Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Finanzmathematik

Exchangeable exogenous shock models

Steffen Schenk

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.

Vorsitzender:		Univ.-Prof. Dr. Daniel Matthes, Technische Universität München
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Matthias Scherer, Technische Universität München
	2.	Prof. Dr. Fabrizio Durante, Freie Universität Bozen, Italien Nur schriftliche Beurteilung
	3.	Univ.-Prof. Dr. Alfred Müller, Universität Siegen

Die Dissertation wurde am 15.10.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 05.02.2016 angenommen.

Abstract

The present thesis provides a comprehensive analysis of exchangeable exogenous shock models. This model class is fully characterized in terms of its survival function, which admits a surprisingly simple and elegant structure. A subclass of exchangeable exogenous shock models based on a new stochastic representation is derived. It is parameterized by a single stochastic process and is used to both construct new families of multivariate distribution functions and characterize the respective process via the general shock model conditions. From a theoretical point of view, the latter characterization results are used to identify a subclass of infinitely divisible probability distributions in a novel way. From a practical perspective, exchangeable exogenous shock models are applied in the definition of new portfolio default models, the pricing of multiname credit derivatives, and the embedding of model uncertainty in risk management applications.

Zusammenfassung

Die vorliegende Arbeit beinhaltet eine umfassende Untersuchung austauschbarer exogener Schockmodelle. Diese Modellklasse wird vollständig anhand ihrer Überlebensfunktion charakterisiert, welche eine überraschend einfache und elegante Struktur aufweist. Eine auf einer neuen, alternativen stochastischen Darstellung basierende Unterklasse austauschbarer exogener Schockmodelle wird hergeleitet. Diese wird durch einen einzelnen stochastischen Prozess parametrisiert und dazu verwendet, sowohl neue multivariate Verteilungsfunktionen zu konstruieren als auch den verwendeten Prozess mittels der allgemeinen Schockmodellbedingungen zu charakterisieren. In theoretischer Hinsicht werden letztere Charakterisierungsergebnisse zur Identifikation einer Unterklasse unendlich teilbarer Wahrscheinlichkeitsverteilungen verwendet. Aus praktischer Perspektive werden austauschbare exogene Schockmodelle für die Definition neuer Portfolioausfallmodelle, die Bewertung mehrdimensionaler Kreditderivate und die Einbettung von Modellunsicherheit in Risikomanagementfragestellungen herangezogen.

Acknowledgements

To begin with, I want to give thanks to my supervisor Matthias Scherer. He not only drew my attention to doing a doctorate in the first place, but provided an absorbing research area it was inspiring to work in during the past few years. Be it operational involvement in the projects, valuable feedback on my elaborations, incitations for further investigations, or the encouragement and freedom to pursue own ideas – I always felt great support, interest, and confidence in what I was doing.

The same gratitude is due to Jan-Frederik Mai who assisted as an informal advisor in all projects and substantially contributed to the achieved results by virtue of his inquisitiveness, ingenuity, and passion. With his Ph.D. thesis serving as a fundamental starting point for my research activities, he truly can be considered the giant on the shoulders of which I was standing on (needless to say he is actually quite huge in person).

In addition, I want to thank Alfred Müller and Fabrizio Durante for functioning as referees for the present thesis. I still remember a conference in Munich in September 2013 where Fabrizio was attending my talk about exchangeable exogenous shock models. After the presentation, he gave me one of his papers which dealt with bivariate shock models and which, as it turned out, stimulated my progress in the sequel considerably.

The conference in question was hosted by the "KPMG Center of Excellence in Risk Management", which also funded my employment at university. Besides the financial support, the organization of various events at the Chair of Financial Mathematics, and the autonomy in conducting my research, I am explicitly grateful to Franz Lorenz, Matthias Mayer, and Daniel Sommer for their outstanding personal commitment and the great collaboration. The cooperation would also not have been possible without chairholder Rudi Zagst, who I want to thank for the constructive and harmonic working atmosphere and the possibility to present my results at international, prestigious conferences.

The great work climate was primarily induced by the whole team and my colleagues at the chair. I want to thank my room mate German Bernhart for several profound and numerous non-serious discussions, Thorsten Schulz for being the most challenging table soccer opponent, Daniela Neykova for processing the homegrown vegetables I didn't manage to eat, and all others for making the doctorate a great and unique experience.

Last but not least, I want to thank my parents for the enduring support, their encouragement to earn the Ph.D. degree and their honest confession that they still have no clue what I have precisely been working on during the past few years.

Contents

1 Introduction

1.1 Motivation

This thesis is located in the field of high-dimensional dependence modeling. Given d dependent random variables $X_1, \ldots, X_d$, the overall aim is to describe their mutual interaction in a "reasonable" way. As a matter of course, reasonability is strongly context-sensitive, and the judgment on a model's adequacy is influenced by many aspects. In general, being an abstraction of reality by definition, any model has to deal with the conflicting objectives of covering stylized empirical facts while maintaining mathematical tractability. In the present manuscript, this area of tension is analyzed within the universe of *exogenous shock models*, i.e. for random vectors $(X_1, \ldots, X_d)$ that can be represented as

$$X_k = \min\{Z^E : k \in E\}, \quad k = 1, \ldots, d, \tag{1.1}$$

for $2^d - 1$ independent real-valued random variables Z^E, $\emptyset \neq E \subseteq \{1, \ldots, d\}$.

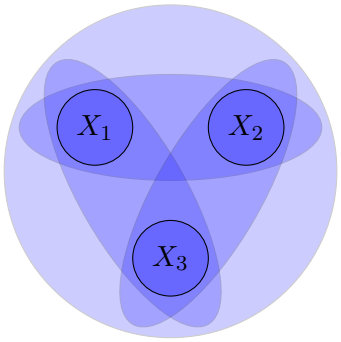

Figure 1.1 Visualization of a three-dimensional exogenous shock model. The blue areas depict seven shocks to the non-empty subsets of the trivariate system (X_1, X_2, X_3).

The construction principle of exogenous shock models is visualized in Figure 1.1 for the three-dimensional case. The key assumption is that in a system of d components, each subset $\emptyset \neq E \subseteq \{1, \ldots, d\}$ is exposed to an independent source of risk Z^E. In the given example, there are 7 shocks, affecting one or several constituents and consisting of idiosyncratic risk factors $Z^{\{1\}}, Z^{\{2\}}, Z^{\{3\}}$ (dark blue inner circles), twofold sources of risk $Z^{\{1,2\}}, Z^{\{1,3\}}, Z^{\{2,3\}}$ (blue ellipses), and a cataclysmic event $Z^{\{1,2,3\}}$ (light blue outer circle). The random variable X_k represents the first time component k is hit by a shock and is thus defined as the minimum of all Z^E with $k \in E$.

By construction, exogenous shock models provide a high degree of flexibility in the sense that interactions between subsets of $\{X_1, \ldots, X_d\}$ can be managed by the respective random variable Z^E. When laying the focus on high-dimensional random vectors, however, the wide range of possible dependence patterns is rather cumbersome. This thesis is motivated by applications in insurance and portfolio credit risk modeling, where $X_k, k = 1, \ldots, d$, might represent, for instance, the lifetimes of policyholders or financial assets. In this context, d is typically a large number ($d = 125$ for multiname credit derivatives, $d \gg 100$ for insurance portfolios), such that calibrating or simulating the $2^d - 1$ shocks Z^E is not practical. For that reason, the present work deals with the subclass of *exchangeable* exogenous shock models. Exchangeability refers to a homogeneity assumption for the random vector $(X_1, \ldots, X_d)$ and, roughly speaking, implies that shocks Z^E that affect the same number of constituents are equal in distribution. In the example in Figure 1.1, this means that $Z^{\{1\}}, Z^{\{2\}}, Z^{\{3\}}$ and $Z^{\{1,2\}}, Z^{\{1,3\}}, Z^{\{2,3\}}$ have the same distribution function, respectively, which essentially reduces the originally seven degrees of freedom to the specification of Z^E for the three possible cardinalities $|E| \in \{1, 2, 3\}$.

Within the class of exchangeable exogenous shock models, this thesis makes a contribution in both theoretical and practical regard. The results can be structured as follows. Firstly, exchangeable exogenous shock models are completely characterized with respect to the survival function of the corresponding random vector $(X_1, \ldots, X_d)$. While it is straightforward to show that the survival function of $(X_1, \ldots, X_d)$ as constructed in Equation (1.1) has a "nice" functional form, it is challenging to prove that this functional form represents a survival function *only if* it represents the survival function of an exchangeable exogenous shock model. This characterization generalizes, extends, and unifies many investigations in the literature that, for the stated reasons of complexity reduction, either solely deal with the bivariate case (i.e. $d = 2$) or specific distributions for the shocks Z^E. On top of that, an alternative stochastic model for a subclass of exchangeable exogenous shock models is presented. The alternative setup consists of a

first-passage time framework for $X_1, \ldots, X_d$ which is parameterized by a single increasing stochastic process $\Lambda = \{\Lambda_t\}_{t \geq 0}$ with independent increments. More precisely,

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d, \tag{1.2}$$

where $E_1, \ldots, E_d$ are independent and identically distributed (i.i.d.) random variables with standard exponential law. On a high level, this distinct perspective on shock models relates the independence of the increments of Λ to the independence of the random variables Z^E in the minimum construction (1.1). Besides, it is obvious that (1.2) not only provides a framework for a fixed set $(X_1, \ldots, X_d)$ of d random variables, but can be raised to higher dimensions, i.e. $X_{d+1}, X_{d+2}, \ldots$, by simply adding further i.i.d. components $E_{d+1}, E_{d+2}, \ldots$ to the construction. Exchangeable random vectors respectively exchangeable exogenous shock models that can be extended in this way are called *extendible*.

What to do with these results? Apart from drawing a line under the seemingly delimited world of exchangeable exogenous shock models, it turns out that these findings open the door for further interesting characterizations that at a first glance have nothing to do with the starting point in Equation (1.1) or random vectors at all. Abstractly speaking, the proceeding works as follows: Leading to an exchangeable exogenous shock model, the stochastic process Λ in the alternative construction (1.2) inevitably has to satisfy the necessary and sufficient conditions derived for the survival function of the general model in (1.1). Possibly, the reverse reasoning works as well, and these conditions may even be used to characterize the stochastic process in concern. This attempt is pursued in the second major outcome of the present manuscript, which emphasizes the manifoldness of the approach and concerns the set of *self-decomposable* probability measures on $[0, \infty)$. These measures correspond to a subset of *infinitely divisible* distributions on the compactified positive half-line, which in turn can be described via *Bernstein functions*. Bernstein functions are used in diverse areas of mathematical application and consist of a linear part (described by constants a, b) and an integral part with respect to a certain measure ν on $(0, \infty)$. Two novel characterizations of self-decomposable Bernstein functions are derived, providing interesting insights into the peculiarity of this subclass. While existing results mainly focus on the shape of the measure ν, the present work analyzes self-decomposable Bernstein functions one layer above in terms of monotonicity of the whole function. In addition to that, self-decomposable probability measures on $[0, \infty)$ are characterized by a specific class of multivariate distribution functions that arises from choosing a *Sato process* for Λ in Equation (1.2).

As a third major contribution of this thesis, the previous results on shock models and

self-decomposability are applied to the pricing of *collateralized debt obbilgations* (CDOs), probably the most famous representatives of multiname credit derivatives. Given a portfolio of d assets with default times $X_1, \ldots, X_d$, the central quantity for pricing these products are call respectively put options on the portfolio loss process $\{L_t\}_{t\geq 0}$, $L_t := \left(\sum_{k=1}^{d} \mathbb{1}_{\{X_k \leq t\}}\right)/d$, which denotes the fraction of defaulted firms in the portfolio in time. Considering the subclass of exchangeable shock models in the first-passage time setup (1.2) for the random vector $(X_1, \ldots, X_d)$, it is pointed out how to derive an approximation formula for these option values based on *inverse Laplace transforms*. In addition to that, a fast and efficient evaluation algorithm for the resulting Laplace inversion integrals is presented, yielding model prices for CDOs within fractions of a second. A generic step-by-step instruction for model calibration to market prices of CDOs is given and illustrated by means of Sato processes for Λ.

Last but not least, the fourth highlight of this manuscript comprises another practical application of exchangeable exogenous shock models and introduces a fundamentally new, universal approach for model risk analysis. The starting point is a company's risk manager who – in order to derive risk measures for the log-return of a given portfolio with respect to (w.r.t.) a certain holding period – typically models granular log-returns $\boldsymbol{R} = \{R_{t_k}\}_{k=1,\ldots,d}$ for d short periods and deduces the requested risk measure from aggregating $R_{t_1} + \ldots + R_{t_d}$. By a classical result in probability theory, any model for $\{R_{t_k}\}_{k=1,\ldots,d}$ can be decomposed into a vector $\boldsymbol{U} := \{U_k\}_{k=1,\ldots,d}$ of i.i.d. random variables – accounting for the randomness in the model – and a d-variate function f – representing the economic reasoning behind. While existing literature mainly addresses model uncertainty by manipulating the function f, we introduce a new philosophy by distorting the source of randomness $\boldsymbol{U}$. The proposed construction involves the first-passage time framework (1.2) for the distorted random vector $\tilde{\boldsymbol{U}} := \{\tilde{U}_k\}_{k=1,\ldots,d}$ and an additive process connected to a *Dirichlet process* for Λ. It is is shown that this proceeding not only satisfies several consistency conditions for a reasonable distortion, but additionally leads to extremely fast simulation algorithms. The universal nature of the methodology is illustrated by means of a case study comparing the impact of the distortion on the aggregated return distribution $R_{t_1} + \ldots + R_{t_d}$ for popular models f.

This thesis is separated into seven chapters. Chapter 1 consists of the present introduction and a brief depiction of this work's incurrence. Chapter 2 elucidates the necessary mathematical background and establishes well-known results on copulas, additive processes, and Bernstein functions, three objects that are of central importance throughout the entire manuscript. Required knowledge on chapter-specific contents such as self-decomposability or Laplace inversion is conveyed in the respective sections themselves.

In Chapter 3, the previously mentioned characterization of exchangeable exogenous shock models is derived. The investigations result in Theorem 3.3.1 (p. 61), which can be considered a major finding of the present thesis. Furthermore, an important subclass of exchangeable shock models additionally exhibiting extendibility is presented in Proposition 3.5.1 (p. 80). The novel characterizations regarding self-decomposable probability measures on the half-line are depicted in Chapter 4 and summarized in Theorem 4.2.1 (p. 91), the second substantial contribution of this work. Among others, self-decomposability is linked to a new class of multivariate distribution functions termed *Sato–frailty copulas*. Chapter 5 focuses on the application of extendible exogenous shock models to the evaluation of CDOs. The crucial result of this part of the thesis is Theorem 5.3.2 (p. 122), which paves the way to an extremely efficient pricing algorithm. Picking up the special case of Sato–frailty copulas, the calibration to market quotes of these multiname credit derivatives is described in detail. In Chapter 6, model uncertainty is addressed in terms of dilution of the stochastic root $\boldsymbol{U} := \{U_k\}_{k=1,\ldots,d}$ of a random vector. The joint distribution function of the originally i.i.d. components in $\boldsymbol{U}$ is changed towards the newly introduced *Dirichlet copula* in Theorem 3.5.3 (p. 83). At long last, Chapter 7 concludes and gives an outlook on potential future research.

1.2 Personal note on the history of origins

The present thesis contains most of the final results I have been working on for the past few years. Yet it tells nothing about their genesis. To my mind, though finally disemboguing in the statement of several conclusions and their validation, scientific progress is typically characterized by a starting point and various directions of movement rather than by a predefined destination. Who knows where the road will end up anyway? By filling this gap and shortly commenting on the milestones of this project, I wish to value and clarify the "scientific pedestal" my work is grounded on, and to point out the new results this thesis places on top. In so doing, I hope to convey a deeper understanding of the current manuscript's structure and to arouse interest for further research on the presented ideas.

I first came across exogenous shock models in the dissertation of Mai (2010) and the related papers by Mai and Scherer (2009b, 2011). There, the authors analyze models of type (1.1) for the special case of exponentially distributed shocks Z^E. One of their main findings is a coherence between a subclass of these models, an algebraic concept called complete monotonicity, and an equivalent alternative stochastic construction for

$(X_1, \ldots, X_d)$ as in (1.2), involving an increasing Lévy process for Λ. Given the suitability of the latter construction for both theoretical deliberations and practical purposes, such as the pricing of multiname credit derivatives, my aim was to investigate this topic in more detail.

As it turned out, considering the alternative construction, yet replacing the Lévy process by a different class of stochastic processes, resulted in a very similar functional form for the joint survival function of $(X_1, \ldots, X_d)$. This observation led to the central questions of this thesis:

1. Is there a reasonable superclass of both survival functions?
2. If so, how can it be characterized and interpreted?
3. How does it relate to the alternative construction with stochastic processes?
4. Do the characterization results for the superclass and the construction with stochastic processes yield a characterization of the stochastic processes themselves?
5. Is the superclass suitable for practical applications?

All of these questions are addressed in this manuscript and the findings have been published in, respectively submitted to, peer-reviewed international journals. The considered superclass turns out to precisely describe the class of exchangeable exogenous shock models. The findings (among others, the answers to questions one to three) are summarized in [Mai, Schenk, Scherer (2015b)] and constitute Chapter 3 of this thesis. The transitive nature of the characterization results (and, thus, question four) is exploited in [Mai, Schenk, Scherer (2015c)] and revealed in Chapter 4. By transferring the necessary and sufficient conditions for the superclass to the alternative construction with Sato processes, it is possible to derive new results for self-decomposable probability measures. In order to put the superclass into play and address question five, Chapter 5 derives a powerful algorithm for efficiently pricing multiname credit derivatives. The procedure is similar to the one in [Mai, Olivares, Schenk, Scherer (2014)][1], yet additionally uses the theoretical results in [Bernhart, Mai, Schenk, Scherer (2015)] to derive a mathematically rigorous evaluation formula. Another demonstration of the manifold practical applicability of the extendible exogenous shock model construction is given in Chapter 6 in a risk management context. The presented model-independent approach for incorporating

[1] This paper builds on the Master's thesis [Schenk (2011)]. Supported by the "Deutsche Gesellschaft für Versicherungs- und Finanzmathematik" – which I profoundly want to give thanks to at this point – I have been funded for six months in order to convert and enhance the thesis' results into a scientific article.

model uncertainty is enhanced by an empirical case study and interesting mathematical byproducts and subsumed in [Mai, Schenk, Scherer (2015a)]. Supplementally to the mentioned references, compact survey articles on fast Fourier transform pricing (see [Schenk (2014)]) and portfolio credit risk modeling (see [Kant and Schenk (2015)]) have been published in a practically oriented, non-peer reviewed journal.

2 Mathematical prerequisites

This chapter recalls mathematical objects that are repeatedly used throughout this thesis. The present manuscript mainly deals with multivariate distribution functions. Thus, Section 2.2 establishes the notion of copulas and motivates their use by means of Sklar's Theorem (see Theorem 2.2.3), which states that an arbitrary d-variate distribution function can be decomposed into the respective marginal distribution functions and a standardized multivariate distribution function on $[0,1]^d$ called copula. Besides giving two equivalent formal definitions of copulas in Section 2.2.1, the chapter introduces related properties such as the concepts of exchangeability and extendibility (see Section 2.2.2), both being of central importance for the study of exchangeable respectively extendible exogenous shock models. Moreover, Section 2.2.3 introduces the most common dependence measures for bivariate copulas.

The second major building block of this manuscript are d-monotone sequences and functions and their relationship to certain copula families. Given a sequence (respectively sufficiently smooth function), these algebraic concepts are introduced in Section 2.3.1 and refer to a homogeneity property of the sequence's finite differences (respectively the function's derivatives). For some popular copula families being parameterized by a sequence respectively single function – among others the Archimedean copulas presented in Section 2.3.2 and the exchangeable Marshall–Olkin copulas depicted in Section 2.3.3 – it is known that the d-monotonicity conditions represent necessary and sufficient requirements for well-definedness. A similar relationship can be derived for the extendible exogenous shock models in Chapter 4, which is why we review this framework in more detail.

Functions that are d-monotone for any natural number d are called *completely monotone*. Completely monotone functions are closely connected to Bernstein functions (see Bernstein (1929)), which are defined in Section 2.4. Being linked to infinitely divisible distributions on the positive half-axis, Bernstein functions provide a natural transition to increasing additive processes, which are captured in Section 2.4 as well. Additive processes are stochastic processes that are essentially characterized by the property of

independent increments. The most prominent example are Lévy processes (see Section 2.4.2), which additionally feature stationarity, i.e. equality in distribution of increments corresponding to periods of equal length. A less common, yet crucial class for Chapters 4 and 6 of this thesis are Sato processes and transformed Dirichlet processes as introduced in Sections 2.4.3 and 2.4.4. Furthermore, Section 2.4.5 indicates several examples of Bernstein function that are useful for later chapters.

2.1 Notations and definitions

Table 2.1 depicts some of the symbols and abbreviations which are used throughout this thesis without further explanation. In addition to that, the following conventions are established:

- **Equality in law**: Equality in law between two random vectors $\boldsymbol{X} = (X_1, \dots, X_d)$ and $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ is denoted by $\boldsymbol{X} \overset{d}{=} \boldsymbol{Y}$. We recall that equality in law means $\mathbb{E}[f(X_1, \dots, X_d)] = \mathbb{E}[f(Y_1, \dots, Y_d)]$ for all bounded, continuous functions $f : \mathbb{R}^d \to \mathbb{R}$, where the expectation values $\mathbb{E}$ are taken on the respective probability spaces of $(X_1, \dots, X_d)$ and $(Y_1, \dots, Y_d)$, which might be different. Equality in law $X \overset{d}{=} Y$ for two stochastic processes $X = \{X_t\}_{t\in\mathbb{R}}$ and $Y = \{Y_t\}_{t\in\mathbb{R}}$ means that $(X_{t_1}, \dots, X_{t_d}) \overset{d}{=} (Y_{t_1}, \dots, Y_{t_d})$ for arbitrary $d \in \mathbb{N}$ and $t_1, t_2, \dots, t_d \in \mathbb{R}$.

- **Generalized inverse**: We use the definition in Embrechts and Hofert (2013). For an increasing function $f : \mathbb{R} \to \mathbb{R}$, the generalized inverse $f^{-1} : \mathbb{R} \to [-\infty, \infty]$ is given by

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

  where $\inf \emptyset := \infty$. Analogously, for a decreasing function f, we denote

$$f^{-1}(y) := \sup\{x \in \mathbb{R} : f(x) \geq y\}, \quad y \in \mathbb{R},$$

  and set $\sup \emptyset := -\infty$.

- **Laplace transform**: The Laplace transform of a non-negative function f, probability measure π, respectively random variable X is denoted by $\mathcal{L}$, meaning that

Label	Symbol	Description
Cardinality	$\lvert . \rvert$	For a set A, $\lvert A \rvert$ denotes the number of elements in A.
Cartesian product	$\times$	For two sets A and B, $A \times B := \{(a, b) : a \in A, b \in B\}$.
Derivative	$^{(k)}$	For $k \in \mathbb{N}_0$ and a function f, $f^{(k)}$ denotes the k-th derivative of $f : \mathbb{R} \mapsto \mathbb{R}$; $f^{(1)}$ is sometimes abbreviated as f'.
Expected value	$\mathbb{E}$	$\mathbb{E}[X]$ denotes the expected value of the random variable X on the respective probability space.
Exponential distribution	$\text{Exp}(\lambda)$	$X \sim \text{Exp}(\lambda)$ for a random variable X and a parameter $\lambda > 0$ means that X is exponentially distributed with parameter λ, i.e. $\mathbb{P}(X > x) = \exp(-\lambda\, x)$ for $x \geq 0$.
Gamma function	Γ	For $x > 0$, the Gamma function Γ is defined as $\Gamma(x) := \int_0^\infty t^{x-1} \exp(-t)\, \mathrm{d}t$.
Identity on $[0, 1]$	$\text{id}_{[0,1]}$	Identity mapping on $[0, 1]$, i.e. $\text{id}_{[0,1]}(u) = u, u \in [0, 1]$.
Law of random vector	$\sim$	Writing $\mathbf{X} = (X_1, \ldots, X_d) \sim F$ for a d-dimensional random vector $\mathbf{X}$ means that the distribution function of $\mathbf{X}$ is given by F.
Number sets	$\mathbb{N}\,(\mathbb{N}_0)$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$	Set of natural (extended by $\{0\}$), rational, real, and complex numbers, respectively.
Poisson distribution	$\text{Poi}(\lambda)$	$X \sim \text{Poi}(\lambda)$ for a random variable X and a parameter $\lambda > 0$ means that X is Poisson distributed with probability mass function $\mathbb{P}(X = n) = \lambda^n / n! \exp(-\lambda), n \in \mathbb{N}_0$.
Probability space	$(\Omega, \mathcal{F}, \mathbb{P})$	Ω denotes the event space, $\mathcal{F}$ the σ-algebra, and $\mathbb{P}$ the probability measure.
Subset	$\subset, \subseteq$	For two sets A and B, $A \subset B$ means that A is a strict subset of B. $A \subseteq B$ additionally allows A to equal B.
Uniform distribution	$\mathcal{U}([a, b])$	$X \sim \mathcal{U}([a, b])$ for a random variable X and $a, b \in \mathbb{R}, a < b$, means that X is uniformly distributed on $[a, b]$, i.e. $\mathbb{P}(X \leq x) = (x-a)/(b-a), x \in [a, b]$.

Table 2.1 Abbreviations/symbols that are repeatedly used in this thesis.

for $z \in \mathbb{C}$,

$$\begin{aligned}\mathcal{L}[f](z) &:= \int_0^\infty e^{-z\,s}\, f(s)\,\mathrm{d}s,\\ \mathcal{L}[\pi](z) &:= \int_0^\infty e^{-z\,s}\,\pi(\mathrm{d}s),\\ \mathcal{L}[X](z) &:= \mathbb{E}[e^{-z\,X}],\end{aligned}$$

provided the respective integrals exist. Consequently, for the probability measure π corresponding to X, the expressions $\mathcal{L}[\pi]$ and $\mathcal{L}[X]$ are used substitutively. The same equivalency holds for $\mathcal{L}[f]$ if π possesses a density f. For certain non-negative probability measures π respectively random variables X, it is convenient to characterize the Laplace transform via its *Laplace exponent* Ψ, which is defined as

$$\Psi(z) := -\log\Big(\mathcal{L}[\pi](z)\Big) = -\log\Big(\mathcal{L}[X](z)\Big), \quad z \in \mathbb{C},$$

provided the respective expression exists.

- **Margins/Marginal distributions**: For a random vector $\mathbf{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d, \mathbf{X} \sim F$, with multivariate distribution function $F : \mathbb{R}^d \to \mathbb{R}$, the univariate distribution functions of $X_1, \ldots, X_d$ are called *margins* or *marginal distributions* of $\mathbf{X}$ and are denoted by $F_1, \ldots, F_d$. The same reasoning applies to the *marginal survival functions* of $\mathbf{X}$, which are typically written as $\bar{F}_1, \ldots, \bar{F}_d$.

- **Monotonicity**: Whenever talking about an *increasing* function f, we refer to the function as being non-decreasing, i.e.

$$f(y) - f(x) \geq 0 \text{ for all } x \leq y \text{ in the domain of } f.$$

If the difference $f(y) - f(x), x < y$, is always greater than zero, we call f *strictly increasing*. The same logic is applied to (strictly) decreasing functions.

- **Order relations**: Order relations between two vectors are to be understood componentwise, i.e. for $\mathbf{x} := (x_1, \ldots, x_d), \mathbf{y} := (y_1, \ldots, y_d) \in \mathbb{R}^d$, $\mathbf{x} \leq \mathbf{y}$ means $x_k \leq y_k$ for all $k = 1, \ldots, d$. Furthermore, we denote by $x_{(1)} \leq \ldots \leq x_{(d)}$ the ordered list of $(x_1, \ldots, x_d)$.

- **Random variable**: When referring to a random variable X without further supplement, we always assume that X is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and is a real-valued, measurable mapping $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, with $\mathcal{B}(\mathbb{R})$ denoting the Borel σ-algebra on $\mathbb{R}$. For subsets $A \subset \mathbb{R}^d$, $\mathcal{B}(A)$ denotes the respective restriction of the Borel σ-algebra on $\mathbb{R}^d$ to A.

2.2 Copulas

This section introduces copulas and demonstrates their suitability for the study of multivariate distribution functions. A broad overview of copulas and their origin is given in Durante and Sempi (2010). The subsequent review is far from complete, however introduces the basic concepts that are relevant for the present thesis. Among others, it addresses questions such as "What are copulas?", "Which classes can they be divided into?", "Which properties are there to analyze?", or "How do examples look like?", all of which prove helpful for the analysis of exogenous shock models.

2.2.1 Probabilistic and analytic definition

When having a look at standard literature on copulas (e.g. Nelsen (2006); Joe (1997); McNeil et al. (2005)), one is likely to find two different, yet equivalent notions. The first one is a rather analytic perspective and introduces copulas as mappings from the unit cube to the unit interval fulfilling certain technical conditions. The most crucial one of those conditions (in the sense that it is the most difficult one to check for a given mapping) is related to *d-boxes*.

Definition 2.2.1 (d-box)
The Cartesian product $I_1 \times \ldots \times I_d$ of finite sets $I_k = [x_k, y_k] \in \mathbb{R}, -\infty < x_k \leq y_k < \infty$, $k = 1, \ldots, d$, is called d-box. If $x_k = y_k$ for at least one $k \in \{1, \ldots, d\}$, it is called degenerated. For vectors $\boldsymbol{x} := (x_1, \ldots, x_d), \boldsymbol{y} := (y_1, \ldots, y_d)$, $\boldsymbol{x} \leq \boldsymbol{y}$, d-boxes are abbreviated by

$$[\boldsymbol{x}, \boldsymbol{y}] := \times_{k=1}^{d} [x_k, y_k] = I_1 \times \ldots \times I_d.$$

Now we can proceed with the definition of a copula from an analytic perspective.

Definition 2.2.2 (Copula, analytic)
A function $C : [0,1]^d \to [0,1]$ is called a d-dimensional copula if C exhibits

(i) groundedness, i.e. $C(u_1, \ldots, u_d) = 0$ if $u_k = 0$ for a $k \in \{1, \ldots, d\}$,

(ii) uniform margins, i.e. $C(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k$ for all $u_k \in [0,1]$, $k \in \{1, \ldots, d\}$,

(iii) d-increasingness, i.e. for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ with $\boldsymbol{u} \leq \boldsymbol{v}$,

$$\mathrm{d}C([\boldsymbol{u}, \boldsymbol{v}]) := \sum_{(w_1, \ldots, w_d) \in \times_{i=1}^{d} \{u_i, v_i\}} (-1)^{|\{i \,:\, w_i = u_i\}|} C(w_1, \ldots, w_d) \geq 0.$$

The meaning of the respective conditions is clarified in the sequel. The importance of studying copulas in the context of dependence modeling is emphasized by Sklar's Theorem, which establishes a connection between arbitrary multivariate distribution functions and copulas.

Theorem 2.2.3 (Sklar's Theorem, 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 $(x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big). \tag{2.1}$$

If $F_1, \ldots, F_d$ are continuous, C is unique. Conversely, if C is a d-dimensional copula and $F_1, \ldots, F_d$ are univariate distribution functions, the function F defined by (2.1) is a d-dimensional distribution function with margins $F_1, \ldots, F_d$.

Proof
See (Nelsen, 2006, Theorem 2.10.9, p. 46). □

Described verbally, any multivariate distribution function can be decomposed into its margins and a copula, and, reversely, plugging arbitrary univariate distribution functions into a copula is a valid approach for constructing new multivariate distribution functions. A similar version of Sklar's Theorem is available for survival functions.

Theorem 2.2.4 (Sklar's Theorem for survival functions)
Let $\bar{F}$ be a d-dimensional survival function with marginal survival functions $\bar{F}_1, \ldots, \bar{F}_d$. Then there exists a d-dimensional copula $\hat{C}$ such that for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$\bar{F}(x_1, \ldots, x_d) = \hat{C}\big(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d)\big). \tag{2.2}$$

If $\bar{F}_1, \ldots, \bar{F}_d$ are continuous, $\hat{C}$ is unique. Conversely, if $\hat{C}$ is a d-dimensional copula and $\bar{F}_1, \ldots, \bar{F}_d$ are univariate survival functions, the function $\bar{F}$ defined by (2.2) is a d-dimensional survival function with marginal survival functions $\bar{F}_1, \ldots, \bar{F}_d$.

Proof
See (McNeil et al., 2005, p. 195-196). □

Throughout this thesis, we mostly deal with random vectors $\mathbf{X} \sim F$ having continuous marginal distribution respectively survival functions. In this case, the corresponding copulas C respectively $\hat{C}$ in (2.1) and (2.2) are unique and are termed *the* copula respectively *the* survival copula of $\mathbf{X}$.

From a stochastic point of view (as we will see in the sequel), copulas are standardized distribution functions on the unit cube. Therefore, it is important to note that both copula and survival copula of a random vector are *distribution* functions. In order to shine a light on the link between the probabilistic interpretation of copulas and Definition 2.2.2, the following well-known result on the probability of a union of sets is required.

Lemma 2.2.5 (Principle of inclusion and exclusion)
For arbitrary sets $A_1, \dots, A_n \in \mathcal{F}$, it holds that

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{\emptyset \neq I \subseteq \{1,\dots,n\}} (-1)^{|I|+1} \mathbb{P}\Big(\bigcap_{j \in I} A_j\Big).$$

Proof
See (Billingsley, 1995, p. 24). □

Now we are able to give a "stochastic" view on copulas.

Proposition 2.2.6 (Copula, probabilistic)
A function $C : [0,1]^d \to [0,1]$ is a d-dimensional copula if and only if there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random vector $(U_1, \dots, U_d)$ such that $U_k \sim \mathcal{U}[0,1]$ for all $k = 1, \dots, d$ and

$$C(u_1, \dots, u_d) = \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d), \quad u_1, \dots, u_d \in [0,1]. \tag{2.3}$$

Proof
Let $C : [0,1]^d \to [0,1]$ satisfy (2.3) for a random vector $(U_1, \dots, U_d)$ with $U_k \sim \mathcal{U}[0,1]$ for all $k = 1, \dots, d$. It is trivial to check that C exhibits groundedness and uniform margins in Definition 2.2.2. The remaining d-increasingness condition can be shown in the spirit of (Mai and Scherer, 2012, p. 8). For vectors $\boldsymbol{u} = (u_1, \dots, u_d)$, $\boldsymbol{v} = (v_1, \dots, v_d) \in [0,1]^d$, $\boldsymbol{u} \leq \boldsymbol{v}$, and a subset $\emptyset \neq I \subseteq \{1, \dots, d\}$, define

$$w_k^{(I)} := \begin{cases} u_k, & k \in I, \\ v_k, & k \notin I, \end{cases} \qquad k = 1, \dots, d.$$

It holds that

$$0 \leq \mathbb{P}\Big(\bigcap_{k=1}^{d} \{u_k \leq U_k \leq v_k\}\Big) = \mathbb{P}\Big(\Big(\bigcap_{k=1}^{d} \{U_k < v_k\}\Big) \setminus \Big(\bigcup_{k=1}^{d} \{U_k < u_k\}\Big)\Big)$$

$$
\begin{aligned}
&= \mathbb{P}\Big(\bigcap_{k=1}^{d}\{U_k < v_k\}\Big) - \mathbb{P}\Big(\Big(\bigcap_{k=1}^{d}\{U_k < v_k\}\Big) \cap \Big(\bigcup_{k=1}^{d}\{U_k < u_k\}\Big)\Big) \\
&= \mathbb{P}\Big(\bigcap_{k=1}^{d}\{U_k < v_k\}\Big) - \mathbb{P}\Big(\bigcup_{k=1}^{d}\Big(\{U_k < u_k\} \cap \bigcap_{k=1}^{d}\{U_k < v_k\}\Big)\Big) \\
&\overset{(*)}{=} \mathbb{P}\Big(\bigcap_{k=1}^{d}\{U_k < v_k\}\Big) - \sum_{\emptyset \neq I \subseteq \{1,\dots,d\}} (-1)^{|I|+1}\mathbb{P}\big(\{U_k < u_k, k \in I\} \cap \{U_k < v_k, k \notin I\}\big) \\
&= C(v_1,\dots,v_d) + \sum_{\emptyset \neq I \subseteq \{1,\dots,d\}} (-1)^{|I|} C\big(w_1^{(I)},\dots,w_d^{(I)}\big) \\
&= \sum_{(w_1,\dots,w_d) \in \times_{i=1}^{d}\{u_i,v_i\}} (-1)^{|\{i\,:\,w_i = u_i\}|}\, C(w_1,\dots,w_d),
\end{aligned}
$$

where Lemma 2.2.5 has been applied in $(*)$.

Conversely, let C *be a copula according to Definition 2.2.2. By Theorem 2.2.3, using* $F_1 = \ldots = F_d = \mathrm{id}_{[0,1]}$, C *is a* d*-dimensional distribution function on* $[0,1]^d$ *with uniform margins.* □

Interpreting the probabilistic perspective on copulas in the context of Sklar's Theorem, (Embrechts and Hofert, 2013, Proposition 3.2) shows that given a d-dimensional distribution function F with continuous margins $F_1,\dots,F_d$, the random vector $\boldsymbol{X} = (X_1,\dots,X_d) \sim F$ has the copula C if and only if C is the distribution function of $\big(F_1(X_1),\dots,F_d(X_d)\big)$. Analogously, $\bar{C}$ is the distribution function of $\big(\bar{F}_1(X_1),\dots,\bar{F}_d(X_d)\big)$ for the marginal survival functions $\bar{F}_1,\dots,\bar{F}_d$ corresponding to $\boldsymbol{X}$.

Proposition 2.2.6 implies that given $\boldsymbol{U} = (U_1,\dots,U_d) \sim C$ for a d-dimensional copula C, the distribution function of the subvector $(U_{j_1},\dots,U_{j_i})$, $i \in \{1,\dots,d\}$, $1 \le j_1 < \ldots < j_i \le d$, is a copula (denoted by $C_{j_1,\dots,j_i}$) as well. For fixed $J := \{j_1,\dots,j_i\}$, $u_{j_1},\dots,u_{j_i} \in [0,1]$, and

$$
w_k^{(J)} := \begin{cases} u_k, & k \in J, \\ 1, & k \notin J, \end{cases} \qquad k = 1,\dots,d,
$$

it is given by

$$
C_{j_1,\dots,j_i}(u_1,\dots,u_i) = C\big(w_1^{(J)},\dots,w_d^{(J)}\big). \tag{2.4}
$$

Furthermore, Proposition 2.2.6 indicates that any d-dimensional copula C induces a probability measure $\mathrm{d}C$ on $[0,1]^d$ defined by

$$
\mathrm{d}C(B) := \mathbb{P}\big((U_1,\dots,U_d) \in B\big), \quad B \in \mathcal{B}\big([0,1]^d\big),
$$

called the *probability measure associated with the copula* C. The proof of the proposition provides a vivid interpretation of the d-increasingness condition in Definition 2.2.2. It ensures that $\mathrm{d}C$ assigns non-negative mass to any d-box $[\mathbf{u}, \mathbf{v}]$ in the unit cube.

Two very simple examples of copulas are given by the *independence copula* and the *comonotonicity copula*. As the name suggests, the independence copula is the distribution function of i.i.d. random variables $U_1, \ldots, U_d$ with uniform distribution on the unit interval. It is denoted by Π and given by

$$\begin{aligned}\Pi(u_1, \ldots, u_d) &= \mathbb{P}(U_1 \leq u_1, \ldots, U_d \leq u_d) = \mathbb{P}(U_1 \leq u_1) \cdot \ldots \cdot \mathbb{P}(U_d \leq u_d) \\ &= u_1 \cdot \ldots \cdot u_d, \quad u_1, \ldots, u_d \in [0,1].\end{aligned}$$

In contrast to that, the comonotonicity copula M describes the distribution function of "perfectly dependent" uniformly distributed random variables $U_1, \ldots, U_1$ on $[0,1]$, i.e.

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

Another famous representative is the *Gaussian copula*. Denoting by $\Phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}} : \mathbb{R}^d \to [0,1], \boldsymbol{\mu} \in \mathbb{R}^d, \boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$, the multivariate normal distribution function with mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ and symmetric, positive definite covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,\ldots,d}$, the Gaussian copula is given by

$$C(u_1, \ldots, u_d) = \Phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\big(\Phi^{-1}_{\mu_1,\sigma_{11}}(u_1), \ldots, \Phi^{-1}_{\mu_d,\sigma_{dd}}(u_d)\big), \quad u_1, \ldots, u_d \in [0,1].$$

The Gaussian copula allegorically demonstrates how copulas can be deduced from multivariate distribution functions F. By Sklar's Theorem (see Equation (2.1)), if $F : \mathbb{R}^d \to [0,1]$ exhibits continuous marginal distribution functions $F_1, \ldots, F_d$, the associated unique copula C satisfies

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big).$$

Replacing x_k by $F_k^{-1}(u_k), u_k \in [0,1], k = 1, \ldots, d$, and using the fact that $F_k \circ F_k^{-1}$ is the identity for continuous F_k (see (Embrechts and Hofert, 2013, Proposition 2.3, (4))), it follows that C is given by

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

The same reasoning can be applied to the survival copula, which in case of continuous marginal survival functions $\bar{F}_1, \ldots, \bar{F}_d$ of the joint survival function $\bar{F}$ equals

$$\hat{C}(u_1, \ldots, u_d) = \bar{F}\big(\bar{F}_1^{-1}(u_1), \ldots, \bar{F}_d^{-1}(u_d)\big).$$

Having introduced copulas and survival copulas, we briefly comment on how to extract one from the other.

Proposition 2.2.7 (Coherence copula ↔ survival copula)
For a given random vector $\mathbf{U} \sim C$ *with copula* $C : [0,1]^d \to [0,1]$*, let* $\hat{C}$ *be the corresponding survival copula of* $\mathbf{U}$*. The relationship between* C *and* $\hat{C}$ *is given by*

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

for all $u_1, \ldots, u_d \in [0,1]$.

Proof
On the one side, by Equation (2.2)*, we know that*

$$\begin{aligned} \hat{C}(u_1, \ldots, u_d) &= \mathbb{P}(1 - U_1 \le u_1, \ldots, 1 - U_d \le u_d) \\ &= \mathbb{P}(1 - u_1 \le U_1 \le 1, \ldots, 1 - u_d \le U_d \le 1). \end{aligned}$$

On the other side, the proof of Proposition 2.2.6 has shown that

$$\mathbb{P}\Big(\bigcap_{k=1}^{d} \{1 - u_k \le U_k \le 1\} \Big) = \sum_{(w_1,\ldots,w_d) \in \times_{i=1}^{d} \{1-u_i, 1\}} (-1)^{|\{i \,:\, w_i = u_i\}|} \, C(w_1, \ldots, w_d).$$

Combining those two observations and recalling the definition of $C_{j_1,\ldots,j_k}$ *in* (2.4) *yields the claim.* □

Another coherence between copula and survival copula can be given via strictly decreasing transformations of the random vector in concern.

Corollary 2.2.8 (Strictly decreasing transformations)
Let $(X_1, \ldots, X_d)$ *be a random vector with continuous marginal distribution functions and copula* C*. For strictly decreasing functions* $h_k : \mathbb{R} \to \mathbb{R}$*,* $k = 1, \ldots, d$*, the copula of* $\big(h_1(X_1), \ldots, h_d(X_d)\big)$ *is the survival copula of* $(X_1, \ldots, X_d)$*.*

Proof
See (Mai and Scherer, 2012, Corollary 1.1, p. 22). □

As a very similar result, it is easy to show that for strictly increasing transformations $h_1, \ldots, h_d$, the copula of $\big(h_1(X_1), \ldots, h_d(X_d)\big)$ equals the copula of $(X_1, \ldots, X_d)$. Summing up, copulas are multivariate distribution functions on the unit cube that are subject to additional normalization constraints. Groundedness and d-increasingness stem from the general requirements of an arbitrary multivariate distribution function and can be seen as multivariate analogues of the limit and monotonicity behavior of univariate distribution functions. Property (ii) in Definition 2.2.2 is copula-specific and ensures that

in addition to the right-continuity of general multivariate distribution functions, copulas provide continuity and uniform marginal distributions[1]. One can even show (see (Schweizer and Sklar, 1983, Lemma 6.1.9, p. 82)) that a copula $C : [0,1]^d \to [0,1]$ is Lipschitz-continuous with constant one, i.e.

$$|C(v_1,\dots,v_d) - C(u_1,\dots,u_d)| \leq \sum_{k=1}^{d} |v_k - u_k|, \quad u_1,\dots,u_d,v_1,\dots,v_d \in [0,1].$$

Due to the continuous marginal distributions, the associated probability measure $\mathrm{d}C$ assigns zero mass to both single points $(u_1,\dots,u_d) \in [0,1]^d$ (as $\mathbb{P}(U_1 = u_1,\dots,U_d = u_d) = 0$) and hyperplanes of the form $\{(u_1,\dots,u_d) \in [0,1]^d : u_k = c\}$, $c \in [0,1]$, $k \in \{1,\dots,d\}$ (as $\mathbb{P}(U_k = c) = 0$). Therefore, one can deduce that the probability mass assigned to a d-box consisting of a set of *non-overlapping* d-boxes equals the sum of probability masses assigned to each individual d-box.

Definition 2.2.9 (Non-overlapping d-boxes)
Two d-boxes $[\boldsymbol{x}^{(1)}, \boldsymbol{y}^{(1)}]$, $[\boldsymbol{x}^{(2)}, \boldsymbol{y}^{(2)}]$ are called non-overlapping if $[\boldsymbol{x}^{(1)}, \boldsymbol{y}^{(1)}] \cap [\boldsymbol{x}^{(2)}, \boldsymbol{y}^{(2)}]$ is either a degenerated d-box or the empty set. A set of d-boxes $[\boldsymbol{x}^{(l)}, \boldsymbol{y}^{(l)}], l = 1,\dots,m$, is called non-overlapping if any two of the d-boxes $[\boldsymbol{x}^{(l)}, \boldsymbol{y}^{(l)}]$, $l = 1,\dots,m$, are non-overlapping.

Lemma 2.2.10 (Additivity property of associated probability measure)
Let C be a d-dimensional copula, $\mathrm{d}C$ the associated probability measure, and $[\boldsymbol{u}, \boldsymbol{v}]$ a d-box in $[0,1]^d$. For a decomposition $[\boldsymbol{u}, \boldsymbol{v}] = \cup_{l=1}^{m} [\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}]$ of $[\boldsymbol{u}, \boldsymbol{v}]$ into m non-overlapping d-boxes $[\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}], l = 1,\dots,m$, it holds that

$$\mathrm{d}C\big([\boldsymbol{u}, \boldsymbol{v}]\big) = \sum_{l=1}^{m} \mathrm{d}C\big([\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}]\big).$$

Proof
By the principle of inclusion and exclusion (see Lemma 2.2.5), we have

$$\begin{aligned}
\mathrm{d}C\big([\boldsymbol{u}, \boldsymbol{v}]\big) &= \mathrm{d}C\Big(\bigcup_{l=1}^{m} [\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}]\Big) \\
&= \sum_{l=1}^{m} \mathrm{d}C\big([\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}]\big) + \sum_{\substack{\emptyset \neq J \subseteq \{1,\dots,d\} \\ |J| \geq 2}} (-1)^{|J|+1} \mathrm{d}C\Big(\bigcap_{j \in J} [\boldsymbol{u}^{(j)}, \boldsymbol{v}^{(j)}]\Big).
\end{aligned}$$

[1] The normalization of marginals to the uniform distribution on $[0,1]$ is arbitrary to some extent, see for instance Embrechts (2009) and the reference therein for alternative specifications. As an example, early publications of Wassilij Hoeffding refer to the interval $[-0.5, 0.5]$ instead.

The non-overlapping property of the d-boxes $[\boldsymbol{u}^{(l)}, \boldsymbol{v}^{(l)}]$ *implies that* $\cap_{j\in J}[\boldsymbol{u}^{(j)}, \boldsymbol{v}^{(j)}]$ *for* $|J| \geq 2$ *is either empty or a degenerated d-box. In any case, there exists a d-box of the form* $I = [0,1] \times \ldots \times [0,1] \times [w,w] \times [0,1] \times \ldots \times [0,1]$ *with a degenerated interval* $[w,w], w \in [0,1]$*, such that*

$$\mathrm{d}C\Big(\bigcap_{j\in J}[\boldsymbol{u}^{(j)}, \boldsymbol{v}^{(j)}]\Big) \leq \mathrm{d}C(I) = \mathbb{P}(U_1 = w) = 0. \qquad \square$$

Though being easy to prove, the decomposition of a given d-box into non-overlapping d-boxes is an essential tool in the context of multivariate distribution functions. Checking whether a d-variate function C on the unit cube exhibits d-increasingness is a very difficult task in general, with complexity rising exponentially for increasing d. Therefore, it is often crucial to find a function-specific decomposition of arbitrary d-boxes that allows to reduce the analysis to simpler problems. A striking example is given in Section 3.3, where the considered class of multivariate distribution functions is characterized w.r.t. the decomposition described in Lemma 3.3.7.

As a last remark, despite of the fact that $\mathrm{d}C$ has no atoms and assigns zero mass to certain hyperplanes, it is important to note that $\mathrm{d}C$ does not have to be absolutely continuous. By Lebesgue's decomposition theorem (see (Klenke, 2006, p. 158)), $\mathrm{d}C$ can be uniquely split into a sum $\mathrm{d}C = \mathrm{d}C^{\mathrm{abs}} + \mathrm{d}C^{\mathrm{sing}}$ of measures, where $\mathrm{d}C^{\mathrm{abs}}$ is absolutely continuous w.r.t. the Lebesgue measure and $\mathrm{d}C^{\mathrm{sing}}$ is a singular measure, meaning that there exists a Lebesgue null set $B \in \mathcal{B}(\mathbb{R}^d)$ such that $\mathrm{d}C^{\mathrm{sing}}([0,1]^d \setminus B) = 0$. The absolutely continuous part can be calculated by

$$\mathrm{d}C^{\mathrm{abs}}(u_1 \ldots, u_d) = \int_0^{u_1} \ldots \int_0^{u_d} \frac{\partial^d}{\partial s_1 \ldots \partial s_d} C(s_1, \ldots s_d)\, \mathrm{d}s_1 \ldots \mathrm{d}s_d, \tag{2.5}$$

where the mixed partial derivatives of C exist almost everywhere in $[0,1]^d$ (this follows from the monotonicity of C in each argument, see also Durante et al. (2013)). If $\mathrm{d}C = \mathrm{d}C^{\mathrm{abs}}$, C is said to be absolutely continuous and the mixed partial derivative of order d in (2.5) is called the *copula density*. In contrast to that, if $\mathrm{d}C = \mathrm{d}C^{\mathrm{sing}}$, C is called *singular*. In this thesis, we mostly deal with copulas that exhibit both an absolutely continuous and a singular part.

2.2.2 Properties and classes

There are various terminologies how to classify, distinguish, or delimit the universe of copula functions. For the present thesis, we restrict ourselves to the separation of copulas

w.r.t. the following aspects. On the one hand, we consider the extent of a copula's homogeneity expressed in terms of *exchangeability* or *extendibility* of the corresponding random vector. On the other hand, we analyze these multivariate distribution functions regarding a limit condition termed *extreme-value property* and a symmetry feature called *radial symmetry*.

This thesis is devoted to *exchangeable* shock models. Exchangeability relates to a uniformity property of a random vector guaranteeing that the law of any subset of the vector's components solely depends on the size, not the constituents of the subset.

Definition 2.2.11 (Exchangeability)

- *A random vector $(X_1, \ldots, X_d)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called exchangeable if for all permutations $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$, one has*

$$(X_1, \ldots, X_d) \stackrel{d}{=} (X_{\sigma(1)}, \ldots, X_{\sigma(d)}).$$

- *An infinite sequence $\{X_k\}_{k\in\mathbb{N}}$ of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called exchangeable if for each $d \in \mathbb{N}$ and $1 \leq i_1 < i_2 < \ldots < i_d$, the random vector $(X_{i_1}, \ldots, X_{i_d})$ is exchangeable.*

- *A distribution function F is called exchangeable if it is invariant w.r.t. permutation of its arguments, i.e. if for all permutations $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$, it holds that*

$$F(x_1, \ldots, x_d) = F(x_{\sigma(1)}, \ldots, x_{\sigma(d)}), \quad x_1, \ldots, x_d \in \mathbb{R}.$$

For a deeper study of exchangeable sequences, we refer to Aldous (1985). A subclass of exchangeable distribution functions respectively random vectors which also plays a major role in the present thesis are extendible laws respectively random sequences. Extendibility demands for a continuation of the exchangeability concept to higher dimensions and is defined as follows.

