tag{4.4}$$

where, as before, $u_{(1)} \leq u_{(2)} \leq \ldots \leq u_{(d)}$ denotes the ordered list of $u_1,\ldots,u_d \in [0,1]$. C_Ψ is called a *Lévy-frailty copula*.

Recall from Corollary 3.5.6 that C_Ψ defined by (4.4) actually defines a proper copula. The subsequent theorem justifies the nomenclature "Lévy-frailty copula". The term "Lévy" corresponds to the underlying Lévy subordinator. The term "frailty" is chosen, since Theorem 4.2.2 shows that the dependence structure of a Lévy-frailty copula is induced by a latent factor, sometimes called a *frailty*, see e.g. [Oakes (1989)]. Theorem 4.2.2 may be considered a key result of this dissertation, since it opens the door to

several applications. Firstly, it allows to derive more distributional properties of the class eMO, some of which are difficult to prove directly. Secondly, it allows to derive sampling strategies in large dimensions. Thirdly, it naturally implies a generalization to hierarchical dependence structures which are still viable with regard to sampling applications (Chapter 5). And last but not least, the construction below may be applied to construct a multivariate default model, which is well-suited for the pricing of portfolio credit derivatives (Chapter 6).

Theorem 4.2.2 (Construction of Lévy-Frailty Copulas)
Let τ be a random variable on $[0,1]$ and let Ψ denote the Laplace exponent of the Lévy subordinator $\Lambda := \mathcal{H}^{-1}(\tau)$. A random vector $(U_1,\ldots,U_d)'$ with joint distribution function C_Ψ can be constructed as follows. Let Λ be defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Independently of Λ, let $E_1,\ldots,E_d$ be i.i.d. $Exp(1)$-distributed random variables. The random vector $(\tau_1,\ldots,\tau_d)'$, defined by

$$\tau_i := \inf\big\{t \geq 0 \,:\, \Lambda_t \geq E_i\big\}, \quad i = 1,\ldots,d,$$

has survival copula C_Ψ and each τ_i is $Exp(1)$-distributed. It follows that the random vector $(U_1,\ldots,U_d)'$, defined by $U_i := \exp(-\tau_i)$ for $i = 1,\ldots,d$, has joint distribution function C_Ψ.

Proof
According to Definition 2.4.2 of a (killed) Lévy subordinator we denote

$$\Lambda_t = \tilde{\Lambda}_t + \infty \cdot \mathbf{1}_{\{N_t \geq 1\}}, \quad t \geq 0,$$

where $\tilde{\Lambda}$ is a classical (real-valued) Lévy subordinator and N is an independent Poisson process. Referring to the Lévy-Khinchin representation, see Theorem 2.4.3, we split the Laplace exponent Ψ of Λ into two parts via

$$\Psi(x) = \mu\, x + \int_{(0,\infty)} \big(1 - e^{-t\,x}\big)\, \nu(dt) + \nu(\{\infty\})\, \mathbf{1}_{\{x>0\}} =: \tilde{\Psi}(x) + \nu(\{\infty\})\, \mathbf{1}_{\{x>0\}},\ x \geq 0,$$

where $\tilde{\Psi}$ denotes the Laplace exponent of $\tilde{\Lambda}$ and $\nu(\{\infty\})$ is the intensity of the Poisson process N. Recall that if $\nu(\{\infty\}) = 0$ this is conveniently interpreted as $\Lambda = \tilde{\Lambda}$, i.e. "$N$ never jumps". For arbitrary $t_1, t_2, \ldots, t_d \in [0,\infty)$ with ordered list $t_{(1)} \leq \ldots \leq t_{(d)}$ and

$t_{(0)} := 0$ it is verified that

$$\sum_{i=1}^{d}(d+1-i)\,\big(\tilde{\Lambda}_{t_{(i)}} - \tilde{\Lambda}_{t_{(i-1)}}\big) = \sum_{i=1}^{d}(d+1-i)\,\tilde{\Lambda}_{t_{(i)}} - \sum_{i=0}^{d-1}(d-i)\,\tilde{\Lambda}_{t_{(i)}} = \sum_{i=1}^{d}\tilde{\Lambda}_{t_i}.$$

$\tilde{\Lambda}$ being a classical Lévy subordinator implies that the random vector of increments

$$(\tilde{\Lambda}_{t_{(d)}} - \tilde{\Lambda}_{t_{(d-1)}}, \ldots, \tilde{\Lambda}_{t_{(1)}} - \tilde{\Lambda}_{t_{(0)}})'$$

has independent components and the component $\tilde{\Lambda}_{t_{(i)}} - \tilde{\Lambda}_{t_{(i-1)}}$ has the same distribution as $\tilde{\Lambda}_{t_{(i)}-t_{(i-1)}}$. Hence, one obtains

$$\mathbb{E}\Big[e^{-\sum_{i=1}^{d}\tilde{\Lambda}_{t_i}}\Big] = \prod_{i=1}^{d}\mathbb{E}\Big[e^{-(d+1-i)\,\tilde{\Lambda}_{(t_{(i)}-t_{(i-1)})}}\Big] = \prod_{i=1}^{d}\exp\Big(-(t_{(i)}-t_{(i-1)})\,\tilde{\Psi}(d+1-i)\Big).$$

Furthermore, since N is a Poisson process with intensity $\nu(\{\infty\})$, it follows with a telescope argument that

$$\mathbb{P}(N_{t_{(d)}} = 0) = e^{-\nu(\{\infty\})\,t_{(d)}} = \prod_{i=1}^{d}\exp\Big(-(t_{(i)}-t_{(i-1)})\,\nu(\{\infty\})\Big).$$

From this, using conditional independence of events in the third equality (conditioned on the σ-algebra $\sigma(\Lambda_t : t \geq 0)$), and the convention $\exp(-\infty) = 0$ in the fourth, it is straightforward to compute

$$\begin{aligned}
\bar{F}(t_1,\ldots,t_d) &:= \mathbb{P}\big(\tau_1 > t_1, \tau_2 > t_2, \ldots, \tau_d > t_d\big) = \mathbb{P}\big(E_1 > \Lambda_{t_1}, E_2 > \Lambda_{t_2}, \ldots, E_d > \Lambda_{t_d}\big) \\
&= \mathbb{E}\left[\prod_{i=1}^{d} e^{-\Lambda_{t_i}}\right] = \mathbb{E}\left[\mathbf{1}_{\{N_{t_{(d)}}=0\}}\, e^{-\sum_{i=1}^{d}\tilde{\Lambda}_{t_i}}\right] + 0 = \mathbb{P}(N_{t_{(d)}} = 0)\,\mathbb{E}\left[e^{-\sum_{i=1}^{d}\tilde{\Lambda}_{t_i}}\right] \\
&= \prod_{i=1}^{d}\exp\Big(-(t_{(i)}-t_{(i-1)})\,\big(\tilde{\Psi}(d+1-i) + \nu(\{\infty\})\big)\Big) \\
&= \prod_{i=1}^{d}\exp\Big(-(t_{(i)}-t_{(i-1)})\,\Psi(d+1-i)\Big).
\end{aligned}$$

In the univariate case, one obtains by the same argument for $i = 1, \ldots, d$ and $t \geq 0$ that

$$\bar{F}_i(t) := \mathbb{P}\big(\tau_i > t\big) = \mathbb{P}\big(E_i > \Lambda_t\big) = e^{-t\,\Psi(1)} = e^{-t}.$$

Thus, the τ_i are $Exp(1)$-distributed. By the analog of Sklar's Theorem for survival copulas, see Theorem 2.2.5, there exists a unique copula $\hat{C}$, which satisfies

$$\bar{F}(t_1,\dots,t_d) = \hat{C}\big(e^{-t_1},\dots,e^{-t_d}\big).$$

On the other hand, since $t \mapsto \exp(-t)$ is decreasing, equation (4.4) in Definition 4.2.1 implies

$$\begin{aligned} C_\Psi\big(e^{-t_1},\dots,e^{-t_d}\big) &= \prod_{i=1}^{d} e^{-t_{(i)}\,\left(\Psi(d+1-i)-\Psi(d-i)\right)} = \prod_{i=1}^{d} e^{-t_{(i)}\,\Psi(d+1-i)} \prod_{i=1}^{d-1} e^{t_{(i)}\,\Psi(d-i)} \\ &= \prod_{i=1}^{d} e^{-t_{(i)}\,\Psi(d+1-i)} \prod_{i=1}^{d} e^{t_{(i-1)}\,\Psi(d+1-i)} = \bar{F}(t_1,\dots,t_d). \end{aligned}$$

Thus, by uniqueness of $\hat{C}$, it holds that $\hat{C} = C_\Psi$. To finally see that the random vector $\big(\exp(-\tau_1),\dots,\exp(-\tau_d)\big)'$ has joint distribution function C_Ψ it suffices to observe that

$$\mathbb{P}\big(e^{-\tau_1} \le u_1,\dots,e^{-\tau_d} \le u_d\big) = \mathbb{P}\big(\tau_1 > -\log(u_1),\dots,\tau_d > -\log(u_d)\big) = C_\Psi(u_1,\dots,u_d).$$

Notice that the continuity of the exponential law allows one to replace "$\ge$" by "$>$" in the first equality. The claim is established. □

From a theoretical point of view it is remarkable that the same distribution C_Ψ can be constructed using two quite different approaches: on the one hand via the original model by [Marshall, Olkin (1967)], compare (2.11), on the other hand via Theorem 4.2.2 above. When investigating distributional properties, this allows to choose the construction among both which is more convenient to work with. The probabilistic construction idea of Theorem 4.2.2 is visualized in Figure 4.1.

Remark 4.2.3 (Exchangeability and Extendibility)
Apparently, the probabilistic construction of Theorem 4.2.2 is possible in any dimension $d \ge 2$. For instance, if an infinite sequence $E_1, E_2, \dots$ of i.i.d. exponential trigger variables is considered instead of a finite sequence, then the respective infinite sequence $\{\tau_k\}_{k\in\mathbb{N}}$ of first passage times can be defined. This sequence is *exchangeable*, meaning that each finite subvector $(\tau_{i_1},\dots,\tau_{i_d})'$ of length $d \ge 2$ is exchangeable. A seminal theo-

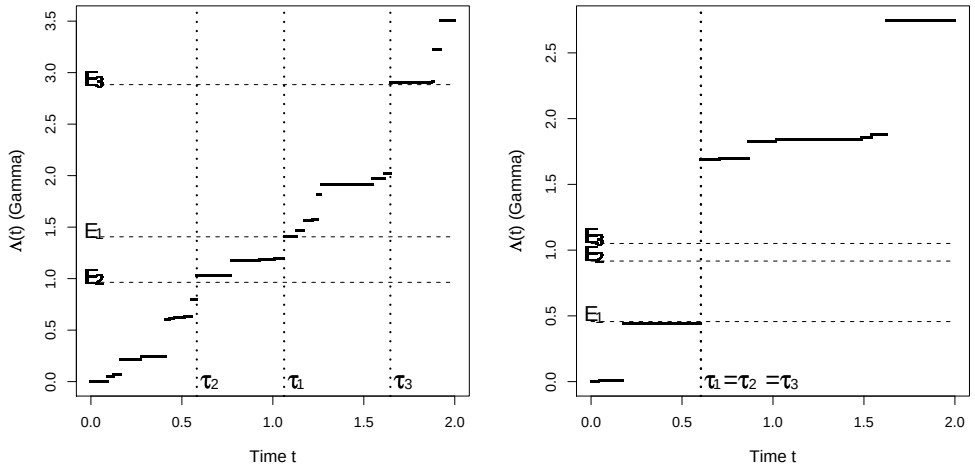


Figure 4.1 Illustration of two realizations of the random vector $(\tau_1, \tau_2, \tau_3)'$ according to Theorem 4.2.2. The simulated Lévy subordinator is a Gamma subordinator with parameters $(\beta, \eta) = (1/\log 2, 1)$. The horizontal dotted lines indicate the trigger variables E_1, E_2, E_3. The vertical dotted lines indicate the first passage times τ_1, τ_2, τ_3.

rem of De Finetti[1] states that every exchangeable sequence of random variables $\{\tau_k\}_{k\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is *conditionally i.i.d.*, i.e. there exists a σ-algebra $\mathcal{G} \subset \mathcal{F}$ such that

$$\mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d \,|\, \mathcal{G}) = \prod_{i=1}^{d} \mathbb{P}(\tau_1 > t_i \,|\, \mathcal{G}), \quad d \geq 2,\, t_1, \ldots, t_d \geq 0.$$

Indeed, $\mathcal{G}$ can be identified as $\sigma(\Lambda_t : t \geq 0)$ in the Lévy-frailty construction, and the conditional survival function is given by $\mathbb{P}(\tau_1 > t \,|\, \mathcal{G}) = \exp(-\Lambda_t)$, $t \geq 0$. In general, a d-dimensional random vector $(\tau_1, \ldots, \tau_d)'$ with eMO-survival copula (not necessarily of Lévy-frailty type) need not be extendible to an infinite exchangeable sequence. Postulating that an extension $(\tau_1, \ldots, \tau_d, \tau_{d+1})'$ of $(\tau_1, \ldots, \tau_d)'$ has an eMO-survival copula is equivalent to postulating that the parameters $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$ of the latter vector can be extended to a sequence in $\mathcal{M}_{d+1}$. For an infinite extension, one has to find a completely monotone sequence extending $(a_0, \ldots, a_{d-1})'$. As Example 3.3.3 shows, this

[1]The original reference is [De Finetti (1937)]. [Aldous (1985)] is one of the standard textbooks on the subject and provides an extensive treatment of exchangeability.

is not always possible. In this regard, Lévy-frailty copulas can be considered as precisely the subclass of eMO-copulas which are extendible. To find useful "extendibility-criteria" is a difficult problem in general, see e.g. [Aldous (1985), Problems (1.11) and (1.12), p. 9-10].

Remark 4.2.4 (Bayesian Point of View)
The probabilistic construction of Theorem 4.2.2 can also be formulated as a Bayesian two-step procedure: in a first step, a random path of the Lévy subordinator Λ is drawn, defining the distribution function $F : t \mapsto 1 - \exp(\Lambda_t)$, $t \geq 0$. Given this pre-determined distribution function, the random variables $\tau_1, \ldots, \tau_d$ are then drawn independently from F in a second step. Interestingly, the jumps of the Lévy subordinator correspond to atoms of the distribution defined by F. One effect of these atoms is the fact that events such as $\{\tau_1 = \tau_2\}$ have positive probabilities, corresponding to the singularities of the Marshall-Olkin distribution.

As a first application of Theorem 4.2.2, a specific example of a Lévy-frailty copula is now provided, which is well-known in the academic literature and serves to demonstrate the usefulness of the alternative construction. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which d i.i.d. random variables $E_1, \ldots, E_d$ with $E_1 \sim Exp(1)$ are defined. Furthermore, let $N = \{N_t\}_{t \geq 0}$ be an independent Poisson process with intensity $1/\alpha$ for $\alpha \in (0, 1)$. Define a compound Poisson subordinator Λ by $\Lambda_t := -\log(1-\alpha)\, N_t$, $t \geq 0$. The Laplace exponent Ψ of Λ is given by (see (2.19) with $\mu = 0$, $\beta = 1/\alpha$ and $J_1 \equiv -\log(1-\alpha)$)

$$\Psi(x) = \frac{1 - (1-\alpha)^x}{\alpha}, \quad x \geq 0.$$

In particular, $\Psi(1) = 1$. Moreover, for each $k \in \mathbb{N}$ it holds that $\Psi(k) - \Psi(k-1) = (1-\alpha)^{k-1}$. Notice that this implies $\mathcal{H}(\Lambda) \equiv 1 - \alpha$. By Theorem 4.2.2 the random vector $\big(\exp(-\tau_1), \ldots, \exp(-\tau_d)\big)'$, where

$$\tau_i := \inf\big\{t \geq 0 \,:\, -\log(1-\alpha)\, N_t \geq E_i\big\}, \quad i = 1, \ldots, d, \tag{4.5}$$

has joint distribution function

$$C_\alpha(u_1, \ldots, u_d) := \prod_{i=1}^{d} u_{(i)}^{(1-\alpha)^{i-1}}. \tag{4.6}$$

This one-parametric family of copulas is the exchangeable special case of a copula family first introduced in [Cuadras, Augé (1981)], hence C_α is called an *exchangeable Cuadras-Augé copula*. However, in the original reference [Cuadras, Augé (1981)] there is no probabilistic interpretation given for the multivariate distribution (4.6). Now this gap is filled. Furthermore, the probabilistic model allows to prove the following lemma, which generalizes a bivariate result from [Cuadras, Augé (1981)] regarding the singular component of the copula C_α. Lemma 4.2.5 is generalized to arbitrary eMO-copulas in Theorem 4.4.5 below. In the proof of this more general result Lemma 4.2.5 will be applied, so it is not redundant.

Lemma 4.2.5 (Singular Component of C_α)
Assume $(U_1, \ldots, U_d)'$ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and has joint distribution function C_α from equation (4.6) for some $\alpha \in [0,1]$. Then it holds that

$$\mathbb{P}\big(U_1 = U_2 = \ldots = U_d\big) = \frac{\alpha^d}{1-(1-\alpha)^d}.$$

Proof
For $\alpha \in \{0,1\}$ it holds that $C_\alpha \in \{\Pi, M\}$ and the statement is clear. For the case $\alpha \in (0,1)$, consider the vector $(\tau_1, \ldots, \tau_d)'$ as defined in equation (4.5), and denote $J := -\log(1-\alpha)$ the jump size of the corresponding Lévy subordinator Λ. One uses in the following order: the strict monotonicity of the exponential function; the construction principle and the observation that different τ_i's are identical if their corresponding E_i's are within the same interval $\big((l-1)\cdot J, l\cdot J\big]$; the disjoint union of events and the i.i.d.-property of the E_i's; the $Exp(1)$-distribution of E_1; and some algebraic manipulations.

$$\begin{aligned}
&\mathbb{P}\big(U_1 = U_2 = \ldots = U_d\big) = \mathbb{P}\big(\exp(-\tau_1) = \exp(-\tau_2) = \ldots = \exp(-\tau_d)\big) \\
&= \mathbb{P}\big(\tau_1 = \tau_2 = \ldots = \tau_d\big) = \mathbb{P}\Big(\bigcup_{l=1}^{\infty}\bigcap_{m=1}^{d}\big\{(l-1)\,J < E_m \le l\,J\big\}\Big) \\
&= \sum_{l=1}^{\infty}\mathbb{P}\big((l-1)\,J < E_1 \le l\,J\big)^d = \sum_{l=1}^{\infty}\Big(\int_{(l-1)\,J}^{l\,J} e^{-t}\,dt\Big)^d = (e^J-1)^d\sum_{l=1}^{\infty}\big(e^{-d\,J}\big)^l \\
&= (e^J-1)^d\Big(\frac{1}{1-e^{-d\,J}}-1\Big) = \frac{(e^J-1)^d\,e^{-d\,J}}{1-e^{-d\,J}} = \frac{\big(1-e^{-d\,J}\big)^d}{1-e^{-d\,J}}.
\end{aligned}$$

Finally, using the equality $J = -\log(1-\alpha)$ yields the claimed result. □

4.3 Concrete Examples

In the following we give a list of concrete examples of Lévy-frailty copulas. Some of them are studied in the literature from an alternative perspective. Others are included, since they are applied in Chapter 6 in the context of financial modeling. Table 4.1 at the end of this section provides a list of some underlying Lévy subordinators.

Example 4.3.1 (A Global Shock Copula)
Consider a random variable τ which is Bernoulli-distributed with success probability $\alpha \in (0,1)$. It follows that $\mathbb{E}[\tau^k] = \alpha + (1-\alpha)\,\mathbf{1}_{\{k=0\}}$. Denoting $\Lambda := \mathcal{H}^{-1}(\tau)$ it follows from Example 4.1.4 that the drift of Λ is given by α and the Lévy measure ν concentrates all mass at infinity, where $\nu\big(\{\infty\}\big) = 1-\alpha$. Its Laplace exponent is thus given by $\Psi(x) = \alpha\, x + (1-\alpha)\,\mathbf{1}_{\{x>0\}}$, $x \geq 0$, and the corresponding Lévy-frailty copula is

$$C_\Psi(u_1,\ldots,u_d) = u_{(1)}\Big(\prod_{i=2}^{d} u_{(i)}\Big)^{\alpha}.$$

This family of copulas appears in [Durante et al. (2007)] and in [Falk et al. (2004), Example 4.3.2, p. 124]. However, in these references it is constructed differently. In the Lévy-frailty construction, the Lévy subordinator grows linearly with constant drift α until it jumps to infinity. Hence, this dependence structure may be used to model d companies' default times, where the firms are all affected by one global shock which eliminates them simultaneously. In such a simplistic model, no default clusters are observed until all companies suddenly default at a time. This model - formulated in terms of the original Marshall-Olkin model without Lévy subordinators - is applied in the context of portfolio credit risk modeling in [Andersen, Sidenius (2005), Burtschell et al. (2009)].

Example 4.3.2 (Families obtained from Discrete Random Variables)
If τ is an arbitrary discrete random variable taking values in $\{x_k\}_{k\in\mathbb{N}}$, $x_k \in (0,1)$, with corresponding probabilities $\{p_k\}_{k\in\mathbb{N}}$, then by Example 4.1.4 the corresponding Lévy subordinator $\mathcal{H}^{-1}(\tau)$ has zero drift and a discrete Lévy measure ν. More precisely, the mass of ν is concentrated on the set $\{-\log x_k\}_{k\in\mathbb{N}}$ with corresponding weights $\{p_k/(1-x_k)\}_{k\in\mathbb{N}}$. This is a weighted sum of Poisson processes. If $\mathbb{P}(\tau = 1) > 0$, the Lévy subordinator has an additional drift, and if $\mathbb{P}(\tau = 0) > 0$ it has an additional positive killing rate.

Example 4.3.3 (A Family based on Stable Subordinators)
Consider an α-stable subordinator Λ with Laplace exponent $\Psi(x) = x^\alpha$, $x \geq 0$, for a parameter $\alpha \in [0, 1]$. Obviously $\Psi(1) = 1$, and the corresponding Lévy-frailty copula is given by

$$C_\Psi(u_1, \ldots, u_d) = u_{(1)} \prod_{i=2}^{d} u_{(i)}^{i^\alpha - (i-1)^\alpha}.$$

Example 4.3.4 (A Family based on the Gamma Subordinator)
Let $\tilde{\Lambda}$ be a Gamma subordinator with parameters $\eta > 0$ and $\beta_\eta := 1/\log(1 + 1/\eta) > 0$. From Subsection 2.4.1 we know that its Laplace exponent $\tilde{\Psi}$ is given by

$$\tilde{\Psi}(x) = \frac{\log\left(1 + \frac{x}{\eta}\right)}{\log\left(1 + \frac{1}{\eta}\right)}, \quad x \geq 0,$$

and hence $\tilde{\Psi}(1) = 1$. Furthermore, denote by $I = \{I_t\}_{t \geq 0}$ the trivial Lévy subordinator $I_t = t$, $t \geq 0$, with Laplace exponent $\Psi_I(x) = x$, $x \geq 0$. For $\beta \in (0, \beta_\eta]$ we may define a new Laplace exponent Ψ as a convex combination of Ψ_I and $\tilde{\Psi}$ via

$$\Psi(x) := \left(1 - \frac{\beta}{\beta_\eta}\right) \Psi_I(x) + \frac{\beta}{\beta_\eta}\, \tilde{\Psi}(x) = \left(1 - \frac{\beta}{\beta_\eta}\right) x + \beta \log\left(1 + \frac{x}{\eta}\right), \quad x \geq 0.$$

The corresponding Lévy subordinator is the sum of $(1 - \beta/\beta_\eta)\, I$ and a Gamma subordinator with parameters (β, η), i.e. a Gamma subordinator with additional drift. Moreover, since Ψ is a convex combination of two Laplace exponents $\tilde{\Psi}$ and Ψ_I with fixpoint 1, also $\Psi(1) = 1$. The corresponding Lévy-frailty copula has the form

$$C_\Psi(u_1, \ldots, u_d) = u_{(1)} \prod_{i=2}^{d} u_{(i)}^{1 - \beta \log\left(1 + \frac{1}{\eta}\right) + \beta \log\left(1 + \frac{1}{\eta + i - 1}\right)}.$$

Notice that the two parameters β, η of this family of copulas are restricted by $\eta > 0$ and $0 < \beta \leq 1/\log(1 + 1/\eta)$. This example is applied in Chapter 6, Section 6.5. See also family (9) in Table 4.1.

Example 4.3.5 (A Family based on the Inverse Gaussian Subordinator)
Following Subsection 2.4.1, let $\tilde{\Lambda}$ be an Inverse Gaussian subordinator with parameters $\eta > 0$ and $\beta_\eta := 1/(\sqrt{2+\eta^2} - \eta) > 0$. Its Laplace exponent $\tilde{\Psi}$ is given by

$$\tilde{\Psi}(x) = \frac{\sqrt{2\,x+\eta^2} - \eta}{\sqrt{2+\eta^2} - \eta}, \quad x \geq 0,$$

and hence $\tilde{\Psi}(1) = 1$. As in the previous example, denote by $I = \{I_t\}_{t\geq 0}$ the trivial Lévy subordinator with Laplace exponent $\Psi_I(x) = x$, $x \geq 0$. Letting $\beta \in (0, \beta_\eta]$, we define a new Laplace exponent Ψ as a convex combination of Ψ_I and $\tilde{\Psi}$ via

$$\Psi(x) := \Big(1 - \frac{\beta}{\beta_\eta}\Big)\,\Psi_I(x) + \frac{\beta}{\beta_\eta}\,\tilde{\Psi}(x) = \Big(1 - \frac{\beta}{\beta_\eta}\Big)\,x + \beta\,\big(\sqrt{2\,x+\eta^2} - \eta\big), \quad x \geq 0.$$

The corresponding Lévy subordinator is the sum of $(1-\beta/\beta_\eta)\,I$ and an Inverse Gaussian subordinator with parameters (β, η), i.e. an Inverse Gaussian subordinator with additional drift. Furthermore, since Ψ is a convex combination of two Laplace exponents $\tilde{\Psi}$ and Ψ_I with fixpoint 1, also $\Psi(1) = 1$. The corresponding Lévy-frailty copula has the form

$$C_\Psi(u_1, \ldots, u_d) = u_{(1)} \prod_{i=2}^{d} u_{(i)}^{1-\beta\,(\sqrt{2+\eta^2}-\eta)+\beta\,\left(\sqrt{2\,i+\eta^2}-\sqrt{2\,(i-1)+\eta^2}\right)}.$$

Notice that the two parameters β, η of this family of copulas have to satisfy $\eta > 0$ and $0 < \beta \leq 1/(\sqrt{2+\eta^2} - \eta)$. This family of copulas is also applied in Chapter 6, Section 6.5. See also family (10) in Table 4.1.

Example 4.3.6 (A Family based on the Beta Distribution)
Consider a Beta-distributed random variable $\tau \sim \mathit{Beta}(p, q)$ with $p, q > 0$, i.e. the density of τ is given by

$$f_\tau(x) = \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)}\,x^{p-1}\,(1-x)^{q-1}\mathbf{1}_{\{x\in(0,1)\}}.$$

One obtains from Example 4.1.5 that $\mathcal{H}^{-1}(\tau)$ has zero drift and Lévy measure

$$\nu(dt) = \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)}\,e^{-p\,t}\,(1-e^{-t})^{q-2}\,\mathbf{1}_{\{t>0\}}\,dt.$$

In this case, according to [Gupta, Nadarajah (2004), p. 35], the sequence of moments of

τ, which determines the corresponding Lévy-frailty copula, is given by

$$\mathbb{E}[\tau^k] = \frac{\Gamma(p+k)\,\Gamma(p+q)}{\Gamma(p)\,\Gamma(p+q+k)}, \quad k \in \mathbb{N}_0.$$

In the case $q = 2$ the Lévy subordinator $\mathcal{H}^{-1}(\tau)$ is a compound Poisson process. Its jump intensity is $1 + p$, the jump size distribution is an exponential distribution with parameter p. See also family (4) in Table 4.1.

As a special case, the distribution $Beta(\alpha, 1-\alpha)$, for $\alpha \in (0,1)$, is the so-called *generalized arcsine law*. A random variable $\tau \sim Beta(\alpha, 1-\alpha)$ can be constructed as

$$\tau := \sup\{s < 1 \,:\, \exists\, t \geq 0 \text{ such that } \Lambda_t = s\},$$

for an α-stable subordinator Λ, see [Bertoin (1999), Proposition 3.1, p. 23]. However, $\mathcal{H}^{-1}(\tau)$ is not an α-stable subordinator. The case $\alpha = 1/2$ is of particular interest, since in this case the sequence of moments of τ is given by

$$a_k := \mathbb{E}[\tau^k] = \frac{1}{2^{2\,k}} \binom{2\,k}{k}, \quad k \in \mathbb{N}_0,$$

and in a certain sense lies in the "center" of the set of all completely monotone sequences, see [Dette, Studden (1997), p. 303-306]. Furthermore, the (random) set

$$\left\{1 - e^{-y} \in [0,1] \,\middle|\, \exists\, t \geq 0 \text{ such that } \mathcal{H}^{-1}(\tau)_t = y\right\},$$

which is a transformation of the range of the associated Lévy subordinator $\mathcal{H}^{-1}(\tau)$, is identical in distribution with the set of zeros of a Brownian bridge on $[0,1]$, see [Gnedin, Pitman (2005), p. 469].

Example 4.3.7 (Families obtained from non-negative Random Variables)
Consider a Gamma-distributed random variable $\tilde{\tau} \sim \Gamma(b,p)$ for $b, p > 0$, i.e. $\tilde{\tau}$ is defined via the density $f_{\tilde{\tau}}(y) = b^p\, y^{p-1} e^{-b\,y}\, \mathbf{1}_{\{y>0\}}/\Gamma(p)$. Choosing $\tau := \exp(-\tilde{\tau})$, the sequence of moments of τ is given by $a_k = b^p/(b+k)^p$, $k \in \mathbb{N}_0$. The associated Lévy subordinator $\mathcal{H}^{-1}(\tau)$ has zero drift and by a density transformation and Example 4.1.5 its Lévy measure is given by

$$\nu(dt) = \frac{b^p}{\Gamma(p)} \frac{e^{-t\,b}}{1 - e^{-t}} t^{p-1}\, \mathbf{1}_{\{t>0\}}\, dt.$$

$\mathcal{H}^{-1}(\tau)$ is of compound Poisson type for $p > 1$ and the jump intensity of this process in the case $b = 1$ equals the value of the Riemann zeta function at $p > 1$. This can be seen from the fact that

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = \sum_{k=0}^{\infty} \frac{1}{(1+k)^p} = \sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \Psi(k+1) - \Psi(k) = \lim_{n\to\infty} \Psi(n) = \nu\big((0,\infty]\big).$$

The last equality follows from the dominated convergence theorem. Recall that $\nu\big((0,\infty]\big)$ gives precisely the intensity of a compound Poisson process by virtue of (2.18). Moreover, if $b = p = 1$, τ is uniformly distributed and the Laplace exponent of $\mathcal{H}^{-1}(\tau)$ interpolates the harmonic series.

Generalizing the Gamma distribution to the case when $\tau := \exp(-\tilde{\tau})$, where $\tilde{\tau}$ is an arbitrary absolutely continuous random variable on $(0,\infty)$ with density $f_{\tilde{\tau}}$, $\mathcal{H}^{-1}(\tau)$ has zero drift and its Lévy measure is given by $\nu(dt) = f_{\tilde{\tau}}(t)/(1-\exp(-t))\,dt$.

Table 4.1 Some parametric families of Lévy subordinators with $\Psi(1) = 1$.

