Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Finanzmathematik

Advances in financial engineering: Bondesson densities, the construction of MSMVE distributions, and the modeling of discrete cash dividends

German Bernhart

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. Noam Berger Steiger
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Matthias Scherer
	2.	Univ.-Prof. Dr. Frank Seifried, Universität Trier
	3.	Prof. Dr. Marcos Escobar, Ryerson University Toronto/Kanada (schriftliche Beurteilung)

Die Dissertation wurde am 23.06.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 21.09.2015 angenommen.

Abstract

This thesis covers three different but interconnected topics in the broad field of financial engineering. It is concerned with representations of univariate distributions on the positive real line, the construction of specific multivariate distributions, and the modeling of discrete cash dividends for stock prices. Based on Laplace inversion methods, a convenient representation for the density of distributions of a large subclass of infinitely divisible distributions on the positive real line is developed. This result can be used to numerically derive the density when closed algebraic expressions are unknown or difficult to evaluate. Similar results are proven for distribution functions and option-like derivatives. Furthermore, new families of min-stable multivariate exponential distributions are constructed using a first-passage time construction. These can be of particular use for portfolio credit risk models as they are tractable also in high dimensions. Finally, a new and flexible approach for the modeling of stock prices with discrete cash dividends is presented. It allows to incorporate non-deterministic dividend payments into almost any stock price model while retaining tractability.

Zusammenfassung

Diese Arbeit behandelt drei verschiedene, aber miteinander verknüpfte Themen aus dem umfassenden Gebiet der Finanzmathematik, welche alle konkreten Anwendungsbezug aufweisen. Konkret beschäftigt sie sich mit Repräsentationen von Verteilungen auf den positiven reellen Zahlen, der Konstruktion einer bestimmten Klasse von multivariaten Verteilungen und der Modellierung diskreter Dividendenzahlungen in Aktienmodellen. Basierend auf Laplace Inversions-Methoden wird eine handhabbare Repräsentation der Dichte von Verteilungen einer speziellen Unterklasse der positiven, unendlich teilbaren Verteilungen hergeleitet. Dieses Ergebniss kann für die numerische Berechnung dieser Dichte verwendet werden, falls keine geschlossene oder praktikable Darstellung bekannt ist. Ähnliche Ergebnisse für Verteilungsfunktionen und Derivate mit Optionsstruktur werden abgeleitet. Des Weiteren werden neue Familien sogenannter minimum-stabiler multivariater Exponentialverteilungen konstruiert. Für deren Konstruktion wird ein Zusammenhang zwischen einer bestimmten Familie stochastischer Prozesse und dieser Klasse multivariater Verteilungen ausgenutzt. Die resultierenden Verteilungen sind insbesondere nützlich für die Modellierung von großen Kreditportfolios, da sie auch in hohen Dimensionen leicht zu handhaben sind. Abschließend wird ein neuer und flexibler

Ansatz zur Modellierung von Aktien mit diskreten Dividendenzahlungen vorgestellt, der es erlaubt, nahezu jedes Aktienmodell um nicht deterministische Dividendenzahlungen zu erweitern.

Acknowledgements

First of all, I would like to thank my supervisor Matthias Scherer for giving me the possibility to write this thesis and all his support over the last years. In countless discussions, he always provided helpful feedback and valuable input, and encouraged me to pursue the ideas that attracted my attention. Furthermore, I am deeply grateful to Jan-Frederik Mai who served as an informal co-advisor and who was heavily involved in all the projects that converged into this thesis. He continuously produced new ideas, challenged existing approaches, and almost instantaneously provided valuable feedback on everything.

Furthermore, I am very grateful to Frank Seifried and to Marcos Escobar for agreeing to serve as referees for this thesis.

I would also like to thank Rudi Zagst and the whole team at the Chair of Mathematical Finance for the great working atmosphere. In particular, I would like to thank my office mate Steffen Schenk for myriads of interesting discussions, be it of academic or secular content, which always sharpened my understanding of the considered problems or simply provided welcome distraction.

Finally, I would like to thank my family and Bianca for their continuing support and encouragement in all situations.

Contents

1 Introduction

Financial mathematics is a wide field, covering diverse areas ranging from mathematical foundations to concrete applications. During my research, I had the great opportunity to work on several topics, which, though interlinked, were from quite different areas of this discipline. As a direct consequence, this thesis does not revolve around one common topic. Instead, three different topics are covered, their similarity being their applicability to financial engineering. The thesis is structured in two parts, a theoretically motivated one consisting of Chapters 3 and 4 (with Chapter 2 introducing the mathematical background), and a practically motivated one consisting of Chapter 5.[1] Whereas the theoretically motivated part is concerned with representations of distributions on the positive real line and the construction of specific multivariate distributions, the second part deals with the modeling of discrete cash dividends for stock prices.

Chapter 3 deals with distributions on the positive real line, which can be characterized by means of several concepts, among others probability densities and integral transforms such as the Laplace transform. Although Laplace transforms represent an elegant way to deal with distributions, probability densities allow for a more intuitive understanding and are necessary in many applications. If only the Laplace transform of a distribution is known, Laplace inversion methods provide a natural starting point to represent the probability density in terms of its Laplace transform. However, it is well known that the standard inversion integrals (which are path integrals in the complex plane) are difficult to evaluate numerically due to their often highly oscillating behavior and unboundedness. The standard method to cope with the first issue consists of transforming the path of the integral in the complex plane. The mathematical idea behind that is Cauchy's Theorem. Although simple on an abstract level, showing the admissibility of such path transforms for a given problem is difficult. Proving the admissibility of a certain path transform for a very large class of distributions (see Theorem 3.3) and illustrating its

[1]Thus, the structure of this thesis mirrors in some sense my then two-pronged work activities. While working as a research and teaching assistant at university, I also worked part-time in the financial industry as a financial engineer.

tractability constitutes our main contribution in Chapter 3. The class considered is the Bondesson class, a large subclass of positive infinitely divisible distributions, including many important distributions such as the positive stable law, the inverse Gaussian law, the Gamma law, and the Hartman–Watson law. In addition, as a side product, similar results for distribution functions and prices of certain financial products can be derived. Since positive infinitely divisible distributions are characterized by Bernstein functions, those functions play a crucial role in our derivation. Thus, Section 2.2 provides an introduction to positive infinitely divisible distributions and their link to Bernstein functions.

The second part of our theoretically motivated work, Chapter 4, is located in the area of multivariate distributions. While there is no direct link to our work on one-dimensional probability densities, there are mathematical concepts that are relevant for both. Bernstein functions, in particular, also play a crucial role in Chapter 4, and can thus be seen as the theoretical link between both chapters. Min-stable multivariate exponential (MSMVE) distributions represent an important class of multivariate distributions and are thoroughly studied on an abstract level. Recently, Mai and Scherer (2014) stated an elegant correspondence between the extendible subclass of these multivariate distributions and a certain class of one-dimensional stochastic processes. Our main contribution can be seen as complementing this abstract correspondence with concrete examples, thereby constructing new classes of MSMVE distributions, allowing for helpful stochastic representations (see Theorems 4.4 and 4.12). As the number of existing MSMVE models which are tractable in high dimensions is limited, this result is of importance on its own and can be of particular use for portfolio credit risk models. The interconnectedness of the diverse mathematical concepts involved, such as stochastic processes, multivariate distributions, and integral transform representations, makes this chapter appealing from a theoretical point of view. The necessary mathematical concepts are introduced in Sections 2.3 and 2.4.

When valuing stock derivatives, i.e. financial products whose value depends on the development of an underlying stock, most research focuses on capturing the most prominent features of the stock price evolution. Considerable progress has been made in incorporating effects such as stochastic/local volatility, price jumps, and default risk. However, in some situations, normally secondary features also come into play, and their appropriate modeling becomes vital. Such a feature, namely the existence of discrete dividend payments, is considered and analyzed in Chapter 5, and a potential approach to deal with it is presented.[2] The general idea is quite vivid, yet (see the representation in Theorem 5.6) it provides a new, elegant, and flexible approach to deal with the problem considered. Its applicability is illustrated in a small case study.

Chapter 2 introduces the mathematical background needed for Chapters 3 and 4. Taking account of the many areas concerned, starting from scratch is not possible, so we confined ourselves to an introduction of the most vital concepts. Everything that is not essential for a clear understanding is omitted to keep the exposition as elegant as possible. On the other hand, connections between different elements are highlighted and short proofs are added whenever they foster the understanding of underlying ideas. Subsequently, Chapters 3 and 4 present the theoretically motivated results on densities of distributions of the Bondesson class and MSMVE distributions. Finally, Chapter 5 presents the practically motivated part on the modelling of discrete cash dividends. No extensive mathematical background is needed and the necessary concepts and definitions are introduced there. As each of the three main topics is motivated separately, each of the main chapters starts with the motivation of its distinct aim.

The work presented resulted in several academic papers, see Bernhart and Mai (2014a); Bernhart and Mai (2015); Bernhart et al. (2015a); Bernhart et al. (2015b), which have been published in peer-reviewed journals or conference proceedings. Consequently, some chapters are strongly related to the respective publications, which we will always state explicitly at the beginning of the chapter.

[2]This chapter is considerably influenced by my work as a financial engineer at XAIA Investment GmbH.

2 Mathematical background

In this chapter, the most relevant mathematical concepts are introduced. As Chapter 3 deals with positive infinitely divisible distributions, this class of distributions and their link to Bernstein functions, which are needed in Chapter 4 as well, is presented in Section 2.2. Subsequently, a class of stochastic processes, called IDT subordinators, is introduced in Section 2.3, as these processes play a central role for the results in Chapter 4. This is due to their relation with min-stable multivariate exponential distributions, which are introduced in Section 2.4.

2.1 Notations and definitions

First of all, let us introduce the necessary notations and definitions.

- Let $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ denote, respectively, the sets of natural, rational, real, and complex numbers, with $\mathbb{Q}_+$ and $\mathbb{R}_+$ the collections of their non-negative elements, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For a complex number z, $\Re(z)$ represents its real part and $\Im(z)$ its imaginary part. $|B|$ denotes the number of elements in a set B. For two real numbers x_1 and x_2, $x_1 \wedge x_2$ denotes their minimum. $\mathbb{R}^d$ is the d-dimensional real coordinate space of column vectors $\mathbf{x} = (x_1, \ldots, x_d)^\intercal$, with $x_i \in \mathbb{R}, i = 1, \ldots, d$, and "$\intercal$" the symbol for the transpose of a matrix. $\mathbb{R}^d_+$ and similar expressions mean the respective subsets of $\mathbb{R}^d$. Furthermore, let $\mathbf{e}_1 := (1, 0, \ldots, 0)^\intercal, \ldots, \mathbf{e}_d := (0, , \ldots, 0, 1)^\intercal$ denote the d vectors of the canonical basis and $\mathbf{0} := (0, \ldots, 0)^\intercal$ the zero vector. Given a vector $\mathbf{x} = (x_1, \ldots, x_d)^\intercal$, $x_{(1)} \leq \ldots \leq x_{(d)}$ represents the ordered list of its elements.

- $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-algebra on $\mathbb{R}^d$ and for all subsets $A \subset \mathbb{R}^d$, we define $\mathcal{B}(A)$ simply as the restriction of $\mathcal{B}(\mathbb{R}^d)$ to A. Unless stated explicitly, when considering a measure μ on some set, it is always assumed that the corresponding measure space consists of the respective Borel σ-algebra. Its Laplace transform $\hat{\mu}$ is defined by

$$\hat{\mu}(x) := \int_{[0,\infty)} e^{-x\,u}\mu(\mathrm{d}u), \quad \text{for } x > 0.$$

The convolution of two finite measures μ_1 and μ_2 is denoted by $\mu_1 \star \mu_2$. δ_x denotes the Dirac measure centered on the point x.

- For a function $g : \mathbb{R} \to \mathbb{R}$, the expression $g^{(n)}$ represents the n-th derivative of this function, with $g^{(0)} = g$ and the alternative expression $g^{(1)} = g'$. The sets C^k denote the sets of k times continuously differentiable functions, with $C^\infty := \bigcap_{k\in\mathbb{N}} C^k$. Furthermore, $g(a+) := \lim_{h\searrow 0} g(a+h)$ and $g(a-) := \lim_{h\searrow 0} g(a-h)$, whenever existent.

- $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a probability space with event space Ω, σ-Algebra $\mathcal{F}$, and probability measure $\mathbb{P}$. By $\mathbb{E}$, the corresponding expectation operator is denoted. Whenever considering a random object (random variable, random vector, or stochastic process), this is assumed to be defined on a probability space, though we will omit that statement in most cases. If it is clear from the context, several random objects are assumed to be defined on the same probability space. $\mathcal{L}(.)$ represents the induced measure of a random object, i.e. for example for a random variable $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $\mathcal{L}(.)$ denotes the measure $\mathbb{P}(X^{-1}(.))$ defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. a.s. stands for "almost surely", a.e. for "almost everywhere" w.r.t. the Lebesgue measure, and iid for "independent and identically distributed". A stochastic process H denotes a family of real random variables $\{H_t\}_{t\geq 0}$, the function $t \mapsto H_t(\omega)$ with $\omega \in \Omega$ is called a path. Furthermore, $\overset{d}{=}$ denotes equality in distribution and $\overset{P}{\to}$ convergence in distribution.

- For a random variable X, $X \sim \mathrm{Exp}(\lambda)$ with $\lambda > 0$ means that $\mathcal{L}(X)(A) = \int_A \lambda \exp(-\lambda\, s)\,\mathrm{d}s$ for $A \in \mathcal{B}(\mathbb{R}_+)$, i.e. the random variable has an exponential distribution with parameter λ. Analogously, $X \sim \mathrm{U}([a,b])$ with $a < b$ means that $\mathcal{L}(X)(A) = \int_A 1/(b-a)\,\mathrm{d}s$ for $A \in \mathcal{B}([a,b])$, i.e. the random variable is uniformly distributed on $[a,b]$. Furthermore, $X \sim \mathcal{N}(\mu, \sigma^2)$ denotes a normally distributed random variable with mean μ and variance σ^2, i.e. its distribution is given by $\mathcal{L}(X)(A) = \int_A 1/\sqrt{2\,\pi\,\sigma^2}\,\exp(-(s-\mu)^2/(2\sigma^2))\,\mathrm{d}s$ for $A \in \mathcal{B}(\mathbb{R})$. Finally, a random variable X is said to be $\Gamma(c,d)$-distributed with $c, d > 0$, if its distribution is given by $\mathcal{L}(X)(A) = \int_A d^c/\Gamma(c)\, s^{c-1}\exp(-d\,s)\,\mathrm{d}s$ for $A \in \mathcal{B}(\mathbb{R}_+)$.

- $\Gamma : (0, \infty) \to \mathbb{R}$ denotes the Gamma function defined by $\Gamma(t) := \int_0^\infty s^{t-1} \exp(-s)\,\mathrm{d}s$. It has an analytic extension to $\mathbb{C}\backslash\{0, -1, -2, \ldots\}$. Furthermore, $\gamma : (0, \infty)^2 \to \mathbb{R}$ denotes the incomplete Gamma function defined by $\gamma(t, x) := \int_0^x s^{t-1} \exp(-s)\,\mathrm{d}s$.

2.2 Positive infinitely divisible distributions

Infinitely divisible (ID) distributions constitute an important class of distributions as they arise naturally in many applications. Since our applications are restricted to positive ID distributions, we will focus on these, which has the convenient side effect that the corresponding theory can be proven in a very elegant way. In Section 2.2.1, we will introduce the general theory and give a simple and elegant proof of the central Lévy–Khintchine representation. This representation is crucial as the resulting correspondence with Bernstein functions will be needed in the subsequent chapters. In Section 2.2.2, we will present well-known subclasses of positive ID distributions and famous examples.

2.2.1 General theory

We formulate the crucial property in terms of probability measures here, whereas in the literature it is sometimes also described using random variables. Most of this section follows along the lines of Schilling et al. (2010), while the proof of the Lévy–Khintchine representation also relies on (Steutel and van Harn, 2004, Chapter III).

Definition 2.1 (ID probability measures)
A probability measure μ is called ID if, for any $n \in \mathbb{N}$, there exists a probability measure μ_n such that

$$\mu_n^{n\star} := \underbrace{\mu_n \star \ldots \star \mu_n}_{n \text{ times}} = \mu. \tag{2.1}$$

ID distributions have been awarded a lot of attention, especially in the second half of the twentieth century. A central tool in this context is the characterization of these distributions via their characteristic functions. When restricting ourselves to positive ID distributions, the Laplace transform (see Widder (1946); Doetsch (1976) for the standard textbooks on the theory of Laplace transforms) assumes the central role of the characteristic function and there is a link to so-called Bernstein functions. We will derive the main result, relying almost exclusively on the powerful Bernstein's Theorem. To state it, the notion of complete monotonicity is needed.

Definition 2.2 (Completely monotone functions)
A function $g : (0, \infty) \to \mathbb{R}$ is a completely monotone (c.m.) function if $g \in C^\infty$ and

$$(-1)^n g^{(n)}(x) \geq 0, \quad \forall n \in \mathbb{N}_0, x > 0. \tag{2.2}$$

We will state Bernstein's Theorem in its general form for arbitrary measures.

Theorem 2.3 (Bernstein's Theorem)
Let $g : (0, \infty) \to \mathbb{R}$ be a c.m. function. Then it is the Laplace transform of a unique measure μ on $[0, \infty)$, i.e. for all $x > 0$,

$$g(x) = \hat{\mu}(x) = \int_{[0,\infty)} e^{-x\,u} \mu(\mathrm{d}u). \tag{2.3}$$

Conversely, whenever $\hat{\mu}(x) < \infty$ for every $x > 0$, $x \mapsto \hat{\mu}(x)$ yields a c.m. function.

Proof
For an elegant proof, see (Schilling et al., 2010, Theorem 1.4). □

Corollary 2.4 (Bernstein's Theorem for probability measures)
Let $g : (0, \infty) \to \mathbb{R}$ be a c.m. function with $g(0+) = 1$. Then it is the Laplace transform of a unique probability measure μ on $[0, \infty)$. Conversely, for every probability measure μ on $[0, \infty)$, $\hat{\mu}$ yields a c.m. function with $\hat{\mu}(0+) = 1$.

Proof
Follows easily from Theorem 2.3 and the observation that $\hat{\mu}(0+) = \mu\big([0, \infty)\big)$ via monotone convergence. □

Another important class of functions are so-called Bernstein functions, which will be used to characterize possible Laplace transforms of positive ID probability measures. For a whole book on this topic, see Schilling et al. (2010). We will adopt a slightly restricted definition.

Definition 2.5 (Bernstein functions)
A function $\Psi : (0, \infty) \to \mathbb{R}$ is a Bernstein function if $\Psi \in C^\infty, \Psi(x) \geq 0$ for all $x > 0$, $\Psi(0+) = 0$, and

$$(-1)^{n-1} \Psi^{(n)}(x) \geq 0, \quad \forall n \in \mathbb{N}, x > 0. \tag{2.4}$$

It admits a unique representation via

$$\Psi(x) = \xi\, x + \int_{(0,\infty)} \left(1 - e^{-x\,u}\right)\, \nu(\mathrm{d}u), \tag{2.5}$$

with $\xi \geq 0$, *the so-called drift term, and* ν *a measure on* $(0,\infty)$ *satisfying* $\int_{(0,\infty)}(1 \wedge u)\,\nu(\mathrm{d}u) < \infty$, *the so-called Lévy measure. Furthermore, for each such pair* (ξ,ν), *Equation* (2.5) *defines a Bernstein function.*

Remark 2.6
Note that Bernstein functions are usually defined without the restriction $\Psi(0+) = 0$, *which then yields an additional positive constant in the representation. This term corresponds to the related measures being sub-probability measures on* $[0,\infty)$, *respectively probability measures on* $[0,\infty]$, *having positive mass at* ∞. *When talking about corresponding Lévy subordinators, this is usually referred to as the killing of a subordinator.*

Remark 2.7
The representation of Bernstein functions in Definition 2.5 follows from applying Bernstein's Theorem (Theorem 2.3) to the first derivative of Ψ, *see (Schilling et al., 2010, Theorem 3.2).*

With these definitions at hand, we are now able to state the main characterization of positive ID distributions via their Laplace exponent, known as the Lévy–Khintchine formula, which is attributed to Lévy (1934) in the general case. As it is essential for our work and as it can be proven without requiring much more than Bernstein's Theorem, we give a proof here. It follows along the lines of the proof in (Steutel and van Harn, 2004, Chapter III).

Theorem 2.8 (Lévy–Khintchine representation)
A probability measure μ *on* $[0,\infty)$ *is ID if and only if*

$$\hat{\mu}(x) = e^{-\Psi(x)}, \quad x > 0, \tag{2.6}$$

with Ψ *a Bernstein function.*

Proof
Starting from a positive ID probability measure, one can define $\Psi(x) := -\log\left(\hat{\mu}(x)\right)$ *for* $x > 0$, *as* $\hat{\mu}(x) > 0$ *for all* $x > 0$. *From Corollary 2.4 follows that* $1 = \hat{\mu}(0+) \geq \hat{\mu}(x)$ *for all* $x > 0$ *and thus* $\Psi(0+) = 0, \Psi(x) \geq 0$ *for all* $x > 0$. *It remains to show that* Ψ' *is c.m., as can be seen from Equation* (2.4). *From* μ *being ID it follows that for every* $n \in \mathbb{N}$

there exists a positive measure μ_n *with* $\mu_n^{n\star} = \mu$ *or, equivalently,* $\hat{\mu}_n^n = \hat{\mu} = \exp(-\Psi)$. *One can rewrite*

$$e^{-\frac{1}{n}\Psi(x)} = \hat{\mu}_n(x), \quad x > 0,$$

which is thus c.m. as the Laplace transform of a positive distribution (Corollary 2.4). Considering the m-th convolution with $m \in \mathbb{N}$, *one can conclude that* $\exp(-(m/n)\Psi)$ *is c.m. for arbitrary* $m, n \in \mathbb{N}$ *and thus* $\exp(-q\Psi)$ *is c.m. for any* $q \in \mathbb{Q}_+$. *As every* $t > 0$ *can be written as the limit of a sequence of rational numbers,* $\exp(-t\Psi)$ *is c.m. for every* $t > 0$ *as the pointwise limit of c.m. functions, see the continuity property (Schilling et al., 2010, Corollary 1.7) or (Steutel and van Harn, 2004, Proposition A.3.7(iv)). The result could be derived using (Schilling et al., 2010, Theorem 3.6 (iii)* $\Rightarrow$ *(i)) or alternatively, using the continuity property again, since*

$$\Psi'(x) = \lim_{t \searrow 0} -\frac{1}{t}\frac{\partial}{\partial x}\hat{\mu}^t(x) = \lim_{t \searrow 0} -\frac{1}{t}\frac{\partial}{\partial x}e^{-t\Psi(x)},$$

representing Ψ' *as the pointwise limit of c.m. functions. For this, one has to observe that* $-\frac{\partial}{\partial x}\exp(-t\Psi(x))$ *is c.m. as the derivative of a c.m. function multiplied by* (-1).

Conversely, starting from a Bernstein function Ψ, *it follows that*

$$g_t(x) := e^{-t\Psi(x)}, \text{ is c.m. for each } t > 0 \text{ with } g_t(0+) = 1.$$

This is proven in (Schilling et al., 2010, Theorem 3.6 (i) $\Rightarrow$ *(iii)) or (Steutel and van Harn, 2004, Proposition A.3.7(vi)), showing that the composition of a c.m. function (here* $\exp(-t\,x)$*) and a Bernstein function is always c.m.. Using Corollary 2.4, we can define for each* $n \in \mathbb{N}$ *a probability measure* μ_n *corresponding to* $\hat{\mu}_n = g_{1/n} = \exp(-1/n\,\Psi)$. *For this, it follows that*

$$\mu_n^{n\star} = \mu, \text{ via } \hat{\mu}_n^n = \hat{\mu} = \exp(-\Psi),$$

proving the claim. □

Remark 2.9
The crucial result for the derivation is again Bernstein's Theorem, using its representation in Corollary 2.4. Apart from that, one only needs two smaller results, namely the continuity property (Schilling et al., 2010, Corollary 1.7), stating that the pointwise limit of a c.m. function is again c.m., and (Schilling et al., 2010, Theorem 3.6 (i) $\Rightarrow$ *(iii)),*

stating that the composition of a c.m. function and a Bernstein function is again c.m.. In total, one can observe that the proof is significantly easier than for ID distributions on the real line, where an important tool is the approximation of distributions via compound Poisson distributions, see, e.g., (Sato, 1999, Chapter 2). For a more probabilistic proof of Theorem 2.8, see, e.g., (Bertoin, 1999, Theorem 1.2).

Remark 2.10

As the Laplace transform $\hat{\mu}$ can also be defined for $x = 0$ and as Ψ can be continuously extended via $\Psi(0) := \Psi(0+) = 0$, we will in the following often consider Bernstein functions as functions on $[0, \infty)$.

2.2.2 Classes of positive infinitely divisible distributions

Having introduced the Lévy–Khintchine representation, we are now able to define several classes of positive ID distributions via the pair (ξ, ν) in Equation (2.5) describing the corresponding Bernstein function. As the drift term ξ only represents a constant additive term, it will be ignored in the following. Denoting by $\mathcal{M}$ the set of all Lévy measures ν, we can define the following subsets:

$$\begin{aligned} U &:= \{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = g(x)\mathrm{d}x,\ g \text{ non-increasing}\}\,, \\ L &:= \{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = g(x)\mathrm{d}x,\ x\,g(x) \text{ non-increasing}\}\,, \\ BO &:= \{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = g(x)\mathrm{d}x,\ g \text{ c.m.}\}\,, \\ T &:= \{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = g(x)\mathrm{d}x,\ x\,g(x) \text{ c.m.}\}\,. \end{aligned}$$

The class of distributions corresponding to U is called "Jurek class" or class of "s-self-decomposable distributions", see Jurek (1985), restricted to distributions on $\mathbb{R}_+$. They have a representation as limit distributions of sums of "shrunken" random variables, see (Jurek, 1985, Remark 2.1). For a couple of equivalent conditions in terms of the Lévy measure, see (Jurek, 1985, Theorem 2.2). All other defined classes are subclasses of U, thus, the following Lemma applies to all of them. It will be helpful in Chapter 3, where we are concerned with the numerical computation of probability densities.

Lemma 2.11 (Absolute continuity - I)

A distribution of the Jurek class is absolutely continuous with respect to the Lebesgue measure if and only if

$$\int_{(0,\infty)} g(x)\mathrm{d}x = \infty.$$

Proof

The first direction follows from (Steutel and van Harn, 2004, Proposition III.4.16). Conversely, if $\int_{(0,\infty)} g(x)\mathrm{d}x < \infty$ *holds, it is easy to recognize the corresponding distribution as a compound Poisson distribution which has positive mass at* $\{0\}$ *(respectively* $\{\xi\}$*).*$\square$

The class of distributions corresponding to L is called class of "selfdecomposable distributions", again restricted to distributions on $\mathbb{R}_+$. Their investigation has a long tradition, for a nice account of which see, e.g., (Sato, 1999, p. 117f). An alternative definition of this class can be given as limit distributions of sums of independent random variables forming a null array, see (Sato, 1999, Theorem 15.3). They constitute a subclass of U, as from $x\,g(x)$ non-increasing it follows that g, being the product of a non-increasing function and $1/x$, is non-increasing as well.

We will call class T the "Thorin class" as it was first investigated and introduced by Thorin (1977a,b) as the smallest class of distributions closed under convolution and (weak) convergence which contains all Gamma distributions. The definition gives rise to the alternative name "generalized Gamma convolutions". The corresponding Bernstein functions are called Thorin–Bernstein functions, see (Schilling et al., 2010, Chaper 8). Famous examples of this class are the Pareto, Weibull, log-normal, and F-distribution as is shown, e.g., in Bondesson (1979), who proves that probability densities of a specific form correspond to distributions of the Thorin class.

The class BO, called the "Bondesson class", is the most relevant one for us as Chapter 3 is devoted to the numerical computation of the density of distributions of this class. This class has been considered for the first time in Bondesson (1981) under the name generalized convolutions of mixtures of exponential distributions ("g.c.m.e.d."). Probabilistically, it can be introduced as the smallest class of distributions closed under (weak) convergence and convolution containing mixtures of exponential distributions, see (Sato, 1999, Definition 51.9). The corresponding Bernstein functions are called complete Bernstein functions, see (Schilling et al., 2010, Chapter 6), and, using Bernstein's Theorem for the densities of their Lévy measures, they allow for an alternative representation via

$$\Psi(x) = \xi\, x + \int_0^\infty \frac{x}{x+t}\,\sigma(\mathrm{d}t), \quad x \geq 0, \tag{2.7}$$

with σ a measure on $(0,\infty)$ satisfying $\int_{(0,\infty)} 1/(1+t)\,\sigma(\mathrm{d}t) < \infty$, called the Stieltjes measure, see (Schilling et al., 2010, Theorem 6.2(ii)). Conversely, if a representation with a Stieltjes measure as given in Equation (2.7) exists, Ψ can be shown to be a complete Bernstein function. The relation between the measures ν and σ is given by $g(s) = \int_{(0,\infty)} \exp(-t\,s)\,t\,\sigma(\mathrm{d}t)$. We can state Lemma 2.11 in terms of the Stieltjes measure, which coincides with (Bondesson, 1981, Theorem 6.1).

Corollary 2.12 (Absolute continuity - II)
A distribution of the Bondesson class is absolutely continuous with respect to the Lebesgue measure if and only if $\sigma\big((0,\infty)\big) = \infty$.

Proof
It is enough to observe that from the stated relation between the measures ν and σ follows the equality $\int_{(0,\infty)} g(x)\mathrm{d}x = \int_{(0,\infty)} \sigma(\mathrm{d}t)$. The rest follows from Lemma 2.11. $\square$

An easy way to derive a new Bondesson distribution from a given Bondesson distribution is provided by exponential tilting (see Rosiński (2007)). The resulting distribution can be defined via its Lévy measure or via its probability density. It is also of the Bondesson class, as the corresponding Lévy density is derived from the initial Lévy density g by $\tilde{g}(s) = \exp(-h\,s)\,g(s)$ with $h > 0$ and the product of two completely monotone functions is again completely monotone, see (Schilling et al., 2010, Corollary 1.6). The Bernstein function of the exponentially tilted distribution is given by $\tilde{\Psi}(x) := \Psi(x+h) - \Psi(h)$ and the probability density is given by multiplying the original density with $\exp(-h\,s+\Psi(h))$. For the same reason, replacing $\exp(-h\,s)$ by an arbitrary completely monotone function q with $q(0+) = 1$, proper tempered distributions as investigated in Rosiński (2007), restricted to our (positive, one-dimensional) setting, are also part of the Bondesson class.

For an illustration of the connection between all four classes of distributions see the so-called Venn diagram in Figure 2.1, which is borrowed from (Schilling et al., 2010, p. 90). Integral representations for all classes presented will be given in Remark 2.29 in the next section. We end this section with a list of distributions of the Bondesson class, on the one hand to illustrate how large it is and on the other hand as they will be needed in Chapter 3. For the most part, it is along the lines of the introduction of Bernhart et al. (2015a).

Positive stable distribution

The class of positive, strictly stable distributions can be characterized by $\Psi^{\mathrm{St}}_{\alpha,\beta}(x) = \beta\,x^{\alpha}, 0 < \alpha < 1, \beta > 0$. It is obvious that $\Psi^{\mathrm{St}}_{\alpha,\beta}$ fulfills the defining properties of a Bern-

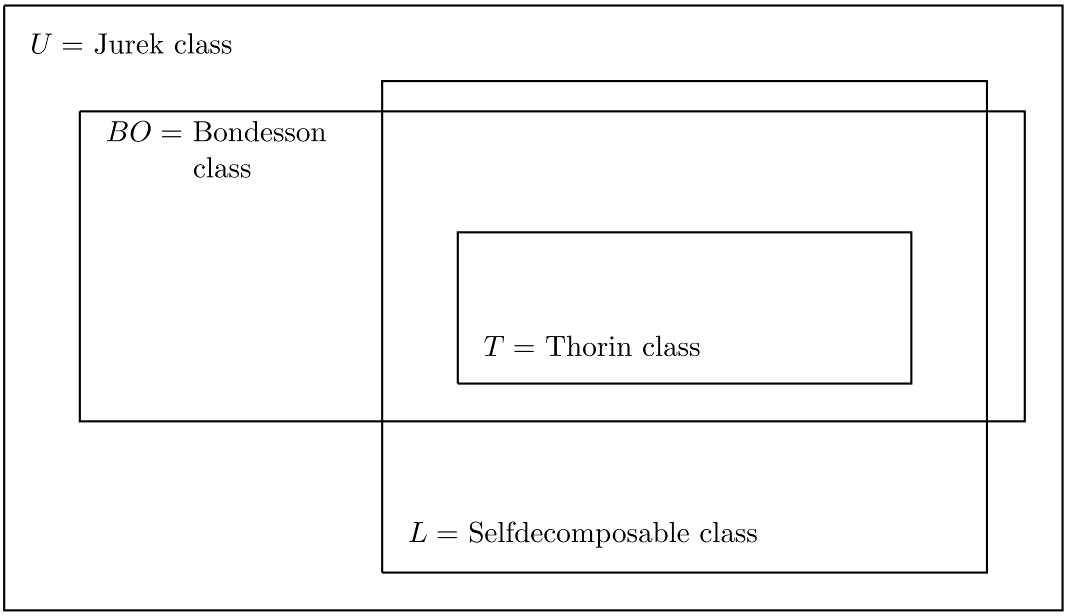


Figure 2.1 Venn diagram illustrating the relation between the different classes of positive ID distributions defined on page 11.

stein function in Definition 2.5 and one can show that it allows for the representations

$$\Psi^{\mathrm{St}}_{\alpha,\beta}(x) = \beta \int_0^\infty (1 - e^{-s\,x})\, \frac{\alpha\, s^{-1-\alpha}}{\Gamma(1-\alpha)}\, \mathrm{d}s = \beta \int_0^\infty \frac{x}{x+s}\, \frac{\sin(\alpha\,\pi)}{\pi}\, s^{\alpha-1}\, \mathrm{d}s, \quad x \geq 0.$$

Consequently (those results are found in (Schilling et al., 2010, Chapter 15)),

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

It is known that the related probability density $f^{\mathrm{St}}_{\alpha,\beta}$ is a C^∞-function, see (Nolan, 2012, Theorem 1.9), though the latter is in general not known in closed form. This class of distributions is of great importance in physics, see Penson and Górska (2010) and the many references therein. In that paper, a closed-form expression for the density for rational numbers $\alpha = k/l, k \leq l$, is derived. This expression, however, is stated in terms of hypergeometric functions which are not easy to evaluate numerically. The importance of this distribution in finance results from the fact that it represents one of the standard examples of heavy-tailed distributions. For instance, $\int_0^\infty x^c \mu(\mathrm{d}x) = \infty$ for $c > \alpha$, see Wolfe (1975), i.e., intuitively speaking, the tails are so heavy that not even the expected value exists (as $\alpha < 1$). A famous approach for deriving the density for general stable distributions is presented in Nolan (1997), which is based on a distribution-specific contour transformation. Using the relation

$$f^{\mathrm{St}}_{\alpha,\beta}(x) = \beta^{-\frac{1}{\alpha}} \, f^{\mathrm{St}}_{\alpha,1}\left(\beta^{-\frac{1}{\alpha}} \, x\right), \tag{2.8}$$

the formula can be written as

$$\begin{aligned} f^{\mathrm{St}}_{\alpha,1}(x) &= \mathbb{1}_{\{x>0\}} \, \frac{\alpha \left(\frac{x}{\gamma}\right)^{\frac{1}{\alpha-1}}}{\gamma \, \pi \, (1-\alpha)} \int_{-\pi/2}^{\pi/2} g_\alpha(u) \, e^{-\left(\frac{x}{\gamma}\right)^{\frac{\alpha}{\alpha-1}} g_\alpha(u)} \, \mathrm{d}u, \qquad (2.9)\\ g_\alpha(u) &= \Big(\cos\big(\frac{\pi \, \alpha}{2}\big)\Big)^{\frac{1}{\alpha-1}} \Big(\frac{\cos u}{\sin(\alpha(\frac{\pi}{2}+u))}\Big)^{\frac{\alpha}{\alpha-1}} \, \frac{\cos(\frac{\pi}{2}\alpha + (\alpha-1)u)}{\cos u},\\ \gamma &= \Big(\cos\big(\frac{\pi \, \alpha}{2}\big)\Big)^{\frac{1}{\alpha}}. \end{aligned}$$

Alternatively, a series representation is known, which, according to Penson and Górska (2010), was derived by Humbert (1945). For an impression of the general shape of the stable density for different values of α, Figure 2.2 is included.

Furthermore, so-called exponentially tilted stable distributions (see Barndorff-Nielsen and Shephard (2001)) with Laplace exponent $\Psi^{\mathrm{tSt}}_{\alpha,\beta,h}(x) = \beta\,((x+h)^\alpha - h^\alpha) = \Psi^{\mathrm{St}}_{\alpha,\beta}(x+h) - \Psi^{\mathrm{St}}_{\alpha,\beta}(h)$ are also in the Bondesson class with density $f^{\mathrm{tSt}}_{\alpha,\beta,h}(x) = f^{\mathrm{St}}_{\alpha,\beta}(x)\,\exp(-h\,x + \Psi^{\mathrm{St}}_{\alpha,\beta}(h))$. The same holds true for proper tempered stable distributions as introduced in Rosiński (2007), restricted to our setting.

Inverse Gaussian distribution

The Inverse Gaussian (IG) distribution (for a detailed introduction, see Seshadri (1993)) constitutes another famous Bondesson distribution. It is characterized by the Bernstein function

$$\Psi^{\mathrm{IG}}_{\beta,\eta}(x) = \beta\,(\sqrt{2x+\eta^2} - \eta), \quad \beta, \eta > 0, x \geq 0.$$

The corresponding density is known and given by

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

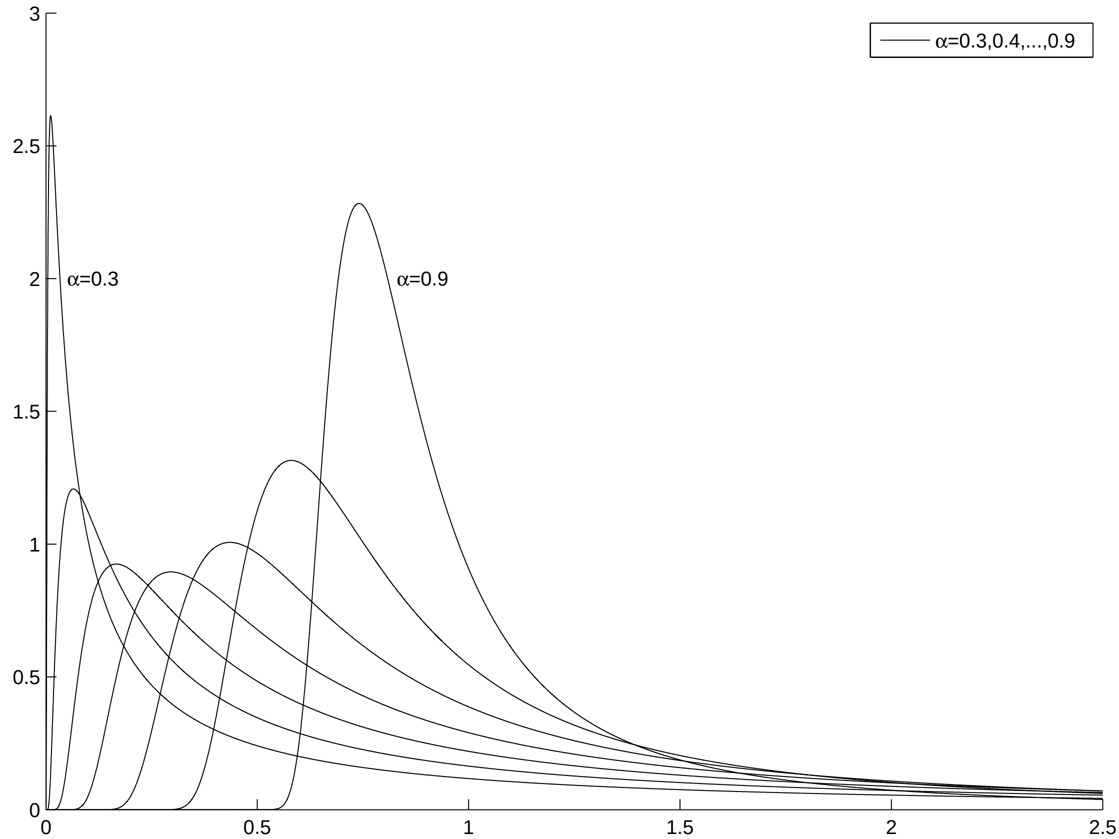


Figure 2.2 Density of the stable distribution for different values of α, computed with the approach presented in Chapter 3.

Shuster (1968) derives a closed-form expression for its cumulative distribution function in terms of the standard normal cumulative distribution function. The name of this distribution is motivated by the fact that a random variable with such a distribution can be constructed via

$$\inf\{t > 0 : \eta\, s + W_s = \beta\},$$

with $W = \{W_s\}_{s \geq 0}$ a standard Brownian motion, see, e.g., (Applebaum, 2004, p. 51)[1]. One can observe that this distribution represents a specific example of an exponentially tilted stable distribution, with $\alpha = 0.5$, $h = \eta^2/2$, ($\alpha = 0.5$ is one of the few examples where the stable probability density has a simple form) and the related measures can thus be derived from the previous paragraph:

$$\nu(\mathrm{d}s) = \sqrt{2}\,\beta\, \frac{1}{2\,\Gamma(0.5)}\, s^{-1.5}\, e^{-\frac{\eta^2}{2} s}\, \mathrm{d}s, \quad \sigma(\mathrm{d}s) = \sqrt{2}\,\beta\, \frac{\sin(0.5\,\pi)}{\pi}\, \frac{(s - \eta^2/2)^{0.5}}{s}\, \mathbb{1}_{\{s > \eta^2/2\}}\, \mathrm{d}s.$$

Gamma distribution

As another example with known density, the Gamma distribution is presented. It is characterized by the Bernstein function

$$\Psi^{\mathrm{Ga}}_{\beta,\eta}(x) = \beta\, \log(1 + \frac{x}{\eta}), \quad \beta, \eta > 0, x \geq 0,$$

and its density equals

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

The related measures are (see (Schilling et al., 2010, Chapter 15))

$$\nu(\mathrm{d}s) = \beta\, \frac{e^{-\eta\, s}}{s}\, \mathrm{d}s, \quad \sigma(\mathrm{d}s) = \beta\, \frac{1}{s}\, \mathbb{1}_{\{s > \eta\}}\, \mathrm{d}s,$$

where the expression for σ can easily be verified by computing the related integral and the expression for ν follows from the given relation between the two measures. The easiest way to notice that $\Psi^{\mathrm{Ga}}_{\beta,\eta}$ is indeed a Bernstein function is via the given Stieltjes representation or via the observation that $\Psi'^{\mathrm{Ga}}_{\beta,\eta} = \beta/(\eta + x)$.

[1] Actually, it is shown there that the corresponding Lévy subordinator (see Lemma 2.17 below) can be constructed via $\Lambda_t := \inf\{t > 0 : \eta\, s + W_s = \beta\, t\}$.

Non-central χ^2 distribution

The non-central chi-squared respectively χ^2 distribution appears very often in the literature, in statistics as well as in various financial applications. Its relevance in finance stems from its relation with so-called Bessel processes. Because of this, it will also show up in Chapter 5. The non-central χ^2 distribution with β degrees of freedom and non-centrality parameter η, sometimes abbreviated by $\chi^2(\beta,\eta)$, can be defined in terms of its probability density, which is given by

$$f^{\chi^2}_{\beta,\eta}(x) = \sum_{i=0}^{\infty} \frac{e^{-\eta/2}\,(\eta/2)^i}{i!} f^{\mathrm{Ga}}_{\beta/2+i,1/2}(x), \quad \beta,\eta,x>0. \tag{2.10}$$

One can recognize that this distribution may also be regarded a Poisson mixture of Gamma distributions.[2] Its importance in statistics stems from its relation to the normal distribution, as, given $X_1,\ldots,X_k$ independent random variables with $X_i \sim \mathcal{N}(m_i,1)$, $m_i>0, i=1,\ldots,k$, the random variable $\sum_{i=1}^k X_i^2$ has a non-central χ^2 distribution with $\beta=k$ and $\eta=\sum_{i=1}^k m_i^2$. For other stochastic representations and a nice account of its appearances in mathematical finance, see Mai (2014b). The related distribution function can be easily computed from its density, yielding

$$F^{\chi^2}_{\beta,\eta}(x) = \int_{[0,x]} f^{\chi^2}_{\beta,\eta}(s)\,\mathrm{d}s = \sum_{i=0}^{\infty} \frac{e^{-\eta/2}\,(\eta/2)^i}{i!} \frac{\gamma(\beta/2+i,x/2)}{\Gamma(\beta/2+i)}, \quad x>0. \tag{2.11}$$

There are many alternative representations for this expression, see, e.g., Larguinho et al. (2013). Furthermore, the corresponding Laplace transform can be computed from the density, using the results for Gamma distributions, as

$$\hat{\mu}(x) = e^{-\Psi^{\mathrm{Ga}}_{\beta/2,1/2}(x)}\, e^{-\frac{\eta\,x}{1+2\,x}} = e^{-\Psi^{\mathrm{Ga}}_{\beta/2,1/2}(x)-\frac{\eta\,x}{1+2\,x}}, \quad x\geq 0. \tag{2.12}$$

[2] Another very interesting class of Bondesson distributions can be defined by the similar mixture

$$f_{\beta,\eta}(x) = \sum_{i=0}^{\infty} \frac{e^{-\beta/2}\,(\beta/2)^i}{i!} f^{\mathrm{Ga}}_{\beta/2+i,\eta/2}(x), \quad \beta,\eta,x>0.$$

This class of distributions yields the pre-image of Gamma distributions under the integral transform Φ_{f_1} introduced below in Remark 2.29. As Gamma distributions are part of T and thus also L, Remark 4.3 applies, which shows that all these distributions have a representation via the convolution of $\mathcal{L}(\Lambda_1)$ and $\mathcal{L}(\int_0^\infty \exp(-s)\,\mathrm{d}\Lambda_s)$, with Λ a compound Poisson subordinator with exponentially distributed jumps.