Definition 2.2.12 (Extendibility)

- *An exchangeable random vector $(X_1, \ldots, X_d)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called extendible if there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ supporting an infinite exchangeable sequence $\{\tilde{X}_k\}_{k\in\mathbb{N}}$ such that $(X_1, \ldots, X_d) \stackrel{d}{=} (\tilde{X}_1, \ldots, \tilde{X}_d)$.*

- *An exchangeable distribution function F (respectively a copula C) is called extendible, if on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random vector $(X_1, \ldots, X_d) \sim F$ (respectively $(U_1, \ldots, U_d) \sim C$) is extendible.*

The crucial consequence for an extendible random vector $\{X_k\}_{k=1,\dots,d}$ is the existence of a random distribution function $\{F_t\}_{t\in\mathbb{R}}$ on $\mathbb{R}$ such that, conditioned on F, the random variables $\{X_k\}_{k=1,\dots,d}$ are i.i.d. with respective distribution function $\{F_t\}_{t\in\mathbb{R}}$. This finding goes back to De Finetti (1937).

Theorem 2.2.13 (De Finetti's Theorem)
An infinite sequence of random variables $\{X_k\}_{k\in\mathbb{N}}$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ is exchangeable (i.e. any subsequence $\{X_{i_1},\dots,X_{i_d}\}$, $1\le i_1<\dots<i_d$, is exchangeable for arbitrary $d\ge 2$) if and only if it is conditionally i.i.d., i.e. there exists a sub-σ-algebra $\mathcal{G}\subset\mathcal{F}$ such that

$$\mathbb{P}(X_1\le x_1,\dots,X_d\le x_d|\mathcal{G})=\prod_{k=1}^{d}\mathbb{P}(X_1\le x_k|\mathcal{G}),\quad x_1,\dots,x_d\in\mathbb{R}.$$

Proof
See De Finetti (1937). □

In De Finetti's Theorem above, the random distribution function $\{F_t\}_{t\in\mathbb{R}}$ is given by $F_t=\mathbb{P}(X_1\le t|\mathcal{G})$, $t\in\mathbb{R}$. A limit concept different to extendibility is provided by extreme-value copulas. An overview on their origin and usage is provided in Gudendorf and Segers (2010). Starting with a sample $\boldsymbol{X^{(i)}}=(X_1^{(i)},\dots,X_d^{(i)}), i=1,\dots,n$, of i.i.d. random vectors with joint continuous distribution function F, extreme-value copulas arise as limiting copulas of $\big(\max\{X_1^{(i)},i=1,\dots,n\},\dots,\max\{X_d^{(i)},i=1,\dots,n\}\big)$ for n approaching infinity. The resulting set of copulas can be characterized as follows.

Definition 2.2.14 (Extreme-value copula)
A copula $C:[0,1]^d\to[0,1]$ is called extreme-value copula if it satisfies

$$C(u_1,\dots,u_d)^t=C(u_1^t,\dots,u_d^t)\quad\text{for all } t>0,\quad u_1,\dots,u_d\in[0,1]. \tag{2.6}$$

In the characterization in Gudendorf and Segers (2010), the authors solely require that $C(u_1^t,\dots,u_d^t)=C(u_1,\dots,u_d)^t$ holds for $t=1/m$, $m\in\mathbb{N}$ $(*)$. If $(*)$ is valid, however, any positive rational number p/q, $p,q\in\mathbb{N}$ satisfies

$$\begin{aligned}C(u_1,\dots,u_d)^{\frac{p}{q}}&=\Big(C(u_1,\dots,u_d)^{\frac{1}{q}}\Big)^p\overset{(*)}{=}C\Big(u_1^{\frac{1}{q}},\dots,u_d^{\frac{1}{q}}\Big)^p\\&=C\Big(\big(u_1^{\frac{p}{q}}\big)^{\frac{1}{p}},\dots,\big(u_d^{\frac{p}{q}}\big)^{\frac{1}{p}}\Big)^p\overset{(*)}{=}C\Big(u_1^{\frac{p}{q}},\dots,u_d^{\frac{p}{q}}\Big).\end{aligned}$$

By continuity of C, the same relationship holds for positive irrational numbers as well such that both definitions of extreme-value copulas coincide.

Another property a given copula C can be checked for is radial symmetry. In the univariate case, a real-valued random variable $X \sim F$ is called symmetric about $\mu \in \mathbb{R}$ if $X - \mu \overset{d}{=} \mu - X$ (or equivalently $X \overset{d}{=} 2\,\mu - X$). Roughly speaking, this means that deviations of X to the left or right of μ are "equally likely". As an example, a uniformly distributed random variable U on the unit interval is symmetric about $1/2$. Lifting the symmetry concept to the multivariate level leads to the definition of radial symmetry for random vectors.

Definition 2.2.15 (Radial symmetry)

1. *A d-dimensional random vector $(X_1, \ldots, X_d)$ is called radially symmetric about the vector $(\mu_1, \ldots, \mu_d) \in \mathbb{R}^d$ if*

$$(X_1, \ldots, X_d) \overset{d}{=} (2\,\mu_1 - X_1, \ldots, 2\,\mu_d - X_d).$$

2. *A copula $C : [0,1]^d \to [0,1]$ is called radially symmetric if $\mathbf{U} \sim C$ is radially symmetric (about the vector $(1/2, \ldots, 1/2)$).*

Regarding radially symmetric copulas, two peculiarities are worth mentioning. First of all, radial symmetry of a copula C necessarily requires radial symmetry of $\mathbf{U} = (U_1, \ldots, U_d) \sim C$ about $(1/2, \ldots, 1/2)$. This is due to the fact that the uniform law on $[0,1]$, which corresponds to the marginal distribution functions of $\mathbf{U}$, is symmetric about $1/2$. Secondly, applying Corollary 2.2.8, the copula of the vector $(1 - U_1, \ldots, 1 - U_d)$, which at the same time is its distribution function due to $1 - U_k \overset{d}{=} U_k, k = 1, \ldots, d$, is the survival copula of $(U_1, \ldots, U_d)$. Consequently, a copula C is radially symmetric if and only if it equals its survival copula $\hat{C}$. Recall that the coherence between both objects is given in Proposition 2.2.7.

2.2.3 Concordance and dependence measures

There are various measures capturing the interaction of random variables respectively the dependence induced by a given multivariate distribution function. In many cases, the interrelationship of a random vector is compressed into a single number in order to simplify the interpretation and facilitate the comparison among different distribution functions. An extensive overview on dependence concepts is given in (Joe, 1997, Chapter 2). In this thesis, we analyze exchangeable exogenous shock models with respect to three well-known and widespread dependence coefficients, each of which refers to the interrelationship of a bivariate random vector respectively copula. Supplements and

multivariate extensions of these measures are given in Joe (1990), Schmid and Schmidt (2007), or Schmid et al. (2010).

A first prominent measure of extremal dependence are tail dependence coefficients (see (Nelsen, 2006, p. 214 f.)), which, loosely speaking, represent the likelihood of one random variable being extremely small (large), given that the other one is extremely small (large).

Definition 2.2.16 (Tail dependence)
Let X_1, X_2 be continuous random variables with distribution functions $X_1 \sim F_1, X_2 \sim F_2$, copula C, and survival copula $\bar{C}$. The lower tail dependence coefficient λ_L of (X_1, X_2) is defined by

$$\begin{aligned}\lambda_L :&= \lim_{u \searrow 0} \mathbb{P}(X_2 \leq F_2^{-1}(u) \big| X_1 \leq F_1^{-1}(u)) \overset{(*)}{=} \lim_{u \searrow 0} \mathbb{P}(F_2(X_2) \leq u \big| F_1(X_1) \leq u) \\ &= \lim_{u \searrow 0} \frac{\mathbb{P}(F_2(X_2) \leq u, F_1(X_1) \leq u)}{\mathbb{P}(F_1(X_1) \leq u)} = \lim_{u \searrow 0} \frac{C(u,u)}{u},\end{aligned}$$

where $(*)$ is valid due to (Embrechts and Hofert, 2013, Proposition 2.3, (4)). In a similar way, the upper tail dependence coefficient λ_U is given by

$$\begin{aligned}\lambda_U :&= \lim_{u \nearrow 1} \mathbb{P}(X_2 > F_2^{-1}(u) \big| X_1 > F_1^{-1}(u)) = \lim_{u \nearrow 1} \frac{\hat{C}(1-u, 1-u)}{1-u} \\ &= \lim_{u \nearrow 1} \frac{C(u,u) - 2\,u + 1}{1-u},\end{aligned}$$

provided the respective limits exist.

Being defined as the limit of a conditional distribution function, it naturally holds that $0 \leq \lambda_L, \lambda_U \leq 1$. Besides tail dependence, which measures dependence in extreme scenarios, the two most famous dependence coefficients are Kendall's tau and Spearman's rho. Both concepts can be assigned to the class of concordance measures. Loosely speaking, concordance of a bivariate random vector (X_1, X_2) means that large (low) values of X_1 are likely to concur with large (low) values of X_2. An axiomatic specification of bivariate concordance measures is given in Scarsini (1984).

Definition 2.2.17 (Kendall's tau)
Let $(U_1, U_2), (V_1, V_2) \sim C$ be i.i.d. random vectors with bivariate copula C. Kendall's tau τ_C of C is defined as

$$\tau_C := \mathbb{P}\big((U_1 - V_1)\,(U_2 - V_2) > 0\big) - \mathbb{P}\big((U_1 - V_1)\,(U_2 - V_2) < 0\big).$$

Equivalently (see (Nelsen, 2006, p. 159 ff.)), one can use the representations

(i) $\tau_C = 4\,\mathbb{E}[C(U_1, U_2)] - 1$,

(ii) $\tau_C = 1 - 4\int_0^1 \int_0^1 \frac{\partial}{\partial u_1} C(u_1, u_2) \frac{\partial}{\partial u_2} C(u_1, u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2$.

Definition 2.2.18 (Spearman's rho)
Let $(U_1, U_2), (V_1, V_2), (W_1, W_2) \sim C$ *be i.i.d. random vectors with bivariate copula* C. *Spearman's rho* ρ_C *of* C *is defined as*

$$\rho_C := 3\Big(\mathbb{P}\big((U_1 - V_1)(U_2 - W_2) > 0\big) - \mathbb{P}\big((U_1 - V_1)(U_2 - W_2) < 0\big)\Big).$$

Equivalently (see (Mai and Scherer, 2012, p. 32 ff.)), one can use the representations

(i) $\rho_C = \mathrm{Corr}(U_1, U_2)$, *where "*Corr*" denotes the Pearson correlation coefficient.*

(ii) $\rho_C = 12\int_0^1 \int_0^1 C(u_1, u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 - 3 = \frac{\int_0^1 C(u_1,u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 - \int_0^1 u_1\,u_2\,\mathrm{d}u_1\,\mathrm{d}u_2}{\int_0^1 \min\{u_1,u_2\}\,\mathrm{d}u_1\,\mathrm{d}u_2 - \int_0^1 u_1\,u_2\,\mathrm{d}u_1\,\mathrm{d}u_2}$.

For a multivariate extension of these measures (and general measures of concordance), the reader is referred to Taylor (2007).

2.3 Copula characterizations via monotonicity of sequences and functions

Up to now, we still lack concrete examples of non-trivial copulas, which is made up for in the present section. The core of this thesis is the characterization of exchangeable exogenous shock models. This is achieved by analyzing the copula linked to the respective multivariate distribution function. Characterization results for some popular and commonly used copula families have already been derived in the literature. Indeed, such characterizations might have worked as a driver for the copulas' popularity in the first place. This section recalls these results for two families called *Archimedean* and *exchangeable Marshall–Olkin copulas*. The families are parameterized by a single function on $[0, \infty)$ (in the Archimedean case) respectively a non-negative sequence of numbers (in the Marshall–Olkin case). It turns out that the characterization of these distribution functions is linked to a monotonicity behavior of the function respectively sequence, which on a high level can be interpreted as a multivariate analogue to the increasingness of a univariate distribution function. Keeping in mind the d-increasingness condition in the Definition 2.2.2 of copulas, such a coherence seems to be natural, though it is generally very hard to derive explicitly. As this thesis reveals similar results for a quite general class of copulas, the concepts are introduced in the present section.

2.3.1 Monotone sequences and functions

We introduce a monotonicity concept for both sequences and functions by the notion of *d-monotonicity.*

Definition 2.3.1 (d-monotone sequence)
A real-valued sequence $\{a_0, \dots, a_{d-1}\}$ *is called d-monotone if*

$$\sum_{i=0}^{j-1} (-1)^i \binom{j-1}{i} a_{k+i} \geq 0 \quad \text{for all } k \in \mathbb{N}_0, j \in \mathbb{N} : k + j \leq d.$$

Definition 2.3.2 (d-monotone function)
A function $\psi : [0, \infty) \to \mathbb{R}$ *is called d-monotone,* $d \geq 2$*, if it is* $d-2$ *times differentiable on* $(0, \infty)$*, if the derivatives satisfy*

$$(-1)^k \, \psi^{(k)}(x) \geq 0, \quad k = 0, 1, \dots, d-2, \quad x > 0,$$

if further $(-1)^{d-2}\psi^{(d-2)}$ *is decreasing and convex on* $(0, \infty)$*, and if* ψ *is continuous at zero.*

Interestingly, d-monotone functions have a stochastic representation by means of certain distributional transforms. Establishing a connection between copulas and d-monotonicity in the following, this link provides a powerful tool in interpreting, constructing, and simulating the respective multivariate distribution functions.

Definition 2.3.3 (Williamson d-transform, Williamson (1956))
Let X *be a positive random variable. For* $d \geq 2$*, the Williamson d-transform* $\psi_{d,X}$ *of* X *is given by*

$$\psi_{d,X}(x) := \mathbb{E}\Big[\max\Big(1 - \frac{x}{X}, 0\Big)^{d-1} \Big], \quad x \geq 0.$$

The following theorem shows that the Williamson d-transform uniquely determines the law of a positive random variable and outlines the correspondence with d-monotone functions.

Proposition 2.3.4 (Properties of the Williamson d-transform)

1. *The distribution of a positive random variable* X *is uniquely determined by its Williamson d-transform* $\psi_{d,X}$*.*
2. *For fixed* $d \geq 2$*, the sets* $\{\psi_{d,X} : X \text{ positive random variable}\}$ *of Williamson d-transforms and* $\{\phi : \phi \;\; d\text{-monotone with } \lim_{x\to\infty} \phi(x) = 0 \text{ and } \phi(0) = 1\}$ *of d-monotone functions coincide.*

Proof
See (McNeil and Nešlehovà, 2009, Proposition 3.1), which goes back to a result by Williamson (1956). □

Extending the concept of d-monotonicity from fixed to arbitrary $d \geq 2$ leads to the class of *completely monotone* sequences and functions. Sections 2.3.2 and 2.3.3 reveal their meaning and demonstrate that the transition from exchangeable to extendible Archimedean and Marshall–Olkin copulas (recall Definitions 2.2.11 and 2.2.12) goes hand in hand with the transition from d-monotone to completely monotone sequences and functions.

Definition 2.3.5 (Completely monotone sequence)
A sequence $\{a_k\}_{k\in\mathbb{N}_0}$ is called completely monotone if

$$\sum_{i=0}^{j-1}(-1)^i \binom{j-1}{i} a_{k+i} \geq 0 \quad \text{for all } k \in \mathbb{N}_0, j \in \mathbb{N}.$$

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

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

For both completely monotone sequences and functions, there are classical results providing probabilistic interpretations.

Theorem 2.3.7 (Hausdorff's Theorem, Hausdorff (1921))
The sequence $\{a_k\}_{k\in\mathbb{N}_0}$ is completely monotone with $a_0 = 1$ if and only if there exists a random variable $X : \Omega \to [0,1]$ such that $a_k = \mathbb{E}[X^k]$ for all $k \in \mathbb{N}_0$. Moreover, the random variable X is uniquely determined by its moments.

Proof
See (Feller, 1966, p. 225 ff.). □

In contrast to the Williamson d-transform, the corresponding distributional transform for completely monotone functions is very well-known: It is the Laplace transform of a positive random variable.

Theorem 2.3.8 (Bernstein's Theorem, Bernstein (1929))
A function $\psi : [0,\infty) \to [0,1]$ with $\psi(0) = 1$ and $\lim_{x\to\infty} \psi(x) = 0$ is completely monotone if and only if ψ is the Laplace transform of a positive random variable M, i.e. $\psi(x) = \mathbb{E}[\exp(-x\,M)]$ and $\mathbb{P}(M > 0) = 1$.

Proof
See (Feller, 1966, p. 439 f.). □

A connection between completely monotone sequences and functions is given in the following lemma.

Lemma 2.3.9
Let $\psi : [0,\infty) \to \mathbb{R}$ be continuous with $\psi(0) = 1$. Then ψ is completely monotone if and only if the sequence $\{\psi(q\,k)\}_{k\in\mathbb{N}_0}$ is completely monotone for every $q \in \mathbb{Q} \cap [0,\infty)$.

Proof
This proof is adapted from (Widder, 1946, p. 162 f.), who uses the same argument to prove Bernstein's Theorem as a corollary to Hausdorff's Theorem.

"$\Leftarrow$": For all $m \in \mathbb{N}$, by Hausdorff's Theorem (Theorem 2.3.7), there exists a unique random variable X_m with values in $[0,1]$ such that $\mathbb{E}[X_m^k] = \psi(k/m)$, $k \in \mathbb{N}_0$. Consequently,

$$\mathbb{E}\big[X_1^k\big] = \psi(k) = \mathbb{E}\big[X_m^{m\,k}\big] = \mathbb{E}\big[(X_m^m)^k\big],$$

where by uniqueness in Hausdorff's Theorem, X_1 is a random variable with $X_1 \stackrel{d}{=} X_m^m$, $m \in \mathbb{N}$. As a result,

$$\psi\Big(\frac{k}{m}\Big) = \mathbb{E}[X_m^k] = \mathbb{E}\Big[(X_m^m)^{\frac{k}{m}}\Big] = \mathbb{E}\Big[X_1^{\frac{k}{m}}\Big], \quad k \in \mathbb{N}_0, m \in \mathbb{N}.$$

Therefore, for all $q \in \mathbb{Q} \cap [0,\infty)$, one has $\psi(q) = \mathbb{E}[X_1^q]$. As both ψ and $x \mapsto \mathbb{E}[X_1^x]$ are continuous on $[0,\infty)$, we conclude that $\psi(x) = \mathbb{E}[X_1^x]$ for all $x \in [0,\infty)$. Thus, by Bernstein's Theorem (Theorem 2.3.8), $x \mapsto \mathbb{E}[X_1^x] = \psi(x)$ is completely monotone, since it is the Laplace transform of the random variable $-\log(X_1) \in [0,\infty)$ (by continuity of ψ at zero, $\mathbb{P}(X_1 = 0) = 0$).

"$\Rightarrow$": Let ψ be completely monotone. By Bernstein's Theorem, there exists a unique random variable X on $[0,\infty)$ (again, since ψ is continuous, $\mathbb{P}(X = \infty) = 0$) such that $\psi(x) = \mathbb{E}[\exp(-x\,X)]$ for $x \in [0,\infty)$. As a consequence,

$$\{\psi(q\,k)\}_{k\in\mathbb{N}_0} = \Big\{\mathbb{E}\big[\big(e^{-q\,X}\big)^k\big]\Big\}_{k\in\mathbb{N}_0}$$

is the moment sequence of the random variable $X_q := \exp(-q\,Y) \in [0,1]$. By Hausdorff's Theorem, it follows that the sequence $\{\psi(q\,k)\}_{k\in\mathbb{N}_0}$ is completely monotone. □

Both the concepts of d-monotone respectively completely monotone functions and the corresponding transforms of positive random variables become especially important in the characterization and construction of *Archimedean copulas*.

2.3.2 Archimedean copulas

Archimedean copulas are parameterized by a single function $\psi : [0, \infty) \to [0, 1]$, which is termed *Archimedean generator* and is commonly demanded to satisfy some technical conditions.

Definition 2.3.10 (Archimedean generator)
A function $\psi : [0, \infty) \to [0, 1]$ is an Archimedean generator if it satisfies the following conditions:

1. *$\psi(0) = 1$ and $\lim_{x\to\infty} \psi(x) = 0$,*
2. *ψ is continuous,*
3. *ψ is decreasing on $[0, \infty)$ and strictly decreasing on $[0, \inf\{x > 0 : \psi(x) = 0\}]$, with the convention $\inf \emptyset := \infty$.*

Now we can proceed to the definition of Archimedean copulas.

Definition 2.3.11 (Archimedean copula)
A copula $C : [0, 1]^d \to [0, 1]$ is an Archimedean copula if C has the form

$$C(u_1, \ldots, u_d) = \psi\big(\psi^{-1}(u_1) + \ldots + \psi^{-1}(u_d)\big), \quad u_1, \ldots, u_d \in [0, 1], \tag{2.7}$$

for an Archimedean generator ψ.

Due to the "nice" analytical structure and the parameterization in terms of the single, yet flexible generator ψ, Archimedean copulas have been extensively studied in the literature. Genest and Rivest (1993) deal with statistical inference in the bivariate case. Results on tail properties (among others the tail dependence coefficients in Definition 2.2.16) of multivariate Archimedean copulas are given in Charpentier and Segers (2009). Studies on the construction, sampling, and application of standard and nested Archimedean copulas (the latter referring to a certain combination of several copulas of Archimedean type) can be found in Hering et al. (2010); Hofert (2008); Hofert and Scherer (2011); McNeil (2008).

By the definition above, Archimedean copulas are exchangeable. Permuting the input arguments $u_1, \ldots, u_d$ has no influence on the value of the joint distribution function C as the generator ψ is evaluated at the sum of the generalized inverses ψ^{-1} for $u_1, \ldots, u_d$. As a note of caution, it has to be emphasized that the requirements of an Archimedean generator ψ are just necessary conditions for C in (2.7) to define a copula. Sufficient

criteria in dimension d are for instance imposed in Genest and Rivest (1989); Müller and Scarsini (2005), where the first d derivatives of ψ are demanded to exist and satisfy $(-1)^k \psi^{(k)} \geq 0$ for $k \in \{1, \ldots, d\}$. These conditions already hint at the characterization results for Archimedean copulas in arbitrary, yet fixed dimension $d \geq 2$, which are given in (McNeil and Nešlehovà, 2009, Theorem 2.2).

Theorem 2.3.12 (Characterization of Archimedean copulas)
Let ψ be an Archimedean generator and define a function $C : [0,1]^d \to [0,1]$ by (2.7). Then C is a copula if and only if ψ is d-monotone.

Proof
See (McNeil and Nešlehovà, 2009, Theorem 2.2). □

Consequently, the generator of an Archimedean copula can be linked to the Williamson d-transform in Definition 2.3.3. In McNeil and Nešlehovà (2009), this connection is used to construct a stochastic model for a random vector $(U_1, \ldots, U_d)$ with Archimedean copula. A more classical result which refers to the characterization of Archimedean generators that define a copula in any dimension $d \geq 2$ is given in Kimberling (1974).

Theorem 2.3.13 (Characterization of extendible Archimedean copulas)
Let ψ be an Archimedean generator and define a function $C : [0,1]^d \to [0,1]$ by (2.7). Then C is a copula for any $d \geq 2$ if and only if ψ is completely monotone.

Proof
The seminal reference for this result is (Kimberling, 1974, Theorems 1 and 2). Naturally, it also follows in a straightforward way from Theorem 2.3.12. □

By De Finetti's Theorem (see Theorem 2.2.13), we know that an exchangeable structure for an infinite sequence of random variables implies the existence of a sub-σ algebra conditioned on which the random variables are i.i.d. Following the deliberations in Marshall and Olkin (1988), this implicit coherence can be made explicit and a stochastic model for a random vector corresponding to an extendible Archimedean copula C can be derived. More precisely, it is straightforward to show that $C : [0,1]^d \to [0,1]$ is the survival copula of the random vector $(X_1, \ldots, X_d)$ constructed by

$$X_k := \frac{E_k}{M} = \inf\{t \geq 0 : t\, M > E_k\}, \quad k \in \mathbb{N}, \tag{2.8}$$

where M is a positive random variable such that its Laplace transform is given by the Archimedean generator. The existence of such a random variable is guaranteed by

Theorem 2.3.8. Conditioned on (the σ-algebra generated by) M, the $\{X_k\}_{k\in\mathbb{N}}$ are i.i.d. with distribution function

$$\mathbb{P}\big(X_k \le x|M\big) = \mathbb{P}\big(E_k \le x\,M|M\big) = 1 - e^{-x\,M}, \quad x \ge 0.$$

Put another way, each X_k is exponentially distributed with a random parameter M that introduces dependence between the random variables. The stochastic model in (2.8) already bears resemblance to the first-passage time construction

$$X_k := \inf\{t \ge 0 : \Lambda_t > E_k\}, \quad k = 1, \dots, d,$$

of extendible exogenous shock models in Equation (1.2) in the introduction, the difference being that as opposed to $\{\Lambda_t\}_{t\ge0}$, the process $\{t\,M\}_{t\ge0}$ obviously does not exhibit independent increments. The coherence between monotonicity behavior of sequences and functions, extendibility of distribution functions, and first-passage time representations of the underlying random vector turns out to be a core element of the present thesis.

2.3.3 Exchangeable Marshall–Olkin copulas

A second example of copula families is given by (exchangeable) Marshall–Olkin copulas, which arise as survival copulas of the *Marshall–Olkin (MO) distribution*. This class of distributions can be seen as a multivariate analogue of the exponential distribution in the univariate case and is originally introduced in Marshall and Olkin (1967). MO distributions constitute the presumably most widespread and studied representative of exogenous shock models. The underlying stochastic model of a random vector $(X_1, \dots, X_d)$ with d-variate MO distribution is given by the general construction

$$X_k = \min\{Z^E : k \in E\}, \quad k = 1, \dots, d,$$

with the additional condition that the independent shocks $Z^E, \emptyset \ne E \subseteq \{1, \dots, d\}$, are exponentially distributed. It is shown in (Mai and Scherer, 2011, Lemma 2.1) with the help of results in Li (2008) that for exchangeable MO distributions, the corresponding (unique) survival copula of $(X_1, \dots, X_d)$ admits the form

$$C(u_1, \dots, u_d) = \prod_{k=1}^{d} u^{a_{k-1}}, \quad u_1, \dots, u_d \in [0,1], \tag{2.9}$$

for non-negative parameters $a_0, \dots, a_{d-1} \ge 0$ with $a_0 = 1$. Thus, exchangeable MO copulas are parameterized by a sequence $\{a_k\}_{k=0,\dots,d-1}$ compared to the function ψ in

the Archimedean setup. As a consequence, it might be tempting to draw an analogy and call $\{a_k\}_{k=0,\ldots,d-1}$ a *Marshall–Olkin-generating sequence*. However, it is in no sense obvious that any sequence leading to a copula C in (2.9) necessarily induces that C is the survival copula of an exchangeable MO distribution or a general exchangeable exogenous shock model at all.

In fact, this reverse direction is a tough mathematical challenge on its own, which is solved for a very general class of copulas in the present thesis. In the special case of copulas of type (2.9), Mai (2010) and the related publications mentioned in the personal note in Section 1.2 derive three major results:

1. The function C in (2.9) defines a copula if and only if $\{a_0, \ldots, a_{d-1}\}$ is d-monotone with $a_0 = 1$.

2. The set of copulas with form (2.9) coincides with the set of survival copulas of exchangeable exogenous shock models with exponentially distributed shocks.

3. The function C in (2.9) defines a copula for all $d \geq 2$ if and only if $\{a_k\}_{k\in\mathbb{N}_0}$ is a completely monotone sequence with $a_0 = 1$. In this case, C is the survival copula of

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d, \tag{2.10}$$

   where $\Lambda = \{\Lambda_t\}_{t\geq 0}$ is a *killed Lévy subordinator* (see Section 2.4) which is (uniquely in law) related to the sequence $\{a_k\}_{k\in\mathbb{N}_0}$. Conversely, any killed Lévy subordinator in the construction above leads to a survival copula of the form (2.9) with a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ satisfying $a_0 = 1$. Following Mai and Scherer (2009b), the corresponding set of copulas is termed *Lévy–frailty copulas*.

Interestingly (see also (Mai, 2010, p. 87)), the characterization results for Archimedean and exchangeable MO copulas are somewhat similar from a structural point of view. First, Equations (2.7) and (2.9) define a copula in fixed dimension d if and only if the function ψ respectively the sequence $\{a_k\}_{k=0,\ldots,d-1}$ are d-monotone. Second, in both cases, the extendible subclass of copulas can be constructed via an explicitly known first-passage time setup involving an increasing process $\{t\, M\}_{t\geq 0}$ in the Archimedean and a Lévy subordinator $\{\Lambda_t\}_{t\geq 0}$ in the MO case.

2.4 Additive processes and Bernstein functions

Up to now, we have not introduced or further specified the Lévy subordinator appearing in the first-passage time framework for MO copulas in Equation (2.10). However, similar constructions involving the more general class of *increasing additive processes* for $\{\Lambda_t\}_{t\geq 0}$ play a crucial role in this thesis. Therefore, the present section establishes the notion of additive processes and indicates three exemplary subclasses consisting of Lévy processes, Sato processes, and a simple transformation of certain Dirichlet processes. The distinctive property of an additive process $\{A_t\}_{t\geq 0}$ is the independence of its increments, meaning that the change $A_T - A_t$ between two arbitrary points in time $0 \leq t < T$ is independent of the history of $\{A_t\}_{t\geq 0}$ before and its state at time t. When considering increasing additive processes $\{\Lambda_t\}_{t\geq 0}$ starting at $\Lambda_0 = 0$, non-negativity and additivitiy of $\{\Lambda_t\}_{t\geq 0}$ allow for a convenient process characterization in terms of the set of Laplace transformations of Λ_t for all $t \geq 0$. In this context, infinite divisibility of Λ_t provides a link between the set of Laplace transforms and a family $\{\Psi_t\}_{t\geq 0}$ of *Bernstein functions*. Bernstein functions represent the second mathematical object depicted in detail in the present section. Besides the connection to increasing additive processes, we provide some illustrative Bernstein functions that are picked up in later chapters.

2.4.1 Definition and coherence

In the sequel, we rely on the following definition of additive processes.

Definition 2.4.1 (Additive process)
A stochastic process $A = \{A_t\}_{t\geq 0}$ on $\mathbb{R}$ is called additive if it satisfies the following properties:

(i) Start at zero: $A_0 = 0$ almost surely.

(ii) Independent increments: For any choice of $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables $A_{t_0}, A_{t_1} - A_{t_0}, A_{t_2} - A_{t_1}, \ldots, A_{t_n} - A_{t_{n-1}}$ are independent.

(iii) Stochastic continuity: For every $t \geq 0$ and $\epsilon > 0$, it holds that $\lim_{s\to t} \mathbb{P}(|A_s - A_t| > \epsilon) = 0$, i.e. A almost surely has no jumps at fixed points in time.

This is a slightly different formulation compared to the seminal reference (Sato, 1999, Definition 1.6, p. 3). There, the conditions of Definition 2.4.1 solely describe an *additive process in law*. For the term *additive process*, it is additionally required that A possesses càdlàg paths, i.e. that there is an $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that for every $\omega \in \Omega_0$,

$\{A_t(\omega)\}_{t\geq 0}$ is right-continuous in $t \geq 0$ and has left limits in $t > 0$. For our purposes, this distinction is not necessary as we are only interested in the law of A, i.e. the system of its finite-dimensional distributions $\mathbb{P}(A_{t_1} \in B_1, \ldots, A_{t_n} \in B_n)$ for $n \in \mathbb{N}, 0 \leq t_1 < t_2 < \ldots < t_n$, and $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$. This system is uniquely determined by the marginal distributions of the additive process, i.e. the law of A_t for all $t \geq 0$ (see (Sato, 1999, Theorem 9.7 (iii), p. 51)). According to (Sato, 1999, Theorem 11.5, p. 63), an additive process can be modified (without changing its marginal distributions) such that it exhibits càdlàg paths.

In this thesis, we primarily deal with increasing additive processes, called *additive subordinators* and denoted $\Lambda = \{\Lambda_t\}_{t\geq 0}$ in the sequel. In this case, Λ_t is a non-negative random variable for any $t \geq 0$. As such, the law of Λ_t can be uniquely described by its Laplace transform (compare (Doetsch, 1974, Theorem 5.1, p. 21)). We will see that the Laplace transform of an additive subordinator at fixed points in time is closely related to *Bernstein functions*, which form an important building block of the present manuscript.

Definition 2.4.2 (Bernstein function)
A function $\Psi : [0, \infty) \to [0, \infty)$ is said to be a Bernstein function if $\Psi(0) = 0$, Ψ possesses derivatives of all orders on $(0, \infty)$, and $(-1)^{n-1}\,\Psi^{(n)}(x) \geq 0$ for all $n \in \mathbb{N}$ and $x > 0$.

Note that by definition, a non-negative, infinitely often differentiable function $f : [0, \infty) \to \mathbb{R}$ with $f(0) = 0$ is a Bernstein function if and only if its derivative f' (to be exact its continuation at zero) is completely monotone (compare Definition 2.3.6). An excellent textbook on the subject is Schilling et al. (2010). Each infinitely divisible probability law π on the compactified positive half-axis $[0, \infty]$ is uniquely associated with a Bernstein function Ψ via the relation $\mathcal{L}[\pi] = \exp(-\Psi)$, i.e. Ψ is the Laplace exponent of π. It is well-known (see, e.g., (Schilling et al., 2010, Theorem 3.2, p. 15)) that every such Ψ admits a unique *Lévy–Khintchine representation*

$$\Psi(x) = a\,\mathbb{1}_{\{x>0\}} + b\,x + \int_{(0,\infty)} (1 - e^{-x\,t})\,\nu(\mathrm{d}t), \quad x \geq 0, \tag{2.11}$$

with constants $a, b \geq 0$ and a measure ν on $(0, \infty)$ (called the *Lévy measure*) satisfying $\int_{(0,\infty)} \min\{1, t\}\,\nu(\mathrm{d}t) < \infty$. Conversely, any such choice of (a, b, ν) determines a unique Bernstein function Ψ.

For an additive subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$, the independence of increments induces that Λ_t is an infinitely divisible random variable (see (Sato, 1999, Theorem 9.1, p. 47)) on $[0, \infty)$ for any $t \geq 0$. Combining this observation with the fact that an additive process is

(uniquely in law) described by its marginal distributions, the law of Λ can be described by a family $\{\Psi_t\}_{t\geq 0}$ of Bernstein functions via the relation $\mathcal{L}[\Lambda_t] = \exp(-\Psi_t)$. The link to Definition 2.4.1 is established by the following consistency conditions for $\{\Psi_t\}_{t\geq 0}$.

(i) Start at zero: $\Psi_0(x) \equiv 0$ for all $x \geq 0$.

(ii) Independent increments: $\Psi_t - \Psi_s$ is a Bernstein function for all $0 \leq s \leq t$.

(iii) Stochastic continuity: $t \mapsto \Psi_t(x)$ is continuous in $(0, \infty)$ for each $x \geq 0$.

Any family $\{\Psi_t\}_{t\geq 0}$ of Bernstein functions corresponding to an additive subordinator satisfies these properties. Condition (i) simply reflects the fact that $\Lambda_0 = 0$. The second requirement follows from the independence (and, thus, the infinite divisibility) of increments as for $0 \leq s \leq t$, it holds that

$$\exp(-\Psi_t) = \mathcal{L}[\Lambda_t] = \mathcal{L}[\Lambda_s + \Lambda_t - \Lambda_s] = \mathcal{L}[\Lambda_s]\,\mathcal{L}[\Lambda_t - \Lambda_s] = \exp(-\Psi_s)\,\mathcal{L}[\Lambda_t - \Lambda_s].$$

As $\Lambda_t - \Lambda_s$ is infinitely divisible, its Laplace exponent has to be a Bernstein function given by $\Psi_t - \Psi_s$. To recognize Condition (iii), note that $t \mapsto \Psi_t(x)$ is continuous if and only if for arbitrary $x, t > 0$,

$$\begin{aligned} 0 &= \lim_{s\to t} \Psi_s(x) - \Psi_t(x) = \lim_{s\to t} \Big(-\log\Big(\mathbb{E}\big[e^{-x\,\Lambda_s}\big]\Big)\Big) + \log\Big(\mathbb{E}\big[e^{-x\,\Lambda_t}\big]\Big) \\ &\overset{(*)}{=} \log \frac{\mathbb{E}\big[e^{-x\,\Lambda_t}\big]}{\mathbb{E}\big[\lim_{s\to t} e^{-x\,\Lambda_s}\big]} \Leftrightarrow \mathbb{E}\big[e^{-x\,\Lambda_t} - \lim_{s\to t} e^{-x\,\Lambda_s}\big] = 0, \end{aligned}$$

where we have used the continuity of the logarithm and bounded convergence in $(*)$. If the last expected value was larger than zero, the probability

$$\mathbb{P}\Big(e^{-x\,\Lambda_t} - \lim_{s\to t} e^{-x\,\Lambda_s} > 0\Big)$$

would have to be larger than zero as well, implying that there is a positive likelihood for $\lim_{s\to t} \Lambda_s$ to exceed Λ_t. This would contradict the stochastic continuity of Λ. Reversely, it is induced by (Sato, 1999, Theorem 9.7, p. 51) that the properties (i)-(iii) of $\{\Psi_t\}_{t\geq 0}$ already guarantee the existence of a (unique in law) additive subordinator Λ characterized by $\mathcal{L}[\Lambda_t - \Lambda_s] = \exp(-\Psi_t + \Psi_s)$ for all $0 \leq s \leq t$.

As a side remark (and note of caution), it has to be kept in mind that there might still be other stochastic processes outside the universe of additive subordinators corresponding to $\{\Psi_t\}_{t\geq 0}$. It is shown in (Mai and Scherer, 2014, Theorem 1.1) that a family $\{\Psi_t\}_{t\geq 0}$ of Bernstein functions satisfying $\Psi_t = t\,\Psi, t \geq 0$, with a Bernstein function Ψ and, thus, fulfilling Conditions (i)-(iii) above corresponds to an increasing stochastic process

$S = \{S_t\}_{t\geq 0}$ that is *infinitely divisible with respect to time* (see the proof of Proposition 4.3.1 for a definition). However, generally, S is not unique in law, which is illustrated by means of an example in Mansuy (2005). Consider the Bernstein function $\Psi(x) = x^\alpha, x \geq 0, \alpha \in (0,1)$, for which the requirements in Definition 2.4.2 are easy to verify. There are two quite different processes $\{S_t\}_{t\geq 0}$ and $\{\Lambda_t\}_{t\geq 0}$ satisfying $\mathbb{E}[\exp(-x\, S_t)] = \mathbb{E}[\exp(-x\, \Lambda_t)] = \exp(-t\, \Psi(x))$. The first one is $S_t := t^{1/\alpha}\, M$ for a random variable M with Laplace transform $\mathcal{L}[M] = \exp(-\Psi)$. The second one is an additive subordinator that features stationary increments, meaning that for an equidistant grid $t_k - t_{k-1} = \Delta > 0, k = 1, \ldots, n$, the increments in Definition 2.4.1 are not only independent, but even identically distributed. Self-evidently, the joint distribution of $\{S_t\}_{t\geq 0}$ at different points in time strongly differs from the one of $\{\Lambda_t\}_{t\geq 0}$. For $0 \leq t_1 < t_2$, the law of $S_{t_2} - S_{t_1}$ fully depends on the increment $S_{t_1} - S_0 = S_{t_1}$, while there is no dependence at all for the increments of $\{\Lambda_t\}_{t\geq 0}$.

2.4.2 Lévy process

In the preceding example, the process $\{\Lambda_t\}_{t\geq 0}$ represents one of the most prominent and best analyzed subclasses of additive processes. It belongs to the class of Lévy processes, which are defined as additive processes $\{A_t\}_{t\geq 0}$ with the additional property of stationary increments, i.e. for any $s, t \geq 0$, the distribution of $A_{s+t} - A_t$ is independent of t. Increasing Lévy processes $\Lambda = \{\Lambda_t\}_{t\geq 0}$ are called *Lévy subordinators*. Referring to the characterization of increasing additive processes via a family $\{\Psi_t\}_{t\geq 0}$ of Bernstein functions with consistency Conditions (i)-(iii) above, stationarity implies that

$$e^{-\Psi_t(x)+\Psi_s(x)} = \mathcal{L}[\Lambda_t - \Lambda_s](x) = \mathcal{L}[\Lambda_{t-s}](x) = e^{-\Psi_{t-s}(x)} \Leftrightarrow \Psi_{t-s}(x) = \Psi_t(x) - \Psi_s(x)\,(*)$$

for all $x \geq 0$. Thus, it holds that for all $t_1, t_2, x \geq 0$,

$$\Psi_{t_1}(x) + \Psi_{t_2}(x) = \Psi_{t_1}(x) + \Psi_{t_1+t_2-t_1}(x) \overset{(*)}{=} \Psi_{t_1+t_2}(x),$$

where $(*)$ is applied for $t = t_1 + t_2, s = t_1$. Defining $\Phi_x(t) := \Psi_t(x)$, we have that $\Phi_x(t_1) + \Phi_x(t_2) = \Phi_x(t_1 + t_2)$. By the Cauchy functional equation, it consequently holds that $\Phi_x(t) = t\, \Phi_x(1)$, which finally yields the distinctive property $\Psi_t(x) = t\, \Psi_1(x)$, $x, t \geq 0$, of Lévy subordinators. A slight extension of a Lévy subordinator $\{\Lambda_t\}_{t\geq 0}$ is given by the notion of a *killed* Lévy subordinator $\{\tilde{\Lambda}_t\}_{t\geq 0}$ (see (Schilling et al., 2010, p. 36)), which is defined via an additional, exponentially distributed random variable E^a

with parameter $a > 0$ (called *killing rate*) by

$$\tilde{\Lambda}_t = \begin{cases} \Lambda_t, & t < E^a, \\ +\infty, & t \geq E^a. \end{cases}$$

Thus, the state space $[0, \infty)$ of a Lévy subordinator is extended to $[0, \infty) \cup \{\infty\}$ in the killed case. It is straightforward to deduce that if $\{\Lambda_t\}_{t \geq 0}$ is a Lévy subordinator with Bernstein family $\{\Psi_t\}_{t \geq 0}$, $\Psi_t = t\,\Psi_1$, the killed process $\{\tilde{\Lambda}_t\}_{t \geq 0}$ satisfies

$$\mathbb{E}[e^{-x\,\tilde{\Lambda}_t}] = e^{-t\,\tilde{\Psi}_1(x)}, \quad x \geq 0,$$

for $\tilde{\Psi}_1(x) := a\,\mathbb{1}_{\{x>0\}} + \Psi_1(x)$. Reversely (see (Schilling et al., 2010, p. 35 ff.)), it can be shown that for an arbitrary Bernstein function $\tilde{\Psi}_1$, there is a (unique in law) killed Lévy subordinator such that the corresponding family $\{\tilde{\Psi}_t\}_{t \geq 0}$ of the process satisfies $\tilde{\Psi}_t = t\,\tilde{\Psi}_1, t \geq 0$. As a final remark, take note of the fact that in the Lévy–Khintchine representation of Ψ_1 in (2.11), the parameter a represents the killing rate of the corresponding Lévy subordinator. Therefore, for classical, non-killed Lévy subordinators, it holds that $a = 0$.

2.4.3 Sato process

The second class of additive processes which especially play a major role in Chapter 4 of this thesis are self-similar additive processes. A stochastic process $\{X_t\}_{t \geq 0}$ is called self-similar if for any $a > 0$, there is a $b > 0$ such that $\{X_{a\,t}\}_{t \geq 0} \overset{d}{=} \{b\,X_t\}_{t \geq 0}$. (Sato, 1999, Theorem 13.11, p. 73) shows that any self-similar process features an $H > 0$ such that $b = a^H$. In the literature, $H > 0$ is often called the self-similarity exponent, while the corresponding self-similar additive process with exponent H is referred to as *H-Sato process*.

Describing increasing self-similar additive processes $\{\Lambda_t\}_{t \geq 0}$ with self-similarity exponent H, termed *H-Sato subordinators* in the sequel, via the corresponding family $\{\Psi_t\}_{t \geq 0}$ of Bernstein functions, one notes that $\Lambda_{a\,t} \overset{d}{=} a^H\,\Lambda_t$ for all $a > 0$ and $t \geq 0$ is equivalent to

$$\Psi_{a\,t}(x) = -\log\left(\mathbb{E}[e^{-x\,\Lambda_{a\,t}}]\right) = -\log\left(\mathbb{E}[e^{-x\,a^H\,\Lambda_t}]\right) = \Psi_t(a^H\,x), \quad a > 0, \quad x, t \geq 0.$$

This implies, yet is not equivalent to $\Psi_t(x) = \Psi_1(t^H\,x)$ for all $x, t \geq 0$. Considering the reverse, for a given Bernstein function Ψ_1, the family $\{\Psi_t\}_{t \geq 0}$ defined for a specific $H > 0$ by $\Psi_t(x) := \Psi_1(t^H\,x)$, $x, t \geq 0$, corresponds to a Sato subordinator if and only

if Ψ_1 is *self-decomposable*. This coherence is treated in more detail in Section 4.1. Self-decomposable Bernstein functions form a subclass of Bernstein functions that results from a special structure of the Lévy measure ν in the Lévy-Khintchine representation (2.11). More precisely, it can be shown (see (Sato, 1999, Corollary 15.11, p. 95)) that a Bernstein function is self-decomposable if and only if in the general Bernstein function representation (2.11), it holds that $a = 0$ and that the Lévy measure ν has a density $\nu(\mathrm{d}t) = k(t)/t\,\mathrm{d}t$ such that $t \mapsto k(t)$ is decreasing.

2.4.4 Dirichlet and neutral-to-the-right process

The third class of additive subordinators which is considered in this thesis originates from the construction of Dirichlet processes, which are initially introduced in Ferguson (1973) in a Bayesian analysis context. Dirichlet processes are not additive processes themselves, however can be connected to the latter by a simple transformation. In general, a Dirichlet process $Z = \{Z_t\}_{t\in\mathbb{R}}$ is a certain random distribution function on the real line. It is parameterized by a pair (c, G), where G denotes a non-random distribution function on $\mathbb{R}$ (the expected path of Z) and $c > 0$ influences the "fluctuations" of the paths of Z around G. In a non-parametric Bayesian context, the popularity of the Dirichlet process stems from the fact that the support of Z (i.e. the set of possible realizations of the random distribution function) is "huge", while its posterior distribution given samples from Z is still a Dirichlet process and, thus, remains tractable. The latter property is immaterial for the present thesis. However, the support condition plays a crucial role in Chapter 6.

The Dirichlet process is intimately connected with the *Dirichlet distribution*, whose definition is briefly recalled in order to proceed with the definition of the process itself. The Dirichlet distribution with parameters $(\alpha_1, \ldots, \alpha_\ell) \in [0, \infty)^\ell$, not all components zero, is the law of the random vector

$$\Big(\frac{E_1}{\sum_{k=1}^{\ell} E_k}, \ldots, \frac{E_\ell}{\sum_{k=1}^{\ell} E_k}\Big),$$

where $E_1, \ldots, E_\ell$ is a list of independent Gamma-distributed random variables with densities (for those components k with $\alpha_k > 0$)

$$f_{E_k}(x) = \frac{e^{-x}\, x^{\alpha_k - 1}}{\Gamma(\alpha_k)}\, \mathbb{1}_{\{x>0\}}, \quad k = 1, \ldots, \ell.$$

The case $\alpha_k = 0$ is interpreted as $E_k \equiv 0$. The Dirichlet process Z with parameters (c, G) is a random distribution function on $\mathbb{R}$ which is characterized by the property that for

any measurable partition[2] $\{B_1, \dots, B_\ell\}$ of $\mathbb{R}$, the random vector $\big(\mathrm{d}Z(B_1), \dots, \mathrm{d}Z(B_\ell)\big)$ has a Dirichlet distribution with parameters $c\,\big(\mathrm{d}G(B_1), \dots, \mathrm{d}G(B_\ell)\big)$. It follows that the random variable Z_t has a Beta-distribution with parameters $\big(c\,G(t), c\,(1-G(t))\big)$, using the standard parameterization[3] of the latter. In particular, this implies that $\mathbb{E}[Z_t] = G(t)$ and $\mathrm{Var}[Z_t] = G(t)\,(1-G(t))/(c+1)$. Summarizing, Z is a random distribution function with "average realization" G, whose variance is controlled by the parameter $c > 0$.

A stochastic representation of the Dirichlet process which manifests the connection to additive processes is given in Ferguson (1974). Define $t_0 := \sup\{t \in \mathbb{R} : G(t) = 0\}$ and $t_1 := \inf\{t \in \mathbb{R} : G(t) = 1\}$, with $t_0 := -\infty$ respectively $t_1 := \infty$ if $G(t) > 0$ respectively $G(t) < 1$ for all $t \in \mathbb{R}$. Let $\Lambda = \{\Lambda_t\}_{t\in\mathbb{R}}$ be a right-continuous, increasing stochastic process with independent increments, whose probability law is uniquely determined by the Laplace transforms of its marginals. For $t < t_0$ respectively $t > t_1$ (provided $t_0 > -\infty$ respectively $t_1 < \infty$), we set $\Lambda_t = 0$ respectively $\Lambda_t = \infty$. For $t \in \mathbb{R} \cap [t_0, t_1]$, we demand

$$\mathbb{E}\Big[e^{-x\,\Lambda_t}\Big] = e^{-\Psi_t(x)}, \quad x \geq 0,\, t \in \mathbb{R} \cap [t_0, t_1],$$

where the family $\{\Psi_t\}_{t\in\mathbb{R}\cap[t_0,t_1]}$ of Laplace exponents has the Lévy–Khintchine representation

$$\Psi_t(x) = \int_0^\infty \big(1 - e^{-x\,u}\big)\,\frac{e^{-u\,c\,(1-G(t))} - e^{-u\,c}}{u\,(1-e^{-u})}\,\mathrm{d}u \tag{2.12}$$

with a constant $c > 0$. Relying on results of Doksum (1974), Ferguson (1974) shows that the random distribution function defined by $Z_t := 1 - \exp(-\Lambda_t)$, $t \in \mathbb{R}$, is a Dirichlet process with parameters (c, G).

It is important to mention that the process Λ characterized by (2.12) generally deviates from an additive subordinator (i.e. an additive process in the sense of Definition 2.4.1 that is increasing) in two regards: On the one hand, it is defined on $\mathbb{R}$ rather than on $[0, \infty)$. On the other hand, when analyzing the continuity of $t \mapsto \Psi_t$ in (2.12), one recognizes that Λ is stochastically continuous if and only if G is a continuous distribution function (see also (Doksum, 1974, Proposition 3.1)). However, when plugging a non-negative continuous distribution function G into (2.12), the corresponding process Λ is an additive process in terms of our definition.

[2] I.e. the $\{B_k\}_{k=1,\dots,\ell}$ are measurable, disjoint, and their union equals $\mathbb{R}$.

[3] A random variable X is said to be Beta(p,q)-distributed, $p, q > 0$, if its density on $(0,1)$ is given by $f(x) = x^{p-1}\,(1-x)^{q-1}/B(p,q)$ and $f(x) = 0$ for $x \notin (0,1)$, where $B(p,q) := \int_0^1 u^{p-1}\,(1-u)^{q-1}\,\mathrm{d}u$ denotes the Beta function.

A generalizing concept of Dirichlet processes is introduced in Doksum (1974) as *neutral-to-the-right priors*. It is motivated in a quite similar way and aims at offering a broader class of random distribution functions in a non-parametric Bayesian context. In particular, the posterior distribution for observed samples from a neutral-to-the-right prior is again neutral-to-the-right, which is convenient for statistical applications. The crucial result for this thesis is (Doksum, 1974, Theorem 3.1), which characterizes separable neutral-to-the-right processes $\{Z_t\}_{t\in\mathbb{R}}$ as random distribution functions of the form

$$Z_t := 1 - e^{-\Lambda_t}$$

for a separable, increasing, right-continuous stochastic process $\{\Lambda_t\}_{t\geq 0}$ with independent increments and $\lim_{t\to-\infty}\Lambda_t = 0, \lim_{t\to\infty}\Lambda_t = \infty$ almost surely. Consequently, neutral-to-right-priors on $[0,\infty)$ that satisfy stochastic continuity can be used to derive specifications of additive subordinators defined on the real line.

2.4.5 Examples of Bernstein functions

The relationship $\Psi = -\log(\mathcal{L}[\pi])$ between Bernstein functions Ψ and infinitely divisible probability laws π on $[0,\infty]$ plays an important role in this thesis. Thus, we consider four examples of popular infinitely divisible distributions on the positive half-axis and the associated Bernstein functions, which are picked up later in various applications.