	$\Psi(x) =$	Parameter(s)	Distribution of Λ
		One-parametric families	
(1)	x^α	$\alpha \in (0,1)$	$\Lambda_t \sim S\big(\alpha, 1, \big(\cos(\pi\,\alpha/2)\,t\big)\big)^{\frac{1}{\alpha}}, 0; 1\big)$
(2)	$\frac{1-(1-\alpha)^x}{\alpha}$	$\alpha \in (0,1)$	$\Lambda = CPP(0, 1/\alpha, \delta_{-\log(1-\alpha)})$
(3)	$\alpha\,x + (1-\alpha)\,\mathbf{1}_{\{x>0\}}$	$\alpha \in (0,1)$	$\mathbb{P}(\Lambda_t = \alpha\,t) = e^{-(1-\alpha)\,t} = 1 - \mathbb{P}(\Lambda_t = \infty)$
(4)	$(1+\alpha)\,\frac{x}{\alpha+x}$	$\alpha > 0$	$\Lambda = CPP(0, \alpha+1, Exp(\alpha))$
(5)	$\log\Big(1+\frac{x}{\alpha}\Big)\Big/\log\Big(1+\frac{1}{\alpha}\Big)$	$\alpha > 0$	$\Lambda_t \sim \Gamma\Big(t/\log\Big(1+\frac{1}{\alpha}\Big), \alpha\Big)$
(6)	$\big(\sqrt{2\,x+\alpha^2}-\alpha\big)\Big/\big(\sqrt{2+\alpha^2}-\alpha\big)$	$\alpha > 0$	$\Lambda_t \sim IG\Big(t/\big(\sqrt{2+\alpha^2}-\alpha\big), \alpha\Big)$
		Two-parametric families	
(7)	$\big(1-\beta\,\big(1-e^{-\eta}\big)\big)\,x + \beta\,\big(1-e^{-x\,\eta}\big)$	$\eta > 0,\ 0 < \beta \le 1/\big(1-e^{-\eta}\big)$	$\Lambda = CPP(1-\beta+\beta\,e^{-\eta}, \beta, \delta_\eta)$
(8)	$\Big(1-\frac{\beta}{\eta+1}\Big)\,x + \frac{\beta\,x}{\eta+x}$	$\eta > 0,\ 0 < \beta \le \eta+1$	$\Lambda = CPP(1-\frac{\beta}{\eta+1}, \beta, Exp(\eta))$
(9)	$\Big(1-\beta\,\log\Big(1+\frac{1}{\eta}\Big)\Big)\,x + \beta\,\log\Big(1+\frac{x}{\eta}\Big)$	$\eta > 0,\ 0 < \beta \le 1/\log\Big(1+\frac{1}{\eta}\Big)$	$\Lambda_t = \Big(1-\beta\,\log\Big(1+\frac{1}{\eta}\Big)\Big)\,t + \Gamma(\beta\,t, \eta)$
(10)	$\big(1-\beta\,(\sqrt{2+\eta^2}-\eta)\big)\,x + \beta\,(\sqrt{2\,x+\eta^2}-\eta)$	$\eta > 0,\ 0 < \beta \le 1/(\sqrt{2+\eta^2}-\eta)$	$\Lambda_t = \big(1-\beta\,(\sqrt{2+\eta^2}-\eta)\big)\,t + IG(\beta\,t, \eta)$

$CPP(\mu, \beta, F)$ denotes a compound Poisson process with drift μ, intensity β, and jump distribution F. δ_x denotes Dirac measure at x.

4.4 Distributional Properties of the Class eMO

This section collects several distributional properties of eMO-copulas, such as the computation of multivariate dependence measures. Some of them may directly be derived from the proof of Theorem 3.5.3 as byproducts. Others however, have to be computed separately. The reason for placing this section here, and not in the previous chapter, is that the proof of the Technical Lemma 4.4.6 below relies on Lemma 4.2.5, which itself is based on the construction with Lévy subordinators. Hence, all properties which rely on this lemma (Theorem 4.4.5 and Corollary 4.4.7 below) could not have been derived in the previous chapter.

To begin with, the proof of Theorem 3.5.3 allows to derive some properties of copulas in the class eMO as byproducts. First of all, copulas of the form $C(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_{(i)}^{a_{i-1}}$ are extreme-value copulas, i.e. $C(u_1^t, \ldots, u_d^t) = C(u_1, \ldots, u_d)^t$ for all $t > 0$, compare Definition 2.3.1. The corresponding Pickands representation from Theorem 2.3.2 is given in Corollary 4.4.1 below. The associated measure δ on the d-dimensional unit simplex S_d is a discrete measure.

Corollary 4.4.1 (Pickands Representation of the Class eMO)
Assume $(a_0, \ldots, a_{d-1})' \in \mathcal{M}_d$. The measure δ on the d-dimensional unit simplex S_d corresponding to the Pickands dependence function of the extreme-value copula $C(u_1, \ldots, u_d) = u_{(1)}^{a_0} \cdot u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$ is given by

$$\delta = \sum_{j=0}^{d-1} \binom{d-1}{j} \nabla^j a_{d-1-j}\, \delta^{(j)},$$

where for $j \in \{0, \ldots, d-1\}$ the measure $\delta^{(j)}$ is discrete and defined by

$$\delta^{(j)}\left(\left\{\frac{1}{j+1}(\vec{e}_{i_0} + \ldots + \vec{e}_{i_j})\right\}\right) = \frac{j+1}{\binom{d-1}{j}}, \quad 1 \le i_0 < \ldots < i_j \le d,$$

$\delta^{(j)}$ zero else, and $\vec{e}_l = (0, \ldots, 0, 1, 0, \ldots, 0)'$ is the l-th unit vector in $\mathbb{R}^d$. In particular, the Pickands dependence function is given by $P(w_1, \ldots, w_d) = \sum_{i=1}^{n} a_{d-i}\, w_{(i)}$, where $w_{(1)} \le w_{(2)} \le \ldots \le w_{(d)}$ is the ordered list of the numbers $w_1, \ldots, w_d$.

Proof
This representation is derived in the proof of sufficiency in Theorem 3.5.3. □

Using the formulas in [Li (2009)], one may apply Corollary 4.4.1 to compute various coefficients of tail dependence. Moreover, Corollary 4.4.1 extends [Falk et al. (2004), Example 4.3.2], where the Pickands dependence function is derived for the case $a_1 = \ldots = a_{d-1} = \alpha$, which corresponds to family (3) in Table 4.1. Next, a probabilistic interpretation of the parameters $(a_0, \ldots, a_{d-1})'$ is given. In the bivariate case $d = 2$ an eMO-copula $u_{(1)}\, u_{(2)}^{a_1}$ boils down to a bivariate Cuadras-Augé copula with parameter $1 - a_1$. Thus, the coefficient of upper-tail dependence is $1 - a_1$ by Example 2.2.10. In some sense the following corollary extends this result to larger dimensions.

Corollary 4.4.2 (Formula for a_k)
A probabilistic interpretation of $(a_1, \ldots, a_{d-1})'$ is available, namely

$$\begin{aligned} a_k &= \lim_{t \uparrow 1} \mathbb{P}\big(U_1 \leq t, \ldots, U_k \leq t \,\big|\, U_{k+1} > t\big) \\ &= 1 - \lim_{t \uparrow 1} \mathbb{P}\big(\max\{U_1, \ldots, U_k\} > t \,\big|\, U_{k+1} > t\big), \quad k = 1, \ldots, d-1, \end{aligned}$$

where $(U_1, \ldots, U_d)'$ is a random vector with joint distribution function $C(u_1, \ldots, u_d) = u_{(1)}^{a_0} \cdot u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Proof
The parameter $a_k = (1/a_0)\, \nabla^0 a_k$ is computed as byproduct of the proof of necessity in Theorem 3.5.3, compare equation (3.6) with $j = 1$, page 80. □

Another direct consequence of the proof of Theorem 3.5.3 is a formula for the upper-extremal dependence coefficient as introduced in Definition 2.2.11.

Corollary 4.4.3 (Upper-Extremal Dependence Coefficient)
Given the copula $C(u_1, \ldots, u_d) = u_{(1)}^{a_0}\ u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$, its upper-extremal dependence coefficient is given by

$$UEDC_C = \frac{\nabla^{d-1} a_0}{a_0 + a_1 + \ldots + a_{d-1}}.$$

Proof
With $c_d := a_0 + a_1 + \ldots + a_{d-1}$ the expression $(1/c_d)\, \nabla^{d-1} a_0$ is computed as byproduct of the proof of necessity in Theorem 3.5.3, compare equation (3.6) with $k = 0$ and $j = d$, page 80. □

It is further possible to compute the multivariate Spearman's Rho, compare Definition 2.2.12, for the copula $C(u_1,\dots,u_d) = u_{(1)}^{a_0}\, u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$.

Theorem 4.4.4 (Spearman's Multivariate Rho)
Suppose one is given the copula $C(u_1,\dots,u_d) = u_{(1)}^{a_0} \cdot u_{(2)}^{a_1} \cdots u_{(d)}^{a_{d-1}}$. Spearman's Multivariate Rho ρ_C^d for C is given by

$$\rho_C^d = \frac{d+1}{2^d-(d+1)} \left(2^{d-1}\, d! \left(\prod_{k=2}^{d} \left(k + \sum_{i=0}^{k-1} a_i \right) \right)^{-1} - 1 \right).$$

Proof
First recall that $a_0 = 1$ by Theorem 3.5.3. For an arbitrary sequence $(a_1,\dots,a_{d-1})'$ of non-negative numbers denote

$$I_d^{(a_1,\dots,a_{d-1})} := d! \int_0^1 u_1 \int_{u_1}^1 u_2^{a_1} \int_{u_2}^1 u_3^{a_2} \dots \int_{u_{d-1}}^1 u_d^{a_{d-1}}\, du_d \dots du_1.$$

Solving the innermost integral it is observed for $d \geq 3$ that

$$\begin{aligned} I_d^{(a_1,\dots,a_{d-1})} &= \frac{d!}{1+a_{d-1}} \int_0^1 u_1 \int_{u_1}^1 u_2^{a_1} \int_{u_2}^1 u_3^{a_2} \dots \int_{u_{d-2}}^1 u_{d-1}^{a_{d-2}} \left(1 - u_{d-1}^{a_{d-1}+1}\right) du_{d-1} \dots du_1 \\ &= \frac{d}{1+a_{d-1}} \left(I_{d-1}^{(a_1,\dots,a_{d-2})} - I_{d-1}^{(a_1,\dots,a_{d-3},a_{d-2}+a_{d-1}+1)} \right). \end{aligned} \tag{4.7}$$

Since it is easily verified that $I_2^{(a_1)} = 1/(a_1+3)$, equation (4.7) implies a recursion for $I_d^{(a_1,\dots,a_{d-1})}$ which is used to prove that

$$2\, I_d^{(a_1,\dots,a_{d-1})} = d! \left(\prod_{k=2}^{d} \left(k + \sum_{i=0}^{k-1} a_i \right) \right)^{-1}. \tag{4.8}$$

Equation (4.8) is easily verified in the case $d = 2$. The general formula is obtained by induction over d. The first equality in the following computation uses the recursion (4.7), the second equality the induction hypothesis (IH).

$$\begin{aligned} 2\, I_{d+1}^{(a_1,\dots,a_d)} &= \frac{d+1}{1+a_d} \left(2\, I_d^{(a_1,\dots,a_{d-1})} - 2\, I_d^{(a_1,\dots,a_{d-2},a_{d-1}+a_d+1)} \right) \\ &\overset{(IH)}{=} \frac{d+1}{1+a_d} \left(\prod_{k=2}^{d} \frac{k}{k+\sum_{i=0}^{k-1} a_i} - \left(\prod_{k=2}^{d-1} \frac{k}{k+\sum_{i=0}^{k-1} a_i} \right) \frac{d}{d+1+a_d+\sum_{i=0}^{d-1} a_i} \right) \end{aligned}$$

$$= \frac{d+1}{1+a_d} \Big(\prod_{k=2}^{d-1} \frac{k}{k+\sum_{i=0}^{k-1} a_i} \Big) \Big(\frac{d}{d+\sum_{i=0}^{d-1} a_i} - \frac{d}{d+1+a_d+\sum_{i=0}^{d-1} a_i} \Big)$$

$$= \frac{d\,(d+1)}{1+a_d} \Big(\prod_{k=2}^{d-1} \frac{k}{k+\sum_{i=0}^{k-1} a_i} \Big) \frac{1+a_d}{\big(d+\sum_{i=0}^{d-1} a_i\big)\,\big(d+1+\sum_{i=0}^{d} a_i\big)}$$

$$= \frac{d+1}{d+1+\sum_{i=0}^{d} a_i} \prod_{k=2}^{d} \frac{k}{k+\sum_{i=0}^{k-1} a_i} = \prod_{k=2}^{d+1} \frac{k}{k+\sum_{i=0}^{k-1} a_i}.$$

Thus, (4.8) is established. Denoting by $\mathcal{P}_d$ the set of all $d!$ permutations on $\{1, 2, \ldots, d\}$, the invariance of C with respect to permutations of its arguments implies

$$\begin{aligned} \int_{[0,1]^d} C(u_1, \ldots, u_d)\, d(u_1, \ldots, u_d) &= \sum_{\pi \in \mathcal{P}_d} \int_{\{u_{\pi(1)} < \ldots < u_{\pi(d)}\}} C(u_1, \ldots, u_d)\, d(u_1, \ldots, u_d) \\ &= d! \int_{\{u_1 < \ldots < u_d\}} C(u_1, \ldots, u_d)\, d(u_1, \ldots, u_d) = I_d^{(a_1, \ldots, a_{d-1})}. \end{aligned}$$

Applying this observation to the definition of ρ_C^d gives

$$\begin{aligned} \rho_C^d &= \frac{d+1}{2^d - (d+1)} \Big(2^d \int_{[0,1]^d} C(u_1, \ldots, u_d)\, d(u_1, \ldots, u_d) - 1 \Big) \\ &= \frac{d+1}{2^d - (d+1)} \Big(2^{d-1}\, 2\, I_d^{(a_1, \ldots, a_{d-1})} - 1 \Big). \end{aligned}$$

The claim then follows from equation (4.8). □

A further distinctive property of Marshall-Olkin distributions is their singular component. In particular, the probability that two or more components of a random vector, whose joint distribution function is a copula in the class eMO, are identical is positive, compare e.g. Lemma 4.2.5. It is possible to compute this probability explicitly in the general case, which is carried out in Theorem 4.4.5 below. It is a generalization of Lemma 4.2.5. Notice that the statement is remarkable not only because the probability of a "global shock" can be computed explicitly, but also because of the connection between the singular component and the upper-extremal dependence coefficient. It again illustrates the usefulness of Lévy-frailty copulas, on which a part of the proof relies.

Theorem 4.4.5 (The Singular Component of the Class *eMO*)
Let $(U_1,\ldots,U_d)'$ have joint distribution function $C(u_1,\ldots,u_d) = u_{(1)}^{a_0}\cdot u_{(2)}^{a_1}\cdots u_{(d)}^{a_{d-1}}$ from the class eMO. Then it holds that

$$\mathbb{P}(U_1 = \ldots = U_d) = \frac{\nabla^{d-1} a_0}{\sum_{i=0}^{d-1} a_i} = UEDC_C.$$

Proof
The second equality of the statement follows from Corollary 4.4.3. The first equality is easier to establish in the original Marshall-Olkin model, so we apply Theorem 3.4.1. Consider the copula $u_{(1)}^{a_0}\cdot u_{(2)}^{a_1}\cdots u_{(d)}^{a_{d-1}}$ for a sequence $(a_0,\ldots,a_{d-1})' \in \mathcal{M}_d$. Using the reparameterization from Theorem 3.4.1, we consider the original parameters $(\lambda_1,\ldots,\lambda_d)' \in [0,\infty)^d \setminus \{(0,\ldots,0)'\}$ defined by $\lambda_k := \nabla^{k-1} a_{d-k}$. Recall the probability space $(\Omega,\mathcal{F},\mathbb{P})$ from the original Marshall-Olkin construction. I.e. for each $\emptyset \neq I \subset \{1,\ldots,d\}$ let $E_I \sim Exp(\lambda_{|I|})$ (or $E_I \equiv \infty$ if $\lambda_{|I|} = 0$), and assume all these $2^d - 1$ random variables are independent. Further recall that the minimum of independent exponentially distributed random variables is exponentially distributed, and the parameter of the minimum equals the sum of the parameters. It follows from this fact that

$$\min\big\{E_I \,\big|\, 1 \le |I| \le d-1\big\} \sim Exp\Big(\sum_{l=1}^{d-1}\binom{d}{l}\lambda_l\Big).$$

Denoting $\tilde{\lambda} := \sum_{l=1}^{d-1}\binom{d}{l}\lambda_l$, it follows that the two random variables

$$E_{\{1,\ldots,d\}} \quad \text{and} \quad \min\big\{E_I \,\big|\, 1 \le |I| \le d-1\big\}$$

are independent and exponentially distributed with means $1/\lambda_d$ and $1/\tilde{\lambda}$, respectively. Hence, using the original probabilistic model defining the Marshall-Olkin distribution, compare (2.11), we know that C is the joint distribution function of the random vector $(U_1,\ldots,U_d)'$ defined by $U_i := \exp(-a_0\,\tau_i)$, $i = 1,\ldots,d$, where the τ_i are constructed from the E_I's as in (2.11), and $a_0 = \sum_{l=0}^{d-1}\binom{d-1}{l}\lambda_{l+1} = 1$ (recall this from the proof of Lemma 3.3.4). It follows that

$$\begin{aligned}\mathbb{P}(U_1 = \ldots = U_d) &= \mathbb{P}\Big(E_{\{1,\ldots,d\}} \le \min\big\{E_I \,\big|\, 1 \le |I| \le d-1\big\}\Big)\\ &= \int_0^\infty\int_x^\infty \tilde{\lambda}\, e^{-\tilde{\lambda}\,y}\,\lambda_d\, e^{-\lambda_d\,x}\,dy\,dx = \frac{\lambda_d}{\sum_{l=1}^{d}\binom{d}{l}\lambda_l} = \frac{\nabla^{d-1}a_0}{\sum_{l=1}^{d}\binom{d}{l}\nabla^{l-1}a_{d-l}}. \qquad (4.9)\end{aligned}$$

Finally, one can simplify the denominator in (4.9) by Lemma 4.4.6 below to obtain the claimed statement. □

Lemma 4.4.6 (Technical Lemma 3)
For arbitrary numbers $a_0, a_1, \ldots, a_{d-1}$ it holds that

$$\sum_{i=1}^{d} \binom{d}{i} \nabla^{i-1} a_{d-i} = \sum_{i=0}^{d-1} a_i.$$

Proof
We start with a re-summation.

$$\begin{aligned}
\sum_{i=1}^{d} \binom{d}{i} \nabla^{i-1} a_{d-i} &= \sum_{i=0}^{d-1} \binom{d}{d-i} \nabla^{d-i-1} a_i \\
&\overset{L.2.5.2}{=} \sum_{i=0}^{d-1} \binom{d}{d-i} \sum_{j=0}^{d-i-1} \binom{d-i-1}{j} (-1)^j \, a_{i+j} \\
&= \sum_{j=0}^{d-1} (-1)^j \sum_{i=0}^{d-j-1} \binom{d}{d-i} \binom{d-i-1}{j} a_{i+j} \\
&\overset{(k=i+j)}{=} \sum_{j=0}^{d-1} (-1)^j \sum_{k=j}^{d-1} \binom{d}{d-k+j} \binom{d-k+j-1}{j} a_k \\
&= \sum_{k=0}^{d-1} a_k \sum_{j=0}^{k} (-1)^j \binom{d}{d-k+j} \binom{d-k+j-1}{j}.
\end{aligned}$$

Denoting

$$\kappa_k := \sum_{j=0}^{k} (-1)^j \binom{d}{d-k+j} \binom{d-k+j-1}{j}, \quad k = 0, \ldots, d-1,$$

left to show is that $\kappa_k = 1$ for all $k \in \{0, \ldots, d-1\}$. To this end, we consider a special sequence, namely $a_k := \alpha^k$ for $\alpha \in (0,1)$. Let $(U_1, \ldots, U_d)'$ be defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and be distributed according to the joint distribution function $C_{1-\alpha}(u_1, \ldots, u_d) := \prod_{i=1}^{d} u_{(i)}^{\alpha^{i-1}} \in eMO$. Notice that $C_{1-\alpha}$ is an exchangeable Cuadras-Augé copula, see (4.6). Lemma 4.2.5 shows that in this case it holds that

$$\mathbb{P}(U_1 = \ldots = U_d) = \frac{(1-\alpha)^d}{1-\alpha^d} = \frac{(1-\alpha)^{d-1}}{\frac{1-\alpha^d}{1-\alpha}} = \frac{\sum_{i=0}^{d-1} \binom{d-1}{i} (-1)^i \, \alpha^i}{\sum_{i=0}^{d-1} \alpha^i} = \frac{\nabla^{d-1} a_0}{\sum_{i=0}^{d-1} \alpha^i}.$$

Combining this with the general formula (4.9) we obtain in this special case that

$$\frac{\nabla^{d-1}a_0}{\sum_{i=0}^{d-1}\alpha^i\,\kappa_i} = \mathbb{P}(U_1 = \ldots = U_d) = \frac{\nabla^{d-1}a_0}{\sum_{i=0}^{d-1}\alpha^i}.$$

In particular, it follows for all $\alpha \in (0,1)$ that

$$\sum_{i=0}^{d-1}\kappa_i\,\alpha^i = \sum_{i=0}^{d-1}\alpha^i.$$

The fundamental theorem of algebra now implies that $\kappa_i = 1$ for all $i \in \{0,1,\ldots,d-1\}$. Thus, the proof is finished. □

As a corollary to Theorem 4.4.5 it is possible to compute a probability involving Lévy subordinators which is not easy to obtain directly.

Corollary 4.4.7 (Application to Lévy Subordinators)
On a probability space $(\Omega,\mathcal{F},\mathbb{P})$, let Λ be an arbitrary Lévy subordinator with Laplace exponent Ψ, drift $\mu \geq 0$, and Lévy measure ν on $(0,\infty]$, excluding the degenerate case $\Lambda_t \equiv 0$. Let $d \geq 2$ and $E_1,\ldots,E_d$ be i.i.d. with $E_1 \sim Exp(1)$, independent of Λ. Then it holds that

$$\begin{aligned}
&\lim_{t\downarrow 0}\mathbb{P}\Big(\Lambda_t > \max\{E_1,\ldots,E_d\}\,\Big|\,\Lambda_t > \min\{E_1,\ldots,E_d\}\Big)\\
&\quad= \mathbb{P}\Big(\lim_{u\uparrow t}\Lambda_u < E_1,\ldots,E_d \leq \Lambda_t \text{ for some } t \geq 0\Big) = \frac{-1}{\Psi(d)}\sum_{i=0}^{d}\binom{d}{i}(-1)^i\,\Psi(i)\\
&\quad= \frac{\int_{(0,\infty]}\left(1-e^{-t}\right)^d\nu(dt)}{\mu\,d+\int_{(0,\infty]}\left(1-e^{-d\,t}\right)\nu(dt)} = \frac{\mathbb{E}\big[\nu\big([\max\{E_1,\ldots,E_d\},\infty]\big)\big]}{\mu\,d+\mathbb{E}\big[\nu\big([E_1/d,\infty]\big)\big]}.
\end{aligned}$$

Proof
Since Λ is not identically zero, $\Psi(1) > 0$, and we can rewrite

$$\bigcup_{t\geq 0}\Big\{\lim_{u\uparrow t}\Lambda_u < E_1,\ldots,E_d \leq \Lambda_t\Big\} = \bigcup_{t\geq 0}\Big\{\lim_{u\uparrow t}\Lambda_{\frac{u}{\Psi(1)}} < E_1,\ldots,E_d \leq \Lambda_{\frac{t}{\Psi(1)}}\Big\}.$$

Thus, the probability in concern agrees with

$$\mathbb{P}\Big(\lim_{u\uparrow t}\tilde{\Lambda}_u < E_1,\ldots,E_d \leq \tilde{\Lambda}_t \text{ for some } t \geq 0\Big),$$

where $\tilde{\Lambda} = \{\tilde{\Lambda}_t\}_{t\geq 0} = \{\Lambda_{t/\Psi(1)}\}_{t\geq 0}$ is a Lévy subordinator with Laplace exponent $\tilde{\Psi}(.) = \Psi(.)/\Psi(1)$ satisfying $\tilde{\Psi}(1) = 1$. Let $(\tilde{U}_1, \ldots, \tilde{U}_d)'$ be defined on $(\Omega, \mathcal{F}, \mathbb{P})$ having the d-dimensional Lévy-frailty copula $C_{\tilde{\Psi}}$ as joint distribution function. Theorem 4.4.5 with parameters $a_k := \tilde{\Psi}(k+1) - \tilde{\Psi}(k)$ for $k = 0, 1, \ldots, d-1$ proves that

$$\begin{aligned}
\mathbb{P}\Big(\lim_{u\uparrow t} \Lambda_u < E_1, \ldots, E_d \leq \Lambda_t \text{ for some } t \geq 0 \Big) &= \mathbb{P}(\tilde{U}_1 = \ldots = \tilde{U}_d) \\
&= \frac{\nabla^{d-1} a_0}{\sum_{i=0}^{d-1} a_i} = UEDC_{C_{\tilde{\Psi}}} = \lim_{u\uparrow 1} \mathbb{P}(\tilde{U}_{(1)} > u \,|\, \tilde{U}_{(d)} > u) \\
&= \lim_{t\downarrow 0} \mathbb{P}\Big(\Lambda_t > \max\{E_1, \ldots, E_d\} \,\Big|\, \Lambda_t > \min\{E_1, \ldots, E_d\} \Big).
\end{aligned}$$

Moreover, it is an easy exercise to compute that

$$\frac{\nabla^{d-1} a_0}{\sum_{i=0}^{d-1} a_i} = \frac{-1}{\Psi(d)} \sum_{i=0}^{d} \binom{d}{i} (-1)^i \, \Psi(i).$$

Finally, the last two claimed equations are established. From the Lévy-Khinchin representation we obtain

$$\begin{aligned}
\sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, \Psi(i) &= \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \left(\mu \, i + \int_{(0,\infty]} \left(1 - e^{-i\,t}\right) \nu(dt) \right) \\
&= \mu \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, i + \int_{(0,\infty]} \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} - \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, e^{-i\,t} \nu(dt) \\
&= \mu \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, i + \int_{(0,\infty]} (-1) \, (1-1)^d + \left(1 - e^{-t}\right)^d \nu(dt) \\
&= \mu \sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, i + \int_{(0,\infty]} \left(1 - e^{-t}\right)^d \nu(dt).
\end{aligned}$$

Moreover, we check that

$$\begin{aligned}
\sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \, i &= \sum_{i=1}^{d} \frac{d!}{i! \, (d-i)!} \, (-1)^{i+1} \, i = \sum_{i=1}^{d} \frac{(d-1)! \, d}{(i-1)! \, (d-1-(i-1))!} \, (-1)^{i+1} \\
&= d \sum_{i=1}^{d} \binom{d-1}{i-1} (-1)^{i+1} \overset{(j=i-1)}{=} d \sum_{j=0}^{d-1} \binom{d-1}{j} (-1)^j = d \, (1-1)^{d-1} = 0.
\end{aligned}$$

This yields the first equation. The second equation then follows from the following

observation:

$$1 - e^{-d\,t} = \mathbb{P}(E_1 \leq d\,t) = \mathbb{P}(E_1/d \leq t) = \mathbb{E}\big[\mathbf{1}_{\{E_1/d \leq t\}}\big],$$

which by Tonelli's theorem implies that

$$\int_{(0,\infty]} \big(1 - e^{-d\,t}\big)\,\nu(dt) = \mathbb{E}\Big[\int_{(0,\infty]} \mathbf{1}_{\{E_1/d \leq t\}}\,\nu(dt)\Big] = \mathbb{E}\big[\nu([E_1/d, \infty])\big].$$

Similarly,

$$\big(1 - e^{-t}\big)^d = \mathbb{P}(\max\{E_1, \ldots, E_d\} \leq t) = \mathbb{E}\big[\mathbf{1}_{\{\max\{E_1,\ldots,E_d\} \leq t\}}\big],$$

which implies (by exactly the same argument) that

$$\int_{(0,\infty]} \big(1 - e^{-t}\big)^d\,\nu(dt) = \mathbb{E}\big[\nu([\max\{E_1, \ldots, E_d\}, \infty])\big].$$

The claim is established. □

If one is willing to compute the probability of the event $\{\lim_{u\uparrow t} \Lambda_u < E_1, \ldots, E_d \leq \Lambda_t \text{ for some } t \geq 0\}$, the explicit formula in terms of the Laplace exponent is numerically challenging to evaluate for large $d \gg 2$. For this reason, the more convenient integral representation is included in the statement of Corollary 4.4.7. Furthermore, the last two equalities imply that the probability in concern decreases with increasing drift μ. This is quite intuitive since larger drift implies that the range $B := \{y \in [0,\infty) \,|\, \Lambda_t = y \text{ for some } t \geq 0\}$ of Λ becomes "larger". Hence, it becomes more likely that $E_1, \ldots, E_d \in B$, implying that they are not all identical. It is remarkable that a probability involving the whole path of a Lévy subordinator (up to the maximum of d independent exponential random variables) is the same as the limit for $t \downarrow 0$ of probabilities that only involve the path of the Lévy subordinator up to time $t > 0$.

Remark 4.4.8 (Link to Regenerative Composition Structures)
The Lévy-frailty construction is closely related to so-called *regenerative composition structures*, as introduced in [Gnedin, Pitman (2005)]. More precisely, let us assume the notation of Theorem 4.2.2 and apply the usual convention $\tau_{(1)} \leq \ldots \leq \tau_{(d)}$ for the ordered list of $\tau_1, \ldots, \tau_d$. Whereas the present thesis studies the distribution of $(\tau_1, \ldots, \tau_d)'$, the reference [Gnedin, Pitman (2005)] studies the distribution of $(P_1, \ldots, P_{K_d})'$, where

$K_d := |\{\tau_1, \ldots, \tau_d\}|$ denotes the number of different first passage times, and

$$\begin{aligned} P_1 &:= \max\{j \in \mathbb{N} \,|\, \tau_{(j)} = \tau_{(1)}\}, \\ P_2 &:= \max\{j \in \mathbb{N} \,|\, \tau_{(P_1+j)} = \tau_{(P_1)}\}, \\ &\vdots \\ P_{K_d} &:= \max\{j \in \mathbb{N} \,|\, \tau_{(P_{K_d-1}+j)} = \tau_{(P_{K_d-1})}\}. \end{aligned}$$

Notice that $P_1, \ldots, P_{K_d} \in \mathbb{N}$ and $P_1 + \ldots + P_{K_d} = d$, hence $(P_1, \ldots, P_{K_d})'$ defines a so-called *random partition* of d, and vast literature can be found on the study of such objects. For instance, it is shown in [Gnedin, Pitman (2005), Theorem 5.2] that

$$\begin{aligned} \mathbb{P}(\tau_{(1)} = \tau_{(m)} < \tau_{(m+1)}) &= \frac{\binom{d}{m} \sum_{i=0}^{m} \binom{m}{i} (-1)^{i+1} \Psi(d-m+i)}{\Psi(d)}, \quad m = 1, \ldots, d-1, \\ \mathbb{P}(\tau_{(1)} = \tau_{(d)}) &= \frac{\sum_{i=0}^{d} \binom{d}{i} (-1)^{i+1} \Psi(i)}{\Psi(d)}. \end{aligned} \tag{4.10}$$

The formulas in (4.10) extend a part of Corollary 4.4.7, which only considers the probability $\mathbb{P}(\tau_{(1)} = \tau_{(d)})$. However, an alternative proof of these formulas - in fact even a generalization to arbitrary eMO-copulas (not necessarily of Lévy-frailty kind) - can be derived in a similar manner as in the proof of Theorem 4.4.5. Since these are of minor importance, their derivation is left as an exercise for the interested reader.

5 Sampling Applications and Non-Exchangeable Structures

"What happens if a big asteroid hits Earth? Judging from realistic simulations involving a sledge hammer and a common laboratory frog, we can assume it will be pretty bad."

D. Barry, American writer and humorist.

Random number generation is crucial for many practical applications. If a probabilistic model is used to describe reality, then it is often important to be able to generate a huge number of scenarios of the model (on a standard PC). For the simulation of univariate random variables there exist quite general simulation algorithms, see e.g. the standard textbook [Devroye (1986)]. Often they are based on the transformation of a simulated standard uniformly distributed random variable $U \sim U[0,1]$. Simulation algorithms for the latter typically rely on algebraic derivations, for a prominent example see [Matsumoto, Nishimura (1998)]. Having such algorithms at hand, the simulation of independent random variables is possible. However, the simulation of random vectors requires one to simulate possibly dependent random variables. More precisely, consider a d-dimensional distribution function F with continuous margins $F_1, \ldots, F_d$ and copula C. To simulate a random vector with distribution function F, one may generate a random vector $(U_1, \ldots, U_d)'$, which has the copula C as joint distribution function, and then transform the components $U_1, \ldots, U_d$ to the desired marginals, setting[1] $\tau_i := F_i^{-1}(U_i)$, $i = 1, \ldots, d$. Simulating $(U_1, \ldots, U_d)'$ with joint distribution function C is referred to as *sampling the copula* C in the sequel. In the bivariate case, there exist sampling algorithms for quite arbitrary copulas, see e.g. [Nelsen (1999), p. 36] and the references therein. However, in larger dimensions, such general sampling strategies become expensive and the efficient sampling of copulas depends on convenient probabilistic constructions. E.g. Theorem 2.3.10 implies an efficient sampling algorithm for some Archimedean copulas.

[1] F_i^{-1} denotes the generalized inverse of F_i as in the proof of Theorem 2.2.5.

The content of this chapter is to outline how the class of Lévy-frailty copulas, and in particular Theorem 4.2.2, can be applied to sample Marshall-Olkin survival copulas in large dimensions. Recall that the original probabilistic construction of the Marshall-Olkin distribution requires one to evaluate a complicated function of $2^d - 1$ independent exponential random variables, see Subsection 2.3.2 or Algorithm 2 below. In large dimensions $d \gg 2$ this approach is inefficient, not to say impossible, on a standard PC. However, in one of the main results of this dissertation (Theorem 4.2.2) we have seen that for a specific subclass (Lévy-frailty copulas) there is an alternative probabilistic model available. This involves only d independent exponential threshold levels $E_1, \ldots, E_d$ and one path of an independent Lévy subordinator up to the first time the maximum $E_{(d)}$ of the threshold levels is exceeded. Lemma 5.1.1 below shows that $d \mapsto \mathbb{E}[E_{(d)}]$ grows like $d \mapsto \log d$, which is "slow". Consequently, the new sampling approach is efficient also in large dimensions $d \gg 2$. A generic sampling algorithm for a Lévy-frailty copula is given below.