Consequently, the Bernstein function $\Psi^{\chi^2}_{\beta,\eta}$ corresponding to a non-central χ^2 distribution is given by

$$\Psi^{\chi^2}_{\beta,\eta}(x) = \Psi^{\mathrm{Ga}}_{\beta/2,1/2}(x) + \frac{\eta\, x}{1+2\,x}, \quad \beta,\eta > 0, x \geq 0.$$

The second summand can be easily recognized as a complete Bernstein function stating its Stieltjes measure $\sigma(\mathrm{d}s) = \eta/2\,\delta_{1/2}(\mathrm{d}s)$, with $\delta_{1/2}(\mathrm{d}s)$ denoting the Dirac measure centered around $1/2$.[3] Consequently, the given expression is indeed a (complete) Bernstein function as these are stable under summation. For reasons of completeness, we state the related measures

$$\nu(\mathrm{d}s) = \left(\beta\,\frac{e^{-s/2}}{2\,s} + \frac{\eta}{4}\,e^{-s/2}\right)\mathrm{d}s, \quad \sigma(\mathrm{d}s) = \frac{\beta}{2\,s}\,\mathbb{1}_{\{s>1/2\}}\,\mathrm{d}s + \eta/2\,\delta_{1/2}(\mathrm{d}s).$$

Hartman–Watson distribution

Finally, we present the Hartman–Watson distribution. The Hartman–Watson distribution arises as the first hitting time of certain diffusion processes, see Kent (1982), and is of paramount interest in mathematical finance in the context of Asian option pricing, see Yor (1992); Barrieu et al. (2004); Gerhold (2011). This is due to the fact that it plays an important role when investigating certain exponential functionals of Brownian motion. The distribution can be defined via its Laplace transform $\hat{\mu}(x) := I_{\sqrt{2\,x}}(r)/I_0(r)$, $r, x > 0$, where $I_\nu(z)$ is called the modified Bessel function of the first kind. For $z, \nu \in \mathbb{C}$, $I_\nu(z)$ is introduced as the solution to a differential equation, has a series representation[4] via

$$I_\nu(z) := \left(\frac{z}{2}\right)^\nu \sum_{m=0}^{\infty} \frac{1}{m!\,\Gamma(m+\nu+1)} \left(\frac{z}{2}\right)^{2\,m}, \tag{2.13}$$

and for each $z \neq 0$, it is an entire function in ν (see (Abramowitz and Stegun, 1965, p. 374f)). It was shown in Hartman (1976) that this distribution is infinitely divisible. Moreover, it follows from a result in Kent (1982) that it is also part of the Bondesson

[3] Actually, this summand represents the Bernstein function corresponding to a compound Poisson distribution with Exp(1/2)-distributed jumps and intensity $\eta/2$.

[4] We restrict ourselves to $\nu \notin -1, -2, \ldots$, to avoid confusion about the definition of the Gamma function.

class. The corresponding complete Bernstein function is consequently given by

$$\Psi_r^{\mathrm{HW}}(x) = -\log\left(I_{\sqrt{2x}}(r)/I_0(r)\right), \quad r > 0, x \geq 0.$$

However, a closed-form expression of the corresponding Lévy density is, to the best of our knowledge, unknown. The same holds true for the related probability density f_r^{HW}, for which Yor (1992) (result originally derived in Yor (1980)) states an integral representation via $f_r^{\mathrm{HW}}(x) = \theta(r,x)/I_0(r)$, with

$$\theta(r,x) := \frac{r\, e^{\frac{\pi^2}{2x}}}{\sqrt{2\pi^3 x}} \int_0^\infty e^{-\frac{y^2}{2x} - r\cosh(y)} \sinh(y) \sin\left(\frac{\pi y}{x}\right) \mathrm{d}y. \tag{2.14}$$

Figure 2.3 illustrates the general form of the density of the Hartman–Watson law.

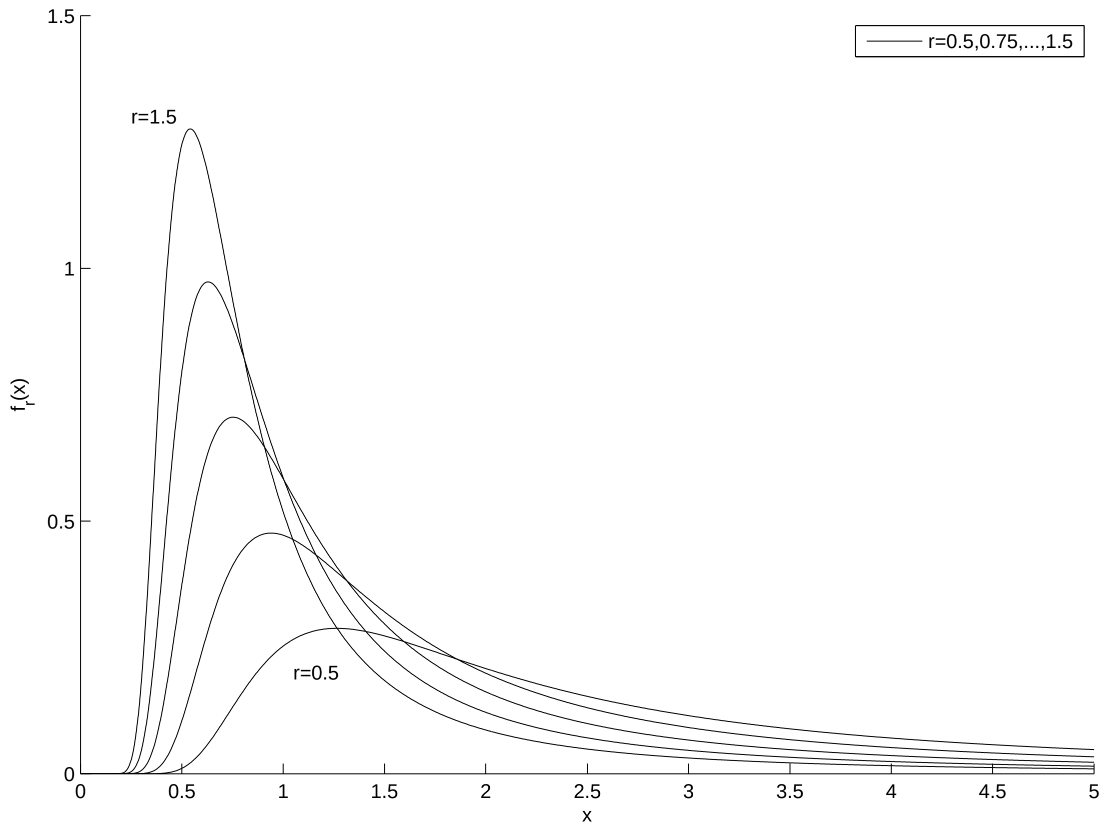


Figure 2.3 Densities of the Hartman–Watson distribution for several values of r, computed with the approach presented in Chapter 3.

2.3 IDT subordinators

IDT subordinators constitute an interesting class of stochastic processes and we will introduce them in this section. The acronym IDT is motivated by the term "infinitely divisible with respect to time".[5] The reason for investigating IDT subordinators in the present thesis is the paper Mai and Scherer (2014), respectively the link between these processes and certain multivariate dependence models established therein. This link will be introduced in Section 2.4.4, and in Chapter 4 we will make use of this connection. In the present section, we define IDT subordinators (Section 2.3.1), investigate their properties (Section 2.3.2), and present a possible construction of large families of IDT subordinators (Section 2.3.4), for which the definition of an integral with respect to Lévy subordinators (Section 2.3.3) is needed.

2.3.1 Definition of IDT subordinators

To the best of our knowledge, IDT processes have been investigated only in few recent papers (see, e.g., Mansuy (2005); Es-Sebaiy and Ouknine (2007); Hakassou and Ouknine (2011)), with Mansuy (2005) being the first work defining them explicitly. It is sometimes differentiated between "strong" and "weak" IDT processes, but here we will restrict our attention to strong IDT processes and omit the prefix strong from here on.

Definition 2.13 (IDT process)
A stochastic process $H = \{H_t\}_{t\geq 0}$ is called an IDT process if it satisfies the condition

$$\{H_t\} \overset{d}{=} \{H^{(1)}_{t/n} + \ldots + H^{(n)}_{t/n}\}, \quad \forall n \in \mathbb{N}, \tag{2.15}$$

where $H^{(1)}, \ldots, H^{(n)}$ are independent copies of H and equality in distribution of processes means equality of all finite-dimensional marginal distributions.

We define IDT subordinators simply as non-decreasing, càdlàg IDT processes starting at 0. The name is motivated by the idea of subordination, where subordination is the transformation of a process by a random (non-decreasing) time change with a process independent of the original process. Originally, the non-decreasing processes used were Lévy subordinators, but here we consider a broader class. The concept of subordination

[5] I became aware of the definition of this class for the first time in an early version of Mai and Scherer (2014) and it caught my attention because of its simplicity. At that point, we did not even know if this class of processes existed in the literature already.

is attributed to Bochner (1949) and it was first used in finance in Clark (1973). Another name for those processes that can be found, e.g., in Barndorff-Nielsen et al. (2006a), is chronometers, again motivated by the idea of time-change.

Definition 2.14 (IDT subordinator)
An a.s. non-decreasing, càdlàg IDT process $H = \{H_t\}_{t\geq 0}$ with $H_0 = 0$, a.s., is called an IDT subordinator.

Remark 2.15
Mai and Scherer (2014) define IDT subordinators with the additional condition that it approaches infinity, a.s.. Their motivation for this is the following: Starting from an exponentially distributed random variable $E \sim \mathrm{Exp}(1)$ independent of an IDT subordinator H, the property "$E < H_t$ for at least one $t > 0$, a.s.," is needed. We will show later on that this is fulfilled for every IDT subordinator except the trivial case $H \equiv 0$. Furthermore, they allow the process to take values in $[0, \infty]$. We will ignore this aspect in our exposition of IDT subordinators as we do not need it for our results, and as the general literature on those processes does not include it. We will explicitly mention this property when stating the main result of Mai and Scherer (2014).

The most famous subclass of IDT processes, respectively IDT subordinators, is given by Lévy processes, respectively Lévy subordinators. For completeness, we recall their definition (for a textbook account of those processes, see, e.g., Sato (1999)) in the one-dimensional case, see (Sato, 1999, Definition 1.6).

Definition 2.16 (Lévy processes and subordinators)
A stochastic process $X = \{X_t\}_{t\geq 0}$ on $\mathbb{R}$ is a Lévy process if the following conditions are satisfied:

(1) For any $0 \leq t_0 < t_1 < \ldots < t_n, n \in \mathbb{N}$, the increments $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent random variables.

(2) $X_0 = 0$, a.s..

(3) For every $t \geq 0$ and $\epsilon > 0$, $\lim_{s\to t} \mathbb{P}\left(|X_s - X_t| > \epsilon\right) = 0$, i.e. X_s converges to X_t in probability for $s \to t$.

(4) $\mathcal{L}(X_{s+t} - X_s)$ does not depend on s, so the increments are homogeneous.

(5) There is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$, such that for all $\omega \in \Omega_0$, $X_t(\omega)$ is càdlàg, which means the process is càdlàg a.s..

A non-decreasing Lévy process is called a Lévy subordinator.

We will always denote Lévy subordinators by Λ and assume that for every $\omega \in \Omega$, $\Lambda_t(\omega)$ is càdlàg, non-decreasing, and $\Lambda_0(\omega) = 0$ (see, e.g. (Sato, 1999, p. 197)), which basically is a convenient indistinguishable modification of a Lévy subordinator. We will not investigate Lévy subordinators in detail here, all relevant properties will be proven in the next section for the superclass of IDT subordinators. However, representing the only well-known subclass, they are crucial for our understanding of IDT subordinators and we will need some of their properties to emphasize the differences. The necessary properties will be stated without proof in the following lemma.

Lemma 2.17 (Uniqueness marginal distribution)
For every ID distribution μ on $\mathbb{R}$, there is a Lévy process $X = \{X_t\}_{t\geq 0}$ with $\mathcal{L}(X_1) = \mu$. It is unique up to identity in law. In particular, X is a subordinator if and only if $\mu\big((-\infty, 0]\big) = 0$, i.e. if μ is a positive ID distribution.

Proof
The first statement follows from (Sato, 1999, Corollary 11.6), the second from (Sato, 1999, Theorem 24.11). $\square$

We start proving that Lévy subordinators indeed are IDT subordinators, which is well known and easy to proof.

Lemma 2.18 (Lévy subordinators as IDT subordinators)
A Lévy subordinator is an IDT subordinator.

Proof
Actually, Lévy processes are IDT processes, as Condition (2.15) *can be shown to hold for all Lévy processes using only basic properties: Starting from a Lévy process X and for $n \in \mathbb{N}$ considering iid copies $X^{(1)}, \ldots, X^{(n)}$, it is clear from Definition 2.16 that $X^{(1)}_{t/n} + \ldots + X^{(n)}_{t/n}$ representing the sum of independent Lévy processes is again a Lévy process. Furthermore,*

$$\begin{aligned} X^{(1)}_{1/n} + \ldots + X^{(n)}_{1/n} &\overset{d}{=} X^{(1)}_{1/n} + X^{(2)}_{2/n} - X^{(2)}_{1/n} + \ldots + X^{(n)}_{n/n} - X^{(n)}_{(n-1)/n} \\ &\overset{d}{=} X^{(1)}_{1/n} + X^{(1)}_{2/n} - X^{(1)}_{1/n} + \ldots + X^{(1)}_{n/n} - X^{(1)}_{(n-1)/n} \overset{d}{=} X_1, \end{aligned}$$

using only independence and homogeneity of increments. From this, Condition (2.15) *follows by the uniqueness of the one-dimensional marginal distribution stated in Lemma 2.17. The claim follows from the fact that Lévy subordinators fulfill all remaining properties by definition.* $\square$

Lemma 2.17 shows that a Lévy process is uniquely determined by its one-dimensional marginal distributions and in particular from positive marginal distributions it follows that it is non-decreasing. While the first property does not hold for general IDT processes, it is not clear on first sight if non-decreasingness follows from positiveness. The following example shows that it does not.

Example 2.19
Let $\{\Lambda_t\}_{t\geq 0}$ denote a Lévy subordinator. Define a process $\{Y_t\}_{t\geq 0}$ via $Y_t := \Lambda_{a\,t} - \Lambda_t$ with $a > 1$. This yields an IDT process which follows directly from the IDT property of a Lévy subordinator. It has positive one-dimensional marginals, but it is not non-decreasing, which can be easily seen when considering, e.g., a simple Poisson process. Furthermore, it can be already seen from this simple example that IDT subordinators do not necessarily possess independent increments.

Instead of subtracting, one could add Λ_t in Example 2.19, which then would yield an IDT subordinator. This was actually the starting point for the development of the idea we will present in Section 2.3.4, where IDT subordinators are constructed from the increments of a Lévy subordinator. We present another well-known example which is of quite pathological structure.

Example 2.20
Let M denote a positive stable random variable, i.e. a random variable with Bernstein function $\Psi(x) = \Psi^{St}_{\alpha,1}(x) = x^{\alpha}$, see Section 2.2.2. Then it is well known, see (Mansuy, 2005, Example 2.1), (Es-Sebaiy and Ouknine, 2007, Example 3.1), or (Mai and Scherer, 2014, Example 3.2), that the definition $H_t := t^{1/\alpha}\, M, t \geq 0$, yields an IDT subordinator.

In the next section, important properties of IDT subordinators are investigated to provide the necessary insights for the following considerations.

2.3.2 Properties of IDT subordinators

There are many interesting properties of IDT subordinators stated in the literature, and many links to other classes of processes are established. However, we will focus on the essential aspects only. The most obvious property, which also represents the connection to Section 2.2, is the infinite divisibility of the margins of an IDT subordinator.

Lemma 2.21 (Laplace transform)
Let H be an IDT subordinator. H_t is positive ID, for all $t \geq 0$. With Ψ the Bernstein function corresponding to H_1, one has

$$\mathbb{E}\left[e^{-x\,H_t}\right] = e^{-t\,\Psi(x)}, \quad t \geq 0, x > 0. \tag{2.16}$$

Proof
The infinite divisibility follows easily from Condition (2.15)*. It is easy to prove Equation* (2.16) *for $t \in \mathbb{Q}_+$. For arbitrary $t \geq 0$, the result follows from the fact that H is a.s. continuous from the right, for a more detailed proof see (Mai and Scherer, 2014, Lemma 3.7).* □

Remark 2.22
Note that stochastic continuity of the process is not needed in the proof of Lemma 2.21 and that it is not stated as a necessary property of an IDT subordinator. Instead, it follows naturally from the càdlàg and non-decreasing property and Condition (2.15)*.*

Remark 2.23
From Lemma 2.21 follows that the one-dimensional marginal distributions of an IDT subordinator H equal the one-dimensional marginal distributions of the Lévy subordinator Λ with $\mathcal{L}(\Lambda_1) = \mathcal{L}(H_1)$, which exists according to Lemma 2.17. This is a result stated quite often in the literature, see, e.g., (Mansuy, 2005, Proposition 4.1).

Not only the one-dimensional marginal distributions of an IDT subordinator are infinitely divisible, but all finite-dimensional marginal distributions are. This can be easily seen from the fact that for any $n \in \mathbb{N}$,

$$\begin{aligned}(H_{t_1}, \ldots, H_{t_n}) &\stackrel{d}{=} \left(\sum_{i=1}^{n} H^{(i)}_{t_1/n}, \ldots, \sum_{i=1}^{n} H^{(i)}_{t_n/n}\right)\\ &\stackrel{d}{=} \sum_{i=1}^{n} \left(H^{(i)}_{t_1/n}, \ldots, H^{(i)}_{t_n/n}\right).\end{aligned}$$

Consequently, IDT subordinators are a subclass of ID processes as defined, e.g., in Maruyama (1970); Barndorff-Nielsen et al. (2006a).[6] In particular, they represent a subclass of chronometers and of infinitely temporally selfdecomposable processes as defined in Barndorff-Nielsen et al. (2006a). The latter fact was already observed in (Mansuy, 2005, Proposition 6.2).

[6]Via that observation we originally discovered the existence of IDT processes in the literature.

Futhermore, it is interesting to note that the scaling property in Equation (2.16) also holds for the multivariate Laplace transform of multidimensional marginals of H, see, e.g., (Mai and Scherer, 2014, proof of Theorem 5.3), which actually represents a condition equivalent to Equation (2.15).

Lemma 2.24 (Alternative characterization)
Consider a non-decreasing, right-continuous process $\{H_t\}_{t\geq 0}$ with $H_0 = 0$, a.s.. H is an IDT subordinator if and only if

$$\mathbb{E}\left[e^{-\sum_{i=1}^{n} x_i H_{t_i s}}\right] = \mathbb{E}\left[e^{-\sum_{i=1}^{n} x_i H_{t_i}}\right]^s, \quad \forall n \in \mathbb{N}, s > 0, x_i > 0, t_i \geq 0, i \in \{1, \ldots, n\}.$$

Proof
When showing necessity, the general idea is analogous to the proof of Lemma 2.21, see also (Mai and Scherer, 2014, proof of Theorem 5.3). Sufficiency follows by setting $s = 1/n$ for $n \in \mathbb{N}$ and using uniqueness of the Laplace transform. □

As mentioned before, IDT subordinators represent a subclass of ID processes. Barndorff-Nielsen et al. (2006a) characterize ID processes via a set of Lévy measures and Maruyama (1970) via a “big” Lévy measure. Using the multivariate scaling property in Lemma 2.24, IDT subordinators can also be uniquely characterized via those characteristics with additional conditions. However, both approaches do not yield a “simple” construction of IDT subordinators, which would have been helpful. Instead, in Barndorff-Nielsen et al. (2006a) existence is proven via application of Kolmogorov’s extension theorem and in Maruyama (1970) via integrals with respect to Poisson random measures on a path space. Nevertheless, these results provide an intuitive understanding of how big the class of IDT subordinators really is.

Another way of looking at IDT subordinators is to consider them as random variables taking values in the path space, see, e.g., (Mansuy, 2005, Section 5). Without going into details, it is obvious from Equation (2.15) that in this space, IDT subordinators (and also IDT processes) would represent infinitely divisible random variables. The converse direction does not hold true in general and (Mansuy, 2005, Lemma 5.1) states sufficient conditions. Actually, IDT processes remind more of a form of stable random variables, as the sum of independent copies equals in distribution a “scaled” (time-scaled) random variable of the original form.

Finally, we want to show that the exceedance property mentioned in Remark 2.15 holds.

Lemma 2.25 (Finite first-passage time)
Let $H = \{H_t\}_{t\geq 0}$ denote a non-trivial IDT subordinator, i.e. $H \not\equiv 0$, and E an Exp(1)*-distributed random variable independent of E. It holds that*

$$\mathbb{P}(E < H_t, \text{ for at least one } t > 0) = 1.$$

Proof
We prove the equivalent statement $\mathbb{P}(E \geq H_t, \forall t > 0) = 0$. Using that H is a.s. non-decreasing, it is enough to show that $\mathbb{P}(E \geq H_n, \text{ for infinitely many } n \in \mathbb{N}) = 0$. This follows using the Borel–Cantelli Lemma, since

$$\sum_{n=1}^{\infty} \mathbb{P}(E \geq H_n) = \sum_{n=1}^{\infty} e^{-n\,\Psi(1)} = \sum_{n=1}^{\infty} \left(e^{-\Psi(1)}\right)^n = \frac{e^{-\Psi(1)}}{1 - e^{-\Psi(1)}} < \infty,$$

where we used that $\mathbb{P}(E \geq H_n) = \mathbb{E}[\exp(-H_n)] = \exp(-n\,\Psi(1))$, see Lemma 2.21, and the fact that $\Psi(1) > 0$ for $H \not\equiv 0$. $\square$

2.3.3 An integral with respect to Lévy subordinators

For the general construction of IDT subordinators investigated in Section 2.3.4, one needs the notion of an integral with respect to a Lévy subordinator. The construction is not new and shows up in the literature in similar form, see, e.g., Mansuy (2005) or Hakassou and Ouknine (2011). However, it is not clear when the expressions used there actually exist, as they are stated under quite vague conditions on an involved measure (the measure has to be a "good" measure).

To make our exposition precise, we have to define the integral of a function f with respect to a Lévy subordinator Λ, giving a meaning to the expression "$\int f \,\mathrm{d}\Lambda$". The most general way would be to use Rajput and Rosinski (1989), who define integrals with respect to an ID "(independently scattered) random measure". It is shown in Sato (2004) that this is equivalent (when defining those ID random measures over $\mathbb{R}_+$ and ensuring stochastic continuity via an additional condition) to integrals with respect to natural additive processes. However, Sato (2004) treats only integrals over bounded intervals and the unbounded case is given by the limit in probability, whereas in Rajput and Rosinski (1989), there is no distinction between both cases. Alternatively, Jurek (1985) defines the integral using pathwise formal integration by parts (again for bounded intervals, treating the unbounded case as the limit in probability).

We are dealing with the comfortable case that the integral can be defined pathwise as a Lebesgue–Stieltjes integral due to the fact that the involved process is non-decreasing and f is positive. It might, however, take the value ∞, which triggers additional considerations based on the above literature. We will start from the definition in Rajput and Rosinski (1989) as it is the most general case, in particular treating unbounded integrals directly. We will show that this definition coincides almost surely with our pathwise definition, enabling us to state sufficient conditions for the existence of the integral and to use all results of Rajput and Rosinski (1989) on its distribution. It is not too difficult to do so but allows us to be precise in our statements.

It is possible to define an ID random measure μ_Λ on $\mathcal{B}([0,\infty))$ with $\mu_\Lambda([0,t]) = \Lambda_t$, for any $t \geq 0$, see (Sato, 2004, Theorem 3.2). We will omit the exact definition of an ID random measure here and instead observe that for every $\omega \in \Omega$, we can define a measure μ_Λ as above via the Stieltjes measure corresponding to the function $t \mapsto \Lambda_t(\omega)$, which yields a simpler way to define the corresponding ID random measure in our case. Intuitively speaking, a random measure is a random process with independent increments, which almost surely fulfills the properties of a measure. We proceed defining the integral as it is done in (Rajput and Rosinski, 1989, Section 2) and denoting it by $\int^r$. For a simple function $f = \sum_{j=1}^n x_j \, \mathbb{1}_{A_j}$ with $A_j \in \mathcal{B}([0,\infty))$ and bounded, for every $j \in \{1,\ldots,n\}$, $n \in \mathbb{N}$, the integral is defined via

$$\int_A^r f \, \mathrm{d}\Lambda := \sum_{j=1}^n x_j \, \mu_\Lambda(A \cap A_j), \tag{2.17}$$

for any $A \in \mathcal{B}([0,\infty))$. A measurable function $f : ([0,\infty), \mathcal{B}([0,\infty))) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is Λ-integrable, if: (i) there exists a sequence of simple functions f_n such that f_n converges to f almost everywhere (with respect to the Lebesgue measure); (ii) for every $A \in \mathcal{B}([0,\infty))$ the sequence of integrals of f_n as defined in Equation (2.17) converges in probability. The integral is then defined as this (well-defined) limit, i.e.

$$\int_{[0,\infty)}^r f_n \, \mathrm{d}\Lambda \xrightarrow{P} \int_{[0,\infty)}^r f \, \mathrm{d}\Lambda.$$

Linking the pathwise Lebesgue–Stieltjes integral to the integral definition of Rajput and Rosinski (1989), the following theorem includes all relevant aspects. It is formulated only for non-negative functions f so that we stay within the cosmos of positive distributions.

Theorem 2.26 (Integral with respect to Lévy subordinators)

(i) For f a measurable, non-negative function and $\Lambda = \{\Lambda_t\}_{t\geq 0}$ a Lévy subordinator, the integral $\int_{[0,\infty)}^r f\,\mathrm{d}\Lambda$ exists if and only if

$$\int_0^\infty \left(\xi\, f(s) + \int_0^\infty (1 \wedge (x\, f(s)))\,\nu(\mathrm{d}x)\right)\,\mathrm{d}s < \infty,$$
$$\int_0^\infty \int_0^\infty (1 \wedge (x^2\, f(s)^2))\,\nu(\mathrm{d}x)\,\mathrm{d}s < \infty, \tag{2.18}$$

with ξ denoting the drift and ν the Lévy measure of the Bernstein function Ψ_Λ corresponding to Λ.

(ii) Assume the conditions in Equation (2.18) to be fulfilled. Denoting by $\int_{[0,\infty)} f\,\mathrm{d}\Lambda$ the pathwise Lebesgue–Stieltjes integral, it holds that

$$\int_{[0,\infty)}^r f\,\mathrm{d}\Lambda = \int_{[0,\infty)} f\,\mathrm{d}\Lambda, \qquad \text{a.s.},$$

and thus $\int_{[0,\infty)} f\,\mathrm{d}\Lambda$ exists. Furthermore, the distribution of $\int_{[0,\infty)} f\,\mathrm{d}\Lambda$ is positive ID with Bernstein function Ψ given by

$$\Psi(x) = \int_0^\infty \Psi_\Lambda\,(x\, f(s))\,\mathrm{d}s, \quad x > 0, \tag{2.19}$$

with Ψ_Λ denoting the Bernstein function corresponding to Λ (see Lemma 2.21).[7]

Proof

Proof of (i): This is the result of (Rajput and Rosinski, 1989, Theorem 2.7), observing that λ is given by the Lebesgue measure, $\rho(s, \mathrm{d}x) = \nu(\mathrm{d}x)$, $a(s) \equiv \xi$ and $\sigma(s) \equiv 0$.
Proof of (ii): Based on (Rajput and Rosinski, 1989, proof of Theorem 2.7), given the conditions in Equation (2.18), there exist simple functions $f_n \geq 0$, $f_n \leq f$, $f_n \to f$ everywhere, such that

$$\int_{[0,\infty)}^r f(s)\,\mathrm{d}\Lambda_s := \lim_{n\to\infty} \int_{[0,\infty)}^r f_n(s)\,\mathrm{d}\Lambda_s,$$

where the limit is taken in probability. Thus, we know there exists a subsequence $\{n_k\}_{k\in\mathbb{N}}$

[7] This result is stated in the literature quite often in various forms, see, e.g., (Sato, 2004, Proposition 4.3), (Jurek and Vervaat, 1983, Lemma 1.1), and also in work focused on applications, see, e.g.,(Eberlein and Raible, 1999, Lemma 3.1).

for which the convergence is a.s., i.e. there is a nullset N such that $\forall \omega \in \Omega \backslash N$

$$\begin{aligned}\infty > \left(\int_{[0,\infty)}^{r} f(s)\,\mathrm{d}\Lambda_s\right)(\omega) &= \lim_{k\to\infty}\left(\int_{[0,\infty)}^{r} f_{n_k}(s)\,\mathrm{d}\Lambda_s\right)(\omega)\\ &= \lim_{k\to\infty}\left(\int_{[0,\infty)} f_{n_k}(s)\,\mathrm{d}\Lambda_s\right)(\omega) = \lim_{k\to\infty}\int_{[0,\infty)} f_{n_k}(s)\,\mathrm{d}(\Lambda_s(\omega))\\ &= \liminf_{k\to\infty}\int_{[0,\infty)} f_{n_k}(s)\,\mathrm{d}(\Lambda_s(\omega)) \geq \int_{[0,\infty)} f(s)\,\mathrm{d}(\Lambda_s(\omega)),\end{aligned}$$

using Fatou's Lemma for the last inequality. The crucial step is the second equality where we use that, for simple functions, both integral definitions coincide. Consequently, the pathwise defined expression is finite a.s. as well. It is easy to see that it coincides with the definition as limit in probability, starting as in the previous computations and using dominated convergence since now we have proven that f is integrable almost surely. The fact that the resulting distribution is ID follows from (Rajput and Rosinski, 1989, Theorem 2.7) and positivity is clear from the construction. The form of the Bernstein function follows with (Rajput and Rosinski, 1989, Proposition 2.6), observing that $K(x,s) = \Psi_\Lambda(x)$. □

Remark 2.27
Part (ii) is the important part of Theorem 2.26, we only took the detour over the integral definition of Rajput and Rosinski (1989) to use their results on existence and properties of the integral for our pathwise definition. Essentially, Theorem 2.26 simply proves that the pathwise (almost sure) limit exists if the limit in probability exists, and that both coincide. However, one has to pay attention to the slightly differing definitions involved. Having proven this, we can work with the pathwise definition in the following.

Remark 2.28
Alternatively, we could have introduced the pathwise integral similar to Jurek and Vervaat (1983), defining it for intervals $[a,b]$, $0 < a < b < \infty$, and searching conditions for its almost sure convergence when a and b approach the corresponding limits (almost sure convergence in this context is equivalent to convergence in distribution, see (Jurek and Vervaat, 1983, Lemma 1.2)). However, we decided to use existing results, starting from the most general available definition and using its well-known properties.

Remark 2.29
Based on Theorem 2.26(ii), for a given suitable function f, the expression $\int_{[0,\infty)} f\,\mathrm{d}\Lambda$ can be considered a mapping from the set of Lévy subordinators to the set of positive ID distributions. Combining this observation with the fact that Lévy subordinators are in

a one-to-one correspondence to positive ID distributions (see Lemma 2.17), the integral expression can be consideread a mapping from the set of positive ID distributions onto itself. It will be denoted by Φ_f*, with*

$$\Phi_f : \mathcal{L}(\Lambda_1) \mapsto \mathcal{L}\left(\int_{[0,\infty)} f \,\mathrm{d}\Lambda\right),$$

and we will use it simultaneously on the level of the involved Lévy measures, i.e. $\Phi_f : \nu_\Lambda \mapsto \nu_\int$*, where* ν_Λ *is the Lévy measure corresponding to the Lévy subordinator* Λ *and* $\nu_\int$ *denotes the Lévy measure corresponding to the distribution of* $\int_{[0,\infty)} f \,\mathrm{d}\Lambda$*. This type of function is well known by the name stochastic integral mapping and investigated, e.g., in Barndorff-Nielsen et al. (2006b); Sato (2006). It is of interest to state the domain and the range of a given mapping* Φ_f *and the pre-images of known subclasses. In Section 4.2.1, we will investigate one such function in more detail. Here, we will only state a couple of well-known results restricted to the case of non-negative distributions. For* $f_0(s) := \exp(-s)$*, Jurek and Vervaat (1983) show that* Φ_{f_0} *is a one-to-one mapping* $\Phi_{f_0} : \mathcal{M}_{\log} \to L$*, with* $\mathcal{M}_{\log} := \{\nu \in \mathcal{M} : \int_{(2,\infty)} \log(u)\nu(\mathrm{d}u) < \infty\}$*. For* $f_1(s) := (1-s)_+ = \max\{1-s, 0\}$*, it is shown in Jurek (1985) that* Φ_{f_1} *is a one-to-one mapping* $\Phi_{f_1} : \mathcal{M} \to U$*. Finally, for* $f_2(s) := \log(1/s)_+ = \max\{\log(1/s), 0\}$*, it is shown in Barndorff-Nielsen et al. (2006b) that this yields a one-to-one function* $\Phi_{f_2} : \mathcal{M} \to BO$ *with* $\Phi_{f_2}(L) = T$*.*

Remark 2.30
From the fact that $\Lambda_0 = 0$*, a.s., it follows that* $\mu_\Lambda(\{0\}) = 0$*, a.s.. Consequently,* $\int_{[0,\infty)} f \,\mathrm{d}\Lambda = \int_{(0,\infty)} f \,\mathrm{d}\Lambda$ *holds and there is no need to distinguish between both cases. Thus, we will often simply denote the integral by* $\int_0^\infty f \,\mathrm{d}\Lambda$*.*

2.3.4 Construction of families of IDT subordinators

The construction of families of IDT subordinators we investigate is based on a known construction of IDT processes, see (Mansuy, 2005, Example 2.2) or (Hakassou and Ouknine, 2011, Example 3.1). Both references consider a slightly different notation than we will, constructing an IDT process $H = \{H_t\}_{t\geq 0}$ from a Lévy process Λ via $H_t := \int_0^\infty \Lambda_{s\,t}\,\sigma(\mathrm{d}s)$, with σ a “good” measure. σ is called a “good” measure if the integral expression exists and consequently, using integration by parts, the alternative representation $H_t = \int_0^\infty \sigma([s/t,\infty))\,\mathrm{d}\Lambda_s$ is assumed to hold. We will start from this alternative representation, replacing $\sigma([s,\infty))$ by a decreasing non-negative function $f(s)$.

Having introduced the integral with respect to Lévy subordinators in the previous section, we are now able to state a construction of IDT subordinators in the same spirit but with concrete conditions on f ensuring existence. A shortened version of this result can be found in Bernhart et al. (2015b).

Lemma 2.31 (A class of IDT subordinators)
Define pathwise

$$H_t := \int_0^\infty f(s/t)\,\mathrm{d}\Lambda_s, \quad t > 0, \tag{2.20}$$

with $H_0 := 0$, for f a measurable, non-negative, non-increasing, left-continuous function, fulfilling

$$\int_0^\infty \left(\xi_\Lambda\, f(s) + \int_0^\infty (1 \wedge (x\, f(s)))\, \nu_\Lambda(\mathrm{d}x) \right) \mathrm{d}s < \infty,$$
$$\int_0^\infty \int_0^\infty (1 \wedge (x^2\, f(s)^2))\, \nu_\Lambda(\mathrm{d}x)\, \mathrm{d}s < \infty,$$

where ξ_Λ and ν_Λ are the drift and the Lévy measure of the subordinator Λ. Then $H = \{H_t\}_{t>0}$ yields an IDT subordinator.

Proof
The two integral conditions stated above are sufficient conditions for the existence of the integral (for $t = 1$) stated in Theorem 2.26. As shown in Lemma 5.13 in Barndorff-Nielsen et al. (2006a), the existence of the integral for $t = 1$ ensures the existence for any $t > 0$. This is clear as replacing $f(s)$ by $f(s/t)$ in the conditions stated in Theorem 2.26 can be reduced, using a change of variable, to above expressions multiplied by t. Thus, for all $\omega \in \Omega$ such that the expression is finite for all $t \in \mathbb{N}$, we will consider the process defined in Equation (2.20), otherwise we set it to zero. Furthermore,

$$H_t = \int_0^\infty f(s/t)\,\mathrm{d}\Lambda_s \geq \int_0^\infty f(s/u)\,\mathrm{d}\Lambda_s = H_u, \quad t \geq u,$$

as f is non-increasing. In addition, for arbitrary $u > 0$

$$\lim_{t \searrow u} H_t - H_u = \lim_{t \searrow u} \int_0^\infty f(s/t) - f(s/u)\,\mathrm{d}\Lambda_s = \int_0^\infty f\big((s/u)-\big) - f(s/u)\,\mathrm{d}\Lambda_s = 0,$$

as f is left-continuous, using dominated convergence (for $u = 0$, one can use the same idea). Finally, for the crucial IDT property, the general idea is as follows: denoting by

$\Lambda^{(1)}, \ldots, \Lambda^{(n)}$ *iid copies of* Λ,

$$\begin{aligned}
H^{(1)}_{t/n} + \ldots + H^{(n)}_{t/n} &= \int_0^\infty f(n\,s/t)\,\mathrm{d}\Lambda^{(1)}_s + \ldots + \int_0^\infty f(n\,s/t)\,\mathrm{d}\Lambda^{(n)}_s \\
&= \int_0^\infty f(n\,s/t)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right) \\
&\overset{d}{=} \int_0^\infty f(n\,s/t)\,\mathrm{d}\Lambda_{n\,s} = \int_0^\infty f(s/t)\,\mathrm{d}\Lambda_s = H_t,
\end{aligned}$$

where the equality in distribution follows from the IDT property of Lévy subordinators, which also holds when considering multivariate marginal distributions. To state this point in a mathematically more rigorous way, note that for every non-negative, left-continuous, decreasing function $f(s)$*, there exists a sequence of functions* f_k *of the form* $f_k(s) = \sum_{i=1}^k c_i\,\mathbb{1}_{(a_i,b_i]}(s), a_i < b_i, c_i \geq 0$*, with* $f_k \nearrow f$ *(applying, e.g., the procedure in the proof of (Elstrodt, 1999, Theorem III.4.13), where such a sequence is constructed). The advantage when using functions of this specific form is that the corresponding integral with respect to a Lévy subordinator only depends on the value of the subordinator at finitely many points. For any finite-dimensional sequence* $t_1, \ldots, t_d, d \in \mathbb{N}$*, one can observe*

$$\begin{aligned}
&\left(H^{(1)}_{t_1/n} + \ldots + H^{(n)}_{t_1/n}, \ldots, H^{(1)}_{t_d/n} + \ldots + H^{(n)}_{t_d/n}\right) \\
&= \left(\int_0^\infty f(n\,s/t_1)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right), \ldots, \int_0^\infty f(n\,s/t_d)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right)\right) \\
&= \lim_{k\to\infty}\left(\int_0^\infty f^{(t_1)}_k(s)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right), \ldots, \int_0^\infty f^{(t_d)}_k(s)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right)\right),
\end{aligned}$$

with $f^{(t_i)}_k, i \in 1, \ldots, d$*, denoting different sequences, where the limit exists a.s. as we have seen above. Furthermore, because of the form of the functions* f_k*, the expression in brackets is a (multivariate) function of the process* $\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right)$ *at finitely many points. As any finite-dimensional marginal distribution of* $\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right)$ *coincides with that of* $\Lambda_{n\,s}$*, which follows from a Lévy subordinator being an IDT subordinator, it follows that*

$$\begin{aligned}
&\left(\int_0^\infty f^{(t_1)}_k(s)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right), \ldots, \int_0^\infty f^{(t_d)}_k(s)\,\mathrm{d}\left(\Lambda^{(1)}_s + \ldots + \Lambda^{(n)}_s\right)\right) \\
&\overset{d}{=} \left(\int_0^\infty f^{(t_1)}_k(s)\,\mathrm{d}\Lambda_{n\,s}, \ldots, \int_0^\infty f^{(t_d)}_k(s)\,\mathrm{d}\Lambda_{n\,s}\right).
\end{aligned}$$

Since the second expression converges a.s. to the integral expressions with respect to $\Lambda_{n\,s}$*,* *the distributions of both limits have to coincide, which yields*

$$\begin{aligned}&\left(H^{(1)}_{t_1/n}+\ldots+H^{(n)}_{t_1/n},\ldots,H^{(1)}_{t_d/n}+\ldots+H^{(n)}_{t_d/n}\right)\\ \overset{d}{=}&\left(\int_0^\infty f(n\,s/t_1)\,\mathrm{d}\Lambda_{n\,s},\ldots,\int_0^\infty f(n\,s/t_d)\,\mathrm{d}\Lambda_{n\,s}\right)\\ =&\left(H_{t_1},\ldots,H_{t_d}\right).\end{aligned}$$

This proves the statement in a mathematically rigorous way. □

2.4 MSMVE distributions

In this section, an important class of multivariate distributions is introduced, the so-called min-stable multivariate exponential (MSMVE) distributions. This sets the stage for Chapter 4, which is devoted to constructing new parametric families of this class. Section 2.4.1 presents the necessary definitions, from which important properties are derived in Section 2.4.2. The relevance of MSMVE distributions is, among others, based on their connection with other important classes of distributions, which will be illustrated in Section 2.4.3. Section 2.4.4 concludes with stochastic representations and the crucial link to IDT subordinators.

2.4.1 Definition of MSMVE distributions

In Esary and Marshall (1974), several definitions of multivariate distributions with exponential minima are presented. All of them can be considered ideas to lift the concepts and properties of the one-dimensional exponential law to higher dimensions when relying on min-stability as characterizing property. The following definition introduces two of the ideas, among them MSMVE distributions.

Definition 2.32 (MSMVE and EM distributions)
A random vector $(X_1,\ldots,X_d)^\intercal$ with support $[0,\infty)^d$ is said to have an MSMVE distribution, if for all non-empty subsets $I\subset\{1,\ldots,d\}$ and $c_i>0, i\in 1,\ldots,d$, $\min_{i\in I}\{c_i\,X_i\}$ has an exponential distribution.
If this property is only required to hold for $c_1=\ldots=c_d$, the random vector is said to have a distribution with exponential minima (EM).

MSMVE distributions can be found in Esary and Marshall (1974) as distributions fulfilling condition "(c)", the name "MSMVE" can be found, e.g., in (Joe, 1990, Chapter 6), whereas in De Haan and Pickands (1986), these distributions are simply called "min-stable", and in Pickands (1989), they are called "(multivariate) negative exponential distributions". EM distributions are defined in Esary and Marshall (1974) as distributions fulfilling condition "(a)", obviously representing a superclass of MSMVE distributions. For reasons of completeness, we also introduce a famous subclass of MSMVE distributions, the so-called Marshall–Olkin (MO) distributions, first defined in Marshall and Olkin (1967) and also considered in Esary and Marshall (1974) as multivariate exponential distributions. A random vector $(X_1, \ldots, X_d)^\intercal$ with support $[0, \infty)^d$ is said to have an MO distribution if its distribution can be constructed via

$$X_i := \min_{\emptyset \neq I \subset \{1,\ldots,d\}:\, i \in I} \{E_I\}, \quad i = 1, \ldots, d, \tag{2.21}$$

with E_I independent and exponentially distributed.[8] It represents a typical shock-model construction with exponential shocks. From this construction and the min-stability of the exponential law, it is obvious that MO distributions are MSMVE distributions as well. In total, denoting the corresponding classes of distributions via the same abbreviations, we have

$$\text{MO} \subsetneq \text{MSMVE} \subsetneq \text{EM},$$

see, e.g., Esary and Marshall (1974). A last concept of Esary and Marshall (1974) we want to introduce is the notion of distributions "marginally equivalent in minimums". Two EM distributions are marginally equivalent in minimums if for the accordingly distributed random vectors $(X_1, \ldots, X_d)^\intercal$ and $(\tilde{X}_1, \ldots, \tilde{X}_d)^\intercal$, it follows

$$\mathcal{L}\left\{\min_{i \in I} \{X_i\}\right\} = \mathcal{L}\left\{\min_{i \in I} \left\{\tilde{X}_i\right\}\right\}, \text{ for all } \emptyset \neq I \subset \{1, \ldots, d\}.$$

In the following, we will focus on MSMVE distributions as Chapter 4 is devoted to the construction of new families of this class of distributions. However, if possible, results are proven for the broader class of EM distributions.

[8] Some of the E_I might have parameter $\lambda_I = 0$, corresponding to $E_I = \infty$. As MO distributions are not of particular relevance for the rest of the work, details are omitted. For a nice treatise on these, see (Mai and Scherer, 2013, Chapter 3).

2.4.2 Properties of MSMVE distributions

The next lemma collects a number of easy to derive but important properties of MSMVE distributions, denoted in terms of their survival functions $\bar{F}(\mathbf{x}) := P(X_1 > x_1, \ldots, X_d > x_d)$, where $\mathbf{x} = (x_1, \ldots, x_d)^\intercal$, with $x_i \geq 0, i \in 1, \ldots, d$.

Lemma 2.33 (Properties of MSMVE distributions)

(i) Let $\bar{F}$ be the survival function of an MSMVE distribution. Then it holds that

$$\bar{F}(\mathbf{x})^t = \bar{F}(t\,\mathbf{x}), \quad \forall \mathbf{x} \in [0,\infty)^d, t > 0. \tag{2.22}$$

(ii) Conversely, if the survival function $\bar{F}$ of a random vector $(X_1, \ldots, X_d)^\intercal$ with $\bar{F}(\mathbf{0}) > 0$ satisfies condition (2.22), the corresponding distribution is an MSMVE distribution.

(iii) Let $\bar{F}$ be the survival function of an MSMVE distribution. Then it has a representation via

$$\bar{F}(\mathbf{x}) = \exp\left(-\ell(\mathbf{x})\right), \forall \mathbf{x} \in [0,\infty)^d, \tag{2.23}$$

with $\ell := -\log\left(\bar{F}\right)$, $\ell : \mathbb{R}_+^d \to \mathbb{R}_+$, a homogeneous function of order 1, i.e. $\ell(t\,\mathbf{x}) = t\,\ell(\mathbf{x})$.

Proof

Property (i) follows directly from the definition. To be precise, one has for $\mathbf{x} \neq \mathbf{0}$:

$$\begin{aligned}\bar{F}(t\,\mathbf{x}) &= P(X_1 > t\,x_1, \ldots, X_d > t\,x_d) = P(X_i/x_i > t, \forall i : x_i > 0)\\ &= P\left(\min_{i:x_i>0}\{X_i/x_i\} > t\right) = P\left(\min_{i:x_i>0}\{X_i/x_i\} > 1\right)^t = \bar{F}(\mathbf{x})^t,\end{aligned}$$

where we have used the characteristic property plus the fact that $\mathbb{P}(X_i > 0) = 1, i = 1, \ldots, d$, as the one-dimensional marginals are exponentially distributed as well. For $\mathbf{x} = \mathbf{0}$, the property is obvious.

Part (ii) follows from the fact that $\bar{F}(\mathbf{0}) > 0$ implies $\bar{F}(\mathbf{0}) = 1$ and $\bar{F}(\mathbf{x}) > 0$ for all $\mathbf{x} \in [0,\infty)^d$. The second implication follows as from $\bar{F}(\mathbf{x}) = 0$ for an arbitrary $\mathbf{x} \in [0,\infty)^d$, $\bar{F}(\mathbf{0}) = 0$ can be derived using Equation (2.22) and continuity from below. The rest follows analogously to the computations in Part (i).

Part (iii) is simply Part (i) rewritten. □

Based on Lemma 2.33(iii), we are able to introduce so-called stable tail dependence functions.

Definition 2.34 (Stable tail dependence function)
The function $\ell : \mathbb{R}_+^d \to \mathbb{R}_+$ in Equation (2.23) is called a stable tail dependence function if we assume the one-dimensional marginals of $(X_1, \ldots, X_d)^\intercal$ to be unit exponentials, i.e. if additionally $\ell(\mathbf{e}_i) = 1$ for all unit vectors $\mathbf{e}_i, i = 1, \ldots, d$.