Positive stable distribution

In general, the density of positive, strictly stable distributions is unknown. Nevertheless, they can be characterized by the Laplace exponent of the related probability law, which is given by

$$\Psi^{\mathrm{St}}_{\alpha,\beta}(x) = \beta\, x^{\alpha}, \quad 0 < \alpha < 1, \beta > 0, \quad x \geq 0,$$

and can be written as

$$\Psi^{\mathrm{St}}_{\alpha,\beta}(x) = \beta \int_0^\infty (1 - e^{-s\,x})\, \frac{\alpha\, s^{-1-\alpha}}{\Gamma(1-\alpha)}\, \mathrm{d}s.$$

Consequently, one derives the Lévy measure of $\Psi^{\mathrm{St}}_{\alpha,\beta}$ in the Lévy–Khintchine representation (2.11) as

$$\nu(\mathrm{d}s) = \beta\, \frac{\alpha}{\Gamma(1-\alpha)}\, s^{-1-\alpha}\, \mathrm{d}s.$$

Inverse Gaussian distribution

The Inverse Gaussian (IG) distribution (for a detailed introduction, see Seshadri (1993))

constitutes another famous example of infinitely divisible laws. It possesses the density

$$f^{\mathrm{IG}}_{\beta,\eta}(x) = \frac{\beta}{\sqrt{2\pi}}\, x^{-\frac{3}{2}}\, \exp\Big(\eta\,\beta - \frac{1}{2}\big(\frac{\beta^2}{x} + \eta^2\, x\big)\Big), \quad x > 0,$$

for parameters $\beta, \eta > 0$. The corresponding Bernstein function is given by

$$\Psi^{\mathrm{IG}}_{\beta,\eta}(x) = -\log\Big(\mathcal{L}[f^{\mathrm{IG}}_{\beta,\eta}]\Big)(x) = \beta\,(\sqrt{2x+\eta^2} - \eta), \quad x \geq 0.$$

The related Lévy measure can be derived as

$$\nu(\mathrm{d}s) = \sqrt{2}\,\beta\,\frac{1}{2\,\Gamma(\frac{1}{2})}\, s^{-\frac{3}{2}}\, e^{-\frac{\eta^2}{2}s}\,\mathrm{d}s.$$

Gamma distribution

As a third example, the Gamma distribution is presented. For parameters $\beta, \eta > 0$, it is characterized by the density

$$f^{\mathrm{Ga}}_{\beta,\eta}(x) = \frac{\eta^\beta}{\Gamma(\beta)}\, x^{\beta-1}\, e^{-\eta\, x}, \quad x \geq 0.$$

The corresponding Bernstein function is given by

$$\Psi^{\mathrm{Ga}}_{\beta,\eta}(x) = \beta\, \log(1 + \frac{x}{\eta}),$$

and the Lévy measure of $\Psi^{\mathrm{Ga}}_{\beta,\eta}$ equals

$$\nu(\mathrm{d}s) = \beta\,\frac{e^{-\eta\, s}}{s}\,\mathrm{d}s.$$

Compound Poisson distribution with exponentially distributed jumps

Consider a sequence $\{E_k\}_{k\in\mathbb{N}}$ of i.i.d. exponentially distributed random variables with parameter $\eta > 0$. For an independent Poisson distributed random variable $N \sim \mathrm{Poi}(\beta), \beta > 0$, the law of

$$X := \sum_{k=1}^{N} E_k$$

is called *compound Poisson distribution* with exponentially distributed jumps (CPE distribution). Conditioning on N and using the i.i.d. property of $\{E_k\}_{k\in\mathbb{N}}$, the Laplace transform of X is given by

$$\begin{aligned}
\mathbb{E}[e^{-z\,X}] &= \sum_{n=0}^{\infty} \mathbb{E}\Big[e^{-z\,\sum_{k=1}^{N} E_k} | N = n\Big]\, \mathbb{P}(N = k) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\Big[e^{-z\,\sum_{k=1}^{n} E_k}\Big]\, \frac{\beta^n}{n!}\, e^{-\beta} \\
&= e^{-\beta} \sum_{n=0}^{\infty} \mathcal{L}[E_1](z)^n\, \frac{\beta^n}{n!}, \quad z \geq 0.
\end{aligned}$$

Exploiting the series representation of the exponential function, this yields

$$\mathbb{E}[e^{-z\,X}] = e^{-\beta}\, e^{\beta\,\mathcal{L}[E_1](z)} = e^{-\beta(1-\mathcal{L}[E_1](z))}. \tag{2.13}$$

It is straightforward to show that the Laplace transform of $E_1 \sim \mathrm{Exp}(\lambda)$ is given by $\mathcal{L}[E_1](z) = \eta/(z+\eta)$ for $z \in [0,\infty)$. As a result, the Bernstein function corresponding to X is given by

$$\Psi^{\mathrm{CPE}}_{\beta,\eta}(x) = \beta(1-\mathcal{L}[E_1](x)) = \beta\big(1-\frac{\eta}{x+\eta}\big), \quad x \geq 0.$$

Furthermore, denoting the distribution function of E_1 by F, (2.13) can be written as

$$\mathbb{E}[e^{-z\,X}] = \exp\Big(-\beta\big(1-\int_0^\infty e^{-z\,x}\,F(\mathrm{d}x)\big)\Big) = \exp\Big(-\int_0^\infty (1-e^{-z\,x})\,\beta\,F(\mathrm{d}x)\Big),$$

indicating that the Lévy measure corresponding to $\Psi^{\mathrm{CPE}}_{\beta,\eta}$ is given by

$$\nu(\mathrm{d}s) = \beta\,F(\mathrm{d}s) = \beta\,\eta\,e^{-\eta\,s}\,\mathrm{d}s.$$

As pointed out in Section 2.4.3, in order to examine the Bernstein functions above for self-decomposability, one has to check whether the Lévy measures ν have a density $\nu(\mathrm{d}t) = k(t)/t\,\mathrm{d}t$ with a decreasing function k. Denoting $k^{\mathrm{St}}_{\alpha,\beta}, k^{\mathrm{IG}}_{\beta,\eta}, k^{\mathrm{Ga}}_{\beta,\eta}, k^{\mathrm{CPE}}_{\beta,\eta}$ for the positive stable, Inverse Gaussian, Gamma, and CPE distribution, respectively, we obtain

$$\begin{aligned}
k^{\mathrm{St}}_{\alpha,\beta}(s) &= \beta\,\frac{\alpha}{\Gamma(1-\alpha)}\,s^{-\alpha},\\
k^{\mathrm{IG}}_{\beta,\eta}(s) &= \sqrt{2}\,\beta\,\frac{1}{2\,\Gamma(0.5)}\,s^{-0.5}\,e^{-\frac{\eta^2}{2}s},\\
k^{\mathrm{Ga}}_{\beta,\eta}(s) &= \beta\,e^{-\eta\,s},\\
k^{\mathrm{CPE}}_{\beta,\eta}(s) &= \beta\,\eta\,s\,e^{-\eta\,s}, \quad s>0.
\end{aligned}$$

It is easy to see that $k^{\mathrm{St}}_{\alpha,\beta}, k^{\mathrm{IG}}_{\beta,\eta}$, and $k^{\mathrm{Ga}}_{\beta,\eta}$ are decreasing. Therefore, the related Bernstein functions are self-decomposable and can be used as Laplace exponents to construct both Lévy and Sato subordinators. In contrary to that, $k^{\mathrm{CPE}}_{\beta,\eta}$ is strictly increasing in a neighborhood sufficiently close to zero, implying that $\Psi^{\mathrm{CP}}_{\beta,\eta}$ is a non-self-decomposable Bernstein function.

3 Exchangeable exogenous shock models

Having the basic mathematical tools at hand, we are turning the focus to exchangeable exogenous shock models in the following. The chapter is primarily based on [Mai, Schenk, Scherer (2015b)], however enhances the results by some of the theoretical findings in [Mai, Schenk, Scherer (2015a)] and a new proof idea of the central Theorem 3.3.1. Section 3.1 defines exchangeable exogenous shock models and shows related work by other authors. In the bivariate case, this class has already been investigated and characterized in literature. The respective findings and an example are outlined in Section 3.2. The centerpiece of this chapter is constituted by the general d-variate characterization results (see Theorem 3.3.1) in Section 3.3. As with many studies of multivariate distribution functions, "true" multidimensionality starts with $d \geq 3$, meaning that the analytic approach in the bivariate proof can not be raised to higher dimension in a straightforward way and demands for alternative concepts. Two distinct proofs of (parts of) Theorem 3.3.1 are given, both being technical at a first glance, yet relying on rather vivid ideas. Having accomplished the characterization, properties of exchangeable exogenous shock models are depicted in Section 3.4. An interesting subclass consisting of extendible exogenous shock models is introduced in Section 3.5. This subclass is represented by an alternative stochastic model involving additive processes. The link between copulas, additive processes, and the related results on Bernstein functions and complete monotonicity proves quite powerful in the remaining chapters. A first "foretaste" is given in Section 3.6, where the recent results on Marshall–Olkin distributions depicted in Section 2.3.3 are recalled and established in an alternative way by means of Theorem 3.3.1.

3.1 Definition and literature overview

Whenever talking about exogenous shock models in the sequel, we refer to random vectors $(X_1, \ldots, X_d)$ that can be represented as follows.

Definition 3.1.1 (Exogenous shock model)
A random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ *is called exogenous shock model if*

$$X_k = \min\{Z^E : k \in E\}, \quad k = 1, \dots, d, \tag{3.1}$$

for $2^d - 1$ *independent real-valued random variables* Z^E *(called shocks),* $\emptyset \neq E \subseteq \{1, \dots, d\}$.

The term "shock model" has already been motivated in the introduction. Thinking of $(X_1, \dots, X_d)$ as a vector of lifetimes of companies/individuals, the shocks Z^E are typically random variables on $[0, \infty)$ and may refer to economic, political, or social disturbances affecting several lifetimes at once. Lifetime X_k is simply defined as the first arrival time of a shock hitting component k.

Exogenous shock models have been extensively analyzed in the literature. The most prominent example is given by the multivariate distribution introduced in Marshall and Olkin (1967) (see also Section 2.3.3). The authors consider a random vector $(X_1, \dots, X_d)$ as constructed in (3.1), where the shocks Z^E are exponentially distributed. Sarhan and Balakrishnan (2007) relax this condition in the bivariate case by considering univariate shocks $Z^{\{1\}}, Z^{\{2\}}$ with generalized exponential distribution. Kundu and Gupta (2009) extend this idea and study the law of the vector (X_1, X_2), where $X_1 = \max\{Z_1, Z_3\}$, $X_2 = \max\{Z_2, Z_3\}$, and both Z_1, Z_2, and Z_3 are independent random variables with generalized exponential distribution[1]. Proceeding in a similar way, Li and Pellerey (2011) consider construction (3.1) with the shocks Z^E having arbitrary distribution functions on $(0, \infty)$, and study the resulting dependence structure in the bivariate case. Relying on a method of the eponymous authors, Shoaee and Khorram (2012) introduce the Block and Basu bivariate generalized exponential distribution, which results from decomposing the joint distribution function in its singular and absolutely continuous part and solely considering the latter. Another direction is pursued by Kundu and Gupta (2014). The authors start with the construction in (3.1) for $d = 2$ and consider Weibull distributed random variables Z^E. Denoting by F the resulting joint distribution function of (X_1, X_2), they study the law of (Y_1, Y_2) given by

$$Y_1 = \min\{X^{\{1,1\}}, \dots, X^{\{1,N\}}\}, \quad Y_2 = \min\{X^{\{2,1\}}, \dots, X^{\{2,N\}}\},$$

where N is a mixing variable with geometric distribution and $(X^{\{1,n\}}, X^{\{2,n\}})_{n \in \mathbb{N}}$ are i.i.d. random vectors with distribution function F.

[1] Keeping in mind Corollary 2.2.8, analyzing the maximum construction in terms of its copula is equivalent to studying the minimum definition by means of the survival copula.

Considering an exogenous shock model for practical applications has two immediate advantages: On the one side, this class is easy to interpret. For an insurance company modeling a portfolio of d policies, $(X_1, \dots, X_d)$ might represent the occurrence times of claim reports. Identifying Z^E with external magnitudes of influence such as natural catastrophes, epidemics, or military conflicts, exogenous shock models offer a natural framework for depicting these systemic effects. On the other hand, the independence of shocks Z^E in the model construction makes this class tractable. When being given the survival functions of the shocks, calculating the joint survival function of $(X_1, \dots, X_d)$ is straightforward as for $x_1, \dots, x_d \in \mathbb{R}$,

$$\begin{aligned}&\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) = \mathbb{P}\big(\min\{Z^E : k \in E\} > x_k, k = 1, \dots, d\big)\\ =&\mathbb{P}\big(Z^E > \max\{x_k, k \in E\}, \emptyset \neq E \subseteq \{1, \dots, d\}\big) = \prod_E \mathbb{P}\big(Z^E > \max\{x_k, k \in E\}\big). \quad (3.2)\end{aligned}$$

The suitability for various applications explains the ongoing research interest concerning exogenous shock models. Marshall and Olkin (1967) show that in their setup with exponentially distributed shocks, the joint survival function of $(X_1, \dots, X_d)$ can be linked to the first jump times of independent Poisson processes affecting either one or several components of the random vector. This derivation is picked up in Linkskog and McNeil (2003), where (X_1, X_2) is identified with the first occurrence time of windstorm losses in France, respectively Germany, that are governed by west (affecting only France), central (affecting only Germany), and pan-European (affecting both countries) windstorms. The idea of modeling insurance events by independent counting processes and mapping the events to one or several claim types is a popular approach in multivariate insurance models (see the extensive overview in Anastasiadis and Chukova (2012)). Alternatively, applying exogenous shock models in a credit risk context, one might also think of Z^E as arrival times of economic catastrophes influencing the default of one or several assets in a portfolio. This idea is applied, e.g., in Giesecke (2003). In Baglioni and Cherubini (2013), the authors rely on the Marshall–Olkin setup (i.e. exponentially distributed shocks) and model cross-country dependencies between European obligors in order to compute the systemic risk of the banking sector in the post-crisis era.

However, the apparent advantages vanish, if not turn into major challenges for large portfolios. Interpreting and simulating the $2^d - 1$ shocks in a reasonable way becomes more and more difficult for increasing dimension d. Determining the distribution of Z^E, for instance by calibration to market data, becomes highly involved, if not impossible. In addition to that, identifying and characterizing reasonable subclasses of exogenous shock models is quite challenging. When analyzing the joint distribution of $(X_1, \dots, X_d)$

and validating its (survival) copula, the crucial condition to verify is d-increasingness (see Definition 2.2.2). However, this property is typically non-trivial to check, with the complexity of the problem increasing exponentially in higher dimensions. Thus, it is not surprising that while the universe of exogenous shock models in the literature is wide, most of them are restrictive in the sense that the distribution functions of the shocks Z^E are specified in a certain way and/or the vector $(X_1, \ldots, X_d)$ is only studied in the bivariate case, i.e. for $d = 2$ (compare e.g. the examples above, Kundu et al. (2014), and the references therein).

Given that we are aiming at applications in high-dimensional portfolios in credit risk and insurance, we are dealing with the subclass of exchangeable exogenous shock models in the present thesis. Before discussing pros and cons of this crucial structural restriction, we show how the exchangeability property stated in Definition 2.2.11 translates to the shock model class in (3.1).

Proposition 3.1.2 (Exchangeability of exogenous shock models)
Let $(X_1, \ldots, X_d)$ be an exogenous shock model as given in (3.1). Denote by $\bar{F}^E$ the survival function of Z^E and by $\bar{F}$ the survival function of $(X_1, \ldots, X_d)$. Then $(X_1, \ldots, X_d)$ is exchangeable if and only if $\operatorname{supp}(\bar{F}) = I^d$ *for $I \subseteq \mathbb{R}$ and*

$$\bar{F}^D(x) = \bar{F}^E(x) \quad \text{for all } x \in I \text{ and } \emptyset \neq D, E \subseteq \{1, \ldots, d\} \text{ with } |D| = |E|.$$

Proof
Let $(X_1, \ldots, X_d)$ be exchangeable. It follows that $\bar{F}$ is invariant w.r.t. its input arguments, which immediately implies the claimed form of the support. For the remainder, we proceed similarly to the proof of Lemma 3.1.1 in Mai (2010), where the special case of exponentially distributed shocks Z^E is considered. Using Equation (3.2) and the exchangeability of $(X_1, \ldots, X_d)$, it holds that for all $j \in \{2, \ldots, d\}$ and $x \in I$,

$$\prod_{E \neq \{1\}} \bar{F}^E(x) = \mathbb{P}(X_2 > x, \ldots, X_d > x) = \mathbb{P}(X_k > x, k \in \{1, \ldots, d\} \setminus \{j\}) = \prod_{E \neq \{j\}} \bar{F}^E(x).$$

Dividing by $\prod_E \bar{F}^E(x) > 0$ and comparing the reciprocals yields $\bar{F}^{\{1\}}(x) = \ldots = \bar{F}^{\{d\}}(x)$. Now assume that $\bar{F}^D(x) = \bar{F}^E(x)$ for all $x \in I$ and $\emptyset \neq D, E \subseteq \{1, \ldots, d\}$ with $|D| = |E| \leq k \in \{1, \ldots, d-2\}$. For an arbitrary subset $E_0 \subseteq \{1, \ldots, d\}$ with $|E_0| = k + 1$, it analogously holds that

$$\begin{aligned} \prod_{E \not\subseteq E_0} \bar{F}^E(x) &= \mathbb{P}(X_k > x, k \in \{1, \ldots, d\} \setminus E_0) \\ &= \mathbb{P}(X_{k+2} > x, \ldots, X_d > x) = \prod_{E \not\subseteq \{1, \ldots, k+1\}} \bar{F}^E(x), \quad x \in I. \end{aligned}$$

Again dividing by $\prod_E \bar{F}^E(x)$ *and comparing the reciprocals results into*

$$\bar{F}^{E_0}(x) \prod_{E \subset E_0} \bar{F}^E(x) = \bar{F}^{\{1,\dots,k+1\}}(x) \prod_{E \subset \{1,\dots,k+1\}} \bar{F}^E(x).$$

By induction hypothesis, $\prod_{E \subset E_0} \bar{F}^E(x)$ *only depends on the cardinality of* E_0*. Consequently,* $\bar{F}^{E_0}(x) = \bar{F}^{\{1,\dots,k+1\}}(x)$ *and one part of the claim is established.*

The converse is straightforward. Simplifying (3.2)*, the joint survival function equals*

$$\begin{aligned}\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) &= \prod_E \bar{F}^E\big(\max\{x_k, k \in E\}\big) \\ &= \prod_{j=1}^{d} \prod_{E:|E|=j} \bar{F}^{\{1,\dots,j\}}\big(\max\{x_k, k \in E\}\big).\end{aligned}$$

Considering the inner expression, there are $l-1$ *over* $j-1$ *sets* E *with cardinality* j *that contain* $x_{(l)}, l \in \{j, \dots, d\}$*, as its largest element. As a result,*

$$\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) = \prod_{j=1}^{d} \prod_{l=j}^{d} \big(\bar{F}^{\{1,\dots,j\}}(x_{(l)})\big)^{\binom{l-1}{j-1}} = \prod_{l=1}^{d} \prod_{j=1}^{l} \big(\bar{F}^{\{1,\dots,j\}}(x_{(l)})\big)^{\binom{l-1}{j-1}}, \tag{3.3}$$

which indicates the independence of the joint survival function w.r.t. permutation of the input arguments. □

Proposition 3.1.2 implies that w.l.o.g., the shock survival functions in exchangeable exogenous shock models can be altered such that $\bar{F}^D(x) = \bar{F}^E(x)$ for all $x \in \mathbb{R}$ and $|D| = |E|$. We always assume that the model is given in that way in the following. Put differently, the class of exchangeable exogenous shock models arises from considering (3.1) with the distribution of Z^E solely depending on the cardinality of E. It is beyond debate that focusing on this subclass is a strong limitation in terms of flexibility. It means that while the number of random sources is still the same ($2^d - 1$ random variables Z^E), their structure simplifies massively and becomes rather homogeneous. All shocks affecting a fixed number of components X_k are identically distributed, ignoring the possibility of asymmetry between the idiosyncratic risk factors in concern. When modeling a small portfolio of various financial instruments (i.e. stocks, bonds, commodities), the exchangeability assumption is justifiably in question.

However, for large portfolios, dealing with the increasing complexity of $(X_1, \dots, X_d)$ demands for simplifying assumptions or alternative approaches. If one is only interested

in functionals $f(X_1, \ldots, X_d)$, $f : \mathbb{R}^d \to \mathbb{R}$, of the random vector, a popular top-down ansatz is to model the distribution of $f(X_1, \ldots, X_d)$ directly (see e.g. Giesecke et al. (2010b,a)). However, in general, it is difficult to derive the induced stochastic model for the underlying vector $(X_1, \ldots, X_d)$, which might be an important aspect for understanding the model and its dependence structure. When starting with a bottom-up approach and modeling the random variables X_k themselves, exchangeability is one possibility to reduce complexity and maintain the interpretability of $(X_1, \ldots, X_d)$. The severity of the disadvantages resulting from homogeneity depends on the application in concern. For insurance portfolios consisting of numerous policies of similar type, homogeneity might not only be acceptable, but even desirable and natural. The same argument can be applied to large credit portfolios of a bank or standardized multiname credit derivatives which are examined later.

In addition to this application-sensitive justification, another reason for considering the subclass of exchangeable exogenous shock models is their mathematical tractability. This manuscript reveals deep results on the connection between exchangeable exogenous shock models and copulas $C : [0,1]^d \to [0,1], d \geq 2$, of the form

$$C(u_1, \ldots, u_d) = \prod_{k=1}^{d} g_k(u_{(k)}), \tag{3.4}$$

where $g_1 = \mathrm{id}_{[0,1]}$ is the identity on $[0,1]$ and the mappings $g_k : [0,1] \to [0,1]$ satisfy $g_k(1) = 1, k = 2, \ldots, d$. Within the framework of exchangeable exogenous shock models, Proposition 3.1.3 shows that the corresponding set of survival copulas has form (3.4). For uniqueness of the copula, we consider the case of continuous marginal distribution functions.

Proposition 3.1.3 (Survival copula of exchangeable exogenous shock models) *Let $\boldsymbol{X} := (X_1, \ldots, X_d)$ be an exchangeable exogenous shock model. Assume that the distribution functions of the random variables Z^E are continuous. The (unique) survival copula C of $\boldsymbol{X}$ has the form* (3.4).

Proof
Let Z^E have survival function $\bar{F}^E$. Due to exchangeability, we can assume that $\bar{F}^E(x) = \bar{F}^{\{1,\ldots,m\}}(x)$ for all $x \in \mathbb{R}$ and shocks Z^E with $|E| = m$. By (3.3)*, the marginal survival functions of $\boldsymbol{X} = (X_1, \ldots, X_d)$ equal*

$$\bar{F}_1(x) := \mathbb{P}(X_k > x) = \prod_{j=1}^{d} \left(\bar{F}^{\{1,\ldots,j\}}(x)\right)^{\binom{d-1}{j-1}}, \quad x \in \mathbb{R}.$$

The survival copula C of $(X_1, \ldots, X_d)$ is given by

$$\begin{aligned} C(u_1, \ldots, u_d) &= \mathbb{P}\big(X_1 \geq \bar{F}_1^{-1}(u_1), \ldots, X_d \geq \bar{F}_1^{-1}(u_d)\big) \\ &\overset{(*)}{=} \prod_{l=1}^{d} \prod_{j=1}^{l} \Big(\bar{F}^{\{1,\ldots,j\}}\big(\bar{F}_1^{-1}(u_{(d+1-l)})\big)\Big)^{\binom{l-1}{j-1}} \\ &= \prod_{l=1}^{d} \prod_{j=1}^{d+1-l} \Big(\bar{F}^{\{1,\ldots,j\}}\big(\bar{F}_1^{-1}(u_{(l)})\big)\Big)^{\binom{d-l}{j-1}}, \quad u_1, \ldots, u_d \in [0,1], \end{aligned}$$

where we have used Equation (3.3) *in* $(*)$. *Defining*

$$g_l : [0,1] \to [0,1], \quad g_l(u) := \prod_{j=1}^{d+1-l} \Big(\bar{F}^{\{1,\ldots,j\}}\big(\bar{F}_1^{-1}(u)\big)\Big)^{\binom{d-l}{j-1}}, \quad l = 1, \ldots, d,$$

the copula has the claimed form. In particular, $g_1 = \mathrm{id}_{[0,1]}$ *by the definition of* $\bar{F}_1^{-1}$. $\square$

The difficult part is to establish the converse and hence to show that any copula of form (3.4) is the survival copula of an exchangeable exogenous shock model. In this regard, the characterizing Theorem 3.3.1 is the main contribution to be derived in the present chapter. To get accustomed to the general proceeding of identifying necessary and sufficient conditions for $g_1, \ldots, g_d$, the following section analyzes copulas of type (3.4) in the simpler bivariate case.

3.2 Characterization in the bivariate case

In Durante et al. (2007), the authors study functions of type (3.4) for the special case $g_2 = g_3 = \ldots = g_d$ and derive necessary and sufficient conditions on g_2 such that (3.4) defines a copula. In the bivariate case, such copulas coincide with the more general class considered in the present thesis. Consequently, it is reasonable to use the available characterization results for $d = 2$ in order to get an idea of how to proceed for arbitrary $d \geq 2$. When checking Definition 2.2.2, it is straightforward to see that groundedness of $C(u_1, u_2) := u_{(1)}\, g_2(u_{(2)})$ is fulfilled for any g_2, while uniform margins require $g_2(1) = 1$. For 2-increasingness, one has to make sure that $\mathrm{d}C$ assigns non-negative mass to any 2-box $[u_1, v_1] \times [u_2, v_2] \subseteq [0,1]^2$. As C is exchangeable by definition such that $\mathrm{d}C\big([u_1, v_1] \times [u_2, v_2]\big) = \mathrm{d}C\big([u_2, v_2] \times [u_1, v_1]\big)$, it suffices to consider 2-boxes with $v_1 \leq v_2$. Thus, one has to distinguish between the three possible cases $(a)\, u_1 < v_1 \leq u_2 < v_2$, $(b)\, u_1 \leq u_2 < v_1 \leq v_2$, and $(c)\, u_2 \leq u_1 < v_1 \leq v_2$.

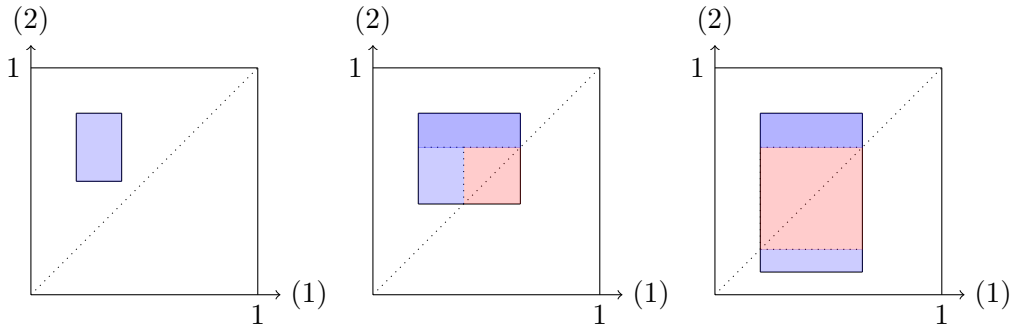

Figure 3.1 Decomposition of 2-boxes $[\boldsymbol{u}, \boldsymbol{v}]$ into squares around the diagonal (red) and 2-boxes exclusively located (blue) in the upper or lower triangular (separated by dotted line) of the unit square. The three plots illustrate the cases (from left to right) $(a)\, u_1 < v_1 \leq u_2 < v_2$, $(b)\, u_1 \leq u_2 < v_1 \leq v_2$, and $(c)\, u_2 \leq u_1 < v_1 \leq v_2$

.

Figure 3.1 visualizes these 2-boxes and highlights a useful decomposition as considered in (Durante et al., 2007, Theorem 3). It becomes obvious (and can be shown analytically) that any 2-box can be split into a non-overlapping combination of squares on the diagonal (marked red in Figure 3.1) and 2-boxes that entirely lie above or below the diagonal (marked blue). Due to exchangeability of C, the latter two cases coincide. Thus, C is a copula if and only if $\mathrm{d}C$ assigns non-negative mass to two types of 2-boxes. For blue boxes of the form $[u_1, v_1] \times [u_2, v_2], v_1 \leq u_2$, respectively the red boxes $[u, v]^2$, one demands

$$(i)\ \ \mathrm{d}C\big([u_1, v_1] \times [u_2, v_2]\big) = (v_1 - u_1)\,\big(g_2(v_2) - g_2(u_2)\big) \geq 0,$$

$$(ii)\ \ \mathrm{d}C\big([u, v]^2\big) = v\, g_2(v) - 2\, u\, g_2(v) + u\, g_2(u) \geq 0.$$

While already providing an analytical characterization of C, these conditions can be translated to continuity and monotonicity requirements for g_2 and used to derive the stochastic model behind the copula. It is clear that (i) is valid if and only if g_2 is increasing on $(0, 1]$. A simple computation (which is carried out later in the proof of Theorem 3.3.1) indicates that (ii) implies g_2 to be strictly positive and continuous on $(0, 1]$. Last but not least, the proof of (Durante et al., 2008, Theorem 4) shows that in order for (ii) to hold, $u \mapsto g_2(u)/u$ must be decreasing on $(0, 1]$. Characterizing exchangeable exogenous shock models in the bivariate case, we are briefly reciting the computations.

The first step in Durante et al. (2008) is to artificially rewrite Condition (*ii*) as

$$\frac{g_2(u)}{u} \geq \frac{g_2(v)}{v} - g_2(v)\,\frac{(v-u)^2}{u^2\,v}. \tag{3.5}$$

If $u \mapsto g_2(u)/u$ is not decreasing, there exist $0 < u_0 < v_0 \leq 1$ such that

$$\frac{g_2(v_0)}{v_0} - \frac{g_2(u_0)}{u_0} \geq a(u_0, v_0)\,(v_0 - u_0)$$

for a positive constant $a(u_0, v_0) > 0$. With g_2 being continuous, it follows that for an arbitrary $\epsilon \in (0, v_0 - u_0]$, there are $u_\epsilon, v_\epsilon \in [u_0, v_0]$, $u_\epsilon = v_\epsilon - \epsilon$, satisfying

$$\frac{g_2(v_\epsilon)}{v_\epsilon} - \frac{g_2(u_\epsilon)}{u_\epsilon} \geq a(u_0, v_0)\,(v_\epsilon - u_\epsilon) = a(u_0, v_0)\,\epsilon.$$

Plugging u_ϵ, v_ϵ into (3.5) and replacing $g_2(v_\epsilon)/v_\epsilon$ by the latter inequality yields

$$\begin{aligned} & \frac{g_2(u_\epsilon)}{u_\epsilon} \geq \frac{g_2(v_\epsilon)}{v_\epsilon} - g_2(v_\epsilon)\,\frac{\epsilon^2}{u_\epsilon^2\,v_\epsilon} \geq \frac{g_2(u_\epsilon)}{u_\epsilon} + a(u_0, v_0)\,\epsilon - g_2(v_\epsilon)\,\frac{\epsilon^2}{u_\epsilon^2\,v_\epsilon} \\ \Leftrightarrow \quad & g_2(v_\epsilon)\,\frac{\epsilon^2}{u_\epsilon^2\,v_\epsilon} \geq a(u_0, v_0)\,\epsilon. \end{aligned}$$

As a final step, noting that $g_2(v_\epsilon)/(u_\epsilon^2\,v_\epsilon) \leq g_2(v_0)/u_0^3$, the left-hand side of the last inequality is quadratic in ϵ, while the one on the right-hand side is linear. Consequently, for sufficiently small ϵ, the demanded inequality is violated and $u \mapsto g_2(u)/u$ necessarily has to be decreasing.

This monotonicity requirement can be naturally interpreted from a probabilistic point of view: In (Durante et al., 2007, Remark 4), it is shown that C is the copula of $\mathbf{U} = (U_1, U_2)$, where

$$\begin{aligned} U_1 &= \max\{V^{\{1\}}, V^{\{1,2\}}\}, \\ U_2 &= \max\{V^{\{2\}}, V^{\{1,2\}}\}, \end{aligned}$$

and $V^{\{1\}}, V^{\{2\}}, V^{\{1,2\}}$ are independent random variables with distribution functions[2] $V^{\{1\}}, V^{\{2\}} \sim g_2$ and $V^{\{1,2\}} \sim g_1/g_2$. To verify this stochastic model, note that for $u_1, u_2 \in [0,1]$,

$$\begin{aligned} \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) &= \mathbb{P}(\max\{V^{\{1\}}, V^{\{1,2\}}\} \leq u_1, \max\{V^{\{2\}}, V^{\{1,2\}}\} \leq u_2) \\ &= \mathbb{P}(V^{\{1\}} \leq u_1, V^{\{2\}} \leq u_2, V^{\{1,2\}} \leq u_{(1)}) \\ &= \mathbb{P}(V^{\{1\}} \leq u_1)\,\mathbb{P}(V^{\{2\}} \leq u_2)\,\mathbb{P}(V^{\{1,2\}} \leq u_{(1)}) \\ &= g_2(u_1)\,g_2(u_2)\,\frac{g_1(u_{(1)})}{g_2(u_{(1)})} = g_1(u_{(1)})\,g_2(u_{(2)}). \end{aligned}$$

[2] Strictly speaking, in order for g_2 and g_1/g_2 to be proper distribution functions on $[0,1]$, one has to consider their right-continuous extensions at zero.

For an arbitrary continuous survival function $\bar{F}$ on $\mathbb{R}$, the generalized inverse $\bar{F}^{-1}$ is strictly decreasing on $[0,1]$ (see (Embrechts and Hofert, 2013, Proposition 2.3)). Defining $Z^E := \bar{F}^{-1}(V^E)$ and applying Corollary 2.2.8 on strictly monotone transformations, C is the survival copula of (X_1, X_2), where

$$\begin{aligned} X_1 &:= \bar{F}^{-1}(U_1) = \min\left\{\bar{F}^{-1}\big(V^{\{1\}}\big), \bar{F}^{-1}\big(V^{\{1,2\}}\big)\right\} = \min\{Z^{\{1\}}, Z^{\{1,2\}}\}, \\ X_2 &:= \bar{F}^{-1}(U_2) = \min\{Z^{\{2\}}, Z^{\{1,2\}}\}. \end{aligned}$$

This insight complements Proposition 3.1.3 in the bivariate case and shows that any copula of type (3.4) is the survival copula of an exchangeable exogenous shock model such that (3.4) not only comprises, but precisely consists of the set of survival copulas corresponding to exchangeable exogenous shock models.

Several dependence properties of C can be derived in closed form. For instance (see also (Nelsen, 2006, Theorem 5.4.2, p. 214)), the lower and upper tail dependence coefficients λ_L and λ_U equal

$$\begin{aligned} \lambda_L &= \lim_{u \searrow 0} \frac{C(u,u)}{u} = \lim_{u \searrow 0} g_2(u), \\ \lambda_U &= \lim_{u \nearrow 1} \frac{C(u,u) - 2\,u + 1}{1-u} = 1 - g_2'(1-), \end{aligned}$$

where $g_2'(1-)$ denotes the left-sided derivative of g_2 at $u = 1$, which exists by monotonicity of g_2. Further dependence properties are derived in Durante (2006); Durante et al. (2007), including measures of association and extremal dependence coefficients. For instance, the authors show that Kendall's tau τ_C respectively Spearman's rho ρ_C are given by

$$\begin{aligned} \tau_C &= 4 \int_0^1 u\, g_2^2(u)\, \mathrm{d}u - 1, \\ \rho_C &= 12 \int_0^1 u^2\, g_2(u)\, \mathrm{d}u - 3. \end{aligned}$$

Furthermore, given $(U_1, U_2) \sim C$, the stochastic model for (U_1, U_2) allows to derive the probability $\mathbb{P}(U_1 = U_2) > 0$, showing that the induced measure $\mathrm{d}C$ of C possesses a

singular component. It holds that[3]

$$\begin{aligned}\mathbb{P}(U_1 = U_2) &= \mathbb{P}\Big(\max\{V^{\{1\}}, V^{\{1,2\}}\} = \max\{V^{\{2\}}, V^{\{1,2\}}\}\Big) \\ &= \mathbb{P}\Big(\underbrace{\max\{V^{\{1\}}, V^{\{2\}}\}}_{\sim g_2^2} \leq \underbrace{V^{\{1,2\}}}_{\sim g_1/g_2}\Big) \\ &= g_2^2(0) + \int_0^1 \Big(1 - \frac{u}{g_2(u)}\Big)\, 2\, g_2(u)\, \mathrm{d}g_2(u) \\ &= g_2^2(0) + 2\int_0^1 (g_2(u) - u)\, \mathrm{d}g_2(u).\end{aligned}$$

By simple application of integration by parts, it follows that

$$\mathbb{P}(U_1 = U_2) = 2\int_0^1 g_2(u)\, \mathrm{d}u - 1. \tag{3.6}$$

In addition to that, (Durante, 2006, Theorem 3) points out that C is radially symmetric if and only if $C = \alpha\,\Pi + (1-\alpha)\,M$ for some $\alpha \in [0,1]$, i.e. if and only if C is a linear combination of the independence and the comonotonicity copula (recall Section 2.2.1 for the definition of these copulas). A generalization of this radial symmetric copula to higher dimensions is given via the Dirichlet copula introduced in Section 3.5.

As an example of C, define $g_2 : [0,1] \to [0,1]$ by

$$g_2(u) := \min\{a\,u + b, 1 - c + c\,u\}, \quad a > 1, b > 0, c > 0, \quad b + c \leq 1.$$

Specified in that way, g_2 starts at b with slope a, has a kink at $u = (1-(b+c))/(a-c)$, continues to increase with slope c, and ends at $g_2(1) = 1$. Applying the tail dependence formulas, it follows that $\lambda_L = \lim_{u \searrow 0} g_2(u) = b$ and $\lambda_U = 1 - g_2'(1-) = 1 - c$. Thus, copulas of type (3.4) can admit both positive upper and lower tail dependence with arbitrary values in $[0,1]$.

Concerning simulation, arbitrary bivariate copulas C can be simulated by *conditional sampling*. Being monotonic in each argument as a distribution function, C is partially differentiable almost everywhere. For $(U_1, U_2) \sim C$ and $u_1, u_2 \in [0,1]$, it holds that

$$\begin{aligned}\mathbb{P}(U_1 \leq u_1 | U_2 = u_2) &:= \lim_{\epsilon \searrow 0} \mathbb{P}(U_1 \leq u_1 | U_2 \in [u_2 - \epsilon, u_2 + \epsilon]) \\ &= \lim_{\epsilon \searrow 0} \frac{C(u_1, u_2 + \epsilon) - C(u_1, u_2 - \epsilon)}{\epsilon} = \frac{\partial}{\partial\, u_2} C(u_1, u_2),\end{aligned}$$

[3] Once again, we consider the right-continuous extension of g_2 and g_1/g_2 at zero to guarantee proper distribution functions on $[0,1]$.

provided the derivative exists, meaning that the conditional distribution function of U_1 given U_2 is the (right-continuous extension of the) partial derivative of C (see Darsow et al. (1992)). Consequently, a simple two-step routine can be carried out as follows.

Algorithm 3.2.1 (Simulating arbitrary bivariate copulas)

1. *Sample a uniformly distributed random variable U_2 on $[0,1]$.*
2. *Sample a uniformly distributed random variable $\tilde{U}_1$ on $[0,1]$, independent of U_2. Define the function $F_{U_1|U_2} : [0,1] \to [0,1]$, $F_{U_1|U_2}(u) := \partial/(\partial\, u_2)\, C(u,u_2)_{|u_2=U_2}$, set $U_1 := F^{-1}_{U_1|U_2}(\tilde{U}_1)$, and return the vector $(U_1, U_2) \sim C$.*

Applying the algorithm to a copula of the form $C(u_1,u_2) = g_1(u_{(1)})\, g_2(u_{(2)})$ yields

$$\frac{\partial}{\partial\, u_2} C(u,u_2) = \begin{cases} g_2(u), & u_2 \leq u, \\ u\, g_2'(u_2), & u < u_2, \end{cases}$$

wherever g_2' is well-defined. As a consequence, $F_{U_1|U_2}$ has a jump at U_2 with jump height $g_2(U_2) - U_2\, g_2'(U_2)$. The induced probability of U_1 and U_2 taking the same values on $[0,1]$ is thus given by

$$\mathbb{P}(U_1 = U_2) = \int_0^1 \big(g_2(u_2) - u_2\, g_2'(u_2)\big)\, \mathrm{d}u_2 = 2 \int_0^1 g_2(u_2)\, \mathrm{d}u_2 - 1,$$

establishing the link to Equation (3.6).

3.3 General characterization

The essential property of a d-variate function C to serve as distribution function of a probability law is d-increasingness or, equivalently, complete monotonicity when viewed as a mapping from the semigroup $([0,\infty)^d, \wedge)$ to $\mathbb{R}$, see Ressel (2011), both properties being non-trivial to verify in general. Though having identified the meaning of copulas with form (3.4) in the bivariate case, deriving a general characterization for arbitrary $d \geq 2$ turns out to be rather complicated for two major reasons. On the one hand, finding a suitable decompositon of d-boxes as in Figure 3.1 is not straightforward. While it seems reasonable to consider cubes around the diagonal by analogy with the (red) squares in dimension two, there is no obvious way how to treat the remaining building blocks. On the other hand, even when assuming to have found a useful fragmentation along with necessary and sufficient non-negativity conditions of the form $\mathrm{d}C([\mathbf{u},\mathbf{v}]) \geq 0$, it is not clear how to translate these inequalities into monotonicity properties of the

functions $g_1, \dots, g_d$ (if possible at all) and how to relate them to a stochastic model for the copula. The existing results in the bivariate case are based on purely analytical reasoning and are hard to adapt in arbitrary dimension. We overcome both difficulties in the following theorem, which characterizes copulas of type (3.4) and establishes the connection to general exchangeable exogenous shock models. For notational brevity, we introduce a particular set of distribution functions on $[0,1]$, denoted $\mathcal{D}$ and defined by

$$\mathcal{D} := \Big\{ F : [0,1] \to [0,1] : F \text{ continuous and increasing}, F(1) = 1, \\ F \text{ strictly positive on } (0,1] \Big\}.$$

Theorem 3.3.1 (Characterization of exchangeable exogenous shock models) *Let $C : [0,1]^d \to [0,1]$ have form (3.4). The following statements are equivalent:*

(i) C is a copula, i.e. a multivariate distribution function.

(ii) For all $0 < u < v \leq 1, k \in \mathbb{N}_0, j \in \mathbb{N}$, with $k + j \leq d$, it holds that

$$G_{j,k}(u,v) := \sum_{i=0}^{j} \binom{j}{i} (-1)^i \prod_{l=1}^{i} g_{l+k}(u) \prod_{l=i+1}^{j} g_{l+k}(v) \geq 0.$$

(iii) For all $k \in \mathbb{N}_0, j \in \mathbb{N}$, with $k + j \leq d$, it holds that $H_{j,k} \in \mathcal{D}$, where

$$H_{j,k}(u) := \begin{cases} \prod_{i=0}^{j-1} g_{k+1+i}^{(-1)^i \binom{j-1}{i}}(u), & u \in (0,1], \\ \lim_{v \searrow 0} \prod_{i=0}^{j-1} g_{k+1+i}^{(-1)^i \binom{j-1}{i}}(v), & u = 0. \end{cases}$$

(iv) For all $m \in \{1, \dots, d\}$, it holds that $H_{m,d-m} \in \mathcal{D}$.

In this case, C is the distribution function of $(U_1, \dots, U_d)$, where

$$U_k := \max\{V^E : k \in E\}, \quad k = 1, \dots, d,$$

and V^E are independent random variables with $V^E \sim H_{m,d-m}$ for all subsets E with cardinality $|E| = m$.

Remark 3.3.2 (Implications of Theorem 3.3.1)

1. *Theorem 3.3.1 consists of three crucial achievements: First of all, in (ii), copulas of type (3.4) are characterized purely analytically and it is shown that the functions $g_k, k = 1, \dots, d$, have to satisfy certain inequality conditions. Second, in (iii) respectively (iv), these conditions are translated to monotonicity requirements, showing that certain functionals of the g_k have to yield distribution functions on*

$[0,1]$. *Last but not least, these functionals are interpreted from a probabilistic point of view by introducing a stochastic model, which manifests the connection to shock models.*

2. *It is important to note that the $H_{m,d-m}$ defined in Theorem 3.3.1.(iii) can be solved for the g_k, yielding*

$$g_k = \prod_{m=1}^{d+1-k} H_{m,d-m}^{\binom{d-k}{m-1}}, \quad k = 1, \dots, d.$$

Consequently, $H_{m,d-m} \in \mathcal{D}$, $m = 1, \dots, d$, can be arbitrary distribution functions, provided that the normalization constraint

$$g_1 = \prod_{m=1}^{d} H_{m,d-m}^{\binom{d-1}{m-1}} = \mathrm{id}_{[0,1]},$$

which solely stems from the formulation of the theorem in terms of copulas rather than general multivariate distribution functions, is fulfilled.

3. *The $g_k, k = 1, \dots, d$, can be interpreted as conditional distribution functions. More precisely, for $u \in (0,1]$ and $(U_1, \dots, U_d) \sim C$,*

$$\begin{aligned} g_k(u) &= \frac{\prod_{i=1}^{k} g_i(u)}{\prod_{i=1}^{k-1} g_i(u)} = \frac{\mathbb{P}(U_1 \leq u, \dots, U_k \leq u)}{\mathbb{P}(U_1 \leq u, \dots, U_{k-1} \leq u)} \\ &= \mathbb{P}(U_1 \leq u, \dots, U_k \leq u | U_1 \leq u, \dots, U_{k-1} \leq u). \end{aligned}$$

Note that for $k \geq 3$, g_k corresponds to the ratio between two diagonal sections of copulas. As such, due to the Lipschitz continuity of copulas mentioned in Section 2.2, it is the ratio between a k-Lipschitz and a $(k-1)$-Lipschitz function.

4. *The distribution functions $H_{j,k}$ defined in Theorem 3.3.1.(iii) can be constructed via the "monotonicity triangle" visualized in Figure 3.2. The first row of the triangle simply consists of the functions $g_1, \dots, g_d$. The cells in the subsequent rows arise as the quotient of the top and top-right elements in the row above. Depicted like that, the function $H_{j,k}$ corresponds to the j-th row and $(k+1)$-th column of the triangle. Consequently, the functions $H_{m,d-m}, m = 1, \dots, d$, equal the colored diagonal in the respective figure.*

Theorem 3.3.1.(ii) indicates that for functions of form (3.4), the d-increasingness conditions can be massively simplified and reduce to the verification of $G_{j,k}(u,v) \geq 0$ for certain indices $k \in \mathbb{N}_0, j \in \mathbb{N}$, and certain pairs $(u,v) \in [0,1]^2$, $u \leq v$. Apart from some

g_1 g_2 g_3 $\cdots$ g_d $|E| = 1$

g_1/g_2 g_2/g_3 $\cdots$ g_{d-1}/g_d $|E| = 2$

$g_1\, g_3/g_2^2$ $\cdots$ $g_{d-2}\, g_d/g_{d-1}^2$ $|E| = 3$

$\vdots$

$\prod_{i=0}^{d-1} g_{i+1}^{\binom{d-1}{i}(-1)^i}$ $|E| = d$

Figure 3.2 Distribution functions $H_{j,k}$ in Theorem 3.3.1.(iii) in terms of the functions g_k. The blue cells mark the distribution functions $H_{m,d-m}$ appearing in the stochastic model for $(U_1, \ldots, U_d) \sim C$.

technical lemmata, the proof of Theorem 3.3.1 provides valuable insights into the structure of the objects $G_{j,k}$ and $H_{j,k}$ and their relation to the stochastic model in Equation (3.1). We are going to show that $(iv) \Rightarrow (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv)$. The central ideas can be summarized as follows.

Remark 3.3.3 (Structure of the proof of Theorem 3.3.1)

$(iv) \Rightarrow (i)$ *Starting with the random vector* $\boldsymbol{U} = (U_1, \ldots, U_d)$ *given in the theorem, one can compute that each* $U_k, k = 1, \ldots, d$, *is uniformly distributed on* $[0, 1]$ *and that* C *is the distribution function of* $\boldsymbol{U}$, *hence a copula.*

$(i) \Rightarrow (ii)$ *Being a copula,* C *induces a probability measure* $\mathrm{d}C$ *on* $[0, 1]^d$. *It can be deduced that* $G_{j,k}(u, v)$ *corresponds to the mass assigned by* $\mathrm{d}C$ *to certain subsets of* $[0, 1]^d$. *Therefore, it has to be greater than or equal to zero.*

$(ii) \Rightarrow (iii)$ *This is the most difficult and lengthy part of the proof. Besides minor technical conditions, the central task is to show that non-negativity of* $G_{j,k}(u, v), u < v$, *implies increasingness of* $H_{j,k}$. *The underlying proof idea is to split up* $G_{j,k}(u, v)$ *into two summands, one involving the difference* $H_{j,k}(v) - H_{j,k}(u)$, *the other one corresponding to the probability mass* $\mathrm{d}C(I)$ *induced by a copula* C *of type* (3.4) *for a subset* $I \subset [0, 1]^d$. *For a sufficiently "small" subset* I, *it is shown that the sign of* $G_{j,k}(u, v)$ *is dominated by the first part, i.e. the difference* $H_{j,k}(v) - H_{j,k}(u)$. *Thus, for* $G_{j,k}(u, v)$ *to be non-negative,* $H_{j,k}(v) - H_{j,k}(u)$ *has to be greater than or equal to zero, which establishes the claimed increasingness of* $H_{j,k}$.

(iii) ⇒ (iv) This is trivial as (iv) is a special case of (iii).

Proof (Proof of Theorem 3.3.1)

(iv) ⇒ (i): If (iv) holds, the functions $H_{m,d-m}$ defined in Theorem 3.3.1.(iii) are valid distribution functions on $[0,1]$, and we can consider the corresponding stochastic model given in the theorem. We will show that the resulting distribution function of $(U_1,\dots,U_d)$ is a copula given by C. First of all, recognize that each $U_k, k=1,\dots,d$, has a uniform marginal distribution due to

$$\begin{aligned}\mathbb{P}(U_k \le u) &= \prod_{E:k\in E} \mathbb{P}\big(V^E \le u\big) \overset{(*)}{=} \prod_{m=1}^{d} H_{m,d-m}^{\binom{d-1}{m-1}}(u)\\ &= \prod_{m=1}^{d}\Bigg(\prod_{i=0}^{m-1} g_{d-m+1+i}^{(-1)^i\binom{m-1}{i}}(u)\Bigg)^{\binom{d-1}{m-1}}\\ &= \prod_{m=1}^{d}\prod_{i=0}^{m-1} g_{d-m+1+i}^{(-1)^i\binom{m-1}{i}\binom{d-1}{m-1}}(u)\\ &\overset{(**)}{=} \prod_{k=1}^{d} g_k^{\sum_{m=d+1-k}^{d}(-1)^{k+m-d-1}\binom{m-1}{k+m-d-1}\binom{d-1}{m-1}}(u) = u.\end{aligned}$$

The equality in $(*)$ stems from the fact that there are $d-1$ over $m-1$ shocks V^E with cardinality m and distribution function $H_{m,d-m}$ that appear in the stochastic construction of U_k. The equality in $(**)$ is yielded by grouping the g_k, i.e. regarding all $g_{d-m+1+i}$ with $d-m+1+i=k$, which is the same as setting $i=k+m-1-d$. For $k=1$, it is apparent that the exponent in the last line is equal to one. For $k\ge 2$, due to

$$\binom{m-1}{k+m-d-1}\binom{d-1}{m-1} = \frac{(d-1)!}{(d-k)!\,(k-1)!}\binom{k-1}{d-m},$$

it follows that the exponent equals

$$\begin{aligned}&\frac{(d-1)!}{(d-k)!\,(k-1)!}\sum_{m=d+1-k}^{d}(-1)^{k+m-d-1}\binom{k-1}{d-m}\\ =&\frac{(d-1)!}{(d-k)!\,(k-1)!}\sum_{m=0}^{k-1}(-1)^m\binom{k-1}{m} = 0.\end{aligned}$$

Secondly, consider the joint distribution function of $(U_1,\dots,U_d)$. For $u_1,\dots,u_d\in(0,1]$, it is given by

$$\mathbb{P}(U_k\le u_k, k=1,\dots,d) = \prod_{\emptyset\neq E\subseteq\{1,\dots,d\}} \mathbb{P}\big(V^E \le \min\{u_k : k\in E\}\big). \tag{3.7}$$

Among all subsets E with cardinality $|E| = m$, there are $d-1$ over $m-1$ elements where $\min\{u_k : k \in E\} = u_{(1)}$. Analogously, there are $d-k$ choose $m-1$ elements where $\min\{u_l : l \in E\} = u_{(k)}$, $k \in \{2, \dots, d-m+1\}$. Thus, (3.7) is equal to

$$\begin{aligned}
&\prod_{m=1}^{d} \prod_{k=1}^{d-m+1} \mathbb{P}\big(V^E \le u_{(k)}, |E| = m\big)^{\binom{d-k}{m-1}} \\
&= \prod_{m=1}^{d} \prod_{k=1}^{d-m+1} H_{m,d-m}^{\binom{d-k}{m-1}}(u_{(k)}), \\
&= \prod_{m=1}^{d} \prod_{k=1}^{d-m+1} \prod_{i=0}^{m-1} g_{d-m+1+i}^{(-1)^i \binom{m-1}{i}\binom{d-k}{m-1}}(u_{(k)}) \\
&= \prod_{k=1}^{d} \prod_{m=1}^{d-k+1} \prod_{i=0}^{m-1} g_{d-m+1+i}^{(-1)^i \binom{m-1}{i}\binom{d-k}{m-1}}(u_{(k)}) \\
&\overset{(*)}{=} \prod_{k=1}^{d} \underbrace{\prod_{n=k}^{d} g_n^{\sum_{m=d+1-n}^{d+1-k} (-1)^{m-1+n-d}\binom{m-1}{m-1+n-d}\binom{d-k}{m-1}}(u_{(k)})}_{\text{should be equal to } g_k(u_{(k)})}.
\end{aligned}$$

Now $(*)$ can be derived by sorting all $g_{d-m+1+i}$ with $d-m+1+i = n$, i.e. $i = m-1+n-d$. For $n = k$, it becomes obvious that the exponent of g_n is equal to one. For $n \in \{k+1, \dots, d\}$, by using the same deliberations as for the derivation of the marginal distributions, the exponent of g_n is given by

$$\begin{aligned}
&\frac{(d-k)!}{(d-n)!\,(n-k)!} \sum_{m=d+1-n}^{d+1-k} (-1)^{m-1+n-d} \binom{n-k}{m-1+n-d} \\
&= \frac{(d-k)!}{(d-n)!\,(n-k)!} \sum_{m=0}^{n-k} (-1)^m \binom{n-k}{m} = 0.
\end{aligned}$$

Summing up, we have

$$\mathbb{P}(U_k \le u_k, \forall k = 1, \dots, d) = \prod_{k=1}^{d} g_k(u_{(k)}) = C(u_1, \dots, u_d),$$

and we can conclude that C is a copula.