Algorithm 1 (Generic Sampling Algorithm: Lévy-Frailty Copula)
To sample the copula $C_\Psi(u_1, \ldots, u_d)$ from Definition 4.2.1, the following steps may be exercised.

(0) Input: a Laplace exponent Ψ satisfying $\Psi(1) = 1$.

(1) Generate $E_1, \ldots, E_d$ i.i.d. with $E_1 \sim Exp(1)$.

(2) Find the maximum $E_{(d)} := \max\{E_1, \ldots, E_d\}$.

(3) Independently of the random variables $E_1, \ldots, E_d$, simulate one path of a Lévy subordinator $\Lambda = \{\Lambda_t\}_{t\in[0,T]}$ with Laplace exponent Ψ, where T is chosen large enough such that $\Lambda_T \geq E_{(d)}$.

(4) Determine the first passage times

$$\tau_i := \inf\big\{t \geq 0 \,:\, \Lambda_t \geq E_i\big\}, \quad i = 1, \ldots, d.$$

(5) Return $(U_1, \ldots, U_d)'$, where $U_i := \exp(-\tau_i)$, $i = 1, \ldots, d$.

For instance, in Section 5.1 Algorithm 1 is demonstrated exemplarily to efficiently sample exchangeable Cuadras-Augé copulas, as given by (4.6).

For some applications the symmetry of the Lévy-frailty copula C_Ψ in its arguments might be too restrictive. The fact that $C_\Psi(u_1, \ldots, u_d)$ is invariant under permutations of its arguments implies a degree of exchangeability of the corresponding random vector which sometimes is not justified in applications. In particular, it follows for $2 \leq k \leq d$ that all k-margins of $C_\Psi(u_1, \ldots, u_d)$ are given by $C_\Psi(u_1, \ldots, u_k)$, and hence are identical in structure. To overcome this drawback, Section 5.2 aims at enlarging the class of Lévy-frailty copulas to a family of copulas which allows for non-exchangeable dependence structures as well, without losing the advantage of efficient sampling.

5.1 Sampling Exchangeable Cuadras-Augé Copulas

This section is concerned with the specific copula $C_\alpha(u_1, \ldots, u_d) = u_{(1)} \prod_{i=2}^{d} u_{(i)}^{(1-\alpha)^{i-1}}$, for some prespecified parameter $\alpha \in [0, 1]$. As mentioned before, this one-parametric family of copulas coincides with the exchangeable special case of a distribution first discussed in [Cuadras, Augé (1981)]. In the case $d = 2$, [Ocana, Ruiz-Rivas (1990)] show how to sample a bivariate Cuadras-Augé copula.[2] Extending their investigations, we show how C_α can efficiently be sampled in arbitrary dimension $d > 2$. The effect of the parameter α on the dependence structure is illustrated in Figure 5.1, using Algorithm 3 below.

Before Algorithm 3 is presented, an alternative simulation algorithm, which is based on the original probabilistic model (2.11), is outlined. By virtue of Theorem 3.4.1, C_α is from the class eMO with parameters $\lambda_k := (1-\alpha)^{d-k}\, \alpha^{k-1}$, $k = 1, \ldots, d$. Hence applying (2.11), a random vector $(U_1, \ldots, U_d)'$ with joint distribution function C_α might be simulated by the following algorithm.

Algorithm 2 (Sampling of Exchangeable Cuadras-Augé Copulas)

(0) The input for the algorithm is the dimension $d \geq 2$ and the parameter $\alpha \in [0, 1]$. If $\alpha = 0$, $U_1, \ldots, U_d$ are independent and therefore straightforward to simulate. If $\alpha = 1$, $U_1, \ldots, U_d$ are completely comonotonic. Hence, one simulates $U_1 \sim U[0, 1]$ and sets $U_2 := \ldots := U_d := U_1$. We proceed with the interesting cases $\alpha \in (0, 1)$.

(1) Compute $\lambda_k := (1-\alpha)^{d-k}\, \alpha^{k-1}$, $k = 1, \ldots, d$.

[2] See also Subsection 2.2.1 for an alternative construction.

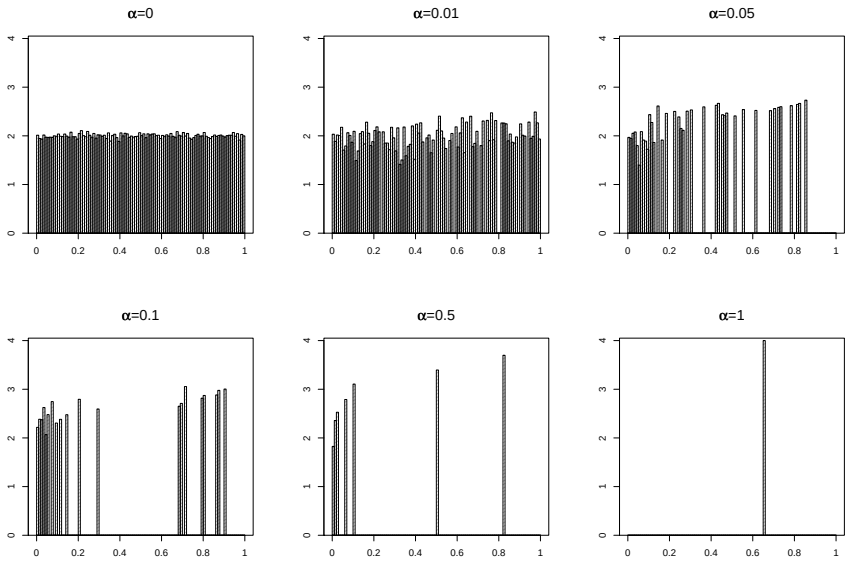


Figure 5.1 Histograms for six simulations of a random vector $(U_1, \ldots, U_d)'$ with joint distribution function $C_\alpha(u_1, \ldots, u_d)$, $d = 10\,000$, using different α's. The j-th of the 100 bars in each plot counts the number of U_i's falling into the interval $[(j-1)/100, j/100)$, which is displayed using a log-scale with base 10. It is observed that the clustering increases in α. For fixed α, moving from the left to the right end of the scale, the number of clusters decreases, whereas the size of clusters increases. This effect is explained by the fact that $\mathbb{E}[E_{(i)} - E_{(i-1)}] = 1/(d-i+1)$, $i = 2, \ldots, d$, see [David, Nagaraja (1970), p. 18].

(2) For each $k = 1, \ldots, d$, generate d choose k i.i.d. random variables $E_I \sim Exp(\lambda_k)$, $I \subset \{1, \ldots, d\}$ with $|I| = k$.

(3) Determine the minima

$$\tau_k := \min\{E_I \,|\, \emptyset \neq I \subset \{1, \ldots, d\},\, k \in I\}, \quad k = 1, \ldots, d.$$

(4) Return $(U_1, \ldots, U_d)'$, where $U_k := \exp(-\tau_k)$ for $k = 1, \ldots, d$.

It follows from Remark 3.4.2 that in the algorithm above $\tau_k \sim Exp(1)$, $k = 1 \ldots, d$, and $(\tau_1, \ldots, \tau_d)'$ has survival copula C_α. Consequently, the random vector $(U_1, \ldots, U_d)'$, defined as a transformation of $(\tau_1, \ldots, \tau_d)'$ in Step (4), has joint distribution function

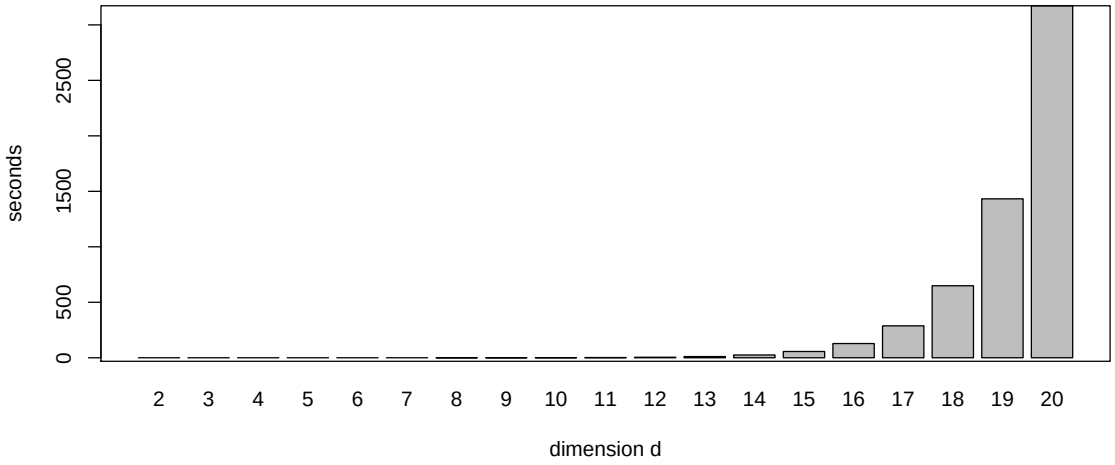


Figure 5.2 The runtimes for the simulation of one random vector $(U_1, \ldots, U_d)' \sim C_{1/2}$ are illustrated in dimension $d = 2, \ldots, 20$ using Algorithm 2, which is based on the orignial probabilistic construction (2.11).

C_α. If the dimension d is large, then in particular Step (3) is very expensive. But even Step (2) is expensive, since storage of the $2^d - 1$ samples E_I might exceed the computational resources, if d is large. This is verified by a brute force implementation of Algorithm 2 in R, which was carried out for $d = 2, \ldots, 20$. As expected, the algorithm runtime increases exponentially in the dimension d. On a Lenovo ThinkPad X200 Model 7458WB6-U (Intel Core 2 Duo CPU) with 2.39 GHz and 4GB RAM already for $d = 15$, the runtime for the simulation of one random vector is 56.63 seconds. For $d = 20$, it is already 3172.34 seconds, see Figure 5.2. The simulation of random vectors in dimension $d > 100$ is thus practically impossible using this approach. In comparison, the alternative Algorithm 3 below, which is based on Theorem 4.2.2, takes only fractions of a second on the same PC in dimensions $d \leq 20$. It requires 0.01 seconds for $d = 100$ and 0.61 seconds for $d = 10\,000$.

As outlined earlier, C_α is a Lévy-frailty copula corresponding to family (2) in Table 4.1, page 106. I.e. C_α may be constructed as the survival copula of the random vector $(\tau_1, \ldots, \tau_d)'$ of first passage times of a Lévy subordinator Λ across independent exponential threshold levels $E_1, \ldots, E_d$. In particular, for the interesting cases $\alpha \in (0, 1)$, Λ

is specified by $\Lambda_t := J\,N_t$, where $J := -\log(1-\alpha) > 0$ and $N = \{N_t\}_{t\geq 0}$ is a Poisson process with intensity $1/\alpha$. Since a Poisson process can be sampled efficiently without discretization bias, see e.g. [Cont, Tankov (2004), Algorithm 6.1, p. 174], the Lévy-frailty construction may be used to derive the following unbiased sampling algorithm.

Algorithm 3 (Improved Sampling of Exchangeable Cuadras-Augé Copulas)

(0) The input for the algorithm is the dimension $d \geq 2$ and the parameter $\alpha \in [0,1]$. If $\alpha = 0$, $U_1, \ldots, U_d$ are independent and therefore straightforward to simulate. If $\alpha = 1$, $U_1, \ldots, U_d$ are completely comonotonic. Hence, one simulates $U_1 \sim U[0,1]$ and sets $U_2 := \ldots := U_d := U_1$. We proceed with the interesting cases $\alpha \in (0,1)$.

(1) Specify the jump size of the process $J\,N$ as $J := -\log(1-\alpha)$.

(2) Generate $E_1, \ldots, E_d$ i.i.d. with $E_1 \sim Exp(1)$, and find the largest element $E_{(d)}$.

(3) Choose $K \in \mathbb{N}$ such that $J\,K \geq E_{(d)}$, e.g. choose K as the smallest integer greater than or equal to $E_{(d)}/J$.

(4) Simulate a path of N up to time T, where T is chosen such that $J\,N_T \geq E_{(d)}$. For this, generate K independent $Exp(1/\alpha)$-distributed random variables $Y_1, \ldots, Y_K$ and compute the j-th jump time T_j of N as $T_j = Y_1 + \ldots + Y_j$, for $j = 1, \ldots, K$.

(5) Obtain a sample of $(\tau_1, \ldots, \tau_d)'$ by setting

$$\tau_i := \sum_{j=1}^{K} \mathbf{1}_{\{(j-1)\,J < E_i \leq j\,J\}}\, T_j, \quad i = 1, \ldots, d.$$

(6) Return $(U_1, \ldots, U_d)'$, where $U_i := \exp(-\tau_i)$ for $i = 1, \ldots, d$.

Choosing J as in Step (1) results in the desired copula C_α. Since the random variables τ_i are constructed as first passage times of $\{J\,N_t\}_{t\geq 0}$ above the levels E_i, one has to generate the random variables E_i in Step (2) as well as a path of N up to time T, where $J\,N_T \geq E_{(d)}$. In Step (3) the number of required jumps of N to cross all E_i is computed and denoted by K. Therefore, one needs to generate a path of N up to its K-th jump time. A sample path of N is completely determined by its jump times $T_1 \leq T_2 \leq \ldots \leq T_K$. This simulation is accomplished in Step (4), where the construction idea (2.13) of a Poisson process is used: the first jump time is exponentially distributed with the Poisson intensity $1/\alpha$ as parameter. The Lévy properties of N guarantee that the increment of

the first to the second jump time is again an $Exp(1/\alpha)$-distributed random variable, and so on. Finally, defining the function $f(s) := \inf\{t \geq 0 : s \leq J\, N_t\}$ gives $f(E_i) = \tau_i$. It is clear that $f(E_i)$ equals the j-th jump time of N when $(j-1)\, J < E_i \leq j\, J$, which corresponds to Step (5).

The required computation time for sampling a d-dimensional copula with Algorithm 3 is discussed in the following. Steps (1) and (3) are computed in $\mathcal{O}(1)$. Finding the maximum of d numbers in Step (2) has computational effort $\mathcal{O}(d)$, since one simply has to consider all E_i and to compare each number with its predecessor. To estimate the expected runtime of Step (4) one needs to know the expected value of K. For this purpose, Lemma 5.1.1 below shows that $E_{(d)}$ has expected size in $\mathcal{O}(\log d)$. Thus, the expected value of K is in $\mathcal{O}(\log d)$, which therefore equals the expected runtime of Step (4). The next step requires the most computational resources, since for each $i = 1, \ldots, d$ on average one has to sum up K numbers. Since K has expected size in $\mathcal{O}(\log d)$, Step (5) has expected complexity $\mathcal{O}(d \log d)$. However, since only one term in the sum differs from zero, a binary tree argument can be used for each i to find the interval $\big((j-1)\cdot J,\, j \cdot J\big]$ containing E_i and speed up the summation this way. Since a bisection procedure for a set with K elements has computational efficiency $\mathcal{O}(\log K)$, see [Sedgewick (1988), p. 198 ff], one finally obtains computational efficiency $\mathcal{O}(d \log\log d)$ for Step (5). Since the runtime of Algorithm 3 has the same complexity as its slowest step, the complexity $\mathcal{O}(d \log\log d)$ is obtained as expected runtime for the algorithm. Compared to the complexity $\mathcal{O}\big(2^d\big)$ of Algorithm 2, this is a remarkable improvement.

Lemma 5.1.1 (Computational Efficiency)
It holds that $\mathbb{E}[E_{(d)}] \in \mathcal{O}(\log d)$, i.e $\lim_{d\to\infty} \mathbb{E}[E_{(d)}]/\log d = 1$.

Proof
The distribution function of $E_{(d)}$ is given by $\big(1-\exp(-t)\big)^d \mathbf{1}_{\{t>0\}}$. Hence, the distribution function of $E_{(d)} - \log d$ is given by $F^{[d]}(t) := \big(1-\exp(-t-\log d)\big)^d \mathbf{1}_{\{t>-\log d\}}$. It follows that $\lim_{d\to\infty} F^{[d]}(t) = \exp\big(-\exp(-t)\big)$, compare [David, Nagaraja (1970), Exercise 2.1.3, p. 22]. Therefore, independently considering for each $d \in \mathbb{N}$ the maximum $E_{(d)}$ of d i.i.d. $Exp(1)$-distributed random variables, the sequence $E_{(d)} - \log d$ converges in distribution to a random variable τ with distribution function $t \mapsto \exp\big(-\exp(-t)\big)$. The mean of τ is easily verified to exist, hence

$$\big|\mathbb{E}[E_{(d)}] - (\log d + \mathbb{E}[\tau])\big| \to 0, \quad d \to \infty.$$

This establishes the claim.

Alternatively, one can easily show that $\mathbb{E}[E_{(d+1)}] = \mathbb{E}[E_{(d)}] + 1/(d+1)$. Thus, $\mathbb{E}[E_{(d)}]$ is the harmonic series which is known to grow as $\log d + \gamma$, where $\gamma = 0.5772\ldots$ is the so-called Euler-Mascheroni constant, see e.g. [Rudin (1998), Exercise 9(a), p. 231]. In fact, $\mathbb{E}[\tau]$ from above is just a representation for this constant. □

5.2 Hierarchical Lévy-Frailty Copulas

We now introduce a hierarchical dependence structure by generalizing the family of Lévy-frailty copulas from Definition 4.2.1. Consider $M \in \mathbb{N}$ groups with $d_j \in \mathbb{N}$ members in group j. A random vector $(\tau_{1,1}, \ldots, \tau_{1,d_1}, \tau_{2,1}, \ldots, \tau_{M,d_M})' \in \mathbb{R}^{\sum_{j=1}^{M} d_j}$ is constructed such that:

1. Each $\tau_{j,i}$ is exponentially distributed with mean one.

2. For each group $j \in \{1, \ldots, M\}$ the random subvector $(\tau_{j,1}, \ldots, \tau_{j,d_j})'$ has a Lévy-frailty survival copula.

3. Two random variables from different groups are not necessarily independent. More precisely, for $j_1 \neq j_2$ and $(i_1, i_2) \in \{1, \ldots, d_{j_1}\} \times \{1, \ldots, d_{j_2}\}$ the bivariate random vector $(\tau_{j_1,i_1}, \tau_{j_2,i_2})'$ has a Cuadras-Augé survival copula.

The construction of such partitioned random vectors is motivated by applications. As we have seen in the previous chapter, Lévy-frailty copulas rely on a model which is exchangeable in the spirit of De Finetti's theorem: all components are equally affected by one common latent factor, a Lévy subordinator. However, in applications one often faces situations, where it is more realistic to assume that there are multiple latent factors. Considering for example a portfolio of d companies, it is natural to subdivide it into groups according to industrial branches. Given this partition, it might be intuitive to assume that the pharmaceutical branch is affected by different market factors than the financial branch. Nevertheless, there might be a global factor which affects all industrial branches, so the branches cannot be modeled independently of each other. Such motivations have led researchers to derive hierarchical extensions of Archimedean copulas, see e.g. [Joe (1997), McNeil (2008), Hering et al. (2010)]. On the contrary, Lévy-frailty copulas have a natural non-exchangeable extension, since they are actually defined as a subclass of (non-exchangeable) Marshall-Olkin survival copulas. However,

the previous paragraph reveals that the original Marshall-Olkin construction is not well-suited for simulations in large dimensions $d \gg 2$. To overcome this drawback, we rather extend the idea of Theorem 4.2.2, which is based on Lévy subordinators. The definition and respective sampling algorithm of this construction is carried out in Subsection 5.2.1 below. Subsection 5.2.2 finally shows that the constructed random vector is still of Marshall-Olkin kind, even though no longer exchangeable.

5.2.1 Construction and Sampling Algorithm

To construct a random vector with the properties postulated above, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which independent Lévy subordinators $\Lambda^{(0)}, \Lambda^{(1)}, \dots, \Lambda^{(M)}$ with corresponding Laplace exponents $\Psi_0, \Psi_1, \dots, \Psi_M$ are defined. All Laplace exponents $\Psi_0, \dots, \Psi_M$ are assumed to have 1 as a fixpoint. Moreover, let $\{E_{j,i} \,|\, j = 1, \dots, M,\, i = 1, \dots, d_j\}$ be i.i.d. with $Exp(1)$-distribution, all being independent of the Lévy subordinators. Finally, let $a \in [0,1]$ be an additional parameter. Then we define

$$\tau_{j,i} := \inf\big\{t \geq 0 \,:\, \Lambda^{(0)}_{a\,t} + \Lambda^{(j)}_{(1-a)\,t} \geq E_{j,i}\big\}, \quad j = 1, \dots, M,\, i = 1, \dots, d_j. \tag{5.1}$$

The main theoretical result about this random vector is Theorem 5.2.1 below.

Theorem 5.2.1 (Lévy-Frailty Copula with Hierarchical Structure)
The survival copula $\hat{C}$ of the random vector $(\tau_{1,1}, \dots, \tau_{1,d_1}, \tau_{2,1}, \dots, \tau_{M,d_M})'$ is given by

$$\hat{C}(\vec{u}) = C_{\Psi_0}(\vec{u})^a \cdot \Big(\prod_{j=1}^{M} C_{\Psi_j}(u_{j,1}, \dots, u_{j,d_j})\Big)^{1-a}, \tag{5.2}$$

where $\vec{u} = (u_{1,1}, \dots, u_{1,d_1}, \dots, u_{M,1}, \dots, u_{M,d_M})' \in [0,1]^{d_1+\dots+d_M}$. Moreover, each component $\tau_{j,i}$ is $Exp(1)$-distributed.

Proof
Let $(t_{1,1}, \dots, t_{1,d_1}, \dots, t_{M,d_M})' \in [0,\infty)^{\sum_{j=1}^{M} d_j}$. Then, using conditional independence in the second equality, one obtains

$$\begin{aligned}
&\mathbb{P}\big(\tau_{1,1} > t_{1,1}, \dots, \tau_{1,d_1} > t_{1,d_1}, \dots, \tau_{M,d_M} > t_{M,d_M}\big) \\
&\quad = \mathbb{P}\big(\Lambda^{(0)}_{a\,t_{1,1}} + \Lambda^{(1)}_{(1-a)\,t_{1,1}} < E_{1,1}, \dots, \Lambda^{(0)}_{a\,t_{1,d_1}} + \Lambda^{(1)}_{(1-a)\,t_{1,d_1}} < E_{1,d_1}, \\
&\qquad \dots, \Lambda^{(0)}_{a\,t_{M,d_M}} + \Lambda^{(M)}_{(1-a)\,t_{M,d_M}} < E_{M,d_M}\big)
\end{aligned}$$

$$= \mathbb{E}\Big[e^{-\Lambda^{(0)}_{a\,t_{1,1}} - \Lambda^{(1)}_{(1-a)\,t_{1,1}}} \cdots e^{-\Lambda^{(0)}_{a\,t_{1,d_1}} - \Lambda^{(1)}_{(1-a)\,t_{1,d_1}}} \cdots e^{-\Lambda^{(0)}_{a\,t_{M,d_M}} - \Lambda^{(M)}_{(1-a)\,t_{M,d_M}}}\Big]$$
$$= \mathbb{E}\Big[e^{-\sum_{j=1}^{M}\sum_{i=1}^{d_j}\Lambda^{(0)}_{a\,t_{j,i}}}\Big] \prod_{j=1}^{M} \mathbb{E}\Big[e^{-\sum_{i=1}^{d_j}\Lambda^{(j)}_{(1-a)\,t_{j,i}}}\Big].$$

In the proof of Theorem 4.2.2 it is shown that for arbitrary numbers $t_1, \ldots, t_d \geq 0$ with order statistics $0 := t_{(0)} \leq t_{(1)} \leq t_{(2)} \leq \ldots \leq t_{(d)}$ it follows for an arbitrary Lévy subordinator Λ with Laplace exponent Ψ having 1 as a fixpoint that

$$\mathbb{E}\Big[e^{-\sum_{i=1}^{d}\Lambda_{t_i}}\Big] = \prod_{i=1}^{d} e^{-(t_{(i)}-t_{(i-1)})\,\Psi(d+1-i)} = C_\Psi(e^{-t_1}, \ldots, e^{-t_d}). \tag{5.3}$$

Since furthermore $\exp(a\,t) = \exp(t)^a$ and Lévy-frailty copulas are extreme-value copulas, applying (5.3) to the previous computation implies

$$\mathbb{P}\big(\tau_{1,1} > t_{1,1}, \ldots, \tau_{1,d_1} > t_{1,d_1}, \ldots, \tau_{M,d_M} > t_{M,d_M}\big)$$
$$= \Big(C_{\Psi_0}(e^{-t_{1,1}}, \ldots, e^{-t_{1,d_1}}, \ldots, e^{-t_{M,d_M}})\Big)^a \Big(\prod_{j=1}^{M} C_{\Psi_j}(e^{-t_{j,1}}, \ldots, e^{-t_{j,d_j}})\Big)^{1-a}.$$

Hence, the claim is established by proving that each $\tau_{j,i}$ is $Exp(1)$-distributed. To this end, observe for $t \geq 0$ that

$$\mathbb{P}\big(\tau_{j,i} > t\big) = \mathbb{P}\big(\Lambda^{(0)}_{a\,t} + \Lambda^{(j)}_{(1-a)\,t} \leq E_{j,i}\big) = \mathbb{E}\Big[e^{-\Lambda^{(0)}_{a\,t} - \Lambda^{(j)}_{(1-a)\,t}}\Big]$$
$$= \mathbb{E}\Big[e^{-\Lambda^{(0)}_{a\,t}}\Big]\,\mathbb{E}\Big[e^{-\Lambda^{(j)}_{(1-a)\,t}}\Big] = e^{-a\,t\,\Psi_0(1)}\,e^{-(1-a)\,t\,\Psi_j(1)} = e^{-t}.$$

The last equality follows from the fact that 1 is a fixpoint of all involved Laplace exponents. □

For a specific group $j \in \{1, \ldots, M\}$ the subvector $(\tau_{j,1}, \ldots, \tau_{j,d_j})'$ has the survival copula $C_{a\,\Psi_0+(1-a)\,\Psi_j}(u_{j,1}, \ldots, u_{j,d_j})$, which is again a Lévy-frailty copula (corresponding to the Lévy subordinator $\{\Lambda^{(0)}_{a\,t} + \Lambda^{(j)}_{(1-a)\,t}\}_{t\geq 0}$ with Laplace exponent $a\,\Psi_0 + (1-a)\,\Psi_j$). Hence, the copula in Theorem 5.2.1 may be considered as a hierarchical Lévy-frailty copula.

Corollary 5.2.2 (Bivariate Cuadras-Augé Coefficients)
Bivariate subpairs of $(\tau_{1,1}, \ldots, \tau_{1,d_1}, \tau_{2,1}, \ldots, \tau_{M,d_M})'$ have a Cuadras-Augé survival copula[3] as dependence structure. The upper-tail dependence parameter $\alpha \in [0, 1]$ of the

[3] See Definition 2.2.2.

corresponding bivariate Cuadras-Augé copula C_α is determined as follows[4]:

1. **Inter sector:** For a pair $(\tau_{j_1,i_1}, \tau_{j_2,i_2})'$ taken from different groups $j_1 \neq j_2$, it is given by $\alpha^{inter} := a\big(2 - \Psi_0(2)\big)$.

2. **Intra sector:** For a pair $(\tau_{j,i_1}, \tau_{j,i_2})'$ taken from the same group j, it is given by $\alpha_j^{intra} := a\big(2 - \Psi_0(2)\big) + (1-a)\big(2 - \Psi_j(2)\big) = \alpha^{inter} + (1-a)\big(2 - \Psi_j(2)\big)$.

The intra sector dependence is always stronger than (or, in the special cases $a = 1$ or $\Psi_j(2) = 2$, equal to) the inter sector dependence. This is quite intuitive from the construction, since components of the same group j are similarly affected by $\Lambda^{(0)}$ and $\Lambda^{(j)}$; whereas components from different groups $i \neq j$ are only dependent through $\Lambda^{(0)}$.

Proof

First consider a pair $(\tau_{j_1,i_1}, \tau_{j_2,i_2})'$ from different groups $j_1 \neq j_2$, w.l.o.g. $j_1 < j_2$. Its survival copula C is obtained from the overall copula (5.2) by plugging in the special vector $\vec{u} = (1, \dots, 1, u_{j_1,i_1}, 1, \dots, 1, u_{j_2,i_2}, 1, \dots, 1)'$. It follows that

$$\begin{aligned}
C(u_{j_1,i_1}, u_{j_2,i_2}) &= \hat{C}(1, \dots, 1, u_{j_1,i_1}, 1, \dots, 1, u_{j_2,i_2}, 1, \dots, 1) \\
&= \big(\min\{u_{j_1,i_1}, u_{j_2,i_2}\} \max\{u_{j_1,i_1}, u_{j_2,i_2}\}^{\Psi_0(2)-1}\big)^a \, u_{j_1,i_1}^{1-a} \, u_{j_2,i_2}^{1-a} \\
&= \min\{u_{j_1,i_1}, u_{j_2,i_2}\} \max\{u_{j_1,i_1}, u_{j_2,i_2}\}^{a(\Psi_0(2)-1)+1-a} \\
&= \min\{u_{j_1,i_1}, u_{j_2,i_2}\} \max\{u_{j_1,i_1}, u_{j_2,i_2}\}^{1-a\,(2-\Psi_0(2))} \\
&= C_{a\,(2-\Psi_0(2))}(u_{j_1,i_1}, u_{j_2,i_2}).
\end{aligned}$$

This proves the first claim. The proof of the second claim works similarly: let now $(\tau_{j,i_1}, \tau_{j,i_2})'$ be a pair from the same group j, w.l.o.g. $i_1 < i_2$. Then its survival copula C is given by

$$\begin{aligned}
C(u_{j,i_1}, u_{j,i_2}) &= \hat{C}(1, \dots, 1, u_{j,i_1}, 1, \dots, 1, u_{j,i_2}, 1, \dots, 1) \\
&= \big(\min\{u_{j,i_1}, u_{j,i_2}\} \max\{u_{j,i_1}, u_{j,i_2}\}^{\Psi_0(2)-1}\big)^a \\
&\qquad \times \big(\min\{u_{j,i_1}, u_{j,i_2}\} \max\{u_{j,i_1}, u_{j,i_2}\}^{\Psi_j(2)-1}\big)^{1-a} \\
&= \min\{u_{j,i_1}, u_{j,i_2}\} \max\{u_{j,i_1}, u_{j,i_2}\}^{a\,(\Psi_0(2)-1)+(1-a)\,(\Psi_j(2)-1)} \\
&= \min\{u_{j,i_1}, u_{j,i_2}\} \max\{u_{j,i_1}, u_{j,i_2}\}^{1-\left(a\,(2-\Psi_0(2))+(1-a)(2-\Psi_j(2))\right)} \\
&= C_{a\,(2-\Psi_0(2))+(1-a)(2-\Psi_j(2))}(u_{j,i_1}, u_{j,i_2}).
\end{aligned}$$

[4] A random vector $(\tau_1, \tau_2)'$ with survival copula C_α has lower-tail dependence parameter α, whereas the copula C_α has upper-tail dependence α, compare Remark 2.2.9.

The claim is established. □

Remark 5.2.3 (Deeper Hierarchical Structures)
Deeper hierarchical structures can similarly be constructed by introducing more Lévy subordinators: instead of considering the convex combination of only two Laplace exponents, one may as well consider a convex combination of $m > 2$ Laplace exponents. Since for the presentation of the general idea and most applications one hierarchy is sufficient, the immediate extension to more levels is left to the reader.

Example 5.2.4 (Hierarchical Cuadras-Augé Copulas)
Suppose that all involved Lévy subordinators $\Lambda^{(0)}, \Lambda^{(1)}, \ldots, \Lambda^{(M)}$ are of compound Poisson type with zero drift and constant jump sizes $J_0, J_1, \ldots, J_M > 0$. This means that $\Lambda_t^{(l)} = J_l N_t^{(l)}$, where $N^{(l)} = \{N_t^{(l)}\}_{t\geq 0}$ are Poisson processes with intensities β_l, $l = 0, \ldots, M$. In order to have one as a fixpoint of the Laplace exponents $\Psi_l(x) = \beta_l\,(1 - \exp(-x\,J_l))$, $x \geq 0$, the jump intensities β_l of the Poisson processes are forced to be given by $\beta_l := 1/(1 - \exp(-J_l))$, $l = 0, \ldots, M$. Thus, this multivariate distribution is fully determined by the parameter $a \in [0, 1]$ and the jump sizes $J_0, J_1, \ldots, J_M > 0$. The corresponding dependence coefficients $0 \leq \alpha^{inter} \leq \alpha_j^{intra} \leq 1$, for $j = 1, \ldots, M$, are given as follows:

- $\alpha^{inter} = a\,\big(1 - \exp(-J_0)\big)$.
- $\alpha_j^{intra} = \alpha^{inter} + (1 - a)\,\big(1 - \exp(-J_j)\big)$, for $j = 1, \ldots, M$.

Figure 5.3 illustrates this distribution by means of pairwise scatterplots in a nine-dimensional example with three groups; Algorithm 4 below is used to sample the respective copula. Note in particular that this distribution can be viewed as an extension of exchangeable Cuadras-Augé copulas as investigated in Section 5.1.