Remark 2.35
One can assign a unique stable tail dependence function to every MSMVE distribution. This is due to the fact that for any MSMVE distributed random vector $(X_1, \ldots, X_d)^\intercal$, there exist constants $c_i > 0$ such that $c_i\, X_i \sim \mathrm{Exp}(1), i = 1, \ldots, d$. Consequently,

$$\begin{aligned}\bar{F}(\mathbf{x}) &= \mathbb{P}\left(X_1 > x_1, \ldots, X_d > x_d\right) = \mathbb{P}\left(c_1\, X_1 > c_1\, x_1, \ldots, c_d\, X_d > c_d\, x_d\right) \\ &= \exp\left(-\ell(c_1\, x_1, \ldots, c_d\, x_d)\right),\end{aligned}$$

with ℓ a stable tail dependence function, which will be called "the" stable tail dependence function of this distribution.

An important consequence of Lemma 2.33 and Remark 2.35 is that every MSMVE distribution can be characterized by the corresponding tail dependence function. Because of that, in the following, we will often treat MSMVE distributions via their stable tail dependence functions. Some necessary conditions for stable tail dependence functions, as, e.g., homogeneity, follow directly from the definition. An interesting affiliated question is to find sufficient conditions for such a function to be a stable tail dependence function. There are several results based on representations via measures on subspaces of $\mathbb{R}^d$, see, e.g., (Resnick, 1987, Proposition 5.11), for a collection of the corresponding results. Furthermore, Molchanov (2008) characterizes them as support functions of specific (normalized) convex sets, so called max-zonoids. The characterization of Hofmann (2009) is refined in (Ressel, 2013, Theorem 6), who proves that a function $\ell : \mathbb{R}_+^d \to \mathbb{R}_+$ is a stable tail dependence function if and only if it is homogeneous of order 1, fully d-max-increasing, and $\ell(\mathbf{e}_1) = \ldots = \ell(\mathbf{e}_d) = 1$. The fully d-max-decreasingness property is basically a rewriting of the d-increasingness property of the associated MSMVE's distribution function after transformation to survival functions and then applying the log-transform.

In the well-studied bivariate case, it follows from homogeneity that a stable tail dependence function ℓ is characterized by the so-called Pickands dependence function

$A : [0,1] \to [1/2,1]$, which is defined by $A(t) := \ell(t, 1-t)$. Sufficient conditions for a function to be a bivariate Pickands dependence function are known, see (Gudendorf and Segers, 2010, Theorem 2.3): A is a bivariate Pickands dependence function if and only if A is convex and $\max\{t, 1-t\} \leq A(t) \leq 1$, for all $t \in [0,1]$. Concerning measures of dependence, like concordance measures and tail dependence coefficients[9], it is well known for a bivariate MSMVE vector (X_1, X_2) that these measures can be computed easily from A (see, e.g., Gudendorf and Segers (2010)), e.g.

$$\begin{aligned}
\tau &= \int_0^1 \frac{t\,(1-t)}{A(t)} \mathrm{d}A'(t), && \text{(Kendall's } \tau\text{)},\\
\rho &= 12 \int_0^1 \frac{1}{(1+A(t))^2}\, \mathrm{d}t - 3, && \text{(Spearman's } \rho\text{)},\\
\lambda_L &= 2\big(1 - A(1/2)\big), && \text{(lower tail-dependence coefficient)},\\
\lambda_U &= \mathbb{1}_{\{A(1/2)=1/2\}}, && \text{(upper tail-dependence coefficient)}.
\end{aligned}$$

From the restrictions on A, it is clear that $\lambda_U = 0$ unless $A(t) = \max\{t, 1-t\}$, which corresponds to $\bar{F}(x_1, x_2) = \exp(-\max\{c_1\, x_1, c_2\, x_2\})$, with c_1, c_2 as in Remark 2.35. Thus, unless $X_2 = c_1/c_2\, X_1$, i.e. perfect dependence, the upper tail dependence coefficient is zero.

Before we introduce the link between MSMVE distributions and other important classes of distributions, we want to present one last property which will give us some intuitive guidance for a result presented later. It can be stated not only for MSMVE distributions, but on the level of EM distributions. We have to introduce the concept of exchangeability first.

Definition 2.36 (Exchangeability)
The distribution of a random vector $(X_1, \ldots, X_d)^\intercal$ is said to be exchangeable if it is invariant under arbitrary permutations, i.e. if $\mathcal{L}\left((X_1, \ldots, X_d)^\intercal\right) = \mathcal{L}\left((X_{\pi(1)}, \ldots, X_{\pi(d)})^\intercal\right)$ for every permutation $\pi : \{1, \ldots, d\} \to \{1, \ldots, d\}$.

[9] To keep this section readable, we only sketch the definition of those measures in a footnote, as they do not represent a crucial aspect of this work. For more information on that topic, the interested reader is referred to Joe (1990), McNeil et al. (2005), or Nelsen (2006). We only consider bivariate random vectors with continuous marginals for reasons of simplicity. Kendall's τ and Spearman's ρ are measures for the dependence of a bivariate random vector, where concordance is used as a concept of dependence. Both are invariant under strictly increasing transformations of the components of the vector and "normalized" to the interval $[-1, 1]$. For a bivariate random vector (X_1, X_2) with distribution function F and margins F_1 and F_2, they are defined via $\tau := 4 \int_{\mathbb{R}^2} F(\mathbf{x}) \mathrm{d}F(\mathbf{x}) - 1 = 4\,\mathbb{E}[F(X_1, X_2)] - 1$, and $\rho := 12 \int_{\mathbb{R}^2} F_1(x_1)\, F_2(x_2) \mathrm{d}F(\mathbf{x}) - 3 = 12\,\mathbb{E}[F_1(X_1)\, F_2(X_2)] - 3$. Tail-dependence in turn is a measure of "dependence in the extremes" of a bivariate vector, which is also invariant under strictly increasing transformations. The lower tail-dependence coefficient λ_L can be defined as $\lambda_L := \lim_{x \searrow 0} \mathbb{P}(F_1(X_1) \leq x | F_2(X_2) \leq x)$, and the upper tail-dependence coefficient λ_U via $\lambda_U := \lim_{x \nearrow 1} \mathbb{P}(F_1(X_1) > x | F_2(X_2) > x)$, whenever these limits exist.

For exchangeable distributions, the distribution of $\min_{i\in I}\{X_i\}$ only depends on the size of I, i.e. on $|I|$. For exchangable MSMVE distributions, we denote the parameters of the corresponding exponential distributions by λ_i, i.e. $\min_{i\in I}\{X_i\} \sim \mathrm{Exp}(\lambda_{|I|})$ for every $\emptyset \neq I \subset \{1,\dots,d\}$. It follows, for example, that the one-dimensional marginals are unit exponentials if $\lambda_1 = 1$. It is possible to derive a necessary condition for the finite sequence $\{\lambda_i\}_{i=1,\dots,d}$. To our knowledge, this has not been stated for EM distributions before.

Lemma 2.37 (Necessary condition)
Let $\{\lambda_i\}_{i=1,\dots,d}$ denote the parameter sequence corresponding to an exchangeable EM vector $(X_1,\dots,X_d)^\intercal$ and define $\lambda_0 := 0$. Then, this sequence has to be d-alternating, i.e.

$$\nabla^{d-k}\lambda_k := \sum_{i=0}^{d-k}(-1)^i\binom{d-k}{i}\lambda_{k+i} \leq 0, \quad k = 0,1,\dots,d-1. \tag{2.24}$$

Proof
One could directly modify the proof of necessity of (Mai and Scherer, 2009a, Theorem 3.1), as it can be generalized to all random vectors with distributions fulfilling $\mathbb{P}(Y_i < y, i \in I) = y^{\lambda_{|I|}}$, $y \in [0,1]$. Such a random vector is defined via the expression $(\exp(-X_1),\dots,\exp(-X_d))^\intercal$. Alternatively, (Esary and Marshall, 1974, Theorem 4.1) proves that for the given EM distribution, there exists an MO distribution marginally equivalent in minimums, i.e. with the same sequence $\{\lambda_i\}_{i=1,\dots,d}$. By direct application of (Mai and Scherer, 2009a, Theorem 3.1) to this MO distribution, it follows that the sequence $a_0 := \lambda_1/\lambda_1 = 1, a_1 := (\lambda_2-\lambda_1)/\lambda_1, \dots, a_{d-1} := (\lambda_d - \lambda_{d-1})/\lambda_1$ is d-monotone, i.e. $\nabla^{d-k-1}a_k \geq 0, k = 0,1,\dots,d$. The claim is established using the equality $\nabla^{d-k}\lambda_k = -\lambda_1\,\nabla^{d-k-1}a_k$. □

Remark 2.38
On first sight, it is not clear why Lemma 2.37 is included here. It will be seen later that for exchangeable sequences of EM random variables, there is a connection to positive ID distributions. This link can already be motivated by the result stated here: Similar as in Mai and Scherer (2009a), one can use (Gnedin and Pitman, 2008, Corollary 4.2), which basically states that if Equation (2.24) *holds for all $d \in \mathbb{N}$, there is a positive ID distribution (here, on $[0,\infty]$) with the corresponding Bernstein function fulfilling $\Psi(i) = \lambda_i, i \in \mathbb{N}$. For an explicit investigation of exactly this connection in a special case, see Mai (2014c).*

2.4.3 Link to other distributions

The importance of MSMVE distributions is to a large part due to their relation to other distributions, in particular multivariate extreme-value distributions. This link will be illustrated in the present section. Multivariate extreme-value distributions are very well studied, see, e.g., Resnick (1987); (Joe, 1997, Chapter 6); Kotz and Nadarajah (2000), and arise in many applications. For their introduction, we follow Resnick (1987). They are defined as the limit distributions of component-wise maxima of iid random vectors, when scaling these appropriately. To be precise, let $(Z_1, \dots, Z_d)^\intercal$ denote a random vector with distribution function F and $(Z_1^{(j)}, \dots, Z_d^{(j)})^\intercal$, $j \in \mathbb{N}$, iid copies of it. The component-wise maxima of the first n copies are defined via $M_i^{(n)} := \max_{j \leq n}\{Z_i^{(j)}\}$ for $i = 1, \dots, d$. Assume now that there exist normalizing sequences $\{a_i^{(n)}\}_{n\in\mathbb{N}}, i = 1, \dots, d$, with all $a_i^{(n)} > 0$, and $\{b_i^{(n)}\}_{n\in\mathbb{N}}, i = 1, \dots, d$, such that for $n \to \infty$,

$$\mathbb{P}\left(\left(M_i^{(n)} - b_i^{(n)}\right)/a_i^{(n)} \leq x_i, i = 1, \dots, d\right) = F\left(a_1^{(n)}\, x_1 + b_1^{(n)}, \dots, a_d^{(n)}\, x_d + b_d^{(n)}\right)^n \\ \to G(\mathbf{x}),$$

weakly for a distribution function G with non-degenerate marginals. By “weakly”, we mean pointwise-convergence for all points of continuity of G and by a “degenerate distribution”, we mean a Dirac distribution. The distributions corresponding to such distribution functions G are called multivariate extreme-value (MEV) distributions. A very helpful result, stated, e.g., in (Resnick, 1987, Proposition 5.10), is the fact that it is possible to standardize the problem to fixed marginal distributions, e.g. to marginal distributions with the distribution function $\Theta_c(x) := \exp(-c/x), x > 0$ with some $c > 0$. The standardization works as described in the following remark, which is stated verbally to avoid introducing cumbersome additional notation.

Remark 2.39 (Resnick (1987), Proposition 5.10)
For any random vector with MEV distribution, there is a component-wise transformation (strictly increasing on the support of the components) such that the resulting vector has an MEV distribution with marginal distributions given by Θ_1.[10] *Furthermore, the corresponding vector $(Z_1, \dots, Z_d)^\intercal$ can also be transformed component-wise such that*

[10] We can explicitly state the component-wise transform. For $(Y_1, \dots, Y_d)^\intercal$ an MEV distribution, each Y_i has to follow a one-dimensional extreme-value distribution, i.e. one of the distributions given in (Resnick, 1987, Proposition 0.3). In particular, all the marginal distribution functions $G_i, i = 1, \dots, d$ have to be continuous and strictly increasing on their support. The vector given by $(-1/\log(G_1(Y_1)), \dots, -1/\log(G_d(Y_d)))^\intercal$ has the desired distribution.

the resulting distribution of maxima converges to the transformed MEV distribution with desired marginals. The converse holds true as well.

From this follows that it is enough to investigate MEV distributions with marginal distributions of the form Θ_1. We want to consider those with marginal distributions of the form Θ_c with arbitrary values for c, which basically corresponds to a linear scaling of the respective marginals. This yields exactly all distributions with non-degenerate marginals that satisfy the condition

$$G(t\,\mathbf{x})^t = G(\mathbf{x}), \quad \forall \mathbf{x} \in [0,\infty)^d, t > 0, \tag{2.25}$$

see (Resnick, 1987, Proposition 5.9) combined with (Resnick, 1987, Proposition 5.10(b)). These distributions coincide with so-called max-stable processes as defined in De Haan (1984) (restricted to non-degenerate marginals). We will also call them max-stable (though max-stable sometimes refers to a broader class, see, e.g., (Resnick, 1987, Section 5.4)).

Definition 2.40 (Max-stable distributions)
A random vector $(Y_1, \ldots, Y_d)^\intercal$ is said to be max-stable, if its marginals are non-degenerate and its distribution function G fulfills Equation (2.25).

Having defined max-stable distributions as a class of standardized MEV distributions, we are able to state their connection to MSMVE distributions.

Lemma 2.41 (Connection between MSMVE and max-stable distributions)
A random vector $(Y_1, \ldots, Y_d)^\intercal$ is max-stable if and only if the random vector $(X_1, \ldots, X_d)^\intercal$ defined via $(X_1, \ldots, X_d)^\intercal := (1/Y_1, \ldots, 1/Y_d)^\intercal$ has an MSMVE distribution.
Furthermore, G is the distribution function of a max-stable distribution if and only if it has the form

$$G(\mathbf{x}) = \exp\left(-\ell\left(\frac{c_1}{x_1}, \ldots, \frac{c_d}{x_d}\right)\right), \quad \mathbf{x} \in (0,\infty)^d,$$

with ℓ a stable tail dependence function and constants $c_i > 0, i = 1, \ldots, d$, such that $Y_i \sim \Theta_{c_i}$.

Proof
The claim follows easily from Lemma 2.33. □

Lemma 2.41 establishes a first link between MSMVE distributions and MEV distributions via max-stable distributions. One can deduce, e.g., why the study of stable tail dependence functions might be crucial for both classes of distributions and how to derive stochastic representations for one class from the other. It is even possible to explicitly state a connection between ℓ and the distribution of the initially introduced random vector $(Z_1, \dots, Z_d)^\intercal$. More precisely, according to, e.g., (Segers, 2012, Equation (2.5)),

$$\ell(\mathbf{x}) = \lim_{n\to\infty} n\,\mathbb{P}\left(F_1(Z_1) > 1 - \frac{x_1}{n}, \text{ or ..., or } F_d(Z_d) > 1 - \frac{x_d}{n}\right), \quad \mathbf{x} \in [0,\infty)^d,$$

with $F_1, \dots, F_d$ denoting the marginal distribution functions of F. When using the language of copulas, the link between MEV and MSMVE distributions might be stated in even more elegant form, but the introduction of additional concepts is required. We will only shortly define copulas and the most relevant properties, for more information, the interested reader is referred to Joe (1997); Nelsen (2006); Mai and Scherer (2013).

Definition 2.42 (Copulas)
A function $C : [0,1]^d \to [0,1]$ is called a (d-dimensional) copula, if there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random vector $(U_1, \dots, U_d)^\intercal$ with $U_i \sim U([0,1]), i = 1, \dots, d$, such that

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

The crucial result in the world of copulas is Sklar's Theorem, which can be found in the next lemma.

Lemma 2.43 (Crucial properties)

(i) *(Sklar's Theorem) Let F be a d-dimensional distribution function with margins $F_1, \dots, F_d$. Then, there exists a (d-dimensional) copula C such that*

$$F(x_1, \dots, x_d) = C\left(F_1(x_1), \dots, F_d(x_d)\right),$$

for all $(x_1, \dots, x_d)^\intercal \in \mathbb{R}^d$. C is unique if $F_1, \dots, F_d$ are continuous. Conversely, with C a copula and $F_1, \dots, F_d$ univariate distribution functions, the above equation defines a d-dimensional distribution function.

(ii) *(Sklar's Theorem for survival copulas) Let $\bar{F}$ be a d-dimensional survival function with marginal survival functions $\bar{F}_1, \dots, \bar{F}_d$. Then, there exists a (d-dimensional) copula C, the so-called survival copula, such that for all $(x_1, \dots, x_d)^\intercal \in \mathbb{R}^d$ it holds* *that*

$$\bar{F}(x_1, \ldots, x_d) = C\left(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d)\right).$$

C is unique if $\bar{F}_1, \ldots, \bar{F}_d$ are continuous. Conversely, with C a copula and $\bar{F}_1, \ldots, \bar{F}_d$ survival functions, the above equation defines a d-dimensional survival function.

(iii) Let $(X_1, \ldots, X_d)^\intercal$ be a random vector with copula C and $T_j : \mathbb{R} \to \mathbb{R}$ functions strictly increasing on the support of X_j for all $j = 1, \ldots, d$. Then, the random vector $(T_1(X_1), \ldots, T_d(X_d))^\intercal$ also has copula C.

(iv) Let $(X_1, \ldots, X_d)^\intercal$ be a random vector with copula C and $T_j : \mathbb{R} \to \mathbb{R}$ functions strictly decreasing on the support of X_j for all $j = 1, \ldots, d$. Then, the random vector $(T_1(X_1), \ldots, T_d(X_d))^\intercal$ has survival copula C.

Proof
For (i) and (ii), see (Mai and Scherer, 2013, Theorem 1.2 and Theorem 1.3). In the case of continuous marginals, the idea is to simply consider $(F_1(X_1), \ldots, F_d(X_d))^\intercal$. Part (iii) exists in many versions, we cited a very precise one from (Embrechts and Hofert, 2013, Proposition 4). Part (iv) follows easily from Part (iii) using that $1 - T_j$ constitute strictly increasing functions and using continuity of the copula to get an expression for survival functions. □

Intuitively speaking, the concept of copulas allows to separate marginal distributions and dependence structure and to analyze them separately. This is the reason for its major importance in theory and practice. Coming back to MSMVE and MEV distributions, one can show that they share the same dependence structure, which is basically only an alternative way to formulate Lemma 2.41 combined with Remark 2.39.

Lemma 2.44 (Connection between MSMVE and MEV distributions)
The set of copulas of MEV distributions (also called extreme-value copulas, see, e.g., Gudendorf and Segers (2010)) coincides with the set of survival copulas of MSMVE distributions.
Consequently, copulas of MEV distributions have the form

$$C(u_1, \ldots, u_d) = \exp\left(-\ell\left(-\log(u_1), \ldots, -\log(u_d)\right)\right), \quad u_1, \ldots, u_d \in (0, 1],$$

with ℓ a stable tail dependence function.

Proof
Starting from an arbitrary MEV distribution, Remark 2.39 states that one gets a max-stable distribution via a component-wise transform which is strictly increasing on the support of each margin. Consequently, both share the same copula with Lemma 2.43(iii). Starting from an arbitrary MSMVE distribution, it is easy to transform it to an MSMVE distribution with unit exponential marginals, see, e.g., Remark 2.35, whereas the survival copula remains unchanged. Now, the claim follows from Lemma 2.41, using that $1/y$ is strictly decreasing and thus, copula and survival copula are interchanged with Lemma 2.43(iv).

The given form follows from the fact that the survival function of an MSMVE distribution with unit exponential marginals can be written as

$$\bar{F}(\mathbf{x}) = \exp\left(-\ell(x_1, \ldots, x_d)\right) = \exp\left(-\ell\left(-\log\left(\bar{F}_1(x_1)\right), \ldots, -\log\left(\bar{F}_d(x_d)\right)\right)\right). \qquad \square$$

An alternative second way to state the link between both classes of distributions would be to represent MSMVE distributions as limit distributions of component-wise minima, see, e.g., Pickands (1989), but we will not include this aspect here.

2.4.4 Stochastic representations

In this section, we present two stochastic representations for MSMVE distributions. The first holds for all of them, the second only for a subclass. Stochastic representations are important for several reasons: (a) they allow for a better understanding, (b) can be helpful for the derivation of further properties, and (c) can serve as a starting point when developing simulation algorithms.

Both stochastic representations can even be given for random sequences $\{X_i\}_{i \in \mathbb{N}}$, not only for random vectors. A stochastic sequence is called MSMVE (or EM), if the crucial property presented in Definition 2.32 holds for all finite subsets $I \in \mathbb{N}$.[11] The first representation, the so-called spectral representation, can be found in De Haan and Pickands (1986) and is basically a slight adaption of (De Haan, 1984, Theorem 2).

[11] One could also require it to hold for infinite subsets if $\mathrm{Exp}(\infty) = \delta_0$ is considered an exponential distribution, where δ_0 denotes the Dirac measure at 0. However, this property is already implied by the property for finite subsets.

Lemma 2.45 (Spectral representation of MSMVE sequences)
$\{X_i\}_{i\in\mathbb{N}}$ denotes an MSMVE sequence if and only if it has a representation via

$$X_i := \inf_{l\in\mathbb{N}} \left\{ \frac{\sum_{k=1}^{l} E_k}{f_i(U_l)} \right\}, \quad i \in \mathbb{N},$$

with $\{E_i\}_{i\in\mathbb{N}}$ a sequence of iid $\mathrm{Exp}(1)$-distributed random variables, $\{U_i\}_{i\in\mathbb{N}}$ a sequence of iid $\mathrm{U}([0,1])$-distributed random variables independent thereof, and $f_i : [0,1] \to \mathbb{R}_+$, with $\int_0^1 f_i(s)\mathrm{d}s < \infty$ for $i \in \mathbb{N}$.

Proof
See (De Haan and Pickands, 1986, Theorem 2.1). There, it is stated with $\{S_l, U_l\}_{l\in\mathbb{N}}$ an enumeration of points of a homogeneous Poisson process with unit intensity on the strip $[0,1] \times \mathbb{R}_+$, which can be constructed as above to avoid introducing the necessary mathematical definitions: It is well known that $\sum_{k=1}^{l} E_k, l \in \mathbb{N}$, yields an enumeration of points of a homogeneous Poisson process on $\mathbb{R}_+$, the rest follows using (Resnick, 1987, Proposition 3.8). □

The second representation, which we call IDT-frailty representation, only holds for exchangeable sequences. A random sequence $\{X_i\}_{i\in\mathbb{N}}$ is called exchangeable if all finite subsets are exchangeable in the sense of Definition 2.36. The famous De Finetti's Theorem proves that exchangeable sequences allow for a conditional independence structure, which was used in Mai and Scherer (2014) to derive the following representation.

Theorem 2.46 (IDT-frailty representation of exchangeable MSMVE seq.)
An exchangeable random sequence $\{X_i\}_{i\in\mathbb{N}}$ is MSMVE if and only if it allows for a representation via

$$X_i := \inf\left\{t > 0 : H_t > E_i\right\}, \quad i \in \mathbb{N}, \tag{2.26}$$

with $\{E_i\}_{i\in\mathbb{N}}$ a sequence of iid $\mathrm{Exp}(1)$-distributed random variables and $\{H_t\}_{t\geq 0}$ an IDT subordinator (on $[0,\infty]$) independent thereof.
For exchangeable MO sequences, the analogous result holds with H a (killed) Lévy subordinator.

Proof
See (Mai and Scherer, 2014, Theorem 5.3). A first-passage time construction as in Equation (2.26) follows from De Finetti's Theorem where H is called the conditional cumulative hazard process. The IDT property is then shown to correspond to the MSMVE

property of the random sequence. The result for MO sequences is shown (formulated on the level of copulas) in Mai and Scherer (2009a). □

Remark 2.47
This representation consequently holds for all random vectors that can be constructed as a part of an exchangeable sequence. A given (necessarily exchangeable) random vector is called extendible if such a sequence exists. For more information on exchangeability, extendibility, and its implications, see, e.g., Aldous (1985).

Remark 2.48
A similar result as in Theorem 2.46 also holds for EM sequences, see (Mai and Scherer, 2014, Theorem 1.1): An exchangeable random sequence is EM if and only if a representation as in Equation (2.26) *holds with H a subordinator that fulfills Equation* (2.15) *only for one-dimensional marginal distributions. These processes are called weak IDT subordinators which are a superclass of IDT subordinators for which, e.g., Lemma 2.21 holds as well. Now, we are able to make sense of Lemma 2.37 and Remark 2.38. If an exchangeable MSMVE sequence is extendible, among others, the parameter sequence $\{\lambda_i\}_{i=1,\dots,d}$ can be extended to a sequence $\{\lambda_i\}_{i\in\mathbb{N}}$ that fulfills Equation* (2.24) *for all $d \in \mathbb{N}$. With (Gnedin and Pitman, 2008, Corollary 4.2), it follows that there is a positive ID distribution (on $[0,\infty]$) with $\lambda_i = \Psi(i)$. This is exactly the marginal distribution of the corresponding (weak) IDT subordinator, as for $|I| = i, i \in \mathbb{N}, t \geq 0$,*

$$e^{-\lambda_i t} = \mathbb{P}\left(\min_{j\in I}\{X_j\} > t\right) = \mathbb{P}\left(\min_{j\in I}\{E_j\} > H_t\right) = \mathbb{E}\left[e^{-i\,H_t}\right] = e^{-t\,\Psi(i)},$$

where the third equality follows from the fact that $\min_{j\in I}\{X_j\} \sim \mathrm{Exp}(i)$.
In particular, EM sequences constructed via a first-passage time construction (as in Equation (2.26)*) are marginally equivalent in minimums if and only if the corresponding (weak) IDT subordinators share the same one-dimensional marginal distribution.*

Remark 2.49
Lemma 2.45 can be formulated slightly different for exchangeable sequences according to (De Haan and Pickands, 1986, Theorem 5.1), replacing the sequence of functions f_i by a function f_0 and a function transformation called Piston, as exchangeability implies strict stationarity.

The implication of Theorem 2.46 which is most important for our work is the fact that, starting from an IDT subordinator, construction (2.26) yields an MSMVE sequence. Based on this result, we will construct new MSMVE families in Chapter 4. Starting from a subordinator and determining the resulting MSMVE distribution, one has to compute quantities of the form

$$\mathbb{P}\left(X_i > x_i, i \in I\right) = \mathbb{E}\left[e^{-\sum_{i\in I} H_{x_i}}\right],$$

for finite $I \subset \mathbb{N}$, $x_i > 0, i \in I$. Basically, one has to be able to compute the multivariate Laplace transform of H restricted to natural numbers.

Furthermore, using Lemma 2.50, which is taken from Mai (2014a), it is possible to construct not only exchangeable sequences, but also non-exchangeable vectors, e.g. based on factor-model motivations. This is important as it also stresses the applicability of these models beyond the cosmos of exchangeable models.

Lemma 2.50 (Multi-factor MSMVE distributions)
Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting $n+1 \in \mathbb{N}$ independent, non-decreasing IDT subordinators $\tilde{H}^{(i)} = \{\tilde{H}_t^{(i)}\}_{t\geq 0}$, $i = 0, \ldots, n$, and an independent iid sequence $E_1, \ldots, E_d$ of exponential random variables with unit mean. Moreover, let $A = (a_{i,j}) \in \mathbb{R}^{d\times(n+1)}$ be an arbitrary matrix with non-negative entries, having at least one positive entry per row. We define the vector-valued stochastic process

$$\mathbf{H}_t = \begin{pmatrix} H_t^{(1)} \\ H_t^{(2)} \\ \vdots \\ H_t^{(d)} \end{pmatrix} := A \cdot \begin{pmatrix} \tilde{H}_t^{(0)} \\ \tilde{H}_t^{(1)} \\ \vdots \\ \tilde{H}_t^{(n)} \end{pmatrix} = \begin{pmatrix} a_{1,0}\,\tilde{H}_t^{(0)} + \ldots + a_{1,n}\,\tilde{H}_t^{(n)} \\ a_{2,0}\,\tilde{H}_t^{(0)} + \ldots + a_{2,n}\,\tilde{H}_t^{(n)} \\ \vdots \qquad\qquad \vdots \\ a_{d,0}\,\tilde{H}_t^{(0)} + \ldots + a_{d,n}\,\tilde{H}_t^{(n)} \end{pmatrix},$$

whose component processes are all IDT subordinators. The random vector $(X_1, \ldots, X_d)^\intercal$ defined via

$$X_k := \inf\{t > 0 \,:\, H_t^{(k)} > E_k\}, \quad k = 1, \ldots, d,$$

has an MSMVE law.[12]

Proof
See (Mai, 2014a, Lemma 4.4). □

[12] Note that by defining the entries of the matrix A appropriately and interpreting the processes $\tilde{H}^{(i)}$ as stochastic drivers, dedicated factor models can be constructed.

3 The density of distributions from the Bondesson class

3.1 Motivation

In this chapter, we derive a convenient representation for the probability density of distributions from the Bondesson class, which was introduced in Section 2.2.2. On the one hand, we have seen that it is a very large class, including important distributions (e.g. the stable distribution and the Hartman–Watson distribution) for which closed-form expressions of the respective probability density functions are unknown. On the other hand, it is easy to define new distributions in terms of the corresponding Laplace exponent, since complete Bernstein functions are well studied (for a list of over one hundred complete Bernstein functions, see (Schilling et al., 2010, p. 218ff)). Moreover, whenever the distribution of the first hitting time of a diffusion is studied, it is of the Bondesson class as well, as is shown in Bondesson (1981). In all cases, a convenient representation for the probability density would be desirable to be able to work with these distributions. The range of possible areas of application is huge: In mathematical finance, they play an important role, in particular due to the direct link to Lévy subordinators. This class of processes has gained considerable attention, among others through the concept of time-change, see, e.g., Clark (1973), Carr and Wu (2004), or Mendoza-Arriaga et al. (2010). An example where these processes, and thus also the corresponding distributions, play a role in dependence modeling can be found in Mai and Scherer (2009b).

There is only a small number of positive, infinitely divisible distributions whose density is known in closed form and one typically treats those distributions by their Laplace transform. Therefore, when looking for convenient representations of the densities, Laplace inversion constitutes a natural starting point, but the basic approach based on integration along the Bromwich contour often suffers from numerical instabilities, as we will see below.

We present an alternative representation relying on the idea of contour transformation of the Bromwich integral. The integration path of Kiesel and Lutz (2011) is considered and it is shown that this helps to circumvent undesirable features of the original integrand, primarily the highly oscillating behavior. Additionally, it allows for a transformation of the integral to a finite integration interval, consequently avoiding truncation errors. The idea of contour transformation is quite popular and there are many papers proposing such transforms, e.g., Talbot (1979), Evans and Chung (2000), Abate and Valko (2004), López-Fernández and Palencia (2004), or Weideman and Trefethen (2007). These approaches are typically formulated very generally. However, checking the admissibility conditions for a given problem or distribution can be quite difficult. In practical applications, these methods are therefore often used heuristically, i.e., without actually checking admissibility.[1] In a joint project with Jan-Frederik Mai, Steffen Schenk, and Matthias Scherer, we investigated this aspect for a whole class of distributions, the Bondesson class, resulting in the paper Bernhart et al. (2015a). The main aspects of this chapter are published in that paper and stem from joint work with the co-authors, whereas some results for the Hartman–Watson law are published in Bernhart and Mai (2014a). Consequently, parts of this chapter exhibit a considerable conformity with these references.

The main contribution of this chapter consists of proving the admissibility of our transformation for the Bondesson class (see Theorem 3.3). The approach is then investigated for a couple of examples and the results outline the exceptional suitability in terms of stability and efficiency.

The remainder of this chapter is structured as follows: Section 3.2 presents the main theoretical result. Remarks on the numerical implementation of the resulting formula and results of numerical tests can be found in Section 3.3. Section 3.4 summarizes the results and concludes.

3.2 Integral representation of f_μ

Let μ denote a distribution from the Bondesson class, as introduced in Section 2.2.2. Consequently, its Laplace exponent is a complete Bernstein function admitting a rep-

[1] I encountered the problem of checking admissibility for the first time during my master thesis, when I considered the pricing of CDO tranches in a default model based on scale mixtures of Marshall–Olkin copulas as introduced in Bernhart et al. (2013).

resentation as given in Equation (2.7). The following property of complete Bernstein functions is crucial when considering contour transformations.

Lemma 3.1 (Analytic extension of complete Bernstein functions)
Assume that Ψ is a complete Bernstein function. Then Ψ has an analytic extension to $\mathbb{C}\setminus(-\infty,0]$, with the representation

$$\Psi(z) = \xi\, z + \int_{(0,\infty)} \frac{z}{z+t}\,\sigma(\mathrm{d}t), \quad z \in \mathbb{C}\setminus(-\infty,0],$$

where σ is the Stieltjes measure on $(0,\infty)$ satisfying $\int_{(0,\infty)}(1+t)^{-1}\,\sigma(\mathrm{d}t) < \infty$.

Proof
See (Schilling et al., 2010, Theorem 6.2 (v) and (vi)). □

Assumption 3.2 (W.l.o.g. we neglect the drift)
From here on, we assume the drift ξ to equal zero, as it only corresponds to a constant additive term, which can be easily incorporated into a density or distribution function.

Furthermore, let f_μ denote the corresponding probability density which we know to exist if and only if $\sigma\big((0,\infty)\big) = \infty$ (see Corollary 2.12). A natural starting point for deriving the density, if only the Laplace exponent is known, is the Bromwich inversion formula, see (Widder, 1946, Theorem 7.3):

$$f_\mu(x) = \lim_{R\to\infty} \frac{1}{2\,\pi\, i} \int_{a-i\,R}^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z, \quad a, x > 0. \tag{3.1}$$

The sole assumption made in the present chapter is that Equation (3.1) holds. Conditions ensuring this for a given x are bounded variation in a neighborhood of x and continuity of $f_\mu(.)$ at x, see (Widder, 1946, Theorem 7.3). This can easily be verified to hold for all $x > 0$ for a specific distribution by investigating its Laplace exponent, e.g., by applying (Sato, 1999, Proposition 28.1), which states that f_μ is in $\mathcal{C}^1$ if $\int_{\mathbb{R}} |\exp(-\Psi(-i\,s))\,s|\mathrm{d}s < \infty$. Alternatively, if the Lévy density g is known in closed form, one can try to verify sufficient conditions starting from (Steutel and van Harn, 2004, Proposition III.4.16), who state that f_μ fulfills the following functional equation[2]:

$$x\, f_\mu(x) = \int_0^x f_\mu(x-u)\, u\, g(u)\,\mathrm{d}u, \quad x > 0.$$

[2] For example, continuity can be shown if $u\,g(u)$ is bounded on every interval $(0,x]$, and $f_\mu \in \mathcal{C}^1$ can be shown along the lines of the proof of (Elstrodt, 1999, Proposition V.3.7) if the derivative of $u\,g(u)$ is uniformly continuous.

One can further simplify Equation (3.1) by

$$
\begin{aligned}
f_\mu(x) &= \lim_{R\to\infty} \frac{1}{2\,\pi\,i} \int_{a-i\,R}^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z \\
&= \lim_{R\to\infty} \frac{1}{2\,\pi\,i} \Big(\int_{a}^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z - \int_{a}^{a-i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z \Big) \\
&= \lim_{R\to\infty} \frac{1}{2\,\pi\,i}\, 2\,i\, \Im\Big(\int_{a}^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z \Big) \\
&= \lim_{R\to\infty} \frac{1}{\pi}\, \Im\Big(\int_{a}^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z \Big),
\end{aligned} \tag{3.2}
$$

where the third equality is valid as the integrand satisfies $\overline{f(\overline{s})} = f(s)$. Consequently, even though we started from a Cauchy principal value (the upper and lower bound simultaneously going to infinity), we end up with a one-sided limit. However, from that we can not derive that the limiting integral in the last equality exists, as this is not guaranteed for its real part. But when stating the resulting formula as the integral over the imaginary part, one gets a non-diverging integral.

The main formula of this chapter is given in the following theorem. It provides an alternative, convenient integral representation for the density of an arbitrary distribution of the Bondesson class, given in terms of its Laplace exponent. For a specific family of complete Bernstein functions, the involved imaginary part can be resolved easily.

Theorem 3.3 (Main representation)
If μ is a distribution of the Bondesson class with Laplace exponent Ψ and if Equation (3.1) *holds for the density f_μ and $x > 0$, then*

$$
f_\mu(x) = \frac{M\,e^{x\,a}}{\pi} \int_0^1 \Im\Big(e^{-x\,M\,\log(v)\,(b\,i-a)}\, e^{-\Psi(a-M\,\log(v)\,(b\,i-a))}\,(b\,i-a) \Big)\,\frac{\mathrm{d}v}{v}, \tag{3.3}
$$

with arbitrary parameters $a, b > 0$ and $M > 2/(a\,x)$. This integral is a proper Riemannian integral as one can show that the integrand vanishes for $v \searrow 0$.

Remark 3.4 (Basic ideas)
This theorem is based on two ideas:

1. *It is well known that integration along the Bromwich contour $\{z(u) = a + i\,u \,:\, 0 \le u < \infty\}$ is numerically challenging, since the integrand is typically highly oscillating due to the first factor $e^{x\,z} = e^{a}\,e^{i\,u}$. For an exemplary integrand when computing the stable density, see Figure 3.1. Therefore, it is often more convenient to perform the Laplace inversion along a different contour ending in the left half plane, as*

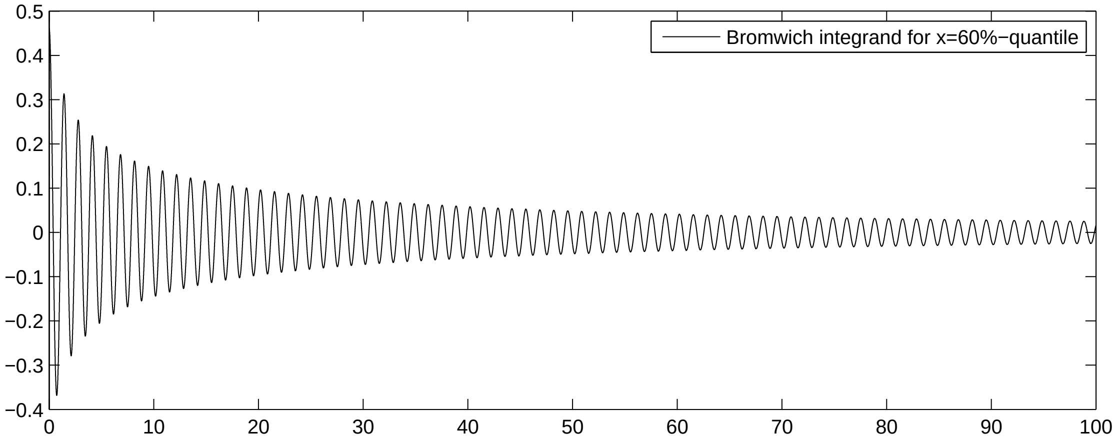


Figure 3.1 Bromwich integrand (i.e. the resulting integrand when evaluating Equation (3.2)) on $[0, 100]$ for x the 60%-quantile of the stable distribution with $\alpha = 0.3$, $a = 1/x$.

along such a contour, the absolute value of the first factor $|e^{x\,z}| = \exp(x\,\Re(z))$ *decreases. This can also be seen in Figure 3.2. A popular example for such a path transform is Talbot (1979). However, changing the integration contour in general requires additional conditions on the involved Laplace transforms, which can be very hard to verify. One can show for instance that positive stable distributions as presented in Section 2.2.2 with* $\alpha > 0.5$ *do not satisfy the sufficient conditions stated in Talbot (1979). To recognize this, consider* $s = y\,i - u, u > 0$*, for an arbitrary* $y > 0$*. In the stable case, we have*

$$\left|e^{-\Psi(s)}\right| = e^{-(y^2+u^2)^{\alpha/2}\cos(\alpha\varphi(u))}, \quad \varphi(u) := \pi - \arctan(y/u).$$

As for $\alpha > 0.5$, $\exists u_0 : \forall u > u_0, \cos(\alpha\varphi(u)) < 0$*, the absolute value of the Laplace transform does not vanish for* $u \nearrow \infty$*, thus violating the sufficient condition* $|\exp(-\Psi(s))| \to 0$ *uniformly in* $\Re(s) \leq 0$ *for* $|s| \to \infty$*, which is stated in Talbot (1979). Kiesel and Lutz (2011) perform the Laplace inversion of an integrated CIR-process along the contour*

$$\gamma(u) = a + u\,(b\,i - a), \quad 0 \leq u < \infty, \quad a, b > 0. \tag{3.4}$$

The first step consists of showing that this contour is always admissible[3] *for our*

[3] We will see in Section 3.3 that it is indeed also more convenient than the Bromwich contour.

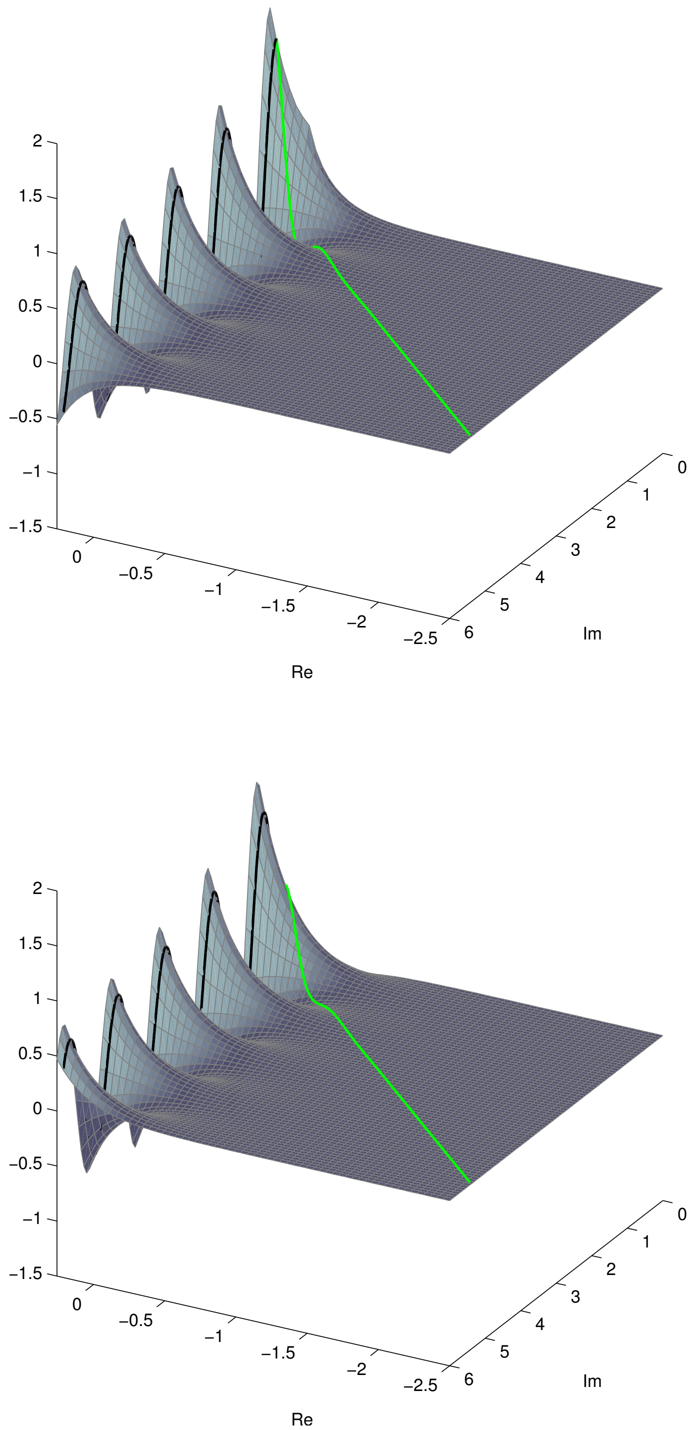


Figure 3.2 $\Re(e^{x\,z}\,e^{-\Psi(z)})$ (top figure) and $\Im(e^{x\,z}\,e^{-\Psi(z)})$ (bottom figure) for the stable distribution with $\alpha = 0.3$ and x the respective 60%-quantile. The Bromwich contour (black line) and the transformed path (green line) for parameters $a = 1/x$ and $b = 2/x$ are plotted.

setup of a distribution of the Bondesson class.