$(i) \Rightarrow (ii)$: Let $(U_1, \dots, U_d)$ be a random vector with copula C in (3.4) as distribution function. Moreover, assume for a moment that the $g_k, k = 2, \dots, d$, are strictly

positive on $(0,1]$. *Then, for* $u, v \in (0,1], u < v$, $G_{j,k}$ *has the representation*

$$G_{j,k}(u,v) = \frac{1}{\prod_{m=1}^{k} g_m(u)} \Big(\mathbb{P}(A_\emptyset) - \sum_{i=1}^{j} (-1)^{i+1} \sum_{\substack{L \subseteq \{k+1,\dots,k+j\}: \\ |L|=i}} \mathbb{P}\Big(\bigcap_{l \in L} A_l \Big) \Big),$$

$$A_l := \Big(\bigcap_{m \in \{1,\dots,k,l\}} \{U_m \le u\} \Big) \cap \Big(\bigcap_{m \in \{k+1,\dots,k+j\} \setminus \{l\}} \{U_m \le v\} \Big)$$

$$A_\emptyset := \{U_1 \le u, \dots, U_k \le u, U_{k+1} \le v, \dots, U_{k+j} \le v\}.$$

Applying the principle of inclusion and exclusion (see Lemma 2.2.5), we have

$$\begin{aligned} G_{j,k}(u,v) &= \frac{1}{\prod_{m=1}^{k} g_m(u)} \Big(\mathbb{P}(A_\emptyset) - \mathbb{P}\Big(\bigcup_{l=k+1}^{k+j} A_l \Big) \Big) \\ &= \frac{1}{\prod_{m=1}^{k} g_m(u)} \, \mathbb{P}\Big(U_1 \le u, \dots, U_k \le u, U_{k+1} \in [u,v], \dots, U_{k+j} \in [u,v] \Big) \\ &= \mathbb{P}\Big(U_{k+1} \in [u,v], \dots, U_{k+j} \in [u,v] \Big| U_1 \le u, \dots, U_k \le u \Big) \ge 0. \end{aligned}$$

Strict positivity of g_k *(which we have assumed so far) as well as continuity on* $(0,1]$ *for* $k = 2, \dots, d$ *can be shown by induction. To begin with, assume that there is a* $u^* := \sup\{u \ge 0 : g_2(u) = 0\} > 0$. *As* C *is a copula and hence continuous, it follows that* $g_2(u^*) = 0$, *such that for* $v > u^*$,

$$\mathbb{P}\Big(U_1 \in [u^*, v], U_2 \in [u^*, v] \Big) = G_{2,0}(u^*, v) = v \, g_2(v) - 2\, u^* \, g_2(v) < 0$$

for v *sufficiently close to* u^*. *This is a contradiction and hence* $g_2(u) > 0$ *for* $u \in (0,1]$. *Similarly, to show continuity, assume that there is a* $v^* \in (0,1]$ *such that* $g_2(v^*-) := \lim_{u \nearrow v^*} g_2(u) < g_2(v^*)$. *Then*

$$\begin{aligned} 0 \le \lim_{u \nearrow v^*} G_{2,0}(u, v^*) &= \lim_{u \nearrow v^*} \big(v^* \, g_2(v^*) - 2\, u \, g_2(v^*) + u \, g_2(u) \big) \\ &= -v^* \, g_2(v^*) + v^* \, g_2(v^*-) < 0, \end{aligned}$$

which is a contradiction. Hence, there is no such v^* *and* g_2 *is left-continuous on* $(0,1]$. *Analogously, if* $g_2(u^*+) := \lim_{v \searrow u^*} > g_2(u^*)$ *for an* $u^* \in (0,1)$, $G_{2,0}(u^*, v)$ *becomes negative for sufficiently small* $v > u^*$. *Consequently,* g_2 *is continuous on* $(0,1]$.

For the induction step $k-1 \mapsto k$, *note that*

$$G_{2,k-1}(u,v) = \mathbb{P}\Big(U_k \in [u,v], U_{k+1} \in [u,v] \Big| U_1 \le u, \dots, U_{k-1} \le u \Big)$$

induces $0 \leq G_{2,k-1}(u,v) = g_{k-1}(v)\, g_k(v) - 2\, g_{k-1}(u)\, g_k(v) + g_{k-1}(u)\, g_k(u)$. *By the same arguments as for the induction start, this implies that* g_k *is both strictly positive and continuous on* $(0,1]$.

$(ii) \Rightarrow (iii)$*: To begin with, we establish an auxiliary lemma that – apart from the outlined interpretation of* $G_{j,k}$ *as a conditional distribution function – provides another perspective on* $G_{j,k}$*. However, note that in the proof of "*$(i) \Rightarrow (ii)$*" above, we solely require that* C *is a copula and do not assume increasingness of* $H_{m,k+j-m}$*, which is why we can not apply the alternative interpretation in the respective case.*

Lemma 3.3.4 (Alternative interpretation of $\boldsymbol{G_{j,k}}$***)***
If $H_{m,k+j-m}$ *in Theorem 3.3.1.(iii) is an element of* $\mathcal{D}$ *for all* $m \in \{1,\dots,j\}$*, one has*

$$\begin{aligned} &G_{j,k}(u,v) = \mathbb{P}(X_{k+1} \in [u,v], \dots, X_{k+j} \in [u,v]), \quad \text{where} \\ &X_l := \max\{Z^E : k \in E\}, \quad l = k+1,\dots,k+j, \quad \text{and} \\ &Z^E \sim H_{m,k+j-m} \text{ for } |E| = m, \end{aligned}$$

with independent random variables $Z^E, \emptyset \neq E \subseteq \{k+1,\dots,k+j\}$.

Proof
Similarly to the proof of "$(iv) \Rightarrow (i)$*" above, the distribution function of* $(X_{k+1},\dots,X_{k+j})$ *is given by*

$$\mathbb{P}(X_{k+1} \leq x_1, \dots, X_{k+j} \leq x_j) = \prod_{i=1}^{j} g_{k+i}(x_{(i)}), \quad x_1,\dots,x_j \in [0,1].$$

Consequently, by applying the principle of inclusion and exclusion in the very same way as in the proof of "$(i) \Rightarrow (ii)$*", the claim follows.* □

Let $G_{j,k}(u,v) \geq 0$ *for all* $0 < u < v \leq 1, k \in \mathbb{N}_0, j \in \mathbb{N}$ *with* $k + j \leq d$*. By the proof of "*$(i) \Rightarrow (ii)$*" above, this induces* g_k*,* $k = 2,\dots,d$ *to be strictly positive and continuous on* $(0,1]$*. The idea of this part of the proof is to establish a connection between* $G_{j,k}$ *and a related stochastic model similar to Lemma 3.3.4 in order to derive reasonable estimates that help to extract the required conditions in* (iii)*. For readability, we are going to proceed by induction.*

Suppose that we have already shown that for a $j-1 \in \{1,\dots,d\}$*, the conditions* $G_{i,k}(u,v) \geq 0$ *for all* $1 \leq i \leq j-1$ *and* $k+i \leq d$ *imply that* $H_{i,k}$ *is increasing for all* $1 \leq i \leq j-1$ *and* $k+i \leq d$*. For* $j = 2$*, this is obviously satisfied and the induction basis is established. In order to carry out the induction step* $j-1 \mapsto j$*, we need to show that* $H_{j,k}$ *is increasing*

for all $k+j \leq d$ ($H_{i,k}$ for $i \leq j-1$ are increasing by induction hypothesis). This is shown in several steps.

Step 1 (Main observation): There is a useful decomposition of $G_{j,k}$ that we require below.

Lemma 3.3.5 (Decomposition of $G_{j,k}$)
Instead of $G_{j,k}$, write $G_{g_{k+1},\dots,g_{k+j}}$ to emphasize the dependence of $G_{j,k}$ on the functions $g_{k+1},\dots,g_{k+j}$. It holds that

$$\begin{aligned} G_{g_{k+1},\dots,g_{k+j}}(u,v) = \tilde{g}_{k+1}(v)\, g_{k+2}(v)\, \dots\, g_{k+j}(v) \left(\frac{g_{k+1}(v)}{\tilde{g}_{k+1}(v)} - \frac{g_{k+1}(u)}{\tilde{g}_{k+1}(u)}\right) \\ + \frac{g_{k+1}(u)}{\tilde{g}_{k+1}(u)}\, G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}(u,v), \quad 0<u<v\leq 1, \end{aligned} \tag{3.8}$$

for an arbitrary function $\tilde{g}_{k+1}$ that is unequal to zero on $(0,1)$.

Proof
The decomposition consists of nothing else than changing the last summand of $G_{g_{k+1},\dots,g_{k+j}}$ (ending up with the last line in Equation (3.8)) and adding the resulting difference as an extra term (corresponding to the first line in Equation (3.8)). The non-zero condition for $\tilde{g}_{k+1}$ is required for well-definedness of the quotients. □

Define $\tilde{g}_{k+1} := g_{k+1}/H_{j,k}$, which is continuous and strictly positive on $(0,1]$ as seen earlier, and note that Lemma 3.3.5 then yields

$$\begin{aligned} 0 \leq G_{g_{k+1},\dots,g_{k+j}}(u,v) = &\tilde{g}_{k+1}(v)\, g_{k+2}(v)\, \dots\, g_{k+j}(v) \Big(H_{j,k}(v) - H_{j,k}(u)\Big) \\ &+ H_{j,k}(u)\, G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}(u,v). \end{aligned} \tag{3.9}$$

We want to conclude that $H_{j,k}(v) \geq H_{j,k}(u)$. Therefore, we have to prove that the second summand is not responsible for non-negativity of $G_{g_{k+1},\dots,g_{k+j}}$. The crucial consequence of (3.9) is that $G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}$ can be related to a stochastic model. To this end, we want to apply Lemma 3.3.4 to $G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}$. In order to do so, one has to make sure that the corresponding functions $\tilde{H}_{m,k+j-m}$ (which are defined just like $H_{m,k+j-m}$ in Theorem 3.3.1.(iii), however with replacing g_{k+1} by $\tilde{g}_{k+1}$) are distribution functions in $\mathcal{D}$ for all $m=1,\dots,j$. Due to the definition of $\tilde{g}_{k+1}$, it holds that

$$\tilde{H}_{m,k+j-m} = \begin{cases} H_{m,k+j-m}, & \text{for } m=1,\dots,j-1, \\ 1, & \text{for } m=j. \end{cases}$$

As $H_{m,k+j-m}$ are distribution functions for $m=1,\dots,j-1$ by induction hypothesis and $\tilde{H}_{m,k+j-m}$ is a degenerated distribution function for $m=j$, the requirements of Lemma

3.3.4 are satisfied.

Step 2 (Stochastic model): As a consequence of Lemma 3.3.4,

$$\begin{aligned}
&G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}(u,v) = \mathbb{P}\Big(X_{k+1},\dots,X_{k+j} \in [u,v]\Big), \quad \text{where}\\
&X_l := \max\Big\{Z^E, \emptyset \neq E \subset \{k+1,\dots,k+j\}, l \in E\Big\}, \quad l = k+1,\dots,k+j,\\
&Z^E \sim H_{m,k+j-m} \quad \text{for } |E| = m,
\end{aligned}$$

with independent random variables $Z^E, \emptyset \neq E \subset \{k+1,\dots,k+j\}$. Thus,

$$\begin{aligned}
&\mathbb{P}\Big(X_{k+1},\dots,X_{k+j} \in [u,v]\Big)\\
=&\mathbb{P}\Big(\underbrace{\bigcap_{l=k+1}^{k+j}\Big\{\max\big\{Z^E, E\cap\{l\}\neq\emptyset\big\} \in [u,v]\Big\}}_{:=A}\Big)
\end{aligned}$$

requires that all Z^E are less than or equal to v and – as there is no common shock with $|E| = j$ due to $\tilde{H}_{j,k}(x) \equiv 1$ – at least two $Z^I, Z^J, I, J \subset \{k+1,\dots,k+j\}, I \neq J$, need to be in the interval $[u,v]$. This implies that

$$A \subset \bigcup_{\substack{\emptyset\neq I,J\subset\{k+1,\dots,k+j\}\\ I\neq J}} \Big\{u \le Z^I, Z^J \le v\Big\}.$$

Moreover, as $\mathbb{P}(\cup_{i=1}^n A_i) \le \sum_{i=1}^n \mathbb{P}(A_i)$ for arbitrary $A_i \in \mathcal{F}$, and as there are (2^j-2) over 2 possibilities to pick $Z^I, Z^J \in [u,v]$, we have

$$\begin{aligned}
G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}(u,v) = \mathbb{P}(A) &\le \sum_{\substack{\emptyset\neq I,J\subset\{k+1,\dots,k+j\}\\ I\neq J}} \mathbb{P}\Big(\big\{u \le Z^I, Z^J \le v\big\}\Big)\\
&\le \underbrace{\binom{2^j-2}{2}}_{:=b} \max_{m=1,\dots,j-1}\Big\{\big(H_{m,k+j-m}(v) - H_{m,k+j-m}(u)\big)^2\Big\}.
\end{aligned} \tag{3.10}$$

Step 3 (Lipschitz-continuity): Using Equation (3.10), we are going to derive Lipschitz-continuity-type results for $G_{\tilde{g}_{k+1},g_{k+2},\dots,g_{k+j}}$. In order to do so, the following lemma is helpful.

Lemma 3.3.6

For $k \in \mathbb{N}_0, j \ge 2$, let $H_{1,k},\dots,H_{j,k} : (0,1] \to (0,1]$ and $H_{1,k+1},\dots,H_{j-1,k+1} : [0,1] \to [0,1]$ be increasing functions with $H_{l,k} = H_{l-1,k}/H_{l-1,k+1}$ for $l \in \{2,\dots,j\}$. Then it

holds that

$$0 \leq H_{j,k}(v) - H_{j,k}(u) \leq \Big(\prod_{l=1}^{j-1} \frac{1}{H_{l,k+1}(u)} \Big) \big(H_{1,k}(v) - H_{1,k}(u) \big).$$

Proof

For $j = 2$ and $k \in \mathbb{N}_0$, we have

$$\begin{aligned} 0 \leq & H_{2,k}(v) - H_{2,k}(u) = \frac{1}{H_{1,k+1}(u)} \Big(\underbrace{\frac{H_{1,k+1}(u)}{H_{1,k+1}(v)}}_{\leq 1} H_{1,k}(v) - H_{1,k}(u) \Big) \\ \leq & \frac{1}{H_{1,k+1}(u)} \big(H_{1,k}(v) - H_{1,k}(u) \big). \end{aligned}$$

For $j \mapsto j+1$, the claim follows by simple induction. □

Applying Lemma 3.3.6 to Equation (3.10), $G_{\tilde{g}_{k+1},\ldots,g_{k+j}}(u,v)$ has an upper bound

$$b \cdot \max_{m=1,\ldots,j-1} \Big\{ \Big(\prod_{l=1}^{m-1} \frac{1}{H_{l,k+j-m+1}(u)} \Big)^2 \big(\underbrace{g_{k+j-m+1}(v) - g_{k+j-m+1}(u)}_{= H_{1,k+j-m}(v) - H_{1,k+j-m}(u)} \big)^2 \Big\}. \tag{3.11}$$

This can be further simplified as

$$\begin{aligned} 0 \leq G_{2,0}(u,v) &= g_1(v)\, g_2(v) - 2\, g_1(u)\, g_2(v) + g_1(u)\, g_2(u) \\ &= g_2(v) \big(g_1(v) - g_1(u) \big) - g_1(u) \big(g_2(v) - g_2(u) \big) \\ \Leftrightarrow\ g_2(v) - g_2(u) &\leq \frac{g_2(v)}{g_1(u)} \big(g_1(v) - g_1(u) \big) = \frac{g_2(v)}{u} \big(v - u \big) \end{aligned}$$

and since g_1 is the identity by definition. Analogously, one has

$$0 \leq G_{2,k}(u,v) \Leftrightarrow\ g_{k+2}(v) - g_{k+2}(u) \leq \frac{g_{k+2}(v)}{g_{k+1}(u)} \big(g_{k+1}(v) - g_{k+1}(u) \big)$$

By induction over k, one can conclude that

$$g_k(v) - g_k(u) \leq \Big(\prod_{l=1}^{k-1} \frac{g_{l+1}(v)}{g_l(u)} \Big) (v - u) \quad \text{for all } k \text{ in concern.}$$

Applying this result, the expression in (3.11) is less than or equal to

$$\begin{aligned} & b \cdot \max_{m=1,\ldots,j-1} \Big\{ \Big(\prod_{l=1}^{m-1} \frac{1}{H_{l,k+j-m+1}(u)} \Big)^2 \Big(\prod_{l=1}^{k+j-m} \frac{g_{l+1}(v)}{g_l(u)} \Big)^2 (v-u)^2 \Big\} \\ = & \, p_{j,k}(u,v)\, (v-u)^2, \end{aligned}$$

with

$$p_{j,k}(u,v) := b \cdot \max_{m=1,\dots,j-1} \Big\{ \Big(\prod_{l=1}^{m-1} \frac{1}{H_{l,k+j-m+1}(u)} \Big)^2 \Big(\prod_{l=1}^{k+j-m} \frac{g_{l+1}(v)}{g_l(u)} \Big)^2 \Big\}.$$

Additionally, due to the monotonicity of the g_k and $H_{j,k}$ appearing in $p_{j,k}$, we can conclude that for any $u_0, v_0 \in (0,1], u_0 < v_0$, it holds that $p_{j,k}(u,v) \leq p_{j,k}(u_0,v_0)$ for all $u, v \in [u_0, v_0], u \leq v$. Combining all those observations, one ends up with

$$0 \leq G_{\tilde{g}_{k+1}, g_{k+2}, \dots, g_{k+j}}(u,v) \leq p_{j,k}(u_0,v_0)\,(v-u)^2, \quad u, v \in [u_0, v_0], u \leq v. \tag{3.12}$$

Step 4 (Proof by contradiction): Finally, we can proceed similarly to the proof in the bivariate case depicted in (Durante et al., 2008, p. 67). Assume that $H_{j,k}$ is not increasing and that there exist $u_0, v_0 \in (0,1], u_0 < v_0$, such that

$$H_{j,k}(v_0) - H_{j,k}(u_0) = -a(u_0,v_0)\,(v_0 - u_0)$$

for a positive constant $a(u_0,v_0) > 0$. Consequently, by continuity of the g_k and hence $H_{j,k}$, for every $\epsilon \in (0, v_0 - u_0]$, there are $u_\epsilon, v_\epsilon \in [u_0, v_0], u_\epsilon = v_\epsilon - \epsilon$, such that

$$H_{j,k}(v_\epsilon) - H_{j,k}(u_\epsilon) \leq -a(u_0,v_0)\,(v_\epsilon - u_\epsilon) = -a(u_0,v_0)\,\epsilon. \tag{3.13}$$

Independently of this assumption, splitting the positive and negative powers in $H_{j,k}$,

$$H_{j,k}(u) = \prod_{i=0}^{j-1} g_{k+1+i}^{(-1)^i \binom{j-1}{i}}(u) = \frac{\prod_{i=0}^{\lfloor \frac{j-1}{2} \rfloor} g_{k+1+2\,i}^{\binom{j-1}{2\,i}}(u)}{\prod_{i=0}^{\lfloor \frac{j-2}{2} \rfloor} g_{k+1+2\,i+1}^{\binom{j-1}{2\,i+1}}(u)}, \quad u > 0,$$

with "$\lfloor \cdot \rfloor$" denoting the floor function. Thus, for $u \in [u_0, v_0]$, it holds by the monotonicity of the g_k that

$$H_{j,k}(u) \leq \frac{\prod_{i=0}^{\lfloor \frac{j-1}{2} \rfloor} g_{k+1+2\,i}^{\binom{j-1}{2\,i}}(v_0)}{\prod_{i=0}^{\lfloor \frac{j-2}{2} \rfloor} g_{k+1+2\,i+1}^{\binom{j-1}{2\,i+1}}(u_0)} =: p_{\max}(u_0,v_0).$$

Plugging u_ϵ, v_ϵ into Equation (3.9) and using all previous results yields

$$\begin{aligned} 0 \leq G_{g_{k+1},\dots,g_{k+j}}(u_\epsilon, v_\epsilon) &= \underbrace{\tilde{g}_{k+1}(v_\epsilon)}_{= \frac{g_{k+1}(v_\epsilon)}{H_{j,k}(v_\epsilon)}}\, g_{k+2}(v_\epsilon) \,\dots\, g_{k+j}(v_\epsilon) \Big(\underbrace{H_{j,k}(v_\epsilon) - H_{j,k}(u_\epsilon)}_{\leq -a(u_0,v_0)\,\epsilon} \Big) \\ &\quad + H_{j,k}(u_\epsilon)\, G_{\tilde{g}_{k+1},\dots,g_{k+j}}(u_\epsilon, v_\epsilon) \\ &\overset{(3.12)}{\leq} \frac{g_{k+1}(u_0)}{p_{\max}(u_0,v_0)}\, g_{k+2}(u_0) \,\dots\, g_{k+j}(u_0)\, (-a(u_0,v_0)\,\epsilon) \\ &\quad + p_{\max}(u_0,v_0)\, p_{j,k}(u_0,v_0)\, \epsilon^2. \end{aligned}$$

Thus, for sufficiently small ϵ, $G_{j,k}$ becomes negative and yields a contradiction. Consequently, $H_{j,k}$ has to be increasing and the induction is complete.

(iii) $\Rightarrow$ (iv): Trivial, as (iv) is a special case of (iii) for $j = m, k = d - m$. $\square$

In order to relate the inequality constraints for $G_{j,k}$ in Theorem 3.3.1.(ii) to necessary and sufficient conditions for the copula-induced measure $\mathrm{d}C$ (as was done in the bivariate case), we provide an alternative proof of (i)$\Leftrightarrow$(ii) in Theorem 3.3.1 by means of a tricky decomposition of d-boxes. The following lemma is tailor-made for checking d-increasingness of functions with representation (3.4) as it splits d-boxes $[\mathbf{u}, \mathbf{v}]$ in a way that helps to exploit the factorial structure in (3.4).

Lemma 3.3.7 (Decomposition of d-boxes)
Consider two vectors $\boldsymbol{u} = (u_1, \ldots, u_d), \boldsymbol{v} = (v_1, \ldots, v_d) \in [0,1]^d, d \geq 2$, with $\boldsymbol{u} < \boldsymbol{v}$, $v_1 \leq \ldots \leq v_d$, and with $m := \min\left\{i \in \{1, \ldots, d\} : v_i > u_{(d)}\right\} \leq d$. Then $[\boldsymbol{u}, \boldsymbol{v}]$ has a decomposition into $2(d-m)+1$ non-overlapping d-boxes

$$
\begin{aligned}
[\boldsymbol{u}, \boldsymbol{v}] &= \bigcup_{l=m+1}^{d} A_l \cup \bigcup_{k=1}^{d-m} B_k \cup C, \quad \textit{where} \\
A_l &= \times_{i=1}^{m}[u_i, v_i] \times_{i=m+1}^{l-1} [u_i, v_m] \times [v_m, v_l] \times_{i=l+1}^{d} [u_i, v_i], \\
B_k &= \times_{i=1}^{m-1}[u_i, v_i] \times_{i=m}^{d} [u_i^{m,d,k}, v_i^{m,d,k}], \\
C &= \times_{i=1}^{m-1}[u_i, v_i] \times [u_{(d)}, v_m]^{d-m+1}, \\
u_i^{m,d,k} &= \begin{cases} u_i, & \textit{if } i \in \{m, \ldots, d\} \setminus \sigma_{m,d}(\{1, \ldots, k-1\}), \\ u_{(d)}, & \textit{else}, \end{cases} \\
v_i^{m,d,k} &= \begin{cases} u_{(d)}, & \textit{if } i = \sigma_{m,d}(k), \\ v_m, & \textit{else}, \end{cases}
\end{aligned}
$$

$\sigma_{m,d} : \{1, \ldots, d-m+1\} \to \{m, \ldots, d\}$ is a bijection ordering $(u_m, \ldots, u_d)$ such that $u_{\sigma_{m,d}(1)} \leq \ldots \leq u_{\sigma_{m,d}(d)}$.

Proof
In a first step, it is shown that the sets A_l, B_k, C appearing above are non-overlapping.

Consider $A_{l_1}, A_{l_2}, l_1 < l_2$.

- *The l_1-th component of A_{l_1} is given by $[v_m, v_{l_1}]$.*
- *The l_1-th component of A_{l_2} is given by $[u_{l_1}, v_m]$.*

Consider $B_{k_1}, B_{k_2}, k_1 < k_2$.

- *The* $\sigma_{m,d}^{-1}(k_1)$*-th component of* B_{k_1} *is given by* $[u_{\sigma_{m,d}^{-1}(k_1)}, u_{(d)}]$.
- *The* $\sigma_{m,d}^{-1}(k_1)$*-th component of* B_{k_2} *is given by* $[u_{(d)}, v_m]$.

Consider A_l, B_k.

- *The* l*-th component of* A_l *is given by* $[v_m, v_l]$.
- *The* l*-th component of* B_k *is a subset of* $[u_{\sigma_{m,d}^{-1}(1)}, v_m]$.

The sets A_l, C *(respectively* B_k, C*) can be treated analogously to the sets* A_l, B_k *(respectively* $B_{k_1}, B_{k_2}, k_1 < k_2$*) and it becomes obvious that all sets in the decomposition of* $[\boldsymbol{u}, \boldsymbol{v}]$ *are non-overlapping. In a second step, it is demonstrated that the union of the sets* A_l, B_k, C *yields the* d*-box* $[\boldsymbol{u}, \boldsymbol{v}]$*. To begin with, note that*

$$C \cup B_{d-m} = \times_{i=1}^{m-1}[u_i, v_i] \times [u_{(d)}, v_m]^{\sigma_{m,d}^{-1}(d-m)-1} \times [u_{\sigma_{m,d}^{-1}(d-m)}, v_m] \times [u_{(d)}, v_m]^{d-m+1-\sigma_{m,d}^{-1}(d-m)}.$$

Proceeding in the same manner, one recognizes that

$$C \cup \bigcup_{k=1}^{d-m} B_k = \times_{i=1}^{m}[u_i, v_i] \times_{i=m+1}^{d} [u_i, v_m].$$

Adding A_d *yields*

$$C \cup \bigcup_{k=1}^{d-m} B_k \cup A_d = \times_{i=1}^{m-1}[u_i, v_i] \times_{i=m}^{d-1} [u_i, v_m] \times [u_d, v_d].$$

Finally, continuing analogously, one ends up with

$$C \cup \bigcup_{k=1}^{d-m} B_k \cup \bigcup_{l=m+1}^{d} A_l = \times_{i=1}^{d}[u_i, v_i] = [\boldsymbol{u}, \boldsymbol{v}]. \qquad \square$$

Figure 3.3 illustrates the described decomposition of 3-boxes with $m \in \{1, 2\}$. Lemma 3.3.7 has a crucial consequence: Any d-box $[\mathbf{u}, \mathbf{v}]$ with $\mathbf{u} < \mathbf{v}$, $v_1 \leq \ldots \leq v_d$, and with $m := \min\{i \in \{1, \ldots, d\} : v_i > u_{(d)}\} \in \{1, \ldots, d\}$ can be decomposed into non-overlapping d-boxes of types A_l, B_k, and C.

- Comparing $\times_{i=1}^{d}[u_i, v_i]$ to A_l, we note that

- u_l is replaced by v_m,
- $v_{m+1}, \ldots, v_{l-1}$ are replaced by $v_m, \ldots, v_m$.

- Comparing $\times_{i=1}^{d}[u_i, v_i]$ to B_k, we note that
  - $k-1$ elements of $\{u_m, \ldots, u_d\}$ are replaced by $u_{(d)}$,
  - one element of $\{v_m, \ldots, v_d\}$ is replaced by $u_{(d)}$, the remaining ones by v_m.

Thus, both A_l and B_k are d-boxes with an increased index for m. By recursively segmenting A_l and B_k further, we solely end up with d-boxes of type C for different values of m. Summing up, C is a copula if and only if $\mathrm{d}C$ assigns non-negative mass to d-boxes of the form

$$\times_{i=1}^{m-1}[u_i, v_i] \times [u_{(d)}, v_m]^{d-m+1}, \quad v_1 \leq \ldots \leq v_d, \quad v_{m-1} \leq u_{(d)} < v_m, \quad m = 1, \ldots, d,$$

with the convention $v_0 := 0$. The special structure of these d-boxes yields

$$\mathrm{d}C\big(\times_{i=1}^{m-1}[u_i, v_i] \times [u_{(d)}, v_m]^{d-m+1}\big) = \mathrm{d}C_{1,\ldots,m-1}\big(\times_{i=1}^{m-1}[u_i, v_i]\big) \cdot G_{d-m+1,m-1}(u_{(d)}, v_m) \tag{3.14}$$

for all $m = 1, \ldots, d$, where we set $\mathrm{d}C_{\emptyset} \equiv 1$. This relationship paves the way for a proof by induction. We have already shown in Section 3.2 that C is a bivariate copula if and only if for all $0 < u < v \leq 1, k \in \mathbb{N}_0, j \in \mathbb{N}$, with $k + j \leq 2$, it holds that $G_{j,k}(u, v) \geq 0$. Given that this relationship holds in dimensions $2, \ldots, d-1$, it follows by (3.14) that C is a copula in dimension d if and only if $\mathrm{d}C_{1,\ldots,m-1}([\mathbf{u}, \mathbf{v}]) \geq 0$ and $G_{d-m+1,m-1}(u, v) \geq 0$ for all $m = 1, \ldots, d$. This shows the claimed equivalence (ii)⇔(i) in Theorem 3.3.1.

3.4 Properties and classes

One may determine the intersection between copulas having form (3.4) and extreme-value copulas. Proposition 3.4.1 shows that extreme-value copulas of type (3.4) correspond to choosing power functions for g_k, where the sequence of exponents must be d-monotone.

Proposition 3.4.1 (Extreme-value copulas of type (3.4)***)***
Let C have the form (3.4)*. C is an extreme-value copula if and only if $g_k(u) = u^{a_{k-1}}, k = 1, \ldots, d$, for $u \in (0, 1]$ and a d-monotone sequence $\{a_0, \ldots, a_{d-1}\}$ with $a_0 = 1$.*

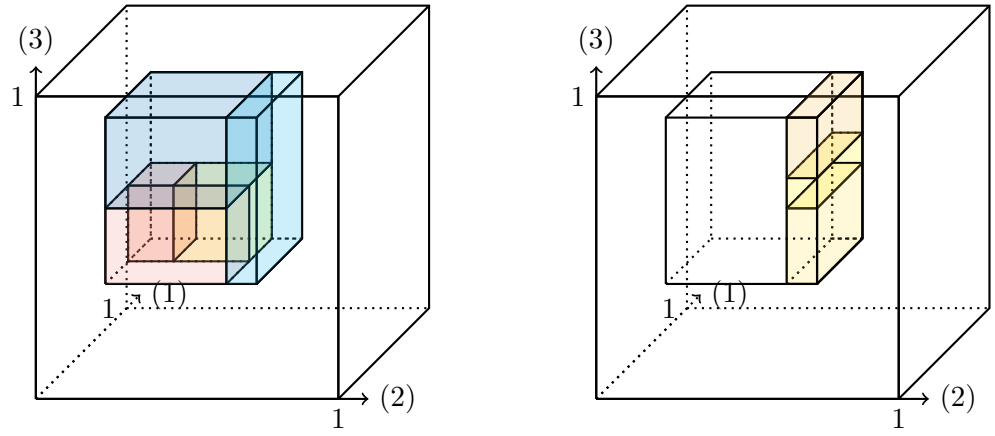

Figure 3.3 Iterative decomposition of d-boxes according to Lemma 3.3.7. The left figure shows a fragmentation of the d-box $[0.1, 0.6] \times [0.2, 0.7] \times [0.35, 0.9]$ with $m = 1$ into $2\,(3-m)+1 = 5$ d-boxes of types A_2, A_3 (marked blue), B_1, B_2 (red), and C (yellow). The right figure indicates the next iteration for A_2 (which now has an index $m = 2$) into 3 d-boxes of type C.

Proof

If $g_k(u) = u^{a_{k-1}}, k = 1, \ldots, d$, *for a sequence* $\{a_0, \ldots, a_{d-1}\}$, *the functions* $H_{j,k}$ *in Theorem 3.3.1.(iii) are given by*

$$H_{j,k}(u) = \prod_{i=0}^{j-1} g_{k+1+i}^{(-1)^i \binom{j-1}{i}}(u) = \prod_{i=0}^{j-1} u^{(-1)^i \binom{j-1}{i} a_{k+i}} = u^{\sum_{i=0}^{j-1} (-1)^i \binom{j-1}{i} a_{k+i}}.$$

Thus, the $H_{j,k}$ *are distribution functions in* $\mathcal{D}$ *(i.e.* C *is a copula) if and only if the sequence* $\{a_0, \ldots, a_{d-1}\}$ *is d-monotone. Moreover, it is apparent that due to the power function structure of the* g_k, C *satisfies the extreme-value property.*

It remains to show that any extreme-value copula of type (3.4) *implies a power function structure for the* g_k. *By setting* $u_1 = u_2 = u \in [0, 1]$ *and* $u_3 = \ldots = u_d = 1$ *in the extreme-value copula definition in* (2.6),

$$g_2(u^t) = g_2(u)^t \quad \text{for all } u \in (0, 1], t > 0.$$

Defining $\Theta : [0, \infty) \to \mathbb{R}, \Theta(x) := g_2(\exp(-x))$, *this is equivalent to*

$$\Theta(t\,x) = \Theta(x)^t \quad \text{for all } x, t > 0. \tag{3.15}$$

Thus, setting $t = n \in \mathbb{N}$ *and* $x = 1/n$, *it holds that*

$$\Theta(1) = \Theta\Big(\frac{1}{n}\Big)^n \Rightarrow \Theta\Big(\frac{1}{n}\Big) = \Theta(1)^{\frac{1}{n}}. \tag{3.16}$$

Consequently, for all $x \in \mathbb{Q} \cap (0,\infty)$, $x = p/q, p, q \in \mathbb{N}$, *we have*

$$\Theta\Big(\frac{p}{q}\Big) \overset{(3.15)}{=} \Theta\Big(\frac{1}{q}\Big)^p \overset{(3.16)}{=} \Theta(1)^{\frac{p}{q}}.$$

By Theorem 3.3.1.(iii), it holds that $H_{1,1} = g_2 \in \mathcal{D}$, *implying that* $g_2(u) > 0$ *for* $u \in (0,1]$. *Therefore, one deduces that* $0 < \Theta(1) = g_2(\exp(-1)) \leq 1$ *and the previous equation yields* $\Theta(x) = \exp(-a_1\, x)$ *for all* $x \in \mathbb{Q} \cap (0,\infty)$, *where* $a_1 = -\log\big(\Theta(1)\big) \geq 0$. *As* $\mathbb{Q} \cap (0,\infty)$ *is dense in* $(0,\infty)$, *it follows that*

$$g_2\big(e^{-x}\big) = \Theta(x) = e^{-a_1\, x} = \big(e^{-x}\big)^{a_1}, \quad x \in (0,\infty).$$

Iteratively, by setting $u_1 = \ldots = u_l = u \in (0,1]$, $u_{l+1} = \ldots = u_d = 1$, *and subsequently raising* l, *we conclude that* $g_k(u) = u^{a_{k-1}}, k = 1, \ldots, d$, *for parameters* $a_0, \ldots, a_{d-1} \geq 0, a_0 = 1$.[4] □

The corresponding class of extreme-value copulas is well-known in the literature and has already been introduced in Section 2.3.3: It is precisely the exchangeable subfamily of Marshall–Olkin survival copulas. In Section 3.6, we will investigate this example in more detail. Besides the extreme-value-property, the subclass of radially symmetric copulas corresponding to exchangeable exogenous shock models can be derived explicitly as well.

Theorem 3.4.2 (Radially symmetric copulas of type (3.4)***)***
Let C *have the form* (3.4) *and assume that the functions* $g_k, k = 1, \ldots, d$, *are differentiable. Then* C *is radially symmetric if and only if*

$$g_k(u) = \frac{c\,u + k - 1}{c + k - 1}, \quad u \in [0,1], k = 1, \ldots, d,$$

for some $c \in [0,\infty]$.

[4] An alternative way to constitute the power function structure of the g_k for extreme-value copulas of type (3.4) is via Pickands' Theorem as considered in Durante and Sempi (2010). Comparing the Pickands representation of C with the functional form in (3.4) on the diagonal $u = u_1 = \ldots = u_d$, it can be shown by induction that $g_k(u) = u^{a_{k-1}}$ for a parameter

$$a_{k-1} = k\, P\Big(\underbrace{\frac{1}{k}, \ldots, \frac{1}{k}}_{k \text{ times}}, 0, \ldots, 0\Big) - (k-1)\, P\Big(\underbrace{\frac{1}{k-1}, \ldots, \frac{1}{k-1}}_{(k-1) \text{ times}}, 0, \ldots, 0\Big)$$

depending on the Pickands dependence function P for fixed values. By the first part of the proof, the claim follows.

Proof

The proof can be found in [Mai, Schenk, Scherer (2015a)]. Assume that C is radially symmetric. Since C is invariant w.r.t. the ordering of its arguments, we may simplify notation by restricting our attention to preordered input arguments $0 \leq u_1 \leq u_2 \leq \ldots \leq u_d \leq 1$. Denoting $C_{1,\ldots,k} := \prod_{i=1}^{k} g_i(u_i)$, so that $C = C_{1,\ldots,d}$, radial symmetry of the exchangeable copula C is equivalent to (see Definition 2.2.15 and Proposition 2.2.7)

$$\begin{aligned} u_1\, g_2(u_2) \cdot \ldots \cdot g_d(u_d) &= C(u_1, \ldots, u_d) = \hat{C}(u_1, \ldots, u_d) \\ &= 1 + \sum_{k=1}^{d} (-1)^k \sum_{1 \leq j_1 < \ldots < j_k \leq d} C_{1,\ldots,k}(1 - u_{j_1}, \ldots, 1 - u_{j_k}). \end{aligned} \tag{3.17}$$

Applying the mixed partial derivative operator $(\partial/\partial u_1) \ldots (\partial/\partial u_d)$ on both sides of the last equation (which is possible due to differentiability of the g_k), apparently all but one summand of the right-hand side vanish, leaving us with

$$g_2'(u_2) \ldots g_d'(u_d) = g_2'(1 - u_{d-1}) \ldots g_d'(1 - u_1).$$

Iteratively, this implies that all $g_k', k = 1, \ldots, d$, are constants, hence g_k is linear with $g_k(1) = 1$, i.e. $g_k(u) = (1 - g_k(0))\, u + g_k(0)$. Thus, it already follows that g_2 has the claimed form $g_2(u) = (c\, u + 1)/(c + 1)$ for some $c := 1/g_2(0) - 1 \geq 0$. If $g_2(0) = 0$, then we conveniently interpret this as $c = \infty$, i.e. $C_{1,2}$ equals the independence copula. Due to Theorem 3.3.1.(iii), we know that for any copula C of the structural form above, $H_{3,k}(u) := \big(g_{k+1}(u)\, g_{k+3}(u)\big)/g_{k+2}^2(u)$ has to be increasing on $(0, 1]$ for all $k \in \{0, \ldots, d-3\}$. If $g_{k+1}(u) = g_{k+2}(u) = u$ for a certain k and if g_{k+3} is linear with $g_{k+3}(0) \geq 0$, increasingness of $H_{3,k}(u) = g_{k+3}(u)/u$ requires $g_{k+3}(0) = 0$ and consequently $g_{k+3}(u) = u$. Iteratively, for $c = \infty$, it follows that C equals the d-dimensional independence copula. Consequently, we will henceforth assume that $c < \infty$ and we are left with the task of computing $g_k(0)$ for $d \geq 3$. Evaluating Equation (3.17) for $u_1 = \ldots = u_d = u$ yields

$$\prod_{k=1}^{d} g_k(u) = 1 + \sum_{k=1}^{d} (-1)^k \binom{d}{k} \prod_{l=1}^{k} g_l(1 - u).$$

Differentiating the left-hand side w.r.t. u results in

$$\frac{\partial}{\partial u} \prod_{k=1}^{d} g_k(u) = \sum_{k=1}^{d} g_k'(u) \prod_{\substack{j=1 \\ j \neq k}}^{d} g_j(u).$$

Analogously, the right-hand side gives

$$\frac{\partial}{\partial u} \Big(1 + \sum_{k=1}^{d} (-1)^k \binom{d}{k} \prod_{l=1}^{k} g_l(1 - u)\Big) = \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \sum_{l=1}^{k} g_l'(1 - u) \prod_{\substack{j=1 \\ j \neq l}}^{k} g_j(1 - u).$$

Plugging in $u = 0$ on both sides and using that $g_1(0) = 0$, $g_k(1) = 1, k = 1, \ldots, d$, we end up with the equality

$$\prod_{k=2}^{d} g_k(0) = \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \sum_{l=1}^{k} g_l'(1). \tag{3.18}$$

Inductively, the only unknowns are $g_d(0)$ on the left-hand side and $g_d'(1)$ on the right-hand side. As we know that $g_d'(1) = 1 - g_d(0)$, this is a linear equation in $g_d(0)$. We are going to show by induction that $g_d(0) := (d-1)/(c+d-1)$ solves this equation. For $d = 3$, the statement is easy to verify. Thus, consider $d \rightsquigarrow d+1$ and set $g_{d+1}(0) := d/(c+d)$. The respective right-hand side of (3.18) then equals[5]

$$\sum_{k=1}^{d+1} (-1)^{k+1} \underbrace{\binom{d+1}{k}}_{=\binom{d}{k}+\binom{d}{k-1}} \sum_{l=1}^{k} g_l'(1) = \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \sum_{l=1}^{k} g_l'(1) + \sum_{k=0}^{d} (-1)^{k} \binom{d}{k} \sum_{l=1}^{k+1} g_l'(1)$$

$$= \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \sum_{l=1}^{k} g_l'(1) + 1 - \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \underbrace{\sum_{l=1}^{k+1} g_l'(1)}_{\sum_{l=1}^{k} g_l'(1) + g_{k+1}'(1)}$$

$$= 1 - \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} g_{k+1}'(1).$$

Summing up, we have to show that

$$\prod_{j=1}^{d} \frac{j}{c+j} = \prod_{k=2}^{d+1} g_k(0) = \sum_{k=1}^{d+1} (-1)^{k+1} \binom{d+1}{k} \sum_{l=1}^{k} g_l'(1) = 1 - \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \frac{c}{c+k}, \tag{3.19}$$

given that the equation is valid for arbitrary $c > 0$ when replacing d by $d-1$. Decomposing the binomial coefficient on the right-hand side of (3.19), it holds that[6]

$$1 - \sum_{k=1}^{d} (-1)^{k+1} \binom{d}{k} \frac{c}{c+k}$$

$$= \underbrace{1 - \sum_{k=1}^{d-1} (-1)^{k+1} \binom{d-1}{k} \frac{c}{c+k}}_{\text{apply (3.19) for } d-1} + \sum_{k=0}^{d-1} (-1)^{k+1} \binom{d-1}{k} \frac{c}{c+1+k}$$

[5] We use the convention $\binom{d}{d+1} = 0$ in the first equality.
[6] We use the convention $\binom{d-1}{d} = 0$ in the first equality.

$$
\begin{aligned}
&= \prod_{j=1}^{d-1} \frac{j}{c+j} - \frac{c}{c+1} + \underbrace{\sum_{k=1}^{d-1} (-1)^{k+1} \binom{d-1}{k} \frac{c}{c+1+k}}_{= \frac{c}{c+1} \sum_{k=1}^{d-1} (-1)^{k+1} \binom{d-1}{k} \frac{c+1}{c+1+k}} \\
&= \prod_{j=1}^{d-1} \frac{j}{c+j} - \frac{c}{c+1} \Big(\underbrace{1 - \sum_{k=1}^{d-1} (-1)^{k+1} \binom{d-1}{k} \frac{c+1}{c+1+k}}_{\text{apply (3.19) for } d-1 \text{ and } \hat{c} = c+1} \Big) \\
&= \prod_{j=1}^{d-1} \frac{j}{c+j} - \frac{c}{c+1} \prod_{j=1}^{d-1} \frac{j}{c+j+1} = \frac{(d-1)!\,(c+d) - c\,(d-1)!}{\prod_{j=1}^{d}(c+j)} \\
&= \frac{d!}{\prod_{j=1}^{d}(c+j)} = \prod_{j=1}^{d} \frac{j}{(c+j)}.
\end{aligned}
$$

□

As a plausibility check, recall that in Section 3.2, we have cited a known result concerning radial symmetry in the bivariate case. It states that a bivariate copula of type (3.4) is radially symmetric if and only if $C = \alpha\,\Pi + (1-\alpha)\,M$ for some $\alpha \in [0,1]$, i.e. $C(u_1,u_2) = \alpha\,u_1\,u_2 + (1-\alpha)\,u_{(1)}$ for $u_1,u_2 \in [0,1]$. By defining c such that $c/(c+1) = \alpha$ ($\alpha = 1$ corresponds to the limiting case $c \to \infty$), one can rewrite

$$
C(u_1,u_2) = \alpha\,u_1\,u_2 + (1-\alpha)\,u_{(1)} = u_{(1)}\,(\alpha\,u_{(2)} + 1 - \alpha) = u_{(1)}\,\frac{c\,u_{(2)}+1}{c+1},
$$

which is precisely the structure of the g_k required in Theorem 3.4.2. The fact that the radial symmetry result in the bivariate case does not demand differentiability of the copula suggests that Theorem 3.4.2 would also work without the differentiability assumption for the g_k. However, we forego this possible generalization as we are mostly dealing with differentiable functions g_k in the subsequent chapters.

3.5 Additive–frailty copulas

Theorem 3.3.1 has shown that any d-dimensional copula of the form (3.4) arises from the stochastic construction in (3.1), involving $2^d - 1$ random variables $Z^E, \emptyset \neq E \subseteq \{1,\ldots,d\}$. This section introduces an alternative construction for a subclass of type (3.4)-copulas based on a first-passage time construction with additive processes, which is useful in at least three regards: First, it allows to easily build new families of exogenous shock models by means of different specifications of the additive process. Second, it

provides an alternative viewpoint to the classical shock construction in (3.1) that can be helpful for interpreting the dependence structure on the one hand and deriving dependence measures and other properties on the other hand. Third, if the additive process can be simulated efficiently and accurately, the first-passage time model facilitates sampling, especially in large dimensions, as essentially only paths of a one-dimensional stochastic process have to be sampled.

Consider an additive subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with $\lim_{t\to\infty} \Lambda_t = \infty$ and define a sequence $\{X_k\}_{k\in\mathbb{N}}$ of random variables by

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k \in \mathbb{N}, \tag{3.20}$$

where $E_k, k \in \mathbb{N}$, are i.i.d. unit exponentially distributed random variables that are independent of $\{\Lambda_t\}_{t\geq 0}$. By construction, $\{X_k\}_{k\in\mathbb{N}}$ is an exchangeable sequence of random variables. The following proposition outlines that the (unique) survival copula of $(X_1, \ldots, X_d)$, denoted $C_{\Lambda,d}$ in the sequel, is of type (3.4) for any $d \geq 2$. In analogy to the terminology "Lévy–frailty copula" in Section 2.3.3, where $\{\Lambda_t\}_{t\geq 0}$ in (3.20) was specified as a Lévy subordinator, we call the more general class $C_{\Lambda,d}$ in Proposition 3.5.1 *additive–frailty copulas*. This term further underlines the interpretation of Λ as a joint risk factor impacting the random variables X_k. When referring to the sequence $\{X_k\}_{k\in\mathbb{N}}$ as a whole or its subvectors $(X_1, \ldots, X_d)$, we call the construction in (3.20) *additive–frailty model*.

Proposition 3.5.1 (Additive–frailty copulas)
Define a sequence $\{X_k\}_{k\in\mathbb{N}}$ of random variables as in (3.20). Let $\{\Psi_t\}_{t\geq 0}$ be the family of Bernstein functions corresponding to the increasing additive process $\Lambda = \{\Lambda_t\}_{t\geq 0}$ via $\mathcal{L}[\Lambda_t] = \exp(-\Psi_t)$ and denote by $\bar{F}_1$ the survival function of X_1. The survival copula $C_{\Lambda,d}$ of $(X_1, \ldots, X_d)$ has the form (3.4) for any $d \geq 2$, with

$$\begin{aligned} g_k(u) :&= \exp\Big(- \Psi_{\bar{F}_1^{-1}(u)}(k) + \Psi_{\bar{F}_1^{-1}(u)}(k-1)\Big), \quad k = 1, \ldots, d, \\ \bar{F}_1(x) &= \exp(-\Psi_x(1)), \quad x \geq 0. \end{aligned}$$

Proof
The survival function of each X_k is given by

$$\begin{aligned} \bar{F}_1(x) :&= \mathbb{P}(X_k > x) = \mathbb{P}(E_k > \Lambda_x) = \mathbb{E}\big[\mathbb{P}(E_k > \Lambda_x | \Lambda_x)\big] \\ &= \mathbb{E}\big[e^{-\Lambda_x}\big] = e^{-\Psi_x(1)}, \quad x \geq 0. \end{aligned}$$

The joint survival function of $(X_1, \ldots, X_d)$ can be derived analogously. For the vector

$\boldsymbol{x} := (x_1, \ldots, x_d) \geq 0$, *with the convention* $x_0 := 0$, *it is given by*

$$\begin{aligned}
\bar{F}(\boldsymbol{x}) :&= \mathbb{P}(X_1 > x_1, \ldots, X_d > x_d) = \mathbb{E}\big[e^{-\sum_{k=1}^{d} \Lambda_{x_k}}\big] \\
&= \mathbb{E}\big[e^{-\sum_{k=1}^{d} (d-k+1)\,(\Lambda_{x_{(k)}} - \Lambda_{x_{(k-1)}})}\big] \\
&= \prod_{k=1}^{d} \mathbb{E}\big[e^{-(d-k+1)\,(\Lambda_{x_{(k)}} - \Lambda_{x_{(k-1)}})}\big] \\
&= \prod_{k=1}^{d} \exp\Big(- \Psi_{x_{(k)}}(d-k+1) + \Psi_{x_{(k-1)}}(d-k+1)\Big) \\
&= \prod_{k=1}^{d} \exp\Big(- \Psi_{x_{(k)}}(d-k+1) + \Psi_{x_{(k)}}(d-k)\Big).
\end{aligned}$$

As outlined in Section 2.4, $x \mapsto \Psi_x(1)$ *is continuous. Thus, the unique survival copula* $C_{\Lambda,d}$ *of* $(X_1, \ldots, X_d)$ *is defined as*

$$\begin{aligned}
C_{\Lambda,d}(u_1, \ldots, u_d) :&= \bar{F}\big(\bar{F}_1^{-1}(u_1), \ldots, \bar{F}_1^{-1}(u_d)\big) \\
&= \prod_{k=1}^{d} \exp\Big(- \Psi_{\bar{F}_1^{-1}(u_{(k)})}(k) + \Psi_{\bar{F}_1^{-1}(u_{(k)})}(k-1)\Big) \\
&= \prod_{k=1}^{d} g_k(u_{(k)}), \qquad (3.21)
\end{aligned}$$

with g_k *as defined in the proposition. By construction,* $g_1 = \mathrm{id}_{[0,1]}$ *and* $g_2(1) = \ldots = g_d(1) = 1$. □

Proposition 3.5.1 shows that any additive–frailty model corresponds to an exchangeable exogenous shock model. By Theorem 3.3.1, $C_{\Lambda,d}$ is the copula (and distribution function) of $(U_1, \ldots, U_d)$, where

$$U_k := \max\{V^E : k \in E\}, \quad k = 1, \ldots, d,$$

and $V^E \sim H_{m,d-m}$ ($H_{m,d-m}$ as defined in Theorem 3.3.1.(iii)) for all subsets E with cardinality $|E| = m$. Analogously to (Embrechts and Hofert, 2013, Proposition 2.3, (7)), one knows that the generalized inverse $\bar{F}_1^{-1}$ of the continuous marginal survival function $\bar{F}_1$ in Proposition 3.5.1 is strictly decreasing. By Corollary 2.2.8, $C_{\Lambda,d}$ is the survival copula of $\big(\bar{F}_1^{-1}(U_1), \ldots, \bar{F}_1^{-1}(U_d)\big)$, where

$$\bar{F}_1^{-1}(U_k) = \bar{F}_1^{-1}\Big(\max\{V^E : k \in E\}\Big) = \min\Big\{\bar{F}_1^{-1}(V^E) : k \in E\Big\}, \quad k = 1, \ldots, d.$$

Last but not least, as the transformed random variable $\bar{F}_1^{-1}(U_k)$ has survival function $\bar{F}_1$ (compare (Embrechts and Hofert, 2013, Proposition 3.1, (2))), we conclude that the additive–frailty model $(X_1, \ldots, X_d)$ is equal in distribution to the exchangeable exogenous shock model setup $\big(\bar{F}_1^{-1}(U_1), \ldots, \bar{F}_1^{-1}(U_d)\big)$.