An explicit sampling algorithm for the copula defined in (5.2), when all involved Lévy subordinators are chosen as compound Poisson processes with drift, is presented below. Recall that a compound Poisson process with drift is determined by its drift $\mu \geq 0$, an intensity parameter $\beta > 0$, and a jump-size distribution function G on $(0, \infty)$. Its Laplace exponent is given by $\Psi(x) = \mu\,x + \beta\,(1 - \mathbb{E}[\exp(-x\,J)])$, $x \geq 0$, where J is a random variable with distribution function G. Since compound Poisson processes have finitely many jumps on bounded intervals, the resulting sampling algorithm is unbiased. Algorithm 4 is first presented and then discussed.

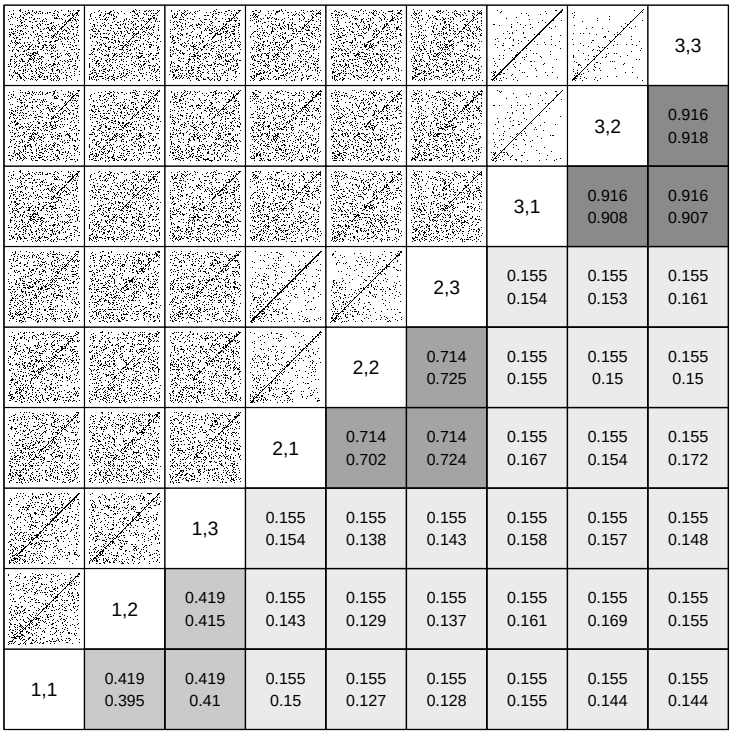


Figure 5.3 This figure illustrates 1 000 samples (by means of pairwise scatterplots) of a nine-dimensional random vector as defined in Example 5.2.4. The components are partitioned into $M = 3$ groups with dimensions $d_1 = d_2 = d_3 = 3$. The parameters are $a = 0.2$, $J_0 = 1.5$, $J_1 = 0.4$, $J_2 = 1.2$, and $J_3 = 3$. The two numbers in each panel of the diagonal denote the subindices j, i of $\tau_{j,i}$. Above the diagonal, the panel in row l and column k illustrates the scatterplot corresponding to the bivariate subvector with the indices which are given in the diagonal in rows l and $9 - k + 1$, respectively. The coefficients $\alpha^{inter}, \alpha_1^{intra}, \alpha_2^{intra}, \alpha_3^{intra}$ are given below the diagonal, depending on whether the two corresponding random variables are in the same or in different groups. The numbers in row l and column k correspond to the samples in row k and column l. The upper value is the theoretical one, whereas the lower one gives the empirical value for the specific panel based on the 1 000 samples. For this estimation the maximum likelihood estimator (2.2) is used. The figure illustrates the several levels of dependence between and within groups.

Algorithm 4 (Sampling of Hierarchical Lévy-Frailty Copulas)

(0) Input: the number of groups $M \in \mathbb{N}$, group sizes $d_1, \ldots, d_M \in \mathbb{N}$, $a \in (0, 1)$

($a \in \{0,1\}$ is trivial[5]), intensities $\beta_0, \beta_1, \ldots, \beta_M > 0$, and jump size distribution functions $G_0, G_1, \ldots, G_M$, subject to the restrictions $\beta_l \left(1 - \mathbb{E}\big[\exp(-J_l)\big]\right) \leq 1$, for $l = 0, \ldots, M$.

(1) Compute the drifts $\mu_l := 1 - \beta_l \left(1 - \mathbb{E}\big[\exp(-J_l)\big]\right)$ for $l = 0, 1, \ldots, M$.

(2) For $k = 0, 1, 2, \ldots$ and each $l \in \{0, 1, \ldots, M\}$ define the following dynamic object. Here, $simExp(\beta)$ denotes a function which generates and returns an exponential random variable (independent of the past) with mean $1/\beta$; $\Delta^{(l)}[.]$ denotes an array, respectively list object.

FUNCTION $\Delta^{(l)}(k)$
  IF $\big(k == 0\big)$ THEN $\Delta^{(l)}[k] := 0$
  ELSE DO
    IF $\big(\Delta^{(l)}[k]$ is undefined$\big)$ THEN $\Delta^{(l)}[k] := \Delta^{(l)}[k-1] + simExp(\beta_l)$
  RETURN $\Delta^{(l)}[k]$

For $k = 1, 2, \ldots$ and each $l = 0, 1, \ldots, M$ define the following dynamic object. $J_l[.]$ denotes an array, respectively list object, $simJ_l()$ denotes a function which generates and returns a sample from the distribution function G_l.

FUNCTION $J_l(k)$
  IF $\big(J_l[k]$ is undefined$\big)$ THEN $J_l[k] := simJ_l()$
  RETURN $J_l[k]$

(3) Generate $\{E_{j,i} \,|\, j = 1, \ldots, M, i = 1, \ldots, d_j\}$ i.i.d. $Exp(1)$-distributed.

(4) For each group $j = 1, \ldots, M$ generate a path of the process $\tilde{\Lambda}^{(j)}$, defined via $\tilde{\Lambda}^{(j)}_t := \Lambda^{(0)}_{a\,t} + \Lambda^{(j)}_{(1-a)\,t}$, as follows: jump times, jump sizes, and values of these processes at the jump times determine the sample paths completely. Hence, it is enough to store them in three lists:

  (a) Determine $E_{j:(d_j)} := \max\{E_{j,i} \,|\, i = 1, \ldots, d_j\}$.

  (b) Set up three lists $\tilde{\Delta}^{(j)}[.]$, $\tilde{\Lambda}^{(j)}[.]$, and $JS^{(j)}[.]$ with the following contents:

[5] $a = 0$ corresponds to independent groups, and $a = 1$ corresponds to an ordinary (exchangeable) Lévy-frailty copula.

- $\tilde{\Delta}^{(j)}[s]$ stores the s-th jump time of the process $\tilde{\Lambda}^{(j)}$.
- $\tilde{\Lambda}^{(j)}[s]$ stores the value of $\tilde{\Lambda}^{(j)}$ right after the s-th jump, i.e $\tilde{\Lambda}^{(j)}_{\tilde{\Delta}^{(j)}[s]}$.
- $JS^{(j)}[s]$ stores the s-th jump size, i.e. $\tilde{\Lambda}^{(j)}_{\tilde{\Delta}^{(j)}[s]} - \lim_{t \nearrow \tilde{\Delta}^{(j)}[s]} \tilde{\Lambda}^{(j)}_t$.

Set the initializing entries for $s = 0$ to zero in all three lists.

(c) Initialize the following values.

- $k = 1$, where k counts the jumps of the process $\Lambda^{(0)}$.
- $h = 1$, where h counts the jumps of the process $\Lambda^{(j)}$.
- $\tilde{\mu}_j = a\,\mu_0 + (1-a)\,\mu_j$, which is the drift of the process $\tilde{\Lambda}^{(j)}$.

(d) Run the following while-loop.

WHILE $\left(E_{j:(d_j)} > \tilde{\Lambda}^{(j)}[k+h-2]\right)$ DO
  $\tilde{\Delta}^{(j)}[k+h-1] := \min\{\Delta^{(0)}(k)/a,\, \Delta^{(j)}(h)/(1-a)\}$
  IF $\left(\Delta^{(0)}(k)/a < \Delta^{(j)}(h)/(1-a)\right)$ THEN
    $\tilde{\Lambda}^{(j)}[k+h-1] := \tilde{\Lambda}^{(j)}[k+h-2] + J_0(k) + \tilde{\Delta}^{(j)}[k+h-1]\,\tilde{\mu}_j$
    $JS^{(j)}[k+h-1] := J_0(k)$
    $k := k+1$
  ELSE DO
    $\tilde{\Lambda}^{(j)}[k+h-1] := \tilde{\Lambda}^{(j)}[k+h-2] + J_j(h) + \tilde{\Delta}^{(j)}[k+h-1]\,\tilde{\mu}_j$
    $JS^{(j)}[k+h-1] := J_j(h)$
    $h := h+1$

(5) For each $j = 1, \ldots, M$ and each $i = 1, \ldots, d_j$ determine the first passage time $\tau_{j,i}$ of $\tilde{\Lambda}^{(j)}$ across $E_{j,i}$ as follows:

(a) Find s such that $\tilde{\Lambda}^{(j)}[s-1] < E_{j,i} \leq \tilde{\Lambda}^{(j)}[s]$.

(b) Return the sample of $\tau_{j,i}$ by distinguishing the following cases:

$$\text{IF } \left(\tilde{\Lambda}^{(j)}[s] - JS^{(j)}[s] \leq E_{j,i}\right) \text{ THEN RETURN } \tau_{j,i} := \tilde{\Delta}^{(j)}[s]$$
$$\text{ELSE RETURN } \tau_{j,i} := \tilde{\Delta}^{(j)}[s-1] + \left(E_{j,i} - \tilde{\Lambda}^{(j)}[s-1]\right)/\tilde{\mu_j}$$

(6) Return the standardized uniform marginals $U_{j,i} := \exp(-\tau_{j,i})$.

Algorithm 4 is briefly discussed in the sequel. First, in Step (0) the general setup for the copula needs to be specified. This includes the sector segmentation and group dimensions, which specify the overall dimension of the copula. The construction of equation (5.1) is based on $M+1$ Lévy subordinators; one for each group, and another one affecting all components simultaneously. The Lévy subordinators are specified as compound Poisson processes with drift, which implies that the jump size distributions and intensities fully specify their distributional properties. The drift of each Lévy subordinator in Step (1) is chosen such that the technical requirement $\Psi_l(1) = 1$ holds for the Laplace exponent Ψ_l of the Lévy subordinator $\Lambda^{(l)}$, $l = 0, \ldots, M$. Step (2) of Algorithm 4 provides functions which generate and store the jump times of all Lévy subordinators. It is proposed to implement this step using dynamic lists that are recursively defined, and created when first called. More precisely, when first called with argument k, the function $\Delta^{(l)}(k)$ generates a new exponential random variable $simExp(\beta_l)$ (the time between jump $k-1$ and jump k of Lévy subordinator l) and appends $\Delta^{(l)}(k-1) + simExp(\beta_l)$ at position k of the dynamic list. If $\Delta^{(l)}(k)$ was called before, the previously stored value is returned. Similar dynamic objects are defined, i.e. $J_l(k)$, to generate and store the jump sizes of each Lévy subordinator. Step (3) generates the independent exponentially distributed trigger variables of each component. For each sector j, Step (4) combines the global Lévy subordinator $\Lambda^{(0)}$ with the sector Lévy subordinator $\Lambda^{(j)}$ to the new sector Lévy subordinator $\tilde{\Lambda}^{(j)}$. In order to avoid unnecessary computations, this iterative construction is stopped when the largest trigger in the respective sector is exceeded. The sample path of $\tilde{\Lambda}^{(j)}$ is completely determined by its jump times, jump sizes, and the systematic drift. In order to facilitate the computation of the $\tau_{j,i}$ we additionally store the values of $\tilde{\Lambda}^{(j)}$ right after each jump. The while-loop in Step (4) (d) iteratively adds jumps and drift of $\Lambda^{(0)}$ or $\Lambda^{(j)}$, depending on which new jump time comes next. Finally, in Step (5) one first determines between which jump times each trigger $E_{j,i}$ lies. Then, one distinguishes whether or not the trigger is exceeded by $\tilde{\Lambda}^{(j)}$ through a jump or by drift; in the latter case $\tau_{j,i}$ is adjusted accordingly.

Regarding the speed of Algorithm 4, a similar argument as in the previous paragraph guarantees that the expected complexity is in $\mathcal{O}(d \log d)$, where $d = d_1 + \ldots + d_M$. In particular, the algorithm can be run in dimensions $d \gg 2$. It thus provides a new possibility to simulate flexible hierarchical Marshall-Olkin structures in large dimensions. Finally, since arbitrary Lévy subordinators can be approximated arbitrarily close by compound Poisson subordinators, see e.g. [Damien et al. (1995)][6], Algorithm 4 also provides an approximate sampling strategy for the general case of Theorem 5.2.1.

5.2.2 Hierarchical Marshall-Olkin Distribution

Let us briefly review what has been done so far: the general Marshall-Olkin distribution was introduced in Chapter 2, Subsection 2.3.2. In Chapter 3 the subclass of exchangeable Marshall-Olkin distributions was determined. And furthermore in Chapter 4 the "subsubclass" of exchangeable and extendible Marshall-Olkin distributions was found. For the latter, we could derive an alternative probabilistic model using Lévy subordinators. The present chapter showed that this alternative construction is extremely convenient for sampling in large dimensions. In particular, it was argued that it is superior to the original model in this regard. Now starting from the convenient Lévy subordinator representation of the "subsubclass", we constructed hierarchical Lévy-frailty copulas to obtain a larger and more flexible family of copulas. A natural question is therefore: is this hierarchical dependence structure again of Marshall-Olkin kind? This short paragraph shows that this is indeed the case.

To verify this, we show that the random vector defined in (5.1) satisfies the lack of memory property (2.9), which characterizes the Marshall-Olkin distribution.

Lemma 5.2.5 (Hierarchical Marshall-Olkin Distribution)
We consider the random vector $(\tau_{1,1}, \ldots, \tau_{1,d_1}, \ldots, \tau_{M,d_M})'$ defined in (5.1) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote the dimension of this random vector by $d := d_1 + \ldots + d_M$. Then this random vector has a (d-dimensional) Marshall-Olkin distribution. Put differently, its survival copula (5.2) is a Marshall-Olkin survival copula.

Proof
It is an easy exercise to check for arbitrary $n \in \mathbb{N}$ (with the convention $C_\Psi(u) = u$ for

[6] An alternative approximate sampling strategy for general Lévy measures can also be found in [Bondesson (1982)].

$n = 1$), Laplace exponent Ψ with $\Psi(1) = 1$, and $t_1, \ldots, t_n, h \geq 0$ that

$$C_\Psi\big(e^{-t_1-h}, \ldots, e^{-t_n-h}\big) = C_\Psi\big(e^{-t_1}, \ldots, e^{-t_n}\big)\, C_\Psi\big(e^{-h}, \ldots, e^{-h}\big).$$

This property obviously carries over to the survival copula (5.2). This implies for $t_{1,1}, \ldots, t_{M,d_M}, h \geq 0$ that

$$\begin{aligned}
&\mathbb{P}\big(\tau_{1,1} > t_{1,1} + h, \ldots, \tau_{M,d_M} > t_{M,d_M} + h\big) \\
&\quad = \hat{C}\big(e^{-t_{1,1}-h}, \ldots, e^{-t_{M,d_M}-h}\big) \\
&\quad = \hat{C}\big(e^{-t_{1,1}}, \ldots, e^{-t_{M,d_M}}\big)\, \hat{C}\big(e^{-h}, \ldots, e^{-h}\big) \\
&\quad = \mathbb{P}\big(\tau_{1,1} > t_{1,1}, \ldots, \tau_{M,d_M} > t_{M,d_M}\big)\, \mathbb{P}\big(\tau_{1,1} > h, \ldots, \tau_{M,d_M} > h\big).
\end{aligned}$$

Put differently, this shows that the random vector $(\tau_{1,1}, \ldots, \tau_{1,d_1}, \ldots, \tau_{M,d_M})'$ satisfies the multivariate lack of memory property (2.9). Since subvectors are identical in structure, it is clear that all subvectors also satisfy the multivariate lack of memory property. Since [Marshall, Olkin (1967)] show that this property characterizes the Marshall-Olkin distribution, the claim is thus established. □

Concluding, Lemma 5.2.5 shows that hierarchical (and therefore non-exchangeable) Marshall-Olkin distributions can also be constructed using Lévy subordinators. Instead of one latent Lévy subordinator, dependence is introduced by one global and several group-specific Lévy subordinators. As outlined earlier, such latent factor representations are not only intuitive, they also provide efficient sampling routines. One open question, and a subject of further research, is to determine precisely which Marshall-Olkin distributions are obtained, i.e. what is the relation between the new parameters given in terms of the Laplace exponents $\Psi_0, \ldots, \Psi_M$ and the original parameters of the Marshall-Olkin distribution.

6 The Lévy-Frailty Default Model

"The purpose of models is not to fit the data but to sharpen the questions."
S. Karlin, April 1983.

This chapter shows how to use a multivariate distribution of Marshall-Olkin type for the pricing of portfolio credit derivatives. More precisely, the vector of default times of d defaultable firms is assumed to have a Lévy-frailty survival copula, see Definition 4.2.1. An outline of the chapter is as follows. Section 6.1 introduces notions and products which are related to portfolio credit risk. Section 6.2 gives an overview over existing portfolio credit risk models and indicates how the present approach sets itself apart from these references. Section 6.3 presents the modeling approach. Section 6.4 computes quantities which are required for pricing applications, Section 6.5 presents a calibration to CDO market quotes, and Section 6.6 compares the presented approach with existing portfolio credit models.

6.1 Portfolio Credit Derivatives

A *bond* is a debt security which is traded on fixed income markets. It works like a loan: the issuer of the bond is the borrower, the bond holder is the lender. As compensation for granting a loan to the issuer, the bond holder receives periodic premium payments, called *coupons*, which may be viewed as interest rate payments. One typically classifies bonds by means of the creditworthiness of the issuer. For example, a government bond such as a German *Bundesschatzbrief* is often considered to be default-free. In contrast, corporate bonds (i.e. the issuer is a company) or government bonds of emerging markets are considered to be subject to the possibility of credit default of the issuer. In this case, the bond is said to be *defaultable*. A *credit derivative* is a financial contract with the purpose to sell risks, arising from defaultable bonds and similar credit-risky assets, to another party. The most important example is a *Credit Default Swap (CDS)*, which is

a financial contract between a protection buyer and a protection seller. The protection buyer makes periodic premium payments - typically quarterly or semi-annually - to the protection seller. In return, the latter is committed to pay a default compensation if a pre-determined reference entity, e.g. a defaultable bond, defaults. In this regard, a CDS is similar to a life insurance contract, replacing a person's lifetime by the "lifetime" of a generic reference entity. Since the reference entity is typically a defaultable bond of a single company, a CDS is called a *single-name* credit derivative.

Contracts which have a whole portfolio of credit-risky assets as underlying are called *portfolio credit derivatives* or *basket credit derivatives*. Unlike single-name credit derivatives, these financial contracts depend on the creditworthiness of a portfolio of $d \in \{2, 3, \ldots\}$ entities. From a mathematical point of view, such contracts are quite challenging to handle, since empirical data typically suggest that the companies in the considered portfolio can not be treated independently. Based on default data published by Standard & Poor's[1], Figure 6.1 illustrates observed company defaults in the last 28 years. One can clearly see that company defaults exhibit clusters in time, i.e. in some years there are many defaults and in some years only a few or even none. This indicates that default events are jointly triggered by economic factors and should not be treated independently. Thus, a mathematical model for the default times of all d considered reference entities is necessary that allows for flexible dependence structures.

Throughout this chapter we always consider a portfolio of d defaultable reference entities, which – without loss of generality – we assume to be corresponding to d companies. In the sequel, the random vector $(\tau_1, \ldots, \tau_d)'$ is supposed to denote the default times of the d firms, i.e. τ_k is the random lifetime of company k. Moreover, for $k = 1, \ldots, d$ we denote by $D^{(k)} = \{D_t^{(k)}\}_{t \geq 0}$ the *default indicator process* of firm k, i.e. $D_t^{(k)} := \mathbf{1}_{\{\tau_k \leq t\}}$, $t \geq 0$. A basket credit derivative is a financial contract whose payment streams depend on the vector of default times corresponding to the reference portfolio. The payments may additionally depend on the *losses given default*, i.e. the fraction of money bond holders get back once a firm has defaulted. This fraction of the principal is usually referred to as *recovery rate*. For simplicity, we always assume that the recovery rates of all considered bond issuers are constant and identical. On a portfolio level, this assumption can sometimes be justified by an average argument. It allows to focus on a model for the

[1] "2008 Annual Global Corporate Default Study And Rating Transitions", published by Standard & Poor's, April 2, 2009. Retrievable from http://www2.standardandpoors.com/spf/pdf/fixedincome/corporate_default_study.pdf.

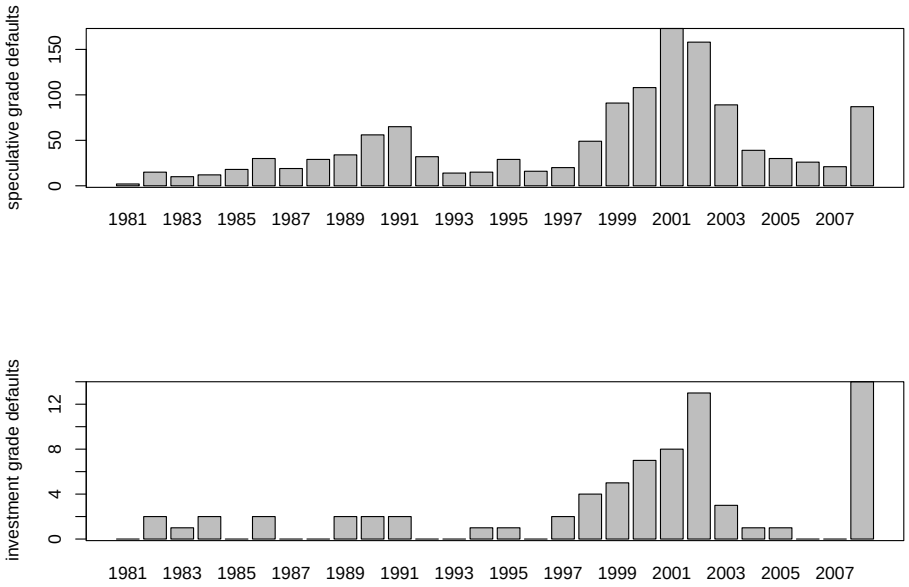


Figure 6.1 The upper graph illustrates speculative grade defaults from 1981 to 2008, i.e. defaults of companies that were rated BB or lower (according to Standard & Poor's) before default. The lower graph illustrates defaults of companies that were rated BBB or higher (according to Standard & Poor's).

default times (and neglect recovery rates), which reduces the mathematical difficulties. Typical examples for basket credit derivatives are *Portfolio Credit Default Swaps (portfolio CDS)*, *Collateralized Debt Obligations (CDOs)* and *n-th-to-default swaps*, which are introduced in the sequel. All are bilateral contracts between an insurance buyer and an insurance seller, and the payment streams depend on the default times corresponding to the reference portfolio. Pricing of the aforementioned contracts corresponds to determining the fair periodic premium payment, typically quoted in terms of a so-called *spread*, such that the expected discounted premium payments are equal to the expected discounted default payments. The idea of this approach is that, on average, the cash flows of insurance seller and insurance buyer are identical, so none of both parties makes a systematic gain. Formulated mathematically, we fix a payment schedule $\mathcal{T} : t_0 = 0 < t_1 < \ldots < t_M = T$ and assume an identical and constant recovery rate $R \in [0, 1]$ for all d companies. Moreover, suppose we are given discount factors $disc_{t_i}$

corresponding to the time points t_i, $i = 0, \ldots, M$. The *zero-recovery relative portfolio loss process* $L = \{L_t\}_{t \geq 0}$ is defined by

$$L_t := \frac{1}{d} \sum_{k=1}^{d} \mathbf{1}_{\{\tau_k \leq t\}} = \frac{1}{d} \sum_{k=1}^{d} D_t^{(k)}, \quad t \geq 0.$$

Thus, L_t takes values in the (finite) set $\{0, 1/d, 2/d, \ldots, 1\}$ and gives the share of defaulted firms in the portfolio up to time t. The remaining nominal of the portfolio at time t is given by $Nom_t = 1 - L_t$. Now let us describe the three aforementioned basket credit derivatives more detailed:

- **Portfolio CDS:** the insurance buyer makes periodic premium payments at the pre-specified time points $t_0, \ldots, t_M$ and, in return, receives default compensations from the insurance seller, resulting from losses in the underlying credit portfolio. Usually, the premium payments are quoted in terms of a so-called *portfolio CDS spread*, which is denoted by s^{CDS}. Based on the overall portfolio losses $(1 - R) L_{t_i}$ up to time points t_i, $i = 0, \ldots, M$, the expected discounted premium and default legs ($EDPL^{CDS}$ and $EDDL^{CDS}$) of the portfolio CDS are assumed to be (approximately) given by

$$EDPL^{CDS} = \mathbb{E}\left[\sum_{i=1}^{M} disc_{t_i}\, s^{CDS}\, \Delta t_i \left(Nom_{t_i} + \frac{Nom_{t_{i-1}} - Nom_{t_i}}{2}\right)\right],$$
$$EDDL^{CDS} = \mathbb{E}\left[\sum_{i=1}^{M} disc_{t_i}\, (1 - R)\, (L_{t_i} - L_{t_{i-1}})\right],$$

  where $\Delta t_i := t_i - t_{i-1}$. Accrued interest is (approximately) considered in the formula for $EDPL^{CDS}$ above by assuming companies to default at the midpoint of two payment dates. The corresponding fair spread s^{CDS} is obtained by equating the expected discounted premium and default legs. It is given by

$$s^{CDS} = \frac{\sum_{i=1}^{M} disc_{t_i}\, (1 - R)\, (\mathbb{E}[L_{t_i}] - \mathbb{E}[L_{t_{i-1}}])}{\sum_{i=1}^{M} disc_{t_i}\, \Delta t_i \left(\mathbb{E}[Nom_{t_i}] + \frac{\mathbb{E}[Nom_{t_{i-1}}] - \mathbb{E}[Nom_{t_i}]}{2}\right)}. \tag{6.1}$$

- **CDO:** CDOs are constructed by partitioning the reference credit portfolio in so-called *tranches* with different seniorities. A tranche of a CDO represents a certain loss piece of the overall portfolio which is defined via its lower (l) and upper

(u) attachment points. These attachment points define a partition of $[0,1]$ in the following manner: $0 = l_1 < u_1 = l_2 < u_2 = l_3 < \ldots < u_{K-1} = l_K < u_K = 1$, where we index the tranches by the numbers $j = 1, \ldots, K$. E.g. tranche 1, called the *equity tranche*, is the most subordinate loss piece, and thus bears the highest risk. The market standard for the partition of the credit portfolio is defined by different synthetic portfolios of the International Index Company, called iTraxx, and the portfolio managed by Dow Jones, called DJ CDX. Table 6.1 lists their respective tranches. The insurance seller, or investor in tranche $j \in \{1, \ldots, K\}$, receives

Table 6.1 The iTraxx and DJ CDX segmentation.

$d = 125$ companies		iTraxx		DJ CDX	
Tranche	j	l_j	u_j	l_j	u_j
Equity	1	0%	3%	0%	3%
Junior Mezzanine	2	3%	6%	3%	7%
Senior Mezzanine	3	6%	9%	7%	10%
Senior	4	9%	12%	10%	15%
Super Senior	5	12%	22%	15%	30%
	6	22%	100%	30%	100%

periodic premium payments at the pre-specified time points t_i, depending on the remaining nominal of tranche j. In return, the insurance buyer is compensated for losses affecting this tranche. For example, if 5% of the notional of the reference portfolio happen to default, an investor in tranche 2 with attachment points $l_2 = 3\%$, $u_2 = 6\%$ has to pay 2% of the notional to the insurance buyer. Let us note that a portfolio CDS may be considered equivalent to a CDO when there is only a single tranche with attachment points $l_1 = 0\%$, $u_1 = 100\%$. The annualized premium payments are called *tranche spreads* and are denoted by s^j, $j \in \{1, \ldots, K\}$, in the sequel. In mathematical terms, the loss affecting tranche j of a CDO up to time t is given by

$$L_{t,j} := \min\big\{ \max\{0, (1-R)\, L_t - l_j\}, u_j - l_j\big\}, \quad j = 1, \ldots, K. \tag{6.2}$$

The remaining nominal of tranche j is denoted by $Nom_{t,j} := u_j - l_j - L_{t,j}$, $j = 1, \ldots, K$. The corresponding expected discounted premium and default legs for tranches $j = 2, \ldots, K$ of a CDO ($EDPL^j$ and $EDDL^j$) are assumed to be given

by

$$EDPL^j = \mathbb{E}\left[\sum_{i=1}^{M} disc_{t_i}\, s^j\, \Delta t_i \left(Nom_{t_i,j} + \frac{Nom_{t_{i-1},j} - Nom_{t_i,j}}{2}\right)\right],$$
$$EDDL^j = \mathbb{E}\left[\sum_{i=1}^{M} disc_{t_i}\, (L_{t_i,j} - L_{t_{i-1},j})\right], \tag{6.3}$$

where again $\Delta t_i := t_i - t_{i-1}$ and accrued interest is considered in the formula for $EDPL^j$ by assuming defaults to occur at the midpoint of two payment dates. For tranches $2, \ldots, K$ the corresponding fair spread is obtained by solving the equation $EDPL^j = EDDL^j$ for s^j, hence

$$s^j = \frac{\sum_{i=1}^{M} disc_{t_i}\, (\mathbb{E}[L_{t_i,j}] - \mathbb{E}[L_{t_{i-1},j}])}{\sum_{i=1}^{M} disc_{t_i}\, \Delta t_i \left(\mathbb{E}[Nom_{t_i,j}] + \frac{\mathbb{E}[Nom_{t_{i-1},j}] - \mathbb{E}[Nom_{t_i,j}]}{2}\right)}, \quad j = 2, \ldots, K. \tag{6.4}$$

For the equity tranche of a CDO it is market standard to assume a running spread of 500 basis points, i.e. $s^1 := 0.05$. Therefore, an *upfront payment* (quoted as a percentage of the nominal of the first loss piece) is introduced to correct for this artificial spread. This upfront payment, denoted by *up* in the sequel, satisfies the relation

$$EDDL^1 = up\,(u_1 - l_1) + \mathbb{E}\left[\sum_{i=1}^{M} disc_{t_i}\, 0.05\, \Delta t_i \left(Nom_{t_i,1} + \frac{Nom_{t_{i-1},1} - Nom_{t_i,1}}{2}\right)\right], \tag{6.5}$$

where $EDDL^1$ is given as in (6.3) for $j = 1$. Accordingly, pricing of the equity tranche of a CDO corresponds to determining the upfront payment *up* such that equality (6.5) holds, i.e.

$$up = \frac{1}{u_1 - l_1}\Bigg(\sum_{i=1}^{M} disc_{t_i}\, (\mathbb{E}[L_{t_i,1}] - \mathbb{E}[L_{t_{i-1},1}]) - \sum_{i=1}^{M} disc_{t_i}\, 0.05\, \Delta t_i \left(\mathbb{E}[Nom_{t_i,1}] + \frac{\mathbb{E}[Nom_{t_{i-1},1}] - \mathbb{E}[Nom_{t_i,1}]}{2}\right)\Bigg). \tag{6.6}$$

- **n-th-to-default swap:** The insurance buyer pays an annualized premium $s^{(n)}$ to the insurance seller at the time points t_i, $i = 0, \ldots, M$. As soon as n assets of the underlying portfolio have defaulted, the contract ends. This means that the insurance buyer stops further premium payments and the insurance seller is committed to pay the insurance buyer the default compensation $1 - R$, where $R \in [0, 1]$ denotes the recovery rate as mentioned earlier. If less than n assets default up to maturity T of the contract, the contract also ends but no default compensation is payed. The annualized premium $s^{(n)}$ is called *n-th-to-default spread*. The expected discounted premium and default legs for an n-th-to-default swap ($EDPL^{(n)}$ and $EDDL^{(n)}$) are given by

$$EDPL^{(n)} = \mathbb{E}\Big[\sum_{i=1}^{M} disc_{t_i}\, s^{(n)}\, \Delta t_i\, \mathbf{1}_{\{\tau_{(n)} > t_i\}}\Big],$$
$$EDDL^{(n)} = \mathbb{E}\Big[\sum_{i=1}^{M} disc_{t_i}\, (1 - R)\, \mathbf{1}_{\{t_{i-1} < \tau_{(n)} \leq t_i\}}\Big],$$

  where again $\Delta t_i := t_i - t_{i-1}$ and $\tau_{(n)}$ denotes the n-th default time, i.e. the n-th smallest element of the set $\{\tau_1, \ldots, \tau_d\}$. Accordingly, the fair spread $s^{(n)}$ is obtained by equating $EDPL^{(n)}$ and $EDDL^{(n)}$. Hence,

$$s^{(n)} = \frac{\sum_{i=1}^{M} disc_{t_i}\, (1 - R)\, \mathbb{P}(t_{i-1} < \tau_{(n)} \leq t_i)}{\sum_{i=1}^{M} disc_{t_i}\, \Delta t_i\, \mathbb{P}(\tau_{(n)} > t_i)}. \tag{6.7}$$

Since we assume a constant recovery rate $R \in [0, 1]$, one observes from equation (6.1) that the pricing of a portfolio CDS requires one only to compute the expected value $\mathbb{E}[L_t]$ of the relative portfolio loss process, which is easily accomplished in most multivariate default models. More complicated is the pricing of the tranche of a CDO, which necessitates the computation of $\mathbb{E}[L_{t,j}]$, the expectation of a non-trivial function of the random variable L_t, compare (6.2) and formulas (6.4) and (6.6) above. Finally, for the pricing of an n-th-to-default swap it is necessary to know the distribution of the n-th default time $\tau_{(n)}$, see equation (6.7), which is mathematically demanding in general.