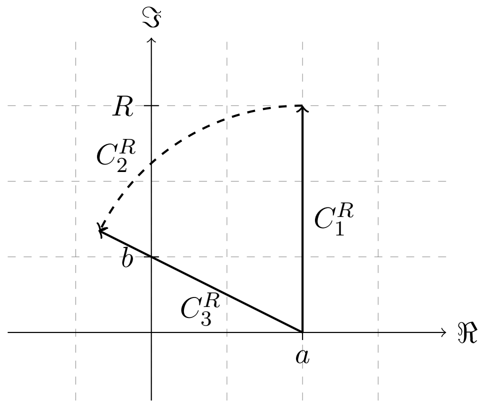
2. *Having transformed the contour, one still faces the problem of truncation, as the integral is indefinite. In a second step, a substitution to a finite integration interval, the interval* $[0,1]$*, is applied. It is shown that the resulting integrand vanishes for* $v \searrow 0$ *and we can thus extend the integrand continuously, yielding a proper Riemannian integral.*

Proof (of Theorem 3.3)
Step (1): *In a first step, we show that it is possible to transform the integral in Equation* (3.2) *to the new path* γ*, as defined by Equation* (3.4)*, which corresponds to*

$$\lim_{R\to\infty} \frac{1}{\pi} \Im\Big(\int_a^{a+i\,R} e^{x\,z}\, e^{-\Psi(z)}\, \mathrm{d}z \Big) = \lim_{R\to\infty} \frac{1}{\pi} \Im\Big(\int_{\gamma^R} e^{x\,z}\, e^{-\Psi(z)}\, \mathrm{d}z \Big), \quad x > 0, \tag{3.5}$$

where γ^R *denotes the path of* γ *for* $0 \le u \le R$*. Using Lemma 3.1, we know that the integrand has an analytic extension to* $\mathbb{C}\setminus(-\infty, 0]$*. The equivalence is thus proven using Cauchy's Theorem (for an elegant introduction, see, e.g., Jänich (2004)): For* $R > 0$*, consider the following closed contour in the upper half plane consisting of three parts* C_1^R, C_2^R*, and* C_3^R*:*

$$\begin{aligned} C_1^R &: a + i\,u, \quad 0 \le u \le R, \\ C_2^R &: a + R\,e^{i\,u}, \quad \frac{\pi}{2} \le u \le \pi - \arctan(b/a), \\ C_3^R &: a + u\,(b\,i - a), \quad 0 \le u \le \frac{R}{\sqrt{a^2+b^2}}. \end{aligned}$$



By Cauchy's Theorem, it holds that

$$0 = \int_{C_1^R} e^{x\,z}\, e^{-\,\Psi(z)}\, \mathrm{d}z + \int_{C_2^R} e^{x\,z}\, e^{-\,\Psi(z)}\, \mathrm{d}z - \int_{C_3^R} e^{x\,z}\, e^{-\,\Psi(z)}\, \mathrm{d}z.$$

Therefore, it remains to show that

$$\lim_{R\to\infty} \int_{C_2^R} e^{x\,z}\, e^{-\Psi(z)}\, \mathrm{d}z = 0.$$

This statement requires a very technical proof that is stated in Lemma 3.5 below. As a result of Step (1), with $\frac{\partial}{\partial u}\gamma(u) = b\,i - a$, we derive the representation

$$f_\mu(x) = \frac{1}{\pi} \int_0^\infty \Im\Big(e^{x\,\gamma(u)}\, e^{-\Psi(\gamma(u))}\,(b\,i - a)\Big)\,\mathrm{d}u. \tag{3.6}$$

Step (2): *When numerically evaluating Equation* (3.6)*, one still faces the problem of truncation as the interval is unbounded. Thus, we apply a substitution to the finite interval* $(0,1]$*, defining* $u = -M\,\log(v), v \in (0,1]$*, with a constant* $M > 0$*. This yields*

$$f_\mu(x) = \frac{M}{\pi} \int_0^1 \Im\Big(e^{x\,(a - M\,\log(v)\,(-a + b\,i))}\, e^{-\Psi(a - M\,\log(v)\,(-a+b\,i))}\,(b\,i - a)\Big)\,\frac{\mathrm{d}v}{v},$$

which as such does not remove the difficulties as this still could represent an improper integral at $v = 0$*. However, we can show that for* $M > 2/(a\,x)$*, the new integrand vanishes for* $v \searrow 0$ *and thus, the integral can be considered a proper integral on a finite interval. Let* h *denote the integrand of the integral before substitution,*

$$h(u) = \Im\Big(e^{x\,\gamma(u)}\, e^{-\Psi(\gamma(u))}\,(b\,i - a)\Big), \quad u > 0.$$

We have

$$|h(u)| \le \sqrt{a^2 + b^2}\,\Big|e^{x\,(a + u\,(b\,i - a))}\, e^{-\Psi(a + u\,(b\,i - a))}\Big|.$$

As for $u > 1$*, we can rewrite* $a + u(b\,i - a) = a + u\,\sqrt{a^2+b^2}\,\exp(i\,\varphi)$*, with* $\varphi = \pi - \arctan(b/a)$*, we can use the additional estimate stated in Lemma 3.5, Part* (iii)*. Consequently, there exists* $u_0 > 0$ *(using the notation of Lemma 3.5, one could write* $u_0 = R_0/\sqrt{a^2+b^2}$*) such that for* $u > u_0$*,*

$$|h(u)| \le \underbrace{\sqrt{a^2+b^2}\, e^{a\,x + 1}}_{:=\hat{c}}\, \exp\Big(\frac{u\,x}{2}\,\underbrace{\sqrt{a^2+b^2}\,\cos(\varphi)}_{=-a}\Big) \le \hat{c}\, e^{-\frac{a\,u\,x}{2}}.$$

Thus, for the integrand after substitution, $h(-M\,\log(v))/v$*, for small values of* v *it holds*

$$\Big|\frac{h(-M\,\log(v))}{v}\Big| \le \hat{c}\,\frac{v^{M\,\frac{a\,x}{2}}}{v} = \hat{c}\,v^{M\frac{ax}{2} - 1}.$$

Therefore, we can conclude that for $M > 2/(a\,x)$ *the new integrand vanishes for* $v \searrow 0$.□

Lemma 3.5 (Technical lemma)
Under the conditions stated in Theorem 3.3 and for the contour C_2^R *defined in the proof*

of that theorem,

$$\lim_{R\to\infty} \int_{C_2^R} e^{x\,z}\, e^{-\Psi(z)}\,\mathrm{d}z = 0, \quad x > 0.$$

Proof

(i) *We start by proving the following helpful statement. For a similar statement, see (Schilling et al., 2010, Proof of Corollary 6.5):*
For $u \in [\pi/2, \pi - \arctan(b/a)]$ and $t > 0$, we have

$$\Big|\frac{1}{a+t+R\,e^{i\,u}}\Big| \le \underbrace{\frac{\sqrt{2}}{\sqrt{1-\cos(\arctan(b/a))}}}_{=:c_{a,b}} \, \frac{1}{R+t+a}.$$

Proof:
Considering the function

$$f_{u,R} : [-(R+a), \infty) \to [0,\infty), \quad t \mapsto \frac{R+t+a}{|a+t+R\,e^{i\,u}|},$$

we want to show that for every $u \in [\pi/2, \pi - \arctan(b/a)]$

$$f_{u,R}(t) = \frac{R+t+a}{\sqrt{(a+t+R\,\cos(u))^2 + (R\,\sin(u))^2}} \le c_{a,b}.$$

It holds that $\lim_{t\to\infty} f_{u,R}(t) = 1$ and $f_{u,R}(\cdot)$ is continuous. Further,

$$\frac{\partial}{\partial t} f_{u,R}(t) = \frac{1}{\sqrt{(a+t+R\,\cos(u))^2 + (R\,\sin(u))^2}} \times \Big(1 - \frac{(R+t+a)(R\,\cos(u)+t+a)}{(a+t+R\,\cos(u))^2 + (R\,\sin(u))^2}\Big).$$

It follows that

$$\begin{aligned} & \frac{\partial}{\partial t} f_{u,R}(t) \le 0, \\ \Leftrightarrow \quad & (R+t+a)(R\,\cos(u)+t+a) \ge (a+t+R\,\cos(u))^2 + (R\,\sin(u))^2, \\ \Leftrightarrow \quad & (t+a)\,R\,(1-\cos(u)) \ge R^2(1-\cos(u)), \\ \Leftrightarrow \quad & t \ge R-a. \end{aligned}$$

Thus, for u in the given interval,

$$\max_{t\in[0,\infty)} f_{u,R}(t) \leq \max_{t\in[-(R+a),\infty)} f_{u,R}(t) = f_{u,R}(R-a) = \frac{2R}{R\,|1+e^{i\,u}|}$$
$$= \frac{2}{\sqrt{2+2\cos(u)}} = \frac{\sqrt{2}}{\sqrt{1+\cos(u)}} \leq \frac{\sqrt{2}}{\sqrt{1-\cos(\arctan(b/a))}} = c_{a,b}.$$

(ii) *Now, we show that we can find $R_0 > 0$ such that $\forall R > R_0$ and $u \in [\pi/2, \pi - \arctan(b/a)]$, it holds that*

$$\exp\left(-a\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)\right) \leq e,$$

and

$$\exp\left(-R\cos(u)\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)\right) \leq \exp\Big(-R\,\cos(u)\,\frac{x}{2}\Big).$$

Proof:
Using the statement in (i), we can show that

$$\lim_{R\to\infty}\left|\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right| = 0,$$

uniformly in $u \in [\pi/2, \pi - \arctan(b/a)]$, which is basically a slight extension of (Schilling et al., 2010, Corollary 6.5), where $a = 0$ is considered. From this, the two claims will be derived. For $u \in [\pi/2, \pi - \arctan(b/a)]$, using the Stieltjes representation of Ψ, it holds that

$$\left|\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right| = \left|\int_0^\infty \frac{1}{a+R\,e^{i\,u}+t}\,\sigma(\mathrm{d}t)\right| \leq \int_0^\infty \left|\frac{1}{a+R\,e^{i\,u}+t}\right|\sigma(\mathrm{d}t)$$
$$\overset{(i)}{\leq} c_{a,b}\int_0^\infty \frac{1}{a+R+t}\,\sigma(\mathrm{d}t).$$

Using $1/(R+t+a) \leq 1/(t+1)$ for $R > 1-a$ and $\int_0^\infty 1/(1+t)\,\sigma(\mathrm{d}t) < \infty$, dominated convergence yields the convergence of the right-hand side to zero for $R \to \infty$. Thus, uniform convergence to zero of the left-hand side is shown. Consequently, we can *find $R_0 > 0$ such that $\forall R > R_0$ and $u \in [\pi/2, \pi - \arctan(b/a)]$, it holds that*

$$\left|\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)\right| \leq \left|\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right| \leq \min\Big(\frac{1}{a}, \frac{x}{2}\Big).$$

From this follows, for all $R > R_0$,

$$\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right) \geq -\frac{1}{a},$$

and thus

$$\exp\left(-a\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)\right) \leq \exp\Big(-a\Big(-\frac{1}{a}\Big)\Big) = e.$$

Furthermore, for all $R > R_0$,

$$\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right) \leq \frac{x}{2},$$

and thus, as $\cos(u) \leq 0$ for $u \in [\pi/2, \pi - \arctan(b/a)]$, it follows that

$$\exp\left(\underbrace{-R\cos(u)}_{\geq 0}\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)\right) \leq \exp\Big(-R\,\cos(u)\,\frac{x}{2}\Big).$$

(iii) For $R > R_0$ from (ii), with $u \in [\pi/2, \pi - \arctan(b/a)]$, one can show the following estimate:

$$\left|e^{x\,(a+R\,e^{i\,u})}\,e^{-\Psi(a+R\,e^{i\,u})}\right| \leq e^{g_a(R)}\,e^{x\,a+1}\,e^{\frac{x}{2}\,R\,\cos(u)},$$

with

$$g_a(R) := -\int_{(0,\infty)} \frac{R^2\,\sin(\pi - \arctan(b/a))^2}{(t+a)^2 + R^2}\,\sigma(\mathrm{d}t).$$

From this, one can also deduce the following statement, which is helpful for Step (2) of Theorem 3.3. Using that $g_a(R) \leq 0$, it holds that

$$\left|e^{x\,(a+R\,e^{i\,u})}\,e^{-\Psi(a+R\,e^{i\,u})}\right| \leq e^{x\,a+1}\,e^{\frac{x}{2}\,R\,\cos(u)}.$$

Proof:
Consider $z := a + R\,e^{i\,u}$ *with* $a, R > 0$ *and* $u \in [\pi/2, \pi - \arctan(b/a)]$. *We can rewrite*

$$\left|e^{-\Psi(z)}\right| = \exp\Big(-\Re\Big(z\,\frac{\Psi(z)}{z}\Big)\Big) = \exp\Big(-\Re(z)\,\Re\Big(\frac{\Psi(z)}{z}\Big) + \Im(z)\,\Im\Big(\frac{\Psi(z)}{z}\Big)\Big).$$

Furthermore, it holds that

$$\begin{aligned}
\Im(z)\,\Im\Big(\frac{\Psi(z)}{z}\Big) &= R\,\sin(u)\int_{(0,\infty)} \underbrace{\Im\Big(\frac{1}{t+a+R\,e^{i\,u}}\Big)}_{=\frac{-R\,\sin(u)}{(t+a+R\,\cos(u))^2+(R\,\sin(u))^2}}\,\sigma(\mathrm{d}t)\\
&= -\int_{(0,\infty)} \frac{R^2\,\sin(u)^2}{(t+a+R\,\cos(u))^2+(R\,\sin(u))^2}\,\sigma(\mathrm{d}t)\\
&\leq -\int_{(0,\infty)} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a+R\,\cos(u))^2+(R\,\sin(u))^2}\,\sigma(\mathrm{d}t)\\
&= -\int_{(0,\infty)} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2+2\,(t+a)\,R\,\cos(u)}\,\sigma(\mathrm{d}t)\\
&\leq -\int_{(0,\infty)} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2}\,\sigma(\mathrm{d}t)\\
&= g_a(R).
\end{aligned}$$

Therefore, we obtain for $R > R_0$ *from (ii)*

$$\begin{aligned}
\left|e^{x\,(a+R\,e^{i\,u})}\,e^{-\Psi(a+R\,e^{i\,u})}\right| &\leq e^{x\,(a+R\,\cos(u))}\,e^{-\Re(a+R\,e^{i\,u})\,\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)}\,e^{g_a(R)}\\
&= e^{x\,(a+R\,\cos(u))}\,e^{-(a+R\,\cos(u))\,\Re\left(\frac{\Psi(a+R\,e^{i\,u})}{a+R\,e^{i\,u}}\right)}\,e^{g_a(R)}\\
&\overset{(ii)}{\leq} e^{x\,(a+R\,\cos(u))}\,e\,e^{-R\,\cos(u)\,\frac{x}{2}}\,e^{g_a(R)}\\
&= e^{g_a(R)}\,e^{x\,a+1}\,e^{\frac{x}{2}\,R\,\cos(u)}.
\end{aligned}$$

(iv) *Considering the initial expression, it follows for* $R > R_0$ *from (ii)*

$$\begin{aligned}
\Big|\int_{C_2^R} e^{x\,z}\,e^{-\Psi(z)}\,\mathrm{d}z\Big| &= \Big|\int_{\pi/2}^{\pi-\arctan(b/a)} e^{x\,(a+R\,e^{i\,u})}\,e^{-\Psi(a+R\,e^{i\,u})}\,R\,i\,e^{i\,u}\,\mathrm{d}u\Big|\\
&\leq R\int_{\pi/2}^{\pi-\arctan(b/a)} \left|e^{x\,(a+R\,e^{i\,u})}\,e^{-\Psi(a+R\,e^{i\,u})}\right|\,\mathrm{d}u
\end{aligned}$$

$$\begin{aligned}
&\overset{(iii)}{\leq} R \int_{\pi/2}^{\pi-\arctan(b/a)} e^{g_a(R)}\, e^{x\,a+1}\, e^{\frac{x}{2}\,R\,\cos(u)}\,\mathrm{d}u \\
&\leq R\, e^{g_a(R)}\, e^{x\,a+1} \int_{\pi/2}^{\pi} e^{\frac{x}{2}\,R\,\cos(u)}\,\mathrm{d}u \\
&= R\, e^{g_a(R)}\, e^{x\,a+1} \int_{0}^{\pi/2} e^{-\frac{x}{2}\,R\,\sin(u)}\,\mathrm{d}u \\
&\leq R\, e^{g_a(R)}\, e^{x\,a+1} \int_{0}^{\pi/2} e^{-\frac{x\,u\,R}{\pi}}\,\mathrm{d}u \\
&\leq e^{g_a(R)}\, e^{x\,a+1}\,\frac{\pi}{x}\left(1-e^{-\frac{x\,R}{2}}\right) \leq e^{g_a(R)}\, e^{x\,a+1}\,\frac{\pi}{x},
\end{aligned}$$

where in the second from last line, we estimate $\sin(u) \geq 2\,u/\pi$.

(v) *If we can show that*

$$e^{g_a(R)} \overset{R\to\infty}{\longrightarrow} 0,$$

then by (iv), the claim follows.
Proof:
We see that

$$\begin{aligned}
0 \leq \lim_{R\to\infty} e^{g_a(R)} &= e^{\lim_{R\to\infty} g_a(R)} \\
&= e^{-\lim_{R\to\infty}\int_{(0,\infty)} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2}\,\sigma(\mathrm{d}t)},
\end{aligned}$$

where the actual existence of these limits follows from the following considerations. Using Fatou's Lemma,

$$\begin{aligned}
\liminf_{R\to\infty} \int_{(0,\infty)} &\frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2}\,\sigma(\mathrm{d}t) \\
&\geq \int_{(0,\infty)} \liminf_{R\to\infty} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2}\,\sigma(\mathrm{d}t) \\
&= \int_{(0,\infty)} \sin(\pi-\arctan(b/a))^2\,\sigma(\mathrm{d}t) \\
&= \sin(\pi-\arctan(b/a))^2\,\sigma((0,\infty)) = \infty.
\end{aligned}$$

Here, Corollary 2.12 is used which states that $\sigma((0,\infty)) = \infty$ *is necessary for the existence of a density. It follows that*

$$\lim_{R\to\infty} \int_{(0,\infty)} \frac{R^2\,\sin(\pi-\arctan(b/a))^2}{(t+a)^2+R^2}\,\sigma(\mathrm{d}t) = \infty,$$

and, thus,

$$\lim_{R\to\infty} e^{g_a(R)} = 0. \qquad \square$$

Corollary 3.6 (Integral representation of distribution function)
If μ is a distribution of the Bondesson class, for every $x > 0$ one has

$$\mu([0,x]) = \frac{M\, e^{x\,a}}{\pi} \int_0^1 \Im\Big(e^{-x\,M\,\log v\,(b\,i-a)}\,\frac{e^{-\Psi(a-M\,\log v\,(b\,i-a))}}{a-M\,\log v\,(b\,i-a)}\,(b\,i-a)\Big)\,\frac{\mathrm{d}v}{v},$$

with arbitrary parameters $a, b > 0$ 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
The cumulative distribution function is obviously of bounded variation and furthermore, following from (Bondesson, 1981, Theorem 6.1), continuous at every $x > 0$. The Laplace transform of $\mu([0,x])$ is given by $\exp(-\Psi(z))/z$, which follows by application of Tonelli's Theorem. Thus, the result follows similarly to the proof of Theorem 3.3. Using the additional factor $1/z$ in Part (iv) of Lemma 3.5 allows to estimate $|R/z| = R/|a + R\exp(i\,u)|$ via a constant. Consequently, it is sufficient to estimate $\exp(g_a(R)) \leq 1$ and one does not need $\sigma((0,\infty)) = \infty$. Hence, the result holds without requiring the existence of a density. $\square$

Remark 3.7
The value of $\mu(\{0\})$ can be easily derived from ν. If $\nu\big((0,\infty)\big) = \sigma\big((0,\infty)\big) = \infty$, $\mu(\{0\}) = 0$ follows from Lemma 2.12. If $\lambda := \nu\big((0,\infty)\big) < \infty$, the corresponding distribution is a compound Poisson distribution and thus, $\mu(\{0\}) = \exp(-\lambda)$, see also (Bondesson, 1981, Theorem 6.1).

The following corollary is useful in the context of call option pricing via Laplace methods in the spirit of Carr and Madan (1999), Raible (2000), or Eberlein et al. (2010). More precisely, it can be used to compute expected values of the form $\mathbb{E}[(\exp(-X) - K)^+] = K\,\mathbb{E}[(\exp(-X - \ln(K)) - 1)^+]$, where the random variable X has a distribution from the Bondesson class. Expected values of the form $\mathbb{E}[(\exp(-X + y) - 1)^+]$ for $y > 0$ can be computed via Laplace inversion of $\exp(-\Psi(z))/(z\,(z-1))$, where Ψ denotes the Laplace exponent corresponding to X. An intuitive explanation for this (which holds in case X exhibits a density) is that the expected value (as a function of y) can be interpreted as the density of a convolution of the density of X and a modified payoff function, whose

Laplace transform is given by $1/(z\,(z-1))$ (for $\Re(z)>1$). Then, the Laplace transform of this convolution is given as the product of the Laplace transforms and we can apply Laplace inversion to get the expected value (for a given value of y, e.g. $y=-\ln(K)$). For a derivation of this result under very general conditions, see Eberlein et al. (2010). The following corollary can be used whenever one is pricing call options on an asset whose distribution has a representation as $\exp(-X)$, with X being distributed according to a Bondesson distribution. It might be less helpful for equity modeling (as the distribution has support $[0,1]$), but there are applications when pricing CDO tranches in models like Mai and Scherer (2009b), as is explicitly shown in Mai (2013).

Corollary 3.8 (Integral representation of an option-like structure)
Let $f(.)$ be a non-negative function with Laplace transform $\exp(-\Psi(z))/(z\,(z-1))$ (for $\Re(z)>1$), where $\Psi(.)$ denotes a complete Bernstein function. If the Bromwich inversion formula in Equation (3.1) (with $a>1$) is valid for $f(.)$, it holds that

$$f(x)=\frac{M\,e^{x\,a}}{\pi}\int_0^1 \Im\Big(\frac{e^{-x\,M\,\log v\,(b\,i-a)}\,e^{-\Psi(a-M\,\log v\,(b\,i-a))}}{(a-M\,\log v\,(b\,i-a))\,(a-1-M\,\log v\,(b\,i-a))}\,(b\,i-a)\Big)\,\frac{\mathrm{d}v}{v},$$

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
The restriction $a>1$ is needed, since we know that the Laplace transform can be analytically extended to $\mathbb{C}\setminus(-\infty,1]$. The result follows similarly to the proof of Theorem 3.3. Analogously to the previous corollary, using the additional factor $1/(z\,(z-1))$, it is sufficient to estimate $\exp(g_a(R))\leq 1$ and one does not need $\sigma((0,\infty))=\infty$, hence the distribution corresponding to Ψ need not be absolutely continuous. $\square$

Remark 3.9 (A remark on hyperbolic contours)
Having a look at the proof of Lemma 3.5, in particular Part (iv), where the absolute value of the integral is estimated versus the integral of the absolute value, it becomes clear that also the integral over a contour which is a subset of C_2^R converges to zero with R approaching ∞. Consequently, one can show analogously to Step (1) in Theorem 3.3 that it is possible to transform the integral to the hyperbolic contours as defined in López-Fernández and Palencia (2004) (and revised in Weideman and Trefethen (2007)). This is due to the fact that those hyperbolic contours always enclose a linear contour as used here and, consequently, the connecting contour in Step (1) of an analogous proof can be defined as a subset of C_2^R.

3.3 Implementation and numerical tests

Having motivated and derived Equation (3.3), it remains to confirm that numerical results justify our deliberations and to show how to choose the free parameters a, b, and M. The issue of parameter choice is dealt with in Section 3.3.1. Furthermore, three numerical tests are conducted to investigate our approach: In Section 3.3.2, the accuracy of the derived representation is tested by comparing the results with analytically available densities, namely those of the Gamma and Inverse Gaussian distribution. Furthermore, in Section 3.3.3, the approach is used to compute the distribution function of the non-central χ^2 distribution and to check the results. In Section 3.3.4, the approach is compared with alternative numerical approaches available for deriving the stable density. Finally, the stability of the approach is examined when considering the Hartman–Watson density in Section 3.3.5.

3.3.1 Remarks on the parameter choice

So far, the choice of the free parameters a, b, and M has not been considered. One wants to choose the free parameters in a way that renders the resulting integral representation convenient for a computation by numerical methods. Though the general representation holds for all admissible parameter constellations, some might yield "nicer" integrands.

Considering a, we follow the argumentation and choice of Kiesel and Lutz (2011): a is chosen as $1/x$, which yields a constant factor $\exp(a\,x)$ for different values of x and, furthermore, prevents the factor from "exploding" for large values of x. Taking this into account, the condition on M ($M > 2/(a\,x)$, see Equation (3.3)) is simplified to $M > 2$ and we always chose $M = 3$, which is also proposed in Kiesel and Lutz (2011).

The remaining and most critical parameter choice is the choice of b. In general, there is a trade-off to acknowledge: On the one hand, small values for b correspond to a "faster" decrease of the factor $\exp(x\,z)$, as the real part is decreasing faster along the new contour. On the other hand, the smaller b, the closer the new contour gets to possible singularities along the negative real line. In most of the examples that we implemented, i.e. for the Inverse Gaussian, Gamma, non-central χ^2, and Hartman–Watson distribution, we found that the choice $b = 2a$ provides a robust algorithm, although we used $b = a$ for the Hartman–Watson distribution, as this yields a slightly smoother result. However, there are families of distributions where further investigations might be necessary. This is due to the fact that, even though a vanishing integrand at $v = 0$ is guaranteed, this does not prevent the integrand from oscillating close to zero, as our estimate in the proof of Theorem 3.3 holds for

$$v < \exp\left(-M \frac{R_0}{\sqrt{a^2+b^2}}\right),$$

for some $R_0 > 0$, only. When considering the stable distribution and choosing the parameter b arbitrarily, such problems can be observed. By using the example of the stable density, we explain considerations that might circumvent such problems.

For a more intuitive presentation, we consider the integrand before the transformation to the interval $[0,1]$. Problems of the final integrand close to zero correspond to the behavior of the previous integrand for large values of u. Considering the absolute value of the Laplace transform of a positive, stable distribution along the new contour, we have for $u > 1$

$$\left|e^{-\Psi(\gamma(u))}\right| = e^{-\Re(\gamma(u)^\alpha)} = e^{-\left((1-u)^2\, a^2+u^2\, b^2\right)^{\alpha/2}\,\cos(\alpha\,\varphi(u))},$$

with $0 < \varphi(u) := \pi - \arctan(b\,u/(a\,(u-1)) < \pi$. In order to prevent this term from "exploding", we seek to have $\cos(\alpha\,\varphi(u)) \geq 0$ for large u, which in this case is equivalent to $\alpha\,\varphi(u) \leq \pi/2$. For $\alpha \leq 0.5$, this is obviously true for every choice of b and consequently, we follow the previous argumentation and choose $b = 2a$. For $\alpha > 0.5$, one can show that

$$\alpha\,\varphi(u) \leq \pi/2, \quad \forall u > 1 \qquad \Leftrightarrow \qquad b \geq a\,\tan\left(\frac{\pi}{\alpha}\,(\alpha - 0.5)\right) := \hat{b}.$$

Consequently, we have shown that for $b \geq \hat{b}$, the absolute value of the Laplace transform considered is bounded by 1 along the corresponding path (at least for $u > 1$). As on the other hand we want to choose a small b, we set $b = \hat{b}$ for $\alpha > 0.5$ in the numerical test presented in Section 3.3.4.

3.3.2 Test against known probability densities

In a first step, the tractability of the derived representation is tested by numerically evaluating it for densities that are known in convenient analytical form. For the quadrature, a simple trapezoid rule is used as more sophisticated algorithms might cover up interesting effects. We consider the Gamma and the Inverse Gaussian distribution and compare the known densities to the numerical results. Both are parameterized by pairs

(β, η) and the combinations $(\beta, \eta) \in \{1, 1.5, 2, \ldots, 10\} \times \{1, 1.5, 2, \ldots, 10\}$ are taken into account. The densities are evaluated at 1000 equidistant points from the 0.5%-quantile to the 99.5%-quantile of the respective distribution. The trapezoid rule for evaluating the integral in Equation (3.3) is employed using 1000 equidistant grid points.

The high accuracy of the results indicated in Table 3.1 suggests a very tractable integrand. Consequently, the smooth behavior of the exemplary integrands depicted in Figure 3.3 is not surprising. Based on these results, it seems promising to investigate other, more "difficult" probability distributions.

Density	Abs. error (max/mean)	Rel. error (max/mean)
Gamma	4.51E-6/4.78E-7	5.57E-5/3.62E-6
IG	7.95E-6/3.50E-7	7.36E-5/1.41E-6

Table 3.1 Density calculation for Inverse Gaussian and Gamma distribution by numerical evaluation of Equation (3.3) and comparison to exact values. Maximum and average absolute and relative errors over all parameter combinations and evaluation points are listed.

3.3.3 Test using the non-central χ^2 distribution function

Having investigated the suitability of the new integral representation when testing against known probabilities, in a second step, we want to test it using the distribution function of the non-central χ^2 distribution. There are closed-form expressions available, e.g. via Equation (2.11), but implementing these is not a trivial task and has triggered some research, see, e.g., Larguinho et al. (2013) for a comparison of several approaches, of which Benton and Krishnamoorthy (2003) performs best. The aim of this section is to show that the approach presented is applicable and its straight-forward implementation (using Corollary 3.6) yields reliable results.

To check for the stability of this approach, we implement the resulting representation using MATLABs quadrature algorithm `quadgk` and compare the results to MATLABs built-in implementation of the non-central χ^2 distribution function, `ncx2cdf`. Regarding the investigated parameters, we follow Benton and Krishnamoorthy (2003) and, for several combinations of x and β, investigate 10000 randomly selected values of η.[4] Using

[4] To be precise, for the (x, β)-combinations $(5, 5), (12, 12), (400, 200), (300, 290)$, the values of η are drawn as absolute values of $\mathcal{N}(2, 1)$-distributed random variables, for the combinations

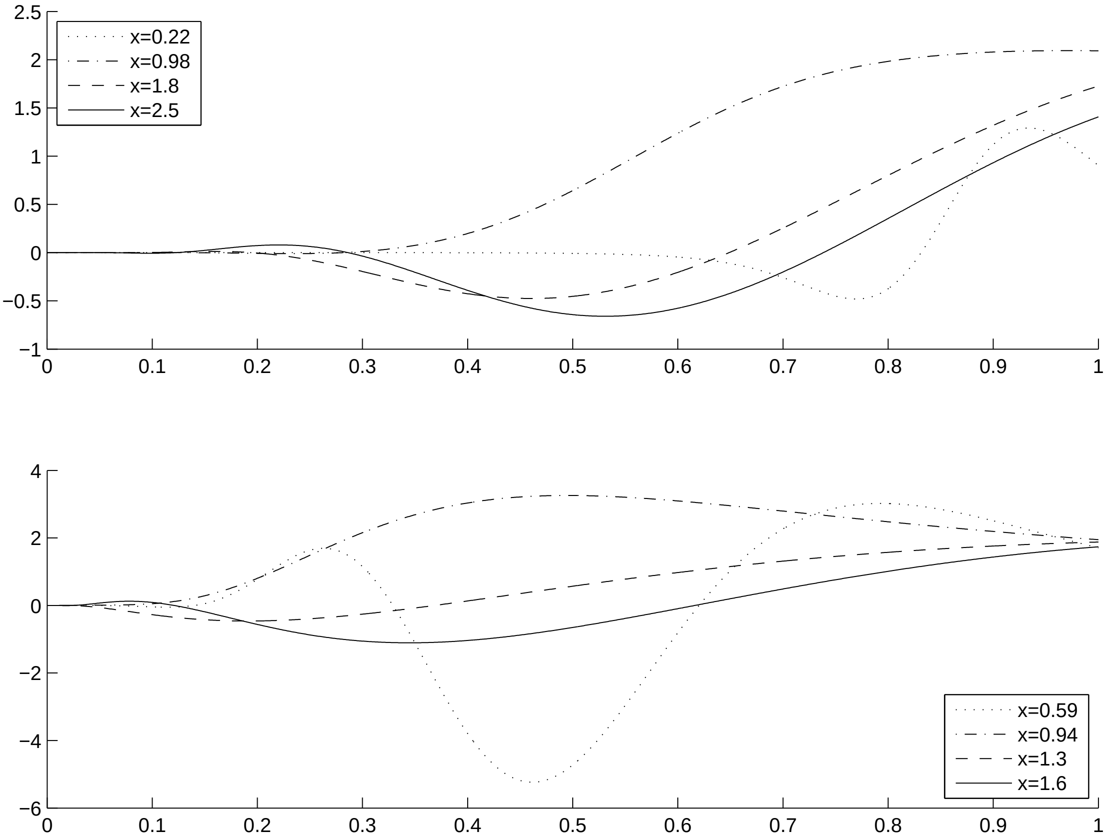


Figure 3.3 Resulting integrand in Equation (3.3) for the Gamma (upper figure) and the Inverse Gaussian (lower figure) distribution, $\beta = 5$, $\eta = 5$, with the values for x chosen as 4 equally spaced points starting, respectively ending, with the 0.5% and 99.5%-quantiles.

`quadgk` with its default settings, over all 160000 evaluations, this results in a maximum relative deviation of 3.97E-06 from MATLABs built-in function. The average relative error amounts to 1.97E-09. Changing the stopping criteria of the quadrature algorithm, the precision can be tuned arbitrarily. However, the lesson should not be to use our representation instead of the especially designed algorithm. Instead, it allows us to observe that even for such a distribution, whose implementation requires quite some attention, our representation yields a simple to implement and reliable algorithm. No additional considerations regarding the implementation were necessary. This might serve as a an encouraging sign with respect to other distributions for which, so far, no robust representations exist.

3.3.4 Test using the stable density

Since the results of the previous sections suggest the tractability of the approach presented, it is now tested in a more complex setup. Therefore, we consider the positive stable distribution. As mentioned in Section 2.2.2, there is in general no simple closed-form expression available for the corresponding density. We compare the results for the density based on the new representation to results based on Nolan's representation (Equation (2.9)), which is tailor-made for this problem.

Using Equation (2.8), it is sufficient to consider $f^{\mathrm{St}}_{\alpha,1}$ which corresponds to $\Psi^{\mathrm{St}}_{\alpha,1}(x) = x^{\alpha}$. All parameter values $\alpha \in \{0.01, 0.02, \ldots, 0.05, 0.10, 0.20, \ldots, 0.90, 0.95, 0.96, \ldots, 0.99\}$ are taken into account. As the distribution function is not known in closed form, Corollary 3.6 is used to compute quantiles. We will again consider 1000 points between the 0.5%-quantile and the 99.5%-quantile of each distribution, however, this time the points are log-spaced to locate more points in the area with higher probability mass, taking into consideration the heavy tails of the distribution. The parameters are set as described in Section 3.3.1 and as a quadrature algorithm for both approaches, the target-oriented Matlab built-in integration algorithm `quadgk` is used.

The results are described in Table 3.2, comprising the time needed for the computations and the maximum relative deviation of the new approach's results from Nolan's inver-

$(40, 20), (60, 40), (220, 200), (340, 280)$, the values of η are drawn as absolute values of $\mathcal{N}(20, 1)$-distributed random variables, for the combinations $(290, 10), (500, 220), (800, 520), (1500, 30)$, the values of η are drawn as absolute values of $\mathcal{N}(280, 1)$-distributed random variables, and for the combinations $(1000, 5), (1200, 200), (1300, 290), (1500, 30)$, the values of η are drawn as absolute values of $\mathcal{N}(1000, 1)$-distributed random variables.

sion formula's results. Two observations can be made: Firstly, the time needed for the computations is similar in both cases. Secondly, the results of the new approach deviate only marginally from the benchmark results based on Equation (2.9).[5] Keeping in mind that Nolan's inversion formula was specifically developed for stable densities, the results of the new representation seem astonishingly good as it is a very general formula applicable to all Bondesson densities. Thus, the results suggest that the corresponding integrand should be tractable and smooth. To illustrate the improvements made over the initial Bromwich integrand, Figure 3.4 depicts how the initially presented integrand in Figure 3.1 changes with the new representation. As expected, it is very smooth and tractable. In Figure 3.5, the integrand for two other parameters and values is depicted, a low value $\alpha = 0.1$, and a high value $\alpha = 0.9$. The parameters are chosen as described in Section 3.3.1, i.e. for $\alpha = 0.9$, the value of b is chosen as $\hat{b}$. If instead $b = 2\,a$ was chosen, for $x = 0.6$ one could observe the difficulties described in Section 3.3.1, i.e. an integrand oscillating considerably close to 0. This effect is more pronounced for even higher values of α in combination with small quantiles.

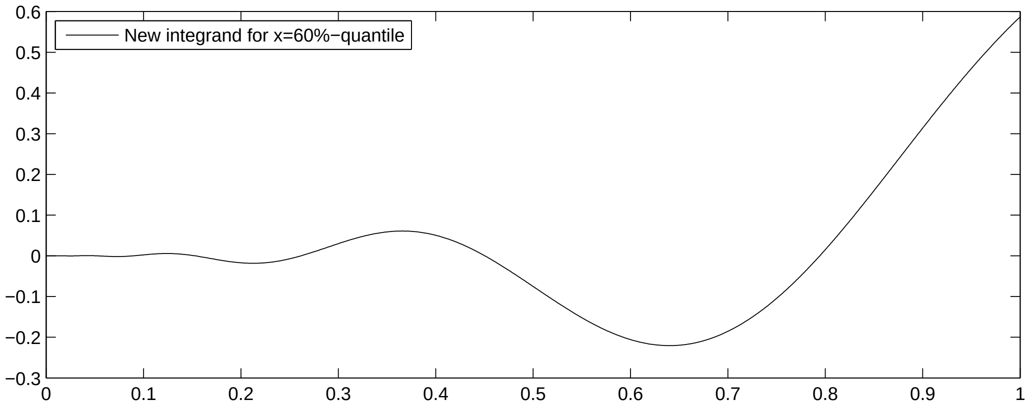


Figure 3.4 Resulting integrand in Equation (3.3) for the same parameter constellation as in Figure 3.1, with x the 60%-quantile of the stable distribution with $\alpha = 0.3$.

Completing the picture, we investigate the behavior of the resulting density outside the given quantile-based range. For that, we consider $\alpha = 0.01$, which obviously has the most extreme quantiles, and compute the density with our approach outside of the given interval. Even there, the computed density exhibits a smooth and regular behavior, as

[5]The comparably higher value for $\alpha = 0.3$ results from one single grid point, where `quadgk` has minor difficulties when integrating over Nolan's contour. Splitting the integral into two parts, this problem is removed.

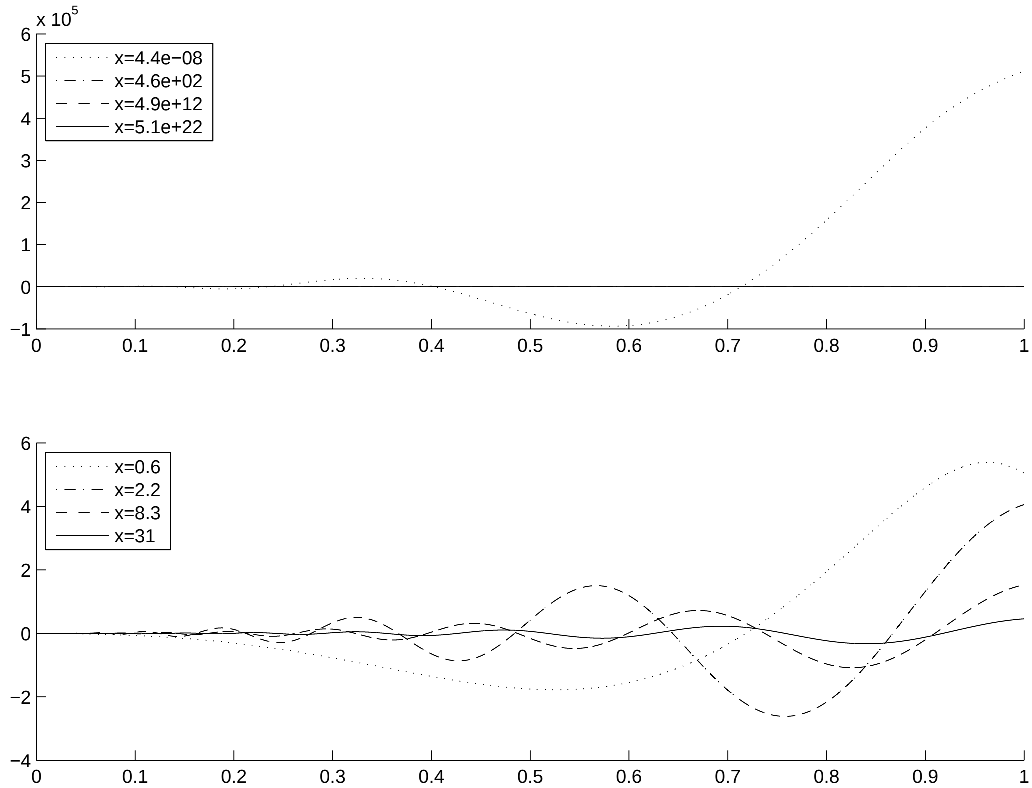


Figure 3.5 Resulting integrand in Equation (3.3) for the stable distribution with $\alpha = 0.1$ (upper figure, three of the lines coincide for the given scale) and $\alpha = 0.9$ (lower figure), with the values for x chosen as 4 log-spaced points starting, respectively ending, with the 0.5% and 99.5%-quantiles.

α	0.5%-quantile	99.5%-quantile	Time (N/T)	Max. relative deviation from N
0.01	2.24E-73	5.49E+229	1.42/1.05	1.40E-05
0.02	3.73E-37	5.49E+114	1.37/1.06	7.01E-06
0.03	4.52E-25	2.53E+76	1.26/1.06	6.47E-05
0.04	5.06E-19	1.71E+57	1.24/1.07	3.50E-06
0.05	2.18E-15	5.37E+45	1.22/1.08	3.15E-06
0.10	4.39E-08	5.14E+22	1.15/1.05	1.39E-06
0.20	2.35E-04	1.48E+11	1.08/1.07	1.47E-06
0.30	0.005	1.95E+07	1.04/1.10	1.00E-03
0.40	0.023	2.08E+05	1.03/1.13	3.11E-07
0.50	0.063	1.27E+04	1.03/1.27	5.65E-07
0.60	0.132	1.82E+03	1.25/1.06	1.20E-09
0.70	0.236	408.350	1.23/1.11	7.99E-09
0.80	0.385	113.914	1.28/1.42	3.12E-07
0.90	0.603	30.779	1.48/1.86	4.63E-07
0.95	0.758	12.708	1.70/2.43	2.11E-06
0.96	0.795	9.990	1.77/2.65	3.29E-06
0.97	0.836	7.491	1.85/2.93	7.60E-06
0.98	0.881	5.180	1.97/3.37	5.31E-06
0.99	0.933	3.028	2.14/4.17	1.07E-05

Table 3.2 Comparison of Nolan's inversion algorithm (N) and integration over the new representation (T) for stable distributions. The quantiles of the respective distribution and the time (CPU time in seconds) required by both methods to evaluate 1 000 density points are listed. Furthermore, the maximum relative deviation of (T) with respect to Nolan's approach is given.

can be seen in Figure 3.6.

3.3.5 Test using the Hartman–Watson density

Having observed the astonishing stability of numerical implementations of the stable density based on the new representation, testing it for other problematic cases seems natural. The Hartman–Watson distribution (see Section 2.2.2) represents an interesting candidate for this. The following aspects are partially published in Bernhart and Mai (2014a), where the approach is additionally compared to other Laplace inversion techniques. Though the Hartman–Watson distribution is of high relevance in particular in mathematical finance, the numerical evaluation of its density still poses a serious problem. Equation (2.14) states the best known representation but its instabilities are well documented, see, e.g., Barrieu et al. (2004); Ishiyama (2005). In this section, we investigate if the new representation suffers from similar problems. As there is no reliable benchmark method, only stability considerations are taken into account.

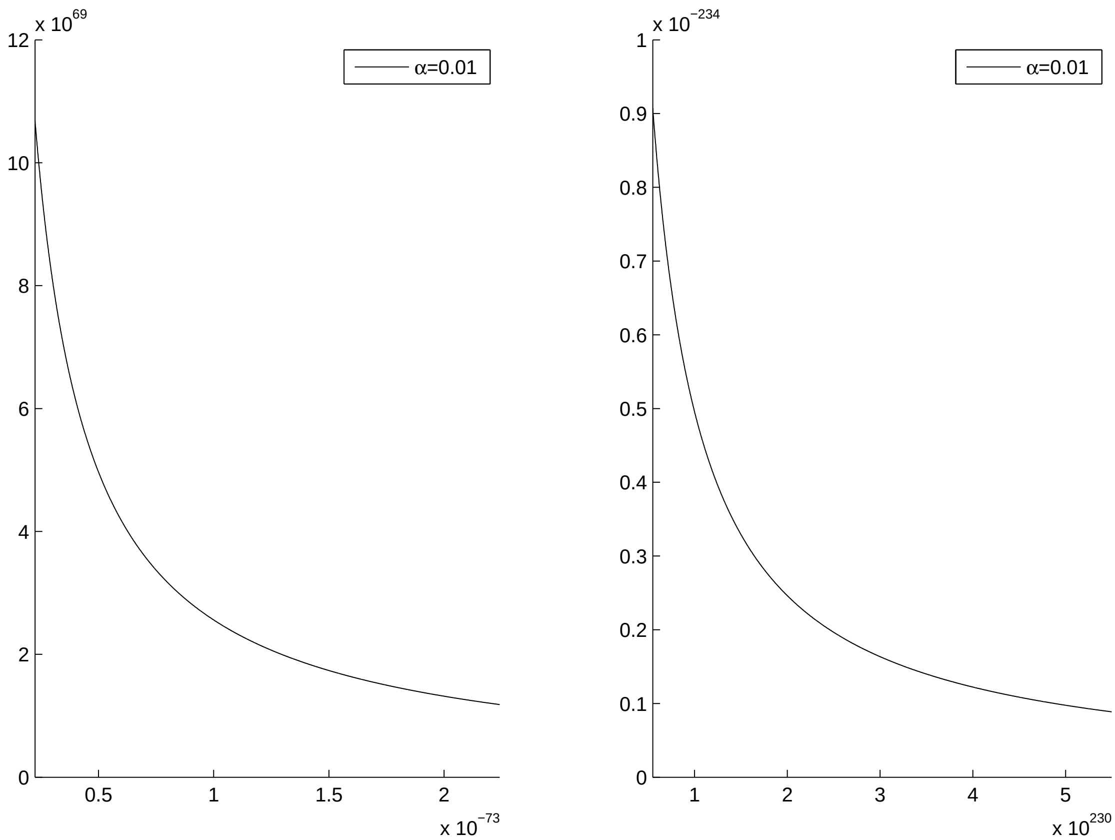


Figure 3.6 Density of the stable stable distribution with $\alpha = 0.01$ computed from the new representation outside of the quantile-based range used in Table 3.2, i.e. left of the 0.05%-quantile (left figure) and right of the 99.5%-quantile (right figure).

One particular challenge with this method is that the modified Bessel function I_ν needs to be evaluated for complex ν. A straight-forward implementation sufficient for our needs is achieved by using the partial sums related to the representation in Equation (2.13). It has the advantage that error bounds can be computed, as for $r > 0$ and $S_n^\nu(r) := \sum_{m=0}^n \frac{1}{m!\,\Gamma(m+\nu+1)} \left(\frac{r}{2}\right)^{2\,m+\nu}$, one can compute

$$|S_n^\nu(r) - I_\nu(r)| \leq \left(\frac{r}{2}\right)^{\Re(\nu)} \sum_{m=n+1}^{\infty} \frac{1}{m!\,|\Gamma(m+\nu+1)|} \left(\frac{r}{2}\right)^{2\,m}.$$

Using the Gamma functional equation $\Gamma(z+1) = \Gamma(z)\,z$, it is easy to see that $|\Gamma(z+1)| \geq |\Gamma(z)|$ for $|z| \geq 1$. Thus, for $n \geq -\Re(\nu) - 1$, the sequence $\{|\Gamma(m+\nu+1)|\}_{m=n+1,n+2,\ldots}$ is increasing, yielding

$$|S_n^\nu(r) - I_\nu(r)| \leq \frac{\left(\frac{r}{2}\right)^{\Re(\nu)}}{|\Gamma(n+\nu+2)|} \sum_{m=n+1}^{\infty} \frac{1}{m!} \left(\frac{r^2}{4}\right)^m,$$

where the series term is the residual of the Taylor expansion of $\exp(-r^2/4)$, which allows for a closed-form estimate. Consequently, one is able to choose n such that the modified Bessel function is approximated up to a required accuracy. Using the Gamma functional equation, one has to compute the complex Gamma function only once which further increases efficiency. The complex Gamma function is computed using the Lanczos approximation, see Lanczos (1964).[6] More evolved approaches for computing the modified Bessel function might be helpful when one is trying to fine-tune the accuracy of the numerical approach. As we are mainly interested in its stability, the implementation presented is sufficient for our needs. As in the previous section, Equation (3.3) is implemented using the MATLAB quadrature algorithm `quadgk`.

For small values of x, stability issues of Equation (2.14) are well known. Figure 3.7 replicates the problems described in Ishiyama (2005) and compares the resulting density to results derived with the new representation. It can be observed that the new approach does not suffer from similar problems. Indeed, the resulting integrand exhibits a tractable structure, as can be seen in Figure 3.8. Figure 3.7 also zooms in on the result for the new representation, illustrating that compared to the results in the previous section, the approach here obviously exhibits minor errors (as the value of the density is negative). However, they are not severe and are caused by the approximation of the modified

[6] We use the implementation of P. Godfrey published on `http://www.mathworks.com/matlabcentral/fileexchange/3572-gamma`.

Bessel function, which can be seen when increasing the accuracy of the approximation. An explanation for the very large quantiles in Figure 3.8 can be found in Figure 2.3, which illustrates that this law exhibits heavy tails for small values of r.