The connection between the additive–frailty construction in Proposition 3.5.1 and the characterization of general exchangeable exogenous shock models in Theorem 3.3.1 turns out to be beneficial in many regards and helps to derive some of the most interesting insights of the present thesis. For additive–frailty copulas $C_{\Lambda,d}$, we now know two quite different construction approaches: One is based on a minimum/maximum construction involving $2^d - 1$ independent random variables, the other one results from a first-passage time setup parameterized by a single additive subordinator Λ. On a high level, the independence of the additive subordinator's increments is reflected by the independence of the random variables in the shock model. A generic sampling algorithm for $C_{\Lambda,d}$ based on the additive–frailty construction can be stated as follows.

Algorithm 3.5.2 (Simulation of $C_{\Lambda,d}$ in Proposition 3.5.1)

1. *Simulate d independent, unit exponentially distributed random variables $E_1, \ldots, E_d$.*
2. *Simulate one path of $\{\Lambda_t\}_{t\geq 0}$ until $\Lambda_t \geq \max\{E_1, \ldots, E_d\}$.*
3. *Compute $X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d$.*
4. *Set $U_k := \bar{F}_1(X_k), k = 1, \ldots, d$, and return $(U_1, \ldots, U_d)$.*

Clearly, the central task is to simulate the paths of the additive process Λ. Provided this can be accomplished efficiently, the algorithm provides a fast sampling routine even in large dimensions. Extending d solely requires the simulation of further i.i.d. exponentially distributed triggers $E_{d+1}, E_{d+2}, \ldots$, possibly supplemented by simulating additional increments of Λ until the largest trigger is exceeded (it is shown in Mai and Scherer (2009a) that $\lim_{d\to\infty} \mathbb{E}[E_{(d)}]/\log d = 1$). One example of easy-to-simulate additive processes are Lévy subordinators of (compound) Poisson type, see (Sato, 1999, p. 17 ff.). Another instance is given in Section 4.3 within the class of Sato subordinators introduced in Section 2.4.3.

As an aside, note that the proof of Proposition 3.5.1 works in the very same way for

$$X_k := \inf\{t \in \mathbb{R} : \Lambda_t \geq E_k\}, \quad k \in \mathbb{N}, \tag{3.22}$$

with additive subordinators defined on the real-line, i.e. satisfying $\lim_{t\to-\infty} \Lambda_t = 0$ instead of the standard condition $\Lambda_0 = 0$ presupposed in Definition 2.4.1. The difference

simply consists in the marginal distribution function of the X_k potentially having support on $\mathbb{R}$ instead of $[0,\infty)$. As the proposition refers to the (survival) copula of $(X_1,\dots,X_d)$, this subtlety is "eradicated" and has no impact. Consequently, Proposition 3.5.1 also applies to additive processes related to stochastically continuous neutral-to-the-right priors as introduced in Section 2.4.4. Considering a transformed Dirichlet process for Λ results in a very tractable additive–frailty copula.

Theorem 3.5.3 (Dirichlet copula)
Let $\{Z_t\}_{t\geq 0}$ be the Dirichlet process introduced in Section 2.4.4 with parameters (c,G), $c>0$, G continuous, and consider the additive subordinator $\Lambda=\{\Lambda_t\}_{t\in\mathbb{R}}, \Lambda_t := -\log(1-Z_t)$. The corresponding additive–frailty copula $C_{\Lambda,d}$ (termed Dirichlet copula in the sequel) is given by

$$C_{\Lambda,d}(u_1,\dots,u_d)=\prod_{k=1}^{d}\frac{c\,u_{(k)}+k-1}{c+k-1},\quad u\in[0,1].$$

Note that $C_{\Lambda,d}$ in the theorem is not termed Dirichlet–frailty, but solely Dirichlet copula. The reason is that the characterizing process Λ is not a Dirichlet process itself. Instead, it is connected to the latter via a simple log-transform, which is why we have chosen the stated denomination.

Proof (Proof of Theorem 3.5.3)
Recall the stochastic representation of the Dirichlet process in Section 2.4.4. We let $\Lambda=\{\Lambda_t\}_{t\in\mathbb{R}}$ be the related additive subordinator on $\mathbb{R}$ defined by $\Lambda_t := -\log(1-Z_t)$ for $t\in\mathbb{R}$. Furthermore, define $t_0:=\sup\{t\in\mathbb{R}: G(t)=0\}$ and $t_1:=\inf\{t\in\mathbb{R}: G(t)=1\}$, with $t_0:=-\infty$ respectively $t_1:=\infty$ if $G(t)>0$ respectively $G(t)<1$ for all $t\in\mathbb{R}$. We observe with the finite geometric series and the Frullani integral formula, cf. Tricomi (1951), that for $k\in\mathbb{N}$, the Bernstein family $\{\Psi_t\}_{t\in\mathbb{R}}$ in Equation (2.12) corresponding to Λ satisfies

$$\begin{aligned}\Psi_t(k)&=\int_0^\infty\frac{1-e^{-k\,u}}{1-e^{-u}}\,\frac{e^{-u\,c\,(1-G(t))}-e^{-u\,c}}{u}\,\mathrm{d}u\\&=\sum_{i=0}^{k-1}\int_0^\infty\frac{e^{-u\,(i+c\,(1-G(t)))}-e^{-u\,(i+c)}}{u}\,\mathrm{d}u\\&=\sum_{i=0}^{k-1}\log\Big(\frac{c+i}{c\,(1-G(t))+i}\Big)=\log\Big(\prod_{i=0}^{k-1}\frac{c+i}{c\,(1-G(t))+i}\Big),\quad t\in\mathbb{R}\cap[t_0,t_1].\end{aligned}$$

Thus, w.r.t. Proposition 3.5.1, one can derive the three quantities

$$\bar{F}_1(x) = \exp(-\Psi_x(1)) = \exp\Big(-\log\big(\frac{1}{1-G(x)}\big)\Big) = 1 - G(x), \quad x \in \mathbb{R} \cap [t_0, t_1],$$

$$\bar{F}_1^{-1}(u) = G^{-1}(1-u), \quad u \in (0,1),$$

$$-\Psi_{\bar{F}_1^{-1}(u)}(k) + \Psi_{\bar{F}_1^{-1}(u)}(k-1) = -\log\Big(\frac{c+k-1}{c\,(1-G(\bar{F}_1^{-1}(u))) + k - 1}\Big) = \log\Big(\frac{c\,u+k-1}{c+k-1}\Big).$$

Applying Proposition 3.5.1, the claimed form of the copula $C_{\Lambda,d}$ follows. □

As an immediate consequence, we derive the following corollary regarding radial symmetry of the Dirichlet copula.

Corollary 3.5.4 (Radial symmetry of the Dirichlet copula)
The Dirichlet copula is radially symmetric. If furthermore the univariate probability law $\mathrm{d}G$ *in the construction of the Dirichlet process* $Z = \{Z_t\}_{t\in\mathbb{R}}$ *in Theorem 3.5.3 is continuous and symmetric about its median, the random vector* $(X_1, \ldots, X_d)$ *defined in* (3.22) *for the additive subordinator* $\Lambda = \{\Lambda_t\}_{t\in\mathbb{R}}$, $\Lambda_t := -\log(1 - Z_t)$, *is radially symmetric for any* $d \geq 2$ *as well.*

Proof
Radial symmetry of the copula immediately follows from Theorem 3.4.2. For the second claim, note that the function G *equals the (unconditional) distribution function of the components* X_k *(compare the survival function* $\bar{F}_1$ *of* X_k *in the proof of Theorem 3.5.3). The continuity assumption on* G *implies that* Z *almost surely has no jump at any given* $x \in \mathbb{R}$, *since*[7]

$$0 = G(x) - G(x-) = \mathbb{E}[Z_x - Z_{x-}]$$

implies that the non-negative random variable $Z_x - Z_{x-}$ *is almost surely zero. Due to radial symmetry of* $\mathrm{d}G$, *there is a number* $\mu \in \mathbb{R}$, *namely the median of* $\mathrm{d}G$, *such that* $\mu - X_k \stackrel{d}{=} X_k - \mu$ *holds for each* k. *Moreover,* G *is point symmetric about* μ *in the sense that* $1 - G(\mu - x) = G(\mu + x)$ *for all* $x \in \mathbb{R}$. *Since* Z *has no jumps at fixed time points and conditioned on* Z, *the* X_k *are i.i.d. with distribution function* Z, *it holds that*

$$\begin{aligned}\mathbb{P}(\mu - X_1 \leq x_1, \ldots, \mu - X_d \leq x_d) &= \mathbb{P}(X_1 \geq \mu - x_1, \ldots, X_d \geq \mu - x_d) \\ &= \mathbb{E}\Big[(1 - Z_{(\mu-x_1)-}) \ldots (1 - Z_{(\mu-x_d)-})\Big] = \mathbb{E}\Big[(1 - Z_{\mu-x_1}) \ldots (1 - Z_{\mu-x_d})\Big].\end{aligned}$$

[7] Dominated convergence is used in the second equality.

Thus, in order to verify radial symmetry of $\{X_k\}_{k=1,\dots,d}$*, it remains to show that*

$$\begin{aligned}&\mathbb{P}(X_1-\mu\le x_1,\dots,X_d-\mu\le x_d)\\ =&\,\mathbb{E}[Z_{\mu+x_1}\,\dots\,Z_{\mu+x_d}]\overset{!}{=}\mathbb{E}\Big[(1-Z_{\mu-x_1})\,\dots\,(1-Z_{\mu-x_d})\Big].\end{aligned}$$

We are going to show an even stronger statement, namely that

$$\{Z_t\}_{t\in\mathbb{R}}\overset{d}{=}\big\{1-Z_{2\,\mu-t}\big\}_{t\in\mathbb{R}}.$$

To this end, it is sufficient to verify that for arbitrary $d\in\mathbb{N}$ *and* $t_1<t_2<\ldots<t_d$*,*

$$\begin{aligned}&(Z_{t_1},Z_{t_2}-Z_{t_1},\dots,Z_{t_d}-Z_{t_{d-1}},1-Z_{t_d})\\ &\quad\overset{d}{=}\big(1-Z_{2\,\mu-t_1},-Z_{2\,\mu-t_2}+Z_{2\,\mu-t_1},\dots\\ &\qquad\dots,-Z_{2\,\mu-t_d}+Z_{2\,\mu-t_{d-1}},Z_{2\,\mu-t_d}\big).\end{aligned}\tag{3.23}$$

Concerning the left-hand side, by definition, the distribution of $(Z_{t_1},Z_{t_2}-Z_{t_1},\dots,Z_{t_d}-Z_{t_{d-1}},1-Z_{t_d})$ *is a Dirichlet distribution with parameters* $c\,\big(G(t_1),G(t_2)-G(t_1),\dots,G(t_d)-G(t_{d-1}),1-G(t_d)\big)$*. Concerning the right-hand side of Equation (3.23), the respective random vector by definition follows a Dirichlet distribution with parameters*

$$\begin{aligned}&c\,\big(1-G(2\,\mu-t_1),G(2\,\mu-t_1)-G(2\,\mu-t_2),\dots\\ &\quad\dots,G(2\,\mu-t_{d-1})-G(2\,\mu-t_d),G(2\,\mu-t_d)\big)\\ &\quad=c\,\big(G(t_1),G(t_2)-G(t_1),\dots,G(t_d)-G(t_{d-1}),1-G(t_d)\big),\end{aligned}$$

where we have used the symmetry property $1-G(\mu-x)=G(\mu+x)$*.* □

Summing up, we have not only identified the subclass of radially symmetric copulas corresponding to exchangeable exogenous shock models in Theorem 3.4.2, but also found the stochastic model behind to be of additive–frailty type. The copula itself is surprisingly simple in the sense that there are only few examples of high-dimensional multivariate distribution functions available in closed form. Furthermore, it allows for the explicit calculation of concordance measures and an extremely fast simulation algorithm. Both issues are treated in more detail in Chapter 6, where the Dirichlet copula is applied in a risk management context. Another class of additive–frailty copulas is introduced in Chapter 4, where Λ is chosen as the Sato subordinator introduced in Section 2.4.3. Both the Dirichlet and the Sato setup show how Proposition 3.5.1 can lead to new, tractable multivariate distribution functions.

3.6 Exchangeable Marshall–Olkin copulas revisited

A first example of the relationship between general shock model construction, first-passage time framework, and corresponding (survival) copula function has already been adumbrated in the characterization of exchangeable MO copulas by means of d-monotone sequences in Section 2.3.3. The aim of the present paragraph is to use the general results in Section 3.3 to put a new complexion on this special case. For convenience, we restate the well-known findings in Mai (2010) and the references therein for copulas of the form

$$C(u_1,\ldots,u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}}, \quad u_1,\ldots,u_d \in [0,1], \tag{3.24}$$

with non-negative parameters $a_0,\ldots,a_{d-1} \geq 0$.

1. (3.24) defines a copula if and only if $\{a_0,\ldots,a_{d-1}\}$ is d-monotone with $a_0 = 1$.

2. The set of copulas with form (3.24) coincides with the set of survival copulas of exchangeable exogenous shock models with exponentially distributed shocks.

3. The function C in (3.24) defines a copula for all $d \geq 2$ if and only if $\{a_k\}_{k\in\mathbb{N}_0}$ is a completely monotone sequence with $a_0 = 1$. In this case, C is the survival copula of

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1,\ldots,d,$$

where $\Lambda = \{\Lambda_t\}_{t\geq 0}$ is a (unique in law) killed Lévy subordinator characterized by a family $\{\Psi_t\}_{t\geq 0}$, $\Psi_t = t\,\Psi_1$, of Bernstein functions with $\Psi_1(k) - \Psi_1(k-1) = a_{k-1}, k \in \mathbb{N}$. Conversely, any killed Lévy subordinator in the construction above leads to a survival copula of the form (3.24) with a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ satisfying $a_0 = 1$.

Note that the third finding above is a slight refinement of the third item in Section 2.3.3, where the coherence between the Lévy subordinator in (2.10) and the completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ has not been specified. We are going to investigate how the general characterization results derived in the present chapter relate to the three observations above.

1. Concerning the first item, by applying Theorem 3.3.1, a function C of type (3.4) with $g_k(u) = u^{a_{k-1}}, k = 1,\ldots,d$, and a real-valued sequence $\{a_0,\ldots,a_{d-1}\}$, $a_0 = 1$, defines a copula in dimension $d \geq 2$ if and only if $H_{j,k} \in \mathcal{D}$ for all $j \in \mathbb{N}, k \in \mathbb{N}_0 : k + j \leq d$. As

discussed in the proof of Proposition 3.4.1, this is naturally equivalent to $\{a_0, \dots, a_{d-1}\}$ being d-monotone.

2. Addressing the second aspect, consider an exchangeable exogenous shock models with exponentially distributed shocks, i.e. (recall the exchangeability condition in Proposition 3.1.2) a vector $(X_1, \dots, X_d)$ that satisfies Equation (3.1) with $Z^E \sim \mathrm{Exp}(\lambda_j)$ for $|E| = j$. By Proposition 3.1.3, the survival copula of $(X_1, \dots, X_d)$ is given by (3.4) with

$$g_k(u) = \prod_{j=1}^{d+1-k} \Big(\bar{F}^{\{1,\dots,j\}}\big(\bar{F}_1^{-1}(u)\big)\Big)^{\binom{d-k}{j-1}}, \quad u \in [0,1], \quad k = 1, \dots, d,$$

where

$$\begin{aligned}
\bar{F}^{\{1,\dots,j\}}(x) &= \mathbb{P}(Z^E > x, |E| = j) = e^{-\lambda_j x}, \\
\bar{F}_1(x) &= \prod_{j=1}^{d} \big(\bar{F}^{\{1,\dots,j\}}(x)\big)^{\binom{d-1}{j-1}} = e^{-x \sum_{j=1}^{d} \binom{d-1}{j-1} \lambda_j}, \quad x > 0.
\end{aligned}$$

Plugging $\bar{F}^{\{1,\dots,j\}}$ and $\bar{F}_1$ into g_k yields

$$g_k(u) = u^{a_{k-1}}, \quad a_{k-1} := \frac{\sum_{j=1}^{d+1-k} \binom{d-1}{j-1} \lambda_j}{\sum_{j=1}^{d} \binom{d-1}{j-1} \lambda_j},$$

proving one part of the second claim above. Conversely, when starting with a copula C of the form (3.4) with $g_k(u) = u^{a_{k-1}}, k = 1, \dots, d$, Theorem 3.3.1 shows that the random variables V^E in the corresponding stochastic model for C with $|E| = m$ have the distribution function $H_{m,d-m} : [0,1] \to [0,1]$,

$$H_{m,d-m}(u) = \prod_{i=0}^{m-1} g_{d-m+1+i}^{(-1)^i \binom{m-1}{i}}(u) = u^{\lambda_m}, \quad \lambda_m := \sum_{i=0}^{m-1} (-1)^i \binom{m-1}{i} a_{d-m+i}.$$

Due to Corollary 2.2.8, C is the survival copula of the transformed vector

$$\big(-\log(X_1), \dots, -\log(X_d)\big) \stackrel{d}{=} \Big(\min\{-\log(V^E) : 1 \in E\}, \dots, \min\{-\log(V^E) : d \in E\}\Big).$$

It is straightforward to show that $-\log(V^E)$ is exponentially distributed, outlining the remaining part of the second claim above.

3. Last but not least, we have to analyze the alternative stochastic model mentioned in point three above. As a technical prerequisite, Theorem 2.5.8 (originally derived by

Gnedin and Pitman (2008)) and Lemma 4.1.1 in Mai (2010) show that a real-valued sequence $\{a_k\}_{k\in\mathbb{N}_0}$ with $a_0 = 1$ is completely monotone if and only if there exists a (unique in law) killed Lévy subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ characterized by $\{\Psi_t\}_{t\geq 0}, \Psi_t = t\,\Psi_1$, such that $a_{k-1} = \Psi_1(k) - \Psi_1(k-1)$ for all $k \in \mathbb{N}$. Now consider construction (3.20) for a killed Lévy subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with $\{\Psi_t\}_{t\geq 0}, \Psi_t = t\,\Psi_1$. W.l.o.g., i.e. without having an impact on the resulting survival copula, we can assume that $\Psi_1(1) = 1$. Applying Proposition 3.5.1, $C_{\Lambda,d}$ is a copula of type (3.4) for all $d \geq 2$, with

$$\begin{aligned} g_k(u) &:= \exp\Big(- \Psi_{\bar{F}_1^{-1}(u)}(k) + \Psi_{\bar{F}_1^{-1}(u)}(k-1)\Big), \quad k \in \mathbb{N}, \\ \bar{F}_1(x) &= e^{-\Psi_x(1)} \overset{(\Psi_x = x\,\Psi_1)}{=} e^{-x\,\Psi_1(1)} \overset{(\Psi_1(1)=1)}{=} e^{-x}. \end{aligned}$$

Substituting $\bar{F}_1^{-1}(u) = -\log(u)$ in g_k, it follows that

$$\begin{aligned} g_k(u) &= \exp\Big(- \Psi_{-\log(u)}(k) + \Psi_{-\log(u)}(k-1)\Big) \\ &= \exp\Big(\log(u)\,\Psi_1(k) - \log(u)\Psi_1(k-1)\Big) = u^{\Psi_1(k)-\Psi_1(k-1)}. \end{aligned}$$

Summarizing, if C in (3.4) exhibits $g_k(u) = u^{a_{k-1}}, k = 1,\ldots,d$, and defines a copula for any $d \geq 2$, it can be constructed in two quite different ways. On the one hand, C is the survival copula of $(X_1,\ldots,X_d)$ in (3.1) with exponentially distributed shocks Z^E. On the other hand, it is the survival copula of the random vector $(X_1,\ldots,X_d)$ defined in (3.20) when choosing $\{\Lambda_t\}_{t\geq 0}$ to be a killed Lévy subordinator. It is well-known (see also Section 2.4) that within the class of increasing additive processes, Lévy subordinators are distinguished by the feature of stationary increments. Besides, the exponential distribution is the only continuous law on $(0,\infty)$ exhibiting the lack-of-memory property (see e.g. Marsaglia and Tubilla (1975)), i.e. $X \sim \text{Exp}(\lambda)$ satisfies the distinctive condition $\mathbb{P}(X > x + t | X > t) = \mathbb{P}(X > x)$ for all $x, t > 0$. Thus, from an algebraic point of view, the coherence between the minimum construction in (3.1) and the first-passage time setup in (3.20) indicates that the i.i.d. property of the Lévy subordinator's increments translates to the memoryless property of the shock distribution functions.

4 Self-decomposability on the half-line

The analysis of extendible Marshall–Olkin distributions in Section 3.6 has shown that there is an intimate relationship between completely monotone sequences and Bernstein functions Ψ. However, the connection has solely referred to the values of Ψ on the discrete grid $\mathbb{N}_0$, which is why it could not be used to describe Bernstein functions distinctively. In the current chapter, we consider a different class of copulas that is parameterized by one single self-decomposable Bernstein function, yet involves this function's values on the whole domain. As a result, it is possible to derive new characterization results located in the area of self-decomposable probability measures on $[0, \infty)$. The chapter builds on [Mai, Schenk, Scherer (2015c)] and is structured as follows: Section 4.1 recalls self-decomposability and reveals different accesses to this concept. In Section 4.2, two novel characterizations of self-decomposable Bernstein functions are derived. One of these characterizations is given in terms of a multivariate distribution function termed Sato–frailty copula. Section 4.3 analyzes this copula class with respect to its properties (see Section 4.3.1), simulation routines (see Section 4.3.2), and some examples involving the (self-decomposable) Gamma distribution (see Section 4.3.3).

4.1 Definition and alternative representations

Self-decomposable laws constitute a proper subfamily of infinitely divisible laws. For a random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, they are characterized by the property (see, e.g., Jurek and Yor (2004))

$$\begin{aligned} & X \sim \pi, \quad \pi \text{ self-decomposable} \\ \Leftrightarrow \quad & \forall c \in (0,1)\, \exists \text{ random variable } Y \text{ independent of } X \text{ s.t. } X \stackrel{d}{=} c\,X + Y. \end{aligned} \tag{4.1}$$

The set of self-decomposable Bernstein functions arises as the set of Laplace exponents of self-decomposable probability laws on $[0, \infty)$, denoted $\mathrm{SD}[0, \infty)$ in the sequel.

There are different ways to access self-decomposable probability laws on the real line. Apart from the classical definition (4.1), Lévy (1937) and Khintchine (1938) introduce the so-called class L distributions (which subsequently have been found to coincide with the set of self-decomposable distributions) arising as limit distributions of shifted and scaled sums of an infinite sequence of independent random variables. More precisely, a distribution function F on $\mathbb{R}^d$ is said to belong to class L if there exists a sequence $\{Z_k\}_{k\in\mathbb{N}}$ of independent random vectors on $\mathbb{R}^d$ and "suitable" sequences $\{b_k\}_{k\in\mathbb{N}}, \{c_k\}_{k\in\mathbb{N}}$ with $b_k > 0, c_k \in \mathbb{R}^d, k \in \mathbb{N}$, such that the distribution function of $c_n + 1/b_n \sum_{k=1}^{n} Z_k$ converges to F in distribution for $n \to \infty$. A special case is the common central limit theorem which states that for i.i.d. random variables $\{Z_k\}_{k\in\mathbb{N}}$ with finite variance $\mathrm{Var}(Z_1) < \infty$, the scaled sum converges to a normal distribution when choosing $b_n = \sqrt{n \,\mathrm{Var}(Z_1)}$ and $c_n = -n\,\mathbb{E}[Z_1]$. More details and the equivalence to self-decomposable probability measures are provided in (Sato, 1999, Theorem 15.3, p. 91).

A different approach relies on additive processes as introduced in Section 2.4. Wolfe (1982) analyzes a generalized Ornstein–Uhlenbeck process $\{O_t\}_{t\geq 0}$ given by the stochastic differential equation $dO_t = -\gamma\, O_t\, \mathrm{d}t + dL_t$, $\gamma > 0$, where $\{L_t\}_{t\geq 0}$ is a real-valued Lévy process. As a main result, he shows that $\{O_t\}_{t\geq 0}$ converges in law to a random variable for $t \to \infty$ if and only if $\mathbb{E}[\max\{\log(|L_1|), 0\}] < \infty$, and that the resulting set of random variables coincides with the set of self-decomposable random variables on the real line. Jurek and Vervaat (1983) generalize these findings to real separable Banach spaces. Related results can also be found in Sato and Yamazato (1984).

The crucial connection for the derivations in the present manuscript is the relation between self-decomposable probability laws and self-similar additive processes (see Section 2.4.3). A key result is (Sato, 1999, Theorem 16.1, p. 99), stating that the marginal distribution of an H-Sato process is self-decomposable at any time and that, conversely, for any self-decomposable probability law π on $\mathbb{R}$ and any $H > 0$, there is a (unique in law) real-valued H-Sato process $\{X_t\}_{t\geq 0}$ such that $X_1 \sim \pi$. Thus, if we focus on self-decomposable probability laws $\pi \in \mathrm{SD}[0, \infty)$ on the half-line, then associated with each pair (π, H), $H > 0$, there exists a unique H-Sato subordinator $\{\Lambda_t\}_{t\geq 0}$ with $\Lambda_1 \sim \pi$. This characterization and the previous approach via stochastic integrals are elegantly unified in Jeanblanc et al. (2002). As an application in practice, Carr et al. (2007) embed Sato processes in stock price models used for option pricing and point out in an empirical study that the fitting performance of the corresponding models is quite encouraging.

Summing up, self-decomposable probability laws in many respects form an interesting subclass of infinitely divisible laws. The present chapter provides new characteriza-

tions of SD$[0,\infty)$, analytically in terms of monotonicity conditions for self-decomposable Bernstein functions Ψ and probabilistically in terms of an exchangeable exogenous shock model. The copula of the corresponding random vector turns out to be quite tractable and is analyzed in more detail. Among others, it may be used to model dependent default times in a multivariate setting. This potential application is picked up in Chapter 5.

4.2 Two novel characterizations

The starting point for the self-decomposability results in the present chapter is Proposition 3.5.1, which shows that additive–frailty copulas constitute an (extendible) subclass of exchangeable exogenous shock models. With the proceeding in the present paragraph potentially being applicable to derive characterizations of other mathematical objects as well, we outline the fundamental line of reasoning underlying the upcoming computations.

(a) Considering an arbitrary increasing additive process $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with $\lim_{t\to\infty} \Lambda_t = \infty$ (for instance one of the families depicted in Section 2.4), we know that $C_{\Lambda,d}$ as defined in Proposition 3.5.1 is a copula corresponding to an exchangeable exogenous shock model for any $d \geq 2$.

(b) Consequently, Λ satisfies the equivalence conditions stated in Theorem 3.3.1.

(c) Reversely, do the conditions in the theorem even characterize the stochastic process Λ in concern?

When regarding H-Sato subordinators $\Lambda = \{\Lambda_t\}_{t\geq 0}$, we already know due to Section 2.4.3 that Λ is (uniquely in law) determined by the corresponding familiy $\{\Psi_t\}_{t\geq 0}$ of Bernstein functions satisfying $\Psi_t(x) := \Psi_1(t^H x)$ for all $x, t \geq 0$ and a self-decomposable Bernstein function Ψ_1. Executing steps $(a) - (c)$ above and translating the conditions for Λ into requirements for Ψ_1 yields the following main finding of the present chapter.

Theorem 4.2.1 (Characterization of self-decomposability on the half-line) *Let $\psi : [0,\infty) \to (0,1]$ be continuous and strictly decreasing with $\psi(0) = 1$ and with $\lim_{x\to\infty} \psi(x) = 0$. The following statements are equivalent:*

(i) $\psi = \exp(-\Psi)$ for a self-decomposable Bernstein function Ψ.

(ii) For all $k \in \mathbb{N}_0$ and $j \in \mathbb{N}$, with $\Psi := -\log(\psi)$, it holds that

$$x \mapsto A_{j,k}(x) := \sum_{i=0}^{j} \binom{j}{i} (-1)^i \, \Psi\big((k+i)\, x\big) \quad \text{is decreasing in } [0,\infty). \tag{4.2}$$

(iii) For every $d \geq 2$, the function $C^{\psi}_{\Lambda,d} : [0,1]^d \to [0,1]$,

$$C^{\psi}_{\Lambda,d}(u_1,\ldots,u_d) := \prod_{k=1}^{d} g_k(u_{(k)}),$$

is a copula, where $g_k : [0,1] \to [0,1], k = 1,\ldots,d$, is defined by

$$g_k(u) := \begin{cases} \dfrac{\psi\big(k\,\psi^{-1}(u)\big)}{\psi\big((k-1)\,\psi^{-1}(u)\big)}, & u > 0, \\ 0, & u = 0. \end{cases}$$

Theorem 4.2.1 provides two interesting insights. On the one side, it extends the correspondence results between multivariate distributions and transforms of univariate probability laws derived for (a) Archimedean copulas and Laplace transforms, see Kimberling (1974); Marshall and Olkin (1988); McNeil and Nešlehovà (2009) and the other references mentioned in Section 2.3.2 respectively Theorem 2.3.13, and (b) Marshall–Olkin distributions and Bernstein functions, see the example in Section 3.6 and Mai and Scherer (2009b); Ressel (2011, 2013). Comparing the findings to Theorem 2.3.13, instead of characterizing extendible Archimedean copulas in terms of completely monotone functions, the new result establishes a connection between extendible copulas of the form

$$C(u_1,\ldots,u_d) = \prod_{k=1}^{d} \frac{\psi\big(k\,\psi^{-1}(u_{(k)})\big)}{\psi\big((k-1)\,\psi^{-1}(u_{(k)})\big)}, \quad u_1,\ldots,u_d \in [0,1],$$

and the shape of the copula-generating function ψ. As an additional constraint compared to generators of extendible Archimedean copulas, ψ not only has to coincide with the Laplace transform of an infinitely divisible, but a self-decomposable random variable on $[0,\infty)$.

On the other side, Theorem 4.2.1.(ii) provides a new perspective towards the dissociation of self-decomposable Bernstein functions from general ones. So far, self-decomposable Bernstein functions Ψ have been described by the special structure of the Lévy measure in (2.11), see Section 2.4.3. This distinction is "lifted up" to the behavior of the function

Ψ itself. It is known (see (Gnedin and Pitman, 2008, Corollary 4.2)) that the sequence $\{\Psi(k)\}_{k\in\mathbb{N}}$ of an arbitrary Bernstein function Ψ is *completely alternating*, meaning that

$$\sum_{i=0}^{j}(-1)^i \binom{j}{i}\Psi(k+i) \leq 0 \quad \text{for all } k \in \mathbb{N}_0, j \in \mathbb{N}.$$

In terms of $A_{j,k}$, this is equivalent to $A_{j,k}(1) \leq 0$ for all $k \in \mathbb{N}_0, j \in \mathbb{N}$. Given that $\Psi(0) = 0$ and therefore $A_{j,k}(0) = 0$, the characterizing property (4.2) establishes that Ψ is self-decomposable if and only if, additionally, $A_{j,k}$ is interpolated in such a way that it becomes decreasing in $[0, \infty)$. Another perspective on (4.2) can be given in terms of complete monotonicity. $A_{j,k}$ being decreasing is equivalent to

$$\sum_{i=0}^{j}\binom{j}{i}(-1)^i\Big(\Psi\big((k+i)\,x\big) - \Psi\big((k+i)\,y\big)\Big) \geq 0, \quad 0 \leq x \leq y.$$

This means that the sequence $\{a_k^{x,y}\}_{k\in\mathbb{N}_0}, a_k^{x,y} := \Psi\big(k\,x\big) - \Psi\big(k\,y\big)$, is completely monotone for any pair $(x, y) \in [0, \infty)^2$ with $x \leq y$.

We now proceed to the proof of Theorem 4.2.1.

Remark 4.2.2 (Structure of the proof of Theorem 4.2.1)

(i) ⇒ (iii) *Any self-decomposable Bernstein function Ψ can be connected to an H-Sato subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$. Considering the corresponding additive–frailty construction in (3.20), Propositon 3.5.1 shows that $C^{\psi}_{\Lambda,d} = C_{\Lambda,d}$.*

(i) ⇒ (ii) *Due to (i) ⇒ (iii), we know that $C^{\psi}_{\Lambda,d}$ is a copula. Therefore, one could directly apply the equivalence (i)⇔(iii) in Theorem 3.3.1 and deduce that in order for the respective functions $H_{j,k}$ to be elements of $\mathcal{D}$, $A_{j,k}$ must be decreasing. However, we will proceed differently. The major distinction between copulas $C^{\psi}_{\Lambda,d}$ and the superclass in (3.4) is that for Ψ being a self-decomposable Bernstein function, the g_k in the definition of $C^{\psi}_{\Lambda,d}$ are differentiable. Due to this structural simplification, we partially replicate the proof of (ii)⇔(iii) in Theorem 3.3.1, yet shorten it massively. As a byproduct, for the case of differentiable functions g_k, this "elegant" version of the proof also verifies the rather complicated general proceeding.*

(ii) ⇒ (i) *Due to decreasingness of $A_{j,k}$, the difference $A_{j,k}(x) - A_{j,k}(y)$ is non-negative for any $0 \leq x < y < \infty$. By relating these differences to completely monotone sequences and by relying on an auxiliary lemma and classical results on the conjunction of completely monotone sequences and absolutely monotonic functions (which are defined in the sequel), the claim follows.*

(iii) ⇒ (i) *The copula-induced measure* $\mathrm{d}C^{\psi}_{\Lambda,d}$ *assigns non-negative mass to any d-box on* $[0,1]^d$. *Similarly to the proof of (ii) ⇒ (i), the resulting inequalities can be affiliated to completely monotone structures and yield the required self-decomposability of* Ψ.

Proof (Proof of Theorem 4.2.1)

(i) ⇒ (iii): Assume that $\psi = \exp(-\Psi_1)$ *with a self-decomposable Bernstein function* Ψ_1. *As pointed out in Section 2.4.3, for every* $H > 0$, *there exists an H-Sato subordinator* $\{\Lambda_t\}_{t\geq 0}$ *with* $\mathcal{L}[\Lambda_t] = \exp(-\Psi_t)$ *for a family* $\{\Psi_t\}_{t\geq 0}$ *of Bernstein functions satisfying* $\Psi_t(x) = \Psi_1(t^H x)$ *for all* $x, t \geq 0$. *Define a sequence* $\{X_k\}_{k\in\mathbb{N}}$ *of random variables by the additive–frailty model*

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k \in \mathbb{N}.$$

By Proposition 3.5.1, the copula $C_{\Lambda,d}$ *has the form* (3.4) *for all* $d \geq 2$, *with* $g_k, k = 1, \ldots, d$, *given by*

$$g_k(u) := \exp\Big(- \Psi_{\bar{F}_1^{-1}(u)}(k) + \Psi_{\bar{F}_1^{-1}(u)}(k-1)\Big), \quad k \in \mathbb{N},$$
$$\bar{F}_1(x) = e^{-\Psi_x(1)} \overset{(\Psi_x(1)=\Psi_1(x^H))}{=} e^{-\Psi_1(x^H)}.$$

Plugging $\bar{F}^{-1}(u) = \big(\Psi_1^{-1}(-\log u)\big)^{1/H}$ *into* g_k *yields*

$$\begin{aligned} g_k(u) &= \exp\Big(- \Psi_{\left(\Psi^{-1}(-\log u)\right)^{1/H}}(k) + \Psi_{\left(\Psi^{-1}(-\log u)\right)^{1/H}}(k-1)\Big) \\ &= \exp\Big(- \Psi_1\big(k\,\Psi_1^{-1}(-\log u)\big) + \Psi_1\big((k-1)\,\Psi_1^{-1}(-\log u)\big)\Big) \\ &= \frac{\psi\big(k\,\psi^{-1}(u)\big)}{\psi\big((k-1)\,\psi^{-1}(u)\big)}. \end{aligned}$$

By definition of g_1, *it holds that* $g_1(0) = 0$. *As the value of* $g_k(0)$ *has no influence on* $C_{\Lambda,d}$ *for* $k \geq 2$, *part (iii) of Theorem 4.2.1 is established.*

(i) ⇒ (ii): Let Ψ *be a self-decomposable Bernstein function and define* $\psi = \exp(-\Psi)$. *By the proof of (i) ⇒ (iii) above,* $C^{\psi}_{\Lambda,d}$ *is a copula for all* $d \geq 2$. *Applying Theorem 3.3.1.(ii), for all* $u, v \in [0,1], u < v$, *it holds that*

$$\sum_{i=0}^{j}(-1)^i \binom{j}{i} \prod_{l=1}^{i} g_{l+k}(u) \prod_{l=i+1}^{j} g_{l+k}(v) \geq 0, \quad k \in \mathbb{N}_0, j \in \mathbb{N}.$$

We can rewrite this inequality as

$$\sum_{i=0}^{j-1}(-1)^i \binom{j-1}{i} \big(g_{k+1+i}(v) - g_{k+1+i}(u)\big) \prod_{l=1}^{i} g_{k+l}(u) \prod_{l=i+2}^{j} g_{k+l}(v) \geq 0.$$

Dividing by $(v-u)\prod_{l=1}^{j} g_{k+l}(u)$, *letting* v *approach* u, *and using the differentiablity*[1] *of* $g_m, m \in \{k+1,\dots,k+j\}$, *it follows that*

$$\begin{aligned}
&\sum_{i=0}^{j-1}(-1)^i \binom{j-1}{i} \frac{\big(g_{k+i+1}(u)\big)'}{g_{k+i+1}(u)} = \sum_{i=0}^{j-1}(-1)^i \binom{j-1}{i} \Big(\log\big(g_{k+i+1}(u)\big)\Big)' \\
=&\Big(\log\Big(\prod_{i=0}^{j-1}\Big(\frac{\psi\big((k+1+i)\,\psi^{-1}(u)\big)}{\psi\big((k+i)\,\psi^{-1}(u)\big)}\Big)^{(-1)\binom{j-1}{i}}\Big)\Big)' \\
=&\Big(\sum_{i=0}^{j-1}(-1)^i \binom{j-1}{i} \Big(-\Psi\big((k+i+1)\,\psi^{-1}(u)\big)+\Psi\big((k+i)\,\psi^{-1}(u)\big)\Big)\Big)' \\
=&\Big(\sum_{i=0}^{j}(-1)^i \binom{j}{i} \Psi\big((k+i)\,\psi^{-1}(u)\big)\Big)' \geq 0.
\end{aligned}$$

As ψ *and, thus,* ψ^{-1} *are continuous and strictly decreasing by assumption, we deduce that*

$$x \mapsto A_{j,k}(x) := \sum_{i=0}^{j}(-1)^i \binom{j}{i} \Psi\big((k+i)\,x\big) \quad \text{is decreasing in } [0,\infty)$$

for all $k \in \mathbb{N}_0, j \in \mathbb{N}$.

$(ii) \Rightarrow (i)$: *Let* $A_{j,k}$ *be decreasing for all* $k \in \mathbb{N}_0, j \in \mathbb{N}$. *Consequently, for arbitrary* $x, y \in [0,\infty), x < y$, *one has*

$$\sum_{i=0}^{j} \binom{j}{i} (-1)^i \Big(\Psi\big((k+i)\,x\big) - \Psi\big((k+i)\,y\big)\Big) \geq 0,$$

such that the sequence $\{\Psi(k\,x) - \Psi(k\,y)\}_{k\in\mathbb{N}_0}$ *is completely monotone. By (Lorch and Newman, 1983, Lemma 1), if* f *is an absolutely monotonic*[2] *function on* $[0,\infty)$, *and* $\{a_k\}_{k\in\mathbb{N}_0}$ *a completely monotone sequence, the transformation* $\{f(a_k)\}_{a\in\mathbb{N}_0}$ *again yields a completely monotone sequence. Thus, by applying the absolutely monotonic function* $x \mapsto \exp(x)$ *to* $\{\Psi(k\,x) - \Psi(k\,y)\}_{k\in\mathbb{N}_0}$, *it follows that the sequence* $\{\psi(k\,y)/\psi(k\,x)\}_{k\in\mathbb{N}_0}, \psi = \exp(-\Psi)$, *is completely monotone for any* $x, y \in [0,\infty), x < y$. *This implies that* $\{\psi(q\,k)/\psi(c\,q\,k)\}_{k\in\mathbb{N}_0}$ *is completely monotone for all* $c \in (0,1)$ *and all* $q \in \mathbb{Q} \cap [0,\infty)$. *Finally, applying Lemma 2.3.9,* $x \mapsto \psi(x)/\psi(c\,x)$ *is a completely monotone function for all* $c \in (0,1)$ *and, according to (Schilling et al., 2010, Definition 5.12), it holds that*

[1] Differentiablity is given by assumption as g_m consists of (self-decomposable) Bernstein functions.

[2] A function $f : \mathbb{R} \to \mathbb{R}$ is called absolutely monotonic in the interval (a,b), $a, b \in \mathbb{R}$, $a < b$, if it has non-negative derivatives of all orders in (a,b). If, in addition, f is continuous in a respectively b, it is called absolutely monotonic in $[a,b)$ respectively $(a,b]$.

$\psi = \exp(-\Psi)$ *for a self-decomposable Bernstein function* Ψ.

(iii) $\Rightarrow$ *(i): Let* $\{U_k\}_{k\in\mathbb{N}}$ *be a sequence of random variables with* $(U_1,\dots,U_d) \sim C^{\psi}_{\Lambda,d}$ *for all* $d \geq 2$. *Applying (i)* $\Leftrightarrow$ *(ii) of Theorem 3.3.1, for all* $u, v \in [0,1], u < v, k \in \mathbb{N}_0, j \in \mathbb{N}$, *it holds that*

$$G_{j,k}(u,v) = \sum_{i=0}^{j} \binom{j}{i} (-1)^i \prod_{l=1}^{i} g_{l+k}(u) \prod_{l=i+1}^{j} g_{l+k}(v) \geq 0.$$

Due to the telescope structure of $\prod_{l=1}^{i} g_{l+k}(u)$ *respectively* $\prod_{l=i+1}^{j} g_{l+k}(v)$ *for the given form of* g_k, *the inequality simplifies to*

$$G_{j,k}(u,v) = \frac{\psi\big((k+j)\,\psi^{-1}(v)\big)}{\psi\big(k\,\psi^{-1}(u)\big)} \sum_{i=0}^{j} (-1)^i \binom{j}{i} \frac{\psi\big((k+i)\,\psi^{-1}(u)\big)}{\psi\big((k+i)\,\psi^{-1}(v)\big)} \geq 0.$$

As $0 < u < v \leq 1$ *implies* $\infty > \psi^{-1}(u) > \psi^{-1}(v) \geq 0$, *we get for arbitrary* $\infty > b > a \geq 0$ *that*

$$\sum_{i=0}^{j} (-1)^i \binom{j}{i} \frac{\psi\big((k+i)\,b\big)}{\psi\big((k+i)\,a\big)} \geq 0.$$

Let $c \in (0,1)$ *and* $q \geq 0$ *be arbitrary and set* $b := q$, $a := c\,q \leq b$. *Then we obtain complete monotonicity of the sequence* $\{\psi(k\,q)/\psi(k\,c\,q)\}_{k\in\mathbb{N}_0}$. *With Lemma 2.3.9, it follows that* $x \mapsto \psi(x)/\psi(c\,x)$ *is completely monotone for* $c \in (0,1)$. *By Definition 5.12 in Schilling et al. (2010), the claim follows.* □

4.3 Sato–frailty copulas

In Section 3.5, the copula $C_{\Lambda,d}$ arising from the first-passage time construction in Proposition 3.5.1 is termed additive–frailty copula, referring to the additive process Λ parameterizing the multivariate distribution function. The proof of Theorem 4.2.1 has shown that the copula $C^{\psi}_{\Lambda,d}$ in (iii) arises in a similar way by considering an H-Sato subordinator for Λ. As a consequence, $C^{\psi}_{\Lambda,d}$ is termed *Sato–frailty copula* in the sequel. The self-similarity exponent H can be dropped in the denomination as we have seen that for a fixed self-decomposable Bernstein function Ψ and the corresponding law π with $\mathcal{L}\pi = \psi := \exp(-\Psi)$, any H-Sato subordinator $\{\Lambda_t\}_{t\geq 0}$ with $\Lambda_1 \sim \pi$ leads to the same copula. The independence of H is simply a result of the copula-inherent normalization

to uniform marginal distributions. Another normalization effect results from the expressions $\psi\big(k\,\psi^{-1}(u)\big)$ appearing in the definition of a Sato–frailty copula: Given that $\psi = \exp(-\Psi)$ for some self-decomposable Bernstein function Ψ, it holds that

$$\psi\big(k\,\psi^{-1}(u)\big) = \exp\Big(-\Psi\big(k\,\Psi^{-1}(-\log u)\big)\Big), \quad u \in (0,1].$$

Defining the stretched/compressed (self-decomposable) Bernstein function$\tilde{\Psi}$ by $\tilde{\Psi}(x) := \Psi(c\,x), x \geq 0$, for a constant $c > 0$, it follows that $\tilde{\Psi}^{-1}(x) = (1/c)\,\Psi^{-1}(x), x \geq 0$. Therefore, when considering the Sato–frailty copula $C^{\tilde{\psi}}_{\Lambda,d}$ with $\tilde{\psi} := \exp(-\tilde{\Psi})$, it holds that

$$\begin{aligned}\tilde{\psi}\big(k\,\tilde{\psi}^{-1}(u)\big) &= \exp\Big(-\tilde{\Psi}\big(k\,\tilde{\Psi}^{-1}(-\log u)\big)\Big) = \exp\Big(-\Psi\big(c\,k\,(1/c)\,\Psi^{-1}(-\log u)\big)\Big)\\ &= \exp\Big(-\Psi\big(k\,\Psi^{-1}(-\log u)\big)\Big) = \psi\big(k\,\psi^{-1}(u)\big), \quad u \in (0,1]. \qquad (4.3)\end{aligned}$$

As a consequence, the copulas $C^{\psi}_{\Lambda,d}$ and $C^{\tilde{\psi}}_{\Lambda,d}$ coincide and we can always stretch respectively compress the Bernstein function in the construction, thus eliminating one degree of freedom in the parameterization of Ψ.

4.3.1 Statistical properties and connection to Archimedean copulas

This section investigates Sato–frailty copulas $C^{\psi}_{\Lambda,d}$ w.r.t. the classes and structural properties introduced in Section 2.2.2 and the concordance measures in Section 2.2.3. Interestingly, as it turns out, Sato–frailty and extendible Archimedean copulas share a variety of similarities. This coherence is elucidated in the present paragraph and illustrated graphically in Section 4.3.3.

By construction of the additive–frailty model behind, we already know that Sato–frailty copulas are extendible. Applying the general result in Proposition 3.4.1 for extreme-value copulas corresponding to exchangeable exogenous shock models, $C^{\psi}_{\Lambda,d}$ is an extreme-value copula for any $d \geq 2$ if and only if

$$g_k(u) := \frac{\psi\big(k\,\psi^{-1}(u)\big)}{\psi\big((k-1)\,\psi^{-1}(u)\big)} = u^{a_{k-1}}, \quad u \in [0,1], k \in \mathbb{N},$$

for a completely monotone sequence $\{a_k\}_{k\in\mathbb{N}_0}$ with $a_0 = 1$. The corresponding family of (extendible) copulas has already been treated in Sections 2.3.3 and 3.6. It coincides with the set of extendible MO copulas, which results from plugging a Lévy subordinator Λ into the additive–frailty construction in Proposition 3.5.1. In (Sato, 1999, Remark 16.2, p. 100), it is outlined that an H-Sato subordinator with corresponding Bernstein family

$\{\Psi_t\}_{t\geq 0}$ is a Lévy subordinator if and only if $H \geq 1$ and $\Psi_1(x) = \beta\, x^{1/H}$, $\beta > 0, x \geq 0$. Recall that this Bernstein function has already been introduced in Section 2.4.5 and corresponds to the positive stable distribution. Therefore, the intersection of Lévy and Sato subordinators is called α-stable (Lévy) subordinator. The following proposition states that $\Psi(x) = \beta\, x^{\alpha}, \alpha \in (0,1]$, is not only sufficient, but also necessary for $C^{\psi}_{\Lambda,d}$ to be of extreme-value type for any $d \geq 2$.

Proposition 4.3.1 (Extreme-value Sato–frailty copulas)
A Sato–frailty copula $C^{\psi}_{\Lambda,d}$ satisfies the extreme-value property for any $d \geq 2$ if and only if $\psi = \exp(-\Psi)$, $\Psi(x) = \beta\, x^{\alpha}, \alpha \in (0,1], \beta > 0, x \geq 0$.

Proof
Before proceeding with the proof, we introduce the following auxiliary lemma in Mai and Scherer (2014) addressing processes that are weakly infinitely divisible with respect to time[3] (weak IDT).

Lemma 4.3.2 (Weak IDT processes and Bernstein functions)
Let $\{H_t\}_{t\geq 0}$ be a stochastic process which is right-continuous, non-negative, and non-decreasing with $H_0 = 0$ and $\lim_{t\to\infty} H_t = \infty$. $\{H_t\}_{t\geq 0}$ is weak IDT if and only if there exists a non-zero Bernstein function $\Psi : [0,\infty) \to [0,\infty)$ such that

$$\mathbb{E}\Big[e^{-x\, H_t}\Big] = e^{-t\, \Psi(x)}, \quad x, t \geq 0.$$

Proof
See (Mai and Scherer, 2014, Theorem 1.1). □

Moreover, Remark (b) on page 6 and Theorem 5.3 in Mai and Scherer (2014) imply that the survival copula $C_{\Lambda,d}$ of $(X_1, \ldots, X_d)$ as constructed in Proposition (3.5.1) is an extreme-value copula for any $d \geq 2$ if and only if there exists a strictly increasing function $h : [0,\infty) \to [0,\infty)$ with $h(0) = 0$ such that $\tilde{\Lambda} = \{\tilde{\Lambda}_t\}_{t\geq 0}$, $\tilde{\Lambda}_t := \Lambda_{h(t)}$, is strong IDT[4]. As additive processes are described uniquely in law by their marginal distributions (see Section 2.4), weak and strong IDT property coincide in the present case. Now consider

[3] $\{\Lambda_t\}_{t\geq 0}$ is called weak IDT if

$$\Lambda_t \overset{d}{=} \Lambda^{(1)}_{t/n} + \ldots + \Lambda^{(n)}_{t/n} \quad \text{for all } t \geq 0 \text{ and } n \in \mathbb{N},$$

where $\{\Lambda^{(k)}_t\}_{t\geq 0}, k = 1, \ldots, n$, are independent copies of $\{\Lambda_t\}_{t\geq 0}$.

[4] A weak IDT process $\{\Lambda_t\}_{t\geq 0}$ is called strong IDT if it additionally satisfies

$$\{\Lambda_t\}_{t\geq 0} \overset{d}{=} \{\Lambda^{(1)}_{t/n} + \ldots + \Lambda^{(n)}_{t/n}\}_{t\geq 0} \quad \text{for all } t \geq 0,\ n \in \mathbb{N}.$$

an H-Sato subordinator $\{\Lambda_t\}_{t\geq 0}$ characterized by $\mathcal{L}[\Lambda_1] = \exp(-\Psi)$. We are going to show that $\{\tilde{\Lambda}_t\}_{t\geq 0}$, $\tilde{\Lambda}_t = \Lambda_{h(t)}$ for a strictly increasing function h with $h(0) = 0$, is a weak IDT process if and only if there exists an $\alpha \in (0,1]$ such that $h(t) = h(1)\, t^{1/(\alpha H)}$ and $\mathbb{E}[\exp(-x\,\Lambda_1)] = \exp(-\beta\, x^\alpha), \beta > 0, x \geq 0$.