6.2 Literature Overview: Default Models

Considering financial contracts whose payment streams are functions of $(\tau_1, \ldots, \tau_d)'$, one has to set up a mathematical model for this random vector of default times. On a high level, every such model pursues two conflicting goals: on the one hand it has to be realistic and flexible, and on the other hand it has to be mathematically viable. By viability we mean analytical properties such as intuitive parameters and closed formulas, but also an ease of implementation for practical applications of the model. It is often the case that cutting back on realism allows to gain more viability, and conversely, more flexible and realistic models are often difficult to handle. From a practical point of view, the challenge is to build a model which can be implemented and is intuitive, but still flexible and realistic enough to be of use.

Vast literature can be found on different modeling approaches for the vector of default times $(\tau_1, \ldots, \tau_d)'$. The first model for the pricing of a defaultable bond is [Merton (1974)]. The standard Black-Scholes setup is considered and a company is defined to be defaulted at time T if its asset value at that time, modeled as a lognormal random variable, is less than or equal to a constant threshold level. This model is quite simplistic since only a single time point T is considered, and a default at earlier times $0 \leq t \leq T$ is not possible. In this regard, the model does not precisely model a default time τ_k, but only its default indicator $D_T^{(k)}$ at time T. A more realistic model is proposed by [Black, Cox (1976)], who define the default of a firm as the first hitting time of the firm's asset value, modeled by a (geometric) Brownian motion, below a constant liability threshold level, which introduces a dynamic aspect to the model of [Merton (1974)]. Both models originated the class of so-called *structural models*, which are maybe the most intuitive default models. Their idea is to define the default time τ_k as a time point when the aggregated asset value of company k falls below the aggregated liabilities of the same firm. In this regard, one has to define models for the evolution of assets and liabilities of all d firms, and then the vector of default times is a function of these stochastic processes. Closed-form solutions for prices in such an approach are difficult to obtain, unless one accepts several simplifying assumptions on the involved stochastic processes. However, properties such as normality or continuity of the Brownian motion in the model of [Black, Cox (1976)] imply unrealistic and inflexible models. In particular, the continuity of the paths of Brownian motion, together with a deterministic liability threshold level, imply that default times are predictable and

that immediate defaults are impossible. The latter property is problematic if the model is applied to the pricing of derivatives with short maturity. Therefore, extensions to random threshold levels or to more sophisticated firm-value processes are studied e.g. in [Zhou (2001), Giesecke (2004), Willemann (2007)].

In contrast to structural models, many recent approaches are of a *reduced-form* type. This means that the vector of default times $(\tau_1, \ldots, \tau_d)'$ is not defined as a function of more or less observable asset and liability values of the firms. Rather, it is directly defined via some multivariate distribution, whose choice is often justified by a good fit to market data. Two lines of reduced-form modeling are carried out in the sequel to demonstrate the most popular approaches: *intensity-based approaches* and *factor-copula models*.

A classical intensity-based model, also called *doubly-stochastic model*, defines the default time of firm k as

$$\tau_k := \inf\Big\{t \geq 0 \,:\, \int_0^t \lambda_s^{(k)}\, ds \geq E_k\Big\}, \tag{6.8}$$

where $E_1, \ldots, E_d$ are i.i.d. with $E_1 \sim Exp(1)$ and $\vec{\lambda} := (\lambda^{(1)}, \ldots, \lambda^{(d)})'$ is a vector of non-negative stochastic processes $\lambda^{(k)} = \{\lambda_t^{(k)}\}_{t\geq 0}$, which are possibly dependent. In the univariate case such an approach is introduced by [Jarrow, Turnbull (1995)] and essentially embossed by [Madan, Unal (1998), Lando (1998), Duffie, Singleton (1999)]. On a multivariate level doubly-stochastic approaches and extensions thereof are studied e.g. in the references [Duffie, Gârleanu (2001), Das et al. (2007), Duffie et al. (2007), Yu (2007)]. Conditioned on the σ-algebra $\sigma(\vec{\lambda}_u \,:\, u \geq 0)$, which is generated by the path of $\vec{\lambda}$, it holds that

$$\mathbb{P}\big(\tau_k > t \,\big|\, \sigma(\vec{\lambda}_u \,:\, u \geq 0)\big) = e^{-\int_0^t \lambda_s^{(k)}\, ds}, \quad t \geq 0. \tag{6.9}$$

Equivalently, one may consider τ_k as the first jump time of an inhomogeneous Poisson process with intensity process $\lambda^{(k)}$, see [Lando (1998)]. The unconditional survival function of τ_k is obtained from (6.9) by taking the expectation. The mathematical viability then depends on the fact whether a closed-form expression for the expectation of the right-hand side is available or not. Classically, one chooses affine processes $\lambda^{(k)}$, e.g. Cox-Ingersoll-Ross processes. The distribution of τ_k in an intensity-based approach may be viewed as an extension of the exponential distribution in the following sense:

the case when $\lambda_t^{(k)} \equiv c > 0$ agrees with an exponential distribution with parameter c. However, evidence is found in [Das et al. (2007)] that doubly-stochastic models fail to explain excess clustering as observed in the markets. Therefore, extensions have been proposed to incorporate for example *direct contagion effects*, i.e. the collapse of one firm directly increases the default intensities of the remaining firms, see e.g. [Yu (2007)]. However, the price one has to pay for incorporating such desirable features is limited viability. For instance, the calibration of a doubly-stochastic multivariate default model to market data usually requires time-consuming Monte Carlo techniques. This is due to the fact that the resulting multivariate distribution of $(\tau_1, \ldots, \tau_d)'$ is defined implicitly. In particular, one usually does not know the copula of the default times, which is required for the derivation of closed-form expressions for prices. Moreover, computing the unconditional distribution of τ_k, one obtains a formula which typically still depends on parameters that affect the distributions of all other default times as well. In this case, when calibrating the model, it is impossible to fit the univariate parameters in a first, and the dependence parameters of the model in a second step. Generally speaking, it is difficult to specify a priori the univariate stochastic processes $\lambda^{(k)}$, for $k = 1, \ldots, d$, and to equip them a posteriori with a dependence structure which does not change their pre-determined univariate distributions, and which is easy to work with; see e.g. [Bielecki et al. (2008)] for a discussion of this topic.

Such a drawback is overcome by copula-based approaches. The most popular model is invented by [Vasicek (1987)] and put into a copula framework by [Li (2000)]. It can be considered a multivariate analog of the structural model of [Merton (1974)] and defines a "risk factor" $A^{(k)}$ for company k by

$$A^{(k)} := \sqrt{\rho}\, X^{(0)} + \sqrt{1-\rho}\, X^{(k)}, \tag{6.10}$$

where $\rho \in [0,1]$ and $X^{(0)}$, $X^{(1)}, \ldots, X^{(d)}$ are i.i.d. standard normally distributed random variables. Denoting by Φ the distribution function of a standard normally distributed random variable and by G_k a pre-specified continuous distribution function for the k-th default time, τ_k is defined via $\tau_k := G_k^{-1}(\Phi(A^{(k)}))$. The crucial point in this ansatz is the fact that each $A^{(k)}$ is again standard normally distributed, independently of ρ. However, $A^{(1)}, \ldots, A^{(d)}$ are dependent since the common factor $X^{(0)}$ affects all d obligors. The larger ρ is, the bigger is the dependence of all companies on the market factor $X^{(0)}$. This tricky model construction allows to determine the marginal default probabilities (via the G_k) in a first step, and the dependence parameter ρ in a second step. The model (6.10)

implies that the random vector $\left(A^{(1)}, \ldots, A^{(d)}\right)'$, and hence also the random vector of default times $(\tau_1, \ldots, \tau_d)'$, has a Gaussian copula. This normality assumption is often criticized, since it does not support extreme scenarios. In mathematical terms, this can be seen from the fact that Gaussian copulas have zero tail dependence parameters. For this reason, the same construction as in equation (6.10) has been applied with several distributions other than the normal. One example is [Hull, White (2004)] who replace the normal distribution by a Student t-distribution. Some other extensions are unified in a generic modeling approach of [Albrecher et al. (2007)] using an arbitrary infinitely divisible distribution. More precisely, an infinitely divisible distribution defines a Lévy process $X = \{X_t\}_{t\in[0,1]}$, see [Sato (1999)] for further information. It is further assumed that $\mathbb{E}[X_1] = 0$ and $\mathrm{Var}[X_1] = 1$. The risk factor of company k is then defined by

$$A^{(k)} := X_\rho^{(0)} + X_{1-\rho}^{(k)}, \tag{6.11}$$

where $X^{(0)}, X^{(1)}, \ldots, X^{(d)}$ are independent copies of X, and $\rho \in [0,1]$. Especially if X is a standard Brownian motion, then (6.11) is equivalent to (6.10). Different models are obtained by choosing different specifications for the Lévy process X, and all these models are called *factor-copula models*. [Moosbrucker (2006)] uses a Variance-Gamma process, [Guégan, Houdain (2005), Kalemanova et al. (2007)] a Normal Inverse Gaussian process, and [Baxter (2006)] the sum of a Brownian motion and a Variance-Gamma process. One major drawback of these approaches is the fact that there is only one common factor $X_\rho^{(0)}$, conditioned on which the default times are independent: this exchangeability assumption simplifies mathematical derivations, but is often not very realistic. Nevertheless, the resulting ease of applicability is a striking argument in favor of these models. If a large homogeneous portfolio assumption is justified (as it is often the case for the pricing of CDOs), one can use the common factor $X_\rho^{(0)}$ to approximate the portfolio loss distribution. Since this allows to calibrate factor-copula models to CDO market data in seconds, they are commonly used in practice.

Last but not least, a modeling approach by [Schönbucher, Schubert (2001)] shares properties of intensity-based models and factor-copula models. The idea is to consider a classical intensity-based approach as in equation (6.8), where the trigger variables $E_1, \ldots, E_d$ are assumed to be dependent to begin with. More precisely, the triggers are equipped with a copula, often of Archimedean type. In the simplest form, when $\vec{\lambda}$ is a deterministic process, then the copula of $(\tau_1, \ldots, \tau_d)'$ is precisely the specified copula of the trigger variables $E_1, \ldots, E_d$. Hence, this approach is equivalent to speci-

fying the distribution of $(\tau_1, \ldots, \tau_d)'$ by its copula (via the copula of the triggers) and its marginal distributions (via the deterministic intensities). The separation of dependence structure and marginal distributions makes this approach well-suited for practical implementations, see e.g. [Hofert, Scherer (2009)] for a specific Monte Carlo pricing algorithm using nested Archimedean copulas. If the specified copula is of Archimedean type with completely monotone generator, one may even derive efficient analytic approximations for the portfolio loss as $d \to \infty$ under a homogeneous portfolio assumption, see [Schönbucher (2002)], an approach which is quickly implemented on a standard PC. This approximation relies on Theorem 2.3.10, which reveals a common latent factor as the source of dependence in such models. In this case the approximation is quite similar to the one in factor-copula models. Therefore, the model of [Schönbucher (2002)] is often said to belong to the class of factor-copula models, even though it is constructed differently. On a theoretical level, our approach below is similar, but the Archimedean copula is replaced with a Lévy-frailty copula, see Subsection 6.6 below for a closer analysis.

Most aforementioned models have been invented for the pricing of CDO contracts, the most liquid derivatives in the context of portfolio credit risk. As the underlying portfolio of a typical CDO is quite large, most of these models rely on simplifying assumptions to derive pricing formulas. For instance, a homogeneous portfolio structure and an infinite number of assets are convenient assumptions which may allow to approximate the portfolio loss distribution via stochastic limit theorems. For instance, factor-copula models of the form (6.11) are quickly implemented if a portfolio of infinitely many firms is assumed. Among them, the standard market model is the Gaussian copula model, see [Li (2000)]. In contrast to typical CDO portfolios, the standard basket size for an n-th-to-default swap is usually rather small, e.g. in the ninth series of the iTraxx Europe, baskets of five names are considered. Thus, for the pricing of portfolio credit derivatives depending on a small portfolio of underlying assets, assumptions such as a homogeneous portfolio are critical and a large portfolio assumption is not justified at all. Generally speaking, closed-form solutions of the portfolio loss distribution for an inhomogeneous portfolio are difficult to obtain, but required for the valuation of portfolio credit derivatives without time-consuming simulations. With regard to the pricing of n-th-to-default swaps, existing literature mainly focuses on the factor-copula models. Again, the standard market model is the Gaussian copula model, see [Li (2000)]. [Joshi, Kainth (2004)] describe an efficient Monte Carlo sampling technique to apply the Gaussian copula model to the pricing of n-th-to-default swaps and generalize this ap-

proach to other elliptic copulas. They especially treat the problem of small default probabilities by means of importance sampling for rare events. A method for the valuation of an n-th-to-default swap without Monte Carlo simulation is presented in [Hull, White (2004)]. By conditioning on a common market factor, conditional default probabilities are calculated recursively in this framework. Afterwards, the conditional probabilities are numerically integrated with respect to the market factor's distribution. This technique is applied to a Gaussian copula framework as well as using a t-copula. Both papers [Joshi, Kainth (2004), Hull, White (2004)] present sensitivity analyses for model prices with respect to input data such as correlations or idiosyncratic hazard rates.

The model presented in the next section shares properties with all aforementioned reduced-form approaches: its definition resembles an intensity-based model, however we achieve a separation of dependence and margins similar to the factor-copula models. We are going to construct the vector of default times $(\tau_1, \ldots, \tau_d)'$ such that it has the survival function $C_\Psi\big(\bar{G}_1(t_1), \ldots, \bar{G}_d(t_d)\big)$, where C_Ψ is an arbitrary Lévy-frailty copula, see Definition 4.2.1. $G_1, \ldots, G_d$ are given continuous and strictly increasing univariate distribution functions, with corresponding survival functions $\bar{G}_k = 1 - G_k$, for $k = 1, \ldots, d$. We will show that the arsenal of Lévy-frailty copulas, developed in Chapter 4, is well-suited for efficient CDO pricing. Our approach inherits all convenient properties of a copula-based approach, since it is quickly and easily calibrated to market data. This is due to the separation of quite arbitrary marginal default probabilities (determined by $G_1, \ldots, G_d$) and dependence parameters (given by Ψ). If a large homogeneous portfolio assumption is justified, calibration to market data is achieved quickly on a standard PC without the need for Monte Carlo simulations. Since the underlying dependence structure is of Marshall-Olkin kind, joint defaults are possible under a homogeneous portfolio assumption. None of the aforementioned models has this property. In fact, for reasons of mathematical convenience other approaches often rely on the fact that the probability of joint defaults is zero, see e.g. [Yu (2007)]. [Laurent, Gregory (2005), p. 4] even claim that a "conditional independence assumption between default times precludes simultaneous defaults"; a statement which we are going to disprove by the introduction of the Lévy-frailty default model. We think a model which supports joint defaults is reasonable, since global economy shocks may affect several firms at the same time. Of course, in practice it does not matter whether two default times are precisely identical or whether they are just close, but from a didactic point of view we think it is

appropriate to model joint defaults. Moreover, especially the recent credit crisis gives rise to believe that it is important to use a modeling approach which allows for "armageddon scenarios". Under a homogeneous portfolio assumption (i.e. $G_1 = \ldots = G_d$), this is reflected in our model by a positive probability that all d companies default at the same time, see e.g. Theorem 4.4.5.

Finally, we do not want to conceal that the idea of using Marshall-Olkin distributions for portfolio credit risk is already outlined in the references [Embrechts et al. (2003), Lindskog, McNeil (2003), Giesecke (2003), Andersen, Sidenius (2005)] as well as [Burtschell et al. (2009)]. However, these approaches are all based on the original construction (2.11). In dimensions $d \gg 2$ - in particular for the pricing of CDOs when typically one has $d = 125$ - this model is of limited practical use. In particular, the formulas derived in [Giesecke (2003)] are impossible to evaluate for large d. Probably for this reason, [Giesecke (2003)] provides an examplary graph for $d = 30$, which is much smaller than the usual CDO portfolio size $d = 125$. The references [Embrechts et al. (2003), Lindskog, McNeil (2003), Giesecke (2003)] indicate how to simulate Marshall-Olkin distributions based on the construction (2.11). But it was already argued in the last chapter that even this algorithm is inefficient in large dimensions $d \gg 2$. [Giesecke (2003), Lindskog, McNeil (2003)] propose to circumvent this problem by restricting the general Marshall-Olkin distribution to hierarchical subclasses. More clearly, it is assumed that almost all $2^d - 1$ original parameters λ_I are zero. For example, the portfolio is subdivided into a small number of groups, and the only parameters that are positive are those corresponding to this partition. Given this assumption, an efficient simulation of the Marshall-Olkin distribution is possible. However, this model only supports joint defaults of whole groups and is therefore quite simplistic. A more elaborate model in the same spirit - in fact a generalization of such models - would be to use hierarchical Lévy-frailty copulas as discussed in the previous chapter. However, the pricing of CDO tranches using this approach still relies on time-consuming Monte Carlo simulations. It is still unclear how to derive a convenient closed-form expression for required quantities. The references [Andersen, Sidenius (2005), Burtschell et al. (2009)] restrict their analysis to the even more simplistic Marshall-Olkin model based on a single "kill-all" shock, corresponding to the global shock copula of Example 4.3.1. Obviously, extending this approach to general Lévy-frailty copulas is a major improvement. The use of Lévy-frailty copulas allows to derive efficient approximations for the distribution of the portfolio loss process $\{L_t\}_{t \geq 0}$, see Section 6.4 below.

6.3 Definition of the Model

Throughout we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Since the model is fitted to market prices later on, we assume from the very start that $\mathbb{P}$ is a given pricing measure, i.e. we put ourselves in a risk-neutral setting. On this probability space let $E_1, \ldots, E_d$ be i.i.d. random variables with $E_1 \sim Exp(1)$. Independent of them, let $\Lambda = \{\Lambda_t\}_{t\geq 0}$ be a Lévy subordinator. The Laplace exponent Ψ of Λ is assumed to satisfy $\Psi(1) = 1$. Furthermore, for each $k = 1, \ldots, d$ let $h_k : [0, \infty) \to [0, \infty)$ be a continuous and strictly increasing function with $h_k(0) = 0$ and $\lim_{t\to\infty} h_k(t) = \infty$. The function h_k is the so-called *cumulative hazard function* of the distribution function

$$G_k(t) := 1 - \exp\big(-h_k(t)\big), \quad t \geq 0.$$

The vector of default times $(\tau_1, \ldots, \tau_d)'$ is defined by

$$\tau_k := \inf\big\{t \geq 0 : \Lambda_{h_k(t)} \geq E_k\big\}, \quad k = 1, \ldots, d. \tag{6.12}$$

Equivalently, the default indicator process $D = \big(D^{(1)}, \ldots, D^{(d)}\big)'$ is given by

$$D^{(k)} = \{D_t^{(k)}\}_{t\geq 0}, \quad D_t^{(k)} = \mathbf{1}_{\{\tau_k \leq t\}} = \mathbf{1}_{\{\Lambda_{h_k(t)} \geq E_k\}}, \quad t \geq 0, \quad k = 1, \ldots, d.$$

Comparing definition (6.12) to definition (6.8) of a classical intensity-based approach, one observes that the integrated intensity process is replaced by a time-changed Lévy subordinator. Since the Lévy subordinator is a jump process, joint defaults become possible. This is impossible in classical doubly-stochastic frameworks, since the integrated intensity process is continuous. Nevertheless, it is demonstrated below that the presented model is quite different from intensity-based approaches and has rather to be considered as a copula-based approach. In particular, it is shown in the sequel that the survival function of $(\tau_1, \ldots, \tau_d)'$ is $C_\Psi\big(\bar{G}_1(t_1), \ldots, \bar{G}_d(t_d)\big)$ - which is well-studied, for instance in earlier chapters of this thesis.

Lemma 6.3.1 (Survival Function of the Default Times)
The joint survival function of $(\tau_1, \ldots, \tau_d)'$ is given by

$$\mathbb{P}\big(\tau_1 > t_1, \ldots, \tau_d > t_d\big) = C_\Psi\big(\bar{G}_1(t_1), \ldots, \bar{G}_d(t_d)\big), \quad t_1, \ldots, t_d > 0,$$

where C_Ψ denotes the Lévy-frailty copula corresponding to Λ.

Proof

One easily verifies for all $k = 1, \dots, d$ and $t > 0$ that

$$\begin{aligned}\mathbb{P}\big(\tau_k > t\big) &= \mathbb{P}\big(\Lambda_{h_k(t)} < E_k\big) = \mathbb{E}\Big[\mathbb{P}\big(\Lambda_{h_k(t)} < E_k \,\big|\, \Lambda_{h_k(t)}\big)\Big] \\ &= \mathbb{E}\Big[e^{-\Lambda_{h_k(t)}}\Big] = e^{-h_k(t)\,\Psi(1)} = e^{-h_k(t)} = \bar{G}_k(t).\end{aligned}$$

Thus, the marginal survival functions are given by $\bar{G}_1, \dots, \bar{G}_d$. With $u_1, \dots, u_d \in (0,1)$ the survival copula $\hat{C}$ of $(\tau_1, \dots, \tau_d)'$ is thus computed via

$$\begin{aligned}\hat{C}(u_1, \dots, u_d) &= \mathbb{P}\big(\tau_1 > \bar{G}_1^{-1}(u_1), \dots, \tau_d > \bar{G}_d^{-1}(u_d)\big) \\ &= \mathbb{P}\big(\Lambda_{h_1\left(\bar{G}_1^{-1}(u_1)\right)} < E_1, \dots, \Lambda_{h_d\left(\bar{G}_d^{-1}(u_d)\right)} < E_d\big) \\ &= \mathbb{P}\big(\Lambda_{-\log(u_1)} < E_1, \dots, \Lambda_{-\log(u_d)} < E_d\big).\end{aligned}$$

The last equality follows from the identity $h_k = -\log(\bar{G}_k)$, for $k = 1, \dots, d$. We know from Theorem 4.2.2 that for all $t_1, \dots, t_d > 0$ it holds that

$$C_\Psi\big(e^{-t_1}, \dots, e^{-t_d}\big) = \mathbb{P}\big(\Lambda_{t_1} < E_1, \dots, \Lambda_{t_d} < E_d\big).$$

Hence, plugging $t_k := -\log u_k$, for $k = 1, \dots, d$, into the last equation implies that the unique survival copula $\hat{C}$ equals C_Ψ. □

It is observed that τ_k has distribution function G_k and the default indicator $D_t^{(k)}$ at time $t > 0$ has a Bernoulli distribution with success probability $G_k(t)$ for $k = 1, \dots, d$. Moreover, the presented construction induces dependence to (almost) arbitrary marginal default distributions $G_1, \dots, G_d$. If for example $\lambda^{(k)} : [0, \infty) \to [0, \infty)$, $i = 1, \dots, d$, are (deterministic) non-negative functions with $\int_0^\infty \lambda^{(k)}(s)\, ds = \infty$, one can define the marginal distribution functions as $G_k(t) = 1 - \exp(-\int_0^t \lambda^{(k)}(s)\, ds)$, $t \geq 0$. Equivalently, the cumulative hazard functions take the form $h_k(t) = \int_0^t \lambda^{(k)}(s)\, ds$.

When a portfolio credit risk model is used in practice, it is often useful and intuitive to consider bivariate pairs of default indicators, see [Lucas (1995)] for a motivation. In particular, classical factor-copula models as defined in (6.11) use a dependence parameter $\rho \in [0, 1]$, which equals the correlation of two firms' risk factors at a fixed time point. In the present model there is an inherent parameter, which plays the role of ρ. Recall that each bivariate pair $(\tau_i, \tau_j)'$ of default times has a bivariate Cuadras-Augé survival copula with parameter $\alpha_\Psi := 2 - \Psi(2) \in [0, 1]$, which follows from Lemma 6.3.1 above together

with Definition 2.2.2. We know from previous chapters that $\alpha_\Psi = 0$ implies independence of the two default times, whereas $\alpha_\Psi = 1$ implies complete comonotonicity. The following lemma collects several properties related to the model-inherent dependence parameter α_Ψ.

Lemma 6.3.2 (Properties of α_Ψ)
Recall for $k = 1, \dots, d$ that $D_t^{(k)} = \mathbf{1}_{\{\tau_k \le t\}} = \mathbf{1}_{\{\Lambda_{h_k(t)} > E_k\}}$, $t \ge 0$.

(a) Denoting by (μ, ν) the characteristics of Λ, it holds that $0 \le \alpha_\Psi \le 1 - \mu \le 1$. Moreover, α_Ψ may be represented as

$$\alpha_\Psi = 2 - \Psi(2) = \int_{(0,\infty]} \left(1 - e^{-t}\right)^2 \nu(dt).$$

(b) Consider a sequence of Lévy subordinators $\{\Lambda^{(k)}\}_{k\in\mathbb{N}}$ with Laplace exponents $\{\Psi_k\}_{k\in\mathbb{N}}$ satisfying $\Psi_k(1) = 1$. Denote by ν_k the Lévy measure of $\Lambda^{(k)}$.

(i) If for a positive null sequence $\{a_k\}_{k\in\mathbb{N}}$ it holds that $\nu_k\big([a_k, \infty]\big) = 0$ for all $k \in \mathbb{N}$, and if $\sup_{k\in\mathbb{N}} \int_{(0,1)} t\, \nu_k(dt) < \infty$, then $\lim_{k\to\infty} \alpha_{\Psi_k} = 0$.

(ii) If $\lim_{k\to\infty} \alpha_{\Psi_k} = 1$, then for all $u_1, \dots, u_d \in [0,1]$ it holds that

$$\lim_{k\to\infty} C_{\Psi_k}(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\}.$$

(c) For $i \ne j$ and $t > 0$ the correlation coefficient of $D_t^{(i)}$ and $D_t^{(j)}$, in the spirit of [Lucas (1995)], is given by

$$\frac{G_i(t) + G_j(t) - 1 + \Big(1 - \big(G_i(t) \wedge G_j(t)\big)\Big)^{1-\alpha_\Psi} \Big(1 - \big(G_i(t) \vee G_j(t)\big)\Big)}{\sqrt{G_i(t) - G_i(t)^2}\,\sqrt{G_j(t) - G_j(t)^2}} - \frac{G_i(t)\, G_j(t)}{\sqrt{G_i(t) - G_i(t)^2}\,\sqrt{G_j(t) - G_j(t)^2}},$$

where for two real numbers x, y we denote $x \wedge y := \min\{x, y\}$ and $x \vee y := \max\{x, y\}$.

(d) Let $i \ne j$. Assume the existence of the limits $G_k'(0) := \lim_{t\downarrow 0} G_k(t)/t = h_k'(0) > 0$ for $k \in \{i, j\}$. Moreover, assume the existence of an $\epsilon > 0$ such that $h_i(t) \le h_j(t)$

for all $t \in [0, \epsilon]$. Then, it holds that

$$\lim_{t \downarrow 0} \operatorname{Corr}\big[D_t^{(i)}, D_t^{(j)}\big] = \sqrt{\frac{G_i'(0)}{G_j'(0)}}\, \alpha_\Psi.$$

Proof

(a) The drift μ of Λ is completely determined by the Lévy measure ν via $\Psi(1) = 1$. It is given by $\mu = 1 - \int_{(0,\infty]} \big(1 - \exp(-t)\big)\, \nu(dt)$. Hence,

$$\begin{aligned}\alpha_\Psi &= 2 - \Psi(2) = 2 - 2\,\mu - \int_{(0,\infty]} (1 - e^{-2\,t})\, \nu(dt) \\ &= 2 - 2 + \int_{(0,\infty]} 2\,(1 - e^{-t}) + (e^{-2\,t} - 1)\, \nu(dt) = \int_{(0,\infty]} (1 - e^{-t})^2\, \nu(dt).\end{aligned}$$

From this it is obvious that α_Ψ is non-negative, and for the upper bound one observes

$$\alpha_\Psi = \int_{(0,\infty]} (1 - e^{-t})^2\, \nu(dt) \leq \int_{(0,\infty]} (1 - e^{-t})\, \nu(dt) = 1 - \mu.$$

(b) Denote by μ_k the drift of $\Lambda^{(k)}$.

(i) Since $\nu_k\big([a_k, \infty]\big) = 0$ it follows from part (a) that

$$\begin{aligned}\alpha_{\Psi_k} &= \int_{(0,a_k)} \big(1 - e^{-t}\big)^2\, \nu_k(dt) \leq \int_{(0,a_k)} \big(1 - e^{-t}\big)\, t\, \nu_k(dt) \\ &\leq \big(1 - e^{-a_k}\big) \int_{(0,a_k)} t\, \nu_k(dt) \leq \big(1 - e^{-a_k}\big) \int_{(0,1)} t\, \nu_k(dt).\end{aligned}$$

The last inequality holds for almost all k, since $\{a_k\}_{k \in \mathbb{N}}$ is a null sequence. By assumption, the remaining integral is bounded by a constant independent of k, and the claim follows.

(ii) Each Ψ_k is a non-decreasing function, hence for $i \geq 2$ it holds that

$$\begin{aligned}0 \leq \Psi_k(i) - \Psi_k(i-1) &= \mu_k + \int_{(0,\infty]} e^{-(i-1)\,t} \big(1 - e^{-t}\big)\, \nu_k(dt) \\ &\leq \mu_k + \int_{(0,\infty]} e^{-t} \big(1 - e^{-t}\big)\, \nu_k(dt) = \Psi_k(2) - \Psi_k(1) = 1 - \alpha_{\Psi_k}.\end{aligned}$$

From this it follows that all differences $\Psi_k(i) - \Psi_k(i-1)$ tend to zero as k tends to infinity. From the definition of Lévy-frailty copulas (see Definition 4.2.1) the claim is now easily established.

(c) First of all, using the Lévy properties of Λ it is verified for positive numbers $a, b > 0$ and $0 < t_1 \leq t_2$ that

$$\begin{aligned}\mathbb{E}\Big[e^{-a\,\Lambda_{t_1} - b\,\Lambda_{t_2}}\Big] &= \mathbb{E}\Big[e^{-(a+b)\,\Lambda_{t_1} - b\,(\Lambda_{t_2} - \Lambda_{t_1})}\Big] \\ &= \mathbb{E}\Big[e^{-(a+b)\,\Lambda_{t_1}}\Big]\,\mathbb{E}\Big[e^{-b\,\Lambda_{t_2 - t_1}}\Big] = e^{-t_1\,\left(\Psi(a+b) - \Psi(b)\right) - t_2\,\Psi(b)}. \end{aligned} \tag{6.13}$$

Obviously, $D_t^{(k)}$ is Bernoulli distributed for $k \in \{i, j\}$ with success probability $G_k(t)$. Hence, it follows that $\mathrm{Var}(D_t^{(k)}) = G_k(t) - G_k(t)^2$. Moreover, by conditional independence, one obtains

$$\begin{aligned}\mathbb{E}\big[D_t^{(i)}\,D_t^{(j)}\big] &= \mathbb{E}\big[\mathbb{P}\big(E_i \leq \Lambda_{h_i(t)} \,\big|\, \sigma(\Lambda_s \,:\, s > 0)\big)\,\mathbb{P}\big(E_j \leq \Lambda_{h_j(t)} \,\big|\, \sigma(\Lambda_s \,:\, s > 0)\big)\big] \\ &= \mathbb{E}\big[(1 - e^{-\Lambda_{h_i(t)}})\,(1 - e^{-\Lambda_{h_j(t)}})\big] \\ &= 1 - e^{-h_i(t)} - e^{-h_j(t)} + \mathbb{E}\big[e^{-\Lambda_{h_i(t)} - \Lambda_{h_j(t)}}\big] \\ &= G_i(t) + G_j(t) - 1 + e^{-(h_i(t) \wedge h_j(t))\,(\Psi(2) - 1) - (h_i(t) \vee h_j(t))}.\end{aligned}$$

In the last equality (6.13) is applied with $a = b = 1$. From these computations the claim follows, since $\exp\big(-h_k(t)\big) = 1 - G_k(t)$ for $k \in \{i, j\}$ and $\Psi(2) - 1 = 1 - \alpha_\Psi$.