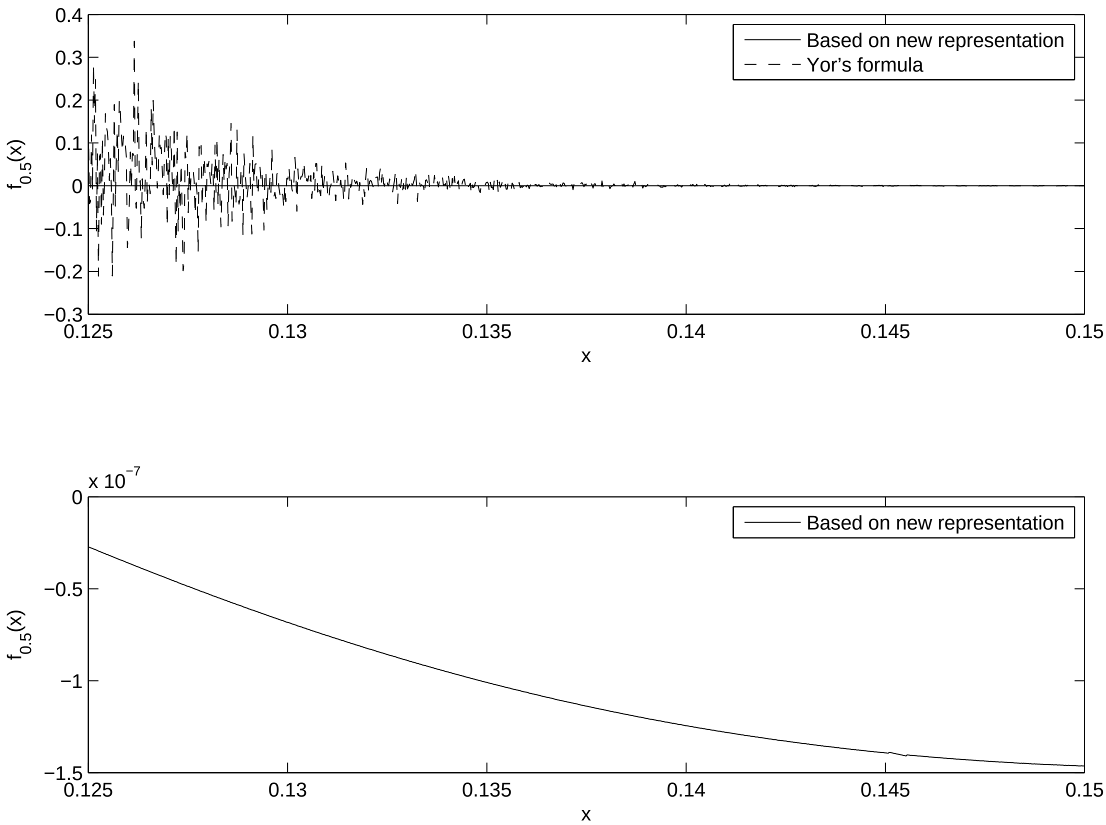


Figure 3.7 A comparison of the values for the Hartman–Watson density for $r = 0.5$ based on the approach presented and Equation (2.14) (upper figure). The lower figure zooms in on the results for the new approach.

3.4 Conclusion

In this chapter, a new integral representation for densities of the Bondesson class in terms of their Laplace exponent was derived. It was derived via a contour transform of the original Laplace inversion, a main contribution of the work presented being the

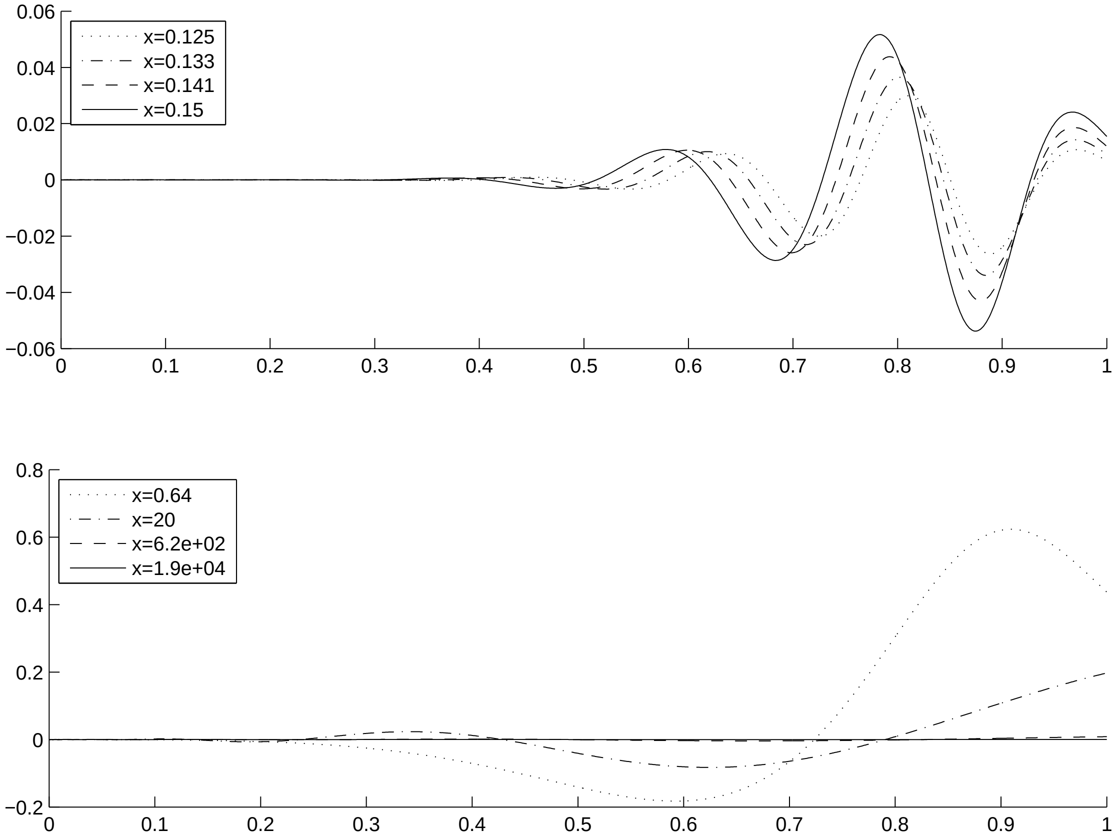


Figure 3.8 Resulting integrand in Equation (3.3) for the Hartman–Watson density for $r = 0.5$, with the values for x chosen as 4 log-spaced points in the interval in Figure 3.7 (upper figure), and starting, respectively ending, with the 0.5% and 99.5%-quantiles (lower figure).

proof of the applicability of the involved contour transform for all Bondesson distributions. Furthermore, as an interesting corollary, it was shown that the approach is also applicable for distribution functions and option-like structures. Numerical tests confirm the tractability and smoothness of the resulting representation. Consequently, building algorithms based on it yields nice results as the involved integrand is of a convenient form.[7] In general, applications might be found in all areas where positive infinitely divisible distributions or Lévy subordinators play a role.

[7] Actually, having implemented the formula for the stable density in several projects, we never experienced any stability issues.

4 Constructing MSMVE distributions from Bernstein functions

4.1 Motivation

In Section 2.4, we presented so-called MSMVE distributions, which constitute an important family of multivariate distributions, among others due to their relation to extreme-value distributions illustrated in Section 2.4.3. Furthermore, as a natural extension of the exponential distribution to higher dimensions, these are also interesting for applications such as credit portfolio models. A natural, widespread, and robust way to model the default times of single components of a credit portfolio (e.g. bonds or loans) is based on the exponential distribution. For a portfolio manager, it is daily business to express the default risk associated with a single credit-risky asset in terms of an exponential rate often called the "credit spread". However, it is not obvious how to model the joint distribution of the default times of multiple credit-risky assets. Though copula methods (see, e.g., Schönbucher and Schubert (2001); Cherubini et al. (2004)) are applicable, using a "true" multivariate exponential model seems to be more promising. In particular as the available data is usually scarce, exponential concepts should be used because such concepts naturally fit the intuitive motivation of modeling lifetimes.

Though the class of MSMVE distributions is very well studied on an abstract level, the number of known parametric families is rather small. Even though Molchanov (2008), Hofmann (2009), and Ressel (2013) have proven necessary and sufficient conditions for a function ℓ to provide a stable tail dependence function, these conditions are not easy to check analytically for a given function. Furthermore, the number of parametric models for which concrete stochastic representations (and thus sampling strategies) are known is even smaller. For most families, it is only known that the spectral representation of De Haan (1984) (see Lemma 2.45) exists or an implicit stochastic representation as a limit distribution (see Section 2.4.3) can be stated.

Motivated by the above considerations, there is some recent work aiming at the construction of new and flexible parametric families in high dimensions, see, e.g., Fougères et al. (2009); Durante and Salvadori (2010); Ballani and Schlather (2011); Segers (2012). In the same spirit, the aim of the present chapter is to develop new parametric models that have a convenient stochastic representation based on the representation stated in Theorem 2.46. We will develop two similar classes of MSMVE distributions, giving rise to a huge quantity of parametric stable tail dependence functions in arbitrary dimensions.[1] Furthermore, the underlying stochastic model can be used for efficient simulations, in particular in high dimensions, to construct non-exchangeable extensions (see Lemma 2.50), and to investigate statistical properties of the associated MSMVE distribution. As a side product, tractable examples of IDT processes are constructed, and related integral transforms are investigated.

To be more precise, using Lemma 2.31, we will construct tractable families of IDT subordinators by specifying two suitable choices of the function f, denoted f_1 and f_2. These families of processes yield MSMVE distributions via Theorem 2.46 respectively the IDT-frailty construction therein. Furthermore, they are tractable enough to allow for the computation of the required quantities characterizing the MSMVE distributions. Having investigated two instances of this general approach, a second aim of this chapter is to present a classification of all MSMVE distributions that can be constructed using IDT subordinators which are defined as in Lemma 2.31. We will show how they arise as limits of an extended shock-model construction. Such results help to improve the intuitive understanding of this class of distributions and the underlying dependence structure.

The class of distributions constructed here seems to be particularly useful for credit portfolio modeling. An intuitive understanding of the dependence structure used is crucial whenever the data available is scarce, as this requires an additional assessment of suitability. Furthermore, those portfolios can become quite large and it is often necessary to simulate from their distribution. Finally, the frailty construction employed has one further advantage: In the case of a large homogeneous portfolio, the relative portfolio loss for any point in time, i.e. the portion of defaulted entities in the portfolio until this point, can be efficiently approximated using a Glivenko–Cantelli type of reasoning, as applied, e.g., in Mai et al. (2013). This results in formulas comparable to the famous

[1] Similar to Joe (1990), the aim is to introduce new classes of models, while an application of these in practice requires a further, very detailed investigation of particular parametric families, which lies outside the scope of this work.

Vasicek formula, see, e.g., Vasicek (2002), which allow an efficient evaluation of more complex securities like loss tranches and an easy computation of risk measures like the portfolio Value-at-Risk.

The remainder of this chapter is structured as follows: In Section 4.2, a specific family of IDT subordinators (based on f_1) is investigated and the corresponding MSMVE distribution is determined. A similar example (based on f_2) is sketched in Section 4.3. A note on simulation can be found in Section 4.4 and in Section 4.5, the general class of distributions attainable by our construction is investigated. Section 4.6 concludes. This chapter is based on work I completed under supervision of Jan-Frederik Mai and Matthias Scherer. A condensed version of the results of this chapter can be found in Bernhart et al. (2015b).

4.2 A first construction based on $f_1(s) = (1-s)_+$

We examine the construction of Lemma 2.31 using the function $f_1(s) = (1-s)_+ := \max\{1-s, 0\}$, which obviously fulfills the conditions stated in Lemma 2.31 for every Lévy subordinator, i.e. we consider the IDT subordinator

$$H_t := \int_0^\infty \left(1 - \frac{s}{t}\right)_+ \mathrm{d}\Lambda_s = \int_0^t \left(1 - \frac{s}{t}\right) \mathrm{d}\Lambda_s, \quad t > 0, \tag{4.1}$$

with $H_0 := 0$. The process H has an alternative representation using integration by parts, namely

$$H_t := \frac{1}{t} \int_0^t \Lambda_s \mathrm{d}s, \quad t > 0, \tag{4.2}$$

which can be seen as some kind of moving average of the increasing process Λ. From Equation (4.1) it can be seen that pathwise, H_t equals a Williamson 2-transform evaluated at $1/t$, see, e.g., Williamson (1956) for the definition of Williamson d-transforms. Consequently, one can deduce that $H_t = \psi(1/t)$ with ψ a (random) convex and non-increasing function. It can be seen from the representation in Equation (4.2) that H_t equals the product of a differentiable function and a function that is a.e. differentiable, a.s.. Consequently, the paths of H are a.e. differentiable, a.s.. Furthermore,

$$\begin{aligned}H_{t+x} - H_x &= \frac{1}{t+x}\int_x^{x+t} \Lambda_s \mathrm{d}s - \frac{t}{x\,(x+t)}\int_0^x \Lambda_s \mathrm{d}s \\ &= \frac{1}{t+x}\int_x^{x+t} (\Lambda_s - \Lambda_x)\,\mathrm{d}s + \frac{t\,\Lambda_x}{x+t} - \frac{t}{x+t}H_x \\ &\overset{d}{=} \frac{1}{t+x}\int_0^t \tilde{\Lambda}_s \mathrm{d}s + \frac{t}{x+t}(\Lambda_x - H_x) = \frac{t}{x+t}\,\tilde{H}_t + \frac{t}{x+t}(\Lambda_x - H_x),\end{aligned}$$

where $\tilde{\Lambda}$ is an independent copy of Λ and $\tilde{H}$ the corresponding independent copy of H. Consequently, the increments, given the path of Λ up to time x, can be decomposed into a stochastic component independent of the previous evolution and a component measurable with respect to $\mathcal{F}_x := \sigma(\Lambda_s, 0 \leq s \leq x)$. However, as the value of Λ_x can not be recovered from H_x, H is not Markovian.

We start with an analysis of the marginal distributions of H that are attainable in this construction in Section 4.2.1, and then investigate the resulting MSMVE distribution in Section 4.2.2.

4.2.1 Attainable marginal distributions

In a first step, we analyze possible marginal distributions of H that can arise from this construction. This will be helpful when investigating the resulting MSMVE distribution in a second step. However, we also consider this of interest in its own right and, thus, present the results in some detail. Let Ψ_Λ be the Laplace exponent of Λ_1, ν_Λ the corresponding Lévy measure, and a_Λ its drift term. For the resulting IDT subordinator H, we denote its associated Bernstein function by Ψ_H with Lévy measure ν_H and drift a_H. Let Φ_{f_1} denote the integral transform considered, i.e. $\Phi_{f_1} : \mathcal{L}(\Lambda_1) \mapsto \mathcal{L}\left(\int_0^1 (1-s)\,\mathrm{d}\Lambda_s\right)$, which we will use simultaneously on the level of corresponding Lévy measures $\Phi_{f_1} : \nu_\Lambda \mapsto \nu_H$, see Remark 2.29. It can be shown that the resulting Lévy measure $\nu_H = \Phi_{f_1}(\nu_\Lambda)$ possesses a non-increasing density.

Lemma 4.1 (Lévy measures associated with H)

(i) The Lévy measure ν_H possesses a density g_H with respect to the Lebesgue measure, given by

$$g_H(y) = \int_y^\infty \frac{\nu_\Lambda(\mathrm{d}x)}{x}, \quad y > 0,$$

and the drift of H is given by $a_H := a_\Lambda/2$.

(ii) For any non-negative, measurable function h

$$\int_0^\infty h(x)\,\nu_H(\mathrm{d}x) = \int_0^\infty \frac{1}{x}\left(\int_0^x h(y)\,\mathrm{d}y\right)\nu_\Lambda(\mathrm{d}x).$$

Proof
We only prove (ii), as (i) can be proven along the same lines (see, e.g., (Barndorff-Nielsen et al., 2008, Example 6.3 (1))). For $H = \int_0^1 (1-s)\,\mathrm{d}\Lambda_s$, Equation (2.19) *in Theorem 2.26 yields*

$$\Psi_H(x) = \int_0^1 \Psi_\Lambda\big(x\,(1-s)\big)\,\mathrm{d}s, \quad x \geq 0,$$

as $f_1(s) = (1-s)_+$ is obviously integrable. One has

$$\nu_H(B) = \int_0^1 \nu_\Lambda\big(B/(1-s)\big)\,\mathrm{d}s = \int_0^1 \nu_\Lambda(B/s)\,\mathrm{d}s, \quad B \in \mathcal{B}(\mathbb{R}_+). \tag{4.3}$$

Thus, for any non-negative, measurable function h,

$$\begin{aligned}\int_0^\infty h(x)\,\nu_H(\mathrm{d}x) &= \int_0^1\int_0^\infty h(x\,s)\,\nu_\Lambda(\mathrm{d}x)\,\mathrm{d}s = \int_0^\infty\int_0^1 h(x\,s)\,\mathrm{d}s\,\nu_\Lambda(\mathrm{d}x)\\ &= \int_0^\infty \frac{1}{x}\left(\int_0^x h(y)\,\mathrm{d}y\right)\nu_\Lambda(\mathrm{d}x),\end{aligned}$$

proving the claim. □

Actually, one can show even more. As mentioned in Remark 2.29, it can be shown that $\Phi_{f_1} : \mathcal{M} \to U$ in a surjective manner, i.e. every Bernstein function possessing a non-increasing density can be reached by the given construction. This result has already been shown in Jurek (1985) for distributions on general Banach spaces, as he shows that all such distributions admit a representation as $\int_0^1 s\,\mathrm{d}\Lambda_s \overset{d}{=} \int_0^1 (1-s)\,\mathrm{d}\Lambda_s$. Even more, he showed that the mapping Φ_{f_1} is one-to-one. Furthermore, it can be seen from Equation (4.3) that Φ_{f_1} is a so-called "Upsilon transform" in the sense of Barndorff-Nielsen et al. (2008), with dilation measure $\gamma(\mathrm{d}x) = \mathbb{1}_{[0,1]}\mathrm{d}x$, from which more results can be derived, e.g., on continuity properties of the transform Φ_{f_1}.

We will present a detailed analysis of the integral transform. Not the complete analysis is necessary for the rest of the chapter, however, we hope that it is of interest on its own. Furthermore, the results allow to explicitly compute the pre-image of important distributions, which is crucial for finding the stochastic representation of the MSMVE laws defined in Section 4.2.2.

Lemma 4.2 (Analysis of Φ_{f_1})

(a) $\Phi_{f_1} : \mathcal{M} \to U$ *and the mapping is one-to-one.*

(b) $\Phi_{f_1}^{-1}\big(g(x)\mathrm{d}x\big) = -x\, g'(x)\mathrm{d}x$ *for g a c.m. Lévy density.*

(c) *It holds that* $\Phi_{f_1}^{-1}(T) \subset BO$ *and*

$$\begin{aligned}\Phi_{f_1}^{-1}(T) &= \Big\{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = f(x)\mathrm{d}x, \\ &\qquad\qquad \textit{with } f(x) = -x\, g'(x),\ \ g \textit{ a Thorin Lévy density}\Big\} \\ &= \left\{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = f(x)\mathrm{d}x, \right. \\ &\qquad\qquad \left. \textit{with } f(x) = \int_{(0,\infty)} \exp(-x\, s)\, \mathrm{d}(s\, w(s)),\ w \in W\right\},\end{aligned}$$

where

$$\begin{aligned}W := \Big\{w : (0,\infty) \to [0,\infty) \textit{ is non-decreasing, left-continuous,} \\ \int_0^\infty (1+t)^{-1}\, t^{-1}\, w(t)\, \mathrm{d}t < \infty\Big\}.\end{aligned}$$

(d) $\Phi_{f_1}(BO) \subset BO$*, but* $\Phi_{f_1}^{-1}(BO) \not\subset BO$*, and*

$$\begin{aligned}\Phi_{f_1}^{-1}(BO) = \Big\{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = f(x)\mathrm{d}x, \\ \textit{with } f(x) = x \int_{(0,\infty)} \exp(-x\, s) s\, \eta(\mathrm{d}s),\ \eta \in V\Big\},\end{aligned}$$

where

$$V := \left\{\eta \textit{ measure on } (0,\infty), \int_0^\infty (s^{-1} \wedge 1)\, s^{-1}\, \eta(\mathrm{d}s) < \infty\right\}.$$

(e)

$$\Phi_{f_1}^{-1}(L) = \left\{\nu \in \mathcal{M} : \nu(\mathrm{d}x) = \frac{f(x)}{x}\,\mathrm{d}x + \mathrm{d}(-f(x)),\ \textit{with } f \in F\right\},$$

where

$$F := \left\{ f : (0,\infty) \to [0,\infty),\ \textit{decreasing},\ \int_0^\infty (s^{-1} \wedge 1)\, f(s)\, \mathrm{d}s < \infty \right\}.$$

Proof

(a) *This is the result in (Jurek, 1985, Theorem 2.6) restricted to distributions on $\mathbb{R}_+$. Here, we only give a short proof of surjectivity as $\Phi_{f_1}(\mathcal{M}) \subset U$ follows from Lemma 4.1 already. Let g be the right-continuous version of a non-increasing Lévy density on $(0,\infty)$, i.e. $\int_0^\infty (1 \wedge u)\, g(u)\, \mathrm{d}u < \infty$ and thus $\lim_{u\to\infty} g(u) = 0$. Therefore, one can define a measure on $(0,\infty)$ by $\tilde{\nu}((u,\infty)) := g(u)$. Setting $\nu(\mathrm{d}x) := x\,\tilde{\nu}(\mathrm{d}x)$ defines the pre-image of the desired distribution as*

$$g(u) = \int_{(y,\infty)} \tilde{\nu}(\mathrm{d}x) = \int_{(y,\infty)} \frac{\nu(\mathrm{d}x)}{x},$$

and as $\nu \in \mathcal{M}$, which follows from

$$\begin{aligned}
\infty > \int_0^\infty (1 \wedge u)\, g(u)\, \mathrm{d}u &= \int_0^\infty (1 \wedge u) \int_{(y,\infty)} \frac{\nu(\mathrm{d}x)}{x}\, \mathrm{d}u \\
&= \int_0^\infty \frac{1}{x} \int_{(0,x)} (1 \wedge u)\, \mathrm{d}u\, \nu(\mathrm{d}x) = \int_0^1 \frac{x}{2}\, \nu(\mathrm{d}x) + \int_1^\infty \frac{x - 0.5}{x}\, \nu(\mathrm{d}x) \\
&\geq \frac{1}{2} \int_0^\infty (1 \wedge x)\, \nu(\mathrm{d}x).
\end{aligned}$$

This shows that $\Phi_{f_1}(\mathcal{M}) = U$.

(b) *Using Lemma 4.1(i), it is easy to see that starting from $\nu(\mathrm{d}x) = -x\, g'(x)\mathrm{d}x$, for the resulting density g_H it holds*

$$g_H(y) = \int_y^\infty \frac{\nu(\mathrm{d}x)}{x} = \int_y^\infty \frac{-x\, g'(x)\, \mathrm{d}x}{x} = g(y).$$

(c) *Each Thorin Lévy density is also c.m., as it can be written as the product of $1/x$ and a c.m. function (and the product of c.m. functions being c.m. again). Consequently, from (b) follows that every $\nu \in \Phi_{f_1}^{-1}(T)$ can be written as $\nu(\mathrm{d}x) = f(x)\mathrm{d}x$ with $f(x) = -x\, g'(x)$ and g a Thorin Lévy density. Since $-x\, g'(x) = g(x) + (-(x\, g(x))')$ is c.m. as the sum of two c.m. functions, we know that $\nu \in BO$ and thus $\Phi_{f_1}^{-1}(T) \subset BO$. Moreover, we use results in (Schilling et al., 2010, Remark 8.3), which state* *that for every Thorin density g there exists a function $w \in W$ with*

$$g(x) = \int_{(0,\infty)} e^{-x\,s} w(s)\mathrm{d}s, \qquad x\,g(x) = \int_{(0,\infty)} e^{-x\,s}\mathrm{d}w(s).$$

Thus,

$$\begin{aligned} f(x) = g(x) - (x\,g(x))' &= \int_{(0,\infty)} e^{-x\,s} w(s)\mathrm{d}s + \int_{(0,\infty)} e^{-x\,s} s\,\mathrm{d}w(s) \\ &= \int_{(0,\infty)} e^{-x\,s} \left(w(s)\mathrm{d}s + s\,\mathrm{d}w(s)\right) = \int_{(0,\infty)} e^{-x\,s} \mathrm{d}\left(s\,w(s)\right). \end{aligned}$$

(d) *For $\nu \in BO$ with c.m. Lévy density g it follows that $\Phi_{f_1}(\nu)$ has a decreasing Lévy density g_H with*

$$g_H(y) = \int_y^\infty \frac{\nu(\mathrm{d}x)}{x} = \int_y^\infty \frac{g(x)\mathrm{d}x}{x}.$$

Thus $g_H'(y) = -g(x)/x$, where $g(x)/x$ is c.m., which proves that g_H is c.m. and $\Phi_{f_1}(\nu) \in BO$.
Again, from (b) follows that every $\nu \in \Phi_{f_1}^{-1}(BO)$ can be written as $\nu(\mathrm{d}x) = f(x)\mathrm{d}x$ with $f(x) = -x\,g'(x)$ and g a c.m. Lévy density. One can deduce that $x\,\exp(-x) \in \Phi_{f_1}^{-1}(BO)$, which is not c.m., proving $\Phi_{f_1}^{-1}(BO) \not\subset BO$. The representation of $\Phi_{f_1}^{-1}(BO)$ follows analogously to the previous steps, as every c.m. Lévy density g can be written as the Laplace transform of a measure $\eta \in V$, see (Schilling et al., 2010, proof of Theorem 6.2).

(e) *Every Lévy density g corresponding to a distribution in L has a representation as $g(x) = f(x)/x$ with $f \in F$. Furthermore, using integration by parts in the third equality below,*

$$\begin{aligned} \int_x^\infty \frac{1}{y}\left(\frac{f(y)}{y}\,\mathrm{d}y + \mathrm{d}(-f(y))\right) &= -\left(\int_x^\infty \frac{1}{y}\,\mathrm{d}f(y) + \int_x^\infty \frac{-1}{y^2}\,f(y)\,\mathrm{d}y\right) \\ &= -\left(\int_x^\infty \frac{1}{y}\,\mathrm{d}f(y) + \int_x^\infty f(y)\,\mathrm{d}(1/y)\right) \\ &= -\left(\lim_{z\to\infty} \frac{f(z)}{z} - \frac{f(x)}{x}\right) = \frac{f(x)}{x} = g(x). \qquad \square \end{aligned}$$

Remark 4.3 (Remark on representation of $\Phi_{f_1}^{-1}(L)$)
The right-continuous modifications of functions f in our representation of $\Phi_{f_1}^{-1}(L)$ have

a meaning in the stochastic representation proven in (Jurek, 1985, Theorem 4.5): Every distribution in $\Phi_{f_1}^{-1}(L)$ can be written as the convolution of an ID measure μ and $\mathcal{L}\big(\int_0^\infty \exp(-s)\,\mathrm{d}\Lambda_s\big)$, where μ has a Lévy measure in $\mathcal{M}_{\log}$ corresponding to $\mathrm{d}(-f)$, $f \in F$, and Λ is a Lévy subordinator with $\mathcal{L}(\Lambda_1) = \mu$.
More precisely, it holds that $\nu = \mathrm{d}(-f) \in \mathcal{M}_{\log} \Leftrightarrow f \in F$ and for a subordinator Λ with Lévy measure ν, the expression $\int_0^\infty \exp(-s)\,\mathrm{d}\Lambda_s$ can be defined if and only if $\nu \in \mathcal{M}_{\log}$, see, e.g., (Jurek and Vervaat, 1983, Theorem 2.3) (they consider the equivalent condition $\int_1^\infty \log(x+1)\,\nu(\mathrm{d}x) < \infty$). Furthermore, the Lévy measure corresponding to the defined integral has Lévy density $f(x)/x$, which follows analogously to the steps in Lemma 4.1

The previous lemma does not only characterize the pre-images of important classes of distributions, it also allows for their explicit computation. Using the previous results, we can further conclude that for every Lévy subordinator with marginal distributions in the Jurek class, we can find another non-decreasing process with the same marginal distributions and a.e. differentiable paths, a.s..

4.2.2 The corresponding MSMVE family

It is known that for all $d \geq 2$, $(X_1, \dots, X_d)^\intercal$ constructed as in Equation (2.26) exhibits an MSMVE distribution which corresponds to a stable tail dependence function ℓ. In the given construction, ℓ can be computed explicitly. In particular, it is a function of the Bernstein function Ψ_H. This constitutes a very flexible class of stable tail dependence functions, since one can plug in any desired Bernstein function of the Jurek class.

Theorem 4.4 (Constructing parametric MSMVE families - I)
For every Bernstein function Ψ_H with drift a_H, Lévy measure $\nu_H \in U$, and $\Psi_H(1) = 1$, the function

$$\begin{aligned}\ell(x_1, \dots, x_d) = &\frac{d}{\sum_{j=1}^d 1/x_{(j)}} \Psi_H(d) \\ &- \sum_{i=1}^{d-1} \left(\frac{d-i+1}{\sum_{j=i}^d 1/x_{(j)}} - \frac{d-i}{\sum_{j=i+1}^d 1/x_{(j)}} \right) \Psi_H \left(d - i - \sum_{j=i+1}^{d} x_{(i)}/x_{(j)} \right)\end{aligned}$$

is a stable tail dependence function for every $d \geq 2$. A random vector $(X_1, \dots, X_d)^\intercal$ with *the respective MSMVE distribution can be constructed via*

$$X_k := \inf\left\{t > 0 : E_k < \int_0^t \left(1 - \frac{s}{t}\right) \mathrm{d}\Lambda_s\right\}, \quad k = 1, \ldots, d,$$

with $\Lambda = \{\Lambda_t\}_{t\geq 0}$ *a Lévy subordinator with drift* $a_\Lambda = 2\, a_H$ *and Lévy measure* $\nu_\Lambda = \Phi_{f_1}^{-1}(\nu_H)$, *and an iid sequence* $\{E_k\}_{k\in\mathbb{N}}$ *of unit exponential random variables independent of* Λ.

Remark 4.5 (Defining ℓ for $x_i = 0$)
Actually, the expression for ℓ *in Theorem 4.4 is only defined for values* $x_1, \ldots, x_d > 0$. *However, since the construction yields* $\ell(x_1, \ldots, x_d) = -\log\big(\mathbb{E}\big[\exp(-H_{x_1} - \ldots - H_{x_d})\big]\big)$ *and* $H_0 = 0$, *it is obvious that the case* $x_i = 0$ *for at least one* $i \in \{1, \ldots, d\}$ *has a simple solution: for* $I := \{i \in \{1, \ldots, d\} : x_i = 0\}$ *with* $k := |I|$, *one has* $\ell(x_1, \ldots, x_d) = \ell(x_{(1)}, \ldots, x_{(d)}) = \ell(0, \ldots, 0, x_{(k+1)}, \ldots, x_{(d)}) = \ell(x_{(k+1)}, \ldots, x_{(d)})$ *and* $\ell(0, \ldots, 0) = 0$. *The same observation holds true for Theorem 4.12 below.*

Proof (of Theorem 4.4)
It follows from Lemma 4.2 that there exists a Lévy subordinator Λ *with drift* $a_\Lambda = 2\, a_H$, $\nu_\Lambda = \Phi_{f_1}^{-1}(\nu_H)$, *such that the marginal distribution of the IDT subordinator* $H_t = \int_0^t (1 - s/t)\, \mathrm{d}\Lambda_s$ *corresponds to the Bernstein function* Ψ_H. *Using this process* H *in the IDT-frailty construction of Theorem 2.46, we observe*

$$\begin{aligned}
\mathbb{P}(X_1 > x_1, \ldots, X_d > x_d) &= \exp\big(-\ell(x_1, \ldots, x_d)\big) = \mathbb{E}\big[\exp(-H_{x_1} - \ldots - H_{x_d})\big] \\
&= \mathbb{E}\left[\exp\left(-\sum_{i=1}^d \int_0^\infty \left(1 - \frac{s}{x_i}\right)_+ \mathrm{d}\Lambda_s\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\int_0^\infty \sum_{i=1}^d \left(1 - \frac{s}{x_i}\right)_+ \mathrm{d}\Lambda_s\right)\right] \\
&= \exp\left(-\int_0^\infty \Psi_\Lambda\left(\sum_{i=1}^d \left(1 - \frac{s}{x_i}\right)_+\right) \mathrm{d}s\right),
\end{aligned}$$

where in the last step, we use Theorem 2.26(ii) again. Following Remark 4.5, we considered $x_1, \ldots, x_d > 0$. *Consequently, we find*

$$\ell(x_1, \ldots, x_d) = a_H \sum_{j=1}^d x_j + \int_0^{x_{(d)}} \int_0^\infty 1 - \exp\left(-u \sum_{i=1}^d \left(1 - \frac{s}{x_i}\right)_+\right) \nu_\Lambda(\mathrm{d}u)\, \mathrm{d}s. \quad (4.4)$$

We proceed with three helpful equalities:

$$-x_{(d)} = -\frac{1}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} - \sum_{i=1}^{d-1} \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right), \tag{4.5}$$

which follows from a telescope argument applied to the right hand side. Furthermore, one has

$$x_{(d)} = \frac{d}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} - \sum_{i=1}^{d-1} \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right), \tag{4.6}$$

as

$$\left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right) = \frac{d-i+1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{d-i}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}},$$

so again, we can use a telescope argument. Finally,

$$\sum_{j=1}^{d} x_j = \frac{d^2}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} - \sum_{i=1}^{d-1} \left(\frac{d-i+1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{d-i}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} \right) \left(d-i - \sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right), \tag{4.7}$$

which can be shown by rearranging

$$\left(\frac{d-i+1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{d-i}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} \right) \left(d-i - \sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right) = \frac{(d-i+1)^2}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{(d-i)^2}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - x_{(i)},$$

so again, we can use another telescope argument.

For the second term in Equation (4.4)*, we can compute, defining* $x_{(0)} := 0$*,*

$$\begin{aligned}
&\int_0^{x_{(d)}} \int_0^{\infty} 1 - \exp\left(-u \sum_{i=1}^{d} \left(1 - \frac{s}{x_i} \right)_+ \right) \nu_\Lambda(\mathrm{d}u)\, \mathrm{d}s \\
&= \int_0^{\infty} x_{(d)} - \sum_{i=1}^{d} \int_{x_{(i-1)}}^{x_{(i)}} \exp\left(-u \sum_{j=i}^{d} \left(1 - \frac{s}{x_{(j)}} \right) \right) \mathrm{d}s\, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^{\infty} x_{(d)} - \sum_{i=1}^{d} \frac{e^{-u\,(d-i+1)}}{u \sum_{j=i}^{d} \frac{1}{x_{(j)}}} \left(e^{u\, x_{(i)} \sum_{j=i}^{d} \frac{1}{x_{(j)}}} - e^{u\, x_{(i-1)} \sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \nu_\Lambda(\mathrm{d}u)
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{u} \Bigg(u\, x_{(d)} - x_{(d)} + \frac{e^{-u\, d}}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} \\
&\qquad + \sum_{i=1}^{d-1} e^{u \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right)} \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \Bigg) \nu_\Lambda(\mathrm{d}u)
\end{aligned}
$$

where now the aim is to bring this in a specific form such that Lemma 4.1(ii) is applicable,

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{u} \Bigg(u\, x_{(d)} + \frac{e^{-u\, d} - 1}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} \\
&\qquad + \sum_{i=1}^{d-1} \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \left(e^{u \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right)} - 1 \right) \Bigg) \nu_\Lambda(\mathrm{d}u)
\end{aligned}
$$

having replaced $-x_{(d)}$ *using Equation* (4.5)*, which yields exactly the missing constants to rewrite the expression as an integral*

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{u} \Bigg(\int_0^u x_{(d)} - \frac{d\, e^{-s\, d}}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} + \sum_{i=1}^{d-1} e^{s \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right)} \\
&\qquad \times \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right) \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \mathrm{d}s \Bigg) \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} - \frac{d\, e^{-u\, d}}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} + \sum_{i=1}^{d-1} e^{u \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right)} \\
&\qquad \times \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right) \left(\frac{1}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} - \frac{1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} \right) \nu_H(\mathrm{d}u)
\end{aligned}
$$

where in the last step, we used Lemma 4.1(ii),

$$
\begin{aligned}
&\overset{(4.6)}{=} \int_0^\infty \frac{d}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} \left(1 - e^{-d\, u} \right) - \sum_{i=1}^{d-1} \left(\frac{d-i+1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{d-i}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} \right) \\
&\qquad \times \left(1 - e^{u \left(\sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} - (d-i) \right)} \right) \nu_H(\mathrm{d}u).
\end{aligned}
$$

From Equation (4.4) *follows that one has to add the term* $a_H \sum_{j=1}^{d} x_j$ *to the above expression in order to compute* ℓ*. According to Equation* (4.7)*,* $a_H \sum_{j=1}^{d} x_j$ *can be* *rewritten as*

$$a_H \sum_{j=1}^{d} x_j = \frac{d}{\sum_{j=1}^{d} \frac{1}{x_{(j)}}} a_H\, d - \sum_{i=1}^{d-1} \left(\frac{d-i+1}{\sum_{j=i}^{d} \frac{1}{x_{(j)}}} - \frac{d-i}{\sum_{j=i+1}^{d} \frac{1}{x_{(j)}}} \right) a_H \left(d - i - \sum_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right),$$

and the claim follows. $\square$

Remark 4.6
By choosing the drift $a_H \in [0,1]$ one can interpolate between a deterministic process ($a_H = 1$ implying $H_t = t$) and a "completely random" process ($a_H = 0$), which for the corresponding multivariate distribution of $(X_1, \dots, X_d)^\intercal$ means interpolating between independence and the maximal dependence attainable by the given random structure. In the following examples, we will always set the drift to zero.

Having a closer look at the resulting dependence structure of $(X_1, \dots, X_d)^\intercal$, it might be interesting to consider the Pickands dependence function $A : (0,1) \to [1/2, 1]$ of the corresponding bivariate distribution. As the distribution is exchangeable, A has to be symmetric and it is sufficient to consider and visualize it for $0 < t \leq 0.5$:

$$A(t) = \ell(t, 1-t) = 2\,t(1-t)\,\Psi_H(2) + (1-t)(1-2\,t)\,\Psi_H\left(1 - \frac{t}{1-t}\right). \tag{4.8}$$

It is obvious that $A \in C^\infty$ on $(0, 1/2)$ (and thus on $(1/2, 1)$), as Ψ_H is. Using the symmetry around $t = 1/2$, A is (continuously) differentiable on $(0,1)$ if the left derivative at $t = 1/2$ is zero, which is shown in Lemma 4.7. Consequently, the bivariate distribution is absolutely continuous.[2] This fact is intuitively clear from the stochastic model as the process H does not exhibit jumps, and thus it can not jump across several trigger variables E_k at the same time.

Lemma 4.7 (Derivative at $t = 1/2$)
For A as defined in Equation (4.8)*, it holds that*

$$\lim_{t \nearrow 1/2} A'(t) = 0.$$

[2] This can be easily seen considering the a.e. existing density $\partial_{x_1}\partial_{x_2} F(x_1, x_2)$, which denotes the second order partial derivative.

Proof
It is easy to see that for $t < 1/2$

$$A'(t) = (2-4\,t)\,\Psi_H(2) + (4\,t-3)\,\Psi_H\left(1-\frac{t}{1-t}\right) - \frac{1-2\,t}{1-t}\,\Psi_H'\left(1-\frac{t}{1-t}\right).$$

Consequently,

$$\lim_{t\nearrow 1/2} A'(t) = -2\lim_{t\nearrow 1/2}(1-2\,t)\,\Psi_H'\left(\frac{1-2\,t}{1-t}\right).$$

The drift term of the Bernstein function can be ignored, yielding

$$\begin{aligned}\lim_{t\nearrow 1/2} A'(t) &= -2\lim_{t\nearrow 1/2}(1-2\,t)\int_0^\infty u\,e^{-u\left(\frac{1-2\,t}{1-t}\right)}\nu_H(\mathrm{d}u)\\ &= -2\lim_{t\nearrow 1/2}\int_0^\infty (1-2\,t)\,u\,e^{-u\left(\frac{1-2\,t}{1-t}\right)}\nu_H(\mathrm{d}u)\\ &= -2\int_0^\infty \lim_{t\nearrow 1/2}(1-2\,t)\,u\,e^{-u\left(\frac{1-2\,t}{1-t}\right)}\nu_H(\mathrm{d}u) = 0,\end{aligned}$$

where we have used dominated convergence, as it holds that

$$(1-2\,t)\,u\,e^{-u\left(\frac{1-2\,t}{1-t}\right)} \le \min(u,1).$$

□

Using the well-known formulas for bivariate upper and lower tail dependence coefficients λ_U and λ_L stated in Section 2.4.2, one can compute

$$\lambda_L = 2\big(1-A(1/2)\big) = 2-\Psi_H(2), \qquad \lambda_U = \mathbb{1}_{\{\Psi_H(2)=1\}} = 0.$$

In our construction, λ_U is always zero as perfect dependence can only be constructed using a process that jumps from 0 directly to ∞. There is a nice relation between the lower tail dependence coefficient and H. As $\lambda_L = 2+\log(\mathbb{E}[\exp(-H_1)^2])$, it is increasing in the variance of $\exp(-H_1)$. This fits with the intuitive interpretation that the higher the variability of H (near zero), the higher the probability of joint early hitting of the exponential triggers, the higher the lower tail dependence.

We present two examples of possible parametrizations. An example with a very simple form is based on the positive α-stable case.

Example 4.8
$\Psi_H(x) = x^\alpha, \alpha \in (0,1)$, *is attainable (respectively part of the Jurek class* U*) as* $\nu_H(\mathrm{d}x) =$

$g(x)\,\mathrm{d}x$, with

$$g(x) = \frac{\alpha}{\Gamma(1-\alpha)}\, x^{-1-\alpha}, \quad x > 0,$$

see Section 2.2.2, which is a decreasing density. It actually is a c.m. function and consequently, the distribution is part of the Bondesson class, i.e. $\nu_H \in BO \subset U$. Using Lemma 4.2, we can compute the density f of $\nu_\Lambda = \Phi_{f_1}^{-1}(\nu_H)$ as $f(x) = -x\,g'(x) = (1+\alpha)\,g(x)$. Consequently, the corresponding Λ is an α-stable subordinator. The resulting bivariate Pickands dependence function is

$$A(t) = t(1-t)\,2^{\alpha+1} + (1-t)^{1-\alpha}\,(1-2\,t)^{1+\alpha}, \quad 0 < t \leq 0.5.$$

For $\alpha \in (0,1)$, this interpolates between complete dependence and independence as can be seen in Figure 4.1. The lower tail dependence coefficient is given by $\lambda_L = 2 - 2^\alpha$. Though A appears to exhibit a kink at $t = 1/2$, we know from previous computations that it is indeed differentiable. In Figure 4.2, the dependence function for the case $d = 3$ and $\alpha = 0.5$ can be found.

Another simple class of Bernstein functions is based on the compound Poisson distribution. We present one specific instance.

Example 4.9
$\Psi_H = (1+a)\,x/(x+a)$, $a > 0$, is attainable as $\nu_H(\mathrm{d}x) = g(x)\,\mathrm{d}x$, with

$$g(x) = (1+a)\,a\,e^{-a\,x}, \quad x > 0,$$

which corresponds to a compound Poisson process with intensity $(1+a)$ and $\mathrm{Exp}(a)$-distributed jumps. The related Lévy process is a compound Poisson process with intensity $(1+a)$ and $\Gamma(2,a)$-distributed jumps, as can be seen from its Lévy measure $\nu_\Lambda(\mathrm{d}x) = \Phi_{f_1}^{-1}(\nu_H(\mathrm{d}x)) = (1+a)\,a^2\,x\,\exp(-a\,x)\,\mathrm{d}x$ (see Lemma 4.2). The resulting bivariate Pickands dependence function is

$$A(t) = (1+a)\,(1-t)\left(\frac{4\,t}{2+a} + \frac{(1-2\,t)^2}{1-2\,t+a\,(1-t)}\right), \quad 0 < t \leq 0.5,$$

and the lower tail dependence coefficient is given by $\lambda_L = 2/(2+a)$, i.e. every value in $(0,1]$ is attainable.

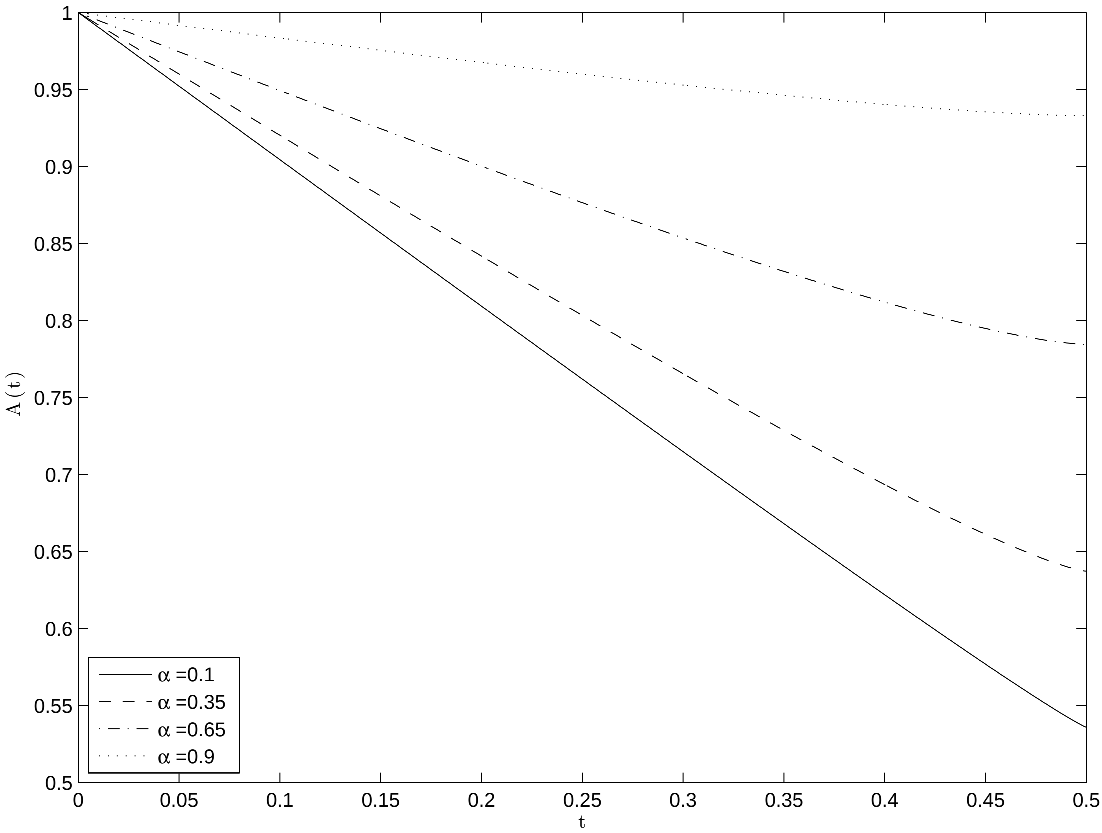


Figure 4.1 Bivariate Pickands dependence functions for Example 4.8 ($\Psi_H(x) = x^\alpha$) and different values of α.