The "if"-part follows from Lemma 4.3.2 since under the stated conditions,

$$\mathbb{E}\Big[e^{-x\,\tilde{\Lambda}_t}\Big] = \mathbb{E}\Big[e^{-x\,\Lambda_{h(t)}}\Big] = \mathbb{E}\Big[e^{-x\,h(t)^H\,\Lambda_1}\Big] = e^{-t\,\beta\,h(1)^{\alpha H}\,x^\alpha}, \quad x, t \geq 0,$$

and $\Psi : [0,\infty) \to [0,\infty), \Psi(x) = \tilde{\beta}\, x^\alpha$, $\tilde{\beta} := \beta\, h(1)^{\alpha H}$, defines a proper Bernstein function for $\alpha \in (0,1]$ and $\tilde{\beta} > 0$. For the "only if"-part, assume that $\{\tilde{\Lambda}_t\}_{t\geq 0}$ is weak IDT. Denote by Ψ the Bernstein function associated with Λ_1, i.e. $\mathbb{E}[\exp(-x\,\Lambda_1)] = \exp(-\Psi(x))$ for $x \geq 0$. Due to Lemma 4.3.2, there exists a Bernstein function $\tilde{\Psi}$ such that

$$e^{-t\,\tilde{\Psi}(x)} = \mathbb{E}\Big[e^{-x\,\tilde{\Lambda}_t}\Big] = \mathbb{E}\Big[e^{-x\,\Lambda_{h(t)}}\Big] = \mathbb{E}\Big[e^{-x\,h(t)^H\,\Lambda_1}\Big] = e^{-\Psi(x\,h(t)^H)}. \tag{4.4}$$

By strict monotonicity of h, the inverse function h^{-1} exists and the function $\tilde{h} : [0,\infty) \to [0,\infty), \tilde{h}(t) := h(t)^H$, possesses an inverse $\tilde{h}^{-1} : [0,\infty) \to [0,\infty), \tilde{h}^{-1}(t) = h^{-1}(t^{1/H})$, for all $t \geq 0$. Thus, for any $u > 0$, it holds that

$$\begin{aligned}\Psi(x) &= \Psi\Big(\frac{x}{u}\,\tilde{h}\big(\tilde{h}^{-1}(u)\big)\Big) = \Psi\Big(\frac{x}{u}\,h\big(h^{-1}(u^{1/H})\big)^H\Big) \overset{(4.4)}{=} \tilde{h}^{-1}(u)\,\tilde{\Psi}\Big(\frac{x}{u}\Big) \\ &\overset{(4.4)}{=} \tilde{h}^{-1}(u)\,\Psi\Big(\frac{x}{u}\,h(1)^H\Big) = \tilde{h}^{-1}(u)\,\Psi\Big(\frac{x}{u}\,\tilde{h}(1)\Big).\end{aligned} \tag{4.5}$$

As shown in Section 2.4.2, there exists a (unique in law) Lévy subordinator $\{L_t\}_{t\geq 0}$ associated with Ψ by $\mathbb{E}[\exp(-x\,L_1)] = \exp(-\Psi(x))$ for $x \geq 0$. In addition, for fixed $u > 0$, the Bernstein function $\Psi^{(u)} : [0,\infty) \to [0,\infty), \Psi^{(u)}(x) = \tilde{h}^{-1}(u)\,\Psi(x\,\tilde{h}(1)/u)$, corresponds to the Lévy subordinator $\{L^{(u)}_t\}_{t\geq 0}, L^{(u)}_t := \tilde{h}(1)/u\, L_{\tilde{h}^{-1}(u)\,t}$. Equation (4.5) and Sato (1999), Theorem 7.10. (iii), imply that

$$\begin{aligned}\{L_t\}_{t\geq 0} &\overset{d}{=} \Big\{\frac{\tilde{h}(1)}{u}\,L_{\tilde{h}^{-1}(u)\,t}\Big\}_{t\geq 0} \\ \overset{(a:=\tilde{h}^{-1}(u))}{\Leftrightarrow}\ \{L_{a\,t}\}_{t\geq 0} &\overset{d}{=} \Big\{\frac{\tilde{h}(a)}{\tilde{h}(1)}\,L_t\Big\}_{t\geq 0} \Leftrightarrow \{L_{a\,t}\}_{t\geq 0} \overset{d}{=} \Big\{\frac{h(a)^H}{h(1)^H}\,L_t\Big\}_{t\geq 0}.\end{aligned} \tag{4.6}$$

It is well-known (see for instance (Sato, 1999, Remark 16.2)) that the only non-decreasing Lévy process $\{L_t\}_{t\geq 0}$ with the selfsimilarity property (4.6) is an α-stable Lévy subordinator with index $\alpha \in (0,1]$ such that $\tilde{h}(a)/\tilde{h}(1) = a^{1/\alpha}$. Therefore, h is given by $h(t) = h(1)\, t^{1/(\alpha H)}$. Last but not least, substituting $\tilde{h}^{-1}(u) = (u/\tilde{h}(1))^\alpha, u \geq 0$, in Equation (4.5) yields

$$\Psi(x) = \Big(\frac{u}{\tilde{h}(1)}\Big)^\alpha\,\Psi\Big(\frac{x}{u}\,\tilde{h}(1)\Big) \overset{(\tilde{u}:=\tilde{h}(1)/u)}{=} \tilde{u}^{-\alpha}\,\Psi(x\,\tilde{u}).$$

Setting $x = 1$, it follows that

$$\Psi(\tilde{u}) = \Psi(1)\,\tilde{u}^{\alpha} \overset{\beta:=\Psi(1)}{=} \beta\,\tilde{u}^{\alpha}, \quad \tilde{u} \geq 0,$$

and the claim is established. □

Recall that our characterization result for Sato–frailty copulas (see Theorem 4.2.1) represents an analogue to Kimberling's Theorem (see Theorem 2.3.13) for extendible Archimedean copulas. While an Archimedean generator ψ defines an Archimedean copula in any dimension $d \geq 2$ if and only if ψ is the Laplace transform of a positive random variable, the function ψ defines an extendible family $C^{\psi}_{\Lambda,d}$ of Sato–frailty copulas via $\mathcal{L}[\Lambda_1] = \psi$ if and only if ψ is the Laplace transform of a self-decomposable positive random variable. Interestingly, the result on extreme-value Sato–frailty copulas yields another connection to Archimedean copulas. As we have just shown, Sato–frailty copulas $C^{\psi}_{\Lambda,d}$ are of extreme-value type in any dimension if and only if $\psi = \exp(-\beta\,x^{\alpha}), \beta > 0, \alpha \in (0,1]$, i.e. ψ is the Laplace transform of a positive stable random variable. The same kind of result exists for Archimedean copulas (see Genest and Rivest (1989)), which also exhibit the extreme-value property for all $d \geq 2$ if and only if the generator ψ is the Laplace transform of a positive stable random variable[5].

Last but not least, the similarity between Sato–frailty and (extendible) Archimedean copulas manifests in another structural property, namely the behavior in the tails. (Charpentier and Segers, 2009, p. 1525) point out that for bivariate Archimedean copulas C^{A}_{2} with generator ψ, the respective lower tail dependence coefficient introduced in Definition 2.2.16 is given by

$$\lambda_L := \lim_{u \searrow 0} \frac{C^{A}_{2}(u,u)}{u} = 2^{-\frac{1}{\theta_0}}, \tag{4.7}$$

where $\theta_0 := -\lim_{s \searrow 0} s\,(\psi^{-1})'(s)/\psi^{-1}(s)$, provided the limit exists in $[0,\infty]$. Here, $2^{-1/\theta_0}$ is defined as 0 or 1 if θ_0 equals 0 or ∞, respectively.

In Equation (2.8), it has been demonstrated that extendible Archimedean copulas arise as survival copulas of the random sequence

$$X_k := \frac{E_k}{M} = \inf\{t \geq 0 : t\,M \geq E_k\}, \quad k \in \mathbb{N},$$

involving i.i.d. unit exponentially distributed random variables E_k that are independent of an arbitrary positive random variable M. Applying a strictly increasing transformation to each X_k does not affect the survival copula of the corresponding random vector.

[5] The corresponding Archimedean copula is called *Gumbel copula*.

Therefore, for any $H > 0$, extendible Archimedean copulas equivalently arise as survival copulas of $(\tilde{X}_1, \dots, \tilde{X}_d)$, where

$$\tilde{X}_k := X_k^{\frac{1}{H}} = \Big(\frac{E_k}{M}\Big)^{\frac{1}{H}} = \inf\{t \geq 0 : t^H M \geq E_k\}, \quad k \in \mathbb{N}.$$

The resemblance to the frailty construction in Section 3.5 becomes apparent at this point. If the distribution of the random variable M above is self-decomposable, there exists an H-Sato subordinator $\{\Lambda_t\}_{t\geq 0}$ such that $M \stackrel{d}{=} \Lambda_1$ and $\Lambda_t \stackrel{d}{=} t^H \Lambda_1$ for all $t \geq 0$. Thus, the construction of $\tilde{X}_k$ above looks similar to the frailty–setup (3.5.1) with H-Sato subordinators $\{\Lambda_t\}_{t\geq 0}$ in the sense that the marginal distribution functions of the processes $\{\Lambda_t\}_{t\geq 0}$ and $\{t^H M\}_{t\geq 0}$ are the same. However, note that there is still a substantial difference, since in general the processes are not equal in distribution. Nevertheless, the following proposition outlines that the extendible Archimedean and Sato construction result in a similar behavior of the resulting survival copula in terms of tail dependence. The resemblance allows for a nice explicit representation of the lower tail dependence coefficient in the Sato–frailty setup with respect to the Lévy measure corresponding to the infinitely divisible random variable Λ_1.

Proposition 4.3.3 (Tail dependence of Sato–frailty copulas)
Let $\Psi := -\log(\psi)$ be a self-decomposable Bernstein function, i.e. Ψ has a representation as in (2.11), with a=0 and a Lévy density $\nu(\mathrm{d}t) = k(t)/t\,\mathrm{d}t$ such that $t \mapsto k(t)$ is decreasing. The lower and upper tail dependence coefficients of the bivariate Archimedean copula C_2^A with generator ψ and the Sato–frailty copula $C_{\Lambda,2}^{\psi}$ coincide. If Ψ additionally satisfies $b = 0$ in (2.11) and $k(0+) := \lim_{t\searrow 0} k(t) < \infty$, the lower tail dependence coefficient λ_L is given by (4.7) with $\theta_0 = 1/k(0+)$.

Proof
Denote by $\lambda_L^{C_2^A}$ ($\lambda_U^{C_2^A}$) and $\lambda_L^{C_{\Lambda,2}^{\psi}}$ ($\lambda_U^{C_{\Lambda,2}^{\psi}}$) the lower (upper) tail dependence coefficients of C_2^A and $C_{\Lambda,2}^{\psi}$, respectively. According to Definition 2.2.16, it holds that

$$\lambda_L^{C_2^A} = \lim_{u\searrow 0} \frac{C_{\Lambda,2}^{\psi}(u,u)}{u} = \lim_{u\searrow 0} \frac{\psi\big(2\,\psi^{-1}(u)\big)}{u} = \lim_{u\searrow 0} \frac{C_{\Lambda,2}^{\psi}(u,u)}{u} = \lambda_L^{C_{\Lambda,2}^{\psi}},$$
$$\lambda_U^{C_2^A} = 2 - \lim_{u\nearrow 1} \frac{1 - C_2^A(u,u)}{1-u} = 2 - \lim_{u\nearrow 1} \frac{1 - C_{\Lambda,2}^{\psi}(u,u)}{1-u} = \lambda_U^{C_{\Lambda,2}^{\psi}},$$

which establishes the equality of the tail dependence coefficients. Now let $\Psi := -\log(\psi)$ satisfy $b = 0$ and $k(0+) < \infty$.

It remains to show that θ_0 *in* (4.7) *is given by* $\theta_0 = 1/k(0+)$. *It holds that*

$$\begin{aligned}
\theta_0^{-1} &= -\lim_{s \searrow 0} \frac{\psi^{-1}(s)}{s\,(\psi^{-1})'(s)} = -\lim_{s \searrow 0} \frac{\psi^{-1}(s)\,\psi'(\psi^{-1}(s))}{s} \\
&= -\lim_{x \to \infty} \frac{x\,\psi'(x)}{\psi(x)} = \lim_{x \to \infty} x\,\Psi'(x) = \lim_{x \to \infty} x\,\frac{\partial}{\partial x} \int_0^\infty (1 - e^{-x\,t})\,\frac{k(t)}{t}\,\mathrm{d}t \\
&= \lim_{x \to \infty} x \int_0^\infty \frac{\partial}{\partial x}\,(1 - e^{-x\,t})\,\frac{k(t)}{t}\,\mathrm{d}t = \lim_{x \to \infty} x \int_0^\infty e^{-x\,t}\,k(t)\,\mathrm{d}t \\
&= \lim_{x \to \infty} \int_0^\infty e^{-t}\,k(t/x)\,\mathrm{d}t = k(0+),
\end{aligned}$$

having used dominated convergence. □

As $k(0+) < \infty$ is equivalent to $\{\Lambda_t\}_{t \geq 0}$ being finitely active[6] (see Carr et al. (2005)), the theorem above states that for finitely active non-decreasing Sato processes without drift, the lower tail dependence coefficient of the resulting copula is fully characterized by $k(0+)$ and equals the one induced by the respective Archimedean setup.

In addition to that, further properties of Sato–frailty copulas such as Kendall's tau, Spearman's rho, and the singular component of $C^{\psi}_{\Lambda,2}$ can be derived in semi-closed form as a simple integral w.r.t. the function ψ.

Proposition 4.3.4 (Kendall's tau and Spearman's rho of Sato–frailty copulas) *Kendall's tau* $\tau_{C^{\psi}_{\Lambda,2}}$ *and Spearman's rho* $\rho_{C^{\psi}_{\Lambda,2}}$ *of* $C_{\psi,2}$ *are given by*

$$\begin{aligned}
\tau_{C^{\psi}_{\Lambda,2}} &= 4 \int_0^1 \frac{\psi\big(2\,\psi^{-1}(u)\big)^2}{u}\,\mathrm{d}u - 1, \\
\rho_{C^{\psi}_{\Lambda,2}} &= 12 \int_0^1 u\,\psi\big(2\,\psi^{-1}(u)\big)\,\mathrm{d}u - 3.
\end{aligned}$$

Moreover, for $(U_1, U_2) \sim C^{\psi}_{\Lambda,2}$, *it holds that*

$$\mathbb{P}(U_1 = U_2) = 2 \int_0^1 \frac{\psi\big(2\,\psi^{-1}(u)\big)}{u}\,\mathrm{d}u - 1.$$

Proof

The results follow from the general formulas in Section 3.2. □

[6] $\{\Lambda_t\}_{t \geq 0}$ being finitely active means that the process jumps finitely often within every finite interval. In contrast, infinite activity induces infinitely many jumps in any interval.

As a mnemonic trick, note that when rewriting

$$\mathbb{P}(U_1 = U_2) = 4 \int_0^1 \frac{\frac{1}{2}\,\psi\big(2\,\psi^{-1}(u)\big)}{u}\,\mathrm{d}u - 1,$$

the probability $\mathbb{P}(U_1 = U_2)$ looks very similar to Kendall's τ, the difference being that $\psi\big(2\,\psi^{-1}(u)\big)$ in the integrand is multiplied with $1/2$ instead of being squared.

Summing up, there is an intimate relationship between Sato–frailty and extendible Archimedean copulas. Similarities between Archimedean and exchangeable Marshall–Olkin copulas of the form $C(u_1, \ldots, u_d) = \prod_{k=1}^{d} u_{(k)}^{a_{k-1}}$ have already been discussed in (Mai, 2010, Section 3.6). However, the results in the present paragraph show that within the class of additive–frailty copulas, the Sato–frailty subclass bears a much more distinct resemblance to the (extendible) Archimedean setup than the Lévy–frailty construction for MO copulas in Section 3.6. Table 4.1 gives an overview on the commonalities. On a high level, some of these analogies can be explained by the fact that the diagonal section[7] S of a Sato–frailty copula $C^{\psi}_{\Lambda,d}$ is given by

$$S(u) = C^{\psi}_{\Lambda,d}(u, \ldots, u) = \prod_{k=1}^{d} \frac{\psi\big(k\,\psi^{-1}(u)\big)}{\psi\big((k-1)\,\psi^{-1}(u)\big)} = \psi\big(d\,\psi^{-1}(u)\big),$$

and, thus, equals the diagonal section of an Archimedean copula C^{A}_{d} with generator ψ. As an immediate consequence, being a function of the diagonal section, lower and upper tail dependence coefficient naturally coincide in both setups.

4.3.2 Simulation

This paragraph illustrates how to simulate copulas of type $C^{\psi}_{\Lambda,d}$ and presents some parametric families, one of which will be used to apply and verify the statistical results derived in the previous section. For Sato–frailty copulas, Algorithm 3.5.2 reads as follows.

Algorithm 4.3.5 (Simulation of d-dimensional copulas $C^{\psi}_{\Lambda,d}$)

1. *Simulate d independent, unit exponentially distributed random variables $E_1, \ldots, E_d$.*
2. *Simulate one path of the 1-Sato subordinator $\{\Lambda_t\}_{t\geq 0}$ corresponding to ψ until $\Lambda_t \geq \max\{E_1, \ldots, E_d\}$.*

[7] Given a copula $C : [0,1]^d \to [0,1]$, the diagonal section $S : [0,1] \to [0,1]$ of C is defined by $S(u) := C(u, \ldots, u)$.

	Archimedean copula	Sato–frailty copula	Source/Reference
Copula function C	$\psi\big(\psi^{-1}(u_1)+\ldots+\psi^{-1}(u_d)\big)$	$\prod_{k=1}^{d} \frac{\psi\left(k\,\psi^{-1}(u_{(k)})\right)}{\psi\left((k-1)\,\psi^{-1}(u_{(k)})\right)}$	Definition 2.3.11, Theorem 4.2.1.(iii)
Generating function ψ	Archimedean generator	Archimedean generator	Definition 2.3.10
Uniqueness of ψ	Up to stretching/compression	Up to stretching/compression	Equation (4.3)
Extendibility of copula	$\psi(x)=\mathbb{E}[\exp(-x\,M)]$, M positive random variable	$\psi(x)=\mathbb{E}[\exp(-x\,\Lambda_1)]$, Λ_1 selfd. positive random variable	Theorem 2.3.13, Theorem 4.2.1
Stochastic model for extendible $(U_1,\ldots,U_d)\sim C$	$U_k:=\psi\big(\inf\{t\geq 0: t\,M\geq E_k\}\big)$, M positive random variable	$U_k:=\psi\big(\inf\{t\geq 0:\Lambda_t\geq E_k\}\big)$, $\{\Lambda_t\}_{t\geq 0}$ 1-Sato subordinator	Equation (2.8), Proof of Theorem 4.2.1
LTD coefficient	$\lim_{u\searrow 0}\psi\big(2\,\psi^{-1}(u)\big)$	$\lim_{u\searrow 0}\psi\big(2\,\psi^{-1}(u)\big)$	Proposition 4.3.3
UTD coefficient	$2-\lim_{u\nearrow 1}\frac{\partial}{\partial u}\Big(\psi\big(2\,\psi^{-1}(u)\big)\Big)$	$2-\lim_{u\nearrow 1}\frac{\partial}{\partial u}\Big(\psi\big(2\,\psi^{-1}(u)\big)\Big)$	Proposition 4.3.3
Extreme-value subclass	$\psi(x)=\exp(-\beta\,x^{\alpha})$, $\beta>0,\alpha\in(0,1]$	$\psi(x)=\exp(-\beta\,x^{\alpha})$, $\beta>0,\alpha\in(0,1]$	Proposition 4.3.1
Singular component	$\mathbb{P}(U_1=\ldots=U_d)=0$	$\mathbb{P}(U_1=\ldots=U_d)>0$	Proposition 4.3.4

Table 4.1 Comparison between extendible Archimedean and Sato–frailty copulas.

3. *Compute* $X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d.$

4. *Set* $U_k := \psi(X_k), k = 1, \ldots, d,$ *and return* $(U_1, \ldots, U_d)$.

The crucial task in Algorithm 3.5.2 is the simulation of $\{\Lambda_t\}_{t\geq 0}$. One possible approach can be deduced from the general results in Marsaglia and Tsang (1984). As an alternative, sampling the independent, yet not necessarily identically distributed increments of a Sato process can be simplified by using the connection to Lévy processes. Following (Jeanblanc et al., 2002, Theorem 1), every H-Sato process $\{\Lambda_t\}_{t\geq 0}$ corresponds to a Lévy process $\{L_t\}_{t\geq 0}$ with $\mathbb{E}[\log(\max\{1, |L_s|\})] < \infty$ for all $s \geq 0$ (and vice versa) via the relationship

$$\Lambda_r = \begin{cases} \int_{\log(1/r)}^{\infty} e^{-t\,H}\, dL_t^{(-)}, & 0 \leq r \leq 1, \\ \Lambda_1 + \int_0^{\log(r)} e^{t\,H}\, dL_t^{(+)}, & r \geq 1, \end{cases} \tag{4.8}$$

where $\{L_t^{(-)}\}_{t\geq 0}, \{L_t^{(+)}\}_{t\geq 0}$ are two independent copies of $\{L_{tH}\}_{t\geq 0}$. In Jurek and Yor (2004), it is pointed out that the coherence between $\{\Lambda_t\}_{t\geq 0}$ and $\{L_{tH}\}_{t\geq 0}$ can be expressed in terms of the Fourier transforms

$$\begin{aligned} \log\left(\mathbb{E}\left[e^{i\,t\,\Lambda_1}\right]\right) &= \int_0^t \log\left(\mathbb{E}\left[e^{i\,v\,L_H}\right]\right)\frac{dv}{v}, \quad t \in \mathbb{R}, \\ \log\left(\mathbb{E}\left[e^{i\,t\,L_H}\right]\right) &= t\,\frac{\partial}{\partial\, t}\left(\log\left(\mathbb{E}\left[e^{i\,t\,\Lambda_1}\right]\right)\right), \quad t \in \mathbb{R}. \end{aligned}$$

On the level of Laplace exponents for Sato and Lévy subordinators, this translates to

$$\begin{aligned} \log\left(\mathbb{E}\left[e^{-t\,\Lambda_1}\right]\right) &= \int_0^t \log\left(\mathbb{E}\left[e^{-v\,L_H}\right]\right)\frac{dv}{v}, \quad t \geq 0, \\ \log\left(\mathbb{E}\left[e^{-t\,L_H}\right]\right) &= t\,\frac{\partial}{\partial\, t}\left(\log\left(\mathbb{E}\left[e^{-t\,\Lambda_1}\right]\right)\right), \quad t \geq 0. \end{aligned} \tag{4.9}$$

The decomposition in Equation (4.8) implies that a Sato process is infinitely active, meaning that it jumps infinitely often on every compact interval, if and only if the underlying Lévy process is. In the infinitely active case, the path of $\{L_t\}_{t\geq 0}$ can not be simulated exactly and has to be approximated. In case of finite activity, the Lévy process is of compound Poisson type and exact simulation is possible, provided that the jump size distribution can be simulated accurately. By means of Equation (4.8), it holds that

$$\begin{aligned} \Lambda_r - \Lambda_{r_0} &= \int_{\log(1/r)}^{\log(1/r_0)} e^{-t\,H}\, dL_t^{(-)} && \text{for } r_0 < r \leq 1, \\ \Lambda_r - \Lambda_1 &= \int_0^{\log(r)} e^{t\,H}\, dL_t^{(+)} && \text{for } r \geq 1. \end{aligned} \tag{4.10}$$

Consequently, the sampling approach can be realized as follows.

Algorithm 4.3.6 (Simulation of finitely active Sato processes)

1. *Sample Λ_{r_0} for very small r_0 (which can be achieved easily if the Fourier, respectively Laplace transform of Λ_{r_0} is known analytically).*
2. *Simulate the jump times and heights of $\{L_t^{(-)}\}_{t\geq 0}$ on $[0, \log(1/r_0)]$.*
3. *If $\max\{E_1, \ldots, E_d\} > \Lambda_1$, subsequently simulate the jump times and heights of $\{L_t^{(+)}\}_{t\geq 0}$ and extract $\{\Lambda_t\}_{t\geq 0}$ for $t > 1$ from (4.10) until $\Lambda_t \geq \max\{E_1, \ldots, E_d\}$.*

4.3.3 Examples

We can insert any self-decomposable Bernstein function Ψ in $\psi = \exp(-\Psi)$ to define a valid Sato–frailty copula $C^{\psi}_{\Lambda,d}$ in arbitrary dimension $d \geq 2$. A wide range of (not necessarily self-decomposable) Bernstein functions can be found in (Schilling et al., 2010, p. 218-277). To begin with, we focus on the Gamma distribution introduced in Section 2.4.5 with corresponding Bernstein function

$$\Psi^{\mathrm{Ga}}_{\beta,\eta}(x) = \int_0^{\infty} (1 - e^{-x\,t}) \underbrace{\beta\, \frac{\exp(-\eta\, t)}{t}\, \mathrm{d}t}_{:=\nu(\mathrm{d}t)} = \beta\, \log\Big(1 + \frac{x}{\eta}\Big), \quad x, \beta, \eta > 0.$$

We have already noted that ν has the form $\nu(\mathrm{d}t) = k(t)/t\, \mathrm{d}t$, with $k(t) := \beta\, \exp(-\eta\, t)$ being decreasing in t, implying $\Psi^{\mathrm{Ga}}_{\beta,\eta}$ to be self-decomposable. Moreover, as pointed out at the beginning of Section 4.3, the associated Sato–frailty copula $C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,\eta})}_{\Lambda,d}$ is invariant w.r.t. stretching or compressing $\Psi^{\mathrm{Ga}}_{\beta,\eta}$. Therefore, the parameter η has no influence on the copula and

$$\begin{aligned} C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,\eta})}_{\Lambda,d}(u_1, \ldots, u_d) &= C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,d}(u_1, \ldots, u_d) \\ &= \Bigg(\prod_{k=1}^{d} \frac{1 + (k-1)\left(u_{(k)}^{-\frac{1}{\beta}} - 1\right)}{1 + k\left(u_{(k)}^{-\frac{1}{\beta}} - 1\right)} \Bigg)^{\beta}, \quad u_1, \ldots, u_d \in [0,1], \end{aligned}$$

which in the bivariate case simplifies to

$$C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}(u_1, u_2) = \frac{u_{(1)}}{\left(2 - u_{(2)}^{\frac{1}{\beta}}\right)^{\beta}}.$$

Interestingly, the Gamma example allows to explicitly derive and interpret the Lévy process connected to the Sato process in (4.8). Applying the relations in Equation (4.9),

the Lévy subordinator $\{L_{tH}\}_{t\geq 0}$ corresponding to the H-Sato subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ is characterized by (see also Barndorff-Nielsen and Sheppard (2001))

$$\Psi^{L_H}(t) := -\log\Big(\mathbb{E}[e^{-t\,L_H}]\Big) = t\,\frac{\partial}{\partial\, t}\Psi^{\mathrm{Ga}}_{\beta,\eta}(t) = t\,\beta\,\frac{1}{\eta+t} = \beta\big(1-\frac{\eta}{\eta+t}\big), \quad t\geq 0.$$

This Bernstein function is well-known. It has already been introduced in Section 2.4.5 and is the Laplace exponent of a Compound Poisson random variable with exponentially distributed jump sizes. Consequently, the associated Lévy process is a Compound Poisson process and can be simulated efficiently and accurately (see (Sato, 1999, p. 18 ff.)). Thus, we can use either conditional sampling (for dimension $d = 2$) or exact path simulation of the Sato subordinator according to Algorithm 4.3.6 by means of the underlying Lévy subordinator (for arbitrary $d \geq 2$) to create realizations of the above copula. Figures 4.1 and 4.2 illustrate some exemplary sample plots for dimensions $d = 2$ and $d = 3$ and for varying values of β.

It becomes obvious that the lower the value of β, the more concentrated the samples of $(U_1, U_2) \sim C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}$ respectively $(U_1, U_2, U_3) \sim C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,3}$ on the diagonal of the unit square respectively cube. This observation is in line with Proposition 4.3.4: Kendall's τ of $C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}$, respectively the singular component $\mathbb{P}(U_1 = U_2), (U_1, U_2) \sim C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}$, are given by

$$\begin{aligned}
\tau_{C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}} &= 4\int_0^1 \frac{\big(2\,u^{-\frac{1}{\beta}}-1\big)^{-2\,\beta}}{u}\,\mathrm{d}u - 1,\\
\mathbb{P}(U_1 = U_2) &= 2\int_0^1 \frac{\big(2\,u^{-\frac{1}{\beta}}-1\big)^{-\beta}}{u}\,\mathrm{d}u - 1.
\end{aligned}$$

Table 4.2 depicts the resulting values for varying parameters of β. Apparently, both Kendall's tau and the probability $\mathbb{P}(U_1 = U_2)$ are decreasing from 1 to 0 with increasing β. The table also indicates the lower tail dependence coefficient of the Sato–frailty copula in the Gamma specification. The general results in Section 3.2 show that the tail dependence coefficients of $C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}$ are given by

$$\begin{aligned}
\lambda_L^{C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}} &= g_2(0+) = 2^{-\beta},\\
\lambda_U^{C^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}_{\Lambda,2}} &= 1 - g_2'(1-) = \lim_{u\nearrow 1}\frac{u^{\frac{1}{\beta}-1}}{(2-u^{\frac{1}{\beta}}))^{\beta+1}} = 0.
\end{aligned}$$

Note that this validates the properties of Sato–frailty copulas derived in Section 4.3.1. As the Lévy measure $\nu(\mathrm{d}t) = k(t)/t\,\mathrm{d}t$, $k(t) = \beta\,\exp(-\eta\,t)$, $t \geq 0$, corresponding to the

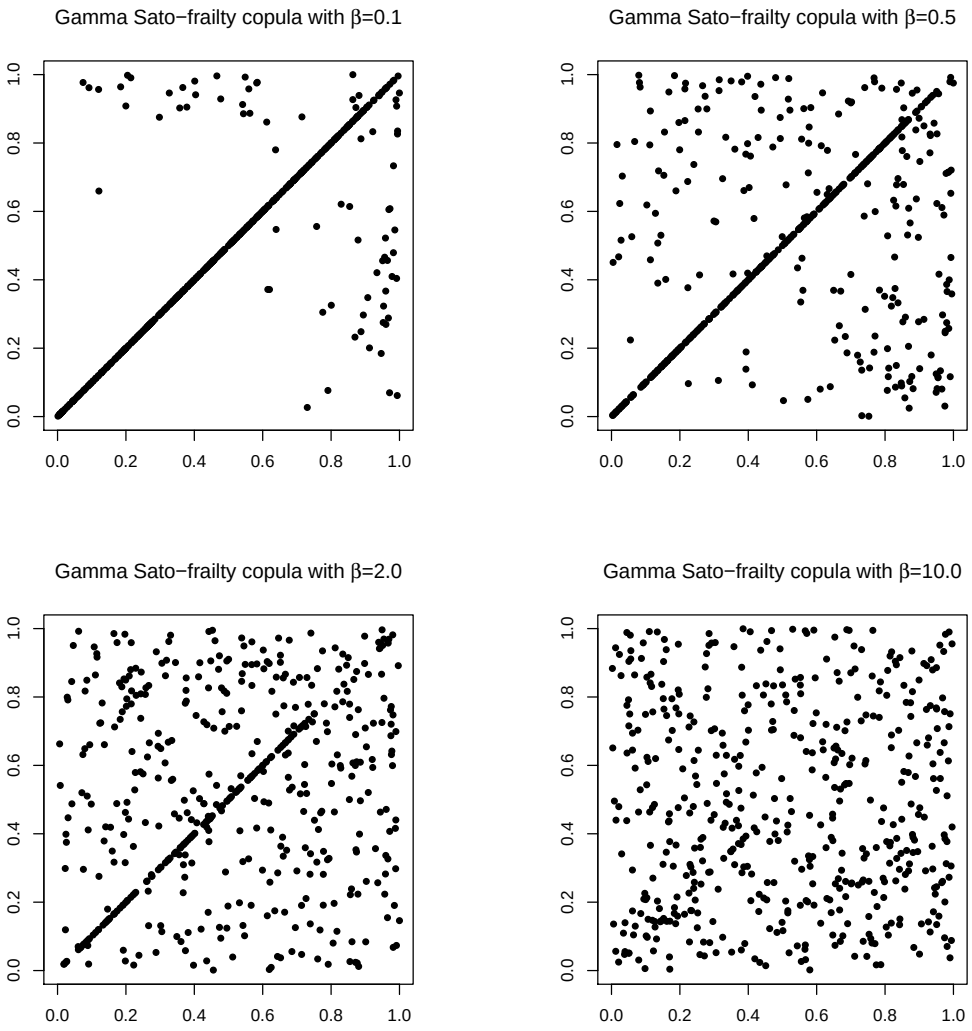

Figure 4.1 500 realizations of $C^{\mathrm{Ga}}_{\psi,2}$ for two different values of β.

Gamma distribution satisfies $k(0+) = \beta < \infty$, Proposition 4.3.3 states that

$$\lambda_U^{C_{\Lambda,2}^{\exp(-\Psi^{\mathrm{Ga}}_{\beta,1})}} = 2^{-k(0+)} = 2^{-\beta}.$$

As a last illustration, we visualize the coherence between extendible Archimedean and Sato–frailty copulas as depicted in Section 4.3.1. We consider the self-decomposable

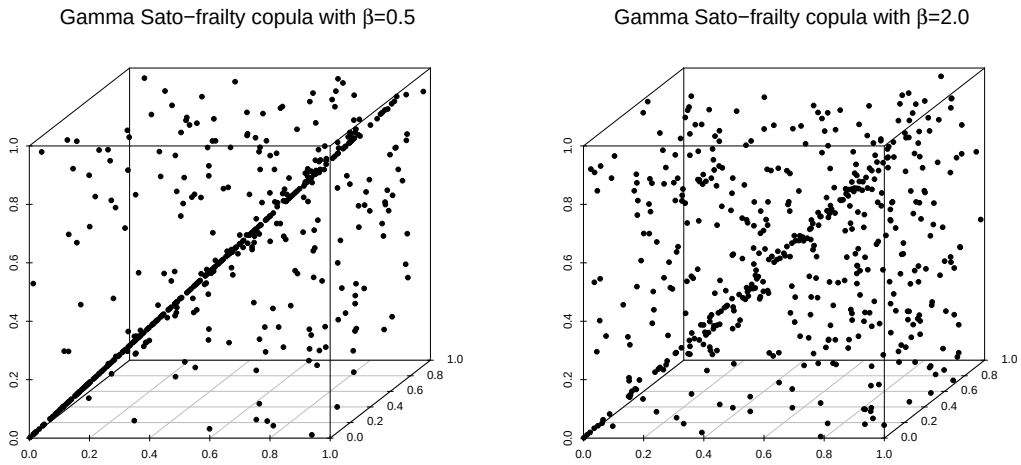

Figure 4.2 500 realizations of $C_{\Lambda,3}^{\exp(-\Psi_{\beta,1}^{\mathrm{Ga}})}$ for varying values of β.

β	$\tau_{C_{\Lambda,2}^{\exp(-\Psi_{\beta,1}^{\mathrm{Ga}})}}$	$\mathbb{P}(U_1 = U_2)$	$\lambda_L^{C_{\Lambda,2}^{\exp(-\Psi_{\beta,1}^{\mathrm{Ga}})}}$
0.01	0.9729	0.9846	0.9931
0.05	0.8762	0.9346	0.9659
0.1	0.7766	0.8762	0.9330
0.3	0.5225	0.6952	0.8123
0.5	0.3863	0.5708	0.7071
1	0.2274	0.3863	0.5000
2	0.1215	0.2274	0.2500
10	0.0250	0.0497	0.0010
100	0.0025	0.0050	0.0000

Table 4.2 Kendall's τ, $\mathbb{P}(U_1 = U_2)$, and $\lambda_L^{C_{\psi,2}^{\mathrm{Gamma}}}$ for the copula $C_{\Lambda,2}^{\exp(-\Psi_{\beta,1}^{\mathrm{Ga}})}$ and varying values of $\beta > 0$. Computations are carried out using standard Matlab integration algorithms.

Bernstein functions corresponding to the positive stable and Inverse Gaussian distribution in Section 2.4.5. Due to the invariance of the Sato–frailty copula $C_{\Lambda,d}^{\psi}$ w.r.t. stretching/compression of ψ, the following normalized Bernstein functions $\Psi := -\log(\psi)$ are

taken into account.

(i) Positive stable: $\Psi^{\mathrm{St}}_{\alpha,1}(x) = x^{\alpha}, \quad 0 < \alpha < 1, \quad x \geq 0.$

(ii) Inverse Gaussian: $\Psi^{\mathrm{IG}}_{1,\eta}(x) = \sqrt{2x + \eta^2} - \eta, \quad \eta > 0, \quad x \geq 0.$

As for the Gamma example, these specifications lead to one-parametric Archimedean and Sato–frailty copulas. Figures 4.3 and 4.4 contrast the resulting bivariate scatterplots in the positive stable and the Inverse Gaussian case. In order to compare Archimedean and Sato–frailty setup, the parameter α respectively η is set to the identical value for both copula families. At a first glance, the Sato–frailty and Archimedean framework look quite diverse, an impression that is strengthened by the difference w.r.t. the singular component, i.e. the diagonal of the unit square. Looking closer, however, one perceives the indicated similarities. On the one hand, the behavior in the tails, i.e. in the left lower and right upper corner of the unit square, is perceptible. On the other hand, making a rough estimate on the points within the squares $[0,u]\times[0,u], u \in [0,1]$, it becomes at least plausible that it is approximately the same in the Archimedean and Sato–frailty case. This guess is quantified by Figure 4.5, which shows the diagonal section for both copula families in the positive stable and Inverse Gaussian example. For a more meaningful and smoother comparison, the latter plot refers to an increased sample size of 10,000.

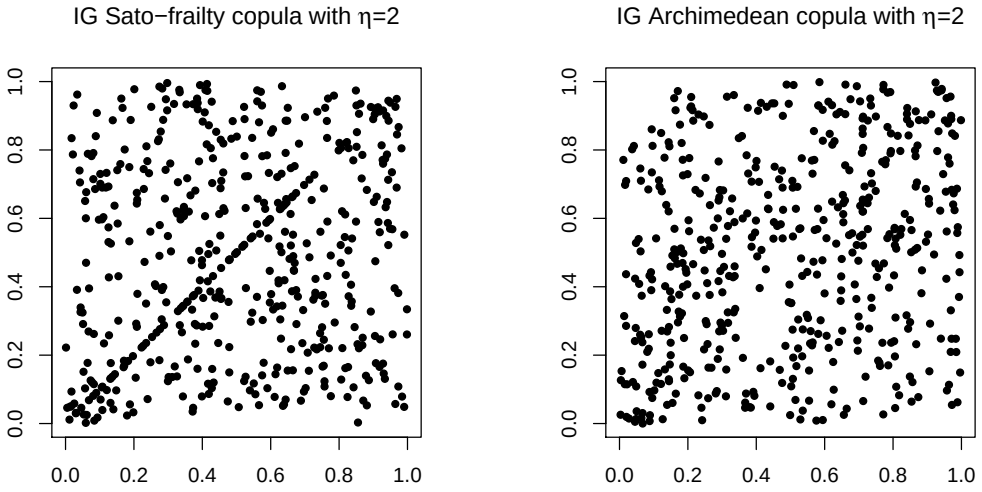

Figure 4.3 500 realizations of $C^{\exp(-\Psi^{\mathrm{IG}}_{1,\eta})}_{\psi,2}$ (left) and the Archimedean copula C_2 (right) for the Inverse Gaussian distribution and $\eta = 2$.

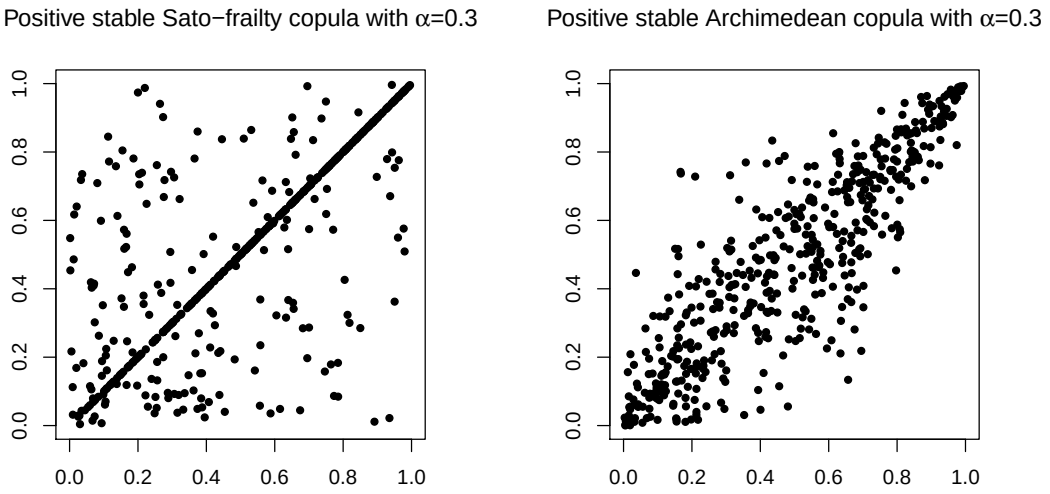

Figure 4.4 500 realizations of $C_{\psi,2}^{\exp(-\Psi_{\alpha,1}^{\mathrm{St}})}$ (left) and the Archimedean copula C_2 (right) for the positive stable distribution and $\alpha = 0.3$.

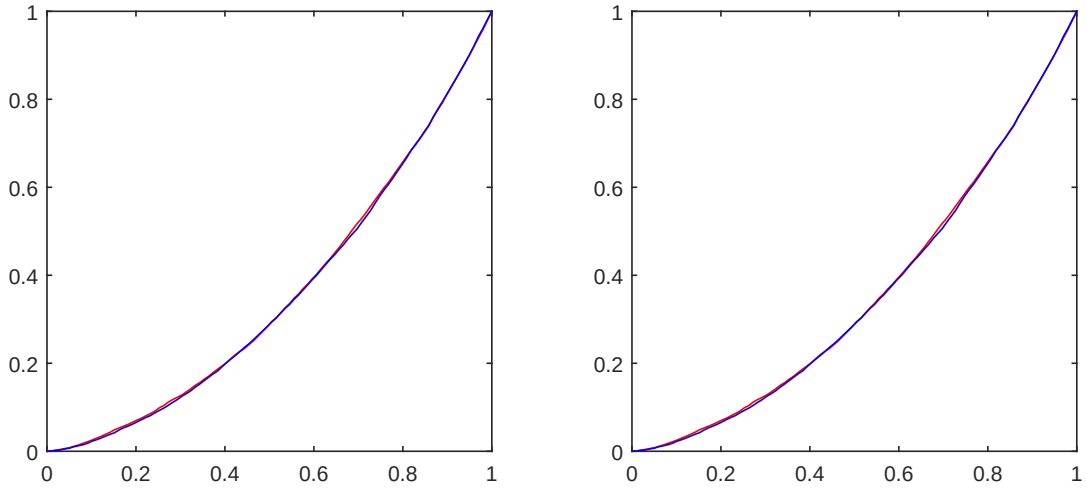

Figure 4.5 Empirical diagonal section $u \mapsto C(u,u), u \in [0,1]$, in the Inverse Gaussian (left) and positive stable (right) case. The parameters are the same as in Figures 4.3 and 4.4, however, the sample size has been set to 10,000. In both plots, the red line indicates the diagonal section of the Sato–frailty specification, while the Archimedean setup is depicted in blue.

5 CDO pricing with additive–frailty copulas

This chapter demonstrates how to apply exchangeable exogenous shock models to the pricing of CDOs in the spirit of Mai and Scherer (2009c). It is premised on [Schenk (2011), Mai, Olivares, Schenk, Scherer (2014)]. The payoff structures of these multiname credit derivatives and the related index credit default swaps (index CDS) are sketched in Section 5.1. Both instruments are basically linked to the cumulated losses in a high-dimensional portfolio of assets. Given the mathematical complexity of the loss distribution for large portfolios, considering the subclass of additive–frailty models in Section 3.5 for modeling the default times is motivated. Providing a useful limit theorem for the portfolio loss process, the special structure of the additive–frailty framework is exploited in Section 5.3 to derive an efficient pricing formula for CDOs when the portfolio size tends to infinity. The implementation of the formula for particular additive–frailty models is based on recent results in [Bernhart, Mai, Schenk, Scherer (2015)] regarding the Laplace inversion of *Bondesson distributions* and, together with the connection to shock models, manifests the main contribution of this chapter. "Operating instructions" for the calibration of additive–frailty models to market prices of CDOs and index CDS are given in Section 5.4.

5.1 Multiname credit derivatives

According to (Hull, 2008, p. 525), credit derivatives refer to "*contracts where the payoff depends on the creditworthiness of one or more companies or countries*". More general, Bielecki and Rutkowski (2001) specify them as "*privately negotiated derivative securities that are linked to a credit-sensitive asset (index) as the underlying asset (index)*" where the reference security "*can be any financial instrument that is subject to risk of default (or, more generally, to the credit risk)*". By credit risk, they pool the "*risk associated with any kind of credit-linked events, such as [...] changes in the credit quality [...], variations of*

credit spreads, and the default event". In short, multiname credit derivatives are bilateral contracts in which two parties agree to exchange payoffs dependent on realized losses in an underlying portfolio. The realized loss resulting from an asset's default is essentially described by three characteristics. First, the default time X indicates the random point in time of the asset's bankruptcy. Second, the constant N represents the nominal invested in the asset. Third, the final loss at time X might be less than the notional N due to a recovery payment of the respective company, denoted by the (possibly random) variable Rec. For a given pool of d assets, the total accumulated loss at time $t \geq 0$, denoted Loss_t, is given as the sum of the individual losses and, thus, equals

$$\text{Loss}_t = \sum_{k=1}^{d} N_k \left(1 - \text{Rec}_k\right) \mathbb{1}_{\{X_k \leq t\}}.$$

To facilitate the tractability of the central object Loss_t, one often neglects the possibility of random, inhomogeneous recovery rates. Though this restriction is not backed by empirical evidence (for instance, recovery rates have repeatedly found to be negatively correlated to default frequencies, see Andersen and Sidenius (2004) and the references therein), we proceed according to market practice and fix $\text{Rec}_k = \text{Rec}, k = 1, \ldots, d$, for a constant $0 \leq \text{Rec} \leq 1$ (typically, one sets $\text{Rec} = 0.4$). In addition to that, when considering identical nominal values and assuming a total investment of one unit to be equally distributed among the d assets, it holds that

$$\text{Loss}_t = (1 - \text{Rec}) \underbrace{\frac{1}{d} \sum_{k=1}^{d} \mathbb{1}_{\{X_k \leq t\}}}_{=:L_t} = (1 - \text{Rec})\, L_t, \quad t \geq 0. \tag{5.1}$$

In this context, the process $\{L_t\}_{t \geq 0}$ is termed *portfolio loss process* and denotes the fraction of defaulted firms in the portfolio in time. In the sequel, we always assume that Loss_t is given in this simplified way.

Many multiname credit derivatives essentially depend on the distribution of L_t at fixed points in time. Two examples are *index credit default swaps* (index CDS) and *collateralized debt obligations* (CDOs). An index CDS can be thought of as an insurance-like contract between two parties, with the assets in the portfolio representing the subject matter of the insurance. Until maturity, the *buyer* of the index CDS periodically pays a fixed percentage (termed *premium*) of the remaining nominal of the portfolio to the *seller* and in exchange receives a monetary compensation whenever defaults occur. In our setup, at the beginning of the contract, the buyer has to pay the premium on the total nominal of one. Once a default occurs, the seller has to make a payment equal to

the associated realized loss. For instance, once asset k defaults, the realized loss at time X_k is given by

$$\text{Loss}_{X_k} - \lim_{\epsilon \searrow 0} \text{Loss}_{X_k - \epsilon} = (1 - \text{Rec}) \left(L_{X_k} - \lim_{\epsilon \searrow 0} L_{X_k - \epsilon} \right) = (1 - \text{Rec})/d.$$

After time X_k, the buyer's premium payments are guided by the remaining notional, which is reduced by an absolute amount of $1/d$ compared to the nominal before default.

Whereas an index CDS guarantees protection against all losses in a given portfolio, the idea of a synthetic CDO is to provide insurance against parts of the portfolio by dividing the total realized loss Loss_t into loss segments (called *tranches*) determined by lower and upper attachment points l_j respectively u_j, with $0 = l_1 < u_1 = l_2 < u_2 = l_3 \ldots < u_J = 1$. In a CDO contract for tranche $[l_j, u_j]$, $j \in \{1, \ldots, J\}$, both premium and default payments at time t are guided by the tranche-specific realized loss $L_{t,j}$, which is given by (see Figure 5.1 for a visualization)

$$L_{t,j} := \min \Big(\max \big(0, (1 - \text{Rec}) L_t - l_j \big), u_j - l_j \Big).$$

The payment procedure is similar to an index CDS. At the beginning, the buyer's

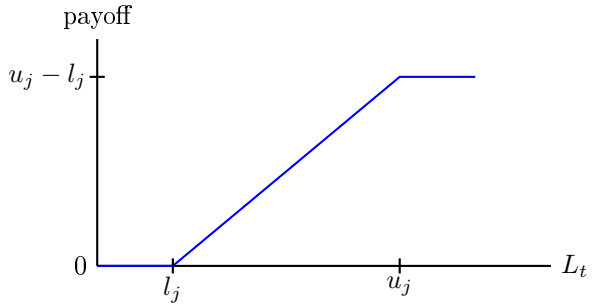

Figure 5.1 Payoff profile of tranche loss $L_{t,j}$ with respect to the portfolio loss process L_t at time t.

premium payment is proportional to the whole tranche nominal $u_j - l_j$. Once the first asset, for example company k, defaults, the realized tranche loss compensated by the CDO seller equals $L_{X_k,j} - \lim_{\epsilon \searrow 0} L_{X_k - \epsilon, j}$. After the default, the premium payments of the buyer refer to the reduced tranche nominal $u_j - l_j - L_{t,j}$. Consequently, except for the first tranche $[l_1, u_1] = [0, u_1]$ (also termed *equity tranche* in practice), initial defaults in the portfolio do not trigger payments in a CDO. As long as the realized tranche loss $L_{t,j}$ does not exceed the lower attachment point l_j, defaults neither reduce the buyer's premium payments nor trigger reimbursements by the seller. Once the realized tranche

loss is above l_j, the payment streams evolve as described until the end of the maturity respectively until the upper attachment point u_j is exceeded as well.

For a detailed depiction of the payment streams, including additional quantities such as discount factors or upfront payments, see Brigo et al. (2007). Further subtleties such as accrued interest are out of scope of the present analysis. From a pricing point of view, index CDS and CDOs essentially depend on the (risk-neutral) distribution of the portfolio loss L_t for $t \geq 0$. In an index CDS, both premium and default payments depend linearly on the portfolio loss process, such that the evaluation reduces to the calculation of risk-neutral expected values

$$\mathbb{E}[L_t] = \mathbb{E}\Big[\frac{1}{d}\sum_{k=1}^{d} \mathbb{1}_{\{X_k \leq t\}}\Big] = \frac{1}{d}\sum_{k=1}^{d} \mathbb{P}\big(X_k \leq t\big), \quad t \geq 0.$$

Thus, when modeling the vector $(X_1, \ldots, X_d)$ of default times, the dependence structure is irrelevant and only the marginal distributions of $X_k, k = 1, \ldots, d$, matter. The evaluation of CDOs is more complex. Here, the central quantity is the tranche-specific realized loss $L_{t,j}$, demanding the computation of $\mathbb{E}[L_{t,j}]$ for $t \geq 0, j \in \{1, \ldots, J\}$. Keeping in mind the payoff profile of $L_{t,j}$ in Figure 5.1, one recognizes that $L_{t,j}$ has a bull spread structure and, therefore, can be decomposed into the difference of two put (alternatively call) options on the portfolio loss L_t by

$$\mathbb{E}[L_{t,j}] = u_j - l_j + \mathbb{E}\Big[\big(l_j - (1 - \mathrm{Rec})\, L_t\big)^+\Big] - \mathbb{E}\Big[\big(u_j - (1 - \mathrm{Rec})\, L_t\big)^+\Big], \tag{5.2}$$

where "$(.)^+$" denotes the maximum between the expression in brackets and zero. Due to this non-linear relationship between $L_{t,j}$ and L_t, CDOs are sensitive w.r.t. the dependence structure of the default times $X_1, \ldots, X_d$. Indeed, the financial crisis has demonstrated the huge impact of default correlation on the evaluation of these multi-name credit derivatives. The challenge in finding a suitable model is exacerbated by the generally large dimensionality of the underlying portfolio. For the major markets, broad indices of investment grade companies have been created, serving as a basis for standardized index CDS and CDOs. For North America, the respective index is abbreviated *CDX.NA.IG*, the European counterpart is termed *iTraxx Europe*. Detailed and up-to-date information on the attachment points l_j, u_j, and further practical conventions can be found in Markit (2014). Both indices comprise $d = 125$ (periodically updated) companies (see Amato and Gyntelberg (2005) for a detailed composition), raising the question how to find a suitable, yet tractable model for the vector $(X_1, \ldots, X_d)$ of default times.