(d) Firstly, by the assumption on the existence and positivity of the limits $G_k'(0)$ for $k \in \{i, j\}$, L'Hospital's rule implies

$$\lim_{t \downarrow 0} \sqrt{\frac{G_i(t) - G_i(t)^2}{G_j(t) - G_j(t)^2}} = \sqrt{\frac{G_i'(0)}{G_j'(0)}}, \quad \lim_{t \downarrow 0} \sqrt{\frac{G_j(t) - G_j(t)^2}{G_i(t) - G_i(t)^2}} = \sqrt{\frac{G_j'(0)}{G_i'(0)}}.$$

Using this and the second assumption $h_i \leq h_j$ near zero, one can again apply L'Hospital's rule to the result in part (c):

$$\begin{aligned}&\lim_{t \downarrow 0} \left(\frac{G_i(t) + G_j(t) - 1 + \Big(1 - \big(G_i(t) \wedge G_j(t)\big)\Big)^{1 - \alpha_\Psi} \Big(1 - \big(G_i(t) \vee G_j(t)\big)\Big) - G_i(t)\,G_j(t)}{\sqrt{G_i(t) - G_i(t)^2}\,\sqrt{G_j(t) - G_j(t)^2}} \right) \\ &= \lim_{t \downarrow 0} \left(\frac{G_i(t) + G_j(t) - 1 + e^{\left(h_i(t) \wedge h_j(t)\right)\left(1 - \Psi(2)\right) - h_i(t) \vee h_j(t)\big)} - G_i(t)\,G_j(t)}{\sqrt{G_i(t) - G_i(t)^2}\,\sqrt{G_j(t) - G_j(t)^2}} \right)\end{aligned}$$

$$= 2 \lim_{t \downarrow 0} \frac{G_i'(t)+G_j'(t)+e^{(\cdots)} \left((1-\Psi(2))\, h_i'(t)-h_j'(t)\right)-G_i'(t)\, G_j(t)-G_j'(t)\, G_i(t)}{\sqrt{\frac{G_i(t)-G_i(t)^2}{G_j(t)-G_j(t)^2}} \left(G_j'(t)-2\, G_j'(t)\, G_j(t)\right)+\sqrt{\frac{G_j(t)-G_j(t)^2}{G_i(t)-G_i(t)^2}} \left(G_i'(t)-2\, G_i'(t)\, G_i(t)\right)}$$

$$= 2\, \frac{G_i'(0) + G_j'(0) + (1-\Psi(2))\, h_i'(0) - h_j'(0)}{\sqrt{\frac{G_i'(0)}{G_j'(0)}}\, G_j'(0) + \sqrt{\frac{G_j'(0)}{G_i'(0)}}\, G_i'(0)}$$

$$= \frac{G_i'(0) + (1-\Psi(2))\, G_i'(0)}{\sqrt{G_i'(0)\, G_j'(0)}} = \frac{G_i'(0)\left(2-\Psi(2)\right)}{\sqrt{G_i'(0)\, G_j'(0)}} = \sqrt{\frac{G_i'(0)}{G_j'(0)}}\, \alpha_\Psi.$$

In the fourth equality we use for $k \in \{i, j\}$ that

$$h_k'(t) = \Big(-\log\big(1-G_k(t)\big)\Big)' = \frac{G_k'(t)}{1-G_k(t)} \quad \Rightarrow \quad h_k'(0) = G_k'(0).$$

Thus, the claim is established. □

The results of Lemma 6.3.2 are interpreted as follows. Part (a) implies that high dependence (i.e. α_Ψ close to one) requires the Lévy subordinator to be of pure jump type. The larger the drift $\mu \in [0,1]$, the smaller the dependence parameter α_Ψ. However, it is important to stress that this statement is not restrictive concerning the level of dependence achievable in the model, since μ is implicitly defined via the condition $\Psi(1) = 1$. Therefore, if the model parameters are fitted to market quotes and large dependence is required, μ automatically becomes small. Part $(b)\,(i)$ states that significant dependence is typically induced by the possibility of large jumps, see also the following Example 6.3.3. Part $(b)\,(ii)$ means that in the limiting case $\alpha_\Psi = 1$ all d default times are comonotonic. In analytic terms, $\alpha_\Psi = 1$ forces the Laplace exponent of the Lévy subordinator to be given by $\Psi(x) = \mathbf{1}_{\{x>0\}}$, $x \geq 0$, which implies that $C_\Psi = M$ equals the upper-Fréchet-Hoeffding bound. Part (c) computes the correlation of two default indicators. In particular, this correlation is a function of the margins G_i, G_j and the dependence parameter α_Ψ. Part (d) shows that the model is able to allow for a positive limit of default correlations at zero; a property which is important for the pricing of correlation driven products with short maturity.

Example 6.3.3 (From Finite to Infinite Activity)
This example demonstrates part $(b)\,(i)$ of Lemma 6.3.2. Let $J > 0$ be a fixed jump size and define the Lévy subordinator as $\Lambda_t := J\, N_t$, $t \geq 0$, where $N = \{N_t\}_{t \geq 0}$ is a Poisson process with intensity $1/\big(1-\exp(-J)\big)$. It follows that the Lévy measure ν is given by $\nu(B) = \mathbf{1}_{\{J \in B\}}/\big(1-\exp(-J)\big)$, for $B \in \mathcal{B}\big((0,\infty]\big)$. The Laplace exponent Ψ

in this case is given by $\Psi(x) = \big(1-\exp(-x\,J)\big)/\big(1-\exp(-J)\big)$, $x \geq 0$, and has 1 as a fixpoint. By part (a) of Lemma 6.3.2 we have $\alpha_\Psi = 1-\exp(-J)$. Hence, α_Ψ tends to zero as J tends to zero, which agrees with part $(b)\,(i)$ of Lemma 6.3.2 (using that $\sup_{J\in(0,1)} J/\big(1-\exp(-J)\big) < \infty$). However, it is interesting to see that the intensity of the Poisson process N increases to infinity as J tends to zero. Thus, for large J the model has few big jumps, and for small J the model has many small jumps. In the limit as J goes to zero, the model tends to an infinite activity model with vanishing jump sizes. Since α_Ψ tends to zero in this case, one may conclude that jump size affects the dependence more than jump frequency.

Remark 6.3.4 (Link to Bernoulli Mixture Models)
The result on the correlation in part (c) of Lemma 6.3.2 can also be stated in terms of the mixing variable Λ:

$$\operatorname{Corr}\big[D_t^{(i)}, D_t^{(j)}\big] = \frac{\operatorname{Cov}\big[1-e^{-\Lambda_{h_i(t)}}, 1-e^{-\Lambda_{h_j(t)}}\big]}{\sqrt{\operatorname{Var}\big[D_t^{(i)}\big]}\,\sqrt{\operatorname{Var}\big[D_t^{(j)}\big]}}.$$

In the homogeneous case $h_1 = \ldots = h_d =: h$, the covariance of the default indicators is given by the variance of the mixing variable $1-\exp(-\Lambda_{h(t)})$. This result is in concordance with a finding by [Frey, McNeil (2001)] in the context of so-called *Bernoulli Mixture Models*.

Equivalently, denoting by $\hat{C}_{\alpha_\Psi}(u_1,u_2) := u_1+u_2-1+C_{\alpha_\Psi}(1-u_1,1-u_2)$ the survival copula of the Cuadras-Augé copula C_{α_Ψ} with parameter α_Ψ, the correlation coefficient may be written as

$$\operatorname{Corr}\big[D_t^{(i)}, D_t^{(j)}\big] = \frac{\hat{C}_{\alpha_\Psi}(G_i(t),G_j(t)) - G_i(t)\,G_j(t)}{\sqrt{G_i(t)-G_i(t)^2}\,\sqrt{G_j(t)-G_j(t)^2}}.$$

Recall that for the bivariate Cuadras-Augé copula the probability mass on the diagonal as well as the upper-tail dependence coefficient are known, see (2.1) and Example 2.2.10. Thus, we may gather several different representations for the important correlation measure α_Ψ, which are summarized as follows:

$$\begin{aligned}
\alpha_\Psi = 2-\Psi(2) &= \int_0^\infty (1-e^{-t})^2\,\nu(dt) && \text{(analytical formula)}\\
&= \lim_{u\downarrow 0} \mathbb{P}\big(\tau_1 \leq G_1^{-1}(u) \,\big|\, \tau_2 \leq G_2^{-1}(u)\big) && \text{(lower-tail dependence)}
\end{aligned}$$

$$= \sqrt{\frac{G_1'(0)}{G_2'(0)}} \lim_{t \downarrow 0} \mathrm{Corr}\big[D_t^{(1)}, D_t^{(2)}\big] \qquad \text{(default correlation at zero)}$$

$$= \frac{2\,\mathbb{P}\big(G_1(\tau_1) = G_2(\tau_2)\big)}{1 + \mathbb{P}\big(G_1(\tau_1) = G_2(\tau_2)\big)}. \qquad \text{(from the singular component)}$$

Notice that the third equality requires the assumptions from Lemma 6.3.2 (d) with $i = 1$, $j = 2$, and the last equality follows from the fact that $\mathbb{P}\big(G_1(\tau_1) = G_2(\tau_2)\big) = \alpha_\Psi/(2 - \alpha_\Psi)$, compare (2.1).

Figure 6.2 shows a simulation of the model with $d = 125$ firms over a period of 20 years. In this example, $h_1(t) = \ldots = h_d(t) = 0.01\,t =: h(t)$, corresponding to an exponential distribution with mean 100 for all default times, and Λ is specified as a 0.8-stable subordinator, i.e. $\Psi(x) = x^{0.8}$, $x \geq 0$. The bars in the graph show the default times that occur before time $t = 20$. The solid line illustrates the path of $\Lambda_{h(t)}$.

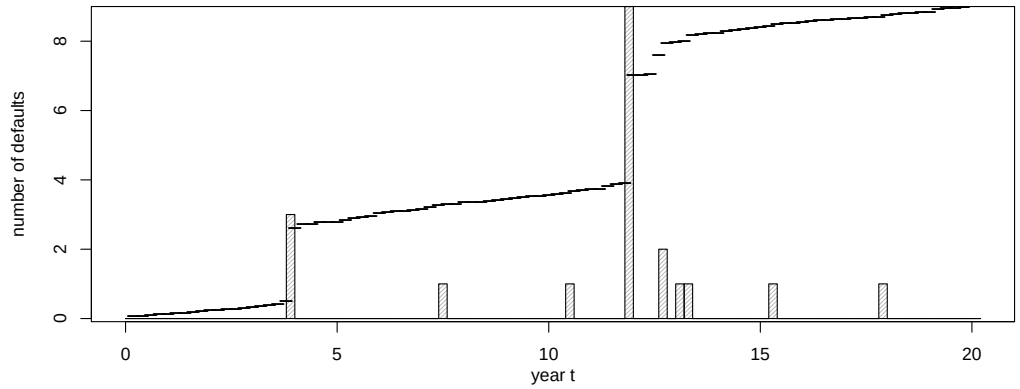


Figure 6.2 The plot shows one realization of dependent defaults in the presented model, together with the path of the time-changed Lévy subordinator.

6.4 The Portfolio Loss Distribution

In this section we establish the mathematical quantities in the Lévy-frailty default model which are required for the pricing of basket credit derivatives. As mentioned earlier, these comprise the distribution of L_t (for the pricing of portfolio CDS and CDOs) and

the distribution of the n-th largest default time $\tau_{(n)}$ (for the pricing of n-th-to-default contracts). The latter distribution is computed in Theorem 6.4.1 below. To handle the distribution of L_t we proceed in two directions:

- We compute the exact (discrete) distribution of L_t in the general case (Theorem 6.4.1 (a)) and in the special case of a homogeneous portfolio, i.e. if $G_1 = \ldots = G_d$ (Theorem 6.4.1 (b)).

- In the special case of a homogeneous portfolio, we derive an approximate distribution for L_t, which is obtained when assuming an infinite portfolio size (Theorem 6.4.3). For pricing CDOs, this is necessary for practical implementations and can be justified, since the portfolio size is large, typically $d = 125$.

The first result comprises the exact distribution of L_t and of $\tau_{(n)}$ in the general and in the homogeneous case.

Theorem 6.4.1 (The Portfolio Loss Distribution)
We introduce the superscript $[i]$ to denote an i-margin of the copula C_Ψ, i.e. we write $C_\Psi^{[i]}(u_1, \ldots, u_i)$, for $i \geq 2$. Additionally, for $i \in \{0, 1\}$ we define $C_\Psi^{[0]} \equiv 1$ and $C_\Psi^{[1]}(u) = u$.

(a) **General Case:** it holds for $k = 1, \ldots, d$ and $t > 0$ that

$$\mathbb{P}\big(d\, L_t = k\big) = \sum_{\substack{1 \leq i_1 < \ldots < i_k \leq d \\ \{j_1, \ldots, j_{d-k}, i_1, \ldots, i_k\} = \{1, \ldots, d\}}} \Big(C_\Psi^{[d-k]}\big(\bar{G}_{j_1}(t), \ldots, \bar{G}_{j_{d-k}}(t)\big) + \sum_{l=1}^{k} (-1)^l \sum_{1 \leq m_1 < \ldots < m_l \leq k} C_\Psi^{[d-k+l]}\big(\bar{G}_{j_1}(t), \ldots, \bar{G}_{j_{d-k}}(t), \bar{G}_{i_{m_1}}(t), \ldots, \bar{G}_{i_{m_l}}(t)\big) \Big). \tag{6.14}$$

(b) **Homogeneous Case:** if a homogeneous portfolio is assumed, i.e. $G_1 = \ldots = G_d =: G$, then (6.14) simplifies to

$$\mathbb{P}\big(d\, L_t = k\big) = \binom{d}{k} \sum_{l=0}^{k} (-1)^l \binom{k}{l} \bar{G}(t)^{\Psi(d+l-k)}.$$

We obtain for each $n = 1, \ldots, d$ the distribution of the n-th default time $\tau_{(n)}$, i.e.

$$\mathbb{P}\big(\tau_{(n)} \leq t\big) = \sum_{k=n}^{d} \mathbb{P}\big(d\, L_t = k\big). \tag{6.15}$$

Proof

(a) It can be shown by induction that for arbitrary numbers $x_1, \ldots, x_d \in \mathbb{R}$ one has

$$\prod_{i=1}^{d} \big(1 - e^{-x_i}\big) = 1 + \sum_{l=1}^{d} (-1)^l \sum_{1 \leq j_1 < \ldots < j_l \leq d} e^{-\sum_{z=1}^{l} x_{j_z}}. \tag{6.16}$$

This induction is carried out as follows: for $d = 1$ the statement is valid. Now by induction hypothesis (IH) the statement is assumed to hold for some $d \geq 1$. Then it follows that

$$\begin{aligned}
\prod_{i=1}^{d+1} \big(1 - e^{-x_i}\big) &= \big(1 - e^{-x_{d+1}}\big) \prod_{i=1}^{d} \big(1 - e^{-x_i}\big) \\
&\overset{(IH)}{=} \Big(1 + \sum_{l=1}^{d} (-1)^l \sum_{1 \leq j_1 < \ldots < j_l \leq d} e^{-\sum_{z=1}^{l} x_{j_z}}\Big) - e^{-x_{d+1}} \\
&\quad + \sum_{l=1}^{d} (-1)^{l+1} \sum_{1 \leq j_1 < \ldots < j_l \leq d} e^{-x_{d+1} - \sum_{z=1}^{l} x_{j_z}} \\
&= 1 \underbrace{-e^{-x_{d+1}} - \sum_{i=1}^{d} e^{-x_i}}_{= \sum_{1 \leq j_1 \leq d+1} (-1)^1 e^{-x_{j_1}}} + \underbrace{\sum_{l=2}^{d} (-1)^l \sum_{1 \leq j_1 < \ldots < j_l \leq d} e^{-\sum_{z=1}^{l} x_{j_z}}}_{= \sum_{l=2}^{d+1} (-1)^l \sum_{\substack{1 \leq j_1 < \ldots < j_l \leq d+1 \\ \text{all } j_i \neq d+1}} e^{-\sum_{z=1}^{l} x_{j_z}}} \\
&\quad + \underbrace{\sum_{l=1}^{d} (-1)^{l+1} \sum_{1 \leq j_1 < \ldots < j_l \leq d} e^{-x_{d+1} - \sum_{z=1}^{l} x_{j_z}}}_{= \sum_{l=2}^{d+1} (-1)^l \sum_{\substack{1 \leq j_1 < \ldots < j_l \leq d+1 \\ \text{at least one } j_i = d+1}} e^{-\sum_{z=1}^{l} x_{j_z}}} \\
&= 1 + \sum_{l=1}^{d+1} (-1)^l \sum_{1 \leq j_1 < \ldots < j_l \leq d+1} e^{-\sum_{z=1}^{l} x_{j_z}}.
\end{aligned}$$

Hence, (6.16) is established. Using conditional independence (conditioned on the information $\sigma(\Lambda_t : t \geq 0)$ about the whole path of the Lévy subordinator) in the

last equation, one computes

$$\begin{aligned}
&\mathbb{P}\Big(\sum_{i=1}^{d} D_t^{(i)} = k\Big) = \mathbb{P}\Big(\sum_{i=1}^{d} \mathbf{1}_{\{E_i \le \Lambda_{h_i(t)}\}} = k\Big) \\
&= \mathbb{P}\Bigg(\bigcup_{\substack{1\le i_1<\ldots<i_k\le d \\ \{j_1,\ldots,j_{d-k},i_1,\ldots,i_k\}=\{1,\ldots,d\}}} \cap_{\sigma=1}^{k} \cap_{\gamma=1}^{d-k} \Big\{ E_{i_\sigma} \le \Lambda_{h_{i_\sigma}(t)}\,,\, E_{j_\gamma} > \Lambda_{h_{j_\gamma}(t)} \Big\} \Bigg) \\
&= \sum_{\substack{1\le i_1<\ldots<i_k\le d \\ \{j_1,\ldots,j_{d-k},i_1,\ldots,i_k\}=\{1,\ldots,d\}}} \mathbb{P}\Big(E_{i_\sigma} \le \Lambda_{h_{i_\sigma}(t)},\, \sigma = 1,\ldots,k,\, E_{j_\gamma} > \Lambda_{h_{j_\gamma}(t)},\, \gamma = 1,\ldots,d-k \Big) \\
&= \sum_{\substack{1\le i_1<\ldots<i_k\le d \\ \{j_1,\ldots,j_{d-k},i_1,\ldots,i_k\}=\{1,\ldots,d\}}} \underbrace{\mathbb{E}\Big[\Big(\prod_{\sigma=1}^{k} \big(1 - e^{-\Lambda_{h_{i_\sigma}(t)}}\big) \Big) \Big(\prod_{\gamma=1}^{d-k} \big(e^{-\Lambda_{h_{j_\gamma}(t)}}\big) \Big) \Big]}_{=:(*)}.
\end{aligned}$$

Making use of equation (6.16) we further expand the inner expectation as follows:

$$\begin{aligned}
(*) &= \mathbb{E}\Bigg[\Big(1 + \sum_{l=1}^{k} (-1)^l \sum_{1\le m_1<\ldots<m_l\le k} e^{-\sum_{z=1}^{l} \Lambda_{h_{i_{m_z}}(t)}} \Big) e^{-\sum_{\gamma=1}^{d-k} \Lambda_{h_{j_\gamma}(t)}} \Bigg] \\
&= \mathbb{E}\Big[e^{-\sum_{\gamma=1}^{d-k} \Lambda_{h_{j_\gamma}(t)}} \Big] + \sum_{l=1}^{k} (-1)^l \sum_{1\le m_1<\ldots<m_l\le k} \mathbb{E}\Big[e^{-\sum_{z=1}^{l} \Lambda_{h_{i_{m_z}}(t)} - \sum_{\gamma=1}^{d-k} \Lambda_{h_{j_\gamma}(t)}} \Big].
\end{aligned}$$

This last term has the claimed copula representation, which can be seen from the proof of Theorem 4.2.2. Hence, the proof is complete.

(b) Conditioned on $\sigma(\Lambda_t : t > 0)$, $d\,L_t$ has a binomial distribution with d trials and success probability $1 - \exp(-\Lambda_{h(t)})$. Thus, one computes

$$\begin{aligned}
\mathbb{P}\Big(\sum_{i=1}^{d} D_t^{(i)} = k\Big) &= \mathbb{E}\Big[\mathbb{P}\Big(\sum_{i=1}^{d} \mathbf{1}_{\{E_i \le \Lambda_{h(t)}\}} = k \,\Big|\, \Lambda_{h(t)} \Big) \Big] \\
&= \mathbb{E}\left[\binom{d}{k} \big(1 - e^{-\Lambda_{h(t)}}\big)^k \big(e^{-\Lambda_{h(t)}}\big)^{d-k} \right] = \binom{d}{k} \mathbb{E}\Big[\big(1 - e^{-\Lambda_{h(t)}}\big)^k \big(e^{-\Lambda_{h(t)}}\big)^{d-k} \Big].
\end{aligned}$$

The term inside the expectation may be expanded using the binomial formula as

$$\big(1 - e^{-\Lambda_{h(t)}}\big)^k \big(e^{-\Lambda_{h(t)}}\big)^{d-k} = \sum_{l=0}^{k} \binom{k}{l} (-1)^l\, e^{-(l+d-k)\,\Lambda_{h(t)}};$$

the claim then follows easily. □

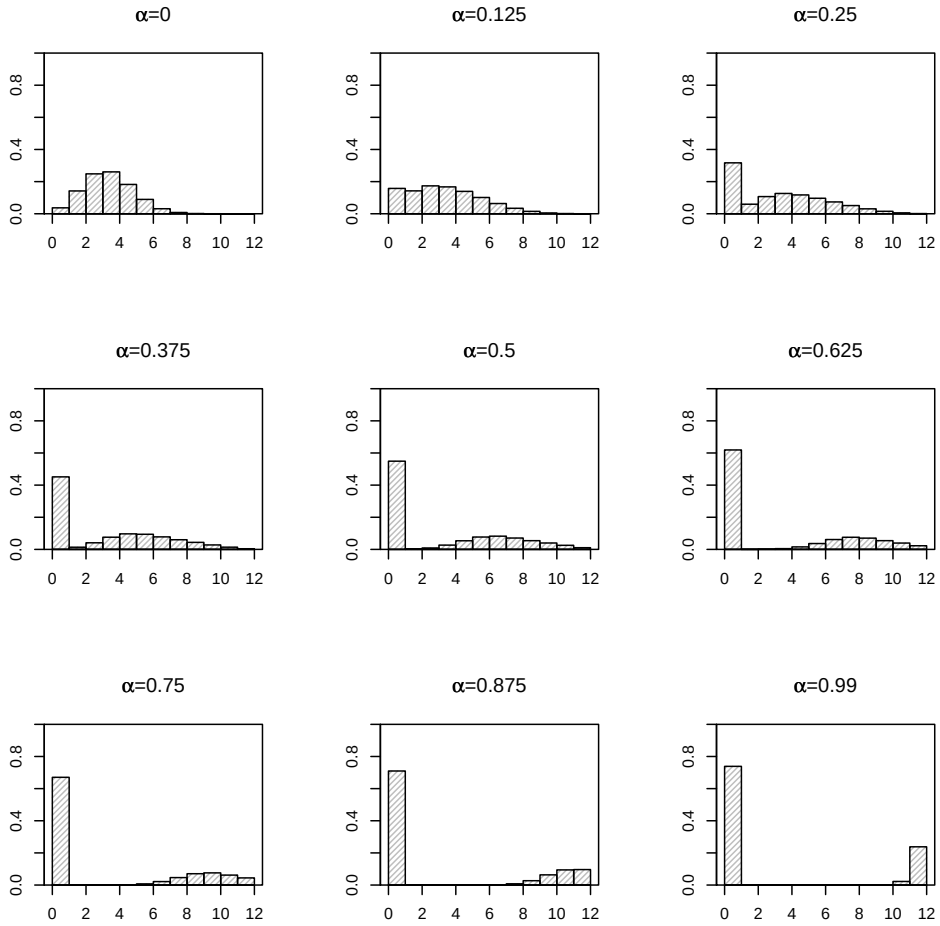


Figure 6.3 Histograms for the distribution of L_{10} with portfolio size $d = 11$. The marginal distribution G is chosen to be exponential with parameter 0.03, resulting in $G(10) \approx 26\%$. The underlying Lévy-frailty copula is of exchangeable Cuadras-Augé type, i.e. C_α as defined in (4.6), for a parameter $\alpha \in [0, 1]$. The parameter α is varied to illustrate several levels of dependence.

In the case of a homogeneous portfolio, Figure 6.3 illustrates the effect of dependence on the distribution of L_{10}, and fixed default probability $G(10) \approx 26\%$. The stronger the dependence among the default times, the more probability mass is assigned to events with many defaults. Moreover, the event of no default becomes quite likely if strong dependence is observed. Mathematically speaking: under independence, $10\,L_{10}$ has a

binomial law with success probability 26%. Under extremal dependence, in 26% of the cases all companies default, and in the other 74% of the cases, no company defaults.

Remark 6.4.2 (Pricing n-th-to-default Swaps)
In general, formula (6.15) from Theorem 6.4.1 can be plugged into the pricing formula (6.7) for the spread of an n-th-to-default swap. Admittedly, the terms in (6.14) become numerically intractable for large d, due to the numerous possibilities to choose k out of d elements, resulting in accordingly many summands. However, standard underlying baskets for n-th-to-default swaps consist of about $d = 5$ names, compare the iTraxx Europe conventions for the ninth series. For such small basket sizes the sums in equation (6.14) can be written out explicitly. The formulas are then evaluated on a standard computer in seconds. Especially, if only the distribution of the first default time is to be computed, the sum boils down to a single copula since for $t > 0$ one has

$$\mathbb{P}\big(\tau_{(1)} > t\big) = \mathbb{P}\big(\tau_1 > t, \tau_2 > t, \ldots, \tau_d > t\big) = C_\Psi\big(\bar{G}_1(t), \ldots, \bar{G}_d(t)\big)$$

by Theorem 6.3.1.

For the pricing of basket credit derivatives with large basket sizes $d \geq 20$, such as CDOs, the formulas of Theorem 6.4.1 become useless from a practical point of view. Formula (6.14) becomes inefficient to evaluate, not to say impossible. Even in the homogeneous case the binomial coefficients become huge and the sum is alternating, which causes inconvenient cancelation effects resulting in numerical imprecision. Thus, for practical implementations one has to approximate the distribution of L_t efficiently. In the homogeneous case, which we assume for the pricing of CDO contracts, this is achieved in the following theorem.

Theorem 6.4.3 (Approximation for Large Homogeneous Portfolios)
Assume $G_1 = \ldots = G_d =: G$ and consequently $h_1 = \ldots = h_d =: h$. As the portfolio size d tends to infinity, the following assertions can be made:

(a) Fix a time horizon $T \in (0, \infty)$. Then $\{L_t\}_{t\in[0,T]}$ tends to $\{1-\exp(-\Lambda_{h(t)})\}_{t\in[0,T]}$ in the Banach space $L^2(\Omega \times [0,T], \mathcal{F}, \mathbb{P})$ of all square-integrable stochastic processes X on $[0,T]$ with the norm

$$||X||_{L^2} := \Big(\mathbb{E}\Big[\int_{(0,T)} X_t^2 \, dt\Big]\Big)^{\frac{1}{2}}.$$

(b) $\sup_{t\geq 0}\left|L_t - \left(1-e^{-\Lambda_{h(t)}}\right)\right|$ tends to zero almost surely.

Proof

According to [Frey, McNeil (2001)], for fixed $t \in [0,T]$ the model is a Bernoulli Mixture Model with mixing variable $1-\exp(-\Lambda_{h(t)})$. From [McNeil et al. (2005), p. 357 ff] we can already derive an almost sure convergence of L_t for fixed $t \in [0,T]$. In contrast, the claimed statements are uniform in $t \in [0,T]$ (respectively $t \in [0,\infty)$) and are derived as follows.

(a) To show L^2-convergence, one computes $\mathbb{E}\big[L_t\big] = G(t)$ as well as (by a simple computation)

$$\mathbb{E}\big[L_t\,(1-e^{-\Lambda_{h(t)}})\big] = \mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big], \quad \mathbb{E}\big[L_t^2\big] = \frac{G(t)}{d} + \frac{d-1}{d}\,\mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big].$$

From this, it follows that

$$\begin{aligned}&\mathbb{E}\Big[\big(L_t - (1-e^{-\Lambda_{h(t)}})\big)^2\Big] = \mathbb{E}\big[L_t^2\big] - 2\,\mathbb{E}\big[L_t\,(1-e^{-\Lambda_{h(t)}})\big] + \mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big]\\ &= \frac{G(t)}{d} + \frac{d-1}{d}\,\mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big] - \mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big],\end{aligned}$$

which tends to zero as d tends to infinity. Hence, for $T>0$ it holds that

$$\begin{aligned}&\lim_{d\to\infty}\int_0^T \mathbb{E}\Big[\big(L_t - (1-e^{-\Lambda_{h(t)}})\big)^2\Big]\,dt\\ &\qquad = \lim_{d\to\infty}\left(\frac{1}{d}\int_0^T G(t)\,dt - \frac{1}{d}\int_0^T \mathbb{E}\big[(1-e^{-\Lambda_{h(t)}})^2\big]\,dt\right) = 0.\end{aligned}$$

Thus, $\{L_t\}_{t\in[0,T]}$ tends to $\{1-\exp(-\Lambda_{h(t)})\}_{t\in[0,T]}$ in $L^2(\Omega\times[0,T],\mathcal{F},\mathbb{P})$.

(b) The second statement follows from the Gliwenko-Cantelli Theorem: conditioned on the path of Λ, the function $t \mapsto L_t$ equals one realization of the empirical distribution function of the distribution function $t \mapsto 1-\exp(-\Lambda_{h(t)})$. Hence,

$$\begin{aligned}&\mathbb{P}\Big(\lim_{d\to\infty}\sup_{t\geq 0}\left|L_t - \left(1-e^{-\Lambda_{h(t)}}\right)\right| = 0\Big)\\ &\qquad = \mathbb{E}\Big[\mathbb{P}\Big(\lim_{d\to\infty}\sup_{t\geq 0}\left|L_t - \left(1-e^{-\Lambda_{h(t)}}\right)\right| = 0 \,\Big|\, \sigma(\Lambda_s\,:\,s\geq 0)\Big)\Big] = \mathbb{E}[1] = 1,\end{aligned}$$

where the second equality follows from the classical Gliwenko-Cantelli Theorem, see e.g. [Loève (1977), p. 20]. □

As a consequence, Theorem 6.4.3 justifies the following useful argument: provided $d \gg 2$ is large and the random variable $\Lambda_{h(t)}$ is absolutely continuous with density $f_{\Lambda_{h(t)}}$ for $t > 0$, the distribution of L_t may be approximated by the absolutely continuous distribution with density $f_t^{(\infty)}$ given by

$$f_t^{(\infty)}(x) := f_{\Lambda_{h(t)}}\big(-\log(1-x)\big)\,\frac{1}{1-x}, \quad x \in (0,1), \tag{6.17}$$

which equals the density of the random variable $1 - \exp(-\Lambda_{h(t)})$.

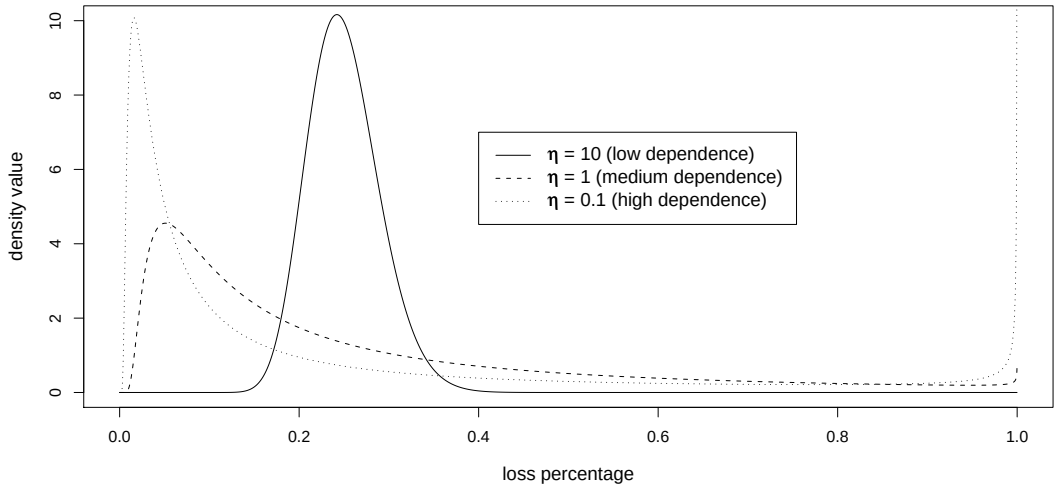


Figure 6.4 Specifying the dependence structure via an Inverse Gaussian subordinator, the approximate density (6.17) of the portfolio loss $L_t \approx 1 - \exp(-\Lambda_{h(t)})$ is illustrated. More precisely, the Laplace exponent is given by $\Psi(x) = (\sqrt{2\,x + \eta^2} - \eta)/(\sqrt{2 + \eta^2} - \eta)$, $x \geq 0$, for a parameter $\eta > 0$ (family (6) in Table 4.1). Parameters used are $h(t) = t\,\log(1/0.75)/5$ and $t = 5$, such that $\mathbb{E}[L_t] = 25\%$, and $\eta \in \{10, 1, 0.1\}$. The dependence measure $\alpha_\Psi = 2 - \Psi(2)$ decreases in η.

The required distribution of $\Lambda_{h(t)}$ is easy to handle if a convenient Lévy subordinator is specified to implement the model. Of particular interest for a calibration of the model to CDO market quotes are Lévy subordinators Λ, such that the density of Λ_t is known in closed form for all $t \geq 0$. If this is the case, then the required expectations $\mathbb{E}[L_{t,j}] =: \mathbb{E}[g(L_t)]$ can be approximated by $\mathbb{E}\big[g\big(1 - \exp(-\Lambda_{h(t)})\big)\big]$ and numerically be computed within fractions of a second. Examples of such Lévy subordinators comprise

the Gamma subordinator, the Inverse Gaussian subordinator and several compound Poisson subordinators (all possibly with additional drifts). Laplace inversion techniques allow to choose general Lévy subordinators, but the computation becomes more expensive in that case. One such example is an α-stable subordinator. In some special cases (e.g. compound Poisson processes with drift and constant or exponential jump sizes), the expectations $\mathbb{E}[L_{t,j}]$ can even be computed analytically and exact formulas are obtained. In the case of an Inverse Gaussian subordinator, Figure 6.4 visualizes the density of the approximated portfolio loss for three different levels of dependence.