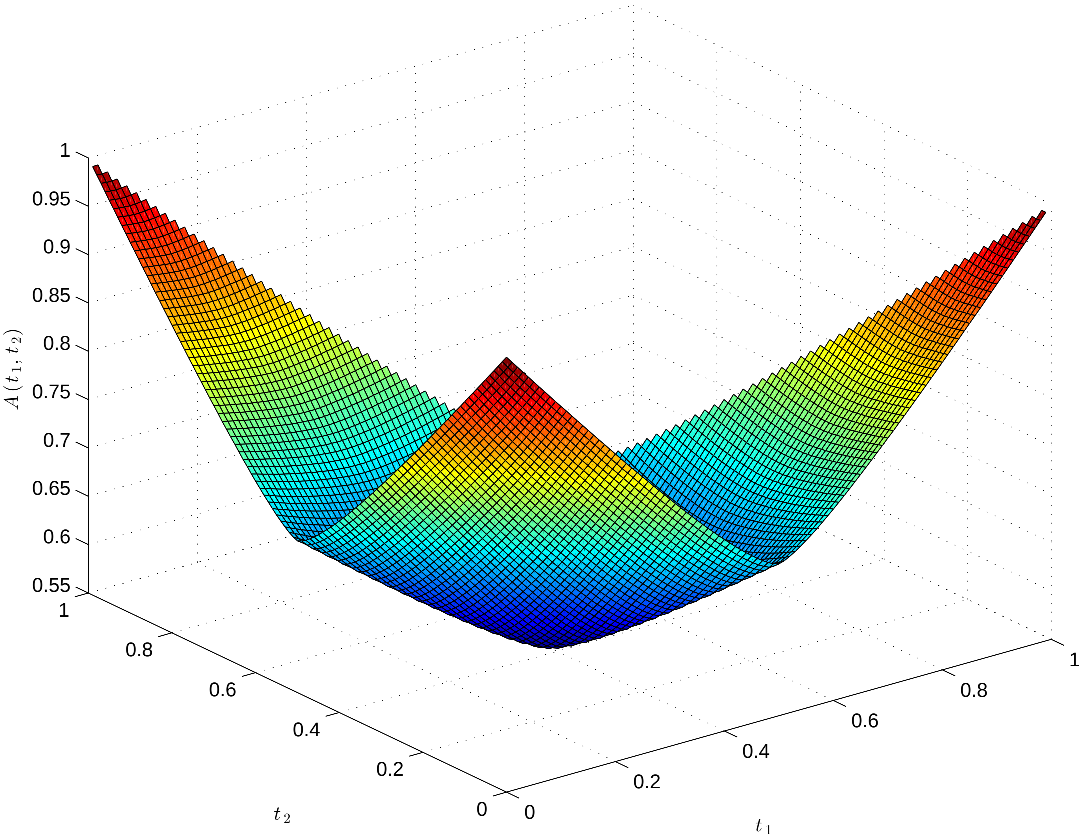


Figure 4.2 Trivariate Pickands dependence function $A(t_1, t_2) := \ell\big(t_1, t_2, 1 - (t_1 + t_2)\big)$, $t_1, t_2 \geq 0$ with $0 \leq t_1 + t_2 \leq 1$, for Example 4.8 ($\Psi_H(x) = x^\alpha$) with $\alpha = 0.5$.

The number of parametric families of attainable Bernstein functions is huge. As mentioned earlier, (Schilling et al., 2010, pp. 218–277) list more than one hundred complete Bernstein functions, which are a proper subclass of the attainable Bernstein functions, see Section 2.2.2. A small selection of interesting examples can be found in Table 4.1. In many cases, one can also compute the corresponding Lévy process. If, e.g., H is distributed according to a compound Poisson distribution, the corresponding subordinator is a compound Poisson process, as, using Lemma 4.1, it follows that

$$\nu_\Lambda\big((0,\infty)\big) = \int_0^\infty \nu_\Lambda(\mathrm{d}x) = \int_0^\infty \frac{1}{x}\left(\int_0^x \mathrm{d}y\right)\nu_\Lambda(\mathrm{d}x) = \int_0^\infty h(x)\,\nu_H(\mathrm{d}x) = \nu_H\big((0,\infty)\big),$$

and a distribution is compound Poisson if and only if its Lévy measure is finite. Figure 4.3 and Figure 4.4 indicate how different attainable shapes look like.

4.3 A second construction based on $f_2(s) = \log_+(1/s)$

As a second family, we examine the construction of Lemma 2.31 using the function $f_2(s) = \log_+(1/s) := \max\{\log(1/s), 0\}$, which results in

$$H_t := \int_0^t \log\left(\frac{t}{s}\right)\mathrm{d}\Lambda_s, \quad t > 0,$$

with $H_0 := 0$. This approach does not yield closed form solutions for arbitrary subordinators, but it allows finding a convenient expression for the corresponding stable tail dependence function such that tractable instances can be constructed easily. From (Barndorff-Nielsen et al., 2006b, Proposition 2.3) it follows that f_2 fulfills the integrability conditions in Lemma 2.31 for every Lévy subordinator. Furthermore, the corresponding integral transform mapping $\mathcal{L}(\Lambda_1)$ to $\mathcal{L}(\int_0^1 \log(1/s)\,\mathrm{d}\Lambda_s)$ is well known and thoroughly investigated in arbitrary dimensions, see Barndorff-Nielsen et al. (2006b) and also Remark 2.29. Denoting the transform restricted to distributions on $\mathbb{R}_+$ by Φ_{f_2}, they show that $\Phi_{f_2}(\mathcal{M}) = BO$, $\Phi_{f_2}(L) = T$, and that Φ_{f_2} is one-to-one. We provide a short proof of the first result as it is helpful for understanding the transform itself. Ψ_H, Ψ_Λ, and the related expressions are defined analogously to Section 4.2.

Table 4.1 Possible choices for Ψ_H together with their Lévy densities g_H, $\nu_\Lambda(\mathrm{d}x) = \Phi_{f_1}^{-1}(g_H \mathrm{d}x)$, and the corresponding type of Λ (where CP denotes a compound Poisson process). The restrictions $\alpha \in (0,1), a > 0, \beta > 0, \eta > 0$ have to hold. All are also complete Bernstein functions. Furthermore, we define $c := \log(1 + 1/\beta)$ and $d := \sqrt{2 + \eta^2} - \eta$ which we need to ensure $\Psi_H(1) = 1$.

Name	Ψ_H	g_H	$\nu_\Lambda(\mathrm{d}x)/\mathrm{d}x$	type of Λ
Stable	x^α	$\alpha/\Gamma(1-\alpha)\, x^{-1-\alpha}$	$(1+\alpha)\, g_H(x)$	Stable
CP1	$(1+a)\, x/(x+a)$	$a \exp(-a\, x)$	$(1+a)\, f_{\Gamma(a,2)}(x)$	CP with $\Gamma(2,a)$-jumps
Gamma	$1/c \log(1 + x/\beta)$	$\exp(-\beta\, x)/(c\, x)$	$\beta/c \exp(-\beta\, x) + g_H(x)$	Sum of ind. CP1 and Gamma
IG	$(\sqrt{2\, x + \eta^2} - \eta)/d$	$\sqrt{0.5}/(d\, \Gamma(0.5))\, x^{-3/2} \exp(-\eta^2/2)$	$1.5\, g_H(x) + \eta/(2\, d)\, f_{\Gamma(0.5,\eta^2/2)}(x)$	Sum of indep. IG and CP

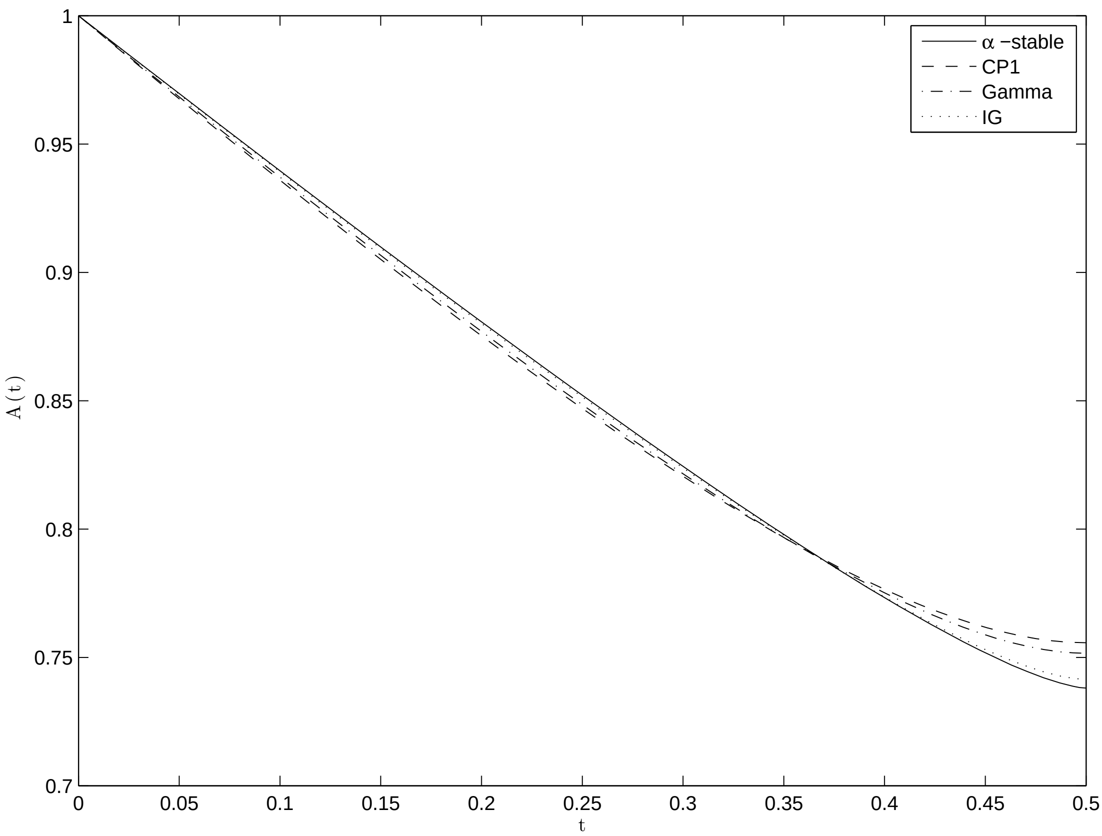


Figure 4.3 The bivariate Pickands dependence functions resulting from Theorem 4.4 for different Ψ_H corresponding to the α-stable case, CP1, the Gamma case, and the IG case as defined in Table 4.1. The parameters are chosen such that all models exhibit a Spearman's ρ of 0.5.

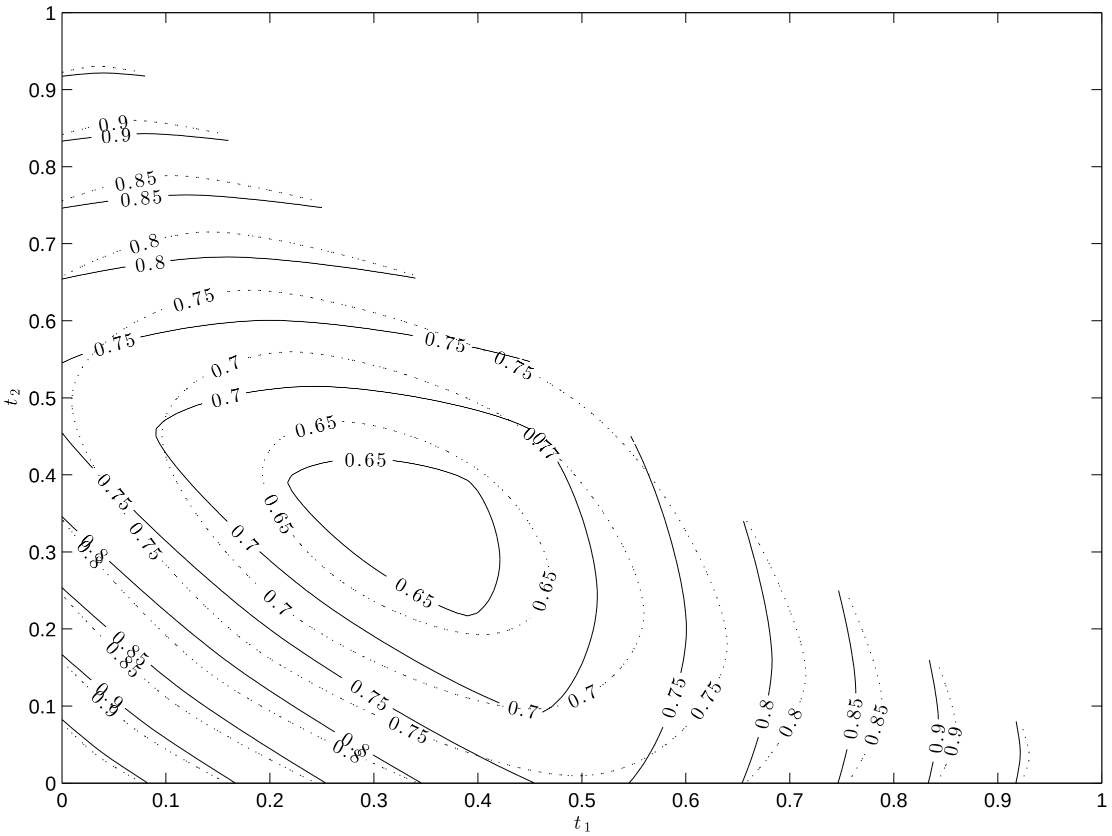


Figure 4.4 A contour plot for the trivariate Pickands dependence function resulting from Theorem 4.4 for different choices of Ψ_H, comparing the α-stable case and CP1. The dotted lines correspond to CP1. Parameters are chosen as in Figure 4.3. Though the two-dimensional models seem to be similar, the three-dimensional dependence structures differ considerably.

Lemma 4.10 (Structure of Ψ_H)
Using f_2, Ψ_H has a representation

$$\Psi_H(x) = a_H\, x + \int_0^\infty \frac{x}{x+t}\,\sigma_H(\mathrm{d}t), \quad x \geq 0,$$

with $a_H = a_\Lambda$ and $\sigma_H(B) := \int_0^\infty \mathbb{1}_B(1/u)\nu_\Lambda(\mathrm{d}u)$, $B \in \mathcal{B}(\mathbb{R})$.
Proof
Similar as in Lemma 4.1, for $x \geq 0$ we see

$$\begin{aligned}
\Psi_H(x) &= \int_0^\infty \Psi_\Lambda\left(x\, f_2(s)\right)\,\mathrm{d}s = \int_0^1 \Psi_\Lambda\left(x\,\log_+(1/s)\right)\,\mathrm{d}s \\
&= a_\Lambda\, x \int_0^1 \log_+(1/s)\,\mathrm{d}s + \int_0^1 \int_0^\infty 1 - \exp\left(-u\, x\, \log_+(1/s)\right)\,\nu_\Lambda(\mathrm{d}u)\,\mathrm{d}s \\
&= a_\Lambda\, x + \int_0^\infty \int_0^1 1 - s^{u\,x}\,\mathrm{d}s\,\nu_\Lambda(\mathrm{d}u) \\
&= a_\Lambda\, x + \int_0^\infty \frac{x}{x+1/u}\,\nu_\Lambda(\mathrm{d}u) = a_\Lambda\, x + \int_0^\infty \frac{x}{x+t}\,\sigma_H(\mathrm{d}t).
\end{aligned}$$

□

Remark 4.11
Using Lemma 4.10, $\Phi_{f_2}(\mathcal{M}) = BO$ follows from the observation that σ_H as defined above is a Stieltjes measure if and only if ν_Λ is a Lévy measure, as

$$\int_0^\infty (1 \wedge u)\,\nu_\Lambda(\mathrm{d}u) = \int_0^\infty \left(1 \wedge \frac{1}{s}\right)\sigma_H(\mathrm{d}s),$$

and

$$\frac{1}{1+s} \leq \left(1 \wedge \frac{1}{s}\right) \leq \frac{2}{1+s}.$$

Consequently, Lemma 4.10 defines a direct connection between the characteristics of Ψ_H and Ψ_Λ as we can write the Stieltjes measure of H in terms of the Lévy measure of Λ. We will make use of this fact below.

What is interesting in this context is that the process H itself has an alternative representation, using integration by parts, as shown in (Barndorff-Nielsen et al., 2006b, Proposition 2.4), via

$$H_t = \int_0^t \frac{\Lambda_s}{s}\mathrm{d}s = \lim_{u \searrow 0} \int_u^t \frac{\Lambda_s}{s}\mathrm{d}s, \tag{4.9}$$

where the limit is a.s.. In the context of the construction in Equation (2.26), this can be interpreted as an intensity model with intensity $\lambda_s := \Lambda_s/s, s > 0$, $\lambda_0 := a_\Lambda$, in the spirit of, e.g., Duffie and Gârleanu (2001). It is defined consistently, as $\lim_{s \searrow 0} \lambda_s = a_\Lambda$ a.s., see (Sato, 1999, p. 351). This is a peculiar construction, as it follows, assuming the first and second moment of Λ to exist, that $\mathbb{E}[\lambda_s] = a_\Lambda + \int_0^\infty x\, \nu_\Lambda(\mathrm{d}x)$ independent of s and $\mathrm{Var}[\lambda_s] = 1/s \left(\int_0^\infty x^2\, \nu_\Lambda(\mathrm{d}x) \right)$ for $s > 0$. Thus, the variance of the intensity is exploding close to 0 and vanishing for large s.

4.3.1 The corresponding MSMVE family

As we have seen in Lemma 4.10 and Remark 4.11, for an arbitrary Stieltjes measures σ_H one can find a corresponding Lévy measure. We will use this fact and state the dependence function of the resulting multivariate distribution in terms of the Stieltjes measures, such that arbitrary Stieltjes measures can be plugged in. Notice that Remark 4.5 also applies to Theorem 4.12.

Theorem 4.12 (Constructing parametric MSMVE families - II)
For every complete Bernstein function Ψ_H with Stieltjes measure σ_H and drift a_H such that $\Psi_H(1) = 1$, the function

$$\ell(x_1, \ldots, x_d) = \sum_{i=1}^{d} x_{(i)} \left(a_H + \int_0^\infty \left(\prod_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right)^{1/s} \frac{s}{(s+d-i+1)\,(s+d-i)} \sigma_H(\mathrm{d}s) \right) \tag{4.10}$$

denotes a stable tail dependence function for every $d \geq 2$. A stochastic representation of an MSMVE distribution $(X_1, \ldots, X_d)^\intercal$ with stable tail dependence function ℓ and unit exponential marginals is given by

$$X_k := \inf \left\{ t > 0 : E_k < \int_0^t \log\left(\frac{t}{s}\right) \mathrm{d}\Lambda_s \right\},$$

with Λ a Lévy subordinator with drift $a_\Lambda = a_H$ and Lévy measure ν_Λ given by $\nu_\Lambda(B) := \int_0^\infty \mathbb{1}_B(1/u)\sigma_H(\mathrm{d}u)$, $B \in \mathcal{B}(\mathbb{R})$, and an iid sequence $\{E_k\}_{k\in\mathbb{N}}$ of unit exponential random variables independent of Λ.

Proof
From Lemma 4.10 and Remark 4.11 we know that there exists a Lévy subordinator Λ with drift $a_\Lambda = a_H$ and Lévy measure ν_Λ, such that $H_t = \int_0^t \log(t/s)\, \mathrm{d}\Lambda_s$ has a marginal dis-

tribution corresponding to Ψ_H. *Using this* H *in the IDT-frailty construction in Theorem 2.46, we obtain*

$$\ell(x_1, \ldots, x_d) = a_H \sum_{j=1}^{d} x_j + \int_0^{x_{(d)}} \Psi_{\nu_\Lambda} \left(\sum_{i=1}^{d} \log_+ \left(\frac{x_i}{s} \right) \right) \mathrm{d}s,$$

where following Remark 4.5, we consider $x_1, \ldots, x_d > 0$. *For the second term, we compute*

$$\begin{aligned}
&\int_0^{x_{(d)}} \Psi_{\nu_\Lambda} \left(\sum_{i=1}^{d} \log_+ \left(\frac{x_i}{s} \right) \right) \mathrm{d}s \\
&= \int_0^\infty \int_0^{x_{(d)}} 1 - \exp \left(-u \left(\sum_{i=1}^{d} \log_+ \left(\frac{x_i}{s} \right) \right) \right) \mathrm{d}s \, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} - \sum_{i=1}^{d} \int_{x_{(i-1)}}^{x_{(i)}} \exp \left(-u \left(\sum_{j=i}^{d} \log \left(\frac{x_{(j)}}{s} \right) \right) \right) \mathrm{d}s \, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} - \sum_{i=1}^{d} \int_{x_{(i-1)}}^{x_{(i)}} \left(\frac{\prod_{j=i}^{d} x_{(j)}}{s^{d-i+1}} \right)^{-u} \mathrm{d}s \, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} - \sum_{i=1}^{d} \left(\prod_{j=i}^{d} x_{(j)} \right)^{-u} \int_{x_{(i-1)}}^{x_{(i)}} s^{u\,(d-i+1)} \mathrm{d}s \, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} - \sum_{i=1}^{d} \left(\prod_{j=i}^{d} x_{(j)} \right)^{-u} \frac{x_{(i)}^{u\,(d-i+1)+1} - x_{(i-1)}^{u\,(d-i+1)+1}}{u\,(d-i+1)+1} \, \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} \frac{u}{u+1} \\
&\quad - \sum_{i=1}^{d-1} \left[\left(\prod_{j=i}^{d} x_{(j)} \right)^{-u} \frac{x_{(i)}^{u\,(d-i+1)+1}}{u\,(d-i+1)+1} - \left(\prod_{j=i+1}^{d} x_{(j)} \right)^{-u} \frac{x_{(i)}^{u\,(d-(i+1)+1)+1}}{u\,(d-(i+1)+1)+1} \right] \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} \frac{u}{u+1} - \sum_{i=1}^{d-1} \left(\prod_{j=i+1}^{d} x_{(j)} \right)^{-u} x_{(i)}^{u\,(d-i)+1} \left[\frac{1}{u\,(d-i+1)+1} - \frac{1}{u\,(d-i)+1} \right] \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} \frac{u}{u+1} - \sum_{i=1}^{d-1} x_{(i)} \left(\prod_{j=i+1}^{d} \frac{x_{(j)}}{x_{(i)}} \right)^{-u} \left[\frac{1}{u\,(d-i+1)+1} - \frac{1}{u\,(d-i)+1} \right] \nu_\Lambda(\mathrm{d}u) \\
&= \int_0^\infty x_{(d)} \frac{1}{1+s} - \sum_{i=1}^{d-1} x_{(i)} \left(\prod_{j=i+1}^{d} \frac{x_{(j)}}{x_{(i)}} \right)^{-1/s} \left[\frac{s}{(d-i+1)+s} - \frac{s}{(d-i)+s} \right] \sigma_H(\mathrm{d}s)
\end{aligned}$$

$$= \sum_{i=1}^{d} x_{(i)} \int_0^\infty \left(\prod_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right)^{1/s} \frac{s}{(s+d-i+1)\,(s+d-i)}\, \sigma_H(\mathrm{d}s).$$

The claim follows. □

At least two approaches are possible when looking for tractable specifications. As there exists a direct link between σ_H and ν_Λ, one can start from both sides. It is, for example, possible to start from σ_H corresponding to a desired Ψ_H and try to compute the expression in Theorem 4.12. Schilling et al. (2010) lists the Stieltjes measures for many of the known complete Bernstein functions. One could also start from a ν_Λ such that the Laplace transform of the measure $\nu_\Lambda(\mathrm{d}u)/(n\,u+1)$ for $n \in \mathbb{N}$ is known in closed form. This can be seen from the third from last line of the computation in the proof of Theorem 4.12.

We present one example of a possible parametrization starting from the Laplace exponent Ψ_H.

Example 4.13

$\Psi_H(x) = (1+a)\,x/(x+a), a > 0$ is attainable and corresponds to a compound Poisson distribution with intensity $(1+a)$ and jump-size distribution $\mathrm{Exp}(a)$. This coincides with the Bernstein function in Example 4.9 in the previous section (also called CP1 in Table 4.1), i.e. it is possible to construct a process H using f_2 which has the same marginal distributions as the process constructed in Example 4.9 using f_1. Thus, the minima of subsets of the two different resulting MSMVE sequences have the same exponential distributions (in particular, the two sequences are marginally equivalent in minimums), though their multivariate distributions differ.

The corresponding Stieltjes measure is determined as $\sigma_H(\mathrm{d}s) = (1+a)\delta_a(s)$. It is easy to see that $\nu_\Lambda = \Phi_{f_2}^{-1}(\sigma_H)$ is given by $\nu_\Lambda(\mathrm{d}s) = (1+a)\delta_{1/a}(s)$, so Λ is a Poisson process with fixed jump-size $1/a$ and intensity $(1+a)$. A closed-form solution for ℓ defined in Equation (4.10) is given by

$$\ell(x_1,\ldots,x_d) = \sum_{i=1}^{d} x_{(i)} \left(\prod_{j=i+1}^{d} \frac{x_{(i)}}{x_{(j)}} \right)^{1/a} \frac{a\,(a+1)}{(a+d-i+1)\,(a+d-i)}.$$

The bivariate Pickands dependence function for $0 < t < 0.5$ can be stated as

$$A(t) = (1-t) + t \left(\frac{t}{1-t} \right)^{1/a} \frac{a}{a+2}.$$

The dependence functions of this model and the one of Example 4.9 are compared in Figure 4.5. It can be observed that both approaches yield considerably different dependence functions.

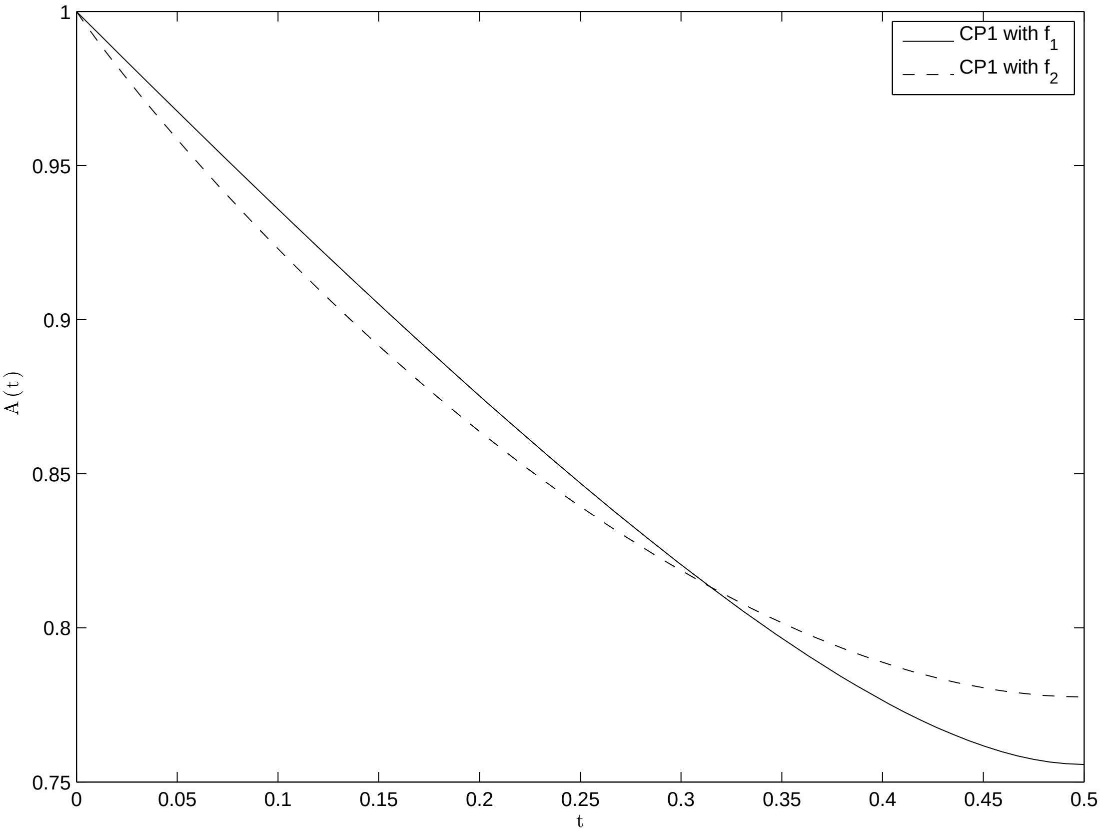


Figure 4.5 The bivariate Pickands dependence functions for Example 4.9 and Example 4.13, which share the same Bernstein function Ψ_H, but are based on different families of IDT subordinators. The parameters are chosen such that both models exhibit a Spearman's ρ of 0.5.

4.4 A note on simulation

As mentioned before, the stochastic representation of $(X_1, \ldots, X_d)^\intercal$ as an IDT-frailty model can be used to develop efficient simulation algorithms. When the involved Lévy subordinators are compound Poisson processes, simulating is straight-forward. Other

Lévy subordinators can be approximated by compound Poisson processes, see, e.g., Damien et al. (1995), or more involved schemes can be developed based on the given representation. We compare Example 4.9 with Example 4.13, which both yield CP1, i.e. $\Psi_H(x) = (1+a)\,x/(x+a), a > 0$, as the desired (complete) Bernstein function for H, i.e. the resulting models are marginally equivalent in minimums (if the same parameter a is chosen). In Example 4.9, this corresponds to $\Lambda^{(1)}$ being a compound Poisson process with intensity $(1+a)$ and $\Gamma(2,a)$-distributed jumps, i.e.

$$H_t^{(1)} = \int_0^t \left(1 - \frac{s}{t}\right) \mathrm{d}\Lambda_s^{(1)}, \quad t > 0,$$

has the desired Laplace exponent. For the family f_2, as described in Example 4.13, this corresponds to $\Lambda^{(2)}$ being a Poisson process with deterministic jump-size $1/a$ and intensity $(1+a)$, i.e.

$$H_t^{(2)} = \int_0^t \log\left(\frac{t}{s}\right) \mathrm{d}\Lambda_s^{(2)}, \quad t > 0,$$

yields a second construction with the desired marginal distribution. Denoting by $\tau_i, i \in \mathbb{N}$, the jump times of a Poisson process with intensity $(1+a)$, one can rewrite

$$H_t^{(1)} = \sum_{\tau_i \le t} G_i \left(1 - \frac{\tau_i}{t}\right), \quad t \ge 0,$$
$$H_t^{(2)} = \frac{1}{a} \sum_{\tau_i \le t} \log\left(\frac{t}{\tau_i}\right), \quad t \ge 0,$$

where $G_i, i \in \mathbb{N}$, are iid $\Gamma(2,a)$-distributed. To illustrate the construction, sample paths are shown in Figure 4.6, where the same jump times are used to emphasize the differences of the resulting paths.

Based on these representations, it is clear how to sample from the construction in Equation (2.26). Exemplary scatterplots can be found in Figure 4.7, where we transformed the marginals to uniform distributions on $[0,1]$ so that samples from the related extreme-value survival copulas are obtained for reasons of better comparability. Example 4.9 yields more samples close to the diagonal, which can be explained by the additional randomness introduced through the random variables G_i. High values of G_i correspond to a steep increase of $H^{(1)}$, which increases the probability of imminent triggering for both components within a short time period.

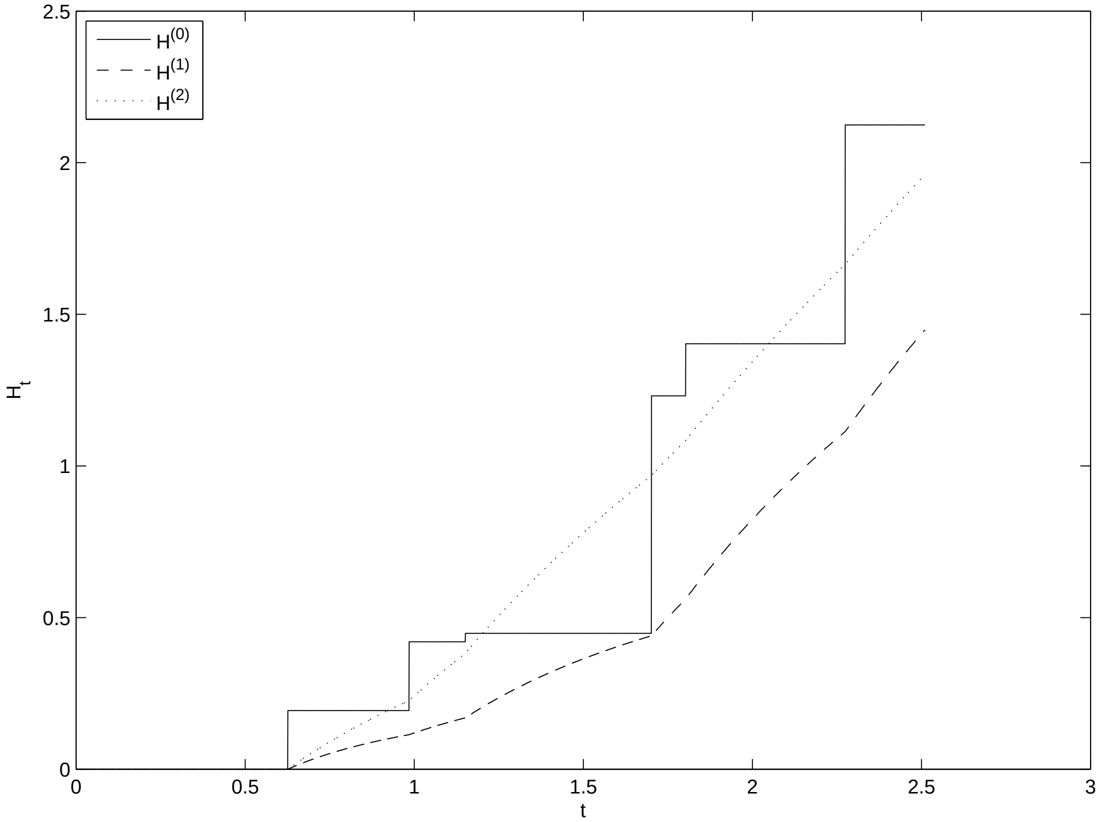


Figure 4.6 Simulated paths of the processes $H^{(1)}$ and $H^{(2)}$ where $a = 2$ is chosen. For comparison, a path of the simple compound Poisson process $H^{(0)}$ with Exp(a)-distributed jumps is added, which has the same marginal distribution. For all processes, the same jump times are used.

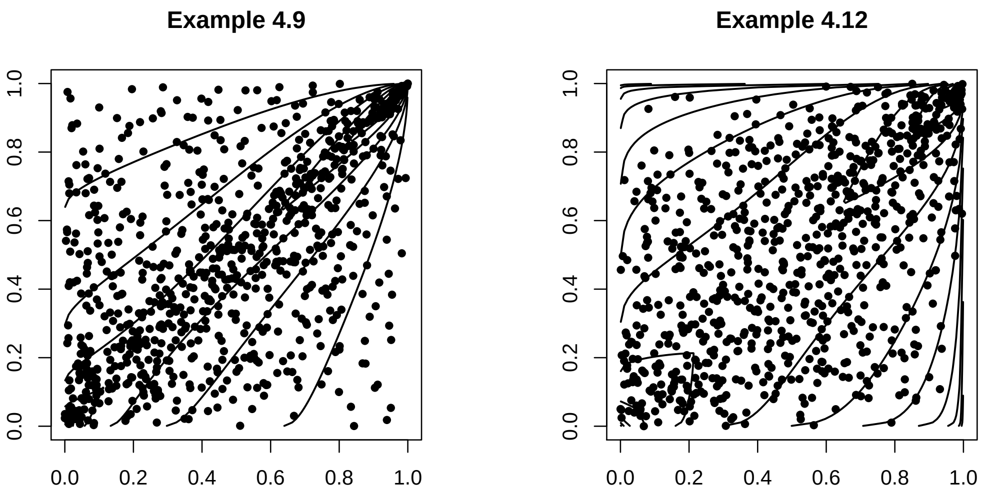


Figure 4.7 Scatterplots with 800 samples of the survival copulas of the MSMVE distribution generated by Example 4.9 and Example 4.13. The corresponding Pickands dependence functions are depicted in Figure 4.5 and the same parameters are chosen. Contour lines of the related densities are added.

4.5 Analysis of the arising subclass

Having presented two tractable instances in Section 4.2 and 4.3, in this section, the class of all distributions arising from the general IDT-frailty construction with IDT subordinators constructed as in Lemma 2.31 is investigated. We call those distributions ILF distributions and define the related class of distributions as follows:

$$\begin{aligned}\mathrm{ILF} &:= \big\{\{X_i\}_{i\in\mathbb{N}} \text{ constructed as in Equation (2.26) with } H \text{ given as in Lemma 2.31}\big\} \\ &= \Big\{\{X_i\}_{i\in\mathbb{N}}: \; X_i := \inf\Big\{t>0 : E_i < \int_0^\infty f(s/t)\mathrm{d}\Lambda_s\Big\}, \text{ with} \\ &\qquad \Lambda \text{ a Lévy subordinator, } f \text{ a suitable function}, E_i \text{ iid Exp}(1)\Big\},\end{aligned}$$

where a "suitable" function f fulfills the conditions in Lemma 2.31. The aim of this section is, on the one hand, to give a better understanding of the underlying dependence structure, and, on the other hand, characterize the resulting subclass of exchangeable MSMVE distributions.

We will illustrate the relation between ILF distributions and a generalization of the Marshall–Olkin shock model, see Equation (2.21). For that, we first have to define so-called scaled minima of Lévy-frailty (SMLF) distributions, where, based on Theorem 2.46, exchangeable MO sequences are also called Lévy-frailty distributions. In this section, to be consistent to the previous sections, we will only consider exchangeable MO sequences that correspond to Lévy subordinators without killing and call them exMO sequences.[3]

$$\mathrm{SMLF} := \Bigg\{\{X_i^{(l)}\}_{i\in\mathbb{N}}: \; X_i^{(l)} := \min\Big\{\frac{X_{l\,(i-1)+1}}{a_1},\dots,\frac{X_{l\,i}}{a_l}\Big\}, i\in\mathbb{N}, \text{with} \\ a_j>0,\, j=1,\dots,l,\, l\in\mathbb{N}, X\in \mathrm{exMO}\Bigg\}.$$

Consequently, SMLF sequences arise from exMO sequences by taking minima over l-dimensional scaled subsets, where l is called the order of the SMLF distribution. The definition is illustrated below, where $\{X_i\}_{i\in\mathbb{N}}$ denotes the exMO or Lévy-frailty sequence

[3] The results and also the construction in Lemma 2.31 could without much effort be extended starting from killed subordinators. However, this would not add much value and is therefore omitted.

and $\{X_i^{(l)}\}_{i\in\mathbb{N}}$ the resulting SMLF sequence.

$$
\begin{array}{l}
X_1,\quad X_2,\quad \ldots,\quad X_l,\qquad X_{l+1},\quad X_{l+2},\quad \ldots \\
\qquad\qquad\qquad\qquad\qquad\qquad\Downarrow \\
\underbrace{\tfrac{X_1}{a_1},\ \tfrac{X_2}{a_2},\ \ldots,\ \tfrac{X_l}{a_l}}_{\substack{\min\\ =:X_1^{(l)}}} \ \Big| \ \underbrace{\tfrac{X_{l+1}}{a_1},\ \tfrac{X_{l+2}}{a_2},\ \ldots\ \tfrac{X_{2\,l}}{a_l}}_{\substack{\min\\ =:X_2^{(l)}}} \ \Big| \ \tfrac{X_{2\,l+1}}{a_1}\ \ldots \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\ldots
\end{array}
$$

It is easy to see that an SMLF sequence is not an exMO sequence unless $a_1 = \ldots = a_l$. Furthermore, it is clear from the definition that SMLF sequences are exchangeable as well. The main result of this section will be that SMLF $\subset$ ILF $\subset$ $\overline{\text{SMLF}}$, where $\overline{\text{SMLF}}$ denotes the set of all distributions that arise as limits in distributions of SMLF distributions. While this result on its own does not provide additional insights into the underlying dependence structure, together with an intuitive representation of SMLF distributions, it becomes very meaningful.

In a first step, the following lemma provides an intuitive representation for SMLF distributions.

Lemma 4.14 (Generalized shock representation for SMLF vectors)
Let $(X_1^{(l)}, \ldots, X_d^{(l)})^\intercal$ denote a random vector which is part of an SMLF sequence of order l. Then, it allows for a representation via

$$X_i^{(l)} = \min_{\emptyset \neq I \subset \{1,\ldots,d\}: i \in I} \left\{ \min_{0 \le j \le l^{|I|}-1} \left\{ \frac{E_I^{(j)}}{a_{\pi_i^I(j)}} \right\} \right\}, \quad i = 1, \ldots, d, \tag{4.11}$$

with all $E_I^{(j)} \sim \mathrm{Exp}(\lambda_{|I|}^{(j)})$ independent, $a_1, \ldots, a_l > 0$ the positive constants in the definition of the SMLF distribution and $\pi_i^I : \{0, 1, \ldots, l^{|I|} - 1\} \to \{1, \ldots, l\}$ for $i \in I$. This is defined via $\pi_i^I(j) = \pi^{|I|}_{|\{m \in I : m \le i\}|}(j)$, where $\pi_i^k(j)$ are given by the unique representation

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

It is furthermore necessary that $\lambda_k^{(j_1)} = \lambda_k^{(j_2)}$, if $\pi_\cdot^k(j_1)$ is a permutation of $\pi_\cdot^k(j_2)$.

Remark 4.15
π_i^I are functions such that $\left(\pi_{i_1}^I(j), \ldots, \pi_{i_{|I|}}^I(j)\right)_{0 \le j \le l^{|I|}-1}$ yields an enumeration of the

points of $\{1,\ldots,l\}^{|I|}$*. One way to define this is as described above. To allow for a more intuitive understanding of the representation in Lemma 4.14, the representation is written out in full for* $d=l=2$*:*

$$\begin{aligned}
X_1^{(2)} &:= \min\left\{\frac{1}{a_1}E_1^{(0)}, \frac{1}{a_2}E_1^{(1)}, \frac{1}{a_1}E_{12}^{(0)}, \frac{1}{a_2}E_{12}^{(1)}, \frac{1}{a_1}E_{12}^{(2)}, \frac{1}{a_2}E_{12}^{(3)}\right\}\\
&= \min\left\{E_1, E_{12}, \frac{1}{a_2}E_{12}^{(1)}, \frac{1}{a_1}E_{12}^{(2)}\right\},\\
X_2^{(2)} &:= \min\left\{\frac{1}{a_1}E_2^{(0)}, \frac{1}{a_2}E_2^{(0)}, \frac{1}{a_1}E_{12}^{(0)}, \frac{1}{a_1}E_{12}^{(1)}, \frac{1}{a_2}E_{12}^{(2)}, \frac{1}{a_2}E_{12}^{(3)}\right\}\\
&= \min\left\{E_2, E_{12}, \frac{1}{a_1}E_{12}^{(1)}, \frac{1}{a_2}E_{12}^{(2)}\right\},
\end{aligned}$$

where in the second form, the random variables affecting the components equally are combined to improve clarity of the exposition, defining

$$\begin{aligned}
E_1 &:= \min\left\{\frac{1}{a_1}E_1^{(0)}, \frac{1}{a_2}E_1^{(1)}\right\}, \quad E_2 := \min\left\{\frac{1}{a_1}E_2^{(0)}, \frac{1}{a_2}E_2^{(0)}\right\},\\
E_{12} &:= \min\left\{\frac{1}{a_1}E_{12}^{(0)}, \frac{1}{a_2}E_{12}^{(3)}\right\}.
\end{aligned}$$

Thus, additional shocks are added to a Marshall–Olkin shock model, which affect the corresponding components scaled by differing constants. This can be seen as an extension of usual shock models, which are completely characterized and analyzed in Mai et al. (2015).

Furthermore, one could also combine shocks which obviously affect components in the same way, as was done in the above example, and rewrite Equation (4.11) *as*

$$X_i^{(l)} = \min_{\emptyset\neq I\subset\{1,\ldots,d\}: i\in I}\left\{E_I, \min_{\substack{0\leq j\leq l^{|I|}-1\\ j\notin\mathbb{N}_0\,(l^{|I|}-1)/(l-1)}}\left\{\frac{E_I^{(j)}}{a_{\pi_i^I(j)}}\right\}\right\}, \quad i=1,\ldots,d.$$

Using that representation, the resemblance to Marshall–Olkin models becomes more clear.

Proof (of Lemma 4.14)

For every subvector $(X_1,\ldots,X_{l\,d})^\intercal$ *of an exMO sequence, there exists an exchangeable Marshall–Olkin shock representation as given in Equation* (2.21)*. Plugging this representation into the definition of the SMLF distribution and consolidating all shocks that only enter components of a specific subvector* I *with a given vector of scalar factors* $\left(a_{\pi_{i_1}^I(j)},\ldots,a_{\pi_{i_{|I|}}^I(j)}\right)^\intercal$*, the stated representation can be found.* □

Remark 4.16
Obviously, Equation (4.11) *defines an exchangeable distribution. It is tempting to generalize this definition to the non-exchangeable case, however, we pass on that due to the already quite cumbersome notation. It is obvious that for every parameter specification, Equation* (4.11) *defines an exchangable distribution in a given dimension* $d \in \mathbb{N}$*. For readers familiar with the paper Mai and Scherer (2009a), it seems reasonable to ask for extendibility of the given distribution. It seems natural to assume that the extendible subclass is exactly given by SMLF. This poses an interesting question but lies outside the scope of this thesis.*

Lemma 4.14 provides some insights into the dependence structure underlying an SMLF distribution. It can be seen as a shock model, where compared to usual shock models, shocks affect the different components scaled by different factors. It is obvious from the given representation that SMLF distributions are MSMVE distributions. We will investigate their relation to the ILF class.

Theorem 4.17 (Relation of ILF and SMLF)
It holds that SMLF $\subset$ *ILF* $\subset \overline{\mathrm{SMLF}}$.