5.2 Additive–frailty model for default times

It is beyond debate that modeling the high-dimensional default time vector $(X_1, \ldots, X_d)$ without losing tractability of the portfolio loss process $L_t = 1/d \sum_{k=1}^{d} \mathbb{1}_{\{X_k \leq t\}}$, i.e. the sum of d random variables, requires severe restrictions on the random vector's dependence structure. Even if it was possible to derive the distribution of L_t for a complex, heterogeneous model, calibration to market data would be almost impossible as the universe of credit derivatives linked to dependencies within each subset of $(X_1, \ldots, X_d)$ is generally sparse.

A common escape from this situation consists in directly modeling the portfolio loss process in a *top-down approach*. The process $\{L_t\}_{t \geq 0}$ is specified as an increasing, right-continuous, piecewise constant stochastic process starting at $L_0 = 0$, with jumps representing an asset's default. In many cases (see for example Giesecke et al. (2010b)), this task is achieved by considering a portfolio loss point process and modeling the dynamics of its conditional arrival rate. Other examples of top-down models are given in Arnsdorf and Halperin (2008), Brigo et al. (2007), Cont and Minca (2008), Giesecke et al. (2010a), Lopatin and Misirpashaev (2008), or Schönbucher (2005). While being appealing due to their flexibility, top-down models usually suffer from a limited knowledge about the underlying default times, i.e. the difficulty of linking the distribution of the aggregated quantity L_t to the individual random variables $X_k, k = 1, \ldots, d$. This deficiency induces a model that is artificial to some extent in the sense that it provides no interpretation for the stochastic origin behind.

This drawback is circumvented when using *bottom-up models*. Starting with a direct definition of the default times X_k, the stochastic construction is given by definition. However, the challenge entails dealing with the distribution of the sum L_t. An extensive overview on bottom-up approaches and corresponding literature is given in (Mai, 2010, Section 6.2) and the references therein. In Giesecke (2003), the default times are modeled via the exogenous shock model introduced in Definition 3.1.1 for exponentially distributed random variables Z^E. Given the characterization results in the present thesis, we are aiming at modeling $(X_1, \ldots, X_d)$ by a general exchangeable exogenous shock model. Assuming the marginal survival function (which is the same for all components due to exchangeability) of $X_k, k = 1, \ldots, d$, to be continuous and given by $\bar{F}$, by Theorem 3.3.1, we already know that the survival copula of $(X_1, \ldots, X_d)$ is given by an explicitly

known copula C of type (3.4), i.e. we have

$$\begin{aligned}\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) &= \mathbb{P}\big(\bar{F}(X_1) \leq \bar{F}(x_1), \dots, \bar{F}(X_d) \leq \bar{F}(x_d)\big) \\ &= C(\bar{F}(x_1), \dots, \bar{F}(x_d)) = \prod_{k=1}^{d} g_k\big(\bar{F}(x_{d+1-k})\big), \quad x_1, \dots, x_d \in \mathbb{R},\end{aligned}$$

for known functions $g_k, k = 1, \dots, d$. Consequently, the distribution of the portfolio loss process can be derived in closed form. For $x = l/d, l \in \{0, \dots, d\}$, it holds that

$$\begin{aligned}\mathbb{P}(L_t = x) &= \mathbb{P}\Big(\sum_{k=1}^{d} \mathbb{1}_{\{X_k \leq t\}} = l\Big) \overset{(*)}{=} \binom{l}{d} \mathbb{P}(X_1 \leq t, \dots, X_l \leq t, X_{l+1} > t, \dots, X_d > t) \\ &= \binom{l}{d} \mathbb{P}\big(\bar{F}(X_1) \geq \bar{F}(t), \dots, \bar{F}(X_l) \geq \bar{F}(t), \bar{F}(X_{l+1}) < \bar{F}(t), \dots, \bar{F}(X_d) < \bar{F}(t)\big), \quad (5.3)\end{aligned}$$

where we have used exchangeability in $(*)$. Given that $\big(\bar{F}(X_1), \dots, \bar{F}(X_d)\big) \sim C$, the proof of Proposition 2.2.6 has shown that the latter probability can be expressed in terms of an alternating sum. More precisely, it holds that

$$\begin{aligned}&\mathbb{P}\big(\bar{F}(X_1) \geq \bar{F}(t), \dots, \bar{F}(X_l) \geq \bar{F}(t), \bar{F}(X_{l+1}) < \bar{F}(t), \dots, \bar{F}(X_d) < \bar{F}(t)\big) \\ = &\sum_{(w_1, \dots, w_d) \in \times_{i=1}^{d} \{u_i, v_i\}} (-1)^{|\{i \,:\, w_i = u_i\}|} \, C(w_1, \dots, w_d), \quad (5.4)\end{aligned}$$

for $\{u_1, v_1\} = \dots = \{u_l, v_l\} = \{\bar{F}(t), 1\}$ and $\{u_{l+1}, v_{l+1}\} = \dots = \{u_d, v_d\} = \{0, \bar{F}(t)\}$. On a theoretical basis, this formula enables us to apply exchangeable exogenous shock models to the pricing of index CDS and CDOs as the distribution of L_t can be used to calculate the expected tranche loss $\mathbb{E}[L_{t,j}]$. From a practical perspective, however, the crucial difficulties in evaluating $\mathbb{P}(L_t = x)$ for large d are the binomial coefficient in (5.3) and cancellation effects in the alternating sum (5.4). For the standardized portfolios with $d = 125$ mentioned in Section 5.1, this coefficient can become huge. Together with the alternating sum in (5.4), the formula for the portfolio loss distribution potentially becomes numerically unstable and prevents the model from being applicable practically.

One way out of this quandary that has been pursued in the literature for other families of distribution functions is to not only require $(X_1, \dots, X_d)$ to be exchangeable, but even extendible (see Definition 6.4). The crucial consequence of this additional restriction is the validity of De Finetti's Theorem (see Theorem 2.2.13), linking the vector of default times to an infinite sequence $\{X_k\}_{k \in \mathbb{N}}$ of conditionally i.i.d. random variables. Conditioned on a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$, the random variables $\{X_k\}_{k \in \mathbb{N}}$ are i.i.d., which

is why this setup is often termed ***c**onditionally **i**ndependent and **i**dentically **d**istributed (CIID)* framework in the literature (for an overview on the subject, see e.g. Mai et al. (2012)). The most famous CIID representative is the one-factor Gaussian copula model, see Vasicek (1987); Li (2000). Extensions and alternative approaches aiming at improving the flexibility in modeling the dependence structure are given in Burtschell et al. (2009). As a major consequence of the CIID structure, well-known limit results for d tending towards infinity can be applied to the distribution function of the portfolio loss L_t. This methodology allows to approximate the law of L_t by its limiting distribution for an infinitely large portfolio and provides an evaluation formula for index CDS and CDOs.

Before elucidating the proceeding in more detail, the initial question for the applicability in our context is: When considering an exchangeable exogenous shock model for $(X_1, \ldots, X_d)$, how do we know that it is even extendible and how can we derive the sub-σ-algebra $\mathcal{G}$? In general, this is an open question that could be addressed in future research (see Chapter 7). However, we have already identified a subclass of extendible exogenous shock models via the additive–frailty models in Section 3.5. Defining the default times by

$$X_k := \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d, \tag{5.5}$$

for an additive subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with $\lim_{t\to\infty} \Lambda_t = \infty$ and independent unit exponentially distributed random variables $\{E_k\}_{k\in\mathbb{N}}$, extendibility is given by construction. For the special case of Λ being a Lévy subordinator, this model has already been analyzed in (Mai and Scherer, 2009c, Section 5). The random vector $(X_1, \ldots, X_d)$ can be extended to an infinite exchangeable sequence $\{X_k\}_{k\in\mathbb{N}}$ by defining $X_{d+1}, X_{d+2}, \ldots$ in the very same way as $X_1, \ldots, X_d$ for additional i.i.d. unit exponentially distributed random variables $E_{d+1}, E_{d+2}, \ldots$ In this case, the sub-σ-algebra $\mathcal{G}$ is precisely the σ-algebra generated by Λ as it holds that

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d|\Lambda) = \mathbb{P}(E_1 \leq \Lambda_{x_1}, \ldots, E_d \leq \Lambda_{x_d}|\Lambda) &= \prod_{k=1}^{d} \mathbb{P}(E_k \leq \Lambda_{x_k}|\Lambda) \\ &= \prod_{k=1}^{d} \left(1 - e^{-\Lambda_{x_k}}\right) = \prod_{k=1}^{d} \mathbb{P}(X_1 \leq x_k|\Lambda), \quad x_1, \ldots, x_d \in [0, \infty).\end{aligned}$$

When assuming the sequence $\{X_k\}_{k\in\mathbb{N}}$ to be defined by an additive–frailty model, the CIID structure provides a useful limit theorem known as *large homogeneous portfolio (LHP) approximation*. In our context, it reads as follows.

Proposition 5.2.1 (LHP approximation for additive–frailty models)
Let $\{X_k\}_{k\in\mathbb{N}}$ be given by the additive–frailty model in (5.5) and consider the portfolio loss process $\{L_t\}_{t\geq 0}$ in (5.1). Then it holds that

$$\mathbb{P}\Big(\lim_{d\to\infty}\sup_{t\geq 0}\big|\big(1-e^{-\Lambda_t}\big)-L_t\big|=0\Big)=1.$$

Proof
See the more general result in (Mai et al., 2012, Lemma 2.2). □

Thus, the portfolio loss L_t almost surely converges to the random distribution function $F_t = 1-\exp(-\Lambda_t)$ for $d\to\infty$. Referring to index CDS and CDOs with $d=125$, this means that for additive–frailty models, we can approximate the portfolio loss distribution at time t by the law of $1-\exp(-\Lambda_t)$. In other words, the bottom-up construction of $(X_1,\ldots,X_d)$ is replaced by a top-down specification of L_t. Yet, in this case, the portfolio loss process is not some artificial object, however provides a probabilistic insight into the underlying dependence structure of the default times.

From a practical point of view, while in general the portfolio loss process $\{L_t\}_{t\geq 0}$ is a sum of d dependent random variables, the CIID structure in additive–frailty models can be exploited to reduce complexity and constrain computations to the one-dimensional stochastic process $\Lambda=\{\Lambda_t\}_{t\geq 0}$. Another desirable advantage of this simplification compared to the portfolio loss distribution in arbitrary exchangeable exogenous shock models is related to the calibration of the model. Instead of having to specify the d shock distribution functions in an exchangeable exogenous shock model, solely the parameters of the stochastic process Λ have to be determined. All in all, representing the default times by an additive–frailty model, two challenges remain. First of all, the computation of the expected (tranche) loss $\mathbb{E}[L_t]$ and $\mathbb{E}[L_{t,j}]$ relevant for index CDS and CDO pricing has to be discussed. Secondly, deriving a consistent calibration of the parameters of Λ to both index CDS and CDOs is a necessary requirement for the model's practical suitability. Both aspects are addressed in the next two sections.

5.3 Laplace inversion formula for expected tranche loss

Consider the additive–frailty model in (5.5) for the default times, i.e.

$$X_k := \inf\{t\geq 0 : \Lambda_t \geq E_k\},\quad k=1,\ldots,d,$$

for an additive subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with $\lim_{t\to\infty}\Lambda_t = \infty$ and (independent of Λ) i.i.d. unit exponentially distributed random variables $E_1, \ldots, E_d$. The large homogeneous portfolio approximation (see Proposition 5.2.1) implies that the portfolio loss L_t converges towards $1-\exp(-\Lambda_t)$ for increasing portfolio size. Denoting by $\{\Psi_t\}_{t\geq 0}$ the family of Bernstein functions corresponding to Λ, one recognizes that

$$\mathbb{E}[L_t] \longrightarrow \mathbb{E}[1-e^{-\Lambda_t}] = 1 - e^{-\Psi_t(1)} \text{ for } d\to\infty. \tag{5.6}$$

Thus, if $\{\Psi_t\}_{t\geq 0}$ is known explicitly, the expected loss $\mathbb{E}[L_t]$ can be approximated in closed-form by $1-\exp\big(-\Psi_t(1)\big)$. For the expected tranche losses $\mathbb{E}[L_{t,j}]$ crucial for evaluating CDOs, the situation is more difficult. For $L_t \approx 1-\exp(-\Lambda_t)$, the expected tranche loss $\mathbb{E}[L_{t,j}]$ can be approximated (compare (5.2)) by

$$\begin{aligned}\mathbb{E}[L_{t,j}] &\approx \mathbb{E}\Big[u_j - l_j + \Big(l_j - (1-\mathrm{Rec})(1-e^{-\Lambda_t})\Big)^+ - \Big(u_j - (1-\mathrm{Rec})(1-e^{-\Lambda_t})\Big)^+\Big] \\ &= u_j - l_j - (1-\mathrm{Rec})\,\mathbb{E}\Big[\Big(e^{-\Lambda_t} - \big(1-\frac{u_j}{1-\mathrm{Rec}}\big)\Big)^+ - \Big(e^{-\Lambda_t} - \big(1-\frac{l_j}{1-\mathrm{Rec}}\big)\Big)^+\Big].\end{aligned} \tag{5.7}$$

Thus, $\mathbb{E}[L_{t,j}]$ is decomposed into two call options with the same fictitious underlying $\exp(-\Lambda_t)$, but different strikes $1-u_j/(1-\mathrm{Rec})$ respectively $1-l_j/(1-\mathrm{Rec})$. For $K>0$, defining

$$C_t^K(\xi) := \mathbb{E}\Big[\big(e^{\xi-\Lambda_t}-K\big)^+\Big] = K\,\mathbb{E}\Big[\big(e^{\xi-\log(K)-\Lambda_t}-1\big)^+\Big] = K\,C_t^1\big(\xi-\log(K)\big),$$

one recognizes that the problem reduces to the computation of call options with strike one. While there is no explicit solution for these call prices in general, it is possible to derive their Laplace transform.

Lemma 5.3.1 (Laplace transform of call options on $\exp(-\Lambda_t)$)
For the additive–frailty setup, it holds that

$$\mathcal{L}[C_t^1](z) = \frac{1}{z(z-1)}\,e^{-\Psi_t(z)}, \quad z\geq 0.$$

Proof
The proof is analogous to the proceeding in (Mai et al., 2014, Section 3.2) and is based on the general results in Raible (2000). □

Deriving the value of C_t^1 via its Laplace transform can be accomplished by means of *inverse Laplace transform algorithms*. Given a non-negative function f with well-defined

Laplace transform $\mathcal{L}[f](s)$ for $s > 0$, a commonly used approach is the Bromwich inversion formula (see (Widder, 1946, Theorem 7.3)), stating that

$$f(x) = \lim_{R\to\infty} \frac{1}{2\,\pi\,i} \int_{a-i\,R}^{a+i\,R} e^{x\,z}\,\mathcal{L}[f](z)\,\mathrm{d}z, \quad a, x > 0. \tag{5.8}$$

Conditions ensuring the validity of the formula are existence, bounded variation, and continuity of f. These conditions can be verified for a given Laplace transform for instance by applying (Sato, 1999, Proposition 28.1), which states that f is continuously differentiable if $\int_{\mathbb{R}} |\mathcal{L}[f](-i\,s)\,s|\mathrm{d}s < \infty$. It is well-known that (5.8) generally induces massive numerical problems due to the oscillating behavior of $\exp(x\,z)$ for varying imaginary part of the argument z. Therefore (see Weideman (2006) and the references therein), attempts have been made to change the integration path (5.8) in a suitable manner. One of the most famous *contour transformations* is the one in Talbot (1979). There, the original path $z = a + i\,R, R \in \mathbb{R}$, is altered such that the imaginary part of z is bounded while for the limiting cases $|R| \to \infty$, its real part tends towards minus infinity. As a result, oscillations of $\exp(x\,z)$ are reduced and dampened by the negative real part of z.

However, though being powerful in many applications, the validity of a given contour transformation can be hard to check. While Talbot (1979) provides technical conditions for his approach to be applicable, verifying them for a specific Laplace transform might be hard to accomplish. The main result of the present chapter is an alternative contour transformation in (5.8), applied to the extraction of C_t^1 from its Laplace transform $\mathcal{L}[C_t^1]$. More precisely, it refers to the inversion of $\mathcal{L}[C_t^1]$ in case of the Bernstein function Ψ_t in Lemma 5.3.1 being *complete*. Completeness of the Bernstein function means that in the Lévy–Khintchine representation of Ψ_t in (2.11), the Lévy measure ν possesses a completely monotone density in the sense of Definition 2.3.6. The subclass of positive, infinitely divisible laws π which is connected to complete Bernstein functions Ψ via $\mathcal{L}[\pi] = \exp(-\Psi)$ is termed *Bondesson class*. An extensive list of more than 100 complete Bernstein function is given in (Schilling et al., 2010, p. 218–277). Our Laplace inversion theorem reads as follows (by "$\Im(z)$", we denote the imaginary part of a complex number $z \in \mathbb{C}$).

Theorem 5.3.2 (Contour transformation for complete Bernstein functions) *If Ψ_t in Lemma 5.3.1 is a complete Bernstein function and if the Bromwich inversion formula* (5.8) *is valid for C_t^1, it holds for $x > 0$ that*

$$C_t^1(x) = \frac{M\,e^{x\,a}}{\pi} \int_0^1 \Im\Big(\frac{e^{-x\,M\,\log v\,(b\,i-a)}\,e^{-\Psi_t(a-M\,\log v\,(b\,i-a))}}{(a - M\,\log v\,(b\,i-a))\,(a - 1 - M\,\log v\,(b\,i-a))}\,(b\,i-a)\Big)\,\frac{\mathrm{d}v}{v},$$

with arbitrary parameters $a > 1, b > 0$ and $M > 2/(a\,x)$. This integral is a proper Riemann integral as one can show that the integrand vanishes for $v \searrow 0$.

Proof
See (Bernhart et al., 2015, Corollary 3.5). □

Though looking complicated at a first glance, the expression for C_t^1 in the above theorem solely comprises a single integral to be evaluated for the interval $[0, 1]$. The crucial steps underlying this contour transformation are as follows.

1. First, instead of using the Bromwich contour $\{z(u) = a + i\,u \,:\, 0 \leq u < \infty\}$, we consider the transformation in Kiesel and Lutz (2011), which is given by

$$\gamma(u) = a + u\,(b\,i - a), \quad 0 \leq u < \infty, \quad a, b > 0.$$

   The main computation in Bernhart et al. (2015) consists in showing admissibility of this contour transformation for the class of complete Bernstein functions.

2. Second, as the resulting integral is still indefinite and, thus, subject to truncation errors, another substitution is carried out by introducing the integration variable $v := \exp(-u/M), u > 0$ for constant $M > 0$. It can be shown that the resulting integrand vanishes for $v \searrow 0$ such that it can be extended continuously at zero, yielding a proper Riemannian integral.

Empirical analyses in Bernhart et al. (2015) point out that the choice $a = 1/x$, $b = 2\,a$, and $M = 3$ generally provides satisfactory results. The peculiarity of Theorem 5.3.2 is its mathematical rigorousness in the sense that the depicted contour transformation does not require further conditions and is indeed viable for complete Bernstein functions.

As a side remark, it should be mentioned that a more general result in this context is actually (Bernhart et al., 2015, Theorem 3.1), which derives a Laplace inversion formula for the density of distributions from the Bondesson class. Therefore, Theorem 5.3.2 is stated as a corollary rather than a theorem in the respective reference. The proof of the general finding is very technical and lengthy and requires some preliminary results from complex analysis. We could elaborate it in detail in the present section as it yields a powerful numerical procedure potentially being relevant for other areas in mathematical finance as well. However, with the present chapter aiming at generalizing the model in Mai and Scherer (2009c) to arbitrary additive–frailty models and combining it with the recent results in Bernhart et al. (2015) rather than profoundly introducing the two building blocks on their own, we omit a detailed discussion of the proof.

Summarizing, note that for both the approximation of index CDS and CDOs by means of the limit theorem $L_t \to 1 - \exp(-\Lambda_t)$ (for $d \to \infty$), it is very convenient to know the Laplace transform of the additive subordinator Λ in play (at fixed points in time). This observation explains the numerical suitability of the Lévy subordinator specification in Mai and Scherer (2009c). In this regard, due to the characterization via Bernstein functions, additive subordinators seem to be a natural generalization allowing to easily build new portfolio default models.

5.4 Calibration of additive–frailty models

By the preceding section, we are able to calculate expected (tranche) losses $\mathbb{E}[L_t]$ and $\mathbb{E}[L_{t,j}]$, and therefore to evaluate index CDS and CDOs for particular additive–frailty model specifications. However, being given market prices of these credit derivatives, a remaining challenge consists in calibrating the model, i.e. finding a specification consistent with observed quotes. In the additive–frailty construction of default times, the only degree of freedom is the additive subordinator Λ, such that fitting its parameters simultaneously to both index CDS and CDOs seems to be a delicate task. As a useful model extension, one can proceed analogously to Mai and Scherer (2009c) and provide the vector $(X_1, \ldots, X_d)$ of default times in (5.5) with a deterministic time-change, i.e. one considers

$$X_k := \inf\{t \geq 0 : \Lambda_{h(t)} \geq E_k\}, \quad k = 1, \ldots, d,$$

for a strictly increasing function $h : [0, \infty) \to [0, \infty)$ with $h(0) = 0$ and $\lim_{t\to\infty} h(t) = \infty$. The crucial consequence is that the function h provides an additional degree of freedom when fitting the model to market data, however, by Corollary 2.2.8 (respectively its counterpart for strictly increasing transformations), does not change the dependence structure of the default times. Put differently, h solely impacts the marginal distributions of $(X_1, \ldots, X_d)$, which are the input for computing $\mathbb{E}[L_t]$ and, thus, evaluating index CDS. Similarly to Equation (5.6), it holds that

$$\mathbb{E}[L_t] \longrightarrow \mathbb{E}[1 - e^{-\Lambda_{h(t)}}] = 1 - e^{-\Psi_{h(t)}(1)} \quad \text{for } d \to \infty.$$

Given the function $c : t \mapsto \Psi_t(1)$, one can define h by $h(t) := c^{-1}\big(-\log\big(1 - F(t)\big)\big)$ for an arbitrary strictly increasing distribution function F on $[0, \infty)$. As a consequence,

$$\mathbb{E}[L_t] \longrightarrow 1 - e^{-\Psi_{h(t)}(1)} = F(t) \quad \text{for } d \to \infty,$$

i.e. the expected loss $\mathbb{E}[L_t]$ is independent of Λ and is only impacted by the distribution function F. This feature facilitates the consistent model calibration to index CDS and CDOs and allows for a convenient two-step procedure:

1. Due to $\mathbb{E}[L_t] = F(t)$, specify the function F such that the model-induced index CDS value equals the observed market price.

2. Having fixed F, calibrate the additive subordinator Λ to observed CDO tranche market prices: For a given specification of Λ with corresponding family $\{\Psi_t\}_{t\geq 0}$ of complete Bernstein functions, use Theorem 5.3.2 (replacing t by $h(t)$ in the inversion integral) to calculate the expected tranche losses $\mathbb{E}[L_{t,j}]$ and to evaluate CDO tranches $j = 1, \ldots, J$.

A possible way to specify the marginal distribution function F in the first step is to orientate oneself towards univariate *intensity-based models*. When defining a single default time X_1 via

$$X_1 := \inf\{t \geq 0 : \int_0^t \lambda_s \mathrm{d}s \geq E_k\}$$

for i.i.d. unit exponentially distributed random variables $E_1, \ldots, E_d$ and a non-stochastic, constant default intensity $\lambda_s = \lambda > 0$, the distribution function of X_1 is given by $F(t) := \mathbb{P}(X_1 \leq t) = \mathbb{P}(E_k \leq \lambda\, t) = 1 - \exp(-\lambda\, t)$. Specifying the time-change $h(t) := c^{-1}\big(-\log\big(1 - F(t)\big)\big)$ as depicted above, the index CDS value essentially becomes a function of the default intensity λ. It can thus be determined (for instance via bisection) such that the model-implied index CDS value is equal to the observed one in the market. When specifying the additive subordinator Λ by calibration to market CDO quotes in the second step, the degrees of freedom (i.e. the parameters of Λ) are generally too sparse to perfectly replicate all CDO tranches $j = 1, \ldots, J$, simultaneously. Thus, an optimization routine, for instance a minimization of the mean-squared error between model and market prices among all tranches, is necessary. However, due to the velocity of the Laplace inversion procedure outlined in Theorem 5.3.2, the run time of such an optimization is a matter of few seconds for many model setups (compare the calibration of a similar model to time series in [Mai, Olivares, Schenk, Scherer (2014)]).

As a last point, we are going to illustrate the CDO pricing methodology for additive–frailty models by applying the approximation formulas to particular Sato–frailty copulas. To begin with, assume that we have specified Λ as an H-Sato subordinator $\Lambda = \{\Lambda_t\}_{t\geq 0}$ with Bernstein family $\{\Psi_t\}_{t\geq 0}, \Psi_t(x) = \Psi_1(t^H\, x), x, t \geq 0$, by the self-decomposable Bernstein functions in Section 2.4.5. Adding a drift component $\mu \geq 0$

to the Bernstein functions in order to gain an additional degree of freedom, we choose $\Psi_1 \in \{\Psi^{\mathrm{St+drift}}_{\alpha,\beta}, \Psi^{\mathrm{IG+drift}}_{\beta,\eta}, \Psi^{\mathrm{Ga+drift}}_{\beta,\eta}\}$, where for $x \geq 0$,

$$\begin{aligned}
\Psi^{\mathrm{St+drift}}_{\alpha,\beta}(x) &:= \mu\, x + \beta\, x^{\alpha}, \quad \alpha \in (0,1), \\
\Psi^{\mathrm{IG+drift}}_{\beta,\eta}(x) &:= \mu\, x + \beta\, (\sqrt{2x+\eta^2} - \eta), \quad \beta, \eta > 0, \\
\Psi^{\mathrm{Ga+drift}}_{\beta,\eta}(x) &:= \mu\, x + \beta\, \log(1 + \frac{x}{\eta}), \quad \beta, \eta > 0.
\end{aligned}$$

Under the LHP assumption in Proposition 5.2.1, the preceding section has shown that the expected portfolio loss satisfies

$$\mathbb{E}[L_t] = 1 - e^{-\Psi_{h(t)}(1)} = 1 - e^{-\Psi_1(h(t)^H)}, \quad t \geq 0.$$

Consequently, when setting $h(t) := \big(\Psi_1^{(-1)}(-\log(1-F(t))\big)^{(1/H)}$ for an arbitrary strictly increasing distribution function F on $[0,\infty)$, it holds that $\mathbb{E}[L_t] = F(t)$, such that the expected portfolio loss becomes independent of both the self-similarity index H and the specification of Ψ_1. Furthermore, looking at the Laplace transform of the call options in Lemma 5.3.1, one recognizes that

$$\Psi_{h(t)}(z) = \Psi_1(h(t)^H\, z) = \Psi_1\Big(\Psi_1^{(-1)}\big(-\log(1-F(t))\, z\Big).$$

As already pointed out in Section 4.3 for general Sato–frailty copulas, this expression is invariant w.r.t. stretching/compression of the Bernstein function Ψ_1. This implies that the expected tranche loss $\mathbb{E}[L_{t,j}]$ (which by Theorem 5.3.2 is essentially a function of $\Psi_{h(t)}$) and, as a consequence, the value of a CDO tranche, are independent of manipulating Ψ_1 by $\tilde{\Psi}_1(x) := \Psi_1(c\,x)$, $c > 0$, as well. Because of this, in the examples given above, we can always assume w.l.o.g. that the drift μ of the Bernstein function Ψ is fixed to one, i.e. $\mu = 1$. All in all, the specification of h allows us to fix some parameters in the Sato–frailty portfolio default model. In this special case, the described two-step procedure reads as follows.

1. Specify the function F such that the model-induced index CDS spread equals the observed market price.

2. W.l.o.g., set $H = 1$ and $\mu = 1$ for the H-Sato subordinator corresponding to $\Psi_1 \in \{\Psi^{\mathrm{St+drift}}_{\alpha,\beta}, \Psi^{\mathrm{IG+drift}}_{\beta,\eta}, \Psi^{\mathrm{Ga+drift}}_{\beta,\eta}\}$. For a given parameterization of the Bernstein function Ψ_1 in concern, i.e. fixed parameters $\alpha \in (0,1), \beta > 0$, respectively $\eta > 0$, the expected tranche loss $\mathbb{E}[L_{t,j}]$ can be computed via Theorem 5.3.2. Determine the parameters such that the values of the CDO tranches $j = 1, \ldots, J$ are consistent with the corresponding market prices.

6 Model uncertainty for sums of random variables

This chapter arises from the article [Mai, Schenk, Scherer (2015a)] and provides a second practical application of exogenous shock models that strongly differs from the multivariate default time context in the preceding paragraph. Considering the (possibly daily) log-returns $\boldsymbol{R} := \{R_{t_k}\}_{k=1,\ldots,d}$ of a portfolio or index, the aim is to incorporate uncertainty into the stochastic model for the return distribution of the aggregated log-return $R_{t_1} + \ldots + R_{t_d}$ in a universal way. The universality of the presented approach consists in its general applicability to arbitrary return specifications, demonstrating that it constitutes a fundamentally new mindset towards the analysis of model risk. Section 6.1 illustrates the idea, which is based on a well-known decomposition of the (dependent) random vector $\boldsymbol{R}$ into a deterministic function and an i.i.d. random source, and which consists in manipulating the latter. In Section 6.2, several conditions for a reasonable distortion are derived and stated in an axiomatic manner. Relying on the Dirichlet copula introduced in Section 3.5, Section 6.3 introduces a manipulated random source that satisfies the required properties. A case study to illustrate the approach and incorporate model uncertainty into different popular return specifications (e.g. i.i.d. normally distributed log-returns $\{R_{t_k}\}_{k=1,\ldots,d}$ as the most basic example) is presented in Section 6.4.

6.1 Risk management in the financial sector

In the financial industry, it is a risk manager's daily routine to quantify the market risk of a given portfolio w.r.t. a certain holding period, e.g. via the Value-at-Risk (VaR) or by means of more elaborate risk measures[1]. This task is typically implemented by modeling granular log-returns $\boldsymbol{R} := \{R_{t_k}\}_{k=1,\ldots,d}$ for d short periods, e.g. for $d = 250$

[1] See Pérignon and Smith (2010) for a survey of VaR methods in the banking industry and Artzner et al. (1999) for risk measures beyond VaR.

trading days, and aggregating them to the log-return $R_{t_1} + \ldots + R_{t_d}$ of the required period, e.g. one year, which the risk measure is deduced from. There are several reasons to do so. On the one hand, the risk manager could have a more distinct intuition about the behavior of the portfolio in the short term and might already have an idea of which distribution to consider for the portfolio log-return. On the other hand, testing and quantifying this intuition by calibrating the distribution to market data demands for a broad data basis and, thus, becomes easier (respectively possible at all) when splitting the entire time horizon into granular time periods. Especially when using discretized time-continuous models, choosing the optimal time horizon for the model specification first and extracting the return distribution for the target holding period subsequently is a natural proceeding.

Another risk management context where a sum $R_{t_1} + \ldots + R_{t_d}$ of random variables has to be calculated is provided in Embrechts et al. (2013). In this reference, the authors address an application in operational risk and $R_{t_1}, \ldots, R_{t_d}$ represent the losses of different business lines or risk types w.r.t. a fixed time period. Regardless of the application in particular, as a matter of course, the possibilities to model the vector $\boldsymbol{R}$ are manifold and the statistical and economic verification of an approach's validity is cumbersome, if not impossible. Consequently, regulatory guidelines such as Basel III[2] or the supervisory guidance of the Federal Reserve Bank[3] acknowledge the relevance of model uncertainty. There already exists considerable work concerning this topic in risk analysis. Alexander and Sarabia (2012) translate VaR model risk into uncertainty of the respective quantile and introduce a distribution function for the latter. Bignozzi and Tsanakas (2014) analyze residual risk, which arises from deviations between a portfolio's true loss distribution and a distribution function extracted from a data set of (i.i.d.) loss observations. In Embrechts et al. (2013), VaR bounds for sums of risks are derived in a Fréchet-type problem, where marginal laws are given, but the dependence structure is uncertain. The impact of small distortions of the data set on the estimation of risk measures is studied in Cont et al. (2010). Other papers are concerned with questions in derivative pricing, see e.g. Avellaneda et al. (1995); Schoutens et al. (2004); Cont (2006); Gupta et al. (2010); Bannör and Scherer (2013).

[2]We cite from (Basel Committee on Banking Supervision, 2010, p. 2): "*The reforms raise both the quality and quantity of the regulatory capital base and enhance the risk coverage of the capital framework. They are underpinned by a leverage ratio that serves as a backstop to the risk-based capital measures, is intended to constrain excess leverage in the banking system and* ***provide an extra layer of protection against model risk*** *and measurement error.*"

[3]Compare Federal Reserve System (2011): "*Model risk should be managed like other types of risk*".

From a structural point of view, many of the mentioned references pursue a similar line of reasoning in terms of the Markov regression representation stated in Rüschendorf and de Valk (1993): According to this, any d-variate random vector $\boldsymbol{R} = \{R_{t_k}\}_{k=1,\dots,d}$ can be depicted as

$$R_{t_k} = f_k(R_{t_1}, \dots, R_{t_{k-1}}, U_k), \quad k = 1, \dots, d, \tag{6.1}$$

for a sequence $\{U_k\}_{k=1,\dots,d}$ of i.i.d. random variables with uniform distribution on the unit interval and a set of measurable functions $f = (f_1, \dots, f_d)$, $f_k : \mathbb{R}^{k-1} \times [0,1] \to \mathbb{R}$, $k = 1, \dots, d$[4]. Representation (6.1) separates any model for the random vector $\boldsymbol{R}$ into a purely deterministic component f and a purely stochastic input $\boldsymbol{U} = \{U_k\}_{k=1,\dots,d}$. The function f itself is found by economic and statistical reasoning and consists of both model parameters and structural assumptions, such as dependence properties between subsequent returns. As will be pointed out in the sequel, for most models in practice, the k-th component of f is known explicitly. The recursive definition of R_{t_k} has a natural time-dynamic interpretation and U_k can be seen as an exogenous risk factor entering the model in time step k. Figure 6.1 illustrates the return generation setup.

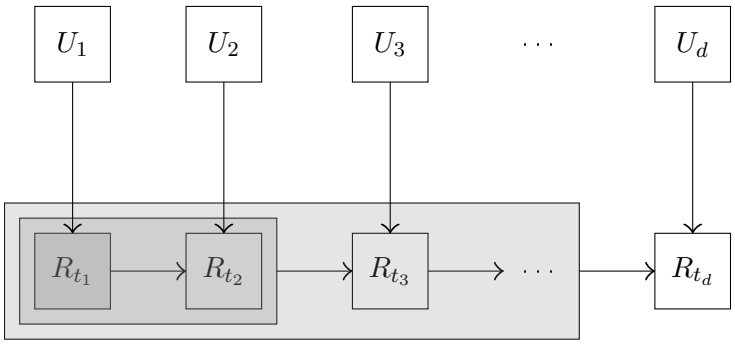

Figure 6.1 Universal stochastic representation of returns $\{R_{t_k}\}_{k=1,\dots,d}$. On the one hand, R_{t_k} can depend on earlier returns $R_{t_1}, \dots, R_{t_{k-1}}$ to account for serial dependence. On the other hand, an exogenous, purely random input U_k represents the uncertainty in time step k – one might think of the "random state of nature" in period k. The stochastic model behind is described by the choice of the function f, resp. f_k, for $k = 1, \dots, d$.

With regards to the universal representation (6.1), the essential question in many of the works concerning model uncertainty is: How to account for model risk w.r.t. a specific model f? The characteristics of the analysis can be manifold, involving both

[4] The representation (6.1) is by no means unique, of course.

identification risk (determination of the correct model), parameter risk (proper choice of parameters), and calibration risk (uncertainty due to noise in market data and inaccuracy of calibration algorithms). However, all these branches already presuppose the use of some structure on f. Furthermore, any model for the uncertainty w.r.t. the specification of f is itself prone to model risk.

In contrast to that, this chapter aims at deriving a universal, model-independent, and easy-to-implement framework for the assessment of a model uncertainty charge to any stochastic model in concern, setting it apart from previous research. We are not manipulating the model function f, but rather evaluate its robustness with respect to a distortion of the distribution function of the i.i.d. random variables $\{U_k\}_{k=1,\dots,d}$. Abstractly speaking, we consider the distorted model $\tilde{\boldsymbol{R}} := (\tilde{R}_1, \dots, \tilde{R}_d)$, $\tilde{R}_k = f_k(\tilde{R}_1, \dots, \tilde{R}_{k-1}, \tilde{U}_k)$, for a random vector $\tilde{\boldsymbol{U}} = \{\tilde{U}_k\}_{k=1,\dots,d}$ that follows a "distorted" uniform distribution on $[0,1]^d$. That is, we assume uncertainty about the i.i.d. uniform distribution of the random number generator. One key advantage of this method is its general applicability, i.e. it may be added to virtually any existing model (and implementation thereof) in a straightforward way, as will be demonstrated in Section 6.4. This is very convenient in practice, since our methodology can be "pulled over" an existing IT infrastructure without major effort. The induced distortion of the stochastic root of the model produces a distortion of the ultimate return variable $\tilde{R}_{t_1} + \dots + \tilde{R}_{t_d}$ in concern without the need of a modeling effort that is specific to f. This procedure intuitively prevents systematic model mis-specifications when modeling the uncertainty itself, because no model-f-dependent economic reasoning enters. Admittedly, this great level of universality comes at the cost of efficiency for model-specific questions. For instance, if it is commonly accepted that f lies in a certain model family, then the present approach might potentially overestimate the degree of model uncertainty. However, this only corresponds to the unavoidable second side of the coin, whose first (and in our view quite appealing) side is the total abstraction from model-dependent, economic reasoning when distorting the given model. An illustration of this procedure's philosophy is given in Figure 6.2.

6.2 Distortion of the stochastic root

6.2.1 Popular return specifications

In order to implement and distort the framework in (6.1), we consider models for which the function f can be given explicitly. Some popular models for $\boldsymbol{R}$ which are used in

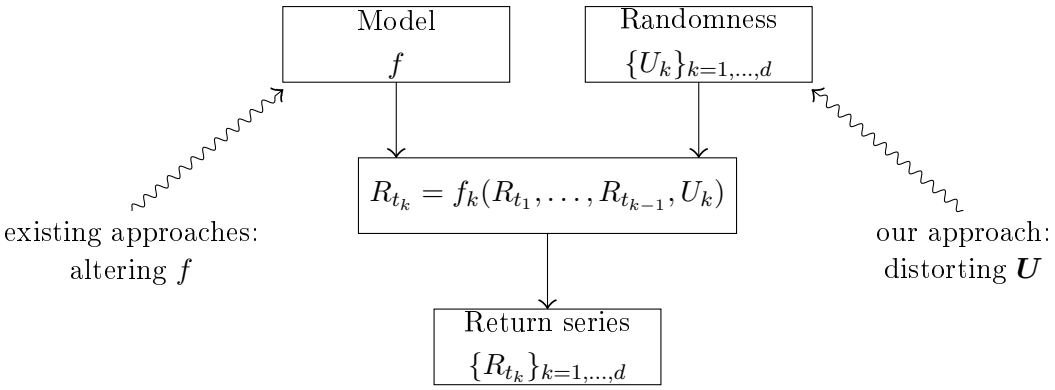

Figure 6.2 Idea of our distortion approach: While common model uncertainty frameworks specifically address the structure of f, we focus on distorting the purely stochastic input represented by the exogenous risk factors $\{U_k\}_{k=1,\ldots,d}$.

risk management for the analysis of $R_{t_1} + \ldots + R_{t_d}$ and which induce a known model function f are listed below. More information and the respective references are provided in Alexander and Sheedy (2008).

(I) $\{R_{t_k}\}_{k=1,\ldots,d}$ *are i.i.d. with univariate distribution function* $G(.)$*:*

This means that the k-th component of f is simply given by $R_{t_k} = f_k(U_k) = G^{-1}(U_k)$, where G^{-1} is the generalized inverse of G. Popular choices for G in practice are standard parametric models like the normal distribution, Student's t distribution, or mixtures of normal distributions. A discrete distribution is also a popular choice for G, because it arises in the case of a historical VaR simulation, which is still among the dominantly used VaR methods applied in practice, according to Pérignon and Smith (2010).

(II) $\{R_{t_k}\}_{k=1,\ldots,d}$ *follow a GARCH model:*

This setup is developed and applied, e.g., in Bollerslev (1986); Alexander and Sheedy (2008). The k-th (mean-adjusted) return has the form $R_{t_k} = \sigma_k Z_k$ for an i.i.d. sequence $\{Z_k\}_{k\in\mathbb{Z}}$ with expected value zero and variance equal to one (called "innovations") and a scaling factor σ_k that depends on both the former realized returns $R_{t_{k-1}}, R_{t_{k-2}}, \ldots$ and past scaling factors $\sigma_{k-1}, \sigma_{k-2}, \ldots$ In this case, just as in (I) above, the i.i.d. innovations Z_k are retrieved by applying the generalized inverse G^{-1} to the exogenous risk factors $U_k, k = 1, \ldots, d$.

It is further natural – and possible for any $\boldsymbol{R}$ – to specify all f_k such that they are increasing (respectively decreasing) w.r.t. their k-th component, i.e. the random input U_k. In this regard, U_k represents the random source of global economic prosperity / depression, which is translated into the behavior of the portfolio in time step k by the model function f_k.

6.2.2 Consistency conditions for the distortion

Our aim is to derive a random number generator that returns a distorted series $\tilde{\boldsymbol{U}} = \{\tilde{U}_k\}_{k=1,\dots,d}$ of random variables to be used instead of the i.i.d. uniform sequence $\boldsymbol{U} = \{U_k\}_{k=1,\dots,d}$ in (6.1) for arbitrary $d \in \mathbb{N}$. However, for a reasonable distortion, three crucial structural properties of the original random vector $\boldsymbol{U}$ have to be maintained.

1. Firstly, the joint distribution of $\{U_k\}_{k=1,\dots,d}$ is invariant w.r.t. permutation of its components, emphasizing the fact that $\boldsymbol{U}$ accounts for a time-homogeneous source of risk separated from the model function f. For an i.i.d. sequence, this condition is trivially satisfied. For the distorted series $\tilde{\boldsymbol{U}} = \{\tilde{U}_k\}_{k=1,\dots,d}$, we have to make sure that the property

$$(\tilde{U}_1, \dots, \tilde{U}_d) \overset{d}{=} (\tilde{U}_{\sigma(1)}, \dots, \tilde{U}_{\sigma(d)}) \quad \text{for any permutation } \sigma \text{ on } \{1, \dots, d\}, \tag{6.2}$$

   i.e. exchangeability of $\tilde{\boldsymbol{U}}$ (see Definition 2.2.11), is still satisfied.

2. Secondly, the distribution of $\{U_k\}_{k=1,\dots,d}$ is also invariant w.r.t. point reflection of all its components at $u = 1/2$, i.e. replacing all U_k by $1 - U_k$ has no influence on the vector's law. This property mirrors a symmetry feature resulting from the marginal uniform distributions of the i.i.d. random source $\boldsymbol{U}$. On a univariate level, it implies that deviations of the exogenous risk factor U_k from its mean value occur in a symmetric manner, i.e. there is no bias towards large or small values, which might be interpreted as good or bad economic conditions. On a multivariate level, it is important that a path $(U_1, \dots, U_d)$ is as likely as the path $(1-U_1, \dots, 1-U_d)$, meaning that there is no inherent bias towards accumulations of either "good" or "bad" outcomes. To preserve this symmetry, we seek for a distorted sequence $\{\tilde{U}_k\}_{k=1,\dots,d}$ that satisfies

$$(\tilde{U}_1, \dots, \tilde{U}_d) \overset{d}{=} (1 - \tilde{U}_1, \dots, 1 - \tilde{U}_d). \tag{6.3}$$

   Put another way, we require $\{\tilde{U}_k\}_{k=1,\dots,d}$ to be radially symmetric about $(1/2,\dots,1/2)$ (see Definition 2.2.15).

3. Last but not least, the random source $\boldsymbol{U} = \{U_k\}_{k=1,\dots,d}$ not only generates the vector $\{R_{t_k}\}_{k=1,\dots,d}$ for some fixed d, but allows for a consistent extension to higher dimensions. Put another way, $(U_1, \dots, U_d)$ can be viewed as the initial part of an infinite sequence $\{U_k\}_{k\in\mathbb{N}}$, and the dimension d merely represents a model time horizon at which we stop the random number generator from spitting out further risk factors. This time horizon should be arbitrarily modifiable, reflecting the economic interpretation of time as a continuum. Thus, we demand that the Markov representation (6.1) can be extended to $d+1, d+2, \dots$, while maintaining the properties in (6.2) and (6.3). For an i.i.d. sequence satisfying these conditions, a consistent extension is straightforward as the series $\{U_k\}_{k=1,\dots,d}$ can simply be complemented by drawing further i.i.d. random variables $U_{d+1}, U_{d+2}, \dots$ Nevertheless, for our distortion, we have to guarantee that it is possible to construct an infinite sequence $\{\tilde{U}_k\}_{k\in\mathbb{N}}$ such that

$$\{\tilde{U}_k\}_{k\in\{1,\dots,d\}} \quad \text{satisfies (6.2) and (6.3) for all } d \geq 2. \tag{6.4}$$

According to Definition 2.2.12, a random vector that satisfies the condition in (6.4) is termed extendible (to an infinite exchangeable sequence). As pointed out in De Finetti's Theorem (see Theorem 2.2.13), the crucial consequence for an extendible random vector $\{\tilde{U}_k\}_{k=1,\dots,d}$ is the existence of a random distribution function $F = \{F_t\}_{t\in\mathbb{R}}$ on $\mathbb{R}$ such that, conditioned on F, the random variables $\{\tilde{U}_k\}_{k=1,\dots,d}$ are i.i.d. with respective distribution function F. This implies that the distorted random source is embedded in a CIID framework, which has already been considered in Chapter 5 in a completely different context. For a sequence of i.i.d. random variables $\{V_k\}_{k\in\mathbb{N}}$ with uniform distribution on the unit interval $[0,1]$ that are independent of F, $\{\tilde{U}_k\}_{k\in\mathbb{N}}$ can be defined as

$$\begin{aligned}
\tilde{U}_k :&= \inf\{t \in \mathbb{R} : F_t \geq V_k\} \\
&= \inf\{t \in \mathbb{R} : \underbrace{-\log(1-F_t)}_{:=M_t} \geq \underbrace{-\log(1-V_k)}_{:=E_k}\} \\
&= \inf\{t \in \mathbb{R} : M_t \geq E_k\}, \quad k \in \mathbb{N}.
\end{aligned} \tag{6.5}$$

With $M = \{M_t\}_{t\in\mathbb{R}}$ being an increasing stochastic process and E_k denoting i.i.d. unit exponentially distributed random variables[5] independent of M, this setup already bears resemblance to the additive–frailty construction in (3.22) (and, thus, to exchangeable exogenous shock models). However, M does not necessarily have to exhibit independent increments.

[5] It can easily be observed that $\mathbb{P}(E_k \leq x) = \mathbb{P}(V_k \leq 1 - \exp(-x)) = 1 - \exp(-x)$ for $x \geq 0$.

6.2.3 Axioms for a suitable random distribution function

By the reasoning in the previous section, the discussion about a reasonable stochastic model for the distorted random input $\tilde{\boldsymbol{U}}$ boils down to the specification of a suitable random distribution function F. As pointed out, by construction, Properties (6.2) and (6.4) are fulfilled for arbitrary F. However, in order to guarantee (6.3) and further consistency conditions, the following Axioms (A1)–(A4) for F are postulated.

(A1) The random distribution function F has mean value $G(t) := \mathbb{E}[F_t] = t$ for $t \in [0,1]$, i.e. the uniform law on $[0,1]$ is the "average realization".

→ In the original model, the vector $\{U_k\}_{k=1,\dots,d}$ is i.i.d. with uniform distribution on $[0,1]$. In the distorted setup, the random variables $\{\tilde{U}_k\}_{k=1,\dots,d}$ are i.i.d. conditioned on F with marginal distribution function F. Thus, the unconditional distribution function of each $\tilde{U}_k$ is $\mathbb{P}(\tilde{U}_k \leq t) = \mathbb{E}[F_t]$, $t \in \mathbb{R}$. Axiom (A1) ensures that the uniform distribution of the random components in the original model is maintained. In other words, the model for $\tilde{\boldsymbol{U}}$ may be viewed as a distortion of the original source $\boldsymbol{U}$, which itself constitutes the "average" realization of $\tilde{\boldsymbol{U}}$.

(A2) There is a single "uncertainty parameter" $\tilde{c} \in [0,1]$ which controls our confidence in the original model f. For $\tilde{c} = 0$, we have full confidence in our original model. As $\tilde{c}$ increases, this confidence shrinks.

→ Axiom (A2) allows us to conveniently control the strength of the distortion; the uncertainty parameter depends on one's risk aversion or may even be set by the regulator.

(A3) The model for F is such that the stochastic input $\{\tilde{U}_k\}_{k=1,\dots,d}$ has the same distribution as $\{1-\tilde{U}_k\}_{k=1,\dots,d}$ for arbitrary d, i.e. the distortion is radially symmetric in any dimension d.

→ Axiom (A3) guarantees consistency with property (6.3). It guarantees that the stochastic input is distorted in such a way that the dilution of the original model's economic meaning contains no systematic bias.

(A4) The support of F should be as large as possible within the space of possible distributions on $[0,1]$.

→ Axiom (A4) reflects a non-parametric nature within the present approach. The distortion from the original model is maximally chaotic and itself "un-modeled". It shall guarantee that we do not neglect potential sources for model mis-specification.

Besides, as a practical requirement, the distorted stochastic input $\{\tilde{U}_k\}_{k\in\mathbb{N}}$, respectively finite subvectors thereof, should be efficient and accurate to simulate from. Provided the original models f in concern can be simulated quickly from the original stochastic input $\boldsymbol{U}$, a rapid generation of samples for $\tilde{\boldsymbol{U}}$ is crucial for a fast simulation of the distorted model when replacing $\boldsymbol{U}$ by $\tilde{\boldsymbol{U}}$ in the general representation (6.1).

6.3 Model for the distorted random source

Having stated a number of crucial properties, the overall question is: Is there a random distribution function F such that Axioms (A1)–(A4) are satisfied? The following theorem shows that a Dirichlet process with parameters $(c, \mathrm{id}_{[0,1]})$ is a valid choice for F in the sense of the postulated conditions. Furthermore, as we know (see Section 2.4.4) that in this case, the stochastic process M appearing in (6.5) defines an additive subordinator, $\tilde{\boldsymbol{U}}$ is connected to an additive–frailty copula.

Theorem 6.3.1 (Distortion via Dirichlet copula)
Let $F = \{F_t\}_{t\in\mathbb{R}}$ be a Dirichlet process introduced in Section 2.4.4 with parameters $(c, \mathrm{id}_{[0,1]})$, $c > 0$, and define $\{\tilde{U}_k\}_{k\in\mathbb{N}}$ via (6.5). Then F satisfies Axioms (A1)–(A4). Moreover, the joint distribution function of $(\tilde{U}_1, \ldots, \tilde{U}_d)$ is given by the Dirichlet copula

$$C_c(u_1, \ldots, u_d) := \prod_{k=1}^{d} \frac{c\, u_{(k)} + k - 1}{c + k - 1}, \quad u_1, \ldots, u_d \in [0, 1].$$

Proof
We start by proving the second claim and compute the distribution function of $(\tilde{U}_1, \ldots, \tilde{U}_d)$. By Theorem 3.5.3, the survival copula of $(\tilde{U}_1, \ldots, \tilde{U}_d)$ is given by C_c. As the Dirichlet copula is radially symmetric (see Corollary 3.5.4), C_c also manifests the copula of the random vector. For a Dirichlet process Z with parameters (c, G), it holds that $\mathbb{E}[Z_t] = G(t)$ for all $t \in \mathbb{R}$. Thus, for the considered parameterization, it holds that $\mathbb{P}(\tilde{U}_k \le u) = \mathbb{E}[F_u] = u$ for $u \in [0, 1]$ and the Dirichlet copula C_c represents the distribution function of $(\tilde{U}_1, \ldots, \tilde{U}_d)$.