6.5 Calibration to CDO Market Quotes

In the sequel we present a calibration of the model to CDO market quotes, using the approximation for the portfolio loss from Theorem 6.4.3. The payment streams for the calibration are in concordance with the iTraxx conventions, i.e. quarter-yearly premium payments and attachment points $[0\%, 3\%]$, $[3\%, 6\%]$, $[6\%, 9\%]$, $[9\%, 12\%]$, and $[12\%, 22\%]$ for the tranches, compare Table 6.1. Further required input are discount factors, which are obtained from risk-free par yields. Market quotes, to which the model is calibrated, comprise the portfolio CDS spreads with maturities three and five years, the upfront payment for the first tranche, and spreads for the remaining tranches; all tranche spreads for contracts maturing in five years. One week of daily data is used from the seventh series of iTraxx Europe, ranging from June 20, 2007, to June 26, 2007, and a calibration is run for each of these days.

Due to the separation of dependence structure and marginal default probabilities, one can proceed in two steps. At first, the marginal distribution G is calibrated, for which a piecewise linear intensity is assumed. To be precise, we let

$$1 - G(t) := e^{-\int_0^t \lambda(s)\,ds}, \quad \lambda(t) := \lambda_3 \min\{t, 3\} + \lambda_5\, t\, \mathbf{1}_{\{t>3\}}, \quad t \geq 0, \tag{6.18}$$

where λ_3 and λ_5 are positive intensity parameters which are calibrated to the portfolio CDS spreads for a three- and a five-year contract, respectively. This calibration is done via two succeeding bisection procedures. Secondly, λ_3 and λ_5 are fixed and the parameters of the Lévy subordinator, specifying the dependence, are calibrated to observed market spreads of the tranches of the CDO. All implemented Lévy subordinators are chosen such that they are parameterized by two parameters (η, β). To calibrate those,

a grid is defined for η. For each given η, the equality $\Psi(1) = 1$ defines an admissibility interval for β. On this interval, β is chosen such that the observed market upfront payment is matched perfectly, if possible. This is done by a bisection procedure again. This procedure is guaranteed to find at most one solution, as the upfront payment is monotone in β. After completion of these steps, among the obtained parameter pairs that perfectly explain the upfront payment, (η, β) is chosen to be the minimizer of the sum of absolute deviations of market to model spreads over all tranches $j \geq 2$. Using the approximation of Theorem 6.4.3, the whole calibration requires only few seconds on a Mac iBook G4 with 1.2 GHz for all implemented choices of Λ. The results of the calibration are shown in Table 6.2.

We implement three different model specifications, i.e. three different Lévy subordinators, respectively Lévy-frailty copulas. All of them are such that the dependence in the model is parameterized by two parameters (η, β):

(i) EXP(η, β): in this specification we choose the Lévy subordinator according to family (8) in Table 4.1. That is we let Λ be a compound Poisson subordinator with drift and exponential jump sizes, i.e.

$$\Lambda_t := \mu(\eta, \beta; EXP)\, t + \sum_{i=1}^{N_t} J_i, \quad N_1 \sim Poi(\beta),\, J_1 \sim Exp(\eta), \quad t \geq 0.$$

In order for the Laplace exponent Ψ of Λ to satisfy $\Psi(1) = 1$, the parameters have to satisfy

$$\eta > 0, \quad 0 < \beta \leq \eta + 1, \quad \mu(\eta, \beta; EXP) := 1 - \frac{\beta}{\eta + 1}.$$

It follows that $\alpha_\Psi = 2\,\beta/(\eta^2 + 3\,\eta + 2)$. In particular, for each $\alpha \in (0, 1)$ there are parameters (η, β) such that $\alpha_\Psi = \alpha$. This can be seen easily by choosing $\beta = \eta + 1$ and exploring the resulting expression for α_Ψ as a function of $\eta > 0$.

(ii) $\Gamma(\eta, \beta)$: in this specification we choose the Lévy subordinator according to family (9) in Table 4.1. That is we let Λ^Γ be a Gamma subordinator with parameters (η, β) and set

$$\Lambda_t := \mu(\eta, \beta; \Gamma)\, t + \Lambda_t^\Gamma, \quad t \geq 0.$$

In order for the Laplace exponent Ψ of Λ to satisfy $\Psi(1) = 1$, the parameters have to satisfy

$$\eta > 0, \quad 0 < \beta \leq \left(\log\left(1 + \frac{1}{\eta}\right) \right)^{-1}, \quad \mu(\eta, \beta; \Gamma) := 1 - \beta \log\left(1 + \frac{1}{\eta}\right).$$

It follows that $\alpha_\Psi = \beta \log\big((\eta+1)^2/(\eta^2 + 2\,\eta)\big)$. In particular, for each $\alpha \in (0,1)$ there are parameters (η, β) such that $\alpha_\Psi = \alpha$. This can be seen easily by choosing $\beta = 1/\log\big(1 + 1/\eta\big)$ and exploring the resulting expression for α_Ψ as a function of $\eta > 0$.

(iii) IG(η, β): in this specification we choose the Lévy subordinator according to family (10) in Table 4.1. That is we let Λ^{IG} be an Inverse Gaussian subordinator with parameters (η, β) and set

$$\Lambda_t := \mu(\eta, \beta; IG)\, t + \Lambda_t^{IG}, \quad t \geq 0.$$

In order for the Laplace exponent Ψ of Λ to satisfy $\Psi(1) = 1$, the parameters have to satisfy

$$\eta > 0, \quad 0 < \beta \leq \big(\sqrt{2 + \eta^2} - \eta\big)^{-1}, \quad \mu(\eta, \beta; IG) := 1 - \beta\,\big(\sqrt{2 + \eta^2} - \eta\big).$$

It follows that $\alpha_\Psi = \beta\,(-\eta + 2\,\sqrt{2 + \eta^2} - \sqrt{4 + \eta^2})$. For each $\alpha \in (0, 2 - \sqrt{2})$ there are parameters (η, β) such that $\alpha_\Psi = \alpha$. This can be seen by choosing $\beta = 1/\big(\sqrt{2 + \eta^2} - \eta\big)$ and exploring the resulting expression for α_Ψ as a function of $\eta > 0$. Even though a dependence coefficient $\alpha_\Psi > 2 - \sqrt{2} \approx 0.586$ cannot be obtained using this model specification, the achievable dependence structures are flexible enough for our purpose.

Table 6.2 on page 168 contains the calibration results. It shows the three- and five-year portfolio CDS spreads $s_{3y}^{CDS}, s_{5y}^{CDS}$ used for the calibration, as well as the fitted parameters λ_3 and λ_5 of the marginal distribution. These are independent of the choice of Lévy subordinator and the market quotes are perfectly matched each time. Concerning the dependence, Table 6.2 also shows the fitted dependence parameters (η, β) for each of the three implemented models, the corresponding upfront payment (in percent), and the tranche spreads (in basis points). Moreover, the market quotes for these quantities are reported, and the absolute deviation of model to market spreads E_a (in basis points) and the implied dependence parameter α_Ψ (in

percent) are given in Table 6.2. It is observed that our criterion of minimizing the sum of absolute errors results in the fact that the second tranche is also matched each time. The reason for this phenomenon might be that the spread of this tranche is much higher than those of tranches 3, 4, and 5. Furthermore, the large clipping from tranche spread 2 to tranche spread 3, as observed in the market in the considered period, causes most fitting problems. This decrease from tranche 2 to tranche 3 seems to be hard to capture for other popular models as well, see e.g. the calibration results in [Kalemanova et al. (2007), Albrecher et al. (2007), Hofert, Scherer (2009)]. The inverse Gaussian subordinator copes best with this phenomenon among the three specifications implemented. Overall, one might conclude that the calibration results are satisfying for the considered period.

Figure 6.5 on page 169 illustrates the calibrated Laplace exponents and the corresponding default probabilities for June 26, 2007. It is observed that the Laplace exponent of the Inverse Gaussian subordinator grows slowest among the three specifications. In probabilistic terms this corresponds to the fact that the corresponding joint default probabilities $p_k := \mathbb{P}(\tau_1 = \ldots = \tau_k)$ are highest for large k. Recall that by Theorem 4.4.5 p_k equals the upper extremal dependence coefficient of an arbitrary k-margin of the Lévy-frailty copula C_Ψ. Equivalently, they represent lower extremal dependence coefficients of k-dimensional subvectors of the vector of default times, compare Remark 2.2.9. In other words, p_k gives the probability that k firms collapse immediately, given at least one firm collapses immediately. Consequently, for large k the number p_k is a measure of extremal dependence in the model, since it quantifies how strongly one default is correlated with defaults of the other firms. Concluding, the Inverse Gaussian subordinator implies the most extremal dependence among the three specifications, and the compound Poisson process the fewest. Notice that Figure 6.5 does not precisely illustrate p_k, but monotonically transformed values $f(p_k)$, since p_k are rapidly decreasing and hence too difficult to visualize otherwise. The computation of the values p_k for large k is accomplished using the integral representation in Corollary 4.4.7.

Table 6.2 Fitted portfolio CDS and CDO tranche spreads (5 years).

Model$(\eta, \beta, \lambda_3, \lambda_5)$	s_{3y}^{CDS}	s_{5y}^{CDS}	up in %	s^2 in bp	s^3 in bp	s^4 in bp	s^5 in bp	E_a in bp	α_Ψ in %
Market 6-20-07	11.50	21.60	7.13	47.00	12.30	5.60	2.10		
EXP$(10.28, 2.48, 0.131\%, 0.162\%)$	11.50	21.60	7.13	47.00	26.85	14.86	4.31	25.94	3.6
$\Gamma(5.48, 1.57, 0.131\%, 0.162\%)$	11.50	21.60	7.13	47.00	24.02	13.49	4.68	22.10	3.8
IG$(2.59, 1.00, 0.131\%, 0.162\%)$	11.50	21.60	7.13	47.00	22.13	12.37	4.66	19.07	3.9
Market 6-21-07	12.59	22.78	8.48	50.47	13.28	6.11	2.43		
EXP$(9.72, 2.42, 0.143\%, 0.168\%)$	12.59	22.78	8.48	50.48	29.73	17.00	5.23	30.13	3.9
$\Gamma(5.05, 1.48, 0.143\%, 0.168\%)$	12.59	22.78	8.48	50.48	26.56	15.33	5.61	25.68	4.1
IG$(2.44, 0.94, 0.143\%, 0.168\%)$	12.59	22.78	8.48	50.28	24.40	14.03	5.58	22.37	4.3
Market 6-22-07	13.00	23.36	9.65	55.08	14.49	6.68	2.69		
EXP$(11.93, 2.93, 0.148\%, 0.171\%)$	13.00	23.36	9.65	55.10	28.87	14.60	3.61	23.25	3.3
$\Gamma(6.57, 1.97, 0.148\%, 0.171\%)$	13.00	23.36	9.65	55.09	26.12	13.62	4.16	20.06	3.5
IG$(2.88, 1.19, 0.148\%, 0.171\%)$	13.00	23.36	9.65	55.07	24.30	12.78	4.34	17.57	3.7
Market 6-25-07	13.56	24.33	10.88	59.00	15.16	6.82	2.90		
EXP$(12.12, 3.03, 0.154\%, 0.178\%)$	13.56	24.33	10.88	58.97	30.63	15.34	3.74	24.83	3.3
$\Gamma(6.66, 2.03, 0.154\%, 0.178\%)$	13.56	24.33	10.88	59.00	27.81	14.39	4.35	21.68	3.5
IG$(2.92, 1.23, 0.154\%, 0.178\%)$	13.56	24.33	10.88	58.96	25.73	13.40	4.47	18.71	3.6
Market 6-26-07	13.70	24.12	11.87	63.70	16.26	7.35	3.18		
EXP$(22.25, 5.87, 0.156\%, 0.175\%)$	13.70	24.12	11.87	63.70	19.84	5.79	0.62	7.70	2.1
$\Gamma(14.65, 5.48, 0.156\%, 0.175\%)$	13.70	24.12	11.87	63.70	18.65	5.95	0.82	6.14	2.2
IG$(4.64, 2.72, 0.156\%, 0.175\%)$	13.70	24.12	11.87	63.70	17.96	6.05	0.99	5.18	2.4

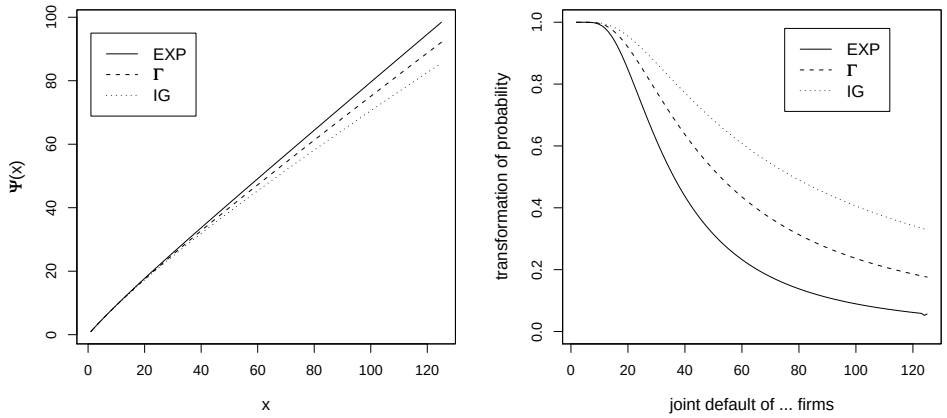


Figure 6.5 The left plot shows the calibrated Laplace exponents for the three implemented models. The right plot illustrates the joint default probabilities $p_k := \mathbb{P}(\tau_1 = \ldots = \tau_k)$ for $k = 2, \ldots, 125$. Since these are hard to distinguish by eye, the monotonically transformed values $1 - \exp\left(-(10^{15}\, p_k)^{1/10}\right)$ are plotted. This transformation is chosen such that one can clearly see the differences for the three calibrated models, when k is large.

6.6 Comparison with Existing Models

This section embeds the presented Lévy-frailty default model into existing literature. Our primary motivation was to construct a portfolio default model which can quickly and efficiently be calibrated to observed CDO market quotes, and in our view is thus of true practical relevance. Therefore, simplifying assumptions such as Lévy properties and a large homogeneous portfolio assumption are essential for the performed calibration in Section 6.5. Generally speaking, if one insists on very realistic assumptions such as e.g. direct contagion effects and a heterogeneous portfolio, a calibration to market data becomes expensive, or even impossible. Among the default models mentioned earlier in Section 6.2, there are many approaches which rely on time-consuming Monte Carlo techniques, when they are calibrated to market quotes. Even though some of them may exhibit desirable properties from a theoretical perspective, we think that comparing them to the presented Lévy-frailty approach is like comparing apples and oranges. Rather, we focus on a closer comparison of our model with other approaches

which allow for efficient calibrations as well: these are the factor-copula models and an approach by [Schönbucher (2002)].

The standard model for the pricing of CDO contracts which is used in the market nowadays is the Gaussian factor-copula model by [Li (2000)]. More generally, the universal framework by [Albrecher et al. (2007)] for most factor-copula models shares the same advantage with regard to an efficient implementation: a large homogeneous portfolio assumption allows to approximate the portfolio loss L_t via a random variable with known density. This enables the practitioner to compute all required expectations of the form $\mathbb{E}[g(L_t)]$ within fractions of a second. To set the present Lévy-frailty approach apart from factor-copula models, the following points can be made:

- The dependence structure in the classical Gaussian factor-copula model, introduced by [Li (2000)], is determined by a single parameter $\rho \in [0, 1]$. Models with only one parameter are often not flexible enough to fit observed market data appropriately. Notice that the implemented examples of Lévy-frailty models from the previous paragraph are chosen two-parametric.

- The dependence structure of the default times $(\tau_1, \ldots, \tau_d)'$ is often not very well studied. Only in the original model by [Li (2000)] (as well as in the model by [Hull, White (2004)]), the underlying copula is quite established in the academic literature: it is the Gaussian copula (resp. the Student t-copula), which has several drawbacks. For instance, the Gaussian copula has zero tail dependence parameters and both families share an unrealistic level of symmetry. In the more general model of [Albrecher et al. (2007)], the resulting copula of the default times is not satisfactorily studied. For instance, distributional properties such as tail dependence parameters or other multivariate dependence measures are not known in closed form. The choice of a particular model is basically justified by a good fit to market data. In contrast, the Lévy-frailty default model relies on the Marshall-Olkin distribution, which is well-studied in the literature (and in the present thesis) and has a number of desirable features (e.g. asymmetric tail dependence coefficients and a singular component).

- Using the large homogeneous portfolio approximation for efficient pricing, the common latent factor in all previously studied factor-copula models is a single random variable. More precisely, using the notation of equation (6.11), the common factor is the random variable $X_\rho^{(0)}$. Given $X_\rho^{(0)}$, the default indicator process $\{D_t\}_{t\geq 0}$ is

known completely. Therefore, the approach is static in the sense that the time t does not affect the latent variable. In contrast, the common factor in the Lévy-frailty model is the path of a Lévy subordinator and the default indicator D_t at time t depends on the complete path of this stochastic process until time t. Even though the Lévy properties are restrictive to some extent, we think that this modeling approach deserves to be considered more dynamic than factor-copula models.

The model which is closest to ours is the approach of [Schönbucher (2002)] in the Archimedean special case, when a large homogeneous portfolio assumption allows to approximate the portfolio loss by a single random variable with known density. The precise model is outlined below. The approximation of the portfolio loss relies on Theorem 2.3.10, which derives a probabilistic construction of Archimedean copulas with completely monotone generator. The marginal distributions can be chosen quite arbitrarily, similar as in the present approach. On a high level, the difference in the Lévy-frailty default model is that the underlying copula is not Archimedean but a Lévy-frailty copula, which has different distributional properties. From this theoretical point of view, comparing our approach to the model by [Schönbucher (2002)] boils down to comparing properties of Archimedean copulas (with completely monotone generator) and Lévy-frailty copulas. In this regard, the following points can be made:

- Lévy-frailty copulas have a singular component, compare Theorem 4.4.5. Translated into the language of default modeling this implies positive probabilities of joint defaults. We think this is a major innovation, and might be an intuitive and desirable feature of a default model. Archimedean copulas with completely monotone generators do not share this property.

- As a subclass of Marshall-Olkin survival copulas, Lévy-frailty copulas can be interpreted from an *environmental shock* point of view: firms in the portfolio are affected by exogenous economy shocks, compare the original construction (2.11). We think this interpretation is quite intuitive, which is also the reason why the references [Giesecke (2003), Lindskog, McNeil (2003)] propose this model in the same context. In contrast, the probabilistic construction of Archimedean copulas in Theorem 2.3.10 in our opinion is not well-suited to give a similar intuitive interpretation of default times. Additionally, Lévy-frailty copulas, as constructed in Theorem 4.2.2, alternatively allow for an *internal damage* interpretation: a com-

pany fails when the cumulated damage, modeled as a time-changed Lévy subordinator, exceeds its exponential trigger variable. The construction of Archimedean copulas can be written in a similar form, see (6.19) below, however this definition appears much more artificial than in the Lévy-frailty case.

- In theory, the class of Archimedean copulas (with completely monotone generator) is more flexible than the class of Lévy-frailty copulas, since it is parameterized by Laplace transforms, which form an infinite-dimensional parameter space. In practice however, all popular Archimedean copulas are parameterized by a single parameter. A one-parametric dependence model might be too restrictive to fit observed CDO market data.

To compare the calibration performance of both approaches, this section calibrates the approach of [Schönbucher (2002)] to the CDO market data from the previous section. To this end, the aforementioned special case of the model of [Schönbucher (2002)] is outlined in more detail.

The model input consists of two components:

(1) A continuous and strictly increasing distribution function G on $(0,\infty)$ which has precisely the same meaning as in the Lévy-frailty default model.

(2) A Laplace transform $\varphi(x) = \mathbb{E}[\exp(-x\,W)]$, $x \geq 0$, of a positive random variable W. Notice that W plays the role of the Lévy subordinator Λ in our approach.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the default times $\big(\tau_1, \ldots, \tau_d\big)'$ are defined for $k = 1, \ldots, d$ via

$$\tau_k := \inf\big\{t > 0 \,:\, W\,\varphi^{-1}\big(1 - G(t)\big) > E_k\big\} = \inf\Big\{t > 0 \,:\, 1 - G(t) > \varphi\Big(\frac{E_k}{W}\Big)\Big\}, \quad (6.19)$$

where $E_1, \ldots, E_d$ are i.i.d. with $E_1 \sim Exp(1)$, independent of W. Notice in particular that the random vector $(U_1, \ldots, U_d)'$, defined by setting $U_k := \varphi(E_k/W)$, has as joint distribution function the Archimedean copula (2.12) with generator φ, see Theorem 2.3.10. More clearly,

$$\mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d) = C\big(1 - G(t_1), \ldots, 1 - G(t_d)\big), \quad t_1, \ldots, t_d \geq 0,$$

where C is given by (2.12). A similar argument as in Theorem 6.4.3 furthermore establishes that the portfolio loss $L_t := \frac{1}{d}\sum_{k=1}^{d} \mathbf{1}_{\{\tau_k \leq t\}}$ converges uniformly in $t > 0$ to

$1 - \exp\big(- W\,\varphi^{-1}(1 - G(t))\big)$, as the portfolio size d tends to infinity. E.g. if W is absolutely continuous, then for large d the distribution of L_t can be approximated in the same spirit as in (6.17) by the density

$$f_t^{(\infty)}(x) := f_W\Big(\frac{-\log(1-x)}{\varphi^{-1}(1-G(t))}\Big)\,\frac{1}{1-x}\,\frac{1}{\varphi^{-1}(1-G(t))}, \quad x \in (0,1),$$

where f_W denotes the density of W. Having calibrated the deterministic function G to portfolio CDS spreads in a first step, in a second step this density can be calibrated to CDO market quotes in a similar fashion as our model is calibrated in the previous paragraph. [Schönbucher (2002)] proposes to use an Archimedean copula of either Clayton- (i.e. W is $\Gamma(1, 1/\theta)$-distributed with $\theta > 0$, see Example 2.3.11) or Gumbel-type (i.e. W is $1/\theta$-stable distributed for $\theta \geq 1$, i.e. $\varphi(x) = \exp(-x^{1/\theta})$). These proposals are due to the fact that these two Archimedean copulas are the most popular ones in the academic literature. However, in the case of a Gumbel copula, the density of W is of a stable kind and, in regard of the formulas (2.21), must be evaluated numerically, making this choice less attractive for applications. We examine an implementation of the Clayton-model (with parameter $\theta > 0$), using the same data as in the previous chapter. The results are shown in Table 6.3. More precisely, Table 6.3 provides the calibration results of three different models:

- **C1**$(\theta, \lambda_3, \lambda_5)$**:** the marginal distribution is determined by G, which is parameterized by $\lambda_3, \lambda_5 > 0$ and given as before by (6.18). The dependence is parameterized by the Clayton parameter $\theta > 0$, i.e. the default times are constructed as in (6.19) with $W \sim \Gamma(1, 1/\theta)$. This implies that the survival copula of the default times is a Clayton copula.

- **C2**$(\theta, \lambda_3, \lambda_5)$**:** the marginal distribution is determined by G, which is parameterized by $\lambda_3, \lambda_5 > 0$ and given as before by (6.18). In contrast to the model C1$(\theta, \lambda_3, \lambda_5)$, the default times now have a Clayton copula, instead of a Clayton survival copula. This is achieved when replacing (6.19) with the similar construction

$$\tau_k := \inf\big\{t > 0 : W\,\varphi^{-1}\big(p(t)\big) < E_k\big\}, \quad k = 1, \ldots, d. \tag{6.20}$$

- **IG**$(\eta, \beta, \lambda_3, \lambda_5)$**:** this is the same model as used before in Table 6.2. The results are repeated here in order to be compared with the results of the two Archimedean

models.

It is important to stress that the marginal distribution of the default times is identical in all three models, only the dependence structure, i.e. the copula, differs. Recall that an eMO-copula exhibits upper-tail dependence, which implies that the default times exhibit lower-tail dependence. In our opinion this is a desirable property, since it means that an early default of one firm is likely to coincide with a collapse of other firms as well. Since the Clayton copula exhibits lower-tail dependence, but no upper-tail dependence, the second model specification C2 is included. The subtle reformulation (6.20) of (6.19) switches from the survival copula to the actual copula of the default times, and accordingly switches upper- and lower-tail dependencies, see Remark 2.2.9. The calibration strategy used is the same as in the previous section: after specifying the marginal distribution via λ_3 and λ_5, the remaining dependence parameter θ is chosen such that the market upfront payment is matched perfectly. The calibration results of model C2 are better than those of model C1, which strengthens our intuition that lower-tail dependence of the default times is a desirable feature. Summarizing, the absolute error E_a is smallest in the two-parametric Lévy-frailty model.

Table 6.3 Fitted portfolio CDS and CDO tranche spreads (5 years), comparison with Clayton model.

Model	s_{3y}^{CDS}	s_{5y}^{CDS}	up in %	s^2 in bp	s^3 in bp	s^4 in bp	s^5 in bp	E_a in bp
Market 6-20-07	11.50	21.60	7.13	47.00	12.30	5.60	2.10	
C1(1.93, 0.131%, 0.162%)	11.50	21.60	7.13	88.90	15.84	2.81	0.17	50.25
C2(0.09, 0.131%, 0.162%)	11.50	21.60	7.13	81.42	18.42	4.91	0.6	42.83
IG(2.59, 1.00, 0.131%, 0.162%)	11.50	21.60	7.13	47.00	22.13	12.37	4.66	19.07
Market 6-21-07	12.59	22.78	8.48	50.47	13.28	6.11	2.43	
C1(1.89, 0.143%, 0.168%)	12.59	22.78	8.48	98.41	18.45	3.42	0.21	58.01
C2(0.09, 0.143%, 0.168%)	12.59	22.78	8.48	90.13	20.96	5.71	0.7	49.47
IG(2.44, 0.94, 0.143%, 0.168%)	12.59	22.78	8.48	50.28	24.40	14.03	5.58	22.37
Market 6-22-07	13.00	23.36	9.65	55.08	14.49	6.68	2.69	
C1(1.71, 0.148%, 0.171%)	13.00	23.36	9.65	96.04	16.42	2.74	0.15	49.37
C2(0.08, 0.148%, 0.171%)	13.00	23.36	9.65	88.34	19.1	4.79	0.53	41.92
IG(2.88, 1.19, 0.148%, 0.171%)	13.00	23.36	9.65	55.07	24.30	12.78	4.34	17.57
Market 6-25-07	13.56	24.33	10.88	59.00	15.16	6.82	2.90	
C1(1.63, 0.154%, 0.178%)	13.56	24.33	10.88	101.82	17.46	2.91	0.16	51.82
C2(0.08, 0.154%, 0.178%)	13.56	24.33	10.88	93.85	20.16	5.04	0.55	44.01
IG(2.92, 1.23, 0.154%, 0.178%)	13.56	24.33	10.88	58.96	25.73	13.40	4.47	18.71
Market 6-26-07	13.70	24.12	11.87	63.70	16.26	7.35	3.18	
C1(1.33, 0.156%, 0.175%)	13.70	24.12	11.87	82.87	10.41	1.22	0.07	34.27
C2(0.07, 0.156%, 0.175%)	13.70	24.12	11.87	77.13	12.92	2.53	0.11	24.65
IG(4.64, 2.72, 0.156%, 0.175%)	13.70	24.12	11.87	63.70	17.96	6.05	0.99	5.18

7 Conclusion

"A conclusion is the place where you got tired of thinking."

A. Bloch, American writer.

The results of this dissertation can be divided into two blocks: the first block (Chapters 3 and 4) contains purely theoretical results, which contribute to the theory of multivariate distributions and the theory of Lévy subordinators. The second block (Chapters 5 and 6) consists of practical applications based on the results of the first block. One application is concerned with the simulation of random vectors with given dependence structure (Chapter 5), the other contributes to portfolio credit risk modeling and pricing (Chapter 6).

The first block can be summarized as follows: the skeletal structure is constituted by the three major findings Theorem 3.4.1 (page 75), Theorem 3.5.3 (page 78), and Theorem 4.2.2 (page 95). Theorem 3.4.1 establishes an alternative representation of exchangeable Marshall-Olkin survival copulas (eMO-copulas). To this end, d-monotone sequences are introduced and investigated. As a consequence of this result, the study of analytical properties of eMO-copulas is facilitated essentially. For instance, in the second major result, Theorem 3.5.3, d-monotone sequences are characterized by eMO-copulas, and vice versa. Direct consequences are an alternative version of Hausdorff's Theorem (Corollary 3.5.6) and interesting analytical analogies between the class eMO and Archimedean copulas (Section 3.6). Last but not least, Theorem 4.2.2 shows how some eMO-copulas, namely Lévy-frailty copulas, can be constructed using Lévy subordinators. Having two alternative construction methods for the same distribution allows to switch between both models whenever one is more convenient than the other. This idea is applied in Section 4.4 to derive some probabilistic properties of eMO-copulas.

The second block consists of two applications of Theorem 4.2.2: Chapter 5 demonstrates how to efficiently sample Lévy-frailty copulas and extensions thereof. It is pointed out that in large dimensions $d \gg 2$ these techniques are necessary, since the original

probabilistic construction of Marshall-Olkin distributions can only be used in small dimensions. Chapter 6 presents a financial application. Lévy-frailty copulas are used to construct a multivariate default model which is well-suited for the pricing of CDO tranches. The model is systematically embedded into existing literature on portfolio credit derivative pricing models, and its distributional properties are investigated thoroughly.

Concluding, the family of Lévy-frailty copulas is defined, and explored from an analytic and from a probabilistic viewpoint. As a subclass of Marshall-Olkin survival copulas, the family is embedded into existing literature. In particular, relations to exchangeable Cuadras-Augé copulas are pointed out. Moreover, Lévy-frailty copulas are demonstrated to be of practical use by an application to portfolio credit risk modeling. Interestingly, some purely mathematical results are obtained as byproducts. For instance, coherences between different mathematical concepts such as Lévy subordinators, copulas, and completely monotone sequences are revealed.

A Appendix

A.1 Proof sketch for Theorem 2.3.2

This proof sketch is based on [Resnick (1987), Proposition 5.11, p. 268 ff].

- First of all, the whole statement is translated into the language of *multivariate extreme-value distributions* by transforming the margins of the copula from $U[0,1]$ to a so-called *Fréchet distribution*. More clearly, one considers the function

$$G_*(t_1,\ldots,t_d) := C\big(e^{-\frac{1}{t_1}},\ldots,e^{-\frac{1}{t_d}}\big), \quad t_1,\ldots,t_d > 0.$$

  The claim is then equivalent to the following statement:
  G_* is a distribution function with margins $t \mapsto \exp(-1/t)$, $t > 0$, satisfying $G_*(t_1,\ldots,t_d) = G_*(t\,t_1,\ldots,t\,t_d)^t$ for all $t > 0$, if and only if there exists a measure δ satisfying the claimed boundary conditions and

$$G_*(t_1,\ldots,t_d) = \exp\Big(-\int_{S_d} \max_{1\le i\le d}\Big\{\frac{u_i}{t_i}\Big\}\,\delta(du_1,\ldots,du_d)\Big).$$

- The necessity of the statement is established as follows: by assumption it follows easily that the functions $G_n := G_*^{1/n}$, $n \in \mathbb{N}$, are distribution functions. Hence, they induce probability measures dG_n, $n \in \mathbb{N}$. It is thus possible to define the measures $\zeta_n := n\,dG_n$, $n \in \mathbb{N}$. It can be shown that, with $n \to \infty$, the measures ζ_n converge *vaguely*[1] to a unique limit measure ζ_* on $[0,\infty]^d \setminus \{(0,\ldots,0)'\}$. The measure δ is finally defined via

$$\delta(A) := \zeta_*\Big(\Big\{\vec{t} \in [0,\infty]^d \setminus \{(0,\ldots,0)'\} \;\Big|\; ||\vec{t}|| > 1,\ ||\vec{t}||^{-1}\,\vec{t} \in A\Big\}\Big), \quad A \in \mathcal{B}(S_d),$$

[1] For details see e.g. [Billingsley (1995), p. 371] or [Resnick (1987)].

where $\vec{t} := (t_1, \ldots, t_d)'$ and $||\vec{t}||$ denotes the Euclidean norm of $\vec{t}$. It can be shown that it has the claimed properties.