Proof
We start proving SMLF $\subset$ *ILF: Let* $X^{(l)} = \{X_i^{(l)}\}_{i\in\mathbb{N}}$ *denote a sequence with an SMLF distribution and* $X = \{X_i\}_{i\in\mathbb{N}}$ *the extendible MO sequence appearing in its definition. From Theorem 2.46 it follows that the extendible MO sequence has a stochastic representation via* $X_i := \inf\{t > 0 : E_i < \Lambda_t\}$*, with* $\{E_i\}_{i\in\mathbb{N}}$ *an iid* $\mathrm{Exp}(1)$ *sequence and* Λ *a Lévy subordinator. Consequently, for an arbitrary finite subset* $I \subset \mathbb{N}$ *and* $x_i > 0, i \in I$*, it holds that*

$$\begin{aligned}
\mathbb{P}\left(X_i^{(l)} > x_i,\, i \in I\right) &= \mathbb{P}\left(\frac{X_{(i-1)\,l+j}}{a_j} > x_i,\, 1 \le j \le l, i \in I\right) \\
&= \mathbb{P}\left(X_{(i-1)\,l+j} > a_j\, x_i,\, 1 \le j \le l, i \in I\right) \\
&= \mathbb{E}\left[\prod_{i\in I}\prod_{j=1}^{l} e^{-\Lambda_{a_j\,x_i}}\right] = \mathbb{E}\left[\prod_{i\in I} e^{-\sum_{j=1}^{l}\Lambda_{a_j\,x_i}}\right] \\
&= \mathbb{P}\left(Y_i > x_i,\, i \in I\right),
\end{aligned}$$

with

$$Y_i := \inf\left\{t > 0 : E_i < \sum_{j=1}^{l} \Lambda_{a_j\,t}\right\}.$$

As $\sum_{j=1}^{l} \Lambda_{a_j t} = \int_0^\infty \sum_{j=1}^{l} \mathbb{1}_{[0,a_j t]}(s)\,\mathrm{d}\Lambda_s = \int_0^\infty \sum_{j=1}^{l} \mathbb{1}_{[0,a_j]}(s/t)\,\mathrm{d}\Lambda_s$*, the first claim follows.*

Now, we want to prove that ILF $\subset \overline{\mathrm{SMLF}}$*: For a given random sequence* $X = \{X_i\}_{i\in\mathbb{N}}$*, we have to find a sequence* $\{X_k\}_{k\in\mathbb{N}}$ *with* $X_k := \{X_{i,k}^{(l_k)}\}_{i\in\mathbb{N}}$ *such that* $X_k \xrightarrow{d} X$ *for* $k \to \infty$*.*
As we know that $\int_0^\infty f(s)\,\mathrm{d}\Lambda_s$ *exists a.s., we know that for every sequence* $f_k \nearrow f$ *pointwise,* $\lim_{k\to\infty} \int_0^\infty f_k(s)\,\mathrm{d}\Lambda_s = \int_0^\infty f(s)\,\mathrm{d}\Lambda_s$ *follows by dominated convergence. Analogously,* $\lim_{k\to\infty} \int_0^\infty f_k(s/t)\,\mathrm{d}\Lambda_s = \int_0^\infty f(s/t)\,\mathrm{d}\Lambda_s$ *follows. This statement holds in particular for the sequence of function* $\{f_k\}_{k\in\mathbb{N}}$ *constructed in (Elstrodt, 1999, Theorem III.4.13), which are of the form*

$$f_k(s) = \sum_{j=0}^{k\,2^k} \frac{k}{2^k} \mathbb{1}_{A_j}(s) = \frac{1}{2^k} \sum_{j=0}^{k\,2^k} \mathbb{1}_{[0,a_j]}(s),$$

and fulfill $f_k \nearrow f$ *pointwise. It follows that*

$$\begin{aligned}\int_0^\infty f(s/t)\mathrm{d}\Lambda_s &= \lim_{k\to\infty} \int_0^\infty f_k(s/t)\,\mathrm{d}\Lambda_s = \lim_{k\to\infty} \int_0^\infty \frac{1}{2^k} \sum_{j=0}^{k\,2^k} \mathbb{1}_{[0,a_j]}(s/t)\,\mathrm{d}\Lambda_s \\ &= \lim_{k\to\infty} \int_0^\infty \sum_{j=0}^{k\,2^k} \mathbb{1}_{[0,a_j]}(s/t)\,\mathrm{d}\left(\frac{1}{2^k}\Lambda_s\right) = \lim_{k\to\infty} H_t^{(k)},\end{aligned}$$

with $H_t^{(k)}$ *defined accordingly. It is clear from the first part of this theorem that the ILF sequences corresponding to* $H^{(k)}$*,* $X_k = \{X_{i,k}^{(l_k)}\}_{i\in\mathbb{N}}$*, are elements of ILF* $\cap$ *SMLF for every* k*. Furthermore, it holds for arbitrary finite subsets* $I \subset \mathbb{N}$ *and* $x_i > 0, i \in I$*, that*

$$\begin{aligned}\mathbb{P}\left(X_{i,k}^{(l_k)} > x_i,\, i \in I\right) &= \mathbb{E}\left[\prod_{i\in I} e^{-\int_0^\infty f_k(s/x_i)\,\mathrm{d}\Lambda_s}\right] = e^{-\int_0^\infty \Psi_\Lambda\left(\sum_{i\in I} f_k(s/x_i)\right)\mathrm{d}s} \\ &\overset{k\to\infty}{\to} e^{-\int_0^\infty \Psi_\Lambda\left(\sum_{i\in I} f(s/x_i)\right)\mathrm{d}s} = \mathbb{P}\left(X_i > x_i,\, i \in I\right),\end{aligned}$$

where the convergence follows from the fact that $f_k \nearrow f$ *and thus, as* Ψ *is increasing and continuous,* $\Psi(\sum_{i\in I} f_k) \nearrow \Psi(\sum_{i\in I} f)$*, so dominated convergence can be applied. Thus, the second part of the claim is established.* $\square$

Corollary 4.18 (Survival function of SMLF vector)
The survival function of an SMLF vector $(X_1^{(l)}, \ldots, X_d^{(l)})^\intercal$ *with* $a_1, \ldots, a_l > 0$*, constructed from an exMO sequence corresponding to a Lévy subordinator* Λ *with Bernstein*

function Ψ_Λ, is given by

$$\mathbb{P}(X_1^{(l)} > x_1, \ldots, X_d^{(l)} > x_d) = \exp\left(-\sum_{j=1}^{d\,l} z_{(j)} \Big(\Psi_\Lambda(d\,l - j + 1) - \Psi_\Lambda(d\,l - j)\Big)\right),$$

with $z_{j\,l+i} := a_j\, x_i$ and $z_{(1)} \leq \ldots \leq z_{(d\,l)}$ the corresponding ordered list.

Proof
This follows from the proof of Theorem 4.17 and the observation therein that the corresponding function f is given by $f(s) = \sum_{j=1}^{l} \mathbb{1}_{[0,a_j]}(s)$. $\square$

From Theorem 4.17 follows that all ILF distributions can be approximated arbitrarily close by SMLF distributions. Consequently, every ILF random vector can be approximated by a shock construction as given in Lemma 4.14. This allows for a better understanding of the class ILF and can also be helpful when looking for approximating simulation schemes etc..

4.6 Conclusion

The present chapter developed two new classes of MSMVE distributions that give rise to many parametric families. The analysis conducted shows that these new classes are quite flexible. One clear advantage of the models presented is the availability of concrete stochastic models allowing for efficient simulation even in high dimensions. As a side product, results on integral transform representations and IDT subordinators were derived. Furthermore, we were able to relate the general class of distributions arising from the construction used to other well-known families, in particular we could give an approximation by extended shock models.

5 Modeling of discrete cash dividends

5.1 Motivation

For the pricing of derivatives with stocks as underlying, one has to model the future behavior of the stock, ideally incorporating the most prominent stylized features of the stock price evolution. Most research has been focused on modeling non-constant volatility, price jumps, or default risk, and considerable progress has been made in this regard. The effect of discrete cash dividends is often considered of secondary importance, i.e. it can not be fully ignored but is often treated using some rough rules of thumb. Thus, dividends are typically assumed to be payed continuously. However, in practice, dividends are payed at discrete dates, in most cases in cash. This has the effect that, loosely speaking, at the same moment the dividend is payed, the value of the stock decreases by the value of the dividend. This is clear as at that moment, holding the stock loses the entitlement to receiving the dividend, i.e. the value of the stock should decrease by the value of this disappeared benefit. We will ignore any kind of tax effects and consequently assume the value of the dividend to equal the size of the cash dividend and the size of the resulting jump in the stock price. Figure 5.1 illustrates an example of such a jump.

It is obvious that this systematic "negative" impact on the stock price evolution may not be ignored when pricing derivatives on the stock. A simple and also quite tractable way to incorporate it is to consider a continuously payed dividend. However, assuming it to be paid continuously does not meet the requirements when trying to consistently price derivatives with different maturities. Consider for example two European call options, which give the holder the right and not the obligation to buy a specific underlying stock at a specific date (the maturity) for a specified price (the strike price). Assume that one of them is maturing the day before the dividend is payed, the other the day after the dividend is payed, and that both have the same strike price. Naturally, the first should be worth considerably more than the second, as it entitles the holder to buy

Figure 5.1 Exemplary stock price development around an ex-dividend date (here, Allianz SE is considered). It is clearly observable that the stock price started trading considerably lower on the ex-dividend date. A dividend of 4.50 EUR was payed which almost coincides with the size of the jump. Source: Bloomberg.

the stock including the dividend at the same price as the second option, where one can buy the stock excluding the dividend payment. However, modeling the dividend to be continuously paid would yield almost the same price for both options. This example illustrates the necessity of properly modeling discrete cash dividends for the purpose of consistently pricing derivatives with different maturities. Similar examples can be constructed using American type options, i.e. options that allow for exercising the option at a freely chosen date before maturity.

Having observed the necessity of modeling discrete cash dividends, one has to consider which features of discrete dividend payments should be reproduced by a model. In particular for derivatives with long maturities, there is a considerable uncertainty about the involved dividend payments. Consequently, it would be desirable to allow for stochastic dividend payments with flexible structures that are able to replicate a trader's view.

Furthermore, in many applications, it is required to consistently price classical stock derivatives (e.g. put and call options) and credit products (i.e. products whose value depends on the creditworthiness of a company, as e.g., corporate bonds, government bonds, credit default swaps, etc.) with differing maturities. For that purpose, more complex credit-equity models have been developed that include the possibility of a default event into the stock price model. Thus, approaches are needed which allow for embedding discrete cash dividends into such quite complex credit-equity models.

However, so far, the existing literature does not provide a sufficient solution for all those requirements. There exist several papers investigating rules of thumb-type approaches to deal with discrete cash dividends in the Black–Scholes framework, see, e.g., Bos and Vandermark (2002); Frishling (2002); Bos et al. (2003); Buryak and Guo (2011), and further work that is concerned with the implementation of "piecewise Black–Scholes models", see, e.g., Haug et al. (2003); Vellenkoop and Nieuwenhuis (2006); Veiga and Wystup (2009); Étoré and Gobet (2011). In addition to that, there exists the paper Korn and Rogers (2005) dealing with proportional dividends in a Lévy process setting, and, rather recently, Buehler (2010) investigated affine dividends in a quite general setting. None of these approaches, however, sufficiently addresses all the requirements presented, which served as a motivation to start to work on that topic. The aim was to find a general framework that is very flexible with respect to the form of the dividends as well as the types of stochastic processes that can be used.

In contrast to the previous two chapters, this chapter is more practically oriented, i.e. with a focus on applications in mathematical finance. No deep results on the involved mathematical concepts are needed and therefore, these concepts are not included in Chapter 2 introducing the mathematical background. Instead, we tried to keep this chapter autonomous, introducing the necessary notation in the next subsection and requiring a basic level of background knowledge as it is included in a standard course on continuous time finance.[1] The results of this chapter are based on joint work with Jan-Frederik Mai. A less mathematical and simplified description of the key points of this chapter is published in Bernhart and Mai (2015).

5.1.1 Notation

As usual, we consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, where the filtration (a family of σ-algebras fulfilling certain consistency conditions) $\{\mathcal{F}_t\}_{t\geq 0}$ represents the available knowledge at each point in time $t \geq 0$. Furthermore, $S = \{S_t\}_{t\geq 0}$ denotes the stock price process, which is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted. At deterministic dates $0 < t_1 < t_2 < \ldots$, the stock is assumed to pay non-negative dividend amounts $D_{t_1}, D_{t_2}, \ldots$, with D_{t_k} an $\mathcal{F}_{t_k}$-measurable random variable for all $k \in \mathbb{N}$. For simplicity, we assume the so-called ex-dividend dates, i.e. the first dates the stock trades without the entitlement to the dividends, to coincide with the payment dates. Furthermore, we ignore any tax effects, i.e. the stock holder is assumed to receive the full dividend payment[2], and consequently, for simple no-arbitrage arguments, the following relation has to hold

$$S_{t_k} = S_{t_k-} - D_{t_k}, \quad \forall k \in \mathbb{N}.$$

A deterministic short rate $\{r_t\}_{t\geq 0}$ is considered, i.e. we ignore any interest rate risk. Consequently, the value of the corresponding bank account can be defined as $N_t = \exp\Big(\int_0^t r_s \,\mathrm{d}s\Big)$ and the discount factor for an interval $[t, T]$ is given by $B(t, T) := N_t/N_T = \exp\Big(-\int_t^T r_s \,\mathrm{d}s\Big)$.

5.1.2 Existing approaches and our contribution

A considerable portion of the existing approaches to deal with discrete cash dividends consists of tractable modifications of the classical Black–Scholes model. Their evident

[1] For example, we will not introduce definitions of put and call options (we only did that in the motivation to keep it understandable for people without background knowledge), the general pricing theory, etc..

[2] Alternatively, one could consider D_{t_k} to represent the after-tax value of the dividend.

purpose is to yield call and put prices that can be computed by changing the input parameters of the well-known Black–Scholes formula. Frishling (2002) introduces two such approaches and warns against their unthinking use. The dividend amounts are assumed to be deterministic in this setup. In the first approach, also called “escrowed model” in Haug et al. (2003) and Vellenkoop and Nieuwenhuis (2006), instead of the stock, the stock minus the present value of all discounted dividends[3] before the maturity T of the derivative considered, i.e.

$$A_t := S_t - \sum_{t<t_k \leq T} B(t,t_k)\, D_{t_k}, \quad t \leq T,$$

is modeled as a geometric Brownian motion. Consequently, $A_T = S_T$, i.e. S_T can be considered the value of a geometric Brownian motion with starting value $A_0 = S_0 - \sum_{t<t_k \leq T} B(t,t_k)\, D_{t_k}$. In the Black–Scholes formula, this is reflected by a change of the spot value. Most of the criticism leveled at this approach can be related to unthinking use of it, as we will see later on.

The second approach, the so-called “forward model”, see Vellenkoop and Nieuwenhuis (2006), models the stock price plus the previously received dividends[4] (reinvested at the riskless rate), i.e.

$$X_t := S_t + \sum_{0<t_k \leq t} D_{t_k}/B(t,t_k), \quad t \geq 0,$$

as a geometric Brownian motion. Similar to the escrowed model, this only corresponds to a modification of the strike price in the Black–Scholes formula for European options.

However, some authors argue that modeling the stock price as a piecewise geometric Brownian motion between the dividend dates, the so-called “piecewise lognormal model”, is closer to reality. Therefore, mixtures between the escrowed and the forward model are introduced in Bos and Vandermark (2002); Bos et al. (2003); Buryak and Guo (2011) to achieve a closer approximation of the option prices in the piecewise lognormal model.

By contrast, Vellenkoop and Nieuwenhuis (2006) directly work on the piecewise lognormal model and present a modification of the tree pricing method to compute option values in this context. Veiga and Wystup (2009) and Étoré and Gobet (2011) propose different approximations based on Black–Scholes prices using Taylor expansions. The

[3]Here, the dividend payments are assumed to be known at $t = 0$.
[4]As for the escrowed approach, dividend payments are assumed to be known at $t = 0$.

work of Haug et al. (2003) is theoretically formulated for a broader model class, but basically yields an iterated integration scheme to price options in the piecewise lognormal model. Furthermore, Haug et al. (2003) is one of the first to allow for a little more flexibility than affine dividends, i.e. dividend payments that do not depend linearly on the stock price. This is actually necessary to ensure the stock price being nonnegative in the piecewise lognormal model.

There exists not much work outside this Black–Scholes cosmos. Korn and Rogers (2005) consider proportional dividends in a setup driven by Lévy processes. We essentially choose the same starting point for our approach, but aim at more generality. Finally, the most elaborate work on discrete cash dividends is Buehler (2010), allowing for almost arbitrary driving processes and considering credit risk as well. However, only affine dividends, i.e. dividends of the form $D_{t_k} = a_k\, S_{t_k} + b_k$, with $a_k, b_k \geq 0$, are considered and it is necessary to make assumptions about D_{t_k} for all $k \in \mathbb{N}$.

To summarize, the existing literature is, with few exceptions, concerned with extensions of the Black–Scholes model. Furthermore, dividends are either constant or of affine form. However, especially for long-dated derivatives, more flexible forms are important that allow for constant dividends in the short run and more evolved parametric forms in the long run, which are able to capture a trader's view. The aim of this work is to develop a general framework which incorporates all these aspects and retains tractability. The underlying idea of our approach is that the stock price is considered as the expected sum over all future discounted dividends, a deliberation motivated by the well-known dividend discount model in economics, see, e.g., Gordon (1959), but also by work in mathematical finance, see, e.g., the starting point of Korn and Rogers (2005). By that, it is already ensured that the model is free of arbitrage, as will be pointed out.

The contribution of the present chapter consists of the following points:

1. For a modeling horizon $[0, T]$, in Section 5.2 we present a generic approach to construct a stochastic model for the stock price process and the dividend payments before T from an arbitrary closed martingale. Our ansatz is inspired by an idea of van Binsbergen et al. (2012), who split the stock price into "short-term assets" and "long-term assets" (in order to extract their respective values from derivative quotes). This ansatz allows to add flexible models for discrete cash dividends to an arbitrary no-dividend-paying stock price model. Regarding the resulting structure, it may also be considered an extension of the escrowed model.

2. It is discussed what is required in order to make our generic modeling approach applicable in practice. Tree pricing for arbitrary, path-dependent derivatives is possible whenever a tree-approximation for the underlying martingale is available. Such approximations are standard for most diffusion-driven models, see, e.g., Appendix F of Brigo and Mercurio (2006), and e.g. also available for Lévy-driven models, see, e.g., Maller et al. (2006). In some rare cases, it is even possible to circumvent tree pricing by transferring closed pricing formulas from the underlying no-dividend-paying stock price model to our setup. Sometimes, these formulas can at least be used in order to speed up tree pricing. Furthermore, we discuss desirable properties that make the specific parametric models convenient for implementation. The approach is flexible enough to allow for various dividend models apart from the often applied assumption of affine dividends. All these practical aspects can be found in Section 5.3.

3. In a case study, the generic framework is applied to the setup of the JDCEV credit-equity model of Carr and Linetsky (2006). We choose this specific model in order to highlight that our approach is well-suited to incorporate flexible discrete cash dividend parameterizations into stock price models that are far outside the Black–Scholes cosmos, distinguishing the present article from many earlier references on dividend modeling. This case study can be found in Section 5.4.

5.2 Consistent modeling of a stock price with discrete cash dividends

In this section, the general modeling framework is presented. In Section 5.2.1, we introduce the general ansatz and its implications. Section 5.2.2 shows how this ansatz can be transformed into a very flexible and tractable approach, representing one of the main contributions of this work.

5.2.1 The general modeling approach

Motivated by the dividend discount model which is well known in economics and often attributed to Gordon (1959), and by work in mathematical finance, see, e.g., Korn and Rogers (2005), the stock price process S_t at time t is considered to equal the expected sum over all future discounted dividend payments. This represents the central starting point for modeling the stock price in the present work, and we will see that it has some considerable advantages with respect to arbitrage-free pricing.

Definition 5.1 (Stock price model)
On the given probability space, the stock price process $S = \{S_t\}_{t\geq 0}$ is defined via

$$S_t := \mathbb{E}\Big[\sum_{t_k>t} B(t,t_k)\, D_{t_k} \,\Big|\, \mathcal{F}_t\Big], \quad t \geq 0, \tag{5.1}$$

where S_0 is assumed to exist, i.e. we assume that $S_0 < \infty$.

Throughout, we consider this ansatz as the most sophisticated and desirable definition of the stock price, and we search for arbitrage-consistent abstractions, respectively simplifications, in order to achieve practical viability. Though Equation (5.1) already imposes some structure on the stock price process, it will be shown later that almost every classical no-dividend model can be embedded into this framework. This is due to the central "truncation"-idea presented in the next section.

A natural next step consists in investigating the absence of arbitrage in this ansatz. We will see that starting from this approach ensures the resulting model to be "free of arbitrage", though this term needs some additional clarification below. This is also explicitly stated as a justification for this approach in Korn and Rogers (2005), who state that "*because we have begun by modeling the process of dividends, we never fall into the kind of inconsistencies that bedevil many common industry approaches*" (Korn and Rogers, 2005, p. 46).

Whenever holding a stock yields additional profits, be it dividend payments or earnings from lending out the stock, a crucial quantity to consider is the wealth that results from holding the stock. A simple way to model this wealth is to assume that the received dividend cash amounts are re-invested at the riskless rate. Additionally discounting this wealth results in the discounted wealth process X.

Definition 5.2 (Discounted wealth)
The process $X = \{X_t\}_{t\geq 0}$ that describes the discounted wealth from investing in the stock and reinvesting received dividends in the bank account is defined by

$$X_t := B(0,t)\Big(S_t + \sum_{0<t_k\leq t} \frac{D_{t_k}}{B(t_k,t)}\Big), \quad t \geq 0. \tag{5.2}$$

It is easy to see that X is a (closed) martingale.

Lemma 5.3 (Martingale property)
The process $X = \{X_t\}_{t\geq 0}$ is a closed martingale.

Proof
A simple reformulation of Equation (5.2) *yields*

$$X_t = \mathbb{E}\Big[\sum_{t_k>0} B(0,t_k)\, D_{t_k}\Big|\,\mathcal{F}_t\Big] = \mathbb{E}\left[X_\infty\Big|\,\mathcal{F}_t\right],$$

with $X_\infty := \sum_{t_k>0} B(0,t_k)\, D_{t_k}$ the present value of a stock investment. Recall that $X_\infty \in \mathrm{L}^1(\Omega,\mathcal{F},\mathbb{P})$ is equivalent to our assumption $S_0 < \infty$. □

This is already sufficient for the model to be free of arbitrage. The term free of arbitrage requires some additional comments in the continuous setting, here we will employ the concept of "no free lunch with vanishing risk" (NFLVR) as defined by Delbaen and Schachermayer (1994), generalizing strategies commonly known as arbitrage opportunities. It is known that - presupposed a positive stock price and ignoring dividend payments - the property (NFLVR) is implied by (and even equivalent to) the existence of a measure equivalent to $\mathbb{P}$ such that under this measure all discounted stock prices are local martingales. That is a version of the fundamental theorem of asset pricing, in its most general form developed by Delbaen and Schachermayer (1998) and here in the form stated in Jarrow et al. (2007), dropping the condition of locally boundedness using that the involved processes are positive. When considering dividends or other proceeds from holding a stock, the wealth process has to be considered instead of the stock price process. According to Lemma 5.3, the probability measure $\mathbb{P}$ under which we defined the processes S and X is already such a measure, so the property (NFLVR) follows directly. Consequently, valuing derivatives in this setup via expectations of discounted payoffs yields arbitrage-free prices.

Remark 5.4 (Incorporation of repo margins)
It would also be possible to include repo margins in this framework, as it is done, e.g., in Buehler (2010). That means one additionally considers the possibility to lend out the stock and, simplifying the mechanics a little, to earn a fee on the stock value in return. This can be modeled by an instantaneous repo margin δ_t (similar to the instantaneous short rate r_t). It has the following meaning: When directly reinvesting the proceeds of lending out the stock into the stock, starting with one unit of the stock and doing so for

a time period $[t,T]$ *results in* $\exp(\int_t^T \delta_s \,\mathrm{d}s)$ *units of the stock. To incorporate this, one would have to change Equation* (5.1) *to*

$$S_t := \mathbb{E}\Big[\sum_{t_k>t} B(t,t_k)\, \exp\Big(\int_t^{t_k} \delta_s \,\mathrm{d}s\Big)\, D_{t_k} \,\Big|\, \mathcal{F}_t\Big], \quad t \geq 0.$$

For X*, when considering the proceeds from lending, this would result in*

$$X_t = \mathbb{E}\Big[\sum_{t_k>0} B(0,t_k)\, \exp\Big(\int_0^{t_k} \delta_s \,\mathrm{d}s\Big)\, D_{t_k} \Big| \mathcal{F}_t\Big].$$

In total, all results would remain true and one would basically change from r_t *to* $r_t - \delta_t$*, however, one has to ensure that the assumption* $S_0 < \infty$ *still holds. If the risk of changes in the repo margin is ignored, i.e.* δ_t *is assumed to be deterministic, this can be done easily. For an analysis of the effect of stochastic repo rates for distressed stocks, see Bernhart and Mai (2014b).*

Remark 5.5
At first sight, it seems that Equation (5.1) *has to hold necessarily for arbitrage-free markets under any pricing measure. Actually, this is not true. (NFLVR) is equivalent to the existence of a measure equivalent to* $\mathbb{P}$ *such that under this measure all discounted assets are local martingales. However, even in complete markets, Jarrow et al. (2007) show that under such an equivalent local martingale measure, Equation* (5.1) *need not hold in general. In fact, following their definition, assuming (NFLVR) to hold and the market to be complete, Equation* (5.1) *holds if the market is "free of bubbles". Equivalently, starting from these assumptions, it is possible to find a measure* $\mathbb{P}$ *such that Equation* (5.1) *holds.*
So the only restriction imposed by starting from Equation (5.1) *is that it excludes the existence of bubbles as defined in Jarrow et al. (2007). However, apart from this restriction our setup is still general enough to comprise the whole battery of classical no-dividend models used in practice, as will be illustrated in the next section.*

5.2.2 A tractable reformulation

The critical observation for the reformulation presented is the fact that for practical applications, a model for the stock price process is only required on a finite time interval $[0,T]$, i.e. one only has to model $\{S_t\}_{t\in[0,T]}$ for some finite modeling horizon $T > 0$. This is due to the fact that all common derivatives have a finite maturity.[5] We will make use of this observation to make Equation (5.1) more tractable.

Obviously, the above quite general formulation is inconvenient due to the infinite series, though tractable special cases can indeed be defined, see, e.g., Korn and Rogers (2005). However, these tractable cases have to rely on very restrictive assumptions such as proportional dividends. In contrast, our approach to render the general formula (5.1) tractable and flexible consists of "truncating" the dividend series at the finite modeling horizon T and allowing for an almost arbitrary process to model the remaining part after T. It is shown below how this embeds classical modeling approaches and allows to extend them to incorporate very flexible dividend models for D_{t_k} with $0 < t_k \leq T$. It is very important to emphasize the dependence of the resulting model on the choice of T to avoid any misunderstandings (this aspect will be examined in more detail in Section 5.3.1).

As mentioned before, we split X_∞ into dividends obtained before and after T:

$$X_\infty = \sum_{0<t_k\leq T} B(0,t_k)\, D_{t_k} + \underbrace{\sum_{t_k>T} B(0,t_k)\, D_{t_k}}_{=:Y_\infty}.$$

Introducing a second closed martingale $\{Y_t\}_{t\geq 0}$ via $Y_t := \mathbb{E}[Y_\infty|\mathcal{F}_t]$ (one could also write $Y_\infty^{(T)}$ and $Y_t^{(T)}$ to emphasize the dependence on T), we obtain from this together with Equation (5.2) that for $t \in [0,T]$

$$\begin{aligned} X_t = \mathbb{E}[X_\infty|\mathcal{F}_t] &= \mathbb{E}\Big[\sum_{0<t_k\leq T} B(0,t_k)\, D_{t_k}\Big|\, \mathcal{F}_t\Big] + Y_t \\ &= \sum_{t_k\leq t} B(0,t_k)\, D_{t_k} + B(0,t)S_t. \end{aligned}$$

From this, the central representation can be easily derived.

Theorem 5.6 (Central reformulation)
On the interval $[0,T]$*, the stock price allows for the representation*

$$S_t = \frac{1}{B(0,t)}\, Y_t + \sum_{t<t_k\leq T} B(t,t_k)\, \mathbb{E}\Big[D_{t_k}\Big|\, \mathcal{F}_t\Big], \quad t \in [0,T], \tag{5.3}$$

[5] There are exceptions like upper Tier 2 bonds, which often are perpetual bonds, but to deal with them, one usually considers them to be finite as well.

with Y the closed martingale introduced above.

This is the main formula we want to use, and it allows for the following two-step procedure to model $\{S_t\}_{t\in[0,T]}$:

(1) Model the process $\{Y_t\}_{t\in[0,T]}$ as your favorite, non-negative closed martingale that is a Markov process[6], with appropriately chosen filtration, e.g.:

- (Non-defaultable case:) Define $\mathcal{F}_t := \sigma(Y_s : s \leq t)$, i.e. consider the natural filtration of $\{Y_t\}_{t\in[0,T]}$ as market filtration.
- (Credit-equity modeling case:) In credit-equity models, such as in the case of a defaultable Markov diffusion (see, e.g., Carr and Linetsky (2006); Linetsky (2006); Bielecki et al. (2011)), $\{Y_t\}_{t\in[0,T]}$ is often modeled via $Y_t = Z_t\,\mathbb{1}_{\{\tau>t\}}$ where $\{Z_t\}_{t\in[0,T]}$ represents the pre-default process and τ the "default time". This default time can be observed in the market and hence needs to be incorporated into $\mathcal{F}_t$, as it is not necessarily measurable with respect to the filtration generated by Z. This issue is usually resolved by defining $\mathcal{F}_t := \sigma(Z_s : s \leq t) \vee \sigma(\mathbb{1}_{\{\tau\leq s\}} : s \leq t)$ respectively its usual augmentation if necessary.

(2) Model $D_{t_k} \geq 0$ for $0 < t_k \leq T$ arbitrary, $\mathcal{F}_{t_k}$-measurable. Here, we will consider $D_{t_k} := f_{t_k}(Y_{t_k})$ for some measurable function f_{t_k} satisfying $f_{t_k}(0) = 0$. The latter is a consistency condition for credit-equity models, which basically says that there are no dividend payments after default. The construction furthermore ensures that D_{t_k} is $\mathcal{F}_{t_k}$-measurable.

For the special case of no dividend payments in $[0, T]$, i.e. $t_1 > T$, it follows easily that $Y_t = X_t$ and Equation (5.3) reduces to

$$S_t = \frac{1}{B(0,t)}\,Y_t, \quad t \in [0, T].$$

This is clearly consistent with any classical model without dividends, i.e. any classical model can be formulated like that. Since the model interval is finite, it suffices to model $\{Y_t\}_{t\in[0,T]}$ such that Y_T is integrable in order to guarantee that the martingale is closed. This is indeed the case in all classical models.

[6] In general, the Markov property is not necessary. For the approach presented, however, it considerably simplifies our considerations with respect to pricing. Therefore, this assumption is made in the following.

It is helpful for understanding the procedure presented that Y_t has an intuitive meaning related to the forward price. Let $F(t,T)$ denote the fair forward strike price, i.e. the price one would agree on at time $t < T$, for buying the stock at T. In case of deterministic (riskless) dividends, a simple replicating strategy respectively no-arbitrage consideration yields

$$F(t,T) = \frac{1}{B(t,T)} \Big(S_t - \sum_{t<t_k\leq T} B(t,t_k)\, D_{t_k} \Big), \quad t \in [t,T].$$

Under the (for long maturities T somewhat artificial) assumption that the considered future dividend payments are traded assets (via dividend swaps or similar derivatives), a replicating strategy yields the same formula for non-deterministic dividend payments, replacing the dividend payments by their market value. Assuming $\mathbb{E}[D_{t_k} | \mathcal{F}_t]$ to coincide with that market value, or directly pricing under $\mathbb{P}$ (in contrast to the previously mentioned argument that is only based on static replication without modeling at all), the following expression for the fair forward price has to hold[7,8]

$$F(t,T) = \mathbb{E}[S_T | \mathcal{F}_t] = \frac{1}{B(t,T)} \Big(S_t - \sum_{t<t_k\leq T} \mathbb{E}\Big[D_{t_k} \Big| \mathcal{F}_t \Big] \Big), \quad t \in [t,T].$$

Consequently, $F(t,T) = Y_t/B(0,T)$, i.e. the martingale $\{Y_t/B(0,T)\}_{t\in[0,T]}$ denotes precisely the forward strike process for an equity forward with maturity T. In other words, we basically model the forward with maturity T and earlier dividends are considered functions of this forward.

Allowing D_{t_k} to depend not only on Y_{t_k} but additionally on t_k is very flexible in the sense that one could, e.g., include already announced dividend payments (e.g. by setting $f_{t_k}(y) = c_{t_k} \mathbb{1}_{\{y>0\}}$ with announced dividend amount $c_{t_k} > 0$). Furthermore, the approach bears some similarities with the so-called escrowed approach, used for example by Roll (1977); Geske (1979); Whaley (1981), as the stock price minus the present value of the dividends is modeled. It could be seen as an extension of that approach, allowing for more flexible dividend structures and almost arbitrary driving processes. However, our derivation explicitly highlights the dependence of the model on T.

[7] This actually yields a possible approach to extract the market value of future dividend payments from other market quotes, as will be mentioned later on.

[8] Note that there is no explicit consideration of credit risk. Due to the replication argument, it does not show up in the S_t-term, and regarding the dividends considered, it is implicitly incorporated by the expected value respectively market value.

Remark 5.7

a) *A similar splitting is applied to the dividends of a whole stock index in van Binsbergen et al. (2012). The first part is called the short-term asset and the second the long-term asset. Market data from derivatives markets is used to extract the price of the first one, even subdivided into different dividend strips.*

b) *The number S_0 is obviously observable. Furthermore, the equation*

$$Y_0 + \sum_{0<t_k\leq T} B(0,t_k)\,\mathbb{E}[D_{t_k}] = S_0,$$

has to hold, i.e. Y_0 and the sum of expected discounted dividend payment amounts on $[0,T]$ must sum up to the initial stock price. Consequently, one of both determines the other. Either, Y_0 can be directly implied from market data due to the relation with the forward value. Or, the expected discounted dividend payments are extracted from derivative quotes or determined otherwise, which in turn determines Y_0. Having found $Y_0 \geq 0$ and having modeled Y as a non-negative process, the non-negativity of S is ensured. This is helpful as thus, no additional restrictions on the dividend payment amounts as used, e.g., in Haug et al. (2003) are needed. A further advantage of this approach over other approaches as, e.g., Korn and Rogers (2005) or Buehler (2010), is that one only has to model a finite number of dividend payments. All dividend payments after T (for which it would be difficult to extract market values from derivatives) are enclosed in the process $\{Y_t\}$. This abstraction allows to concentrate on quite flexible modeling approaches for dividends before T, which will be seen in the next section.

c) *It can not be stressed enough that the resulting model depends on the specific choice of T. The arbitrage opportunities, which the escrowed model allegedly admits, see, e.g., Beneder and Vorst (2002); Haug et al. (2003), follow directly from ignoring the dependence on T, i.e. by inconsistently modeling $Y_t^{(T_1)}$ like $Y_t^{(T_2)}$ with $T_1 < t_k < T_2$. This fact will be highlighted in more detail in the next section.*

d) *Since interest rates are deterministic, the model is driven by only one stochastic factor being $\{Y_t\}_{t\in[0,T]}$. However, it is possible that $\{Y_t\}_{t\in[0,T]}$ itself consists of several stochastic factors, e.g. stochastic volatility models such as, e.g., Heston (1993). However, for applications it is useful to ensure the existence of efficient tree approximations of $\{Y_t\}_{t\in[0,T]}$, which is considerably more complicated when it consists of several factors.*

e) *Some authors argue that the dividend payments D_{t_k} should depend on S_{t_k}, see, e.g., affine dividend structures as used in Villiger (2006) or Buehler (2010). In contrast, we define them dependent on Y_{t_k}, mainly due to viability. This is a small difference but might be considered a theoretical flaw. However, starting from Equation* (5.1)*, it does not seem logically consistent to model the dividend as a function of a quantity the dividend is part of. Furthermore, we think that - apart from the obvious advantage with regards to tractability - our approach also has another advantage. The underlying idea for a dependence of D_{t_k} on S_{t_k} is that the evolution of the stock price is a good indicator for the economic well-being of the company, and thus also for the dividend payment amount. But why should a dividend payment at t_k, causing a drop in the stock price, be an indicator for a lower dividend payment at t_{k+1}? Exactly this effect would be observable for expected future dividend amounts at ex-dividend dates in the case of direct dependence on the stock price. On the contrary, as $\{Y_t\}_{t\in[0,T]}$ does not exhibit a jump triggered by dividend payments, this effect is not observed for our modeling ansatz. Still, the evolution of $\{Y_t\}_{t\in[0,T]}$ can be considered a good indicator for the economic well-being of the firm due to the relation to the forward, so the basic idea is similar.*

As stated before, the modeling approach presented is very flexible with respect to the model choice for $\{Y_t\}_{t\in[0,T]}$. Consequently, everyone should be able to use the model which incorporates all the features considered necessary. However, one recommendation can be deduced from our ansatz to ensure logical consistence: We decompose the stock and explicitly model the value of the dividend payments until T. Assuming the amounts of these payments to be fixed, e.g. if the dividend payment is already announced, these payments represent something like the coupon payments of a corporate bond. Modeling $\{Y_t\}_{t\in[0,T]}$ as a non-defaultable diffusion, such as, e.g., in the Black–Scholes world, those payments contribute to the stock price being discounted with the riskless rate. Thus, their value might be significantly overestimated compared to the value of similar payments due by the same company, traded in bond markets. This is in particular undesirable as in the capital structure, coupon payments should rank higher than dividend payments. A defaultable diffusion model overcomes this problem, which explains why the choice of such a model seems to be more “natural” in the present context. Stated differently, there is a model inherent incentive for using credit-equity approaches, since explicitly considering dividend payments introduces a similarity to debt instruments.

5.3 Practical aspects

We want to explain how the approach presented in the previous section can be put to use. Arising questions, possible answers and solutions to common problems are illustrated. The general procedure for pricing derivatives will usually be as follows:

1. Define the modeling horizon $[0, T]$ and choose a closed martingale $\{Y_t\}_{t\in[0,T]}$. Section 5.3.1 highlights the dependence of the resulting model on the choice of T.

2. Determine $\sum_{0<t_k\leq T} B(0, t_k)\,\mathbb{E}[D_{t_k}]$ respectively the starting value Y_0 of the closed martingale $\{Y_t\}_{t\in[0,T]}$. Consequently, determine $\mathbb{E}[D_{t_k}]$ for all $t_k \leq T$ and choose the functional form f_{t_k} consistent with the previous assumptions. Section 5.3.2 illustrates that procedure.

3. Price derivatives, e.g. as explained in Section 5.3.3.

5.3.1 Dependence on T

Our modeling approach with dividends is naturally equipped with a modeling termination T. The choice of T is crucial for a reasonable implementation of the model: On the one hand, T has to be large enough such that the maturity of every derivative considered is included in $[0, T]$ to allow for a consistent valuation of these products. On the other hand, one wants to choose T as small as possible to reduce the number of dividend payments that have to be modeled (all dividends before T).

Once T and the related model on the interval $[0, T]$ in Equation (5.3) has been fixed, the pricing of every product has to rely on that model. This means in particular that even for derivatives with maturity $T_1 < t_1$, the involved expression for $S_t,\ t \leq T_1$, includes a term representing every dividend payment until T. This might seem unnecessarily complex or even counterintuitive to some, but it is the only way for a consistent modeling of products with maturities in the range considered.[9]

As the so-called escrowed model could be considered a special case of our approach and as it is often claimed that this model allows for arbitrage opportunities, we investigate the related examples and show how these arbitrage opportunities result precisely from

[9] Buehler (2010) argues for a similar structure, even modeling all dividends.

ignoring the advice stated here. The example presented here is constructed similarly to examples found in Beneder and Vorst (2002); Frishling (2002); Haug et al. (2003).

Example 5.8
Consider two American call options on a stock with strike 130 *and maturities* 1 *and* 363/365*,* $S_0 = 100$*, one already fixed dividend payment* $D_{364/365} = 7$*, and a constant short rate* $r = 6\%$*. We choose* $T = 1$*, which is a straight-forward choice based on the previous considerations. Furthermore,* $\{Y_t\}_{t\in[0,1]}$ *is assumed to follow a geometric Brownian motion with* $\sigma = 30\%$*. This yields*

$$S_t = Y_0 e^{(r-\frac{\sigma^2}{2})\,t+\sigma\,W_t} + 7\,e^{-r\,(364/365-t)}\mathbb{1}_{\{t<364/365\}}, \quad t \in [0,1],$$

as a model for the stock, with $\{W_t\}_{t\in[0,1]}$ *a standard Brownian motion and* $Y_0 \approx 93.41$*. Let* $V_A(t,T,S_t,K)$ *denote the price of an American call option with maturity* T *and strike* K *at time t, given the stock value* S_t*, and* V_E *the European equivalent. Applying the pricing approaches presented later, this yields for the two American call option values* $V_A(0,1,100,130) = 4.284$ *and* $V_A(0,363/365,100,130) = 4.267$*.*[10] *These two prices seem reasonable on first sight and there is (of course) no arbitrage opportunity arising from them. However, ignoring the dependence of the model on* T *and wrongly assuming that the given properties of* Y *also hold for the model with maturity* 363/365*, i.e. wrongly assuming* $Y^{(363/365)}$ *to be distributed as* $Y^{(1)}$*, would yield*

$$S_t = 100\,e^{(r-\frac{\sigma^2}{2})\,t+\sigma\,W_t}, \quad t \in [0, 363/365].$$

This approach corresponds to carelessly applying the escrowed model. The price of the shorter American call option arising from this model can easily be computed. It is well known that in such a framework this price equals the price of a European call option given by $V_A(0,363/365,100,130) = V_E(0,363/365,100,130) = 4.883$*. This price, together with the above price for the option with maturity* 1*, would yield an obvious arbitrage opportunity, selling the short-dated option versus the long-dated option. Consequently, the example clarifies the problems that can arise from ignoring the dependence of the model specification on* T*. The arguments brought forward against the escrowed model are based exactly on such examples.*

[10] Alternatively, in this simple setup, for the computation of $V_A(0,1,100,130)$, the previously mentioned Roll–Whaley–Geske approach (see Roll (1977); Geske (1979); Whaley (1981)) is applicable. Furthermore, using basic inequalities, it can be shown that the value of $V_A(0,363/365,100,130)$ can be computed as the value of a European call option in a Black-Scholes model with starting value Y_0, volatility σ, and strike $130 - D * \exp(-r/365)$.

5.3.2 Parametric form of dividend payments

One of the important modeling aspects is the modeling of D_{t_k} for the involved $t_k \leq T$. In a first step, one has to determine the corresponding expected values $\mathbb{E}[D_{t_k}]$. Based on the previously presented link to the forward, $\sum_{0<t_k\leq T} B(0,t_k)\,\mathbb{E}[D_{t_k}]$ respectively Y_0 might be extracted from market prices of derivatives. Either, the forward strike price for the maturity considered is directly observable, or it can be extracted from option prices using the put-call parity. For some stocks, there might even be available values for the individual $\mathbb{E}[D_{t_k}]$ separately, using the whole option surface, several forward strike prices, or dividend futures. Alternatively, one could use simpler approaches, either relying on analyst forecasts or employing simple rules of thumb as $\mathbb{E}[D_{t_k}] = c$ with a constant $c \geq 0$. A reasonable value for c might be $S_0/\sum_{t_k>0} B(0,t_k)$, which is an observable quantity, and is consistent with the untruncated ansatz in Equation (5.1). However, we do not want to elaborate on that aspect here.

Having fixed $\mathbb{E}[D_{t_k}]$ for the relevant k, one has to define f_{t_k} consistently, i.e. one has to make sure that the following equations are fulfilled:

$$\mathbb{E}[D_{t_k}] \stackrel{!}{=} \mathbb{E}[f_{t_k}(Y_{t_k})].$$

A convenient choice is

$$f_{t_k}(y) := \mathbb{E}[D_{t_k}]\frac{h(y)}{\mathbb{E}[h(Y_{t_k})]}, \tag{5.4}$$

with a function $h : [0,\infty) \to [0,\infty)$ satisfying $h(0) = 0$ and having a desired shape.[11] In this context and in particular for the pricing of derivatives later on, a closed form expression for the distribution of Y_{t_k} is highly desirable such that the normalizing denominator can be numerically computed for arbitrary functions h. It is even better to look for specific choices of h and Y_t such that closed form solutions are available.

Regarding the shape of h, there are basically no restrictions. It is reasonable to choose a non-decreasing function, which still leaves space for many different considerations. Three different illustrative dividend specifications will be considered in this document, with the aim to present the flexibility of the model rather than to give an exhaustive overview of possible specifications:

[11] Actually, one could also write h_{t_k} as the choice of the shape might of course be different for different t_k. However, for reasons of clarity of notation, we omitted that additional index.

- $h_1(y) = y$, which corresponds to dividend payments proportional to Y_t,

- $h_2(y) = \mathbb{1}_{\{y>b\}}$ with $b \geq 0$, which corresponds to a dividend payment in case the process Y_t is above a specific threshold, where the dividend amount does not depend on the level of exceedance,

- as an example of a more complex specification, the following function is chosen:

$$h_3(y) = \begin{cases} Y_0 \left(\frac{y}{l\,Y_0}\right)^a, & y \leq l\,Y_0, \\ Y_0, & l\,Y_0 < y \leq u\,Y_0, \\ y - (u-1)Y_0, & y > u\,Y_0, \end{cases}$$

  with $a \geq 1$ and $0 < l < 1 < u$. This corresponds to a constant dividend in an interval (defined by l and u) enclosing the expected value of the process. Above this interval, the dividend payment is increasing linearly with the value of the process. Below the interval, it is decreasing with the speed defined by a.

The form of the three different specifications is sketched in Figure 5.2. Note that h_i only defines the shape of the functions f_{t_k}, the absolute values of the functions h_i are irrelevant as the functions will be normalized in order to guarantee consistency. An example for the resulting normalized functions f_{t_k} can be found in Figure 5.4 below.

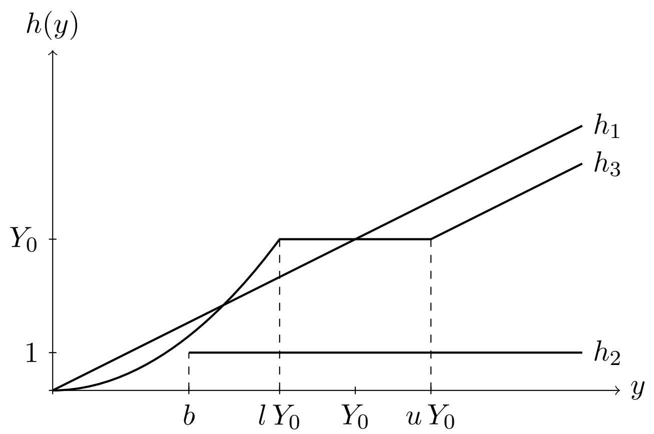


Figure 5.2 A sketch of the form of the three different specifications of h.

5.3.3 Derivatives pricing

In the presented setting, the stock price S_t for $t \in [0,T]$ is given by

$$S_t = \frac{1}{B(0,t)} Y_t + \sum_{t<t_k \leq T} B(t,t_k)\, \mathbb{E}\left[f_{t_k}(Y_{t_k}) \middle| \mathcal{F}_t \right], \tag{5.5}$$

i.e. it is given as the sum of Y_t and conditional expectations of the form $\mathbb{E}[f_{t_k}(Y_{t_k})|\, \mathcal{F}_t] = \mathbb{E}[f_{t_k}(Y_{t_k})|\, Y_t]$. It follows that $S_t = g_t(Y_t),\ t \in [0,T]$, with functions g_t for which ideally closed-form expressions are available or which can be evaluated numerically in an efficient way. This depends on the choice of the functions h, as the computation of such a conditional expectation essentially boils down to computing $\mathbb{E}[h(Y_{t_k})|\, Y_t]$.

The essential implication of this observation is that a derivative on S can be considered a derivative on Y. However, even if there are closed formulas available for those involved conditional expectations, in most cases, it will not be possible to derive closed form expressions for the pricing of derivatives on the stock. Thus, pricing requires tree approximations or comparable techniques. One of the major advantages of the approach described in the present chapter is that tree pricing including dividends is easily possible for every specification of $\{Y_t\}_{t\in[0,T]}$ for which efficient tree approximations are known. This includes most of the standard processes like geometric Brownian motion, exponential Lévy processes (see Maller et al. (2006)) or 1, 5-factor credit-equity models (see, e.g., Carr and Linetsky (2006); Linetsky (2006); Bielecki et al. (2011)), for which standard methods such as presented in Appendix F of Brigo and Mercurio (2006) can be applied. Thus, special dividend adjustment methods for the construction of trees as presented in Vellenkoop and Nieuwenhuis (2006) are not necessary.