The proof of the second claim already implies that properties (A1) (univariate marginal distribution of the U_k) and (A3) (radial symmetry) are fulfilled. Thus, it remains to check properties (A2) and (A4).

(A2) Having fixed the unconditional distribution function $G = \mathrm{id}_{[0,1]}$ of the Dirichlet process in order to satisfy Axiom (A1), the only degree of freedom left in modeling $\tilde{\boldsymbol{U}}$ via

the Dirichlet prior assumption is the parameter $c > 0$, which accounts for the variance of F. It is straightforward to show that for $c \to \infty$, C_c converges to the independence copula, while for $c \searrow 0$, C_c equals the comonotonicity copula. Referring to the interpretation of $\tilde{U}_k$ as random economic well-being / distress entering the model f in time step k, the choice $c < \infty$ introduces an increased probability for a series of either low or high values for all $\tilde{U}_k$. In order to satisfy (A2), we are aiming at a reparameterization involving a single parameter $\tilde{c} \in [0,1]$ that accounts for the severity of the distortion. This can for instance be accomplished by a reparameterization $\tilde{c} = \tilde{c}(c) := \tau_{C_c}$ (respectively $\tilde{c} := \rho_{C_c}$) based on inverting Kendall's τ (respectively Spearman's ρ) in Lemma 6.3.3 below. In this case, $\tilde{c} \to 0$ means "no distortion from i.i.d." and $\tilde{c} = 1$ implies "maximum distortion". During the remainder of the paper, however, we stick with the parameter c for the sake of a consistent notation.

(A4) The following lemma, derived in (Ferguson, 1973, Proposition 3), guarantees that the support of F is a large family of distribution functions on $[0,1]$.

Lemma 6.3.2 (Support of a Dirichlet process $\mathbf{\Lambda}$)
Let $Z = \{Z_t\}_{t \in \mathbb{R}}$ be a Dirichlet process with parameters (c, G). If $\mathbb{Q}$ is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ which is absolutely continuous w.r.t. $\mathrm{d}G$, then, for arbitrary measurable sets $A_1, \ldots, A_m \in \mathcal{B}(\mathbb{R}), m \in \mathbb{N}$ and $\epsilon > 0$, it holds that[6]

$$\mathbb{P}\big(|\mathrm{d}Z(A_i) - \mathbb{Q}(A_i)| < \epsilon \text{ for } i = 1, \ldots, m\big) > 0. \tag{6.6}$$

Intuitively speaking, when setting $F = Z$ in (6.5) for a Dirichlet process Z with parameters $(c, \mathrm{id}_{[0,1]})$, Equation (6.6) states that an arbitrary absolutely continuous probability measure $\mathbb{Q}$ on $[0,1]$ can be approximated arbitrarily close by the set of potential realizations of $\mathrm{d}Z$. Regarding our intention to introduce model uncertainty in a "model-free" way, the large support of F in the Dirichlet construction implies that we are not putting serious structural restrictions on the distorted random vector $\tilde{\boldsymbol{U}}$ – which could be questioned in terms of model robustness – but rather allow for a broad range of model distortions. □

Reparameterizing c via the concordance measures mentioned in the proof of (A2) in Theorem 6.3.1 is substantiated by the following finding.

[6] In Equation (6.6), $\mathrm{d}Z$ denotes the random probability measure induced by Z.

Lemma 6.3.3 (Kendall's tau and Spearman's rho of C_c)
Kendall's tau and Spearman's rho of the bivariate copula $C_c(u_1, u_2) = u_{[1]}\, g_2(u_{[2]})$, where $g_2(u) = (c\,u + 1)/(c + 1)$, are given by

$$\tau_{C_c} = \frac{2\,c + 3}{3\,(c+1)^2}, \qquad \rho_{C_c} = \frac{1}{c+1}.$$

Proof
The results follow from the general formulas stated in Section 3.2. For Kendall's tau, we have

$$\begin{aligned}
\tau_{C_c} &= 4 \int_0^1 u\, g_2(u)^2 \,\mathrm{d}u - 1 = 4 \int_0^1 u \left(\frac{c\,u+1}{c+1}\right)^2 \mathrm{d}u - 1 \\
&= \frac{4}{(c+1)^2} \int_0^1 (c^2\,u^3 + 2\,c\,u + u)\,\mathrm{d}u - 1 \\
&= \frac{4}{(c+1)^2} \left(\frac{1}{4}\,c^2 + \frac{2\,c}{3} + \frac{1}{2}\right) - 1 \\
&= \frac{3\,c^2 + 8\,c + 6 - 3\,(c+1)^2}{3\,(c+1)^2} = \frac{2\,c+3}{3\,(c+1)^2}.
\end{aligned}$$

Spearman's rho is given by

$$\begin{aligned}
\rho_{C_c} &= 12 \int_0^1 u^2\, g_2(u)\,\mathrm{d}u - 3 = 12 \int_0^1 u^2\, \frac{c\,u+1}{c+1}\,\mathrm{d}u - 3 \\
&= \frac{12}{c+1} \int_0^1 (c\,u^3 + u^2)\,\mathrm{d}u - 3 = \frac{12}{c+1} \left(\frac{1}{4}\,c + \frac{1}{3}\right) - 3 \\
&= \frac{3\,c + 4 - 3\,c - 3}{c+1} = \frac{1}{c+1}.
\end{aligned}$$

□

Furthermore, by radial symmetry, the upper and lower tail-dependence coefficients λ_U and λ_L of C_c coincide. Using the general formula in Section 3.2, they are given by

$$\lambda_L = \lim_{u \searrow 0} g_2(u) = \frac{1}{c+1} = \lambda_U = 1 - g_2'(1-).$$

In addition to the fulfilled axioms, a simulation algorithm for the source of randomness $\tilde{\boldsymbol{U}}$, i.e. the copula C_c in arbitrary dimension d, is readily available from the key characteristics of the Dirichlet process. Generally speaking, (Ferguson, 1973, Theorem 1) implies that the random vector $\{\tilde{U}_k\}_{k=1,\ldots,d}$ in (6.5) with a Dirichlet process $F = Z$ with parameters (c, G) may be simulated iteratively as follows:

(1) Simulate $\tilde{U}_1 \sim G$.

(2) For $k = 2, \ldots, d$ perform the following steps to simulate $\tilde{U}_k$:

(a) Simulate a discrete random variable N with distribution given by

$$\mathbb{P}(N=i)=\frac{1}{c+k-1},\quad i=1,\dots,k-1,\quad \mathbb{P}(N=k)=\frac{c}{c+k-1}.$$

(b) If N equals k, simulate $\tilde{U}_k\sim G$, otherwise set $\tilde{U}_k:=\tilde{U}_N$.

Applying this algorithm with $G=\mathrm{id}_{[0,1]}$ yields a sample from C_c. To perform Step (2) for a fixed k, one can (i) divide the interval $[0,1]$ into k non-overlapping subintervals

$$\Big[0,\frac{1}{c+k-1}\Big],\Big[\frac{1}{c+k-1},\frac{2}{c+k-1}\Big],\dots,\Big[\frac{k-2}{c+k-1},\frac{k-1}{c+k-1}\Big],\Big[\frac{k-1}{c+k-1},1\Big],$$

(ii) draw a uniformly distributed random variable U on $[0,1]$, (iii) determine whether U falls into the first $k-1$ intervals or the last one, and (iv) either set $\tilde{U}_k$ equal to one of the previous components $\tilde{U}_1,\dots,\tilde{U}_{k-1}$, or set $\tilde{U}_k=[(c+k-1)/c]\,[U-(k-1)/(c+k-1)]$ to generate an independent uniformly distributed sample on $[0,1]$. Thus, sampling the random vector $\tilde{\boldsymbol{U}}=\{\tilde{U}_k\}_{k=1,\dots,d}$ with (survival) copula C_c can be accomplished by a few lines of code in any programming language. The algorithm is extremely efficient[7], the run time is of order $\mathcal{O}(d\log(d))$ when the involved discrete random variable N is simulated via bisection in $\mathcal{O}(\log(k))$.

For the case study in Section 6.4, we have implemented this algorithm in Matlab. Aiming at reproductability and applicability of the presented method, we present the implementation for generating n samples of the random vector $\{\tilde{U}_k\}_{k=1,\dots,d}$ with (survival) copula C_c in detail.

Algorithm 6.3.4 (Generating n samples of $\{\tilde{U}_k\}_{k=1,\dots,d}\sim C_c$)

```
function Utilde = rDirichletCopula(n,d,c)
    Utilde = zeros(n,d);
    Utilde(:,1)=rand(n,1);
    for k=2:d
        U=rand(n,1);
        cum_p = [0,(1:(k-1))/(c+k-1),1];
        [~,idv] = histc(U,cum_p);
        index=(1:n)'+(idv-1)*n;
        temp=[Utilde(:,1:(k-1)),(c+k-1)/c*(U-(k-1)/(c+k-1))];
```

[7] We have run the routine on a computer with 4 GB RAM and Intel Core 2 Quad CPU 2.4 GHz processor, however not using parallelization. The implementation of Algorithm 6.3.4 below roughly took 0.1 seconds to generate 100 000 samples of the 5-dimensional random vector $(\tilde{U}_1,\dots,\tilde{U}_5)$ for arbitrary values of c.

```
10            Utilde(:,k) = temp(index);
11        end
12    end
```

6.3.1 Alternatives to the Dirichlet process

In the beginning of Section 6.2.3, we have argued that the distorted random source $\{\tilde{U}_k\}_{k\in\mathbb{N}}$ should be constructed via (6.5) with a random distribution function F that, by Axiom (A1), satisfies $\mathbb{E}[F_t] = t$ for $t \in [0,1]$. We proposed the Dirichlet process from non-parametric Bayesian statistics as a natural candidate, and have shown that it satisfies all desirable Axioms (A1)–(A4) postulated in Section 6.2.3. The present section's intention is to outline that alternatives to the Dirichlet process are actually not easy to find, so our choice was rather natural and by no means arbitrary.

By Axiom (A1), one is naturally restricted to copula models for $(\tilde{U}_1, \dots, \tilde{U}_d)$. In order to guarantee the radial symmetry Axiom (A3), one is restricted to radially symmetric copulas. In this regard, one choice for F that might come to one's mind originates from the one-factor Gaussian copula (or elliptical generalizations thereof), which can be generated via (6.5) by defining

$$F_t := \Phi\Big(\frac{\Phi^{-1}(t) - \sqrt{c}\,M}{\sqrt{1-c}}\Big), \quad t \in [0,1], \tag{6.7}$$

where Φ denotes the standard normal distribution function, M a standard normally distributed random variable, and $c \in [0,1)$ the correlation parameter. By symmetry of the normal law, the extendible Gaussian copula satisfies (A3). Being parameterized by a single parameter $c \in [0,1)$, it additionally allows for a natural interpretation and hence satisfies (A2). Simulation is also straightforward and efficient. However, the crucial disadvantage with regards to our intended use is the violation of the support condition in (A4). Vividly speaking, the paths of $\{F_t\}_{t\in[0,1]}$ are generated by the realizations of the single random variable M. Once M is known, i.e. $M = m$ for some $m \in \mathbb{R}$, F is a deterministic distribution function on $[0,1]$ given by

$$\mathbb{P}(\tilde{U}_k \le t | M = m) = \Phi\Big(\frac{\Phi^{-1}(t) - \sqrt{c}\,m}{\sqrt{1-c}}\Big), \quad t \in [0,1].$$

It is evident that the set of distribution functions generated by the possible values for m is rather special and "small". As visualized in Figure 6.3, the paths in the Gaussian case follow a predefined structure guided by the realizations of M and are less "chaotic" than

the Dirichlet-type distortion. This constraint of the Gaussian copula heavily contradicts the spirit of our model-independent distortion attempt and can itself be seen as a quite restrictive structural assumption. The same arguments apply to the more general class of elliptical copulas (see McNeil et al. (2005) for general elliptical distributions), which, to the best of our knowledge, represent the only known well-understood family of distribution functions for $\tilde{\boldsymbol{U}}$ that satisfies the *radial symmetry* property in (A3) in arbitrary dimension d.

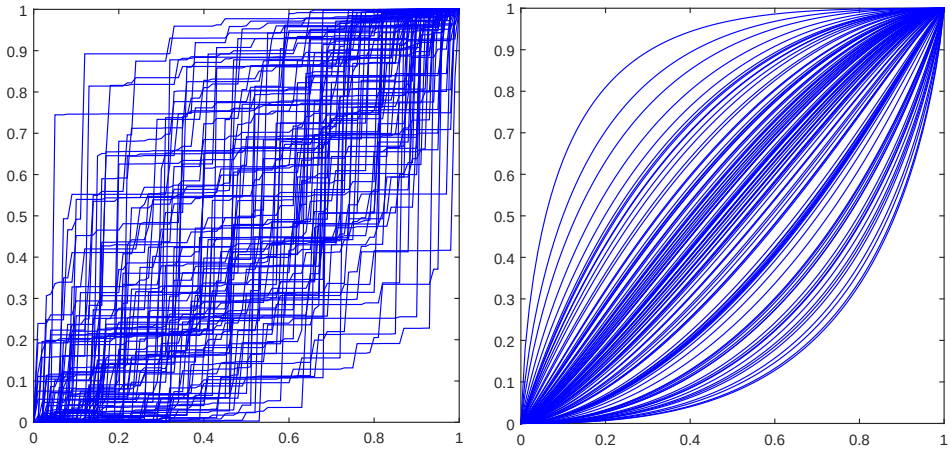

Figure 6.3 100 paths of Dirichlet process with parameters $(c, \mathrm{id}_{[0,1]})$ (left) and the Gaussian path $\{F_t\}_{t\in[0,1]}$ in (6.7) (right). In both the Dirichlet and Gaussian case, the parameter c has been set such that Spearman's rho of the random vector $(\tilde{U}_1, \tilde{U}_2)$ resulting from (6.5) equals 0.3.

Another strategy to finding alternatives to the Dirichlet prior satisfying (A1)–(A4) is the consideration of a more general class of processes for F. In Section 2.4.4, it has been mentioned that any Dirichlet process Z can be constructed by $Z = 1 - \exp(-\Lambda)$ for an increasing, right-continuous process $\Lambda = \{\Lambda_t\}_{t\in\mathbb{R}}$ with independent increments subject to certain constraints on the distribution of the increments. Furthermore, it has been pointed out that if one neglects these constraints, the corresponding process $Z = 1 - \exp(-\Lambda)$ is called neutral-to-the-right prior. When setting $F = Z$ and (in order to fulfill (A1)) assuming that $u = \mathbb{E}[F_u]$, Proposition 3.5.1 implies that the survival copula of $\{\tilde{U}_k\}_{k=1,\dots,d}$ from (6.5) is an additive–frailty copula with structural form

$$C(u_1, \dots, u_d) = \prod_{k=1}^{d} g_k(u_{(k)}), \quad u_1, \dots, u_d \in [0, 1],$$

where the functions g_k can be given explicitly in terms of the Laplace exponents of

Λ. However, even within the larger family of neutral-to-the-right priors, Theorem 3.4.2 shows that for differentiable functions g_k, the copula C is radially symmetric (i.e. satisfies (A3)) if and only if it is equal to the Dirichlet copula. Summing up, when constructing the distorted random vector $\tilde{\boldsymbol{U}}$ via (6.5), to the best of our knowledge, the only valid choice for satisfying Axioms (A1)–(A4) is a Dirichlet process with parameters $(c, \mathrm{id}_{[0,1]})$.

6.4 Empirical case study on model robustness

Though being easily applicable to a broad spectrum of risk models f, one might argue that the method of distorting the source of randomness is ad hoc and not backed by economic reasoning. Yet, we advocate the consideration of this approach for a number of reasons: First of all, any model-f-specific modification of f is itself a presumption subjective to critical debates. Secondly, though having his or her own idea of how to model and stress returns, a company's risk manager – as argued in Alexander and Sheedy (2008) – might be restricted in flexibility, for instance by the existing IT infrastructure. Last but not least, our analysis shows that the distortion parameter c has a universal influence on a variety of popular models for f in the sense that it consistently affects a given model throughout time.

6.4.1 Technical setup and data used

To make our case study reproducible, we assume that our portfolio returns follow the S&P 500 index, for which historical data is easily available, and that we are interested in deriving VaR forecasts for a given confidence level and forecast time horizon. A typical procedure for extracting these estimates is as follows:

(1) Consider a set $\boldsymbol{r} := \{r_{t_k}\}_{k=1,\dots,n}$ of empirical return realizations (training sample).

(2) Fit a model f for the future returns $R_{t_{n+1}}, R_{t_{n+2}}, \dots$ to $\boldsymbol{r}$.

(3) Extract the VaR estimate for the d-step return $R_{t_{n+1}} + \dots + R_{t_{n+d}}$.

Different variations and implementations of procedures (1)–(3) can be found in practice. For instance, Step (2) could be modified by expert opinion or risk-neutral calibration, and Step (3) might be evaluated by either closed-form formulas or simulation, depending on the model f in concern. In the following, however, we focus on the most basic approach described above to illustrate the distortion effect in the most transparent way. Mostly

due to data availability, the training sample is chosen to have daily frequency and to consist of at least $n = 250$ trading days of observations, i.e. about one year of data.

The present section indicates how to account for model uncertainty in the Dirichlet framework using the example of VaR calculation. The core idea is very simple: Instead of computing the VaR for $R_{t_{n+1}} + \ldots + R_{t_{n+d}}$ based on a given model f, one can derive a more conservative estimate by considering the distorted sum $\tilde{R}_{t_{n+1}} + \ldots + \tilde{R}_{t_{n+d}}$, with $\tilde{R}_{t_{n+k}} = f_k(\tilde{R}_{t_{n+1}}, \ldots, \tilde{R}_{t_{n+k-1}}, \tilde{U}_k)$ and $\{\tilde{U}_k\}_{k=1,\ldots,d} \sim C_c$ constructed by Algorithm 6.3.4. By means of backtesting, one can set the distortion parameter c such that the frequency of historically observed VaR-undershootings is lowered to a level the company management, respectively the regulator, prescribes. Recall from, e.g., CESR (2010), that an undershooting is defined as an empirically observed return that has been worse than forecasted by the VaR, and the relative frequency of such undershootings must not exceed the confidence level of the VaR-measurement.

The models we implement in our case study are precisely the ones from Alexander and Sheedy (2008). In this work, eight specifications of f are examined that can be divided into the Groups (I) and (II) from Section 6.2.1. Group (I) consists of four i.i.d. models for the log-returns, meaning that $R_{t_{n+1}}, \ldots, R_{t_{n+d}}$ are assumed to be i.i.d. with univariate distribution functions $G_{(m)}$, $m = 1, \ldots, 4$, respectively. The marginal laws comprise the normal distribution ($m = 1$), a modified version of the empirical distribution function, i.e. historical simulation ($m = 2$), the scaled Student's t-distribution ($m = 3$), and a normal mixture with two normally-distributed components ($m = 4$).

Group (II) consists of GARCH$(1, 1)$-processes for the conditional variance of the mean-adjusted portfolio returns, see Bollerslev (1986) for details. The law of the innovations in the conditional variance process are chosen from (scaled versions of) the i.i.d. model distributions $G_{(m)}$ (denoted $G_{(m),\mathrm{GARCH}}$), yielding another four specifications for f. Following the notation in Alexander and Sheedy (2008), we classify the i.i.d. models from Group (I) as types *unconditional normal*, *unconditional empirical*, *unconditional Student's t*, and *unconditional normal mixture*, while using the terms *conditional normal*, *conditional empirical* etc. when referring to the GARCH models from Group (II). A description of the models' structure and their calibration is outlined in the respective reference. As we are not aiming at discussing model selection and parameter estimation, but rather focus on incorporating an uncertainty add-on into a given specification, we have adopted the fitting methodology from Alexander and Sheedy (2008) and refer to this paper for more details.

The time series we rely on for VaR backtesting consists of the past 25 years of daily S&P 500 log-returns, yielding a total of 6 423 observations $\{r_{t_1}, \ldots, r_{t_{6\,423}}\}$ in the period from November 1989 to November 2014. To derive a series of VaR forecasts and compare them with realized returns, we proceed as follows:

(1) Estimate all models f based on the training sample $\{r_{t_k}\}_{k=1,\ldots,n}$. For the i.i.d. models, this means determining $G_{(m)}$, $m = 1, \ldots, 4$. For the conditional models, this implies estimating both the GARCH parameters and the innovation distributions $G_{(m),\mathrm{GARCH}}$, $m = 1, \ldots, 4$.

(2) Forecast the distorted d-day VaR of $\tilde{R}_{t_{n+1}} + \ldots + \tilde{R}_{t_{n+d}}$ via a Monte Carlo simulation by executing the following steps:

(a) Generate several samples of $\{\tilde{U}_k\}_{k=1,\ldots,d}$ with Algorithm 6.3.4.

(b.GI) Set $\tilde{R}_{t_{n+k}} = G^{-1}_{(m)}(\tilde{U}_k)$, $k = 1, \ldots, d$, for each sample in the i.i.d. case.

(b.GII) Determine the innovations $\tilde{Z}_{n+k} = G^{-1}_{(m),\mathrm{GARCH}}(\tilde{U}_k)$, $k = 1, \ldots, d$, for each sample in the conditional case and set $\tilde{R}_{t_{n+k}} = \sigma_{n+k}\,\tilde{Z}_{n+k}$ according to the GARCH parameters for σ_{n+k}. In general, σ_{n+k} depends on GARCH model parameters as well as previous realizations $\tilde{R}_{t_n}, \ldots, \tilde{R}_{t_{n+k-1}}$.

(c) Compute the sample VaR of $\tilde{R}_{t_{n+1}} + \ldots + \tilde{R}_{t_{n+d}}$ for each model.

(3) Compare the VaR estimate with the empirically observed return $r_{t_{n+1}} + \ldots + r_{t_{n+d}}$.

(4) Repeat Steps (1)–(3) after shifting the time series by d observations, i.e. choose $\{r_{t_{1+d}}, \ldots, r_{t_{n+d}}\}$ as the new training sample and proceed iteratively.

6.4.2 Conclusions of the case study

Using a training sample of size $n = 250$ for model estimation and considering a period of one week (i.e. $d = 5$ trading days) for VaR forecasting, we end up with a total of $(6\,423 - 250)/5 \approx 1\,234$ distorted VaR estimates which we compare to the realized weekly returns. We illustrate the results by means of Figures 6.4 and 6.5.

Figure 6.4 shows the resulting return histograms that are generated in each iteration, more precisely, in Steps (1)–(2) of the methodology described above. The return distribution widens, as the disturbance of the random input leads to an accumulation of extreme events. Due to radial symmetry of the distortion, the stylized shape of the distributions is maintained. The impact of the Dirichlet framework on the VaR and

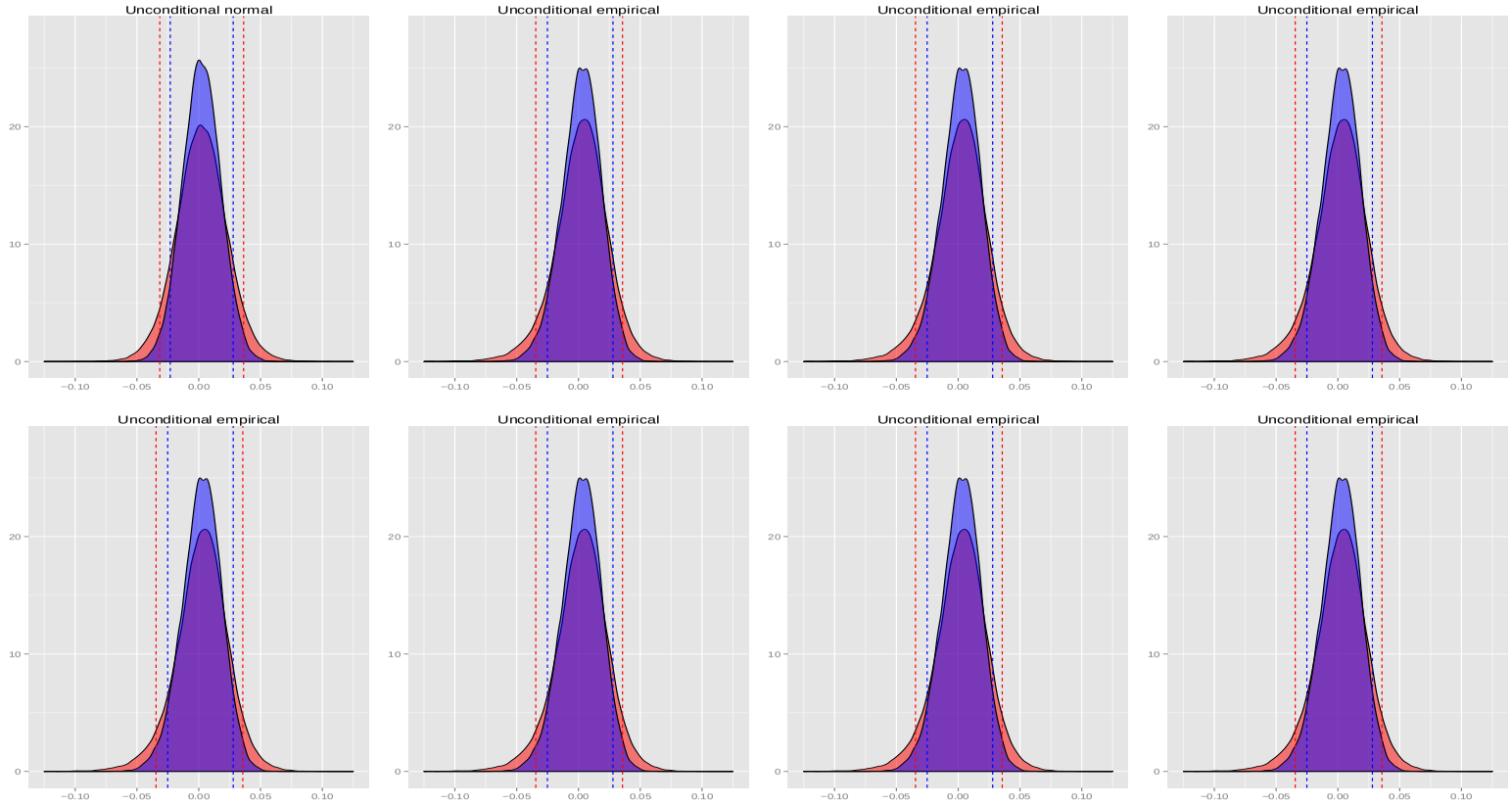

Figure 6.4 Smoothed histograms of the simulated sum $\tilde{R}_{t_{n+1}} + \ldots + \tilde{R}_{t_{n+5}}$ for $N = 100\,000$ samples among all eight models in both the undistorted case (blue, $c = \infty$) and the distorted case (red, $c = 4$). This level of distortion corresponds to $\rho_S = 0.2$, see Lemma 6.3.3. The dotted vertical lines illustrate the lower and upper VaR estimates at level $\alpha = 5\%$ before (blue) and after (red) distortion. The reference date for the values is October 28, 2014, meaning that based on the past $n = 250$ log-returns, return estimates for the next weekly log-return (i.e. $d = 5$ trading days, corresponds to November 4, 2014) are derived.

conditional Value-at-Risk (CVaR) varies among the models. For instance, the CVaR estimates do not double in all four unconditional cases, while they roughly triple for the conditional Student's t model. This means that the conditional Student's t model is less robust with respect to a distortion of the stochastic input than the unconditional models.

When looking at the distortion's influence on a given model over time, however, the impact seems to be rather homogeneous. Figure 6.5 shows three lines of time-varying VaR estimates in the unconditional normal case (corresponding to different levels of distortion) and (as a side information) further indicates the corresponding realized weekly returns. It becomes apparent that the relative VaR increase is quite stable over time and

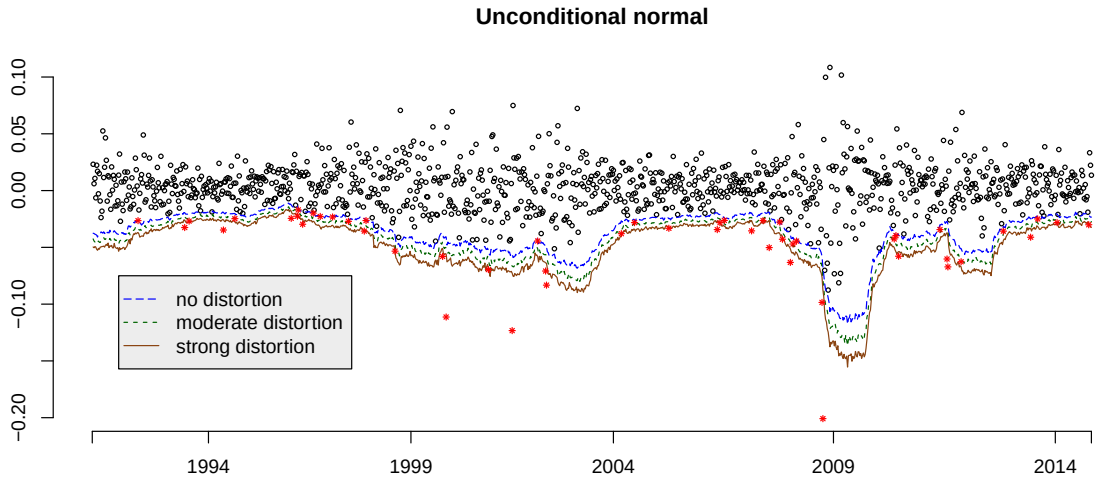

Figure 6.5 Time series of observed weekly returns (circles) and corresponding 5%-VaR estimates for the unconditional normal method (lines). The blue line indicates the estimate in the undistorted case ($c = \infty$), the green one in the moderately distorted case ($c = 9$), and the brown one for a stronger distortion ($c = 4$). The red "$*$"-symbols indicate the observed returns that undercut the VaR-forecast in the undistorted case.

does not seem to depend on the (time-varying) model parameters. A similar observation can be extracted for the other models under consideration.

Last but not least, Table 6.1 shows the relative frequency of 5%-VaR violations for the different models, i.e. the number of times an observed weekly return has undershot the VaR forecast, divided by the total number of weekly returns, w.r.t. the whole S&P 500 time series. As expected, the proportion of VaR violations decreases with increasing

Table 6.1 Total proportion of observed weekly returns that undershot the 5%-VaR forecast derived with different levels of distortion.

	Distortion parameter		
Model	$c = \infty$	$c = 9$	$c = 4$
Unconditional normal	0.045	0.029	0.022
Unconditional empirical	0.037	0.027	0.019
Unconditional Student's t	0.050	0.030	0.026
Unconditional normal mixture	0.034	0.026	0.018
Conditional normal	0.048	0.032	0.019
Conditional empirical	0.041	0.026	0.013
Conditional Student's t	0.050	0.032	0.019
Conditional normal mixture	0.044	0.032	0.019

severity of the distortion as the estimates become more conservative. In the present application, the undistorted estimates seem to capture the 5%-quantile of the weekly return distribution quite accurately. Nevertheless, the company-internal risk management guidelines might prescribe a relative frequency of VaR-undershootings below the confidence level applied in the VaR-computation. Such a guideline might avoid undesired conversations with the regulator, who monitors the relative frequency and requires risk management to review their VaR-model in case of too many undershootings. Distorting the model according to the approach described in the present chapter might be a convenient way to implement such a guideline as it is easy to implement as an add-on to the existing model and can be justified as a charge for model uncertainty.

7 Conclusion and outlook

This thesis has extensively treated exchangeable exogenous shock models $(X_1, \ldots, X_d)$, which can be represented as

$$X_k = \min\{Z^E : k \in E\}, \quad k = 1, \ldots, d,$$

for $2^d - 1$ independent real-valued random variables Z^E, $\emptyset \neq E \subseteq \{1, \ldots, d\}$, whose distribution function solely depends on the cardinality of E.

The central starting point for this thesis' main findings and applications is Chapter 3, which fully characterizes these models by means of the survival copula of $(X_1, \ldots, X_d)$. While Proposition 3.1.3 (p. 54) states that the survival copula necessarily has the form

$$C(u_1, \ldots, u_d) = \prod_{k=1}^{d} g_k(u_{(k)}), \quad u_1, \ldots, u_d \in [0, 1],$$

the fundamental Theorem 3.3.1 (p. 61) establishes the converse and emphasizes that C is a copula only if it can be associated with an exchangeable exogenous shock model. To this end, besides explicitly deriving the stochastic model behind, C is equivalently characterized via both inequality conditions (see Theorem 3.3.1.(ii)) and monotonicity requirements (see Theorem 3.3.1.(iii)) for the functions g_k. Moreover, Proposition 3.5.1 (p. 80) derives a second stochastic representation for an extendible subclass of exchangeable exogenous shock models (termed additive–frailty model) based on a first-passage time construction

$$X_k = \inf\{t \geq 0 : \Lambda_t \geq E_k\}, \quad k = 1, \ldots, d,$$

which is parameterized by an additive subordinator $\Lambda = \{\Lambda_t\}_{t \geq 0}$. Different specifications of Λ can be used to easily construct new, tractable families of multivariate distribution functions with shock model interpretation. A possible challenge for prospective research is to analyze whether the additive–frailty construction coincides with the class of extendible exogenous shock models. As similar results for the special case of exponentially

distributed shocks are already available in the literature, I am confident this can be achieved (though the generality of the result is probably quite demanding).

The powerful connection between the additive–frailty framework for $(X_1, \ldots, X_d)$ and the corresponding shock construction is exploited in the subsequent paragraphs. By construction, the additive subordinator Λ satisfies the inequality and monotonicity requirements for the functions g_k induced by Theorem 3.3.1.(ii)+(iii). Possibly, these conditions are not only necessary, but sufficient to uniquely determine the process Λ. Chapter 4 addresses this question when specifying Λ as a Sato subordinator. With Sato subordinators being linked to self-decomposable probability laws on the positive half-axis, the described procedure results in Theorem 4.2.1 (p. 91) and yields two novel characterizations of self-decomposable Bernstein functions Ψ, one in terms of monotonicity requirements for Ψ and one in terms of multivariate distribution functions. The distribution function (termed Sato–frailty copula) is studied in detail, involving various examples, the computation of several properties, and the comparison to the structurally surprisingly similar class of Archimedean copulas. Concerning the monotonicity characterization of self-decomposable Bernstein functions, a possible application we have thought of is the analysis of the infinitely divisible Hartman–Watson law, which is an important object in mathematical finance due to its appearance in the pricing of Asian options. To the best of our knowledge, the Lévy measure of the associated Bernstein function Ψ is not known and can thus not be used to directly determine whether Ψ is self-decomposable. It might be worth checking whether Ψ satisfies our conditions for self-decomposability, a task which we have briefly worked on up to now, however not pursued in detail.

Regarding the effectiveness and usefulness of exchangeable exogenous shock models in the field of mathematical finance, additive–frailty models are applied in Chapters 5 and 6 in two completely different contexts. Chapter 5 treats the modeling of a high-dimensional portfolio $(X_1, \ldots, X_d)$ of default times in order to price CDOs. In an additive–frailty setup, the loss process connected to the default times can be approximated by the parameterizing additive subordinator Λ. The crucial result in this application is Theorem 5.3.2 (p. 122), which depicts a highly efficient, mathematically rigorous pricing formula based on Laplace inversion. Following the calibration procedure in Section 5.4, goodness-of-fit tests w.r.t. CDO market tranche spreads could be executed for Sato–frailty and other additive–frailty models. As pointed out earlier, we deliberately present the methodology on a conceptual layer. However, we consider an empirical comparison in future research interesting and valuable.

As a second practical application, exchangeable exogenous shock models are used in the

context of model uncertainty in Chapter 6. Based on the "Markov regression representation" of an arbitrary random vector $\boldsymbol{R} = \{R_{t_k}\}_{k=1,\dots,d}$, which yields a decomposition $R_{t_k} = f_k(R_{t_1}, \dots, R_{t_{k-1}}, U_k), k = 1, \dots, d$, we have introduced a new philosophy for model robustness that applies to arbitrary stochastic models. The core idea is to separate model assumptions f from pure randomness $\boldsymbol{U} = \{U_k\}_{k=1,\dots,d}$ and distort the latter. This distortion, however, has to be done in a well considered way, for which we postulate axioms. In Theorem 6.3.1 (p. 135), it is shown that an additive–frailty model based on the Dirichlet process satisfies all postulated axioms. The methodology is finally brought to life by providing an efficient simulation algorithm that is applied in a case study.

To put it in a nutshell, as hinted by this thesis' title, we provide a comprehensive analysis, characterization, and application of exchangeable exogenous shock models. The link to additive–frailty models provides a powerful tool in exploring new copula families, investigating classes of additive processes, and deploying the stochastic framework in practice.

Bibliography

Aldous, D. (1985). Exchangeability and related topics. In *Ecole d'Été de Probabilités de Saint-Flour XIII–1983*, pages 1–198, Berlin. Springer.

Alexander, C. and Sarabia, J. M. (2012). Quantile uncertainty and Value-at-Risk model risk. *Risk Analysis*, 32(8):1293–1308.

Alexander, C. and Sheedy, E. (2008). Developing a stress testing framework based on market risk models. *Journal of Banking and Finance*, 32(10):2220–2236.

Amato, J. and Gyntelberg, J. (2005). CDS index tranches and the pricing of credit risk correlations. *BIS Quarterly Review*, pages 73–87.

Anastasiadis, S. and Chukova, S. (2012). Multivariate insurance models: An overview. *Insurance: Mathematics and Economics*, 51(1):222–227.

Andersen, L. and Sidenius, J. (2004). Extensions to the Gaussian copula: Random recovery and random factor loadings. *Journal of Credit Risk*, 1(1):29–70.

Arnsdorf, M. and Halperin, I. (2008). BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives. *Journal of Computational Finance*, 12(2):77–100.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Avellaneda, M., Levy, A., and Paràs, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88.

Baglioni, A. and Cherubini, U. (2013). Marking-to-market government and guarantees to financial systems - Theory and evidence for Europe. *Journal of International Money and Finance*, 32:990–1007.

Bannör, K.-F. and Scherer, M. (2013). Capturing parameter risk with convex risk measures. *European Actuarial Journal*, 3(1):97–132.

Barndorff-Nielsen, O. E. and Sheppard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial econometrics. *Journal of the Royal Statistical Society*, 63(2):167–241.

Basel Committee on Banking Supervision (2010). A global regulatory framework for more resilient banks and banking systems. `http://www.bis.org/publ/bcbs189.pdf`. Accessed January 29, 2015.

Bernhart, G., Mai, J.-F., Schenk, S., and Scherer, M. (2015). The density for distributions of the Bondesson class. *Journal of Computational Finance*, 18(3):99–128.

Bernstein, S. N. (1929). Sur les fonctions absolument monotones. *Acta Mathematica*, 52(1):1–66.

Bielecki, T. and Rutkowski, M. (2001). *Credit risk: Modeling, valuation, and hedging*. Springer, Berlin.

Bignozzi, V. and Tsanakas, A. (2014). Model uncertainty in risk capital measurement. Forthcoming in Journal of Risk.

Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Mathematical Statistics, New York.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Brigo, D., Pallavicini, A., and Torresetti, R. (2007). Calibration of CDO tranches with the dynamical generalized Poisson loss model. *Risk*, 20(5):70–75.

Burtschell, X., Gregory, J., and Laurent, J.-P. (2009). A comparative analysis of CDO pricing models. *Journal of Derivatives*, 16(4):9–37.

Carr, P., Geman, H., Madan, D., and Yor, M. (2005). Pricing options on realized variance. *Finance and Stochastics*, 9(4):453–475.

Carr, P., Geman, H., Madan, D., and Yor, M. (2007). Self-decomposability and option pricing. *Mathematical Finance*, 17(1):31–57.

CESR (2010). Cesr's guidelines on risk measurement and the calculation of global exposure and counterparty risk for UCITS. `http://www.esma.europa.eu/system/files/10\_788.pdf`. Accessed January 29, 2015.

Charpentier, A. and Segers, J. (2009). Tails of multivariate Archimedean copulas. *Journal of Multivariate Analysis*, 100(7):1521–1537.

Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547.

Cont, R., Deguest, R., and Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6):593–606.

Cont, R. and Minca, A. (2008). Recovering portfolio default intensities implied by CDO quotes. *Mathematical Finance*, 23(1):94–121.

Darsow, W. F., Nguyen, B., and Olsen, E. T. (1992). Copulas and Markov processes. *Illinois Journal of Mathematics*, 36(4):600–642.

De Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Insitut Henri Poincaré*, 7:1–68.

Doetsch, G. (1974). *Introduction to the theory and application of the Laplace transformation*. Springer, Berlin.

Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *The Annals of Probability*, 2(2):183–201.

Durante, F. (2006). A new class of symmetric bivariate copulas. *Journal of Nonparametric Statistics*, 18(7-8):499–510.

Durante, F., Kolesárová, A., Mesiar, R., and Sempi, C. (2008). Semilinear copulas. *Fuzzy Sets and Systems*, 159(1):63–76.

Durante, F., Quesada-Molina, J. J., and Úbeda-Flores, M. (2007). On a family of multivariate copulas for aggregation processes. *Information Sciences*, 177(24):5715–5724.

Durante, F., Sánchez, J. F., and Sempi, C. (2013). A note on the notion of singular copula. *Fuzzy Sets and Systems*, 211:120–122.

Durante, F. and Sempi, C. (2010). Copula theory: An introduction. In Jaworski, P., Durante, F., Härdle, W. K., and Rychlik, T., editors, *Copula Theory and Its Applications*, Lecture Notes in Statistics, pages 3–31. Springer, Berlin.

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance*, 76(3):639–650.

Embrechts, P. and Hofert, M. (2013). A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432.

Embrechts, P., Puccetti, G., and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, 37(8):2750–2764.

Federal Reserve System (2011). Supervisory guidance on model risk management. `http://www.federalreserve.gov/bankinforeg/srletters/sr1107.htm`. Accessed January 29, 2015.

Feller, W. (1966). *An introduction to probability theory and its applications*, volume II. John Wiley and Sons Inc., New York.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209–230.

Ferguson, T. S. (1974). Prior distributions on spaces of probability measures. *The Annals of Statistics*, 2(4):615–629.

Genest, C. and Rivest, L.-P. (1989). A characterization of Gumbel's family of extreme value distributions. *Statistics and Probability Letters*, 8(3):207–211.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88(423):1034–1043.

Giesecke, K. (2003). A simple exponential model for dependent defaults. *Journal of Fixed Income*, 13(3):74–83.

Giesecke, K., Goldberg, L., and Ding, X. (2010a). A top-down approach to multi-name credit. *Operations Research*, 59(2):283–300.

Giesecke, K., Goldberg, L., and Errais, E. (2010b). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.

Gnedin, A. and Pitman, J. (2008). Moments of convex distribution functions and completely alternating sequences. *Probability and Statistics: Essays in honor of David A. Freedman*, 2:30–41.

Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. In Jaworski, P., Durante, F., Härdle, W., and Rychlik, T., editors, *Copulae Theory and its Applications: Proceedings of the Workshop held in Cracow, 25-26 September 2009*, pages 127–146.

Gupta, A., Reisinger, C., and Whitley, A. (2010). Model uncertainty and its impact on derivative pricing. In Böcker, K., editor, *Rethinking Risk Measurement and Reporting*, pages 625–663, London. Risk Books.

Hausdorff, F. (1921). Summationsmethoden und Momentfolgen I. *Mathematische Zeitschrift*, 9(1–2):74–109.

Hering, C., Hofert, M., Mai, J.-F., and Scherer, M. (2010). Constructing nested Archimedean copulas with Lévy subordinators. *Journal of Multivariate Analysis*, 101(6):1428–1433.

Hofert, M. (2008). Sampling Archimedean copulas. *Computational Statistics & Data Analysis*, 52(12):5163–5174.

Hofert, M. and Scherer, M. (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance*, 11(5):775–787.

Hull, J. (2008). *Options, futures and other derivatives*. Pearson Prentice Hall, Upper Saddle River, 7th edition.

Jeanblanc, M., Pitman, J., and Yor, M. (2002). Self-similar processes with independent increments associated with Lévy and Bessel processes. *Stochastic Processes and their Applications*, 100(1–2):223–231.

Joe, H. (1990). Multivariate concordance. *Journal of Multivariate Analysis*, 35(1):12–30.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Jurek, Z. J. and Vervaat, W. (1983). An integral representation for selfdecomposable Banach space valued random variable. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62(2):247–262.

Jurek, Z. J. and Yor, M. (2004). Selfdecomposable laws associated with hyperbolic functions. *Probability and Mathematical Statistics*, 24(1):181–190.

Kant, B. and Schenk, S. (2015). Lévyprozesse in der Portfoliokreditrisikomodellierung. *Risikomanager*, 6:7—11.

Khintchine, A. Y. (1938). Limit laws of sums of independent random variables. *ONTI*. Moscow.

Kiesel, R. and Lutz, M. (2011). Efficient pricing of constant maturity swap spread options in a stochastic volatility LIBOR market model. *Journal of Computational Finance*, 14(3):37–72.

Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae*, 10(2–3):152–164.

Klenke, A. (2006). *Wahrscheinlichkeitstheorie*. Springer, Berlin.

Kundu, D., Franco, M., and Vivo, J.-M. (2014). Multivariate distributions with proportional reversed hazard marginals. *Computational Statistics & Data Analysis*, 77:98–112.

Kundu, D. and Gupta, A. K. (2014). On bivariate Weibull-Geometric distribution. *Journal of Multivariate Analysis*, 123:19–29.

Kundu, D. and Gupta, R. D. (2009). Bivariate generalized exponential distribution. *Journal of Multivariate Analysis*, 100(4):581–593.

Lévy, P. (1937). Théorie de l'addition des variables aléatoires. Paris: Gauthier-Villars.

Li, D. (2000). On default correlation: A copula function approach. *Journal of Fixed Income*, 9(4):43–54.

Li, H. (2008). Tail dependence comparison of survival Marshall–Olkin copulas. *Methodology and Computing in Applied Probability*, 10(1):39–54.

Li, X. and Pellerey, F. (2011). Generalized Marshall–Olkin distributions and related bivariate aging properties. *Journal of Multivariate Analysis*, 102(10):1399–1409.

Linkskog, F. and McNeil, A. J. (2003). Common Poisson shock models. *ASTIN Bulletin*, 33(2):209–238.

Lopatin, A. and Misirpashaev, T. (2008). Two-dimensional Markovian model for dynamics of aggregate credit loss. *Advances in Econometrics*, 22:243–274.

Lorch, L. and Newman, D. J. (1983). On the composition of completely monotonic functions and completely monotonic sequences and related questions. *Journal of the London Mathematical Society*, 28(1):31–45.

Mai, J.-F. (2010). *Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk.* PhD thesis, Technische Universität München.

Mai, J.-F., Olivares, P., Schenk, S., and Scherer, M. (2014). A multivariate default model with spread and event risk. *Applied Mathematical Finance*, 21(1):51–83.

Mai, J.-F., Schenk, S., and Scherer, M. (2015a). Analyzing model uncertainty via distortion of the stochastic root: a Dirichlet prior approach. Working paper.

Mai, J.-F., Schenk, S., and Scherer, M. (2015b). Two novel characterizations of self-decomposability on the half-line. *Forthcoming in Journal of Theoretical Probability.*

Mai, J.-F., Schenk, S., and Scherer, M. (2016). Exchangeable exogenous shock models. *Bernoulli*, 22(2):1278–1299.

Mai, J.-F. and Scherer, M. (2009a). Efficiently sampling exchangeable Cuadras–Augé copulas in high dimensions. *Information Sciences*, 17(5):2872–2877.

Mai, J.-F. and Scherer, M. (2009b). Lévy-frailty copulas. *Journal of Multivariate Analysis*, 100(7):1567–1585.

Mai, J.-F. and Scherer, M. (2009c). A tractable multivariate default model based on a stochastic time-change. *International Journal of Theoretical and Applied Finance*, 12(2):227–249.

Mai, J.-F. and Scherer, M. (2011). Reparameterizing Marshall–Olkin copulas with applications to sampling. *Journal of Statistical Computation and Simulation*, 81(1):59–78.

Mai, J.-F. and Scherer, M. (2012). *Simulating copulas: Stochastic models, sampling algorithms, and applications*. Series in Quantitative Finance, Imperial College Press, London.

Mai, J.-F. and Scherer, M. (2014). Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time. *Extremes*, 17(1):77–95.

Mai, J.-F., Scherer, M., and Zagst, R. (2012). CIID frailty models and implied copulas. In *Copulae in Mathematical and Quantitative Finance, Proceedings of the Workshop held in Cracow, 10-11 July 2012*, Berlin. Springer.

Mansuy, R. (2005). On processes which are infinitely divisible with respect to time. Working paper.

Markit (2014). Markit credit indices – a primer. `http://www.markit.com/assets/en/docs/products/data/indices/credit-index-annexes/Markit%20Credit%20Indices%20Primer.pdf`. Accessed on March 3, 2015.

Marsaglia, G. and Tsang, W. W. (1984). A fast, easily implemented method for sampling from decreasing or symmetric unimodal density functions. *SIAM Journal on scientific and statistical computing*, 5(2):349–359.

Marsaglia, G. and Tubilla, A. (1975). A note on the "lack of memory" property of the exponential distribution. *The Annals of Probability*, 3(2):353–354.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44.

Marshall, A. W. and Olkin, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association*, 83(403):834–841.

McNeil, A. J. (2008). Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation*, 78(6):567–581.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*. Princeton University Press, Princeton.

McNeil, A. J. and Nešlehovà, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Annals of Statistics*, 37(5B):3059–3097.

Müller, A. and Scarsini, M. (2005). Archimedean copulae and positive dependence. *Journal of Multivariate Analysis*, 93(2):434–445.

Nelsen, R. B. (2006). *An introduction to copulas*. Second edition. Springer, New York.

Pérignon, C. and Smith, D. R. (2010). The level and quality of Value-at-Risk disclosure by commercial banks. *Journal of Banking and Finance*, 34(2):362–377.

Raible, S. (2000). *Lévy processes in finance: Theory, numerics, and empirical facts*. PhD thesis, University of Freiburg.

Ressel, P. (2011). Monotonicity properties of multivariate distribution and survival functions – With an application to Lévy-frailty copulas. *Journal of Multivariate Analysis*, 102(3):393–404.

Ressel, P. (2013). Finite exchangeability, Lévy-frailty copulas and higher-order monotonic sequences. *Journal of Theoretical Probability*, 26(3):666–675.

Rüschendorf, L. and de Valk, V. (1993). On regression representations of stochastic processes. *Stochastic Processes and their Applications*, 46(2):183–198.

Sarhan, A. M. and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture. *Journal of Multivariate Analysis*, 98(7):1508–1527.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press, Cambridge.

Sato, K. and Yamazato, M. (1984). Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Processes and their Applications*, 17(1):73–100.

Scarsini, M. (1984). On measures of concordance. *Stochastica*, 8(3):201–218.

Schenk, S. (2011). CIID models – A new multivariate default model based on CGMY-type processes. Master's thesis, TU München.

Schenk, S. (2014). Optionspreisberechnung via Fast Fourier Transform. *Risikomanager*, 23:7–12.

Schilling, R. L., Song, R., and Vondraček, Z. (2010). *Bernstein functions*. De Gruyter, Berlin.

Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, 98(6):1123–1140.

Schmid, F., Schmidt, R., Blumentritt, T., Gaißer, S., and Ruppert, M. (2010). Copula-based measures of multivariate association. In Jaworski, P., Durante, F., Härdle, W., and Rychlik, T., editors, *Copulae Theory and its Applications: Proceedings of the Workshop held in Cracow, 25-26 September 2009*, pages 209–236.

Schönbucher, P. (2005). Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives. Working paper.

Schoutens, W., Simons, E., and Tistaert, J. (2004). A perfect calibration! Now what? *Wilmott Magazine*, pages 66–78.

Schweizer, B. and Sklar, A. (1983). *Probabilistic metric spaces*. North-Holland/Elsevier, New York.

Seshadri, V. (1993). *The inverse Gaussian distribution*. Oxford science publications, Oxford.

Shoaee, S. and Khorram, E. (2012). A new absolute continuous bivariate generalized exponential distribution. *Journal of Statistical Planning and Inference*, 142(7):2203–2220.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231.

Talbot, A. (1979). The accurate numerical inversion of Laplace transforms. *IMA Journal of Applied Mathematics*, 23(1):97–120.

Taylor, M. D. (2007). Multivariate measures of concordance. *Annals of the Institute for Statistical Mathematics*, 59(4):789–806.

Tricomi, F. G. (1951). On the Theorem of Frullani. *The American Mathematical Monthly*, 58(3):158–164.

Vasicek, O. (1987). Probability of loss on loan portfolio. Working paper, KMV Corporation.

Weideman, J. A. C. (2006). Optimizing Talbot's contours for the inversion of the Laplace transform. *SIAM Journal of Numerical Analysis*, 44(6):2342–2362.

Widder, D. V. (1946). *The Laplace transform*. Princeton University Press, Princeton.

Williamson, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Mathematical Journal*, 23(2):189–207.

Wolfe, S. J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \rho\, X_{n-1} + B_n$. *Stochastic Processes and their Applications*, 12(3):301–312.

List of Tables

List of Figures