- To establish sufficiency, one proceeds the other way round: given δ, one first defines a measure ζ_* on $[0,\infty]^d \setminus \{(0,\ldots,0)'\}$ via

$$\zeta_*(A) := \iint_{T(A)} \frac{1}{r^2}\, dr\, \delta(du_1, \ldots, du_d), \quad A \in \mathcal{B}\big([0,\infty]^d \setminus \{(0,\ldots,0)'\}\big),$$

where $T : [0,\infty]^d \setminus \{(0,\ldots,0)'\} \to (0,\infty] \times S_d$ denotes the polar coordinate transformation $\vec{t} \mapsto (||\vec{t}||, \vec{t}/||\vec{t}||)'$. Considering the product measure $dt \times \zeta_*$ of the Lebesque measure and ζ_* on the space $[0,\infty) \times [0,\infty]^d \setminus \{(0,\ldots,0)'\}$, there exists a unique so-called *Poisson random measure* M with mean measure $dt \times \zeta_*$. Quickly explained, M is a random variable with values in the space of all point measures on $[0,\infty) \times [0,\infty]^d \setminus \{(0,\ldots,0)'\}$, say

$$M(A) = \sum_{k\in\mathbb{N}} \mathbf{1}_{\{(t_k, \vec{j}_k)' \in A\}}, \quad A \in \mathcal{B}\big([0,\infty) \times [0,\infty]^d \setminus \{(0,\ldots,0)'\}\big),$$

where the random variables t_k and the random vectors $\vec{j}_k$, $k \in \mathbb{N}$, are subject to certain properties which justify the term "Poisson" in the notion "Poisson random measure"[2], and whose distribution is uniquely determined by the mean measure $dt \times \zeta_*$. Based upon this probabilistic object, G_* is defined as the distribution function of the random vector $\vec{Y}_1$, where

$$\vec{Y}_t := \max\Big\{(0,\ldots,0)', \sup_{k:t_k\le t}\{\vec{j}_k\}\Big\}, \quad t > 0.$$

Notice that the maximum and the supremum are taken componentwise and the supremum might be $(-\infty,\ldots,-\infty)'$ if the set $\{k : t_k \le t\}$ is empty, which explains the occurence of $(0,\ldots,0)'$ in the definition of $\vec{Y}_t$. Finally, one can prove that G_* defined in this way has the claimed properties.

[2]For details see [Resnick (1987)].

A.2 Proof sketch for Theorem 2.4.3

This proof sketch follows the lines of [Bertoin (1999), Theorem 1.2]: first of all, the independence and stationarity of increments can be used to establish

$$\mathbb{E}\Big[e^{-x\,\Lambda_{t+s}}\Big] = \mathbb{E}\Big[e^{-x\,\Lambda_t}\Big]\,\mathbb{E}\Big[e^{-x\,\Lambda_s}\Big], \quad x,s,t \geq 0,$$

which implies the existence of a function Ψ satisfying (2.14), see e.g. [Billingsley (1995), A20, p. 540]. To relate this function to a measure, let $x > 0$ and $E \sim Exp(x)$ be independent of Λ. One computes

$$\begin{aligned}\Psi(x) &= \lim_{n\to\infty} n\big(1-e^{-\Psi(x)/n}\big) = \lim_{n\to\infty} n\,\mathbb{E}\Big[1-e^{-x\,\Lambda_{\frac{1}{n}}}\Big] \\ &= \lim_{n\to\infty} n\,\mathbb{E}\big[\mathbb{P}(\Lambda_{\frac{1}{n}} \geq E \,|\, \Lambda_{\frac{1}{n}})\big] = \lim_{n\to\infty} n\,\mathbb{P}(\Lambda_{\frac{1}{n}} \geq E) = \lim_{n\to\infty} n\,\mathbb{E}\big[\mathbb{P}(\Lambda_{\frac{1}{n}} \geq E \,|\, E)\big] \\ &= x \lim_{n\to\infty} \int_{[0,\infty)} e^{-x\,t}\, n\,\mathbb{P}(\Lambda_{\frac{1}{n}} \geq t)\,dt.\end{aligned}$$

Defining $\tilde{\nu}_n(dt) := n\,\mathbb{P}(\Lambda_{\frac{1}{n}} \geq t)\,dt$, $n \in \mathbb{N}$, the above computation can be used to show that the measures $\tilde{\nu}_n$ converge *vaguely*[3] to a unique limit measure $\tilde{\nu}$, as $n \to \infty$. Since the functions $t \mapsto f_n(t) := n\,\mathbb{P}(\Lambda_{\frac{1}{n}} \geq t)$, $n \in \mathbb{N}$, are non-increasing, $\tilde{\nu}$ is necessarily of the form

$$\tilde{\nu}(A) = \mu\,\mathbf{1}_{\{0\in A\}} + \int_{A\setminus\{0\}} f(t)\,dt, \quad A \in \mathcal{B}\big([0,\infty)\big),$$

with a non-increasing and right-continuous function $f : (0,\infty) \to [0,\infty)$ and $\mu \geq 0$. This can be used to show that

$$\frac{\Psi(x)}{x} = \lim_{n\to\infty} \int_{(0,\infty)} e^{-x\,t}\, f_n(t)\,dt = \mu + \int_{(0,\infty)} e^{-x\,t}\, f(t)\,dt.$$

Defining ν via $\nu((t,\infty]) := f(t)$ for all $t \in (0,\infty)$ (and $\nu(\{\infty\}) := \lim_{t\to\infty} f(t) =: f(\infty)$), clearly $\nu\big((\epsilon,\infty]\big) = f(\epsilon)$ is finite for all $\epsilon > 0$. The claimed representation of Ψ follows from the last equation via integration by parts:

$$\int_{(0,\infty)} f(t)\ \underbrace{x\,e^{-x\,t}\,dt}_{=d\left(1-e^{-x\,t}\right)} + \int_{(0,\infty)} \big(1-e^{-x\,t}\big)\ \underbrace{df(t)}_{\substack{=-d\nu((0,t])\\=-\nu(dt)}} = \Big[\big(1-e^{-x\,t}\big)\,f(t)\Big]_{t=0}^{t=\infty},$$

[3] See e.g. [Billingsley (1995), p. 371] or [Resnick (1987)].

and hence for $x > 0$:

$$\begin{aligned}\Psi(x) &= \mu\, x + \int_{(0,\infty)} x\, e^{-x\, t}\, f(t)\, dt \qquad (A.1)\\ &= \mu\, x + \Big[\big(1 - e^{-x\, t}\big)\, f(t)\Big]_{t=0}^{t=\infty} + \int_{(0,\infty)} \big(1 - e^{-x\, t}\big)\, \nu(dt)\\ &= \mu\, x + f(\infty) + \int_{(0,\infty)} \big(1 - e^{-x\, t}\big)\, \nu(dt) \overset{(*)}{=} \mu\, x + \int_{(0,\infty]} \big(1 - e^{-x\, t}\big)\, \nu(dt).\end{aligned}$$

In $(*)$ it is used that $f(\infty) = \nu(\{\infty\})$. Moreover, $\int_{(0,1]} t\, \nu(dt) < \infty$, since

$$e^{-1} \int_{(0,1]} t\, \nu(dt) = -e^{-1} \int_{(0,1]} t\, df(t) = \int_{(0,1]} e^{-1}\, f(t)\, dt - e^{-1}\, f(1) \leq \Psi(1) - \mu < \infty,$$

where the inequality follows from (A.1) with $x = 1$, and the last equality follows from integration by parts. Uniqueness of μ and ν follows from the uniqueness of the vague limit measure $\tilde{\nu}$.

The converse part of the theorem is proved in a similar manner as the sufficiency part of Theorem 2.3.2: there is a canonical construction of Lévy subordinators based on so-called *Poisson random measures*, which is sketched in the sequel. Since we have already seen how to construct Poisson processes in (2.13) on page 43, it is clear from Definition 2.4.2 that we may without loss of generality assume that $\nu(\{\infty\}) = 0$. The product $dt \times \nu$ of the Lebesgue measure and the given measure ν is easily checked to be σ-finite on $[0,\infty) \times (0,\infty)$, since $\nu\big([\epsilon,\infty)\big) < \infty$ for all $\epsilon > 0$. Consequently, [Resnick (1987), Proposition 3.6(i), p. 130] guarantees the existence of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a Poisson random measure on $[0,\infty) \times (0,\infty)$ with mean measure $dt \times \nu$, denoted by M. I.e. M is a random measure on $[0,\infty) \times (0,\infty)$ subject to certain properties[4]. Based on M, the required Lévy subordinator Λ is defined by

$$\Lambda_t := \mu\, t + \iint_{[0,\infty)\times(0,\infty)} \mathbf{1}_{\{s \in [0,t]\}}\, u\, M(ds, du), \quad t \geq 0.$$

Using well-known statements about the so-called Laplace functional of the Poisson random measure, see [Resnick (1987), Proposition 3.6(ii), p. 130], the defining properties of a Lévy subordinator are easily verified. The only point which requires a slightly more careful guess is the fact that Λ_t is finite $\mathbb{P}$-almost surely for $t > 0$. To this end, recall that a random variable τ on $[0,\infty]$ is almost surely finite if its Laplace transform is

[4] For details see [Resnick (1987)].

continuous at zero, see (the remark after) Theorem 2.3.7. Assumption (2.16) allows to apply the dominated convergence theorem to establish

$$\lim_{x \downarrow 0} \int_{(0,\infty)} \left(1 - e^{-u\,x}\right) \nu(du) = \int_{(0,\infty)} \lim_{x \downarrow 0} \left(1 - e^{-u\,x}\right) \nu(du) = 0,$$

which implies that

$$\lim_{x \downarrow 0} \mathbb{E}\Big[e^{-x\,(\Lambda_t - \mu\,t)}\Big] = \lim_{x \downarrow 0} \exp\Big(\int_{(0,\infty)} \left(1 - e^{-u\,x}\right) \nu(du)\Big) = 1, \quad t > 0.$$

The first equality again follows from [Resnick (1987), Proposition 3.6(ii), p. 130]. This shows that the random variable $\Lambda_t - \mu\,t$, and hence Λ_t, is $\mathbb{P}$-almost surely finite and therefore Λ is well-defined.

List of Figures

2.1 Function plots of bivariate Cuadras-Augé copulas. 18
2.2 Scatterplots from bivariate Cuadras-Augé copulas. 19
2.3 Pickands dependence function of a bivariate Cuadras-Augé copula. . . . 34
2.4 Scatterplots from a two- and three-dimensional Clayton copula. 42
2.5 Paths of a compound Poisson subordinator. 48
2.6 Paths of a Gamma subordinator. 50
2.7 Paths of an Inverse Gaussian subordinator. 51
2.8 Paths of an α-stable subordinator. 53

3.1 Scatterplots from a three-dimensional eMO-copula. 65
3.2 The set $\mathcal{M}_3$. 71

4.1 Illustration of the construction idea of Lévy-frailty copulas. 98

5.1 Histograms for 10 000-dimensional Cuadras-Augé samples. 120
5.2 Simulation runtimes for exchangeable Cuadras-Augé copulas. 121
5.3 Scatterplot matrix from hierarchical Cuadras-Augé copula. 129

6.1 Empirical company defaults. 137
6.2 Realization of the multivariate default model. 156
6.3 Histograms of portfolio loss distribution. 160
6.4 Densities of the approximated portfolio loss. 163
6.5 Calibrated Laplace exponents and transformed joint default probabilities. 169

List of Tables

3.1 Comparison Archimedean copulas vs. copulas from eMO. 87

4.1 Some parametric families of Lévy subordinators with $\Psi(1) = 1$. 106

6.1 The iTraxx and DJ CDX segmentation. 139
6.2 Fitted portfolio CDS and CDO tranche spreads (5 years). 168
6.3 Fitted portfolio CDS and CDO tranche spreads (5 years), comparison with Clayton model. 175

Bibliography

[Albrecher et al. (2007)] H. Albrecher, S.A. Ladoucette and W. Schoutens, A generic one-factor Lévy model for pricing synthetic CDOs, *Advances in Mathematical Finance (Applied and Numerical Harmonic Analysis), Birkhäuser Boston* (2007) pp. 259–278.

[Aldous (1985)] D.J. Aldous, Exchangeability and related topis, *Springer, École d'Été de Probabilités de Saint-Flour XIII-1983. Lecture Notes in Mathematics* **1117** (1985) pp. 1–198.

[Andersen, Sidenius (2005)] L. Andersen and J. Sidenius, CDO pricing with factor models: survey and comments, *Journal of Credit Risk* **1:3** (2005) 71–88.

[Anderson (1991)] W. Anderson, Continuous-time Markov chains, *Springer-Verlag* (1991).

[Applebaum (2004)] D. Applebaum, Lévy processes and stochastic calculus, *Cambridge University Press* (2004).

[Baxter (2006)] M. Baxter, Dynamic modelling of single-name credits and CDO tranches, *Working Paper - Nomura Fixed Income Quant Group, retrievable from http://www.nomura.com/resources/europe/pdfs/dynamiccdomodelling2.pdf* (2006).

[Bernstein (1929)] S. Bernstein, Sur les fonctions absolument monotones, *Acta Mathematica* **52:1** (1929) 1–66.

[Bertoin (1996)] J. Bertoin, Lévy processes, *Cambridge University Press* (1996).

[Bertoin (1999)] J. Bertoin, Subordinators: Examples and Applications, *Springer, École d'Été de Probabilités de Saint-Flour XXVII-1997. Lecture Notes in Mathematics* **1717** (1999) pp. 1–91.

[Bielecki et al. (2008)] T.R. Bielecki, A. Vidozzi and L. Vidozzi, Study of dependence for some stochastic processes, *Stochastic Analysis and Applications* **26:4** (2008) 903–924.

[Billingsley (1995)] P. Billingsley, Probability and measure, *Wiley Series in Probability and Mathematical Statistics* (1995).

[Black, Cox (1976)] F. Black and C. Cox, Valuing corporate securities: some effects of bond indenture provisions, *Journal of Finance* **31:2** (1976) pp. 351–367.

[Bochner (1955)] S. Bochner, Harmonic analysis and the theory of probability, *University of California Press, Berkeley* (1955).

[Bondesson (1982)] L. Bondesson, On simulation from infinitely divisible distributions, *Advances in Applied Probability* **14** (1982) pp. 855–869.

[Burtschell et al. (2009)] X. Burtschell, J. Gregory and J.-P. Laurent, A comparative analysis of CDO pricing models, in *The Definitive Guide to CDOs, G. Meissner (ed.), Chapter 15, Risk Books* (2009) pp. 389–427.

[Charpentier, Segers (2009)] A. Charpentier and J. Segers, Tails of multivariate Archimedean copulas, *Journal of Multivariate Analysis* **100:7** (2009) pp. 1521–1537.

[Clayton (1978)] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial dependency in chronic disease incidence, *Biometrika* **65** (1978) pp. 141–151.

[Cont, Tankov (2004)] R. Cont and P. Tankov, Financial modelling with jump processes, *Chapman and Hall/CRC Financial Mathematics Series* (2004).

[Cuadras, Augé (1981)] C.M. Cuadras and J. Augé, A continuous general multivariate distribution and its properties, *Communications in Statistics - Theory and Methods* **10:4** (1981) pp. 339–353.

[Damien et al. (1995)] P. Damien, P.M. Laud and A.F.M. Smith, Approximate random variate generation from infinitely divisible distributions with applications to Bayesian inference, *Journal of the Royal Statistical Society* **57:3** (1995) pp. 547–563.

[Das et al. (2007)] S.R. Das, D. Duffie, N. Kapadia and L. Saita, Common failings: how corporate defaults are correlated, *Journal of Finance* **62:1** (2007) pp. 93–117.

[David, Nagaraja (1970)] H.A. David and H.N. Nagaraja, Order statistics, *Wiley Series in Probability and Statistics* (1970).

[De Finetti (1937)] B. De Finetti, La prévision: ses lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré* **7** (1937) pp. 1–68.

[De Haan, Resnick (1977)] L. De Haan and S.I. Resnick, Limit theory for multivariate sample extremes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **40:4** (1977) pp. 317–337.

[Dette, Studden (1997)] H. Dette and W.J. Studden, The theory of canonical moments with applications in statistics, probability, and analysis, *Wiley Series in Probability and Statistics* (1997).

[Devroye (1986)] L. Devroye, Non-uniform random variate generation, *Springer-Verlag* (1986).

[Duffie, Singleton (1999)] D. Duffie and K. Singleton, Modeling term structures of defaultable bonds, *Review of Financial Studies* **12:4**, (1999) pp. 687–720.

[Duffie, Gârleanu (2001)] D. Duffie and N. Gârleanu, Risk and valuation of collateralized debt obligations, *Financial Analysts Journal* **57:1**, (2001) pp. 41–59.

[Duffie et al. (2007)] D. Duffie, L. Saita and K. Wang, Multi-period corporate default prediction with stochastic covariates, *Journal of Financial Economics* **83**, (2007) pp. 635–665.

[Durante et al. (2007)] F. Durante, J.J. Quesada-Molina and M. Úbeda-Flores, On a family of multivariate copulas for aggregation processes, *Information Sciences* **177:24** (2007), 5715–5724.

[Efron (1975)] B. Efron, Defining the curvature of a statistical problem (with applications to second order efficiency), *Annals of Statistics* **3:6** (1975), 1189–1242.

[Embrechts et al. (2003)] P. Embrechts, F. Lindskog and A.J. McNeil, Modelling dependence with copulas and applications to risk management, *Handbook of heavy tailed*

distributions in finance, edited by S. Rachev, published by Elsevier/North-Holland, Amsterdam (2003) pp. 329–384.

[Falk et al. (2004)] M. Falk, J. Hüsler and R.-D. Reiss, Laws of small numbers: extremes and rare events, *Birkhäuser Verlag, Basel-Boston-Berlin* (2004).

[Feller (1966)] W. Feller, An introduction to probability theory and its applications, volume II, second edition, *John Wiley and Sons, Inc.* (1966).

[Frahm (2006)] G. Frahm, On the extremal dependence coefficient of multivariate distributions, *Statistics and Probability Letters* **76:14** (2006) pp. 1470–1481.

[Frey, McNeil (2001)] R. Frey and A.J. McNeil, Modelling dependent defaults, *working paper, retrievable from http://e-collection.ethbib.ethz.ch/show?type=bericht&nr=273* (2001).

[Galambos, Kotz (1978)] J. Galambos and S. Kotz, Characterizations of probability distributions, *Lecture Notes in Mathematics (volume 675), Springer-Verlag Berlin - Heidelberg* (1978).

[Giesecke (2003)] K. Giesecke, A simple exponential model for dependent defaults, *Journal of Fixed Income* **13:3** (2003) pp. 74–83.

[Giesecke (2004)] K. Giesecke, Correlated default with incomplete information, *Journal of Banking and Finance* **28:7** (2004) pp. 1521–1545.

[Gnedin, Pitman (2005)] A. Gnedin and J. Pitman, Regenerative composition structures, *Annals of Probability* **33:2** (2005) pp. 445–479.

[Gnedin, Pitman (2008)] A. Gnedin and J. Pitman, Moments of convex distribution functions and completely alternating sequences, *Probability and Statistics: Essays in Honor of David A. Freedman* **2** (2008) pp. 30–41.

[Guégan, Houdain (2005)] D. Guégan and J. Houdain, Collateralized Debt Obligations pricing and factor models: a new methodology using Normal Inverse Gaussian distributions, *Research Report IDHE-MORA, ENS Cachan* **7** (2005).

[Gumbel (1960)] E.J. Gumbel, Bivariate exponential distributions, *Journal of the American Statistical Association* **55:292** (1960) pp. 698–707.

[Gupta, Nadarajah (2004)] A.K. Gupta and S. Nadarajah, Handbook of Beta distribution and its applications, *Marcel Dekker* (2004).

[Hausdorff (1921)] F. Hausdorff, Summationsmethoden und Momentfolgen I, *Mathematische Zeitschrift* **9:3–4** (1921) pp. 74–109.

[Hausdorff (1923)] F. Hausdorff, Momentenproblem für ein endliches Intervall, *Mathematische Zeitschrift* **16** (1923) pp. 220–248.

[Hering et al. (2010)] C. Hering, M. Hofert, J.-F. Mai and M. Scherer, Constructing hierarchical Archimedean copulas with Lévy subordinators, *Journal of Multivariate Analysis* **101:6** (2010) pp. 1428–1433.

[Hofert (2007)] M. Hofert, Sampling Archimedean copulas, *Computational Statistics and Data Analysis* **52:12** (2008) pp. 5163–5174.

[Hofert, Scherer (2009)] M. Hofert and M. Scherer, CDO pricing with nested Archimedean copulas, *Quantitative Finance (in press)* (2009).

[Huff (1969)] B.W. Huff, The strict subordination of a differential process, *Sankhya: The Indian Journal of Statistics, Series A* **31:4** (1969) 403–412.

[Hull, White (2004)] J. Hull and A. White, Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation, *Journal of Derivatives* **12:2** (2004) 8–23.

[Jarrow, Turnbull (1995)] R.A. Jarrow and S. Turnbull, Pricing derivatives on financial securities subject to credit risk, *Journal of Finance* **50:1** (1995) 53–85.

[Joe (1997)] H. Joe, Multivariate models and dependence concepts, *Chapman and Hall/CRC* (1997).

[Joshi, Kainth (2004)] M.S. Joshi and D. Kainth, Rapid and accurate development of prices and Greeks for n-th to default credit swaps in the Li model, *Quantitative Finance* **4:3** (2004) pp. 266–275.

[Joshi, Stacey (2006)] M.S. Joshi and A.M. Stacey, Intensity Gamma: a new approach to pricing portfolio credit derivatives, *Risk* **19:7** (2006) 78–83.

[Kalemanova et al. (2007)] A. Kalemanova, B. Schmid and R. Werner, The Normal Inverse Gaussian distribution for synthetic CDO pricing, *Journal of Derivatives* **14:3** (2007) pp. 80–93.

[Karlin, Shapley (1953)] S. Karlin and L.S. Shapley, Geometry of moment spaces, *Memoirs of the American Mathematical Society* **12** (1953).

[Karlin (1968)] S. Karlin, Total positivity, *Stanford University Press* (1968).

[Khinchin (1937)] A. Khinchin, Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze, *Matematicheskii Sbornik* **44:1** (1937) pp. 79–119.

[Khinchin (1938)] A. Khinchin, Limit laws for sums of independent random variables, *ONTI, Moscow-Leningrad* (1938).

[Kimberling (1974)] C.H. Kimberling, A probabilistic interpretation of complete monotonicity, *Aequationes Mathematicae* **10** (1974) pp. 152–164.

[Lando (1998)] D. Lando, On Cox processes and credit risky securities, *Review of Derivatives Research* **2:2–3** (1998) pp. 99-120.

[Laurent, Gregory (2005)] J.-P. Laurent and J. Gregory, Basket default swaps, CDOs and factor copulas, *Journal of Risk* **7:4** (2005).

[Lévy (1934)] P. Lévy, Théorie de l'addition des variables aléatoires, second edition (first edition 1934), *Gauthier-Villars, Paris* (1954).

[Li (2000)] D.X. Li, On default correlation: a copula function approach, *Journal of Fixed Income* **9:4** (2000) 43–54.

[Li (2008)] H. Li, Tail dependence comparison of survival Marshall-Olkin copulas, *Methodology and Computing in Applied Probability* **10:1** (2008) pp. 39–54.

[Li (2009)] H. Li, Orthant tail dependence of multivariate extreme value distributions, *Journal of Multivariate Analysis* **100:1** (2009) pp. 243–256.

[Lindskog, McNeil (2003)] F. Lindskog and A.J. McNeil, Common Poisson shock models: applications to insurance and credit risk modelling, *Astin Bulletin* **33:2** (2003) pp. 209–238.

[Loève (1977)] M. Loève, Probability theory I, fourth edition, *Springer-Verlag* (1977).

[Lorch, Newman (1983)] L. Lorch and D.J. Newman, On the composition of completely monotonic functions and completely monotonic sequences and related questions, *Journal of the London Mathematical Society* **s2-28:1** (1983) pp. 31–45.

[Lucas (1995)] D.J. Lucas, Default correlation and credit analysis, *Journal of Fixed Income* **4:4** (1995) pp. 76–87.

[Madan, Unal (1998)] D.B. Madan and H. Unal, Pricing the risks of default, *Review of Derivatives Research* **2:2–3** (1998) pp. 121–160.

[Mai, Scherer (2009a)] J.-F. Mai and M. Scherer, Lévy-frailty copulas, *Journal of Multivariate Analysis* **100:7** (2009) pp. 1567–1585.

[Mai, Scherer (2009b)] J.-F. Mai and M. Scherer, A tractable multivariate default model based on a stochastic time-change, *International Journal of Theoretical and Applied Finance* **12:2** (2009) pp. 227–249.

[Mai, Scherer (2009c)] J.-F. Mai and M. Scherer, Pricing k-th-to-default swaps in a Lévy-time framework, *Journal of Credit Risk* **5:3** (2009) pp. 55–70.

[Mai, Scherer (2009d)] J.-F. Mai and M. Scherer, Efficiently sampling exchangeable Cuadras-Augé copulas in high dimensions, *Information Sciences* **179:17** (2009) pp. 2872–2877.

[Mai, Scherer (2010)] J.-F. Mai and M. Scherer, Reparameterizing Marshall-Olkin copulas with applications to sampling, *Journal of Statistical Computation and Simulation (in press)* (2010).

[Marshall, Olkin (1967)] A.W. Marshall and I. Olkin, A multivariate exponential distribution, *Journal of the American Statistical Association* **62:317** (1967) pp. 30–44.

[Marshall, Olkin (1988)] A.W. Marshall and I. Olkin, Families of multivariate distributions, *Journal of the American Statistical Association* **83:403** (1988) pp. 834–841.

[Matsumoto, Nishimura (1998)] M. Matsumoto and T. Nishimura, Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, *ACM Transactions on Modeling and Computer Simulation* **8:1** (1998) pp. 3–30.

[McNeil et al. (2005)] A.J. McNeil, R. Frey and P. Embrechts, Quantitative risk management, *Princeton University Press* (2005).

[McNeil, Nešlehová (2009)] A.J. McNeil and J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions, *Annals of Statistics* **37:5B** (2009) pp. 3059–3097.

[McNeil (2008)] A.J. McNeil, Sampling nested Archimedean copulas, *Journal of Statistical Computation and Simulation* **78:6** (2008) pp. 567–581.

[Merton (1974)] R.C. Merton, On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance* **29:2** (1974) pp. 449–470.

[Moosbrucker (2006)] T. Moosbrucker, Pricing CDOs with correlated Variance Gamma distributions, *Research Report, Department of Banking, University of Cologne* (2006).

[Müntz (1914)] C.H. Müntz, Über den Approximationssatz von Weierstrass, *Festschrift H. A. Schwarz* (1914) pp. 303–312.

[Nelsen (1999)] R.B. Nelsen, An introduction to copulas, *Springer-Verlag* (1999).

[Nolan (1997)] J.P. Nolan, Numerical calculation of stable densities and distribution functions, *Communications in Statistics–Stochastic Models* **13:4** (1997) pp. 759–774.

[Oakes (1989)] D. Oakes, Bivariate survival models induced by frailties, *Journal of the American Statistical Association* **84:406** (1989) pp. 487–493.

[Ocana, Ruiz-Rivas (1990)] J. Ocana and C. Ruiz-Rivas, Computer generation and estimation in a one-parameter system of bivariate distributions with specified marginals, *Communications in Statistics - Simulation and Computation* **19:1** (1990), pp. 37–55.

[Olshen (1974)] R. Olshen, A note on exchangeable sequences, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **28:4** (1974), pp. 317–321.

[Pickands (1981)] J. Pickands, Multivariate extreme value distributions, *Proceedings of the 43rd Session ISI, Buenos Aires* (1981) pp. 859–878.

[Resnick (1987)] S. I. Resnick, Extreme values, regular variation and point processes, *Springer-Verlag* (1987).

[Rudin (1998)] W. Rudin, Analysis (deutsche Ausgabe), *R. Oldenbourg Verlag* (1998).

[Ruiz-Rivas, Cuadras (1988)] C. Ruiz-Rivas and C.M. Cuadras, Inference properties of a one-parameter curved exponential family of distributions with given marginals, *Journal of Multivariate Analysis* **27:2** (1988), pp. 447–456.

[Sato (1999)] K.-I. Sato, Lévy processes and infinitely divisible distributions, *Cambridge University Press* (1999).

[Schmid, Schmidt (2006)] F. Schmid and R. Schmidt, Bootstrapping Spearman's Multivariate Rho, *Proceedings in Computational Statistics, edited by A. Rizzi and M. Vichi* (2006) pp. 759–766.

[Schmid, Schmidt (2007a)] F. Schmid and R. Schmidt, Multivariate conditional versions of Spearman's Rho and related measures of tail dependence, *Journal of Multivariate Analysis* **98:6** (2007) pp. 1123–1140.

[Schmid, Schmidt (2007b)] F. Schmid and R. Schmidt, Nonparametric inference on multivariate versions of Blomqvist's Beta and related measures of tail dependence, *Metrika* **66:3** (2007) pp. 323–354.

[Schönbucher, Schubert (2001)] P.J. Schönbucher and D. Schubert, Copula-dependent risk in intensity models, *Working paper, retrievable from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=301968* (2001).

[Schönbucher (2002)] P.J. Schönbucher, Taken to the limit: simple and not-so-simple loan loss distributions, *Working paper, retrievable from http://www.gloriamundi.org/picsresources/pjs.pdf* (2002).

[Schoutens (2003)] W. Schoutens, Lévy processes in Finance: pricing financial derivatives, *Wiley Series in Probability and Statistics* (2003).

[Schweizer, Sklar (1983)] B. Schweizer and A. Sklar, Probabilistic metric spaces, *New York: North-Holland/ Elsevier* (1983).

[Sedgewick (1988)] R. Sedgewick, Algorithms, second edition, *Addison-Wesley Publishing Company, Second Edition* (1988).

[Shuster (1968)] J. Shuster, On the Inverse Gaussian distribution, *Journal of the American Statistical Association* **63:324** (1968) pp. 1514–1516.

[Sklar (1959)] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de L'Université de Paris* **8** (1959) pp. 229–231.

[Szász (1916)] O. Szász, Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen, *Mathematische Annalen* **77:4** (1916) pp. 482–496.

[Tricomi (1951)] F.G. Tricomi, On the theorem of Frullani, *American Mathematical Monthly* **58:3** (1951) pp. 158–164.

[Vasicek (1987)] O. Vasicek, Probability of loss on loan portfolio, *Working Paper, KMV Corporation* (1987).

[Widder (1946)] D.V. Widder, The Laplace transform, *Princeton University Press* (1946).

[Willemann (2007)] S. Willemann, Fitting the CDO correlation skew: a tractable structural jump-diffusion model, *Journal of Credit Risk* **3:1** (2007) pp. 63–90.

[Williamson (1956)] R.E. Williamson, Multiply monotone functions and their Laplace transforms, *Duke Mathematical Journal* **23:2** (1956) pp. 189–207.

[Wolfe (1975)] S.J. Wolfe, On moments of probability distribution functions, *Lecture Notes in Mathematics (volume 457), Springer-Verlag* (1975) pp. 306–316.

[Wolff (1980)] E.F. Wolff, N-dimensional measures of dependence, *Stochastica* **4:3** (1980) pp. 175–188.

[Yu (2007)] F. Yu, Correlated defaults in intensity-based models, *Mathematical Finance* **17:2** (2007) pp. 155–173.

[Zhou (2001)] C. Zhou, The term structure of credit spreads with jump risk, *Journal of Banking and Finance* **25:11** (2001) pp. 2015–2040.

Index

α-stable subordinator, 51
d-monotone
    function, 38, 41
    sequence, 68, 78
i-margin, 15
n-th-to-default swap, 137

Archimedean copula, 37, 84
Archimedean generator, 37

basket credit derivative, 136
Bernoulli mixture model, 155
Bernstein's Theorem, 38
beta distribution, 103
bond, 135
Bundesschatzbrief, 135

CDO, 137, 164
    tranche, 138
class eMO, 64, 75, 78
Clayton copula, 41
comonotonicity copula, 15
completely alternating sequence, 57
completely monotone, 84
    function, 38, 57
    sequence, 55, 83
compound Poisson process, 47, 128
copula
    definition, 15
coupon, 135
Cuadras-Augé copula, 100, 119, 122
    bivariate, 17, 28, 82, 119, 126
cumulative hazard function, 149

default indicator process, 136, 149
defaultable, 135
difference operator, 54
distribution function, 14
doubly-stochastic model, 143
drift, 45

factor-copula model, 145
finite activity, 48, 154

Gamma subordinator, 48, 102
Gaussian copula, 20
generalized inverse, 22

Hausdorff's Theorem, 56, 83

independence copula, 15
infinite activity, 49, 154
infinitely divisible, 44, 47
intensity-based model, 143
inverse Gaussian subordinator, 49, 103

Lévy measure, 45
Lévy subordinator

classical, 43
killed, 44
Lévy-frailty copula, 94, 149
hierarchical, 125
Lévy-Khinchin Theorem, 45
Laplace exponent, 45, 58, 94
Laplace transform, 38
loss given default, 136
lower-tail dependence, 27

Marshall-Olkin copula, 33
moment problem, 53

Pickands representation, 31, 107
Poisson process, 122
construction, 43
intensity, 43
portfolio CDS, 137
portfolio credit derivative, 136
portfolio loss process, 138, 157

recovery rate, 136
reduced-form model, 143

sampling, 117
scatterplot, 18, 129
singular component, 21, 100, 111
Spearman's rho, 30, 109
spread, 138
structural model, 142
survival copula, 25
survival function, 22, 149

upfront payment, 140
upper-extremal dependence, 29, 108
upper-Fréchet-Hoeffding bound, 15
upper-tail dependence, 27, 108, 126