The pricing procedure is rather simple[12]:

Algorithm 5.9

(i) Build a tree for Y on $[0,T]$, i.e. for a discrete set of time points $0 < u_1 < \ldots < u_N = T$, which includes the dividend payment dates.

(ii) Consider the derivative as a derivative on Y and use standard backward induction techniques. Whenever required, compute the value of $S_{u_k} = g_{u_k}(Y_{u_k}),\ k < N$.

[12]Showing convergence of tree pricing algorithms is by no means a simple task, see, e.g., Müller (2009) for a well-explained illustration of the necessary steps. For convergence of American-type options, see Amin and Khanna (1994). However, if convergence is proven for a given model for Y and a specific derivative, it is only required to extend this convergence result to the derivative given by substituting Y_t by $g_t(Y_t)$.

Depending on the number of derivatives priced, their type, and the computational efforts needed to compute the functions g_{u_k}, this procedure might be quite slow. Therefore, we propose the following approximation of the above pricing procedure:

Algorithm 5.10

(i) Build a tree for Y on $[0,T]$, i.e. for a discrete set of time points $0 < u_1 < \ldots < u_N = T$, which includes the dividend payment dates.

(ii) At the end nodes, representing $t = u_N = T$, set $S_T = Y_T/B(0,T)$.

(iii) Using standard backward induction techniques, the value of $S_{u_k}, k < N$, at every node can be derived using Equation (5.5), since the expected discounted value of future dividend payments can be seen as a derivative itself. Thus, its value at every node is obtained via the tower property of conditional expectation. This can be done while simultaneously pricing the derivatives considered, thus no additional backward induction is required.

For most applications, at least one of these two tree pricing routines should be efficient and easy to apply. However, the following example illustrates that there are some business cases for which further considerations need to be made.

Example 5.11

Assume we want to price an upper Tier 2 bond consistent with a battery of short-dated equity options, i.e. we want to use a credit-equity model. Such bonds are perpetual with call features, i.e. we have infinite maturity and American style features. For this, we can set up a defaultable Markov diffusion model, adapted to our dividend framework. It is standard (and unavoidable) to tackle the infinite maturity problem by truncation, i.e. assume a finite but very long maturity $T < \infty$. Due to the long maturity, tree pricing takes its time. Therefore, and because of data availability, one might have the idea to fit the model to the (much shorter-dated) equity option data, e.g. a battery of puts and calls, and then price the upper Tier 2 bond using the fitted parameters. However, for the pricing of these shorter-dated options using Algorithm 5.9 or 5.10, one still would have to use the "long" tree for the interval $[0,T]$. Hence, the calibration routine might be very tedious.

There is a way to circumvent this issue. In fact, Algorithm 5.9 allows to build a consistent tree for the interval $[0,T_1]$ only, with $T_1 < T$. One can even combine both algorithms, i.e. one can use Algorithm 5.10, in Step (i) building a tree for Y on $[0,T_1]$ only. For that, in Step (ii), one has to compute the value of S_{u_N} using Equation (5.5) as in Algorithm

5.9, i.e. setting $S_{u_N} = g_{u_N}(Y_{u_N})$. Then, one can proceed with Step (iii) as usual. For the purpose presented in Example 5.11, this procedure should be considerably more efficient than using the larger tree for the interval $[0, T]$.

To illustrate the procedure, one step of the corresponding tree for a defaultable diffusion (and basically every other specification allowing for tree approximations) is illustrated in Figure 5.3. We explain how to compute the stock price at the node $Y_{u_i} = x_0$ of the tree in time step u_i from the nodes at time step u_{i+1} which can be reached from the given node. In Figure 5.3, these are four nodes, which can be reached with respective probabilities $p_1, \ldots, p_4$, with $p_1 + \ldots + p_4 = 1$. In the JDCEV model, one transition probability corresponds to defaulting, i.e. $p_4 = \mathbb{P}\big(Y_{u_{i+1}} = 0 \big| Y_{u_i} = x_0\big)$. With

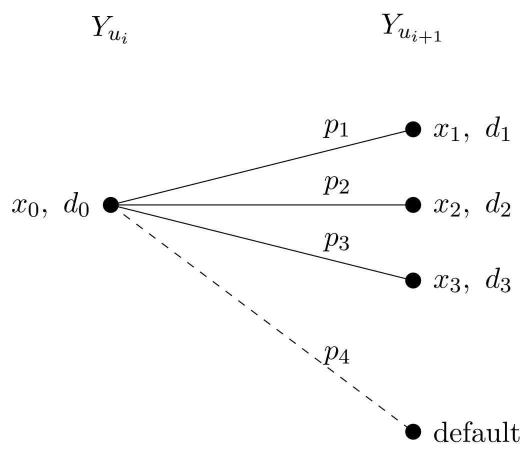


Figure 5.3 A sketch of one step of the corresponding tree.

$$d_j := \mathbb{E}\Big[\sum_{u_{i+1} < t_k \leq T} B(u_{i+1}, t_k) D_{t_k} \Big| Y_{u_{i+1}} = x_j \Big], \quad j = 1, 2, 3,$$

Equation (5.5) implies that the stock price at these nodes is given by $s_j := x_j / B(0, u_{i+1}) + d_j$, $j = 1, 2, 3$. Furthermore, d_0, the expected sum of future dividends at the previous time step u_i, given $Y_{u_i} = x_0$, can be computed either via

$$d_0 = B(u_i, u_{i+1}) \sum_{j=1}^{3} p_j \, d_j,$$

in case u_{i+1} is no dividend date, or via

$$d_0 = B(u_i, u_{i+1}) \sum_{j=1}^{3} p_j \left(d_j + f_{u_{i+1}}(x_j)\right),$$

in case u_{i+1} is a dividend date (with corresponding dividend payoff function $f_{u_{i+1}}$). The stock price at this node is given accordingly via $s_0 = x_0/B(0, u_i) + d_0$. The pricing of derivatives can be performed analogously using the computed values of the stock price at every node, additionally taking into account the path leading to a possible default.

5.4 Case study

In this section, the proposed approach is illustrated by presenting a possible implementation. Consistent with our previous reasoning for a joint modeling of credit and equity components respectively defaultable stock price models, the jump to default extended constant elasticity of variance (JDCEV) model of Carr and Linetsky (2006) is chosen for the closed martingale process $\{Y_t\}_{t\in[0,T]}$. Furthermore, this model choice is able to clearly demonstrate the claimed flexibility of the approach presented - distinguishing it from many of the aforementioned references.

5.4.1 The martingale Y

The JDCEV model represents one of the most popular defaultable stock price models, as it incorporates the most relevant features of the stock price plus a default component while still retaining analytical tractability. Here, a slightly changed notation and some simplifications of the original model are employed. For a thorough analysis of the JDCEV model, the reader is referred to the original paper Carr and Linetsky (2006).

A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which supports a Brownian motion $\{W_t\}_{t\geq 0}$ and an independent random variable $E \sim \text{Exp}(1)$ is considered. For a given maturity T, $\{Y_t\}_{t\in[0,T]}$ is defined as

$$Y_t = Z_t \, \mathbb{1}_{\{\tau > t\}},$$

where $\{Z_t\}_{t\in[0,T]}$ represents a pre-default process and τ the company's default time. The pre-default process is modeled as the solution to the SDE

$$\mathrm{d}Z_t = Z_t \left(\lambda \left(\frac{Z_t}{Z_0} \right)^{2\beta} \mathrm{d}t + \sigma Z_t^{\beta} \mathrm{d}W_t \right), \quad Z_0 = Y_0 > 0, \lambda > 0, \sigma > 0, \beta < 0,$$

where 0 is defined as a killing boundary for those parameter constellations that enable a diffusion to zero. Furthermore, the default time τ is modeled via $\tau := \min\{\hat{\tau}, \tau_0\}$, with

$$\tau_0 := \inf \{t \in [0, T] : Z_t = 0\},$$

and

$$\hat{\tau} := \inf \left\{ t \in [0, T] : \int_0^t \lambda \left(\frac{Z_s}{Z_0} \right)^{2\beta} \mathrm{d}s > E \right\}.$$

As usual, $\inf \emptyset := \infty$, where $\emptyset$ denotes the empty set. Consequently, default is defined as the first time either the pre-default process diffuses to zero or the jump to default governed by $\hat{\tau}$ occurs.[13] The jump to default is modeled using a generic reduced-form approach with the default intensity given as a function of the pre-default process. The considered filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is defined by $\mathcal{F}_t = \sigma(Z_s : s \leq t) \vee \sigma(\mathbb{1}_{\{\tau > s\}} : s \leq t)$. The model can be seen as an extension of the CEV model including a jump to default component which is modeled via a reduced-form approach. Using the given parameterization, the different parameters allow for an intuitive interpretation. The parameter β governs the relation between the level, the volatility, and the default intensity of $\{Y_t\}_{t\geq 0}$. The instantaneous default intensity $\lambda \left(Z_s/Z_0\right)^{2\beta}$ is linked to the evolution of the process, where λ represents its starting value. The drift of the pre-default process includes the default intensity to ensure the martingale property of the process $\{Y_t\}_{t\geq 0}$.

The distinguishing feature of the model is that important building blocks like European claims with no recovery and fixed recovery payments can be priced in closed form. This result is shown by expressing the pre-default process as a time-changed Bessel process and subsequently applying a change of measure. This change of measure removes the typical quantities stemming from the reduced-form approach and under the resulting

[13] In Carr and Linetsky (2006), it is claimed that $\hat{\tau} < \tau_0$, a.s., which would simplify the exposition of the model. However, the results shown in their paper do not depend on that property with exception of Proposition 5.5(ii), which we will not make use of. Consequently, we use the original presentation of the model.

measure, the Bessel process can not diffuse to zero anymore. By that, the involved expectations can be transformed to expectations of functions of a random variable with known density. As we have seen that a known density can be very helpful with regard to specifying a function f_{t_k} for the dividend payment, the following lemma from the original paper is very useful.

Lemma 5.12 (Proposition 5.4 in Carr and Linetsky (2006))
Let $X \sim \chi^2(2(\nu + 1/|\beta|) + 2, Y_0^{2|\beta|}/(t\,\sigma^2|\beta|^2))$, with $\nu := \frac{1}{|\beta|}\left(\frac{\lambda}{\sigma^2} Y_0^{-2\beta} - 0.5\right)$, which denotes a non-central χ^2 distribution as introduced in Section 2.2.2. Then

$$\mathbb{E}\left[h(Z_t)\mathbb{1}_{\{\tau>t\}}\right] = Y_0\,\mathbb{E}\left[\left(|\beta|\sqrt{t\,\sigma^2 X}\right)^{-1/|\beta|} h\left(\left(|\beta|\sqrt{t\,\sigma^2 X}\right)^{1/|\beta|}\right)\right].$$

Consider a function $h : [0,\infty) \to [0,\infty)$ satisfying $h(0) = 0$. As we can write $\mathbb{E}[h(Y_{t_k})] = \mathbb{E}\left[h(Z_{t_k})\mathbb{1}_{\{\tau>t_k\}}\right]$, the quantities required for the construction of f_{t_k} as defined in Equation (5.4) can be computed numerically using the density of the non-central χ^2 distribution. For "nice" h, e.g. of polynomial structure, even closed-form expressions involving Kummer functions are available, see Corollary 5.13. However, for all computations in this thesis we used the integration with respect to the density of a χ^2 distribution, which worked both accurately and quickly.

In the sequel, unless stated otherwise, the following parameter specifications are used: $S_0 = 100$, $\beta = -1$, $\lambda = 150\text{bp} = 0.015$, $\sigma Y_0^{\beta} = 20\%$, $T = 5$, $t_k = k$, $\mathbb{E}[D_{t_k}] = 4$, for $k = 1, \ldots, 5$. The riskless rate is assumed to be given by $r = 4\%$.

5.4.2 The dividend specifications

The three different dividend specifications h_1, h_2, and h_3 presented in Section 5.3.2 will be considered in our case study. For the purpose of normalization in Equation (5.4), one has to compute $\mathbb{E}[h_i(Y_{t_k})]$. Because of the previously mentioned distinguishing feature of the JDCEV model, its tractability, closed form solutions for h_2 and h_3 can be computed based on Lemma 5.12, whereas for h_1 it is obvious that $\mathbb{E}[h_1(Y_{t_k})] = Y_0$.

Corollary 5.13 (Closed form expressions)
With $\Phi^+(p,k;\delta,\eta) := \mathbb{E}\left[X^p\mathbb{1}_{\{X>k\}}\right]$ and $\Phi^-(p,k;\delta,\eta) := \mathbb{E}\left[X^p\mathbb{1}_{\{X\le k\}}\right]$ denoting the truncated p-th moments of a $\chi^2(\delta,\eta)$ distributed random variable X, for which series

representations can be found in Lemma 5.1 in Carr and Linetsky (2006), the following holds:

$$\begin{aligned}
\mathbb{E}\left[h_2(Y_{t_k})\right] &= \frac{Y_0}{\left(|\beta|^2 t_k \sigma^2\right)^{\frac{1}{2|\beta|}}} \Phi^+\left(-\frac{1}{2|\beta|}, \hat{k}(b); \delta, \eta\right), \\
\mathbb{E}\left[h_3(Y_{t_k})\right] &= \frac{Y_0^{2-a}}{l^a} \left(|\beta|^2 t_k \sigma^2\right)^{\frac{a-1}{2|\beta|}} \Phi^-\left(\frac{a-1}{2|\beta|}, \hat{k}(l\, Y_0); \delta, \eta\right) \\
&\quad + \frac{Y_0^2}{\left(|\beta|^2 t_k \sigma^2\right)^{\frac{1}{2|\beta|}}} \Phi^+\left(-\frac{1}{2|\beta|}, \hat{k}(l\, Y_0); \delta, \eta\right) \\
&\quad + Y_0\, \Phi^+\left(0, \hat{k}(u\, Y_0); \delta, \eta\right) \\
&\quad - \frac{u\, Y_0^2}{\left(|\beta|^2 t_k \sigma^2\right)^{\frac{1}{2|\beta|}}} \Phi^+\left(-\frac{1}{2|\beta|}, \hat{k}(u\, Y_0); \delta, \eta\right),
\end{aligned}$$

where $\hat{k}(x) := \frac{x^{2|\beta|}}{t_k |\beta|^2 \sigma^2}$, $\delta = 2\left(\nu + 1/|\beta|\right) + 2$, ν *as in Lemma 5.12, and* $\eta = \frac{Y_0^{2|\beta|}}{t\, \sigma^2 |\beta|^2}$*. For* $b = 0$*, the first expression is reduced to a non-truncated moment which can be computed using the Kummer confluent hypergeometric function.*

Proof

The claim follows easily from Lemma 5.12. The series representations for the truncated moments can be determined computing the integral with respect to the density in Equation (2.10) *term by term.* □

These functions can be very helpful if one wants to speed up calculations. As mentioned before, the absolute values of the functions h_i are irrelevant as these functions are normalized in the definition of f_{t_k} in Equation (5.4). In Figure 5.4, a comparison of the different dividend specifications after normalization for $t_2 = 2$ is depicted. Additionally, the distribution of Y_{t_2} is illustrated as a barplot in the background. The bar on the left illustrates the probability of a default until t_2.

5.4.3 Resulting option prices

Having presented possible tractable specifications for f_{t_k}, the impact on the pricing of derivatives is examined. In particular, we consider stock options. The general impact of the inclusion of dividends on the pricing of those options is well-known, thus allowing us to focus on how the results differ for different dividend specifications. For the pricing, tree techniques as described in Section 5.3.3 are used. The tree for the JDCEV model

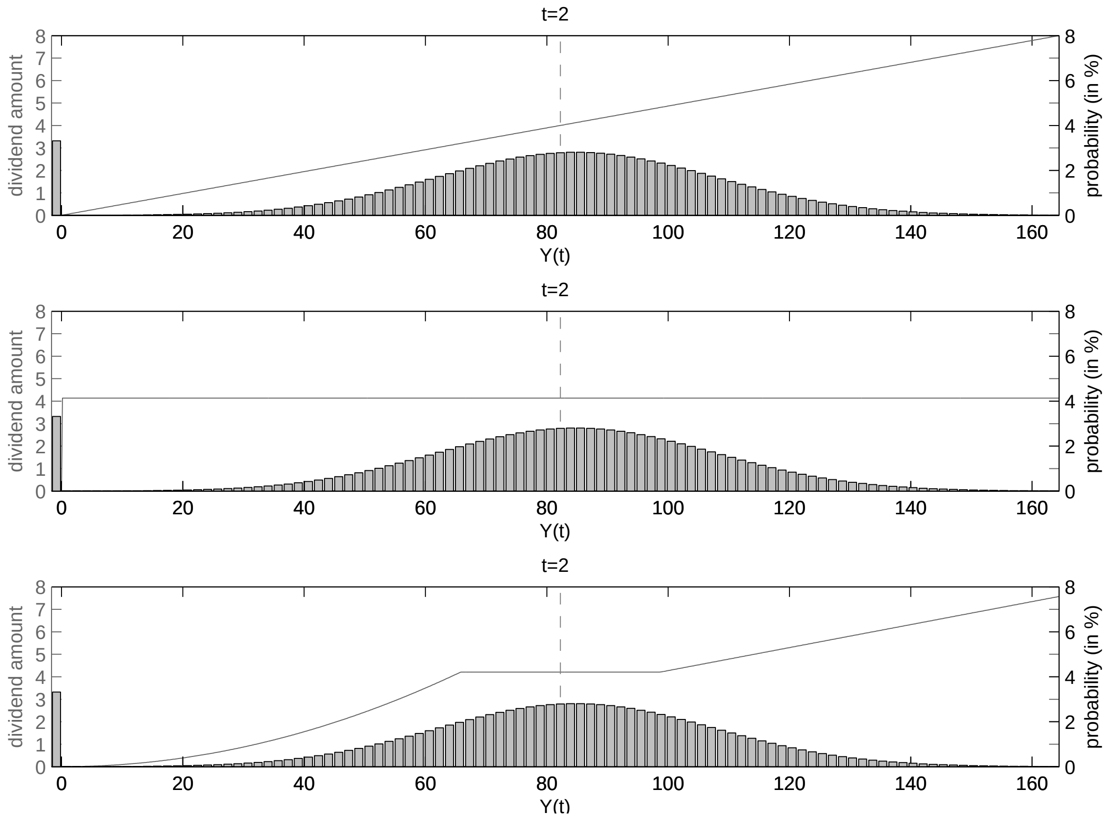


Figure 5.4 The resulting functions f_{t_2} for the different specifications of h, together with the distribution of Y_2. From top to bottom, h_1, h_2 with $b = 0$ and h_3 with $l = 0.8, u = 1.2, a = 2$ are used. The vertical dashed line indicates Y_0. The functions are normalized such that $\mathbb{E}[D_{t_2}] = 4$ holds.

is constructed by applying the techniques described in Appendix F of Brigo and Mercurio (2006) to the Bessel process in Proposition 5.1 of Carr and Linetsky (2006), and subsequently applying the corresponding transformation. Actually, this approach can be applied to price a wide range of complex derivatives, including also hybrid derivatives as, e.g., convertible bonds. As stock options are well understood and the interesting effects can thus be identified easily, the investigation is restricted to this class of derivatives. The typical shape of American and European call option prices for different maturities in the presence of discrete dividend payments can be observed in Figure 5.5 and Figure 5.6.[14] It is worth noting that the pattern observed in Bos et al. (2003) indicating obvious arbitrage for an escrowed model can not be found. Whereas the value of American call options (with fixed strike price) is increasing with increasing maturity, one can observe a downward jump of European call option prices at ex-dividend dates. This is also intuitively clear as after that point, the option does not include the right to receive the dividend payment. In contrast to that, the holder of an American call option has the possibility to receive the dividend by opting for early exercise, in case this is more favorable. The kinks in the value of American call options have a simple intuitive explanation as well. Extending the maturity of an American call option to cover just an additional dividend payment date has almost no impact on the option value. In most cases, the option would rather be exercised before the final dividend date than directly after it because of the downward jump of the stock value. Furthermore, it is clear that the value of European call options with maturity $T_1 = T = 5$ does not depend on the dividend specification, which is also confirmed in Figure 5.6.

More interesting is the relative order of the option values in Figure 5.5 and Figure 5.6 for the different specifications of the dividends. Here, one can observe that the modeling of the dividend payments has a considerable impact. We observe $V_A^{h_1}(0,.,S_0,S_0) \geq V_A^{h_3}(0,.,S_0,S_0) \geq V_A^{h_2}(0,.,S_0,S_0)$, which also holds for other strike prices (see Figure 5.7) and actually also for European call options. One can try to explain this from different perspectives. Starting from a technical deliberation, the stock price S_t can be considered (for a moment neglecting the riskless interest rate) as the sum of Y_t and a

[14] To highlight the distinctions, we additionally include the prices resulting from modeling a continuously paid dividend. Since one needs comparable prices, one has to employ a somewhat artificial construction. Starting from Equation (5.5) and using the same driving stochastic factor, we consider

$$S_t := 1/B(0,t)\, Y_t\, D \exp\left(-\int_0^t q(s)\,\mathrm{d}s\right), \quad t \geq 0$$

with a constant $D > 0$ and a piecewise constant function q. These quantities are chosen such that the expected stock price at dividend dates coincides with the true expected values.

second term representing the value of future dividends before T, which is a function of Y_t. Apart from the jumps at ex-dividend dates, the second term is a deterministic function of Y_t. The higher the change of the second term with a change of Y_t, i.e. the higher the delta, the higher the volatility of S_t and consequently the higher the value of an American call option. Judging from the convexity/concavity of the form of the different cash dividends, it is clear that $V_A^{h_1} \geq V_A^{h_2}$, whereas $V_A^{h_3}$ should lie in between. Another more concrete perspective is linked to the probability of early exercise. From the usual deliberations as, e.g., in Roll (1977), one knows that early exercise of an option is more likely the higher the dividend payment and the deeper in the money the option is just before the ex-dividend date. The higher the probability of early exercise, the higher the value of an American call option. Again, higher dividend values are the most likely for a dividend specification based on h_1 and the least likely for h_2, thus yielding another explanation for the order of the option prices. In total, the differences in prices caused by the different dividend specifications are not negligible.

In Figure 5.7, the implied Black volatility for American call options with maturity $T = 3$, different strike values and the different dividend specifications is shown. One can observe that the relative order between different dividend specifications is retained also for different strike prices. Furthermore, the shape of the implied volatility skew is very similar to the one induced by the original JDCEV model, see, e.g., the skew corresponding to the continuously paid dividend or (Carr and Linetsky, 2006, Figure 3). That means the general form of the implied volatility skew is maintained from the model for Y, slightly increasing the original skew. That observation might be in particular helpful when one has to choose a model for the underlying closed martingale.

5.5 Conclusion

Based on the fundamental ansatz of considering the stock price as the sum of expected discounted dividends, a flexible approach for the modeling of discrete cash dividends was developed. It bears some similarities to the escrowed approach, while allowing for more complex dividend modeling and ensuring arbitrage-free prices. Almost any kind of stochastic process can be used, in particular it is possible to embed defaultable processes, which seems to be important or at least reasonable for discrete cash dividends. Practical implications were discussed and possible dividend specifications were presented. Finally,

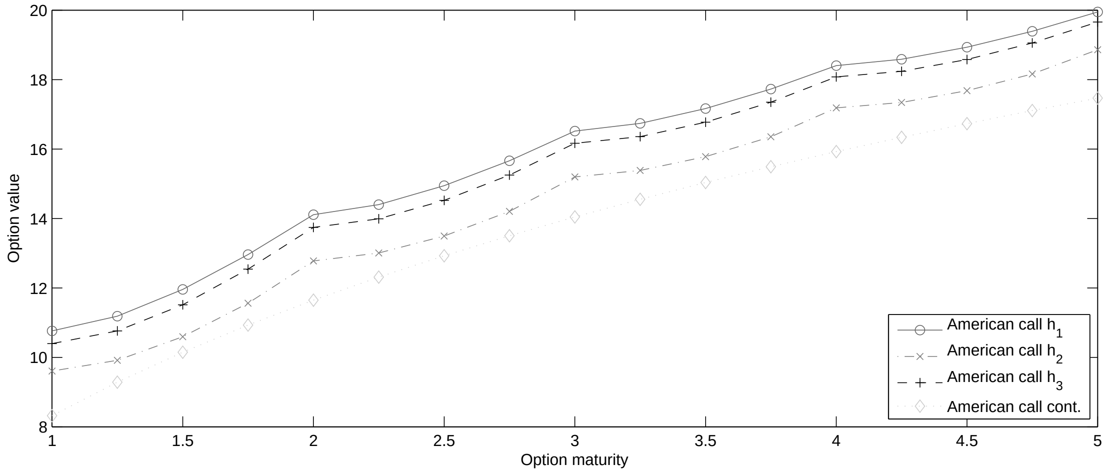


Figure 5.5 The value of American at-the-money (ATM) call options for different maturities and the different dividend specifications. Additionally, the values resulting from modeling a continuously payed dividend are added.

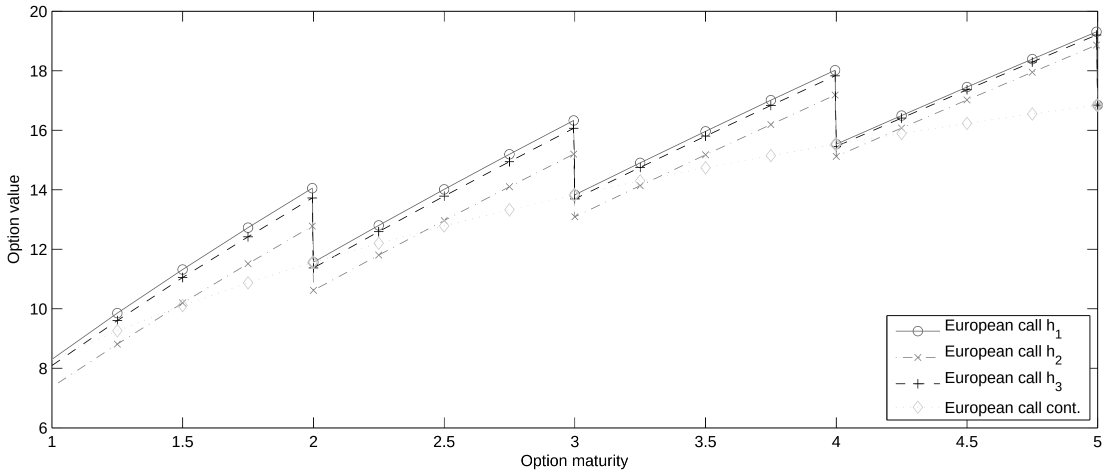


Figure 5.6 The value of European ATM call options for different maturities and the different dividend specifications. Additionally, the values resulting from modeling a continuously payed dividend are added.

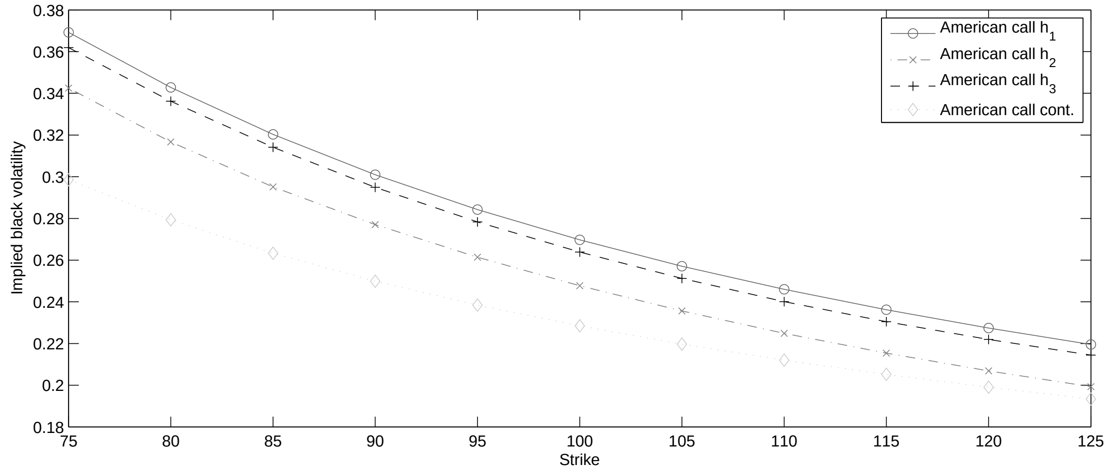


Figure 5.7 The implied Black volatility of American call options with maturity $t = 3$, for different dividend specifications (including a continuously paid dividend) and different strike values.

the impact of different dividend specifications was investigated in a case study in the context of a defaultable Markov diffusion setup.

6 Conclusion

The results of this thesis can be divided into three different parts, mirroring the structure of the thesis. In Chapter 3, a new integral representation for the density of distributions of the Bondesson class was derived. The main contribution consists of the proof of admissibility of the involved contour transform for all these distributions. As a corollary, the applicability of the approach for distribution functions and option-like structures was shown. Numerical tests confirm the tractability and smoothness of the resulting representation.

In Chapter 4, two new and flexible classes of MSMVE distributions were developed, which give rise to many parametric families. The resulting distributions are in particular suitable for the modeling of large credit portfolios, as they are tractable also in high dimensions and allow for an intuitive understanding of the underlying dependence structure. Furthermore, we were able to embed the resulting family of distributions into the context of other well-known distributions.

Finally, in Chapter 5, a new approach for the modeling of discrete cash dividends was presented. It allows to incorporate quite flexible, non-deterministic dividend structures into almost any kind of stock price model. Practical implications were discussed and illustrated in a short case study.

Bibliography

Abate, J. and Valko, P. (2004). Multi-precision Laplace transform inversion. *International Journal for Numerical Methods in Engineering*, 60(5):979–993.

Abramowitz, M. and Stegun, I. A., editors (1965). *Handbook of mathematical functions.* Dover Publications, New York.

Aldous, D. J. (1985). Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour XIII - 1983*, volume 1117 of *Lecture Notes in Mathematics*, Berlin Heidelberg. Springer.

Amin, K. and Khanna, A. (1994). Convergence of American option values from discrete- to continuous-time financial models. *Mathematical Finance*, 4(4):289–304.

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus.* Number 93 in Cambridge Studies in Advanced Mathematics. Cambridge University Press.

Ballani, F. and Schlather, M. (2011). A construction principle for multivariate extreme value distributions. *Biometrika*, 98(3):633–645.

Barndorff-Nielsen, O. E., Maejima, M., and Sato, K.-I. (2006a). Infinite divisibility for stochastic processes and time change. *Journal of Theoretical Probability*, 19(2):411–446.

Barndorff-Nielsen, O. E., Maejima, M., and Sato, K.-I. (2006b). Some classes of multivariate infinitely divisible distributions admitting stochastic integral representations. *Bernoulli*, 12(1):1–33.

Barndorff-Nielsen, O. E., Rosiński, J., and Thorbjornsen, S. (2008). General Υ-transformations. *Alea*, 4:131–165.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Normal modified stable processes. *University of Oxford, Department of Economics Discussion Paper Series*, 72.

Barrieu, P., Rouault, A., and Yor, M. (2004). A study of the Hartman–Watson distribution motivated by numerical problems related to the pricing of Asian options. *Journal of Applied Probability*, 41(4):1049–1058.

Beneder, R. and Vorst, T. (2002). Options on dividend paying stocks. In Yong, J., editor, *Proceedings of the International Conference on Mathematical Finance, Shanghai 2001*, pages 204–217. World Scientific, Singapore, 1st edition.

Benton, D. and Krishnamoorthy, K. (2003). Computing discrete mixtures of continuous distributions: noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient. *Computational Statistics & Data Analysis*, 43(2):249–267.

Bernhart, G., Escobar Anel, M., Mai, J.-F., and Scherer, M. (2013). Default models based on scale mixtures of Marshall–Olkin copulas: properties and applications. *Metrika*, 76(2):179–203.

Bernhart, G. and Mai, J.-F. (2014a). A note on the numerical evaluation of the Hartman–Watson density and distribution function. In Glau, K., Scherer, M., and Zagst, R., editors, *Innovations in Quantitative Risk Management*, Springer Proceedings in Mathematics & Statistics, pages 337–345.

Bernhart, G. and Mai, J.-F. (2014b). On convexity adjustments for stock derivatives due to stochastic repo margins. *XAIA homepage publications*.

Bernhart, G. and Mai, J.-F. (2015). Consistent modeling of discrete cash dividends. *Journal of Derivatives*, 22(3):9–19.

Bernhart, G., Mai, J.-F., Schenk, S., and Scherer, M. (2015a). The density of distributions from the Bondesson class. *Journal of Computational Finance*, 18(3):99–128.

Bernhart, G., Mai, J.-F., and Scherer, M. (2015b). Constructing min-stable multivariate exponential distributions from Bernstein functions. *Dependence Modeling*, 3(1):29–46.

Bertoin, J. (1999). Subordinators: Examples and applications. In Bernard, P., editor, *Lectures on Probability Theory and Statistics*, volume 1717, pages 1–91. Springer, Berlin Heidelberg.

Bielecki, T. R., Crépey, S., Jeanblanc, M., and Rutkowski, M. (2011). Convertible bonds in a defaultable diffusion model. In Kohatsu-Higa, A., Privault, N., and Sheu, S.-J., editors, *Stochastic Analysis with Financial Applications*, volume 65 of *Progress in Probability*, pages 255–298. Springer, Basel.

Bochner, S. (1949). Diffusion equation and stochastic processes. *Proceedings of the National Academy of Sciences of the United States of America*, 35(7):368–370.

Bondesson, L. (1979). A general result on infinite divisibility. *The Annals of Probability*, 7(6):965–979.

Bondesson, L. (1981). Classes of infinitely divisible distributions and densities. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57(1):39–71.

Bos, M. and Vandermark, S. (2002). Finessing fixed dividends. *Risk*, pages 157–158.

Bos, R., Gairat, A., and Shepeleva, A. (2003). Dealing with discrete dividends. *Risk*, pages 109–112.

Brigo, D. and Mercurio, F. (2006). *Interest Rate Models: Theory and Practice*. Springer Finance. Springer, Berlin Heidelberg New York, 2nd edition.

Buehler, H. (2010). Volatility and dividends. Working paper, TU Berlin.

Buryak, A. and Guo, I. (2011). New analytic approach to address put–call parity violation due to discrete dividends. *Applied Mathematical Finance*, 19(1):37–58.

Carr, P. and Linetsky, V. (2006). A jump to default extended CEV model: An application of Bessel processes. *Finance and Stochastics*, 10(3):303–330.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73.

Carr, P. and Wu, L. (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71(1):113–141.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons, Chichester.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155.

Damien, P., Laud, P. W., and Smith, A. F. M. (1995). Approximate random variate generation from infinitely divisible distributions with applications to Bayesian inference. *Journal of the Royal Statistical Society, Series B*, 57(3):547–563.

De Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12(4):1194–1204.

De Haan, L. and Pickands, J. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 72(4):477–492.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520.

Delbaen, F. and Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(2):215–250.

Doetsch, G. (1976). *Einführung in Theorie und Anwendung der Laplace-Transformation.* Number 24 in Lehrbücher und Monographien aus dem Gebiete der exakten Wissenschaften: Math. Reihe. Birkhäuser Verlag, Basel und Stuttgart, 3. edition.

Duffie, D. and Gârleanu, N. (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41–59.

Durante, F. and Salvadori, G. (2010). On the construction of multivariate extreme value models via copulas. *Environmetrics*, 21(2):143–161.

Eberlein, E., Glau, K., and Papapantoleon, A. (2010). Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance*, 17(3):211–240.

Eberlein, E. and Raible, S. (1999). Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1):31–53.

Elstrodt, J. (1999). *Maß- und Integrationstheore.* Springer, Berlin Heidelberg, 2nd edition.

Embrechts, P. and Hofert, M. (2013). A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432.

Es-Sebaiy, K. and Ouknine, Y. (2007). How rich is the class of processes which are infinitely divisible with respect to time? *Statistics & Probability Letters*, 78(5):537 – 547.

Esary, J. and Marshall, A. (1974). Multivariate distributions with exponential minimums. *Annals of Statistics*, 2:84–98.

Étoré, P. and Gobet, E. (2011). Stochastic expansion for the pricing of call options with discrete dividends. *Applied Mathematical Finance*, 19(3):233–264.

Evans, G. A. and Chung, K. (2000). Laplace transform inversions using optimal contours in the complex plane. *International Journal of Computer Mathematics*, 73(4):531–543.

Fougères, A.-L., Nolan, J. P., and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36(1):42–59.

Frishling, V. (2002). A discrete question. *Risk*, pages 115–116.

Gerhold, S. (2011). The Hartman–Watson distribution revisited: asymptotics for pricing Asian options. *Journal of Applied Probability*, 48(3):892–899.

Geske, R. (1979). A note on an analytical valuation formula for unprotected American call options on stocks with known dividends. *Journal of Financial Economics*, 7(4):375–380.

Gnedin, A. and Pitman, J. (2008). Moments of convex distribution functions and completely alternating sequences. In *Probability and Statistics: Essays in Honor of David A. Freedman*, volume 2. Institute of Mathematical Statistics, Beachwood, Ohio, USA.

Gordon, M. J. (1959). Dividends, earnings, and stock prices. *The Review of Economics and Statistics*, 41(2):99–105.

Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. In Jaworski, P., Durante, F., Härdle, W. K., and Rychlik, T., editors, *Copula Theory and Its Applications*, Lecture Notes in Statistics, pages 127–145. Springer, Berlin Heidelberg.

Hakassou, A. and Ouknine, Y. (2011). A contribution to the study of the IDT processes. Working paper, arXiv:1203.0874.

Hartman, P. (1976). Completely monotone families of solutions of n-th order linear differential equations and infinitely divisible distributions. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 3(2):267–287.

Haug, E., Haug, J., and Lewis, A. (2003). Back to basics: A new approach to the discrete dividend problem. *Wilmott magazine*, pages 37–47.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Hofmann, D. (2009). *Characterization of the D-Norm Corresponding to a Multivariate Extreme Value Distribution*. PhD thesis, Julis-Maximilians-Universität Würzburg, http://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/3454.

Humbert, P. (1945). Nouvelles correspondances symboliques. *Bull. Soc. Math. France*, 69:121–129.

Ishiyama, K. (2005). Methods for evaluating density functions of exponential functionals represented as integrals of geometric Brownian motion. *Methodology and Computing in Applied Probability*, 7(3):271–283.

Jänich, K. (2004). *Funktionentheorie*. Springer, Berlin Heidelberg, 6th edition.

Jarrow, R., Protter, P., and Shimbo, K. (2007). Asset price bubbles in complete markets. In Fu, M., Jarrow, R., Yen, J.-Y., and Elliott, R., editors, *Advances in Mathematical Finance*, pages 97–121. Birkhäuser Boston.

Joe, H. (1990). Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistics & Probability Letters*, 9(1):75–81.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Monographs on Statistics and Applied Probability 73. Chapman & Hall/CRC.

Jurek, Z. J. (1985). Relations between the s-selfdecomposable and selfdecomposable measures. *The Annals of Probability*, 13(2):592–608.

Jurek, Z. J. and Vervaat, W. (1983). An integral representation for selfdecomposable Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 62(2):247–262.

Kent, J. T. (1982). The spectral decomposition of a diffusion hitting time. *The Annals of Probability*, 10(1):207–219.

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.

Korn, R. and Rogers, L. (2005). Stocks paying discrete dividends: Modelling and option pricing. *Journal of Derivatives*, 13(2):44–49.

Kotz, S. and Nadarajah, S. (2000). *Extreme value distributions: theory and applications.* Imperial College Press, London.

Lanczos, C. (1964). A precision approximation of the Gamma function. *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, 1:86–96.

Larguinho, M., Dias, J., and Braumann, C. A. (2013). On the computation of option prices and greeks under the CEV model. *Quantitative Finance*, 13(6):907–917.

Lévy, P. (1934). Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 3(3-4):337–366.

Linetsky, V. (2006). Pricing equity derivatives subject to bankruptcy. *Mathematical Finance*, 16(2):255–282.

López-Fernández, M. and Palencia, C. (2004). On the numerical inversion of the Laplace transform of certain holomorphic mappings. *Applied Numerical Mathematics*, 51(2-3):289–303.

Mai, J.-F. (2013). CDO tranche pricing with the stable law - Volume I: LHP assumption. *XAIA homepage publications.*

Mai, J.-F. (2014a). *Mutivariate exponential distributions with latent factor structure and related topics.* Habilitation Thesis, Technische Universität München, https://mediatum.ub.tum.de/node?id=1236170.

Mai, J.-F. (2014b). The non-central χ^2-law in finance: A survey. *XAIA homepage publications.*

Mai, J.-F. (2014c). A note on the Galambos copula and its associated Bernstein function. *Dependence Modeling*, 2(1):22–29.

Mai, J.-F., Schenk, S., and Scherer, M. (2015). Exchangeable exogenous shock models. *Bernoulli.* Forthcoming.

Mai, J.-F. and Scherer, M. (2009a). Lévy-frailty copulas. *Journal of Multivariate Analysis*, 100(7):1567–1585.

Mai, J.-F. and Scherer, M. (2009b). 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. (2013). *Simulating Copulas*, volume 4 of *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. (2013). CIID frailty models and implied copulas. In Jaworski, P., Durante, F., and Härdle, W. K., editors, *Copulae in Mathematical and Quantitative Finance*, Lecture Notes in Statistics, pages 201–230. Springer, Berlin Heidelberg.

Maller, R. A., Solomon, D. H., and Szimayer, A. (2006). A multinomial approximation for American option prices in Lévy process models. *Mathematical Finance*, 16(4):613–633.

Mansuy, R. (2005). On processes which are infinitely divisible with respect to time. Working paper, arXiv:math/0504408.

Marshall, A. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44.

Maruyama, G. (1970). Infinitely divisible processes. *Theory of Probability & Its Applications*, 15(1):1–22.

McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative risk management: concepts, techniques, and tools.* Princeton series in finance. Princeton University Press, Princeton, New Jersey.

Mendoza-Arriaga, R., Carr, P., and Linetsky, V. (2010). Time-changed Markov processes in unified credit-equity modeling. *Mathematical Finance*, 20(4):527–569.

Molchanov, I. (2008). Convex geometry of max-stable distributions. *Extremes*, 11(3):235–259.

Müller, S. (2009). *The Binomial Approach to Option Valuation.* PhD thesis, Technische Universität Kaiserslautern.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer Series in Statistics. Springer, New York.

Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions. *Communications in Statistics-Stochastic Models*, 13(4):759–774.

Nolan, J. P. (2012). *Stable distributions - models for heavy tailed data.* Birkhauser, Boston. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

Penson, K. A. and Górska, K. (2010). Exact and explicit probability densities for one-sided Lévy stable distributions. *Physical Review Letters*, 105(21):210604.

Pickands, J. (1989). Multivariate negative exponential and extreme value distributions. In Hüsler, J. and Reiss, R.-D., editors, *Extreme Value Theory*, volume 51 of *Lecture Notes in Statistics*, pages 262–274, New York. Springer.

Raible, S. (2000). *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts.* PhD thesis, Albert-Ludwigs-Universität Freiburg.

Rajput, B. S. and Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82(3):451–487.

Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes*, volume 4 of *Applied Probability.* Springer, New York.

Ressel, P. (2013). Homogeneous distributions – and a spectral representation of classical mean values and stable tail dependence functions. *Journal of Multivariate Analysis*, 117:246–256.

Roll, R. (1977). An analytic valuation formula for unprotected American call options on stocks with known dividends. *Journal of Financial Economics*, 5(2):251–258.

Rosiński, J. (2007). Tempering stable processes. *Stochastic Processes and their Applications*, 117(6):677–707.

Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions.* Number 68 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge.

Sato, K.-I. (2004). Stochastic integrals in additive processes and application to semi-Lévy processes. *Osaka J. Math.*, 41(1):211–236.

Sato, K.-I. (2006). Two families of improper stochastic integrals with respect to Lévy processes. *Alea*, 1:47–87.

Schilling, R., Song, R., and Vondracek, Z. (2010). *Bernstein functions.* Number 37 in De Gruyter Studies in Mathematics. De Gruyter, Berlin.

Schönbucher, P. J. and Schubert, D. (2001). Copula-dependent defaults in intensity models. Working paper, SSRN eLibrary.

Segers, J. (2012). Max-stable models for multivariate extremes. *REVSTAT – Statistical Journal*, 10(1):61–82.

Seshadri, V. (1993). *The Inverse Gaussian distribution: a case study in exponential families.* Oxford science publications. Clarendon Press.

Shuster, J. (1968). On the Inverse Gaussian distribution function. *Journal of the American Statistical Association*, 63(324):1514–1516.

Steutel, F. W. and van Harn, K. (2004). *Infinite Divisibility of probability distributions on the real line.* Pure and applied mathematics. Marcel Dekker, Inc., New York Basel.

Talbot, A. (1979). The accurate numerical inversion of Laplace transforms. *Journal of the Institute of Mathematics and its Applications*, 23(1):97–120.

Thorin, O. (1977a). On the infinite divisbility of the Pareto distribution. *Scandinavian Actuarial Journal*, 1977(1):31–40.

Thorin, O. (1977b). On the infinite divisibility of the lognormal distribution. *Scandinavian Actuarial Journal*, 1977(3):121–148.

van Binsbergen, J., Brand, M., and Koijen, R. (2012). On the timing and pricing of dividends. *American Economic Review*, 102(4):1596–1618.

Vasicek, O. A. (2002). Loan portfolio value. *Risk*, pages 160–162.

Veiga, C. and Wystup, U. (2009). Closed formula for options with discrete dividends and its derivatives. *Applied Mathematical Finance*, 16(6):517–531.

Vellenkoop, M. and Nieuwenhuis, J. (2006). Efficient pricing of derivatives on assets with discrete dividends. *Applied Mathematical Finance*, 13(3):265–284.

Villiger, R. (2006). Valuation of American call options. *Wilmott Magazine*, pages 64–67.

Weideman, J. A. C. and Trefethen, L. N. (2007). Parbolic and hyperbolic contours for computing the Bromwich integral. *Mathematics of Computation*, 76(259):1341 – 1356.

Whaley, R. E. (1981). On the valuation of American call options on stocks with known dividends. *Journal of Financial Economics*, 9(2):207–211.

Widder, D. V. (1946). *The Laplace transform.* Number 6 in Princeton Mathematical Series. Princeton University Press.

Williamson, R. (1956). Multiply monotone functions and their Laplace transforms. *Duke Mathematical Journal*, 23(2):189–207.

Wolfe, S. J. (1975). On moments of probability distribution functions. In Ross, B., editor, *Fractional Calculus and Its Applications*, volume 457 of *Lecture Notes in Mathematics*, pages 306–316. Springer, Berlin Heidelberg.

Yor, M. (1980). Loi de l'indice du lacet brownien, et distribution de hartman–watson. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 53(1):71–95.

Yor, M. (1992). On some exponential functionals of Brownian motion. *Advances in Applied Probability*, 24(3):509–531.

List of Tables

List of Figures